



Study of the conformational dynamics of prolyl oligopeptidase

Abraham López Asamar

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) i a través del Dipòsit Digital de la UB (diposit.ub.edu) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) y a través del Repositorio Digital de la UB (diposit.ub.edu) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service and by the UB Digital Repository (diposit.ub.edu) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

Programa de Doctorat en Química Orgànica

Universitat de Barcelona

Memòria titulada:

**Study of the conformational dynamics of prolyl
oligopeptidase**

Presentada per:

Abraham López Asamar

Directors de tesi:

Prof. Ernest Giralt i Lledó

Dra. Teresa Tarragó i Clua

Universitat de Barcelona
Facultat de Química
Departament de Química Orgànica

Institute for Research in Biomedicine
Programa de Química i Farmacologia
Molecular



**INSTITUTE
FOR RESEARCH
IN BIOMEDICINE**

Als meus pares
i al meu germà

Agraïments

Doncs, per fer-vos una idea, son les 5 de la matinada del dia abans de portar a imprimir la tesi... i ja intueixo que les tardes interminables tancat a casa escrivint i les nits que s'allarguen massa (davant l'ordinador), poc a poc, van arribant a la seva fi.

Enmig d'aquesta voràgine que ha anat *in crescendo*, m'he parat a pensar sovint que quant defensi aquesta tesi deixaré enrere una etapa molt important de la meua vida entre les parets de l'entranyable *lab 300*. Sembla que fos ahir quan, després dels estudis de Química, vaig decidir treure el nas pel Parc Científic "per veure què s'hi covia". I vaig topar amb el grup de l'Ernest, i vaig pensar: ostres, això és completament diferent al que he vist fins ara!

Crec que serà difícil dedicar un moment a totes aquelles persones que he anat coneixent al llarg d'aquesta trajectòria (havia pensat en incloure també algun dels nostres *simpàtics* equips, com el HPLC-MS antic, però he decidit que millor que no...). Així que seguiré una trajectòria cronològica recordant a les persones (i personatges!) que conegut al llarg d'aquesta etapa.

Lògicament, les primeres persones a qui vaig conèixer van ser els meus *jefes*, l'Ernest i la Tere. Agraïr-vos a tots dos haver-me dedicat part del vostre valuós temps per ensenyar-me tants coneixements. Però no només això: el meu aprenentatge adquirit aquí no s'ha limitat a l'estudi científic d'una proteïna inquieta de 81 KDa: també he rebut uns valors molt importants de treball en equip i de respecte.

Així que també us dedico una gran part d'aquesta tesi a vosaltres. A tu Ernest per ser una autèntica eminència científica, però també a la vegada una persona propera i que sempre inspira confiança. A tu Tere per la teua capacitat (*increïble*) de treball i per la teua atenció amb la gent (no només als de la feina, sinó també als teus!). També aprofito per desitjar-te que vagin molt bé les coses a iProteos!

El que dèiem. Seguint amb l'ordre temporal, el meu primerencontre abans de començar a treballar al lab va ser en el retreat dels nous *PhD students*, a l'hotel Muntanyà... i va ser allà on vaig conèixer algunes de les persones amb qui més *cool-offs*, cafès i viatges andalusos compartiria al llarg de la meua estança al IRB, l'Albert i l'Òscar. Desitjar-vos molta sort a tu Albert amb el *postdog* i a tu Óscar amb la nova empresa (ets un crack!).

Bueno, doncs ara es el torn del *POP team*!! El primero, cómo no, Nessim! Sólo decirte que creo que ya lo tenemos, por fin hemos capturado la forma abierta de la POP! Brindemos! I per descomptat la Mendi, la "baila salsa chica que"! Des d'aquí agraïr-te molt per l'ajuda, i per mil i una anècdotes de lab! I si, reconec que algunes vegades sóc massa agonies! També a l'Albert Puigpinós i a la Marta Mena, per tenir solucions i consells per tot! Ahora, rompiendo la progresión cronológica, me gustaria desear mucha suerte al relevo del *POP*

team, a la Dra. Monica Varese. Ànimo, que ya verás que saldrá algo interesante!! La POP se hace esperar!

Seguint amb ordre cronològic, recordar a les *viejas glorias* del lab 300... em refereixo al Michael, l'Andrey, El Peter i l'Anna-Iris. Es troba a faltar el punt *destroyer* tan característic vostre! Innombrables records, converses o senzillament, conyes de laboratori. I moltes cerveses i concerts!! A veure si ara que tinc temps us escric un dia i m'expliqueu!

També agrair a la gent del lab 100 per haver-me format durant la meva estada d'alumne intern al Departament de Química Orgànica, abans de començar el doctorat; i aquí a l'IRB, per haver-me ajudat i explicat molt *trucos sintètics* en les meves picabaralles amb la trifluorometionina: en Xavi Just, l'Adriana Lorente, Paolo Bruno i por supuesto Janire!

I ara és la tongada dels més propers a la meva promoció! Com per no enrecordarse'n del Dr. Xavi Vila i els *piknics electroniks* (amb els seus corresponents dilluns catastròfics)! Ja t'imagino passejant per la *5th Avenue* amb les teves samarretes hawaianes! Molta sort amb la teva nova etapa a NY! El Pepe, un altre que tall!! Merci per la bona ona que transmetes! I recorda... I'm watching you! Dr. Benjamí, espero que estiguis gaudint d'una petita pausa i gaudint de la muntanya, per seguir amb més força amb la teva carrera científica. Ets un crack! I ara tu Sílvia! Que les nostres tesis seran bessones! Anims que ja queda poc... (el que tothom et diu tot i que sembli que mai s'acabi l'escriptura, eh?). També a el Salva, que ens queden moltes pedalades per fer... ens apuntem a *el Sopiao*? Va, que només és com dos *Tramunbikes* juntes! Pol, anims que tot i que els UPLC no funcionin, el sol surt cada dia... segur? A tu Sonia, moltes gràcies per la teva atenció i per estar sempre disposada a donar un cop de mà. Júlia, ànims amb els *biciclic peptides*, i gràcies per mantenir la vidilla al laboratori!! I també a los coleg@s italianos, Fabio (tengo pendiente una visita en Bologna!) y Roberta, que se os echa de menos! *Mà dai!* Por supuesto no me olvido de Laura Nevola (gracias por tus consejos). I donar molts ànims als que comenceu ara al lab, ja sigui per fer pràctiques, o començar la tesi, sort! I pensant en muntanya i excursions, és el torn de la Cris García. Quines ganes de muntanya m'has fet venir amb les fotos de les teves excursions i escapades!! Espero que segueixis gaudint de la muntanya durant molt de temps!! I com no, agrair també a l'Esther Zurita i a la Eva Poca, per tota l'ajuda rebuda amablement en les comandes, paperassa, etc.

Agrair per descomptat a la Dra. Marta Vilaseca, per haver-me introduït en el món de l'espectrometria de masses (del qual n'era un desconexedor total), sempre amb l'amabilitat i l'atenció que et caracteritza.

Bueno, i ens apropem al torn de recordar la meva estada a Bilbao! Recordo que des del primer moment que em vaig veure amb les maletes enmig del *zazpikale* vaig sentir que la cosa prometia... i no em vaig equivocar! Conservo un record vivíssim de totes les experiències que hi he viscut. A més, va ser allà on la POP va decidir ser benévola amb mi. Poc a poc, en els següents mesos, el misteri de la POP s'anava resolent. Però també va ser

una època de molta bici i muntanya, cosa que ajudava a desconectar i mantenir la tranquil·litat.

Com no, dedicar una part d'aquest treball al Dr. Oscar Millet, un *coco* de la RMN i també excel·lent persona (i molt bon cuiner!). El que més m'ha sorprès gratament de tu ha sigut la teva proximitat i sinceritat. Espero que les coses vagin molt bé pel lab i segueixis amb els teus passejos pel *Muelle Martzana*!! Bueno y por supuesto a ti Gabo! Nuestro querido *cántabru*, muchas gracias por tu inestimable ayuda, se echan de menos los *poteos* y las *calçotades*... También dar las gracias a todos aquellos que me ayudaron durante mi estancia, es imposible nombraros a todos aquí: Paula (y su iniciativa de apuntarse a un bombardeo), Ana, Idoia, Siva, *et al*.

Per descomptat, agrair sincerament al Dr. Pau Bernadó i a la seva dedicació en el projecte de la POP. També una part d'aquesta tesi et va dedicada, gràcies per compartir el teu coneixement també d'una manera molt propera i honesta. Encara recordo el dia (bé, la nit) en què vam fer la mesura del tàndem SEC-SAXS a Hamburg... Deu ni dó, veure sortir el sol des d'un laboratori no té desperdici, però l'esforç va valer molt la pena! Especialmente gracias a ti Fátima, lo siento por insistir en nuestras cadenas de mails eternas sobre si tal o cual *frame*, que si el chi cuadrado... Gracias por ser tan atenta y soportarme! Gracias también a Martin Kotev por tu ayuda con las simulaciones.

Finament, *last but not least*, agrair als meus pares per haver-me donat suport en el dia a dia durant tota aquesta època. També al meu germà, que ja va passar fa no gaire per aquest tràngol (deu ni do, eh?). I especialment a tu Anna, que també m'has fet molt de suport en aquestes últimes estrebades de la tesis, moltes gràcies! Bueno i com no, els de la colla, de Girona, que aquests pla que coneixen els meus canvis d'humor derivats de la tesi. Merci!

Barcelona, juny del 2015

Index

ABBREVIATIONS	1
INTRODUCTION AND OBJECTIVES	3
Structure of prolyl oligopeptidase.....	3
<i>POP structure and organization.....</i>	<i>3</i>
<i>The active site.....</i>	<i>5</i>
Biological function and mechanisms of prolyl oligopeptidase	7
POP inhibitors: drug design and effects.....	11
POP enzymatic mechanism and conformational dynamics.....	13
Objectives.....	18
CHAPTER 1: STUDY OF POP BY NUCLEAR MAGNETIC RESONANCE	19
Introduction: nmr labeling strategies for large proteins and protein complexes	19
Methyl groups in biomolecular NMR	20
Trifluoromethionine POP labeling	21
Use of ¹⁹ F as NMR probe.....	21
Synthesis of trifluoromethionine and expression of trifluoromethionine labeled POP ...	23
Preliminary ¹⁹ F NMR spectra of trifluoromethionine labeled POP	28
Discussion	30
[Methyl-¹³C]-methionine POP labeling: ¹H-¹³C methyl-TROSY HMQC spectra.....	31
Assignment of ¹ H- ¹³ C methyl-TROSY HMQC spectra	32
Analysis of ¹ H- ¹³ C methyl-TROSY HMQC spectra and effect of inhibitors	37
U-[²H], [methyl-¹³C, ¹H]-methionine POP labeling: ¹H-¹³C methyl-TROSY HMQC relaxation dispersion measurements	40
¹ H- ¹³ C methyl-TROSY HMQC relaxation dispersion measurements	40
<i>Synthesis of [methyl-¹³C]-methionine (2, 3, 3, 4, 4 - d₅) and expression of perdeuterated proteins</i>	<i>45</i>
<i>NMR experiments.....</i>	<i>47</i>
<i>Extraction of exchange parameters</i>	<i>49</i>
¹ H- ¹³ C Methyl-TROSY HMQC rotating frame relaxation dispersion experiments	50
Results and discussion.....	52
CHAPTER 2: ANALYSIS OF POP STRUCTURE IN SOLUTION BY SMALL-ANGLE X-RAY SCATTERING	59

Introduction	59
Experimental set-up: tandem size exclusion chromatography-SAXS measurement	62
SAXS data processing	63
Generation of $I(0)$ and R_g SEC-SAXS chromatograms.....	63
Singular value decomposition of monomer region of SEC-SAXS $I(0)$ chromatograms .	66
Multivariate curve resolution by alternating least squares of peak II	67
Results and discussion.....	69
CHAPTER 3: POP MOLECULAR DYNAMICS SIMULATIONS AND CORRELATION WITH EXPERIMENTAL DATA	71
Introduction	71
Principles of molecular dynamics simulations	72
<i>Choosing the initial structure</i>	75
<i>Initial energy equilibration</i>	75
<i>Production run</i>	75
<i>Trajectory analysis</i>	76
Starting structures and homology models of POP	77
MD simulations of free POP.....	78
MD simulations of inhibited POP	80
Analysis of MD trajectories	80
<i>MD1</i>	81
<i>MD2</i>	81
<i>MD3</i>	82
<i>MD4 and MD5</i>	83
Experimental data analysis.....	84
Ensemble optimization fitting of theoretical SAXS scattering profiles	84
<i>EOM of free POP peaks I and II</i>	88
<i>EOM of ZPP-bound POP peak III</i>	89
High resolution structural analysis of POP conformers	89
Correlation with NMR data	93
<i>Root mean square fluctuation</i>	94
<i>Solvent accessible surface area</i>	95
Discussion	96

CHAPTER 4: APPLICATION OF ION MOBILITY MASS SPECTROMETRY IN THE STUDY OF POP CONFORMATIONAL DYNAMICS99

Introduction: mass spectrometry in structural biology99

Soft ionization techniques.....99

Protein ions in the gas phase.....100

Ion mobility mass spectrometry of biomolecules..... 102

Types of ion mobility mass spectrometers 102

Applications of ion mobility mass spectrometry for the study of proteins and protein complexes..... 105

Study of POP structure and dynamics by IMMS 107

Optimization of experimental conditions 108

Optimization of the ionization and ion transmission conditions 108

Optimization of the TWIMS conditions..... 109

IMMS experiments of POP 110

Free and inhibitor-bound native POP..... 110

Collision induced unfolding..... 111

Charge effects and acid denaturing of gaseous POP ions 111

Native POP IMMS measurements in negative mode 112

IMMS data processing 112

Molecular dynamics simulations and theoretical collision cross sections calculations. 115

Replica exchange molecular dynamics simulations in solution 115

Molecular dynamics simulations in the gas phase 116

Theoretical collision cross section calculation 119

Results..... 120

Free and inhibitor-bound POP under native conditions 120

POP conformational landscape in the gas phase: CIU, acid denaturing and reduced charge states 126

Discussion 129

CHAPTER 5: INTERACTION OF POP WITH α -SYNUCLEIN135

Involvement of POP in α -synuclein clearance: possible interaction of POP with α -synuclein 135

Effects of POP inhibition in α -synuclein aggregation 136

Study of protein-protein interactions..... 140

Possible NMR experiments for the detection of the interaction between POP and α -synuclein..... 141

Chemical shift perturbation 142

<i>Approaches based on relaxation properties</i>	143
Preliminary NMR experiments	145
NMR experiments of [methyl- ¹³ C]-methionine labeled POP	149
NMR experiments of U- ¹⁵ N labeled α-synuclein	151
<i>Optimization of NMR conditions and data treatment</i>	151
¹ H- ¹⁵ N HSQC experiments of α-synuclein in the presence of free POP	152
¹ H- ¹⁵ N HSQC experiments of α-synuclein in the presence of ZPP-bound POP.....	155
Discussion	155
GENERAL DISCUSSION	159
MATERIALS AND METHODS	163
Protein expression	163
POP expression protocols	163
<i>Expression of native POP</i>	163
<i>Expression of L-TFM labeled POP</i>	166
<i>Expression of [methyl-¹³C]-methionine labeled POP</i>	167
<i>Expression of perdeuterated and U-[²H], [methyl-¹³C, ¹H]-methionine labeled POP</i>	167
Expression and purification of WT α-synuclein.....	168
Synthetic procedures	170
Synthesis of L-trifluoromethionine	170
Synthesis of [methyl- ¹³ C]-methionine (2, 3, 3, 4, 4-d ₅).....	172
NMR experiments	180
¹ H- ¹³ C methyl-TROSY HMQC experiments of [methyl- ¹³ C]-methionine labeled POP ..	180
RD experiments of U-[² H], [methyl- ¹³ C, ¹ H]-methionine labeled POP	180
<i>R_{1ρ}</i> experiments of perdeuterated [methyl- ¹³ C]-methionine labeled POP.....	181
Processing of ¹ H- ¹⁵ N HSQC experiments of U- ¹⁵ N α-synuclein.....	181
Size exclusion chromatography coupled to SAXS measurements	182
Theoretical simulations	182
Molecular dynamics simulations of POP	182
Molecular dynamics simulations for IMMS experiments and theoretical collision cross sections calculations	183
Experimental procedure for IMMS experiments	184

BIBLIOGRAPHY	185
CONCLUSIONS	199
RESUM	201
APPENDIX	223
Appendix I: relaxation in NMR	223
NMR basics.....	223
Relaxation in NMR	225
CPMG RD experiments	228
$R_{1\rho}$ RD experiments	230
Appendix II: low resolution structural characterization of biomolecules by small-angle X-ray scattering	231
Theory	232
SAXS scattering profiles of biomolecules in solution	234
<i>Structural information from SAXS data</i>	<i>235</i>
<i>Ab initio modelling and calculation of theoretical scattering profiles</i>	<i>237</i>
Bibliography	238
Appendix III: publications	239

Abbreviations

ACh: acetylcholine

ADME: absorption, distribution, metabolism and excretion

CCS: collision cross section

CIU: collision induced unfolding

CNS: central nervous system

CSP: chemical shift perturbation

DTT: dithiothreitol

EOM: ensemble optimization method

FPLC: fast protein liquid chromatography

HB: hydrogen bond

HMQC: heteronuclear multiple quantum correlation

HSQC: heteronuclear single quantum correlation

IDP: intrinsically disordered protein

IMMS: ion mobility mass spectrometry

$I(0)$: forward scattering

L-TFM: *L*-trifluoromethionine

MD: molecular dynamics

MES: 2-(*N*-morpholino)ethanesulfonic acid

Met: methionine

MS: mass spectrometry

NMR: nuclear magnetic resonance

NOE: nuclear Overhauser effect

PAGE/SDS-PAGE: polyacrylamide gel electrophoresis/polyacrylamide gel electrophoresis using sodium dodecylsulfate as denaturing

PD: Parkinson's disease

PDB: protein data bank

POP: prolyl oligopeptidase

PPI: protein-protein interaction

$P(r)$: pair-distance distribution function

Ps, ns, μ s, ms: pico-, nano-, micro-, milliseconds

RD: relaxation dispersion

REMD: replica exchange molecular dynamics simulation

R_g : radius of gyration

RMSD: root mean square deviation

RMSF: root mean square fluctuation

$R_1, R_2, R_{1\rho}$: longitudinal, transverse, rotating frame relaxation rates

SASA: solvent accessible surface area

SAXS: small-angle X-ray scattering

TRIS (Tris): tris(hydroxymethyl)aminomethane

TROSY: transverse relaxation optimized spectroscopy

WT: wild-type

ZPP: benzyloxycarbonyl prolyl proline

Introduction and objectives

STRUCTURE OF PROLYL OLIGOPEPTIDASE

Prolyl Oligopeptidase (POP; EC 3.4.21.26) is a monomeric 81-KDa cytosolic serine endopeptidase which hydrolyses peptides under 30 residues at the carboxyl side of proline [1]. POP is ubiquitous in mammals, but a relatively high concentration of this protein is found in the central nervous system (CNS). The first X-ray structure was obtained by Fulop *et al.* (PDB entries 1QFS and 1QFM) from porcine muscle POP [2]. The crystallographic structure showed that this enzyme has an overall cylindrical shape, constituted by two domains: the α/β -hydrolase and the β -propeller (Figure 1).

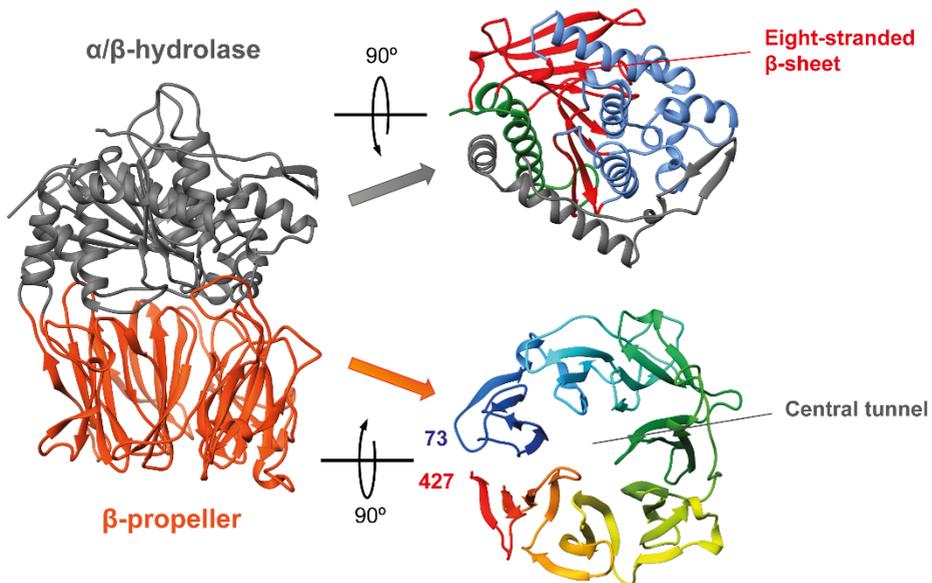


Figure 1: Global structure of POP. The α/β -hydrolase is depicted in gray and the β -propeller in orange. The two domains are shown separately at the right. In the α/β -hydrolase, the eight-stranded β -sheet is depicted in red, while the α -helices are depicted in light blue or green according to the side of the β -sheet in which they are located. In turn, the blades of the β -propeller are colour-ramped blue to red from the first to the last residue. The central tunnel of the β -propeller is also shown.

POP structure and organization

The α/β -hydrolase domain (residues from 1 to 72 and 428 to 710) contains the catalytic triad Ser 554, His 680 and Asp 641 which is responsible for the catalytic activity (Figure 2

A). This domain is formed by a significantly twisted central eight-stranded β -sheet, with all the sheets in a parallel orientation except the second one. The central β -sheet core is surrounded by six α -helices in one side and by two α -helices in the other (Figure 1). The α/β -hydrolase fold is common in all serine peptidases [3].

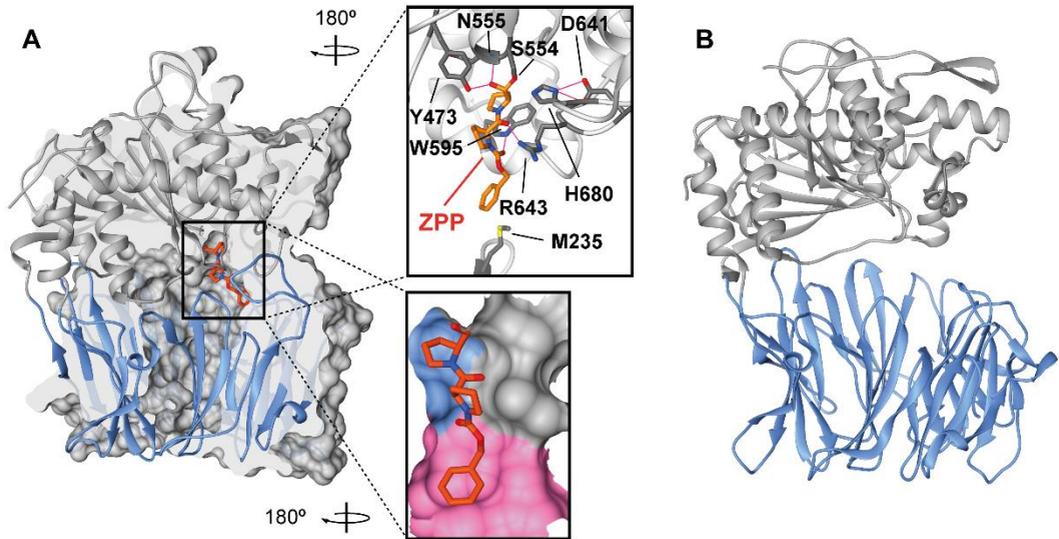


Figure 2: X-ray structures of POP and inhibitor binding. A) X-ray structure of the closed conformation of POP bound to ZPP (PDB entry 1QFS). Insets show magnified views of the active site interacting with ZPP (orange); the catalytic triad and other residues involved in binding are marked. Salt bridges and HB are depicted in magenta, and binding sites S1 and S3 coloured in blue and pink, respectively (down). B) *Aeromonas punctata* POP in the open conformation (PDB entry 3IUJ). In both structures, the α/β -hydrolase is depicted in gray and the β -propeller in blue.

However, the β -propeller domain (residues from 73 to 427) is exclusive for the prolyl oligopeptidase family. The β -propeller is formed by seven blades adopting a radial distribution; each blade is formed by a four-stranded antiparallel β -sheets hold together mainly by hydrophobic forces [2, 4]. This arrangement creates a central tunnel which connects the active site with a small pore opened to the solvent in the opposite end (Figures 1 and 2 A). The β -propeller domain of POP is a relatively stable structure; in fact, it can be expressed as a soluble monomeric protein [5]. However, contrarily to what occurs in most of β -propellers found in other proteins [6], the first and seventh blades are interacting only by weak hydrophobic interactions. This feature confers a particular flexibility of the β -propeller which affects interdomain dynamics (see below).

The two ends of the β -propeller are connected to the α/β -hydrolase by two flexible polypeptide linkers. The crystal structure of the porcine enzyme (PDB entry 1QFM) showed that the two domains are interacting via hydrophobic forces and 27 hydrogen bonds and salt bridges [2]. As a result, both domains are kept in contact in a closed conformation which protects the active site from the solvent (Figure 2 A). Moreover, all the X-ray structures of the mammalian free POP and POP bound with inhibitors/substrates have been crystallized in a similar closed conformation (i.e. PDB entries 1H2W and 2XDW) suggesting that POP exclusively exists in a closed conformation.

In 2005, Shan *et al.* crystallized *Sphingomonas capsulata* POP in a conformation showing a separation between the α/β -hydrolase and β -propeller domains (PDB entry 2BKL) [7]. Few years later, Li and co-workers crystallized a second bacterial POP from *Aeromonas punctata* adopting a similar interdomain separation (Figure 2 B, PDB entries 3IUL and 3IUJ) [8]. Interestingly, soaking the pre-formed crystals of *Aeromonas punctata* POP in the open conformation with substrates and the active site-directed covalent inhibitor benzyloxycarbonyl prolyl proline (ZPP, [9]) caused the shift towards the closed conformation. Although the alignment of bacterial and porcine POP sequences displays poor correlation (only 40 % similarity), their closed conformations show similar tertiary structures (RMSD of 1.653 Å). Hence, several studies of other groups proposed that mammalian POP undergo interdomain dynamics exploring different open conformational states (this question will be explained later in this introduction).

The active site

The active site of POP is buried between the β -propeller and the α/β -hydrolase. It has been characterized in detail after the crystallization of several POP and POP mutants complexed with transition state analogue inhibitors or substrates (for instance, see references [2, 10, 11]). These studies showed that the conformation and interactions of bound inhibitors and peptide substrates were highly similar. Figure 2 A shows the inhibitor ZPP covalently bound to the nucleophile Ser 554 through a hemiacetal bond. This residue is placed in a sharp turn in close proximity to the proton-donor His 680 [10]; at the same time, His 680 is hydrogen bonded to Asp 641. Overall, the disposition of the catalytic triad in the presence of ZPP resembles the transition state; several active site-directed covalent inhibitors have been found to act as transition state analogues.

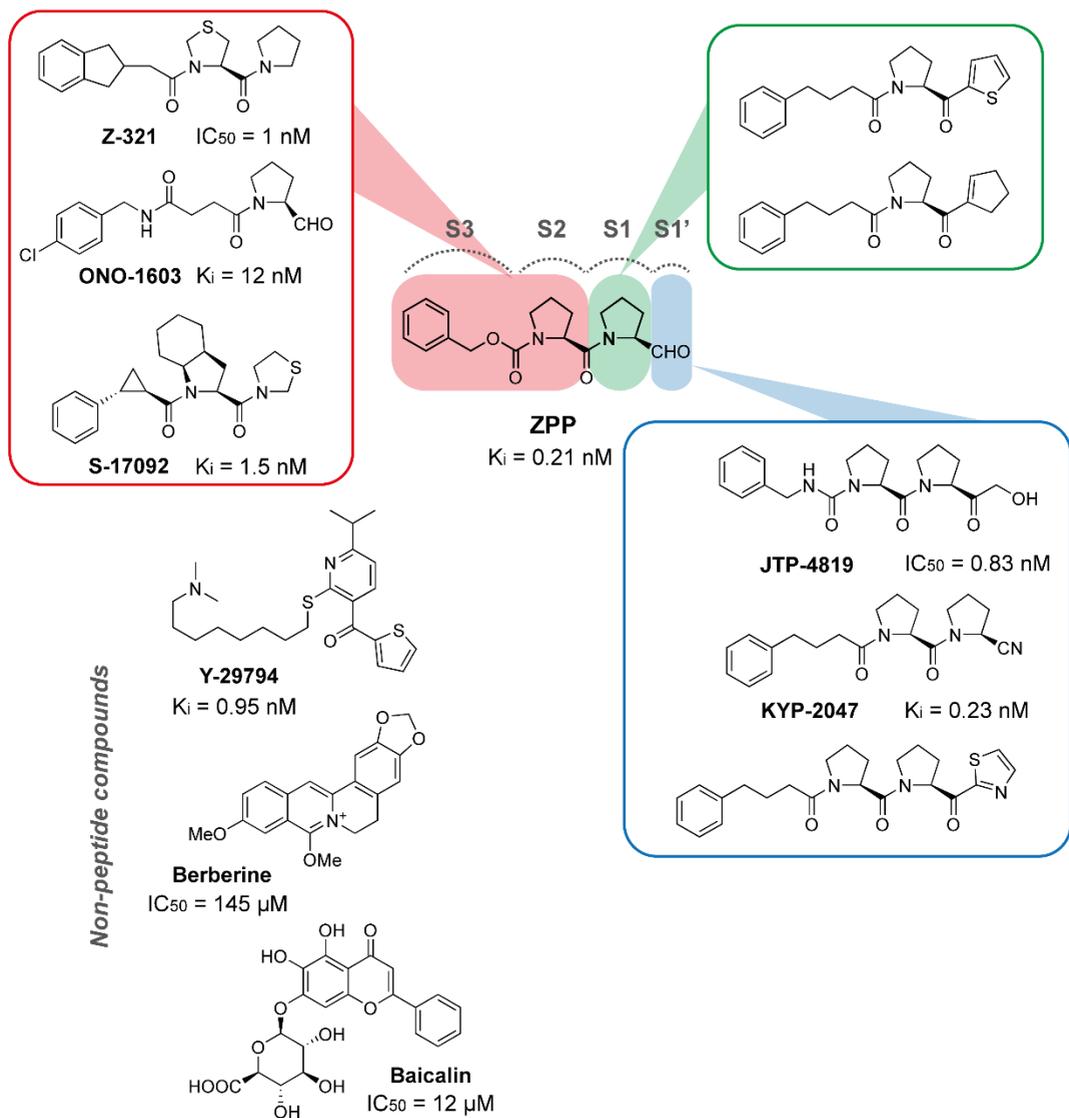


Figure 3: Representative POP inhibitors. The canonical inhibitor ZPP is divided according to the binding sites S1', S1, S2 and S3. Inhibitors have been classified on the basis of the modifications in P1', P1 and P2-P3 positions. Three non-peptide compounds with POP inhibitory activity are also shown. In most cases, the company names and potencies are included.

The binding of the substrate/inhibitor involves several hydrophobic binding sites, referred to S1, S2 and S3 according to the Schechter and Bergen notation [12] (Figure 2 A and 3). Placed before P1, the oxyanion binding site is responsible for the accommodation of the carbonyl oxygen of the scissile bond, which become a negatively charged leaving group

after cleavage [13]. Oxyanion binding site is constituted by the backbone NH of Asn 555 and the hydroxyl group of Tyr 473. S1 is responsible for the binding of the proline ring of ZPP, which is stacked against the indole group of Trp 595; moreover, other residues favour the tight packing of proline ring, namely Phe 476, Val 644, Val 580, and Tyr 599. S2 has no significant structural properties, and S3 is a large hydrophobic pocket which accommodates bulky hydrophobic groups. This hydrophobic pocket is formed by Phe 173, Met 235, Cys 255, Ile 591 and Ala 594 [2]. Moreover, several hydrogen bonds (HB) have important roles in the fixation of the carbonyl groups of substrates/inhibitors in the proper orientation. Arg 643 interacts through a HB with the carbonyl group between P1 and P2 positions of ZPP, and Trp 595 is bound through a HB with the carbonyl group between P2 and P3 [2].

Most POP inhibitors are peptidomimetic compounds derived from the canonical compound ZPP, which act as transition state or substrate analogues (Figure 3). Most of the modifications have been carried out in P1', P1 and P3 positions in order to increase the potency, selectivity and ADME properties. P1' modifications usually include different electrophilic groups that react with Ser 554, for instance, nitriles (e.g. KYP-2047 [14]) or ketones activated by electron-withdrawing groups [15]; these compounds form imino ether [11] or hemiketal bonds with Ser 554, respectively. Modifications in P1 include several saturated and unsaturated cyclic structures that mimic the natural proline ring [16]. A large number of modifications have been developed in P3 position, typically based on aromatic groups separated by an aliphatic spacer chain [17]. Finally, few non-peptide compounds with POP inhibitory activity have been described, for instance Y-29794 [18] and the natural products baicalin [19] and berberine [20]. Nevertheless, the binding mode of these compounds have not been described. A deeper description of POP inhibitors will not be included here since several extensive reviews have been published [21, 22].

BIOLOGICAL FUNCTION AND MECHANISMS OF PROLYL OLIGOPEPTIDASE

Proline is a natural amino acid with particular chemical and conformational properties derived from its cyclic structure. Proline-containing peptides cannot be recognized and hydrolysed by conventional peptidases, and therefore they are protected from general proteolytic degradation. For this reason, most of peptides with proline content are involved in specific physiological processes. Organisms have developed proline-specific peptidases

in order to recognise and degrade these peptides. Proline-specific peptidases usually play a crucial role in the regulation of biological processes and the dysfunction or deficiency of these peptidases is related to diseases. Herein, some of these enzymes are potential therapeutic targets [1].

Substrate	Sequence
Arginine-vasopressin (AVP)	CYFQNC PRG
Substance P (SP)	RPK PQQFFGLM
Oxytocin	CYIQNCPLGNH2
Thyrotropin releasing hormone: precursor peptide (THR)	KRQH PG QR
Neurotensin	pQLYENK PRR PVIL
Angiotensin II	DRVYI HPF
Bradykinin	RPPGF SPFR
Humanin	MAPR GSFCLLLLTSEIDL PVK RRA

Table 1: Some neuropeptide and peptide hormones degraded by POP. The cleavage site is shown in bold.

The biological function of POP was early related to mental processes. Preliminary studies carried out by Reichelt *et al.* in the beginning of the 80s [23], and later by Maes *et al.* [24], found an altered POP activity in the serum of patients suffering from mental disorders. The first hypothesis about the physiological role of POP in the CNS suggested the catabolism of neuropeptides and peptide hormones present in the brain (Table 1 shows some of POP substrates [25]). According to this hypothesis, increased POP hydrolytic activity would decrease neuropeptide levels, affecting mnemonic and cognitive processes. Consequently, these levels might be restored by inhibiting POP activity. However, this theory was discordant with the spatial association between the extracellular neuropeptide substrates and this intracellular peptidase. In this regard, several *in vivo* studies performed in animal models treated with POP inhibitors did not reach solid conclusions about the effect of POP inhibition on neuropeptide levels (Table 2) [26, 27].

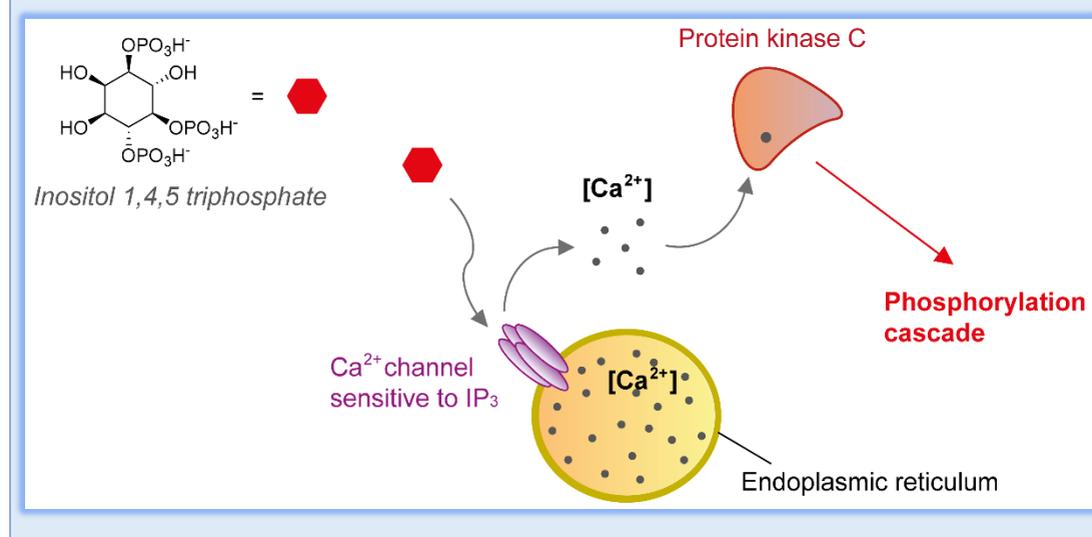
More solid results were found in the involvement of POP in the regulation of the phosphoinositide signalling pathway. A study of Harwood and co-workers showed that lithium-resistant mutants of *Dictyostelium discoideum* lacked the POP gene [28]. Moreover, these mutants presented increased levels of inositol (1, 4, 5)-triphosphate (IP₃), a second messenger involved in the transduction of neuropeptide signalling (Box 1). Increased levels of IP₃ were also observed in the presence of POP inhibitors. Of interest, POP inhibition was found to reverse the effects of lithium, valproic acid and carbamazepine (three mood-

stabilizing compounds based on inositol depletion) on the IP₃ levels, supporting the intervention of POP in the IP₃ signalling pathway [29] (Figure 4 A).

Inhibitor	Animal	Brain region	SP	ACh	AVP	TRH	DA
JTP-4819 S-17092	Rat	Hippocampus and cortex	↑	↑	ND	ND	ND
JTP-4819	Rat	Hippocampus and cortex	=	ND	=	↑	ND
KYP-2047 JTP-4819	Rat	Striatum	ND	=	=	ND	=

Table 2: *In vivo* effects of POP inhibitors on the levels of neuropeptides, peptide hormones and neurotransmitters in specific brain regions. SP = substance P; ACh = acetylcholine; AVP = arginine-vasopressin; TRH = thyrotropin releasing hormone; DA = dopamine.

Box 1: involvement of IP₃ in neuropeptide signal transduction. The presence of IP₃ in the cytosol binds to the corresponding receptors in the endoplasmic reticulum and releases Ca²⁺ from the intracellular stores. The release of Ca²⁺ activates protein kinase cascades, which have direct effects in learning and memory.



Of interest, further experiments performed by Schulz *et al.* in astrogloma cell line U343 were coherent with an indirect involvement of POP in this second messenger pathway [30]. In this study, cell lines with reduced POP expression showed an inverse correlation between POP expression and IP₃ levels. The administration of a POP inhibitor to the WT cells also led to a delayed increase of IP₃ (Figure 4 B). In addition, Brandt and co-workers detected

several endogenous POP substrates in the intracellular space that were related to the phosphoinositide pathway, which might connect POP activity an IP₃ pathway [31]. However, the *in vivo* analysis of IP₃ levels after POP inhibition by KYP-2047 showed no significant effects on this second messenger [32].

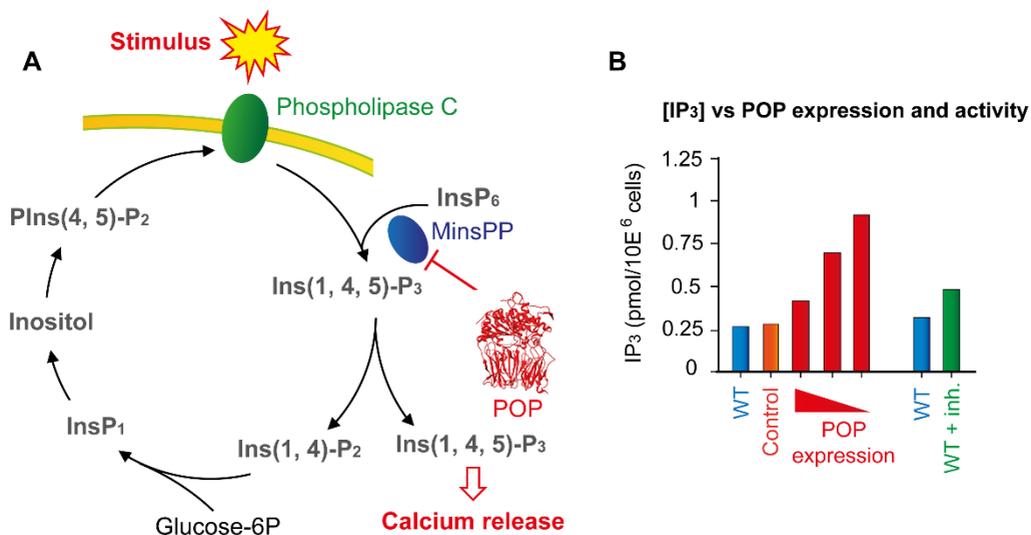


Figure 4: Involvement of POP in the IP₃ pathway. A) Schematic representation of the phosphoinositide cycle. Upon stimulus, phospholipase C hydrolyses phosphatidylinositol (4, 5)-diphosphate (PIns(4, 5)-P₂) to IP₃, which causes an increase of intracellular Ca²⁺ levels. IP₃ can be dephosphorylated to inositol (1, 4)-diphosphate (Ins(1, 4)-P₂), inositol monophosphate (InsP₁), and finally to inositol, which is recycled in this pathway. IP₃ can be incorporated to the cycle by dephosphorylation of high order inositol phosphates (InsP₆). This process is controlled by multiple inositol polyphosphate phosphatase (MinsPP), which can be modulated by POP. B) Effects of POP expression on IP₃ levels in U343 cells. Decreased POP expression causes an increase of IP₃ concentration. A similar effect is observed in the presence of KYP-2047 inhibitor (adapted from [30]).

Together, these evidences indicate that Ca²⁺ signalling by IP₃ might be the main mechanism mediating between POP and neurotransmitter systems. In this regard, several studies related POP with GABAergic neurons in cortex and striatum. Moreover, the expression of POP was raised in hippocampal and cortical glutamatergic neurons, suggesting the participation of POP in this neurotransmission [33]. The administration of two POP inhibitors were found to slightly decrease acetylcholine (ACh) levels in the striatum, suggesting a possible weak effect of POP in cholinergic system [34]. On the contrary, no significant effects were observed in dopaminergic and serotonergic neurons [35].

In another line, Di Daniel *et al.* studied the effects of POP in the mood-stabilizing effects of lithium, valproic acid and carbamazepine [36]. For this purpose, they generated a POP null mutant mice and studied the growth cone response of neurons to these drugs. They observed that POP null-mutant neurons presented alterations in the growth cones, showing larger growth area and less collapsed cones. The WT phenotype of neurons was restored by the transfection with cDNA encoding WT POP as well as the catalytically inactive POP mutant S554A, indicating that the hydrolytic activity was not necessary for the biological function of POP. In the same study, yeast-two-hybrid assays detected growth associated protein-43 (GAP-43) and α -synuclein as possible POP interacting proteins. This work was the first to demonstrate that the hydrolytic activity of POP was not necessary for the biological role in the CNS, indicating that POP probably acts through protein-protein interactions (PPIs).

The interaction of POP with α -synuclein, the major component of Lewi bodies in Parkinson's disease (PD) [37], is potentially relevant for the treatment of this neurodegenerative disease. This topic is discussed in detail in chapter 5 and for this reason it will not be explained here. In turn, the interaction of POP with GAP-43 would partially explain the involvement of POP in cognitive and mnemonic processes. This 25-KDa protein is involved in neuronal growth cone dynamics, axon guidance and Ca^{2+} -dependent synaptic plasticity. Interestingly, high levels of this protein are found in the neocortex and the limbic system, which are involved in mood control [38]. However, *in vitro* studies disclosed only a weak interaction between POP and GAP-43 [39].

In summary, several evidences indicate the involvement of POP in mnemonic and cognitive processes. Probably, these effects are originated by regulation of IP_3 pathway, and proceed through protein-protein interactions. Nevertheless, these mechanisms are not clear, and further studies are required to elucidate the exact role of POP in health and disease.

POP INHIBITORS: DRUG DESIGN AND EFFECTS

As mentioned, changes in POP activity causes alterations in cognitive and memory processes, and might be involved in the pathogenesis of PD. For these reasons, the modulation of POP activity by inhibitors has been a subject of interest during decades; in fact, the study of the *in vivo* POP mechanism have been parallel to the development of POP inhibitors.

The first studies of the effects of POP inhibitors were carried out in impaired animal models. The inhibitor JTP-4819 was tested in scopolamine-treated rats [40], rats with middle cerebral occlusion [41], and naturally aged rodents [27]. In all cases, a slight improvement of spatial memory and cognitive tasks was observed (Table 3). However, pharmacokinetic studies carried out by Jalkanen *et al.* showed that this compound suffered from poor blood-brain-barrier permeability [42]. On the contrary, the same study found that KYP-2047 presented a good brain permeability. Of interest, 5-day treatment of α -synuclein A30P/A53T mutant transgenic mice with this compound reduced the levels of soluble and aggregated α -synuclein in the brain [43]. Moreover, the administration of POP inhibitor S-17092 to monkeys treated with 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (a PD model that causes the loss of dopaminergic neurons) increased the performance of cognitive tasks [44]. These results highlighted the potential of POP inhibitors for the treatment of PD (a more extended explanation for the *in vivo* and *in vitro* effects of POP inhibitors in α -synuclein is found in chapter 5).

Inhibitor	Test	Study	Indication	Result
JTP-4819	Rodents Humans	Phase II	Cognitive disorders and Alzheimer's disease	Improvement in spatial memory Poor BBB permeability
KYP-2047	Rodents	Preclinical	Treatment of cognitive disorders Treatment of PD	Improvement in spatial memory Reduction of α -synuclein levels Increased α -synuclein clearance Good BBB permeability
ONO-1603	Rodents Humans	Phase II	Cognitive disorders	Neuroprotective effects Improvement in memory and learning
S-17092	Rodents Monkeys Humans	Phase I	Cognitive deficits associated with aging	Improvement of cognitive tasks and verbal function No toxicity Good BBB permeability
Y-29794	Rodents Dogs Monkeys	Preclinical	Cognitive disorders and Alzheimer's disease	Good BBB permeability Inhibition of A β deposits formation
Z-321	Rodents	Phase I	Cognitive disorders and Alzheimer's disease	No toxicity Good pharmacokinetics and pharmacodynamics

Table 3: Summary of preclinical and clinical studies carried out with POP inhibitors (adapted from [45]).

Although a plethora of compounds with high POP inhibitory activity and selectivity have been synthesized and patented [22, 46], only few compounds reached clinical trials and

were tested in humans. These inhibitors showed acceptable pharmacodynamics, pharmacokinetics and no toxicity. However, no marked improvements in memory and cognitive tasks were reported. A detailed description of the results of preclinical and clinical studies carried out with all POP inhibitors is out of the scope of this introduction; for a brief summary, see Table 3.

The lack of knowledge about the biological role of POP has been one of the main drawbacks for the development of inhibitors. Moreover, all the animal models used for the *in vivo* studies of POP inhibitors were not suitable to evaluate cognitive improvements [47]; in most cases, brain injured animals were used. In turn, the design of POP inhibitors has been focused mainly on potency using a common peptide pharmacophore. However, the capacity of inhibitors to alter the native conformational landscape of POP has not been exploited to date. This property has a strong potential in the modulation of POP recognition events involved in biological processes.

POP ENZYMATIC MECHANISM AND CONFORMATIONAL DYNAMICS

Briefly, the catalytic cycle of POP starts with the proper orientation of Ser 554 respect to the scissile peptide bond and to His 680 (Figure 5). The nucleophile attack of Ser 554 over the carbonyl group leads to the formation of the tetrahedral intermediate, together with the simultaneous stabilization of the negatively charged oxygen atom by the oxyanion binding site and the acceptance of Ser 554 proton by His 680. Afterwards, the elimination of the amine yields the acyl-enzyme complex and the release of the first cleaved moiety of the peptide. Finally, hydrolysis of the acyl-enzyme complex leads to the release of the hydrolysed peptide and the recovery of the enzyme [2, 10].

Although the chemical step of the cycle is well described by different X-ray structures of POP bound to substrates and inhibitors, the full enzymatic mechanism of POP remains not fully understood. As explained above, all the mammalian X-ray structures of POP and POP bound to inhibitors have been obtained so far in a closed conformation. For this reason, no major interdomain motions are expected to occur in POP. But structural rigidity does not explain how substrates and products circulate between the solution and the buried active site. Hence, several hypothesis have been formulated for the circulation of substrates and products.

The first studies about the elucidation of the enzymatic mechanism of POP were based on activity assays of POP. Polgár showed that the activity of POP presented a doubly sigmoidal pH rate profile, suggesting two different pH-dependent enzyme forms. One of these forms was centred at neutral pH, and the second, at pH value of approximately 5. Moreover, a considerable dependence of the enzyme kinetics with the ionic strength was observed, indicating that charged residues could be involved in the rate-limiting step. Finally, the absence of kinetic isotope effect in POP pointed that a conformational transition was the rate-limiting step of POP at neutral pH [48, 49].

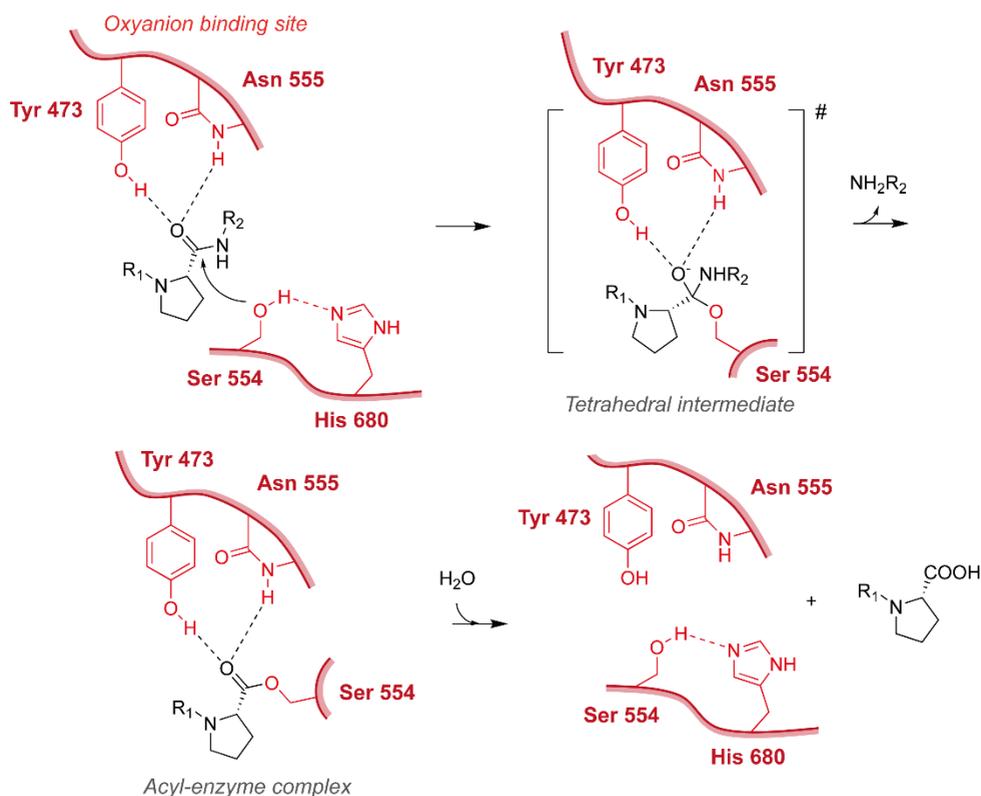


Figure 5: Enzymatic mechanism of POP. Residues involved in the chemical step (Ser 554 and His 680) as well as residues Tyr 473 and Asn 555 of the oxyanion binding site are coloured in red.

Beyond these phenomenological observations, the first X-ray structure of POP obtained by Fülöp and co-workers indicated that the entrance of substrate/inhibitor might occur through the open pore at the bottom of the β -propeller [2]. Afterwards, the same group generated two mutants with reduced β -propeller flexibility through the introduction of non-native disulfide bridges. The first mutant included a disulfide bridge between the first and the last

blades of the β -propeller [50]. In the second one, the α/β -hydrolase and the β -propeller domains were covalently linked by a disulfide bridge facing the hinge region (Figure 6 A) [51]. In both cases, the formation of disulfide bridges rendered inactive enzymes, indicating that oscillations of the β -propeller and the interdomain region were necessary for the catalytic activity.

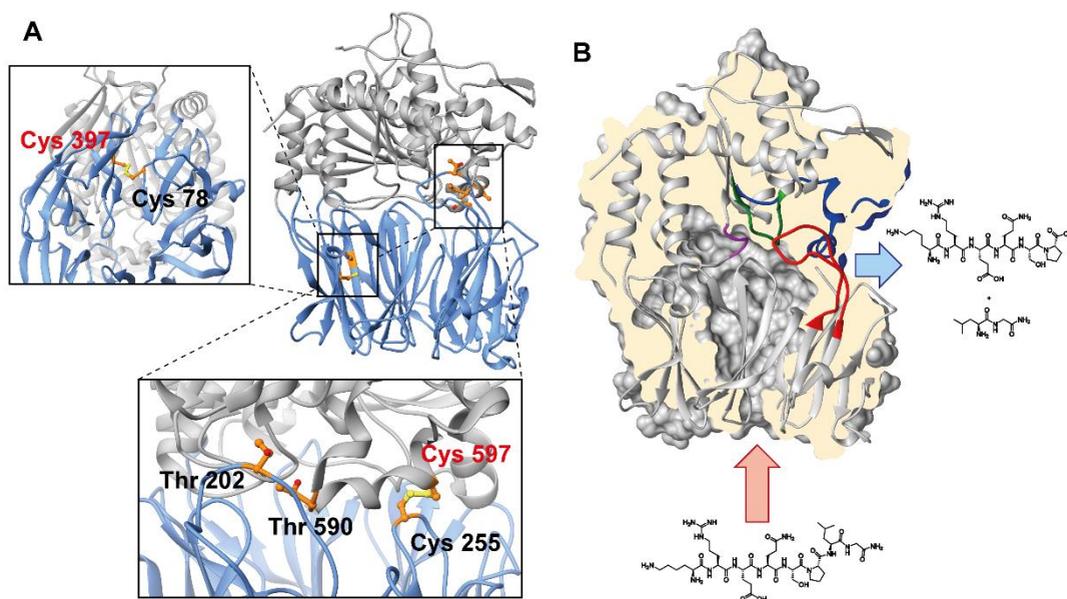


Figure 6: Hypothesis of substrate and products circulation during the enzymatic cycle of POP. A) Three different disulfide bridges were engineered in order to connect the first and the last blades of the β -propeller (Cys 78 – Cys 397), α/β -hydrolase and β -propeller domains (Cys 597 and Cys 255) and α/β -hydrolase with loop A (Cys 202 – Cys 590). In all cases, introducing disulfide bridges resulted in a loss of catalytic activity. Insets show the disulfide bridges observed in the X-ray structures of the POP mutants, except in the case of Cys 202 – Cys 590 (mutated residues are shown in red). B) A possible entrance pathway for the circulation of substrates might occur through the open pore of the β -propeller; hydrolysed fragments would be subsequently released through the flexible loops surrounding the active site (coloured ribbons) [52].

The crystallization of two bacterial POPs in an open conformation [7, 8] demonstrated that interdomain opening occur in the prokaryotic enzyme. In the open conformation, residues of the active site showed an inactive arrangement. Moreover, pre-formed crystals of the open conformation collapsed to the closed conformation after soaking with substrates or ZPP. This long-range conformational transition led to local rearrangements of the residues of the active site, pointing to an induced-fit mechanism [8]. The comparison between all

POP orthologs disclosed that most of the residues involved in interdomain interactions were highly conserved, strongly suggesting that opening and closing mechanism might be common for all prolyl oligopeptidases from prokaryotes and eukaryotes. In this regard, a study of *Trypanosoma brucei* oligopeptidase B found a similar opening and closing mechanism, in which the complete closing of the enzyme was necessary for the competent arrangement of the active site. Of interest, most of the key residues of *Trypanosoma brucei* oligopeptidase B were conserved in mammalian POP, suggesting that both enzymes might share a common mechanism [53].

A second research line pointed towards the involvement of several flexible loops surrounding the active site in a gating mechanism. A combination of MD simulations and crystallographic studies of POP mutants showed that loops A (residues 189-209), B (residues 577-608), C (residues 636-646) and His 680 loop (residues 677-681) were highly flexible [11]. Motions in these areas would favour the recruitment of the substrates, especially the detachment of loop A. In the presence of inhibitors, the interactions between loops A and B strengthen and impede catalysis. Trypsin cleavage assays supported the high flexibility of loop A in free but not inhibitor-bound POP. Further activity assays performed by Szeltner and co-workers found that the artificial introduction of a disulfide bridge between loops A and B by mutagenesis resulted in a decreased catalytic activity, especially in the case of larger substrates (Figure 6 A) [54]. Moreover, dynamic light experiments reflected the structural heterogeneity of POP in solution, which was reduced in the presence of inhibitors [54]. Umbrella sampling simulations indicated that the loop region was the most energetically favourable exit pathway of ZPP [55]; this was further supported by a computational study carried out in our group using protein energy landscape exploration method [56]. These simulations suggested that the entrance of inhibitors in the closed conformation would occur through the open pore at the bottom of the β -propeller domain, while the loop region would be involved in the release of cleaved products (Figure 6 B) [52].

Previous studies in our laboratory pointed towards a conformational sampling mechanism of free POP in solution. The first evidence of a pre-existing conformational equilibrium of POP was obtained from native polyacrylamide gel electrophoresis (PAGE) experiments. According to this experiment, three bands of free POP were observed, which would correspond to three conformers with different electrophoretic properties; upon inhibitor binding, these bands collapsed to a single one. In the same paper, low resolution structures of POP were obtained by electron microscopy. Interestingly, free POP showed a new side cavity (Figure 7 A). Moreover, in the presence of a large inhibitor, the opening at the bottom

of the β -propeller collapsed [57]. Together, these results indicated significant structural changes towards more compact structures in the presence of covalently bound inhibitors. However, the first solid evidences about the highly dynamic nature of POP in solution were obtained by Dr. Nessim Kichik in our group [58]. Robust ^{15}N relaxation NMR experiments of U- ^{15}N labeled POP demonstrated that free POP undergo pervasive μs -ms dynamics; of interest, the presence of inhibitors decreased the amplitude of these motions. In the same paper, relaxation dispersion (RD) experiments of [indole- ^{15}N]-Trp labeled POP [59] showed that the residues experiencing higher μs -ms exchange were located in the interdomain region (Figure 7 B). These results pointed that free POP exists in a conformational sampling in solution, resulting from long-range conformational transitions between the two domains.

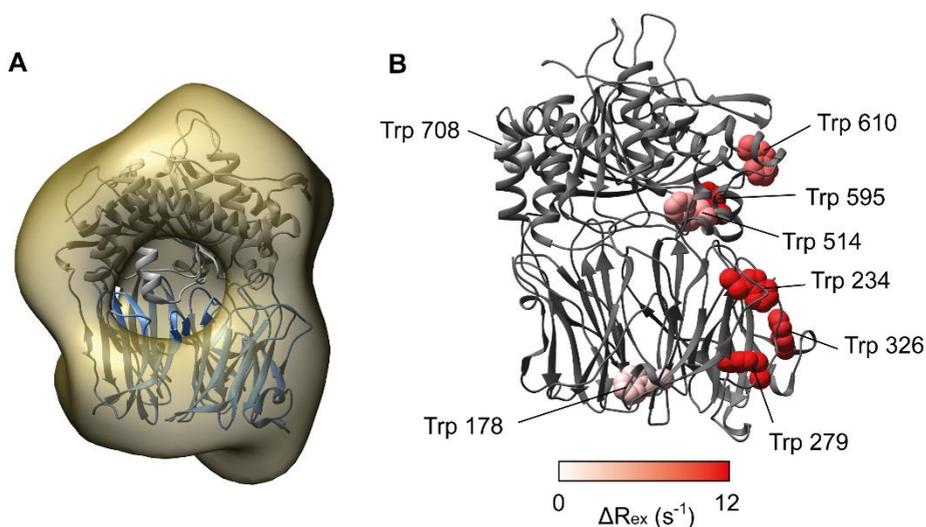


Figure 7: Previous studies of POP structure and dynamics carried out in our group. A) Large side opening observed in free POP by electron microscopy experiments. B) Conformational exchange of Trp residues in the μs -ms time scale obtained by RD experiments of [indole- ^{15}N]-labeled POP. Trp residues are coloured in POP structure according to their ΔR_{ex} value.

In summary, two main theories about the enzymatic mechanism of mammalian POP are considered nowadays. According to the first theory, POP is present exclusively in a closed active conformation with several flexible loops acting in a gating mechanism. The second theory proposes that POP can exist in an open conformation. Nevertheless, in this case it is not clear if the conformational shift is induced by substrates or inhibitors (induced fit mechanism), or if it arises from a conformational sampling. To date, any consensus regarding these theories has not been reached. It is evident that solid studies using robust

biophysical techniques are required to elucidate the structural heterogeneity of POP in solution.

Hence, this PhD thesis pretends to gain a deeper knowledge of the structure and dynamics of human POP in solution. Specifically, this PhD thesis has the following objectives.

OBJECTIVES

1. The main objective is to **use a combination of robust NMR and SAXS experiments, supported by MD simulations, in order to elucidate the conformational dynamics of human POP in solution.** The structure of POP has been studied mainly by crystallography to date. However, only vague hints of POP dynamics have been obtained by indirect observations. For this reason, this PhD thesis aims to provide a breakthrough on the description of POP dynamics. To achieve this goal, we will rely on biophysical methods that contemplate time dimension and describes structures as ensembles of conformations.
 - 1.1. Moreover, we will evaluate the application of ion mobility mass spectrometry (IMMS) for the study of large dynamic proteins. IMMS is an emerging biophysical tool in structural biology, which has been successfully applied in the case of protein complexes and disordered proteins. Therefore, we will use POP as a model of large dynamic protein to be studied by IMMS.
2. **To study the effects of active site-directed inhibitors on the conformational landscape of POP** using the same experimental approach described at point 1. Although POP inhibitors are of high therapeutic interest, there is a lack of knowledge of the effects of these compounds on POP structure and dynamics.
3. **To characterize the possible interaction of POP with α -synuclein by NMR experiments.** It has been recently described that POP inhibitors increase the clearance of α -synuclein *in vivo*. Probably, the involvement of POP in this neurodegenerative disease is originated by a direct interaction with α -synuclein, but this interaction has not been characterized by robust biophysical methods.

Chapter 1: study of POP by nuclear magnetic resonance

INTRODUCTION: NMR LABELING STRATEGIES FOR LARGE PROTEINS AND PROTEIN COMPLEXES

NMR is one of the most powerful biophysical techniques for the elucidation of protein structure, dynamics and biomolecular interactions. Nevertheless, the sensitivity of NMR experiments is compromised by the relaxation properties of nuclear spins. The increased T_2 of large systems caused by the slower correlation times leads to dramatic line broadening. Thus, NMR has been typically limited to biomolecules not exceeding 40 kDa, thus avoiding the study of larger proteins or complexes. Moreover, a second limitation is the significant signal overlapping as a result of the increased number of residues on high MW proteins. The development of transverse-relaxation optimized spectroscopy experiments (TROSY, described later in this chapter) and the use of labeling strategies successfully contributed to overcome these limitations, respectively. These approaches have expanded the frontiers in terms of size and complexity of biomolecular NMR.

Protein labeling strategies have benefitted from the advances in molecular biology applied to the protein expression. Apart from the well-established ^{15}N and ^{13}C uniform labeling of protein backbone (which will not be explained here) other procedures have been developed to selectively label the side chains of certain residues. Two different approaches can be used for the isotopic labeling of protein side chains. The simplest procedure involves the expression of the protein in minimal media supplemented with the labeled amino acid, using the corresponding auxotrophic cells. However, a second procedure for the expression of labeled proteins is based on the supplementation of minimal media with labeled amino acid precursors.

The first approach relies on auxotrophic strains. In these strains, the biosynthetic pathway of a certain amino acid is knocked out. Hence, the capacity of the cell to synthesize the corresponding amino acid is disabled. If the corresponding amino acid (or, in this case, analogues or isotopomers) is supplemented, it can be incorporated to cell metabolism by selective pressure. Although the use of auxotrophic strains is well established (especially in crystallography), in some cases the expression takes place in poor yields. The disruption of

amino acid biosynthetic pathways can be detrimental to cell metabolism, which leads to decreased bacterial growth or protein expression.

In contrast, the approach involving labeled amino acid precursors is less aggressive. Specifically labeled precursors are incorporated to cell metabolism and finally transformed to the corresponding amino acids according to amino acidic biosynthetic pathways. Nevertheless, this approach totally depends on the biosynthetic pathway, and for this reason, it is not always useful for all labeling strategies. For instance, isoleucine-leucine-valine (ILV) labeling has been successfully achieved by supplementation of bacterial media with ^{13}C -labeled α -ketobutyrate and α -ketoisovalerate [60, 61], or [indole- ^{15}N]-Trp labeling has been carried out by supplementation with ^{15}N -indole [58, 59]. Additionally, both approaches accept the possibility of perdeuteration if combined with perdeuteration protocols. Replacement of non-labile ^1H by ^2H is highly relevant because it result in significant gains in sensitivity in the NMR spectrum [62] (as it will be explained later in this chapter).

Methyl groups in biomolecular NMR

Methyl groups have excellent NMR properties: on the one hand, the three equivalent fast rotating protons gives rise to a single, intense signal; on the other, the relaxation properties of methyl group are particularly useful for the application of the TROSY principle [62]. The substitution of ^1H by ^{19}F in methyl groups is also possible in the case of biomolecules, and allows the acquisition of ^{19}F NMR experiments.

In the particular case of the methyl group of Met, additional relevant features must be considered. First, Met is present in the buried areas of the protein core; for this reason, this residue is a good reporter of conformational changes and dynamics of the protein. As well, Met is often localized in surface hydrophobic patches, which are potential recognition areas. Moreover, the thioether bond result in a particular chemical shift of the methyl group of Met signals, which appear in a particular region of the ^1H - ^{13}C 2D heterocorrelation spectra. Hence, Met signals are segregated and not overlapped with the other methyl signals of the rest of residues (ILV signals). Finally, the low natural abundance of Met in protein sequence minimizes overlapping between Met signals.

As said before, previous studies in our group obtained successful results with the [indole- ^{15}N]-Trp labeling of POP [59]. Moreover, ^{15}N RD experiments of U-[^{15}N]-POP detected pervasive μs -ms dynamics of POP backbone, which was decreased in the presence of

inhibitors. RD measurements of [indole-¹⁵N]-Trp labeled POP showed significant dynamics localized in the interdomain space [58]. The NMR project of POP continued in this thesis using Met labeling strategies. Here we exploited different types of Met labeling, which allowed us to perform several NMR experiments. Given the excellent yields and the reliability of protein expression using *E.coli* Met auxotrophic cells, we used exclusively Met auxotrophic cells for the expression of labeled POP. Our objective was to perform two completely different Met labeling of POP: *L*-trifluoromethionine (*L*-TFM) and [methyl-¹³C]-Met labeling. The first approach would open up the opportunity to perform monodimensional ¹⁹F NMR spectroscopy of POP, with a direct application in the screening of compounds with POP inhibitory activity. The second labeling would provide a robust framework to study the ¹H-¹³C 2D heterocorrelation spectra of POP, which would allow us to quantify POP conformational dynamics by relaxation dispersion measurements.

TRIFLUOROMETHIONINE POP LABELING

Use of ¹⁹F as NMR probe

Fluorine have been extensively used in protein engineering during decades. Although fluorine is the most electronegative element, the moderate polarity of the aliphatic C-F bond avoids the formation of non-native hydrogen bonds between C-F and H-X pairs (in which X is a heteroatom). Fluorine is almost isosteric with hydrogen (covalent radius of 1.35 Å with respect to 1.2 Å of the hydrogen [63], Figure 8 A), which minimizes steric hindrance in protein tertiary structure. Therefore, monofluorinated non-canonical amino acid are not expected to cause significant alterations of protein structure.

In turn, ¹⁹F has good NMR properties, which have been detailed in several reviews elsewhere (e.g. [64]). This isotope has a ½ spin and 100 % natural abundance, and an 83 % sensitivity compared to ¹H. The interval of chemical shifts is near 100 times greater than ¹H, which ensures good signal dispersion in a simple 1D spectra. Another factor favouring the signal dispersion is the presence of lone-pair electrons in ¹⁹F. These lone-pairs highly contribute to the nuclear shielding, and are extremely sensitive to the surrounding steric and electronic properties. Hence, ¹⁹F chemical shift reflects subtle changes around the Van der Waals environment of this atom. Obviously, this element is not present in natural proteins. This characteristic ensures that observed signals are specific to the fluorinated residues. Hence, the versatility of ¹⁹F monodimensional spectra resides in the simplicity and the acquisition time, which is much faster than 2D experiments and requires less sample

concentration. However, contrarily to what happens in the case of ^1H , ^{13}C and ^{15}N , ^{19}F NMR signals suffer from line broadening at higher field strengths, due to the anisotropy of the chemical shift tensor. Typically, protein ^{19}F NMR spectra are recorded between 400 – 600 MHz in order to avoid excessive signal broadening.

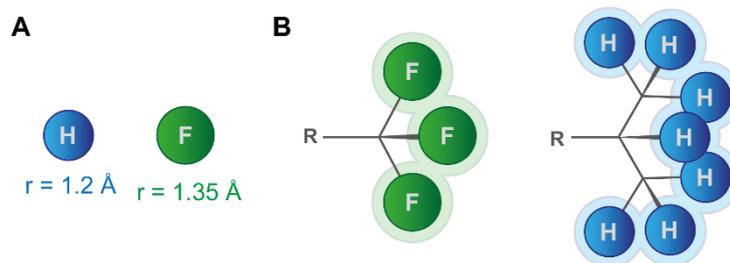


Figure 8: Steric properties of fluorine. A) Comparison between the Van der Waals radii of hydrogen and fluorine. B) Trifluoromethyl group is almost isosteric to isopropyl group.

These good characteristics have been benefited in several NMR studies of proteins incorporating monofluorinated amino acids. For instance, the use of 5-fluoro-tryptophan and 3-fluoro-phenylalanine for the labeling of galactose-binding protein (a protein involved in chemosensing), disclosed the conformational changes induced by the binding of galactose, which extended from the binding cleft to other secondary structural elements of the protein. These conformational changes have been remained elusive by other biophysical techniques, and confirmed the utility of ^{19}F NMR in the study of protein structure [65]. Another case studied the in-cell NMR spectra of six proteins (α -synuclein, ubiquitin, chymotrypsin inhibitor-2, calmodulin, green fluorescent protein and histidinol dehydrogenase) labeled with 3-fluoro-tyrosine or with trifluoromethyl-phenylalanine, which suggested that this approach might be valid to monitor proteins in living organisms [66].

However, the small differences in the atomic radii and the hydrophobicity of fluorine compared to ^1H are intensified in the case of incorporation of polyfluorinated groups. This might represent a limitation in the case of polyfluorinated non-canonical amino acids. On the one hand, the higher hydrophobic contribution of polyfluorinated groups would affect protein stability. On the other, the steric impediment of the slightly bulky polyfluorinated groups would difficult its accommodation into the protein tertiary structure; for instance, trifluoromethyl group is much similar in size to an isopropyl group rather than a methyl group [67] (Figure 8 B). Finally, a practical limitation arise due to the high toxicity of polyfluorinated amino acids to bacteria, which reduces the yield of the protein expression.

Trifluoromethyl group is affected by the limitations described above. Hence, the use of trifluoromethylated amino acid analogues are expected to significantly decrease protein yields, and to result in partial incorporation of the labeled residue [68]. Nevertheless, some successful studies have been published in the particular case of *L*-TFM. For instance, the expression of bacteriophage λ lysozyme (18 KDa, three Met residues) incorporating a 70 % of *L*-TFM was achieved by Duewel and co-workers [69]. More recently, the expression of a 63 KDa N-terminal truncated DNA I polymerase of *Thermus aquaticus* with 82 % of the 14 Met residues substituted by *L*-TFM expanded the frontier of proteins labeled with trifluoromethylated amino acid analogues. Of interest, this DNase retained the enzymatic activity, and some of the 9 observed signals of the ^{19}F NMR spectrum showed significant changes in their chemical shifts if the DNA template or the substrate 2',3'-deoxycytidine-5'-triphosphate was present [70].

Although the possible limitations of the yields of the expression and the stability of trifluoromethyl-containing proteins, we decided to perform the labeling of POP with *L*-TFM. The suitable NMR properties of trifluoromethyl group, together with our expertise in [methyl- ^{13}C]-Met labeling, encouraged us to take this approach. Moreover, the alternative POP labeling based on ^{19}F would represent an important breakthrough for the screening of compounds with POP inhibitory activity: the simplicity and fast acquisition of 1D spectra would facilitate the detection of possible conformational changes caused by molecules altering POP enzymatic activity.

Synthesis of trifluoromethionine and expression of trifluoromethionine labeled POP

The commercial availability of *L*-TFM was expensive and only limited to few companies upon request. This prompted us to synthesize this compound. Most of the synthetic procedures to obtain trifluoromethyl thioethers published up to the 90s proceed by radical mechanisms involving harsh conditions and poor to moderate yields (i.e. see the work of Duewel and co-workers [69]). According to recent literature, two alternative synthetic approaches with improved the synthetic yields have been described (Figure 9 A and B): A) nucleophilic trifluoromethylation of disulfides, and B) electrophilic trifluoromethylation of free thiols.

Nucleophilic trifluoromethylation of disulfides was developed by Ruppert and Prakash [71] with the use of trifluoromethylated organosilicon compounds (typically (trifluoromethyl) trimethylsilane, which is known as "Rupert's reagent" [72]). However, half of the disulfide is wasted in this procedure (Figure 9 A). This approach was improved by Medebielle and

Dolbier [73, 74]. In fact, the improved procedure was constituted by two tandem reactions: the first step involves the generation of CF_3^- anion by reduction of CF_3I by *tetrakis*-(dimethylamino)ethylene (TDAE, Figure 10 A). The trifluoromethyl anion attacks the disulfide by a radical nucleophilic substitution ($\text{S}_{\text{RN}}1$), generating the first equivalent of trifluoromethyl thioether. The second half of the disulfide reacts with the excess of CF_3I , yielding the second equivalent of trifluoromethyl thioether [73]. Although this reaction is especially suitable for aryl disulfides, it is also described for alkyl disulfides [74]. However, our attempts to react protected homocystine (Figure 10 B) failed systematically. For that reason, we moved to the electrophilic trifluoromethylation approach.

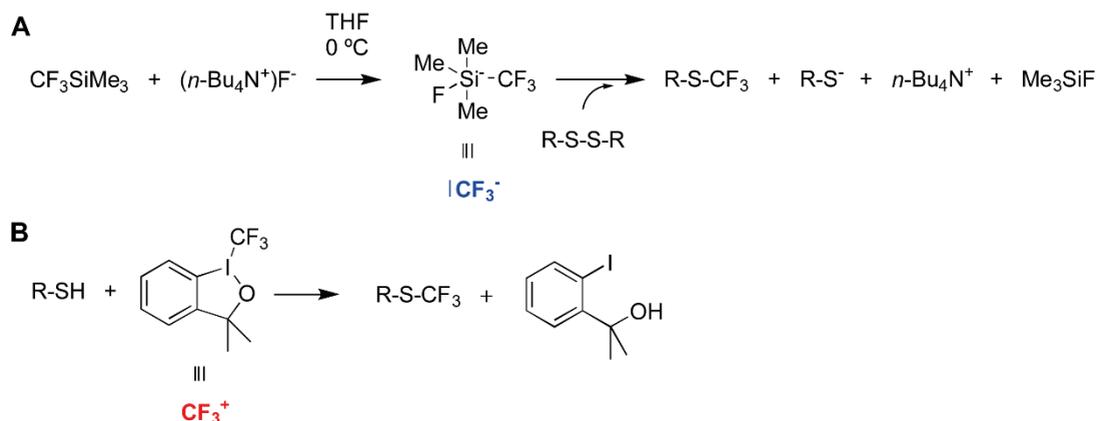


Figure 9: Synthetic approaches to trifluoromethyl thioethers. A) Nucleophilic trifluoromethylation of disulfides by the use of Ruppert's reagent (CF_3^- synthon is shown in blue); note that half of the disulfide is wasted as thiolate, and consequently, synthetic yields will never exceed 50 %. B) Electrophilic trifluoromethylation of free thiols by the use of Togni's reagent (CF_3^+ synthon is in red).

Regarding the electrophilic trifluoromethylation, this approach became possible with the development of hypervalent iodine (III) benzenodioxoles with a trifluoromethyl group attached to the iodine atom by Togni and co-workers [75]. These reagents (specifically, 3,3-dimethyl-1-(trifluoromethyl)-1,2-benziodoxole is commercially known as "Togni's reagent") can transfer a trifluoromethyl group to soft nucleophiles such as free thiols, among others (Figure 9 B); formally, it can be considered that the trifluoromethyl group behaves here as an *umpolung* CF_3^+ synthon compared to the CF_3^- synthon of nucleophilic trifluoromethylation. Several features are favourable to the particular case of *L*-TFM synthesis: first, the high specificity of this reagent allows the presence of several unprotected functional groups; second, the simple and soft experimental conditions would avoid racemization; finally, all the reagents are commercially available. Moreover, several

examples of trifluoromethylation of mono- and di-protected Cys were described [76], as well as trifluoromethylation of Cys residues in α and β peptides [77].

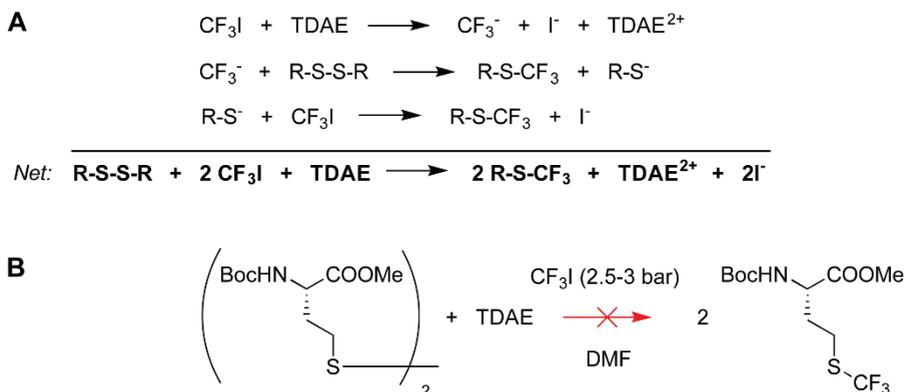


Figure 10: Nucleophilic trifluoromethylation of alkyl disulfides. A) Mechanism of the reaction, as described by Medebielle *et al.* [74] (R = alkyl). B) The reaction with *N*-Boc homocysteine methyl ester did not take place.

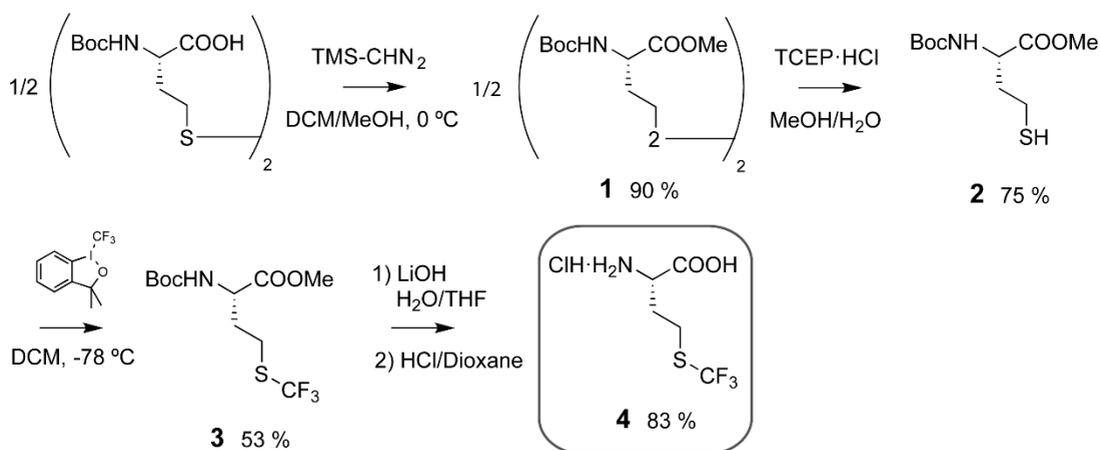


Figure 11: Synthetic scheme for the preparation of *L*-TFM by electrophilic trifluoromethylation of fully protected homocysteine.

The synthesis was successfully performed from the commercially available Di-Boc-*L*-homocysteine (Figure 11). Although it is described that carboxylic groups did not affect electrophilic trifluoromethylation, we protected the carboxyl group as methyl ester **1** to avoid side reactions and to improve the solubility in organic solvents. Afterwards, the disulfide was reduced with *tris*(2-carboxyethyl)phosphine hydrochloride (TCEP·HCl). The resulting protected homocysteine **2** was reacted directly with Togni's reagent in DCM at -

78 °C; trifluoromethylation took place with moderate yield (53 %). The protected *L*-TFM **3** was deprotected by conventional procedures to yield *L*-TFM hydrochloride **4**. The total synthesis involved 4 steps, and took place in an overall yield of 30 %. For experimental details, see Materials and methods.

The expression of *L*-TFM labeled POP failed systematically in our several attempts to grow the B834 *E. coli* auxotrophic cells in minimal media supplemented only with 1 mM *L*-TFM. In order to avoid the toxic effects of *L*-TFM on bacterial growth [69], we also performed the separate growth of bacteria in LB media, followed by harvesting and resuspension of the pelleted living bacteria in minimal media supplemented with *L*-TFM. Nevertheless, this procedure also resulted in a residual yield of POP expression, probably due to the toxicity of high concentrations of *L*-TFM. In order to overcome this serious limitation, we analysed the toxicity of *L*-TFM at different conditions, especially at different *L*-TFM/Met ratios (Figure 12) [68]. Although sources of natural amino acids should be avoided in the expression of proteins containing non-canonical amino acids, in this case a compromise between the toxic *L*-TFM and the natural amino acid were beneficial. As it can be seen, 1 mM *L*-TFM/0.01 mM Met was a good balance between acceptable bacterial growth and high *L*-TFM/Met ratio.

<i>L</i> -Met (mM)	0.1	0	0	0	0	0.01	0.01	0.01	0.05	0.05	0.05
<i>L</i> -TFM (mM)	0	0	0.1	0.5	1	0.1	0.5	1	0.5	1	2
Ratio [TFM]/[Met]	0	Blanc	TFM	TFM	TFM	10	50	100	10	20	40
<i>A</i> ₆₀₀ (3 h)	0.079	0.077	0.076	0.074	0.075	0.081	0.081	0.085	0.084	0.086	0.078
<i>A</i> ₆₀₀ (6 h)	0.123	0.085	0.085	0.088	0.089	0.132	0.109	0.1	0.12	0.119	0.106
<i>A</i> ₆₀₀ (21 h)	1.255	0.085	0.08	0.09	0.095	0.193	0.201	0.242	0.838	0.887	0.773

1:100 dilution

<i>L</i> -Met (mM)	0.1	0	0	0	0.001	0.001	0.005	0.005	0.01	0.01
<i>L</i> -TFM (mM)	0	0	0.5	1	0.5	1	0.5	1	0.5	1
Ratio [TFM]/[Met]	0	Blanc	TFM	TFM	500	1000	100	200	50	100
<i>A</i> ₆₀₀ (4 h)	0.073	0.071	0.069	0.071	0.071	0.068	0.069	0.075	0.072	0.074
<i>A</i> ₆₀₀ (7 h)	0.095	0.074	0.069	0.07	0.081	0.072	0.077	0.079	0.081	0.078
<i>A</i> ₆₀₀ (24 h)	1.224	0.074	0.071	0.071	0.084	0.082	0.172	0.177	0.278	0.274

1:100 dilution

<i>L</i> -Met (mM)	0.1	0	0.005	0.01	0.01
<i>L</i> -TFM (mM)	0	0	1	0.5	1
Ratio [TFM]/[Met]	0	Blanc	200	50	100
<i>A</i> ₆₀₀ (6 h)	0.071	0.067	0.071	0.066	0.071
<i>A</i> ₆₀₀ (21 h)	1.29	0.067	0.17	0.277	0.243

Figure 12: Optimization of the expression conditions for *L*-TFM labeled POP. The toxicity of *L*-TFM was evaluated by growing *E. coli* bacteria in normal minimal media with different proportions of *L*-TFM and Met. Three successive expressions were performed from 1:100 dilutions from the previous culture, in order to detect possible adaptation of *E. coli* to *L*-TFM. No adaptation was observed.

However, several scale-up trials showed that these conditions also hampered cell growth and resulted in poor expression yields. For this reason, culture growth was performed in a separate flask containing minimal media with the minimal amount possible of Met (0.055 mM); in order to improve the bacterial growth, the media was also supplemented with the rest of 19 amino acids at a concentration of 50 mg/l [70]. After growth, *E.coli* cells were harvested, washed and resuspended in the expression media, containing 0.5 mM *L*-TFM/0.025 mM Met. Hence, the final scaled culture was performed in a 0.5 l batch, resulting in a final yield of 3 mg of protein/liter of culture. Based on MS analysis, the final POP labeled with *L*-TFM was a heterogeneous mixture of isomers incorporating an average of 6 *L*-TFM molecules randomly distributed in a total of 12 Met positions (Figure 13). Moreover, preliminary activity assays demonstrated that *L*-TFM partially labeled POP retained enzymatic activity (not shown).

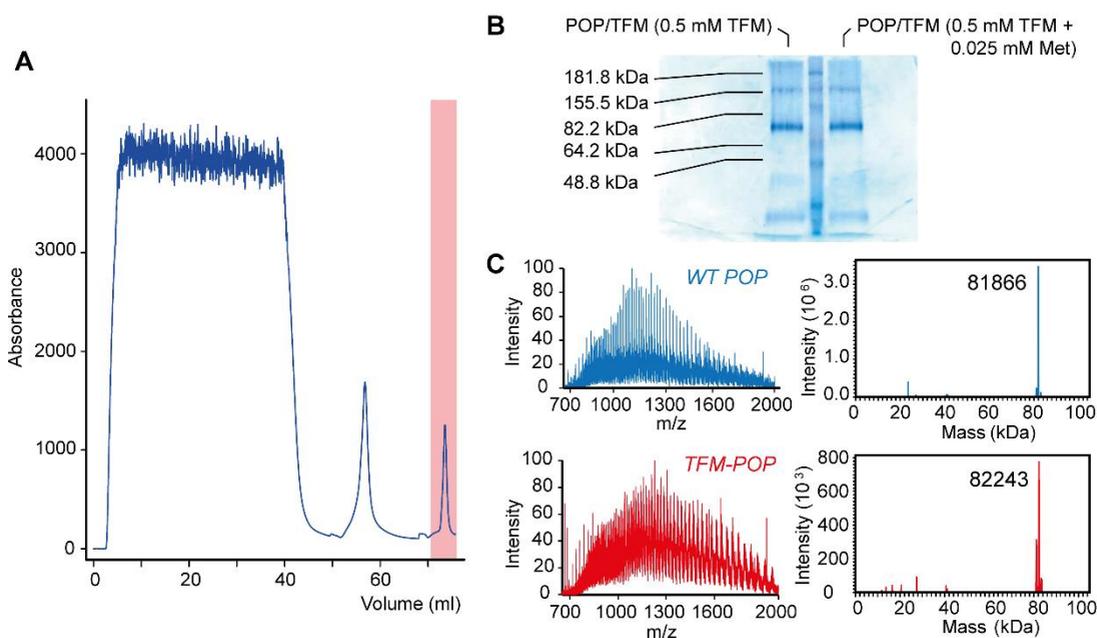


Figure 13: Expression of *L*-TFM labeled POP. A) FPLC elution profile from the HisTrap Ni²⁺ ion affinity column. The area highlighted in red correspond to the elution of *L*-TFM labeled POP by 500 mM imidazole buffer. B) SDS-PAGE of purified protein, showing the major band at 80 KDa. Two samples are shown, one obtained by supplementing the expression media only with 0.5 mM *L*-TFM, and another with 0.5 mM *L*-TFM + 0.025 mM *L*-Met. C) Mass spectra and deconvoluted masses of WT POP (blue) and POP labeled with *L*-TFM (red). The global mass increase of 377 Da correspond to an average of 6 Met substitutions.

Preliminary ^{19}F NMR spectra of trifluoromethionine labeled POP

Monodimensional ^{19}F NMR experiments of *L*-TFM labeled POP were carried out at 400 and 600 MHz, at 35 °C. In both cases, the same sample was used, and it consisted in 31 μM of *L*-TFM labeled POP in Tris-HCl 50 mM pH = 8 buffer with 10 % D_2O . The carrier was placed at -40 ppm, and the spectral window was of 40 ppm. *L*-TFM labeled POP spectra were referenced to the TFA signal of a spectra of aqueous TFA recorded immediately before. NMR acquisition parameters at 400 and 600 MHz are summarized in Table 4. Spectra were processed with MestreC and Topspin softwares for experiments recorded at Varian and Bruker spectrometers, respectively. In all cases, a line broadening of 5 Hz was used, and spectra were scaled according to their S/N ratio.

	Varian 400 MHz	Bruker 600 MHz
Probe	Conventional	Cryoprobe
Acquisition time (s)	1.00	1.45
Interscan delay (s)	0.1	0.1
90° pulse (μs)	7.0	14.3
Number of scans	12000	10000

Table 4: Acquisition parameters for monodimensional ^{19}F spectrum of *L*-TFM labeled POP at 400 and 600 MHz fields.

Figure 14 A shows the ^{19}F spectra of *L*-TFM labeled free POP recorded at 600 MHz showed two signal regions, centred at -41.59 ppm and at -41.86 ppm respectively, probably formed by broadened and overlapped signals. Additionally, three possible peaks were speculated at -40.68 ppm, -40.80 ppm and -41.16 ppm. In order to detect possible structural and dynamic changes in POP, we performed the ^{19}F NMR spectrum at 600 MHz in the presence of 10 equivalents of the canonical active site-directed inhibitor ZPP (this experiment was carried out 7 days after that of free POP). The monodimensional ^{19}F spectra of *L*-TFM labeled POP bound with ZPP disclosed significant changes: first, the broadened region present at -41.86 ppm in free POP was shifted to -42.00 ppm, while the broadened region at -41.59 ppm was maintained; second, a sharp and intense signal at -42.06 ppm appeared, together with a minor sharp signal at -42.00 ppm. The high intensity and the low peak width at half height of the most intense signal (34 Hz) at -41.06 ppm suggested that this residue was in a poorly structured environment; probably, it corresponded to Met 1 placed in the loose N-terminal region.

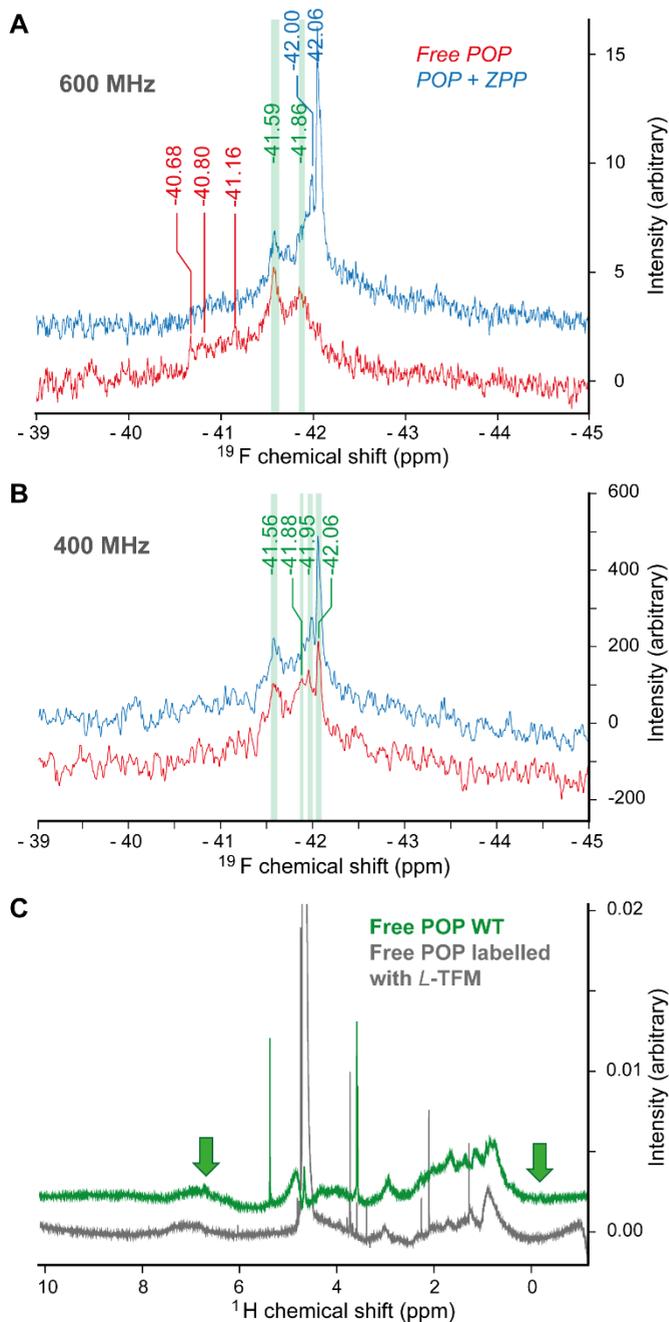


Figure 14: ¹⁹F and ¹H NMR spectra of *L*-TFM labeled POP. Red and blue spectra correspond to free POP and to ZPP-bound POP, respectively. Signals are labeled with the same colour code; in turn, signals labeled in green are common for free and ZPP-bound POP. A) Spectra recorded at 600 MHz. B) Spectra recorded at 400 MHz. C) ¹H spectrum of free POP WT (non-labeled, green) overlaid with the spectrum of free POP labeled with *L*-TFM (gray). Green arrows show the peaks upfield 0 ppm characteristic of folded proteins as well as the aromatic region of proteins. Note that both regions are not completely defined in the case of POP labeled with *L*-TFM.

¹⁹F NMR experiments of *L*-TFM labeled POP at 400 MHz were highly similar between the free and inhibitor-bound forms (the spectra of both forms were recorded 7 days after that of free POP at 600 MHz), and also with the spectra of POP bound to ZPP recorded at 600 MHz (Figure 14 B). In the spectra at 400 MHz, both forms showed peaks at -41.56, -41.88, -

41.95 and -42.06 ppm (these chemical shifts corresponded to free POP resonances). Only small changes occurred between the populations of peaks -41.88 and -41.95 ppm. No significant effects in the resolution were observed between the experiments performed at 400 and 600 MHz, indicating that both fields are valid for these experiments. However, the use of the more sensitive cryoprobe of the 600 MHz spectrometer is much recommended. Finally, ^1H spectra with water pre-saturation were acquired in order to check the folding and the stability of the labeled protein. Unfortunately, the ^1H spectra of POP recorded after 7 days showed a non-optimal folding (Figure 14 C).

Discussion

The expression of proteins including polyfluorinated non-canonical amino acids involves certain risks, which might result in the total failure of the protein expression. As mentioned in the introduction of this section, the incorporation of non-canonical polyfluorinated amino acids with different steric and electronic properties into proteins might result in dramatic alterations on their structural integrity. Moreover, some of these polyfluorinated compounds might be toxic to the bacteria used in the protein expression, therefore limiting bacterial growth.

In our case, successive attempts to express fully substituted *L*-TFM labeled POP failed. Probably, a combination between disfavoured bacterial growth due to bacterial toxicity of *L*-TFM, together with the difficulty to accommodate the relatively bulky *L*-TFM in POP structure, was the cause of this lack of success. In order to overcome these limitations, bacterial growth had to be carried out in minimal media supplemented with a minimal amount of Met; in turn, expression was carried out in a different media containing low concentrations of *L*-TFM and a reduced concentration of natural Met. With this protocol, only a heterogeneous mixture of POP species with random incorporation of *L*-TFM was obtained in a moderate yield. However, this partially labeled protein retained the native tertiary structure necessary for the enzymatic activity, but the stability of the native structure was limited to few days.

Regarding the NMR experiments, the poor yields of the expression of *L*-TFM labeled POP limited the final concentration of the NMR sample (31 μM) and severely compromised the intensity of the signals. Moreover, the partial incorporation of *L*-TFM resulted in a heterogeneous labeling, which contributed to additional line broadening. Together, these major drawbacks were probably the causes for the lack of success in recording monodimensional ^{19}F NMR spectra of *L*-TFM labeled POP. In spite of these inconveniences,

several signals of the total 12 Met residues were observed in the spectra of *L*-TFM labeled free POP recorded at 600 MHz. Some of these signals probably corresponded to broadened and overlapped signals. In the presence of ZPP inhibitor the overall shape of the spectra was maintained, but a sharp and intense signal appeared at -41.06 ppm, together with a minor sharp signal at -42.00 ppm. The low peak width at half height of the signal at -41.06 ppm indicates that it corresponds to an unstructured residue. However, the sample stability is a major concern for the quality of the experiments (Figure 14 C). The ^{19}F spectra of ZPP-bound POP was recorded 7 days after that of free POP, and the quality of the sample was not optimal.

In summary, limitations of the expression avoided the preparation of the protein necessary for good NMR experiments; moreover, only a partial and heterogeneous *L*-TFM substitution was possible. The ^{19}F NMR spectra of *L*-TFM labeled POP at 600 MHz only disclosed some signals and groups of signals, but some differences between the spectra of the free and ZPP-bound POP were observed. In our opinion, the lack of success in the ^{19}F NMR of *L*-TFM labeled POP arises from methodological limitations, rather than from the NMR technique. For this reason, we suggest that another fluorine POP labeling involving less toxic and bulky monofluorinated amino acid analogues might be successful for this protein (e.g. 5-fluorotryptophan). The decrease on the signal intensity due to the minor intensity of a single ^{19}F nucleus compared to the trifluoromethyl group would be further compensated by the much higher and homogeneous incorporation of the monofluorinated label, by the increased concentration of the sample, and by the higher protein stability.

[METHYL- ^{13}C]-METHIONINE POP LABELING: ^1H - ^{13}C METHYL-TROSY HMQC SPECTRA

TROSY principle has been one of the most important milestones in biomolecular NMR [78]. The application of this principle pushed on the size boundaries of systems studied by NMR, providing important information about protein structure and dynamics of large systems at the atomic level [79]. Moreover, the spread of isotope labeling protocols of proteins contributed to the development of NMR experiments benefiting from the TROSY principle.

TROSY experiments relies in the cancellation of certain transverse relaxation mechanisms, typically dipole-dipole coupling (DD) and chemical shift anisotropy (CSA). The components of the cross-correlated multiplet (e.g. the $^{15}\text{N}^1\text{H}$ spin system of the backbone NH on a ^{15}N labeled protein, Figure 15 A) have different combinations of these mechanisms,

originating a fast (narrow) and slow (broad) relaxing components: therefore, TROSY experiments only select the narrow component, in which the destructive interference between relaxation mechanisms takes place [78]. For a more detailed explanation, see Appendix I.

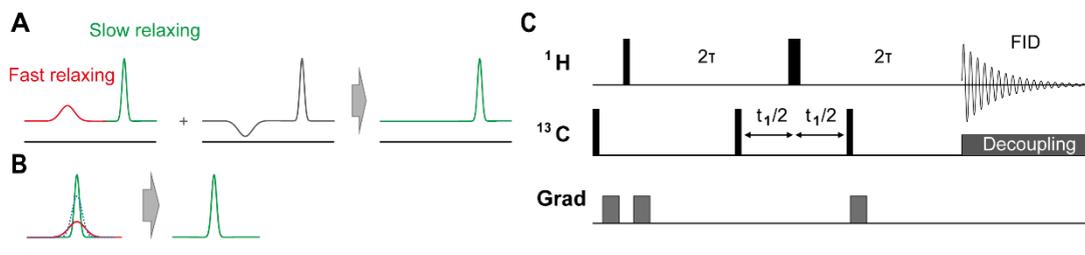


Figure 15: The TROSY effect of large biomolecules. A) Cancellation of the fast relaxing components of the ^1H - ^{15}N doublet. B) Cancellation of the internal dipolar relaxation of the methyl group. C) Pulse sequence of the methyl-TROSY HMQC experiment described by Tugarinov *et al.* [80].

In the particular case of ^1H - ^{13}C methyl-TROSY HMQC experiments used in this chapter, the internal ^1H - ^1H and ^1H - ^{13}C DD couplings of the methyl group are cancelled [80]; in fact, ^1H - ^{13}C HMQC experiments can be considered inherent TROSY experiments by itself [81]. This can be explained by carefully analysing the individual density matrix elements of the $^{13}\text{C}^1\text{H}_3$ system during the pulse sequence. Tugarinov and co-workers demonstrated that the coherence transfer pathways of the multiple quantum experiment maintained the slowly relaxing elements of the $^{13}\text{C}^1\text{H}_3$ system which give rise to the narrow components, while in the case of the HSQC version these elements were inevitably mixed (Figure 15 B) [80]. The pulse sequence of ^1H - ^{13}C methyl-TROSY HMQC experiment is shown in Figure 15 C.

The optimization of the expression protocol of [methyl- ^{13}C]-Met labeled POP and the setting of the spectroscopic parameters were carried out by Dr. Teresa Tarragó. The expression of the labeled protein proceeded without problems and yielded a high quality labeled protein; in turn, the acquisition of ^1H - ^{13}C methyl-TROSY HMQC experiments were successful and yielded an heterocorrelation spectra showing intense and well-dispersed signals. Details are summarized in Materials and methods section.

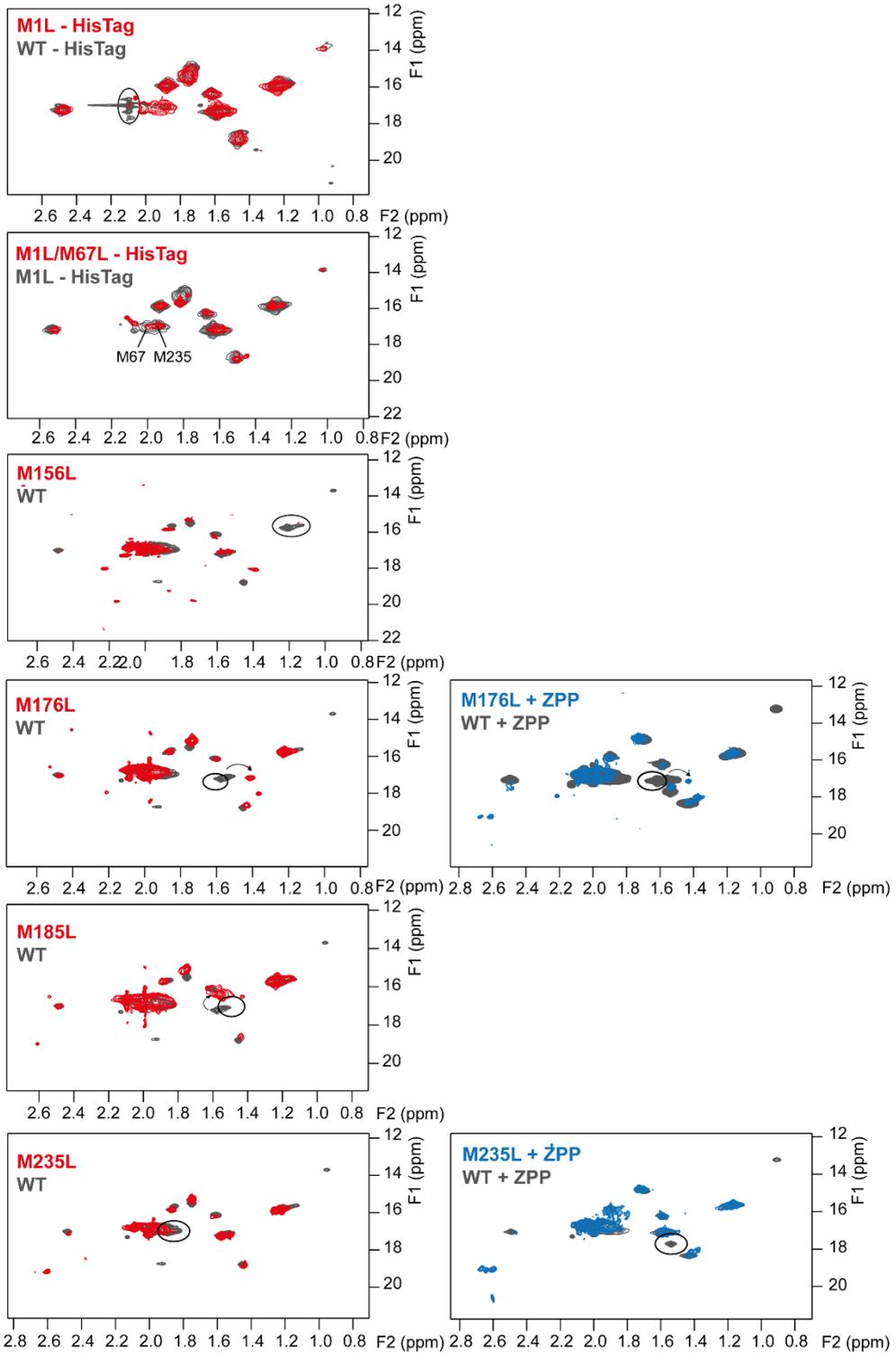
Assignment of ^1H - ^{13}C methyl-TROSY HMQC spectra

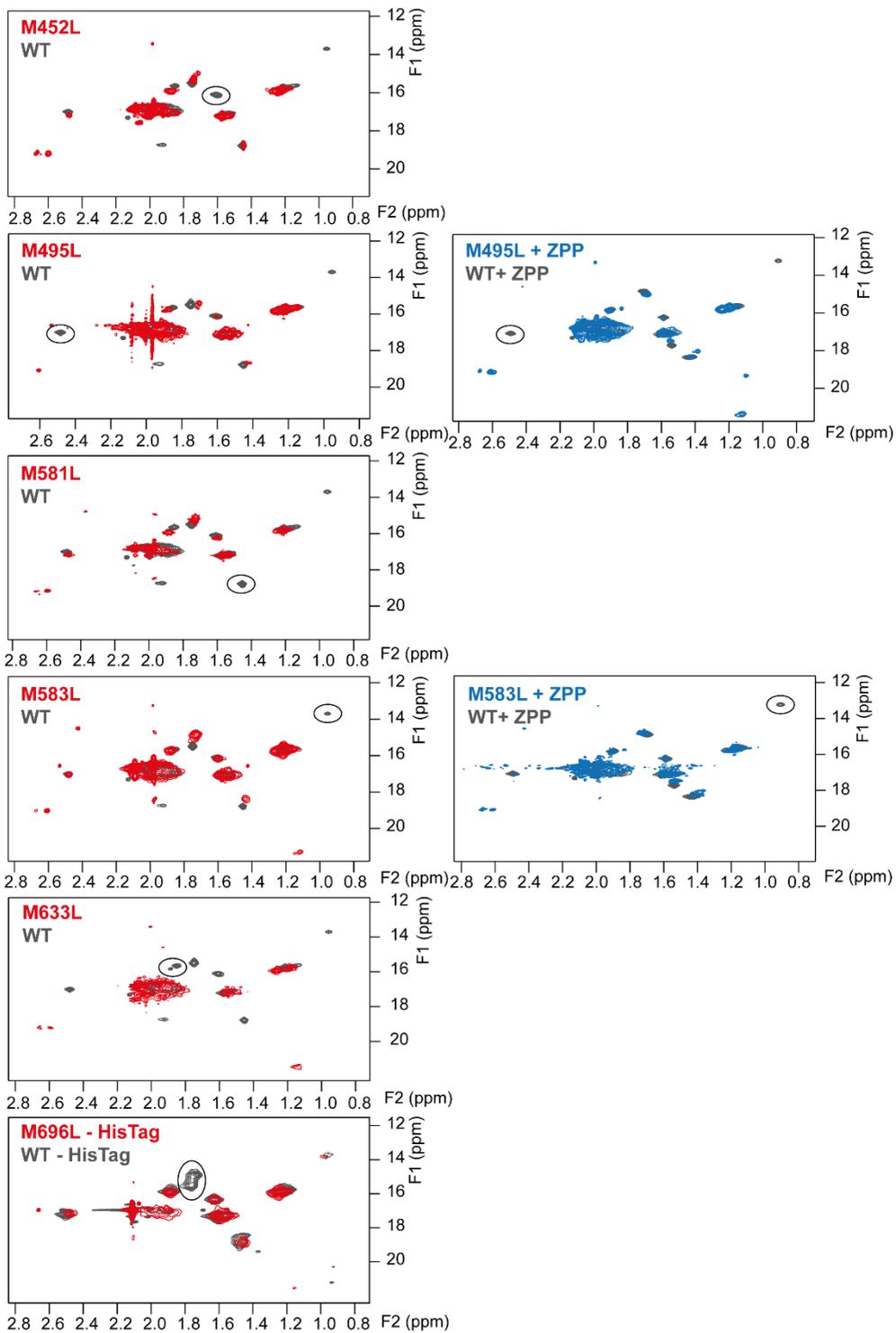
The assignment of Met signals was carried out by site-directed mutagenesis, which consists on the mutation of Met/Leu residues. The signal which is missing on the ^1H - ^{13}C methyl-TROSY HMQC spectrum is assigned to the corresponding mutated residue (Figure 16). The

rest of Met signals of the spectrum are expected to remain unaffected, since the structural alterations induced by Met/Leu replacement are minimal.

