POMPEU FABRA UNIVERSITY
Department of Experimental and Health Sciences

ASSESSING HUMAN GROWTH HORMONE VARIANTS TO DETERMINE THEIR POTENTIAL RELEVANCE IN ANTI-DOPING AND CLINICAL ANALYSIS.

PhD thesis

Gerard Such Sanmartín

Neuropsychopharmacology Programme
Municipal Institute of Medical Research
(IMIM-Hospital del Mar)

Barcelona, November 2009

POMPEU FABRA UNIVERSITY
Department of Experimental and Health Sciences

Doctoral