The production of mutated plasmids was performed with the Quickchange mutagenesis kit (Stratagene, La Jolla, USA). Most of mutants were expressed in pETM-10 plasmids, except those in which HisTag Met were overlapped with the signal: this occurred in M1L, M1L/M67L and M696L mutants. In those cases, pET-11 plasmids were used followed by HisTag cleavage. Similarly, if POP M1 also interferes with the assignment, the double mutant M1L/MXL was produced in pET-11 plasmid, followed by HisTag cleavage (X corresponds to the residue of interest). Regarding the ^1H - ^{13}C methyl TROSY HMQC spectrum of ZPP-bound POP, the mutants of Met residues displaying high chemical shift perturbations as a consequence of inhibitor binding were recorded for further validation.

Figure 16 summarizes all the spectra considered for the assignment of free POP Met signals. In turn, Figure 17 A and B shows the assigned ^1H - ^{13}C methyl TROSY HMQC spectra of U- $[\text{}^2\text{H}]$, [methyl- ^{13}C , ^1H]-Met labeled POP, in the free and inhibitor-bound states (ZPP and KYP-2047 inhibitors). Residues Met 156, 452, 495, 581 and 696 were assigned unambiguously. Met 67 and 235 signals of free POP were broad, and probably overlapped with Met 1 and HisTag Met. For that reason, a double mutant M1L/M67L with HisTag cleaved was expressed. Double mutant did not clarified the position of M67 signal, which remained elusive. The assignment of Met 235 in the free form of the double mutant was possible, although it was slightly ambiguous; however, Met 235 became well resolved in the inhibitor-bound POP spectra as a consequence of the direct interaction between the P3 hydrophobic position of the inhibitor and Met 235 (see below). The poor stability of Met 633 mutant did not allowed the acquisition of a high quality spectrum; for this reason, it was assigned by exclusion. In the case of Met 583 residue, the signal was significantly broad; the concentration and the quality of the sample were crucial for observing this signal.





(From the previous page) **Figure 16:** Assignment of Met signals of free POP. Red spectra correspond to free POP Met mutants; spectra of non-mutated POP is overlaid in gray. Met mutants of ZPP-bound POP are also included in the case of ambiguous Met signals. Met 1, 156, 452, 495, 581, 633 and 696 were unambiguously assigned. Met 67 is hidden by Met 1 and HisTag Met and is overlapped with Met 235 on the free enzyme; however, Met 235 is resolved on the inhibitor-bound POP. Assigned signals are marked with a circle.

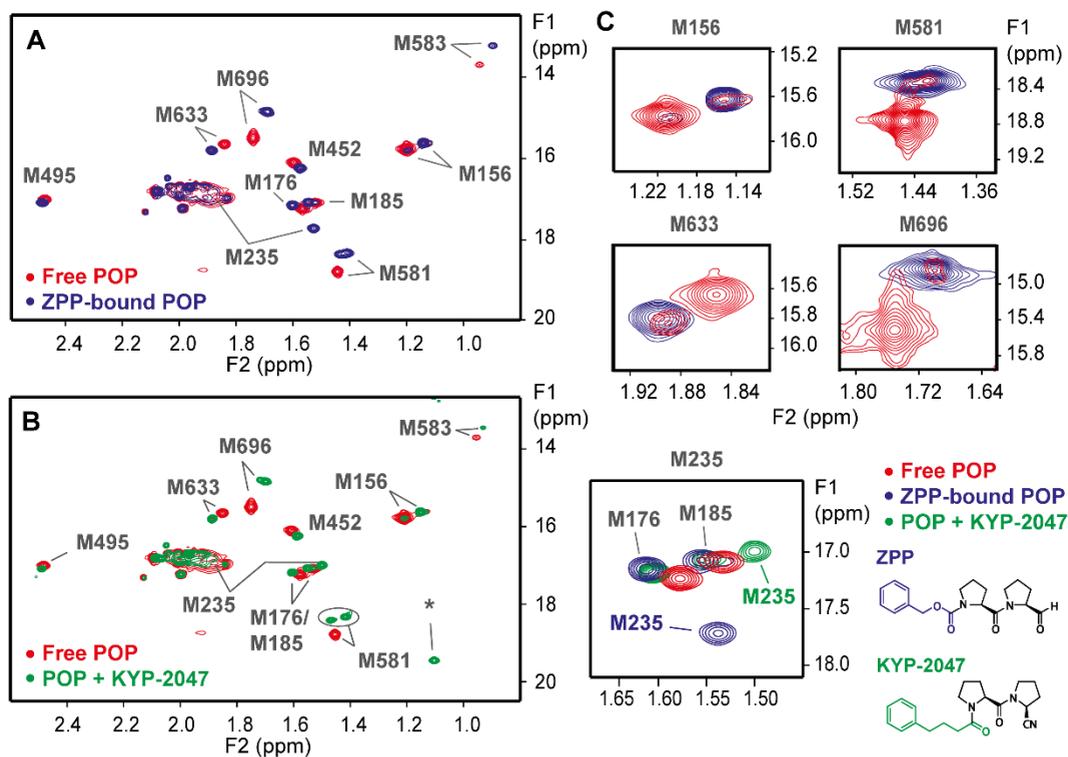


Figure 17: Assigned ^1H - ^{13}C methyl TROSY HMQC spectra of free and inhibitor-bound POP labeled with [methyl- ^{13}C]-Met. A) Spectra of free POP (red) overlaid with ZPP-bound POP spectra (blue). B) Spectra of free POP (red) overlaid with POP bound to KYP-2047 spectra (green). A non-protein signal is marked with an asterisk. C) Details of Met 156, 581, 633 and 696 signals showing the second set of signals. The signals of ZPP-bound POP are overlaid in blue. Zoomed view of Met 235 shows the chemical shift differences of this residue between ZPP-bound POP and POP bound with KYP-2047; the P3 positions of inhibitors are highlighted in the chemical structures depicted at the right.

Analysis of ^1H - ^{13}C methyl-TROSY HMQC spectra and effect of inhibitors

The ^1H - ^{13}C methyl TROSY HMQC spectra of U- $[\text{2H}]$, [methyl- ^{13}C , ^1H]-Met labeled free POP at 25 °C was analysed in detail. The spectrum disclosed 10 signals of a total of 12 Met residues, with good signal to noise ratio and excellent signal dispersion. A close inspection of the signals revealed the presence of a second set of signals for the residues 156, 581, 633 and 696 (Figure 17 C), thus evidencing a minor form in slow equilibrium with the major state in solution. The nature of this second set of signals was dependent with the experimental conditions; particularly, the resolution of the signals significantly improved in highly deuterated samples. The ^1H traces of the signals in equilibrium allowed us to estimate the time scale of the exchange process (Table 5). As it is seen, exchanging signals with small differences in their resonance frequencies ($\Delta\omega = \omega_A - \omega_B$) are resolved in the spectrum. This indicates that the two signals are in a slow exchange in the ms time scale.

	ω_A (ppm)	ω_B (ppm)	$\Delta\omega$ (Hz)
Met 156	1.213	1.136	61.6
Met 581	1.454	1.453	0.8
Met 633	1.851	1.888	29.6
Met 696	1.750	1.699	40.8

Table 5: differences in the resonance frequencies between pairs of signals observed in the ^1H - ^{13}C methyl-TROSY HMQC spectrum of free POP.

In order to monitor the effects of experimental conditions on this slow equilibrium, several ^1H - ^{13}C methyl TROSY HMQC spectra of [methyl- ^{13}C]-Met labeled free POP were carried out at different temperatures and pH values. Figure 18 A shows the spectra recorded at different temperatures ranging from 5 to 35 °C: clearly, all pairs of signals showed no coalescence during the whole range of temperatures, indicating a relatively high energetic barrier for the exchange process. In contrast, significant changes are seen in the signals showing satellites as a function of pH (Figure 18 B). Residues 156, 581, 633 and 696 experience a shift towards the second signal at low pH values. In fact, the signals at pH = 6.5 closely correspond to the satellite signals at pH = 8. However, it should be mentioned that the buffering agent of the sample at pH = 6.5 was MES-Na 20 mM instead of Tris-HCl 50 mM (the concentration of the rest of the components of the buffer was the same).

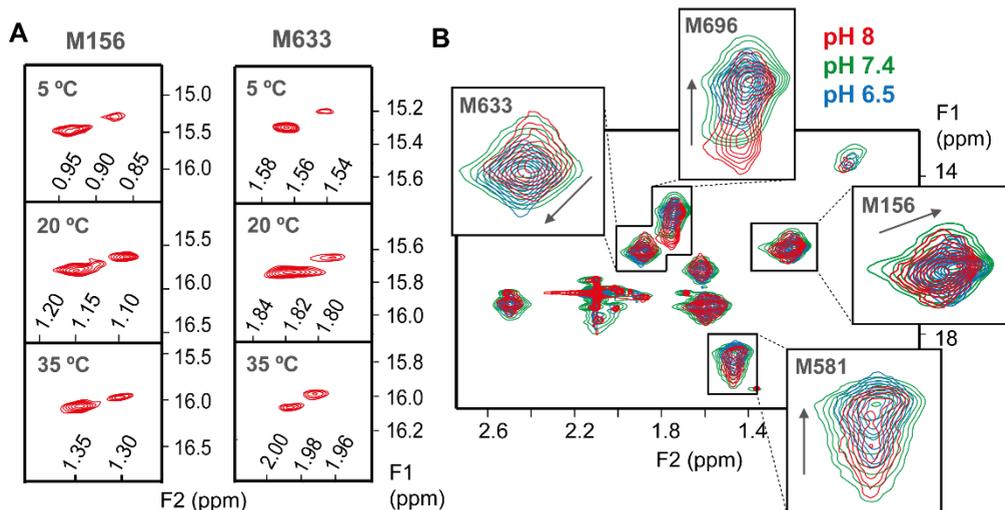


Figure 18: Dependence of free POP NMR spectrum with experimental conditions. A) Effect of the temperature on representative residues Met 156 and 633, from 5 °C to 35 °C. No significant changes in these sets of signals are observed within this temperature range. B) ¹H-¹³C methyl TROSY HMQC spectra of [methyl-¹³C]-Met labeled POP at pH = 8 (red), 7.4 (green) and 6.5 (blue). Insets show magnified views of the signals more affected by the pH change (arrows show the direction of the shift, from pH =8 to pH = 6.5).

The effect of active site-directed inhibitors on the ¹H-¹³C methyl-TROSY HMQC spectra was analysed for inhibitors ZPP and KYP-2047. The superimposed spectra of U-[²H], [methyl-¹³C, ¹H]-Met labeled free and ZPP-bound POP (Figure 17 A) disclosed significant and extensive perturbations in the chemical shifts of the signals, indicating large conformational rearrangements upon inhibitor binding. Remarkably, the signals of inhibitor-bound POP showed high similarity to the second set of signals of free POP (Figure 17 C). This suggests that the species of free POP that originates the second set of signals and ZPP-bound POP have similar tertiary structures. The chemical shift perturbation (CSP) between free and ZPP-bound POP was calculated according to the formula (Equation 1):

$$\Delta\delta = \sqrt{(\Delta\delta \text{ } ^1\text{H})^2 + (\alpha^{13\text{C}} \cdot \Delta\delta^{13\text{C}})^2} \quad \text{Equation 1}$$

Where ΔX is the chemical shift difference between signals (in ppm), and $\alpha^{13\text{C}}$ is the scaling factor for ¹³C chemical shifts, obtained as the ratio of chemical shift intervals between the extreme signals of ¹H and ¹³C dimensions, respectively. The CSP analysis disclosed that Met 581, 583, 633, 696, and to a lesser extent, Met 156, presented a high CSP values. The structural distribution of CSP located most of the residues displaying higher perturbations

(Met 581, 583, 633 and 696) exclusively in the α/β -hydrolase, thus suggesting that this domain is involved in large-scale conformational rearrangements upon inhibitor binding (Figure 19 A). Met 235 displays the highest CSP value, but this residue represents a particular case: the hydrophobic side chain of this residue is involved in the S3 binding pocket of the enzyme [2]. Therefore, Met 235 directly interacts with the P3 position of the inhibitor (Figure 19 B), and as a consequence, the ring current of P3 aromatic substituent deeply alters the chemical environment of this residue.

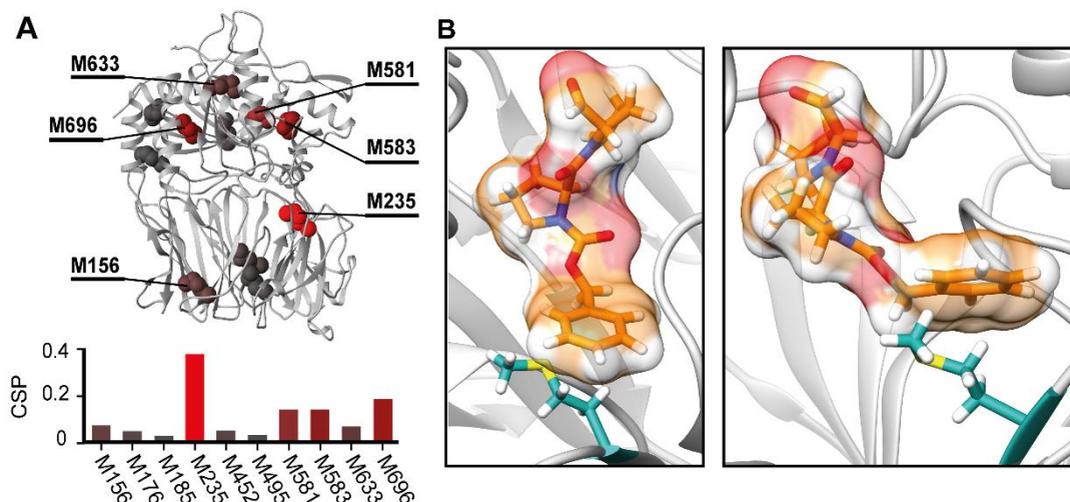


Figure 19: Binding of active site-directed covalent inhibitors. A) CSP of Met residues upon ZPP binding, and distribution on POP structure. B) Details of the direct hydrophobic interaction between Met 235 (light blue) and the aromatic P3 position of ZPP (orange).

In the case of U-[^2H], [methyl- ^{13}C , ^1H]-Met labeled POP bound with KYP-2047, a similar spectrum was obtained (Figure 17 B), suggesting that common analogous effects would be expected for all transition state analogues. According to what discussed previously, the chemical shift of Met 235 in the spectrum of POP bound with KYP-2047 is not corresponded with that of ZPP-bound POP: the particular steric and electronic properties of P3 position of KYP-2047 (Figure 19 B) cause different alterations in the chemical environment of Met 235. It should be mentioned that two resonances are observed for Met 581 in the spectrum of POP bound with KYP-2047. Probably, the presence of this inhibitor causes loop B (where Met 581 is located) to explore two different local conformations.

U-[²H], [METHYL-¹³C, ¹H]-METHIONINE POP LABELING: ¹H-¹³C METHYL-TROSY HMQC RELAXATION DISPERSION MEASUREMENTS

¹H-¹³C methyl-TROSY HMQC relaxation dispersion measurements

The high quality of the ¹H-¹³C methyl-TROSY HMQC spectra of POP labeled with [methyl-¹³C]-Met confirmed the potential of NMR in the study of the conformational dynamics of POP. For this reason, we applied more robust NMR experimental approaches in order to extract kinetic and thermodynamic information of POP conformational equilibrium in solution. Specifically, NMR experiments for the measurement of protein dynamics rely on the quantification of relaxation parameters. In the particular case of Met labeling, ¹H-¹³C methyl-TROSY HMQC relaxation dispersion (RD) experiments are the most indicated approach. In these experiments, Met probes give atomic details of μ s-ms time scale exchange events. For a better comprehension, the theory and the experimental approaches for the measurement of relaxation are summarized in Appendix I.

From the whole protein energy landscape, the mid-barrier transitions that result in μ s-ms dynamics are usually related to biologically relevant processes, such as catalysis or binding events (Figure 20). Several studies based on RD experiments found strong correlations between local μ s-ms dynamics obtained by RD and the catalytic turnover of enzymatic reactions [82-85]. In turn, RD experiments contributed to the study of the implication of μ s-ms dynamics in recognition processes [86-88]. The coupling between dynamics and recognition stressed the role of μ s-ms dynamics in the regulation of protein interactions with ligands, peptides and other proteins. Hence, RD experiments would provide valuable information of POP μ s-ms dynamics intimately connected with the biological function.

The most spread RD experiments rely on CPMG element for the measurement of R_2 [89, 90] (see Appendix I). The CPMG block consists in a series of inversion pulses separated by a delay T_{CPMG} , which modulates the contribution of μ s-ms exchange to R_2 . CPMG-based RD experiments allow the extraction of kinetic and thermodynamic parameters k_{ex} , $\Delta\omega$, p_A and p_B of a system in dynamic equilibrium between states A and B (here k_{ex} is the exchange rate constant, $\Delta\omega$ is the difference between the resonance frequencies of the states A and B , and p_A , p_B are their respective populations) (Figure 21). This approach is especially

indicated for the detection of minor populated excited states in slow to intermediate exchange that are elusive by average techniques.

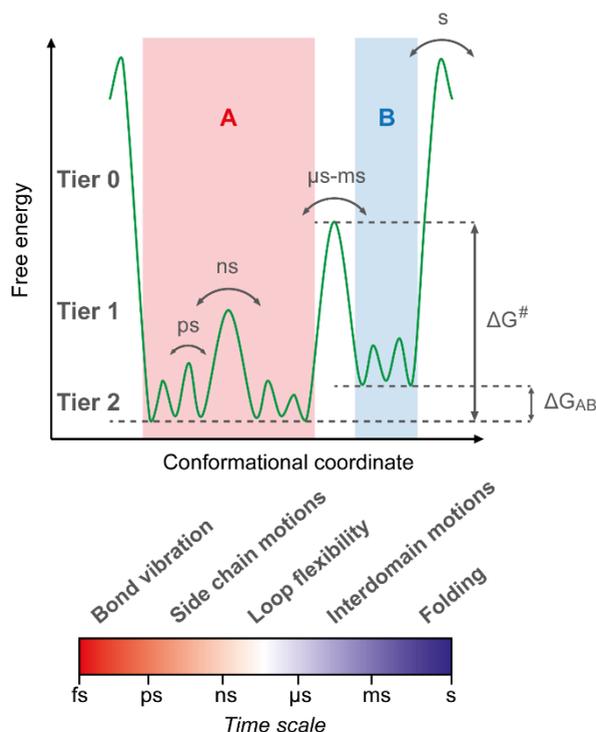
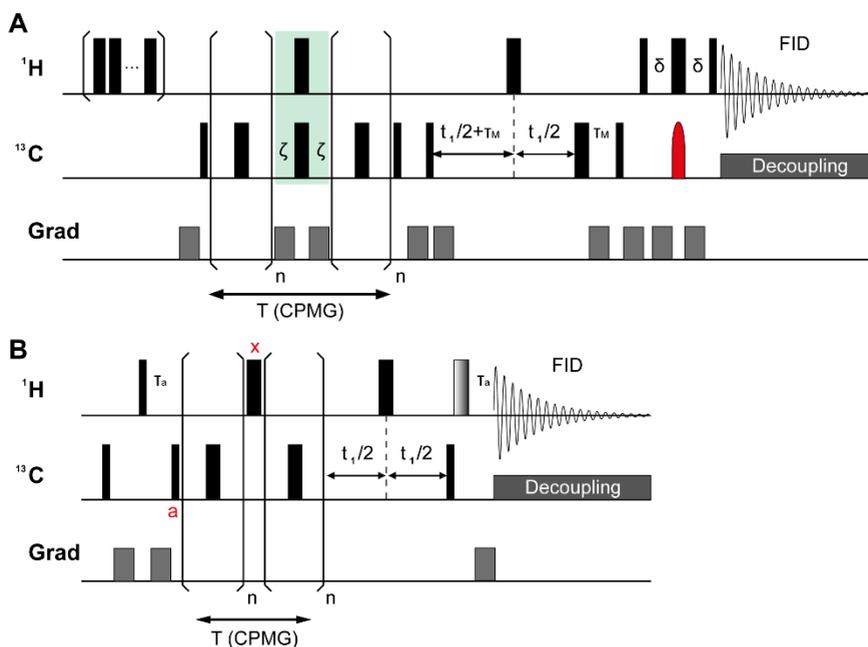
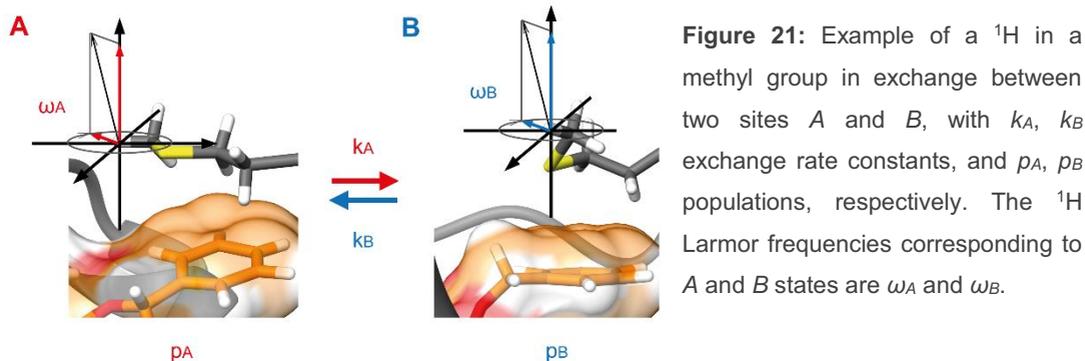


Figure 20: Protein energy landscape (green). Tier 0 transitions involve conformational transitions between A and B states in the μs - ms time range; the population of these states is determined by ΔG_{AB} , and the exchange rates, by ΔG^\ddagger . Tier 1 and tier 2 transitions give faster motions in the ns and ps time scale, respectively. Transitions at tier 0 are usually related to the biological function of proteins. Adapted from Henzler-Wildman *et al.* [91].

Most of RD experiments are mainly focused on protein backbone. However, at the beginning of this century, a methodology for the quantification of slow μs - ms dynamics of side chain methyl groups was developed by Ishima and Torchia [92]. Later, Skrynnikov and co-workers used RD experiments to extract side chain methyl dynamics [93]. However, these experiments were based on ^1H - ^{13}C HQC sequence including a constant-time CPMG block, using non-deuterated samples. This complicated the extraction of the relaxation parameters; on the one hand, the ^1H dipoles interfere with the relaxation properties of ^{13}C . On the other, the number of inversion pulses of the CPMG sequence influenced the scalar coupling effects of the methyl group. In order to minimize the effects of external protons in the relaxation of ^{13}C , the pulse sequences implemented an element which equalizes the dipolar contribution of protons (Figure 22 A).



substituted by the ^1H 180° pulse shown in gray at the end of the sequence. I_0 was obtained from the same experiment performed without the CPMG block.

Nevertheless, the HMQC version of this experiment, developed by Korzhnev *et al.*, effectively cancelled the contribution of the dipole-dipole interactions of the methyl group analogously to the TROSY experiments [60]. Moreover, the use of highly deuterated protein samples minimized spin flips and dipolar interactions with the rest of the protons of the system, which contaminates the measurement of R_2 . The absence of protons in the side chain of the amino acid is crucial for the quality of the results (see Results and discussion section of this chapter). Of interest, fully deuteration resulted in increased spectral sensitivity and more accuracy in the measurement of the exchange contribution to R_2 . The pulse sequence of ^1H - ^{13}C methyl-TROSY HMQC RD experiments is shown in Figure 22 B. In the first part of the sequence, magnetization is created in ^1H and is transferred to ^{13}C : at a point *a*, both zero and double quantum ^1H - ^{13}C coherences exist. In the case of small proteins, a purge element is highly recommended in order to eliminate all the magnetizations arising from the different spin states of the methyl protons; however, this is not necessary for big proteins (such as POP) in which these magnetizations decay rapidly and do not interfere with the experiment. In the case of large proteins, this purge element is substituted by a simple 180° ^1H pulse placed at the end of the sequence. Dephasing of ^{13}C magnetization as a consequence of R_2 is originated during the CPMG element, and afterwards is transferred back to ^1H for its detection. In order to refocus the ^1H chemical shift evolution during the CPMG train, a ^1H 180° inversion pulse is placed at the centre of the CPMG element (point *x*).

The ^1H - ^{13}C methyl-TROSY HMQC RD experiment is sensitive only to the slowly relaxing components of the magnetization. Zero and double quantum coherences are exchanged during the CPMG element, which allows the extraction of $\Delta\omega$ of both ^1H and ^{13}C . In contrast, the HSQC version of the experiment is only sensitive to the ^{13}C frequency. However, the double dependence of the relaxation with the ^1H and ^{13}C frequencies complicates the mathematic description of the system and data processing.

Processing of RD data starts with conventional processing of NMR spectra. Afterwards, RD profiles are calculated by the formula shown in Equation 2 A, in which T is the CPMG length (in seconds), I_0 and I_{CPMG} are the peak intensities without and with the CPMG element, respectively. Decay curves disclose the decrease of the exchange contribution to R_2 as a function of the effective field strength (ν_{CPMG} , related to the number of CPMG pulses by the

expression $v_{CPMG} = 1/(4T_{CPMG})$; for this reason, the transverse relaxation calculated by Equation 2 A is apparent ($R_{2,eff}$):

$$R_{2,eff}(v_{CPMG}) = \frac{-1}{T} \ln \left(\frac{I(v_{CPMG})}{I_0} \right) \quad \text{Equation 2 A}$$

$$R_{2,eff}(v_{CPMG}) = \frac{1}{2} \left(Re(\lambda_1) - \frac{1}{4n\delta} \ln(Q) \right) \quad \text{Equation 2 B}$$

The extraction of the exchange parameters proceeds by the adjustment of the NMR relaxation data to the Carver-Richards equation adapted to the multiple quantum experiment (Equation 2 B) [60]. This equation independently fits NMR data of each residue to a two-site exchange model, in which the relaxation properties of both states in equilibrium (A and B) are assumed to be equivalent. Briefly, $Re(\lambda_1)$ is equal to the Carver-Richards equation of a single quantum experiment; however, the second term Q of the equation depends on v_{CPMG} , and must be included in the fitting in case of constant-time CPMG multiple quantum experiments. In order to obtain a more robust and reliable fitting, RD measurements should be performed at different static fields [94]. Nevertheless, a detailed analysis of the fitting is out of the scope of this thesis.

Fitting of NMR data yields k_{ex} , $\Delta\omega_H$, $\Delta\omega_C$, p_B (and $p_A = 1-p_B$) and R_2 in the fast-pulsing limit ($R_{2MQ\infty}$). $R_{2MQ\infty}$ allows the extraction of another important parameter, calculated as the difference between the intensities at slow and fast pulsing limits ($\Delta R_{ex} = R_{2MQ\rightarrow 0} - R_{2MQ\rightarrow \infty}$). Typically, ΔR_{ex} is used to quantify the amplitude of the motion.

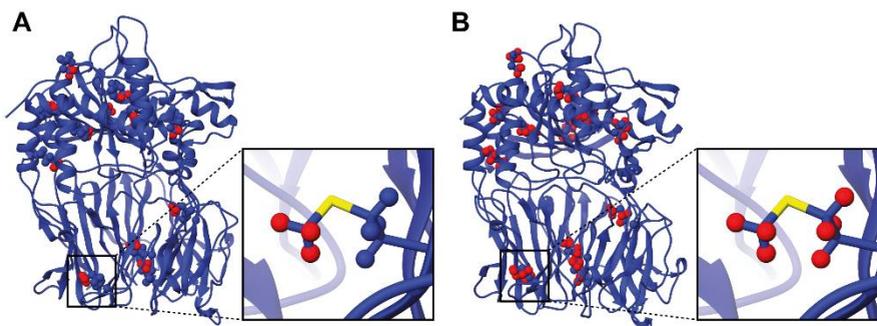


Figure 23: POP labeling with [methyl- ^{13}C]-Met. Blue correspond to deuterium atoms; protons are shown in red spheres; sulfur atom is depicted in yellow. Insets show zoomed views of Met residues for more clarity. A) U- $[\text{2H}]$, [methyl- ^{13}C , ^1H]-Met labeled POP. B) Perdeuterated POP labeled with commercial [methyl- ^{13}C]-Met (side chain protonated).

Synthesis of [methyl-¹³C]-methionine (2, 3, 3, 4, 4 - d₅) and expression of perdeuterated proteins

However, the promising NMR RD experiments require the production of a fully deuterated protein sample with the corresponding ¹H/¹³C labeled methyl probe, which is not a simple task. The preparation of a U-[²H], [methyl-¹³C, ¹H]-Met labeled POP (Figure 23 A) was carried out with *E.coli* B834 Met auxotrophic cells supplemented with [methyl-¹³C]-Met (2, 3, 3, 4, 4 - d₅).

The first step involved the chemical synthesis of [methyl-¹³C]-Met (2, 3, 3, 4, 4 - d₅). Several well established strategies for the enantioselective synthesis of amino acids and amino acid analogues [95] have been reported. However, the retrosynthetic analysis of [methyl-¹³C]-Met (2, 3, 3, 4, 4 - d₅) was highly restricted by the limited pool of commercially available deuterated building blocks. Moreover, undesired D/H exchange in labile positions should be avoided in the synthetic scheme (even during the work-up and purification of products). From the large amount of published studies, we selected 4 methodologies:

1. Enantioselective introduction of electrophilic side chain precursor. This approach is described for methionine, using commercial chiral phase transfer reagents [96, 97] (Figure 24 A).
2. Enantioselective introduction of nucleophilic side chain precursor, via proline organocatalyzed Mannich-type reaction [98] (Figure 24 B).
3. Enantioselective introduction of carboxyl group to imines via Strecker reaction, using readily synthesizable chiral organocatalysts [99] (Figure 24 C).
4. Electrophile aminohydroxylation of terminal alkenes, using chiral organocatalysts. This procedure is described for the preparation of homoserine (an advanced precursor of methionine). However, the synthesis involves the use of the toxic K₂OsO₄ as activating metal of the double bond [100] (Figure 24 D).

We focused on the first approach, since the preparation of Met isotopomers was described in the literature by Siebum *et al.* [96]. The scheme described there was adapted to the synthesis of [methyl-¹³C]-Met (2, 3, 3, 4, 4 - d₅) (Figure 25). First, the labile α protons of glycine iminic ester were exchanged by ²H by an acid/base mechanism, yielding **5**. Afterwards, the chiral S_N2 step took place with the deuterated side chain precursor **4**. The mechanism of this reaction starts with the formation of glycine iminic ester anion in the aqueous phase by deprotonation of the relatively acidic α ²H with base (in this case, 40 %

KOD in D₂O and toluene as heterogeneous binary solvent mixture was used). An ion pair is formed between glycine iminic ester anion and the chiral phase transfer catalyst, in the interphase of the aqueous and organic phases; this chiral ion pair induces enantioselectivity in the S_N2 step. Once the reaction is complete, the ion pair is broken and the catalyst is recovered for further catalytic cycles, while the product remains in the organic phase. Subsequent synthetic steps proceed essentially as described in the paper, but using commercial NaBD₄ in the reductive ozonolysis of the alkene.

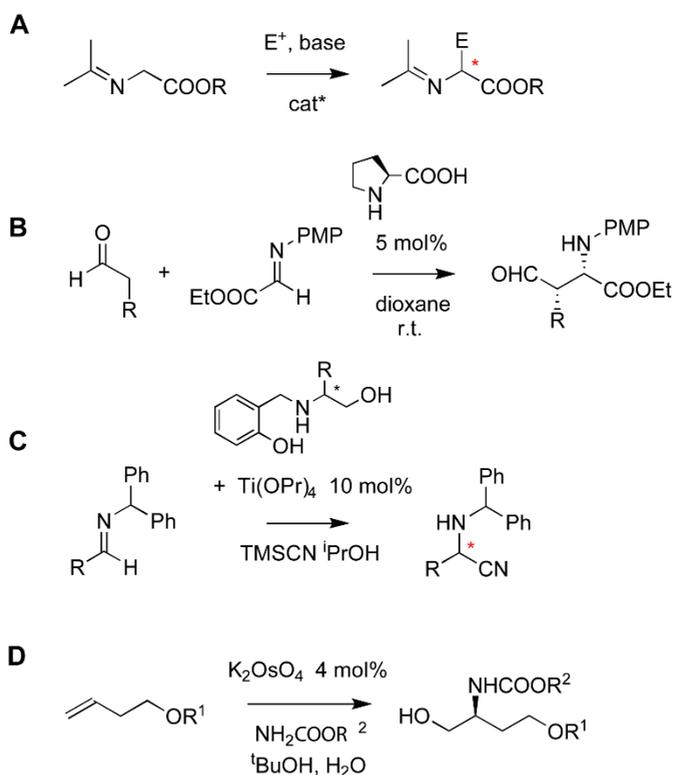


Figure 24: Strategies for the asymmetric synthesis of α -amino acids.

The deuterated side chain precursor **4** was synthesized in four steps. The retrosynthetic analysis was limited due to the poor commercial availability of deuterated alkanes; thus, we selected the reduction of the deuterated α/β -unsaturated ester **2** by commercial LiAlD₄, followed by conventional bromination of the resulting alcohol **3** as the best option. Unfortunately, the preparation of deuterated α/β -unsaturated ester by a Horner-Wadsworth-Emmons reaction was problematic. For this reason, we prepared **2** from the formal reduction with ²H from a totally unsaturated C₃ skeleton. From all the possible options of unsaturated C₃ molecules, 1-ethoxyalkynyl alcohols could be easily converted to

α/β -unsaturated esters by the Meyer-Schuster rearrangement with good yields. The analysis of the mechanism of this rearrangement indicates that the hydroxyl ^1H of 1-ethoxyalkynyl alcohols is transferred to the α -position; therefore, this position must be deuterated. 1-ethoxyalkynyl alcohol **1** was obtained by the nucleophilic attack of the ethoxyacetylene anion to a ketone, followed by quenching with D_2O (this ensures the deuteration of the hydroxyl group). Cyclohexanone was used as a carbonyl compound due to the lower volatility of the resulting products compared in the case of using acetone, which facilitated work-up operations. The synthesis proceeded in a total of 11 steps and an overall yield of 10 %; for more details, see Materials and methods.

Expression of perdeuterated POP was carried out by *E.coli* B834 Met auxotrophic cells according to a modification of our standard protocol (see Materials and methods). 100 % D_2O was used during cell growth and expression, and *D*-glucose (1, 2, 3, 4, 5, 6, 6 - d_7) was used as a carbon source in order to minimize the presence of ^1H [61, 101]. An initial adaptation of *E.coli* cells to D_2O was included by growing in an intermediate pre-culture of 50 % D_2O . The expression media was supplemented with [methyl- ^{13}C]-Met (2, 3, 3, 4, 4 - d_5) in order to incorporate the methyl NMR probe. Final yield was decreased to 1.5 mg of protein per litre of culture. For this reason, HisTag cleavage was avoided (Met of N-terminal HisTag only partially overlaps with Met 67 and Met 235, and it does not cause any alterations of POP structure and dynamics). Several trials were performed in which commercial, protonated [methyl- ^{13}C]-Met was used instead (Figure 23 B).

NMR experiments

This work was performed under the supervision of Dr Oscar Millet (CIC bioGUNE, Vizcaya, Spain). All RD experiments were carried out in the NMR facilities of CIC bioGUNE, Vizcaya, Spain. RD experiments were first optimized using non-definitive samples: the high economic cost of the synthesis and the expression of U- ^{2}H], [methyl- ^{13}C , ^1H]-Met labeled POP only allowed the production of a single definitive batch. Hence, before the definitive experiment, several preliminary trials were carried out to set up the NMR conditions. ^1H - ^{13}C HSQC RD experiments of perdeuterated POP labeled with protonated [methyl- ^{13}C]-Met were performed at 800 MHz with the pulse scheme described by Skrynnikov *et al.* [93] (Figure 22 A). However, the shaped ^{13}C pulse was eliminated because only the methyl ^1H - ^{13}C correlations occur in [methyl- ^{13}C]-Met labeled proteins.

The second set of preliminary trials consisted on ^1H - ^{13}C methyl-TROSY HMQC RD measurements of perdeuterated POP labeled with protonated [methyl- ^{13}C]-Met performed

at 800 MHz, with the pulse sequence described by Korzhnev and co-workers [60] (Figure 22 B).

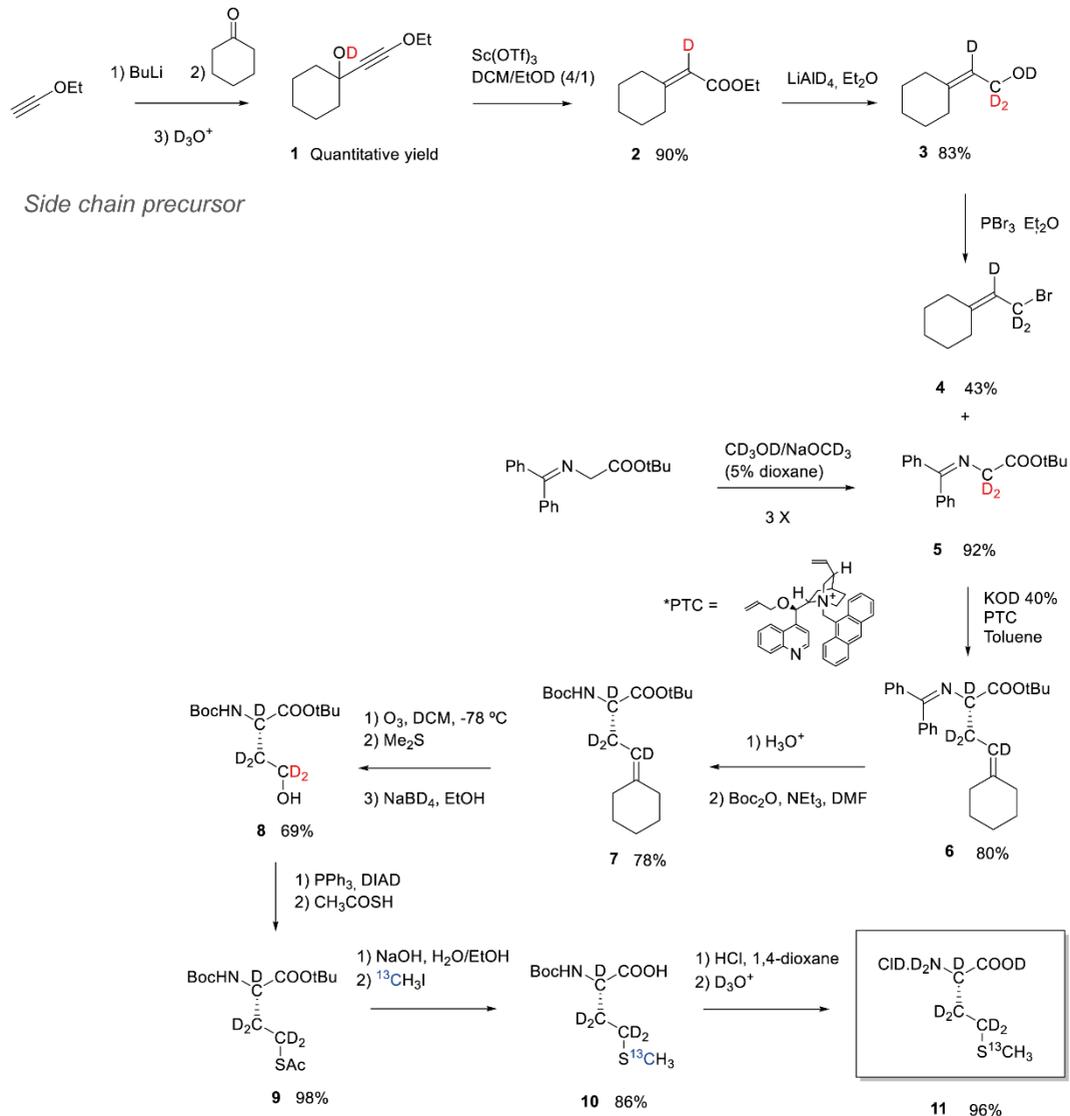


Figure 25: Synthetic scheme for the preparation of [methyl-¹³C]-methionine (2, 3, 3, 4, 4-d₅). Deuterium atoms introduced in each synthetic step are shown in red; ¹³C is shown in blue.

Finally, after these optimization steps, the definitive ¹H-¹³C methyl-TROSY HMQC RD experiments of U-[²H], [methyl-¹³C, ¹H]-Met labeled POP were performed. After the

experiment of free POP, the sample was split in two batches which were incubated with two different covalent inhibitors, ZPP and KYP-2047; these samples were used for the RD measurement of inhibitor-bound POP.

Extraction of exchange parameters

Raw data of definitive ^1H - ^{13}C methyl-TROSY HMQC RD experiments consisted in sets of HMQC spectra recorded at different ν_{CPMG} . RD profiles for each Met residue were obtained from the peak integrations with Equation 2 A (integration areas were constant so the integration ratio was equivalent to the intensity ratio of Equation 2 A)

The fitting of NMR data was performed by Dr Oscar Millet (CIC bioGUNE, Vizcaya, Spain). First, RD data was adjusted to the modified Carver-Richards (Equation 2 B) independently for each methyl group. That is:

Methyl 1: $\Delta\omega_{H1}, \Delta\omega_{C1}, K_{ex1}, \rho_{B1}, R_{2MQ\infty1} (\chi^2_1)$

Methyl 2: $\Delta\omega_{H2}, \Delta\omega_{C2}, K_{ex2}, \rho_{B2}, R_{2MQ\infty2} (\chi^2_2)$

Methyl 3: $\Delta\omega_{H3}, \Delta\omega_{C3}, K_{ex3}, \rho_{B3}, R_{2MQ\infty3} (\chi^2_3) \dots$

In this particular case, RD experiments were performed at 800 MHz. NMR RD data at a single static field allows the extraction of the individual exchange contribution (k_{ex}) and the product $\Delta\omega \cdot \rho_B$ that accounts for the population and the chemical shift of the second state B , which is ultimately related to the amplitude of the motion. The fitting of NMR data on a per-residue basis was successful; nevertheless, it should be kept in mind that the deconvolution between $\Delta\omega$ and ρ_B might experience a certain degree of inaccuracy, given that RD data was acquired at a single field. In order to overcome this handicap, several fittings were performed separately, and the results were compared between them. The reproducibility between successive results was high, and allowed the deconvolution of both parameters; only $\Delta\omega_C$ values were accepted, since the reliability of $\Delta\omega_H$ extracted from the fitting was low.

After performing the adjustment independently for all residues, the global fitting of NMR data was tested. This presupposes a global two-state exchange that affects all Met residues with the same frequency and population (same K_{ex} and ρ_B). Herein, global fitting is an indicator of long-range concerted structural changes. The system is described as:

All methyl groups: $\Delta\omega_{H1}, \Delta\omega_{C1}, \Delta\omega_{H2}, \Delta\omega_{C2}, \Delta\omega_{H3}, \Delta\omega_{C3} (\dots) R_{2MQ\infty1}, R_{2MQ\infty2}, R_{2MQ\infty3} (\dots) K_{ex}, \rho_B.$

Here, the system has more degrees of freedom, but also more parameters. In order to determine if the global fitting describes the system better than the independent one, global fitting was performed by adjusting the spectral density of pairs of residues by least-squares. Subsequently, the quality of the fitting (given by the respective χ^2 value) was evaluated against independent fitting by an F -test:

Methyl 1 (independent fitting): $\Delta\omega_{H1}, \Delta\omega_{C1}, K_{ex1}, P_{B1}, R_{2MQ\infty1}$ (χ^2_1)

Methyl 1, 2 (global fitting): $\Delta\omega_{H1}, \Delta\omega_{C1}, \Delta\omega_{H2}, \Delta\omega_{C2}, K_{ex}, P_B, R_{2MQ\infty1}, R_{2MQ\infty2}$ (χ^2_{12})

F -test: $\chi^2_{12} < \chi^2_1$?

The results of the fitting are described in the Results and discussion section.

¹H-¹³C Methyl-TROSY HMQC rotating frame relaxation dispersion experiments

In order to detect if fast μ s dynamics is occurring in POP, we performed ¹H-¹³C Methyl-TROSY HMQC rotating frame ($R_{1\rho}$) RD experiments. These experiments are based on the application of a spin-lock field which tilts the magnetization out of the z -axis (with an angle θ with respect to z -axis, Figure 26 A). The measurement of the relaxation ($R_{1\rho}$) in this new coordinate frame (named *rotating frame*) at different spin-lock fields (ω_e) allows the quantification of R_1 and $R_{2, eff}$. The theory of this experiment is summarized in Appendix I.

$R_{1\rho}$ RD experiments were developed for the detection of fast μ s backbone motions in U-¹⁵N-labeled proteins. However, Korzhnev and co-workers developed the ¹H-¹³C HSQC version of $R_{1\rho}$ experiment for the detection of ¹³C side chain dynamics [102]. Two different approaches can be applied to $R_{1\rho}$ RD experiments, depending on the offset ($\Delta\omega$) between the transmitter frequency (ω) and the frequencies of interest (ω_0): the on-resonance experiment, in which $\theta > 70^\circ$; and the off-resonance experiment, in which $\theta < 70^\circ$ [103].

In our case we used the off-resonance experiment, which varies ω or ω_1 (the amplitude of the spin-lock field) in order to modulate ω_e . Although it is preferable to perform the experiments by varying ω [104], we carried out the measurements as a function of ω_1 . This might result in a loss of accuracy due to inhomogeneous alignment of magnetization along the effective field and to the imprecise modulation of spin-lock field strength [102]; however, this approach is simple and valid to provide a qualitative view of fast μ s POP dynamics.

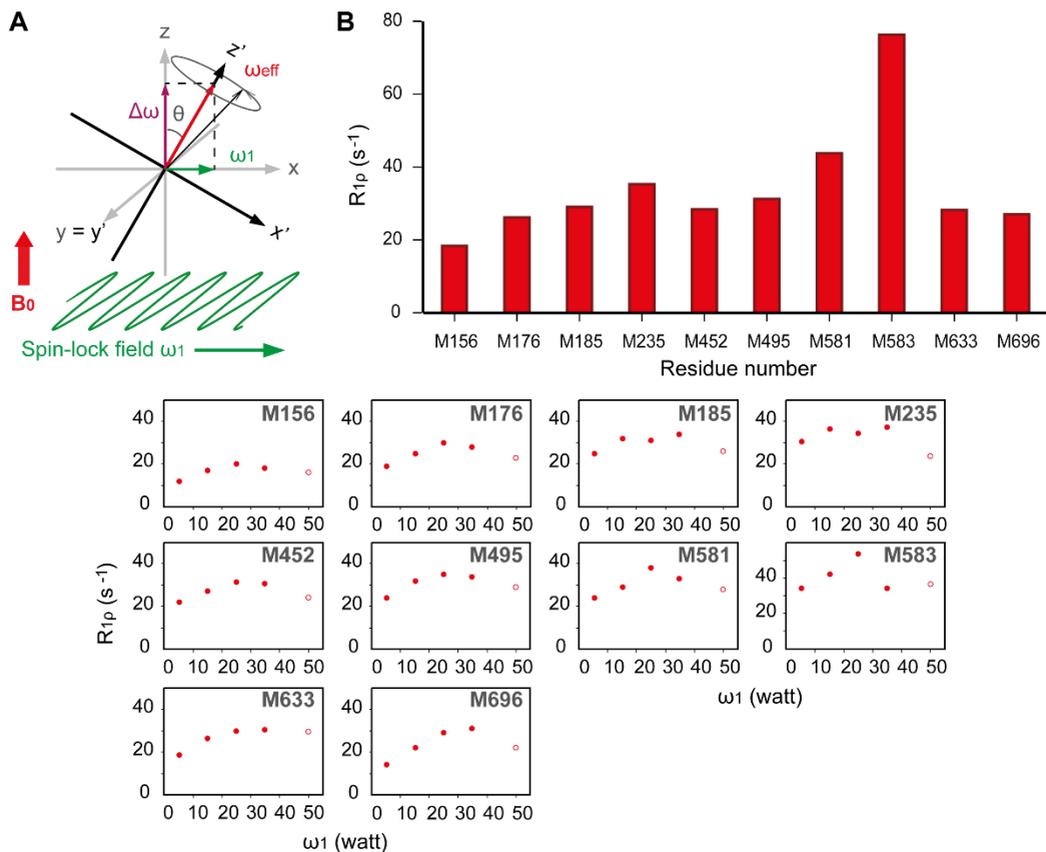


Figure 26: Relaxation in the rotating frame. A) A spin-lock field is applied in the transverse plane (green), which tilts the magnetization out of the transverse plane according to a θ angle; the new coordinate frame defined by x' , y' , z' . B) $R_{2, eff}$ values extracted from 1H - ^{13}C methyl-TROSY HMQC $R_{1\rho}$ RD experiments, corrected by the offset. The RD profiles for all methionine residues are depicted below; data points at the highest spin-lock field (50 watt, open circles) were significantly affected by the offset, and therefore, they were excluded from the analysis.

1H - ^{13}C Methyl-TROSY HMQC $R_{1\rho}$ experiments were performed with a perdeuterated POP labeled with commercial [methyl- ^{13}C]-Met (Figure 23 B). The absence of side chain deuteration of Met residues results in a slight contamination by the dipole-dipole interaction with surrounding 1H ; nevertheless, these limitations did not interfere with the qualitative analysis of fast μs motions. Experiments were performed at 800 MHz (see Materials and methods), and $R_{1\rho}$ profiles were extracted from NMR datasets by Equation 3, analogous of Equation 2 A:

$$R_{1\rho} = \frac{-1}{T} \ln \left(\frac{I(\omega_e)}{I_0} \right) \quad \text{Equation 3}$$

However, the $R_{1\rho}$ values extracted directly from Equation 3 are affected by the offset between the carrier wave and the resonances of interest. In order to correct the off-resonance effect in the measurement of $R_{1\rho}$, the following correction was used [104]:

$$R_2 + R_{ex} = \frac{R_{1\rho} + R_1 \cos^2 \theta}{\sin^2 \theta} \quad \text{Equation 4 A}$$

Since θ is near 70° , $\cos^2 \theta$ can be roughly approximated to 0. Hence, Equation 4 A is simplified to:

$$R_2 + R_{ex} = \frac{R_{1\rho}}{\sin^2 \theta} \quad \text{Equation 4 B}$$

θ was calculated from the equation:

$$\theta = \arctan \frac{\omega_1}{\Delta\omega} \quad \text{Equation 5}$$

Where $\Delta\omega$ is the offset (in Hz) and ω_1 is the spin-lock field strength (in rad/s). $R_2 + R_{ex}$ was used directly to estimate the contribution of fast μ s motions to the transverse relaxation.

RESULTS AND DISCUSSION

The ^1H - ^{13}C methyl-TROSY HMQC spectra of POP labeled with [methyl- ^{13}C]-Met showed good S/N ratio and signal dispersion. A detailed view of the free POP spectra showed that some residues display a second set of signals, indicating a second, well-differentiated POP conformation in slow equilibrium. The energetic barrier for the exchange between the two forms is relatively high, according to the poor dependence of the pairs of signals with the temperature.

The analysis of the spectra of inhibitor-bound POP disclosed high CSP upon inhibitor binding. These values indicated that POP undergo a large-scale conformational transition in the presence of active site-directed covalent inhibitors. Moreover, most of the residues displaying high CSP are located in the α/β -hydrolase domain, pointing that this domain is more affected by the conformational change. Of interest, the chemical shift of this second set of signals is highly similar to that of inhibitor-bound POP, suggesting that both forms resemble tertiary structures.

In addition, certain experimental factors altered the equilibrium between the two conformations of free POP, especially pH: decreasing pH shifted the signals towards the second set. This result suggests that charged residues are involved in the stabilization of the tertiary structure. In coherence, a doubly sigmoidal pH rate profile was observed by Polgár, which was coherent with the existence of two pH-dependent POP forms [48, 105].

Preliminary ^1H - ^{13}C HSQC and ^1H - ^{13}C methyl-TROSY HMQC RD measurements of perdeuterated POP labeled with [methyl- ^{13}C]-Met displayed no decay curves in the RD profiles (Figure 27 A and B). In the former case, the poor sensitivity due to the lack of TROSY effect was the responsible for the absence of decay. In the case of ^1H - ^{13}C methyl-TROSY HMQC RD experiment of POP labeled with [methyl- ^{13}C]-Met, the decreased sensitivity due to the dipolar interaction with neighbouring ^1H also abolished the detection of decay; for this reason, this experiment was used as a negative control for the effect of side chain deuteration on the RD profiles (see below).

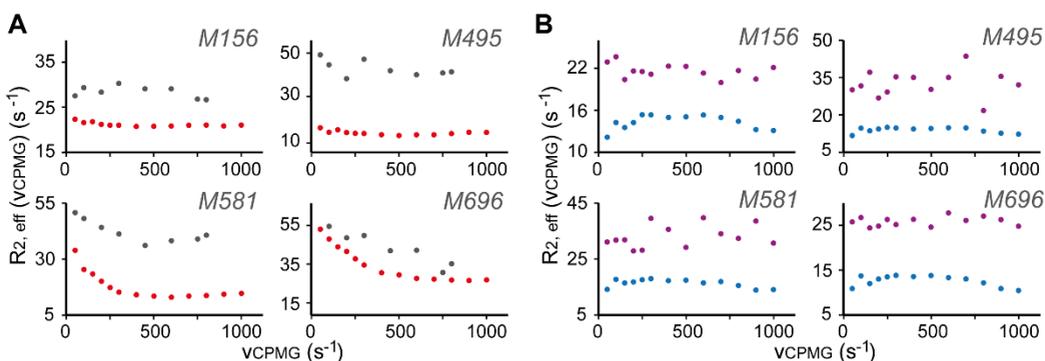


Figure 27: Negative control RD experiments. A) The significant gains in sensitivity of ^1H - ^{13}C methyl-TROSY HMQC RD experiments yields decay curves in the RD profiles (red) as a consequence of the μs -ms POP conformational exchange contribution to R_2 . Contrarily, this does not happen in ^1H - ^{13}C HSQC RD experiments, which yield flat profiles (gray). B) Met side chain deuteration is crucial to eliminate the dipolar contribution of external protons to R_2 ; the RD profiles of ZPP-bound POP (not affected by μs -ms dynamics) systematically presents increased R_2 values in the case of perdeuterated POP labeled with commercial [methyl- ^{13}C]-Met (purple) compared to U- ^{2}H], [methyl- ^{13}C , ^1H]-Met labeled POP (blue).

The definitive experiment was carried out with U- ^{2}H], [methyl- ^{13}C , ^1H]-Met labeled POP at a concentration of 262 μM . The RD profiles for this sample are shown in Figure 28 A. Remarkably, Met 581, 633 and 696, and to a minor extent Met 156, 176, 452 and 495 show decay curves. This indicates that the chemical environment of these residues is affected by

dynamic events in the μ -ms range. In contrast, Met residues 185, 235 and 583 displayed flat profiles, pointing that are hardly affected by dynamics on this time scale. As it will be seen later, the unusually high $R_{2MQ\infty}$ values of Met 235 and 583 (between 40 and 50 s^{-1}) reflects exchange contributions to R_2 in faster time scales.

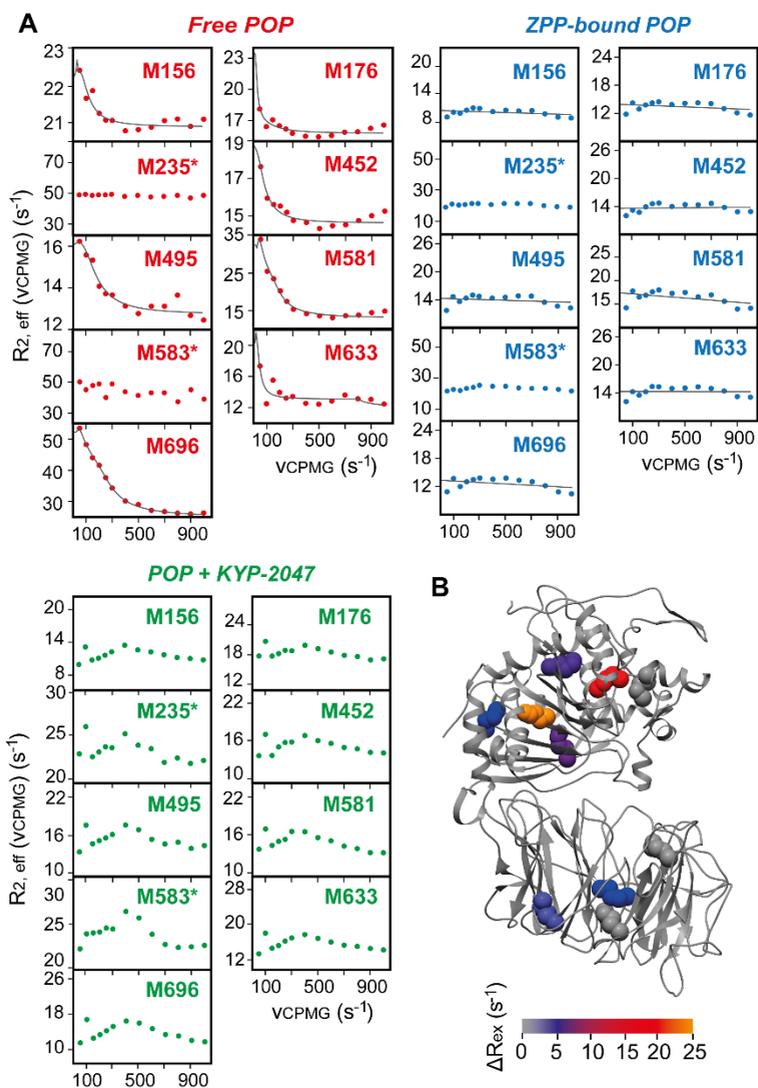


Figure 28: RD experiments of U- $[^2H]$, [methyl- ^{13}C , 1H]-Met labeled POP. A) RD profiles of free POP Met residues (red) showing a significant decay; residues marked with an asterisk show unusually high R_2 values. The RD profiles of POP bound to ZPP (blue) and to KYP-2047 (green) do not show decay curves, and the R_2 values are normalized for all Met (especially for residues marked with an

asterisk). The theoretical profiles obtained from two-state independent fitting are overlaid in grey. B) Structure distribution of free POP ΔR_{ex} values obtained from the independent fitting.

Surprisingly, RD profiles of inhibitor-bound POP disclosed the opposite scenario. A striking decrease in the dynamic behaviour of POP was observed when active site-directed covalent inhibitors were present: ZPP and KYP-2047 inhibitors completely abolished the dispersion decay for all Met residues, resulting in flat RD profiles (Figure 28 A). The significant decrease of $R_{2MQ\infty}$ in the case of Met 235 and 583 (from 40-50 s^{-1} to 20-30 s^{-1}) was also highly relevant, which indicates that fast μs motions contributing to R_2 are also cancelled in the presence of inhibitors. Hence, the generalized decrease of POP μs -ms conformational dynamics reflects the thermodynamic stabilization of a certain POP conformation upon inhibitor binding. Another important observation is the overall decrease in the $R_{2MQ\infty}$ values of inhibitor-bound POP with respect to the free form, which is a consequence of decreased correlation time. Increased tumbling rates of inhibitor-bound POP in solution is a consequence of a compaction of POP structure derived from inhibitor binding.

Met side chain deuteration is crucial for the quality of the results. Comparing the RD profiles of the multiple quantum RD measurements of perdeuterated free POP labeled with the deuterated side chain [methyl- ^{13}C]-Met (Figure 28 A) with the negative control experiment (Figure 27) discloses two relevant evidences: first, only in the former case the $R_{2, eff}$ values are modulated as a function of v_{CPMG} ; second, $R_{2MQ\infty}$ values are systematically increased in the negative control as a consequence of the dipolar contribution of side chain protons to R_2 . Together, these observations stress the relevance of using the HMQC version of the experiment described by Korzhnev and co-workers [60], together with the use of the total deuteration labeling strategy with [methyl- ^{13}C]-Met (2, 3, 3, 4, 4, - d_5).

The independent fitting of the RD data of this experiment to the modified Carver-Richards equation provided an accurate measurement of the exchange parameters, which are summarized in Table 6 (Met 185, 235 and 583 were not fitted due to the flat RD profiles). Table 6 shows that exchanging residues are affected by motions in the μs -ms time scale, with k_{ex} values spanning from 38 s^{-1} to 167 s^{-1} . Moreover, the population of the two states *A* and *B* in equilibrium is close to 50 %. Together, the exchange parameters indicates that the satellite signals of residues Met 156, 581, 633 and 696 observed in the 1H - ^{13}C methyl-TROSY HMQC spectra correspond to the partially resolved signals of the second POP conformation. In the case of inhibitor-bound POP, RD profiles did not show any measurable exchange behaviour, and therefore the fitting of RD data by the Carver-Richards equation was not possible.

The structural distribution of the Met dynamics of free POP revealed a relevant functional organization of μ s-ms dynamics of the whole enzyme (Figure 28 B). The distribution of the ΔR_{ex} values in the tertiary structure of POP clearly evidences that residues with a larger ΔR_{ex} are located in the α/β -hydrolase and the interdomain region, while those with a lower ΔR_{ex} are found in the β -propeller. Hence, this result points that significant interdomain motions of POP occur in the μ s-ms time scale, and that these events involve flexible distortions in the α/β -hydrolase; in contrast, the higher structural stability of the β -propeller [5] prevents extensive flexibility in this domain.

Residue	k_{ex} (s ⁻¹)	p_B (%)	$\Delta\omega$ (¹³ C)/Hz	$\Delta\omega$ (¹ H- ¹³ C-HMQC)/Hz
156	45 ± 7	49	±19	-30
176	50 ± 9	49	±51	-18
452	38 ± 11	51	±17	29
495	167 ± 21	45	±16	15
581	52 ± 18	46	±45	-92
633	50 ± 13	50	±15	+31
696	112 ± 14	47	±104	-128

Table 6: Independent fitting of ¹H-¹³C methyl-TROSY HMQC RD data to the modified Carver-Richards equation. k_{ex} and p_B ($p_A = 1-p_B$) values are shown in the second and third columns, respectively. The predicted ¹³C frequency intervals for the states in equilibrium ($\Delta\omega$ (¹³C)) are also shown in the fourth column. Experimental differences between the ¹³C resonances of ZPP-bound POP and the ¹³C resonances of free POP are displayed in the last column.

Although the k_{ex} values obtained from the independent fittings fell in a narrow range of values, the global fitting did not pass the F -test. This indicates that μ s-ms time scale dynamics is better described by independent motions with distinct amplitudes and frequencies. However, the global fitting forces the equilibrium between two well-defined states, which might not be a realistic description of the behaviour of POP in solution. Hence, this result is compatible with extensive opening and closing events resulting from the composition of several motions in the μ s-ms range.

In turn, $R_{1\rho}$ RD experiments show that faster time scale events are also taking place. Figure 26 B shows the corrected $R_{1\rho}$ values: Met 235 and 581 participate in fast μ s time scale events, as well as Met 583. In the latter case, the low ¹³C chemical shift of this signal results in a significant offset, which might cause a slight overestimation of $R_{1\rho}$. Hence, further

quantification of the $R_{1\rho}$ RD data will not be reliable, and these results might be taken qualitatively. The origin of the dynamics in Met 581 and 583 might be the flexibility of loop B, which contains the two residues; in the case of Met 235, the proximity of loop A probably causes significant alterations in the chemical environment of this residue. Taken together, these data confirm POP motions occurring in fast μs time scale, which also probably participate in interdomain motions.

In summary, these observations indicate that POP exists in solution in a slow μs -ms equilibrium between two main conformations, nearly 50 % populated. The different POP conformers are originated by large-scale conformational motions, and are better described by a composition of independent motions. Moreover, POP also presents local fluctuations in the fast μs time scale. Active site-directed inhibitors shift the equilibrium towards one of the pre-existing conformations, and consequently, the μs -ms dynamics is totally abolished.

Chapter 2: Analysis of POP structure in solution by small-angle X-ray scattering

INTRODUCTION

Given the promising results obtained by NMR, we subsequently focused on the elucidation of the structures of POP conformers in dynamic equilibrium in solution. We had a great interest on this goal, since all the high-resolution structures of mammalian POP obtained so far correspond to crystallographic forms displaying minimal conformational diversity. Gaining structural knowledge of POP conformers in equilibrium would be particularly relevant for the study of the enzymatic mechanism and the recognition events.

Nevertheless, this task represented a challenge at the same time. As it was discussed in the previous chapter, free POP displays a high degree of conformational heterogeneity in multiple time scales. This complicates the study of POP conformers by ensemble-averaged experimental techniques. Therefore, alternative approaches relying on the deconvolution of macroscopic experimental data should be used. These approaches must accomplish several conditions. First, the experimental techniques must be sensitive to conformational changes in a wide range of amplitudes. Moreover, the experimental methods must be sensitive to a large window of time scales. Having an accurate theoretical model of the experiment is especially recommended in order to provide an exhaustive data analysis and to confirm the reliability of the results.

From the repertoire of available biophysical tools, small-angle X-ray scattering (SAXS) stands as an excellent technique. This methodology is based on the low-angle scattering of radiation (X-ray and neutron beams) by the electronic density of the molecule (Figure 29 A). As a result of the molecular tumbling in solution, SAXS data is averaged for all molecular orientations. This entails a loss of resolution of this technique, contrarily to what happens in the X-ray diffraction of ordered crystals. However, the low-angle scattering profiles of biomolecules contains information of the overall size, shape and surface; of interest, the experiment is recorded in solution (a wide range of buffer compositions are allowed) with minimal sample consumption and without the need of labeling. For a more detailed description of the SAXS principle, the reader is referred to Appendix II.

A SAXS experiment yields a scattering profile (or SAXS curve), which is the representation of $\log I(s)$ vs s (Figure 29 A). $I(s)$ is the scattered intensity recorded in the detector, and s is the momentum transfer of the scattered radiation:

$$s = 4\pi \sin\left(\frac{\theta}{\lambda}\right)$$

This magnitude can be formally treated as a magnitude related to the dispersion angle.

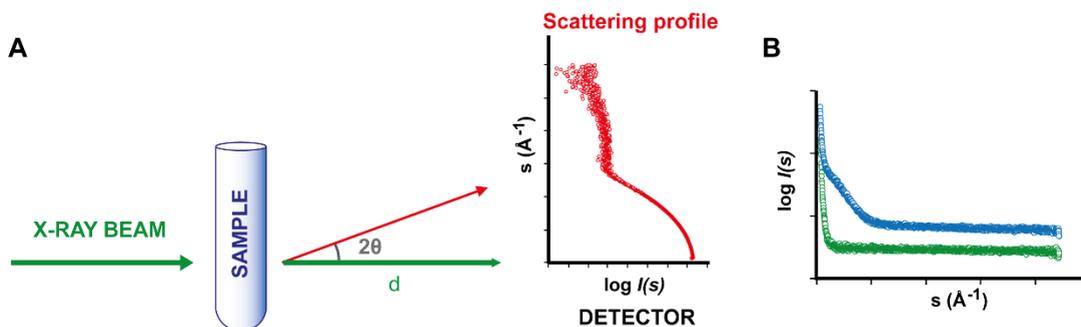


Figure 29: Scheme of SAXS principle. A) A collimated X-ray radiation (green) is scattered by the sample; the scattered intensity $I(s)$ (red arrow) is recorded in a detector at a distance d as a function of the angle 2θ . Finally, the logarithm of the scattered intensity is plotted against the momentum transfer ($\log I(s)$ vs s), obtaining the scattering profile (red). The graphic has been rotated 90° for more clarity B) The scattering profile of the biomolecule (shown in red in A) is obtained by subtracting the buffer signal (green) to the sample signal (biomolecule + buffer, blue). The curve of the sample has been displaced vertically for better clarity.

Computationally, it is relatively straightforward to generate theoretical scattering profiles from the electron densities obtained from high resolution X-ray structures, NMR structures or MD simulations. Robust software has been developed for the systematic calculation of theoretical profiles from PDB files. The most used during the last decade has been CRY SOL program of the ATSAS data analysis software [106]. Of interest, the correspondence between theoretical and experimental profiles is directly related with the structural similarity between the sample and the template structures (for more information, see Appendix II). This approach has strongly contributed to the consolidation of SAXS as a powerful biophysical tool in structural biology.

In the case of heterogeneous mixtures, experimental SAXS scattering profiles are ensemble-averages for all the K individual components of the mixture $I_k(s)$, weighted according to their populations (Figure 30):

$$I(s) = \sum_{k=1}^K u_k I_k(s)$$

Where u_k is the volume fraction of the component. Hence, theoretical scattering profiles of mixtures can be obtained from the average of individual theoretical profiles of all the components of the mixture, weighted according to their populations. The comparison between the experimental and the theoretical scattering profiles of the mixture provides the degree of fitting between both profiles, which indicates the capacity of the mixture to describe the system. Hence, together with the development of statistical tools in SAXS data analysis, this approach has been successfully exploited for the study of systems in conformational equilibrium [107, 108] or in aggregation processes [109, 110].

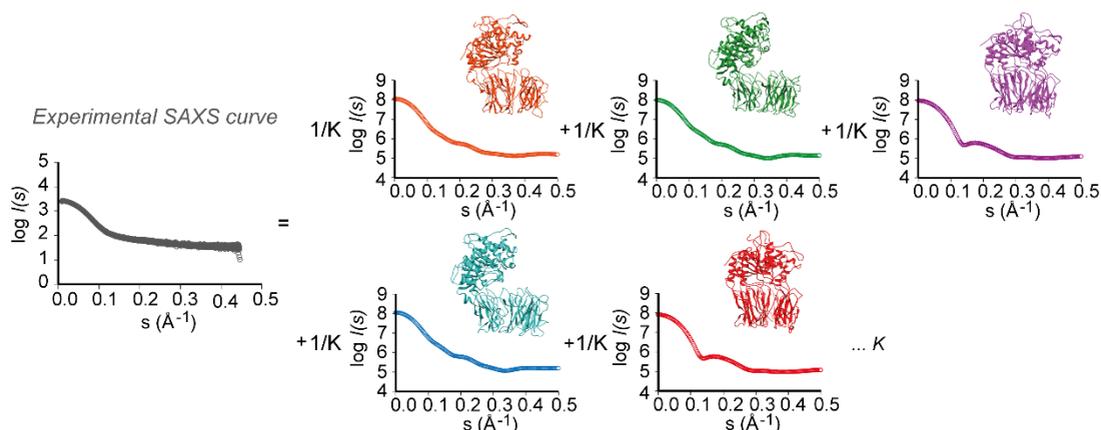


Figure 30: The scattering profile of a mixture of species (gray) is the average of all individual curves of the K components of the mixture. The example displayed here shows the experimental SAXS curve of free POP as an average of conformers with different degree of interdomain separation.

Our objective here was to extract free and inhibitor-bound POP SAXS scattering profiles, and compare them with the theoretical profiles obtained from X-ray structures, homology models, or from MD simulations. The work is described in two separate chapters (2 and 3): the present chapter describes the experimental part of the study; in turn, chapter 3 is focused on the extraction of structures that collectively fit with the experimental scattering profiles. The experimental part consisted in the measurement of experimental scattering profiles. **This work was performed in collaboration with Dr Pau Bernadó and Fátima Herranz (Centre de Biochimie Structurale (CNRS), Montpellier, France), at P12 beamline facilities of the storage ring PETRA III (DESY-EMBL, Hamburg, Germany).**

EXPERIMENTAL SET-UP: TANDEM SIZE EXCLUSION CHROMATOGRAPHY-SAXS MEASUREMENT

As mentioned previously, SAXS scattering profiles are generated from the scattering of X-ray radiation by the electrons of the sample. This includes the sample molecules, solvent and buffering agents, but also heterogeneous aggregates or oligomers; pure SAXS profiles require the subtraction of the blank buffer scattering profile (Figure 29 B), but aggregates or oligomers cannot be subtracted *a priori*. Herein, the interference of aggregates or oligomers is especially problematic in SAXS, due to the fact that large systems with high electron densities yield intense scattering profiles.

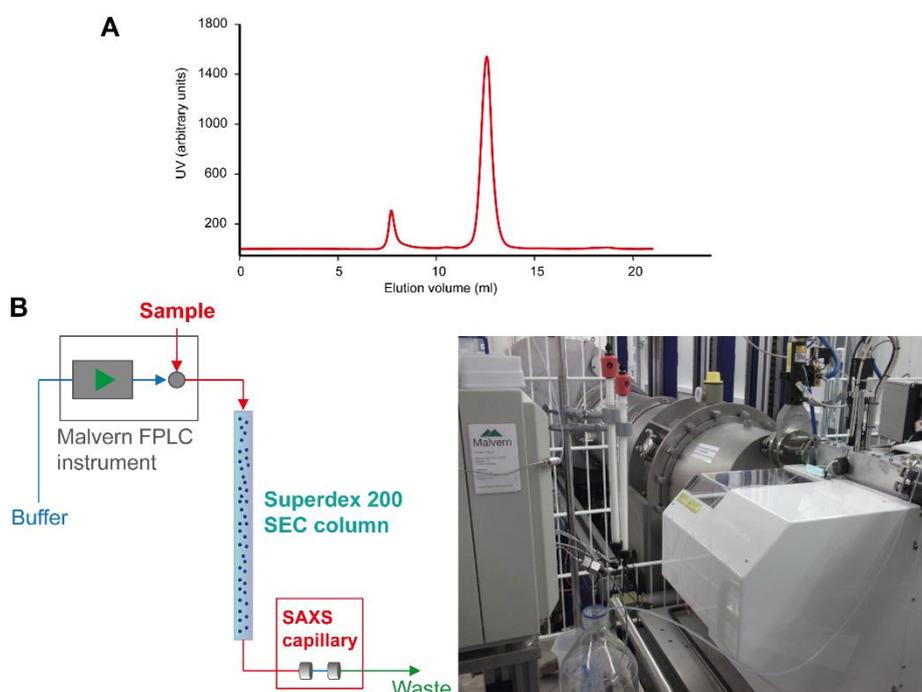


Figure 31: Size exclusion chromatography coupled to SAXS. A) Representative SEC chromatogram of free POP at room temperature, using a Superdex 200 column and standard POP buffer (50 mM Tris-HCl pH=8, 20 mM NaCl). B) Experimental scheme of the system (left) and picture of the real mounting at P12 beamline facilities of the storage ring PETRA III (DESY-EMBL, Hamburg, Germany, right).

In the case of POP, aggregation takes place during the sample preparation (the SEC chromatogram of POP is shown in Figure 31 A). Gel filtration successfully removes POP aggregates, but spontaneous aggregation slowly takes place in solution, especially in

concentrated samples. The significant amount of irreversible aggregates would strongly interfere in direct SAXS measurements of concentrated POP samples (i.e. batch measurement). Hence, an online SEC-SAXS measurement was the best solution to overcome this limitation [111, 112] (Figure 31 B); the separation between POP aggregates and monomers was excellent by conventional Superdex 200 gel filtration columns, and moreover, the facility of P12 beamline allowed the option of connecting a Malvern FPLC instrument to the SAXS capillary (Figure 31 B). Moreover, the combination of a chromatographic separation with a SAXS detector sensitive to size and shape, could disclose the partial resolution of slowly interconverting POP conformers. The detailed experimental conditions are summarized in Materials and methods section.

SAXS DATA PROCESSING

Typically, SAXS data analysis of batch measurements consists simply in the subtraction of buffer signal in order to obtain the pure SAXS scattering profile of the sample [113]. Especial attention must be given to this point: exactly the same buffer stock solution has to be used to obtain the buffer SAXS curve. However, in the present case a tandem SEC-SAXS experiment was performed. In order to minimize the slight fluctuations of the buffer composition along the SEC run, we averaged the buffer curve for the whole run. Hence, data processing was performed in two steps: first, we proceed to the identification of frames corresponding to the buffer; second, we calculated the average buffer curve and we subtracted it from protein frames. These yielded the subtracted experimental profiles.

Generation of $I(0)$ and R_g SEC-SAXS chromatograms

The preliminary step for obtaining the forward scattering ($I(0)$) and the radius of gyration (R_g) consisted in the visual inspection of all SAXS frames. This procedure introduces a time-consuming intervention of the user into the data analysis, but SAXS frames randomly present artefacts derived from large particles in suspension (e.g. dust). These artefacts were easily identified by visual inspection due to their anomalous scattering profiles, and were eliminated by removing the corresponding frame.

For the detection of buffer frames, the whole SEC-SAXS chromatogram was represented by plotting the integration of each SAXS curve from 50 to 100 data points (Figure 32) as a function of frame number (an absolute time scale was not possible due to the manual injection of the sample and other experimental limitations). This operation was carried out

for free and ZPP-bound POP SAXS data. The resulting integration chromatograms clearly showed the regions where scattering occurs (Figure 31). According to integration chromatograms, the following buffer regions were selected: from frame 810 to 1260 and from 4220 to 4500 (free POP), and from frame 0 to 1250 and from 3100 to 4500 (ZPP-bound POP). All buffer frames comprised in these regions were averaged in order to obtain the buffer blank profiles of free and ZPP-bound POP respectively. The buffer blanks were subsequently subtracted to all frames of their corresponding chromatograms. These operation were carried out with PRIMUS program [114] (ATSAS data analysis software).

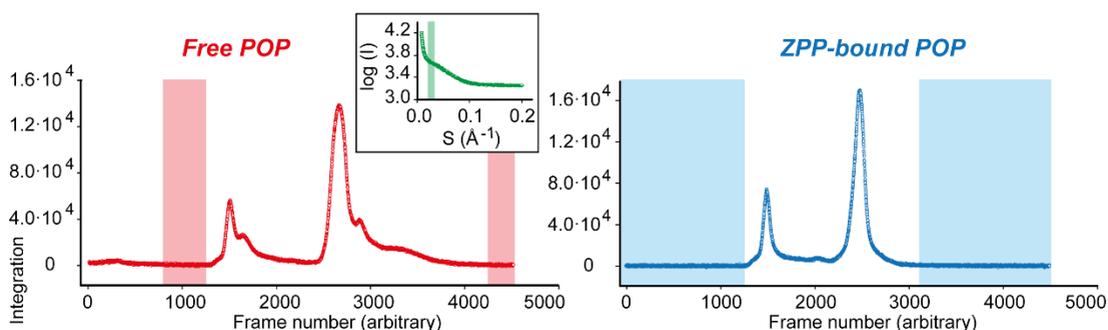


Figure 32: SEC-SAXS integration chromatograms of free POP (red) and ZPP-bound POP (blue). Red and blue regions delimit the intervals used for the averaging of buffer curve in the case of free POP and ZPP-bound POP, respectively. Inset box shows the region of the scattering profiles which was integrated to obtain the integration chromatograms (green area).

Afterwards, $I(0)$ and R_g for the subtracted frames were calculated from the Guinier approximation (see Appendix II). Guinier plots were calculated with PRIMUS program [114] (ATSAS data analysis software). Finally, $I(0)$ and R_g were plotted against SAXS frame number, which can be assigned to an arbitrary elution time; these representations were known as $I(0)$ and R_g chromatograms, respectively. $I(0)$ and R_g chromatograms reflected simultaneously the concentration and the size and of the eluting species. For this reason, these plots were highly useful to monitor the tandem SEC-SAXS measurement.

Figure 33 A shows the regions of SEC-SAXS $I(0)$ chromatograms corresponding to the monomers of free and ZPP-bound POP. As it is seen, the monomer region of free POP is found approximately between frames 2350 and 3000, and the monomer region of ZPP-bound POP, between frames 2200 and 2850. Interestingly, the SEC-SAXS $I(0)$ chromatogram of free POP disclosed two peaks coexisting in slow exchange at room temperature, referred as peaks I and II. According to R_g values, peak I presented the largest R_g . This inconsistency on the R_g values indicated that probably the two peaks corresponded

to different species. In contrast, $I(0)$ chromatogram of ZPP-bound POP displayed a single Gaussian-shaped peak, named peak III. R_g values were constant within this peak, suggesting that this peak was composed by only one specie. Therefore, the presence of different free POP conformers eluting in the monomer interval required an additional method to detect the regions corresponding to pure species.

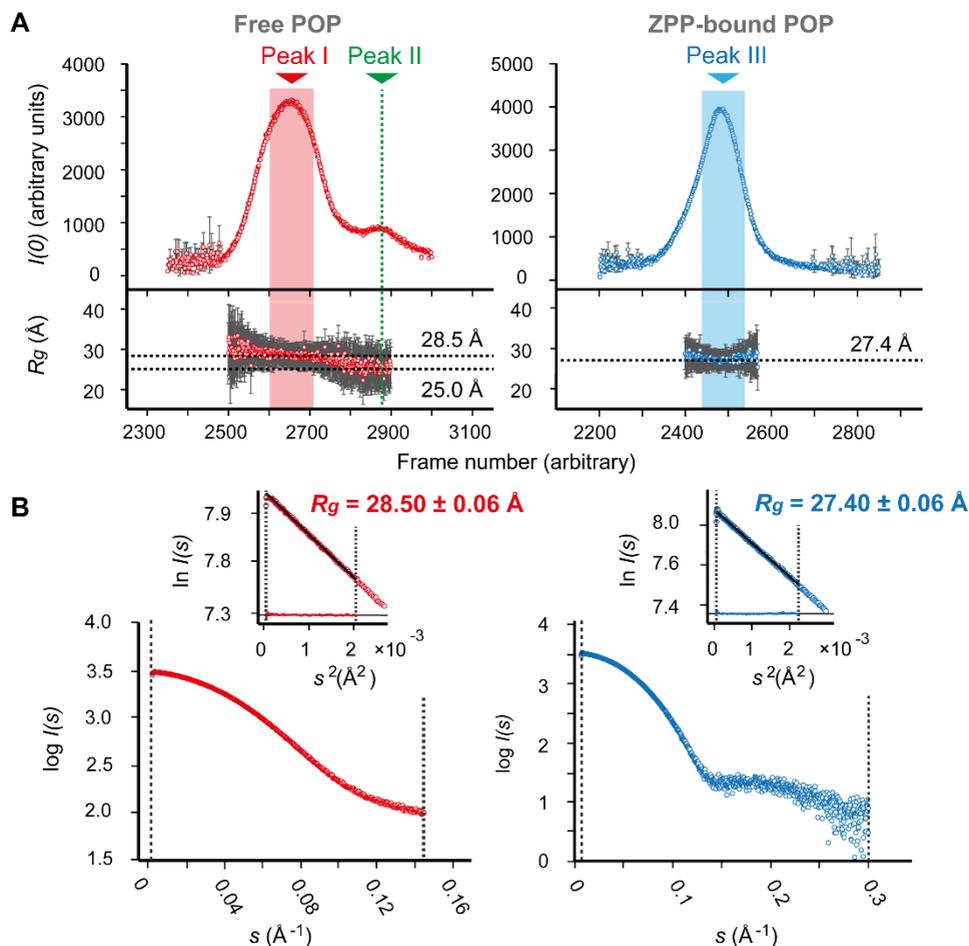


Figure 33: SEC-SAXS $I(0)$ chromatograms and averaged experimental scattering profiles. A) SEC-SAXS $I(0)$ chromatograms of the monomer frames of free POP (red) and ZPP-bound POP (blue). The red and blue areas corresponds to peak I and peak III respectively, as localized by SVD (see below). Green dotted line shows the center of peak II. R_g plot is shown below; R_g derived from the pure SAXS curves are marked in black dotted lines: 28.50 Å for peak I, 25.00 Å for peak II, and 27.40 Å for peak III. B) Scattering profiles of peak I with $s < 0.15$ Å⁻¹ (red), and of peak III with $s < 0.3$ Å⁻¹ (blue). Insets show the respective Guinier plots and R_g values.

Singular value decomposition of monomer region of SEC-SAXS $I(0)$ chromatograms

In a nutshell, singular value decomposition (SVD) consists in the factorization of a matrix. In data analysis, the factorization of a data matrix transforms the original data set described by an infinite number of variables into a data set described by a weighted number of orthogonal (i.e., uncorrelated) variables, which are known as eigenvectors.

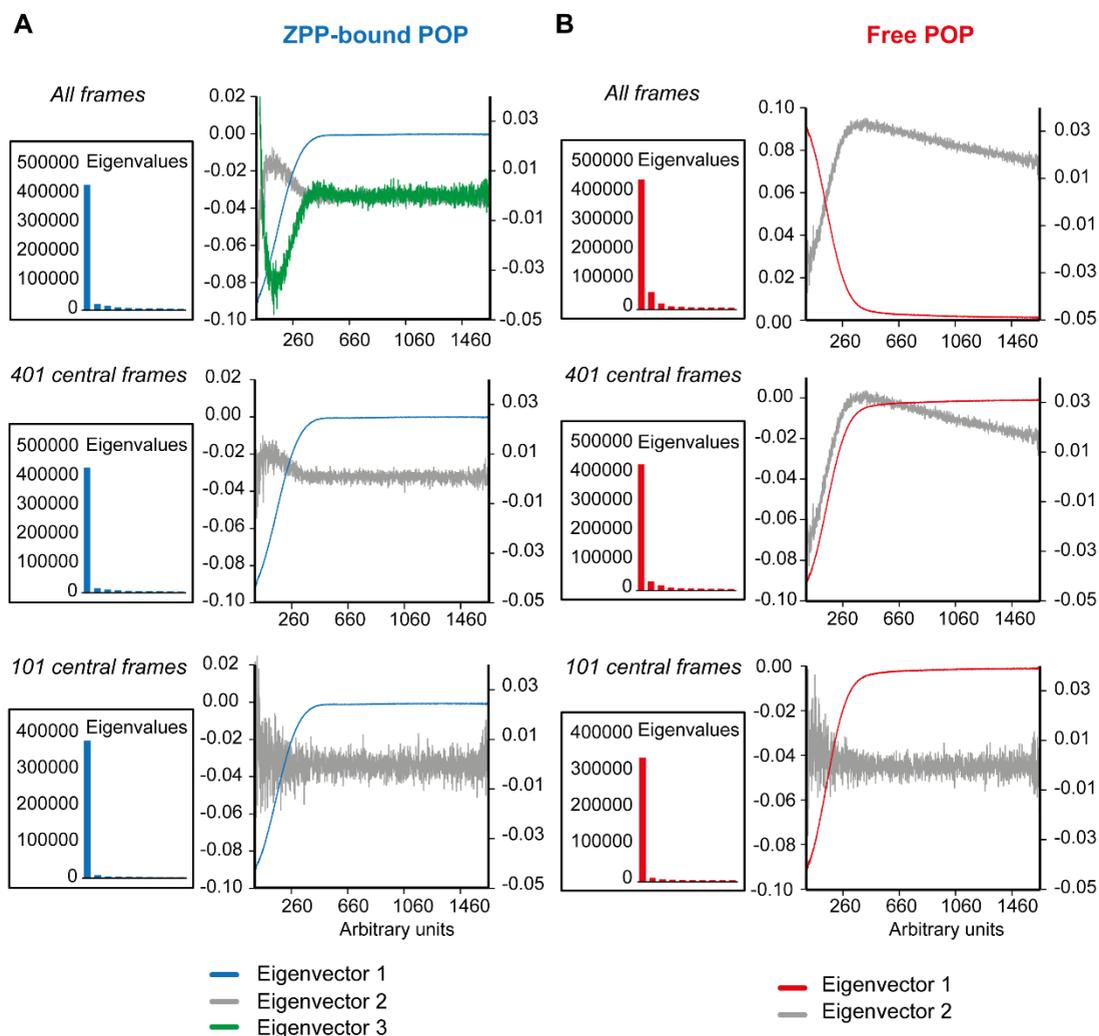


Figure 34: Singular value decomposition (SVD) of the monomer region of SEC-SAXS $I(0)$ chromatograms at different intervals; eigenvalues are displayed in the insets. A) SVD calculations of peak III of ZPP-bound POP show only one predominant eigenvector and eigenvalue, independently of the interval used in the algorithm. In the case of the SVD of all monomer region, a third eigenvector

is shown for more clarity. B) SVD of the full monomer region of peak I of free POP shows a significant contribution of a second eigenvector, as well as in the case of the 401 central frames. Only, in the case of 101 frames centred at peak I, a unique eigenvector is present.

In the present case, SEC-SAXS $I(0)$ chromatograms correspond to a second order data matrix: one order is the elution dimension (i.e. the chromatogram), and the second is the spectral order (i.e. the SAXS curves). Thus, eigenvectors will describe the chromatogram, and their corresponding eigenvalues will give an approximate idea of the weighting of the eigenvectors. Hence, chromatograms were subjected to SVD in increasing intervals centred at peaks I and III (frames 2655 and 2488, respectively). Peak III was the simplest case: here, only one significant eigenvector and eigenvalue were obtained independently of the interval subjected to SVD (Figure 34 A). This result shows that peak III is constituted exclusively by a single species. Regarding peak I, only the SVD of 101 central frames yielded one significant eigenvector and eigenvalue. If this region was expanded (for instance, in the case of considering the 401 central frames, Figure 34 B) the participation of secondary eigenvalues and eigenvectors became significant. These secondary eigenvectors were attributed to the contamination with species of peak II. Therefore, the 101 central frames of peaks I and III were averaged to produce high-quality SAXS curves corresponding to pure species (Figure 33 B). Moreover, the interval with $s \cdot R_g < 1.3$ of the Guinier plots of the pure curves of peaks I and II corresponded to a linear region, indicating that no significant interparticle interactions or radiation damage were present (insets of Figure 33 B). However, it was not possible to apply the approach described above for peak II due to chromatographic tailing from peak I, and a different strategy based on multivariate curve resolution by alternating least squares was used.

Multivariate curve resolution by alternating least squares of peak II

In order to extract the pure SAXS curve corresponding to peak II, all the frames were decomposed with multivariate curve resolution-alternating least-squares (MCR-ALS) [115]. Briefly, this chemometrics method was developed for the decomposition of spectroscopic mixtures [116]. In this case, our data consisted in a collection of scattering profiles (i.e. the SEC-SAXS chromatogram) of a two-component mixture, in which each SAXS frame is a mixture of the scattering profiles of species corresponding to peaks I and II at variable concentrations (Figure 35). The starting point is the second order data matrix, which is subjected to a factorization (typically by SVD). The obtained eigenvalues and eigenvectors

are subsequently optimized by alternating least squares according to certain constraints. This method resolves concentration and spectral profiles, which represent the individual spectra which constitutes the initial mixture.

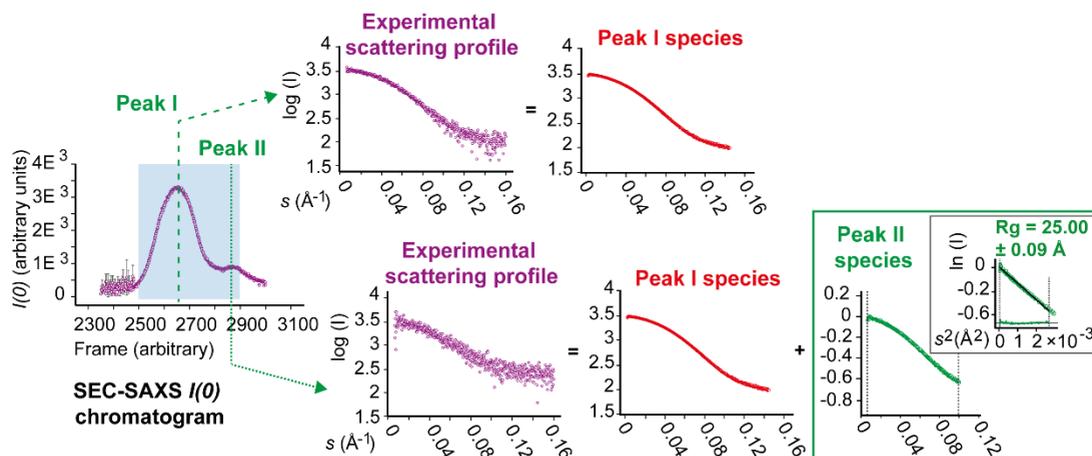


Figure 35: Application of MCR-ALS to SEC-SAXS $I(0)$ chromatogram of free POP. The chromatogram shows the contribution of the two species to the scattering profiles at different regions of the chromatogram. While peak I shows only one species, peak II species is contaminated with species of peak I; for this reason, the interval of SEC-SAXS $I(0)$ chromatogram highlighted in blue was subjected to MCR-ALS in order to extract the SAXS curve of peak II (green). Inset box shows the corresponding Guinier plot and R_g .

The whole monomer region of free POP was subjected to MCR-ALS (frames from 2500 to 2900, Figure 35). In our case, unimodality was used as a restraint for the MCR. Regarding ALS, the monomer region was simplified in a two-component mixture, formed by different proportions of the species of peak I and peak II. This simplification used the pure averaged SAXS curve of peak I obtained by SVD as a common order for all slices. With these conditions, MCR-ALS successfully solved the common elution and spectral profiles of both species, i.e. the concentration and the pure SAXS curves of the species corresponding to peaks I and II. Unfortunately, the low intensity of SAXS curves corresponding to peak II only allowed the deconvolution of the lowest resolution part of scattering profile ($s < 0.1 \text{ \AA}^{-1}$). The concentration of peak I was estimated to be 0.18 mg/ml, according to the comparison of intensities with a BSA standard solution. The Guinier plot of the interval with $s \cdot R_g < 1.3$ of this curve was linear, indicating high quality of the data (Figure 35).

RESULTS AND DISCUSSION

The tandem SEC-SAXS experiment allowed the isolation of the monomeric species of free POP and ZPP-bound POP, and the concomitant measurement of the monomer scattering profiles without any oligomer contamination. In order to monitor the elution of POP conformers through the gel filtration column, SEC-SAXS $I(0)$ chromatograms were generated: in the case of free POP, two different enzyme forms were partially resolved by the tandem SEC-SAXS. This is the first time we observed different species of free POP by gel filtration chromatography. However, the slow exchange regime between the two free POP species observed here might not reflect the real situation. It should be stressed that species eluting in the gel filtration matrix are in an altered environment, with high local crowding and strong interactions with the stationary matrix [117].

The SVD delimited the pure regions of peaks I and III. By averaging the SAXS curves of the pure regions, high quality scattering profiles for both peaks were obtained. These SAXS curves were then used to fit the theoretical data obtained by MD (see chapter 3). The R_g extracted from the Guinier approximation corresponding to peaks I and III were $28.50 \pm 0.06 \text{ \AA}$ and $27.40 \pm 0.06 \text{ \AA}$, respectively (Figure 33 A). The discordant R_g values between the species of peaks I and III indicates that they exist in different conformational states. In this regard, the R_g of peak I is much higher compared to that calculated for the X-ray closed structure 1QFS (25.82 \AA), suggesting that large conformational rearrangements are occurring in the species eluting at peak I.

In the case of the scattering profile of peak II extracted by MCR-ALS, the corresponding radius of gyration was $25.00 \pm 0.09 \text{ \AA}$ (Figure 35). This result is consistent with the order of elution observed in the online SEC-SAXS, and also discloses significant structural changes between the two species. Unfortunately, the low intensity of the scattering profile of peak II obtained by MCR-ALS precluded the fitting of theoretical data. The R_g value of this peak is slightly smaller than the POP crystallographic structure 1QFS. We speculate that this secondary POP conformer arise from a tight arrangement of β -propeller blades [2]; the collapse of the internal tunnel of this domain would result in smaller R_g . However, further work is required to elucidate the structure this minor form.

Finally, in order to have an estimation of the spatial arrangement of atoms of species in peaks I and III, the pair-distance distribution function $P(r)$ was calculated for the pure SAXS curves of the two species with GNOM program [118] (ATSAS data analysis software, Figure 36). This distribution is sensitive to small changes in the protein structure, which alter the

distribution of atoms around the centre of masses and lead to a distortions in the $P(r)$ distribution (see Appendix II). In the case of free POP, the pure curve corresponding to peak I yielded a multimodal distribution, pointing that the tertiary structure of this specie features an irregular global shape. In contrast, for the pure curve of peak III, a Gaussian-like distribution was obtained. Hence, ZPP-bound POP resembles much more globular shape.

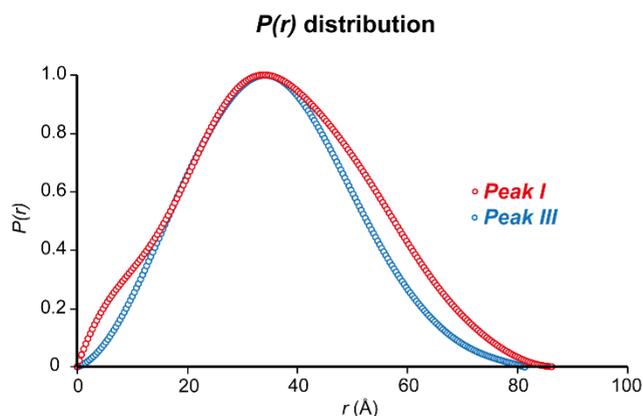


Figure 36: $P(r)$ function of species corresponding to peaks I (red) and III (blue). The multimodal distribution of peak I discloses the structural heterogeneity of this species.

Taken together, these successful and promising results yielded unprecedented evidences of the conformational equilibrium of POP in solution, and the effects of inhibitors in this process. The preliminary analysis of SEC-SAXS $I(0)$ chromatograms, R_g values and $P(r)$ functions confirmed that POP undergo important conformational changes in the presence of inhibitors. Of interest, the high quality of the SAXS data brought us the possibility to structurally characterize the POP species in dynamic equilibrium with the use of simulated data (see chapter 3).

Chapter 3: POP molecular dynamics simulations and correlation with experimental data

INTRODUCTION

At this point, we obtained a large amount of high quality experimental data showing that free POP exists in a μs -ms dynamic equilibrium involving large conformational rearrangements between domains. However, given the nature of the experimental approaches, it was not possible to extract the high-resolution structures of the conformers in equilibrium.

In order to overcome this limitation, we decided to explore the POP conformational landscape by generating of a pool of POP conformers by computational methods. Importantly, the use of statistical and analytical methods would allow us to compare this microscopic data obtained by computational approaches with the ensemble-averaged properties measured by experimental methods. The similarity between experimental and computational data would give us the opportunity to the extract valuable information at the atomic level. Therefore, the cooperation between experimental and computational approaches would significantly complement the information of POP conformational dynamics much more than using the two approaches separately.

However, the quality of the description of experimental data by computational approaches is highly dependent on the initial pool of states (in this case, conformations) subjected to study. For this reason, it is recommended to use deterministic computational methods in order to avoid artefacts or unrealistic conclusions. Here, we relied in MD simulations in order to explore the conformational landscape of POP.

Computational calculations were carried out by Dr. Martin Kotev, a member of our laboratory, under the direction of Dr. Victor Guallar (members of joint BSC-CRG-IRB Research Program in Computational Biology, Barcelona Supercomputing Center).

Principles of molecular dynamics simulations

Quantum mechanics defines the total energy of microscopic systems by the Hamiltonian operator (\hat{H}). This operator is constituted by the sum of kinetic and potential components (\hat{T} and \hat{V} , respectively):

$$\hat{H} = \hat{T} + \hat{V}$$

Applying this operator to the energy function gives the analytical expression for the thermodynamic properties of molecules; however, this is only possible in simple cases. Increasing the complexity of the system avoids the exact description of the forces between particles and the potential energy. This limitation forces the use of approximations in order to determine atomic positions and energetic properties computationally.

Experimental approaches yield macroscopic properties measured as ensemble- and time-averaged data. Hence, the expression of macroscopic properties (A) from microscopic coordinates is crucial to simulate theoretical data. In order to define A , it is necessary to describe the microscopic positions $\mathbf{r}^N(t)$ and momenta $\mathbf{p}^N(t)$ of all N particles of the ensemble. In the case of time average, the averaging is achieved by integrating the property over an infinite period of time:

$$\bar{A} = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_{t=0}^{\tau} A(\mathbf{p}^N(t), \mathbf{r}^N(t)) dt$$

However, performing this integral for a macroscopic population is not feasible due to the large number of particles. To overcome this limitation, Boltzmann and Gibbs alternatively developed the statistical mechanics, which connects the microscopic quantum mechanics with the ensemble-averaged observables. This theory supposes that all microscopic systems are forming part of an ensemble (Box 2), and their properties are distributed according to the Boltzmann function (Figure 38 A). The Boltzmann distribution allows to change the time average of a system of N particles by the average of a large number of A replicate calculations of a single particle. This is the basis of Monte Carlo and MD simulations (Figure 38 B). In turn, MD simulation is a deterministic method which allows the prediction of future states of the system from a given initial state by calculating the positions and trajectories of the particles of the system at different time intervals $A_1, A_2, A_3...$ [119] (Figure 38 C).

Box 2: Ensembles in statistical mechanics. Statistical mechanics suppose large ideal collectives of identical microscopic systems under certain conditions, which are used to derive macroscopic properties: these collectives are called ensembles. Depending on the conditions assumed for the equilibrium, different types of ensembles in equilibrium can be defined:

- Microcanonical ensemble: the number of particles (N), volume (V) and energy (E) are constant. This ensemble reaches the equilibrium at maximum entropy.
- Canonical ensemble: here, N , V and temperature (T) are constant.
- Grand canonical ensemble: in this case, N , T and the chemical potential (μ) are constant.

The calculations used in MD simulations are based on the Newton's principles. The second Newton's law connects the motion of a particle of mass m along x coordinate under the influence of an external force F_x :

$$\frac{d^2x}{dt^2} = \frac{F_x}{m}$$

In order to simplify the calculations, collisions between particles are supposed elastic (the collision result in no loss of energy). The conservation of the momentum on particle collisions is achieved by assuming Lennard-Jones potentials between particles; Lennard-Jones model calculates the potential between particles as a function of interparticle distance (Box 3, Figure 39), providing a good description of interparticle collisions. However, this model requires the recalculation of the potential energy at small time intervals, according to the updated position and velocities of the particles. The successive calculations of particle positions and momenta at increasing time interval constitutes the MD trajectories.

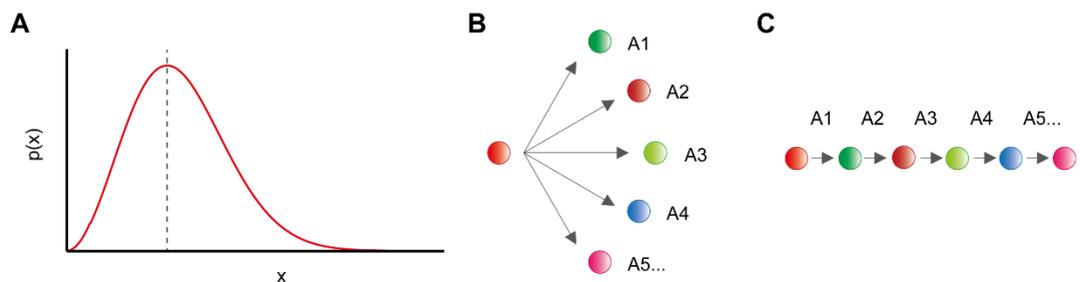


Figure 38: Statistical mechanics and ensemble-averaged observables. A) Example of a hypothetical property (x) distributed according to Boltzmann function $p(x)$. Dashed line indicates the most probable value of x . B) Scheme of Monte Carlo simulations, in which replicate calculations of A property are used to extract ensemble-averaged data. C) Scheme of a MD simulation, in which A is predicted at different time intervals as a function of the position and momentum of the particle.

Historically, the conditions in which these calculations were performed forced constant number of particles (N), volume (V) and energy (E). These conditions correspond to the microcanonical ensemble, which defines all the possible states of the system that are defined by the same N , V and E values.

Box 3: Lennard-Jones potential. The Lennard-Jones potential is an improved simplification for calculating the interaction forces between two particles as a function of their relative position (e.g. two atoms approximating). This approach is more realistic than the hard sphere potential model, which supposes completely elastic collisions only when the two particles physically interact (this is the typical “billiard ball” example). The Lennard-Jones potential between two neutral atoms separated by a distance r considers the Van der Waals attraction between two distant atoms (which is a function of $1/r^6$) and the repulsive forces that appear between the electron clouds when two atoms are close in the space (which is a function of $1/r^{12}$). The combination of the two contributions yields the Lennard-Jones potential between particles i and j (Figure 39). The force between the two atoms is given by the following equation:

$$\mathbf{f}_{ij} = \frac{\mathbf{r}_{ij}}{|\mathbf{r}_{ij}|} \cdot \frac{24\varepsilon}{\sigma} \cdot \left[2 \left(\frac{\sigma}{r_{ij}} \right)^{13} - \left(\frac{\sigma}{r_{ij}} \right)^7 \right]$$

Where ε is the minimum of the potential energy well, and σ is the finite distance where the final force reaches 0.

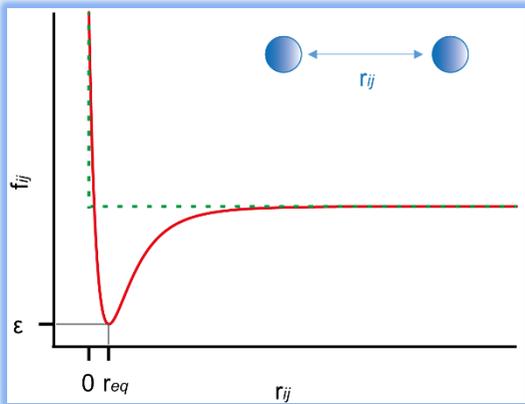


Figure 39: Hard-sphere model (dashed green lines) versus Lennard-Jones potential model (red line) of two particles of negligible radius interacting at a distance r_{ij} . The equilibrium position is achieved at r_{eq} , with a minimal energy of ε .

However, MD simulations can be adapted to mimic the properties of other ensembles. Typically, one of the most preferred ensemble is the isothermal-isobaric or NPT ensemble, with constant temperature (T) and pressure (P). In order to force the thermal equilibrium, an external bath is coupled to the system [120]. At each step, the temperature of the system is equalized to that of the bath, and the microscopic velocities are scaled according

to ΔT . In turn, in order to ensure the isobaric conditions, the volume of the system is changed until the pressure equilibrates to that of the external bath. After the volume fluctuation, the particle coordinates are scaled according to ΔT .

The first step before performing an MD simulation of a biomolecule is to choose the force field. Without going in details (force fields are implemented in MD simulation software packages), force field is the model that describes the interparticle interactions and parameters (e.g. solvent and electrostatic potential). Afterwards, the MD procedure can be performed in 4 main steps: 1) choosing the initial structure; 2) initial energy equilibration; 3) production run; 4) trajectory analysis. These steps will be briefly summarized.

Choosing the initial structure

The starting configuration is critical to ensure the quality of the MD simulation. It is necessary to select structures near the equilibrium configuration in order to avoid artefacts or failure in the MD simulation. At the same time, the initial structure must display the required structural features (e.g. the desired conformation). Commonly, initial structures of biomolecules are provided from experimental techniques, typically from X-ray or NMR. Afterwards, a solvent box is created in order to simulate the solution conditions, and the solvent molecules displaced by the molecule are eliminated. The introduction of a solvent box creates unrealistic boundaries on the system; therefore, defining the correct boundaries is also highly important to simulate the solution conditions. Finally, the initial velocities of the atoms are assigned with the use of the Maxwell-Boltzmann distribution at the desired temperature.

Initial energy equilibration

The equilibration phase allows the system to reach the equilibrium under the particular conditions of the MD simulation. This is especially relevant for initial structures obtained from X-ray structures: changing the solid state lattice environment for the solvent entails certain structural rearrangements. This equilibration procedure is monitored by different parameters, which reach a plateau value when the system is equilibrated.

Production run

The production run corresponds to the trajectories of all atoms as a function of time. All the atoms at the equilibrium configuration are supplemented with the initial velocity, and the system evolves during the production run according to the defined conditions. Production

run calculates bonded interactions (i.e. bond lengths, bond angles and bond dihedral angles) and non-bonded interactions (for instance, repulsion between atoms or electrostatic interactions). The latter involves the calculation of all the pair-wise distances between the N atoms, which scales as N^2 . These calculations are highly computationally demanding, and therefore require an approximation. Approximations establish a distance cut-off value between atoms, since non-bonded interactions strongly depend on the distance. However, establishing distance cut-offs is not trivial and depends on different nature of non-bonding interactions. The dependence of the interactions with the interparticle distance (r) spans from $1/r$ for electrostatic interactions, to $1/r^3$ for dipole-dipole interactions; several approaches have been developed to evaluate these interactions keeping the computational demand to the minimum, but a detailed description is out of the scope of this manuscript. Finally, several constraints of internal coordinates are applied in order to simplify the calculations, while the rest of coordinates remain free. One of the most spread algorithm for constraining bond geometries is the SHAKE algorithm [121].

Trajectory analysis

MD simulations of biomolecules generates huge amounts of data. Consequently, data must be processed to facilitate the analysis by the user and to monitor the quality of the run. One of the most used representations of MD trajectories is the plot of the root mean square deviation (RMSD) as a function of the simulation time. RMSD is defined as the average distance between atoms between several optimally superimposed structures (it has units of distance, typically expressed in Å). In the case of proteins, the RMSD is calculated on the α -carbon coordinates at a given time t according to the following formula:

$$RMSD(t) = \sqrt{\frac{1}{N} \sum_{i=1}^N (r_i(t) - r_i^{ref})^2}$$

In which N is the number of pairs of atoms subjected to the calculation, and r_i , r_i^{ref} are the positions of the i atom at a time t and its reference value, respectively. Another useful parameter to characterize the flexibility of the system is the root mean square fluctuation (RMSF). This parameter is the time-averaged position of a certain atom (x_i) respect a reference position (x_i^*) within all the MD simulation trajectory (T):

$$RMSF_i = \sqrt{\frac{1}{T} \sum_{t=1}^T (x_i(t) - x_i^*)^2}$$

Typically, RMSF is also expressed in distance units (usually Å). RMSF eliminates the time dependence and reflects the fluctuations of the atom in an interval of time: consequently, this parameter provides valuable structural information of flexibility.

In order to evaluate the contact of the biomolecule (or a given region of the biomolecule) with the solvent, MD simulation software packages include algorithms to calculate the solvent accessible surface area (SASA). This parameter is the area which is in contact with the solvent, and it has units of surface (typically Å²). The classical method to calculate the solvent accessibility (developed by Shrake and Rupley [122]) simplified the solvent molecules as spheres of estimate radius. The solvent sphere samples the surface grid, which is calculated according to the Van der Waals radius of the atoms. The number of points that are detected to be immediately outside the grid is related to the SASA (Figure 40). More modern approaches use the linear combinations of pairwise overlaps (LCPO) algorithm [123]. According to this algorithm, an overlap function is calculated by simplifying the solvent and a given atom as hard spheres. Afterwards, the overlap functions of all atoms are lineally combined to yield the SASA.

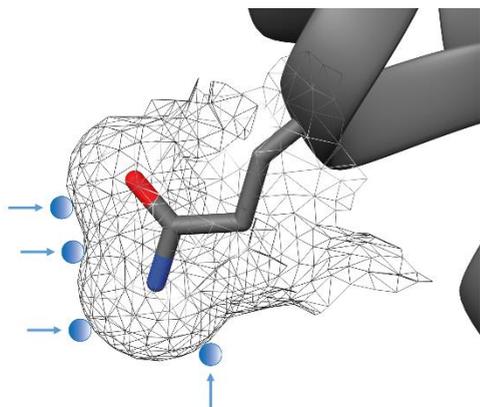


Figure 40: Example of SASA calculation by the classical method of Shrake and Rupley. The approximated solvent spheres that account for the SASA calculation (blue) are limiting with the Van der Waals grid of a glutamine residue.

STARTING STRUCTURES AND HOMOLGY MODELS OF POP

According to our experimental data, POP undergoes long-range conformational rearrangements in a slow time scale which are not accessible by MD simulations. Hence, the exploration of POP conformational space by MD simulations from a unique starting

structure might not be representative. For this reason, several MD simulations starting from different structures of free and inhibitor-bound POP were performed. Because most of the structures on the PDB were porcine, all the MD simulations used porcine POP structures instead of human. However, the structural and sequence similarity between porcine and human POP is extremely high (>97 % sequence similarity). In total, 3 MD simulations of free POP and 2 MD simulations of inhibitor-bound POP were performed.

MD simulations of free POP

With the objective of capturing spontaneous interdomain separations of free POP, we performed a low- μ s MD simulation of the closed form of free POP, named MD1. Explicit solvent was used in all MD simulations, with the TIP3P solvent model. Although there are several crystallized structures of free POP in the PDB database (e.g. 1QFM), MD1 simulation started from the X-ray closed structure of porcine POP bound with ZPP (PDB entry 1QFS) in which the inhibitor was manually removed. The choice of the starting structure was made on the basis of the low B-factor of the X-ray structure of ZPP-bound POP. However, negligible structural differences were found between the two crystalized forms (RMSD between 1QFM and 1QFS structures of 0.2 Å). Production run was carried out up to 1.8 μ s (Figure 41 A).

However, we were highly interested on having high quality models of putative open structures of free POP. Unfortunately, to date there are no available crystal structures of mammalian free POP in the open state. For this reason, we generated the porcine POP homology model of the X-ray structure of *Aeromonus punctata* POP in the open conformation (PDB entry 3IUJ). Although very different in sequence (40 % sequence similarity), porcine and *Aeromonus punctata* POPs displayed high tertiary structure similarity between their closed structures (RMSD of 1.653 Å between 1QFS and 3IVM structures). Homology modelling was performed automatically with I-TASSER server [124], using 3IUJ as a template structure.

After obtaining the porcine POP homology model, two MD simulations were performed. The first MD simulation of the free POP homology model was performed without restraints, during 510 ns (MD2, Figure 41 B). As it will be shown later, the open structure was not stable during the whole trajectory, and interdomain closing occurred in the first part of the trajectory. In order to avoid this closing, a second MD simulation was performed in which the initial structure was allowed to relax for a significant period of time (MD3, Figure 41 C). The relaxation was only allowed for side chains, while the backbone was fixed in the open

conformation by applying α -carbon restraints during the first 90 ns of the simulation. After this relaxation, α -carbon restraints were eliminated and the production run was kept up to 300 ns.

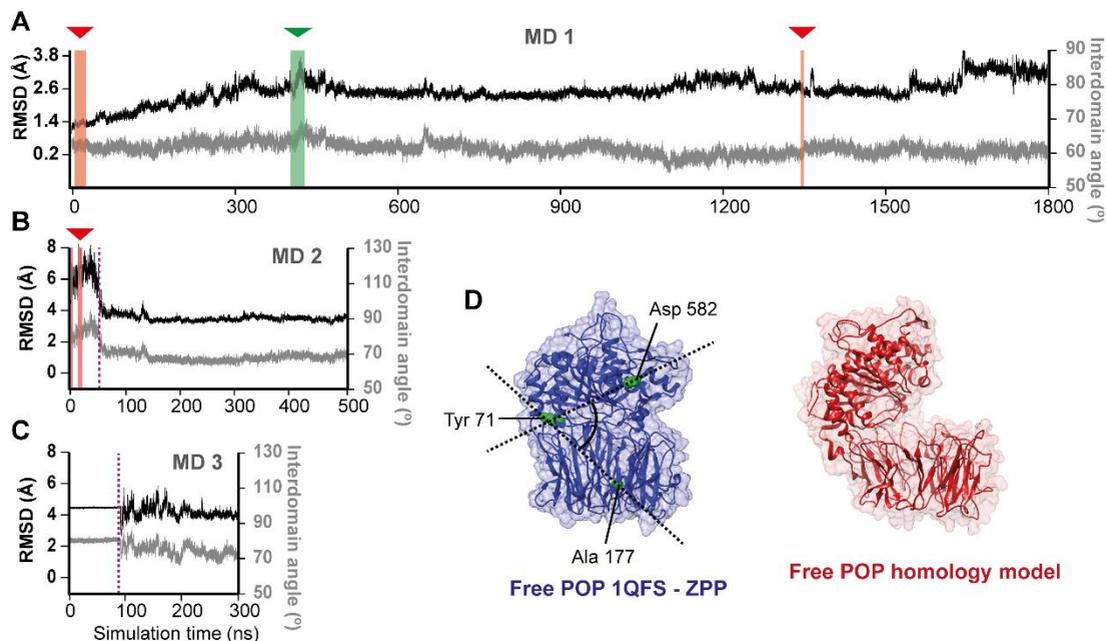


Figure 41: MD simulations of free POP. A) 1.8- μ s trajectory of MD1 simulation of the closed form of free POP. B) MD2 simulation of the homology model of free POP without constraints. Dashed line show the interval which has been selected for further data analyses (0-50 ns). C) MD3 simulation of the homology model of free POP with α -carbon constraints during the first 90 ns of the trajectory (dashed line, excluded from data analyses). RMSD (in Å) with respect to the initial structure is shown in black, and interdomain angle between residues 582, 71 and 177 is shown in gray. The sections in red and green correspond to the intervals selected by EOM of curves corresponding to peaks I and II, respectively (see chapter 2). D) The starting X-ray structure of POP in a closed conformation with the inhibitor removed is shown in blue (PDB entry 1QFS); the porcine POP homology model of *Aeromonus punctata* POP in an open conformation (PDB entry 3IUJ) is shown in red. The residues used for the determination of interdomain angle are displayed as green spheres.

MD simulations of inhibited POP

The first MD simulation of inhibited POP (MD4) started from the crystal structure of POP bound with ZPP (PDB entry 1QFS). In order to avoid the parametrization of the covalent bond between the inhibitor ZPP and the active centre Ser 554, the hemiacetal bond was

removed. However, the rest of the chemical structure of ZPP was maintained, especially the non-covalent contacts with the active site residues of POP [2]. This MD simulation was performed during 50 ns (Figure 42 A).

However, slight changes in the active centre configuration of POP might result in significant alterations of interdomain interactions which might lead to structural rearrangements. For this reason, we decided to perform a second MD simulation of inhibited POP keeping the covalent bond between the inhibitor ZPP and Ser 554 of the active centre (MD5). The covalent bond was described using Desmond software [125], by parametrizing a new residue with the side chain including the hemiacetal bond with ZPP. The production run of MD5 was of 40 ns (Figure 42 B).

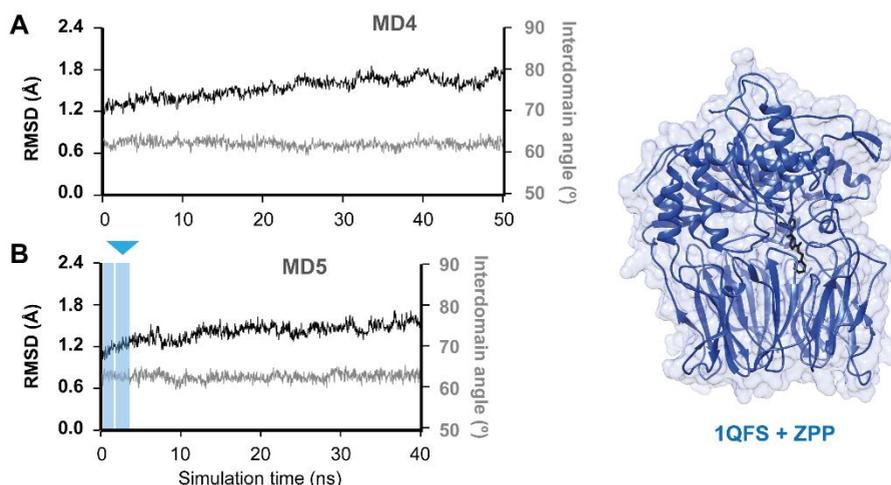


Figure 42: MD simulations of inhibited POP. A) MD4 simulation of X-ray structure 1QFS with the hemiacetal bond between Ser 554 and ZPP removed. B) MD5 simulation of 1QFS structure with ZPP inhibitor covalently bound. RMSD (in Å) is shown in black, and interdomain angle is shown in gray. Blue sections in MD5 simulation correspond to the intervals selected by EOM of the SAXS curve of peak III. 1QFS structure is shown in blue; ZPP inhibitor is depicted in black.

ANALYSIS OF MD TRAJECTORIES

The α -carbon RMSD and interdomain angle of all MD simulations were plotted for the whole simulation time. Interdomain angle allowed us to monitor interdomain separations, and it was calculated as the angle formed between α -carbons of residues 582 (α/β -hydrolase), 71 (interdomain polypeptide hinge) and 177 (β -propeller, Figure 41 D).

MD1

The long trajectory of MD1 simulation showed the evolution of the closed conformer of POP. In a previous study of Kaszuba and co-workers [11], the closed conformer of free POP was not found to undergo conformational rearrangements during a 100-ns MD simulation. However, in our case, we observed several RMSD peaks in MD1 (Figure 41 A). In order to detect if these RMSD peaks corresponded to interdomain opening, we compared RMSD and interdomain angle plots: the two first RMSD peaks (approximately at 422 and 651 ns) were correlated with peaks in the interdomain angle plot. The maximum interdomain angle achieved in these peaks were of 70° and 68°, respectively; these values were higher compared to that of the average interdomain angle of the whole trajectory ($61 \pm 2^\circ$).

A global view of the structures displaying maximal interdomain angles showed that they corresponded to a general, yet slight, interdomain separation (Figure 43 A). Moreover, the highly flexible loops A and B were displaced from their original position; this favoured the existence of a significant cavity in the interdomain loop region which exposed the catalytic centre. At the same time, a second small cavity appeared near His 680 loop and the first polypeptide hinge. The exposure of buried areas have significant impacts in the global SASA of POP: the solvent exposure also reached maximal values simultaneously to interdomain separation (Figure 43 B). The maximum SASA were of 30100 Å² and 29100 Å² for the peaks at 422 and 651 ns, respectively. Taken together, these results highlights the high solvent exposition of the semi-open POP conformers. The reiteration of two consecutive interdomain angle maximums during the same MD simulation trajectory indicate that this breathing of POP structure might occur periodically.

MD2

The first MD simulation starting from the homology model of free POP is shown in Figure 41 B. The initial part of the trajectory (50 ns) displayed a significant RMSD and interdomain angle fluctuations, but the overall value of the interdomain angle indicated that the open structure of free POP was maintained during this interval (Figure 41 B). However, the RMSD and interdomain angle values drop after this first interval of 50 ns, probably as a result of a spontaneous closing of the structure. From 160 ns to the end of the trajectory (510 ns) only the closed conformation was exclusively present. The interdomain angle of the closed conformer generated in the second part of the trajectory was found to be higher compared to that of MD1 simulation ($67 \pm 6^\circ$ vs $61 \pm 2^\circ$, respectively); nevertheless, this difference was attributed to the differences in the starting structures.

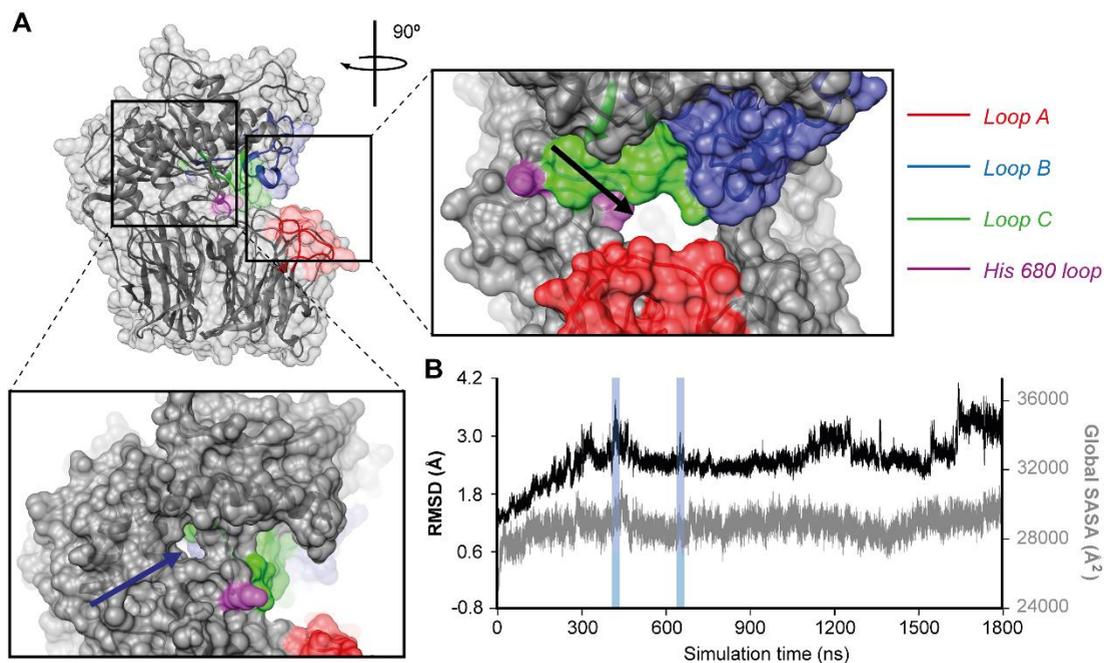


Figure 43: Partial opening of POP structure in MD1 simulation. A) Structure at 422 ns, corresponding to the maximum interdomain separation of the trajectory. Loops are coloured according to the colour code described at the right. Insets show magnified views of the pores that appear by displacement of loop A and His 680 loop (black and blue arrows, respectively). B) Peaks of maximum interdomain angle at 422 and 651 ns are correlated with global exposition to the solvent (blue regions).

MD3

The second MD simulation starting from POP homology model was carried out including α -carbon restraints during the first 90 ns of the trajectory in order to allow side chain relaxation. Obviously, RMSD and interdomain angle during this relaxation interval remained fixed and were not considered in further analyses. After the elimination of the α -carbon restraints, RMSD and interdomain angle underwent marked fluctuations, which reflected the high degree of flexibility of the open structure (Figure 41 C). Of interest, the interdomain angle (Figure 44 A) and the global SASA (32200 \AA^2 , compared to 28900 \AA^2 averaged for MD1 simulation, Figure 44 B) were coherent with a long-range interdomain separation. Together, these results indicated that the homology model of the relaxed open conformation was relatively stable. The initial side chain relaxation of the homology model structure was found to be a critical factor for the stability of the open conformer of free POP in MD simulations.

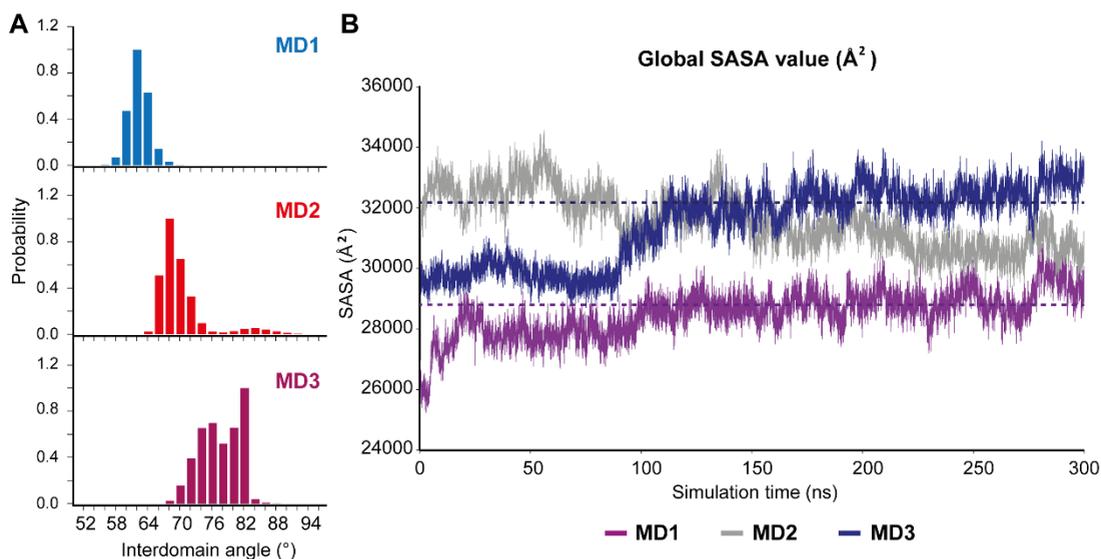


Figure 44: Comparison of MD1, MD2 and MD3 simulations indicate large interdomain separations of free POP. A) Distribution of interdomain angles during the whole trajectories of MD1, MD2 and MD3 simulations. B) Global SASA plots of the first 300 ns of MD1, MD2 and MD3 simulations. Average SASA values of MD1 (28900 Å², purple dashed line) and MD3 (32200 Å², blue dashed line) are shown.

MD4 and MD5

The two MD simulations of inhibited POP shared closely similar features. By comparing the RMSD and interdomain plots, both trajectories reflected high structural stability and poor flexibility (Figure 42 A and B). During the initial equilibration period, the RMSD slightly of MD4 and MD5 simulations increased to a plateau approximately at 1.6 Å and 1.5 Å, respectively. At this point, only thermal motions of small amplitude occurred during the rest of the production run. Interdomain angles remained constant during the whole trajectory, indicating the absence of interdomain separations. In spite of the high degree of similarity between the two MD simulations, only MD5 simulation was chosen for further studies due to the presence of the covalent bond between the inhibitor and Ser 554 of the active site. Although this characteristic did not affect the general evolution of the structure during the MD simulation, it might influence the local configuration of the active site environment. This would lead to wrong conclusions when analysing the active site configuration at atomic detail.

EXPERIMENTAL DATA ANALYSIS

At this point we combined the high-quality experimental data with the theoretical information obtained by the MD simulations. Due to the different experimental methodologies used in this PhD thesis, we performed different data analyses for SAXS and NMR methods.

Ensemble optimization fitting of theoretical SAXS scattering profiles

As explained in Appendix II, SAXS is an experimental technique which allows the direct measurement of low angle X-ray scattering profiles of biomolecules in solution. The scattering profiles contains valuable information about the low-resolution structural features of the biomolecules in diluted solutions without any external perturbations (if no X-ray radiation damage occur). Moreover, SAXS accepts a wide range of molecular sizes and structural complexity. However, the structural information that can be extracted directly from experimental data is limited only to $I(0)$, R_g , and qualitative information from Porod and Kratky plots. For this reason, we performed the fitting of theoretical SAXS profiles obtained from the structures of the MD simulations with the experimental ones. **This study was carried out in collaboration with Dr. Pau Bernadó and Fátima Herranz-Trillo (Centre de Biochimie Structurale (CNRS), Montpellier, France).**

The calculation of theoretical SAXS profiles from computational structures was carried out with CRY SOL program [106], implemented in ATSAS data analysis software. The scattering profiles were calculated up to $s = 0.5 \text{ \AA}^{-1}$, with a total of 201 data points. The parameters of CRY SOL program were a maximum order of 30 harmonics and of 20 Fibonacci grids. Given the extreme number of structures generated by MD1, MD2, MD3 and MD5 simulations, the generation of theoretical SAXS curves for all structures would have been tedious and largely time-consuming. For this reason, only certain structures were periodically sampled and processed by CRY SOL program (Table 7 and Figure 41). The pool of free POP frames was generated by joining the frames obtained from MD1, MD2 and MD3 simulations as described in Table 7; in turn, the pool of ZPP-bound POP frames was formed exclusively by those calculated from MD5 simulation.

MD simulation	Production run interval (ns)	Sampled structures
MD1	0-1800	1 of every 20
MD2	0-50	All
MD3	90-300	1 of every 2
MD5	0-40	1 of every 2
TOTAL GENERATED FRAMES		14785

Table 7: Intervals and structures from MD simulations sampled by CRYSQL program. Exactly the same number of simulated SAXS scattering profiles were obtained.

The fitting of the theoretical curves with experimental scattering profiles by the ensemble optimization method (EOM) was carried out by Fátima Herranz, and under the supervision of Dr. Pau Bernadó.

First of all, a preliminary fitting of of free POP curves with the averaged curve of peak I of SEC-SAXS $I(\theta)$ chromatogram was carried out according to the static model. This model performs a χ^2 test between pairs of curves, which indicates the degree of similarity between curves (Figure 45). This fitting was performed separately with the averaged SAXS curves of each MD simulation and the experimental curve of peak I, in order to estimate the predominant conformation present at peak I. The preliminary static model fitting of averaged SAXS curves suffered from poor quality: all the obtained χ^2 values were not acceptable (0.282, 0.396 and 0.131 for MD1, MD2 and MD3 simulations, respectively), which indicated that the experimental curves did not correspond to a single structure. Rather, they consisted on averaged of curves of well-differentiated conformers.

For this reason, a more robust methodology was used in order to extract the ensemble of conformations that collectively described the SAXS experimental data: specifically, the EOM was chosen [107, 116]. EOM is an optimization method based on a genetic algorithm especially developed to characterize flexible proteins from SAXS scattering profiles (Box 4)

Box 4: Genetic algorithm: a genetic algorithm is a process governed by certain operators and conditions, which allows the solution of a given complex function. The operations and parameters of this process are inspired on the evolutionary laws of biology, i.e. generations involving crossings and mutations under a certain criterion. For this reason, these algorithms are known as genetic algorithms.

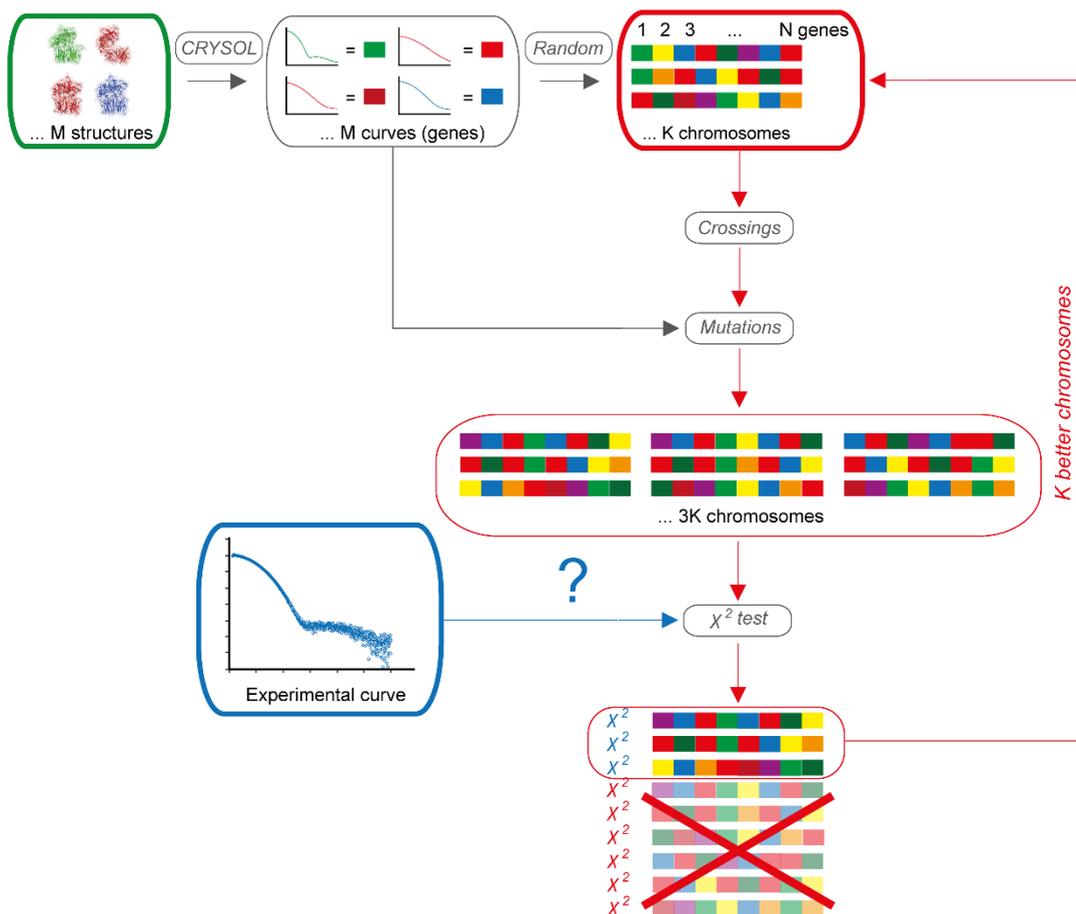


Figure 45: EOM fitting of theoretical SAXS scattering profiles. A pool of M conformers was generated by MD simulations (green box). The corresponding scattering profiles were obtained with CRYSQL software. Curves were randomly selected to generate an initial sub-ensemble of curves (genes) containing K chromosomes formed by N genes (red box). The EOM cycle, or generation (red arrows), starts from this sub-ensemble, performing crossings between chromosomes and mutations from genes of the initial pool. This yield a population of $3K$ chromosomes; each chromosome is compared to the experimental SAXS curve by a χ^2 test. Only the K chromosomes with smallest χ^2 are chosen for further generations.

Briefly, EOM optimizes a small ensemble of conformers obtained from a vast pool of M conformers by using the χ^2 with experimental curves as the driving force of the optimization process (Figure 45). The number of conformers M must be much greater than N , which is assumed to be the minimal number of conformers necessary to describe the experimental

data. In other words, if $I(s)$ is the SAXS profile as a function of the momentum transfer (s), the N -average $I_{theor}(s)$ must be closely similar to $I_{exp}(s)$:

$$I_{exp}(s) \cong I_{theor}(s) = \frac{1}{N} \sum_{n=1}^N I_{theor,n}(s)$$

In particular, EOM uses the precomputed SAXS curves in the algorithm instead of using the structures directly. The pool of curves obtained by MD was formed by 13766 frames in the case of free POP (MD1 + MD2 + MD3 simulations, Table 7) and 1019 frames in the case of ZPP-bound POP (MD5 simulation); these numbers thoroughly accomplish $M \gg N$.

According to the EOM nomenclature, a *gene* is a system which has to be resolved. In this case, genes are the SAXS scattering profiles corresponding to individual structures. N genes are combined to give a *chromosome*, which correspond to the average $I_{theor}(s)$. The EOM starts with the random generation of K chromosomes (Figure 45). Afterwards, different cycles of genetic operations inspired in the evolutionary laws of biology take place in order to optimize these chromosomes: each cycle is known as a *generation*.

The genetic operations used in EOM are *mutations* and *crossings*. Mutations change genes from the rest of $K-1$ chromosomes and from the pool of genes (the proportion of mutations allowed at each generation, as well as the ratio of genes changed between chromosomes and the pool, are determined by the user). In turn, crossing operations exchanges a minimum of two genes between pairs of chromosomes.

After one generation, K' chromosomes are obtained ($K' > K$). The quality of the fitting is calculated at this point by minimizing the discrepancy between the theoretical average and experimental curves, carried out by the χ^2 test:

$$\chi^2 = \frac{1}{K-1} \sum_{j=1}^K \left[\frac{\mu \cdot I_{theor}(s_j) - I_{exp}(s_j)}{\sigma(s_j)} \right]^2$$

The K' chromosomes are ordered by increasing value of χ^2 , and the first K chromosomes with higher similarity (lower χ^2) are selected for further generations.

Typically, several hundreds of independent EOM algorithms are performed. The optimized sub-ensembles of conformations are compared between EOM calculations. The EOM is considered successful if all the independent optimizations show a high degree of similarity.

Finally, the most frequent chromosome from all these independent optimizations is chosen as the optimal result.

EOM requires few user inputs; however, some parameters are critical for the quality of the final result. The most important parameter is N , which is related to the flexibility of the system. The inappropriate setting of this value might lead to incomplete results. N must be high if the system is largely flexible, especially in the case of intrinsically disordered proteins (a value of $N = 50$ is recommended for unfolded systems). However, the conformational fluctuations expected for POP are restricted to interdomain separation and local loop flexibility. For this reason N was set at 20; the use of $N = 50$ was discarded because it did not improve the quality of the result. The typical number of chromosomes K is 50, which was also selected in our case. Each EOM run consisted in 1500 generations, and one hundred independent EOM runs were performed.

EOM of free POP peaks I and II

The EOM of frames from MD1 + MD2 + MD3 simulations was performed using the similarity with the experimental curve of peak I of free POP as a driving force. The fitting of these curves was carried out for data points with $s < 0.15 \text{ \AA}^{-1}$. An excellent fitting was obtained with the EOM (χ^2 of 0.021, Figure 46 A), indicating the suitability of this method to describe experimental curves of flexible proteins. The EOM selected 45 % of the frames exclusively from MD1 and 55 % from MD2 (Figure 41 A and B). The selected frames corresponded either to completely open or completely closed forms, with averaged separation angles of 86° and 62° and theoretical R_g values of 29.17 \AA and 26.03 \AA for the open and closed forms, respectively. These results are coherent with the R_g extracted directly from the experimental curve ($R_g = 28.50 \pm 0.06 \text{ \AA}$), and shows that free POP in solution exists in a dynamic equilibrium between fully open and closed forms.

In case of peak II of free POP $I(0)$ chromatogram, the low intensity of the scattering profile obtained by SVD was a serious limitation for the success of the EOM. Given the low R_g value associated to this peak ($25.00 \pm 0.09 \text{ \AA}$), it is unlikely that open structures are present in peak II. For this reason, we performed a preliminary EOM of frames corresponding only to MD1, restricted to the initial part of the curve ($s < 0.1 \text{ \AA}^{-1}$). This optimization provided a good fitting (χ^2 of 0.547, Figure 46 B), and all 20 selected curves were centred approximately at 422 ns of MD1 simulation (Figure 41 A). This region corresponded to the maximum interdomain separation occurring in this simulation. Analysis of the structures around this point revealed side openings caused by the loosening of loops A, B, C, and His

680 loop, which resulted in slight interdomain fluctuations (average interdomain angle of 64° , Figure 43 A). However, the inconsistency between the R_g values derived from the experimental and the theoretical curves ($25.00 \pm 0.09 \text{ \AA}$ vs 27.10 \AA) did not support the quality of the EOM, which was severely compromised due to the low intensity of the experimental curve. Consequently, the EOM of peak II was discarded.

EOM of ZPP-bound POP peak III

The curves obtained from MD5 simulation were subjected to EOM with the experimental curve of peak III of ZPP-bound POP as a driving force. Given the high quality of the experimental scattering profiles, this analysis was performed for data points with $s < 0.3 \text{ \AA}^{-1}$. Again, an excellent fitting was obtained between the experimental and the averaged theoretical curve selected by the EOM (χ^2 of 0.054, Figure 46 C); in particular, all the 20 theoretical curves selected by the EOM closely resembled the initial 1QFS X-ray structure of inhibited POP (maximum RMSD value of 1.226 \AA , Figure 42 B). Therefore, this result shows that ZPP-bound POP exists in solution in a highly stabilized closed conformation.

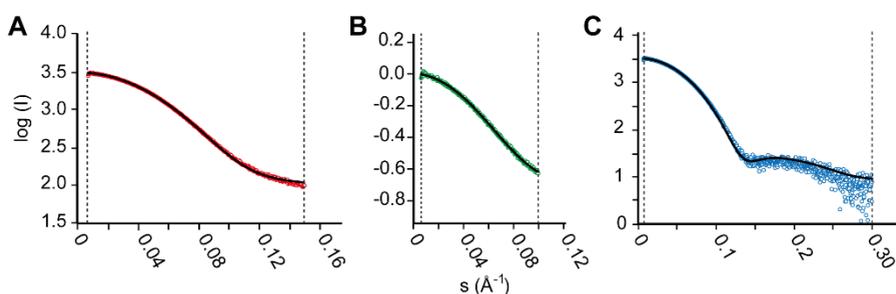


Figure 46: Fitting of theoretical SAXS curves to experimental scattering profiles by EOM. Fitted curves are shown as black curves for peak I of free POP (A), peak II of free POP (B), and peak III of ZPP-bound POP (C).

High resolution structural analysis of POP conformers

Once we obtained the sub-ensembles of high resolution structures that describe the SAXS data of free and ZPP-bound POP, we proceed to the fine structural analysis of these structures. We focused our efforts on the study of the open conformation of free POP; under our knowledge, this was the first time that experimental data explicitly pointed towards an open structure of mammalian POP. The hybrid approach between *bona fide* MD simulations of the homology model together with high quality experimental data encouraged us to study the structure of the open conformation of POP at atomic level. Therefore, the 11 open

structures selected by EOM of peak I were analysed and compared with the 20 closed structures selected for peak III. Particularly, we selected the ZPP-bound POP as a model for the closed conformation due to the high stability of the structure [11].

The analysis of the interactions between residues (salt bridges and HB) was performed with the default parameters of VMD software [126]. Weak salt bridges and hydrogen bonds present in less than 15 % of the structures were not taken into account. Finally, only the interactions corresponding to the interdomain region and to the flexible loops (loop A, B, C, and His 680 loop) were taken into account, since the rest of the elements were not expected to participate in the interdomain opening.

According to this analysis, we observed that the open state of free POP presents 12 % less of stabilizing interactions between the two domains and the flexible loops compared to ZPP-bound POP. In order to carry out a visual inspection of the change in the pattern of stabilizing interactions as a consequence of interdomain opening, we generated the intramolecular interaction matrix. Given that half of the matrix is redundant, we coupled half matrixes corresponding to the open and closed conformation. The mixed matrix allowed a rapid visual appreciation of the effects of interdomain opening in the stabilizing interactions (Figure 47). The most relevant observations are summarized below:

1. Stabilizing HB interactions between loop A and loop B are present only in the closed ZPP-bound POP. These interactions occur between Ser 203/Asn 205 - Thr 590 and between Asn 205 - His 593, indicating that these regions are tied in the inhibited form. This result differs to what described in the work of Kaszuba and co-workers, which described a spontaneous detachment of loop A in the MD simulation of POP bound with ZPP [11]. As it will be seen in the next section, the release of loops A and B in the open conformation have significant repercussions on the dynamics of these regions.
2. Although most of stabilizing interactions of loop A are broken in the open conformation, loop B conserves more interactions in the open conformation; the HB between Asp 598 - Leu 520 and the salt bridge between Asp 598 - His 515/Lys 523 are preserved in both conformations. These interactions connects loop B with the α/β -hydrolase in the open conformation.

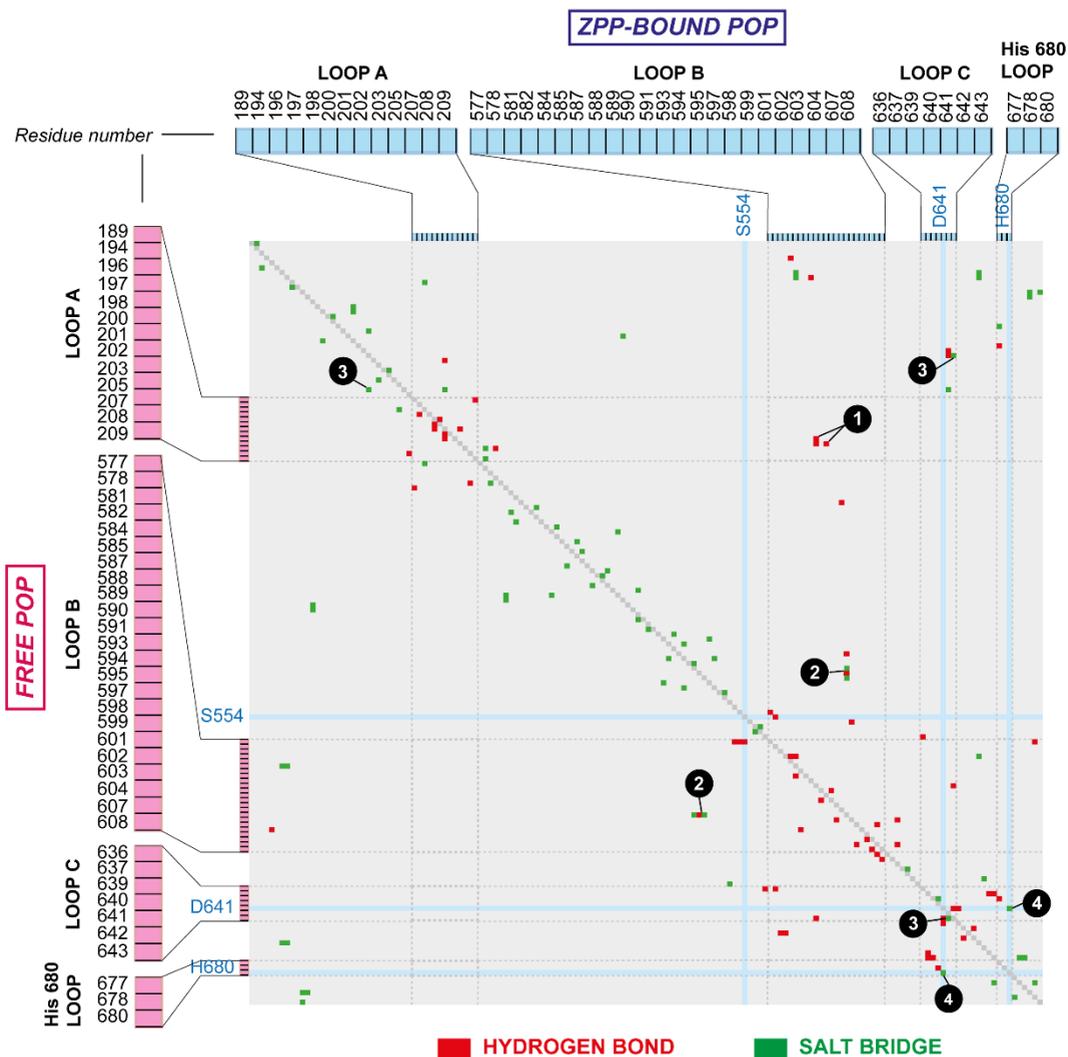


Figure 47: Mixed intramolecular interaction matrix of free POP in the open conformation (left axis) and the closed form of POP bound with ZPP (upper axis) derived from the sub-ensembles of conformations selected by EOM from MD2 and MD5, respectively. Only residues of loops A, B, C, His 680 loop and Ser 554, Asp 640 and His 680 have been taken into account. Hydrogen bonds and salt bridges are described according the colour code shown in the bottom of the figure.

3. As a consequence of interdomain separation, the key interdomain salt bridges between Asp 149 (β -propeller) - Arg 643 (loop C) and Lys 172 (β -propeller) - Asp 642 (loop C) in the closed conformation of ZPP-bound POP are permuted by two intradomain pairs in the open form of free POP. Hence, Asp 149 forms a salt bridge with the neighbouring Lys 172 (both in loop C), while Asp 642 interacts with Arg

643 (in the β -propeller). This permutation probably has important repercussions in POP catalytic activity (see below).

4. The salt bridge between Asp 641 (loop C) and His 680 (His 680 loop) of the catalytic triad is present in the open and the closed conformations. The preservation of this interaction indicated that loop C and His 680 are connected and move synchronously, independently of the conformational state. This interaction is highly relevant for POP hydrolytic activity.

Taken together, these observations provide valuable information about the role of interdomain opening in POP catalytic activity. As it was concluded in the first crystallographic studies of POP by Fulop and co-workers, the conformation of the imidazole ring of His 680 determines the catalytic activity of the enzyme [2, 10, 13]. According to our results, we also observed in MD simulations that the active centre of POP bound with ZPP was frozen in the transition state conformation. His 680 displayed the distance and orientation with respect to the nucleophile Ser 554 required for catalytic activity (distance between N ϵ 2 of His 680 and O γ of Ser 554 of 3.29 Å, Figure 48).

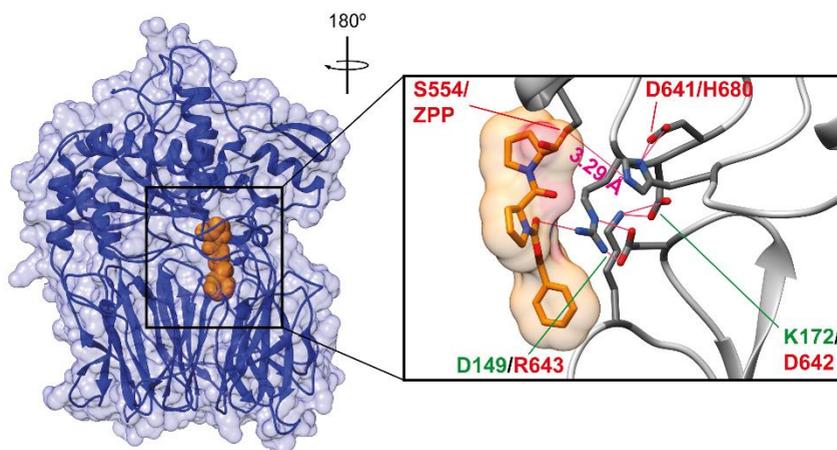


Figure 48: Representative closed structure of POP bound with ZPP selected by EOM of MD5. ZPP inhibitor surface is depicted in orange. Inset box shows the catalytic triad in the transition state configuration; the distance between N ϵ 2 of His 680 and O γ of Ser 554 is shown in magenta. Residues responsible for the active conformation of the active site are also shown. Residues labeled in red correspond to α/β -hydrolase; residues labeled in green, to β -propeller.

However, this scenario is deeply altered if interdomain opening takes place. As a consequence of interdomain separation, the interactions between certain residues of the β -propeller and loop C are broken. Specifically, the salt bridges between Asp 149-Arg 643

and Lys 172-Asp 642, which are present in the closed conformation, are permuted by two intradomain interactions in the open conformation (Asp 149 – Lys 172 and Asp 642 – Arg 643, Figure 49). Losing the interdomain interactions results in a rearrangement of loop C to an occluded conformation towards α/β -hydrolase domain. Since loop C and His 680 loop are connected through the highly stable salt bridge between Asp 641 and His 680, the new conformation of loop C is transmitted to His 680 loop in a mechanism similar to that of *Trypanosoma brucei* oligopeptidase B [53]. The new configuration of His 680 of the catalytic triad displays an inappropriate distance and orientation with respect to Ser 554 (distance between N ϵ 2 of His 680 and O γ of Ser 554 of 6.04 Å). Hence, these observations support the involvement of interdomain motions in switching of POP catalytic activity between the open and closed forms.

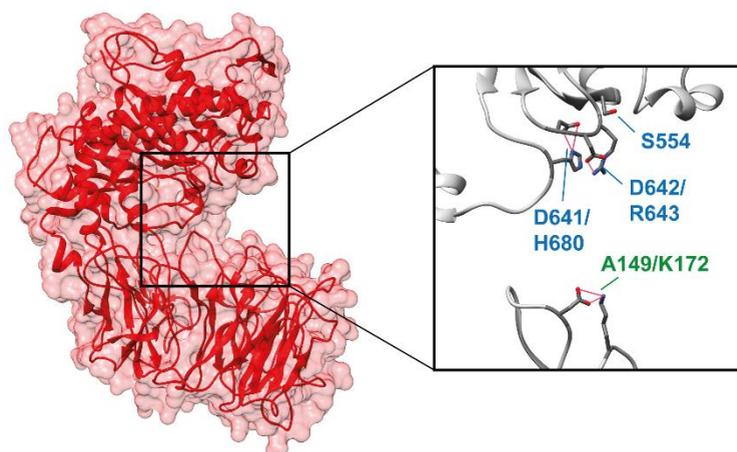


Figure 49: Representative open structure of free POP selected by EOM of MD2. Inset box shows a magnified view of the intradomain salt bridges formed after the separation of α/β -hydrolase and β -propeller domains. Residues labeled in blue correspond to α/β -hydrolase; residues labeled in green, to β -propeller.

Correlation with NMR data

^1H - ^{13}C methyl-TROSY HMQC RD and $R_{1\rho}$ RD measurements provided valuable kinetic and thermodynamic information (k_{ex} , p_A and p_B) about the conformational equilibrium of free POP. However, only qualitative structural insights can be deduced by comparing the frequencies between the two conformers in equilibrium ($\Delta\omega$) extracted by independent fitting to the Carver-Richards equation (Table 6). Similarly, analysing the CSP values between free and inhibitor-bound POP discloses structural changes derived from inhibitor

binding. In both cases, the α/β -hydrolase was clearly affected by the open/close equilibrium (Figure 19), but the exact structural transitions cannot be inferred from the NMR data alone.

However, the complementation of NMR data with MD simulations allowed us to extract structural events which contributed to the changes in Met resonances and dynamics. Since the chemical environment of Met residues is the main factor which determines the chemical shift of NMR signals, the analysis performed here was carried out by considering two local parameters of Met residues: backbone RMSF and methyl SASA.

Root mean square fluctuation

Although MD simulations time scale fall in a different time scale with respect to CPMG and $R_{1\rho}$ RD experiments, fast and slow dynamics can be coupled to some extent. For this reason, we compared the fast dynamics obtained by MD with the slow μ s-ms dynamics derived by experimental NMR techniques. α -carbon RMSF of Met residues were calculated and averaged for all free POP MD simulations (MD1, MD2 and MD3).

Figure 50 shows the backbone RMSF values of Met residues compared to the $R_{2, eff}$ and ΔR_{ex} for $R_{1\rho}$ and ^1H - ^{13}C methyl-TROSY HMQC RD measurements, respectively. As it is seen, the RMSF of Met residues does not correlate with ^1H - ^{13}C methyl-TROSY HMQC RD data. However, in the case of the faster $R_{1\rho}$ RD experiments, residues Met 581 and 583, and to a minor extent, Met 235, are partially correlated. Of interest, this result further validates the high flexibility of loop A (adjacent to Met 235) and loop B (including Met 581 and 583) in dynamics events at ns and μ s range.

In order to detect the effect of interdomain opening on fast dynamics, we compared the backbone RMSF of the first 40 ns of MD1 vs MD2 for all residues. This analysis disclosed a notably higher distribution of backbone RMSF amplitudes of the open conformation (Figure 51 A). This might be a consequence of the increased degree of freedom of the open conformation, which presents fewer stabilizing interactions compared to the closed one. The structural distribution of backbone RMSF of MD2 again showed the high flexibility of loop A, B, C and His 680 loop in the open conformation of free POP (Figure 51 B). The presence of faster dynamic events might be one of the causes of the failure of the global fitting of ^1H - ^{13}C methyl-TROSY HMQC RD data to the Carver-Richards equation. Finally, the backbone RMSF of the closed structure of POP bound with ZPP (MD5 simulation) was also calculated. The distribution of RMSF values was similar to that of MD1, but slightly displaced to lower values. This result indicated the lower contribution of the closed conformations of free and ZPP-bound POP to fast ns flexibility, in contrast to the open conformation.

Solvent accessible surface area

Changes in the solvent exposition of methyl groups of Met residues might result in significant variations on their chemical shifts. For this reason, the methyl SASA of Met residues were calculated for all MD simulations. Of interest, the methyl SASA of MD1 trajectory disclosed periodic fluctuations of high amplitude for residues Met 235 and 696 (Figure 52 A). These oscillations in the solvent exposure pointed that these residues have tendency to explore distinct environments in the closed conformation of free POP.

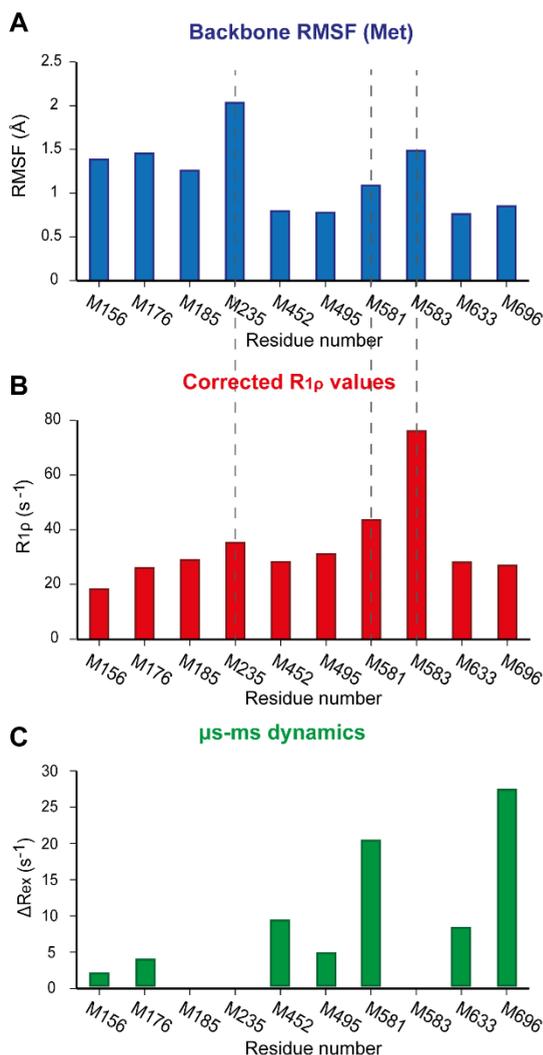


Figure 50: Correlation between fast dynamics calculated from MD simulations and slow dynamics obtained from NMR experiments. A) Backbone RMSF of Met residues averaged from MD1, MD2 and MD3 simulations. B) Fast μ s dynamics of Met residues obtained from $R_{1\rho}$ RD experiments. Residues Met 581 and 583 (loop B), and to a minor extent Met 235, are partially correlated with ns flexibility (dashed gray lines). C) μ s-ms dynamics obtained from ^1H - ^{13}C methyl-TROSY HMQC RD measurements.

In order to estimate the effects of interdomain opening in the local environment of Met residues, methyl SASA of Met were averaged for the whole trajectories: in the case of the

open conformation of free POP, SASA values from MD3 simulation were averaged; in turn, SASA values from MD5 were averaged for the closed conformer of ZPP-bound POP. Figure 52 B shows that Met 156, 235 and 696 underwent the highest increase in their solvent exposure as a consequence of interdomain opening. This significant change in the solvent accessibility might be one of the main causes for the high CSP observed in residues Met 196 and 696 in the NMR spectrum of [methyl- ^{13}C]-Met labeled POP upon inhibitor binding (the high CSP of Met 235 is originated by the direct interaction with the inhibitor, Figure 2 A).

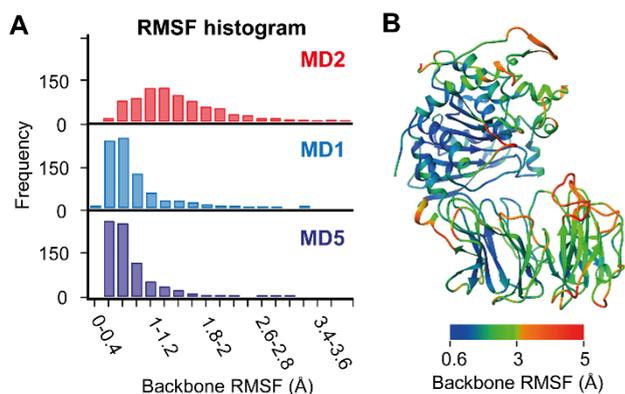


Figure 51: Fast dynamics of different POP conformations. A) RMSF histograms of the first 40 ns of MD1, MD2 and MD5 simulations. B) Structural distribution of backbone RMSF of the whole MD2 simulation trajectory.

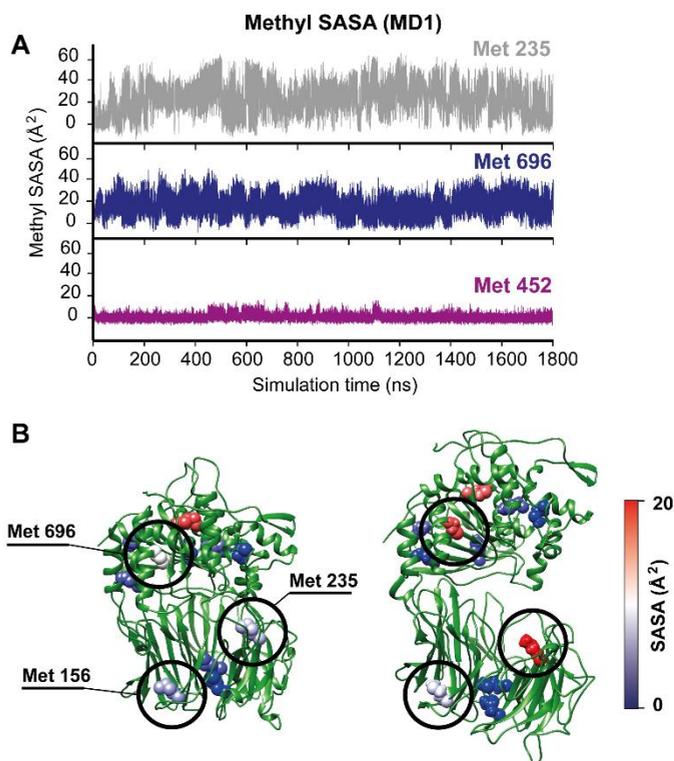
DISCUSSION

In summary, NMR and SAXS experimental data complemented by MD simulations of different POP structures showed that free POP exists in a μs -ms equilibrium between completely open and closed conformers. According to our experimental data, this long-range opening and closing transition consists in a composition of several motions of different amplitudes rather than a single hinge motion. In contrast, inhibitor-bound POP appears exclusively in a closed conformation.

Overall, the analysis of MD simulations correlated with experimental data stressed the highly dynamic nature of free POP at different time scales. MD simulations also showed that interdomain opening have profound repercussions in POP structure, exposing the active site and other buried areas to the solvent. MD simulations indicated faster events involving loops surrounding the active site. The detailed analysis of these unstructured regions disclose the mechanistic role of residues Asp 149 – Lys 172 and Asp 642 – Arg 643 in the active configuration of the catalytic center of POP. Of interest, the active configuration of

the active site only occurred in the closed conformation, pointing that interdomain separation switches between active/inactive conformations.

Figure 52: Exposition of methyl groups of Met residues to the solvent. A) Marked oscillations of methyl SASA of Met 235 and 696 during MD1 simulation. For an easier comparison, methyl SASA of the buried residue Met 452 has also been included. B) Methyl SASA averaged for MD5 and MD3 simulation (closed and open conformations, respectively). Interdomain opening especially affect the exposition of Met 196, 235 and 696.



From a thermodynamic point of view, the increased conformational entropy observed in MD simulations of the highly dynamic open form could partially compensate the enthalpic penalty caused by the disruption of the stabilizing interactions of the closed form. From a functional point of view, the conformational heterogeneity of free POP might be crucial for completing the enzymatic cycle. In this regard, the catalytic turnover of POP obtained by activity assays had a k_{cat} value of $35.0 \pm 3.6 \text{ s}^{-1}$ using Z-Gly-Pro-AMC as substrate (a k_{cat} of $32.5 \pm 1.2 \text{ s}^{-1}$ has been reported using Z-Gly-Pro-Nap as substrate [127]). These turnover numbers are agreement with the k_{ex} values summarized in Table 6. This observation suggest that the opening and closing mechanism of POP could be coupled to the catalytic cycle, and might be the rate limiting step of the hydrolysis reaction.

Chapter 4: Application of ion mobility mass spectrometry in the study of POP conformational dynamics

INTRODUCTION: MASS SPECTROMETRY IN STRUCTURAL BIOLOGY

The development of biophysical techniques for the study of dynamic systems strongly contributes to the progression of structural biology. Among the repertoire of structural elucidation techniques, mass spectrometry (MS) has experienced one of the most outstanding evolution during the past decades [128]. The recent developments and applications on this field are expanding the reliability and versatility of MS, pushing the limit of mass analysis and quantification.

The development of soft ionization methods, such as electrospray ionization (ESI) [129] and nanoflow electrospray ionization (nanoESI) [130], was one of the most important milestones in MS over the past 20 years. These methods opened the door to native MS experiments in structural biology [131, 132]. For instance, hydrogen-deuterium exchange allowed the study of slow dynamic events of complex systems [133], and tandem MS with collision-induced dissociation (CID) became a reliable method for the determination of the stoichiometry of protein complexes [134, 135]. One of the most promising methodologies for the structural characterization of biomolecules is ion mobility mass spectrometry (IMMS) [21, 136].

Soft ionization techniques

Soft ionization was crucial to generate gaseous protein ions for native MS. Both ESI and nanoESI benefit from several important features. First, they allow the analysis of relatively low protein concentrations and tolerate the use of aqueous buffers with a wide range of compositions. Second, they preferentially generate low charge states. Finally, they are less aggressive to the non-covalent dipolar and hydrophobic interactions during the ionization process.

ESI consists in a tip held at high electric potential which circulates the analyte solution (Figure 53) [129]. Charged ions are accumulated at the exit of the tip, with an overall

charge of opposed sign to the applied potential. Simultaneously, an inert gas stream generates charged droplets of r radius at ambient pressure and temperature. According to the charge residue model, the electrostatic repulsion of surface charge q_{Ry} can be equal to the surface tension (σ) at a certain limit [137, 138]. This limit is known as the Rayleigh limit, given by the following equation:

$$q_{Ry} = 8\pi\sqrt{\epsilon_0\sigma r^3}$$

Where ϵ_0 is the electric permittivity in vacuum. At this point, the droplet blow up into smaller droplets, and this process is repeated successively until gaseous protein ions are finally obtained. In the case of nanoESI, a microscopic tip is used instead of the capillary. The overall mechanism is essentially the same, but the use of a smaller tip minimizes sample consumption, and lower voltages and gas flows are required [130].

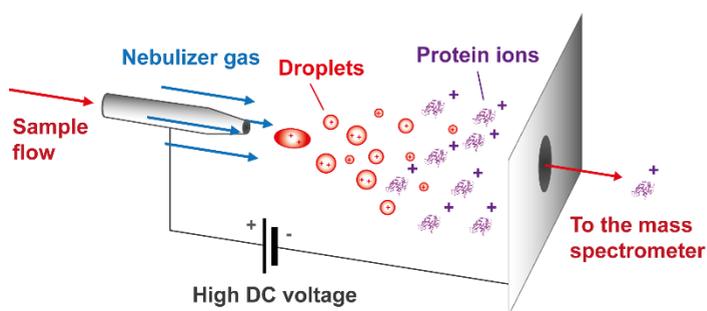


Figure 53: Soft ionization techniques: ESI ion source. The analyte solution circulates through a tip under a high electric potential. Droplets are sprayed by a stream of gas, and charged protein ions are generated. These ions are driven to the vacuum regions of the mass spectrometer.

Protein ions in the gas phase

Nevertheless, the physics underlying the last stages of the ionization process are not completely understood. This raises the question of to what extent the native solution structures are maintained in the gas phase. Gaseous protein ions are in a non-native environment: the energetic balance of attractive intramolecular interactions, charge self-solvation, and repulsive electrostatic forces are dramatically altered in vacuum [139]. Specifically, electrostatic and dipolar contributions are strengthened in the absence of solvent, while hydrophobic forces are weakened [140]. Hence, metastable protein ions generated during the fast ionization (in the time scale of ns) evolve during time in order to adapt to this new solvent-free environment. In general, the preservation of intramolecular weak interactions in the gas phase ensures the native folding.

A model for the temporal evolution of ions during ionization was proposed by Breuker and co-workers, after the analysis of several published studies (Figure 54) [141]. According to

this model, the first ns of the ionization process involve the loss of the monolayer of water molecules, accompanied by a decrease of the microscopic thermal energy (or “cooling”) of the ion. Water molecules surrounding charged groups are the last to be eliminated, favouring solvent shielding of surface charges and minimizing Coulomb repulsion (see below) [142]. During few picoseconds after the elimination of all water molecules, charged groups minimize the electrostatic repulsion by collapsing with other neighbouring charged residues of the opposite sign, or interacting with the backbone [143].

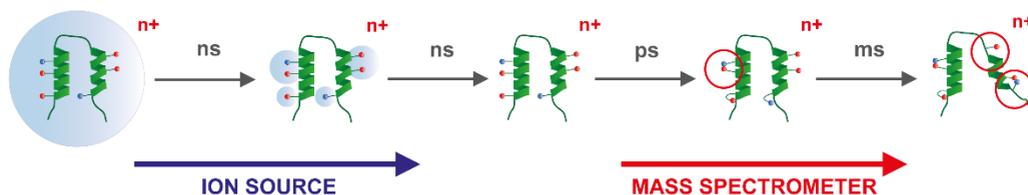


Figure 54: Evolution of protein structure during a MS experiment. Desolvation takes place during the first ns of the experiment, yielding anhydrous protein ions. The surface charged residues of these dehydrated structures rapidly collapse with other with opposite sign or with the backbone (few ps). Afterwards, conformational rearrangements occur in order to adapt the energetic balance of the gaseous environment; this takes place inside the mass spectrometer, during the experiment (ms). Ultimately, gas-phase stable structures are obtained, which are stable up to minutes in vacuum.

Once protein ions have been transferred to the gas phase, they undergo slower structural rearrangements to yield gas-phase stable structures. These rearrangements strongly depends on the charge state of the ions. A study of gaseous ubiquitin ions showed that low charged native ions were stable during a period of 100 ms, while highly charged states rapidly unfold (<10 ms) [144].

The main reason for the unfolding at higher charge states is the increased Coulomb repulsion that dominates over stabilizing internal forces in the absence of solvent shielding [139, 145]. Since the Coulomb interaction depends on $1/r^2$, the repulsion is minimized by increasing the distance between charged groups. This causes the expansion of the protein secondary structure in the gas phase, which leads to different degrees of unfolding; in some published studies, this unfolding has been compared to the acid denaturation of highly protonated structures in solution [146]. In the case of non-covalent protein complexes, Coulomb repulsion might ultimately lead to the destruction of the complex by stripping of one or more of the monomers [147].

In the opposite way, lower charge states can favour side chain collapse. The excellent work of Clemmer and Jarrold with coupled IMMS devices allowed the monitoring of structural

changes in gaseous protein ions as a function of time [148]. They found that low charged protein ions evolve during the first tens of milliseconds to metastable structures. Some of the charged groups on these structures self-solvated via hydrogen bonding with the backbone carbonyls; in turn, polar side chains also undergo more tight packing [139]. However, the native backbone of these low charged metastable protein ions was overall maintained. Only in particular cases of proteins containing internal cavities, long-range rearrangements can occur, which result in more tightly packed structures [149].

ION MOBILITY MASS SPECTROMETRY OF BIOMOLECULES

Although the ion mobility phenomenon was discovered at the beginning of the 20th century [150], it was not until the sixties that ion mobility drift tubes were coupled to mass spectrometers by McDaniel *et al.* [151]. This combination opened the possibility for the simultaneous measurement of drift times and masses; hence, gaseous ions could be resolved not only by their mass to charge ratio, but also by their size and shape. Successive technical advances increased the sensitivity and resolution of this technique continuously up to the modern commercial IMMS spectrometers [152-154].

Ion mobility mass spectrometers feature an IM cell before the mass analyser (Figure 55 A). Gaseous ions are injected into the cell and accelerated by a weak electric field. The presence of a certain pressure of buffer gas in the cell creates low energy collisions with the buffer gas. The number of collisions depend on the collision cross sections (CCS) of the ion, which is intimately related to its size and shape. As a consequence of the collisions, the kinetic energy of the ions is decreased and they take longer to cross the IM cell before they are mass analysed (the time spent to travel the cell is known as “drift time”, t_d). The composition of the three-dimensional spectrum including t_d , m/z and relative intensity is known as ion mobility spectrum.

Types of ion mobility mass spectrometers

Here we will give a brief and general description of the configuration of commercial IMMS spectrometers. A detailed description is out of the scope of this manuscript, since extensive reviews have been published about IMMS instrumentation.

After ionization (typically by nanoESI), ions are selected by mass in a quadrupole, and stored for a few μ s in an ion trap. After this accumulation time, ions are released to the IM cell with a certain initial kinetic energy, which is controlled by the voltage of the transfer

gate. Accumulation of ions and releasing them on discontinuous packs increases the resolution and the sensitivity of the IMMS experiment.

Although all commercial IM spectrometers are based on travelling-wave IM cells (TWIMS), the physical principle is the same that on the case of drift tubes. For this reason, the theory of drift tubes is explained here. Drift tubes consists on a chamber with a low pressure of buffer gas subjected to a weak electric field E (Figure 55 B). Ions are injected in one end and accelerated according to E ; if the ratio between E and the number density of buffer gas N is low (the so-called low field limit), the drift velocity (V_d) of the ion become proportional to the electric field:

$$V_d = KE$$

Where K is the mobility, a parameter related to the overall size and shape of the ion. K is the responsible for the IM separation, which allows the direct extraction of the CCS according to the Mason-Schamp formula (Equation 6):

$$V_d = \frac{L}{t_d} = \frac{3e}{16N} \left(\frac{2\pi}{\mu K_B T} \right)^{\frac{1}{2}} \frac{KE}{\Omega} \quad \text{Equation 6}$$

Equation 6 allows the direct determination of the CCS (Ω) from the drift velocity (the ratio between the drift tube length and t_d), charge of the ion (e), temperature (T), the reduced mass of the ion and the buffer gas (μ) [155].

TWIMS cells are formed by successive stacked electrode rings operating between 0.5 and 1 mbar (Figure 55 C). Alternate voltage pulses are applied between successive rings, resulting in a "travelling wave" (T-wave). T-waves causes ions to "surf" or "roll over" the potential waves, depending on its mobility. Moreover successive rings improve the radial confining of ions and minimize the ion loss during the IM cycle. These features significantly increase the resolution and sensitivity of TWIMS with respect to conventional drift tubes [152, 154].

However, the application of a non-constant electric field rules out the use of the Mason-Schamp law, avoiding the direct calculation of CCS from t_d . Shvartsburg and Smith [156] established a model to describe TWIMS cells and the dependence of t_d and resolution as function of experimental parameters. Mason-Schamp law (Equation 6) can be rewritten as:

$$\Omega = \frac{3e}{16N} \left(\frac{2\pi}{\mu K_B T} \right)^{\frac{1}{2}} A t_d^B \quad \text{Equation 7}$$

Here, E is substituted by a power relationship that compensates the non-linearity of the electric field. The charge and mass dependence of the expression was eliminated, obtaining the general Equation 8:

$$\Omega' = A' t_d^B \quad \text{Equation 8}$$

Where parameters A' and B are obtained experimentally by calibration. The development of a calibration protocol adapted to native biomolecules was parallel to intensive experimental work on determining CCS with TWIMS spectrometers [157, 158]. The suitable calibration for native proteins require native ions as standards with similar masses and mobilities, which is crucial to ensure the correct interpolation of analyte ions. In other words, calibration must be performed in terms of mobilities and not of CCS.

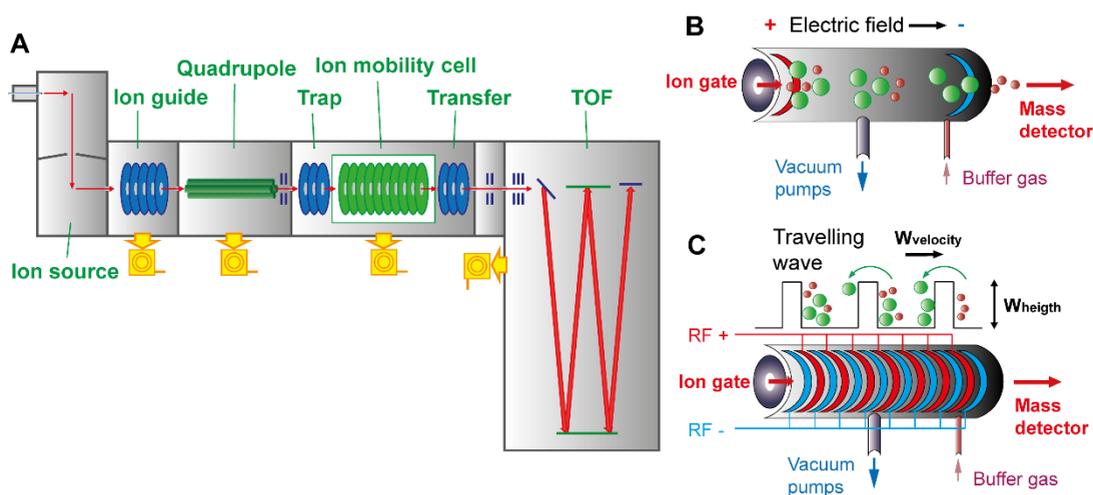


Figure 55: Ion mobility mass spectrometers. A) Scheme of a commercial SYNAPT G1 HDMS spectrometer (Waters Corp, Massachusetts, USA) [152]. The most important parts of the instrument are shown in the picture. B) Drift tube IM cell. The ring electrodes at the two ends of the tube are shown in red and blue. C) Travelling wave IM cell. Successive electrode rings with alternate potentials are shown in red and blue. The potential profile is shown in the top, with smaller ions “surfing” the waves and bigger ions “rolling over” the waves (green arrows). The two parameters which describe the T-wave (wave velocity and wave height) are also shown. In B) and C), charged analyte ions of different sizes are represented as green and brown spheres.

Applications of ion mobility mass spectrometry for the study of proteins and protein complexes

Ion mobility spectra provides the simultaneous measurement of mass and CCS [159]. This characteristic caused an important breakthrough of mass spectrometry in structural biology from the late 90s. Together with the spread of commercial IMMS spectrometers SYNAPT G1 [152], G2 [153] and G2S [154] (Waters Corp, Massachusetts, USA), the studies of biomolecules in the gas phase by IMMS have grown steadily. The large number of publications of high molecular weight protein complexes, heterogeneous systems, and self-aggregating proteins confirms the reliability and versatility of IMMS [21].

One of the most attractive applications of IMMS is the study of large protein complexes. Keeping the quaternary structures during ionization was one of the main challenges for the study protein complexes in the gas phase, but the development of the instrumentation, specific protocols [160] and the use of stabilizing or charge reducing agents, successfully overcome this drawback [149, 161-163]. This allowed Heck and co-workers to investigate the topology and the assembly pathways of virus capsids in the gas phase by IMMS [164, 165]. In turn, the Robinson group achieved the ionization of membrane complexes followed by stripping of detergent molecules in the gas phase [166]. For instance, this methodology allowed the study of the interaction between subunits of V-type ATPase [167]. The development of coarse-grained molecular modelling strongly contributed to the assignation of the complex topology by correlation between experimental and simulated data [168].

The versatility of IMMS found a powerful application for the study of intrinsically disordered proteins (IDPs), proteins containing unstructured regions and self-aggregating proteins. The extremely high structural heterogeneity of these systems difficult the study by ensemble-average biophysical techniques. Instead, the separation power of IMMS together with the simultaneous measurement of mass and size allows the quantification of frozen populations of conformers and aggregates. In this field, Ashcroft and co-workers benefited from IMMS to capture oligomeric species of β_2 -microglobulin; moreover, they found that the formation of β_2 -microglobulin oligomers was prevented in the presence of rifamicyn B [169, 170]. In another case, the binding of cognate DNA on the flexible regions of p53 protein disclosed significant structuring upon DNA binding, while the functional centre of p53 responsible for DNA binding remained folded in vacuum. Given the high complexity of this system, the different compositions of p53 domains and subcomplexes were studied separately [171].

All of these advances of IMMS in structural biology have been possible thanks to the development of methods for the calculation of theoretical CCS. The calculation of theoretical CCS from high-resolution structures of biomolecules or coarse-grained models has been crucial to correlate the low-resolution experimental data with theoretical information. The most important software is MOBCAL, which was developed by Jarrold and Shvarstburg [172, 173]. Briefly, this program calculates the orientationally averaged silhouettes of the protein. These silhouettes are obtained by interactions between buffer gas and the protein ions (Figure 56 A and B); consequently, a realistic description of the interaction model is crucial for the quality of the CCS calculations. According to the interaction model, MOBCAL software contains three different methods:

- Projection approximation (PA): this is the simplest model, which estimates the CCS as an average “shadow” of all orientations of the ion. Since this method completely ignores the long-range interactions and scattering processes between ions and buffer gas particles, it always underestimates the CCS.
- Exact hard sphere scattering (EHSS): this model includes the scattering phenomenon which occur between the incident buffer gas particle and the interacting ion [172]. The scattering modifies the trajectory of the buffer gas, and consequently, it affects the resulting CCS value.
- Trajectory method (TM): a part from scattering, TM also evaluates the long-range interactions between interacting particles, ion-induced dipole interactions, and multiple collisions with buffer gas particles [173]. This results in a more accurate CCS calculation, with values up to 20 % higher compared to that of PA method. The complete description of surface phenomena between ions and buffer gas makes this method especially useful for large systems, proteins displaying surface cavities and biomolecules with irregular shapes.

However, all these models use He as *in silico* buffer gas, while the commercial SYNAPT G1 IMMS spectrometers operate with N₂ as a buffer gas. This leads to inconsistencies in the CCS values, which might hamper the correlation between experimental and theoretical results. However, lower CCS values are systematically expected for experimental CCS with respect to theoretical ones [174].

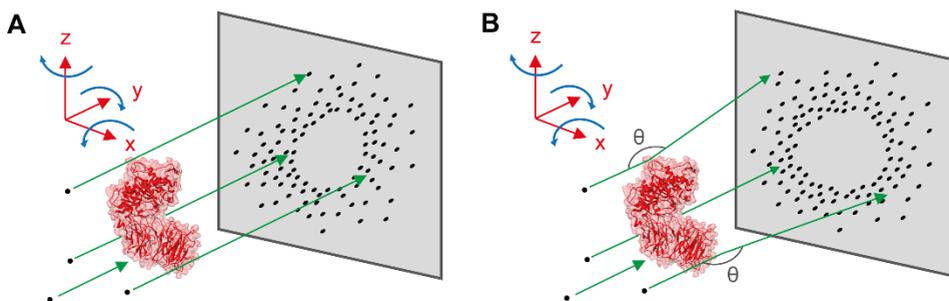


Figure 56: Theoretical calculation of gaseous CCS by MOBCAL software. A) Projection approximation calculates the theoretical CCS as the orientationally averaged “shadow” of the protein in a flux of buffer gas particles (black spheres). B) Trajectory method includes a more complete description of the interaction between buffer gas particles and protein ions, which modifies the direction of the interacting particles (θ , green arrows). This result in larger estimations of CCS.

STUDY OF POP STRUCTURE AND DYNAMICS BY IMMS

As explained in the previous section, during the last decade IMMS studies progressed towards larger and more complex systems. However, few examples of the application of IMMS for the structural characterization of dynamic proteins in the μs -ms time scale are found in the literature. One of the main reasons is the lack of consensus of to what extent the solution structures are transferred to the gas phase. In this regard, several opinions suggest that the conformational populations in exchange can be effectively transferred to the gas phase [148, 175]. In order to assess this question, we decided to analyse the conformational landscape of POP by IMMS.

The objective of this chapter was to determine the capacity of IMMS to detect the different POP conformers existing in solution, as well as the effects of inhibitors on the structures and populations in the gas phase. Moreover, the experimental results were complemented by POP MD simulations in solution and in vacuum.

All ion mobility experiments were carried under the supervision of Dr. Marta Vilaseca (Mass Spectrometry Core Facility Manager, Institute for Research in Biomedicine, Barcelona, Spain). Experiments were done with a SYNAPT G1 HDMS mass spectrometer (Waters Corp, Massachusetts, USA) [152], at the Mass Spectrometry Core Facility at Institute for Research in Biomedicine.

Optimization of experimental conditions

The optimization of the ionization and ion transmission conditions is crucial to obtain gaseous protein ions in the native state. However, mild instrumental conditions are in compromise with sensitivity; for this reason, instrumental parameters were empirically determined by monitoring signal intensity. Setting the TWIMS conditions is also important in order to achieve good chromatographic separation and an accurate CCS measurement.

Optimization of the ionization and ion transmission conditions

A commercial chip-based Triversa Nanomate nanoESI ion source (Advion Inc., New York, USA) was used in all these experiments (ESI ionization of POP required important amounts of protein and high voltages, and suffered from reiterated precipitation in the capillary). The commercial nanoESI device was a robust and reliable ion source, which provided a stable flux of ions. Positive ionization was used in all cases; however, we performed some trials in the negative mode as well.

Definitive voltages for the capillary and the cone were tuned according to the threshold values (Table 8). These parameters were optimized to 1.5-1.75 kV and 20 V for the capillary and the cone, respectively. In order to facilitate ion transmission in the trap region, collision energy (CE) and bias were kept at the minimum value (6 V and 15 V, respectively); higher values caused unfolding of gaseous protein ions [176]. Finally, the backing pressure was minimized to 3 mbar in order to avoid possible structural alterations of POP gaseous ions [158]; lower values than those recommended by Ruotolo *et al.* were chosen [160]. In addition, all these voltages and pressures were compatible for native calibrating protein ions.

50 mM NH₄AcO pH = 8 was used as POP native buffer. Ammonium acetate is an excellent volatile buffering agent for native mass spectrometry, which favours evaporation and stabilizes the native structures [177]. In addition, activity assays showed that POP catalytic activity was maintained in this buffer. The use of higher NH₄AcO concentrations, or the addition of small amounts of NaCl to increase the ionic strength, did not improved the result. The use of organic solvents, such as methanol, severely compromised the stability of the protein in solution. POP concentration was optimal between 10-15 μM; higher concentrations were detrimental for the experiments. In our preliminary experiments, we observed interferences of the N-terminal HisTag in the CCS calculation; for this reason, the HisTag was cleaved. In general, the quality of the buffers and the sample preparation were

critical for the reproducibility of the results (see Materials and methods). All of the considerations described in Materials and methods were crucial in order to keep the purity of POP and avoid the presence of salts, residual proteins from the purification, and highly ionisable molecules in the sample; such contaminants dramatically interfered with the ionization of POP.

Capillary voltage (kV)	1.75
Sampling cone voltage (V)	20
Extraction cone voltage (V)	1
Trap CE (V)	6
Trap DC bias	15
Transfer CE (V)	6
Backing pressure (mbar)	3
IMS gas flow (ml/min)	24
IM wave velocity (m/s)	300
IM wave height (V)	8, 9, 10

Table 8: Definitive instrumental conditions for native POP experiments in SYNAPT G1 TWIMS spectrometer.

Optimization of the TWIMS conditions

Ion mobility parameters were optimized in order to achieve good resolution for all POP charge states, and to cover the drift time range of all native calibrating protein ions. Definitive values are summarized in Table 8. Ion mobility pressures, T-wave velocities and T-wave heights (Figure 55 C) were tuned together [160]. N₂ was used as buffer gas in all IMMS experiments. Given the different polarizability of N₂ compared to He, the CCS obtained in He are always between 5-10 % smaller compared to that obtained in N₂ [157]. However, performing the experiments in N₂ prevents the increased field heating caused by He, which might lead to partial unfolding [178].

IM pressure was chosen according to the values recommended by Ruotolo *et al.*; higher IM pressures did not improve the resolution [160]. IM T-wave velocity was also estimated taking a reference value from the same publication. In order to avoid the effects of the electric field heterogeneities of TWIMS cell, three different T-wave heights were chosen (8, 9 and 10 V); these wave height values covered the full range of mobilities for POP and standard ions [179]. Finally, automatic release mode was used in all measurements. Experimentally, it was observed that the use of manual release at low time intervals (25 μ s) caused artefacts in the mobilograms.

IMMS experiments of POP

Free and inhibitor-bound native POP

Experiments were performed by triplicate protein batches expressed and purified independently in order to minimize the sample dependence. In turn, three measurements at different wave heights were performed for each batch (8, 9 and 10 V). After data processing, all the results were averaged, and the error was obtained by calculating the standard error of the mean. Two additional replicate experiments were performed few months later in identical instrumental conditions, obtaining similar results. This confirmed the robustness of the experiments, but these results were not averaged due to the lower quality of the protein batches (Figure 57 A).

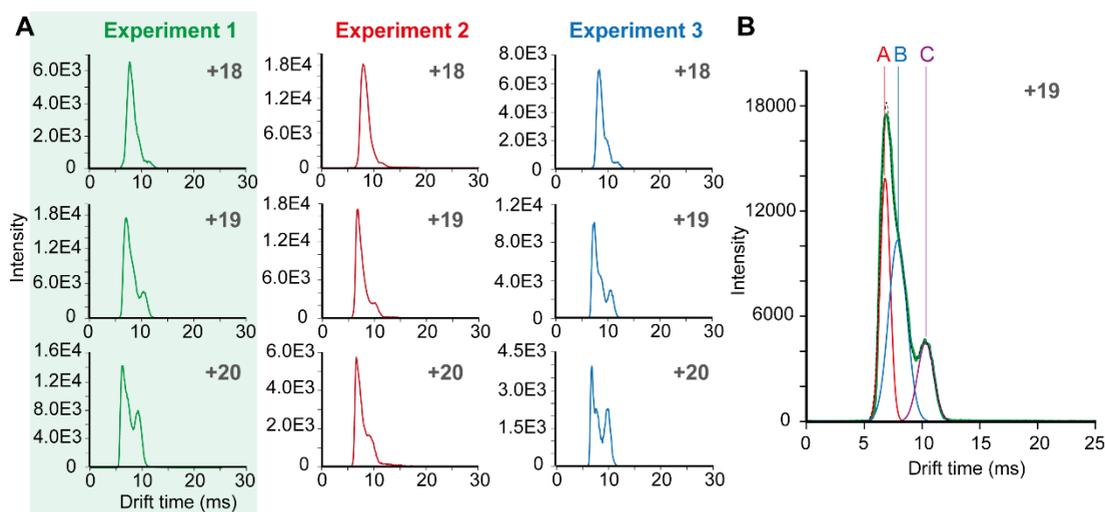


Figure 57: Native IMMS experiments of free POP. A) The mobilograms of +18, +19 and +20 charge states at a wave height of 9 V are shown for the different experiments. All the definitive data was extracted from Experiment 1 (highlighted in green). The differences observed in Experiments 2 and 3 were due to the lower quality of the protein sample. B) Mobilograms of the higher charge states show three different Gaussian distributions, named as A, B and C. The mobilogram displayed in green correspond to +19 charge state of experiment 2. The theoretical mobilogram generated by the sum of A, B and C Gaussian functions is shown in black dashed line.

The calibration was performed with native protein standard ions closely bracketing the range of mobilities of all POP charge states [157]. The standards were transthyretin tetramer (56 KDa), bovine serum albumin (69 KDa), and concanavalin A (103 KDa). Calibration was performed before and after IM experiments, and the values were averaged.

Collision induced unfolding

Ions moving under the electric field of TWIMS cell undergo low energy collisions with the buffer gas, which results in a loss of kinetic energy and a concomitant increase on the internal energy. The excess of internal energy caused by increased collisional activation is dissipated into the structure by high energy transitions. Such transitions involve structural rearrangements, such as conformational changes, and ultimately unfolding. Herein, forcing collisional activation reflects the stability of the folded state [176]. In commercial SYNAPT instruments, collision induced unfolding experiments (CIU) are produced by increasing the acceleration of ions in the trap chamber (trap collision energy, CE) before injection in the TWIMS cell.

In this case, POP unfolding was forced by applying trap CE voltages of 15, 20, 25, 30 and 35 V. In order to simplify these experiments, only a single wave height was used (9 V), and only the +19 charge state was monitored.

Charge effects and acid denaturing of gaseous POP ions

The total charge of ions influences the stability of structures in the gas phase. For this reason, we explored the populations of native POP structures for all the accessible charge states. This task was performed according to several recent studies, in which charge states distributions were displaced with the use of supercharging or charge reducing agents [149, 180].

Certain compounds with low vapour pressures cause supercharging of protein ions if they are present in solution. Such compounds are less volatilized during the evaporation of the droplet, and consequently, they become enriched during droplet evaporation. These compounds prevents evaporative cooling, and therefore, the microscopic temperature of the ion increases; this favours more extended conformations [181]. Typically, these agents are sulfolane or *meta*-nitrobenzyl alcohol (*m*-NBA) at low concentrations (< 1 %) [182]. In turn, charge reducing is produced with highly diluted organic bases, for instance, imidazole or 1,8-diazabicycloundec-7-ene. Similarly, these small molecules are enriched during the solvent evaporation of droplets. In the last stages of evaporation, these basic molecules capture protons from the charged groups of the protein ions; this yields an overall reduction of the final charge state [183]. In both cases, the low concentrations of such agents do not affect the stability of native proteins in solution.

Our initial trials of supercharging POP ions in NH₄AcO 50 mM buffer with 0.5 % *m*-NBA resulted in a dramatic drop of the ionization yield and of the quality of the mass spectra. For this reason, supercharging POP ions was discarded. On the contrary, the addition of 0.05 % imidazole to NH₄AcO 50 mM buffer resulted in an impressive modification of mass spectra towards reduced charge states (this minimal content of imidazole had no effects on POP catalytic activity). Finally, in order to estimate the CCS values of denatured POP conformations, we performed IMMS measurements of free POP in NH₄AcO 50 mM buffer with 1 % formic acid. In this acidic conditions (approximate pH 3), activity assays showed the complete absence of catalytic activity of POP.

Native POP IMMS measurements in negative mode

According to the isoelectric point of POP (5.4), this protein is negatively charged in solution at neutral pH. In order to perform IMMS measurements of POP in a charge state similar to the native state in solution, we performed preliminary IMMS experiments in the negative ionization mode. These experiments would minimize the Coulomb repulsion occurring in positively charged ions.

Ionization in negative mode by nanoESI using the Advion source was not successful. For this reason, ESI ionization was used in these experiments. The rest of the instrumental conditions were optimized analogously to POP experiments in positive ionization mode; definitive parameters are summarized in Table 9.

Capillary voltage (kV)	2
Sampling cone voltage (V)	20
Extraction cone voltage (V)	3
Trap CE (V)	6
Trap DC bias	15
Transfer CE (V)	4
Backing pressure (mbar)	5.3
IMS gas flow (ml/min)	22
IM wave velocity (m/s)	300
IM wave height (V)	11, 12, 13

Table 9: Definitive instrumental conditions for native POP experiments in negative mode.

IMMS data processing

After recording IMMS experiments, *m/z* protein charged species in the mass spectra were deconvoluted to the zero charged averaged masses with MaxEnt 1 algorithm and the

integrated Mass Measure tool (Micromass MassLynx software package, Waters Inc., Massachusetts, USA). Secondly, mobility traces were manually selected for each charge state using the DriftScope program (Waters Inc., Massachusetts, USA), and exported as numerical list.

Native mobilograms of POP disclosed several conformational populations coexisting as a function of the charge state (Figure 57): for this reason, mobilograms were individually subjected to a Gaussian decomposition with GraphPad Prism software (GraphPad Software Inc., La Jolla, CA, USA). Visual systematic inspection of all mobilograms showed that 3 Gaussian distributions were comprised in each mobilogram, named *A*, *B* and *C* (from lower to higher t_d values, Figure 57 B). The centres of the Gaussian functions were taken as t_d value for the calculation of the CCS, and the approximate populations were extracted from the areas. In order to confirm the quality of the fitting, the three Gaussian functions were summed and compared to the corresponding experimental mobilogram. This procedure was carried out for free POP and POP bound to inhibitors, as well as in charge reduced POP experiments. In the case of CIU experiments, new Gaussian functions (*D* and *E*) were assigned to emerging peaks (see below).

Afterwards, an excel sheet was prepared for the calibration of experimental CCS (Figure 58). This sheet contained the formula to eliminate the time offset of experimental t_d values arising from the travelling wave velocity and the synchronization with the TOF detector. The corrected drift (known as t_d'') is obtained by Equation 9:

$$t_d'' = t_d - 31t_t - 61t_m - 0.041 \sqrt{\frac{m}{z}} \quad \text{Equation 9}$$

In which t_t and t_m are time spent by the T-wave to be displaced from one electrode ring to the next in the transfer and mobility regions, respectively; these input values were manually entered in the excel sheet. 0.041 is an empirical factor of the TOF synchronization; and m/z is the mass to charge ratio of the ion (Figure 58 A).

Next, the standard protein ion m/z and t_d values were introduced. With the calculated mass, charge and theoretical CCS value of the standard ions tabulated in [157], it was possible to calculate the corrected CCS (Ω'):

$$\Omega' = \frac{\Omega}{z \sqrt{\frac{1}{m_i} + \frac{1}{m_N}}} \quad \text{Equation 10}$$

Where z and m_i is the charge and mass of the ion, respectively; m_N is the mass of the buffering gas. Given that TWIMS cell of SYNAPT G1 operates with N_2 , we used the tabulated CCS values for this gas (Figure 58 B).

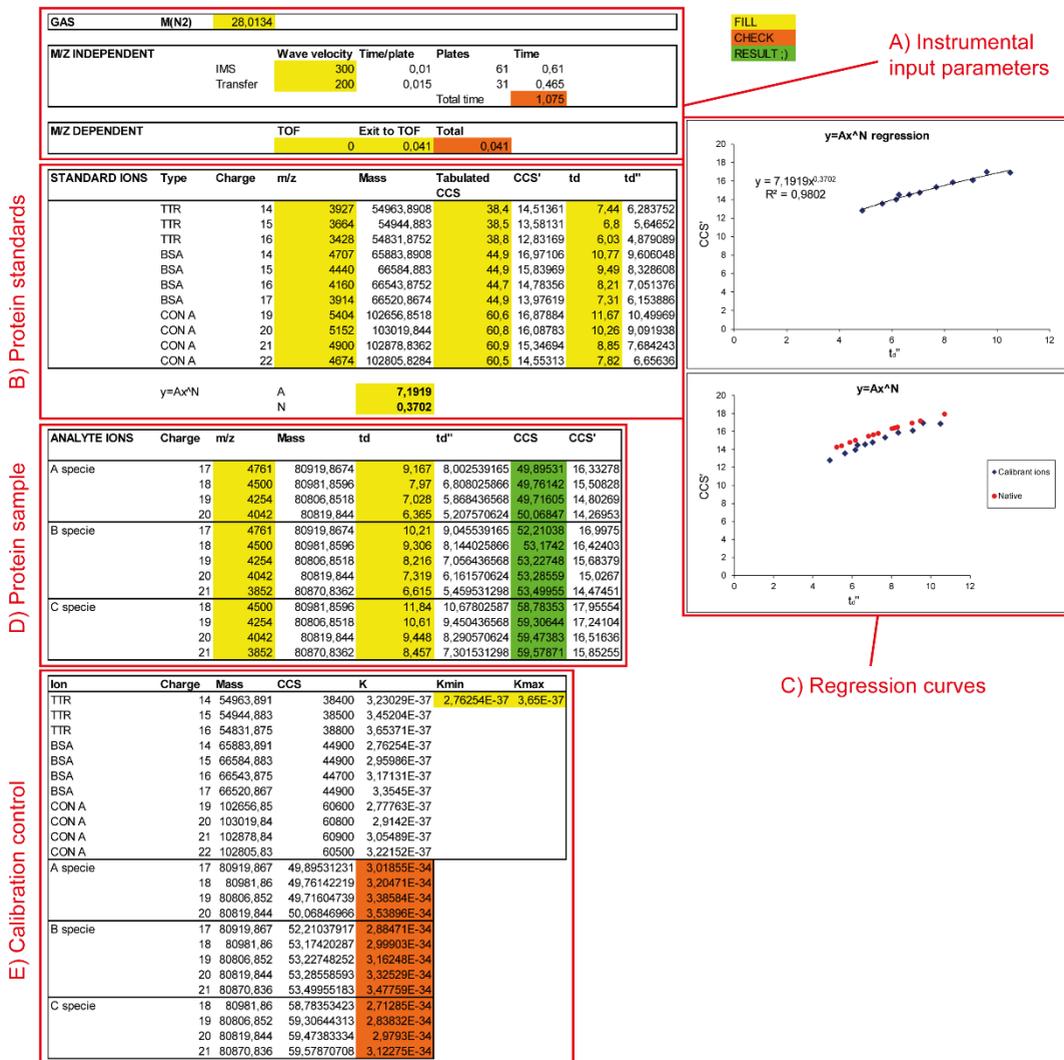


Figure 58: Excel sheet for the CCS calibration and calculation. According to the nomenclature used in the main text, panels A, B C, D and E delimit the different parts of the process.

With all these values the corrected CCS vs t_d'' was plotted, and it was fitted to the power law of Equation 8 (Figure 58 C). A and B values in Equation 8 were extracted from this fitting.

The last step was to introduce the z , m_i , and t_d values from analyte ions (t_d were automatically subjected to the correction of Equation 9). The corrected CCS were obtained with Equation 8, and the inverse operation of Equation 10 yielded the definitive CCS value in N_2 (Figure 58 D):

$$\Omega = (t_d'')^B A z \sqrt{\frac{1}{m_i} + \frac{1}{m_N}} \quad \text{Equation 11}$$

Finally, all mobility values were automatically calculated and checked against the standard ions in order to validate the interpolation of the analyte ions within the calibration curve. A final visual inspection of the plots of CCS' vs t_d'' was carried out (Figure 58 C and E).

Molecular dynamics simulations and theoretical collision cross sections calculations

The use of computational methods to obtain theoretical CCS from high resolution protein structures obtained by X-ray, NMR or MD simulations is crucial for the interpretation of IMMS experiments [174]. For this reason, several CCS calculations of POP structures in solution and in the gas phase were performed. The conformational landscape of POP was studied in explicit solvent using replica exchange molecular dynamics (REMD), and in vacuum using classical molecular dynamics (MD). **All of the computational simulations of the IMMS chapter were carried out by Dr. Sergio Madurga (Department of Physical Chemistry of the University of Barcelona, Barcelona, Spain).**

Replica exchange molecular dynamics simulations in solution

As a first approach, the conformational landscape of POP in solution was explored by REMD simulations. Briefly, REMD calculations consists on several MD simulation replicas of the system performed at different temperatures (i.e. with different internal energies). At certain regular intervals, the structures of MD simulation replicas are exchanged (Figure 59). This exchange facilitates high energy transitions between energy minimized conformations at different temperatures, and allows a broad exploration of the conformational space [184]. For this reason, the simulation time is not realistic, and REMD simulations are usually presented as histograms.

REMD simulations were performed starting from the closed conformation of porcine POP in the free form (PDB entry 1H2W). Three different REMD simulations were performed: in

REMD I, the protonation state corresponding to a neutral solution was selected, and the simulation was kept up to 8.4 ns (Figure 60 A). Afterwards, the initial structure was prepared using the protonated state for all Asp, Glu and His residues. This condition simulated the behaviour of the protein in low (denaturing) pH conditions, and favoured larger conformational rearrangements. REMD simulation was performed during 20 ns (REMD II), and in the replica with the highest temperature (327 °K), loop A repeatedly detached from the protein surface. The POP structure with the Loop A detached was used as a starting point for a third REMD simulation. REMD III was carried out during 20 ns, using the protonated state of charged residues corresponding to the neutral pH (Figure 60 B). This last simulation showed that loop A detachment occurring at denaturing conditions also took place in neutral conditions. For further REMD details, see Materials and methods section.

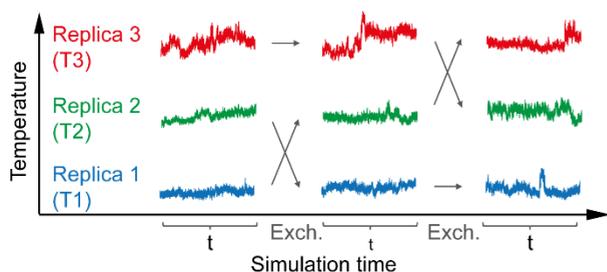


Figure 59: Replica exchange MD simulations. Three simulation replicas at different temperatures (T_1 , T_2 and T_3) are performed during a time interval t . After this interval, the structures are exchanged.

Molecular dynamics simulations in the gas phase

The stability of gaseous POP ions was studied by MD simulations in vacuum. Given that performing MD simulations in a fixed total charge is not trivial, the calculations were restricted to the representative charge state +20. This procedure started with the determination of the appropriated protonation state of acidic or basic residues compatible with a given total charge. In order to take into account all several microstates contributing to the experimental ensemble-averaged results, a Monte Carlo procedure was used to generate a pool of microstates with different protonation distribution compatible with a determined total charge. The Monte Carlo procedure takes into account the protonation constant (K_i) of each residue i . In a Monte Carlo step, the protonation and deprotonation processes are tested, and the free energy associated to these processes is given by the expression:

$$\Delta G = \pm k_B T (\ln 10) (pH - \log K_i)$$

In which the sign + is used for protonation, and – for deprotonation. K_i is the microscopic protonation constant which correspond to the inverse of the microscopic acid/base constant of the acidic or basic residues (Asp, Glu, Arg, Lys, Cys, His, Tyr). K_i values were obtained with the PropKa program through the web server [185, 186], which calculates protein pKa values according to the structural arrangement of residues and other empirical parameters. For this reason, the use of this algorithm is highly reliable to detect the surface charged residues.

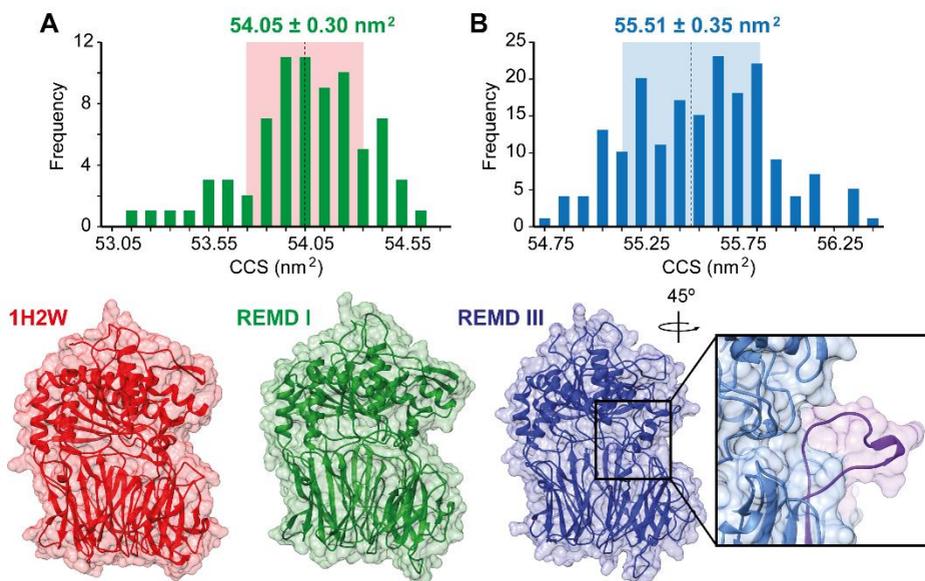


Figure 60: Replica exchange MD simulations of free POP in solution. A) 8.4-ns long REMD I with the protonation state at neutral pH. B) 20-ns long REMD III starting from the loop A detached structure, obtained previously under acidic conditions. In both cases, the average CCS is marked by the dashed line; the confidence interval is limited by the highlighted region. The structures shown below are representative of REMD I and III (green and blue, respectively). Inset of REMD III structure shows a detail of loop A detachment (purple). The initial X-ray structure 1H2W has been included for comparison (red).

The Monte Carlo algorithm generated a distribution of protonation states for the POP that are compatible with a +20 charge state, starting from 1H2W structure. Monte Carlo simulations were performed at pH = 5, in a total of 100000 steps. Afterwards, 10 selected charge distributions of +20 total charge were subjected to MD simulation in a length of 6 ns. Structures were sampled at 11 time points of during the production run, from 1.5 to 6 ns in 500 ps intervals (Figure 61 A).

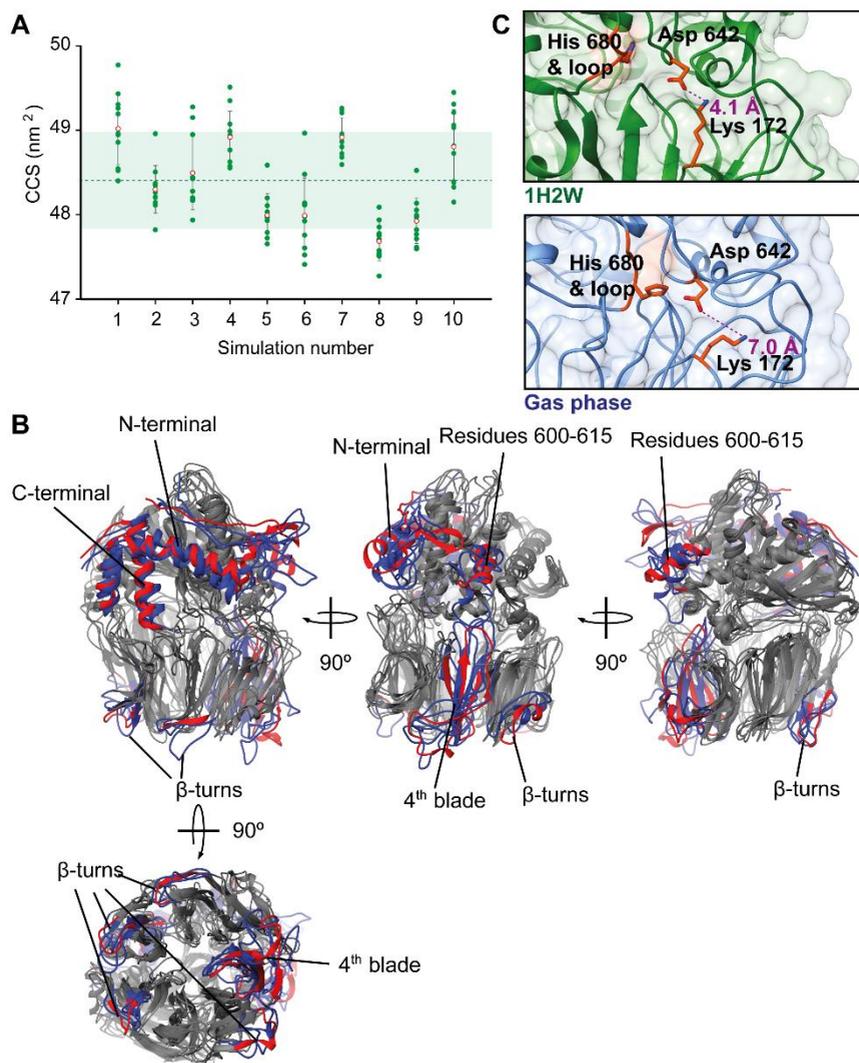


Figure 61: MD simulations in the gas phase. A) CCS obtained from gaseous MD simulations of free POP, in 10 different charge distributions. Each charge distribution consists of 11 data points sampled in 500 ps intervals. The green dotted line shows the average value, and the green area correspond to the standard deviation. B) Different perspectives of three structures of REMD III at 6 ns (blue) overlaid with 1H2W X-ray structure (red). For better clarity, only the most affected regions have been coloured (see labels). C) Details of some residues surrounding the active site (orange) obtained from MD simulations in the gas phase (blue) compared to 1H2W X-ray structure (green). The protonation of Arg 642 breaks the salt bridge with Lys 172, which is pushed towards the surface (the distance between the two residues is shown in magenta). His 680 loop adopts a distorted configuration.

Temperature is an important aspect of MD simulations in the gas phase. Although IMMS experiments are performed at room temperature, collisions of accelerated ions with buffer gas molecules lead to an increase in the microscopic temperature of protein ions. Several studies described how to calculate the gain of internal energy caused by collisions (the so-called field heating). In the case of TWIMS cells, the Mason-Schamp law (Equation 6) can be adapted to provide field heating [187]:

$$T_{eff} = T_{gas} + T_{field} = 298 + \frac{M_N}{3k_B} \left[\frac{2v_d s^2}{s + v_d} \right] \quad \text{Equation 12}$$

Where M_N is the mass of a molecule of buffer gas (in kg), k_B is the Boltzmann constant, s is the TWIMS wave velocity (300 m/s), and v_d is the drift velocity of an analyte ion. Here we calculated v_d by dividing an average t_d'' of 7 ms with the length of the TWIMS cell commercial SYNAPT (185 mm). Equation 12 gave us an effective temperature of 313 °K, which was used in our calculations.

A preliminary analysis of the structures at 6 ns was performed for the 10 different charge distributions. Overall, the tertiary structure of these simulated forms resembled to the initial 1H2W X-ray structure. Comparing the MD simulations carried out in chapter 3, the interdomain angle was similar to the average of MD1 simulation of 1QFS X-ray structure with the inhibitor removed ($62 \pm 4^\circ$ versus $61 \pm 2^\circ$, respectively). However, the average RMSD of 3.2 Å indicates that several local changes take place. Figure 61 B shows that N and C-terminal regions, the external β -turns of the β -propeller, and the fourth blade of the β -propeller are the most affected areas. The analysis of residues surrounding the active site disclosed that the salt bridge between Lys 172 – Asp 642 is not formed in all structures (Figure 61 C). Under the simulation conditions, Asp is in its carboxylic form; consequently, the salt bridge does not take place and the charged Lys 172 protrudes to the surface. In addition, this rearrangement affect His 680 loop, which adopts a non-native conformation. Together, these structural artefacts derived from gaseous environment would cause the compaction observed on POP structure in the MD simulations performed in vacuum.

Theoretical collision cross section calculation

Theoretical CCS obtained from these simulations were calculated with MOBCAL program using the trajectory method [173] (see Materials and methods). The numerical values are summarized in Table 10. Moreover, in order to compare these values with those of POP open conformations obtained by SAXS, the 3 more open structures selected by the EOM of

MD2 (with interdomain separation angles of 89, 91 and 91°) were also included in the calculation.

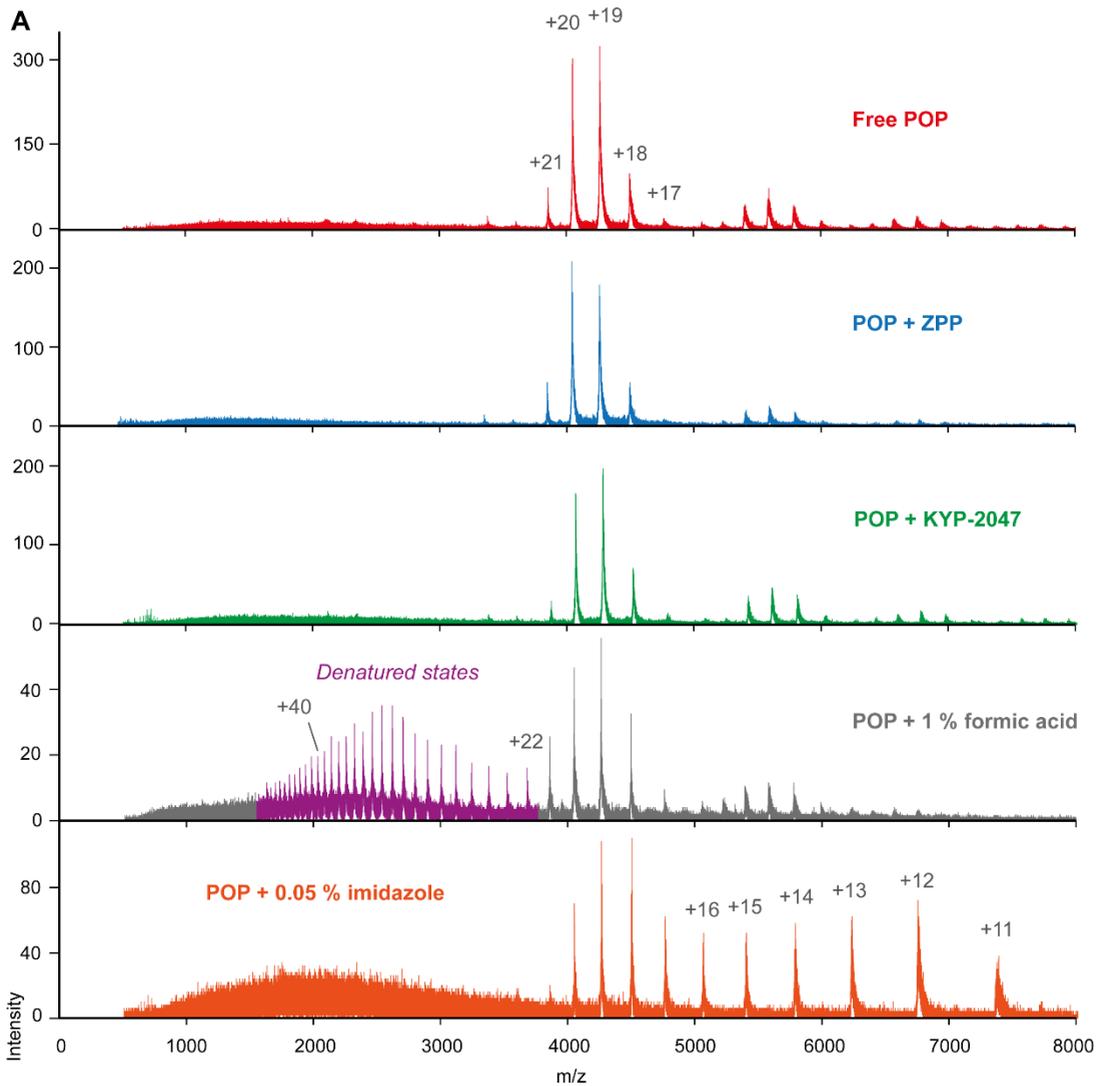
EXPERIMENTAL	CCS species A (nm ²)		CCS species B (nm ²)		CCS species C (nm ²)	
	Average	Interval	Average	Interval	Average	Interval
Free POP	50.4	50.1 - 50.8	53.8	53.2 - 54.2	59.3	58.3 - 59.9
POP + ZPP	49.8	49.6 - 50.1	53.1	52.6 - 53.6	58.7	57.9 - 59.1
POP + KYP-2047	50.6	50.2 - 50.9	54.1	53.7 - 54.6	59.5	58.5 - 60.3
POP + Imidazole	50.9	49.8 - 52.5	54.2	52.9 - 55.5	59.0	58.0 - 60.1
SIMULATED						
	Average CCS (nm ²)	σ	Production run (ns)	Structures		
Gas phase MD	48.41	0.57	1.5 to 6 (10 x)	11 (10 x)		
REMD I	54.05	0.30	-	75		
REMD III	55.51	0.35	-	184		
EOM of MD2	58.53	0.15	-	3		

Table 10: Collision cross sections of POP. The experimental CCS values were obtained by averaging the CCS at all charge states; CCS intervals were delimited by the minimum and maximum values within the charge state interval. The simulated CCS values of REMD simulations were obtained by averaging all replicas; in the case of gaseous MD simulations, CCS were averaged for the 10 simulations from different charge distributions.

Results

Free and inhibitor-bound POP under native conditions

The mass spectra of free POP and POP bound to ZPP and KYP-2047 showed +17, +18, +19, +20 and +21 charge state species (Figure 62 A). The most populated states were the +18, +19 and +20. In the case of native POP, in POP with 0.05 % imidazole, and with 1 % of formic acid, all the masses approximately corresponded to the theoretical value (table of Figure 62 A). Regarding inhibitor-bound POP, the masses showed the mass increase corresponding to the respective inhibitors: this confirmed the quantitative covalent binding of the inhibitors. However, native MS can result in inaccurate measurements of the mass due to poor desolvation of ions in aqueous buffers.



TYPE	Average mass (Da)	SEM	Mass difference (Da)
POP (theoretical)	80828.6	-	-
Native free POP	80824.7	6.5	-3.9
Native POP + ZPP	81261.4	6.2	+ 432.8 (Theor. = 330.38)
Native POP + KYP-2047	81155.0	0.9	+ 326.4 (Theor. = 339.43)
POP + 0,05 % imidazole	80817.1	0.4	-11.5
POP + 1 % formic acid	80821.9	0.7	-6.7

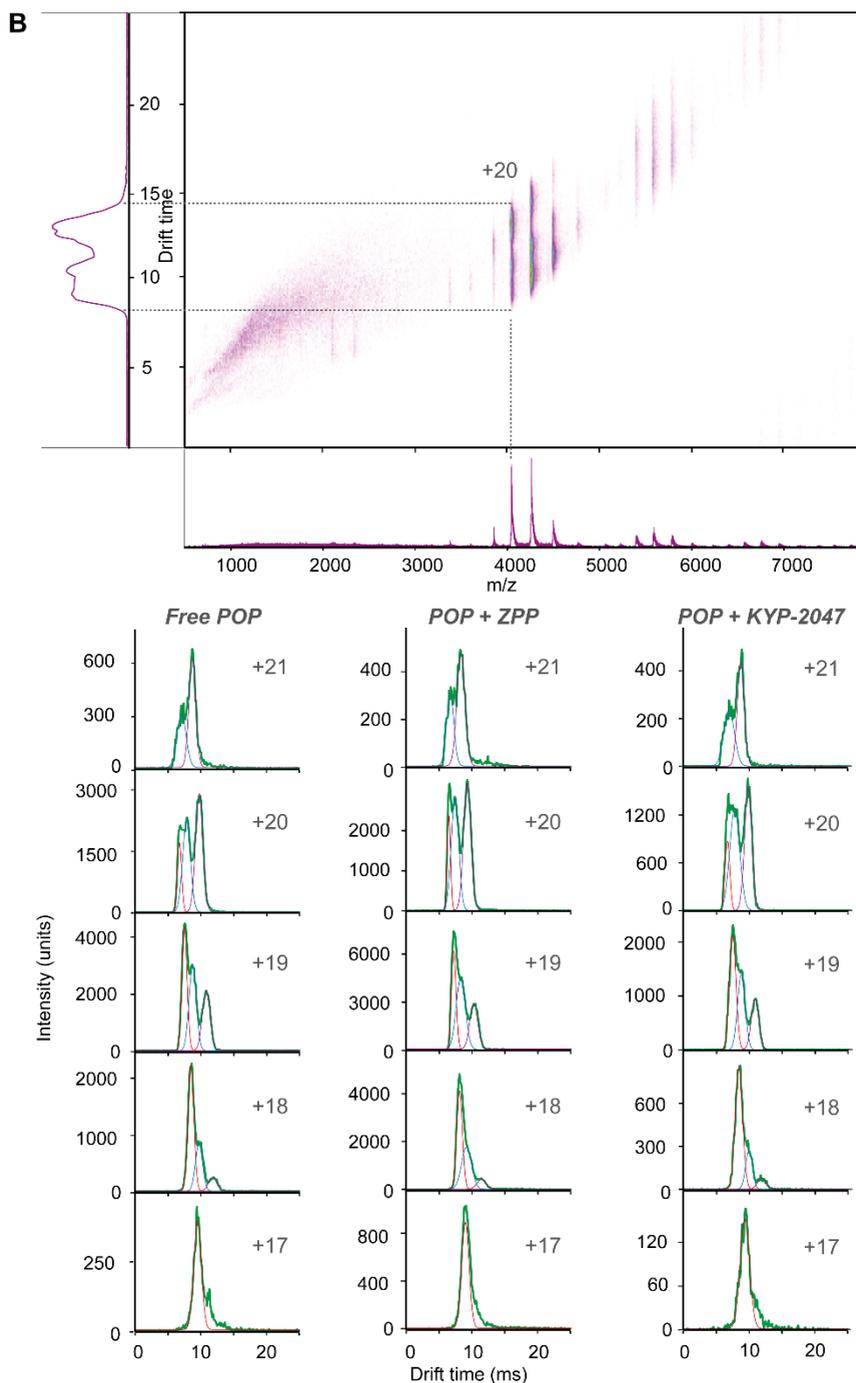


Figure 62: Native IMMS experiments of free POP and POP bound to inhibitors. A) Mass spectra of free POP, ZPP-bound POP, POP bound to KYP-2047, POP + 1 % formic acid and POP + 0.05 % imidazole (from top to down). The experimental masses of these forms of POP are shown in the table

below. The differences with respect to the theoretical mass of POP are shown in the right column. B) Example of a bidimensional ion mobility spectra of free POP, showing the mass (horizontal axis) and ion mobility (vertical axis) dimensions. The +20 state has been marked as example. The representative mobilograms (green) of free POP and POP bound to inhibitors for charge states from +17 to +21 are shown below. The Gaussian distributions of the different species *A*, *B* and *C* are depicted in red, blue and purple, respectively.

Native POP mobilograms disclosed different species as a function of the charge state (Figure 62 B). In the case of +17 charge state a major population was present (*A* species); however, a second population at higher t_d can be perceived, although the Gaussian decomposition failed systematically (*B* species). In the case of the +18 charge state, the data could be adjusted to three Gaussian functions assigned to *A*, *B* and *C* species. In the case of +19 and +20 charge states the data was also nicely described by the three Gaussian functions, but with increasing proportions of *B* and *C* species. Finally, only *B* and *C* species were found to adjust to the mobilogram of +21 charge state. With small differences, this result was reproduced for POP bound to active-site directed inhibitors ZPP and KYP-2047. The experimental CCS of the different species of free POP, POP bound to ZPP and POP bound to KYP-2047 calculated from the calibration are summarized in Table 10 and Figure 63.

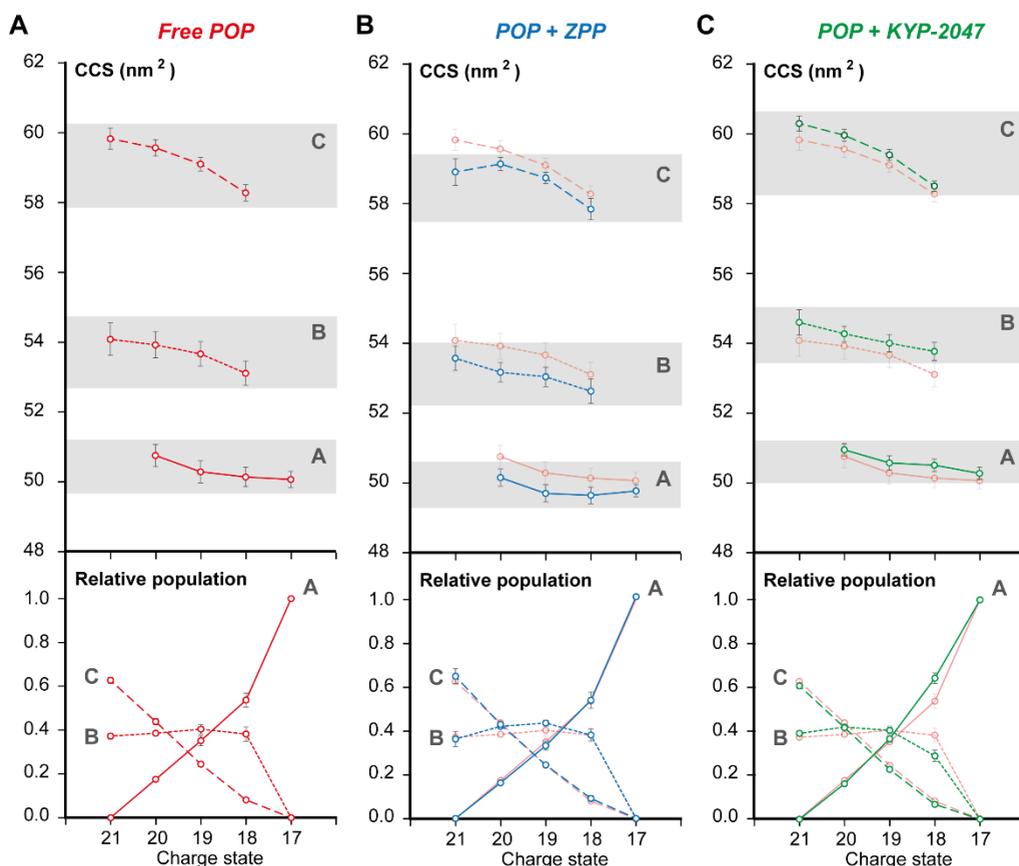
The population of the three species *A*, *B* and *C* was highly dependent on the charge state. For this reason, we calculated the relative population of a given species *X* according to the following formula:

$$\text{Relative population (X)} = \frac{A_X}{A_A + A_B + A_C}$$

Where A_X , A_A , A_B and A_C are the absolute areas of the *X* species and the *A*, *B* and *C* species, respectively. The plot of the relative population as a function of charge state allowed to monitor the evolution of the species with the charge state. Figures 410 and 411 A shows that free POP species with the largest CCS (*C*) became more populated at higher charge states, while *A* population decreased, and finally disappeared at +21 state. Contrarily, *B* species was approximately constant between +18 and +21 states. An almost identic behaviour was observed for the relative populations of *A*, *B* and *C* species of POP bound to inhibitors (Figure 63 B and C).

Figure 63 (*next page*): Experimental CCS (top) and relative populations (down) of *A*, *B* and *C* species as a function of charge state. The results are shown for free POP (*A*, red), POP bound to ZPP (*B*,

blue) and POP bound to KYP-2047 (C, green). The plots of free POP are overlapped in all cases for an easier comparison.



IMMS experiments performed in the negative mode yielded free POP ions from -11 to -18 charge state (Figure 64 A). Of interest, only two distributions were observed in the native mobilograms of free POP (Figure 64 B). The species with lower t_d (smaller CCS) was predominant in the -11 state, while in the -12 state the two species were present. From -13 to -18 states, only the species with higher t_d (larger CCS) was present. Unfortunately, no quantitative data were extracted from these experiments; the absence of a calibration framework for negative ions avoided the calculation of exact CCS values.

Of interest, active site-directed inhibitor ZPP caused significant alterations in the populations of these two species. In -12 and -13 states, the presence of inhibitor favoured the population of the species with lower t_d (smaller CCS) with respect to free POP (Figure 64 C).

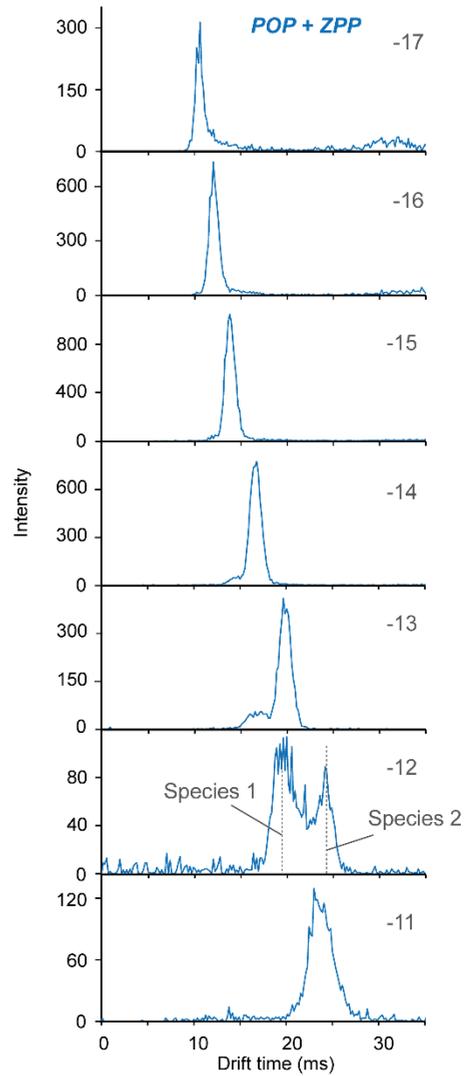
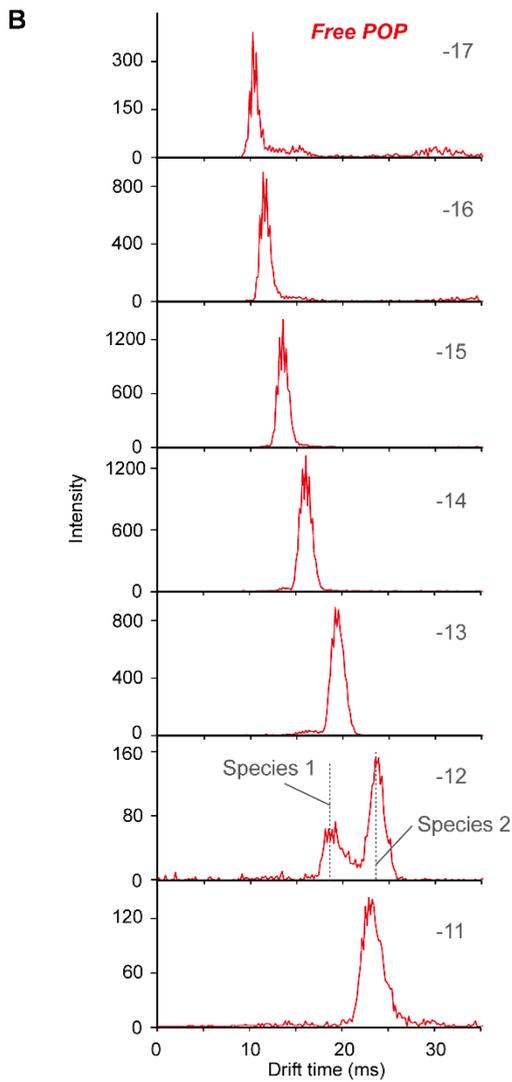
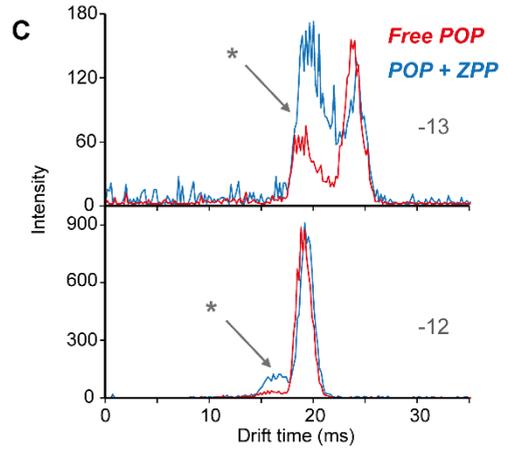
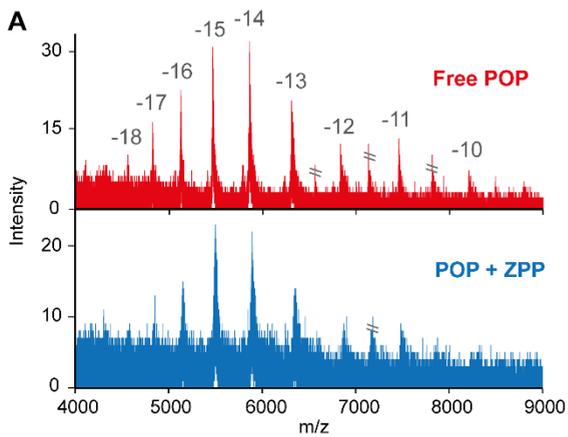


Figure 64 (previous page): IMMS experiments in negative mode. A) Mass spectra of free and ZPP-bound POP, showing the charge state distribution. B) Mobilograms of free (red) and ZPP-bound POP (blue), for charge states from -11 to -17, recorded in a wave height of 11. -12 state shows the two species of different t_d (marked in dotted line for more clarity). C) Effect of active site-directed covalent inhibitor ZPP on the population of the species with lower t_d (asterisk), for -12 and -13 states.

POP conformational landscape in the gas phase: CIU, acid denaturing and reduced charge states

In order to have a qualitative estimation of the thermodynamic stability of the species observed by IMMS, we carried out CIU experiments by increasing the trap CE voltages. The relative populations of the species of the representative state +19 were monitored as a function of trap CE voltages (i.e. the CIU fingerprint). Figure 65 A and B shows the CIU fingerprint of free POP. The relative populations of *A* and *B* species decreased as a function of the voltage until they totally interconverted into other species at 25 V and 30 V, respectively. In contrast, *C* species population steadily growth until 20 V. However, at 25 V a fourth species appeared (*D*, with a high CCS value of $65.8 \pm 0.4 \text{ nm}^2$), and at 30 V, a fifth species raised up (*E*). From 25 to 35 V, the relative populations of these new species increased, while *C* decreased significantly. Unexpectedly, we did not observe any significant change in the CIU fingerprint of POP bound to active site-directed inhibitors ZPP and KYP-2047 (Figure 65 B); usually, thermodynamic stabilization by ligands is reflected by CIU fingerprint [161, 163].

Acidic pH (1 % formic acid, pH approximately of 3) had profound impacts on the mass and ion mobility spectra of POP. A new distribution of highly charge states appeared beyond +40 state (Figure 62 A). Moreover, all charge states showed a single broadened peak in the mobilogram, which indicates large structural heterogeneity of POP at acidic pH (Figure 66 A). The centre of the peaks were directly taken as t_d ; the CCS from +17 to +21 states were comprised 50.8 ± 0.1 to $53.9 \pm 0.5 \text{ nm}^2$, respectively (Figure 66 B). Interestingly, the CCS from +22 to +27 states growth prominently up to $66.8 \pm 0.7 \text{ nm}^2$, indicating a dramatic loss of tertiary structure in highly charged species. Unfortunately, the calculation of CCS could not exceed +27 state due to calibration extrapolation.

In the presence of 0.05 % imidazole, the charge state distribution of free POP was shifted from +11 to +21 (Figure 62 A). Moreover, this charge reducing agent caused slight differences on the relative populations of *A*, *B* and *C* species as a function of charge state (Figure 67 A and B). *A* was the only species at +15 and +16 states. *B* species was more

populated in +17 state in contrast to what happens in free POP (this allowed the quantification of *B* by Gaussian decomposition at +17 state). Finally, *A*, *B* and *C* species were observed from +18 to +20 states. In spite of the tiny increase in the population of *B* species at +17 charge state, the overall relative populations of all species between +17 to +20 states did not show relevant differences compared to free POP.

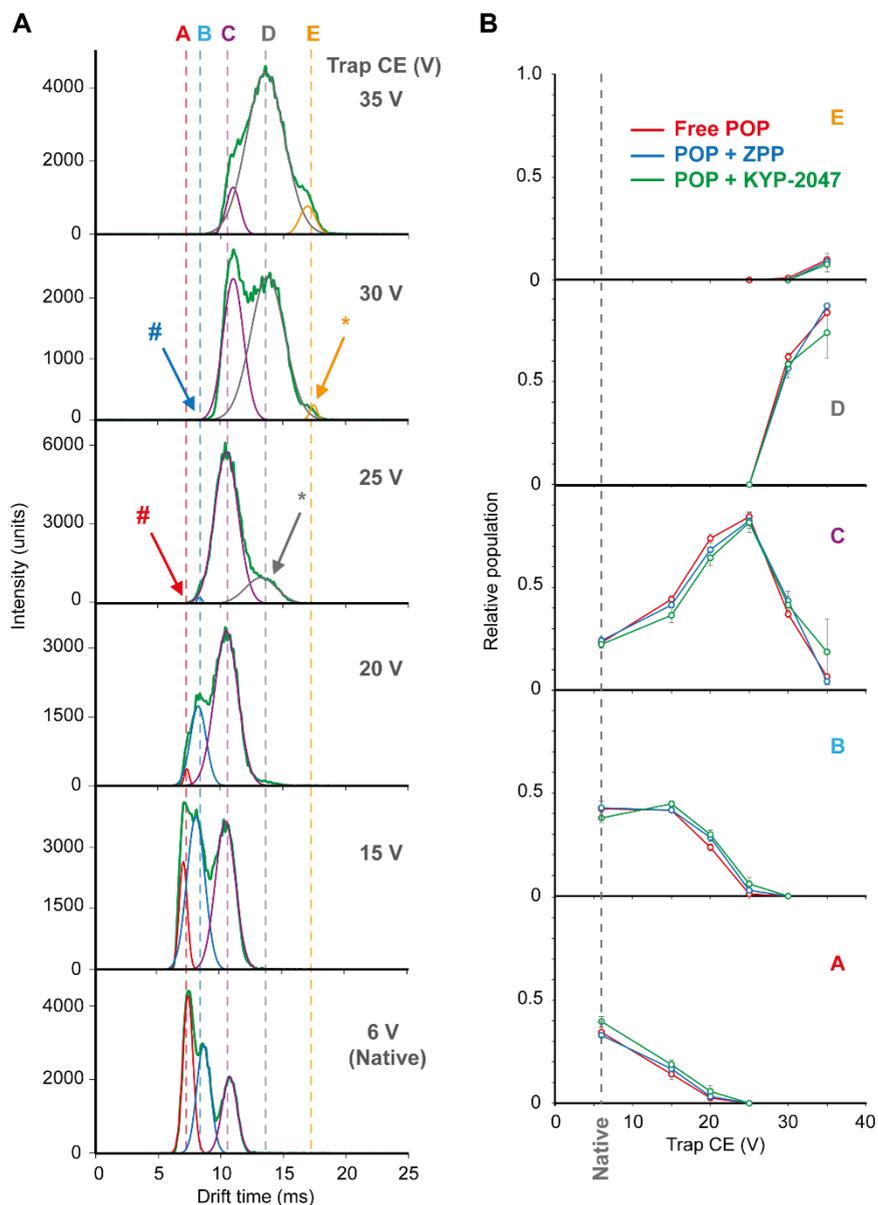


Figure 65: Collision induced unfolding experiments. A) Representative mobilograms of +19 state of free POP (green) at wave height 9 and at different trap CE voltages. The Gaussian distributions of

the species are displayed according to the colour scale shown in the top of the figure. The arrows correspond to the conditions where species appear (asterisk) or disappear (hash). Vertical dashed lines mark the approximate drift time of each species. B) Relative populations of all species of +19 state as a function of trap CE voltage (CIU fingerprint). The results are shown for free POP (red), ZPP-bound POP (blue) and POP bound to KYP-2047 (green). The trap CE voltage corresponding to the native state (6 V) is marked in gray dashed line.

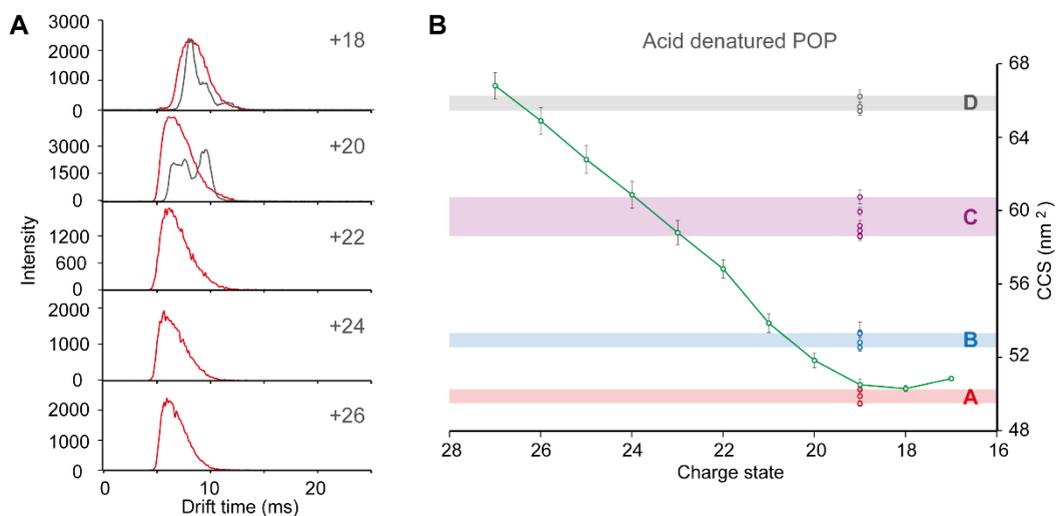


Figure 66: IMMS experiments of POP at acidic conditions. A) Some representative mobilograms of acid denatured POP at different charge states (red), showing a single and broadened peak. Native mobilograms of free POP have been included for comparison (gray). B) The CCS from +17 to +27 states are depicted in green. The CCS of A, B, C and D species of the +19 state obtained by CIU experiments are overlaid according to the colour scale shown in the figure; the coloured regions delimits the interval between minimum and maximum values.

However, the small amount of imidazole caused significant alterations on the structure of gaseous ions. Calibration allowed us to extract the CCS from +15 to +20 (the low intensity of +21 state did not result in reliable t_d values). As it can be seen in Table 10 and Figure 67 C, the average CCS of A, B and C species were similar to that corresponding of free POP. However, the dependence of the CCS with the charge was much higher if imidazole was present. At +15 state the CCS of A species is smaller compared to that of higher charge states, indicating a more compact structure. At +16, +17, +18 and +19 charge states, the CCS of A, B and C species resembles to that of free POP without imidazole. However, at +20 state, CCS of all species were significantly higher compared than in the absence of imidazole.

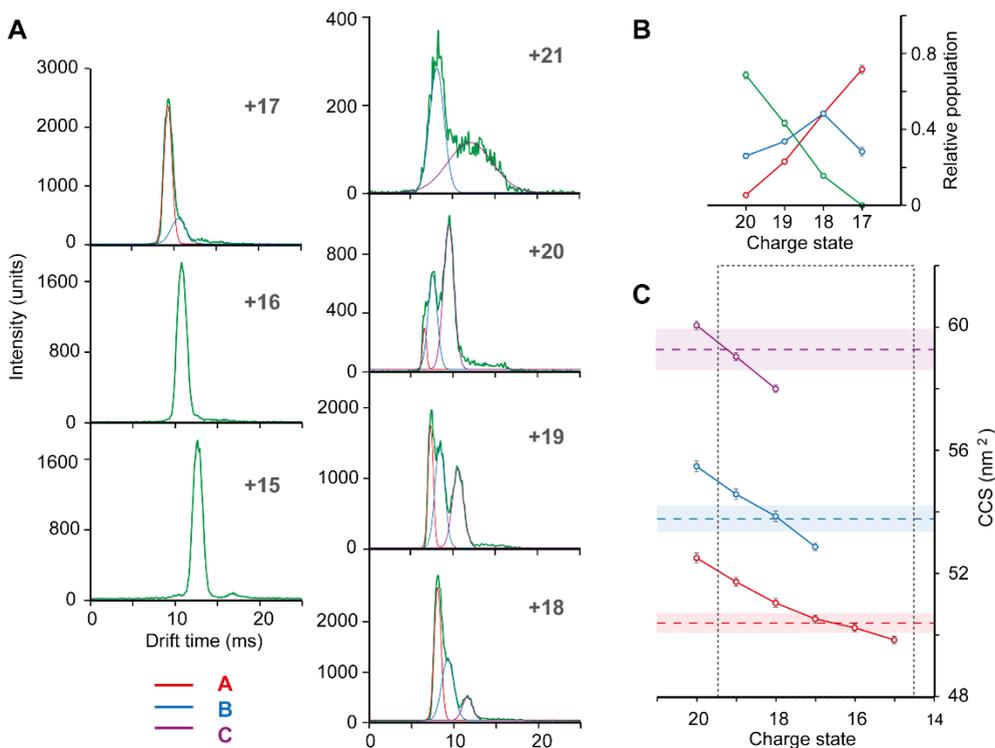


Figure 67: Charge reducing experiments of free POP with 0.05 % imidazole. A) Mobilograms of +15 to +21 states at wave height 9 (green), showing the Gaussian distributions of A, B and C species according to the colour scale displayed at the bottom. B) Relative populations of all species as a function of the charge state. C) CCS of all species as a function of the charge state. Dashed lines mark the CCS derived from native experiments averaged for all states (the coloured regions correspond to the standard deviation). States from +15 to +19 were chosen as reliable for charge reducing experiments (gray dotted box).

DISCUSSION

IMMS experiments of free POP and POP bound to inhibitors (ZPP and KYP-2047) showed three species *A*, *B* and *C* depending on the charge state. The CCS of these species indicated that they correspond to three different POP gaseous conformations. All species displayed highly similar CCS values between free POP and POP bound to KYP-2047, indicating that they resemble tertiary structures. In the case of ZPP-bound POP, slightly smaller CCS were systematically observed in all charge states, suggesting that minor rearrangements are favouring the compaction of POP structure. However, given the small variations of the final

CCS (approximately 1 %), these rearrangements are not expected to have relevant impact on POP structure.

In general, the CCS of all species increased steadily as a function of the charge state. This phenomenon is originated by the higher Coulomb repulsion experienced by gaseous ions in the absence of solvent shielding, which causes a gradual expansion of the POP structure as the overall charge increases [139, 188]. The comparison between the CCS values of free POP obtained experimentally and from simulations shows small divergences [174]; these artefacts are probably derived from the approximations of the computational method used for the calculation of theoretical CCS [173].

Table 10 shows that the CCS of *B* species is similar to that of the closed structure of POP obtained by REMD I simulations. Hence, this species would correspond to the metastable closed conformation of POP in solution. In turn, the CCS of *A* species is comprised between the values of gas-phase simulated POP ions and solution simulated structures. This indicates that *A* might correspond to an intermediate structure between the native and the gas phase stabilized structures. The analysis of gas-phase POP ions disclose that the overall tertiary structure is stable in vacuum, but several local changes take place. Especially, structural rearrangements occur in the N- and C-terminal regions and the β -propeller. Hence, it can be speculated that *A* correspond to a partially collapsed conformation in the gas phase. Finally, *C* species presents higher CCS compared to that of REMD III, but in good agreement to the open conformations derived from SAXS data. This indicates that *C* corresponds to a folded conformer displaying an interdomain separation similar to that found in solution.

The population of *A* decreased as a function of the charge state, and persisted only up to +20 state. Probably, the partial collapse of *A* species does not occur due to the increased Coulomb repulsion at +20 state. The cooling mechanism of imidazole during ionization supports this observation: in the presence of this compound, the charge reduced +15 and +16 states are exclusively formed by this collapsed form [149, 183]. However, the results obtained by charge reducing experiments are only valid at lower charge states. The presence of imidazole destabilizes higher charge states, causing significant structural expansions of POP ions. This behaviour has been documented in the literature for different charge reducing solution additives [189].

In turn, the proportion of *C* increased with the charge state as a consequence of the higher exposition of charged residues in extended conformations [139, 144, 190]. The electrostatic repulsion at high charge states would favour the interdomain opening occurring in *C*

species. In this regard, it can be hypothesized that the protonation of a certain residue(s) triggers determined conformational rearrangements which lead to interdomain separation. Of interest, this would be coherent with the salt bridge network between Asp 149 - Arg 643 and Lys 172 - Asp 642 proposed in chapter 2 on the basis of SAXS and MD simulations.

CIU experiments expanded the energetic barriers of POP conformational landscape and allowed the detection of unfolded forms. A hypothetical conversion pathway of gaseous POP structures is proposed with these data [175]. *A* and *B* species would convert first to *C* under collisional activation. Raising the collisional activation of *C* leads to the highly extended *D* and *E* species. The high CCS of *D* and *E* species indicates that certain regions or domains of POP undergo unfolding. Together, the harsh conditions required for unfolding under collisional activation supports the stability of the solution-like structures *B* and *C* in the gas phase, and corroborates the mild conditions used in native IMMS experiments.

The unfolding processes observed in gaseous ions by CIU experiments can be correlated to some extent to what occurs in acid denatured POP. The large distribution of highly charged ions in the mass spectra of POP at acidic pH is characteristic of elongated, unfolded conformers showing a large number of exposed charged residues [139, 190]. The CCS is also coherent with this observation, especially at the highest charge states: in the case of +27 state, the measured CCS was $66.8 \pm 0.7 \text{ nm}^2$, indicating a dramatic unfolding as a result of acidic conditions and the intense Coulomb repulsion. Of interest, this CCS value is closely similar to that of *D* species obtained by CIU experiments ($65.8 \pm 0.2 \text{ nm}^2$ for +19 state), suggesting that acidic conditions and collisional activation might lead to similar unfolded states. Regarding +17 to +21 states, the comparison of CCS with those obtained in native conditions shows that these species are formed by molten globule structures reminiscent of *A* and *B* species; the CCS of denatured species span from 50.8 ± 0.1 to $53.9 \pm 0.5 \text{ nm}^2$, while average CCS of native POP are 50.4 and 53.8 nm^2 for *A* and *B*, respectively. Moreover, the significant broadening of the mobilograms indicate that POP at acidic pH is a heterogeneous ensemble of loosely packed structures. The irreversible loss of native POP structure and dynamics is coherent with the absence of catalytic activity on acid denatured POP.

Unexpectedly, inhibitors did not cause any relevant differences in POP structures and populations derived from native IMMS experiments. The relative populations of *A*, *B* and *C* species in the gas phase remained unaltered in the presence of such compounds, which is not coherent with the alterations of POP conformational dynamics caused in solution by inhibitors observed by NMR and SAXS experiments. CIU experiments have been used in the

literature to study the stability of protein tertiary structures in the presence of ligands or cofactors [161, 163, 176, 191]; however, inhibitors did not modify CIU fingerprint of inhibitor-bound POP. A possible explanation for this behaviour might be the altered balance of hydrophobic and electrostatic effects in the gaseous environment, which would be hidden within the thermodynamic stabilization by active site-directed inhibitors. In this regard, our preliminary results of POP bound to ZPP obtained by IMMS experiments in the negative mode showed a certain stabilization of the more compact conformer if the inhibitor was present. This result would be coherent by the stabilization of the closed conformers by active site-directed inhibitors observed in solution by SAXS and NMR experiments.

In summary, in this chapter we have successfully set up the conditions to analyse native-like POP ions in the gas phase by IMMS. Our results partially supports the capacity of IMMS to study the conformational landscape of POP: under the mild conditions used in our native experiments, two native-like species, *B* and *C*, preserved the structural features resembling the closed and open conformations of free POP in μ s-ms exchange in solution, respectively. However, sample preparation and instrumental parameters strongly influenced the quality of the data. For this reason, the careful optimization of the experimental IMMS conditions was crucial. In turn, the use of a proper calibration framework was necessary to avoid calibration errors in the CCS calculation. Experimental CCS values were fundamental to interpret the data, comparing them with theoretical CCS values obtained with MD simulations or with other experimental techniques. IMMS was not exempt of artefacts derived from the gaseous environment; although it was not possible to eliminate completely these artefacts, CIU experiments, together with charge reducing and acid denaturing of POP gaseous ions, facilitated their detection. However, the most important disadvantage of IMMS for the study of POP conformational landscape was the lack of structural and thermodynamic effects in the presence of active site-directed inhibitors. We believe that this drawback arises from the altered balance of hydrophobic and electrostatic forces in the gas phase, rather than the capacity of IMMS to discern different conformers and conformational populations derived from the binding of ligands.

According to recent literature, the use of IMMS in the study of dynamic systems is still under development [192]. Most of the investigations have been centred in IDPs (e.g. [169]), and few examples are found for dynamic structured proteins. Several positive characteristics stand the versatility of IMMS for this application. First, the relatively fast time-scale allows the rapid measurement of low resolution structures with minimal exchange. Second, IMMS is a highly sensitive separation technique which allows the

detection of minor populated species. Moreover, this technique is not limited by structural size and complexity. A wide range of dynamic biomolecules and complexes are tolerated by IMMS, but the use of theoretical approaches becomes essential for the interpretation of the results as the complexity of the system increases. Finally, the relatively fast experimental measurement and the low sample consumption stand the potential of IMMS for the screening of ligands.

Nevertheless, the main limitations of this technique for the study of dynamic systems are the intrinsic gas-phase derived artefacts. The altered balance between hydrophobic and electrostatic forces in the gaseous environment causes structural and thermodynamic rearrangements of different extension [171]. Hence, careful optimization of experimental conditions is crucial for the reliability of the results. In turn, extending the use of stabilizing buffering agents or ions will contribute to the preservation of solution native structures in the gas phase [193]. Regarding the measurement of non-covalent interactions, biased results have been reported as a consequence of the modified balance of binding energy in the gas phase [140]. Measuring quantitative thermodynamic parameters in the gas phase is inexorably affected by the energetic balance of the gaseous environment. Hence, the detection of non-covalent interactions of different affinities in the gas phase is one of the main challenges in native MS that will be addressed during the next years [194-196].

Chapter 5: interaction of POP with α -synuclein

The study of the interaction of POP with α -synuclein was the closing project of this PhD thesis. Here we applied all of the knowledge acquired on the conformational dynamics of POP to assess the biologically relevant question of its interaction with α -synuclein. Of interest, the involvement of POP in Parkinson's disease (PD) has gained much attention during the last years. A growing number of studies have been published indicating a possible direct interaction between POP and α -synuclein, the main protein component of Lewy bodies in PD. For this reason, we decided to explore the possible interaction between the two proteins *in vitro* by NMR, and describe the effects of active site-directed inhibitors in this recognition event.

INVOLVEMENT OF POP IN α -SYNUCLEIN CLEARANCE: POSSIBLE INTERACTION OF POP WITH α -SYNUCLEIN

α -synuclein is a 14 kDa intrinsically disordered protein (IDP) which is mainly present in the brain. The biological role of this cytosolic protein is not clear, but probably is involved in synaptic transmission, axonal transport and the regulation of the dopaminergic pathway [197]. This protein has 140 residues, and the sequence can be divided in three different sections: the N-terminal amphipathic region (residues 1-60), the central non-amyloid- β component region (NAC region, a highly aggregation-prone fragment formed by residues 61-95) and the highly acidic C-terminal region, rich in proline residues (residues 96-140) [198].

However, α -synuclein is mainly involved in the pathogenesis of a family of neurodegenerative diseases called synucleinopathies, the most spread being PD [199]. Lewy bodies are the pathological hallmark of PD, which consists in inclusions of insoluble fibrils formed majorly by α -synuclein [37]. These filamentous aggregates show high antiparallel β -sheet content, and similar structural features to amyloid fibrils [200]. Lewy bodies are accumulated in the intracellular space, and probably represent the last stage of the toxic processes involving α -synuclein oligomers.

α -synuclein is also found in pathophysiological conditions in a variety of low- and high-molecular weight soluble oligomers [201]. Several evidences suggest that these soluble and dynamic oligomers constitute the *in vivo* neurotoxic species [202, 203], and Lewy bodies represent the last stage of α -synuclein oligomerization. The causes that trigger α -synuclein accumulation and subsequent aggregation are still unclear, but probably involve

alterations in the synthesis, aggregation and clearance of α -synuclein [204]. Especially, three mutations were found to trigger the early onset of familial PD: A30P [205], A53T [206] and E46K [207]. The kinetic mechanism of α -synuclein fibrillogenesis is a nucleation-dependent polymerization process [208]. It has been found that several post-translational modifications of α -synuclein alter the oligomerization and aggregation rates, and consequently, the toxicity of these modified species (for a comprehensive review, see [209]). Moreover, it is described that the aggregation of α -synuclein can be modulated by interactions with other proteins. Of interest, POP is among these interacting proteins [210].

Brandt *et al.* searched for possible endogenous POP substrates in porcine brain homogenates, and they found a small peptide YQDYEP as one of the POP cleavage products [31]. This fragment matched well with residues 133-138 of the C-terminal part of α -synuclein, YQDY**PEA** (the cleavage site is shown in bold). Although further studies of the same group showed no significant hydrolytic activity of POP on α -synuclein, they found an accelerated α -synuclein aggregation in the presence of minor concentrations of POP [210]. Of interest, the acceleration of the aggregation was prevented if the inactive POP mutant S554A or two active site-directed inhibitors were present, indicating that the interaction would be sensitive to changes on the active site. This study suggested that a direct interaction might occur between the two proteins, facilitating the nucleation α -synuclein. The hypothesis of an *in vivo* interaction between POP and α -synuclein in the pathogenesis of PD was reinforced when Hannula and co-workers found that the two proteins colocalized in the substantia nigra of post-mortem brain samples of PD patients [211].

Effects of POP inhibition in α -synuclein aggregation

During the last three years, several papers were published in which the effects of POP inhibition on cell survival and α -synuclein levels were analysed *in vitro* and *in vivo*. The study carried out by Myöhänen and co-workers described the effects of POP inhibition on the aggregation of WT α -synuclein and A30P, A53T mutants in cell lines and transgenic mouse strains [43]. Particularly, aggregation was induced in cells by forcing oxidative stress conditions, followed by KYP-2047 administration. This compound prevented the colocalization between POP and α -synuclein and decreased the levels of aggregates, which improved cell survival. Similarly, transgenic mice expressing the α -synuclein mutant A30P showed decreased protein levels and aggregation after administration of KYP-2047. Of interest, these results suggested an involvement of POP in the *in vivo* clearance mechanisms of α -synuclein.

This hypothesis was later supported by the study of Savolainen *et al.* in transgenic mice including the A30P mutation in the α -synuclein coding gene [212]. The chronic administration of KYP-2047 inhibitor reduced α -synuclein aggregation. Monitoring of the autophagosome markers suggested that this reduction was caused by increased autophagy (Box 5): the levels of the marker LC3BII were increased in the treated animals, and beclin I levels (a protein regulator of autophagy [213]) raised in HEK-293 cells treated with the same inhibitor; moreover, it was observed that the accumulation of α -synuclein impaired the proteasomal activity. A further study was performed in neuroblastoma cell lines overexpressing the α -synuclein mutant A53T and exposed to the mitochondrial toxin rotenone [214]. Rotenone induces mitochondrial damage by increasing reactive oxygen species (ROS), which results in a significant α -synuclein aggregation. Of interest, in these conditions KYP-2047 inhibitor decreased the presence of α -synuclein aggregates and reduced the formation of ROS.

However, the very recent results of Savolainen and co-workers demonstrated the direct interaction between free POP and α -synuclein *in vitro* and in living cells. Using protein-fragment complementation (PCA, Box 6), they shown that WT POP and the catalytically inactive mutant S554A increased α -synuclein dimerization, contrarily to the preliminary results obtained by Brandt *et al.* [210]. Moreover, in this recent publication it was showed that the inhibitor KYP-2047 reduced dimerization only in the WT protein. The binding affinity obtained by microscale thermophoresis (MST, Box 6) was found in the low μ M range, and the C-terminal fragment of α -synuclein was not required for binding. The authors speculated that the α -synuclein dimers formed by the interaction with POP would act as nucleation centre, similarly to what was proposed by Brandt *et al.* in 2008 [210].

According to all of these evidences, it was postulated that POP would have a dual effect in α -synuclein aggregation (Figure 71). On the one hand, the free form of POP would promote dimerization of α -synuclein, which could act as a nucleation point for aggregation. The accumulation of α -synuclein aggregates would interfere *in vivo* with the UPS clearance system, hampering the catabolism of α -synuclein and leading to higher accumulations of this protein. On the other, the dimerization of α -synuclein is not likely to occur in the presence of POP inhibitors, and besides, the *in vivo* clearance of α -synuclein aggregates would be favoured. Together, these facts stress the potential chaperone activity of POP and its involvement in α -synuclein aggregation, which can be modulated by small molecule inhibitors.

Box 5: Clearance mechanisms of pathogenic proteins.

The accumulation of misfolded or aberrant proteins in neurons might lead to neurodegenerative diseases. In the case of PD, the accumulation of α -synuclein aggregates triggers pathogenic conditions. Two cellular mechanisms are responsible for eliminating pathogenic proteins: the ubiquitin-proteasome system (UPS, Figure 68 A) and the autophagy-lysosomal pathway (ALP, Figure 68 B).

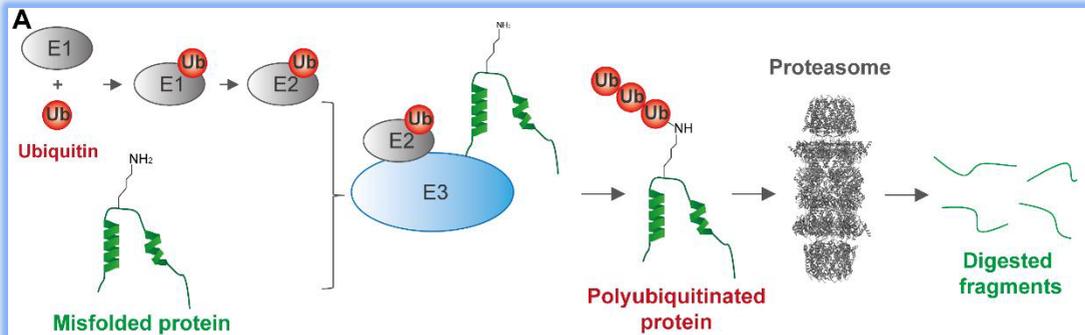


Figure 68: A) UPS degrades soluble proteins in the cytoplasm, nucleus and endoplasmic reticulum. This process starts with the ubiquitination of the protein with three different enzymes (namely E1, E2 and E3). Afterwards, ubiquitinated proteins are recognized by the proteasome, which hydrolyses them to small peptides.

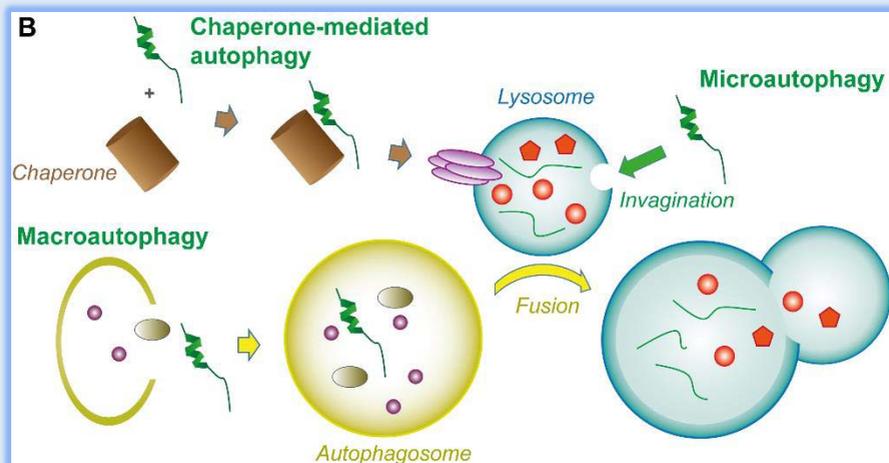


Figure 68: B) ALP degrades bulky cytosolic components in the lysosomes, which are vesicles containing proteases, lipases and glycosidases (represented by orange symbols). The components to be degraded are captured by three different mechanisms: microautophagy (direct invagination), macroautophagy (via organelles called autophagosomes) or through chaperone-mediated autophagy.

Box 6: Interaction of POP with α -synuclein detected by protein-fragment complementation (PCA, Figure 69) and microscale thermophoresis (MST, Figure 70) [215].

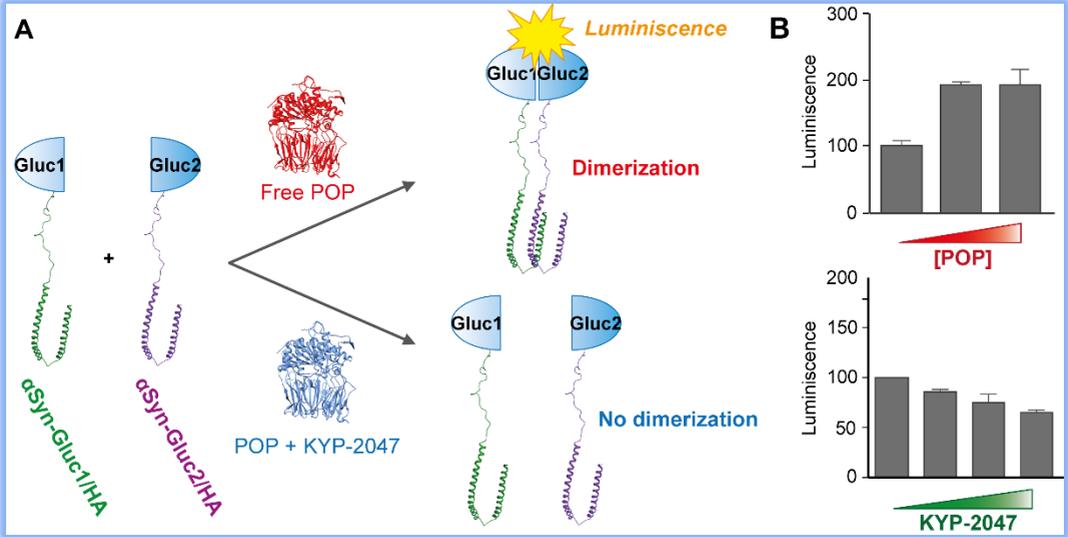


Figure 69: A) PCA is based on the measurement of luminescence originated from the contact between two complementary fragments of a protein reporter attached to the two partners under study. In the present case, two complementary fragments of *Gaussia princeps* luciferase (Gluc 1 and Gluc 2) were attached to α -synuclein. B) In the presence of free POP, dimerization was favoured. If POP was inhibited with KYP-2047, the dimerization was prevented.

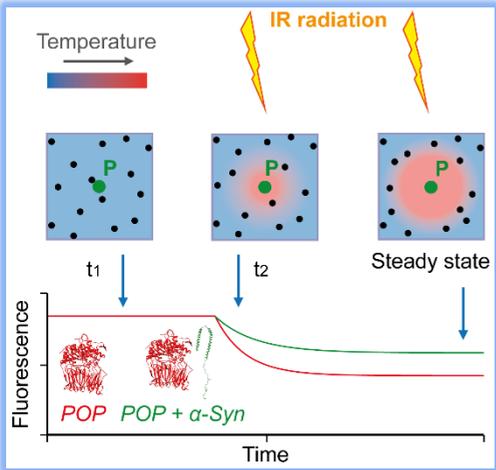


Figure 70: MST is based on the movement of particles in a concentration gradient, typically induced by a local heating by and IR radiation. This movement is directly related with the solvent environment and the size of the system, which are altered if binding occurs. Performing a titration assay by thermophoresis allows to determine the binding affinity. The local concentration is typically monitored by attaching a fluorescent label in one of the interacting partners, and the fluorescence is measured at a point *P* as a function of irradiation time.

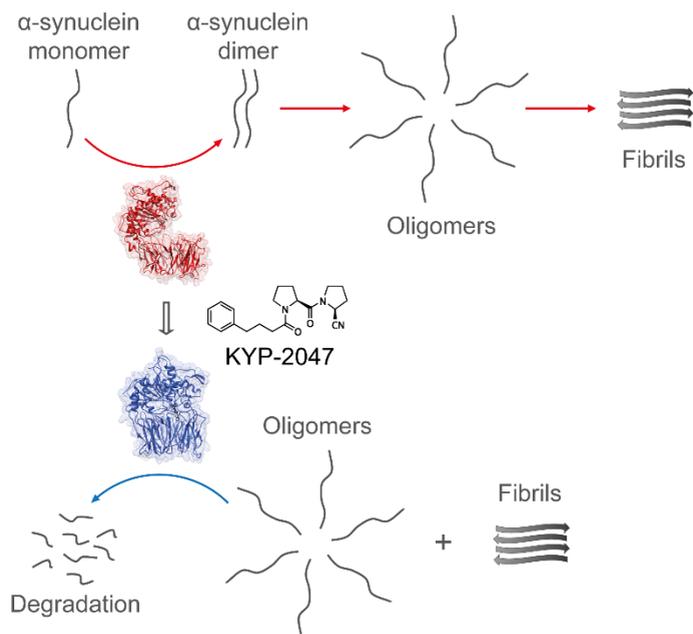
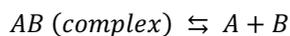


Figure 71: Suggested dual effect of POP in aggregation of α -synuclein. The free enzyme would promote α -synuclein aggregation and fibrillization, while the inhibited form of POP would facilitate the clearance of α -synuclein aggregates.

STUDY OF PROTEIN-PROTEIN INTERACTIONS

Recognition processes between proteins are crucial in living systems, especially those involved in the regulation and signalling of biological processes. For this reason, recognition processes have been intensively studied during decades. However, PPIs span a wide range of affinities, as well as structural and dynamic complexity. For this reason, the study of the interactions between proteins might represent a challenging task.

The properties of the molecular recognition processes can be divided according to the affinity of the complex, which is dictated by the dissociation constant K_d . If we suppose a system formed by A and B species, K_d is defined as follows:



$$K_d = \frac{[A][B]}{[AB]} = \frac{k_{off}}{k_{on}}$$

Where k_{on} and k_{off} are the rate constants of the association and dissociation processes, respectively. Hence, the interactions span from strong and permanent ($K_d < 10^{-9}$ M) to weak and transient ($K_d > 10^{-4}$ M). According to the study of Savolainen and co-workers, the interaction of free POP with α -synuclein would be weak (K_d value of 1.41 μ M). For this reason, we will focus our discussion on this kind of interactions.

In turn, the dynamics of the binding events also influences the properties of the system. For this reason, the study of PPIs involving natively unfolded proteins represent a particular case. Instead of discrete conformational states, IDPs are sampling a conformational landscape formed by a continuous distribution of microstates; as a consequence, the binding affinity of IDPs with other partners is rather correlated with the dissociation rate [216]. In some cases, the binding of disordered proteins to structured proteins acts in an induced fit mechanism, involving a coupled binding and folding events. As a result, IDPs adopt a partially defined conformation in the bound form (these recognition events are known as disorder-to-order transitions [217], Figure 72). In contrast, in other cases the complex might retain the intrinsic dynamics of the interacting partners [218]. The study of these dynamic complexes, which are commonly known as "fuzzy complexes" [219], require more elaborated methodologies compared to that of classical static complexes (Figure 72). However, the dynamic and transient nature of fuzzy complexes might represent a technical challenge for the detection *in vivo* or *in vitro*.

Possible NMR experiments for the detection of the interaction between POP and α -synuclein

Several experimental methods have been used to detect structural changes of α -synuclein as a consequence of interactions with different partners, for instance micelles or lipids. Some of these techniques were single-molecule fluorescence resonance energy transfer (FRET) [201], circular dichroism and Fourier-transform IR spectroscopy [220], among many others. In this work we used NMR due to the high versatility for the *in vitro* detection of PPIs.

NMR allows the analysis of recognition events in a wide range of affinities, providing information of the contact at atomic resolution [221]. Here we summarize the possible NMR experiments that were considered for the study of the interaction between POP and α -synuclein. In our bibliographic search, we focused on NMR approaches sensitive to weak and transient interactions. However, we restricted our search and discarded solid-state NMR

approaches and other elaborated NMR applications involving paramagnetic probes, which are out of the research line of our laboratory.

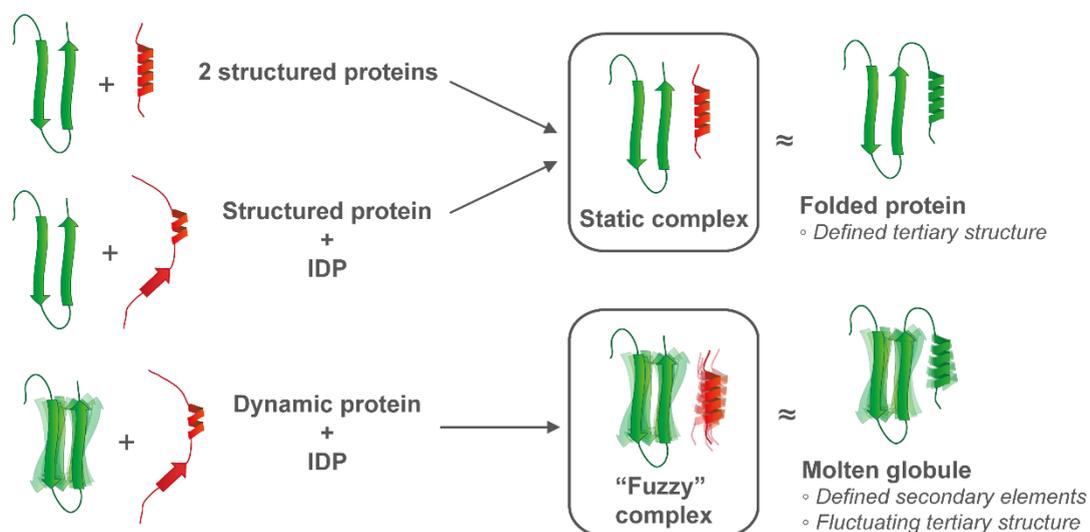


Figure 72: Different types of complexes according to the dynamic properties of the interacting partners and the resulting complex. Complexes involving structured proteins, or undergoing disorder to order transitions, yield static complexes. Such structured complexes can be assimilated to folded proteins. However, interactions involving IDPs and dynamic proteins or proteins with disordered regions, might result in “fuzzy” complexes. These complexes show secondary structure elements but fluctuating tertiary structure, similar to molten globule structures [218].

Chemical shift perturbation

The most useful and direct experiment for the detection of PPIs is chemical shift perturbation (CSP) mapping. This method benefits from the alteration of the chemical environment of certain atoms as a consequence of the interaction, which causes differences in the chemical shift of the NMR signals. The formula to calculate CSP is shown in Equation 1 (chapter 2). The representation of CSP as a function of residue number or over the tertiary structure discloses the regions involved in binding.

The most common approach involves the ^{15}N -labeling of one of the partners, which is titrated with the unlabeled partner. In the case of weak PPIs in fast exchange (high k_{off}) the CSP might be hardly detectable; in these cases CSP can be forced by using a high excess of the unlabeled partner. In IDPs, the low chemical shift dispersion of the signals might hamper the detection of CSP. In order to overcome this limitation, the simultaneous

measurement of CSP involving ^1H , ^{15}N and ^{13}C obtained by 3D NMR experiments is a good solution. For instance, the CSP mapping have been successfully used for the detection of α -synuclein regions interacting with aggregation inhibitors [222].

However, if a weak interaction occurs, CSP might not be optimal due to the low and transient concentration of the complex. For this reason, a modification of this approach can be carried out by mapping the ratio of intensities between the complexed and the uncomplexed signals (I/I_0 ratio). I/I_0 ratio discloses alterations in the populations as a result of a second, not observed, resonance of the complex. Of interest, this method has been widely used in the case of α -synuclein [223, 224].

Approaches based on relaxation properties

Several methods for the detection of weak PPIs are based on the measurement of the NMR relaxation properties of one of the binding partners. In the case of highly dynamic IDPs, these methods provide a higher degree of sensitivity and accuracy in the detection of weak PPIs compared to the CSP approach.

The most spread method consists in the measurement of backbone R_1 , R_2 relaxation and steady-state NOE. The analysis of differential dynamics between complexed and uncomplexed forms disclose alterations in the relaxation parameters at a residue level as a consequence of the interaction [225].

The nature of R_1 , R_2 is described in Appendix I and will not be commented here. The NMR experiments are classical spin-lattice and CPMG measurements respectively. However, special careful must be taken to remove cross-correlation between dipolar and CSA relaxation mechanisms [226] and to eliminate the interference of water without interfering with the measurement [225].

In order to extract quantitative data of the binding dynamics, the NMR data can be analysed by the reduced spectral density formalism. This model provides an estimation of the rotational motions of the NH bond vectors at a frequency $\omega = 0$, ω_N and ω_H (ω_X is the Larmor frequency of the corresponding nucleus) [227, 228]; however, a deeper description of this models is out of the scope of this manuscript.

In turn, NOE effect arises from the coupling between magnetic dipoles. If we suppose a ^1H - ^{15}N spin system in which one of the spin transition is saturated by a radiofrequency pulse during a long time, additional transitions take place to redistribute the spin populations (Figure 73). Such mechanisms strongly depends with the distance in a r^{-6} law (r is the ^1H -

^{15}N distance). The difference spectra between the NOE experiment and blank spectra is called $[^1\text{H}-^{15}\text{N}]$ steady-state NOE experiment. This experiment allows to extract fast motions in the ps-ns time scale, which can be affected by weak interactions. However, these experiments are significantly affected by the saturation of water, and the solvent signal has to be eliminated by pulsed gradients [229].

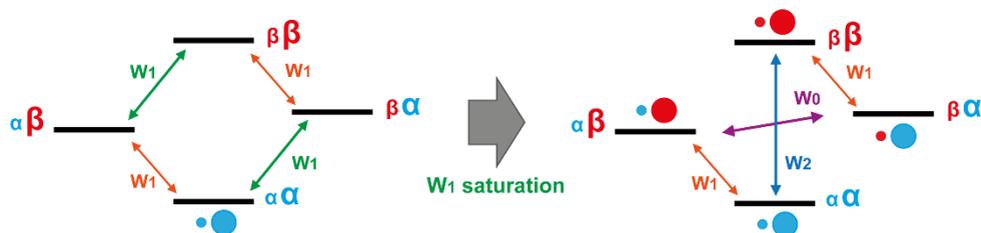


Figure 73: Steady state NOE. A dipolar coupled spin system of two different atoms (represented by spheres of different size) has only W_1 spin transitions in normal conditions (which corresponds to T_1 relaxation). If one of the W_1 transitions is irradiated the populations of $\alpha\alpha/\beta\alpha$ (and $\alpha\beta/\beta\beta$) are equalized. Therefore, W_0 and W_2 transitions takes place to redistribute the spin populations: this effect is the steady-state NOE.

Of interest, R_1 , R_2 and NOE have been successfully used by Bussell and co-workers to quantify the dynamics of WT, A30P and A53T α -synuclein [230], and later by Wu *et al.* [231]. Moreover, this methodology allowed the analysis of the interaction of α -synuclein with micelles [232].

Finally, more complex experiments benefits from the isotopic labeling. Particularly, cross-saturation experiments (Figure 74) [233] stand as a good candidate for the study of the interaction between POP and α -synuclein. This approach benefits from the ^{15}N labeling and the perdeuteration of one of the binding partners. The experiment consists on the non-selective irradiation of the protons of the protonated sample, which become fully saturated by spin diffusion. As a consequence of the interaction, the saturation is transferred to the deuterated counterpart, and the intensity of the signals is affected. However, cross-saturation is limited only to the backbone amide protons directly involved in the interaction. Nevertheless, the use of high D_2O concentrations in the NMR buffer and slightly acidic pH are necessary in order to avoid intramolecular spin diffusion and ^1H exchange, respectively. This method has been successfully applied for the mapping of the interacting surfaces in large and transient protein-protein [233] and protein-nucleic acid complexes [234].

Taken together, it is evident that some of the possible NMR approaches for the study of POP interaction with α -synuclein described above require complex experiments and expensive labeling strategies. Although these methods are robust and highly sensitive, they would require a time-consuming optimization of the experimental conditions and an elaborated data analysis. For this reason, we chosen CSP and intensity analysis of ^1H - ^{15}N HSQC experiments of U- ^{15}N α -synuclein as starting preliminary experiments for the detection of the interaction between POP and α -synuclein by NMR.

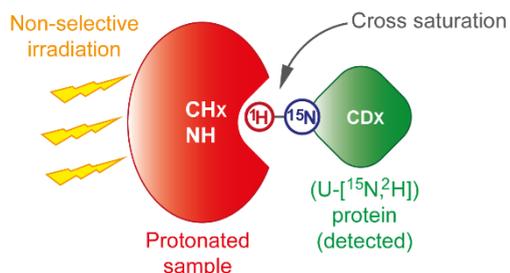


Figure 74: Cross-saturation experiments, involving a conventional protonated protein and a U- ^{15}N , ^2H -labeled partner. The protonated protein is non-selectively irradiated, and the saturation is transferred specifically to the backbone ^1H of the perdeuterated partner which are involved in the interaction.

PRELIMINARY NMR EXPERIMENTS

In our laboratory there was no previous experience in working with α -synuclein. Hence, it was necessary to optimize the expression and purification protocols. Moreover, IDPs have particular properties compared to folded proteins, and for this reason they are not trivial to handle. A detailed search in the literature was necessary before proceeding to the experiments.

α -synuclein expression and NMR sample preparation

Given the elevated prices of commercial α -synuclein, we decided to express the recombinant protein in our laboratory. The protocol for the expression of α -synuclein (see Materials and methods section) was adapted from the procedure described by Carlos W. Bertoncini [235]. The pT7-7 plasmid containing α -synuclein gene was a gift from him.

The expression proceed without difficulties. However, it should be stressed that the addition of Q solution (which has a high content of metals) to the expression media probably caused the partial oxidation of certain positions during the first expressions [236]; for this reason, this solution was not used. Expression and purification of ^{15}N -labeled α -synuclein took place in good yield (30 mg of protein per litre of culture) and good purity (Figure 75 A).

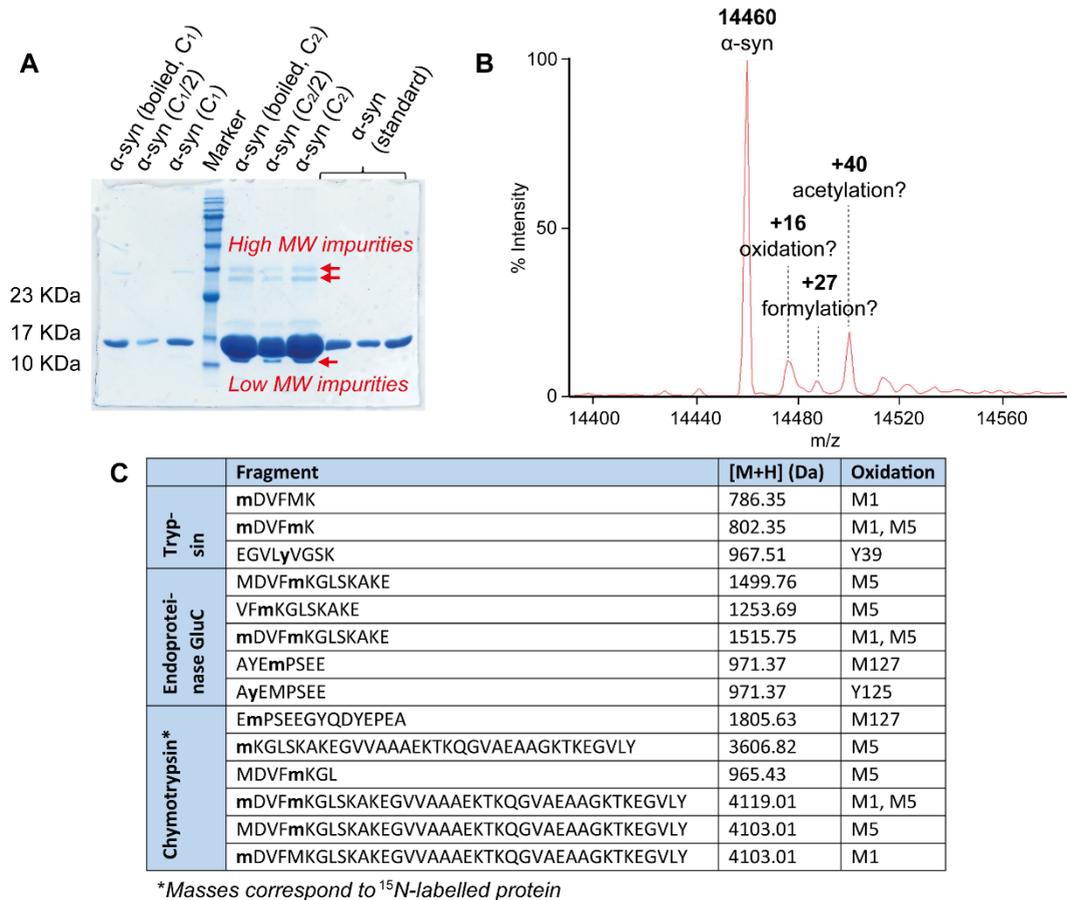


Figure 75: Expression of WT and U-¹⁵N labeled α -synuclein. A) SDS-PAGE of purified α -synuclein at two different concentrations C_1 and C_2 , being $C_2 > C_1$ (some samples were boiled before loading to the gel). Lanes at C_2 show low and high MW impurities, which were partially removed by subsequent steps of filtration and concentration (see Materials and methods). B) Several mass increases are seen in the deconvoluted LC-MS spectrum of α -synuclein, which might correspond to oxidations and other modifications. C) Different samples of α -synuclein were digested by trypsin, endoproteinase GluC and chymotrypsin and subjected to a bottom-up analysis. This analysis showed that all Met were oxidized in a minor proportion (it should be noted that enzyme digestions might cause Met oxidations). Of interest, Tyr 39 (and probably Tyr 125) were found also oxidized. **All MS experiments were carried out by Dr. Marta Vilaseca and Dr. Mireia Díaz (Mass Spectrometry Core Facility, Institute for Research in Biomedicine, Barcelona, Spain).**

The first step of the purification consisted in boiling the whole cell extract obtained by sonication; in order to prevent degradation of α -synuclein by proteases, PMSF was added before sonication. Afterwards, cytoplasmic materials were removed by precipitation, and

finally, α -synuclein was precipitated with concentrated ammonium sulphate. The α -synuclein obtained by this procedure was finally dialysed against the storage buffer, since several papers used purification steps which do not involve chromatographic separation [235, 237]. Moreover, in order to avoid sample heterogeneity, we decided not to lyophilize the final protein. Instead, α -synuclein aliquots were flash frozen in N_2 and stored at $-80\text{ }^\circ\text{C}$ until needed.

However, a minor proportion of oxidized species were always present in the α -synuclein obtained by this protocol according to MS analysis (Figure 75 B and C). Oxidations mainly affected Met residues, but Tyr 39 (and probably Tyr 125) were also found to undergo oxidation. The cause of these oxidations still remains unknown, but this undesired reaction was minimized by using degassed buffers during the purification. The addition of 1 mM *DL*-Met in the purification buffers, which is reported to prevent Met oxidation by radicals [238], did not avoid oxidations.

The scheme for the preparation of samples for NMR experiments is described in Figure 76. Sample preparation started by thawing the sample in ice, and centrifuging it during few minutes in order to remove possible insoluble aggregates. The supernatant was subjected to a buffer exchange to the NMR buffer using PD-10 columns (GE Healthcare Life Sciences, Uppsala, Sweden). Initial experiments used the conditions described by Binolfi *et al.* [224] (MES-Na 20 mM pH 6.5, NaCl 100 mM). Although this method yielded high S/N ratio due to lower amide exchange, we preferred to use a pH more close to physiological conditions. Moreover, slightly acidic pH were found to affect the conformational equilibrium of POP (see chapter 1). According to several publications, we used pH 7.4 [231, 239] in Tris d_{11} -50 mM HCl pH=7.1, NaCl 20 mM, DTT d_6 1 mM, NaN_3 0.03 %, 10 % D_2O buffer (the effect of the temperature on the pH was empirically corrected considering an increment of 0.3 units that were added to the observed pH-meter reading). In parallel, POP was subjected to a buffer exchange analogously to α -synuclein. Afterwards, the two proteins were concentrated with centrifugal filter concentrators to the final concentration required for the NMR experiments, and the two protein stocks were mixed. 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) was added as internal standard. For the experiments with inhibited POP, the same sample of POP with α -synuclein was used. The adjustment of pH was crucial to avoid changes in the pH of the sample due to traces of acid from the inhibitor stock solution; pH was corrected by three consecutive dilution and concentration operations.

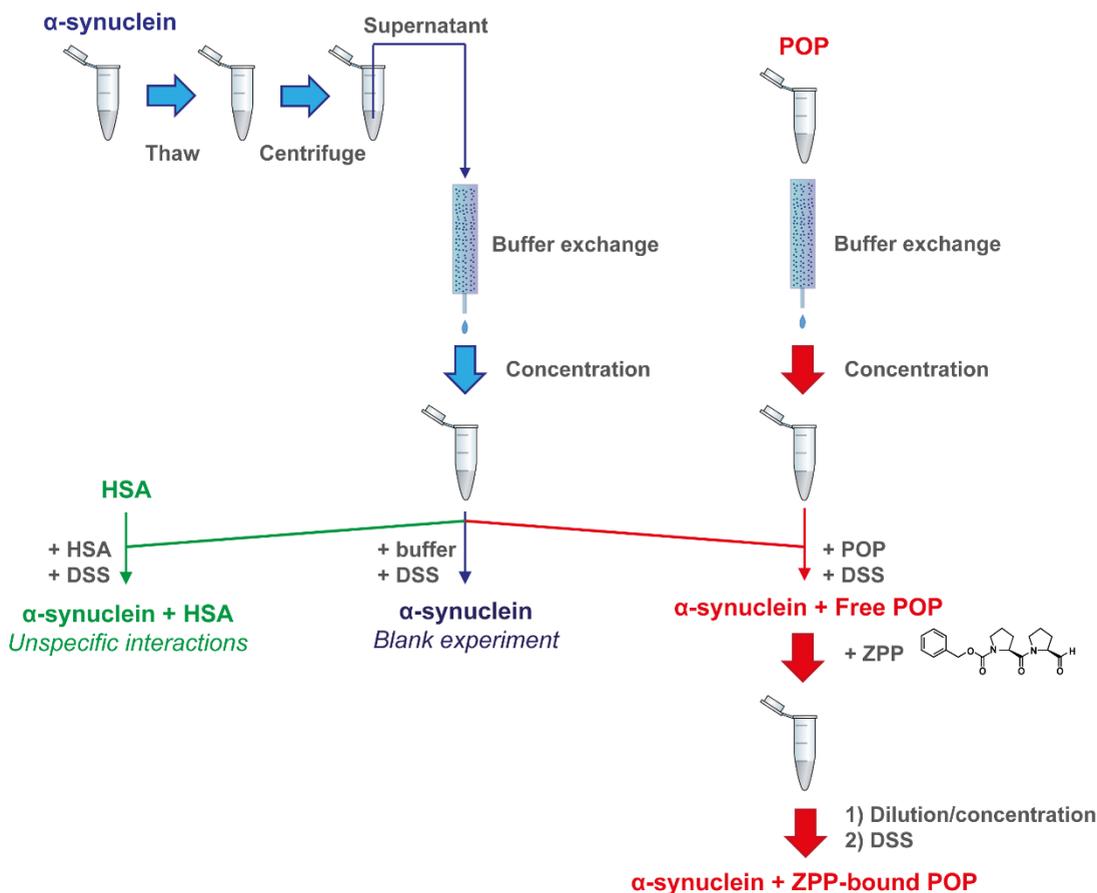


Figure 76: Schematic procedure for the preparation of α -synuclein samples for NMR experiments.

In order to detect if the incubation of POP with α -synuclein caused significant conformational changes, we performed preliminary native polyacrylamide gel electrophoresis (native PAGE) before proceeding to NMR experiments. Previous results in our laboratory disclosed three bands of POP in native PAGE corresponding to different POP conformers. In the presence of inhibitors, these bands collapsed to a single band corresponding to a main conformer [57]. In the present case, native PAGE were performed using 200 mM Tris-HCl pH 8 and 10 % of acrylamide, according to the standard conditions described by Gallagher *et al.* [240]. 10 μ l of approximately 1-2 mg/ml of POP in Tris buffer in 5 % sucrose and 10 μ g/ml bromophenol blue were loaded to each well, and gels were run at 100 V. Native protein ladder was loaded as standard (NativeMark protein ladder, Thermo Fisher Scientific Inc., Waltham, USA). Figure 77 A shows that the electrophoretic band of α -synuclein is comprised between 20 and 66 kDa. This result is coherent to what found by Burré and co-workers [241], and is a consequence of the particular electrophoretic properties of α -synuclein. In turn, Figure 77

B shows native PAGE of free and inhibitor-bound POP in the absence and in the presence of 1 equivalent of α -synuclein; no major effects were observed in the electrophoretic patterns of the two proteins, suggesting that the conformational changes upon binding might not be detectable by native PAGE.

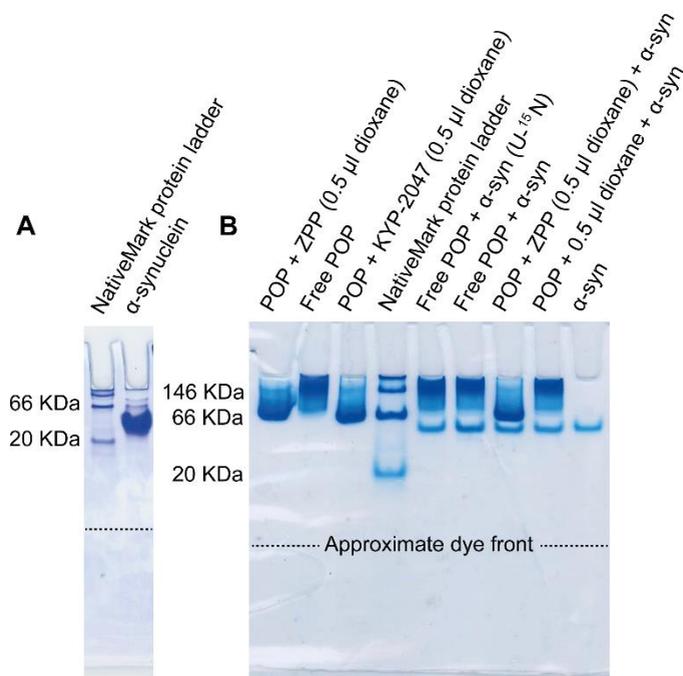


Figure 77: Native PAGE experiments. A) α -synuclein (14 kDa) presents a single band running between 20 and 66 kDa markers (15 % acrylamide). B) Native PAGE (10 % acrylamide) of POP and α -synuclein. The sample conditions are described at the top of the lane; no changes are seen if α -synuclein is present. In A) and B), the dye front is marked by a dotted line.

NMR experiments of [methyl-¹³C]-methionine labeled POP

Given our knowledge on the NMR spectroscopic features of POP, we decided to perform initial ¹H-¹³C-methyl TROSY HMQC experiments of [methyl-¹³C]-Met labeled POP in the presence of α -synuclein. These experiments allowed us to monitor if POP undergo conformational rearrangements induced by α -synuclein easily by observing the 11 Met signals.

Figure 78 shows the spectrum of 180 μ M of free POP overlaid with free POP with 1 equivalent of α -synuclein. Several signals of α -synuclein appeared, corresponding to Met residues with ¹³C natural abundance (α -synuclein has a total of 4 Met residues). However, no generalized

changes were observed in the spectrum, except a slight displacement of Met 583 signal. In the case of the spectrum of ZPP-bound POP overlaid with ZPP-bound POP with 1 equivalent of α -synuclein, no significant changes were detected in the spectrum (including Met 583).

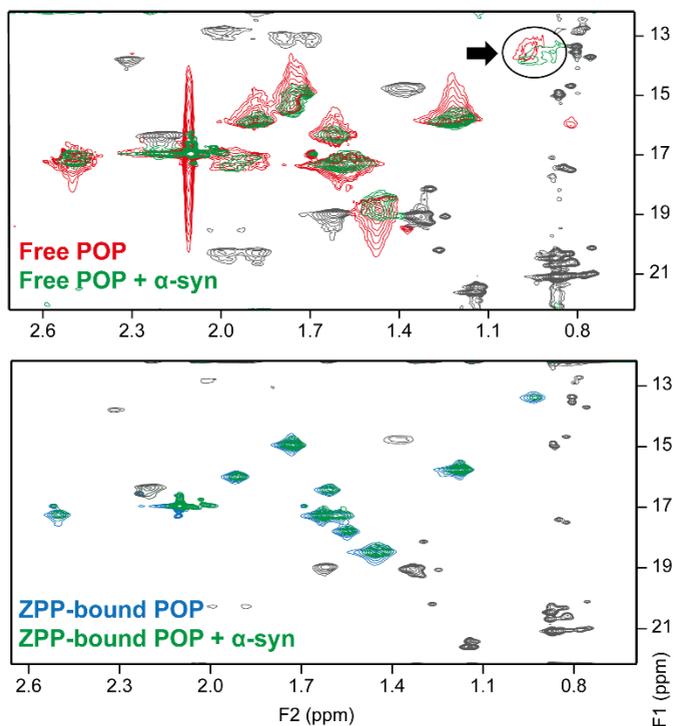


Figure 78: ^1H - ^{13}C methyl-TROSY HMQC experiments of POP in the presence of 1 equivalent of α -synuclein. Top: the overlaid spectra of free POP (red) and free POP with α -synuclein (green) shows no changes in the chemical shifts of the signals, except a minor perturbation of Met 583 signal (black arrow). Down: overlaid spectra of POP bound to ZPP (blue) and POP bound to ZPP in the presence of α -synuclein. No changes are observed. In both cases, the signals corresponding to the Met residues of α -synuclein with ^{13}C natural abundance are shown in gray.

Overall, these results were coherent to what observed by native PAGE. Probably, the interaction of POP with α -synuclein is weak and transient, which precludes the detection by these methods and conditions. Moreover, the 1:1 ratio of POP: α -synuclein might be not optimal for favouring the complex formation. For this reason, we moved to the more sensitive and quantitative ^1H - ^{15}N HSQC experiments of α -synuclein in the presence of POP.

NMR experiments of U-¹⁵N labeled α -synuclein

The high sensitivity and resolution of ¹H-¹⁵N HSQC experiments of U-¹⁵N labeled α -synuclein would facilitate the detection of minor changes in the local environment of α -synuclein residues as a consequence of a weak interaction with POP. Given all the possible NMR experiments to detect protein-protein interactions which were presented in the introduction of this chapter, we started with conventional ¹H-¹⁵N HSQC followed by CSP and intensity analysis. However, before proceeding to the definitive NMR experiments in the presence of free and ZPP-bound POP, the experimental NMR conditions of α -synuclein were optimized.

This work was partially carried out in collaboration with Dr. Monica Varese.

Optimization of NMR conditions and data treatment

The structural and dynamic characteristics of IDPs affect its NMR properties. For this reason, the optimization of the NMR conditions must contemplate sensitivity, resolution and sample stability.

The high solvent exposition of amide protons of IDPs requires the use of slightly acidic pH in ¹H-¹⁵N HSQC experiments in order to avoid ¹H exchange. In this case, however, the low stability of POP at slightly acidic pH forced to raise pH to 7.4 [231, 239]. This pH was suitable for POP and only caused a slight peak broadening in the ¹H-¹⁵N HSQC spectrum of α -synuclein, without affecting the S/N ratio and resolution (Figure 79 A). Amide ¹H exchange is also minimized by decreasing temperature, and for this reason low temperatures increase the quality of ¹H-¹⁵N HSQC spectra. According to several examples in the literature, we found a good S/N ratio in the ¹H-¹⁵N HSQC spectrum of α -synuclein at 15 °C [239]; moreover, POP was also stable at this temperature.

Spectral resolution was a crucial factor for obtaining quantitative data. Especially, the low signal dispersion of IDPs required high resolution spectra in order to avoid signal overlapping. For this reason we performed the experiments acquiring 500, 2048 data points (F1, F2), in a total number of 18 scans. These values yielded a spectral resolution of 7.62, 7.83 Hz (F1, F2), and a moderate total acquisition time.

In all cases, a constant concentration of approximately 30 μ M of α -synuclein were used for the acquisition of ¹H-¹⁵N HSQC spectra. Since POP and α -synuclein are in a weak and transient interaction, it was crucial to favour the formation of the complex by using excess concentrations of the unlabeled partner (e.g. free or ZPP-bound POP). However, this was limited by the sample availability and the difficulties to reach POP concentrations higher

than 250 μM . The possible artefacts of α -synuclein derived from changes in the viscosity of the media or unspecific binding were evaluated by performing a control experiment using an excess of human serum albumin (HSA, Sigma-Aldrich, St Louis, Missouri, USA. Figure 76).

Finally, spectra were processed as described in Materials and methods section. Signals were assigned according to the reference spectra published by Rao and co-workers [239] (BioMagResBank entry number 16300). Most of the signals were unambiguously assigned manually; in the cases where assignment was not possible, signals were excluded from further analysis. Finally, chemical shifts and intensities obtained from the peak piking were used to calculate the CSP and I/I_0 ratios, respectively.

^1H - ^{15}N HSQC experiments of α -synuclein in the presence of free POP

^1H - ^{15}N HSQC spectra of α -synuclein in the presence of 6 equivalents of free POP did not show important alterations (Figure 79 B). The visual inspection showed that most of the signals were unaffected. The signal of His 50 systematically showed small changes that were attributed to small pH fluctuations; for this reason it was rejected from further analysis.

The I/I_0 plot (Figure 80 A) showed a slight decrease in the C-terminal region of α -synuclein (approximately from residues Ser 129 to Ala 140) This alteration was not present in the I/I_0 plot of the HSA control experiment (Figure 80 C). Moreover, a significant decrease in the intensity of Met 5 and Lys 21 of the N-terminal region was also observed (the decrease of intensity of Lys 10 was not considered since it also occurred in the HSA control experiment). These results were partially reproduced in the CSP plot (Figure 80 A): the chemical shift of residues Val 3, Phe 4, Val 16 and Thr 22 were affected; on the contrary, only residue Glu 126 of the C-terminal showed unusually high CSP.

Although previous studies discarded α -synuclein as a POP substrate [210], it is worth comment that sometimes we did observe the presence of degraded fragments (e.g. in Figure 79 C). The LC-MS analysis of NMR samples indicated that the fragments corresponded mainly to α -synuclein truncations in the N-terminal region (Figure 79 D). However, given that these truncations appeared after relatively long periods of incubation (around 2-3 days), we cannot affirm if they corresponded to proteolytic fragments or to degraded α -synuclein.

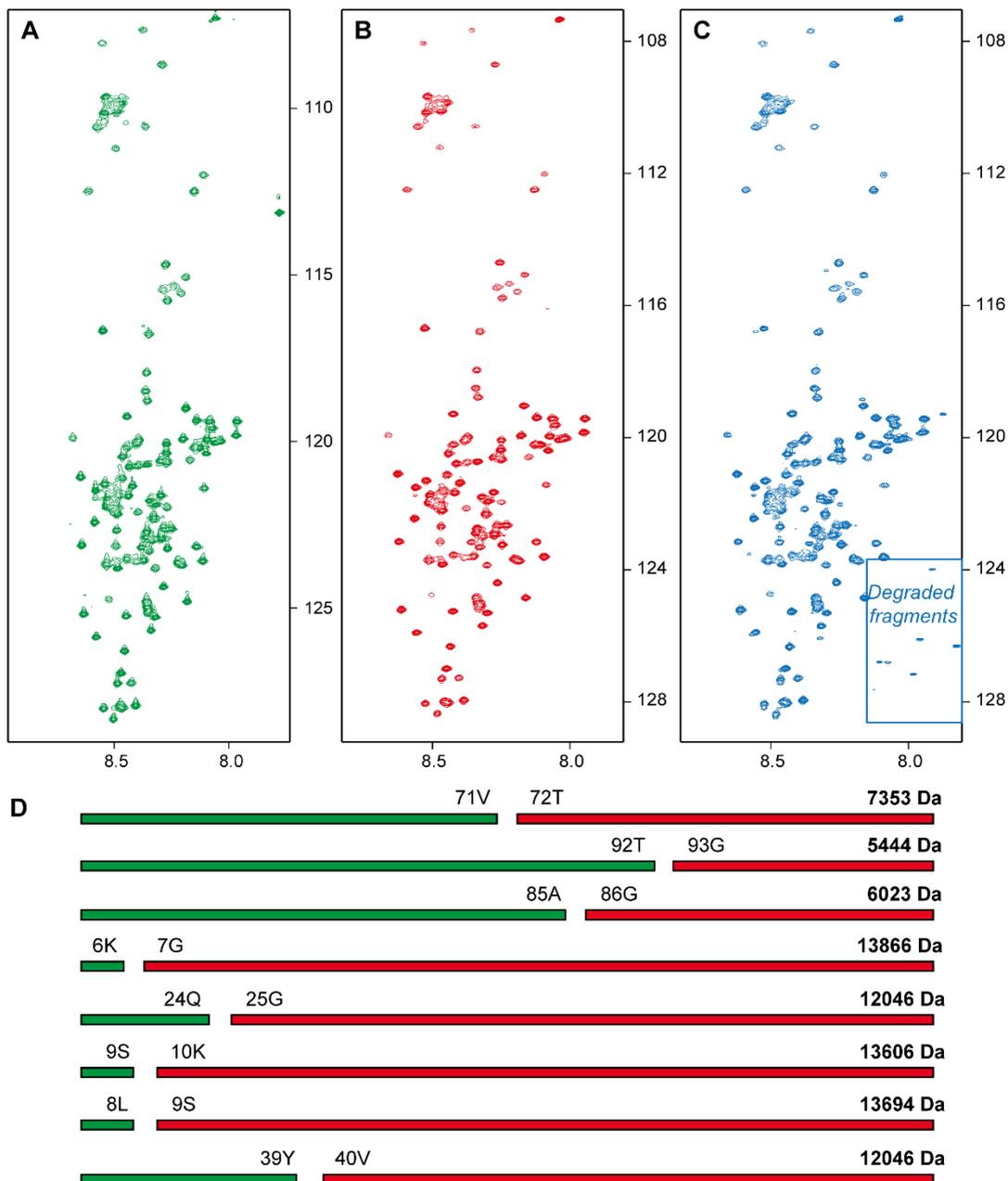


Figure 79: NMR experiments of U-¹⁵N labeled α-synuclein. A) ¹H-¹⁵N HSQC of 33 mM α-synuclein. B) ¹H-¹⁵N HSQC of 33 μM α-synuclein with 6 equivalents of free POP. C) ¹H-¹⁵N HSQC of 33 μM α-synuclein with 6 equivalents of ZPP-bound POP. D) Principal α-synuclein fragments detected by LC-MS analysis (red) after a set of NMR experiments (2-3 days). The missing fragments are shown in green.

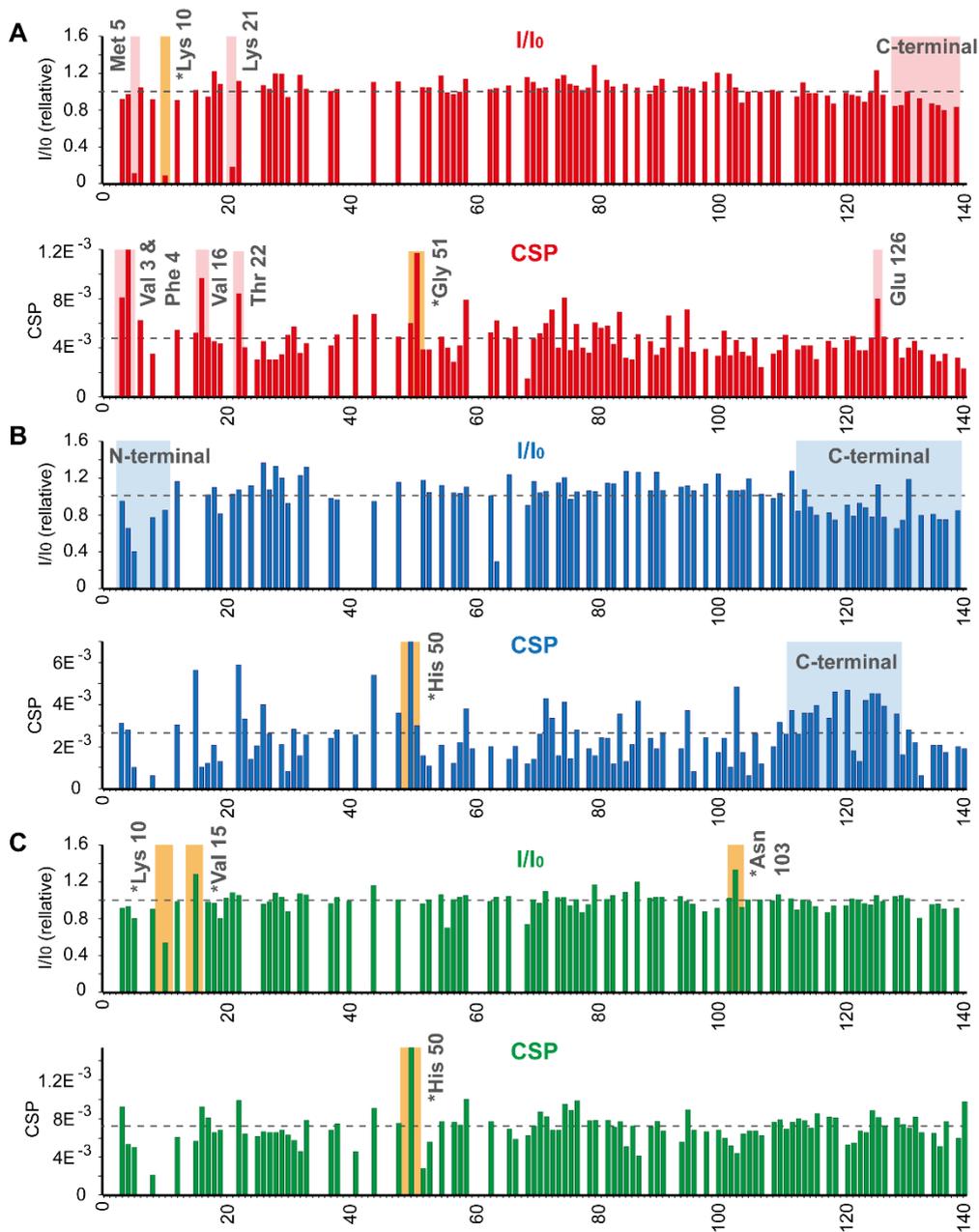


Figure 80: I/I_0 and CSP plots of ^1H - ^{15}N HSQC spectra of α -synuclein in the presence of: A) Free POP; B) ZPP-bound POP; C) Human serum albumin (HSA). In I/I_0 plots, the average value is marked with a dotted line (all I/I_0 values were normalized in order to eliminate the effect of the concentration); in CSP plots, the average value is marked with a dotted line. Residues showing more variations are highlighted; residues marked in orange are considered artefacts and have been discarded.

¹H-¹⁵N HSQC experiments of α -synuclein in the presence of ZPP-bound POP

The visual inspection of the ¹H-¹⁵N HSQC spectra of α -synuclein in the presence of POP bound to ZPP did not show generalized changes (Figure 79 C). However, the I/I_0 plot (Figure 80 B) showed two areas where the intensities decreased: the N- and C-terminal regions. These changes were more important compared than in the presence of free POP, especially in the C-terminal region (residues from Leu 113 to Ala 140). Regarding the N-terminal region, residues from Phe 4 to Lys 10 showed a decrease of the intensities between 50 and 75 %. Of interest, a wide interval of residues of the C-terminal systematically presented higher CSP values (from Ile 112 to Ser 129); in turn, the CSP plot did not show concerted perturbations in the N-terminal part. Together, these results indicated variations in chemical the shift and intensity especially in the C-terminal region.

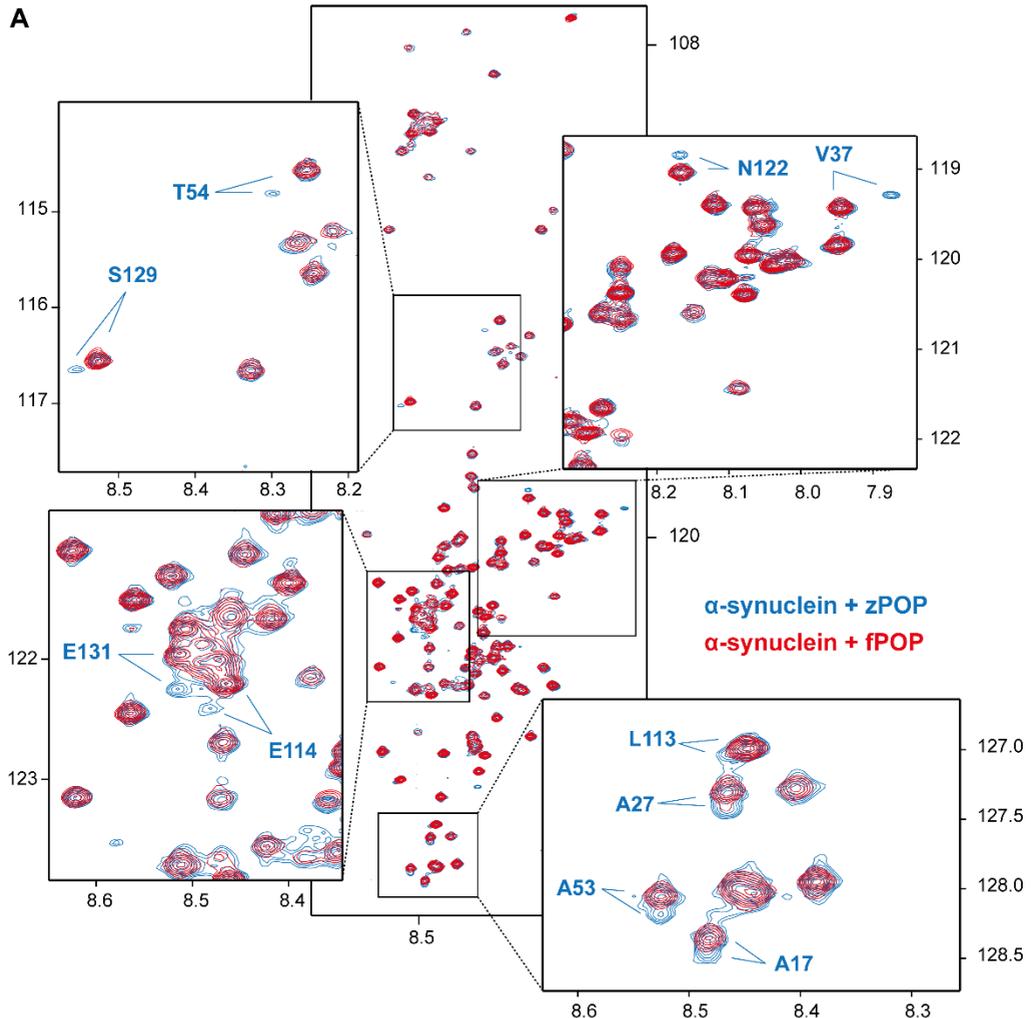
Interestingly, a careful examination of the spectra disclosed some residues displaying a secondary set of signals (Figure 81 A). These residues were present in the internal part of the N-terminal region (residues Ala 17, Ala 27, Val 37, Ala 53 and Thr 54) and in the C-terminal region (Leu 113, Glu 114, Asn 122, Ser 129, Glu 131).

For a better comprehension, Figure 81 B and C summarizes all the effects of free and ZPP-bound POP on the spectra of α -synuclein, respectively.

DISCUSSION

Taken together, the results presented above suggest a possible weak and transient interaction of POP with α -synuclein. The presence of free and ZPP-bound POP caused minor changes in the ¹H-¹⁵N HSQC spectra of α -synuclein, mainly affecting the intensities and the chemical shifts of the signals. In contrast, no changes were seen in the corresponding ¹H-¹³C methyl-TROSY HMQC spectra.

Our data is more coherent with an affinity much weaker compared to the low μ M value published by Savolainen *et al.* [215]. In a 1:6 ratio of α -synuclein:POP, the interaction of free POP with α -synuclein mainly affected few residues of the N- and C-terminal regions in the ¹H-¹⁵N HSQC spectra of α -synuclein. However, in the case of the interaction with ZPP-bound POP, the effects on I/I_0 and CSP were more intense and affected a broad section of the C-terminal region of α -synuclein.



B ¹MD³V⁴F⁵MKGLSKAKEGV¹⁶VAAAE²¹K²²TKQGVAAEAAAGKTKEGVLYVGSKTKEGVVHGVATV
AEKTKEQVTNVGGAVVTGVTAVAQKTVEGAGSIAAATGFVKKDKLQKNEEGAPQEGILEDM
PVDPDNEAY¹²⁶EMP¹²⁹SEEGYQDYEP¹⁴⁰A

C ¹MDV⁷FMKGLS¹⁰KAKEGVV¹⁷AAAEKTKQGV²⁷AEAAGKTKEG³⁷VLYVGSKTKEGVVHGV⁵³A⁵⁴
TVAEKTKEQVTNVGGAVVTGVTAVAQKTVEGAGSIAAATGFVKKDKLQKNEEGAPQEG¹¹²I¹¹³
L¹¹⁴EDMPVDPD¹²²NEAYEMP¹²⁹SE¹³¹EGYQDYEP¹⁴⁰A

Figure 81: Effects of ZPP-bound POP on the ¹H-¹⁵N HSQC spectra of α -synuclein. A) Overlaid ¹H-¹⁵N HSQC spectra of α -synuclein in the presence of free POP (red) and POP bound to ZPP (blue). Zoomed areas show the presence of a secondary set of signals of certain residues, which appear only in the case of α -synuclein in the presence of POP bound to ZPP. B) and C) Summary of the effects of the presence of free POP and ZPP-bound POP in the ¹H-¹⁵N HSQC spectra of α -synuclein, respectively. Residues highlighted in green display changes in the intensity, while those highlighted

in yellow show the appearance of a secondary set of signals. Residues displayed in red show alterations in the chemical shift.

Of interest, the effects of the presence of free POP or POP bound to ZPP in the ^1H - ^{15}N HSQC spectra were not completely correlated. Hence, it can be hypothesized that the binding motif of α -synuclein depends on the conformation of POP. The population shifts induced by inhibitors would alter the populations of conformers showing the required binding motif. This hypothesis might be to some extent related to the dual mechanism of POP suggested for the *in vivo* aggregation and clearance of α -synuclein.

Moreover, the secondary set of signals observed in the case of α -synuclein interacting with ZPP-bound POP deserves special attention. These satellites might correspond to partially resolved signals of the complex, or to a different conformation of α -synuclein arising from the interaction with inhibited POP. Ala 53 and Thr 54 are among the residues showing satellites; of interest, a familial mutation in this position (A53T) leads to an increased aggregation of α -synuclein, indicating that alterations in this area can promote or prevent α -synuclein aggregation. Another residue showing a satellite is Ser 129, which is found after a Pro residue. The second signal can be attributed to the *cis* conformation of the Pro 128 residue, which might be stabilized as a consequence of the interaction with ZPP-bound POP. Moreover, Ser 129 has been widely found phosphorylated *in vivo* [242]. It has been described that this post-translational modification prevents fibrillization [243] and modifies the metal-binding properties [244], indicating the relevance of this residue in the aggregation propensity of α -synuclein.

In summary, our findings suggest a possible weak and transient interaction of POP with α -synuclein which affects the N- and C-terminal regions, and is more intense in the case of inhibitor-bound POP. Nevertheless, the data presented here is only preliminary and precludes the extraction of definitive conclusions. One of the main limitations for the study of the interaction of POP with α -synuclein has been the lack of sensitivity to detect subtle changes derived from the weak binding. Moreover, experimental limitations prevented to reach high α -synuclein:POP ratios. These drawbacks avoided the detailed characterization of the interaction by straightforward CSP and intensity analyses. For this reason, further research will be carried out in our laboratory in order to describe this interaction in more detail. More robust NMR experiments will be performed, especially R_1 , R_2 and steady-state NOE measurements, as well as cross-saturation experiments. These are good candidate experiments to shed light to the transient interaction between POP and α -synuclein

General discussion

Proteins participate in an intricate network of cellular processes in living organisms through enzymatic reactions, structural arrangements and interactions with other partners or ligands. For this reason, proteins are inherent dynamic entities. According to the energy levels of the conformational landscape, dynamic events span a wide range of time scales [91]. Hence, collective motions result in slow events in the time scale of μ s-ms; in turn, local flexibility and side chain dynamics originate fast dynamics in the ps-ns range, which is associated with the conformational entropy of the system.

Slow dynamic events are intimately connected with protein function. In particular, enzymes benefit from μ s-ms dynamics in order to provide all of the substates necessary for catalysis. The work of Eisenmesser and co-workers strongly contributed to this finding; with the use of NMR relaxation methods, they nicely demonstrated that the turnover number of Cyclophilin A was correlated to the frequencies of dynamic motions in the active site [82]. For this reason, the study of protein dynamics has become a field of intense research for the elucidation of enzymatic mechanisms and for the design of inhibitors (in this regard, a JACS select issue has been recently dedicated to protein dynamics [245]).

Protein dynamics is also involved in protein recognition events, facilitating the fine-tuning of protein-protein and protein-ligand interactions [86]. This phenomena is potentially relevant for regulatory processes, but the exact role of dynamics in molecular recognition is still not clear to date. Initially, it was postulated that recognition involving dynamic systems takes place in a conformational selection of those conformations presenting the binding motifs required for the interaction [246]. For instance, in the case of the C-terminal SH2 domain of the regulatory subunit p85 α of the phosphotyrosine 3-kinase, the μ s-ms dynamics was abolished upon interaction with a phosphotyrosine-containing peptide [88]. However, other recent studies points that protein dynamics dictates the dissociation rates, which ultimately determines the affinity of the interaction [247].

POP is a therapeutic target for the treatment of cognitive deficits and for the preventive therapy of PD. However, the *in vivo* role of POP and the mechanism of action of inhibitors are not fully understood. Recent hypothesis points out that POP acts *in vivo* through PPIs rather than through its peptidase activity [36]. Little is known about the interactome of POP and the regions of the protein involved in recognition events. Gaining knowledge about the structure of POP in solution will contribute to unravel the biological function of POP. In turn, understanding the effects of inhibitors in the conformational landscape of POP will be of

high interest for the development of new therapeutic molecules targeting unexplored binding surfaces of the enzyme.

Several studies indicated that POP exists only in a closed active conformation with several flexible loops involved in substrate recruitment. Other research lines propose that POP presents an equilibrium between closed and open conformations. To date, this question still remains unclear. Probably, these hypothesis are not completely excluding; rather, they correspond to partial observations of a more complex scenario involving large structural rearrangements and local motions at different time scales. The elucidation of POP structure and dynamics have remained elusive as a consequence of the large size and structural complexity of POP, which makes difficult the use of conventional biophysical techniques. Most of the studies of POP structure and dynamics have been based on X-ray crystallography, MD simulations starting from X-ray structures, or activity assays of engineered mutants, among others. Hence, extensive studies using robust biophysical methods are lacking.

Experimental methods usually provide ensemble averaged data or static snapshots, which do not represent the whole conformational landscape of dynamic proteins. For this reason, the study of dynamic proteins must rely on biophysical methods that contemplate time dimension and represent the structures as ensembles of conformations [91, 248]. Here, we circumvent the methodological limitations associated with the molecular size and complexity of POP with a combination of biophysical techniques, complemented with accurate data analysis and theoretical simulations. In particular, we studied the conformational dynamics of POP in solution using ^1H - ^{13}C methyl-TROSY HMQC and SAXS experiments combined with MD simulations in order to obtain *bona fide* high resolution structures.

The results obtained by ^1H - ^{13}C methyl-TROSY HMQC RD experiments disclosed high dynamics of free POP arising from interdomain flexibility in the μs -ms time scale. In particular, the most dynamic residues in this time scale were found on the α/β -hydrolase domain and in the interdomain region. The fitting of RD data in a global two-site exchange indicates that μs -ms interdomain separation is better described by independent motions with distinct amplitudes and frequencies rather than by a single concerted movement. In this regard, $R_{1\rho}$ RD experiments reveal that POP dynamics is not exclusive of μs -ms, but also participates in fast μs local motions. Analyses of RMSF and SASA values derived from MD simulations disclose high ns flexibility of loop A, B, C and His 680 loop in the open conformation of free POP. This high loop flexibility might be a consequence of the increased

degree of freedom of the open conformation, which presents fewer stabilizing interactions compared to the closed conformation. The increased conformational entropy of the open form partially compensates the enthalpic penalty caused by the loss of stabilizing interactions. From a functional point of view, the conformational heterogeneity of free POP probably gives rise to all the substates required for the circulation of substrate and products and for the execution of the chemical step.

R_g value obtained from online SEC-SAXS experiments was higher than the calculated for the X-ray structure of POP in the closed conformation, and the $P(r)$ distribution indicated a structure probably originated from interdomain separations. Moreover, an additional highly compact form of free POP was detected in the SEC-SAXS tandem experiments. Probably, this secondary POP conformer might be originated from a tight arrangement of β -propeller blades [2]. Further work is required to elucidate the structure of this minor form.

The EOM performed over a pool of open and closed simulated conformers using the SAXS scattering profile of free POP as a driving force selected approximately 55 % of open and 45 % of closed structures. Together with the NMR data, these results demonstrate that free POP exists in a μ s-ms dynamic equilibrium in solution between open and closed conformations.

However, this scenario is radically altered in the presence of active site-directed covalent inhibitors. RD experiments performed in the presence of inhibitors showed a total quenching of POP dynamics in the μ s-ms time scale. Furthermore, ns flexibility derived from MD simulations also decreased in amplitude if ZPP was covalently bound to POP. SEC-SAXS experiments indicated that active site-directed covalent inhibitors stabilized the closed conformation of POP with the active site configuration similar to the transition state [2, 11]. In contrast, similar to what has been described for *Trypanosoma brucei* oligopeptidase B [53], the active arrangement of the active site does not occur in the open form of free POP. Hence, we propose that the interdomain separation of POP switches the catalytic activity between active and inactive forms.

Abolishing the μ s-ms dynamics of POP by inhibitors prevents the completion of the enzymatic cycle. But more significant for the *in vivo* role of POP, modifying the conformational landscape might be an essential element for the modulation of the PPIs of the enzyme. Our preliminary results of the study of the interaction between POP and α -synuclein are coherent with this hypothesis. In particular, ^1H - ^{15}N HSQC experiments of α -synuclein show that POP and α -synuclein interact with a weak affinity, affecting few

residues of the N- and C-terminal regions of α -synuclein. But in the presence of POP inhibitors, the affinity of the interaction is slightly increased, and mainly affects a broad section of the C-terminal region of α -synuclein. Of interest, these differences in the interaction pattern observed in the presence of inhibitors suggest that the binding motif of α -synuclein depends on the conformation of POP.

The high conformational diversity of POP remarks the potential regulatory role of this enzyme in molecular recognition events. Very recently, it has been described that a POP inhibitor, KYP-2047, increase autophagy activation markers [212]. According to this, POP ligands could operate as conformational switches that enhance autophagy-lysosomal protective pathways in cells [249]. It can be hypothesized that the modification of the conformational landscape of POP could alter the recognition events with putative interacting partners; only certain conformations of POP would provide the structural binding motifs required for binding to POP-interacting proteins. Hence, the drastic energetic and dynamic rearrangements caused by active site-directed POP inhibitors would influence the affinity of the interactions, promoting the *in vivo* effects [215]. Moreover, the capacity to alter the native conformational landscape of POP might not be exclusive of inhibitors. Endogenous POP ligands or substrates might also induce significant conformational changes. In turn, our results showed that pH also influences the conformational equilibrium of POP, suggesting that small variations of cytosolic pH could alter the recognition events.

In summary, our findings supports that the conformational dynamics of POP is inherent for the recognition events responsible for the *in vivo* function. In turn, active site-directed inhibitors severely modifies the conformational landscape of POP, altering the recognition pattern. Interestingly, this opens up new possibilities for the design of novel POP inhibitors conceived as conformational switches with capacity to modulate the PPIs of the enzyme.

Materials and methods

PROTEIN EXPRESSION

POP expression protocols

In this section, four protocols are described for different labeling strategies:

- No labeling (native POP)
- *L*-TFM labeled POP
- [methyl-¹³C]-Met labeled POP
- Perdeuterated and U-[²H], [methyl-¹³C, ¹H]-Met labeled POP

In all cases, expression batches of 3-4 litres were typically used, except otherwise indicated. Fresh POP was used in all experiments, since freeze-thawing the sample resulted in an irreversible aggregation and decrease of the enzymatic activity. Fresh POP was always stored at 4 °C and used in an interval of no more than 5 days.

Expression of native POP

Note: the composition of media, stock solutions and buffers is summarized at the end of the protocol.

1. Transformation of a 50 µl aliquot of *E.coli* BL21(DE3) competent cells with pETM11-POP plasmid by heat shock:
 - Keeping the cells in ice, add a volume of plasmid (usually between 1 and 2 µl). Keep the cells in ice during 15 min.
 - Heat shock the cells during 45 seconds at 42 °C.
 - Return the cells to the ice bath, and after 5 min, add 500 µl of SOC media.
 - After 30 min at 37 °C, transfer the cells to an LB-agar Petri capsule with 50 µg/ml kanamycin, under sterile conditions.
 - Incubate the Petri capsule overnight at 37 °C in an inverted position. Abundant spread colonies should be observed after incubation. Transformed colonies must be used in no longer 3-4 days.
2. Inoculation of 25 ml of pre-culture:

- Inoculate 25 ml of normal minimal media + 50 mg/l of kanamycin (NMM) with one single colony of the Petri capsule, under sterile conditions. Typically, one pre-culture is used for each litre of final culture.
 - Incubate the pre-culture O/N at 37 °C under vigorous agitation.
3. Large-scale culture and expression:
- Add a volume of pre-culture to 1 l NMM batches up to a final OD₆₀₀ of approximately 0.1.
 - Grow 1 l batches at 37 °C with vigorous agitation up to OD₆₀₀ between 0.6 - 0.8 (3-5 h). Afterwards, decrease the temperature to 18 °C.
 - Induce protein expression by adding 1 mM IPTG (from a 1000X stock solution). Expression takes place O/N at 18 °C under agitation.
4. Extraction of the soluble protein extract and first step of purification:
- Harvest the cells from the expression media by centrifugation at 4500 g during 15 min at 4 °C.
 - Resuspend the bacterial pellet with 25 ml of binding buffer, and add a little amount of DNase I.
 - Lysate the resuspended cells with cell disruptor (Constant Systems Ltd, Northants, UK) at 20 Kpsi and at 4 °C.
 - Remove bacterial debris by centrifugation at 40000 g during 30 min at 4 °C. Filter the supernatant through 0.45 µm filter.
5. Ni²⁺ affinity chromatography:
- Load the solution into a superloop mounted in FPLC instrument and inject it to a HisTrap HP 5 ml column (GE Healthcare Life Sciences, Uppsala, Sweden) equilibrated with binding buffer.
 - After injection, wash the column with binding buffer and subsequently with washing buffer.
 - Elute POP by circulating elution buffer.
6. HisTag cleavage by tobacco etch virus (TEV) protease (*TEV is a serine protease which recognizes the ENLYFQG motif included in the pETM11 plasmid*):
- Exchange elution buffer by SEC buffer using PD-10 columns (GE Healthcare Life Sciences, Uppsala, Sweden). Add DTT and EDTA from stock solutions to a final concentration of 1 and 0.5 mM, respectively.
 - Add TEV to POP sample (typically 1 A₂₈₀ of TEV protease per 100 A₂₈₀ of substrate). Digestion takes place during several hours at room temperature (typically 4-5 h) under soft agitation.

7. TEV and HisTag removal and size exclusion chromatography:

- Centrifuge the digestion mixture at 4500 g during 15 min in order to precipitate proteins in suspension, and filter through 0.45 µm filter.
- Load the mixture into a HiPrep 26/10 desalting column (GE Healthcare Life Sciences, Uppsala, Sweden) equilibrated with SEC buffer in a FPLC instrument.
- Collect the fractions eluting during the void volume of the HiPrep column, and inject them into a HisTrap HP 5 ml column, previously equilibrated with SEC buffer. Collect the flow through and concentrate it to 5-7 ml with 30 KDa cut-off membrane concentrators.
- Inject the fraction into a Superdex 200 16/600 HiLoad SEC column (GE Healthcare Life Sciences, Uppsala, Sweden), equilibrated with SEC buffer. Perform the SEC with the same buffer at 4 °C or at room temperature (temperature affects the proportion of oligomers). Take the fractions of the monomer (eluted around 70-80 ml) and concentrate them with 30 KDa cut-off membrane concentrators.
- Quantify concentrated POP stock solutions by measuring the A_{280} using Nanodrop spectrophotometer (Thermo Fisher Scientific Inc., Waltham, USA) and an Abs 0.1 % value of 1.587. Assess the purity by SDS-PAGE.

Media and buffers:

- Normal minimal media (NMM): dissolve 6 g of Na_2HPO_4 , 3 g of KH_2PO_4 and 0.5 g of NaCl in 1 l of distilled water in a 3 l baffled Erlenmeyer flask, and autoclave it. Afterwards, add 1 g of solid NH_4Cl and 4 g of solid *D*-glucose. Add 1 ml aliquots of sterile stock solutions of MgSO_4^* (1M), kanamycin sulfate* (50 mg/ml), Q solution*, biotin* (50 mg/ml) and thiamine* (100 mg/ml).
- Binding buffer: 50 mM Tris-HCl pH=8, 500 mM NaCl, 5 mM imidazole.
- Washing buffer: 50 mM Tris-HCl pH=8, 500 mM NaCl, 50 mM imidazole.
- Elution buffer** : 50 mM Tris-HCl pH=8, 500 mM NaCl, 500 mM imidazole.
- SEC buffer** : 50 mM Tris-HCl pH=8, 20 mM NaCl.

Filter all buffers except NMM through 0.22 µm filter.

* MgSO_4 , kanamycin sulfate, biotin and thiamine sterile stock solutions were prepared by dissolving the corresponding amounts in NMM and filtering through 0.22 µm filter under sterile conditions. Q solution was prepared by dissolving 5 g of $\text{FeCl}_2 \cdot 4\text{H}_2\text{O}$, 184 mg of

CaCl₂·2H₂O, 64 mg of H₃BO₃, 18 mg of CoCl₂·6H₂O, 4 mg of CuCl₂·2H₂O, 380 mg of ZnCl₂, 605 mg of Na₂MoO₄·2H₂O, 40 mg of MnCl₂·4H₂O and acidifying with 4 ml of HCl 37 %. The suspension was diluted to 1 l and autoclaved.

** Special care must be taken in the pH adjustment of these solutions. The pH must be as close as possible to 8.

Expression of L-TFM labeled POP

The expression protocol is essentially the same as described in the case of native POP, but including these modifications in the following steps:

1. Use *E.coli* B834(DE3) auxotrophic cells, and transform them with pETM10-POP plasmid.
2. 0.1 mM of Met was added to the pre-culture.
3. Grow the large-scale bacterial culture as follows:
 - Add a volume of pre-inoculum to 500 ml of NMM including 0.055 mM Met and 50 mg/L of the 19 amino acid except Met.
 - Grow the culture overnight at 25 °C, then the next day at 37 °C until A₆₀₀ = 0.8 – 1.
 - Transfer the culture to a centrifuge tube (sterilized under UV), and harvest cells by centrifugation at 2500 g during 15 min at 4 °C.
 - Carefully wash the cell pellet and the tube at 4 °C with 250 ml of NMM without Met.
 - Carefully resuspend cells were in 500 ml of expression media, and transfer it to a sterile Erlenmeyer flask. Expression media consisted of NMM supplemented with 0.5 mM L-TFM, 0.025 mM Met and 50 mg/L of the 19 amino acid except Met. Temperate cells 15 min at 37 °C, then 15 min at 25 °C.
 - Induce protein expression by adding 1 mM IPTG (from a 1000X stock solution). Expression takes place during 24 h at 25 °C under agitation.
4. Steps 4 and 5 proceed essentially as described in the case of native POP. After eluting the protein from the HisTrap HP column, exchange the buffer to Tris-HCl 50 mM pH = 8, 10 % D₂O buffer with PD-10 columns. Concentrate protein with 30 KDa cut-off membrane concentrators.

Expression of [methyl-¹³C]-methionine labeled POP

The expression protocol is essentially the same described in the case of native POP, but including modifications in the following steps:

1. Use *E.coli* B834(DE3) auxotrophic cells instead of *E.coli* BL21(DE3) cells.
2. NMM media must contain 80 mg/l of [methyl-¹³C]-Met.
3. NMM media must contain 80 mg/l of [methyl-¹³C]-Met.
7. After SEC, exchange against NMR buffer using PD-10 columns and concentrate with 30 KDa cut-off membrane concentrators. Measure final concentration after concentration.

Expression of perdeuterated and U-[²H], [methyl-¹³C, ¹H]-methionine labeled POP

The expression protocol is adapted from the protocol of [methyl-¹³C]-Met labeled POP.

Note: the preparation of deuterated medium is summarized at the end.

1. Transform *E.coli* B834(DE3) auxotrophic cells with pETM10-POP plasmid.
2. Adapt cells to D₂O:
 - Prepare 50 ml of deuterated minimal medium with 50 % D₂O and inoculate with some colonies. Grow the pre-culture at 37 °C under vigorous agitation during 6 hours.
 - Harvest cells by centrifugation at 4500 g during 15 min. After that, resuspend cells in 100 ml of deuterated minimal medium pre-culture prepared with 100 % D₂O, and grow O/n at 37 °C under vigorous agitation.
3. Large-scale culture and expression:
 - Harvest cells by centrifugation at 4500 g during 15 min, and resuspend them to 1 l of deuterated minimal media prepared with 100 % D₂O to an initial OD₆₀₀ slightly higher than 0.1.
 - Grow 1 l batches at 37 °C with vigorous agitation up to an OD₆₀₀ of approximately 0.4. More density is detrimental to protein expression. Afterwards, decrease temperature to 18 °C, and temper batches during several minutes.
 - Induce protein expression by adding 1 mM IPTG (from a 1000X stock solution). Expression takes place O/N at 18 °C under agitation.

6. Steps 4 and 5 proceed as described by the native POP. Size exclusion chromatography is performed at this point:

- Inject POP into a Superdex 200 16/600 HiLoad SEC column (GE Healthcare Life Sciences, Uppsala, Sweden), equilibrated with SEC buffer. Perform SEC at 4 °C. Join monomer fractions and concentrate with 30 KDa cut-off membrane concentrators.
- Quantify POP stock solutions by measuring the A_{280} using Nanodrop spectrophotometer (Thermo Fisher Scientific Inc., Waltham, USA) and an Abs 0.1 % value of 1.587. Exchange POP buffer against NMR buffer using PD-10 columns and concentrate with 30 KDa cut-off membrane concentrators. Measure final concentration after concentration.

Deuterated minimal media (for 1 l): dissolve 6 g of Na_2HPO_4 , 3 g of KH_2PO_4 0.5 g of NaCl and 0.25 g of MgSO_4 in a small amount of D_2O . Lyophilize the solution and dissolve again into a final volume of 1 l. Filter this solution through 0.22 μm filter under sterile conditions, and transfer it to a dry Fernbach flask previously sterilized by UV radiation. After that, add 1 g of NH_4Cl and 2 g of D-glucose (1, 2, 3, 4, 5, 6, 6 - d_7), together with 1 ml aliquots of sterile stock solutions of kanamycin sulfate, Q solution, biotin and thiamine (prepared as described above but dissolving it in deuterated minimal medium). Finally, add 80 mg/l of [methyl- ^{13}C]-Met for the expression of perdeuterated [methyl- ^{13}C , ^1H]-Met labeled POP; in the case of the expression of U- ^2H], [methyl- ^{13}C , ^1H]-Met labeled POP, use 50 mg/l of [methyl- ^{13}C]-Met (2, 3, 3, 4, 4, - d_5).

Expression and purification of WT α -synuclein

E. coli BL21 cells were transformed with pT7-7 plasmid containing α -synuclein gene as described in the case of POP. After transformation, a pre-culture of 10 ml of LB media with 100 $\mu\text{g/ml}$ ampicillin was grown at 37 °C for each litre of final culture.

Protein expression was performed in normal minimal medium supplemented with $^{15}\text{NH}_4\text{Cl}$ in the absence of Q solution (this is strictly necessary to prevent Met oxidations promoted by metal cations [236]), using 100 $\mu\text{g/ml}$ ampicillin as antibiotic. Typical, expression took place in batches of 3-4 litres. The following substances were necessary to prepare 1 litre of medium:

- 6 g of Na_2HPO_4
- 3 g KH_2PO_4

- 500 mg NaCl
- 2 g glucose
- 1 g $^{15}\text{NH}_4\text{Cl}$
- 1 ml MgSO_4 1000X (1 M)
- 1 ml CaCl_2 1000X (0.1 M)
- 1 ml biotin 1000X (50 mg/ml)
- 1 ml thiamine 1000X (100 mg/ml)

Large scale culture was inoculated with an initial A_{600} near 0.1, and allowed to grow until to an A_{600} of 0.6-0.8. Induction of protein expression took place with 1 mM IPTG, during 3-4 h at 37 °C. After this time, cells were harvested and the pellet was frozen in liquid N_2 , if necessary.

The bacterial pellet was resuspended in 15 ml of 10 mM Tris-HCl, pH 8, 1 mM EDTA and 1 mM PMSF (per litre of culture), and lysed by multiple freeze-thaw cycles and sonication (30 second sonication + 30 second wait, 80 % intensity, at 0 °C). The cell lysate was boiled during 20 minutes, and after that, it was centrifuged at 20000 g during 30 minutes at 4 °C in order to precipitate the heat-sensitive proteins (α -synuclein remain soluble). Streptomycin sulphate was added to the supernatant from the previous step (final concentration 10 mg/ml) to precipitate DNA. The mixture was stirred at 4 °C for 15 minutes, followed by centrifugation at 20000 g during 30 minutes at 4 °C. α -synuclein was then precipitated by addition of ammonium sulphate (361 mg/ml) to the supernatant from previous centrifugation. The solution was stirred at 4°C during 15 minutes and centrifuged at 20000 g during 30 minutes at 4 °C.

α -synuclein pellet was resuspended in 20 ml of 25 mM Tris-HCl, pH 7.7. The mixture was filtered through a 0.45 μm filter and desalted with a HiPrep 26/10 desalting column equilibrated with Tris buffer. After that, protein was loaded onto a HiTrap Q HP anion exchange column (GE Healthcare Life Sciences, Uppsala, Sweden) previously equilibrated with Tris buffer. Full length α -synuclein was eluted with a salt gradient at 300 mM NaCl (performing a gradient with 25 mM Tris-HCl, 300 mM NaCl, pH 7.7).

The fractions containing protein (determined by SDS-PAGE) were pooled together and dialyzed overnight against 2 litres of 25 mM Tris-HCl, pH 7.4, 100 mM NaCl, at room temperature using a 6000-8000 MWCO membrane. After that, the protein was filtered through a carefully cleaned concentrator (100.000 MWCO, Vivaspin 20, GE Healthcare Life Sciences, Uppsala, Sweden), in order to eliminate oligomers and higher molecular weight

species. Filtrated solution was concentrated with a 10.000 MWCO concentrator (Amicon Ultra-15 Ultracel cellulose membrane, MerckMillipore, Darmstadt, Germany) to a final protein concentration of 0.7 mM. Protein was aliquoted, flash-frozen in liquid N₂ and stored at -80 °C. Protein expression yields reached up to 30 mg of protein per litre of culture; all of the buffers used during the purification were degassed.

SYNTHETIC PROCEDURES

Synthesis of *L*-trifluoromethionine

L-TFM was synthesized in a total of 4 steps, in an overall yield of 30 %. All reagents were purchased from Sigma-Aldrich (St Louis, Missouri, USA), except Di-Boc-*L*-homocystine, which was purchased from Santa Cruz Biotechnology (Dallas, Texas, USA). ¹H and ¹³C NMR experiments were performed in a Varian Mercury 400 MHz spectrometer. NMR data is reported as follows: chemical shift (ppm) (multiplicity, integration, coupling constant, assignation). All signals were referenced to the residual solvent signal. SiO₂ thin layer chromatographies were monitored by UV light and by the ninhydrin test. HPLC were performed in a Waters Alliance 2695 PDA using an Xbridge C₁₈ 2.5 μM, 4.6 x 75 mm column. Exact masses were obtained in an LTQ-FT Ultra (Thermo Scientific) mass spectrometer.

ACN: acetonitrile; *DCM*: dichloromethane; *EtOAc*: ethyl acetate; *THF* tetrahydrofurane; *TLC*: tin layer chromatgraphy.

Note: the reaction steps are numbered according to the scheme of Figure 11.

1-Di-Boc-*L*-homocystine methyl ester (1): 1.01 g (2.15 mmol) of commercial Di-Boc-*L*-homocystine were dissolved in 25 ml of DCM/MeOH 3/2. The mixture was cooled to 0 °C and 2.4 ml (2.2 equivalents) of (trimethylsilyl)diazomethane (TMS-CHN₂) were added dropwise under stirring, until a persistent pale yellow colour was observed. The reaction was allowed to complete during 30 min. After this time, the excess of TMS-CHN₂ was quenched by the careful addition of acetic acid until the pale yellow colour disappeared. The solvent was evaporated under vacuum, and the crude product was purified by a short column chromatography of basic alumina (Hexane/AcOEt 1/1). 959 mg of pure Di-Boc-*L*-homocystine methyl ester was obtained (1.93 mmol, 90 % yield). White solid. HPLC: *rt* = 7.101 min (30-100 gradient). *M*(calc) = 496.20, *M*(exp) = 497.2001 [M+H]⁺. ¹H NMR (CDCl₃): δ = 1.43 (s, 18 H, Boc); δ = 1.979 (m, 2 H, CH β position); δ = 2.23 (m, 2 H, CH

β position); $\delta = 2.70$ (m, 4 H, CH₂ γ position); $\delta = 3.74$ (s, 6 H, methyl ester); $\delta = 4.39$ (m, 2 H, CH α position); $\delta = 5.14$ (d, $^3J_{\text{H-H}} = 7$ Hz, 2 H, NH). ^{13}C NMR (CDCl₃): $\delta = 28.25$ (CH₃ Boc); $\delta = 32.51$ (CH₂ β position); $\delta = 34.46$ (CH₂ γ position); $\delta = 52.45$ (CH₃ methyl ester + CH α position); $\delta = 80.07$ (Boc quaternary carbon); $\delta = 155.30$ (CO Boc); $\delta = 172.59$ (CO ester).

2-Boc-L-homocysteine methyl ester (2): 959 mg of Di-Boc-L-homocystine methyl ester were weighed and dissolved in a flask purged with N₂ in 60 ml of milliQ water and 120 ml of degassed MeOH (prepared by bubbling a soft stream of N₂ to commercial MeOH, and subsequent sonication under vacuum). 1.11 g of TCEP-HCl (3.87 mmol, 2 equivalents) dissolved in 15 ml of degassed MeOH were added to the previous solution under inert atmosphere. The mixture was stirred during 2 h at room temperature. Afterwards, the solvent was evaporated under vacuum, and 100 ml of citrate buffer pH = 4 were added. The suspension was extracted three times with degassed DCM (prepared analogously to MeOH). Organic layers were dried with MgSO₄ and the solvent was evaporated under vacuum. 725 mg of the pure Boc-L-homocysteine was obtained, without the need of any further purification (2.91 mmol, 75 % yield). Yellow oil. HPLC: rt = 5.079 min (30-100 gradient). M(calc) = 249.11, M(exp) = 250.1110 [M+H]⁺. ^1H NMR (CDCl₃): $\delta = 1.44$ (s, 9 H, Boc); $\delta = 1.59$ (s, 1 H, SH); $\delta = 1.94$ (m, 1 H, CH β position); $\delta = 2.10$ (m, 1 H, CH β position); $\delta = 2.58$ (m, 2 H, CH₂ γ position); $\delta = 3.75$ (s, 3 H, methyl ester); $\delta = 4.46$ (m, 1 H, CH α position); $\delta = 5.07$ (d, $^3J_{\text{H-H}} = 5.4$ Hz, 1 H, NH). ^{13}C NMR (CDCl₃): $\delta = 20.70$ (CH₂ γ position); $\delta = 28.28$ (CH₃ Boc); $\delta = 37.23$ (CH₂ β position); $\delta = 52.24$ (CH α position); $\delta = 52.46$ (CH₃ methyl ester); $\delta = 80.15$ (Boc quaternary carbon); $\delta = 155.39$ (CO Boc); $\delta = 172.79$ (CO ester).

3-Boc-L-trifluoromethionine methyl ester (3): 725 mg (2.91 mmol) of **2** were taken and dissolved in 20 ml of degassed DCM (prepared analogously as described for MeOH), and cooled to -78 °C under inert atmosphere. 1152 mg of Togni's reagent (3,3-Dimethyl-1-(trifluoromethyl)-1,2-benziodoxole) (4.49 mmol, 1.2 equivalents) were weighed and dissolved in 12 ml of degassed DCM. Afterwards, this solution was added dropwise to the solution of **2**. The mixture was kept under stirring at -78 °C during 2 hours. After this time, the reaction was allowed to reach room temperature and it was kept overnight at this temperature. The solvent was removed under vacuum, and 8 g of bulk sorbent (C₁₈, 40-60 μm) in DCM were added. The solvent was removed under vacuum, and the solid residue was loaded to a CombiFlash chromatographer (Teledyne Isco, Lincoln, USA). The product was purified using a RediSep RF gold reverse phase C₁₈ column (50 g, Teledyne Isco,

Lincoln, USA) and a gradient of H₂O/ACN. ACN was eliminated under vacuum, and the aqueous mixture was lyophilized to yield 492 mg of pure product (1.56 mmol, 53 % yield). Pale yellow oil. HPLC: rt = 6.722 min (30-100 gradient). M(calc) = 317.10, M(exp) = 318.0990 [M+H]⁺. ¹H NMR (CDCl₃): δ = 1.44 (s, 9 H, Boc); δ = 2.00 (m, 1 H, CH β position); δ = 2.26 (m, 1 H, CH β position); δ = 2.93 (m, 2 H, CH₂ γ position); δ = 3.77 (s, 3 H, methyl ester); δ = 4.42 (m, 1 H, CH α position); δ = 5.13 (d, ³J_{H-H} = 5.2 Hz, 1 H, NH). ¹³C NMR (CDCl₃): δ = 25.94 (CH₂ γ position); δ = 28.23 (CH₃ Boc); δ = 33.32 (CH₂ β position); δ = 52.36 (CH α position); δ = 52.62 (CH₃ methyl ester); δ = 80.36 (Boc quaternary carbon); δ = 142.69 (CF₃); δ = 155.36 (CO Boc); δ = 172.18 (CO ester). ¹⁹F NMR (CDCl₃, TFA as internal standard): δ = -41.87 (s, CF₃).

4-L-trifluoromethionine hydrochloride: 491 mg of **3** (1.55 mmol) were dissolved in 20 ml of THF. In a separate flask, 651 mg (15.5 mmol, 10 equivalents) of LiOH were dissolved in 35 ml of water. The basic solution was added over the organic phase and the mixture was kept under stirring during 1 hour at room temperature. After this time, the organic solvent was removed under vacuum, and the aqueous solution was brought to a slightly acidic pH with formic acid 1 M. This solution was extracted 4 times with EtOAc, and the organic layers were dried with anhydrous MgSO₄. The solvent was removed under vacuum, and the oily residue was dissolved in 4 ml of HCl in dioxane (4 M). The mixture was stirred during 1 hour and 30 min. After this time the solvent was removed under vacuum, and the oily residue was dissolved in water acidified with HCl. The aqueous mixture was extracted with DCM, and afterwards it was lyophilized. 309 mg of L-trifluoromethionine hydrochloride **4** were obtained (1.29 mmol, 83 % yield). In order to be used in bacterial cultures, the product was dissolved in water acidified with HCl to a final concentration of 0.1 M and filtered through 0.22 μm sterile filters under sterile conditions. The stock solution was kept frozen until needed. White powder. HPLC: rt = 5.908 min (0-40 gradient). M(calc) = 203.03, M(exp) = 204.0301 [M+H]⁺. ¹H NMR (D₂O, TSP as internal standard): δ = 2.34 (m, 2 H, CH β position); δ = 3.14 (m, 2 H, CH₂ γ position); δ = 4.10 (m, 1 H, CH α position). ¹³C NMR (D₂O, TSP as internal standard): δ = 28.26 (CH₂ γ position); δ = 33.59 (CH₂ β position); δ = 55.01 (CH α position); δ = 134.06 (CF₃); δ = 174.43 (COOH). ¹⁹F NMR (D₂O, TFA as internal standard): δ = -42.10 (s, CF₃).

Synthesis of [methyl-¹³C]-methionine (2, 3, 3, 4, 4-d₅)

[Methyl-¹³C]-Met (2, 3, 3, 4, 4, - d₅) was synthesized in 11 steps, in an overall yield of 10 %. The synthesis provided an excellent stereoselectivity and deuterium incorporation (>

97 % deuteration). All conventional reagents were purchased from Sigma-Aldrich (St Louis, Missouri, USA); deuterated reagents were purchased from Cambridge Isotope Laboratories (Tewksbury, Massachusetts, USA). ^1H and ^{13}C NMR experiments were performed in a Varian Mercury 400 MHz spectrometer, with an interscan interval of 25 ms. ^2H NMR experiments were performed in an Innova 300 MHz spectrometer. NMR data is reported as follows: chemical shift (ppm) (multiplicity, integration, coupling constant, assignment). All signals were referenced to the residual solvent signal. Deuteration yield was quantified by integration of the NMR signals. SiO_2 thin layer chromatographies were monitored by UV light and by the ninhydrin test. HPLC were performed in a Waters Alliance 2695 PDA using an Xbridge C18 2.5 μM , 4.6 x 75 mm column. Exact masses were obtained in an LTQ-FT Ultra (Thermo Scientific) mass spectrometer.

Note: the reaction steps are numbered according to the scheme of Figure 602.

1-Ethoxyalkynyl cyclohexyl alcohol-(d_1) (1): the protocol was adapted from [250]. 4.38 g of commercial ethoxyacetylene solution (40 % w/v, 25 mmol) were dissolved in 75 ml of anhydrous THF, in dry material and under inert atmosphere. The solution was cooled to $-78\text{ }^\circ\text{C}$ in a dry ice/acetone bath, and a 11 ml of a commercial BuLi solution (2.5 M in hexanes, 27.5 mmol, 1.1 equivalents) was added dropwise during 30 min. After this time, 2.85 ml of cyclohexanone were added (27.5 mmol, 1.1 equivalents) and the mixture was stirred during 1 h at $-78\text{ }^\circ\text{C}$. Mixture was allowed to reach room temperature and it was kept during an additional hour at this temperature. Afterwards, the mixture was cooled in an ice bath, and D_2O was added dropwise, followed by a diluted solution of D_3PO_4 (obtained by dissolving anhydrous P_2O_5 in cold D_2O), until an approximate $\text{pH} = 3$. Mixture was extracted 3 times with DCM, organic layers were dried with anhydrous MgSO_4 , and the solvent was evaporated under vacuum. The crude product was purified by silica gel column chromatography in a gradient from 9/1 to 5/1 of hexane/EtOAc, including a 0.5 % of EtOD. 4.27 grams of pure product were obtained (25 mmol, quantitative yield). Pale yellow oil. $R_f = 0.17$ (Hexane/EtOAc 7/1). $M(\text{calc}) = 169.24$, $M(\text{exp}) = 191.1036$ [$\text{M}-1+\text{Na}$] $^+$ as a consequence of D/H exchange of labile hydroxyl D. ^1H NMR (CDCl_3): $\delta = 1.353 - 1.845$ (m, 10 H, CH_2 cyclohexyl); $\delta = 1.371$ (t, 3 H, $^3J_{\text{H-H}} = 7.2$ Hz, CH_3 ethyl); $\delta = 4.083$ (q, 2 H, $^3J_{\text{H-H}} = 7.2$ Hz, CH_2 ethyl). ^{13}C NMR (CDCl_3): $\delta = 14.49$ (CH_3 ethyl); $\delta = 23.73$, $\delta = 25.47$, $\delta = 40.95$ (CH_2 cyclohexyl); $\delta = 42.14$ (C quaternary alkyne); $\delta = 68.92$ (C alcohol); $\delta = 74.64$ (CH_2 ethyl); $\delta = 93.56$ (C quaternary alkyne). ^2H NMR (CDCl_3): $\delta = 6.50$ (s, hydroxyl).

2-Cyclohexylidene-acetic acid ethyl ester-(1 d₁) (2): the Meyer-Schuster rearrangement procedure was adapted from [251]. 4.23 g of the ethoxyalkynyl alcohol **1** (25 mmol) were dissolved in 200 ml of anhydrous DCM and 50 ml of EtOD, in dry material and under inert atmosphere. Mixture was stirred during 15 min at room temperature, after which time 123 mg of scandium triflate were added (0.25 mmol, 0.01 equivalents). Mixture was kept under stirring during 4 h at room temperature. Afterwards, solvent was evaporated under vacuum and the crude was purified by silica gel column chromatography (hexane/AcOEt gradient from 9/1 to 5/1). 3.88 g of α/β -unsaturated ester were obtained (22.93 mmol, 90 % yield). 95 % deuterium incorporation at the vinylic position. Scented colourless syrup. $R_f = 0.51$ (Hexane/EtOAc 7/1). $M(\text{calc}) = 169.24$, $M(\text{exp}) = 170.1282$ $[M+H]^+$. $^1\text{H NMR}$ (CDCl_3): $\delta = 1.267$ (t, 3 H, $^3J_{\text{H-H}} = 7.2$ Hz, CH₃ ethyl); $\delta = 1.594 - 2.834$ (m, 10 H, CH₂ cyclohexyl); $\delta = 4.134$ (q, 2 H, $^3J_{\text{H-H}} = 7.2$ Hz, CH₂ ethyl). $^{13}\text{C NMR}$ (CDCl_3): $\delta = 14.46$ (CH₃ ethyl); $\delta = 26.41$, $\delta = 27.94$, $\delta = 28.76$, $\delta = 29.93$ (CH₂, cyclohexyl); $\delta = 59.57$ (CH₂ ethyl); $\delta = 112.85$ (t, $^1J_{\text{C-D}} = 25$ Hz, CD alkene); $\delta = 163.60$ (C quaternary alkene); $\delta = 166.99$ (C ester). $^2\text{H NMR}$ (CDCl_3): $\delta = 4.13$ (s, D vinyl).

2-Cyclohexylidene ethyl alcohol-(1, 1, 2-d₃) (3): 1.26 g of LiAlD₄ (30 mmol, 1.2 equivalents) were dissolved in 170 ml of anhydrous diethyl ether. The α/β -unsaturated ester **2** (4.23 g, 25 mmol) was added dropwise to the suspension at room temperature. After completion of the addition, the mixture was kept reacting under stirring during 4 h. After this time, mixture was cooled in an ice bath, and the reaction was quenched by the slow addition of D₃PO₄ (obtained as described in step 1) until slightly acidic pH. Afterwards, D₂O was added to the mixture, and it was extracted three times with DCM. The organic layers were dried with anhydrous MgSO₄ and the solvent was eliminated under vacuum. Crude product was purified by silica gel column chromatography, with a gradient of hexane/AcOEt from 9/1 to 7/1 (with 5 % EtOD). 2.69 of a pale yellow oil were obtained (20.7 mmol, 83 % yield). 99 % deuterium incorporation in the allylic position. Pale yellow oil. $R_f = 0.2$ (Hexane/Diethyl ether 3/1). $M(\text{calc}) = 130.22$, $M(\text{exp}) = 130.1590$ $[M-1+H]^+$ as a consequence of D/H exchange of labile hydroxyl D. $^1\text{H NMR}$ (CDCl_3): $\delta = 0.88 - 2.19$ (m, 10 H, cyclohexyl). $^{13}\text{C NMR}$ (CDCl_3): $\delta = 26.41$, 26.83, 28.02, 28.54, 28.95, 37.12 (s, CH₂ cyclohexyl); $\delta = 57.90$ (qui, $^1J_{\text{C-D}} = 22$ Hz, CD₂ allyl), $\delta = 119.92$, (t, $^1J_{\text{C-D}} = 24$ Hz, CD vinyl). $^2\text{H NMR}$ (CDCl_3): $\delta = 1.35$, (OD); $\delta = 4.03$ (CD₂ allylic); $\delta = 5.32$ (CD vinylic).

(2-Bromo-ethylidene)-cyclohexane-(1, 2, 2-d₃) (4): the bromination protocol was adapted from [252]. 2.96 g (20.7 mmol) of the allylic alcohol **3** were dissolved in 71 ml of anhydrous diethyl ether. The mixture was cooled to -20 °C in a ternary bath of acetone/dry

ice/water. 0.97 ml of PBr_3 were added (10.3 mmol, 0.5 equivalents), and the mixture was kept under stirring during 3 h; temperature was allowed to reach 0 °C progressively. Afterwards, reaction was quenched by the addition of 120 ml of cold water. Mixture was extracted three times with diethyl ether, organic layers were washed twice with saturated NaHCO_3 and dried with MgSO_4 . Solvent was evaporated under vacuum, and the crude product was purified by silica gel column chromatography (hexane-hexane/ AcOEt 3/1). 1.70 g of pure product were obtained (8.82 mmol, 43 % yield). Brown oil. $R_f = 0.71$ (Hexane/Diethyl ether 3/1). $M(\text{calc}) = 192.11$, $M(\text{exp}) = 112.1200$ [$\text{M} - \text{Br}$] $^+$. ^1H NMR (CDCl_3): $\delta = 0.88 - 2.22$ (m, 10 H, CH_2 cyclohexyl). ^{13}C NMR (CDCl_3): $\delta = 26.72$, $\delta = 27.66$, $\delta = 28.46$, $\delta = 28.71$, $\delta = 37.06$ (s, CH_2 cyclohexyl); $\delta = 117.13$ (t, $^1J_{\text{C-D}} = 24$ Hz, CD vinylic). ^2H NMR (CDCl_3): $\delta = 3.94$ (s, CD_2 allylic); $\delta = 5.42$ (s, CD vinylic).

Glycine iminic ester (2, 2-d₂) (5): 1.75 g of commercial glycine iminic ester were dissolved in 60 ml of CD_3OD and 3 ml of dioxane, in a dry flask and under inert atmosphere. Separately, a 5M stock solution of NaOCD_3 in CD_3OD was prepared by dissolving carefully Na in the corresponding volume of CD_3OD at 0 °C. 0.1 equivalents of the basic methanolic solution were added, and the mixture was kept under stirring during 12 h at r.t. After this time, 450 ml of DCM were added, and the mixture was passed through a short silica column previously equilibrated by circulating three times 80 ml DCM + 10 ml of CD_3OD . Filtrated solution was evaporated under vacuum, dissolved again in DCM, and the procedure was repeated two additional times. After that, crude product was purified by silica gel column chromatography (DCM/ CD_3OD 99/1), previously conditioned with the same solvent. 1.63 g of the pure product were obtained (5.47 mmol, 92 % yield, 98 % deuterium incorporation). White solid. HPLC: $rt = 5.962$ min (20-100 gradient). $M(\text{calc}) = 297.39$, $M(\text{exp}) = 298.1773$ [$\text{M} + \text{H}$] $^+$. ^1H NMR (CDCl_3): $\delta = 1.46$ (s, 9 H, tBu); $\delta = 7.18 - 7.82$ (m, 10 H, aromatic). ^{13}C NMR (CDCl_3): $\delta = 28.25$ (s, CH_3 tBu); $\delta = 81.20$ (s, quaternary carbon tBu); $\delta = 127.88 - 139.5$ (s, C aromatic); $\delta = 169.93$ (s, C imine), $\delta = 171.76$ (s, C ester). ^2H NMR (CDCl_3): $\delta = 3.98$ ppm (CD_2).

tert-Butyl (2S)-2-(benzhydrylideneamino)-4-(cyclohexylidene)-butanoate (2, 3, 3, 4-d₄) (6): the protocol of the enantioselective alkylation of the glycine iminic ester with a chiral phase transfer reagent was adapted from [253]. 750 mg (2.52 mmol) of the deuterated glycine iminic ester obtained in step 5 were dissolved in 16 ml of toluene. 153 mg of the chiral phase transfer reagent (O-allyl-N-(9-anthracenylmethyl)cinchonidinium bromide, 0.1 equivalents, 0.25 mmol) were added to the solution. The mixture was cooled to 0 °C, and 3.6 ml of 40 % KOD in D_2O (w/v) were added (25.2 mmol, 10 equivalents).

Mixture was stirred vigorously, and 553 mg (2.77 mmol, 1.1 equivalents) of the allylic bromide **4** were added dropwise at 0 °C. The dispersion was kept under vigorous stirring at 0 °C during 24 h. After this time, water was added, and the mixture was extracted 4 times with AcOEt. Organic layers were dried with MgSO₄, and the solvent was evaporated under vacuum. Crude product was purified by silica gel column chromatography (hexane/AcOEt 7/1). 868 mg (2.13 mmol, 84 % yield) of pure product were obtained. Colourless oil. HPLC: rt = 4.217 min (40-100 gradient). M(calc) = 407.58, M(exp) = 408.2840 [M+H]⁺. ¹H NMR (CDCl₃): δ = 1.45 (s, 9 H, tBu); δ = 1.26 – 2.17 (m, 10 H, cyclohexyl); δ = 7.16 – 7.82 (m, 10 H, aromatic). ¹³C NMR (CDCl₃): δ = 28.238 (s, CH₃ tBu); δ = 26.98, 27.76, 28.62, 28.95 (s, CH₂ cyclohexyl); δ = 66.49 (t, ¹J_{C-D} = 20 Hz, CD α position); δ = 80.91 (quaternary carbon tBu); δ = 116.55 (t, ¹J_{C-D} = 23 Hz, CD alkene); δ = 128.07 – 137.74 (C aromatic); δ = 142.05 (s, alkene quaternary carbon); δ = 169.75 (s, imine); δ = 171.37 (ester). ²H NMR (CDCl₃): δ = 2.58 (D β position), δ = 3.93 (D vinyl); δ = 5.03 (D α position).

tert-Butyl (2S)-2-(tert-Butoxycarbonylamino)-4-(cyclohexylidene)-butanoate (2, 3, 3, 4-d₄) (7): the removal of dibenzylideneamine group and introduction of Boc protecting group to **6** was carried out as described in [96]. 2.18 g (5.34 mmol) of the product **6** were dissolved in 18 ml of THF, and the same volume of a solution of 10 % (w/v) citric acid was added. The mixture was stirred at room temperature, and the reaction was monitored by TLC. After the disappearance of the starting reagent, solvent was partially eliminated under vacuum, and a volume of aqueous solution of 15 % (v/v) formic acid was added. Aqueous phase was washed with DCM, and after washing, the pH was raised to 12 by careful addition of 0.1 M NaOH at 0 °C. The aqueous phase was extracted 4 times with AcOEt, organic layers were dried with MgSO₄, and the solvent was evaporated under vacuum.

The resulting oil was dissolved in 9 ml of DMF, and 746 μl of NEt₃ were added (5.34 mmol, 1 equivalent) After that, 1.28 g of di-*tert*-butyl dicarbonate (5.88 mmol, 1.1 equivalents) were added. The mixture was kept under stirring at room temperature during 6 hours. DMF was evaporated under vacuum, and the crude product was dissolved in AcOEt. The mixture was washed three times with an aqueous solution of KHSO₄ pH = 2. The organic layer was dried with MgSO₄ and the solvent was evaporated under vacuum. The crude product was purified by silica gel column chromatography (hexane/AcOEt 7/1). 1.43 g of pure product were obtained (4.16 mmol, 78 % yield). Colourless oil. HPLC: rt = 7.730 min (40-100 gradient). M(calc) = 343.49, M(exp) = 344.2736 [M+H]⁺. ¹H NMR (CDCl₃): δ = 1.43 (s, 9

H, tBu ester); δ = 1.45 (s, 9 H, Boc); δ = 1.26 – 2.05 (m, 10 H, cyclohexyl); δ = 5.00 (s, 1 H, NH). ^{13}C NMR (CDCl_3): δ = 28.17 (s, CH_3 tBu ester); δ = 28.43 (s, CH_3 Boc); δ = 26.95, 27.98, 28.69, 28.98, 37.38 (s, CH_2 cyclohexyl); δ = 32.65 (qui, $^1\text{J}_{\text{C-D}}$ = 27 Hz, CD_2 β position); δ = 79.54 (s, Boc quaternary carbon); δ = 81.71 (s, tBu ester quaternary carbon); δ = 114.23 (t, $^1\text{J}_{\text{C-D}}$ = 22 Hz, CD alkene); δ = 143.94 (s, alkene quaternary carbon); δ = 155.30 (s, CO Boc); δ = 171.62 (s, CO ester). ^2H NMR (CDCl_3): δ = 2.30 (CD_2 β position); δ = 4.00 (CD vinylic); δ = 4.87 (CD α position).

***N*-Boc homoserine *tert*-butyl ester (2, 3, 3, 4, 4- d_5) (8):** the protocol of the reductive ozonolysis was adapted from Siebum *et al.* [96]. 1.43 g of product **6** (4.16 mmol) were dissolved in 200 ml of DCM, in a two neck pear-shaped flask. The mixture was cooled to -78 °C, and ozone (generated *in situ* by an ozone generator) was bubbled through the solution until a persistent light blue colour was observed. The reaction was followed by TLC until complete disappearance of the starting reagent. At this point, the system was purged with N_2 and 460 μl of dimethyl sulfide (6.25 mmol, 1.5 equivalents) dissolved in 10 ml of DCM were added dropwise. The mixture was allowed to reach 0 °C gradually during 2 hours. After this time, the solvent was evaporated under vacuum, and the resulting crude aldehyde was dissolved in 100 ml of ethanol at 0 °C. 450 mg of NaBD_4 (10.8 mmol, 2.6 equivalents) were added to the previous ethanolic solution. The reaction was controlled by TLC, and after completion it was quenched by the addition of cold water. Subsequently, cold diluted HCl was added until the pH was approximately 6. The ethanol was eliminated under vacuum, and more water was added. The aqueous suspension was extracted with diethyl ether, and the organic layer was neutralized with saturated KHCO_3 . The organic layers were washed three additional times with water. The organic layer was dried with MgSO_4 and the solvent was evaporated under vacuum. The crude product was purified by silica gel column chromatography (hexane/AcOEt gradient 3/1 to 2/1). 803 mg of pure product were obtained (2.86 mmol, 69 % yield). 97 % deuterium incorporation at γ position. Colourless oil. HPLC: t_r = 5.795 min (0-100 gradient). $M(\text{calc})$ = 280.37, $M(\text{exp})$ = 281.1723 [M+H]. ^1H NMR (CDCl_3): δ = 1.44 (s, 9 H, tBu ester); δ = 1.46 (s, 9 H, tBu Boc); δ = 5.32 (s, 1 H, NH). ^{13}C NMR (CDCl_3): δ = 28.13 (s, CH_3 tBu ester); δ = 28.41 (s, CH_3 Boc); δ = 36.10 (qui, $^1\text{J}_{\text{C-D}}$ = 12 Hz, CD_2 β position); δ = 50.60 (t, $^1\text{J}_{\text{C-D}}$ = 21 Hz, CD α position); δ = 57.61 (qui, $^1\text{J}_{\text{C-D}}$ = 19 Hz, CD_2 γ position); δ = 80.48 (s, Boc quaternary carbon); δ = 82.44 (s, tBu ester quaternary carbon); δ = 156.78 (s, CO Boc); δ = 172.17 (s, CO tBu ester). ^2H NMR (CDCl_3): δ = 2.10 (CD_2 β position); δ = 3.53 (CD_2 γ position); δ = 4.16 (CD α position).

S-Acetyl-N-Boc homocysteine *tert*-Butyl ester (2, 3, 3, 4, 4-d₅) (9): the procedure of the Mitsunobu reaction was adapted from Siebum *et al.* [96]. 1.39 g of triphenylphosphine (5.31 mmol, 2 equivalents) were dissolved in 30 ml of anhydrous THF, in dry material and under inert atmosphere. The mixture was cooled to 0 °C and 1.05 ml of diisopropyl azodicarboxylate (DIAD, 5.31 mmol, 2 equivalents) were added to the previous triphenylphosphine solution. After 10 min, 744 mg of protected homoserine (2, 3, 3, 4, 4-d₅) (2.65 mmol) dissolved in 7 ml of anhydrous THF were added. The suspension was kept under stirring at 0 °C during 45 min. After this time, 395 µl of thioacetic acid (5.31 mmol, 2 equivalents) were added to the previous pale pink suspension, and the mixture was kept under stirring at 0 °C during 1.5 h. Afterwards, temperature was raised to room temperature and the mixture was kept under stirring at these conditions during 2.5 h. Subsequently, diethyl ether was added, and the mixture was washed with water and brine. The organic phase was dried with MgSO₄, and the solvent was evaporated under vacuum. The crude product was purified by silica gel column chromatography (hexane/AcOEt gradient from 7/1 to 5/1). 879 mg of pure S-Acetyl-N-Boc homocysteine *tert*-Butyl Ester (2, 3, 3, 4, 4-d₅) were obtained (2.60 mmol, 98 % yield). White solid. HPLC: rt = 4.973 min (40-100 gradient). M(calc) = 338.47, M(exp) = 339.1997 [M+H]⁺. ¹H NMR (CDCl₃): δ = 1.44 (s, 9 H, CH₃ tBu ester); δ = 1.48 (s, 9 H, CH₃ Boc); δ = 2.32 (s, 3 H, CH₃ acetyl); δ = 5.11 (s, 1 H, NH). ¹³C NMR (CDCl₃): δ = 28.11 (s, CH₃ tBu ester); δ = 28.47 (s, CH₃ Boc); δ = 30.74 (s, CH₃ acetyl); δ = 32.74 (qui, ¹J_{C-D} = 14 Hz, CD₂ β position); δ = 79.65 (s, Boc quaternary carbon); δ = 82.35 (tBu ester quaternary carbon); δ = 155.50 (s, CO Boc); δ = 171.27 (s, CO tBu ester); δ = 195.61 (s, CO acetyl). ²H NMR (CDCl₃): δ = 1.93 (CD₂ β position); δ = 2.91 (CD₂ γ position); δ = 4.24 (CD α position).

N-Boc-[methyl-¹³C]-methionine (2, 3, 3, 4, 4-d₅) (10): hydrolysis of the thioester and alkylation of the free thiol proceed as described by Siebum *et al.* [96]. 868 mg of S-Acetyl-N-Boc homocysteine *tert*-Butyl ester (2, 3, 3, 4, 4-d₅) (**8**, 2.56 mmol) were dissolved in 35 ml of degassed ethanol. 15 ml of aqueous solution of NaOH 2M were added under inert atmosphere, and the mixture was stirred at r.t. The reaction was followed by TLC. After the complete hydrolysis of the thioester, 322 µl of ¹³CH₃I (5.13 mmol, 2 equivalents) were added, and the mixture was stirred overnight at room temperature. Afterwards, the solvent was partially evaporated under vacuum, and EtOAc was added. The organic solution was extracted with water, and the aqueous layer was cooled in an ice bath. KHSO₄ solution (pH = 2) was added until neutral pH was reached; afterwards, pH was lowered to 3 with 10 % citric acid. The aqueous mixture was extracted with EtOAc, acidified with HCl 1M to pH = 1, and extracted again with EtOAc. The organic layer was dried with MgSO₄, and the solvent

was evaporated under vacuum. Small amounts of toluene were added to the remaining oil followed by evaporation under vacuum in order to remove remaining traces of $^{13}\text{CH}_3\text{I}$ and thioacetic acid. 557 mg of pure *N*-Boc-[methyl- ^{13}C]-methionine (2, 3, 3, 4, 4- d_5) were obtained (2.19 mmol, 86 % yield). White solid. The product was used in the next step without any further purification. HPLC: $t_r = 5.367$ min (0-100 gradient). $M(\text{calc}) = 255.14$, $M(\text{exp}) = 256.1455$ $[\text{M}+\text{H}]^+$. ^1H NMR (CD_3OD): $\delta = 1.45$ (s, 9 H, CH_3 Boc); $\delta = 2.08$ (d, 3 H, $^1J_{\text{C-H}} = 136$ Hz, $^{13}\text{CH}_3$). ^{13}C NMR (CD_3OD): $\delta = 15.09$ (s, CH_3 thioether); $\delta = 28.71$ (s, CH_3 Boc); $\delta = 80.52$ (s, Boc quaternary carbon); $\delta = 158.18$ (s, CO Boc); $\delta = 175.93$ (s, CO (carboxyl)). ^2H NMR (CD_3OD): $\delta = 2.61$ (CD_2 β position); $\delta = 3.27$ (CD_2 γ position); $\delta = 4.06$ (CD α position).

[Methyl- ^{13}C]-methionine (2, 3, 3, 4, 4- d_5) hydrochloride (11): 5 ml of HCl 4 M in dioxane were added to 557 mg of *N*-Boc-[methyl- ^{13}C]-methionine (2, 3, 3, 4, 4- d_5) (9), and the solution was stirred during 1 h at room temperature. After this time, the solvent was evaporated under vacuum, and the crude product was dissolved in water acidified with HCl. The aqueous solution was extracted with DCM, and after that it was lyophilized. The crude product was dissolved in a small volume of water, and it was loaded into a Porapak Rxn RP reverse-phase cartridge (Waters Corp., Milford, USA). The product was eluted with a $\text{H}_2\text{O}/\text{ACN}$ gradient from 100/0 to 95/5 (acidified with 1 % of HCl). Fractions containing the pure product were lyophilized. The pure product was redissolved in D_2O + DCl and lyophilized three times to exchange all labile protons. 402 mg of pure [methyl- ^{13}C]-methionine (2, 3, 3, 4, 4- d_5) were obtained (2.10 mmol, 96 % yield). White solid. HPLC: $t_r = 2.464$ min (0-50 gradient). $M(\text{calc}) = 191.06$, $M(\text{exp}) = 156.0931$ $[\text{M}-\text{Cl}]^+$. ^1H NMR (D_2O): $\delta = 2.09$ (d, 3 H, $^1J_{\text{C-H}} = 140$ Hz, $^{13}\text{CH}_3$). ^{13}C NMR (D_2O): $\delta = 13.73$ (s, CH_3 thioether); $\delta = 28.71$ (s, CH_3 Boc); $\delta = 27.65$ (qui, $^1J_{\text{C-D}} = 21$ Hz, CD_2 γ position); $\delta = 51.24$ (t, $^1J_{\text{C-D}} = 23$, CD α position); $\delta = 80.52$ (s, Boc quaternary carbon); $\delta = 158.18$ (s, CO Boc); $\delta = 175.93$ (s, CO (carboxyl)). ^2H NMR (D_2O): $\delta = 2.15$ (CD_2 β position); $\delta = 2.66$ (CD_2 γ position); $\delta = 4.10$ (CD α position).

The pure [methyl- ^{13}C]-methionine (2, 3, 3, 4, 4- d_5) was dissolved in D_2O to a final concentration of 50 g/l (1000 X stock), filtered through a sterile filter of 0.22 μm and stored at -20 $^\circ\text{C}$ until needed.

NMR EXPERIMENTS

All NMR experiments were performed at 25 °C in a Bruker 800 MHz spectrometer equipped with a cryoprobe. POP samples were in Tris d_{11} -HCl 50 mM pH 8, NaCl 20 mM, DTT d_6 1 mM, NaN_3 0.03 %, and 100 % D_2O buffer. Spectra were centred at 2, 15 ppm, with spectral windows of 3 and 10 ppm (F2, F1). High field 5 mm tubes or Shigemitsu NMR tubes were used, depending on the volume of the sample.

1H - ^{13}C methyl-TROSY HMQC experiments of [methyl- ^{13}C]-methionine labeled POP

For these experiments, POP samples in a concentration typically between 150 - 200 μM were used. Experiments were carried out with the pulse sequence described by Tugarinov *et al.* [80] (Figure 14 C). Spectra comprised 128, 512 data points (F1, F2), with 240 scans per FID separated by a delay of 1.5 s. Spectra were processed by Toppin software, and automatic baseline correction was applied to direct and indirect dimensions.

RD experiments of U- ^{2}H], [methyl- ^{13}C , 1H]-methionine labeled POP

The pulse sequence of 1H - ^{13}C methyl-TROSY HMQC RD experiments was adapted from Korzhnev and co-workers [60] (Figure 21 B). A CPMG element of 40 ms, with 0, 2, 4, 6, 8, 10, 12, 16, 20, 24, 28, 32, 36 and 40 inversion pulses was used. Experiments with different inversion pulses were randomly ordered. Spectra were recorded using 100, 512 data points (F1, F2), accumulating 24 scans separated by an interval of 1.5 s. NMRPipe software [254] was used for spectral processing. First, NMR data was converted to NMRPipe format by the convert script. Subsequently, converted data was processed by the process script. Conversion scripts for Bruker data and processing scripts are shown in Figure 82 A and B, respectively: automatic zero filling and forward-backward linear prediction were chosen for the processing. Finally, the integration of the peaks were calculated with the integration script (since the integration box was the same for all peaks, the ratio between integrations was equivalent to the ratio between intensities, I_{VCPMG}/I_0). The calculation of effective decay rates and the analysis of RD data was carried out as described in the main text.

$R_{1\rho}$ experiments of perdeuterated [methyl- ^{13}C]-methionine labeled POP

Off-resonance $R_{1\rho}$ experiments were adapted from Korzhnev *et al.* [102]. ^{13}C spin-lock field strengths of 6, 8, 9.6 and 13 KHz were used. Spectra consisted of 100, 1024 data points (F1, F2), with 8 scans separated by a delay of 1.3 s. NMR data was processed analogously to ^1H - ^{13}C methyl-TROSY HMQC RD experiments.

```
A
#!/bin/csh

foreach xx (9 10 11 12 13 14 15 16 17 18 19 20 21 22 23)

bruk2pipe -in ../fids/$xx/ser \
  -bad 0.0 -aswap -DMX -decim 8330.6666666667 -dspfvs 20 -
  grpdlly 67.9856719970703 \
  -xN 512 -yN 100 \
  -xT 256 -yT 50 \
  -xMODE DQD -yMODE States-TPPI \
  -xSW 2400.77 -ySW 2011.96 \
  -xOBS 800.132 -yOBS 201.20 \
  -xCAR 2 -yCAR 15 \
  -xLAB 1H -yLAB 13C \
  -ndim 2 -aq2D States-TPPI \
  -out ./$xx.fdc -verb -ov

sleep 5

end

B
#!/bin/csh

foreach xx (9 10 11 12 13 14 15 16 17 18 19 20 21 22 23)

nmrPipe -in ./$xx.fdc \
  #| nmrPipe -fn POLY -time \
  \
  | nmrPipe -fn SP -off 0.35 -end 1.0 -pow 2 -c 1.0 \
  | nmrPipe -fn ZF -auto \
  | nmrPipe -fn FT -verb \
  \
  #| nmrPipe -fn EXT -left -sw \
  \
  | nmrPipe -fn PS -p0 -200.2 -p1 2.2 -
  di -verb\
  | nmrPipe -fn TP -auto \
  \
  | nmrPipe -fn SP -off .37 -end 0.98 -pow 2 -c 1.0 \
  | nmrPipe -fn ZF -auto -verb \
  \
  | nmrPipe -fn LP -fb \
  | nmrPipe -fn FT -auto \
  \
  #| nmrPipe -fn REV -sw \
  | nmrPipe -fn PS -p0 -12.2 -p1 0.0 -
  di -verb\
  | nmrPipe -fn TP -auto \
  \
  #| nmrPipe -fn POLY -auto \
  \
  -out ./$xx.ft2 -ov

end
```

Figure 82: Scripts used for the conversion of Bruker data and for spectral processing. The highlighted sections were introduced by the user. A) Convert script. The second line indicates the number of the experiments to be converted; the fourth and fifth correspond to a header obtained from the nmrDraw program; the rest of the sections correspond to spectral parameters. B) Process script. The second line indicates the number of the experiments to be processed; lines 12 and 23 are the phasing angles of ^1H and ^{13}C , respectively.

Processing of ^1H - ^{15}N HSQC experiments of U- ^{15}N α -synuclein

The fids of ^1H - ^{15}N HSQC experiments were converted from Bruker to NMRPipe format with the same convert and process scripts used in the processing of RD experiments. For the conversion to NMRPipe format, basic frequencies, spectral windows and carrier waves were adapted to ^1H - ^{15}N HSQC experiments (Figure 82 A); spectral reference was also included in the carrier wave in order to reference the spectra. Afterwards, peak piking was carried out using the default values of the NMRDraw interface; raw data was extracted from

archive.tab files and transferred to an excel file. After manual peak assignment, peak integrations and chemical shifts were used to calculate I/I_0 and CSP, respectively.

SIZE EXCLUSION CHROMATOGRAPHY COUPLED TO SAXS MEASUREMENTS

The experimental set-up consisted on a Superdex 200 10/300 SEC column (GE Healthcare Life Sciences, Uppsala, Sweden) coupled to EMBL P12 beamline of the storage ring PETRA III (DESY, Hamburg, Germany), using a PILATUS 2M pixel detector (DECTRIS, Switzerland). The X-ray beam wavelength was 1.24 Å, and the range of momentum transfer covered was $0.007 < s < 0.444 \text{ \AA}^{-1}$. POP was in 50 mM Tris-HCl pH 8, 20 mM NaCl, 1 mM DTT buffer, and 50 µl of protein sample at 245 µM concentration were injected to the instrument. The column was run at 0.35 ml/min using a Malvern instrument, acquiring 1 SAXS frame per second.

THEORETICAL SIMULATIONS

Molecular dynamics simulations of POP

The MD simulations of POP used to describe the experimental SAXS data were performed by Dr. Martin Kotev, under the direction of Dr. Victor Guallar (members of joint BSC-CRG-IRB Research Program in Computational Biology in Barcelona Supercomputing Center). Calculations were performed with MinoTauro NVIDIA GPU cluster, at BSC.

The three MD simulations of free POP were performed with AMBER12 software [255]. The parm99 with backbone torsion modifications force field was used [256], and waters were incorporated as TIP3P water model [257]. The POP protein structures were neutralized and additional number of sodium and chlorine ions were added to simulate physiological saline solution. Then protein plus ions were solvated in pre-equilibrated water molecules in a truncated octahedron solvent box. After minimization of only ions and water molecules, the same procedure was repeated for the entire system, and the temperature was progressively raised to 300 K using a constant pressure dynamics. All three simulations were done with a time step of 2.0 fs in the production run and SHAKE algorithm was used to constrain bond lengths involving hydrogen atoms [121, 258]. Structures were saved in every 20 ps. Non-

bonded interactions were explicitly evaluated for distances below 9 Å. All production simulations were performed in NPT ensemble (1 bar and 298 K) with Berendsen barostat and thermostat, respectively [120]. Particle mesh Ewald [259] method was used to treat long-range electrostatic interactions.

In the case of MD simulations of inhibited POP, Desmond molecular dynamics program was used [125, 260]. The parameters from OPLS-AA (optimized parameters for liquid simulations-all atoms) force field and TIP3P water model were used [257]. Inhibited POP structures with non-covalently and covalently bound ZPP were solvated in an orthorhombic box of water molecules and ions were added to neutralize and create an approximate physiological solution of NaCl. The default relaxation protocol in Desmond was used, followed by 50 ns and 40 ns production run, respectively, in the NPT ensemble using the Nose-Hoover thermostat and the Martyna-Tobias-Klein barostat [261, 262].

All trajectory analysis and data processing were done by CPPTRAJ and VMD software packages [126, 263]. Salt bridges and hydrogen bonds were analysed with VMD software, according to default parameters. Solvent exposure was calculated according to LCPO algorithm [123].

Molecular dynamics simulations for IMMS experiments and theoretical collision cross sections calculations

All the simulations used to analyse POP data obtained by IMMS were carried out by Dr. Sergio Madurga (Department of Physical Chemistry of the University of Barcelona, Barcelona, Spain).

All REMD in solution were performed with an ionic strength of 0.1 M of NaCl. Periodic boundary conditions were applied in a cube of about 10 nm edge length. First, the energy of the system was minimized and equilibrated by means of 100 ps of a molecular dynamics calculation with protein heavy atoms restrained, in a total of 1000 ps. REMD simulations were performed using 18 replicas distributed over a temperature range from 300 to 327 °K, where multiple copies (or replicas) of identical systems were simulated in parallel at different temperatures. Exchange occurred every 400 ps at each temperature.

All MD simulations in the gas phase were performed using the GROMACS software package version 4.5.5 [264, 265] and the Amber03 force field, starting from the protonated structures derived from the Monte Carlo procedure. Water molecules were incorporated with the TIP3P model. An integration time step of 2 fs was used, with all bonds constrained

using the LINCS algorithm [266]. A cut-off value of 1.0 nm was established for Van der Waals interactions. Electrostatic interactions were treated with the particle mesh Ewald (PME) method applying a real-space cut-off of 1.0 nm [259]. No cut-off was used for either electrostatic or Lennard-Jones interactions in the gas phase. The system temperature was controlled with the V-rescale procedure [267] and no pressure coupling was used. The system was equilibrated during the first 1000 ps after the production run.

The theoretical determination of collision cross sections were performed with the MOBCAL [172, 173] program. The trajectory method (TM) was used to calculate the CCS values. For CCS corresponding to simulations in solution, water molecules were removed, and the POP structure was assigned a charge of + 20.

EXPERIMENTAL PROCEDURE FOR IMMS EXPERIMENTS

Native POP expression was performed as described in the Protein expression section.

For the preparation of MS buffers, ultrapure water and reagents (analytical grade) were used; pH adjustment was done by taking aliquots in order to avoid direct contact of the electrode with the solution. Low-binding eppendorfs, autoclaved material and carefully cleaned glass labware were used in order to avoid contamination with polymers, plastic additives, detergents or salts.

Prior to the injection to the mass spectrometer, samples were buffer exchanged using Micro Bio-Spin P-30 columns, following the commercial protocol. Two consecutive Bio-Spin P-30 columns were performed. In the case of POP bound with inhibitors, incubation proceed as habitual in Tris-HCl buffer; the excess of inhibitor was completely removed by the double buffer exchange operation.

Individual samples were charged in the sample well plate of the Triversa Nanomate nanoESI source (Advion Inc., New York, USA) and were immediately delivered to the spectrometer. Experiments were monitored by checking the intensity and stability of the detector signal; typically, one IMMS run was manually recorded during 1 min.

In the case of negative ionization experiments, the sample was loaded in a Hamilton syringe and injection was carried out using the automatic sample deliverer of the SYNAPT spectrometer. The sample flow was manually adjusted in order to achieve good signal intensity.

Bibliography

1. Cunningham, D.F. and B. O'Connor, *A study of prolyl endopeptidase in bovine serum and its relevance to the tissue enzyme*. Int J Biochem Cell Biol, 1998. **30**(1): p. 99-114.
2. Fulop, V., Z. Bocskei, and L. Polgar, *Prolyl oligopeptidase: an unusual beta-propeller domain regulates proteolysis*. Cell, 1998. **94**(2): p. 161-70.
3. Ollis, D.L., et al., *The alpha/beta hydrolase fold*. Protein Eng, 1992. **5**(3): p. 197-211.
4. Polgar, L., *The prolyl oligopeptidase family*. Cell Mol Life Sci, 2002. **59**(2): p. 349-62.
5. Juhasz, T., et al., *Unclosed beta-propellers display stable structures: implications for substrate access to the active site of prolyl oligopeptidase*. J Mol Biol, 2005. **346**(3): p. 907-17.
6. Baker, S.C., et al., *Cytochrome cd1 structure: unusual haem environments in a nitrite reductase and analysis of factors contributing to beta-propeller folds*. J Mol Biol, 1997. **269**(3): p. 440-55.
7. Shan, L., Mathews, II, and C. Khosla, *Structural and mechanistic analysis of two prolyl endopeptidases: role of interdomain dynamics in catalysis and specificity*. Proc Natl Acad Sci U S A, 2005. **102**(10): p. 3599-604.
8. Li, M., et al., *Induced-fit mechanism for prolyl endopeptidase*. J Biol Chem, 2010. **285**(28): p. 21487-95.
9. Yoshimoto, T., et al., *Comparison of inhibitory effects of proline-containing peptide derivatives on prolyl endopeptidases from bovine brain and Flavobacterium*. J Biochem, 1985. **98**(4): p. 975-9.
10. Fulop, V., et al., *Structures of prolyl oligopeptidase substrate/inhibitor complexes. Use of inhibitor binding for titration of the catalytic histidine residue*. J Biol Chem, 2001. **276**(2): p. 1262-6.
11. Kaszuba, K., et al., *Molecular dynamics, crystallography and mutagenesis studies on the substrate gating mechanism of prolyl oligopeptidase*. Biochimie, 2012. **94**(6): p. 1398-411.
12. Schechter, I. and A. Berger, *On the size of the active site in proteases. I. Papain*. Biochem Biophys Res Commun, 1967. **27**(2): p. 157-62.
13. Szeltner, Z., et al., *Electrostatic effects and binding determinants in the catalysis of prolyl oligopeptidase. Site specific mutagenesis at the oxyanion binding site*. J Biol Chem, 2002. **277**(45): p. 42613-22.
14. Venalainen, J.I., et al., *Binding kinetics and duration of in vivo action of novel prolyl oligopeptidase inhibitors*. Biochem Pharmacol, 2006. **71**(5): p. 683-92.
15. Tsutsumi, S., et al., *Synthesis and structure-activity relationships of peptidyl alpha-keto heterocycles as novel inhibitors of prolyl endopeptidase*. J Med Chem, 1994. **37**(21): p. 3492-502.
16. Tsuru, D., et al., *Thiazolidine derivatives as potent inhibitors specific for prolyl endopeptidase*. J Biochem, 1988. **104**(4): p. 580-6.
17. Morain, P., et al., *S 17092: a prolyl endopeptidase inhibitor as a potential therapeutic drug for memory impairment. Preclinical and clinical studies*. CNS Drug Rev, 2002. **8**(1): p. 31-52.
18. Nakajima, T., et al., *Y-29794--a non-peptide prolyl endopeptidase inhibitor that can penetrate into the brain*. Neurosci Lett, 1992. **141**(2): p. 156-60.
19. Tarrago, T., et al., *Baicalin, a prodrug able to reach the CNS, is a prolyl oligopeptidase inhibitor*. Bioorg Med Chem, 2008. **16**(15): p. 7516-24.

20. Tarrago, T., et al., *The natural product berberine is a human prolyl oligopeptidase inhibitor*. ChemMedChem, 2007. **2**(3): p. 354-9.
21. Lopez, A., et al., *Applications and future of ion mobility mass spectrometry in structural biology*. New Journal of Chemistry, 2013. **37**(5): p. 1283-1289.
22. Lawandi, J., et al., *Inhibitors of prolyl oligopeptidases for the therapy of human diseases: defining diseases and inhibitors*. J Med Chem, 2010. **53**(9): p. 3423-38.
23. Reichelt, K.L., et al., *Biologically active peptide-containing fractions in schizophrenia and childhood autism*. Adv Biochem Psychopharmacol, 1981. **28**: p. 627-43.
24. Maes, M., et al., *Lower serum prolyl endopeptidase enzyme activity in major depression: Further evidence that peptidases play a role in the pathophysiology of depression*. Biological Psychiatry, 1994. **35**(8): p. 545-552.
25. Garcia-Horsman, J.A., P.T. Mannisto, and J.I. Venalainen, *On the role of prolyl oligopeptidase in health and disease*. Neuropeptides, 2007. **41**(1): p. 1-24.
26. Toide, K., et al., *Effect of a novel prolyl endopeptidase inhibitor, JTP-4819, on prolyl endopeptidase activity and substance P- and arginine-vasopressin-like immunoreactivity in the brains of aged rats*. J Neurochem, 1995. **65**(1): p. 234-40.
27. Toide, K., et al., *Effect of a novel prolyl endopeptidase inhibitor, JTP-4819, on spatial memory and central cholinergic neurons in aged rats*. Pharmacol Biochem Behav, 1997. **56**(3): p. 427-34.
28. Williams, R.S., et al., *Loss of a prolyl oligopeptidase confers resistance to lithium by elevation of inositol (1,4,5) trisphosphate*. EMBO J, 1999. **18**(10): p. 2734-2745.
29. Williams, R.S., et al., *A common mechanism of action for three mood-stabilizing drugs*. Nature, 2002. **417**(6886): p. 292-5.
30. Schulz, I., et al., *Modulation of inositol 1,4,5-triphosphate concentration by prolyl endopeptidase inhibition*. Eur J Biochem, 2002. **269**(23): p. 5813-20.
31. Brandt, I., et al., *Search for substrates for prolyl oligopeptidase in porcine brain*. Peptides, 2005. **26**(12): p. 2536-2546.
32. Jalkanen, A.J., et al., *Beneficial Effect of Prolyl Oligopeptidase Inhibition on Spatial Memory in Young but Not in Old Scopolamine-Treated Rats*. Basic & Clinical Pharmacology & Toxicology, 2007. **100**(2): p. 132-138.
33. Myohanen, T.T., et al., *Cellular and subcellular distribution of rat brain prolyl oligopeptidase and its association with specific neuronal neurotransmitters*. J Comp Neurol, 2008. **507**(5): p. 1694-708.
34. Jalkanen, A.J., et al., *The effect of prolyl oligopeptidase inhibition on extracellular acetylcholine and dopamine levels in the rat striatum*. Neurochem Int, 2012. **60**(3): p. 301-9.
35. Peltonen, I., T.T. Myohanen, and P.T. Mannisto, *Association of prolyl oligopeptidase with conventional neurotransmitters in the brain*. CNS Neurol Disord Drug Targets, 2011. **10**(3): p. 311-8.
36. Di Daniel, E., et al., *Prolyl oligopeptidase binds to GAP-43 and functions without its peptidase activity*. Mol Cell Neurosci, 2009. **41**(3): p. 373-82.
37. Spillantini, M.G., et al., *Alpha-synuclein in Lewy bodies*. Nature, 1997. **388**(6645): p. 839-40.
38. Benowitz, L.I. and A. Routtenberg, *GAP-43: an intrinsic determinant of neuronal development and plasticity*. Trends Neurosci, 1997. **20**(2): p. 84-91.
39. Szeltner, Z., et al., *GAP43 shows partial co-localisation but no strong physical interaction with prolyl oligopeptidase*. Biochim Biophys Acta, 2010. **1804**(12): p. 2162-76.
40. Yoshimoto, T., et al., *Specific inhibitors for prolyl endopeptidase and their anti-amnesic effect*. J Pharmacobiodyn, 1987. **10**(12): p. 730-5.

41. Shinoda, M., A. Matsuo, and K. Toide, *Pharmacological studies of a novel prolyl endopeptidase inhibitor, JTP-4819, in rats with middle cerebral artery occlusion.* Eur J Pharmacol, 1996. **305**(1-3): p. 31-8.
42. Jalkanen, A.J., et al., *Brain pharmacokinetics of two prolyl oligopeptidase inhibitors, JTP-4819 and KYP-2047, in the rat.* Basic Clin Pharmacol Toxicol, 2011. **109**(6): p. 443-51.
43. Myohanen, T.T., et al., *A prolyl oligopeptidase inhibitor, KYP-2047, reduces alpha-synuclein protein levels and aggregates in cellular and animal models of Parkinson's disease.* Br J Pharmacol, 2012. **166**(3): p. 1097-113.
44. Schneider, J.S., M. Giardiniere, and P. Morain, *Effects of the prolyl endopeptidase inhibitor S 17092 on cognitive deficits in chronic low dose MPTP-treated monkeys.* Neuropsychopharmacology, 2002. **26**(2): p. 176-82.
45. Lopez, A., et al., *Peptide POP inhibitors for the treatment of the cognitive symptoms of schizophrenia.* Future Med Chem, 2013. **5**(13): p. 1509-23.
46. Lopez, A., T. Tarrago, and E. Giralt, *Low molecular weight inhibitors of Prolyl Oligopeptidase: a review of compounds patented from 2003 to 2010.* Expert opinion on therapeutic patents, 2011. **21**(7): p. 1023-44.
47. Lambeir, A.M., *Translational research on prolyl oligopeptidase inhibitors: the long road ahead.* Expert opinion on therapeutic patents, 2011. **21**(7): p. 977-81.
48. Polgar, L., *pH-dependent mechanism in the catalysis of prolyl endopeptidase from pig muscle.* Eur J Biochem, 1991. **197**(2): p. 441-7.
49. Polgar, L., *Prolyl endopeptidase catalysis. A physical rather than a chemical step is rate-limiting.* Biochem J, 1992. **283 (Pt 3)**: p. 647-8.
50. Fulop, V., Z. Szeltner, and L. Polgar, *Catalysis of serine oligopeptidases is controlled by a gating filter mechanism.* EMBO Rep, 2000. **1**(3): p. 277-81.
51. Szeltner, Z., et al., *Concerted structural changes in the peptidase and the propeller domains of prolyl oligopeptidase are required for substrate binding.* J Mol Biol, 2004. **340**(3): p. 627-37.
52. Kotev, M., et al., *Unveiling Prolyl Oligopeptidase Ligand Migration by Comprehensive Computational Techniques.* Biophysical Journal, 2015. **108**(1): p. 116-125.
53. Canning, P., et al., *Crystal structures of Trypanosoma brucei oligopeptidase B broaden the paradigm of catalytic regulation in prolyl oligopeptidase family enzymes.* PLoS One, 2013. **8**(11): p. e79349.
54. Szeltner, Z., et al., *The loops facing the active site of prolyl oligopeptidase are crucial components in substrate gating and specificity.* Biochim Biophys Acta, 2013. **1834**(1): p. 98-111.
55. St-Pierre, J.-F.o., et al., *Use of Umbrella Sampling to Calculate the Entrance/Exit Pathway for Z-Pro-Prolinal Inhibitor in Prolyl Oligopeptidase.* Journal of Chemical Theory and Computation, 2011. **7**(6): p. 1583-1594.
56. Borrelli, K.W., et al., *PELE: Protein Energy Landscape Exploration. A Novel Monte Carlo Based Technique.* Journal of Chemical Theory and Computation, 2005. **1**(6): p. 1304-1311.
57. Tarragó, T., et al., *A new side opening on prolyl oligopeptidase revealed by electron microscopy.* FEBS Lett, 2009. **583**(20): p. 3344-3348.
58. Kichik, N., et al., *¹⁵N relaxation NMR studies of prolyl oligopeptidase, an 80 kDa enzyme, reveal a pre-existing equilibrium between different conformational states.* Chembiochem, 2011. **12**(18): p. 2737-9.
59. Tarragó, T., et al., *A Cost-Effective Labeling Strategy for the NMR Study of Large Proteins: Selective ¹⁵N-Labeling of the Tryptophan Side Chains of Prolyl Oligopeptidase.* Chembiochem, 2009. **10**(17): p. 2736-2739.

60. Korzhnev, D.M., et al., *Probing Slow Dynamics in High Molecular Weight Proteins by Methyl-TROSY NMR Spectroscopy: Application to a 723-Residue Enzyme*. Journal of the American Chemical Society, 2004. **126**(12): p. 3964-3973.
61. Goto, N., et al., *A robust and cost-effective method for the production of Val, Leu, Ile (δ 1) methyl-protonated ^{15}N -, ^{13}C -, ^2H -labeled proteins*. J Biomol NMR, 1999. **13**(4): p. 369-374.
62. Gardner, K.H. and L.E. Kay, *THE USE OF ^2H , ^{13}C , ^{15}N MULTIDIMENSIONAL NMR GTO STUDY THE STRUCTURE AND DYNAMICS OF PROTEINS*. Annual Review of Biophysics and Biomolecular Structure, 1998. **27**(1): p. 357-406.
63. Pauling, L., *The Nature of the Chemical Bond*. Ithaca: Cornell Univ. Press, 1960. **3**: p. 644.
64. Danielson, M.A. and J.J. Falke, *Use of ^{19}F NMR to Probe Protein Structure and Conformational Changes*. Annual Review of Biophysics and Biomolecular Structure, 1996. **25**(1): p. 163-195.
65. Luck, L.A. and J.J. Falke, *(^{19}F) NMR Studies of the d-Galactose Chemosensory Receptor. 1. Sugar Binding Yields a Global Structural Change*. Biochemistry, 1991. **30**(17): p. 4248-4256.
66. Li, C., et al., *Protein ^{19}F NMR in Escherichia coli*. Journal of the American Chemical Society, 2009. **132**(1): p. 321-327.
67. Hiyama, T., *Organofluorine compounds*. Springer, 2000: p. 3-3.
68. Budisa, N., et al., *Efforts towards the Design of 'Teflon' Proteins: In vivo Translation with Trifluorinated Leucine and Methionine Analogues*. Chemistry & Biodiversity, 2004. **1**(10): p. 1465-1475.
69. Duewel, H., et al., *Incorporation of Trifluoromethionine into a Phage Lysozyme: Implications and a New Marker for Use in Protein ^{19}F NMR*. Biochemistry, 1997. **36**(11): p. 3404-3416.
70. Holzberger, B., et al., *A Highly Active DNA Polymerase with a Fluorous Core*. Angewandte Chemie International Edition, 2010. **49**(7): p. 1324-1327.
71. Krishnamurti, R., D.R. Bellew, and G.K.S. Prakash, *Preparation of trifluoromethyl and other perfluoroalkyl compounds with (perfluoroalkyl)trimethylsilanes*. The Journal of Organic Chemistry, 1991. **56**(3): p. 984-989.
72. Ruppert, I., K. Schlich, and W. Volbach, *Die ersten CF_3 -substituierten organyl(chlor)silane*. Tetrahedron Letters, 1984. **25**(21): p. 2195-2198.
73. Pooput, C., W.R. Dolbier, and M. Médebielle, *Nucleophilic Perfluoroalkylation of Aldehydes, Ketones, Imines, Disulfides, and Diselenides*. The Journal of Organic Chemistry, 2006. **71**(9): p. 3564-3568.
74. Pooput, C., M. Medebielle, and W.R. Dolbier, *A New and Efficient Method for the Synthesis of Trifluoromethylthio- and Selenoethers*. Organic Letters, 2003. **6**(2): p. 301-303.
75. Eisenberger, P., S. Gischig, and A. Togni, *Novel 10-I-3 Hypervalent Iodine-Based Compounds for Electrophilic Trifluoromethylation*. Chemistry – A European Journal, 2006. **12**(9): p. 2579-2586.
76. Kieltsch, I., P. Eisenberger, and A. Togni, *Mild Electrophilic Trifluoromethylation of Carbon- and Sulfur-Centered Nucleophiles by a Hypervalent Iodine(III)- CF_3 Reagent*. Angewandte Chemie International Edition, 2007. **46**(5): p. 754-757.
77. Capone, S., et al., *Electrophilic S-Trifluoromethylation of Cysteine Side Chains in α - and β -Peptides: Isolation of Trifluoro-methylated Sandostatin® (Octreotide) Derivatives*. Helvetica Chimica Acta, 2008. **91**(11): p. 2035-2056.
78. Pervushin, K., et al., *Attenuated T(2) relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution*. Proc Natl Acad Sci U S A, 1997. **94**(23): p. 12366-12371.

79. Xu, Y. and S. Matthews, *TROSY NMR spectroscopy of large soluble proteins*. Top Curr Chem, 2013. **335**: p. 97-119.
80. Tugarinov, V., et al., *Cross-Correlated Relaxation Enhanced 1H–13C NMR Spectroscopy of Methyl Groups in Very High Molecular Weight Proteins and Protein Complexes*. Journal of the American Chemical Society, 2003. **125**(34): p. 10420-10428.
81. Ollerenshaw, J.E., V. Tugarinov, and L.E. Kay, *Methyl TROSY: explanation and experimental verification*. Magnetic Resonance in Chemistry, 2003. **41**(10): p. 843-852.
82. Eisenmesser, E.Z., et al., *Enzyme Dynamics During Catalysis*. Science, 2002. **295**(5559): p. 1520-1523.
83. Masterson, L.R., et al., *Dynamics connect substrate recognition to catalysis in protein kinase A*. Nat Chem Biol, 2010. **6**(11): p. 821-828.
84. Loria, J.P., R.B. Berlow, and E.D. Watt, *Characterization of Enzyme Motions by Solution NMR Relaxation Dispersion*. Accounts of Chemical Research, 2008. **41**(2): p. 214-221.
85. Masterson, L.R., et al., *Dynamically committed, uncommitted, and quenched states encoded in protein kinase A revealed by NMR spectroscopy*. Proceedings of the National Academy of Sciences, 2011. **108**(17): p. 6969-6974.
86. Andrew Atkinson, R. and B. Kieffer, *The role of protein motions in molecular recognition: insights from heteronuclear NMR relaxation measurements*. Progress in Nuclear Magnetic Resonance Spectroscopy, 2004. **44**(3-4): p. 141-187.
87. Lu, J., D.P. Cistola, and E. Li, *Two Homologous Rat Cellular Retinol-binding Proteins Differ in Local Conformational Flexibility*. J Mol Biol, 2003. **330**(4): p. 799-812.
88. Kristensen, S.M., et al., *Backbone dynamics of the C-terminal SH2 domain of the p85 α subunit of phosphoinositide 3-kinase: effect of phosphotyrosine-peptide binding and characterization of slow conformational exchange processes*. J Mol Biol, 2000. **299**(3): p. 771-788.
89. Carr, H.Y. and E.M. Purcell, *Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments*. Physical Review, 1954. **94**(3): p. 630-638.
90. Meiboom, S. and D. Gill, *Modified Spin-Echo Method for Measuring Nuclear Relaxation Times*. Review of Scientific Instruments, 1958. **29**(8): p. 688-691.
91. Henzler-Wildman, K. and D. Kern, *Dynamic personalities of proteins*. Nature, 2007. **450**(7172): p. 964-72.
92. Ishima, R., J.M. Louis, and D.A. Torchia, *Transverse 13C Relaxation of CHD2 Methyl Isotopomers To Detect Slow Conformational Changes of Protein Side Chains*. Journal of the American Chemical Society, 1999. **121**(49): p. 11589-11590.
93. Skrynnikov, N.R., et al., *Probing Slow Time Scale Dynamics at Methyl-Containing Side Chains in Proteins by Relaxation Dispersion NMR Measurements: Application to Methionine Residues in a Cavity Mutant of T4 Lysozyme*. Journal of the American Chemical Society, 2001. **123**(19): p. 4556-4566.
94. Millet, O., et al., *The Static Magnetic Field Dependence of Chemical Exchange Linebroadening Defines the NMR Chemical Shift Time Scale*. Journal of the American Chemical Society, 2000. **122**(12): p. 2867-2877.
95. Nájera, C. and J.M. Sansano, *Catalytic Asymmetric Synthesis of α -Amino Acids*. Chemical Reviews, 2007. **107**(11): p. 4584-4671.
96. Siebum, Arjan H.G., et al., *Access to Any Site-Directed Isotopomer of Methionine, Selenomethionine, Cysteine, and Selenocysteine – Use of Simple, Efficient Modular Synthetic Reaction Schemes for Isotope Incorporation*. European Journal of Organic Chemistry, 2004. **2004**(13): p. 2905-2913.
97. O'Donnell, M.J., W.D. Bennett, and S. Wu, *The stereoselective synthesis of α -amino acids by phase-transfer catalysis*. Journal of the American Chemical Society, 1989. **111**(6): p. 2353-2355.

98. Córdova, A., et al., *A Highly Enantioselective Route to Either Enantiomer of Both α - and β -Amino Acid Derivatives*. *Journal of the American Chemical Society*, 2002. **124**(9): p. 1866-1867.
99. Banphavichit, V., et al., *A highly enantioselective Strecker reaction catalyzed by titanium-*N*-salicyl- β -aminoalcohol complexes*. *Tetrahedron*, 2004. **60**(46): p. 10559-10568.
100. Harding, M., et al., *Substrate Control in the Asymmetric Aminohydroxylation of -Monosubstituted Alkenes: The Enantioselective Synthesis of GABOB and -Homoserine Derivatives*. *Synlett*, 2005. **2005**(18): p. 2829-2831.
101. Tugarinov, V. and L. Kay, *An Isotope Labeling Strategy for Methyl TROSY Spectroscopy*. *J Biomol NMR*, 2004. **28**(2): p. 165-172.
102. Korzhnev, D.M., et al., *An NMR Experiment for the Accurate Measurement of Heteronuclear Spin-Lock Relaxation Rates*. *Journal of the American Chemical Society*, 2002. **124**(36): p. 10743-10753.
103. Palmer Iii, A.G., C.D. Kroenke, and J. Patrick Loria, [10] - *Nuclear Magnetic Resonance Methods for Quantifying Microsecond-to-Millisecond Motions in Biological Macromolecules*, in *Methods Enzymol*, V.D. Thomas L. James and S. Uli, Editors. 2001, Academic Press. p. 204-238.
104. Akke, M. and A.G. Palmer, *Monitoring Macromolecular Motions on Microsecond to Millisecond Time Scales by R1 ρ -R1 Constant Relaxation Time NMR Spectroscopy*. *Journal of the American Chemical Society*, 1996. **118**(4): p. 911-912.
105. Polgar, L., *Prolyl oligopeptidases*. *Methods Enzymol*, 1994. **244**: p. 188-200.
106. Svergun, D., C. Barberato, and M.H.J. Koch, *CRY SOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates*. *Journal of Applied Crystallography*, 1995. **28**(6): p. 768-773.
107. Bernadó, P., et al., *Structural Characterization of Flexible Proteins Using Small-Angle X-ray Scattering*. *Journal of the American Chemical Society*, 2007. **129**(17): p. 5656-5664.
108. Bernadó, P., *Effect of interdomain dynamics on the structure determination of modular proteins by small-angle scattering*. *Eur Biophys J*, 2010. **39**(5): p. 769-80.
109. Vestergaard, B., et al., *A Helical Structural Nucleus Is the Primary Elongating Unit of Insulin Amyloid Fibrils*. *PLoS Biology*, 2007. **5**(5): p. e134.
110. Bernadó, P., et al., *Structural characterization of unphosphorylated STAT5a oligomerization equilibrium in solution by small-angle X-ray scattering*. *Protein Sci*, 2009. **18**(4): p. 716-26.
111. Watanabe, Y. and Y. Inoko, *Further application of size-exclusion chromatography combined with small-angle X-ray scattering optics for characterization of biological macromolecules*. *Analytical and Bioanalytical Chemistry*, 2011. **399**(4): p. 1449-1453.
112. Watanabe, Y. and Y. Inoko, *Size-exclusion chromatography combined with small-angle X-ray scattering optics*. *Journal of chromatography. A*, 2009. **1216**(44): p. 7461-5.
113. Skou, S., R.E. Gillilan, and N. Ando, *Synchrotron-based small-angle X-ray scattering of proteins in solution*. *Nat. Protocols*, 2014. **9**(7): p. 1727-1739.
114. Konarev, P.V., et al., *PRIMUS: a Windows PC-based system for small-angle scattering data analysis*. *Journal of Applied Crystallography*, 2003. **36**(5): p. 1277-1282.
115. Tauler, R., *Multivariate curve resolution applied to second order data*. *Chemometrics and Intelligent Laboratory Systems*, 1995. **30**(1): p. 133-146.
116. Blobel, J., et al., *Low-Resolution Structures of Transient Protein-Protein Complexes Using Small-Angle X-ray Scattering*. *Journal of the American Chemical Society*, 2009. **131**(12): p. 4378-4386.

117. Wen, J., T. Arakawa, and J.S. Philo, *Size-Exclusion Chromatography with On-Line Light-Scattering, Absorbance, and Refractive Index Detectors for Studying Proteins and Their Interactions*. Analytical Biochemistry, 1996. **240**(2): p. 155-166.
118. Svergun, D., *Determination of the regularization parameter in indirect-transform methods using perceptual criteria*. Journal of Applied Crystallography, 1992. **25**(4): p. 495-503.
119. Leach, A.R., *Molecular Modelling. Principles and Applications*. Pearson Prentice Hall, 2001.
120. Berendsen, H.J.C., et al., *Molecular dynamics with coupling to an external bath*. The Journal of Chemical Physics, 1984. **81**(8): p. 3684-3690.
121. Ryckaert, J.-P., G. Ciccotti, and H.J.C. Berendsen, *Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes*. Journal of Computational Physics, 1977. **23**(3): p. 327-341.
122. Shrake, A. and J.A. Rupley, *Environment and exposure to solvent of protein atoms. Lysozyme and insulin*. J Mol Biol, 1973. **79**(2): p. 351-371.
123. Weiser, J., P.S. Shenkin, and W.C. Still, *Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO)*. Journal of Computational Chemistry, 1999. **20**(2): p. 217-230.
124. Roy, A., A. Kucukural, and Y. Zhang, *I-TASSER: a unified platform for automated protein structure and function prediction*. Nature protocols, 2010. **5**(4): p. 725-738.
125. D. E. Shaw Research, N.Y., *Desmond Molecular Dynamics System, version 3.1*. 2012.
126. Humphrey, W., A. Dalke, and K. Schulten, *VMD: visual molecular dynamics*. J Mol Graph, 1996. **14**(1): p. 33-8, 27-8.
127. Szeltner, Z., et al., *Substrate-dependent competency of the catalytic triad of prolyl oligopeptidase*. J Biol Chem, 2002. **277**(47): p. 44597-605.
128. Benesch, J.L. and B.T. Ruotolo, *Mass spectrometry: come of age for structural and dynamical biology*. Current Opinion in Structural Biology, 2011. **21**(5): p. 641-9.
129. Fenn, J.B., et al., *Electrospray ionization for mass spectrometry of large biomolecules*. Science (New York, N.Y.), 1989. **246**(4926): p. 64-71.
130. Wilm, M.S. and M. Mann, *Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last?* International Journal of Mass Spectrometry and Ion Processes, 1994. **136**(2-3): p. 167-180.
131. Ruotolo, B.T. and C.V. Robinson, *Aspects of native proteins are retained in vacuum*. Current Opinion in Chemical Biology, 2006. **10**(5): p. 402-408.
132. Heck, A.J., *Native mass spectrometry: a bridge between interactomics and structural biology*. Nature methods, 2008. **5**(11): p. 927-33.
133. Eyles, S.J. and I.A. Kaltashov, *Methods to study protein dynamics and folding by mass spectrometry*. Methods (San Diego, Calif.), 2004. **34**(1): p. 88-99.
134. Rostom, A.A., et al., *Detection and selective dissociation of intact ribosomes in a mass spectrometer*. Proceedings of the National Academy of Sciences, 2000. **97**(10): p. 5185-5190.
135. Uetrecht, C., et al., *High-resolution mass spectrometry of viral assemblies: Molecular composition and stability of dimorphic hepatitis B virus capsids*. Proceedings of the National Academy of Sciences, 2008. **105**(27): p. 9216-9220.
136. Uetrecht, C., et al., *Ion mobility mass spectrometry of proteins and protein assemblies*. Chemical Society Reviews, 2010. **39**(5): p. 1633-1655.
137. Kebarle, P. and M. Peschke, *On the mechanisms by which the charged droplets produced by electrospray lead to gas phase ions*. Analytica Chimica Acta, 2000. **406**(1): p. 11-35.
138. De La Mora, J.F., et al., *The current emitted by highly conducting Taylor cones*. Journal of Fluid Mechanics, 1994. **260**: p. 155-184.

139. Shelimov, K.B., et al., *Protein Structure in Vacuo: Gas-Phase Conformations of BPTI and Cytochrome c*. Journal of the American Chemical Society, 1997. **119**(9): p. 2240-2248.
140. Daniel, J.M., et al., *Quantitative determination of noncovalent binding interactions using soft ionization mass spectrometry*. International Journal of Mass Spectrometry, 2002. **216**(1): p. 1-27.
141. Breuker, K. and F.W. McLafferty, *Stepwise evolution of protein native structure with electrospray into the gas phase, 10–12 to 102 s*. Proceedings of the National Academy of Sciences, 2008. **105**(47): p. 18145-18152.
142. Steinberg, M.Z., et al., *The dynamics of water evaporation from partially solvated cytochrome c in the gas phase*. Physical chemistry chemical physics : PCCP, 2007. **9**(33): p. 4690-7.
143. Steinberg, M.Z., et al., *Early Structural Evolution of Native Cytochrome c after Solvent Removal*. Chembiochem, 2008. **9**(15): p. 2417-2423.
144. Wyttenbach, T. and M.T. Bowers, *Structural Stability from Solution to the Gas Phase: Native Solution Structure of Ubiquitin Survives Analysis in a Solvent-Free Ion Mobility–Mass Spectrometry Environment*. The Journal of Physical Chemistry B, 2011. **115**(42): p. 12266-12275.
145. Hoaglund-Hyzer, C.S., A.E. Counterman, and D.E. Clemmer, *Anhydrous protein ions*. Chem Rev, 1999. **99**(10): p. 3037-80.
146. Shelimov, K.B. and M.F. Jarrold, *Conformations, Unfolding, and Refolding of Apomyoglobin in Vacuum: An Activation Barrier for Gas-Phase Protein Folding*. Journal of the American Chemical Society, 1997. **119**(13): p. 2987-2994.
147. Ruotolo, B.T., et al., *Ion mobility-mass spectrometry reveals long-lived, unfolded intermediates in the dissociation of protein complexes*. Angewandte Chemie (International ed. in English), 2007. **46**(42): p. 8001-4.
148. Badman, E.R., C.S. Hoaglund-Hyzer, and D.E. Clemmer, *Monitoring Structural Changes of Proteins in an Ion Trap over ~10–200 ms: Unfolding Transitions in Cytochrome c Ions*. Analytical Chemistry, 2001. **73**(24): p. 6000-6007.
149. van Duijn, E., et al., *Chaperonin complexes monitored by ion mobility mass spectrometry*. Journal of the American Chemical Society, 2009. **131**(4): p. 1452-9.
150. Langevin, P., *L'ionization des gaz*. Ann. Chim. Phys., 1903. **28**: p. 289-384.
151. McDaniel, E.W., D.W. Martin, and W.S. Barnes, *Drift tube-mass spectrometer for studies of low-energy ion-molecule reactions*. Review of Scientific Instruments, 1962. **33**(1): p. 2-7.
152. Pringle, S.D., et al., *An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument*. International Journal of Mass Spectrometry, 2007. **261**(1): p. 1-12.
153. Wallace, A., *A high-resolution ion mobility mass spectrometry platform for breakthrough discoveries in life science research and the pharmaceutical industry*. American Laboratory, 2010. **42**: p. 13-17.
154. Zhong, Y., S.-J. Hyung, and B.T. Ruotolo, *Characterizing the resolution and accuracy of a second-generation traveling-wave ion mobility separator for biomolecular ions*. Analyst, 2011. **136**(17): p. 3534-3541.
155. Mason, E.A. and E.W. McDaniel, *Transport properties of ions in gases*. Wiley, New York., 1988.
156. Shvartsburg, A.A. and R.D. Smith, *Fundamentals of Traveling Wave Ion Mobility Spectrometry*. Analytical Chemistry, 2008. **80**(24): p. 9689-9699.
157. Bush, M.F., et al., *Collision Cross Sections of Proteins and Their Complexes: A Calibration Framework and Database for Gas-Phase Structural Biology*. Analytical Chemistry, 2010. **82**(22): p. 9557-9565.

158. Smith, D.P., et al., *Deciphering drift time measurements from travelling wave ion mobility spectrometry-mass spectrometry studies*. European journal of mass spectrometry (Chichester, England), 2009. **15**(2): p. 113-30.
159. Hoaglund, C.S., et al., *Three-Dimensional Ion Mobility/TOFMS Analysis of Electrosprayed Biomolecules*. Analytical Chemistry, 1998. **70**(11): p. 2236-2242.
160. Ruotolo, B.T., et al., *Ion mobility-mass spectrometry analysis of large protein complexes*. Nat. Protocols, 2008. **3**(7): p. 1139-1152.
161. Han, L., S.-J. Hyung, and B.T. Ruotolo, *Bound Cations Significantly Stabilize the Structure of Multiprotein Complexes in the Gas-phase*. Angewandte Chemie (International ed. in English), 2012. **51**(23): p. 5692-5695.
162. Freeke, J., C.V. Robinson, and B.T. Ruotolo, *Residual counter ions can stabilise a large protein complex in the gas phase*. International Journal of Mass Spectrometry, 2010. **298**(1-3): p. 91-98.
163. Freeke, J., et al., *Gas-phase protein assemblies: Unfolding landscapes and preserving native-like structures using noncovalent adducts*. Chemical Physics Letters, 2012. **524**(0): p. 1-9.
164. Uetrecht, C., et al., *Stability and Shape of Hepatitis B Virus Capsids In Vacuo*. Angewandte Chemie (International ed. in English), 2008. **47**(33): p. 6247-6251.
165. Uetrecht, C., et al., *Interrogating viral capsid assembly with ion mobility-mass spectrometry*. Nat Chem, 2011. **3**(2): p. 126-132.
166. Barrera, N.P., et al., *Mass spectrometry of membrane transporters reveals subunit stoichiometry and interactions*. Nature methods, 2009. **6**(8): p. 585-7.
167. Zhou, M., et al., *Mass spectrometry of intact V-type ATPases reveals bound lipids and the effects of nucleotide binding*. Science (New York, N.Y.), 2011. **334**(6054): p. 380-385.
168. Politis, A., et al., *Integrating Ion Mobility Mass Spectrometry with Molecular Modelling to Determine the Architecture of Multiprotein Complexes*. PLoS One, 2010. **5**(8): p. e12080.
169. Smith, D.P., S.E. Radford, and A.E. Ashcroft, *Elongated oligomers in β (2)-microglobulin amyloid assembly revealed by ion mobility spectrometry-mass spectrometry*. Proc Natl Acad Sci U S A, 2010. **107**(15): p. 6794-6798.
170. Woods, L.A., et al., *Ligand binding to distinct states diverts aggregation of an amyloid-forming protein*. Nature chemical biology, 2011. **7**(10): p. 730-9.
171. Pagel, K., et al., *Intrinsically disordered p53 and its complexes populate compact conformations in the gas phase*. Angewandte Chemie (International ed. in English), 2013. **52**(1): p. 361-5.
172. Shvartsburg, A.A. and M.F. Jarrold, *An exact hard-spheres scattering model for the mobilities of polyatomic ions*. Chemical Physics Letters, 1996. **261**(1-2): p. 86-91.
173. Mesleh, M.F., et al., *Structural Information from Ion Mobility Measurements: Effects of the Long-Range Potential*. The Journal of Physical Chemistry, 1996. **100**(40): p. 16082-16086.
174. Jurneczko, E. and P.E. Barran, *How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase*. Analyst, 2011. **136**(1): p. 20-8.
175. Koeniger, S.L. and D.E. Clemmer, *Resolution and structural transitions of elongated states of ubiquitin*. Journal of the American Society for Mass Spectrometry, 2007. **18**(2): p. 322-31.
176. Hopper, J.S. and N. Oldham, *Collision induced unfolding of protein ions in the gas phase studied by ion mobility-mass spectrometry: The effect of ligand binding on conformational stability*. Journal of the American Society for Mass Spectrometry, 2009. **20**(10): p. 1851-1858.

177. Hernandez, H. and C.V. Robinson, *Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry*. Nature protocols, 2007. **2**(3): p. 715-26.
178. Jurneczko, E., et al., *Effects of Drift Gas on Collision Cross Sections of a Protein Standard in Linear Drift Tube and Traveling Wave Ion Mobility Mass Spectrometry*. Analytical Chemistry, 2012. **84**(20): p. 8524-8531.
179. Michaelevski, I., M. Eisenstein, and M. Sharon, *Gas-Phase Compaction and Unfolding of Protein Structures*. Analytical Chemistry, 2010. **82**(22): p. 9484-9491.
180. Sterling, H.J., et al., *Effects of supercharging reagents on noncovalent complex structure in electrospray ionization from aqueous solutions*. Journal of the American Society for Mass Spectrometry, 2010. **21**(10): p. 1762-74.
181. Sterling, H.J. and E.R. Williams, *Origin of supercharging in electrospray ionization of noncovalent complexes from aqueous solution*. Journal of the American Society for Mass Spectrometry, 2009. **20**(10): p. 1933-43.
182. Lomeli, S.H., et al., *New Reagents for Increasing ESI Multiple Charging of Proteins and Protein Complexes*. Journal of the American Society for Mass Spectrometry, 2010. **21**(1): p. 127-131.
183. Catalina, M.I., et al., *Decharging of globular proteins and protein complexes in electrospray*. Chemistry, 2005. **11**(3): p. 960-8.
184. Sugita, Y. and Y. Okamoto, *Replica-exchange molecular dynamics method for protein folding*. Chemical Physics Letters, 1999. **314**(1-2): p. 141-151.
185. Li, H., A.D. Robertson, and J.H. Jensen, *Very fast empirical prediction and rationalization of protein pKa values*. Proteins, 2005. **61**(4): p. 704-21.
186. Olsson, M.H.M., et al., *PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions*. Journal of Chemical Theory and Computation, 2011. **7**(2): p. 525-537.
187. Morsa, D., V. Gabelica, and E. De Pauw, *Effective Temperature of Ions in Traveling Wave Ion Mobility Spectrometry*. Analytical Chemistry, 2011. **83**(14): p. 5775-5782.
188. Jarrold, M.F., *Unfolding, Refolding, and Hydration of Proteins in the Gas Phase*. Accounts of Chemical Research, 1999. **32**(4): p. 360-367.
189. Bornschein, R.E., S.J. Hyung, and B.T. Ruotolo, *Ion mobility-mass spectrometry reveals conformational changes in charge reduced multiprotein complexes*. Journal of the American Society for Mass Spectrometry, 2011. **22**(10): p. 1690-8.
190. Clemmer, D.E. and M.F. Jarrold, *Ion Mobility Measurements and their Applications to Clusters and Biomolecules*. Journal of Mass Spectrometry, 1997. **32**(6): p. 577-592.
191. Faull, P.A., et al., *Gas-phase metalloprotein complexes interrogated by ion mobility-mass spectrometry*. International Journal of Mass Spectrometry, 2009. **283**(1-3): p. 140-148.
192. Lanucara, F., et al., *The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics*. Nat Chem, 2014. **6**(4): p. 281-294.
193. McLean, J.A. and B.T. Ruotolo, *24th Sanibel Conference on mass spectrometry: Mass spectrometry technologies for structural biology*. Journal of the American Society for Mass Spectrometry, 2012. **23**(12): p. 2208-12.
194. Bich, C., et al., *Probing the Hydrophobic Effect of Noncovalent Complexes by Mass Spectrometry*. Journal of the American Society for Mass Spectrometry, 2010. **21**(2): p. 286-289.
195. Hilton, G.R. and J.L. Benesch, *Two decades of studying non-covalent biomolecular assemblies by means of electrospray ionization mass spectrometry*. J R Soc Interface, 2012. **9**(70): p. 801-16.

196. Sharon, M. and C.V. Robinson, *A quantitative perspective on hydrophobic interactions in the gas-phase*. Current proteomics, 2011. **8**(1): p. 47-58.
197. Surguchov, A., *Molecular and cellular biology of synucleins*. Int Rev Cell Mol Biol, 2008. **270**: p. 225-317.
198. Uéda, K., et al., *Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease*. Proc Natl Acad Sci U S A, 1993. **90**(23): p. 11282-11286.
199. Baba, M., et al., *Aggregation of alpha-synuclein in Lewy bodies of sporadic Parkinson's disease and dementia with Lewy bodies*. The American journal of pathology, 1998. **152**(4): p. 879-84.
200. Conway, K.A., J.D. Harper, and P.T. Lansbury, Jr., *Fibrils formed in vitro from alpha-synuclein and two mutant forms linked to Parkinson's disease are typical amyloid*. Biochemistry, 2000. **39**(10): p. 2552-63.
201. Cremades, N., et al., *Direct Observation of the Interconversion of Normal and Toxic Forms of α -Synuclein*. Cell. **149**(5): p. 1048-1059.
202. Conway, K.A., et al., *Acceleration of oligomerization, not fibrillization, is a shared property of both α -synuclein mutations linked to early-onset Parkinson's disease: Implications for pathogenesis and therapy*. Proc Natl Acad Sci U S A, 2000. **97**(2): p. 571-576.
203. Danzer, K.M., et al., *Different species of alpha-synuclein oligomers induce calcium influx and seeding*. J Neurosci, 2007. **27**(34): p. 9220-32.
204. Opazo, F., et al., *Accumulation and clearance of alpha-synuclein aggregates demonstrated by time-lapse imaging*. J Neurochem, 2008. **106**(2): p. 529-40.
205. Kruger, R., et al., *Ala30Pro mutation in the gene encoding alpha-synuclein in Parkinson's disease*. Nature genetics, 1998. **18**(2): p. 106-8.
206. Polymeropoulos, M.H., et al., *Mutation in the alpha-synuclein gene identified in families with Parkinson's disease*. Science (New York, N.Y.), 1997. **276**(5321): p. 2045-7.
207. Zarranz, J.J., et al., *The new mutation, E46K, of alpha-synuclein causes Parkinson and Lewy body dementia*. Ann Neurol, 2004. **55**(2): p. 164-73.
208. Wood, S.J., et al., *alpha-synuclein fibrillogenesis is nucleation-dependent. Implications for the pathogenesis of Parkinson's disease*. J Biol Chem, 1999. **274**(28): p. 19509-12.
209. Uversky, V.N., *Neuropathology, biochemistry, and biophysics of alpha-synuclein aggregation*. J Neurochem, 2007. **103**(1): p. 17-37.
210. Brandt, I., et al., *Prolyl oligopeptidase stimulates the aggregation of α -synuclein*. Peptides, 2008. **29**(9): p. 1472-1478.
211. Hannula, M.J., et al., *Prolyl oligopeptidase colocalizes with alpha-synuclein, beta-amyloid, tau protein and astroglia in the post-mortem brain samples with Parkinson's and Alzheimer's diseases*. Neuroscience, 2013. **242**: p. 140-50.
212. Savolainen, M.H., et al., *The beneficial effect of a prolyl oligopeptidase inhibitor, KYP-2047, on alpha-synuclein clearance and autophagy in A30P transgenic mouse*. Neurobiol Dis, 2014. **68**: p. 1-15.
213. Ebrahimi-Fakhari, D., L. Wahlster, and P.J. McLean, *Protein degradation pathways in Parkinson's disease: curse or blessing*. Acta Neuropathol, 2012. **124**(2): p. 153-72.
214. Dokleja, L., M.J. Hannula, and T.T. Myohanen, *Inhibition of prolyl oligopeptidase increases the survival of alpha-synuclein overexpressing cells after rotenone exposure by reducing alpha-synuclein oligomers*. Neurosci Lett, 2014. **583C**: p. 37-42.
215. Savolainen, M.H., et al., *Prolyl Oligopeptidase Enhances alpha-synuclein Dimerization via Direct Protein-Protein Interaction*. J Biol Chem, 2015.

216. Prakash, M.K., *Insights on the Role of (Dis)order from Protein-Protein Interaction Linear Free-Energy Relationships*. Journal of the American Chemical Society, 2011. **133**(26): p. 9976-9979.
217. Fong, J.H., et al., *Intrinsic Disorder in Protein Interactions: Insights From a Comprehensive Structural Analysis*. PLoS Comput Biol, 2009. **5**(3): p. e1000316.
218. Mittag, T., L.E. Kay, and J.D. Forman-Kay, *Protein dynamics and conformational disorder in molecular recognition*. Journal of molecular recognition : JMR, 2010. **23**(2): p. 105-16.
219. Tompa, P. and M. Fuxreiter, *Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions*. Trends in biochemical sciences, 2008. **33**(1): p. 2-8.
220. De Franceschi, G., et al., *Structural and morphological characterization of aggregated species of alpha-synuclein induced by docosahexaenoic acid*. J Biol Chem, 2011. **286**(25): p. 22262-74.
221. Vinogradova, O. and J. Qin, *NMR as a unique tool in assessment and complex determination of weak protein-protein interactions*. Top Curr Chem, 2012. **326**: p. 35-45.
222. Lamberto, G.R., et al., *Structural and mechanistic basis behind the inhibitory interaction of PcTS on alpha-synuclein amyloid fibril formation*. Proc Natl Acad Sci U S A, 2009. **106**(50): p. 21057-62.
223. Binolfi, A., et al., *Site-specific interactions of Cu(II) with alpha and beta-synuclein: bridging the molecular gap between metal binding and aggregation*. J Am Chem Soc, 2008. **130**(35): p. 11801-12.
224. Binolfi, A., et al., *Interaction of α -Synuclein with Divalent Metal Ions Reveals Key Differences: A Link between Structure, Binding Specificity and Fibrillation Enhancement*. Journal of the American Chemical Society, 2006. **128**(30): p. 9893-9901.
225. Farrow, N.A., et al., *Backbone Dynamics of a Free and a Phosphopeptide-Complexed Src Homology 2 Domain Studied by ^{15}N NMR Relaxation*. Biochemistry, 1994. **33**(19): p. 5984-6003.
226. Kay, L.E., et al., *Pulse sequences for removal of the effects of cross correlation between dipolar and chemical-shift anisotropy relaxation mechanisms on the measurement of heteronuclear T_1 and T_2 values in proteins*. Journal of Magnetic Resonance (1969), 1992. **97**(2): p. 359-375.
227. Peng, J.W. and G. Wagner, *Mapping of the spectral densities of N-H bond motions in eglin c using heteronuclear relaxation experiments*. Biochemistry, 1992. **31**(36): p. 8571-86.
228. Farrow, N., et al., *Spectral density function mapping using ^{15}N relaxation data exclusively*. J Biomol NMR, 1995. **6**(2): p. 153-162.
229. Grzesiek, S. and A. Bax, *The Importance of Not Saturating H_2O in Protein NMR. Application to Sensitivity Enhancement and NOE Measurements*. Journal of the American Chemical Society, 1993. **115**(26): p. 12593-12594.
230. Bussell, R. and D. Eliezer, *Residual Structure and Dynamics in Parkinson's Disease-associated Mutants of α -Synuclein*. Journal of Biological Chemistry, 2001. **276**(49): p. 45996-46003.
231. Wu, K.P., et al., *Characterization of conformational and dynamic properties of natively unfolded human and mouse alpha-synuclein ensembles by NMR: implication for aggregation*. J Mol Biol, 2008. **378**(5): p. 1104-15.
232. Bussell, R., T.F. Ramlall, and D. Eliezer, *Helix periodicity, topology, and dynamics of membrane-associated α -Synuclein*. Protein Sci, 2005. **14**(4): p. 862-872.
233. Takahashi, H., et al., *A novel NMR method for determining the interfaces of large protein-protein complexes*. Nat Struct Mol Biol, 2000. **7**(3): p. 220-223.

234. Ramos, A., et al., *Mapping the Interfaces of Protein–Nucleic Acid Complexes Using Cross-Saturation*. Journal of the American Chemical Society, 2000. **122**(46): p. 11311-11314.
235. Hoyer, W., et al., *Dependence of alpha-synuclein aggregate morphology on solution conditions*. J Mol Biol, 2002. **322**(2): p. 383-93.
236. Cole, N.B., et al., *Metal-catalyzed Oxidation of α -Synuclein: HELPING TO DEFINE THE RELATIONSHIP BETWEEN OLIGOMERS, PROTOFIBRILS, AND FILAMENTS*. Journal of Biological Chemistry, 2005. **280**(10): p. 9678-9690.
237. Reynolds, N.P., et al., *Mechanism of Membrane Interaction and Disruption by α -Synuclein*. Journal of the American Chemical Society, 2011. **133**(48): p. 19366-19375.
238. Lam, X.M., J.Y. Yang, and J.L. Cleland, *Antioxidants for prevention of methionine oxidation in recombinant monoclonal antibody HER2*. Journal of Pharmaceutical Sciences, 1997. **86**(11): p. 1250-1255.
239. Rao, J.N., et al., *Effect of Pseudorepeat Rearrangement upon α -Synuclein Misfolding, Vesicle Binding and Micelle Binding*. J Mol Biol, 2009. **390**(3): p. 516-529.
240. Gallagher, S.R., *One-dimensional SDS gel electrophoresis of proteins*. Curr Protoc Mol Biol, 2006. **Chapter 10**: p. Unit 10 2A.
241. Burre, J., et al., *Properties of native brain alpha-synuclein*. Nature, 2013. **498**(7453): p. E4-6; discussion E6-7.
242. Mbefo, M.K., et al., *Phosphorylation of Synucleins by Members of the Polo-like Kinase Family*. Journal of Biological Chemistry, 2010. **285**(4): p. 2807-2822.
243. Oueslati, A., M. Fournier, and H.A. Lashuel, *Role of post-translational modifications in modulating the structure, function and toxicity of alpha-synuclein: implications for Parkinson's disease pathogenesis and therapies*. Prog Brain Res, 2010. **183**: p. 115-45.
244. Lu, Y., et al., *Phosphorylation of alpha-Synuclein at Y125 and S129 alters its metal binding properties: implications for understanding the role of alpha-Synuclein in the pathogenesis of Parkinson's Disease and related disorders*. ACS Chem Neurosci, 2011. **2**(11): p. 667-75.
245. Gruebele, M., *Protein Dynamics in Simulation and Experiment*. Journal of the American Chemical Society, 2014. **136**(48): p. 16695-16697.
246. Boehr, D.D., R. Nussinov, and P.E. Wright, *The role of dynamic conformational ensembles in biomolecular recognition*. Nature chemical biology, 2009. **5**(11): p. 789-96.
247. Seo, M.-H., et al., *Protein conformational dynamics dictate the binding affinity for a ligand*. Nat Commun, 2014. **5**.
248. Bernado, P. and M. Blackledge, *Structural biology: Proteins in dynamic equilibrium*. Nature, 2010. **468**(7327): p. 1046-8.
249. Bahr, B.A., *A single pathway targets several health challenges of the elderly*. Rejuvenation Res, 2014. **17**(4): p. 382-4.
250. Nishizawa, M., et al., *Virtually Complete E-Selective α,β -Unsaturated Ester Synthesis by Hg(OTf)₂-Catalyzed Hydration of sec-Ethoxyalkynyl Acetate*. Organic Letters, 2007. **9**(26): p. 5577-5580.
251. Engel, D.A., S.S. Lopez, and G.B. Dudley, *Lewis acid-catalyzed Meyer–Schuster reactions: methodology for the olefination of aldehydes and ketones*. Tetrahedron, 2008. **64**(29): p. 6988-6996.
252. Snyder, S.A., D.S. Treitler, and A.P. Brucks, *Simple Reagents for Direct Halonium-Induced Polyene Cyclizations*. Journal of the American Chemical Society, 2010. **132**(40): p. 14303-14314.
253. Lygo, B. and L.D. Humphreys, *Enantioselective synthesis of α -carbon deuterium-labelled l- α -amino acids*. Tetrahedron Letters, 2002. **43**(37): p. 6677-6679.

254. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes*. J Biomol NMR, 1995. **6**(3): p. 277-93.
255. Case, D.A., et al., *AMBER 12*, University of California, San Francisco. 2012.
256. Hornak, V., et al., *Comparison of multiple Amber force fields and development of improved protein backbone parameters*. Proteins, 2006. **65**(3): p. 712-25.
257. Jorgensen, W.L., et al., *Comparison of simple potential functions for simulating liquid water*. The Journal of Chemical Physics, 1983. **79**(2): p. 926-935.
258. Miyamoto, S. and P.A. Kollman, *Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models*. Journal of Computational Chemistry, 1992. **13**(8): p. 952-962.
259. Darden, T., D. York, and L. Pedersen, *Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems*. The Journal of Chemical Physics, 1993. **98**(12): p. 10089-10092.
260. Bowers, K.J., et al., *Scalable algorithms for molecular dynamics simulations on commodity clusters*, in *Proceedings of the 2006 ACM/IEEE conference on Supercomputing2006*, ACM: Tampa, Florida. p. 84.
261. Hoover, W.G., *Constant-pressure equations of motion*. Physical Review A, 1986. **34**(3): p. 2499-2500.
262. Martyna, G.J., D.J. Tobias, and M.L. Klein, *Constant pressure molecular dynamics algorithms*. The Journal of Chemical Physics, 1994. **101**(5): p. 4177-4189.
263. Roe, D.R. and T.E. Cheatham, *PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data*. Journal of Chemical Theory and Computation, 2013. **9**(7): p. 3084-3095.
264. Berendsen, H.J.C., D. van der Spoel, and R. van Drunen, *GROMACS: A message-passing parallel molecular dynamics implementation*. Computer Physics Communications, 1995. **91**(1-3): p. 43-56.
265. Hess, B., et al., *GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation*. Journal of Chemical Theory and Computation, 2008. **4**(3): p. 435-447.
266. Hess, B., et al., *LINCS: A linear constraint solver for molecular simulations*. Journal of Computational Chemistry, 1997. **18**(12): p. 1463-1472.
267. Bussi, G., D. Donadio, and M. Parrinello, *Canonical sampling through velocity rescaling*. The Journal of Chemical Physics, 2007. **126**(1): p. 014101.

Conclusions

In view of our initial objectives, and according to the results presented in the previous chapters, this PhD thesis concludes as following:

1. We have successfully applied a combination of robust biophysical tools (NMR and SAXS) together with MD simulations in order to study the conformational dynamics of POP. Our findings demonstrates that **free POP exists in solution in a μ s-ms dynamic equilibrium between open and closed conformations**. This equilibrium provides all the substates involved in catalysis: the open conformation exposes the active site and allows the recruitment of substrates; in turn, the closed conformation presents the proper arrangement of residues of the active site.

In order to overcome the size limitation of NMR experiments, we have benefitted from the labeling of side chain methyl groups of Met residues. In order to obtain the accurate measurements of the exchange parameters of POP, we have developed a labelling strategy in order to produce U-[2 H], [methyl- 13 C, 1 H]-Met-labeled POP. This approach required the chemical synthesis of a highly deuterated [methyl- 13 C]-Met analog. This labelling strategy made possible to perform multiple quantum RD experiments.

The structures in equilibrium were obtained from ensemble-averaged SAXS data using the EOM over a pool of POP structures generated by MD simulations. The EOM was driven by similarity with the experimental SAXS data.

Moreover, the following observations have been also extracted:

- 1.1. MD simulations of the open conformation showed an increased flexibility in the ns time scale. The higher conformational entropy partially compensates the enthalpic penalty caused by the disruption of the stabilizing interdomain interactions.
- 1.2. SAXS data shows an additional highly compact form of free POP. Probably, this secondary POP conformer might be originated from a tight arrangement of β -propeller blades, which result in smaller R_g . Further work is required to elucidate the structure this minor form.
- 1.3. We have found that the native open/close equilibrium of POP is retained in the gas phase. **Our results demonstrate that IMMS is suitable technique for the study of large dynamic structured proteins**. However, **we emphasize that the accurate optimization of the experimental conditions is crucial to minimize the artefacts derived from the gas-phase environment**. Unfortunately, the effects of inhibitors in the conformational equilibrium of POP

were not detected by IMMS, probably as a consequence of the altered balance of hydrophobic and electrostatic forces in the gas phase.

2. NMR and SAXS experiments combined with MD simulations showed that **active site-directed covalent inhibitors completely displace the equilibrium towards the stabilized closed conformation. This causes the total suppression of μ s-ms conformational dynamics of POP.**

- 2.1. In the presence of inhibitors, the residues of the active site become stabilized in a configuration similar to the transition state. This impedes the conformational fluctuations necessary for the completion of the enzymatic cycle.

- 2.2. The alteration of the conformational equilibrium of POP caused by inhibitors probably modifies the recognition events responsible for the biological function. This can be exploited for the design of new POP inhibitors envisaged as conformational shifters.

3. The analysis of ^1H - ^{15}N HSQC experiments of α -synuclein shows that **POP and α -synuclein experience a weak and transient interaction**, affecting few residues of the N- and C-terminal regions of α -synuclein. **This interaction is slightly more intense in the case of POP bound to inhibitors, affecting a broad section of the C-terminal region of α -synuclein.** The differences between the results obtained in the presence of free POP and in the presence of inhibitor-bound POP suggests that **the binding motif of α -synuclein depends on the conformation of POP.**

Resum

INTRODUCCIÓ

La prolil oligopeptidasa (POP) és un enzim d'un pes molecular de 81 KDa constituït per dos dominis: la α/β -hidrolasa i la β -turbina. El primer domini conté la tríada catalítica, formada pels residus Ser 554, Asp 641 i His 680; el segon està format per set hèlix de fulles β disposades de forma circular, creant així una estructura cilíndrica amb un túnel intern. La funció enzimàtica de la POP és la d'hidrolitzar pèptids curts amb contingut en prolina.

Ambdós dominis es troben units en un extrem per dues cadenes polipeptídiques. A més a més, nombroses forces no covalents mantenen units els dos dominis, donant lloc a una conformació de l'enzim tancada. En aquesta conformació tancada el centre actiu de la POP es troba ocult a la zona interdominial. Diverses estructures de raigs X de la POP suggereixen que l'enzim es troba sempre en aquesta conformació tancada. No obstant, recentment es van cristal·litzar dues POPs bacterianes en les quals els dos dominis es trobaven separats a causa d'un moviment de tipus xarnera. Aquests estudis van suggerir que la conformació oberta de la POP dels mamífers també podria existir.

Com s'ha comentat abans, el centre actiu de la POP es troba ocult entre els dos dominis, i és difícilment accessible pel solvent i els substrats. Aquesta qüestió obre el debat sobre quina és la via de circulació dels substrats i els productes durant el cicle catalític de la POP. Per una part, la separació dels dos dominis pot exposar periòdicament el centre actiu, fent-lo accessible durant un cert interval de temps als substrats; no obstant, aquesta separació no s'ha detectat per a la POP dels mamífers. Per l'altra, certs estudis han demostrat que algunes cadenes polipeptídiques flexibles que envolten el centre actiu poden desprendre's periòdicament i així exposar el centre actiu al solvent. No obstant, avui dia cap de les dues hipòtesis ha estat demostrada amb una tècnica biofísica adequada. En aquesta direcció, un estudi dut a terme al nostre grup amb anterioritat demostra que la POP és altament dinàmica, indicant que existeixen fluctuacions estructurals importants en l'escala dels μ s-ms.

El rol biològic de la POP està íntimament lligat amb processos mentals i cognitius. Aquesta peptidasa s'expressa especialment al cervell, i des del seu descobriment s'ha associat l'alteració de la seva activitat enzimàtica amb els trastorns mentals. No obstant, el seu paper i mecanisme exacte romanen dubtosos avui dia. Les primeres hipòtesis suggerien que aquest enzim actuava controlant els nivells extracel·lulars de neuropèptids (molts d'ells substrats *in vitro* de la POP), afectant així als processos cognitius i a la memòria. Però aquesta hipòtesi estava en contradicció amb la localització intracel·lular de l'enzim, i per tant es va descartar.

Una de les teories que més ha progressat és la de la implicació de la POP en la ruta de senyalització de l'inositol trifosfat. Aquest missatger secundari està implicat en l'alliberament de calci intracel·lular, que al seu torn desencadena una cascada de fosforilacions que repercuteix en els processos cognitius. Certs estudis van demostrar que la presència de POP causava un descens en els nivells d'inositol trifosfat, i que aquest efecte podia ser revertit gràcies a l'administració d'inhibidors de la POP. Ara bé, el mecanisme exacte d'aquesta implicació no està clar actualment.

Recentment, s'argumenta que la POP realitza la seva activitat biològica a través de interaccions proteïna-proteïna (IPP). Especialment, s'ha trobat que la POP pot prevenir l'agregació de l' α -sinucleïna. L' α -sinucleïna és una proteïna intrínsecament desestructurada de 14 KDa, que compon els cossos de Lewi presents als cervells dels pacients de la malaltia de Parkinson. Es creu que la POP pot accelerar el procés d'agregació de l' α -sinucleïna, mentre que la presència de POP inhibida indueix l'efecte contrari i retarda la seva agregació. Probablement, el mecanisme té lloc a través d'una interacció directa entre la POP i l' α -sinucleïna. *In vivo*, s'ha demostrat que l'administració d'inhibidors de la POP a diversos models cel·lulars i animals incrementa la degradació de l' α -sinucleïna i els seus agregats. Aquest descobriment ha revolucionat la POP com a diana d'interès terapèutic per al tractament d'aquesta malaltia.

El disseny d'inhibidors de la POP ha estat una línia de recerca important. La majoria d'aquests compostos són peptidomimètics derivats a partir del compost canònic benziloxycarbonil prolil prolinal (ZPP). Aquest compost mimetitzava els substrats naturals

de la POP; per la seva part, la presència de l'aldehid C-terminal crea un enllaç covalent tipus hemiacetal amb la Ser 554 del centre actiu de la POP. D'aquesta manera, aquest compost actua com un anàleg de l'estat de transició de la POP. Un altre inhibidor àmpliament estudiat ha estat el KYP-2047; en aquest cas, l'aldehid està substituït per un nitril, que dóna lloc a un enllaç iminoèter.

Tot i haver-se realitzat grans esforços per al desenvolupament d'inhibidors de la POP altament potents, la majoria dels assajos *in vivo* en models animals només han resultat en lleugeres millores en algunes capacitats cognitives. Malauradament, pocs compostos han arribat a la fase clínica, on han demostrat ser no tòxics i tenir bones propietats farmacològiques; però cap d'aquests inhibidors de la POP han progressat al mercat. En aquesta direcció, es requereixen més estudis emprant models animals adequats i una indicació terapèutica definida.

OBJECTIUS

Com s'ha comentat, la POP és una peptidasa d'interès terapèutic. No obstant, es coneix poc sobre la seva estructura en solució, així com l'efecte que hi poden causar els inhibidors.

Per una part, certs estudis suggereixen que la POP es troba en solució en un equilibri conformacional entre confòrmers tancats i oberts; per contra, altres línies de recerca són més partidàries de la implicació de certes cadenes polipeptídiques en l'exposició periòdica del centre actiu, sense que la POP experimenti separacions entre els dos dominis. D'acord amb la literatura que s'ha publicat fins al moment referent al camp de l'estructura i mecanisme de la POP, està clar que calen estudis estructurals sòlids usant tècniques biofísiques robustes.

D'aquesta manera, ampliar el coneixement sobre l'estructura de la POP en solució i estudiar l'efecte dels inhibidors facilitarà l'elucidació del seu rol biològic, així com el desenvolupament de nous inhibidors més efectius. Així, la present tesi doctoral planteja els següents objectius:

1. **Assolir un millor coneixement de l'estructura i la dinàmica de la POP humana en solució.** Per tal d'assolir aquest objectiu, s'empraran diferents

tècniques biofísiques. A l'hora d'interpretar i analitzar els resultats, també s'utilitzaran simulacions computacionals.

A més a més, donat que la POP és una proteïna dinàmica de grans dimensions, s'aprofitarà per tal de *valorar l'aplicació de l'espectrometria de masses de mobilitat iònica (EMMI) per a l'estudi de proteïnes dinàmiques estructurades.*

2. Especialment, es **prestarà atenció als efectes dels inhibidors sobre la dinàmica conformacional de la POP**, mitjançant la mateixa metodologia descrita al punt anterior.
3. Recentment, s'ha demostrat que els inhibidors de la POP són capaços d'incrementar la degradació de l' α -sinucleïna *in vivo*. Probablement, aquest efecte sigui degut a una interacció directa entre la POP i l' α -sinucleïna. Per tal d'entendre aquesta possible implicació de la POP, **s'estudiarà la possible interacció entre la POP i l' α -sinucleïna.**

CAPÍTOL 1: ESTUDI DE LA POP PER RESSONÀNCIA MAGNÈTICA NUCLEAR

La ressonància magnètica nuclear (RMN) és una de les tècniques més emprades en química i biologia estructural per a la determinació de l'estructura, dinàmica i interaccions de molècules. Aquesta tècnica es basa en el desdoblament energètic que experimenten els *spins* nuclears en presència d'un camp magnètic extern. Depenent de l'entorn químic dels nuclis, aquest desdoblament es veu més o menys afectat, i conseqüentment, les seves freqüències associades. Aquestes freqüències són registrades per l'equip i donen lloc als espectres de RMN. En principi, cada tipus de nucli actiu dona lloc a un espectre (en biomolècules, els nuclis més típics són el ^1H , ^{15}N i ^{13}C). No obstant, els experiments de RMN es poden dissenyar de tal manera que observin simultàniament dos nuclis; en aquesta tesi s'usen principalment els experiments d'heterocorrelació ^1H - ^{13}C .

Malauradament, els experiments de RMN estan limitats per la mida de les molècules. Per una part, el solapament de les senyals obtingudes pot arribar a ser catastròfic, i per l'altra, les propietats de relaxació dels sistemes grans causen un eixamplament dels senyals que comprometen la relació senyal/soroll. El llindar per a l'estudi de les

molècules per RMN s'estableix en 30 KDa. Per això es va usar el marcatge selectiu de les metionines de la POP. Segons aquest marcatge, la mostra de POP només conté nuclis actius a la RMN als residus de metionina (en total, 12 residus).

Marcatge de la POP amb L-trifluorometionina

El ^{19}F és un nucli amb excel·lents propietats de RMN. Té una bona sensibilitat, una mida similar al ^1H i molta sensibilitat a l'entorn químic. El marcatge de proteïnes amb aminoàcids fluorats no naturals ha donat bons resultats en algunes ocasions. En aquest cas, es va decidir marcar la POP amb L-trifluorometionina per tal de poder realitzar experiments ràpids de RMN de ^{19}F .

Per a tal fi, es va sintetitzar amb èxit la L-trifluorometionina d'acord amb una reacció de trifluorometilació electròfila. Aquest derivat d'aminoàcid va ser usat per suplementar bacteris d'*E.coli* auxòtrofs a la metionina, o dit d'una altra manera, que requereixen de metionina externa per a créixer. Malauradament, la L-trifluorometionina va resultar força tòxica per als bacteris i per a tal fi es va haver d'usar una mescla de L-trifluorometionina i metionina. Això va ocasionar que la incorporació del marcatge no superés el 50 %. Aquest fet, unit al baix rendiment de l'expressió i a la baixa estabilitat de la mostra, van ocasionar que els espectres de RMN de ^{19}F de la POP marcada amb L-trifluorometionina tinguessin una molt baixa relació senyal/soroll. Així, aquesta estratègia de marcatge es va abandonar.

Marcatge de la POP amb [metil- ^{13}C]-metionina

El segon pas fou utilitzar el marcatge selectiu de la POP amb [metil- ^{13}C]-metionina. Per tal d'expressar aquesta POP marcada, es van usar bacteris d'*E.coli* auxòtrofs a la metionina suplementats amb [metil- ^{13}C]-metionina. Aquest marcatge va permetre usar experiments tipus TROSY (concretament, ^1H - ^{13}C metil-TROSY HMQC). Els experiments TROSY es basen en la cancel·lació dels mecanismes de relaxació que causen l'eixamplament de les senyals, possibilitant així l'estudi de sistemes de gran pes molecular (com es el cas de la POP).

L'espectre d'heterocorrelació ^1H - ^{13}C de la POP va ser d'una gran qualitat, i mostrava 10 senyals molt ben distribuïdes d'un total de 12 residus de metionina. Mitjançant l'ús

de mutagènesi dirigida, es van aconseguir assignar inequívocament aquestes senyals. La optimització dels experiments d'heterocorrelació i l'assignació de les senyals van ser dutes a terme per la Dra. Teresa Tarragó.

L'anàlisi de l'espectre de la POP lliure va desvetllar que algunes de les senyals mostraven un segon joc de ressonàncies, indicant que un segon confòrmer de la POP molt diferenciat estructuralment es trobava en equilibri lent. A més, la presència d'inhibidors (en concret, ZPP i KYP-2047) va causar desplaçaments generalitzats en les senyals de l'espectre. Es va veure que les senyals de la POP unida a inhibidors coincidien amb el segon joc de senyals de la POP lliure, indicant que els inhibidors, de fet, seleccionaven una conformació pre-existent de la POP. També es va observar un altre resultat important després de calcular la pertorbació del desplaçament químic (una senzilla fórmula que permet avaluar els desplaçaments que pateixen els senyals). Segons aquest càlcul, els residus més afectats per la unió dels inhibidors es situaven al domini α/β -hidrolasa, indicant que el canvi conformacional causat pels inhibidors repercutia majoritàriament en aquest domini.

Aquests resultats apuntaven cap a un equilibri conformacional lent (en l'escala de μ -ms) de la POP, que en presència d'inhibidors es desplaçava totalment cap a un únic confòrmer. D'aquesta manera, es va decidir aprofitar tot el potencial de la RMN i del marcatge de les metionines per tal de quantificar els paràmetres d'aquest equilibri. Per a tal fi, es van realitzar experiments de dispersió de relaxació (DR). Breument, els experiments de DR permeten quantificar la dinàmica en l'escala de μ s-ms. Es tracta d'experiments de RMN dissenyats de tal manera que són capaços de modular la contribució dels fenòmens de bescanvi a la relaxació dels senyals. D'aquesta manera, modulant aquest element s'aconsegueix que la relació senyal/soroll dels senyals variï i sigui possible quantificar la relaxació.

No obstant, aquests experiments requereixen l'absència total de ^1H a la mostra (lògicament exceptuant als metils de les metionines), ja que els ^1H interfereixen amb la relaxació i per tant amb les mesures de DR. Aquest fet va implicar un total redisseny del protocol d'expressió de la POP marcada. Per una part, es va usar el protocol d'expressió de proteïnes perdeuterades, que permet la substitució dels ^1H per ^2H (per

a tal finalitat, es va usar aigua i glucosa deuterades per a la preparació dels medis de cultiu). Per l'altra, es va suplementar el medi de cultiu amb [metil- ^{13}C]-metionina-(2, 3, 3, 4, 4- d_5), per tal de garantir la total incorporació de ^2H també als residus de metionina.

Aquest anàleg d'aminoàcid no és comercial, amb la qual cosa es va haver de dissenyar un procediment sintètic per tal de sintetitzar la [metil- ^{13}C]-metionina-(2, 3, 3, 4, 4- d_5). Aquest procediment consistia en la síntesi d'un precursor electròfil de la cadena lateral totalment deuterat a partir de molècules orgàniques senzilles i reactius deuterats comercials. Mitjançant una reacció de substitució nucleòfila catalitzada per un agent quiral, es va aconseguir obtenir un precursor de la metionina enantiomèricament pur i totalment deuterat a les posicions requerides. Amb poques etapes sintètiques més, es va obtenir finalment la [metil- ^{13}C]-metionina-(2, 3, 3, 4, 4- d_5).

L'expressió de la POP totalment deuterada marcada als metils de les metionines va procedir amb èxit, i acte seguit els experiments de DR es van dur a terme. Aquesta part de l'estudi es va dur a terme en col·laboració amb el Dr. Oscar Millet (CIC bioGune, Vizcaya). Aquests experiments van demostrar ser molt robustos tot i les grans dimensions de la POP, i van permetre obtenir unes mesures de DR d'una gran qualitat. Les mesures de DR de la POP lliure demostraven que força residus de metionina estaven afectats per una dinàmica en l'escala dels μs - ms de gran amplitud. Curiosament, la majoria d'aquests residus dinàmics es trobaven al domini α/β -hidrolasa, indicant un cop més la implicació d'aquest domini en els canvis conformacionals. No obstant, en el cas de la POP unida a inhibidors (concretament, ZPP i KYP-2047) les mesures de DR no van detectar processos dinàmics en l'escala dels μs - ms . Aquest resultat refermava que els inhibidors estabilitzaven un dels confòrmers de la POP, congelant així tot el bescanvi conformacional en aquesta escala de temps.

Així i tot, aquestes dades eren merament quantitatives. Donat que la qualitat de les mostres de la POP ho permetia, es van ajustar les dades de DR a un model de bescanvi de dos estats. Aquest ajust va indicar que el bescanvi de la POP estava millor descrit

per un conjunt de moviments de diferents freqüències i amplituds més aviat que per un únic moviment generalitzat. Les constants de bescanvi dels residus obtingudes per l'ajust estaven acotades entre els 40 i els 170 s⁻¹ aproximadament, indicant que es tractava d'un equilibri lent. A més, les poblacions entre els dos estats conformacionals es situava a prop del 50 %.

Adicionalment, es van dur a terme experiments de DR en el sistema rotatori ($R_{1\rho}$) emprant una mostra de POP perdeuterada marcada amb [metil-¹³C]-metionina amb la cadena lateral no deuterada. Aquests experiments, sensibles a dinàmiques en l'escala ràpida dels μ s, van detectar que les metionines situades a les zones pròximes a les cadenes polipeptídiques que envolten el centre actiu estaven fortament afectades per dinàmiques en aquesta escala de temps.

Sumari del capítol 1

En resum, en aquest capítol s'ha detectat que la POP es troba en un equilibri conformacional lent (en l'escala dels μ s-ms) entre dos estats poblats aproximadament al 50 %. Les transicions estructurals que donen lloc a aquest equilibri són de gran extensió, i impliquen majoritàriament al domini α/β -hidrolasa. A més, es va veure que les cadenes polipeptídiques que envolten el centre actiu també eren dinàmiques en l'escala de temps ràpida dels μ s. Curiosament, els experiments de DR van evidenciar que els inhibidors de la POP desplaçaven totalment l'equilibri conformacional cap a un dels confòrmers pre-existents, congelant així la dinàmica de la POP.

CAPÍTOL 2: ANÀLISI DE L'ESTRUCTURA DE LA POP EN SOLUCIÓ PER DISPERSIÓ DE RAIGS X DE BAIX ANGLE

Un cop es va haver constatat la diversitat conformacional de la POP en solució, es va procedir a determinar quines eren les estructures dels seus confòrmers. Per a tal fi, es va escollir la dispersió de raigs X de baix angle (SAXS, l'acrònim en anglès de *small-angle X-ray scattering*).

Aquesta tècnica biofísica és especialment útil per a la determinació d'estructures de baixa resolució en solució de sistemes dinàmics complexos. Breument, es basa en la dispersió de radiació col·limada pels electrons de les molècules. Aquesta dispersió

implica un canvi en la direcció de la radiació, però no de la seva energia: és a dir, el seu nombre d'ona no es veu afectat. Malauradament, a conseqüència del moviment lliure de les molècules en solució, aquest perfil de dispersió només conté informació de baixa resolució. Una propietat important es que els perfils de dispersió obtinguts a partir de mescles heterogènies són un promig dels perfils individuals de tots els components de la mescla. D'aquesta manera, mitjançant l'ús de diferents algorismes, és possible de desxifrar quins compostos es troben presents en la mescla (aquest principi s'aplicarà més endavant).

Per tal de dur a terme les mesures de SAXS de la POP, es va realitzar una col·laboració amb el Dr. Pau Bernadó i la Fátima Herranz (Centre de Biochimie Structurale (CNRS), Montpellier, França). Els experiments es van realitzar al sincrotró PETRA III (DESY-EMBL, Hamburg, Alemanya). Especialment crític per a la qualitat de les mesures de SAXS va ser evitar la presència d'agregats de proteïna, la senyal dels quals podien interferir amb la mesura. Així, es va realitzar una cromatografia d'exclusió molecular acoblada al detector de SAXS, enregistrant una corba de SAXS a cada segon.

Aquest muntatge va permetre monitoritzar l'elució del monòmer de la POP, evitant així la contaminació per agregats residuals. Aquest experiment es va repetir també en el cas de la POP unida a ZPP. En ambdós casos es va parlar una especial atenció al tractament de dades. Aquest procés va començar amb la detecció de les zones d'elució de la solució tamponant pura. Tots els perfils es van promitjar, i el blanc es va sostreure a totes les corbes de SAXS. Acte seguit, mitjançant un programa informàtic específic i fent ús de l'aproximació de Guinier, es va extreure el radi de gir R_g i la intensitat a l'origen $I(0)$ a partir de cada corba de SAXS (segons la llei de Guinier, la part inicial de les corbes de SAXS es poden ajustar per regressió lineal per així extreure el radi de gir R_g i la intensitat a l'origen $I(0)$ de les molècules dispersants). Aquests dos paràmetres es varen representar en forma de cromatograma en funció del temps d'elució.

Així, es va poder constatar que la POP mostrava un R_g major que l'esperat per a l'estructura cristal·logràfica de la POP en la conformació tancada, i que estava acompanyada d'una segona espècie minoritària de menor R_g . En el cas de la POP unida

al ZPP, es va veure que només hi havia una sola espècie, d'un R_g força similar al de l'estructura cristal·logràfica.

No obstant, moltes de les corbes de SAXS estaven contaminades per les corbes del pic minoritari (anomenat pic II). Per a detectar les corbes corresponents a l'espècie majoritària, es va realitzar la descomposició de valors singulars del cromatograma de la POP. Concretament, aquest algoritme permet detectar quants components formen una certa senyal experimental: en aquest cas va permetre delimitar que els 100 corbes de SAXS més properes al màxim corresponien a les corbes pures de la POP. A més, la descomposició de valors singulars també va demostrar que les 100 corbes de SAXS més intenses de la POP unida al ZPP corresponien a corbes pures. Per tal d'obtenir les corbes pures de SAXS de la POP i de la POP unida al ZPP, aquestes 100 corbes més intenses es van promitjar respectivament. Acte seguit, mitjançant l'aproximació de Guinier, es va poder extreure un R_g de $28.50 \pm 0.06 \text{ \AA}$ i $27.40 \pm 0.06 \text{ \AA}$ per a la POP lliure i en presència de ZPP, respectivament. La distribució de distàncies d'aquestes corbes (un indicador de la homogeneïtat estructural) van indicar que la POP mostrava una forma irregular, mentre que la POP unida a inhibidors tenia una forma globular més regular.

Referent a la corba del pic II, es va haver d'emprar un mètode anomenat resolució de corbes amb multivariables mitjançant mínims quadrats. Es tracta d'un mètode capaç factoritzar una matriu de segon ordre, que en aquest cas estava formada per les variables cromatogràfiques i per les corbes de SAXS. Així, es va obtenir la corba corresponent al pic II, que malauradament va ser molt poc intensa deguda a la baixa concentració d'aquesta espècie de POP minoritària. No obstant, es va poder extreure el seu R_g de $25.00 \pm 0.09 \text{ \AA}$, que va resultar ser una mica inferior a l'esperat per a la forma cristal·logràfica de la POP en la conformació tancada (25.82 \AA).

Sumari del capítol 2

Els experiments de SAXS de la POP van permetre obtenir unes corbes de dispersió d'una excel·lent qualitat, gràcies al disseny experimental de cromatografia d'exclusió molecular acoblada a un detector de SAXS. D'aquesta manera, es va observar que la POP eluïa acompanyada d'una espècie minoritària (pic II) mentre que la POP unida al

ZPP constava d'un únic tipus d'espècies. Mitjançant diversos mètodes quimiomètrics, es van obtenir les corbes pures de la POP lliure i unida al ZPP, així com la corba corresponent al pic II. Aquestes corbes permeteren extreure els R_g de totes les espècies, i es va veure com en el cas de la POP aquest valor era superior al corresponent a l'estructura cristal·logràfica; no obstant, el valor del pic II era molt similar. Referent a la POP unida al ZPP, el R_g era una mica superior a l'obtingut per a la forma cristal·logràfica. Finalment, la distribució de distàncies de la POP indicava que aquesta tenia una forma irregular (probablement a conseqüència de separacions entre els dominis), mentre que la POP inhibida mostrava una estructura globular.

CAPÍTOL 3: SIMULACIONS DE DINÀMICA MOLECULAR DE LA POP I CORRELACIÓ AMB DADES EXPERIMENTALS

Fins al moment, es disposava d'unes dades experimentals de gran qualitat que ens havien permès extreure informació sobre la dinàmica, així com alguns detalls estructurals de baixa resolució. No obstant, les estructures de la POP d'alta resolució encara romanien sense resoldre. Per a aquest fi, es va decidir dur a terme diferents simulacions de dinàmica molecular amb la intenció de tenir un col·lectiu de possibles conformacions que podia adoptar la POP. Així, aquestes dades teòriques permetrien correlacionar i interpretar els resultats experimentals, especialment en el cas de les mesures de SAXS.

Sense entrar en detalls, les simulacions de dinàmica molecular estan basades en les lleis de la mecànica quàntica, i permeten extreure propietats termodinàmiques de sistemes microscòpics. D'acord amb els postulats de la mecànica estadística, és possible extrapolar aquestes propietats microscòpiques als sistemes macroscòpics, i per tant, mesurables. D'aquesta manera, les simulacions de dinàmica molecular combinen aquestes lleis per tal de predir l'evolució de cert sistema en funció del temps. Les simulacions de dinàmica molecular consisteixen en un model determinista que permet predir les posicions i velocitats dels àtoms d'una molècula donades unes condicions inicials.

Per tant, aquestes simulacions depenen molt de les estructures i condicions inicials; en el cas de biomolècules, és molt important escollir les conformacions de partida (en

molts casos s'usen estructures experimentals obtingudes per raigs X o RMN). Si les condicions de partida són les adequades, l'anàlisi de les trajectòries de dinàmica molecular aporten informació d'alta resolució sobre l'evolució de l'estructura en funció del temps (típicament, entre els ps i els ns).

D'aquesta manera, es van dur a terme 3 simulacions de dinàmica molecular (anomenades MD1, MD2 i MD3), conjuntament amb una simulacions més per a la POP unida al ZPP (MD5). Aquests càlculs es van realitzar pel Dr. Martin Kotev sota la direcció del Dr. Víctor Guallar (membres del programa de recerca conjunt BSC-CRG-IRB Barcelona Supercomputing Center, Barcelona).

En el cas de la simulació MD1, es va usar l'estructura de la POP cristal·logràfica en la conformació tancada. Aquesta simulació va tenir lloc durant gairebé 2 μ s. Malauradament, l'anàlisi de la trajectòria no va mostrar separacions importants entre els dos dominis, tret d'alguns desplaçaments esporàdics. Donat que la separació interdominial de la POP implica el moviment de molts àtoms, és poc probable que aquesta transició estructural ocorri en l'escala de temps d'una simulació de dinàmica molecular. D'aquesta manera, es van dur a terme dues simulacions més directament a partir d'una conformació de la POP oberta. Al no haver-hi cap estructura cristal·logràfica per a la POP de mamífer en una conformació oberta, es va obtenir el model d'homologia porcina de la POP bacteriana cristal·litzada en la conformació oberta. Usant aquest model d'homologia com a estructura inicial, a la primera simulació (MD2) es va veure que l'estructura mantenia la separació interdominial durant la part inicial de la trajectòria, seguidament d'un tancament progressiu. Així, es va realitzar una segona simulació de la conformació oberta (MD3) en què es va incloure un període inicial de relaxació. Un cop l'estructura va ser relaxada, aquesta conformació oberta es va mantenir durant un període significatiu de temps de simulació. En el cas de la POP unida al ZPP, es va realitzar una simulació curta (MD5) a partir de l'estructura cristal·logràfica de la POP unida covalentment a aquest inhibidor. Al llarg de tota la simulació, l'estructura va demostrar molt poques fluctuacions.

Finalment, un cop es van haver realitzat aquestes simulacions, es van calcular les corbes de SAXS simulades de les estructures generades mitjançant un programa específic.

Ajust de les dades simulades amb les dades experimentals de SAXS

Un cop es va disposar d'aquest extens col·lectiu de possibles conformacions de la POP conjuntament amb les seves corbes de SAXS simulades, es van usar els perfils experimentals de SAXS per tal de determinar quines corbes simulades descrivien millor les dades experimentals. Aquests anàlisis es va realitzar en col·laboració amb el Dr. Pau Bernadó i la Fátima Herranz (Centre de Biochimie Structurale (CNRS), Montpellier, França).

Per a tal finalitat, es va usar el mètode d'optimització de col·lectius. Aquest mètode es basa en l'optimització d'un sub-col·lectiu mitjançant un algoritme genètic, seguint un determinat criteri de qualitat. Per tal de dur a terme aquesta optimització, es duen a terme operacions inspirades en les lleis de la genètica (fonamentalment, en els conceptes de creuament i mutació). En el cas de la POP, es va pretendre optimitzar un sub-col·lectiu format per 20 corbes de SAXS simulades segons la seva capacitat per descriure el perfil experimental.

Per a la POP lliure es van combinar les corbes de SAXS simulades obtingudes a la simulació MD1, a la part inicial de la simulació MD2, i a tota la simulació MD3. Del col·lectiu generat, es va seleccionar aleatòriament un sub-col·lectiu de 20 corbes i es va sotmetre al mètode d'optimització de col·lectius emprant la corba de SAXS obtinguda per a la POP com a criteri de qualitat. Aquest procediment va seleccionar satisfactòriament 20 corbes simulades que col·lectivament descrivien amb excel·lent fidelitat a la corba de SAXS experimental. Remarcablement, es va veure com 11 de les 20 corbes simulades corresponien a confòrmers totalment oberts, mentre que la resta es tractava de confòrmers tancats.

En el cas de la POP unida al ZPP, es van usar les corbes simulades de la MD5. El mètode d'optimització de col·lectius es va dur a terme anàlogament al cas de la POP lliure, però fent servir la corba experimental de SAXS de la POP inhibida amb ZPP com a

critèri de qualitat. També en aquest cas, el mètode d'optimització de col·lectius va proporcionar un conjunt de 20 corbes simulades que reproduïen les dades experimentals amb un elevat grau de correspondència; totes les estructures corresponien a confòrmers tancats de la POP.

La comparació entre les estructures de POP obertes i tancades obtingudes segons aquesta estratègia van demostrar que la separació entre dominis causava un gran increment en l'exposició al solvent. Especialment, aquests canvis en l'exposició afectaven alguns residus de metionina que es correlacionaven amb la pertorbació de desplaçament químic. A més, l'anàlisi detallat dels residus del centre actiu i la regió interdomini van assenyalar que la forma oberta de la POP era catalíticament inactiva com a resultat de l'alteració d'una xarxa de ponts salins propera al centre actiu.

Sumari del capítol 3

Resumint, en aquest capítol es van dur a terme diverses simulacions de dinàmica molecular que van generar un gran col·lectiu de confòrmers de la POP lliure i en presència d'inhibidors. Mitjançant el mètode d'optimització de col·lectius, es van obtenir aquelles estructures simulades que col·lectivament descriuen amb major fidelitat les corbes experimentals de SAXS. En el cas de la POP lliure, es va obtenir que aproximadament el 50 % de la POP es trobava en una conformació oberta, mentre que la resta es trobava en una conformació tancada. En el cas de la POP inhibida, es va trobar que només la conformació lliure hi era present.

CAPÍTOL 4: ESTUDI DE LA POP PER ESPECTROMETRIA DE MASSES DE MOBILITAT IÒNICA

En els capítols anteriors s'havia obtingut una descripció acurada de la dinàmica conformacional de la POP en solució mitjançant diverses tècniques instrumentals i simulacions. Es va detectar que la POP era altament dinàmica, i presentava un equilibri conformacional entre estructures obertes i tancades en l'escala de μ s-ms. Donades aquestes propietats, es va decidir aplicar l'espectrometria de masses de mobilitat iònica (EM-MI) per a observar aquesta dinàmica conformacional de la POP.

L'espectrometria de masses està basada en l'anàlisi de les masses dels ions en fase gas. Per a tal fi, els espectròmetres de masses consten d'una font que genera ions d'analit a partir de la mostra en diferents estats (per exemple, en solució o en fase gas). En el cas de proteïnes, l'ús de tècniques d'ionització suaus permet obtenir ions de proteïna a partir de solucions aquoses que poden mantenir la majoria de característiques de la seva estructura nativa. Els més emprats són l'electrosprai i el nano-electrosprai.

Concretament, l'EM-MI és una tècnica biofísica emergent basada en el principi de mobilitat dels ions carregats en el sí d'un gas. Segons aquest principi, els ions carregats que es troben a certa velocitat en un medi gasós experimenten una pèrdua d'energia cinètica deguda a les col·lisions amb les molècules de gas (un fenomen assimilable al fregament). La mobilitat depèn fortament de la mida i la forma dels ions; així, els espectròmetres de masses de mobilitat iònica incorporen una cambra de mobilitat iònica abans de l'analitzador de masses. D'aquesta manera, l'EM-MI permet la separació simultània d'ions d'acord amb la seva forma i mida, conjuntament amb l'anàlisi de la seva massa. Aquesta tècnica ha permès separar, per exemple, diferents espècies oligomèriques de proteïnes amb la mateixa relació càrrega/massa, o bé complexes de proteïnes amb diferents topologies. Tot i haver-hi diversos instruments d'EM-MI comercials, la seva aplicació a proteïnes dinàmiques no és molt extensa. Per això es va decidir analitzar la POP segons aquesta tècnica.

Aquests experiments es van realitzar en col·laboració amb la Dra. Marta Vilaseca (directora de la plataforma d'espectrometria de masses de l'*Institute for Research in Biomedicine*, Barcelona). Es va usar un instrument d'EM-MI SYNAPT G1 HDMS (Waters Corp, Massachusetts, USA).

El primer pas va consistir en optimitzar la ionització de la POP. Donat que la qualitat dels resultats depenen fortament d'aquesta etapa, es van destinar molts esforços per tal d'evitar artefactes derivats de la ionització. Es va escollir el nano-electrosprai en mode positiu com a font d'ionització, ja que consumia una quantitat molt reduïda de mostra, demostrava unes condicions d'ionització reproduïbles, i generava un flux d'ions molt estable. Es va decidir escollir els voltatges de la font més baixos possibles que

donessin senyal estable. D'altra banda, la preparació de les mostres de POP també van ser un punt clau per a l'èxit dels resultats. La puresa de les mostres havia de ser molt alta, així com la composició i qualitat del tampó (el tampó més adequat fou l'acetat d'amoni aquós a una concentració de 50 mM).

D'aquesta manera es va aconseguir generar uns bons espectre de masses i de mobilitat iònica de la POP. Els estats de càrrega en què es trobava la POP en aquestes condicions van ser el +17, +18, +19, +20 i +21. En el cas dels estats +18 al +21, els espectres de mobilitat iònica corresponents van desvetllar diferents pics amb poblacions variables. Mitjançant un ajust de corbes Gaussianes, es van poder ajustar tots espectres de mobilitat iònica a tres Gaussianes, corresponents a diferents espècies. Gràcies a una calibració amb ions de proteïnes natives de seccions conegudes, va ser possible extreure uns valors de secció per a aquestes tres espècies de POP. Finalment, els experiments de la POP unida a inhibidors (ZPP i KYP-2047) donaren uns resultats pràcticament idèntics.

Arribats a aquest punt, resultava ambigu assignar les espècies de POP en fase gas amb diferents estats conformacionals només basant-se en aquestes dades experimentals. Així, es van realitzar diferents simulacions computacionals de la POP en fase gas per part del Dr. Sergio Madurga (Departament de Química Física de la Universitat de Barcelona). En primer lloc, es van fer unes simulacions curtes en solució per tal de determinar quines serien les dimensions dels ions de POP en la conformació tancada si s'analitzessin en unes condicions similars a les d'un experiment d'EM-MI. Acte seguit, es van fer unes simulacions per determinar com es comportaria la POP en absència de solvent. Per acabar, es van trobar les seccions dels confòrmers oberts de la POP obtinguts en el capítol anterior pel mètode d'optimització de col·lectius. D'aquesta manera, es va trobar una correlació molt clara entre aquestes tres simulacions, i les tres espècies de POP obtingudes per EM-MI. Dues de les espècies tenien unes seccions que es corresponien amb els confòrmers oberts i tancats de la POP. La tercera, que tenia una secció menor que l'esperada per al confòrmer tancat, es tractava d'una espècie amb un empaquetament més compacte degut a l'absència del solvent.

Sumari del capítol 4

Segons els resultats obtinguts amb la POP, l'EM-MI va demostrar la seva capacitat per resoldre equilibris conformacionals lents implicant transicions estructurals importants. No obstant, aquesta tècnica encara està en desenvolupament i calen més estudis per tal de confirmar la seva versatilitat en l'anàlisi de proteïnes dinàmiques. No obstant, l'optimització de les condicions experimentals és crucial per tal de garantir la qualitat dels resultats i evitar artefactes derivats de la fase gas. Per tal d'interpretar els resultats, és molt recomanable disposar de simulacions computacionals o dades experimentals provinents d'altres tècniques.

CAPÍTOL 5: ESTUDI DE LA INTERACCIÓ ENTRE LA POP I L' α -SINUCLÈÏNA

Introducció

Com es va comentar a la introducció d'aquest resum, recentment s'ha descobert que la POP i l' α -sinucleïna, el major component dels cossos de Lewi present als cervells dels afectats per la malaltia de Parkinson, poden interaccionar. Abans de començar, es farà una breu introducció de la implicació de l' α -sinucleïna en la malaltia de Parkinson.

L' α -sinucleïna és una proteïna desestructurada de 14 KDa expressada sobretot en les neurones, que té una gran densitat de residus carregats i de prolina a la seva zona C-terminal. La zona central, en canvi, té tendència a formar interaccions intermoleculares que acaben donant lloc a oligòmers, i finalment, a agregats de proteïna insolubles. Precisament, aquests agregats insolubles són els que majoritàriament formen els cossos de Lewi.

Avui dia no és clar si les espècies neurotòxiques són els oligòmers solubles o bé les espècies agregades, ja que en mostres de pacients afectats per la malaltia de Parkinson i en nombrosos models animals es troben multitud d'espècies oligomèriques i agregades. Tampoc és evident quin és la causa que desencadena aquest procés. S'ha trobat que tres mutacions puntuals en la seqüència de l' α -sinucleïna donen lloc al desenvolupament prematur de la malaltia de Parkinson, però molt probablement el desenvolupament d'aquesta malaltia parteix d'un desajust entre l'expressió i la degradació d'aquesta proteïna, així com en condicions d'estrès oxidatiu en les

neurons. En tot aquest escenari s'han trobat certes proteïnes que interaccionen amb l' α -sinucleïna i permeten modular la seva agregació; la POP és una d'elles.

En aquesta línia de recerca, s'ha vist que l'administració d'inhibidors de la POP en cultius neuronals i en models animals que expressen una de les formes mutades d' α -sinucleïna redueix la quantitat d'agregats d' α -sinucleïna, causant així una millor supervivència de les neurones. Estudis posteriors indicaren que, molt probablement, els inhibidors de la POP afavorien les vies naturals de degradació de l' α -sinucleïna. Aquest efecte podia ser degut a una interacció directa entre la POP i l' α -sinucleïna, ja que un estudi molt recent ha demostrat que les dues proteïnes poden interaccionar amb una afinitat en l'ordre de μM .

Així, donada la nostra experiència en l'estudi de l'estructura i la dinàmica de la POP, es va procedir a estudiar aquesta interacció amb més detall.

Estudi de la interacció entre la POP i l' α -sinucleïna per RMN

La tècnica biofísica emprada per a l'estudi d'aquesta interacció va ser la RMN. Aquesta tècnica permet l'estudi de les interaccions entre dues molècules en un ampli ventall d'afinitats i mides moleculars. Existeixen diferents experiments de RMN depenent de l'afinitat de la interacció. L'afinitat es mesura a partir de la constant de dissociació (K_d) del complex format: així, es poden distingir interaccions fortes i permanent (K_d per sota l'ordre dels nM) o bé dèbils i transitòries (K_d majors que 10^{-4} M). En el primer cas, en un espectre bidimensional convencional, s'observaria el joc de senyals de RMN corresponent al complex, mentre que en el segon es veuria una senyal promitja. En el cas de la POP i l' α -sinucleïna (K_d en l'ordre dels μM), la situació es trobaria més propera al segon cas.

Per tal de dur a terme l'estudi de la interacció entre la POP i l' α -sinucleïna, els experiments preliminars més apropiats serien els espectre d'heterocorrelació enregistrats amb diferents estequiometries de POP i α -sinucleïna. Seguidament, l'anàlisi de pertorbació del desplaçament químic (explicat anteriorment) i l'anàlisi d'intensitats de les senyals (que consisteix en calcular la intensitat relativa d'una

senyal respecte a la d'un espectre de referència) permetrien detectar canvis subtils en els senyals causats per la interacció.

Abans de procedir als experiments de RMN, es va haver d'optimitzar l'expressió d' α -sinucleïna, ja que mai abans s'havia treballat amb aquesta proteïna al nostre laboratori. Durant l'expressió amb medi mínim (necessari per a obtenir α -sinucleïna marcada uniformement amb ^{15}N) es va evitar la presència de traces de metalls pesants; així mateix, durant la purificació es va treballar escrupolosament amb solucions tamponants desgasades. Aquestes precaucions van ser necessàries per evitar l'oxidació indesitjada d'alguns residus; tot i així, va ser impossible evitar totalment la presència d'espècies d' α -sinucleïna oxidades. Deixant de banda aquestes impureses minoritàries, l' α -sinucleïna obtinguda segons el protocol optimitzat va resultar d'una qualitat adequada per als experiments de RMN.

Els primers experiments de RMN que es van dur a terme van ser espectres d'heterocorrelació ^1H - ^{13}C de la POP marcada amb [metil- ^{13}C]-metionina en presència d'1 equivalent d' α -sinucleïna. Aquest experiment es va repetir anàlogament per al cas de la POP unida al ZPP. Malauradament, en ambdós casos no es van apreciar diferències significatives en els espectres. Probablement, els canvis conformacionals de la POP arran d'aquesta interacció són difícils de detectar; a més, la baixa concentració de les dues proteïnes no afavoreix la interacció.

Per aquesta raó, es va optar per a realitzar experiments d'heterocorrelació ^1H - ^{15}N de l' α -sinucleïna marcada uniformement amb ^{15}N (és a dir, amb tots els nitrògens de les amides substituïts per ^{15}N). Per tal d'assignar els senyals, es van fer servir espectres assignats de la base de dades BioMagResBank (www.bmrb.wisc.edu), que van permetre l'assignació inequívoca de la majoria de senyals. Un cop assignat l'espectre, es va procedir a enregistrar-lo en presència de 6 equivalents de POP lliure, POP inhibida amb ZPP, i en presència de seroalbúmina humana (control negatiu). Segons aquesta estratègia, l'anàlisi d'intensitats i de pertorbació del desplaçament químic van desvetllar alguns canvis significatius en l'espectre de l' α -sinucleïna en presència de POP inhibida amb ZPP. Especialment, les intensitats d'una zona de la regió C-terminal van patir un descens sistemàtic, que es correlacionava amb una pertorbació del

desplaçament químic elevada en la mateixa zona. A més, una visió detallada de l'espectre va ressaltar que algunes senyals de l'extrem C-terminal mostraven un petit senyal satèl·lit.

Sumari del capítol 5

En resum, a la vista dels resultats obtinguts per RMN, podem dir que aquesta tècnica està al límit de detecció de la interacció entre la POP i l' α -sinucleïna. Probablement, aquesta interacció sigui dèbil i transitòria, o bé les concentracions emprades no afavoreixin la seva interacció. No obstant, en el cas de l' α -sinucleïna en presència de POP inhibida amb ZPP, es van detectar alguns canvis significatius en la zona C-terminal de l' α -sinucleïna. Curiosament, aquest resultat indica que el patró d'interacció de la POP es modifica per la presència d'inhibidors, probablement degut als canvis conformacionals que ocasionen aquestes molècules. Aquest estudi ha servit per la posta a punt de la metodologia de treball per a futurs experiments en aquesta direcció en el nostre laboratori.

DISCUSSIÓ GENERAL

La POP és una diana terapèutica per al tractament dels dèficits cognitius associats a les malalties mentals. Recentment, s'estan desenvolupant inhibidors de la POP per al tractament preventiu de la malaltia de Parkinson. No obstant, tot i que l'estructura cristal·logràfica d'aquesta proteïna de 81 KDa és ben coneguda, algunes qüestions sobre el seu mecanisme i funció biològica romanen encara sense resoldre.

Es desconeix quines són les fluctuacions estructurals que permeten que aquesta peptidasa capturi els seus substrats i expulsi els productes hidrolitzats, ja que el centre actiu es troba en una regió inaccessible. A més, l'efecte dels inhibidors en aquestes transicions estructurals tampoc és conegut. Recentment es sap que l'activitat hidrolítica de la POP no és la responsable de la seva funció biològica, i per tant aquestes fluctuacions estructurals poden ser importants per al reconeixement molecular amb altres proteïnes.

Certes línies d'investigació apunten que els dos dominis de la POP es mantenen sempre en contacte total donant lloc exclusivament a una conformació tancada. Per tal de

permetre la circulació de substrats i productes, les cadenes polipeptídiques que envolten el centre actiu es desprendrien periòdicament, exposant així el centre actiu a l'exterior. No obstant, altres evidències suggereixen que els dos dominis es poden separar, donant lloc a una conformació de la POP oberta en equilibri amb la conformació tancada. Aquesta conformació oberta, però, només s'ha pogut observar en algunes estructures de raigs X de POP bacteriana. Estudis previs en el nostre grup van demostrar que la POP és altament dinàmica en l'escala dels μ s-ms, i que aquesta dinàmica es centra en la regió interdomini. Per tant, calen estudis fiables per tal de clarificar aquestes hipòtesis.

En aquesta tesi doctoral s'ha emprat una metodologia mixta basada en tècniques biofísiques robustes (RMN i SAXS) complementades amb simulacions computacionals. S'ha demostrat que la POP existeix en solució en un equilibri dinàmic lent en l'escala dels μ s-ms entre conformacions obertes i tancades. Els inhibidors de la POP estableixen la conformació tancada, desplaçant totalment l'equilibri conformacional.

S'ha pogut observar com la forma oberta de la POP és catalíticament inactiva i exposa grans regions al solvent, mentre que la tancada presenta la disposició dels residus adequada per a la catàlisi. Probablement, l'obertura i tancament de la POP són els passos limitants per al seu cicle enzimàtic, tal i com ho demostra la correlació amb alguns paràmetres cinètics.

No obstant, molt possiblement aquesta riquesa conformacional de la POP ha sigut aprofitada pels organismes per tal de modular funcions biològiques a través d'interaccions amb altres proteïnes abans que a través de la seva activitat enzimàtica. És un fet demostrat en nombrosos casos que l'equilibri conformacional de les proteïnes intervé en la regulació bi-direccional de la transducció de senyal.

Recentment, s'ha demostrat que una possible interacció entre la POP i l' α -sinucleïna pot modular els nivells d'agregats d' α -sinucleïna *in vivo*. Segons els nostres resultats de RMN sobre aquesta possible interacció, s'ha vist que la presència de POP inhibida causa alteracions en la zona C-terminal de l' α -sinucleïna, segurament a conseqüència d'una interacció feble i transitòria. Tenint en compte les nostres observacions sobre la dinàmica conformacional de la POP, podem fer la hipòtesi que els inhibidors de la POP

potencien la interacció amb l' α -sinucleïna tot afavorint la conformació tancada, que possiblement presenta una major afinitat amb l' α -sinucleïna.

D'aquesta manera, d'acord amb els objectius plantejats a l'inici d'aquesta tesi, es poden extreure les següents conclusions:

1. **S'ha demostrat que la POP es troba en solució en un equilibri en l'escala dels μ s-ms entre una conformació oberta i tancada.** Aquest equilibri té implicacions directes en el cicle enzimàtic de la POP. A més, s'ha pogut aplicar la EM-MI amb èxit per a l'anàlisi de la POP, fet que assenyala que *la EM-MI és tècnica emergent útil per a l'estudi de proteïnes estructurades dinàmiques.*
2. **Els inhibidors de la POP desplacen totalment l'equilibri estabilitzant la conformació tancada.** Aquest fet bloqueja l'activitat enzimàtica, però molt més important per als sistemes biològics, el desplaçament de l'equilibri natiu de la POP pot modificar el seu patró de reconeixement molecular. Així, aquest resultat encoratja a dissenyar nous inhibidors de POP concebuts com a "interruptors" conformationals.
3. **Els experiments ^1H - ^{15}N HSQC preliminars suggereixen que la POP i l' α -sinucleïna poden interactuar de manera dèbil i transitòria, especialment en el cas de la POP inhibida.** Probablement aquesta interacció involucra una zona de la regió C-terminal de l' α -sinucleïna. Aquest resultat suggereix que el reconeixement entre les dues proteïnes depèn de l'estat conformational de la POP, fet que confirma que la riquesa conformational de la POP és la responsable per a la seva activitat biològica.

Appendix

APPENDIX I: RELAXATION IN NMR

This section provide a schematic description of the relaxation concepts used throughout the main text. Before explaining the relaxation mechanisms, a brief overview of the NMR basics is included as an introduction.

NMR basics

In the absence of an external magnetic field, the magnetic moments of the nuclear spins are randomly ordered throughout the sample; as a consequence, this situation yields a complete cancellation of all the magnetic moments (Figure 83 A). However, if the sample is under the influence of an external magnetic field (B_0), the nuclear spin magnetic moments adopt a preferential orientation along the z-axis. This preferential alignment is partially disrupted by the thermal motions of the molecule, which favour the random orientations. However, the presence of transiently oriented magnetizations originates a macroscopic magnetization along the z-axis (M_0 , Figure 83 B). This orientation along the z-axis is the equilibrium situation of the bulk magnetization in the presence of B_0 .

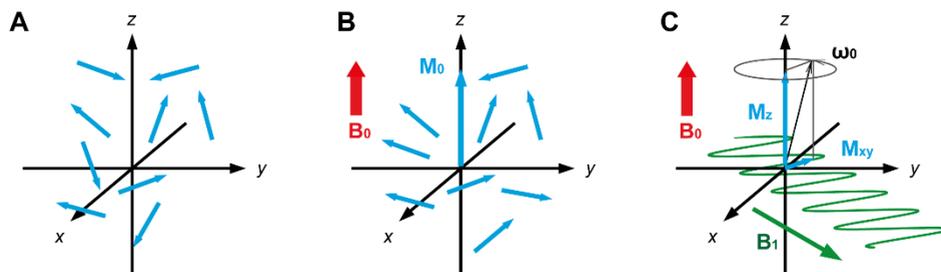


Figure 83: Magnetic moments of nuclear spins in the absence or in the presence of an external magnetic field (B_0). A) Random orientation of the magnetic moments in the absence of B_0 . B) Preferential orientation of nuclear spins in the presence of B_0 (parallel to the z-axis), which yields the bulk magnetization M_0 . C) Tilt of M_0 out of the z-axis by the effect of a radiofrequency pulse in the transverse plane (B_1). This generates a magnetization component in the transverse plane (M_{xy}) which oscillates according to its Larmor frequency (ω_0).

However, this equilibrium situation can be perturbed. In the case of an NMR spectrometer, the magnetization is perturbed by applying a radiofrequency (rf) pulse in the transverse

plane (B_1 , Figure 83 C). This rf pulse interacts with the magnetic moments, and tilts the magnetization out of the z-axis. The magnetization out of the equilibrium precesses according to its Larmor frequency (ω_0), calculated with the following expression:

$$\omega_0 = -\gamma \cdot B_0$$

In which γ is the gyromagnetic constant of the corresponding nucleus.

At this point, it is of great utility to introduce the rotating coordinate frame. This coordinate frame eliminates the time dependence of the oscillating B_1 field. Therefore, the Larmor frequency is expressed in the rotating coordinate frame as:

$$\Omega = \omega_0 - \omega_{rf}$$

Where ω_{rf} is the frequency of the applied field (also known as the transmitter frequency), and Ω is known as the offset. The offset can be used to define the reduced field (ΔB), according to:

$$\Omega = -\gamma \cdot \Delta B$$

The reduced field allows us to analyse more easily the tilt of the magnetization caused by B_1 . Hence, ΔB and B_1 are added vectorially to give rise to the effective field, B_{eff} (Figure 84).

$$B_{eff} = \sqrt{\Delta B^2 + B_1^2}$$

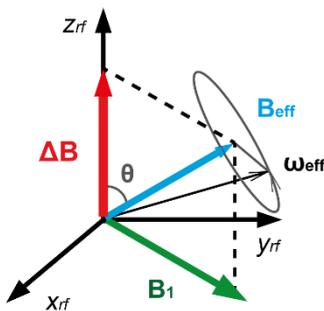


Figure 84: Interaction between reduced field (ΔB) and applied field (B_1) in the rotating coordinate frame, which give rise to the effective field (B_{eff}); B_{eff} forms an angle θ with respect to the z-axis. In the coordinate frame, the magnetization precesses around B_{eff} .

Under the rotating coordinate frame, the magnetization precesses around B_{eff} in the same way that it precessed around the z-axis in the laboratory frame. The tilt angle of the magnetization (θ) can therefore be defined as the angle between ΔB and B_1 . The following trigonometric equations allow to correlate all of the components of the rotating frame. As

it can be seen, placing the transmitter frequency close to ω_0 (i.e. setting the offset close to 0) minimizes ΔB , which result in higher tilt angles.

$$\sin \theta = \frac{B_1}{B_{eff}} ; \cos \theta = \frac{\Delta B}{B_{eff}} ; \tan \theta = \frac{B_1}{\Delta B}$$

The processes that take place after tilting the bulk magnetization give rise to the detectable NMR signal. After the rf pulse, the magnetization in the transverse plane returns to the equilibrium position along the z-axis. During this process, the precessing magnetization emits an electromagnetic signal, which is detected by the spectrometer receptor coils and generates the NMR signal. The process by which this magnetization returns to the z-axis equilibrium position is known as relaxation. This phenomenon is responsible for the free-induction decay of the NMR signal over time. Relaxation is a highly complex process, but with several approximations it can be easily understood.

Relaxation in NMR

Basically, relaxation proceeds by two components: the longitudinal relaxation (T_1) and the transversal relaxation (T_2 , Figure 85). Strictly, T_1 corresponds to the return of the magnetization to the equilibrium orientation towards z-axis, while T_2 corresponds to the loss of coherence in the transverse plane.

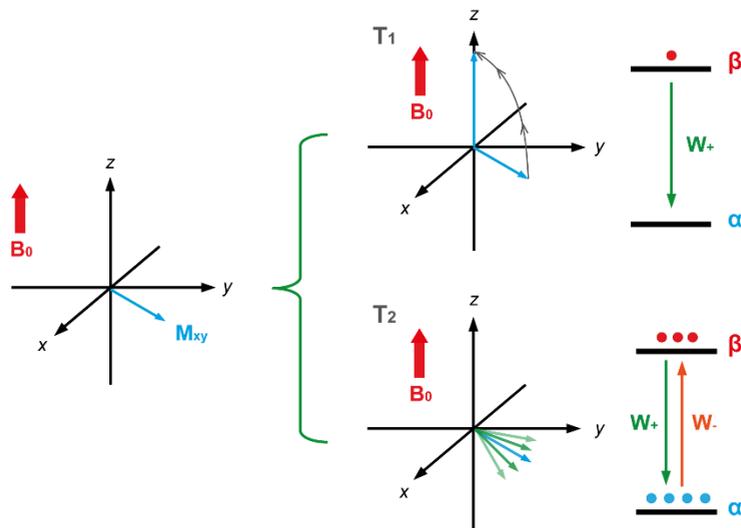


Figure 85: Description of T_1 and T_2 relaxation mechanisms. T_1 involves the return of the transverse magnetization back to the z-axis equilibrium orientation via W_+ transitions. T_2 is originated from the loss of coherence of the transverse magnetization due multiple paired W_+ and W_- transitions.

T_1 is strongly dependent on the molecular tumbling, and it can be conceived as the thermodynamic-driven component of the relaxation. If we suppose a nucleus with spin $\frac{1}{2}$ (state α) which is excited by a rf pulse, the $-\frac{1}{2}$ spin (state β) will become populated. After the radiation, spin β will evolve to the most stable spin α spontaneously due to longitudinal relaxation (Figure 85). According to the Bloembergen-Purcell-Pound (BPP) theory, the transitions between spins are associated to a given probability W_+ and W_- . This theory is useful to describe the spin populations at the different energy levels, but is not able to express the relaxation times.

Two different mechanisms mainly contribute to T_1 : the dipole-dipole interaction (DD), and the chemical shift anisotropy (CSA). DD interaction is originated from the interference between the different nuclear spins of the atoms of the molecule. As a consequence of the molecular mobility, the orientation of the molecule with respect B_0 changes, and consequently, the relative positions and orientations of spins. This causes fluctuating interactions between them, which favour the relaxation mechanism (Figure 86 A). In turn, the same spatial mobility of the molecule results in fluctuating shielding of nucleus by the surrounding electronic density with respect to B_0 . This constitutes the CSA mechanism (Figure 86 B).

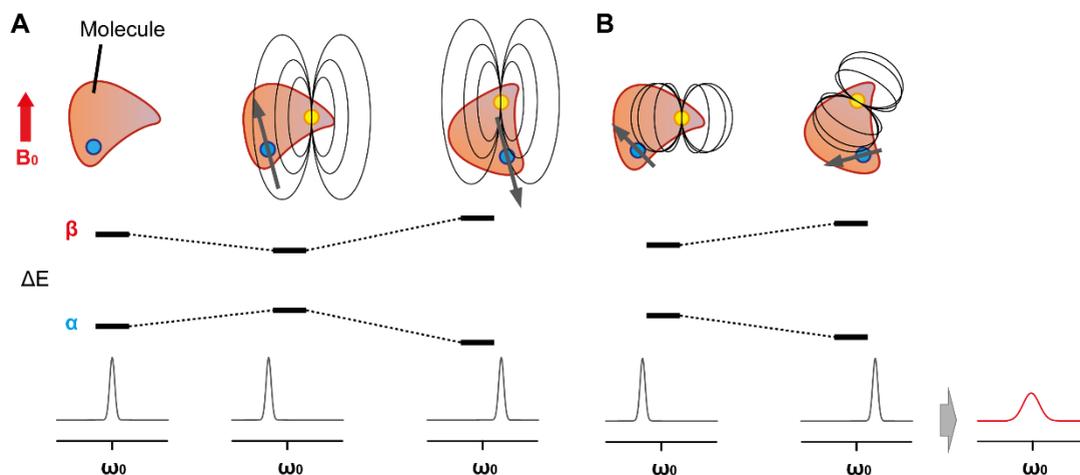


Figure 86: Mechanisms contributing to T_2 . A) Dipole-dipole interaction between the magnetic field generated by the spin of the yellow atom and the observed blue atom. This interaction fluctuates due to the molecular tumbling, and as a consequence, the spin energy levels are altered. B) Chemical shift anisotropy mechanism originated by the fluctuating shielding of the blue atom by the electronic cloud of the yellow atom. Together, these two mechanisms lead to NMR signal broadening.

Given that DD and CSA are originated by molecular tumbling, these two relaxation mechanisms are cross-correlated. The cross-correlation mechanism (η_{XY}) affects distinctly the two components of a doublet signal of systems with high MW. In the backbone amide ^1H - ^{15}N pair, one of the components is broadened as a consequence of the constructive interaction of η_{XY} ; in contrast, the other component is not affected given that η_{XY} is cancelled (Figure 87 A). In a conventional ^1H - ^{15}N HMQC experiment, the two components are mixed along the phase cycle and yield a broadened signal. In TROSY experiments [A1], however, the broadened component is cancelled during the phase cycle and the resulting signal is narrowed. Although this cancellation involves a loss of signal, it results in an important gain of resolution. A similar strategy can be used in the case of the methyl groups, but cancelling the intramethyl dipolar mechanisms (Figure 87 B). This is the basis of the methyl-TROSY HMQC experiments used in this PhD thesis.

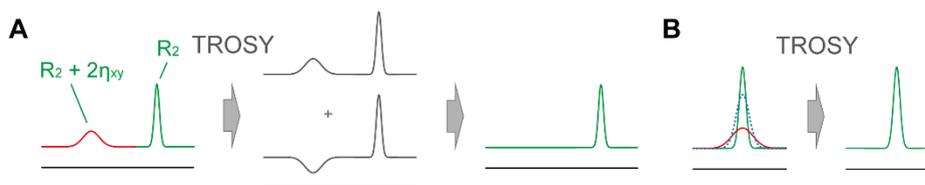


Figure 87: Cross-correlation between DD and CSA and TROSY principle. A) Scheme of a ^1H - ^{15}N doublet of a high MW system; as a consequence of the cross-correlation, one of the components of the doublet is broadened (red). The TROSY principle eliminates the fast relaxing component, resulting in significant gains of resolution. B) Scheme of a ^{13}C -decoupled methyl group of a high MW system (orange dotted line), formed by slow (green) and fast (red) relaxing components. Methyl-TROSY HMQC experiments select only the slow relaxing components.

T_2 is the entropic-driven component of relaxation. According to the BPP theory, several spins in α and β state can undergo to transitions between states without altering the global distribution of states (Figure 85). If the same number of W_+ and W_- transitions occur, the system undergo to changes in entropy but not in energy. With the help of the vector model, the microscopic mechanism can be explained in terms of fluctuations originated by the internal thermal motions of the molecule. These internal motions cause slight alterations of the chemical environment of the atoms, which results in dephasing of transverse magnetization and subsequent loss of coherence. This occurs progressively during the chemical shift and scalar coupling evolution of the magnetization, leading to a loss of signal in the transverse plane.

Given that T_2 strongly depends on the internal motions of molecules, the accurate measurement of T_2 is of great interest in the study of dynamics of molecules. The measurement of T_2 is typically sensitive to μs - ms time scale, which matches well with protein motions of functional relevance (e.g. the catalytic cycle or binding events). One of the most spread procedures for the measurement of T_2 are relaxation dispersion (RD) experiments. RD is defined as the dependence of the effective transverse relaxation rate ($R_{2, \text{eff}}$) with an applied effective field ω_e . The suppression of the dephasing caused by the exchange is modulated by ω_e , resulting in a decay curve of $R_{2, \text{eff}}$ by increasing ω_e (e.g. CPMG and $R_{1\rho}$ RD experiments, see below). $R_{2, \text{eff}}$ can be defined as:

$$R_{2, \text{eff}} = R_2 + R_{\text{ex}} \quad \text{Equation 13}$$

In which R_2 is the transverse relaxation rate in the absence of exchange, and R_{ex} is the exchange contribution to transverse relaxation.

CPMG RD experiments

Consider a nucleus of a protein in exchange between two states A and B, in slow to intermediate exchange ($k_{\text{ex}} \leq \Delta\omega$, typically $k_{\text{ex}} > 10 \text{ s}^{-1}$) (Figure 88):

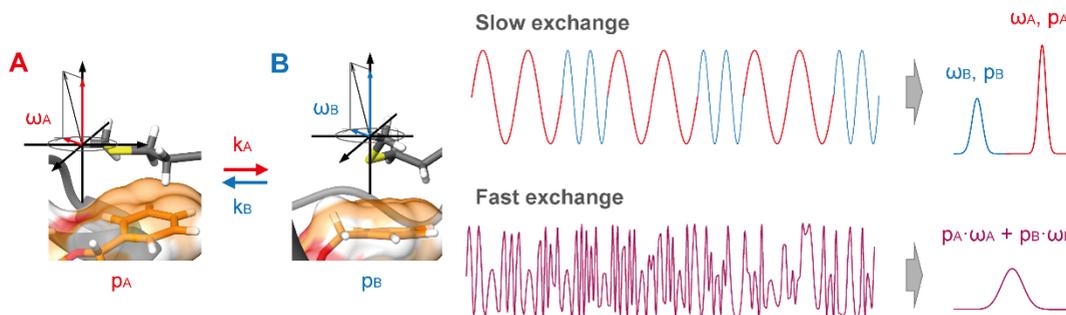


Figure 88: (Left) hypothetical ^1H atom of a methyl group in exchange between two sites A and B, with k_A and k_B exchange rate constants and p_A , p_B populations; the Larmor frequencies of each state are ω_A and ω_B . The figure at the right shows the ^1H resonance in the case of slow exchange ($k_{\text{ex}} < \Delta\omega$), which result in two separate peak resonances (red and blue). In the case of fast exchange ($k_{\text{ex}} > \Delta\omega$), the resonance is a complex composition of the A and B resonances and the exchange rate. As a consequence, a broadened peak centred in the population-weighted frequency is obtained.

Where k_{ex} is the exchange constant ($k_{\text{ex}} = k_{AB} + k_{BA}$), p_A ($p_B = 1 - p_A$) are the populations of the respective states, and ω_A and ω_B ($\omega_A - \omega_B = \Delta\omega$) are the corresponding resonance frequencies.

The CPMG sequence (an acronym of Carr-Purcell-Meiboom-Gill, the developers of this experiment in the 1950s [A2, A3]) allows the extraction of the exchange parameters. This experiment consists in a train of successive spin echoes (Figure 89 A) separated by a time lapse τ_{cp} . The more time between inversion pulses (τ_{cp}), the more probability of irreversible dephasing magnetization according to exchange, which results in a neat loss of signal.

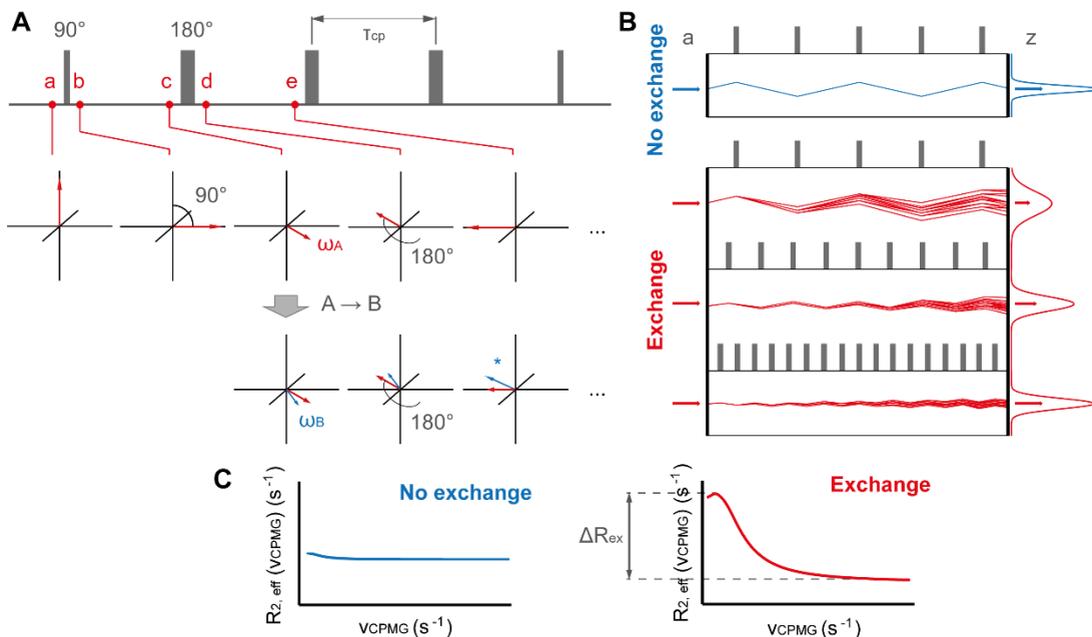


Figure 89: Measurement of the μ s-ms exchange between A and B states by CPMG-type RD experiments. A) CPMG element, formed by successive inversion pulses separated by a delay τ_{cp} . The initial magnetization (point a) is transferred to the transverse plane by a 90° pulse (point b). The magnetization evolves due to the chemical shift; if exchange to B occurs, the evolution takes place according to the chemical shift of B up to point c. After the 180° pulse (point d), the magnetization is inverted and the chemical shift evolution of A is completely refocused at a point e; contrarily, the magnetization of B is not refocused and is lost (asterisk). B) Constant-time CPMG experiments. In the absence of exchange (blue), the yield of the refocusing is ideal and no signal is lost. However, in the presence of exchange (red), the coherence is partially lost during the refocusing, resulting in broadened signals. Increasing the CPMG frequency improves the signal resolution and intensity. C) RD profiles obtained by CPMG experiments in the absence (blue) and in the presence of exchange (red). The extraction of ΔR_{ex} is indicated in the latter.

However, an improved version of the experiment (constant-time CPMG) is more suitable for the accurate measurement of R_2 . Constant-time CPMG varies the number of inversion pulses while keeping constant the period of time of the CPMG element (hence varying τ_{cp}).

The yield of refocusing the magnetization in exchange increases with the number of inversion pulses, leading to a much narrow and intense signal (Figure 89 B). $R_{2,eff}$ can be obtained from the signal intensity according to Equation 14:

$$R_{2,eff} \left(\frac{1}{2\tau_{CP}} \right) = \frac{-1}{T} \ln \left(\frac{I_{CPMG}}{I_0} \right) \quad \text{Equation 14}$$

In which $R_{2,eff}$ is the effective transverse relaxation rate including the exchange contribution to R_2 , T is the CPMG element length (in seconds), I_{CPMG} is the intensity of the signal of the experiment including the CPMG element, and I_0 is the intensity of the same signal in the experiment without CPMG train.

The plot of $R_{2,eff}$ as a function of the frequency of inversion pulses ($1/\tau_{cp}$) yields a decay curve, which is used for the extraction of exchange parameters (Figure 89 C). The amplitudes and frequencies of the exchange can be approximately quantified by the parameter ΔR_{ex} , which is calculated by the difference between $R_{2,eff}$ values at low and at fast pulsing limits (i.e. at low or at fast CPMG pulsing frequencies, Equation 15):

$$\Delta R_{ex} = R_{2,eff} \left(\frac{1}{\tau_{CP}} \rightarrow 0 \right) - R_{2,eff} \left(\frac{1}{\tau_{CP}} \rightarrow \infty \right) \quad \text{Equation 15}$$

If the populations in equilibrium are highly skewed, a relatively simple equation can be used for the calculation of the phenomenological R_2 as a function of τ_{cp} for all time scales [A4]:

$$R_{2,eff} \left(\frac{1}{\tau_{CP}} \right) = R_{2\infty} + \frac{p_A \cdot p_B \cdot \Delta\omega^2 \cdot k_{ex}}{k_{ex}^2 + \left(p_A^2 \cdot \Delta\omega^4 + \frac{144}{\tau_{CP}^4} \right)^{1/2}} \quad \text{Equation 16}$$

Finally, the fitting of RD data to two-site exchange model can be carried out with the use of the Carver-Richards equation, as described in the main text. This fitting allows the extraction of the numerical values of exchange parameters.

$R_{1\rho}$ RD experiments

Rotating frame relaxation dispersion experiments relies on the application of a transverse radiofrequency pulse which locks the magnetization out of the transverse plane (spin-lock). In analogy to the measurement of R_2 in the laboratory frame, $R_{1\rho}$ experiments measure the orthogonal component of the magnetization respect to the effective field in the rotating frame (see Figure 26 of the main text). In the presence of a spin-lock field, R_1 and $R_{2,eff}$

are related by the trigonometric relation shown in Equation 17 [A5]. If the spin-lock field is varied, the relative contributions of R_1 and $R_{2, eff}$ to $R_{1\rho}$ change and can be quantified.

$$R_{1\rho} = R_1 \cdot \cos^2\theta + (R_2 + R_{ex}) \cdot \sin^2\theta \quad \text{Equation 17 A}$$

$$R_{ex} = \frac{p_A \cdot p_B \cdot \Delta\omega^2 \cdot k_{ex}}{k_{ex}^2 + \omega_e^2} \quad \text{Equation 17 B}$$

Where ω is the frequency of the spin-lock field, ω_0 is the averaged chemical shift of the magnetization, ω_1 is the amplitude of the spin-lock field, and ω_e is the effective field in the rotating frame (calculated by $\omega_e = \sqrt{\omega_1^2 + \Delta\omega^2}$), being $\Delta\omega = \omega - \omega_0$). The tilt angle of ω_e is defined by $\tan\theta = \omega_1/\Delta\omega$.

Depending on the offset between ω_1 and the spin-lock frequency, the experiments can be classified as on-resonance $R_{1\rho}$ (if $\Delta\omega = 0$) or off-resonance (if $\Delta\omega \neq 0$, $\theta < 70^\circ$). In the case of the off-resonance $R_{1\rho}$ experiment (used in this work) the tilt angle is varied by modulating ω or ω_1 . Therefore $R_{1\rho}$ can be measured and plotted as a function of ω_e^2 . With these data, the exchange parameters can be extracted by applying Equation 17.

APPENDIX II: LOW RESOLUTION STRUCTURAL CHARACTERIZATION OF BIOMOLECULES BY SMALL-ANGLE X-RAY SCATTERING

The scattering of radiation is a physical phenomenon which was widely studied during the first decades of the 20th century. The application of the X-ray scattering to the study of non-crystalline particles was carried out by André Guinier at the late 1930s. However, the use of X-ray scattering to study molecules and biomolecules in solution did not take place until the 1970s. From this time, the application of X-ray scattering for the elucidation of the structural properties of biological macromolecules has grown constantly [A6].

Overall, two different methodologies based on the scattering of radiation. The most spread is small-angle X-ray scattering (SAXS), which is based on the scattering of collimated X-ray radiation at low angles by the electron envelope of a molecule; these scattering profiles contains low resolution structural information about the molecule. A closely related method is small-angle neutron scattering (SANS), in which neutron are scattered by the nuclei or

unpaired electrons of the molecule. For simplicity, only SAXS will be discussed in this summary.

Theory

Scattering is defined as the elastic interaction between radiation and a particle. Elastic interactions implies changes in the direction of the radiation, but not in its energy (i.e. wavelength). The scattering of radiation $E(t)$ by a single electron placed in the origin of coordinates is described by an observer placed at a point \mathbf{r} (Figure 90 A) by the following equation:

$$E(\mathbf{r}, t) = r_0 \cdot \frac{\sin \Psi}{r} \cdot \frac{\omega^2}{\omega_0^2 - \omega^2} \cdot E(t) \quad \text{Equation 18 A}$$

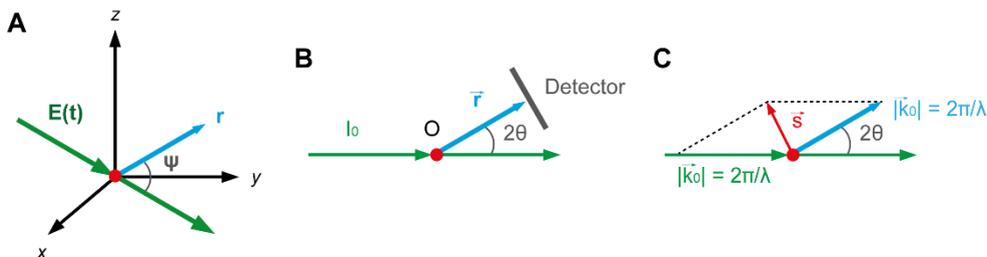


Figure 90: Scattering by a single electron. A) Scattered wave at a point \mathbf{r} by a single electron placed in the origin of coordinates (red). B) Scattering of a single electron recorded at a distance r as a function of the angle formed with the main beam (2θ). C) Momentum transfer \mathbf{s} of the scattered radiation by an electron, which is formed by the difference between the scattered and the incident wavevectors (\mathbf{k}_1 and \mathbf{k}_0 , respectively).

Where $E(\mathbf{r}, t)$ describes the scattered radiation as a function of the observer position, Ψ is the angle formed with the non-scattered radiation, and r_0 is the classical electron radius. The third term of Equation 18 A is a scaling factor between the incident frequency ω and the frequency of electron vibration ω_0 : in the case of SAXS, $\omega_0 \ll \omega$ and therefore there is no change in frequency. Hence, Equation 18 A can be simplified:

$$E(\mathbf{r}, t) = -r_0 \cdot \frac{\sin \Psi}{r} \cdot E(t) \quad \text{Equation 18 B}$$

A SAXS measurement is a record of the intensity of the scattered beam as a function of the angle 2θ (Figure 90 B). The intensity is extracted from the Thomson equation:

$$I(2\theta) = r_0^2 \cdot \frac{1 + \cos^2(2\theta)}{2} \cdot \frac{1}{r^2} \cdot I_0$$

Where I_0 is the intensity of the incident beam. However, this represents the simplest case of the scattering of radiation by a single electron. The scattering by assemblies of electrons is much more complex. For the study of these systems, it is necessary to introduce the momentum transfer between the incident and the scattered beam (Figure 90 C). The modulus of this vector can be expressed as q or s (the latter nomenclature has been used in this manuscript):

$$\mathbf{s} = \mathbf{k}_1 - \mathbf{k}_0$$

$$s = |\mathbf{k}_1 - \mathbf{k}_0| \quad |\mathbf{s}| = \frac{4\pi \sin \theta}{\lambda} \quad \text{Equation 19}$$

The momentum transfer allows to consider the effect of the phase difference when adding all the scattered waves from each individual electron. Hence, the scattering by a continuous electron density cloud $\rho(\mathbf{r})$ of N electrons can be described from the Fourier transform $F(\mathbf{s})$ of the electron density:

$$F(\mathbf{s}) = \int_V \rho(\mathbf{r}) \cdot e^{i \cdot \mathbf{r} \cdot \mathbf{s}} dV \quad \text{Equation 20 A}$$

However, Equation 20 A can be applied only if the distance between scatterers is fixed. Given that the position between scatterers is random in solution, only the addition of intensities is possible. In this case, the following equation is used:

$$i(s) = \langle I(\mathbf{s}) \rangle = \langle F(\mathbf{s}) \cdot F^*(\mathbf{s}) \rangle \quad \text{Equation 20 B}$$

The mathematical description of this formula is out of the scope of this manuscript, but Equation 20 B provides the spherical average of the scattered intensity of all the scatterers in solution. Notice that the introduction of a spherical averaging mathematical operation at this point reduces the three-dimensional scattering to a scalar function. In practice, this limitation is originated from the free molecular tumbling in solution, which avoids the extraction of three-dimensional data. This inconvenience entails a significant loss of information in SAXS.

Finally, with the use of the Fourier transform and other mathematic properties, it is possible to obtain the equation that describes the experimental scattering profiles as a function of s (Equation 20 C):

$$i(s) = 4\pi \int_0^{\infty} P(r) \cdot \frac{\sin(s \cdot r)}{s \cdot r} dr \quad \text{Equation 20 C}$$

The term $P(r)$ of this equation is the pair-wise distance distribution function, which accounts for the distance between all pairs of points of the molecule weighted by its electron densities (a more detailed description is explained later). This equation assumes ideal conditions of monodispersity (only one type of scatterer, e.g. single conformation) and absence of interparticle interactions.

SAXS scattering profiles of biomolecules in solution

Typically, X-ray radiation has wavelength from 1 to 2 Å, and the scattering phenomena occurs between 0.1° and 10°. A SAXS experiment consists in the measurement of the intensities of the scattered beam as a function of the angle. Usually, data is presented as the logarithm of the scattered intensity ($\log I(s)$) plotted against the momentum transfer (s); for an example see Figure 29 of the main text. In order to eliminate the scattering of the buffer, a blank experiment containing only the buffer is subtracted. The concentration is typically calculated by comparison with a known standard solution.

The experimental procedure of SAXS is not complicated but this technique requires sophisticated instrumentation and facilities. SAXS measurements needs highly stable and collimated radiation in order to increase the sensitivity and resolution. The most suitable radiation is the X-ray radiation obtained in synchrotron beamlines. Usually, the sample containers consists of tiny quartz capillaries.

Figure 91 A shows the scattering profiles of two systems of different size. A small protein (human lysozyme) yields a scattering profiles at large s values; in contrast, a big protein complex (20S proteasome) causes scattering at low s values. *A priori*, there is no size limit for the measurement of SAXS scattering profiles; only the properties of the sample (stability, absence of interparticle interactions and monodispersity) are the main determining factors for the quality of SAXS experiments (see below). In some cases, radiation damage of the sample by the incident beam might occur, which entails an important drawback.

Attractive interparticle interactions cause an increase of the scattering at low angles, which strongly interferes with the measurement of scattering profiles: usually, interparticle interactions can be detected by performing SAXS measurements at different concentrations (Figure 91 B). In turn, sample polydispersity yields average scattering profiles of all species

present in the solution, which results in a loss of resolution and poor quality of the data. However, a special comment must be done at this point: the averaging of scattering profiles resulting from species in dynamic equilibrium provides precious information about dynamic systems in equilibrium. The development of statistical and analytical tools allowed the deconvolution of individual scattering profiles from the SAXS curves of mixtures (the deconvolution of SAXS curves of mixtures by chemometric methods is described in the main text. For a reference, see Tauler *et al.* [A7].

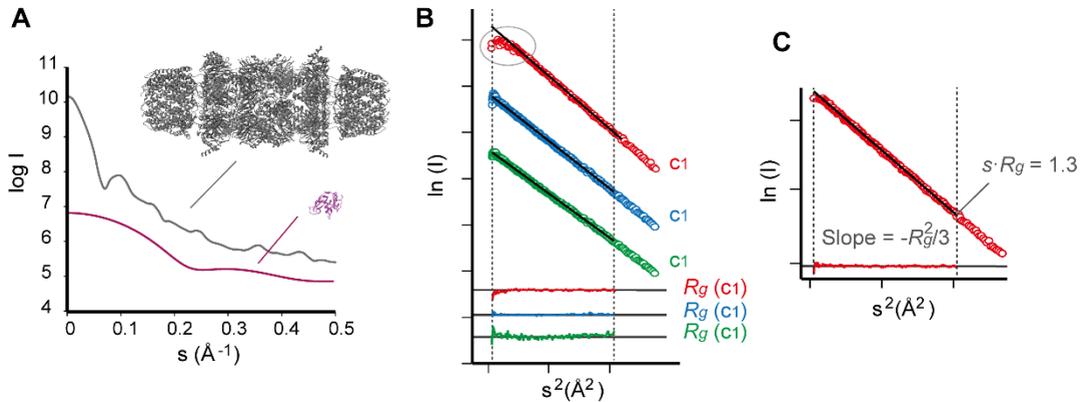


Figure 91: SAXS of biological macromolecules. A) Example of SAXS scattering profiles of two different systems: 20S proteasome (759 kDa, gray) and human lysozyme (15 kDa). B) Guinier plots of SAXS measurements at three increasing concentrations $c_1 < c_2 < c_3$. Attractive interparticle interactions can be seen at c_3 measurements at low angles (gray circle). The plots have been displaced vertically for better clarity. C) Example of a Guinier plot ($\ln(I)$ vs s^2). The Guinier region for biomolecules is typically limited by $s \cdot R_g < 1.3$.

Structural information from SAXS data

The most straightforward procedure which allows the extraction of structural properties directly from scattering profiles is the Guinier approximation. This approximation, developed by A. Guinier at late 1930s [A8], supposes that the scattering profiles at low angles (specifically the region delimited by $s \cdot R_g < 1.3$, named Guinier region) can be approximated to a Gaussian curve. Therefore:

$$I(s) = I(0) \cdot \exp\left(-\frac{1}{3} \cdot R_g^2 \cdot s^2\right) \quad \text{Equation 21 A}$$

By taking the logarithm, it is possible to obtain a linear expression between $\ln(I(s))$ and s^2 (Equation 21 B). The representation of $\ln(I(s))$ vs s^2 is the Guinier plot (Figure 91 C). This

plot allows to obtain the radius of gyration, R_g , directly from the slope. R_g is the spatial average of the radius of the particle, and therefore it gives information about the size of the system. In turn, y-axis intercept gives the forward scattering $I(0)$. This magnitude is related to the number of electrons of the particle. Together with the use of a protein standard of known MW and concentration, it is possible to extract the MW of the sample.

$$\ln[I(s)] = \ln[I(0)] - \frac{R_g^2}{3} \cdot s^2 \quad \text{Equation 21 B}$$

However, more sophisticated methods based on the indirect Fourier transform [A9] can be applied in Equation 20 C to extract more accurate R_g , $I(0)$ values, as well as the par-wise distribution function, $P(r)$. Indirect Fourier transform expresses $P(r)$ as a linear combination of orthogonal basic functions (φ_n), limiting r in a finite interval delimited by D_{\max} ($0 < r < D_{\max}$). The $P(r)$ distribution function is obtained by optimizing the coefficients c_n :

$$p(r) = \sum_{n=1}^M c_n \cdot \varphi_n(r)$$

Of interest, $P(r)$ reflect the shape of the particle: for instance, Gaussian-like functions correspond to globular proteins, while deviations from the Gaussian-like profile indicate elongated shapes; multimodal distribution is characteristic of multidomain proteins, in which the different distributions correspond to interdomain and intradomain distances (Figure 92 A). Choosing the appropriate D_{\max} is crucial for obtain a reliable $P(r)$ distribution.

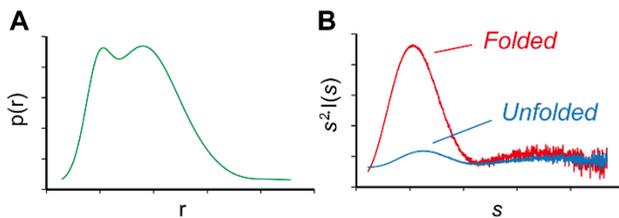


Figure 92: Analysis of SAXS data. A) Example of $P(r)$ distribution of a bidomain protein; the bidomain distribution reflect the interdomain distances. B) Kratky plot of hypothetical folded (red) and unfolded (blue) proteins.

Finally, the representation of $s^2 I(s)$ vs s (Kratky plot) is a particularly useful representation in the case of unfolded system or proteins with highly flexible unstructured regions [A10]. Kratky plot reflects the compactness of a protein: the presence of a prominent maximum at low angles indicates stable tertiary structures; on the contrary, a flat plateau is characteristic of unfolded systems (Figure 92 B).

Ab initio modelling and calculation of theoretical scattering profiles

The scattering profiles of biomolecules in solution provides low resolution information about the overall size and shape. Therefore, the application of SAXS for the structural elucidation of biomolecules requires the complementation with theoretical data. The extraction of three-dimensional structures with best agreement with experimental data is performed by *ab initio* computational rigid body modelling of structures [A6]; one of the most powerful modelling methods is bead-modelling [A11].

An alternative approach for the qualitative appreciation of tertiary structures can be carried out by comparing the experimental SAXS curves with theoretical scattering profiles calculated from high-resolution structures (for the calculation of theoretical scattering profiles, see next paragraph). Such structures are obtained by other biophysical tools, for instance, X-ray or NMR. However, the generation of high-resolution structures is possible by a combination of deterministic computational simulations (typically MD simulations) and SAXS data. This hybrid approach relies in the calculation of theoretical scattering profiles from simulated structures and the comparison with the experimental ones; this approach has been used in this thesis and will not be discussed here.

In the case of large biomolecules in solution, the most useful computational method for the calculation of theoretical scattering profiles is based on spherical harmonics. These mathematical functions describe the electron envelope of the molecule, and incorporate information of the excluded volume (the volume which is not accessible for the solvent) and the hydration layer.

Of interest, this approximation is used in CRY SOL, the most spread program for the calculation of theoretical scattering profiles (ATSAS data analysis software [A12]). The expression of the intensity according to the spherical harmonics approximation is given by:

$$I(s) = 2\pi^2 \sum_{l=0}^L \sum_{m=-l}^l |A_{lm}(s)|^2$$

This equation (expressed in spherical coordinates) is limited by L , which is the truncation value of the series that dictates the quality of the calculation. $A_{lm}(s)$ are the partial amplitudes associated to the spherical harmonics; therefore, the scattering profile of a given structure can be simply calculated as the combination of independent spherical harmonics.

Bibliography

- A1. Pervushin, K., et al., *Attenuated $T(2)$ relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution*. Proc Natl Acad Sci U S A, 1997. **94**(23): p. 12366-12371.
- A2. Carr, H.Y. and E.M. Purcell, *Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments*. Physical Review, 1954. **94**(3): p. 630-638.
- A3. Meiboom, S. and D. Gill, *Modified Spin-Echo Method for Measuring Nuclear Relaxation Times*. Review of Scientific Instruments, 1958. **29**(8): p. 688-691.
- A4. Ishima, R. and D. Torchia, *Estimating the time scale of chemical exchange of proteins from measurements of transverse relaxation rates in solution*. J Biomol NMR, 1999. **14**(4): p. 369-372.
- A5. Akke, M. and A.G. Palmer, *Monitoring Macromolecular Motions on Microsecond to Millisecond Time Scales by $R_{1\rho}$ - R_1 Constant Relaxation Time NMR Spectroscopy*. Journal of the American Chemical Society, 1996. **118**(4): p. 911-912.
- A6. Mertens, H.D.T. and D.I. Svergun, *Structural characterization of proteins and complexes using small-angle X-ray solution scattering*. Journal of Structural Biology, 2010. **172**(1): p. 128-141.
- A7. Tauler, R., *Multivariate curve resolution applied to second order data*. Chemometrics and Intelligent Laboratory Systems, 1995. **30**(1): p. 133-146.
- A8. Guinier, A., *La diffraction des rayons X aux tres petits angles: applications a l'etude de phenomenes ultramicroscopiques*. Annales de physique, 1939. **11.Sér. 12.1939**: p. 161-237.
- A9. Glatter, O., *A new method for the evaluation of small-angle scattering data*. Journal of Applied Crystallography, 1977. **10**(5): p. 415-421.
- A10. Doniach, S. and J. Lipfert, *Use of small angle X-ray scattering (SAXS) to characterize conformational states of functional RNAs*. Methods Enzymol, 2009. **469**: p. 237-51.
- A11. Chacón, P., et al., *Low-Resolution Structures of Proteins in Solution Retrieved from X-Ray Scattering with a Genetic Algorithm*. Biophysical Journal. **74**(6): p. 2760-2775.
- A12. Svergun, D., C. Barberato, and M.H.J. Koch, *CRY SOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates*. Journal of Applied Crystallography, 1995. **28**(6): p. 768-773.

APPENDIX III: PUBLICATIONS

Expert Opinion

1. Introduction
2. Peptide-like and peptidomimetic inhibitors and their therapeutic applications
3. Heteroaryl ketones as POP inhibitors
4. Alkaloids having POP inhibitory activity
5. Conclusion
6. Expert opinion

informa
healthcare

Low molecular weight inhibitors of Prolyl Oligopeptidase: a review of compounds patented from 2003 to 2010

Abraham López^{*†}, Teresa Tarragó^{*} & Ernest Giralt^{*†}

^{*}Institute for Research in Biomedicine, Barcelona Science Park, Barcelona, Spain and [†]Department of Organic Chemistry, University of Barcelona, Barcelona, Spain

Introduction: Prolyl Oligopeptidase (POP) is a serine peptidase that cleaves post-proline bonds in short peptides. Besides the direct hydrolytic regulation function over peptides, neuropeptides and peptide hormones, POP is probably involved in the regulation of the inositol pathway and participates in protein-protein interactions. Experimental data show that POP inhibitors have neuroprotective, anti-amnesic and cognition-enhancing properties. These compounds are considered therapeutic agents of interest for the treatment of cognitive deficits related to neuropsychiatric and neurodegenerative diseases, such as Alzheimer's disease and Parkinson's disease. Recent findings pointed to the involvement of POP in angiogenesis, although the exact mechanism is still under study.

Areas covered: This review comprises patents and patent applications involving POP inhibitors patented between 2003 and 2010, classified as peptidomimetics, heteroaryl ketones and alkaloids. The binding processes and the mechanisms of inhibition of these inhibitors are also discussed, together with their *in vivo* effects.

Expert opinion: The major part of the repertory of POP inhibitors derived from systematical modification of the canonical compound benzyloxycarbonyl-prolyl-proline (ZPP). Nevertheless, only two of them have progressed into the clinical trials. One possible reason for this failure is the lack of studies concerning pharmacodynamics, pharmacokinetics and toxicity, together with the absence of suitable animal models. Moreover, POP is still not a well-defined therapeutic target. Further studies are required for the elucidation of the biological role of POP and to validate the therapeutic action of inhibitors in cognitive processes. In contrast, the involvement of POP in protein-protein interactions together with the recent evidences in angiogenesis opens alternative approaches to the traditional active site-directed inhibitors, as well as new therapeutic applications.

Keywords: Alzheimer's disease, cognitive disorders, neurological disorders, Parkinson's disease, Prolyl Oligopeptidase, Prolyl Oligopeptidase inhibitors

Expert Opin. Ther. Patents [Early Online]

1. Introduction

1.1 Structure and mechanism

Prolyl Oligopeptidase (POP; EC 3.4.21.26) is a cytosolic serine endopeptidase (81 KDa) that hydrolyzes post-proline bonds of small peptides (≤ 30 residues). This monomeric protein was crystallized for the first time from porcine muscle by Fulop *et al.* [1]. It has an overall cylindrical shape constituted by two main domains. The catalytic domain is a typical α/β peptidase domain that is formed from residues

Article highlights.

- Prolyl Oligopeptidase (POP) is a serine peptidase that cleaves post-proline bonds of small peptides. This enzyme is ubiquitous in mammals, but a relatively high concentration is found in CNS tissues. In addition, the observation that most of the neuropeptides contains proline points to the involvement of this enzyme in neurological processes. However, its exact *in vivo* role is still unclear. The participation of POP in the regulation of inositol pathway, as well as specific protein-protein interactions, has recently been hypothesized. Experimentally, POP inhibitors might result in cognition-enhancing, neuroprotective and anti-amnesic properties, being accepted as therapeutic agents for the treatment of cognitive deficits. Nevertheless, other evidences for POP function far from cognitive processes are currently being explored. One of the most important subjects is the peripheral involvement in angiogenesis or in neoplastic tissues.
- Peptidomimetics are the most important family of POP inhibitors. One of the first inhibitors, ZPP, was taken as a canonical compound for the development of new peptidomimetics. These substances act as transition state or substrate analogues, depending on their chemical features. Thus, potent peptidomimetic POP inhibitors were designed and patented. In spite of representing a vast family of highly potent inhibitors, to our knowledge only two compounds have reached clinical trials.
- Four patents covered in this review belong to the peptidomimetic group. Only two are not classifiable as peptidomimetics, although the binding mode is not described. One compound is a heteroaryl ketone and the other is an alkaloid.
- Although a vast repertoire of POP inhibitors is available, *in vivo* pharmacodynamic, pharmacokinetic and toxicological studies are lacking. Therefore, suitable animal models are required for the correct development of these inhibitors. Potency is one of the most important parameters in the development of these inhibitors, although other parameters such as blood-brain barrier permeability are not always optimized. Given that the biological role of this enzyme is still unclear, the therapeutic value of POP should be pinpointed: thus, the search for improved POP inhibitors as cognitive enhancers or as antiangiogenic agents should be more directed. New drug discovery approaches in this direction need to be used to find innovative POP ligands or inhibitory scaffolds.

This box summarizes key points contained in the article.

peptidase and the propeller domains was responsible for the substrate entry [4]. Molecular dynamics and molecular framework studies were carried out to analyze the flexibility of the enzyme and to confirm this idea [5]. Recent studies isolated the crystal structure of the open form of bacterial POP, which collapses to the closed form in the presence of inhibitors [6], and an induced-fit mechanism for bacterial POP was also proposed. In contrast, other observations in mammalian POP are more consistent with the notion of equilibrium between different conformations in solution. This equilibrium may be altered by the presence of POP inhibitors [7,8].

1.2 Biological function of POP

Proline is a natural amino acid with unusual properties. First, the chemical behavior of the secondary amide bond differs from that of typical peptide bonds. Second, proline conformational freedom is constrained due to the cyclic structure. These features protect proline-containing peptides from general proteolytic degradation. Therefore, all organisms require specific proline peptidases to control pathways involving proline-containing peptides [9].

POP is present in bacteria, protozoa, plants and animals, as shown by cDNA sequencing [10-17]. It is highly distributed in all mammalian organs, including the CNS, where possibly it plays a crucial role in different neuronal systems [18-20]. POP location and activity inside the neurons depends on their developmental stage. Although the main location is in the cytoplasm and neuronal projections, POP tends to concentrate in the nucleus in young and old neurons, and shows an increased activity in the early developmental stages of neurons [21]. However, far away from this involvement in neuron cycle and differentiation, the function of this peptidase is suggested to be different in the case of having a peripheral distribution [19].

1.2.1 *In vitro* POP activity on neuropeptides and peptide hormones containing proline: first hypotheses about the biological role of POP

Most neuropeptides and peptide hormones contain proline and are POP substrates *in vitro* (Table 1) [22]. Therefore, it was hypothesized that POP has a direct role in neuropeptide catabolism and regulation. Altered POP activity in the serum of patients with mood disorders was initially reported by Reichelt *et al.* [23], and later by Maes *et al.* [24,25]. Later studies performed with rat brains showed that administration of POP inhibitor JTP-4819 (6, Table 2) increases the levels of substance P and arginine-vasopressin in the hippocampus and frontal cortex [26], leading to the release of acetylcholine. This finding pointed to the involvement of cholinergic neurotransmission in the mechanism of POP inhibition [27,28].

These observations established the involvement of POP in cognitive processes and neurological disorders, and highlighted this enzyme as a therapeutic target of interest [29,30]. Nevertheless, these first hypotheses do not explain how an intracellular enzyme degrades extracellular signaling peptides [31].

1 to 72 and 428 to 710. The catalytic triad is formed by Ser554, His680 and Asp641 [2]. The structural domain (from residues 73 to 427) is a seven-bladed β -propeller structure with nonconventional properties, which acts as an empty cylinder that restricts the size and the orientation of the substrate [1].

It was proposed that concerted movements of the propeller and the peptidase domain are crucial for enzymatic activity [3]. Later on, it was hypothesized that the opening between the

Table 1. *In vitro* POP substrates: some examples of neuropeptides and peptide hormones.

Substrate	Sequence
Arginine-vasopressin	CYFQNC PRG
Substance P	RP KPKQFFGLM
Oxytocin	CYIQNC PLGNH2
Thyrotropin releasing hormone: precursor peptide	KRQH PGQR
Neurotensin	pQLYENK PRRPVIL
Angiotensin II	DRVYI HPF
Bradykinin	RPPGF SPFR
Humanin	MAP RG SFCLLLLLSEIDL PV KRRA

Cleavage site is shown in bold [19].

POP: Prolyl Oligopeptidase.

Recent evidences pointed to the involvement of POP in cancer and tumor growth. The first hint of the function of POP in the endogenous production of the anti-fibrotic and angiogenic peptide Ac-SDKP from the precursor thymosin- β 4 was carried out by Cavasin *et al.* [32]. Further studies confirmed this hypothesis and suggested an auto regulated mechanism for this release, which was minimized with POP inhibition [33]. On the other hand, POP activity was studied in different neoplastic tissues, and an increased activity was seen in such tissues [34]. Therefore, this is the first outline that shows how modulating POP activity might be important in cancer treatment.

1.2.2 Involvement in the inositol triphosphate pathway

In other research lines, it was found that lithium-resistant mutants of *Dyctiostelium discoideum* do not have the homologous POP gene and present increased inositol triphosphate (IP₃) levels. POP inhibitors were found to have the same effect on IP₃ levels. These findings showed the involvement of POP in the IP₃ metabolic pathway and indicated a key role of this endopeptidase in the receptor-mediated signaling cascade of neuropeptides [35]. IP₃ is a second messenger involved in the transduction of neuropeptide signaling. In the presence of certain neuropeptides, the amount of IP₃ is increased and Ca²⁺ is released as a consequence of its interaction with a receptor in the membrane of the endoplasmic reticulum. This release of Ca²⁺ is known to play a crucial role in memory and learning [36].

It has been proposed that the mood-stabilizing drugs lithium, carbamazepine and valproic acid have a common effect on inositol signaling and that inositol depletion is the cellular basis for mood disorders [37]. Moreover, the finding that POP inhibitors reverse the effect of these drugs [38] also links POP to the phosphoinositide signaling pathway.

The involvement of POP in the IP₃ pathway was later supported by the observation of an inverse correlation between POP expression and IP₃ concentration in antisense human

astrogloma cell lines with reduced POP levels. A similar effect was detected by treating wild-type glioma cells with a POP inhibitor, suggesting an interaction between cytosolic POP and the signal transduction cascades of neuropeptides [39]. In addition, other evidence in agreement with this intracellular role of POP was presented by Brandt *et al.* Endogenous POP substrates were identified in the intracellular space, and some of such substrates were common with the IP₃ pathway [40].

1.2.3 POP and protein-protein interactions

To find new POP mechanisms of action, Di Daniel *et al.* [41] generated null mutant POP mice and found that their neurons presented alterations in growth cone dynamics. Moreover, both native POP and catalytically inactive POP mutant restored the wild-type phenotype of the neurons. This observation supported the involvement of POP in protein-protein interactions, independently of its catalytic action. Therefore, it was proposed that POP regulates synaptic functions independently of its peptidase activity. According to this, POP inhibitors may induce changes in the tertiary structure of the enzyme. These modifications modify its protein-protein interactions and may finally affect phosphoinositide synthesis. In the same study, Growth Associated Protein 43 (GAP43) was identified as a POP-interacting protein [41]. GAP43 is expressed in neurons and it shows a clear co-localization with POP in HeLa cells [42]. This finding is more coherent with the cytosolic location of POP.

GAP43 is involved in neuronal growth cone formation, axon guidance and synaptic plasticity. Interestingly, these mechanisms are more relevant in areas involved in mood control (specifically in the neocortex and the limbic system) [43]. Nevertheless, it has recently been shown that the interaction of POP and GAP43 is weak and transient [42].

Finally, α -synuclein was also found to interact with POP. This interaction was suggested to accelerate the aggregation of this intrinsically disordered protein, a process that is reversed by the inhibition of POP [44]. α -synuclein fibrils are found in Lewy bodies of Parkinson's patients, and participate in the pathology in this disease, in which POP may also be involved.

1.3 POP inhibitors have neuroprotective, anti-amnesic and cognition-enhancing properties

Several studies have been done in different impaired animal models to show the anti-amnesic effects of POP inhibitors, as well as their memory and cognition-enhancing properties.

Earlier studies revealed the anti-amnesic and memory-enhancing effect of the POP inhibitor JTP-4819 (6, Table 2) in the passive avoidance learning test in scopolamine-induced amnesia in rats [45]. Further studies showed memory enhancement in these animals. This effect was attributed to the activated cholinergic transmission caused by POP inhibition [27]. The same group performed additional behavioral and pharmacological studies in middle cerebral occlusion-damaged rats and demonstrated the improvement in spatial memory [46], while

Table 2. Inhibitors and patents developed up to 2003.

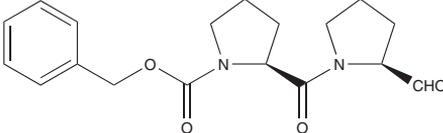
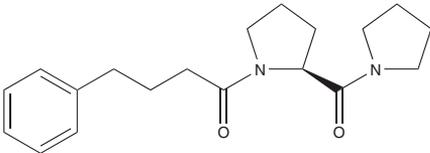
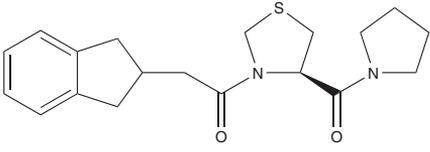
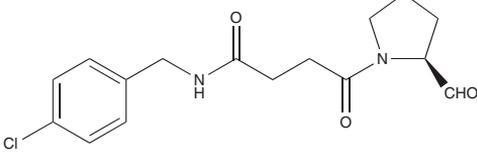
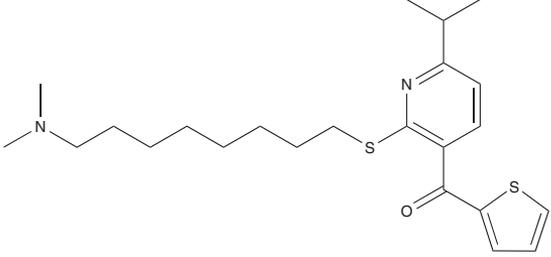
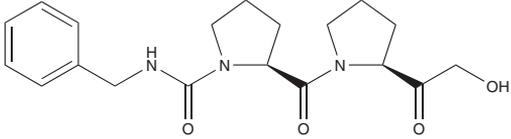
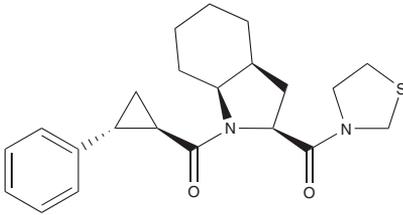
Number	Name/internal code	Research group/Co.	Chemical structure and potency
1	Z-prolyl-prolinal (ZPP)	Nagasaki University	 <p>$K_i = 0.21 - 5 \cdot 10^{-9}$ M (bovine brain) [49]</p>
2	SUAM-1221	Suntory	 <p>$K_i = 0.97 \cdot 10^{-9}$ M (human) [75]</p>
3	Z-321	Zeria Pharmaceuticals	 <p>$IC_{50} = 10^{-9}$ M (canine) [57]</p>
4	ONO-1603	Ono Pharmaceuticals	 <p>$K_i = 12 \cdot 10^{-9}$ M (rat cortex) [86]</p>
5	Y-29794	Yoshitomi Pharmaceuticals	 <p>$K_i = 0.95 \cdot 10^{-9}$ M (rat brain) [89]</p>

Table 2. Inhibitors and patents developed up to 2003 (continued).

Number	Name/internal code	Research group/Co.	Chemical structure and potency
6	JTP-4819	Japan Tobacco	 <p>$IC_{50} = 0.83 \cdot 10^{-9}$ M (rat) [24]</p>
7	S-17092	Servier Research Institute	 <p>$K_i = 1.5 \cdot 10^{-9}$ M (human) [95]</p>

treatment of naturally aged rats with JTP-4819 led to an improvement in spatial memory [27]. Nevertheless, later studies demonstrated the high complexity of this mechanism in memory functions. The low reproducible results, poor dose-dependency and non-predictable behavior of POP inhibition were incoherent with a simple direct effect [30,47,48].

The effects of POP inhibition were also tested in the treatment of cognitive deficits related to neurodegenerative processes. Parkinson's disease was generated in monkeys by treatment with 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP), a neurotoxin that produces depletion of substance P. Subsequent treatment with S-17092 (7), a potent POP inhibitor, increased the performance of cognitive tasks [49]. As mentioned earlier, it was found that POP inhibition prevents the oligomerization of α -synuclein *ex vivo* [45]. In the case of Alzheimer's disease, several *in vivo* experiments in animal models showed that POP inhibition led to neuroprotective and cognition-enhancing effects [50,51]. Neuroprotective effects were originally observed by Katsube's group, when cortical and cerebellar granule cells were prevented from age-induced apoptosis by treatment with the POP inhibitor ONO-1603 (4) [52].

Finally, clinical trials with POP inhibitors in the treatment of cognitive deficits have been performed only in a few cases. The most reported compound was S-17092 (7) [53]. It showed cognition-enhancing properties in healthy elderly subjects and a clear dose-dependency; moreover, no adverse effects were detected. Later studies suggested additional slight mood-stabilizing properties for this compound [54]. The pharmacokinetics and toxicology of Z-321 (3) were also tested in humans [55].

2. Peptide-like and peptidomimetic inhibitors and their therapeutic applications

This section, which provides a brief introduction to the studies related to the search for new POP inhibitors, is necessary for a correct understanding of the present review; however, it does not intend to provide a complete view of the all studies performed. For more detailed information, see DeNanteuil *et al.* [56] and Lawandi *et al.* [57].

2.1 The pharmacophore of peptidomimetic inhibitors

Peptidomimetic inhibitors are designed by systematic modification of the reference compound benzyloxycarbonyl-prolyl-prolinal (ZPP) (1, Table 2) [58]. This molecule acts as a transition state analogue, as shown by NMR [59] and X-ray [1] studies, and it was taken as a starting point in the ligand-based drug design for subsequent peptidomimetic inhibitors.

The first pharmacophore was described by Yoshimoto *et al.* [60], but later SAR studies disclosed new shape features in order to achieve higher selectivity over other prolyl peptidase family proteases, mainly fibroblast activating protein (FAP) and dipeptidyl peptidase IV (DPP-IV) [57,61]. In addition, the influence of the electrostatic environment of POP active site in the binding processes was also studied by Szeltner *et al.* [62].

A careful comparison of all the known inhibitors allowed the description of an optimal pharmacophore to achieve potent human POP inhibition and selectivity [57]. The pharmacophore is summarized in Figure 1, following the nomenclature proposed by Berger and Schechler [63,64].

Bulky hydrophobic groups are well fitted at S3. In contrast, charged groups result in a decrease in potency. A hydrogen bond is generated between the bridge carbonyl group and the indole ring of Trp595 [1].

Structural requirements at P2 are poor. Proline, proline derivatives [58,60,65-69] or natural α -amino acids, both neutral and charged, are fitted [70-72]. Moreover, a wide variety of cyclic or bicyclic scaffolds are compatible [73-75] while open chains can also be fitted. The carbonyl group between this position and the subsequent P1 is crucial for the formation of a hydrogen bond with Arg643 [1], and S stereochemistry is preferred over planar or R stereochemistry [67,76].

At S1, proline, pyrrolidine and pyrrolidine derivatives [73] or thiazolidine [60,65] rings are preferably fitted. α - β Unsaturated carbocycles or heteroaryl structures are well adapted in this environment, by mimicking the planar P2-P1 amide bond, and favoring π -stacking interactions with Trp 595 [77].

Finally, covalent inhibitors mainly share electrophilic functional groups at P1'. These electrophilic groups react with the nucleophilic hydroxyl group of Ser554 [1]. Therefore, the formed hemiacetalic bond often results in an increase of the potency [59]. High stereospecificity to S configuration is required in these substituents in order to form the chemical bond and to achieve the convenient hydrogen bonding network [1]. Aldehydes were early on assumed to form a hemiacetal [78], acting in a reversible manner. Acetals [79] or ketones were also found to form hemiacetal or hemiketal bonds, respectively. In the latter case, α -hydroxymethyl ketones [27], α -keto heterocycles [80,81] or polyfluoroalkyl ketones [82,83] favor the formation of the chemical bond by activating the carbonyl group. Moreover, heterocycles containing an sp^2 nitrogen atom in the beta position result in an increase in potency, as a consequence of the formation of a hydrogen bond between this atom and the histidine of the catalytic triad [81]. Nitriles [68,84] were proposed to form imino ethers with the hydroxyl group of the serine, and were also studied as reversible covalent inhibitors [85,86].

In contrast, chloromethyl ketones behave as irreversible inhibitors [87]. Similarly, 3-acylisoxazoles also react with the serine, leading to ring opening and the formation of the corresponding ester [88].

2.2 POP inhibitors: evolution and patents up to 2006

After several studies on substrate analogue-type inhibitors, the first POP inhibitors were obtained by Yoshimoto at Nagasaki University. ZPP (1, Table 2), a compound with potent inhibitory effect on mammalian POP, was discovered early on [58]. The aldehyde functionality was found to be required for correct binding.

One of the first companies to show interest in new POP inhibitors was Suntory. The first studies performed

by this company revealed the relevance of P3 for the selectivity of mammalian POP, since short arylalkyl or aryloxyalkyl chains were necessary for correct binding [70,71,89]. SUAM-1221 (2) was one of the most active products patented by this company.

A series of structural modifications of SUAM-1221 was produced by Zeria Pharmaceuticals in the following years, thus leading to a family of analogues with improved potency and stability. Other cyclic and bicyclic systems were included at N-terminal position, as well as modifications in the proline or pyrrolidine rings, as exemplified by Z-321 (3) [67,68,90-92].

Since 1986, ONO Pharmaceuticals has patented several agents with POP inhibitory activity [93-97]. In the most potent agent, ONO-1603 (4), the central proline ring was substituted by a succinic moiety, thus giving a more flexible structure compared with other typical prolylpyrrolidine derivatives.

In 1993, Yoshitomi Pharmaceuticals carried out research projects to develop new non-peptide POP inhibitors. Therefore, Y-29794 (5), a pyridine derivative, was developed. Further studies showed high selectivity and permeability to the blood-brain barrier, as well as high accumulation in the brain cortex and hippocampus [98-100]. It is noteworthy that this is one of the few studies that moved away from the systematic modification of dipeptide structures, and it disclosed an innovative structure with drug-like properties.

Japan Tobacco, in collaboration with Yoshitomi Pharmaceuticals, addressed the dipeptide structure and disclosed JTP-4819 (6) as a potent and specific POP inhibitor [27,28]. This family of proline derivatives and their pharmaceutical use [101] was covered in successive patents [102-104].

Further studies into novel POP inhibitors and their therapeutic applications were done at the Servier Research Institute. Modifications of the dipeptide structure were made in order to achieve higher potency, specificity, and stability of the compounds [73]. The best candidate was S-17092 (7) [105], which has a perhydroindole ring in the central P2 position and a rigidified (2-phenylcyclopropyl)carbonyl chain. Moreover, this compound showed high potency, selectivity and cell permeability [106] and reached clinical trials [53].

2.3 Peptidomimetic inhibitors patented by Orion Corp

Orion Corp. filed two patents and several extensions between 2003 and 2006, protecting two series of POP inhibitors. According to different previous studies [46,107], they claimed the effects of POP inhibition on the fixation of memory, as well as on the prevention of β -amyloid protein neurotoxicity mediated by substance P. Hence, these compounds were protected to improve learning and mnemonic functions in the treatment of neurodegenerative diseases, such as Alzheimer's disease or senile dementia.

The patents consist of two peptidomimetics with IC_{50} values in the nanomolar range, high stability and good selectivity against the dipeptidyl peptidase family.

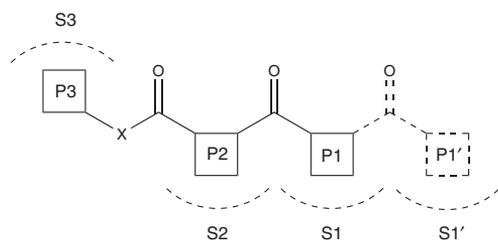


Figure 1. General scheme for peptidomimetic inhibitors, showing the complementary binding sites of the enzyme.

2.3.1 WO2004060862 (A2 and A3) [108]

This patent protects a family of inhibitors that shares typical alkyl or alkenyl chains at P3, including aryl or heteroaryl moieties or substitutions such as carboxyl, acyl or amino groups. Nitrogen or oxygen atoms are claimed at X position. The compounds disclosed in this patent are characterized on the basis of the proline mimetic fitted at S2 site (Figure 2). The first proline mimetic is substituted at carbon 5, within a wide variety of possibilities, including alkyl or alkenyl chains (also substituted with halogens, alkoxy, amino, carboxyl or acyl groups). The *t*-butyl group slightly decreased the potency (compare the effect in 13 versus the methyl group in 11, Figure 3) but conveniently increased the lipophilicity to achieve blood-brain barrier permeability. For example, logP values for compound 13 is 3.3 units, compared with 1.8 in the case of the similar compound SUAM-1221 (2, Table 2). The most suitable stereochemistry of *t*-butyl group was R [109].

The *t*-butyl group in position 5 of P2 proline ring allowed the presence of a Boc group at P3 portion. The authors suggested that bulky groups in P2 (such as *t*-butyl in R configuration) stabilized the *cis* conformation of the P3-P2 amide bond, thus resulting in a conformationally rigid compound. This effect is stronger over the more rigid Boc group than over the flexible phenylbutanoyl chain; in this latter case a methyl group is preferred (compare 11, 13 and 14) [109].

A cyclopent-2-enecarbonyl group, a second innovative proline mimetic, is also described in P2. Although other compounds with a cyclopent-1-enecarbonyl moiety were previously patented, the potency of the inhibitors was, in general, decreased [110,111]. Cyclopent-2-enecarbonyl, in contrast, was found to be a suitable substitution for the proline ring, as shown by the IC₅₀ values (Figure 3). This novel modification led to an improvement in lipophilicity (e.g., logP value increased from 0.2 units for the classical compound JTP-4819 to 0.6 for the patented compound 8), but have the inconvenience of including a Michael acceptor site [57,76]. This site shows affinity for nucleophiles, thus leading to low stability, undesirable reactivity with other substrates or poor selectivity. This disadvantage is minimized in the case of α/β unsaturated amides, because of having an

electron-donating nitrogen atom in X position. The stereochemistry was not controlled, but R stereochemistry is proposed to be more suitable [76].

Optional substituents at P1 are also reported, typically cyano, halogens, amino, hydroxyl or alkoxy groups; carboxyl, acyl groups and their derivatives are other protected functionalities. Lower alkyl, carbocyclic and heterocyclic scaffolds are extended to P1' binding site. Hydrogen, cyano, formyl and hydroxyacetyl groups are exemplified in this patent (Figure 3).

Finally, later studies published by the same group showed that the *t*-butyl group at P2, as well as the cyano and hydroxyacetyl group at P1, led to an increase in the dissociation half-life of the complex. This promising aspect increased the therapeutic value of these agents [86].

2.3.2 WO03004468 (A1) [112]

Orion Corp. reported a family of compounds that is structurally characterized by two *L*-prolyl-pyrrolidine moieties connected symmetrically with a short dicarboxylic acid linker at P3 (Figure 4). The authors found a lead compound with POP inhibitory potency in the low nanomolar range [comparable to that of SUAM-1221 (2) or JTP-4819 (6)] but with increased lipophilicity, thereby making it more susceptible to crossing the blood-brain barrier [113].

Although centrosymmetrical or almost centrosymmetrical, this class of compounds was expected to bind in the same way as the typical inhibitors. S1, S2, S3, S4 and S5 cavities are filled depending on the chemical features of the inhibitor (Figure 4). This is, to our knowledge, the first study to propose the new S4 and S5 binding sites [113].

The most important central dicarboxylic acids placed in P3 are succinic, glutaric (16, in Figure 5), 3,3-dimethylglutaric (17), adipic (18), terephthalic and isophthalic (19). The latter showed the highest potency and lipophilicity and was extensively studied [84-86,113,114]. Nevertheless, Q covers other arylene, cycloalkylene, cycloalkenylene groups. Alkylene and alkenylene chains are also included (which can incorporate the aforementioned cyclic scaffolds as a chain member). In addition, the patent protects a wide variety of substituents in this position.

P2 and P4 are usually proline moieties, but modifications were carried out by including other acyclic amino acids. Some examples are glycine, alanine, phenylalanine and methionine (19). All of these led to a decrease in the potency regardless of whether P4, P2 or both were substituted (compare, for instance, 19 and 20) [113]. Independently, in P4, heterocycles containing amino, alkyl amino, aryl amino or arylalkyl amino groups are also covered (which can be conveniently substituted). In those cases, P5 is empty (18).

In positions P1 and P5, the patent claims pyrrolidine (20) or cyclopentane ring (21). This avoidance of the N atom led to a loss of potency (as it can be seen by comparing IC₅₀ values of 20 with 21) but to a significant increase in the logP value, from -0.2 to 2.7, respectively [113]. The patent is also extended to a wide variety of compounds, mainly substituted alkyl or alkenyl chains, carbocycles or heterocycles, such as

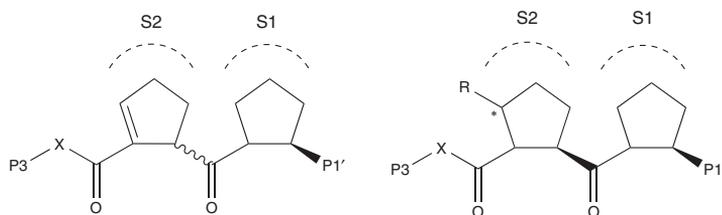


Figure 2. General structures for compounds claimed in patent WO2004060862 (A2 and A3).

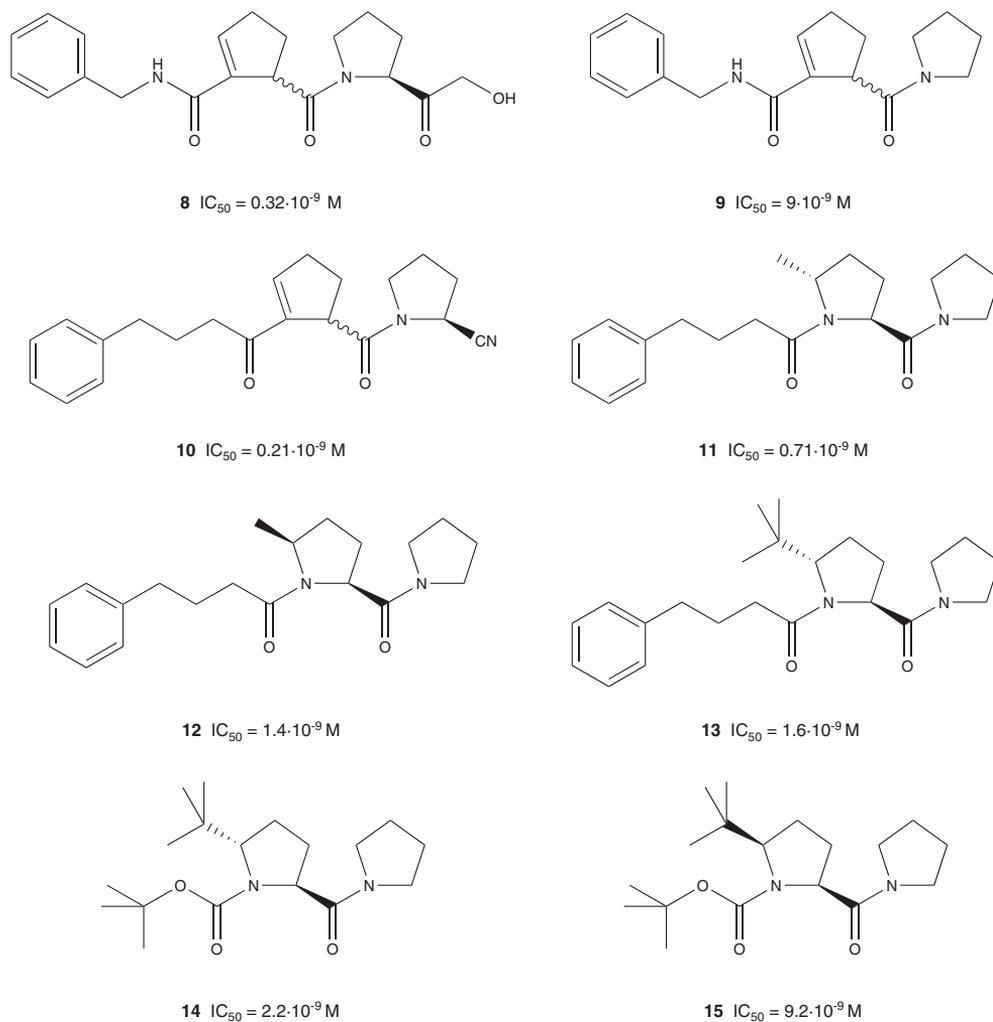


Figure 3. Compounds exemplified in patent WO2004060862 (A2 and A3) showing potency (pig brain extract).

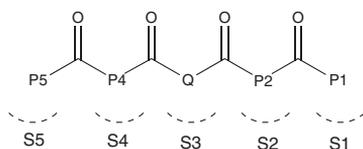


Figure 4. General structure for compounds protected in patent WO03004468 (A1).

tetrazoles. These cyclic structures are also included as a chain member of the mentioned alkyl or alkenyl chains.

More specifically, in P1, a wide variety of substituents in the pyrrolidine ring are included. Cyano (22) and hydroxyacetyl (23) groups in the S configuration were found to be the most suitable by interacting with the serine of the active site (potency in compound 20 was increased from $26 \cdot 10^{-9}$ M to $1.5 \cdot 10^{-9}$ and $0.4 \cdot 10^{-9}$ M in the compounds 21 and 22, respectively) [84]. Nevertheless, the hydroxyacetyl group led to a marked decrease in logP value [113]. Moreover, as mentioned before, the cyano and hydroxyacetyl groups also increased the dissociation half-life of the complex [86].

P5 was not found to contribute to the binding [as demonstrated by the wide range of groups allowed (16, 18, 24)], the cycloalkyl and phenyl groups being the best candidates [113].

2.4 Peptidomimetic inhibitors of Probiodrug AG

Probiodrug AG filed POP inhibitor patents for the treatment of several kinds of amnesia and for neurological disorders, especially mild cognitive impairment, Alzheimer's, Parkinson's and Huntington's diseases, Down syndrome, as well as other pathogenic psychotic conditions. Patents were also extended as inhibitors of microbial growth, in the treatment of post-operative resistance to transplanted organs or tissues, and as anticancer agents.

The three patents described in this review include the co-administration of POP inactivators with inhibitors of other enzymes, agonists or antagonists of certain receptors, as well as amyloid peptide deposition inhibitors. Within these groups, glutamyl cyclase, dipeptidyl peptidases, monoamine oxidases, beta and gamma secretases and ACE are mentioned, among others.

2.4.1 WO2006058720 (A2 and A3) [115]

This patent comprises a family of compounds derived directly from the structure of the dipeptide inhibitors. The different moieties of these compounds correspond with P3, P2, P1 and P1' positions (as shown in Figure 6), and several chemical and structural modifications are included to achieve a higher complementarity with the peptide pharmacophore. Therefore, high potencies in the nanomolar range were achieved.

Typical hydrophobic groups are placed at P3, mainly alkyl, alkenyl or alkynyl chains, optionally ended with a carbocycle or heterocycle. Amino acids are also included as P3 groups.

These moieties are linked to the peptide backbone through amide, carbamate or urea bonds, depending on X atom (compounds are exemplified in Table 3).

A linear amino acid moiety is placed at P2, without any kind of restriction on the side chain. In the case of acidic amino acids (Glu and Asp), the side chain is optionally linked to an additional amino acid through a peptide bond.

A huge variety of groups are protected at the S1 binding site. In general, these groups consist of cyclic and bicyclic structures that mimic a proline ring, but non-cyclic amino acid side chains are also allowed. The structures summarized in Figure 6 are selected for being characteristic of the P1 binding site. Although bicyclic or fused cyclic structures have been reported at P2 [73,74], this is the first time, to our knowledge, that bicyclic structures have been included at P1. R is alkyl, alkenyl, alkynyl chains, carbocycles or heterocycles. The patent covers heteroatoms, halogens and a wide variety of substituents in this position as well. Unusual reactive, acid or charged functional groups are covered as R'. For instance, boronic or sulfonic acid groups, sulfonamide or sulfone groups, phosphonic acid or esters, phosphoric acid or esters and thiols or trifluoromethyl groups are mentioned. In Z position, substituted carbon or nitrogen atoms, oxygen or sulfur atoms, and sulfoxide or sulfone groups are claimed.

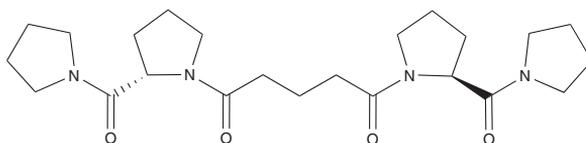
Finally, heterocyclic structures placed at P1' are claimed. The increase in the potency, by analogy with other heterocyclic compounds, may be attributed to the carbonyl activating function of the heterocyclic moiety, as well as to the formation of a stabilizing hydrogen bond network with the histidine residue of the catalytic triad, as previously described [80,81].

2.4.2 WO2009007415 (A2 and A3) [116]

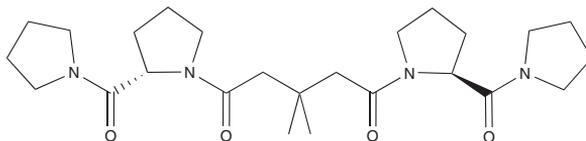
This patent protects compounds mimicking dipeptide inhibitors (Figure 7). Nevertheless, the mode of binding of these compounds is not described in the patent or in the literature. A proposed mode is described here, although assigning P1 and P1' sites is not easily predictable (as mentioned earlier, the S1 site accepts a wide variety of structures further from pyrrolidine, typically unsaturated carbocycles, heterocycles and aromatic heterocycles) [60,65,73,77]. By similarity with the peptide pharmacophore, and comparing the features with a number of extensively studied compounds, a hypothetical mode of binding could be inferred. W may fit nicely at the S2 binding site, as a wide variety of structures that are allowed at P2. The central amino acid residue should accommodate at the S1 site, although there is a highly specific site for proline or pyrrolidine rings [57]. Finally, Z may interact with the S1' site.

Typical alkyl or alkenyl chains completed with aryl or heteroaryl structures are claimed at W position (some compounds are exemplified in Figure 7). X comprises oxygen, sulfur or nitrogen atoms, thus yielding carbamate, thio-carbamate or urea bond, respectively. Moreover, the patent covers the nitrogen atom of the amino acid residue in

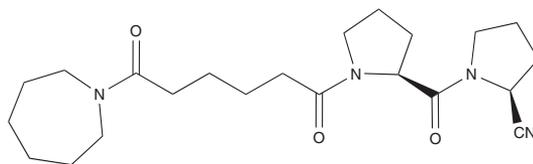
Low molecular weight inhibitors of Prolyl Oligopeptidase: a review of compounds patented from 2003 to 2010



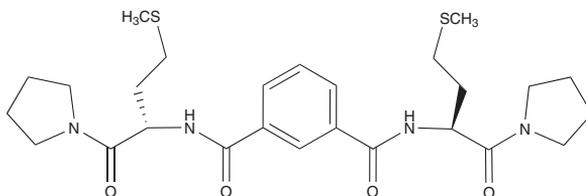
16 $IC_{50} = 48 \cdot 10^{-9} M$



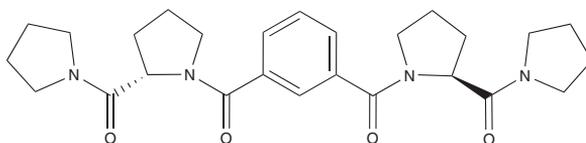
17 $IC_{50} = 13 \cdot 10^{-9} M$



18 $IC_{50} = 0.8 \cdot 10^{-9} M$

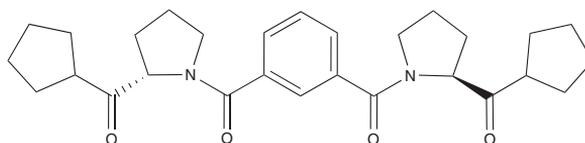


19 $IC_{50} = 110 \cdot 10^{-9} M$

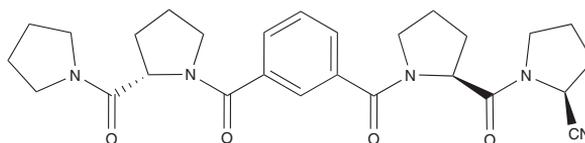


20 $IC_{50} = 26 \cdot 10^{-9} M$

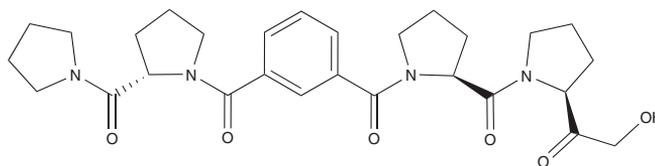
Figure 5. Some examples of compounds claimed in patent WO03004468 (A1), showing the potency (pig brain extract).



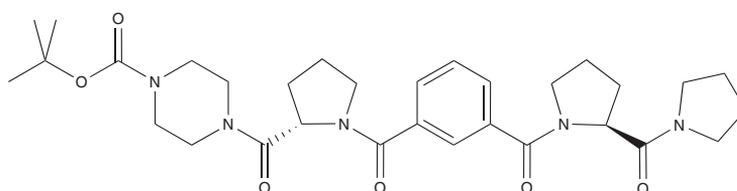
21 $IC_{50} = 78 \cdot 10^{-9}$ M



22 $IC_{50} = 1.5 \cdot 10^{-9}$ M



23 $IC_{50} = 0.4 \cdot 10^{-9}$ M



24 $IC_{50} = 170 \cdot 10^{-9}$ M

Figure 5. Some examples of compounds claimed in patent WO03004468 (A1), showing the potency (pig brain extract) (continued).

its free form, protected as *N*-methyl amide, or forming a proline ring. These modifications are possible because this amide group is not essential in the formation of hydrogen bonds [57].

The patent claims a huge variety of residues and derivatives at hypothesized P1. In general, these are natural polar or hydrophobic non-charged amino acid residues (compounds exemplified in Figure 7). Modifications comprise nitro (31), halogen, alkyl, hydroxyl and hydroxyalkyl groups as substituents on phenyl rings of phenylalanine; alkyl chains substituting certain positions in indole tryptophan rings,

and substituting hydroxyl or thiol groups in serine, threonine and cysteine residues. In the case of Asn and Gln, side chains are claimed as secondary or tertiary alkylamide derivatives, as well as esterified or acylated derivatives in the case of acidic or basic side chains, respectively.

Finally, at Z position, aryl and heteroaryl structures are included. Alkyl, alkenyl, alkynyl or carbocycles, alkoxy and other ethers, sulfone, acyl, carboxyl and amide groups and halogens, amino and cyano groups are protected as substituents of these aryl and heteroaryl structures.

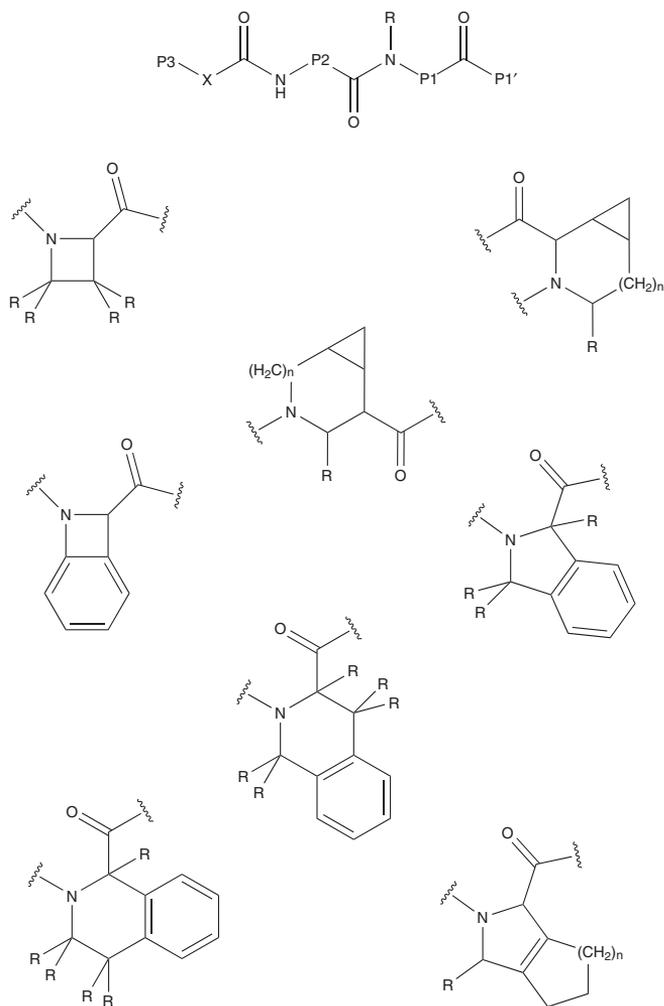


Figure 6. General structure for compounds protected by patent number WO2006058720 (A2 and A3). Cyclic and bicyclic scaffolds claimed at P1 position are joined below.

To our knowledge, this would be one of the few cases described where the proline or pyrrolidine ring in P1 position is not present. In contrast with peptidomimetic compounds WO2004060862, WO03004468 and WO2006058720, the potency of the present series of compounds fall into the micromolar range, probably due to a partial fit with POP binding sites. Although the mechanism and the mode of binding are not clear, the aromatic heterocycles at Z position may be involved in the activation of the adjacent carbonyl group, thus favoring the reactivity with Ser544. Moreover, in most of the compounds exemplified there is an sp^2 nitrogen

atom in the α position of these heterocycles, which can result in a possible hydrogen bond with His680 at the catalytic site [80,81].

3. Heteroaryl ketones as POP inhibitors

3.1 WO2006120104 (A1) [117], Probiodrug AG

The compounds covered by this patent are characterized by a general heteroaryl ketone core. *A priori*, the chemical and structural features are not attributable to any peptide scaffold; in contrast, they correspond more to a non-peptide inhibitor,

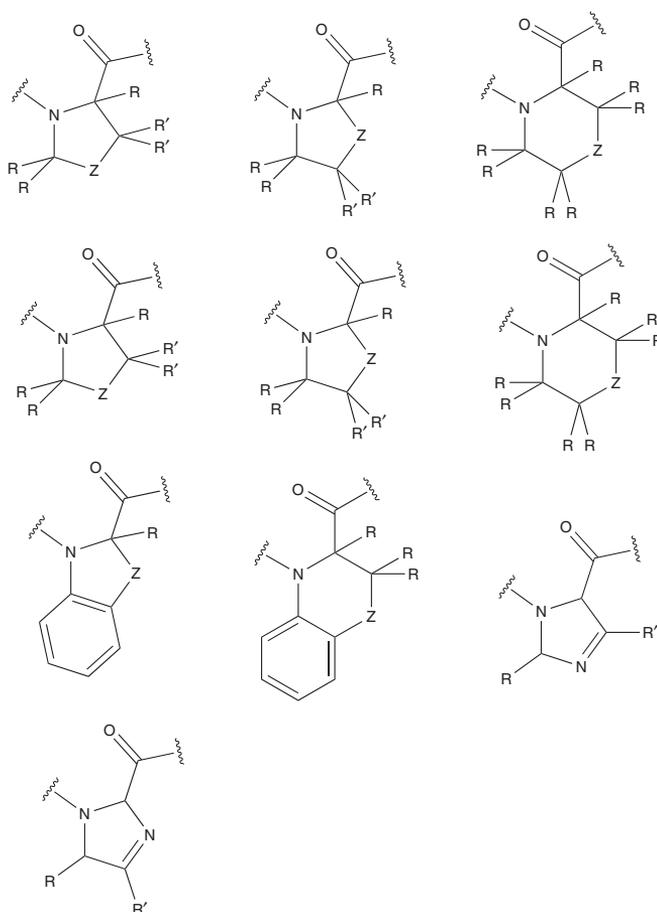


Figure 6. General structure for compounds protected by patent number WO2006058720 (A2 and A3) (continued). Cyclic and bicyclic scaffolds claimed at P1 position are joined below.

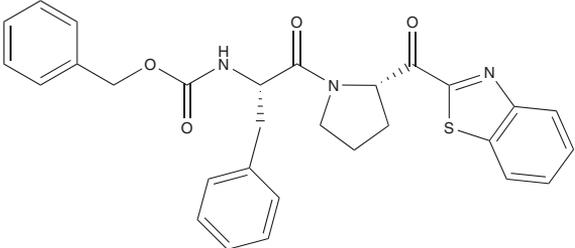
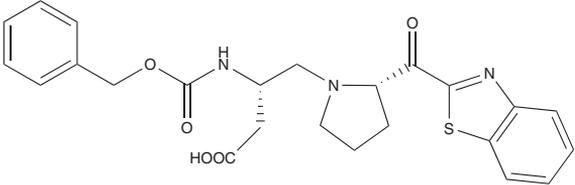
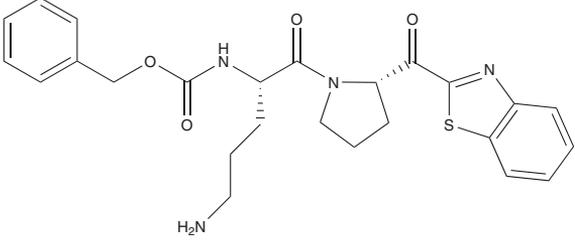
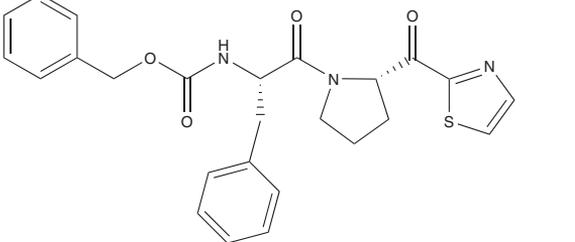
which shares a similar heteroaryl ketone moiety with Y-29794 [100], a previously described non-peptide compound (Figure 8A). The authors did not offer any details about the mode of binding of this family of compounds, thus making it impossible to predict whether they bind to the active site or an alternative binding site [57]. Moreover, the IC_{50} values move up to the low micromolar range, a higher value compared with those of well-defined peptidomimetic compounds (patent numbers WO2004060862, WO03004468 and WO2006058720, with IC_{50} values in the nanomolar range).

The general heteroaryl ketone scaffold for compounds protected by this patent is disclosed in Figure 8B. In A position, furyl and thienyl moieties (which can be substituted by halogens, cyano, nitro, hydroxyl or alkoxy groups, or lower

alkyl chains) are claimed. In the case of supposing a similar mode of binding than the peptidomimetic compounds, this moiety might interact with the S1 binding site. The carbonyl group might keep the necessary hydrogen bond network, and the planar sp^2 carbon in a position from carbonyl group may mimic the N atom of the classical pyrrolidine moiety. Heteroaryl structure is supposed to enhance the π -stacking interaction with Trp 595 [77].

Higher structural diversity is found in position B. Here, cycloalkyl structures up to seven-membered rings, aryl and heteroaryl structures summarized later are protected. Y (from 1 to 5) is a carbon or a nitrogen atom independently. R, from 1 to 10, is absent (in the case of having a nitrogen atom), or can be comprised between hydrogen, hydrocarbon

Table 3. Examples claimed in patent WO2006058720 (A2 and A3) and their potency.

Number	Compound exemplified	IC ₅₀ (recombinant human POP)
25		6·10 ⁻⁹ M
26		17·10 ⁻⁹ M
27		2·10 ⁻⁹ M
28		15·10 ⁻⁹ M

POP: Prolyl Oligopeptidase.

chains, carbocycles or heterocycles (charged species formed if Y is a nitrogen atom also taken into account). Z can be an oxygen, a sulfur, a nitrogen or a carbon atom, in these last two cases with the same structural claims as given earlier. X is a nitrogen or an oxygen atom. In the case of nitrogen, the patent protects R as hydrogen, alkyl, alkenyl or alkynyl chains, as well as carbocyclic or heterocyclic structures.

Finally, as C group, alkyl, alkenyl or alkynyl chain is claimed. Optional substituted carbocyclic or heterocyclic structures are also included as a chain member. Compounds exemplified are summarized in Figure 9.

4. Alkaloids having POP inhibitory activity

4.1 WO2008077978 (A1) [118], University of Barcelona

Natural products are now reemerging as an important source of active compounds and new scaffolds for future drugs [119]. With this in mind, the Giralt research group focused their efforts on the discovery of POP inhibitors using the pool of natural products employed in traditional Chinese medicine (TCM).

Novel ¹⁹F NMR-based screening techniques using fluorinated substrates as probes were developed [120]. By this

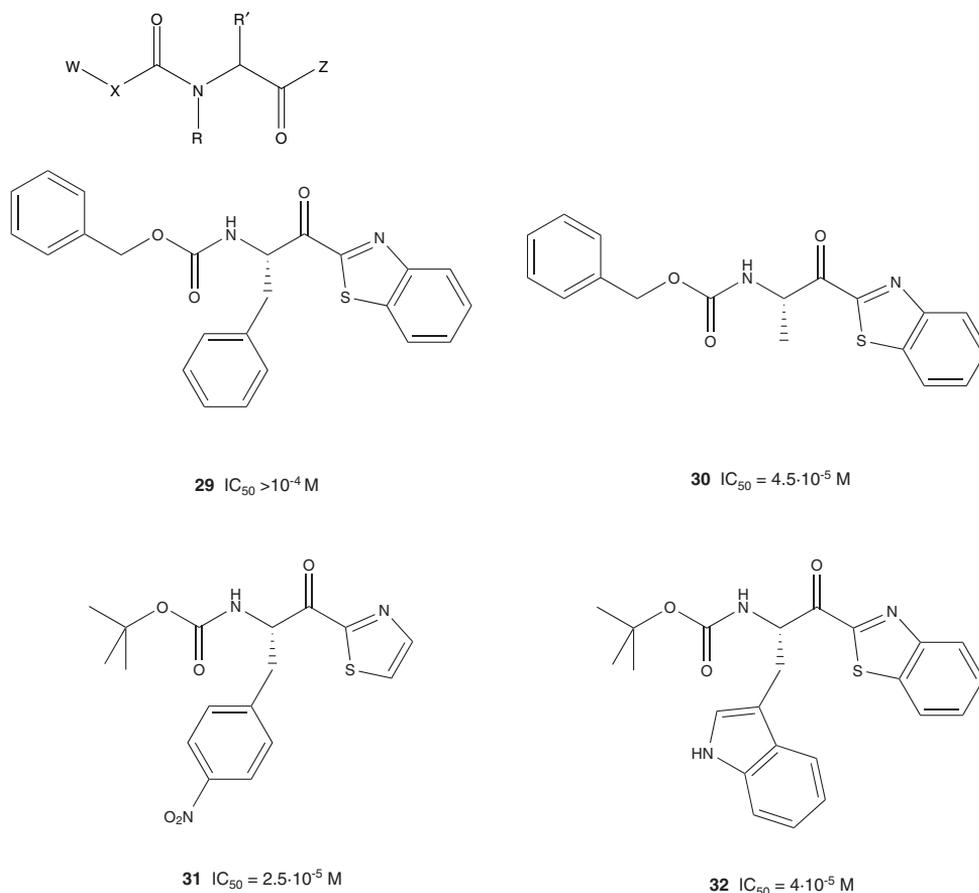


Figure 7. General structure for patent WO2009007415 (A2 and A3). Some examples are shown below (inhibitory activity was found in recombinant human POP).

POP: Prolyl Oligopeptidase.

means, the effect of various plant extracts of TCM against POP enzymatic activity was analyzed. POP inhibitory activity was found in *Rhizoma coptidis* aqueous extract [121,122]. Subsequent fractionation and analysis of the fractions allowed the isolation of berberine as one of the most active products [123]. This compound is included in a family of alkaloids that are isomers of 5,6-dihydro-9,10-dimethoxybenzo[g]-1,3-benzodioxolo[5,6-a]quinazolinium cation (Figure 10).

Berberine was patented as POP inhibitor for the treatment of bipolar affective disorder in humans, since it may be useful for the treatment of the cognitive deficits of these patients. The administration of berberine is safe in humans and has already reached clinical trials in two cases [124,125].

Regarding the activity of berberine, a probably non-covalent compound, a poor IC_{50} value of $145 \cdot 10^{-6}$ M

in recombinant human POP was found [123]. In spite of this, other advantages were found for this compound: permeability to the blood-brain barrier, accumulation in the hippocampus [126] and penetration inside neurons [127]. Finally, berberine was selective for POP and did not show inhibitory activity against other two proteases from different families, the aspartyl protease HIV-1 and the serine protease trypsin [123].

5. Conclusion

POP inhibition is correlated with an improvement of cognitive processes, as well as with neuroprotective and anti-amnesic effects. However, the exact mechanism of action of this intracellular enzyme is still unclear. Other

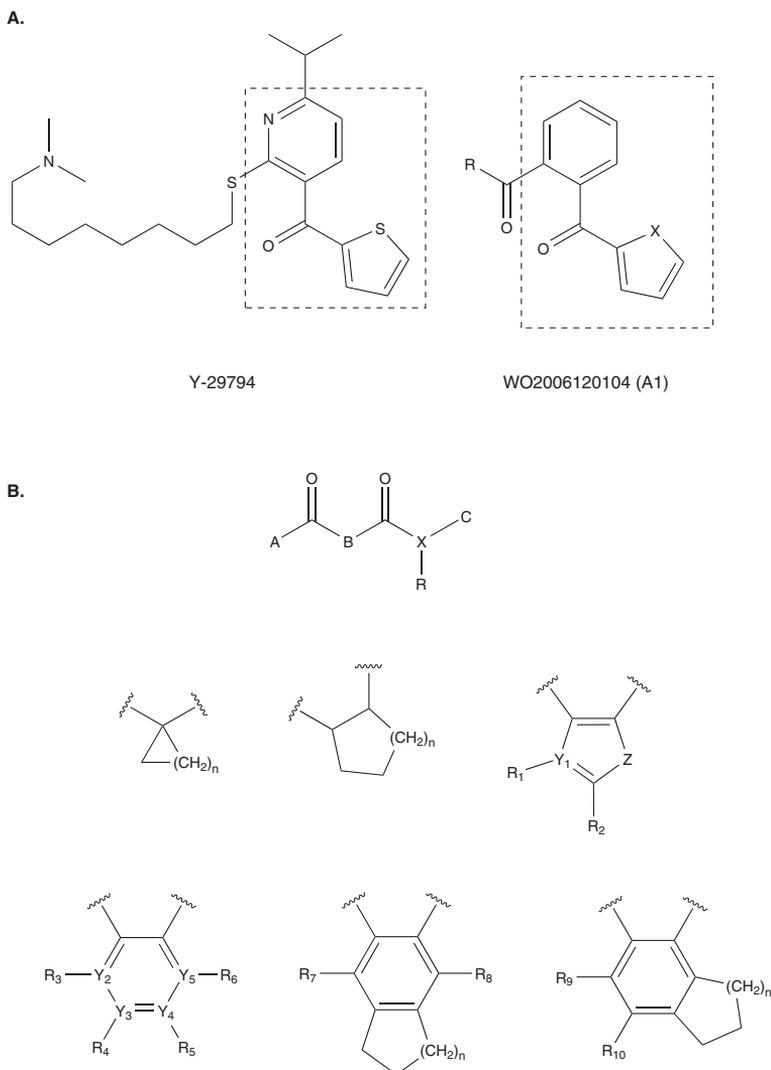


Figure 8. Patent number WO2006120104 (A1). A. Comparison between non-peptide inhibitor Y-29794 and the heteroaryl ketone moiety claimed. **B.** General formula for patent WO2006120104 (A1) and structures disclosed at B position.

involvements in protein aggregation and angiogenic processes are currently being hypothesized. The most important requirements for POP inhibitors, in addition to potency, are selectivity against other proteases and the capacity to cross the blood-brain barrier to penetrate and distribute inside neurons. From 2003 to 2010, compounds patented as POP inhibitors consist mainly of peptidomimetics, which have been widely studied. Nevertheless, heteroaryl ketone inhibitors and alkaloids are also represented. Finally, no

pharmacological studies for these patented compounds have been performed to date.

6. Expert opinion

When the involvement of POP in neuropeptide processing was proposed, parallel work began to address inhibitors and to determine the biological role of this protein. Several *in vivo* studies were performed in animals to evaluate the

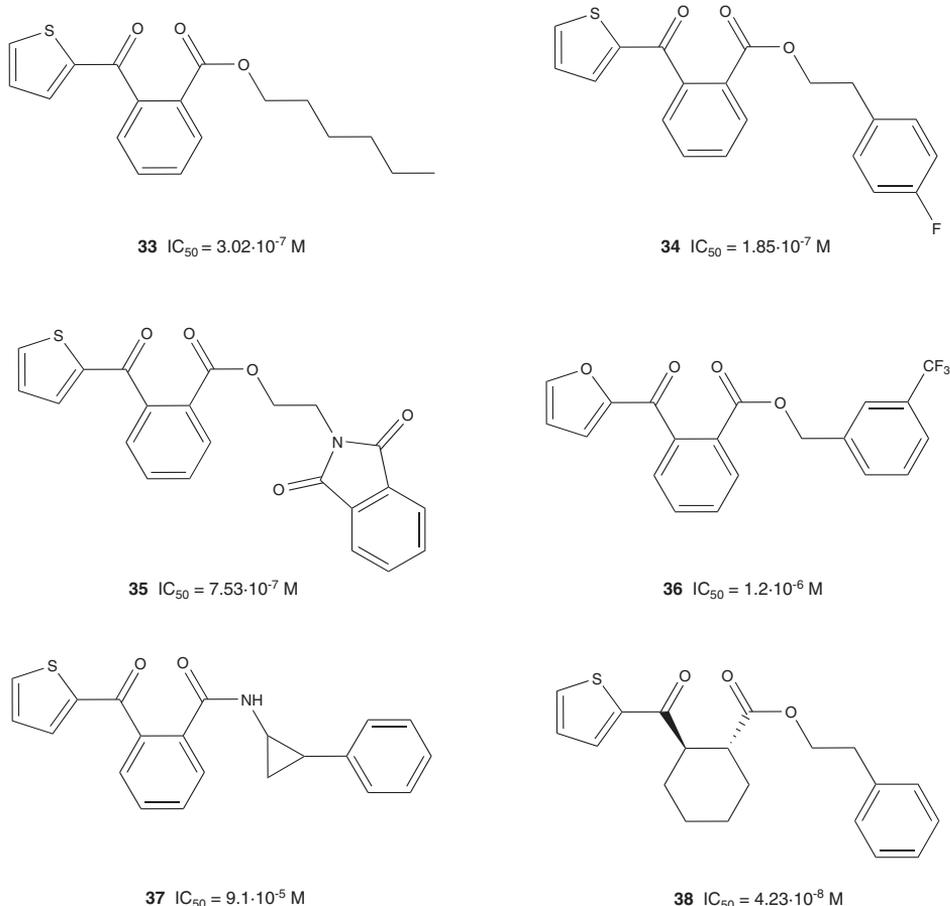


Figure 9. A selection of compounds claimed in patent WO2006120104 (A1). Inhibitory activity was measured in recombinant human POP.

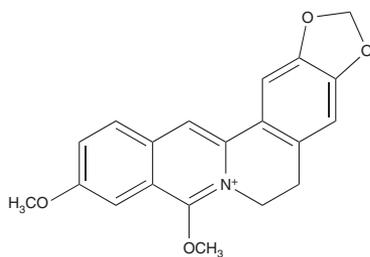
POP: Prolyl Oligopeptidase.

anti-amnesic, cognition-enhancing and neuroprotective properties of these inhibitors. Subsequently, POP was recognized as an important target for the treatment of cognitive disorders in neurodegenerative processes. Research groups, both from academy and industry, devoted efforts to the development of new POP inhibitors and, therefore, a wide variety of compounds were patented.

The research of these inhibitors has been focused mainly on the search for new peptidomimetics derived from the canonical compound ZPP. Patents disclosed in this review show how chemical modifications improve the potency, the specificity in the targeting and sometimes modulate their lipophilicity [108,112,115,116]. Surprisingly, only two

patents move away from the ZPP-derived peptidomimetic structures. In the first case (WO2006120104 (A1) [117]), heteroaryl ketone derivatives were disclosed as suitable scaffolds for POP inhibition. In the second, an alkaloid (WO2008077978 (A1) [118]) was claimed as POP inhibitor. In addition, there are a few published examples of innovative inhibitory scaffolds [98-100,128,129].

To the best of our knowledge, only two compounds [S-17092 (7) and Z-321 (3)] have reached clinical trials in humans. S-17092 improved cognitive tasks, presented good pharmacokinetic properties and did not show toxicity [53]. The reasons why S-17092 and Z-321 [54] have not progressed into further phases of clinical trials are not clear. It is



39 $IC_{50} = 145 \cdot 10^{-6} M$

Figure 10. Chemical structure of berberine and its IC_{50} value (measured in recombinant human POP).

POP: Prolyl Oligopeptidase.

also noteworthy that none of the companies or groups cited in the text continued with the clinical trials of their patented compounds.

Given the vast repertoire of inhibitors, it would be logical to ask why POP inhibitors have not yet found their application in the treatment of memory disorders in humans. The reason is not potency, as many of the POP inhibitors (such as S-17092 or JTP-4819) have K_i s in the low nanomolar range [30]. With few exceptions of *in vivo* animal [85,86] and human [53,54,58] pharmacodynamic studies, the lack of data concerning toxicology, blood-brain barrier permeability and mainly human pharmacokinetic and pharmacodynamic studies hinders any assessment. Moreover, a plethora of animal models has been used to evaluate the *in vivo* effects of inhibitors [27,46,44,47-52]; however, it seems that the control parameters are not properly defined. Hence, choosing the optimal cognition-enhancing test model may be one of the key points in the future. The development of improved animal models to evaluate cognition enhancers is clearly needed.

In addition, the lack of knowledge about the biological function of POP is currently one of the main drawbacks in the development of new inhibitors of this protein. Advancing in this direction would define and establish the importance

of POP as a therapeutic target. In the same way, the *in vivo* pharmacological properties of POP inhibitors should be confirmed as an effect of correct POP targeting, and not as a consequence of the unspecific activity of these compounds over other pathways. On one hand, the finding that POP is involved in protein-protein interactions [41] implies that its activity may be modulated by structurally different compounds other than the traditional active site-directed POP inhibitors. Therefore, by simply binding to certain POP surfaces, novel inhibitors could modify POP-protein interactions. This new possibility to regulate POP activity independently of its proteolytic activity has not been explored to date and should not be neglected. On the other hand, the involvement of POP in cancer and tumor processes should be taken into account. The first hints of novel roles of POP in angiogenesis are currently being explored, but deeper studies in this direction are needed. Therefore, it will provide a new therapeutic applications for this peptidase and new applications for the already known POP inhibitors.

The lack of innovative structures should encourage scientists to undertake new drug discovery approaches. Fragment-based, high-throughput or natural product screenings are potentially useful methodologies to discover new scaffolds. In this regard, the discovery of a new pool of ligands and compounds with POP inhibitory activity will facilitate the study and understanding of the function of this enzyme *in vivo* and its recognition events.

Declaration of interest

The authors declare no conflict of interest. Mr. Abraham López is employed by the Department of Organic Chemistry, University of Barcelona, and has received funding from the Institute for Research in Biomedicine, the Instituto de Salud Carlos III and the Ministerio de Ciencia e Innovación. Dr. Teresa Tarragó has received funding from the Institute for Research in Biomedicine, and Fundació La Marató de TV3. Dr. Ernest Giral is employed by the Department of Organic Chemistry, University of Barcelona, and has received funding from the Institute for Research in Biomedicine, MCYT-FEDER (Bio2008-00799), and the Generalitat de Catalunya (XRB and Grup Consolidat).

Bibliography

- Fulop V, Bocskei Z, Polgar L. Prolyl oligopeptidase: an unusual beta-propeller domain regulates proteolysis. *Cell* 1998;94(2):161-70
- Polgar L. Structural relationship between lipases and peptidases of the prolyl oligopeptidase family. *FEBS Lett* 1992;311(3):281-4
- Szeltner Z, Rea D, Juhasz T, et al. Concerted structural changes in the peptidase and the propeller domains of prolyl oligopeptidase are required for substrate binding. *J Mol Biol* 2004;340(3):627-37
- Juhasz T, Szeltner Z, Fulop V, Polgar L. Unclosed beta-propellers display stable structures: implications for substrate access to the active site of prolyl oligopeptidase. *J Mol Biol* 2005;346(3):907-17
- Fuxreiter M, Magyar C, Juhasz T, et al. Flexibility of prolyl oligopeptidase: molecular dynamics and molecular framework analysis of the potential substrate pathways. *Proteins* 2005;60(3):504-12
- Li M, Chen C, Davies DR, Chiu TK. Induced-fit Mechanism for Prolyl Endopeptidase. *J Biol Chem* 2010;285(28):21487-95
- Tarrago T, Martin-Benito J, Sabido E, et al. A new side opening on prolyl oligopeptidase revealed by electron microscopy. *FEBS Lett* 2009;583:3344-48
- Tarrago T, Claesen B, Kichik N, et al. A cost-effective labeling strategy for the NMR study of large proteins: selective ¹⁵N-labeling of the tryptophan side chains of prolyl oligopeptidase. *ChemBioChem* 2009;10:2736-9
- Cunningham DF, O'Connor B. Proline specific peptidases. *Biochim Biophys Acta* 1997;1343(2):160-86
- Rennex D, Hemmings BA, Hofsteenge J, Stone SR. cDNA cloning of porcine brain prolyl endopeptidase and identification of the active-site seryl residue. *Biochemistry* 1991;30(8):2195-203
- Yoshimoto T, Kanatani A, Shimoda T, et al. Prolyl endopeptidase from *Flavobacterium meningosepticum*: cloning and sequencing of the enzyme gene. *J Biochem* 1991;110(6):873-8
- Chevallier S, Goeltz P, Thibault P, et al. Characterization of a prolyl endopeptidase from *Flavobacterium meningosepticum*. Complete sequence and localization of the active-site serine. *J Biol Chem* 1992;267(12):8192-9
- Kabashima T, Fujii M, Meng Y, et al. Prolyl endopeptidase from *Sphingomonas capsulata*: isolation and characterization of the enzyme and nucleotide sequence of the gene. *Arch Biochem Biophys* 1998;358(1):141-8
- Ishino T, Ohtsuki S, Homma K, Natori S. cDNA cloning of mouse prolyl endopeptidase and its involvement in DNA synthesis by Swiss 3T3 cells. *J Biochem* 1998;123(3):540-5
- Robinson KA, Bartley DA, Robb FT, Schreier HJ. A gene from the hyperthermophile *Pyrococcus furiosus* whose deduced product is homologous to members of the prolyl oligopeptidase family of proteases. *Gene* 1995;152(1):103-6
- Shirasawa Y, Osawa T, Hirashima A. Molecular cloning and characterization of prolyl endopeptidase from human T cells. *J Biochem* 1994;115(4):724-9
- Vanhoof G, Goossens F, Hendriks L, et al. Cloning and sequence analysis of the gene encoding human lymphocyte prolyl endopeptidase. *Gene* 1994;149(2):363-6
- Wilk S. Prolyl endopeptidase. *Life Sci* 1983;33(22):2149-57
- Myohanen TT, Garcia-Horsman JA, Tenorio-Laranga J, Mannisto PT. Issues about the physiological functions of prolyl oligopeptidase based on its discordant spatial association with substrates and inconsistencies among mRNA, protein levels, and enzymatic activity. *J Histochem Cytochem* 2009;57(9):831-48
- Irazusta J, Larrinaga G, Gonzalez-Maeso J, et al. Distribution of prolyl endopeptidase activities in rat and human brain. *Neurochem Int* 2002;40(4):337-45
- Moreno-Baylach MJ, Felipe V, Mannisto PT, Garcia-Horsman JA. Expression and traffic of cellular prolyl oligopeptidase are regulated during cerebellar granule cell differentiation, maturation, and aging. *Neuroscience* 2008;156(3):580-5
- Garcia-Horsman JA, Mannisto PT, Venalainen JI. On the role of prolyl oligopeptidase in health and disease. *Neuropeptides* 2007;41(1):1-24
- Reichelt KL, Hole K, Hamberger A, et al. Biologically active peptide-containing fractions in schizophrenia and childhood autism. *Adv Biochem Psychopharmacol* 1981;28:627-43
- Maes M, Goossens F, Scharpe S, et al. Lower serum prolyl endopeptidase enzyme activity in major depression: further evidence that peptidases play a role in the pathophysiology of depression. *Biol Psychiatry* 1994;35(8):545-52
- Maes M, Goossens F, Scharpe S, et al. Alterations in plasma prolyl endopeptidase activity in depression, mania, and schizophrenia: effects of antidepressants, mood stabilizers, and antipsychotic drugs. *Psychiatry Res* 1995;58(3):217-25
- Toide K, Okamiya K, Iwamoto Y, Kato T. Effect of a novel prolyl endopeptidase inhibitor, JTP-4819, on prolyl endopeptidase activity and substance P- and arginine-vasopressin-like immunoreactivity in the brains of aged rats. *J Neurochem* 1995;65(1):234-40
- Toide K, Iwamoto Y, Fujiwara T, Abe H. JTP-4819: a novel prolyl endopeptidase inhibitor with potential as a cognitive enhancer. *J Pharmacol Exp Ther* 1995;274(3):1370-8
- Toide K, Shinoda M, Fujiwara T, Iwamoto Y. Effect of a novel prolyl endopeptidase inhibitor, JTP-4819, on spatial memory and central cholinergic neurons in aged rats. *Pharmacol Biochem Behav* 1997;56(3):427-34
- Rosenblum JS, Kozarich JW. Prolyl peptidases: a serine protease subfamily with high potential for drug discovery. *Curr Opin Chem Biol* 2003;7(4):496-504
- Mannisto PT, Venalainen J, Jalkanen A, Garcia-Horsman JA. Prolyl oligopeptidase: a potential target for the treatment of cognitive disorders. *Drug News Perspect* 2007;20(5):293-305
- Brandt I, Scharpe S, Lambeir AM. Suggested functions for prolyl

Low molecular weight inhibitors of Prolyl Oligopeptidase: a review of compounds patented from 2003 to 2010

- oligopeptidase: a puzzling paradox. *Clin Chim Acta* 2007;377(1-2):50-61
32. Cavin MA, Rhaleb NE, Yang XP, Carretero OA. Prolyl oligopeptidase is involved in release of the antifibrotic peptide Ac-SDKP. *Hypertension* 2004;43(5):1140-5
 33. Myohanen T, Tenorio-Laranga J, Jokinen B, et al. Prolyl oligopeptidase induces angiogenesis both in vitro and in vivo in a novel regulatory manner. *Br J Pharmacol* 2010. Doi: 10.1111/j.1476-5381.2010.01146.x
 34. Larinaga G, Perez I, Blanco L, et al. Increased prolyl endopeptidase activity in human neoplasia. *Regul Pept* 2010;163(1-3):102-6
 35. Williams RS, Eames M, Ryves WJ, et al. Loss of a prolyl oligopeptidase confers resistance to lithium by elevation of inositol (1,4,5) trisphosphate. *EMBO J* 1999;18(10):2734-45
 36. Komatsu Y. GABAB receptors, monoamine receptors, and postsynaptic inositol trisphosphate-induced Ca²⁺ release are involved in the induction of long-term potentiation at visual cortical inhibitory synapses. *J Neurosci* 1996;16(20):6342-52
 37. Harwood AJ, Agam G. Search for a common mechanism of mood stabilizers. *Biochem Pharmacol* 2003;66(2):179-89
 38. Williams RS, Cheng L, Mudge AW, Harwood AJ. A common mechanism of action for three mood-stabilizing drugs. *Nature* 2002;417(6886):292-5
 39. Schulz I, Gerhartz B, Neubauer A, et al. Modulation of inositol 1,4,5-trisphosphate concentration by prolyl endopeptidase inhibition. *Eur J Biochem* 2002;269(23):5813-20
 40. Brandt I, De Vriendt K, Devreese B, et al. Search for substrates for prolyl oligopeptidase in porcine brain. *Peptides* 2005;26(12):2536-46
 41. Di Daniel E, Glover CP, Grot E, et al. Prolyl oligopeptidase binds to GAP-43 and functions without its peptidase activity. *Mol Cell Neurosci* 2009;41(3):373-82
 42. Szeltner Z, Morawski M, Juhasz T, et al. GAP43 shows partial co-localisation but no strong physical interaction with prolyl oligopeptidase. *Biochim Biophys Acta* 2010;1804(12):2162-76
 43. Benowitz LI, Routtenberg A. GAP-43: an intrinsic determinant of neuronal development and plasticity. *Trends Neurosci* 1997;20(2):84-91
 44. Brandt I, Gerard M, Sergeant K, et al. Prolyl oligopeptidase stimulates the aggregation of alpha-synuclein. *Peptides* 2008;29(9):1472-8
 45. Yoshimoto T, Kado K, Matsubara F, et al. Specific inhibitors for prolyl endopeptidase and their anti-amnesic effect. *J Pharmacobiodyn* 1987;10(12):730-5
 46. Shinoda M, Matsuo A, Toide K. Pharmacological studies of a novel prolyl endopeptidase inhibitor, JTP-4819, in rats with middle cerebral artery occlusion. *Eur J Pharmacol* 1996;305(1-3):31-8
 47. Peltonen I, Jalkanen AJ, Sinerva V, et al. Different effects of scopolamine and inhibition of prolyl oligopeptidase on mnemonic and motility functions of young and 8- to 9-month-old rats in the radial-arm maze. *Basic Clin Pharmacol Toxicol* 2010;106(4):280-7
 48. Jalkanen AJ, Puttonen KA, Venalainen JI, et al. Beneficial effect of prolyl oligopeptidase inhibition on spatial memory in young but not in old scopolamine-treated rats. *Basic Clin Pharmacol Toxicol* 2007;100(2):132-8
 49. Schneider JS, Giardiniere M, Morain P. Effects of the prolyl endopeptidase inhibitor S 17092 on cognitive deficits in chronic low dose MPTP-treated monkeys. *Neuropsychopharmacology* 2002;26(2):176-82
 50. Kato A, Fukunari A, Sakai Y, Nakajima T. Prevention of amyloid-like deposition by a selective prolyl endopeptidase inhibitor, Y-29794, in senescence-accelerated mouse. *J Pharmacol Exp Ther* 1997;283(1):328-35
 51. Toide K, Shinoda M, Miyazaki A. A novel prolyl endopeptidase inhibitor, JTP-4819—its behavioral and neurochemical properties for the treatment of Alzheimer's disease. *Rev Neurosci* 1998;9(1):17-29
 52. Katsube N, Sunaga K, Aishita H, et al. ONO-1603, a potential antimentia drug, delays age-induced apoptosis and suppresses overexpression of glyceraldehyde-3-phosphate dehydrogenase in cultured central nervous system neurons. *J Pharmacol Exp Ther* 1999;288(1):6-13
 53. Morain P, Robin JL, De Nanteuil G, et al. Pharmacodynamic and pharmacokinetic profile of S 17092, a new orally active prolyl endopeptidase inhibitor, in elderly healthy volunteers. A phase I study. *Br J Clin Pharmacol* 2000;50(4):350-9
 54. Morain P, Boeijinga PH, Demazieres A, et al. Psychotropic profile of S 17092, a prolyl endopeptidase inhibitor, using quantitative EEG in young healthy volunteers. *Neuropsychobiology* 2007;55(3-4):176-83
 55. Umemura K, Kondo K, Ikeda Y, et al. Pharmacokinetics and safety of Z-321, a novel specific orally active prolyl endopeptidase inhibitor, in healthy male volunteers. *J Clin Pharmacol* 1999;39(5):462-70
 56. De Nanteuil G, Portevin B, Lepagnol J. Prolyl endopeptidase inhibitors: a new class of memory enhancing drugs. *Drugs Future* 1998;23(2):167-79
 57. Lawandi J, Gerber-Lemaire S, Juillerat-Janneret L, Moitessier N. Inhibitors of prolyl oligopeptidases for the therapy of human diseases: defining diseases and inhibitors. *J Med Chem* 2010;53(9):3423-38
 58. Yoshimoto T, Kawahara K, Matsubara F, et al. Comparison of inhibitory effects of prolyl-containing peptide derivatives on prolyl endopeptidases from bovine brain and *Flavobacterium*. *J Biochem* 1985;98(4):975-9
 59. Kahyaoglu A, Haghjoo K, Guo F, et al. Low barrier hydrogen bond is absent in the catalytic triads in the ground state but is present in a transition-state complex in the prolyl oligopeptidase family of serine proteases. *J Biol Chem* 1997;272(41):25547-54
 60. Yoshimoto T, Tsuru D, Yamamoto N, et al. Structure activity relationship of inhibitors specific for prolyl endopeptidase. *Agric Biol Chem* 1991;55(1):37-43
 61. Tran T, Quan C, Edosada CY, et al. Synthesis and structure-activity relationship of N-acyl-Gly-, N-acyl-Sar- and N-blocked-boroPro inhibitors of FAP, DPP4, and POP. *Bioorg Med Chem Lett* 2007;17(5):1438-42
 62. Szeltner Z, Rea D, Renner V, et al. Electrostatic environment at the active site of prolyl oligopeptidase is highly influential during substrate binding. *J Biol Chem* 2003;278(49):48786-93

63. Schechter I, Berger A. On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun* 1967;27(2):157-62
64. Abramowitz N, Schechter I, Berger A. On the size of the active site in proteases. II. Carboxypeptidase-A. *Biochem Biophys Res Commun* 1967;29(6):862-7
65. Tsuru D, Yoshimoto T, Koriyama N, Furukawa S. Thiazolidine derivatives as potent inhibitors specific for prolyl endopeptidase. *J Biochem* 1988;104(4):580-6
66. Kanai K, Erd S, Susan E, et al. Prolyl endopeptidase inhibitors: N-acyl derivatives of L-thioprolin-pyrrolidine. *Bioorg Med Chem Lett* 1997;7(13):1701-4
67. Arai H, Nishioka H, Niwa S, et al. Synthesis of prolyl endopeptidase inhibitors and evaluation of their structure-activity relationships: in vitro inhibition of prolyl endopeptidase from canine brain. *Chem Pharm Bull (Tokyo)* 1993;41(9):1583-8
68. Tanaka Y, Niwa S, Nishioka H, et al. New potent prolyl endopeptidase inhibitors: synthesis and structure-activity relationships of indan and tetralin derivatives and their analogues. *J Med Chem* 1994;37(13):2071-8
69. Kanai K, Aranyi P, Bocskai Z, et al. Prolyl oligopeptidase inhibition by N-acyl-pro-pyrrolidine-type molecules. *J Med Chem* 2008;51(23):7514-22
70. Saito M, Hashimoto M, Kawaguchi N, et al. Synthesis and inhibitory activity of acyl-peptidyl-prolinalderivatives toward post-proline cleaving enzyme as nootropic agents. *J Enzyme Inhib* 1990;3(3):163-78
71. Saito M, Hashimoto M, Kawaguchi N, et al. Synthesis and inhibitory activity of acyl-peptidyl-pyrrolidine derivatives toward post-proline cleaving enzyme; a study of subsite specificity. *J Enzyme Inhib* 1991;5(1):51-75
72. Li J, Wilk E, Wilk S. Inhibition of prolyl oligopeptidase by Fmoc-aminoacylpyrrolidine-2-nitriles. *J Neurochem* 1996;66(5):2105-12
73. Portevin B, Benoist A, Remond G, et al. New prolyl endopeptidase inhibitors: in vitro and in vivo activities of azabicyclo[2.2.2]octane, azabicyclo[2.2.1]heptane, and perhydroindole derivatives. *J Med Chem* 1996;39(12):2379-91
74. Vendeville S, Bourel L, Davioud-Charvet E, et al. Automated parallel synthesis of a tetrahydroisoquinolin-based library: potential prolyl endopeptidase inhibitors. *Bioorg Med Chem Lett* 1999;9(3):437-42
75. Bakker AV, Daffeh J, Jung S, et al. Novel in vitro and in vivo inhibitors of prolyl endopeptidase. *Bioorg Med Chem Lett* 1991;1(11):585-90
76. Jarho EM, Venalainen JI, Huuskonen J, et al. A cyclopent-2-enecarbonyl group mimics proline at the P2 position of prolyl oligopeptidase inhibitors. *J Med Chem* 2004;47(23):5605-7
77. Jarho EM, Venalainen JI, Poutiainen S, et al. 2(S)-(Cycloalk-1-enecarbonyl)-1-(4-phenyl-butanoyl)pyrrolidines and 2(S)-(aroyl)-1-(4-phenylbutanoyl)pyrrolidines as prolyl oligopeptidase inhibitors. *Bioorg Med Chem* 2007;15(5):2024-31
78. Wilk S, Orłowski M. Inhibition of rabbit brain prolyl endopeptidase by n-benzyloxy-carbonyl-prolyl-prolinal, a transition state aldehyde inhibitor. *J Neurochem* 1983;41(1):69-75
79. Shishido Y, Furushiro M, Tanabe S, et al. ZTTA, a postproline cleaving enzyme inhibitor, improves cerebral ischemia-induced deficits in a three-panel runway task in rats. *Pharmacol Biochem Behav* 1996;55(3):333-8
80. Joyeau R, Maoulida C, Guillet C, et al. Synthesis and activity of pyrrolidinyl- and thiazolidinyl-dipeptide derivatives as inhibitors of the Tc80 prolyl oligopeptidase from *Trypanosoma cruzi*. *Eur J Med Chem* 2000;35(2):257-66
81. Tsutsumi S, Okonogi T, Shibahara S, et al. Synthesis and structure-activity relationships of peptidyl alpha-keto heterocycles as novel inhibitors of prolyl endopeptidase. *J Med Chem* 1994;37(21):3492-502
82. Hoechst AG. Novel pyrrolidine-2-(1,3-dicarbonyl) derivatives, a process for their preparation, agents containing them, and their use. EP0286927 (A2 and A3); 1988
83. Merrell Pharma, Inc. Prolyl endopeptidase inhibitors. US161106; 1996
84. Wallen EA, Christiaans JA, Forsberg MM, et al. Dicarboxylic acid bis(L-prolyl-pyrrolidine) amides as prolyl oligopeptidase inhibitors. *J Med Chem* 2002;45(20):4581-4
85. Venalainen JI, Juvonen RO, Garcia-Horsman JA, et al. Slow-binding inhibitors of prolyl oligopeptidase with different functional groups at the P1 site. *Biochem J* 2004;382(Pt 3):1003-8
86. Venalainen JI, Garcia-Horsman JA, Forsberg MM, et al. Binding kinetics and duration of in vivo action of novel prolyl oligopeptidase inhibitors. *Biochem Pharmacol* 2006;71(5):683-92
87. Vendeville S, Goossens F, Debreu-Fontaine MA, et al. Comparison of the inhibition of human and *Trypanosoma cruzi* prolyl endopeptidases. *Bioorg Med Chem* 2002;10(6):1719-29
88. Bal G, Van der Veken P, Antonov D, et al. Prolylisoxazoles: potent inhibitors of prolyl oligopeptidase with antitrypanosomal activity. *Bioorg Med Chem Lett* 2003;13(17):2875-8
89. Suntory Ltd. Novel thiazolidine derivative and use thereof. JP1301671 (A); 1989
90. Zeria Pharmaceuticals Co. Ltd. Phenylalkanoylamine derivatives, process for their preparation and their uses. WO 8910918; 1989
91. Zeria Pharmaceuticals Co. Ltd. Condensed benzene derivative. US5028604 (A); 1991
92. Zeria Pharmaceuticals Co. Ltd. Thiazolidine derivative and medicine containing the same. WO1995001352; 1995
93. Ono Pharmaceutical Co. Novel prolinal derivatives. EP0268190 (A1 and B1); 1988
94. Ono Pharmaceutical Co. Thiazolidine derivatives. EP0280956 (A1 and B1); 1988
95. Ono Pharmaceutical Co. Prolinal derivatives and pharmaceutical compositions thereof. US 5262431 (A); 1993
96. Ono Pharmaceutical Co. Prolinal derivatives useful for treating amnesia. US5340832 (A) US5254550 (A); 1994
97. Katsube N, Tateishi N, Matsushita Y, et al. Anti-amnesic effect of ONO-1603,

Low molecular weight inhibitors of Prolyl Oligopeptidase: a review of compounds patented from 2003 to 2010

- a novel prolyl endopeptidase inhibitor. *Jpn J Pharmacol* 1989;49:91-8
98. Yoshitomi Pharmaceutical. Benzene compound. JP3255080 (A); 1991
99. Yoshitomi Pharmaceutical; Japan Tobacco, Inc. Diarylketone compound and its medicinal use. JP7126229 (A); 1995
100. Nakajima T, Ono Y, Kato A, et al. Y-29794-a non-peptide prolyl endopeptidase inhibitor that can penetrate into the brain. *Neurosci Lett* 1992;141(2):156-60
101. Japan Tobacco, Inc. Yoshitomi Pharmaceutical. New proline derivative. JP5025125 (A); 1993
102. Japan Tobacco, Inc. Yoshitomi Pharmaceutical. Proline derivative. JP5186498 (A); 1993
103. Japan Tobacco, Inc. Yoshitomi Pharmaceutical. New proline derivative. JP5201970 (A); 1993
104. Japan Tobacco, Inc. Yoshitomi Pharmaceutical. Compound with prolyl endopeptidase inhibitor activity and pharmaceutical use thereof. WO9412474 (A1); 1994
105. ADIR et Cie. New nitrogenous bicyclic derivatives, process for their preparation and the pharmaceutical compositions which contain them. FR2681864 (A1); 1993
106. Barelli H, Petit A, Hirsch E, et al. S 17092-1, a highly potent, specific and cell permeant inhibitor of human proline endopeptidase. *Biochem Biophys Res Commun* 1999;257(3):657-61
107. Kowall NW, Beal MF, Busciglio J, et al. An in vivo model for the neurodegenerative effects of beta amyloid and protection by substance P. *Proc Natl Acad Sci USA* 1991;88(16):7247-51
108. Orion Corp. Compounds having prolyl oligopeptidase inhibitory activity. WO2004060862 (A2 and A3); 2005
109. Wallen EA, Christiaans JA, Saarinen TJ, et al. Conformationally rigid N-acyl-5-alkyl-L-prolyl-pyrrolidines as prolyl oligopeptidase inhibitors. *Bioorg Med Chem* 2003;11(17):3611-19
110. Ono Pharmaceutical Co. Novel prolinol derivative, production thereof and anti-amnesia containing said derivative. JP1156957 (A); 1989
111. Pfizer. Aromatic pyrrolidine and thiazolidine amides. WO9118891(A1); 1991
112. Orion Corp. Compounds having prolyl oligopeptidase inhibitory activity, methods for their preparation and their use. WO03004468 (A1); 2003
113. Wallen EA, Christiaans JA, Jarho EM, et al. New prolyl oligopeptidase inhibitors developed from dicarboxylic acid bis(L-prolyl-pyrrolidine) amides. *J Med Chem* 2003;46(21):4543-51
114. Jarho EM, Wallen EA, Christiaans JA, et al. Dicarboxylic acid azacycle l-prolyl-pyrrolidine amides as prolyl oligopeptidase inhibitors and three-dimensional quantitative structure-activity relationship of the enzyme-inhibitor interactions. *J Med Chem* 2005;48(15):4772-82
115. Probiodrugs AG. Novel compounds for the treatment of neurological disorders. WO2006058720 (A2 and A3); 2006
116. Probiodrugs AG. Novel compounds. WO2009007415 (A2 and A3); 2009
117. Probiodrugs AG. Inhibitors of prolyl endopeptidase. WO2006120104 (A1); 2006
118. Universitat de Barcelona; Giralt E; Tarrago T. Therapeutic agent for treatment of bipolar affective disorder in mammals. WO2008077978 (A1); 2008
119. Harvey AL. Natural products in drug discovery. *Drug Discov Today* 2008;13(19-20):894-901
120. Tarrago T, Frutos S, Rodriguez-Mías RA, Giralt E. Identification by 19F NMR of traditional Chinese medicinal plants possessing prolyl oligopeptidase inhibitory activity. *Chembiochem* 2006;7(5):827-33
121. Kondo Y, Kondo F, Asanuma M, et al. Protective effect of oren-gedoku-to against induction of neuronal death by transient cerebral ischemia in the C57BL/6 mouse. *Neurochem Res* 2000;25(2):205-9
122. Xu J, Murakami Y, Matsumoto K, et al. Protective effect of Oren-gedoku-to (Huang-Lian-Jie-Du-Tang) against impairment of learning and memory induced by transient cerebral ischemia in mice. *J Ethnopharmacol* 2000;73(3):405-13
123. Tarrago T, Kichik N, Seguí J, Giralt E. The natural product berberine is a human prolyl oligopeptidase inhibitor. *ChemMedChem* 2007;2(3):354-9
124. Zeng XH, Zeng XJ, Li YY. Efficacy and safety of berberine for congestive heart failure secondary to ischemic or idiopathic dilated cardiomyopathy. *Am J Cardiol* 2003;92(2):173-6
125. Kong W, Wei J, Abidi P, et al. Berberine is a novel cholesterol-lowering drug working through a unique mechanism distinct from statins. *Nat Med* 2004;10(12):1344-51
126. Wang X, Wang R, Xing D, et al. Kinetic difference of berberine between hippocampus and plasma in rat after intravenous administration of Coptidis rhizoma extract. *Life Sci* 2005;77(24):3058-67
127. Wang X, Xing D, Wang W, et al. The uptake and transport behavior of berberine in Coptidis Rhizoma extract through rat primary cultured cortical neurons. *Neurosci Lett* 2005;379(2):132-7
128. Tarrago T, Masdeu C, Gomez E, et al. Benzimidazolium salts as small, nonpeptidic and BBB-permeable human prolyl oligopeptidase inhibitors. *ChemMedChem* 2008;3:1558-65
129. Tarrago T, Kichik N, Claasen B, et al. Baicalin, a prodrug able to reach the CNS, is a prolyl oligopeptidase inhibitor. *Bioorg Med Chem* 2008;16:7516-24

Affiliation

Abraham López^{*,†1,2}, Teresa Tarragó^{*1} & Ernest Giralt^{*,†1,2}

^{†,*}Authors for correspondence

¹Institute for Research in Biomedicine, Barcelona Science Park, Baldiri Reixac, 10, E-08028, Barcelona, Spain

Tel: +34 934037123; Fax: +34 934037126;

E-mail: teresa.tarrago@irbbarcelona.org

²Department of Organic Chemistry,

University of Barcelona,

Martí Franqués, 1, E-08028, Barcelona,

Spain

Applications and future of ion mobility mass spectrometry in structural biology

Cite this: *New J. Chem.*, 2013, **37**, 1283

Abraham López,^{ab} Teresa Tarragó,^b Marta Vilaseca^b and Ernest Giralt^{*ab}

Received (in Montpellier, France)
19th November 2012,
Accepted 18th March 2013

DOI: 10.1039/c3nj41051j

www.rsc.org/njc

Ion mobility coupled to mass spectrometry (IMMS) is a technique that simultaneously separates gaseous ions on the basis of their mass, shape and size. The continuous advances in native mass spectrometry applied to IMMS have prompted its application in the structural study of biomolecules. Hence, challenging systems in terms of molecular size, complexity and heterogeneity can be transferred to the gas phase and their structural properties can be analyzed. Although IMMS is a low resolution technique, *in silico* theoretical models or simulations can provide valuable microscopic information about the structure, dynamics and population of the species. Given these unprecedented advantages of IMMS, this emerging method is becoming an established, reliable and versatile tool for structural biology.

Introduction

Structural biology plays a key role in life sciences, and as such this discipline requires the development of structural elucidation techniques. Mass spectrometry (MS) is a well-established approach for quantitative mass analysis. Nevertheless, extracting structural information from mass relies on indirect measurements, for

instance hydrogen–deuterium exchange.¹ Thus, new MS developments aim to detect structural and dynamical features.² In this regard, one of the most promising and emerging cutting-edge methodologies is ion mobility mass spectrometry (IMMS).

In this technique, gaseous ions are separated not only on the basis of their mass to charge ratio, but also of their size and shape.³ Continuous technical advances have increased the sensitivity and resolution power of IMMS, and the development of soft ionization techniques has prompted the structural study of biomolecules. These applications have been growing steadily ever since.

^a Institute for Research in Biomedicine (IRB Barcelona), Barcelona, Spain.

E-mail: ernest.giralt@irbbarcelona.org

^b Department of Organic Chemistry, University of Barcelona, Barcelona, Spain



Abraham López

Abraham López was born in Girona (Spain). He obtained his license in chemistry at the University of Barcelona in 2009. He completed the European master in advanced organic chemistry in the same faculty. Currently he is pursuing his PhD in organic chemistry in Ernest Giralt's laboratory (Institute for Research in Biomedicine and University of Barcelona). His thesis project combines ion mobility mass spectrometry and nuclear magnetic resonance.



Ernest Giralt

Prof. Ernest Giralt was born in Barcelona. He was a postdoctoral researcher at the University of Montpellier, France. He has been a Full Professor in Organic Chemistry at the University of Barcelona since 1986. He was a Visiting Professor at the University of California, San Diego, and a Research Associate at the Scripps Research Institute, USA, in 1991. Since 2002, he has been the Head of the Chemistry and Molecular Pharmacology

Programme in the Institute for Research in Biomedicine (IRB Barcelona). His major interests lie in the field of peptide synthesis and structure determination, using NMR and MS in relation to molecular-recognition processes.

Mobility of gaseous ions: types of ion mobility mass spectrometers

The ion mobility phenomenon was discovered at the beginning of the 20th century.⁴ Later on, McDaniel *et al.*⁵ coupled drift tubes to a mass spectrometer for the simultaneous measurement of drift times and mass, similarly to the modern IMMS.

Ion mobility mass spectrometers feature an IM cell before the mass analyzer. Gaseous ions are injected into the cell and accelerated by a weak electric field. The presence of a certain pressure of buffer gas in this cell yields low energy collisions with the buffer gas: the higher the collision cross sections (CCS) of the ion, the greater the number of collisions with the buffer gas. As collisions increase, there is an energy loss, and consequently ions take a longer time to cross the IM cell (the “drift time”). After that, ions are injected into the mass analyzer, achieving a simultaneous separation on the basis of the CCS to charge ratio and the m/z ratio. The three-dimensional spectrum obtained consists of mass, drift time and relative intensity⁶ (Fig. 1d, lower panel). Depending on the configuration of the IM cell, there are three main classifications, which are described in the following section. Nevertheless, these classifications are not intended to be an exhaustive description of the methodologies.

Drift tube

The simplest IM cell is the drift tube, which is formed by a tube containing a low pressure buffer gas that is subjected to a weak electric field. Ions are injected and accelerated throughout the cell by the electric field (Fig. 1a). The theoretical model is

applicable only when the ratio between the electric field and the buffer gas number density (E/N) is low (the so-called “low field limit”). In this case, the drift velocity (V_d) is proportional to the electric field E . The constant, K , which is known as mobility, includes information about the buffer gas, shape, size and mass to charge ratio.

$$V_d = KE \quad (1)$$

For an individual charge state and defined experimental conditions, eqn (1) can be developed to eqn (2), whereby:⁷

$$V_d = \frac{3e}{16N} \left(\frac{2\pi}{\mu k_B T} \right)^{\frac{1}{2}} \frac{KE}{\sigma} \quad (2)$$

This equation allows the direct determination of the CCS (σ) from the drift velocity, charge of the ion (e), temperature (T), the reduced mass of the ion and the buffer gas (μ), and the buffer gas particle density (N).

High field asymmetric waveform IM

The main characteristic of this IM chamber is that it operates at pressures near atmospheric pressure. The separation takes place by a high field rather than a low field, and thus there is no exact mathematical description for the overall process. Ions injected into the cell are subjected to an alternate asymmetric waveform potential between two electrodes. Ions adopt a zigzag trajectory that depends on the balance between the high and low field mobilities. By applying a potential (the “compensation voltage”) only certain ions adopt a straight overall trajectory and cross the cell (Fig. 1b). The scanning of this “compensation voltage” yields the spectrum.^{8,9}

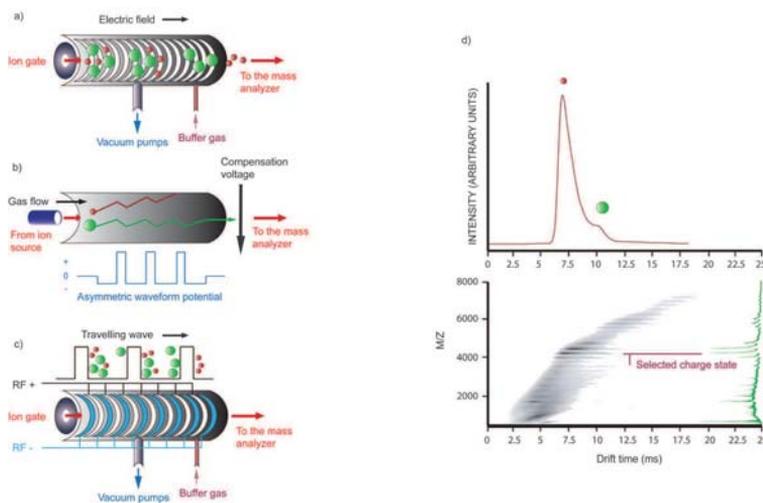


Fig. 1 Separation of gaseous ions on the basis of ion mobility. (a) Drift tube, consisting of a tube filled with buffer gas and subjected to an electric field.⁶ (b) High field asymmetric waveform IM, where ions are carried with the help of a gas flow. The separation takes place by applying an asymmetric waveform potential, which is cancelled with a compensation voltage characteristic of each ion.⁸ (c) Traveling wave IM, in which ions are pushed by consecutive voltage pulses. The ions that show more resistance with the buffer gas are more likely to surf the wave.¹⁴ (d) Right, a three dimensional mobilitygram showing the time domain (abscissae) and the mass domain (ordinate). The ion mobility spectrum for a certain charge state is shown at the top, where two hypothetical isobaric species were separated on the basis of size.

Travelling wave IM (TWIM)

TWIM spectrometers are the most common nowadays, since the three commercialized IM instruments operate with these cells. The Synapt G1 HDMS,¹⁰ G2¹¹ and G2S (launched by Waters in 2006, 2009 and 2011 respectively) show similar characteristics, although the newest systems have higher MS resolution and sensitivity as well as higher IM resolution power.¹² The resolution of the first commercial IMMS instruments was comparable to that of drift tube devices; however, their sensitivity was markedly increased as a result of the high duty cycle.¹⁰ The cell, which operates between 0.5 and 1 mbar,^{10,13} is formed by successive stacked electrode rings. Voltage pulses are applied between successive rings, radially confining analyte ions. Displacement of these “waves” to the adjacent rings makes these ions surf the waves. Those with lower mobility are more likely to roll over these waves and therefore take a longer time to cross the cell^{14–16} (Fig. 1c). Several mass spectra are acquired simultaneously for each push in the orthogonal acceleration TOF mass analyzer. Shwartsburg and Smith provided a good model to partially explain the theory of ion transit velocity and resolution as a function of mobility and experimental parameters.¹⁶ Later on, several experimental studies were developed to facilitate the optimization of the process.¹⁷ Bearing in mind the limitations of this model, the only way to extract the CCS from the drift time is by calibration.^{16–18} The currently accepted calibration is based on native standards of known CCS spanning similar mass and mobility, yielding the correct interpolation and avoiding extrapolation errors. Nevertheless, the size range covered by calibrant ions limits the study of large complexes.¹⁹

Configuration of an IMMS spectrometer

The construction of an IMMS spectrometer should ensure simultaneous separation of ions on the basis of size and mass, without compromising their transmission through all the vacuum regions.

The first step is the generation of gaseous ions, which starts in the ionization source. Electrospray ionization (ESI) and nanoflow electrospray ionization (nESI) are formed by a tip held at a high electric potential. A flow of analyte solution is sprayed through the tip and, after several physical processes, the gaseous charged analyte ions are formed. The development of soft ionization techniques is of paramount importance for structural biology. Both ESI²⁰ and nESI²¹ preserve protein tertiary structures and the non-covalent interactions of complexes or aggregates.²² In particular, nESI is the most suitable ion source for biomolecules, because of its low sample consumption and its suitability for aqueous buffers.

Once formed, ions are usually mass selected in a mass filter, typically a quadrupole. Subsequently, the release of packs of ions into the IM cell and the simultaneous recording of MS spectra⁶ allow the composition of the three-dimensional IM spectrum.

Biomacromolecular ions in the gas phase

Several factors affect gaseous ions during IM experiments. Although each particular system behaves in a different manner

depending on its structural and dynamic properties, data analysis must observe these effects in order to avoid artifacts.

Once vaporized, solution-like metastable ions tend to adopt gas phase stable structures.^{23–25} In the absence of solvent, the balance of hydrophobic, hydrogen bonding and electrostatic interactions is reorganized. Consequently, structural rearrangements take place to adapt the new gaseous environment.^{22,26–28}

Electrostatic forces strengthen as a result of lack of solvent shield between neighboring charges. Adjacent charges of the same sign suffer electrostatic repulsions that cause an expansion of the tertiary structure, especially at higher charge states.^{3,26–30} In contrast, low energy collisions with buffer gas molecules during the IM experiments increase the internal energy of the ions.^{12,31} This increase can activate certain vibrational modes that lead to structural transitions.²⁶ Side chain reorganization is the main rearrangement in solvent exposed areas,²² although the backbone and the hydrogen bond networking are globally maintained.²⁹ Nevertheless, the absence of solvent can have a greater effect if internal cavities are present in the system. The structure may partially collapse as a result of the instability under vacuum.^{30,32,33}

Residual water molecules, adducts of buffer components or counter ions can resist desolvation and ionization and remain bound to the gaseous ions. When present at specific positions, these substances help to stabilize the native-like structures and protect them against the gas phase distortions mentioned above.^{23,32,34}

Theoretical models for IM measurements

IM is a low resolution technique, which provides information about the macroscopic structure and global structural processes. In order to establish a comparison with the high resolution data provided by X-ray or nuclear magnetic resonance (NMR), *in silico* theoretical models are crucial.

Several methodologies have been proposed for CCS calculation; however, currently the most widely accepted ones are those included in the MOBCAL software. The simplest calculation of the cross section is the rotational average of the geometric cross section of the ion, derived from ion–buffer gas molecule collisions.³⁵ This method, called “hard sphere projection approximation”, or simply “projection approximation”, is the fastest calculation. Nevertheless, it presents distortions with respect to the experimental CCS. These distortions are attributed to the fact that the method does not take into account the microscopic interactions between the ions and the gas molecules. The first microscopic interaction corrected was surface scattering, giving the “exact hard sphere scattering model”.³⁶ Nevertheless, the most accurate procedure is the “trajectory method”.³⁷ Here, the long-range potentials between the ion and the buffer gas are considered, giving approximately 20% higher values than the “projection approximation”.³⁸

In order to show better correlation with the experimental data, and also to decipher the microscopic mechanisms and processes that take place in the gas phase, the X-ray or NMR coordinates can be subjected to molecular dynamics or

molecular mechanics. Simulations are now a routine task to limit the energetically accessible conformational space, instead of a single static structure.³⁷

Nevertheless, the main limitation is that atomistic simulations are too computationally demanding to obtain significant time-scales for the simulation of complexes. For such large systems, the simplest approach is to approximate each subunit as a sphere and perform molecular dynamics of the quaternary arrangement on these rigid spheres.¹⁸ This approach is known as "coarse grain modelling". Hybrid methodologies incorporating partial atomic information for monomers, together with global oligomeric constraints, are the most widely used methods, since they derive both the gas-phase topology and the conformational arrangements for large complexes.³⁹ The complete atomistic models are useful for the detailed study of systems, for instance in docking or unfolding processes.^{17,33,40}

Examples of IM studies of biomolecules

This section is intended to give an overall perspective of the studies carried out by IM, especially those focusing on the field of biomolecules. For a complete description of the experiments, we suggest the reader refer to the literature.

The first proof of concept of coupled IMMS instruments for the study of biomolecules came together with the development of soft ionization techniques. In that regard, Jarrold and Clemmer confirmed the conformational stability of monomeric protein ions in the gas phase,³ and later several groups studied the conformational space of peptides and small proteins.^{3,6,10,13,29} Similarly, the increase in conformational stability as a result of ligand or cofactor binding was also successfully evaluated in various systems.^{30,31} For that purpose, the ions were subjected to collision activation prior to the ion mobility section. This energetic increment induced a certain degree of unfolding,⁴¹ which was partially blocked when small binding molecules were present.^{30,31}

A research line of particular interest in recent years has been the IMMS analysis of intrinsically disordered proteins (IDPs) or proteins containing disordered regions, since the study of these molecules is compromised by other structural elucidation techniques. Analysis of the high mobility group A (HMGA) protein by IMMS is a clear example of how this approach can correlate the structural changes caused by post-translational modifications with the final biochemical function. HMGA, which is formed by a DNA binding domain and an intrinsically disordered acidic C-terminal tail, presents different degree of tertiary structure compacity depending on the type of post-translational modifications. The IMMS analysis of the C-terminal partial truncations, together with the phosphorylations on this acidic tail, indicated that the DNA binding domain was hindered to some extent by the C-terminal region. Consequently, the interactions with other proteins and with DNA were disrupted.⁴² Similarly, p53 protein was also interrogated by IMMS. Other methodologies suggested that the disordered regions of the free complex adopt a compact structure when cognate DNA is present. To clarify the mechanism by which p53

binds to DNA, the isolated domains and subcomplexes, as well as several constructs, were analyzed separately. Taken together, the data showed an unprecedented compaction of the tertiary structure in the gas phase, which was attributed to the collapse of the disordered regions. In contrast, the more structured internal core, responsible for the organization of the complex, remains folded in the gaseous environment.⁴³ Several other studies performed on IDPs reported the versatility of IMMS to discriminate species in complex mixtures on the basis of their conformation and oligomeric state. Furthermore, the high sensitivity of the methodology reflected the conformational redistribution when the ligand was present.^{44,45}

IMMS is of particular interest in the case of large non-covalent protein complexes, since other approaches suffer experimental limitations that severely compromise their use for this kind of molecules. Immediately after the first intact complexes were transferred to the gas phase by ESI, several groups performed IMMS studies to determine the topology and the conformational changes of complexes with the help of theoretical models.^{46,47} The remarkable study by Heck and Uetrecht in virus capsids⁴⁸ and intermediate nucleation oligomers⁴⁹ allowed the unprecedented detection of gas phase virus conformers and revealed their assembly pathways. To shed light on the gaseous stability of complexes, collisional ion activation was extensively applied. The common decomposition mechanism in many systems proceeds by unfolding of the monomers prior to their stripping from the complex, rather than by a dissociation of the folded monomer from the complex.^{33,40,50} Nevertheless, one of the main limitations that must be addressed in the study of intact large complexes is the gas phase stability of ions. Currently, the search for charge reducing agents, which minimize coulombic repulsion by favoring lower charge states,⁴⁷ is an area of great interest. Certain buffer additives or counterions are able to dissipate the excess energy of the ions by an evaporative cooling mechanism during the ionization step, thereby stabilizing the solution structures.^{32,34,50} However, the ionization of whole membrane complexes has been particularly challenging. The development of ionization strategies for these complexes involves the vaporization of micelles formed by a non-ionic detergent containing the complex. Afterwards, the activation of the micelle forces the concomitant release of the complex and the stripping of the detergent molecules.⁵¹ Thus, a mass spectrometry approach designed by Robinson and coworkers disclosed that tightly bound lipids are required for the integrity of the complex and subunit interactions for V-type ATPase. Subsequent IMMS analysis revealed a higher degree of structural heterogeneity when subunit I was present, stating that this dynamics was necessary for the mechanism of the H⁺ channel.⁵²

The millisecond time-scale of IM separation, combined with the high sensitivity and resolution of MS, was exploited for the direct detection and quantification of transient species in protein aggregation processes. The valuable work done by Ashcroft and co-workers on β_2 microglobulin sheds light on amyloidogenic aggregation and fibrillogenesis. After studying the conformational space of the monomeric β_2 microglobulin⁵³ (Fig. 2a), they detected

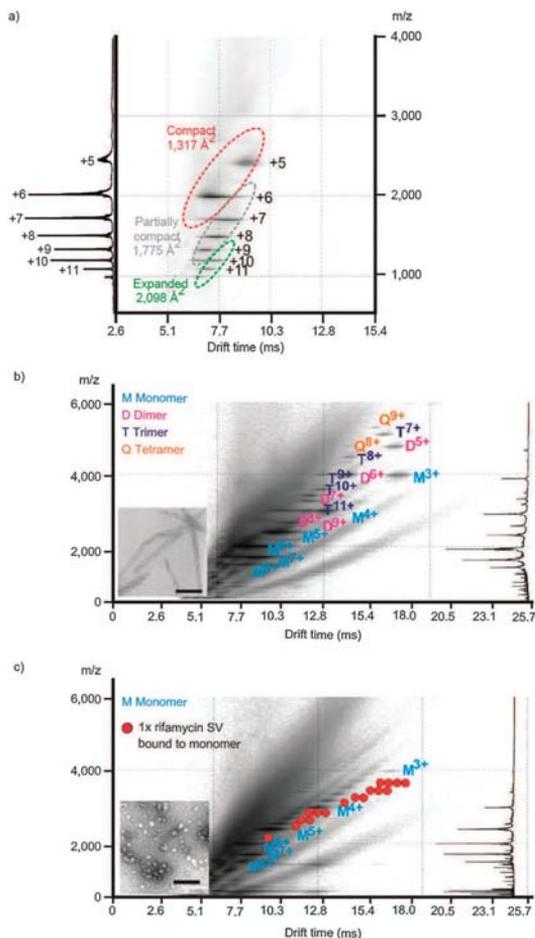


Fig. 2 (a) Conformational distribution of monomeric β_2 -microglobulin at pH = 2.5, showing the compact (CCS = 1.317 \AA^2), partially compact (CCS = 1.775 \AA^2) and expanded species (CCS = 2.098 \AA^2). Note that the extended forms have more exposed sites and therefore displace towards higher charge states. (b) Oligomeric distribution after a certain incubation period, showing from the monomer to the tetramer. (c) Oligomeric distribution after the incubation period in the presence of 10 equivalents of the fibrillogenic inhibitor rifamycin SV, in which disassembly of oligomers can be seen. A further analysis disclosed that rifamycin SV binds to all of the monomeric conformers, thus preventing the elongated conformers to act as nucleation cores. Additionally, the stabilization of the monomer-bound states thermodynamically drives the disruption of oligomers.⁵⁵ In (b) and (c) transmission electron microscopy images of the fibrils are also shown (with permission of Nature Publishing Group).

oligomeric species in the first stages of amyloid formation⁵⁴ (Fig. 2b). With the support of other methodologies, they proposed a mechanism for the fibrillogenic inhibitor rifamycin B, which binds to certain conformational states, thus preventing nucleation⁵⁵ (Fig. 2c). Monomers and oligomers of the amyloid peptides $\alpha\beta 40$ and $\alpha\beta 42$, which are involved in Alzheimer's disease, have also been addressed by IMMS. The analysis of the oligomeric distribution indicated a possible oligomerization mechanism.^{56,57}

Along a different line, the Bowers group applied an IMMS approach for the aggregation mechanism of α -synuclein, the protein responsible for the formation of Lewy bodies in Parkinson's disease and other neurodegenerative disorders. The dependency of α -synuclein conformers on pH pointed out that compact structures at low pH have more tendency to aggregate.⁵⁸ Recent studies by other groups disclosed certain autoproteolytic fragments that can act as key intermediates in aggregation.⁵⁹

IMMS has also been successfully used to study oligonucleotides. These nonglobular, highly charged and flexible biomolecules presented drastic structural and dynamic differences compared to most of the proteins. In spite of this, IMMS detected chemically modified DNA and DNA-drug complexes,⁶⁰ but was unfortunately limited by the low resolution of the technique. Interestingly, the study of an oligonucleotide complexed with a ruthenium-based ligand not only provided the CCS of the complex, but also disclosed the binding site.⁶¹ Furthermore, the versatility of ESI-TWIMS was applied to detect quadruplex DNA.⁶² Similarly, the first direct experimental detection of DNA triplex structures was carried out by Orozco and co-workers. The experiments were coherent with a flexible ensemble of structures that showed a dynamic behavior in solution⁶³ (Fig. 3). Nevertheless, the robust computational approach was crucial for developing the model structures.

Conclusions and future outlook

IMMS is becoming a well-established technique in structural biology, although it is still under constant development. This methodology allows the simultaneous measurement of the mass and the shape of ions.³ The development of soft ionization techniques (ESI and nESI) has allowed the study of biomacromolecules by IMMS. Nonetheless, several artifacts derived from the balance of electrostatic and hydrophobic interactions in the gas phase can affect the integrity of gaseous ions.

The progress of IMMS has been made possible, at least in part, by the implementation of theoretical approaches in data treatment and analysis. This has allowed researchers to gain microscopic insights beyond the IMMS low resolution. The most straightforward way is to compare the CCS value from the X-ray or NMR structures.³⁸ More coherent with the experimental data, the simulation of the conformational space for gaseous ions under experimental conditions provides information at the atom level for conformational changes, rearrangements and unfolding processes.

The applications of IMMS are currently expanding in terms of size and complexity. The major concern is the stability of gaseous ions during ionization and the IM measurement,^{22,26,27} especially for high molecular weight complexes, membrane proteins, and viral components. To overcome these limitations, many research efforts are being channeled into the search for solution additives^{32,34,50} and charge shifting reagents.⁴⁷ As Brian T. Chait stated at the ASMS Sanibel Conference 2012, robust methods for freezing both high and low affinity *in vivo* interactions and for isolating endogenous protein complexes are among the challenges that should be further addressed by MS.

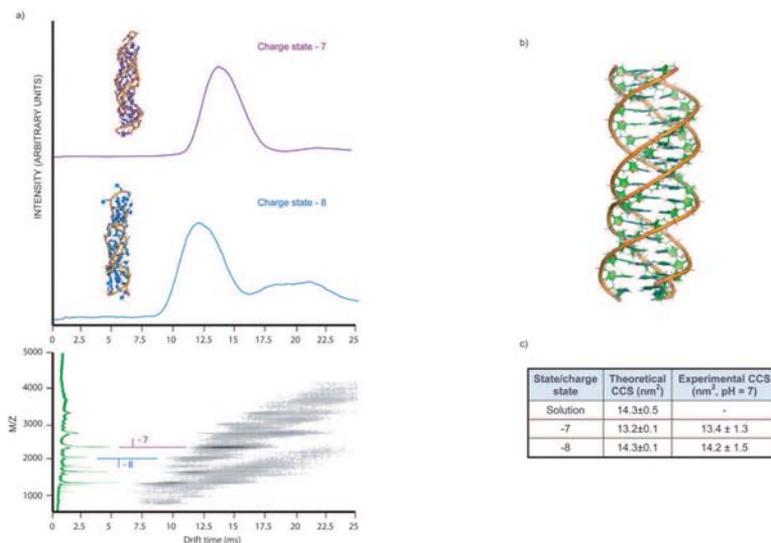


Fig. 3 (a) Ion mobility spectrum for the -7 and -8 charge states of the DNA triplex 18-mer. The peak broadening is due to the conformational flexibility and the charge distributions that contribute to the same charge state. (b) Simulated solution DNA triplex 18-mer in solution. Comparison of the simulated solution structural descriptors with the gas phase ones provided microscopic confirmation of the maintenance of the gas phase structure (data not shown). (c) Simulated CCS in gas phase compared with the experimental value. Although the gas phase CCS differs slightly from the solution CCS, the values are in good agreement. The DNA triplex structures simulated in the gas phase are depicted over the corresponding ion mobility spectra.⁶³

The advantages of a rapid experimental time scale and high sensitivity set the high potential of IMMS in good stead for the analysis of heterogeneous systems. Proof of this is the numerous studies carried out on IDPs^{54,55,57} or systems containing disordered regions.^{42,43} These systems are of particular interest since they are involved in interaction networks or self-aggregation processes.

Finally, IMMS is sensitive to the effects of low molecular weight ligands over the conformational space and populations. This property points to a promising future for IMMS in the screening of ligands in a wide variety of systems, such as membrane complexes and self-aggregating proteins.

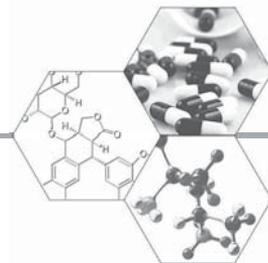
Acknowledgements

The authors want to thank Annalisa Arcella for supplying the DNA triplex images. Mr Abraham López is a PhD student at the Organic Chemistry Department of the University of Barcelona and the Institute for Research in Biomedicine, and has received funding from the Instituto de Salud Carlos III and the Ministry of Science and Innovation. Dr Teresa Tarragó is a Research Associate at the Institute for Research in Biomedicine and has received funding from La Marató de TV3 Foundation. Dr Ernest Giralt is employed by the Department of Organic Chemistry, University of Barcelona, and has received funding from the Institute for Research in Biomedicine, MCYT-FEDER (Bio2008-00799), and the Generalitat de Catalunya (XRB and Grup Consolidat). Dr Marta Vilaseca is a mass spectrometry core facility manager at the Institute for Research in Biomedicine and has received funding from the Institute for Research in Biomedicine.

References

- 1 S. J. Eyles and I. A. Kaltashov, *Methods*, 2004, **34**, 88.
- 2 J. L. P. Benesch and B. T. Ruotolo, *Curr. Opin. Struct. Biol.*, 2011, **21**, 641.
- 3 D. E. Clemmer and M. F. Jarrold, *J. Mass Spectrom.*, 1997, **32**, 577.
- 4 P. Langevin, *Ann. Chim. Phys.*, 1903, **28**, 433.
- 5 E. W. McDaniel, W. S. Barnes and D. W. Martin, *Rev. Sci. Instrum.*, 1962, **33**, 2.
- 6 C. S. Hoaglund, S. J. Valentine and C. R. Spörle, *Anal. Chem.*, 1998, **70**, 2236.
- 7 E. A. Mason and E. W. McDaniel, *Transport properties of ions in gases*, Wiley, New York, 1988.
- 8 R. W. Purves, R. Guevremont and S. Day, *Rev. Sci. Instrum.*, 1998, **69**, 4094.
- 9 R. Guevremont, *J. Chromatogr., A*, 2004, **1058**, 3.
- 10 S. D. Pringle, K. Giles and J. L. Wildgoose, *et al.*, *Int. J. Mass Spectrom.*, 2007, **261**, 1.
- 11 A. Wallace, *Am. Lab.*, 2010, **42**, 13.
- 12 Y. Zhong, S. J. Hyung and B. T. Ruotolo, *Analyst*, 2011, **136**, 3534.
- 13 D. P. Smith, T. W. Knapman and I. Campuzano, *et al.*, *Eur. J. Mass Spectrom.*, 2009, **15**, 113.
- 14 K. Giles, S. D. Pringle and K. R. Worthington, *et al.*, *Rapid Commun. Mass Spectrom.*, 2004, **18**, 2401.
- 15 K. Thalassinou, S. E. Slade and K. R. Jennings, *et al.*, *Int. J. Mass Spectrom.*, 2004, **236**, 55.
- 16 A. A. Shvartsburg and R. D. Smith, *Anal. Chem.*, 2008, **80**, 9689.
- 17 B. T. Ruotolo, J. L. P. Benesch and A. M. Sandercock, *et al.*, *Nat. Protoc.*, 2008, **3**, 1139.

- 18 B. T. Ruotolo, K. Giles and I. Campuzano, *et al.*, *Science*, 2005, **310**, 1658.
- 19 M. F. Bush, Z. Hall and K. Giles, *et al.*, *Anal. Chem.*, 2010, **82**, 9557.
- 20 J. B. Fenn, M. Mann and C. K. Meng, *et al.*, *Science*, 1989, **246**, 64.
- 21 M. S. Wilm and M. Mann, *Int. J. Mass Spectrom.*, 1994, **136**, 167.
- 22 B. T. Ruotolo and C. V. Robinson, *Curr. Opin. Chem. Biol.*, 2006, **10**, 402.
- 23 T. Wyttenbach and M. T. Bowers, *J. Phys. Chem. B*, 2011, **115**, 12266.
- 24 E. R. Badman, S. Myung and D. E. Clemmer, *J. Am. Soc. Mass Spectrom.*, 2005, **16**, 1493.
- 25 E. R. Badman, C. S. Hoaglund-Hyzer and D. E. Clemmer, *Anal. Chem.*, 2001, **73**, 6000.
- 26 K. B. Shelimov, D. E. Clemmer and R. R. Hudgins, *et al.*, *J. Am. Chem. Soc.*, 1997, **119**, 2240.
- 27 M. F. Jarrold, *Annu. Rev. Phys. Chem.*, 2000, **51**, 179.
- 28 M. F. Jarrold, *Acc. Chem. Res.*, 1999, **32**, 360.
- 29 C. A. Scarff, K. Thalassinou and G. R. Hilton, *et al.*, *Rapid Commun. Mass Spectrom.*, 2008, **22**, 3297.
- 30 P. A. Faull, K. E. Korkeila and J. M. Kalapothakis, *et al.*, *Int. J. Mass Spectrom.*, 2009, **283**, 140.
- 31 J. T. S. Hopper and N. J. Oldham, *J. Am. Soc. Mass Spectrom.*, 2009, **20**, 1851.
- 32 L. Han, S. J. Hyung and J. J. S. Mayers, *et al.*, *J. Am. Chem. Soc.*, 2011, **133**, 11358.
- 33 Z. Hall, A. Politis and M. F. Bush, *et al.*, *J. Am. Chem. Soc.*, 2012, **134**, 3429.
- 34 J. Freeke, C. V. Robinson and B. T. Ruotolo, *Int. J. Mass Spectrom.*, 2010, **298**, 91.
- 35 E. Mack, *J. Am. Chem. Soc.*, 1925, **47**, 2468.
- 36 A. A. Shvarstburg and M. F. Jarrold, *Chem. Phys. Lett.*, 1996, **261**, 86.
- 37 M. F. Mesleh, M. F. Hunter and A. A. Shvartsburg, *et al.*, *J. Phys. Chem.*, 1996, **100**, 16082.
- 38 E. Jurneczko and P. E. Barran, *Analyst*, 2011, **136**, 20.
- 39 A. Politis, A. Y. Parh and S. J. Hyung, *et al.*, *PLoS One*, 2010, **5**, 12080.
- 40 B. T. Ruotolo, S. J. Hyung and P. M. Robinson, *et al.*, *Angew. Chem., Int. Ed.*, 2007, **46**, 8001.
- 41 S. L. Koeniger and D. E. Clemmer, *J. Am. Soc. Mass Spectrom.*, 2007, **18**, 322.
- 42 E. Maurizio, L. Cravello and B. Brady, *et al.*, *J. Proteome Res.*, 2011, **10**, 3283.
- 43 K. Pagel, E. Natan, Z. Hall, A. R. Fersht and C. V. Robinson, *Angew. Chem., Int. Ed.*, 2013, **52**, 361.
- 44 S. R. Harvey, M. Porrini and C. Stachl, *et al.*, *J. Am. Chem. Soc.*, 2012, **47**, 19384.
- 45 F. Canon, F. Ballivian and R. Chirot, *et al.*, *J. Am. Chem. Soc.*, 2011, **133**, 7847.
- 46 J. A. Loo, B. Berhane and C. S. Kaddis, *et al.*, *J. Am. Soc. Mass Spectrom.*, 2005, **16**, 998.
- 47 E. Van Duijn, A. Barendregt and S. Synowsky, *et al.*, *J. Am. Chem. Soc.*, 2009, **131**, 1452.
- 48 C. Uetrecht, C. Versluis and N. R. Watts, *et al.*, *Angew. Chem., Int. Ed.*, 2008, **47**, 6247.
- 49 C. Uetrecht, I. M. Barbu and G. K. Shoemaker, *et al.*, *Nat. Chem.*, 2011, **3**, 126.
- 50 J. Freeke, F. Matthew and C. V. Robinson, *et al.*, *Chem. Phys. Lett.*, 2012, **524**, 1.
- 51 N. P. Barrera, S. C. Isaacson and M. Zhou, *et al.*, *Nat. Methods*, 2009, **6**, 585.
- 52 M. Zhou, N. Morgner and N. P. Barrera, *et al.*, *Science*, 2011, **334**, 380.
- 53 D. P. Smith, K. Giles and R. H. Bateman, *et al.*, *J. Am. Soc. Mass Spectrom.*, 2007, **18**, 2180.
- 54 D. P. Smith, S. E. Radford and A. E. Ashcroft, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 6794.
- 55 L. A. Woods, G. W. Platt and A. L. Hellewell, *et al.*, *Nat. Chem. Biol.*, 2011, **7**, 730.
- 56 S. L. Bernstein, T. Wyttenbach and A. Baumketner, *et al.*, *J. Am. Chem. Soc.*, 2005, **127**, 2075.
- 57 S. L. Bernstein, N. F. Dupuis and N. D. Lazo, *et al.*, *Nat. Chem.*, 2009, **1**, 326.
- 58 S. L. Bernstein, D. F. Liu and T. Wyttenbach, *et al.*, *J. Am. Soc. Mass Spectrom.*, 2004, **15**, 1435.
- 59 C. Vlad, C. Lindner and S. Karreman, *et al.*, *ChemBioChem*, 2011, **12**, 2740.
- 60 J. M. Koomen, B. T. Ruotolo and K. J. Gillig, *et al.*, *Anal. Bioanal. Chem.*, 2002, **373**, 612.
- 61 J. P. Williams, J. A. Lough and I. Campuzano, *et al.*, *Rapid Commun. Mass Spectrom.*, 2009, **23**, 3563.
- 62 K. C. Porter and J. L. Beck, *Int. J. Mass Spectrom.*, 2011, **304**, 195.
- 63 A. Arcella, G. Portella and M. Luz Ruiz, *et al.*, *J. Am. Chem. Soc.*, 2012, **134**, 6596.



For reprint orders, please contact reprints@future-science.com

Peptide POP inhibitors for the treatment of the cognitive symptoms of schizophrenia

Schizophrenia is a serious life-long disease that affects a significant part of the adult population. Although there is considerably effective medication for the positive symptoms of the disease, none are available for the associated cognitive deficits. These deficits are a core feature of schizophrenia, and they severely impair the functionality and social integration of patients. POP is a promising target for the treatment of the cognitive deficits of schizophrenia. Inhibitors of this peptidase show cognition-enhancing properties, act through a complex mechanism and have suitable pharmacological properties. Nevertheless, several studies must be carried out in order to improve the design and clinical evaluation of these substances. Permeability to the brain, appropriate animal models and suitable indications are the main issues that must be addressed. However, current information supports the potential of POP as an interesting drug target for the treatment of the cognitive deficits related to schizophrenia.

Cognitive impairment in schizophrenia: an unmet medical need

Schizophrenia is the term used to describe a brain disorder that impairs mental function [1]. It is a heterogenic disease that presents several common features in patients. According to the WHO, seven in every 1000 adults are affected by schizophrenia, which represents almost 24 million people worldwide who require constant medication as no cure is yet available.

Schizophrenia is of genetic origin. The onset of this disease depends on the alteration of key genes, but it is also related to environmental factors [1]. This combination of development and genetic origin makes schizophrenia a complex disease. Whether schizophrenia is a single disorder presenting variations among patients, or rather a term that the scientific and medical community has agreed to use to cover a group of mental diseases with similar symptomatology, is still a subject of debate. Onset of the disease occurs between adolescence and early adulthood. In childhood, some minor symptoms may appear, but it is not until puberty that schizophrenic features become more evident. In this period, changes in the environment, such as variations in lifestyle, promote the appearance of marked symptoms and severely hinder the functionality of patients [1].

For decades, the pathophysiology of schizophrenia was believed to point to imbalances in the dopaminergic system [2]. Nevertheless, during recent years this hypothesis has been modified to also include neurochemical alterations in other neurotransmitter systems, which

can indirectly affect the dopaminergic system. For instance, serotonergic [3], glutamatergic [4], central cholinergic [5] and γ -aminobutyric [4,6] systems are involved in the pathophysiology of schizophrenia.

The symptomatology of this mental disorder comprises various cognitive and behavioral dysfunctions (FIGURE 1, left and central columns). The classification is phenomenological and sometimes the symptoms overlap with those present in other mental diseases. Positive symptoms are the most pronounced in schizophrenia and include disorganization, hallucinations and delusions. Negative symptoms imply a lack of normal emotional responses and are characterized by apathy, anhedonia and a general loss of motivation. In addition to positive and negative symptoms, alterations in cognitive functions are also considered to be a core feature of schizophrenia, since several unbalanced conditions might precede the psychotic state in some patients during their infancy and youth [7], and working memory and attention are specifically altered in schizophrenic patients [8]. This complex symptomatology of schizophrenia makes the greatest contribution to the functional disability of patients [1,9]. Moreover, current pharmacological treatment of this disease has strong extrapyramidal side effects that may indeed promote cognitive deficits [10]. Finally, other symptoms of the disease may include mood disorders, such as depression.

Treatment for schizophrenia consists of a combination of drug and psychosocial therapies. In the field of pharmacology (FIGURE 1, right

Abraham López^{1,2}, Laura Mendieta², Roger Prades³, Soledad Royo³, Teresa Tarragó^{2,3} & Ernest Giralt^{*1,2}

¹Department of Organic Chemistry, University of Barcelona, Barcelona, Spain

²Institute for Research in Biomedicine (IRB Barcelona), Barcelona, Spain

³proteosSL, Barcelona, Spain

*Author for correspondence:

E-mail: ernest.giralt@irbbarcelona.org

**FUTURE
SCIENCE**

part of
fsg

Schizophrenia	Positive symptoms	Disorganized behavior Hallucinations Delusions	Antipsychotic drugs	Treatment
	Negative symptoms	Apathy Loss of motivation	Antipsychotic drugs	
	Mood disorders	Depression Anxiety Aggression	Antidepressants Antipsychotic drugs	
	Neurocognitive disorders	Impaired behavior Deficit in social cognition	No current treatment	

Figure 1. Classification of schizophrenia symptoms. The main pathological conditions are summarized in the left and central columns, and the updated current treatments are listed in the right column.

column), antipsychotic drugs (APDs) have been widely used to ameliorate positive symptoms. Classical APDs are designed to restore neurochemical imbalances and are therefore based on dopamine (DA) D_2 receptor antagonists [11]. The first reported antipsychotic effect of chlorpromazine in schizophrenic patients was made by Delay and Deniker early in 1952 [9]. Since then, much research effort has been devoted to APD drug discovery. Non- D_2 APDs have recently been addressed, but no effective results have been demonstrated to date [12]. Despite the efficacy of treatment with APDs, 30% of patients do not overcome the symptomatology of schizophrenia and over 60% continue to have residual symptoms [9]. In addition, the percentage of patients who abandon treatment must also be taken into account. The discontinuity of medication causes not only the reappearance of positive symptoms, but also a loss of APD effectiveness. This implies that patients who had previously responded to the treatment may not be able to overcome the symptomatology again [1].

One of the main drawbacks that limit schizophrenia therapies is the lack of efficacy of APDs for the treatment of cognitive impairment [13]. Patients are unable to improve their task capacity and abilities because of cognitive deficits, and consequently present difficulties in executive functions, working memory and attention levels. These lasting effects have a significant negative social and functional impact on patients. Moreover, another underlying inconvenience arises as a result of APD treatment, since the extra pyramidal side effects of these substances

can accentuate the brain damage and therefore favor secondary negative symptoms and cognitive deficits. To overcome these problems, researchers have proposed the combination of drugs [14]. Compounds targeting symptoms other than the positive ones could be complemented with APDs in order to fully cover all schizophrenic features. Cognition-enhancing drugs may be suitable for such combination treatment, overcoming these side effects by specifically modifying procognitive targets [11]. Although research pursuing the identification of such targets has increased in recent years, no target has yet been validated. Progress in this direction will improve schizophrenia treatment and the development of selective and effective drugs for the treatment of its symptoms. Among the new targets, POP (EC3.4.21.26) emerges as an innovative candidate for the treatment of the cognitive deficits associated with schizophrenia.

Pop: a promising target for the treatment of cognitive deficits of schizophrenia

In contrast to most of the APDs currently available, the inhibitors developed for POP show cognition-enhancing properties, and thus their use for treatment of schizophrenic patients could improve their functionality and social integration. In terms of pharmacodynamics and pharmacokinetics, the non- D_2 specific mechanism and good brain penetration of POP inhibitors are advantageous for the treatment of the cognitive symptoms associated with schizophrenia. Although the physiological role of this peptidase is still not clear, research into POP inhibitors may fill the gap regarding the untreated symptoms of schizophrenia.

■ Biological function of POP

POP is an 81-kDa monomeric serine peptidase. This enzyme specifically cleaves the post proline bonds in small peptides (usually ≤ 3 kDa). Although its crystal structure is well known [15], the enzymatic mechanism is still not fully understood. Nevertheless, recent studies indicate that this enzyme shows dynamic behavior and participates in a conformational sampling mechanism [16,17].

The physiological role of POP is elusive. Although this protease is ubiquitous in the human body, there is an increased concentration in the CNS [18,19]. Most *in vitro* POP substrates are proline-containing neuropeptides or peptide hormones present in the brain [20].

To some extent, the presence of proline protects these substances from general proteolytic degradation. For this reason, the first hypothesis on the *in vivo* role of POP proposed a specific prolyl hydrolytic function as the main mechanism of action. Indeed, dysregulation of POP hydrolytic activity was detected in the serum of patients suffering from mood disorders [21–23]. This observation provided the first indication of POP involvement in mental processes.

Nevertheless, these studies were only preliminary and did not provide any explanation for the neurochemical mechanism of POP. Although many *in vitro* POP substrates (substance P, neurotensin, arginine-vasopressin or thyrotropin-releasing hormone, among others) are involved in learning and mnemonic functions [24], it appears that the hydrolytic function of POP is not the main responsible for its physiological role. The initial hypothesis proposed that the hydrolytic activity of this protease controls neuropeptide levels, which, in turn, interact with the corresponding receptors. Therefore, POP inhibitors would prevent the catabolism of these neurotransmitters by blocking POP activity. The resulting increase in neuropeptide levels is in fact responsible for the effect on the cognitive functions [20,25,26]. In order to confirm this hypothesis, a series of studies were carried out in rats that had been treated with POP inhibitors. The administration of two POP inhibitors led to an increase in the concentration of substance P in the hippocampal and cortical tissue of these animals [27,28]. Raising the concentration of this neuropeptide might alter cholinergic neurotransmission by stimulating the release of acetylcholine (ACh) in these areas. Furthermore, thyrotropin-releasing hormone was decreased in cerebral damaged rats, while POP inhibitors reestablished the levels of this hormone [29]. Together, these studies showed poor effects over the neuropeptide levels [29], or even reached contradictory results [27,28]. The main reason for the divergence of the results is the discordant spatial association between the intracellular POP and its extracellular substrates, a question that still remains to be clarified [30]. In a different direction, behavioral experiments in rats showed incoherent results in terms of spatial memory: the increase in spatial memory in young scopolamine-induced amnesic rats was not evidenced in aged animals [31]. More recently, the effect of inhibitors was found

to be strongly dependent on the brain region analyzed [32]. The *in vivo* extracellular levels of ACh and DA in rat striatum were evaluated after the administration of two distinct POP inhibitors. Surprisingly, in contrast to what had been observed in the hippocampus and cortex, ACh levels were only slightly decreased after the administration of these compounds; neither did DA levels show any significant variation. To the best of our knowledge, this is the only study to dismiss any major role of the cholinergic and dopaminergic systems in the cognitive response to POP inhibition published so far. Taken together, these observations do not show any direct relationship between POP inhibition and neuropeptide-mediated receptor systems. Nevertheless, other studies performed over other neurotransmitter systems, such as GABA, glutamatergic or 5-HT, are lacking. These would provide new insights into POP *in vivo* mechanism.

In another research line, a study of Harwood and co-workers on the effect of lithium on the inositol triphosphate (IP_3) pathway found that lithium-resistant mutants of *Dictyostelium discoideum* (a soil-living amoeba) lacked the POP gene [33]. This observation was accompanied by an increased IP_3 concentration, thereby suggesting a regulatory role of POP in IP_3 signaling. Moreover, it was further demonstrated that the inhibition of POP abolished the effects of lithium, carbamazepine and valproic acid (three mood stabilizers that are based on inositol depletion [34]). IP_3 is a second messenger involved in neuropeptide signal transduction. The increased concentration of IP_3 in the cytosol favors the binding to the corresponding receptors in the endoplasmic reticulum and increases the release of Ca^{2+} from intracellular stores. In turn, this activates the protein kinase cascades, which finally affect learning and mnemonic functions. Further studies in astrogloma U343 cell lines found an inverse correlation between POP proteolytic activity and IP_3 concentration, thus confirming this hypothesis. Proteolytic activity was modulated by generating POP antisense cell lines, or administering the POP inhibitor Fmoc-Ala-Pyrr-CN [35]. Only one *in vivo* study exploring the correlation between IP_3 levels and inhibition of POP has been performed after administration of POP inhibitor KYP-2047. Unfortunately, no significant variation of IP_3 levels upon inhibition of POP was found in the cortex and the hippocampus of rats [31].

■ The role of POP in protein–protein interactions

The interaction of POP with other proteins, independent of its catalytic activity, can explain, to some extent, some of the observed effects on neuron growth, synaptic plasticity and protein secretion. In a study performed with the above-mentioned U343 cells, immunofluorescence showed that POP binds to tubulin. This observation supports the notion that POP participates in microtubule processes, such as protein secretion. POP antisense cell lines showed a marked decrease in protein secretion, while secretion was less reduced when this enzyme was inhibited. This finding suggests that the interaction of POP with tubulin, independent of the catalytic activity of the protease, is the main factor responsible for the regulation of protein secretion [36]. Nevertheless, the localization, activity and mRNA levels of POP showed several inconsistencies [30]. Although the main distribution of POP mRNA was in areas rich in neuropeptide receptors within the brain nuclei [37], a high concentration of mRNA was found in the hypothalamus and cerebellum. Nevertheless, high POP activity was not detected in those regions [38]. The elevated activity of POP in the cortex and other more evolved brain areas points to the involvement of this protease in mnemonic, cognitive and learning functions.

According to preliminary studies of Williams and co-workers, POP inhibitors reverse the effects of valproic acid, lithium and carbamazepine (widely used mood stabilizers) in growth cone dynamics [34]. These results prompted further research to elucidate the function of POP in neuron growth. They generated POP knock-out mice and examined *in vivo* the effect of POP-deficiency/absence on neuronal development. POP-null mutant mice showed altered cone growth dynamics, with a greater growth area and fewer collapsed cones. This effect was observed in wild-type-neurons when mood stabilizers were administered and was reversed by treatment with POP inhibitors. Moreover, POP catalytic activity was found not to be essential for the effects observed on neuron growth, since transfection with either catalytically active or inactive POP restored the wild-type phenotype. Changes in POP dynamics and conformation as a consequence of inhibitor binding likely led to alterations in the interaction patterns with other proteins [17], a fact which is independent on the enzymatic

activity. Furthermore, GAP43 has been found to be an interaction partner of POP. This protein is a regulator of growth cone dynamics and calcium-dependent synaptic plasticity. This finding is particularly relevant since it is plausible that the interaction between GAP43 and POP alters signaling in mature neurons in the neocortex and limbic system, which are involved in mood control [39]. Nevertheless, the interaction between the two proteins is weak in physiological conditions, and the effects on growth cones probably arise from axonal transport or through the phosphoinositide pathway [40]. Interestingly, in other research lines, the expression of POP in cerebellar granule cells was found to be concentrated mainly around the cytosol and neuronal projections. Curiously, in early stages of development, POP expression around the nucleus was upregulated. Expression peaked before cellular differentiation, but it also increased in late stages and concentrated around the nucleus. These observations highlighted the involvement of POP in neuronal differentiation, as well as the neuroprotective effects of POP inhibitors [41].

Finally, it has recently been reported that POP interacts with α -synuclein, one of the main components of Lewy bodies present in Parkinson's disease [42,43], and also colocalizes with α -synuclein, tau protein and β -amyloid (A β) [44]. Although this effect on α -synuclein aggregation is out of the scope of this review, this research line is receiving increasing interest for novel applications of POP inhibitors for the treatment of synucleinopathies.

■ Reported cognition-enhancing properties of POP inhibitors in animal models & clinical trials

The next section describes the POP inhibitors that have been tested in animals or in clinical trials in humans (TABLES 1 & 2). The physiological effects of these inhibitors, together with the cognitive, mnemonic and behavioral improvements induced, are systematically reported. The updated development status for each case has been obtained from the Thomson Reuters Integrity™ database. Pharmacodynamics, pharmacokinetics and toxicity are also mentioned, when appropriate. For more detailed information, the reader is referred to the literature.

JTP-4819

As of 1997, this compound has entered Phase II clinical trials [45,201]. The safety and

Table 1. Preclinical and clinical studies carried out with POP inhibitors.

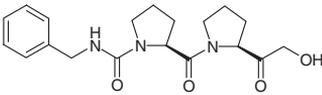
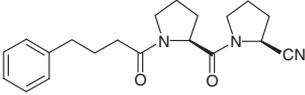
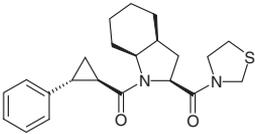
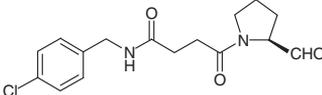
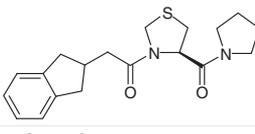
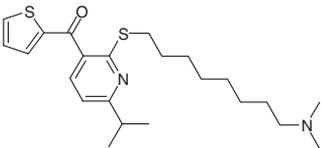
POP inhibitor	Model	Test	Result
Preclinical			
KYP-2047 Treatment of cognition disorders, University of Eastern Finland	<i>In vitro</i> and <i>in vivo</i>	BBB permeability	Good BBB permeability High plasma clearance High plasma protein unspecific binding
	Scopolamine-treated rats	Morris water maze task	Improvement in spatial memory Slight improvement in motor abilities
	Transgenic mice (A30P/A53T)	α -synuclein aggregation	Reduction of α -synuclein aggregation Reduction of soluble α -synuclein
Y-29794 Treatment of cognition disorders, Mitsubishi Tanabe Pharma	Rats, dogs and monkeys	BBB permeability	Good BBB permeability
	Senescence accelerated mouse	β -amyloid plaques	Inhibition of β -amyloid deposits formation
Phase I			
S-17092 Treatment of cognition disorders, Servier	Neurotoxin MPTP-treated monkeys	Variable delayed reponse Delayed alteration Delayed matching-to-sample Visual discrimination tasks	Improvement of cognitive task performance
	Phase I study	Plasma levels Electroencephalogram	No toxic effect Good toleration Inhibitor penetration to the brain Improvement in verbal function
Z-321 Treatment of Alzheimer's dementia, Zeria	Phase I study	Plasma levels	No accumulation in blood
Phase II			
JTP-4819 Treatment of cognition disorders, Mitsubishi Tanabe Pharma – Japan Tobacco	<i>In vitro</i> and <i>in vivo</i>	BBB permeability	Poor BBB permeability
	Scopolamine-treated rats	Passive avoidance	Improvement in special memory
	Scopolamine-treated rats	Passive avoidance Morris water maze task	Reduced escape time
ONO-1603	<i>In vitro</i>	Neuron survival Apoptosis of cortical and cerebellar granule cells	Increase of neuron survival upon induced cell death Delay in age-induced apoptosis
	Scopolamine-treated cells	–	Improvement in mnemonic and learning

BBB: Blood–brain barrier.

pharmacokinetics of this substance was examined in 28 healthy male volunteers, showing no important adverse effects. Analysis of plasma and urine only detected an increase of cholinesterase activity, probably attributable to the POP inhibitory activity. In general, JTP-4819 showed acceptable pharmacodynamics and pharmacokinetic properties. However, there have been no reports on the results derived from these studies. JTP-4819 has been extensively studied in several rodent models. The first experiments were carried out by Toide and co-workers in scopolamine (an anticholinergic agent)-induced amnesic rats. In that study, the one-trial passive avoidance test showed an improvement in spatial memory [27]. Furthermore, middle cerebral artery occlusion was caused surgically. The administration of JTP-4819 increased latency in the passive avoidance test, whereas the escape time in the Morris

water maze task was reduced [29]. Moreover, the effects on cortical and hippocampal neuropeptide levels and amyloid metabolism, with a clear focus on the prevention of Alzheimer's disease were also determined [46]. Shinoda and co-workers suggested that the APP hydrolytic fragments cleaved by the β -secretases might be POP substrates, which would represent a new route for the generation of A β (1–42). They supported this hypothesis by measuring the cleavage of N- and C-terminal fragments of A β in NG108-15 cells. Interestingly, it was found that this process was decreased after treatment of cells with JTP-4819, highlighting the utility of this compound preventing the generation of A β in Alzheimer's disease patients [47]. The pharmacokinetics of this inhibitor and of KYP-2047 (see below) was assessed by determining their *in vitro* and *in vivo* blood–brain barrier (BBB)

Table 2. Summary of POP inhibitors that have been tested in animals or have reached clinical trials.

Name	Chemical structure	Indication	Highest Phase	Tested in	Ref.
JTP-4819		Treatment of cognition disorders: cognitive enhancer in scopolamine-induced amnesia, neuroprotection, Alzheimer's disease	Phase II	Rodents, humans	[27,45]
KYP-2047		Treatment of cognition disorders: cognitive enhancer in scopolamine-induced amnesia; treatment of Parkinson's disease: cognitive deficits associated with Parkinson disease, reduction of α -synuclein aggregation process	Preclinical	Rodents	[31,48]
S-17092		Treatment of cognition disorders: cognitive deficits associated with aging and/or neurodegenerative diseases	Phase I	Rodents, monkeys, humans	[50,52,82]
ONO-1603		Treatment of cognition disorders: cognitive enhancer in scopolamine-induced amnesia, neuroprotection	Phase II	Rodents	[55,201]
Z-321		Treatment of cognition disorders: dementia associated to Alzheimer's disease	Phase I	Rodents, humans	[58,59]
Y-29794		Treatment of cognition disorders: treatment of Alzheimer's disease	Preclinical	Rodents, dogs, monkeys	[60,61]

permeability. JTP-4819 showed poor BBB permeability. On the contrary, KYP-2047 showed satisfactory brain permeability, similar to that of other brain-targeted drugs, but also considerable plasma clearance and plasma protein unspecific binding [48]. To the best of our knowledge, this is the only study to address the BBB permeability of POP inhibitors.

KYP-2047

Few *in vivo* experiments have been carried out with this inhibitor. The first experiment showed only a slight improvement in the spatial memory of young scopolamine-induced amnesic rats, as shown by the results from the Morris water maze task [31]. Furthermore, this study detected inconsistencies between the levels of several neuropeptides and IP_3 in response to treatment with this inhibitor. Further studies using the radial arm maze have been performed to clarify the effect of a single dose of KYP-2047 in scopolamine-treated rats. In that test, no

significant evidence of mnemonic improvement was observed; in fact, only a slight enhancement of motor skills was detected [49]. As mentioned above, Jalkanen *et al.* have found that treatment with KYP-2047 has little or no effect on striatal ACh levels [32]. This evidence has led to a new hypothesis to explain the observed positive cognitive effects which, according to the authors, would not be mediated through the cholinergic system. This contradicts the hypothesis introduced by Toide and co-workers on the restoration of ACh levels after POP inhibition [27]. In another line of research, recent *in vitro* studies have demonstrated that POP accelerates the aggregation of α -synuclein. The treatment of transgenic mice (A30P/A53T) with KYP-2047 for 5 days reduced the amount of α -synuclein immunoreactivity and soluble α -synuclein protein in the brain. These observations support the applicability of this inhibitor for the prevention of synucleinopathies rather than as a cognition enhancer [43]. This new application of POP

inhibitors for the treatment of Parkinson's disease, although out of the scope of this review, might represent a promising treatment for synucleinopathies.

S-17092

This substance reached Phase I clinical trials in humans, which consisted of a double-blind, randomized, placebo-controlled study [50]. The pharmacokinetics of this inhibitor was evaluated in the plasma of aged healthy humans by LC and MS, after single and repeated oral administration. No toxic effects and good toleration were found. In order to elucidate the pharmacodynamics of this compound, a quantitative electroencephalogram (EEG) of aged healthy patients was performed, revealing significant alterations in subjects receiving S-17092, in contrast to those on the placebo. This finding suggests that the inhibitor enters the brain. Nevertheless, no dose-dependence was observed in the EEG. A further study measured the quantitative EEG during a 10-day treatment with S-17092 [51], and only slight improvements in certain verbal functions were observed. To test S-17092 as an agent for the treatment of the cognitive deficits related to Parkinson's disease, this substance was also used to palliate the cognitive impairment induced by chronic administration of neurotoxin MPTP in monkeys. In particular, the aim of the study was to find an alternative treatment for cognitive deficits associated with Parkinson's disease, since classical therapies targeting only the dopaminergic system are not effective. Interestingly, S-17092 improved cognitive task performance in these animal models, although the study was not specific enough to distinguish between attention or mnemonic functions [52]. On a different direction, Petit and co-workers performed studies in human cell lines in the presence of this inhibitor in order to test the effect of S-17092 in the production of A β [53]. Contrarily to what had been previously suggested for JTP-4819 [46], POP was demonstrated to be unable to cleave β APP *in vitro*. Moreover, the presence of S-17092 did not have any significant effect on the A β 40 and A β 42 formation in HEK293 cells overexpressing wild-type β APP [53].

ONO-1603

According to the information available on Thomson Reuters Integrity, this POP inhibitor has reached Phase II clinical trials for Alzheimer's [201]. However, no information has been reported on the results of these trials. Besides

that, this inhibitor was tested in brain-damaged rats [54]. Amnesia was induced in these animals by treatment with scopolamine. The mnemonic and learning lesions were ameliorated after administration of the inhibitor. Moreover, Katsube and co-workers studied the neuroprotective effects of ONO-1603 against age-induced apoptosis of cortical and cerebellar granule cells [55,56]. In particular, they found a considerable delay in the apoptosis of these cells, which was probably mediated by the downregulation of GADPH mRNA and protein expression. This result was coherent with previous studies by the same group in this neuronal model, in which ONO-1603 promoted neuronal survival against potassium deprivation-induced cell death [57].

Z-321

The Phase I safety evaluation of this compound was carried out on young and healthy male volunteers [58]. The plasma and urine concentration of the inhibitor, together with its main metabolites, were determined after single and multiple dose experiments. Regarding the multiple dosing, maximum concentration for this compound was reached at 1.92 h for the first dose and 1.33 h for the last dose (dose 13). Other parameters, such as the maximal concentration and half-life, were not significantly altered during the administration. These observations suggest no influence of the dosage on the pharmacokinetics of Z-321. Moreover, no plasma drug accumulation trend was observed. Nevertheless, plasma cholinesterase activity was increased, similarly to what occurred after the administration of JTP-4819 [45]; this finding was attributed to the POP inhibitory capacity of Z-321. Finally, the plasma levels of thyroid hormones and neuropeptides were examined after inhibitor administration. No significant alterations were detected [59]. In summary, these results indicated acceptable pharmacodynamics and pharmacokinetics for Z-321, as well as nontoxicity. Unfortunately, to our knowledge, no cognitive enhancement or learning improvements for Z-321 have been reported yet.

Y-29794

Initial studies with this substance showed good penetration to the brain in rats, dogs and monkeys [60]. The previously observed effects of POP inhibitors on neuropeptide levels had pointed to the possible involvement of POP function in amyloid progression in the hippocampal region. In order to validate this hypothesis,

senescence-accelerated mice were produced, and the granular structures were observed as a function of time. Immunostaining revealed a significant reduction in the A β -positive cells after the administration of this inhibitor, thus indicating that this non-peptide inhibitor prevents the formation of A β deposits [61]. To our knowledge, this is the only experiment carried out with a POP inhibitor on an animal model of Alzheimer's disease. Nevertheless, no cognitive performance tests have been performed with Y-29794.

After reviewing all the cognitive studies mentioned above, it is clear that additional *in vivo* investigations are still required to understand and clarify the cognition-enhancing properties of POP inhibitors. The molecular interactions of POP in the brain, both in normal and pathological conditions, are poorly understood, thus clearly hindering the selection of the most appropriate animal model and cognitive assays. In addition, an exhaustive biodistribution study of the inhibitors under development is required in order to detect the main tissues targeted [62].

Many of the animal models used to date are based on the induction of amnesia with scopolamine or by causing brain injuries (e.g., temporal middle cerebral artery occlusion), thus providing a model with cognitive impairment; however, these systems are far from simulating a true pathological state. The indication of inhibitors that underwent Phase II clinical trials was for the treatment of Alzheimer's disease, a highly disabling disorder associated with great cognitive impairment. The clinical trials required for Alzheimer's disease are highly complex because of the intrinsic nature of the disease, thus slowing down the development of new drugs.

As it can be appreciated, none of the studies performed to date are specifically related to cognitive deficits associated with schizophrenia or with any similar psychotic condition. Therefore, translational animal models based on the same neurobiological systems altered during schizophrenia would be particularly useful in the translation of this disease to humans [63]. Although efficient APDs are focused on the dopaminergic D₂ systems (known as first-generation antipsychotics), the dysfunction of other systems might play a crucial role in the cognitive deficits associated with schizophrenia. Therefore, some compounds, called second-generation antipsychotics, have been developed to act in other nonspecific systems in addition to D₂. In this regard, the inhibition of glutamatergic receptors produces a

psychotic state similar to that of schizophrenia [64]. Nevertheless, behavioral studies based on glutamatergic receptors have not shown significant advantages over the traditional D₂ models. Neurodevelopmental models can explain some of the neurochemical alterations observed in schizophrenia [65]. A more complete view of the genetic and environmental factors of schizophrenia could be provided by combining these neurodevelopmental models with genetic models that target characteristic genes involved in the pathophysiology [63].

Peptide & peptidomimetic POP inhibitors as drug-like compounds for cognitive enhancement

Peptides are a family of compounds that are widely present in all living organisms, and have broad biological functions. Due to their nature, peptides can readily mimic the protein surface or cavity with which they interact [66], and thus represent a straightforward approach for the design of ligands. Peptides allow multiple anchor points on a protein, which, in the present case, implies a clear advantage for greater selectivity and specificity between proteases of the same subfamily. Both the chemical and stereochemical space available to peptide ligands are enormous. The side chains of natural amino acids represent a wide variety of functional groups, different degrees of polarity, potential for electrostatic interactions and other types of interactions. Moreover, the use of non-natural or D-amino acids, isosteres or conformationally restricted building blocks, allows further exploration of the chemical space and the best fitting with the pharmacophore or the binding site. From a conformational point of view, due to their plasticity, peptide molecules are able to adopt a range of distinct structures. An additional advantage of these molecules is the availability of methods and techniques for their solid-phase synthesis, which offer great synthetic versatility and provide access to a variety of ligands without intensive synthetic efforts.

Most POP inhibitors are peptidomimetic compounds that fit into the active site and the surrounding cavities [15]. Therefore, correct complementarity with the pharmacophore is one of the main requirements for the successful design of POP inhibitors. The structure–activity relationship ensures satisfactory potency and selectivity against other serine peptidases [67,68]; however, for acceptable pharmacokinetics, lipophilicity must provide appropriate BBB

permeability and low clearance. The exploration of other POP binding sites might bring about innovative scaffolds. Therefore, the design of peptide compounds that interact with POP might not be directed only by ligand-based design. For example, natural product screening found some small molecules that show POP inhibitory activity, which is exerted possibly by interaction with an allosteric site [69,70], while others interact with other cavities surrounding the catalytic center [71].

The pharmacophore of POP inhibitors can be described as a function of the enzymatic cavities involved in the interaction, according to the nomenclature of Berger and Schechter [72,73].

Thus, the P1', P1, P2 and P3 positions describe the POP pharmacophore (FIGURE 2). One of the first POP inhibitors discovered was Z-prolyl-prolinal (ZPP), a molecule formed by two consecutive proline residues terminated with a hydrophobic group in the N-terminus. The carbonyl groups of the amide bonds are required to establish hydrogen bonds with several residues of the enzyme, and the C-terminal formyl group forms a hemiacetal with the Ser554 of the catalytic center, yielding a covalent reversible inhibitor that acts as a transition state analog [15]. This structure has been taken as a reference compound for the ligand-based development of most POP inhibitors, thereby maintaining a similar

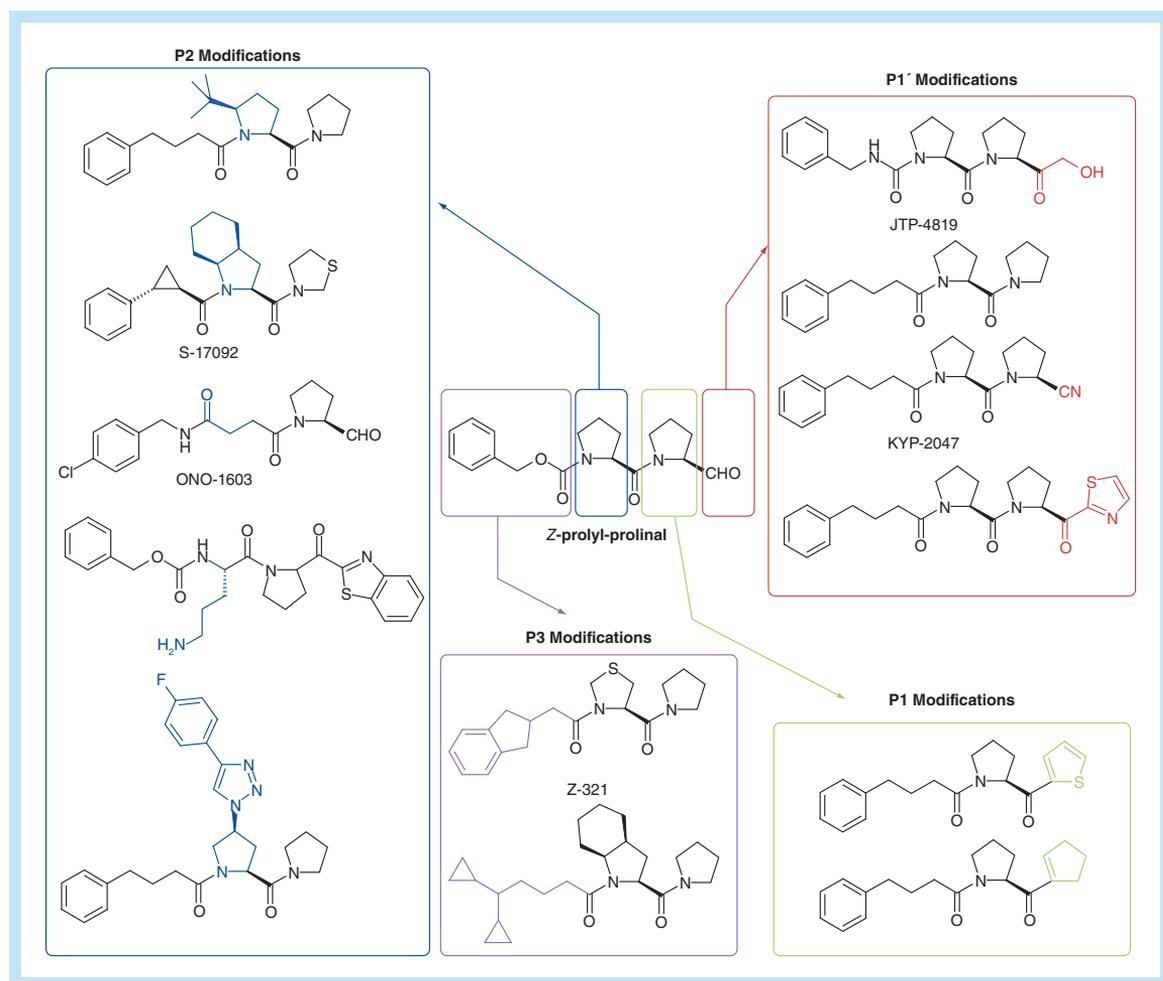


Figure 2. Examples of POP inhibitors obtained through modifications of the canonical compound Z-prolyl-prolinal in different positions.

binding mode. In some cases, the original peptide structure has been so extensively modified that they are difficult to classify (e.g., Y-29794). Nevertheless, this section does not intend to elucidate the chemical structures of all compounds in detail, as complete descriptive reviews have been published [67,74,75]. Instead, only the general structural requirements for each pharmacophore position will be discussed.

Electrophilic groups in the P1' position forms a covalent bond with Ser554 in the catalytic center. Less electrophilic groups in P1', such as ketones, need to be activated by electron-withdrawing groups, for instance α -hydroxymethylketones [27] or heterocycles [76,77]. Fluorinated chains, which might also improve the BBB permeability of the compounds, have also been studied [101,102]. A reactive group of particular interest is cyano, since it decreases lipophilicity but forms an imino ether bond with the serine residue; nevertheless, the dissociation constant is significantly increased compared with that of carbonyl compounds [78,79]. Other groups lead to irreversible inhibitors, through nucleophilic displacement of halogen atoms. However, reactive groups in P1' affect the stability of the inhibitors in serum and cause unspecific binding and poor selectivity.

The P1 position is specific for cyclic structures, typically proline, pyrrolidine or thiazolidine [80], which probably stack with the Trp595 of the catalytic cavity [15]. Studies carried out by Tsuru *et al.* pointed out that high potency can be achieved by placing a thiazolidine ring in the P1 position [81]. The substitution by an α,β -unsaturated ring or an aromatic ring also leads to compounds with good inhibitory properties [82].

P2 position is not sterically or electronically demanding. The literature reveals toleration of a wide variety of structures, for instance linear amino acids (both charged and uncharged), proline, and simple or double rings [81]. Recently, the modification of the proline P2 position with new substituents has afforded potent and selective compounds. Orion Corp. studied the rigidification of the P3–P2 peptide bond towards the *cis* conformation by the presence of bulky substituents in the 5 position of the P2 proline and developed potent inhibitors [83,103]. Other groups have addressed substituents that extend into a new binding site beyond P2 and provide a feasible alternative to the typical reactive groups in P1'. Moreover, in those cases, improved physicochemical properties are achieved (solubility and lipophilicity) [84].

Finally, P3 positions are usually occupied by hydrophobic groups, usually aromatic compounds separated by an aliphatic spacer chain. Morain *et al.* developed new scaffolds suitable for P3, for instance alkyldicycpropylcarbonyl or 2-phenylcycpropylcarbonyl groups [103], which is exemplified by compound S-17092, tested in clinical trials [50]. Orion Corp. patented a series of symmetric or pseudosymmetric compounds, consisting of the aminoacylpyrrolidinediamides of dicarboxylic acids, which showed good potencies. Extensive examples of structures in P3 can be found in several reviews [67,74].

■ BBB: a bottleneck for the development of brain-targeted compounds

As discussed, all the indications of the POP inhibitors developed so far are mainly for the treatment of CNS-related disorders. This implies that these inhibitors must have the capacity to cross the BBB in order to exert the desired effect. The BBB is a highly selective biological and physical barrier that controls and regulates the passage of compounds from the blood stream to the CNS, thus insulating the CNS from the other parts of the body and providing an optimal environment for neuronal function. This protective role is the major bottleneck for the delivery of therapeutic agents to the brain.

The POP inhibitors published to date generally have a relatively low MW (from 314 to 420 Da) and are usually highly lipophilic (their logP ranges from 0.66 to 6). Two features that make them potentially good candidates to overcome the BBB by passive diffusion include a non-energy dependent and non-saturable transport mechanism. However, BBB permeability has not been optimized for most of these inhibitors and, more importantly, *in vivo* pharmacodynamic, pharmacokinetic and exhaustive toxicological studies have either not been performed or are insufficient.

The implementation, in the preclinical Phase, of a number of *in vitro* tools such as logP measurements, immobilized artificial membrane chromatography [85], and permeability testing using parallel artificial membrane permeability assay (PAMPA) [70,86] or cell based assays [87], definitely could contribute in the optimization and refinement process of the absorption, distribution, metabolism and excretion properties of the molecules, thus maximizing the success in the development of new POP inhibitors.

Nevertheless, these tools are just an initial approach to the complexity of *in vivo* experiments. The obtained results can be used as a

cut-off to select the most relevant compounds regarding to its permeability across the biological barriers, especially across the BBB, for further experiments. After this first approach, *in vivo* permeability assays of the most promising compounds should be performed in order to provide information about the bioavailability, distribution and clearance of the molecules of interest into the CNS and the areas of the brain in which the compound is accumulated. This will bring light about the putative effect of the POP inhibitor being tested.

Future perspective

An overview of the current treatments for schizophrenia highlights that some important medical needs are still unmet. One of the main issues is the lack of adequate treatment for cognitive deficits associated with schizophrenia. The search for new therapeutic targets for the development of treatments for these symptoms is of great interest. In this regard, inhibition of POP may be promising.

Nevertheless, extensive testing in order to validate POP as a drug target for the treatment of cognitive deficits of schizophrenia is still required. Although the discovery and development of POP inhibitors have been a hot topic for years (first published data are from 1989), and positive outcomes have been obtained for some inhibitors regarding their use as cognitive enhancers, it is surprising that none of these molecules has yet reached the market. Most of the POP inhibitors described have been discontinued in the preclinical development stage. S-17092 and Z-321 have undergone Phase I clinical trials, and JTP-4819 and ONO-1603 have reached Phase II. Unfortunately, little public information is available regarding the concerns and reasons that led to the discontinuation of these clinical trials. In addition, no specific study linking POP inhibition and cognitive improvement in schizophrenia has been performed so far, thus remaining an unexplored field.

In our opinion, POP inhibitors have failed because of a combination of factors during their development and evaluation, rather than the lack of true positive biological effects [62] namely:

- Known POP inhibitors comprise a narrow family of compounds, mainly peptidomimetics that have been designed from the modification of the canonical compound ZPP;
- Until now, potency has been the main factor to be optimized. Other parameters, especially

BBB permeability, have not always been taken into account;

- The lack of suitable animal models hampers the evaluation of cognitive improvement through administration of POP inhibitors;
- The indication selected for clinical studies has, in our opinion, limited the progression of those POP inhibitors that have entered clinical phases.

Regarding drug discovery, the development of POP inhibitors has been based almost exclusively on the modification of the canonical peptidomimetic compound ZPP in order to achieve better fitting with the pharmacophore. This strategy does not take into account the development of other possible innovative (peptide) scaffolds with POP inhibitory activity.

In another context, potency and selectivity against other prolylpeptidases have been the main parameters subjected to optimization. These parameters are poor and unrealistic simplifications for an intracellular peptidase targeted in the brain. Several other relevant factors must be taken into account. First of all, poor BBB permeability probably prevents these compounds from entering the brain, thus explaining the apparently incoherent *in vitro*–*in vivo* effects and accumulation of POP inhibitors in different brain areas. In this regard, biodistribution studies of the inhibitors under development may clarify the main tissues targeted. More significantly, few animal models are suitable for the study of the effects of these putative pro-cognition compounds. The models most frequently used to date comprise mainly brain-injured rodents, which are submitted to behavioral and cognitive tasks after administration of a POP inhibitor. In these pathological conditions, the anatomical structure of the brain is highly impaired and therefore the poor enhancement of cognition observed might not be representative.

The development of animal models of schizophrenia represents a huge challenge. Moreover, for the negative symptoms and the cognitive deficits related to schizophrenia, the modeling is even more complex because these symptoms are essentially observed only in humans. Translational animal models would be particularly useful to confirm whether a given drug candidate is the most appropriate to undergo clinical trials for this indication and consequently increase the chance of success. Therefore, we encourage the study with appropriate animal models for schizophrenia, in

order to provide a preclinical proof of concept of POP inhibitors in the treatment of cognitive deficits associated to schizophrenia.

The clinical trials performed so far with POP inhibitors have shown acceptable pharmacokinetics and pharmacodynamics and demonstrated no toxic effects. Nevertheless, these clinical trials were designed for Alzheimer's disease. In our opinion, this neurodegenerative disease is not a suitable indication for these cognitive enhancers. Alzheimer's disease is a highly disabling disease, which involves considerable and irreversible neuronal damage as a consequence of amyloidogenesis. The improvement of the cognitive functions is severely limited, and is not expected to occur through the administration of POP inhibitors. Proof of this is the lack of success of POP inhibitors during the clinical trials for the treatment of Alzheimer's disease. Therefore, other indications, such as mild cognitive impairment, may allow more promising results in the clinical trials. Selection of the appropriate indication for POP inhibitors is a difficult task because the molecular interactions of POP in the brain, both in normal and pathological conditions, are not fully understood.

In spite of these difficulties, POP inhibitors may have a promising future in the treatment of the cognitive deficits associated with schizophrenia. Efficient APDs have been developed focused on the dopaminergic D₂ system, as well as on other compounds that in addition act in nonspecific systems. Nevertheless, these APDs do not show cognition-enhancing properties. In fact, clinical antipsychotic trials in intervention effectiveness studies have shown that none of these substances are useful for the treatment of the cognitive impairment associated with schizophrenia. Moreover, they show strong extrapyramidal side effects that can aggravate the cognitive and negative symptoms [88].

Taken together, these observations highlight the need of cognitive enhancers for the treatment of schizophrenia. Novel nootropic agents should not be based solely on D₂ system, since the cognitive deficits associated with schizophrenia arise as a result of a complex dysfunction in which the dopaminergic system is not the main mechanism. The mechanism of action causing cognitive

deficits is complex and probably involves more than one receptor system. Despite the limitations regarding animal models and the need for further studies to elucidate the role of POP in other neurotransmitter systems, we believe POP inhibitors are promising agents for the treatment of the cognitive deficits associated with schizophrenia. Moreover, the combined administration of typical APDs and POP inhibitors as cognitive enhancers may be tested for the simultaneous treatment of both positive and cognitive symptoms. Furthermore, since POP inhibitors are involved in neural development, the co-administration of APDs and POP inhibitors could have neuroprotective effects. This would minimize the extrapyramidal effects related to the administration of neuroleptics based on DA D₂ antagonists [89].

In summary, the rational design of peptidic or peptidomimetic ligands, together with effective preclinical studies and adequate animal models, are expected to validate POP as a robust therapeutic target for the treatment of the cognitive deficits of schizophrenia.

Financial & competing interests disclosure

A López is employed by the Department of Organic Chemistry, University of Barcelona, and has received funding from the Instituto de Salud Carlos III and the Ministry of Science and Innovation. L Mendieta is employed by the Institute for Research in Biomedicine and has received funding from the Institute for Research in Biomedicine. R Prades and S Royo are researchers at Iproteos, and have received funding from Spanish Ministry of Economy (Inncorpora-TU, Inncorpora Torres Quevedo, Neotec from CDTI). T Tarragó is Iproteos CEO and employed by the Institute for Research in Biomedicine, and has received funding from the Institute for Research in Biomedicine and from Iproteos. E Giral is Professor at the Department of Organic Chemistry, University of Barcelona, and has received funding from the Institute for Research in Biomedicine, MCYT-FEDER (Bio2008-00799), and the Generalitat de Catalunya (XRB and Grup Consolidat). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

References

- Lewis DA, Lieberman JA. Catching up on schizophrenia: natural history and neurobiology. *Neuron* 28(2), 325–334 (2000).
- Howes OD, Kapur S. The dopamine hypothesis of schizophrenia: version III – the final common pathway. *Schizophr. Bull.* 35(3), 549–562 (2009).
- Abi-Dargham A, Laruelle M, Aghajanian GK, Charney D, Krystal J. The role of serotonin in the pathophysiology and treatment of schizophrenia.

- J. Neuropsychiatry Clin. Neurosci.* 9(1), 1–17 (1997).
- 4 Carlsson A, Waters N, Holm-Waters S, Tedroff J, Nilsson M, Carlsson ML. Interactions between monoamines, glutamate, and GABA in schizophrenia: new evidence. *Annu. Rev. Pharmacol. Toxicol.* 41, 237–260 (2001).
 - 5 Martin LF, Freedman R. Schizophrenia and the alpha7 nicotinic acetylcholine receptor. *Int. Rev. Neurobiol.* 78, 225–246 (2007).
 - 6 Benes FM, Berretta S. GABAergic interneurons: implications for understanding schizophrenia and bipolar disorder. *Neuropsychopharmacology* 25(1), 1–27 (2001).
 - 7 Merlotti E, Piegari G, Galderisi S. Cognitive impairment as a core feature of schizophrenia. *Minerva Psichiatr.* 14(1), 67–78 (2000).
 - 8 Elvevag B, Goldberg TE. Cognitive impairment in schizophrenia is the core of the disorder. *Crit. Rev. Neurobiol.* 14(1), 1–21 (2000).
 - 9 Delay J, Deniker P. Réactions biologiques observées au cours du traitement par l'chlorhydrate de déméthylaminopropyl-N-chlorophénothiazine. *CR Congr. Méd. Alién Neurol.* 50, 514–518 (1952).
 - 10 Burton S. Symptom domains of schizophrenia: the role of atypical antipsychotic agents. *J. Psychopharmacol.* 20(6 Suppl.), 6–19 (2006).
 - 11 Miyamoto S, Miyake N, Jarskog LF, Fleischhacker WW, Lieberman JA. Pharmacological treatment of schizophrenia: a critical review of the pharmacology and clinical effects of current and future therapeutic agents. *Mol. Psychiatry* 17(12), 1206–1227 (2012).
 - 12 Kucharska-Pietura K, Mortimer A. Can antipsychotics improve social cognition in patients with schizophrenia? *CNS Drugs* 27(5), 335–343 (2013).
 - 13 Kitchen H, Rofail D, Heron L, Sacco P. Cognitive impairment associated with schizophrenia: a review of the humanistic burden. *Adv. Ther.* 29(2), 148–162 (2012).
 - 14 Moller HJ. Clinical evaluation of negative symptoms in schizophrenia. *Eur. Psychiatry* 22(6), 380–386 (2007).
 - 15 Fulop V, Bocskei Z, Polgar L. Prolyl oligopeptidase: an unusual beta-propeller domain regulates proteolysis. *Cell* 94(2), 161–170 (1998).
 - 16 Kaszuba K, Rog T, Danne R *et al.* Molecular dynamics, crystallography and mutagenesis studies on the substrate gating mechanism of prolyl oligopeptidase. *Biochimie* 94(6), 1398–1411 (2012).
 - 17 Kichik N, Tarrago T, Claasen B, Gairi M, Miller O, Giralte E. 15N relaxation NMR studies of prolyl oligopeptidase, an 80 kDa enzyme, reveal a pre-existing equilibrium between different conformational states. *Chembiochem* 12(18), 2737–2739 (2011).
 - 18 Wilk S. Prolyl endopeptidase. *Life Sci.* 33(22), 2149–2157 (1983).
 - 19 Goossens F, De Meester I, Vanhoof G, Scharpe S. Distribution of prolyl oligopeptidase in human peripheral tissues and body fluids. *Eur. J. Clin. Chem. Clin. Biochem.* 34(1), 17–22 (1996).
 - 20 Garcia-Horsman JA, Mannisto PT, Venalainen JI. On the role of prolyl oligopeptidase in health and disease. *Neuropeptides* 41(1), 1–24 (2007).
 - 21 Reichelt KL, Hole K, Hamberger A *et al.* Biologically active peptide-containing fractions in schizophrenia and childhood autism. *Adv. Biochem. Psychopharmacol.* 28, 627–643 (1981).
 - 22 Maes M, Goossens F, Scharpe S, Meltzer HY, D'Hondt P, Cosyns P. Lower serum prolyl endopeptidase enzyme activity in major depression: further evidence that peptidases play a role in the pathophysiology of depression. *Biol. Psychiatry* 35(8), 545–552 (1994).
 - 23 Maes M, Goossens F, Scharpe S, Calabrese J, Desnyder R, Meltzer HY. Alterations in plasma prolyl endopeptidase activity in depression, mania, and schizophrenia: effects of antidepressants, mood stabilizers, and antipsychotic drugs. *Psychiatry Res.* 58(3), 217–225 (1995).
 - 24 Kovacs GL, De Wied D. Peptidergic modulation of learning and memory processes. *Pharmacol. Rev.* 46(3), 269–291 (1994).
 - 25 Cunningham DF, O'Connor B. Proline specific peptidases. *Biochim. Biophys. Acta* 1343(2), 160–186 (1997).
 - 26 Mannisto PT, Venalainen J, Jalkanen A, Garcia-Horsman JA. Prolyl oligopeptidase: a potential target for the treatment of cognitive disorders. *Drug News Perspect.* 20(5), 293–305 (2007).
 - 27 Toide K, Iwamoto Y, Fujiwara T, Abe H. JTP-4819: a novel prolyl endopeptidase inhibitor with potential as a cognitive enhancer. *J. Pharmacol. Exp. Ther.* 274(3), 1370–1378 (1995).
 - 28 Bellemere G, Morain P, Vaudry H, Jegou S. Effect of S 17092, a novel prolyl endopeptidase inhibitor, on substance P and alpha-melanocyte-stimulating hormone breakdown in the rat brain. *J. Neurochem.* 84(5), 919–929 (2003).
 - 29 Shinoda M, Matsuo A, Toide K. Pharmacological studies of a novel prolyl endopeptidase inhibitor, JTP-4819, in rats with middle cerebral artery occlusion. *Eur. J. Pharmacol.* 305(1–3), 31–38 (1996).
 - 30 Myohanen TT, Garcia-Horsman JA, Tenorio-Laranga J, Mannisto PT. Issues about the physiological functions of prolyl oligopeptidase based on its discordant spatial association with substrates and inconsistencies among mRNA, protein levels, and enzymatic activity. *J. Histochem. Cytochem.* 57(9), 831–848 (2009).
 - 31 Jalkanen AJ, Puttonen KA, Venalainen JI *et al.* Beneficial effect of prolyl oligopeptidase inhibition on spatial memory in young but not in old scopolamine-treated rats. *Basic Clin. Pharmacol. Toxicol.* 100(2), 132–138 (2007).
 - 32 Jalkanen AJ, Piepponen TP, Hakkarainen JJ, De Meester I, Lambeir AM, Forsberg MM. The effect of prolyl oligopeptidase inhibition on extracellular acetylcholine and dopamine levels in the rat striatum. *Neurochem. Int.* 60(3), 301–309 (2012).
 - 33 Williams RS, Eames M, Ryves WJ, Viggars J, Harwood AJ. Loss of a prolyl oligopeptidase confers resistance to lithium by elevation of inositol (1,4,5) trisphosphate. *EMBO J.* 18(10), 2734–2745 (1999).
 - 34 Williams RS, Cheng L, Mudge AW, Harwood AJ. A common mechanism of action for three mood-stabilizing drugs. *Nature* 417(6886), 292–295 (2002).
 - 35 Schulz I, Gerhartz B, Neubauer A *et al.* Modulation of inositol 1,4,5-trisphosphate concentration by prolyl endopeptidase inhibition. *Eur. J. Biochem.* 269(23), 5813–5820 (2002).
 - 36 Schulz I, Zeitschel U, Rudolph T *et al.* Subcellular localization suggests novel functions for prolyl endopeptidase in protein secretion. *J. Neurochem.* 94(4), 970–979 (2005).
 - 37 Bellemere G, Vaudry H, Mounien L, Boutelet I, Jegou S. Localization of the mRNA encoding prolyl endopeptidase in the rat brain and pituitary. *J. Comp. Neurol.* 471(2), 128–143 (2004).
 - 38 Irazusta J, Larrinaga G, Gonzalez-Maeso J, Gil J, Meana JJ, Casis L. Distribution of prolyl endopeptidase activities in rat and human brain. *Neurochem. Int.* 40(4), 337–345 (2002).
 - 39 Di Daniel E, Glover CP, Grot E *et al.* Prolyl oligopeptidase binds to GAP-43 and functions without its peptidase activity. *Mol. Cell. Neurosci.* 41(3), 373–382 (2009).
 - 40 Szeltner Z, Morawski M, Juhasz T *et al.* GAP43 shows partial co-localisation but no strong physical interaction with prolyl

- oligopeptidase. *Biochim. Biophys. Acta* 1804(12), 2162–2176 (2010).
- 41 Moreno-Baylach MJ, Felipe V, Mannisto PT, García-Horsman JA. Expression and traffic of cellular prolyl oligopeptidase are regulated during cerebellar granule cell differentiation, maturation, and aging. *Neuroscience* 156(3), 580–585 (2008).
 - 42 Brandt I, Gerard M, Sergeant K *et al.* Prolyl oligopeptidase stimulates the aggregation of alpha-synuclein. *Peptides* 29(9), 1472–1478 (2008).
 - 43 Myohanen TT, Hannula MJ, Van Elzen R *et al.* A prolyl oligopeptidase inhibitor, KYP-2047, reduces alpha-synuclein protein levels and aggregates in cellular and animal models of Parkinson's disease. *Br. J. Pharmacol.* 166(3), 1097–1113 (2012).
 - 44 Hannula MJ, Myohanen TT, Tenorio-Laranga J, Mannisto PT, Garcia-Horsman JA. Prolyl oligopeptidase colocalizes with alpha-synuclein, beta-amyloid, tau protein and astroglia in the post-mortem brain samples with Parkinson's and Alzheimer's diseases. *Neuroscience* 242, 140–150 (2013).
 - 45 Umemura K, Kondo K, Ikeda Y, Kobayashi T, Urata Y, Nakashima M. Pharmacokinetics and safety of JTP-4819, a novel specific orally active prolyl endopeptidase inhibitor, in healthy male volunteers. *Br. J. Clin. Pharmacol.* 43(6), 613–618 (1997).
 - 46 Toide K, Okamiya K, Iwamoto Y, Kato T. Effect of a novel prolyl endopeptidase inhibitor, JTP-4819, on prolyl endopeptidase activity and substance P- and arginine-vasopressin-like immunoreactivity in the brains of aged rats. *J. Neurochem.* 65(1), 234–240 (1995).
 - 47 Shinoda M, Toide K, Ohsawa I, Kohsaka S. Specific inhibitor for prolyl endopeptidase suppresses the generation of amyloid beta protein in NG108–15 cells. *Biochem. Biophys. Res. Comm.* 235(3), 641–645 (1997).
 - 48 Jalkanen AJ, Hakkarainen JJ, Lehtonen M *et al.* Brain pharmacokinetics of two prolyl oligopeptidase inhibitors, JTP-4819 and KYP-2047, in the rat. *Basic Clin. Pharmacol. Toxicol.* 109(6), 443–451 (2011).
 - 49 Peltonen I, Jalkanen AJ, Sinerva V, Puttonen KA, Mannisto PT. Different effects of scopolamine and inhibition of prolyl oligopeptidase on mnemonic and motility functions of young and 8- to 9-month-old rats in the radial-arm maze. *Basic Clin. Pharmacol. Toxicol.* 106(4), 280–287 (2010).
 - 50 Morain P, Robin JL, De Nanteuil G, Jochemsen R, Heidet V, Guez D. Pharmacodynamic and pharmacokinetic profile of S 17092, a new orally active prolyl endopeptidase inhibitor, in elderly healthy volunteers. A Phase I study. *Br. J. Clin. Pharmacol.* 50(4), 350–359 (2000).
 - 51 Morain P, Boeijinga PH, Demazieres A, De Nanteuil G, Luthringer R. Psychotropic profile of S 17092, a prolyl endopeptidase inhibitor, using quantitative EEG in young healthy volunteers. *Neuropsychobiology* 55(3–4), 176–183 (2007).
 - 52 Schneider JS, Giardiniere M, Morain P. Effects of the prolyl endopeptidase inhibitor S 17092 on cognitive deficits in chronic low dose MPTP-treated monkeys. *Neuropsychopharmacology* 26(2), 176–182 (2002).
 - 53 Petit A, Barelli H, Morain P, Checler F. Novel proline endopeptidase inhibitors do not modify Abeta40/42 formation and degradation by human cells expressing wild-type and swedish mutated beta-amyloid precursor protein. *Br. J. Pharmacol.* 130(7), 1613–1617 (2000).
 - 54 Katsube N, Tateishi N, Matsushita Y. Anti-amnesic effect of ONO-1603, a novel prolyl endopeptidase inhibitor. *Jpn J. Pharmacol.* 49, 91–98 (1989).
 - 55 Katsube N, Sunaga K, Aishita H, Chuang DM, Ishitani R. ONO-1603, a potential antedementia drug, delays age-induced apoptosis and suppresses overexpression of glyceraldehyde-3-phosphate dehydrogenase in cultured central nervous system neurons. *J. Pharmacol. Exp. Ther.* 288(1), 6–13 (1999).
 - 56 Katsube N, Ishitani R. A review of the neurotrophic and neuroprotective properties of ONO-1603: comparison with those of tetrahydroaminoacridine, an antedementia drug. *CNS Drug Rev.* 6, 21–34 (2000).
 - 57 Katsube N, Sunaga K, Chuang DM, Ishitani R. ONO-1603, a potential antedementia drug, shows neuroprotective effects and increases m3-muscarinic receptor mRNA levels in differentiating rat cerebellar granule neurons. *Neurosci. Lett.* 214(2–3), 151–154 (1996).
 - 58 Umemura K, Kondo K, Ikeda Y *et al.* Pharmacokinetics and safety of Z-321, a novel specific orally active prolyl endopeptidase inhibitor, in healthy male volunteers. *J. Clin. Pharmacol.* 39(5), 462–470 (1999).
 - 59 Miura N, Shibata S, Watanabe S. Increase in the septal vasopressin content by prolyl endopeptidase inhibitors in rats. *Neurosci. Lett.* 196(1–2), 128–130 (1995).
 - 60 Nakajima T, Ono Y, Kato A, Maeda J, Ohe T. Y-29794 – a non-peptide prolyl endopeptidase inhibitor that can penetrate into the brain. *Neurosci. Lett.* 141(2), 156–160 (1992).
 - 61 Kato A, Fukunari A, Sakai Y, Nakajima T. Prevention of amyloid-like deposition by a selective prolyl endopeptidase inhibitor, Y-29794, in senescence-accelerated mouse. *J. Pharmacol. Exp. Ther.* 283(1), 328–335 (1997).
 - 62 Lambeir AM. Translational research on prolyl oligopeptidase inhibitors: the long road ahead. *Expert Opin. Ther. Pat.* 21(7), 977–981 (2011).
 - 63 Pratt J, Winchester C, Dawson N, Morris B. Advancing schizophrenia drug discovery: optimizing rodent models to bridge the translational gap. *Nat. Rev. Drug Discov.* 11(7), 560–579 (2012).
 - 64 Krystal JH, Karper LP, Seibyl JP *et al.* Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans. Psychotomimetic, perceptual, cognitive, and neuroendocrine responses. *Arch. Gen. Psychiatry* 51(3), 199–214 (1994).
 - 65 Flagstad P, Mork A, Glenthøj BY, Van Beek J, Michael-Titus AT, Didriksen M. Disruption of neurogenesis on gestational day 17 in the rat causes behavioral changes relevant to positive and negative schizophrenia symptoms and alters amphetamine-induced dopamine release in nucleus accumbens. *Neuropsychopharmacology* 29(11), 2052–2064 (2004).
 - 66 Giralt E, Peczuł MW, Salvatella X. *Protein Surface Recognition: Approaches for Drug Discovery*. John Wiley and Sons, Ltd, Chichester, UK (2011).
 - 67 Lawandi J, Gerber-Lemaire S, Juillerat-Jeanneret L, Moitessier N. Inhibitors of prolyl oligopeptidases for the therapy of human diseases: defining diseases and inhibitors. *J. Med. Chem.* 53(9), 3423–3438 (2010).
 - 68 Tran T, Quan C, Edosada CY *et al.* Synthesis and structure–activity relationship of *N*-acyl-Gly-, *N*-acyl-Sar- and *N*-blocked-boroPro inhibitors of FAP, DPP4, and POP. *Bioorg. Med. Chem. Lett.* 17(5), 1438–1442 (2007).
 - 69 Tarrago T, Kichik N, Segui J, Giralt E. The natural product berberine is a human prolyl oligopeptidase inhibitor. *ChemMedChem* 2(3), 354–359 (2007).
 - 70 Tarrago T, Kichik N, Claasen B, Prades R, Teixido M, Giralt E. Baicalin, a prodrug able to reach the CNS, is a prolyl oligopeptidase inhibitor. *Bioorg. Med. Chem.* 16(15), 7516–7524 (2008).
 - 71 Tarrago T, Masdeu C, Gomez E, Isambert N, Lavilla R, Giralt E. Benzimidazolium salts as small, nonpeptidic and BBB-permeable human prolyl oligopeptidase inhibitors. *ChemMedChem* 3(10), 1558–1565 (2008).
 - 72 Schechter I, Berger A. On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Comm.* 27(2), 157–162 (1967).

- 73 Abramowitz N, Schechter I, Berger A. On the size of the active site in proteases. II. Carboxypeptidase-A. *Biochem. Biophys. Res. Comm.* 29(6), 862–867 (1967).
- 74 De Nanteuil G, Portevin B, Lepagnol J. Prolyl endopeptidase inhibitors: a new class of memory enhancing drugs. *Drugs Fut.* 23(2), 167 (1998).
- 75 Lopez A, Tarrago T, Giralt E. Low molecular weight inhibitors of prolyl oligopeptidase: a review of compounds patented from 2003 to 2010. *Expert Opin. Ther. Pat.* 21(7), 1023–1044 (2011).
- 76 Joyeau R, Maoulida C, Guillet C *et al.* Synthesis and activity of pyrrolidinyl- and thiazolidinyl-dipeptide derivatives as inhibitors of the Tc80 prolyl oligopeptidase from *Trypanosoma cruzi*. *Eur. J. Med. Chem.* 35(2), 257–266 (2000).
- 77 Tsutsumi S, Okonogi T, Shibahara S *et al.* Synthesis and structure–activity relationships of peptidyl alpha-keto heterocycles as novel inhibitors of prolyl endopeptidase. *J. Med. Chem.* 37(21), 3492–3502 (1994).
- 78 Venalainen JI, Juvonen RO, Garcia-Horsman JA *et al.* Slow-binding inhibitors of prolyl oligopeptidase with different functional groups at the P1 site. *Biochem. J.* 382(Pt 3), 1003–1008 (2004).
- 79 Venalainen JI, Garcia-Horsman JA, Forsberg MM *et al.* Binding kinetics and duration of *in vivo* action of novel prolyl oligopeptidase inhibitors. *Biochem. Pharmacol.* 71(5), 683–692 (2006).
- 80 Tsuru D, Yoshimoto T, Koriyama N, Furukawa S. Thiazolidine derivatives as potent inhibitors specific for prolyl endopeptidase. *J. Biochem.* 104(4), 580–586 (1988).
- 81 Portevin B, Benoist A, Remond G *et al.* New prolyl endopeptidase inhibitors: *in vitro* and *in vivo* activities of azabicyclo[2.2.2]octane, azabicyclo[2.2.1]heptane, and perhydroindole derivatives. *J. Med. Chem.* 39(12), 2379–2391 (1996).
- 82 Jarho EM, Venalainen JI, Poutiainen S *et al.* 2(S)-(cycloalk-1-enecarbonyl)-1-(4-phenylbutanoyl)pyrrolidines and 2(S)-(aroyl)-1-(4-phenylbutanoyl)pyrrolidines as prolyl oligopeptidase inhibitors. *Bioorg. Med. Chem.* 15(5), 2024–2031 (2007).
- 83 Wallen EA, Christiaans JA, Saarinen TJ *et al.* Conformationally rigid *N*-acyl-5-alkyl-L-prolyl-pyrrolidines as prolyl oligopeptidase inhibitors. *Bioorg. Med. Chem.* 11(17), 3611–3619 (2003).
- 84 Van Der Veken P, Fulop V, Rea D *et al.* P2-substituted *N*-acylprolylpyrrolidine inhibitors of prolyl oligopeptidase: biochemical evaluation, binding mode determination, and assessment in a cellular model of synucleinopathy. *J. Med. Chem.* 55(22), 9856–9867 (2012).
- 85 Pidgeon C, Venkataram UV. Immobilized artificial membrane chromatography: supports composed of membrane lipids. *Anal. Biochem.* 176(1), 36–47 (1989).
- 86 Di L, Kerns EH, Fan K, McConnell OJ, Carter GT. High throughput artificial membrane permeability assay for blood–brain barrier. *Eur. J. Med. Chem.* 38(3), 223–232 (2003).
- 87 Garberg P, Ball M, Borg N *et al.* *In vitro* models for the blood–brain barrier. *Toxicol. In Vitro* 19(3), 299–334 (2005).
- 88 Lieberman JA, Stroup TS. The NIMH-CATIE schizophrenia study: what did we learn? *Am. J. Psychiatry* 168(8), 770–775 (2011).
- 89 Tost H, Braus DF, Hakimi S *et al.* Acute D2 receptor blockade induces rapid, reversible remodeling in human cortical-striatal circuits. *Nat. Neurosci.* 13(8), 920–922 (2010).

■ Patents

- 101 Hoechst AG: EP-0286927-A2 (1988).
 102 Merrell Pharma, Inc.: US161106 (1996).
 103 Orion Corp: WO2004060862 (2005).

■ Website

- 201 Thomson Reuters Integrity. Copyright 2013 – Prous Science.
<http://thomsonreuters.com/integrity/#modal>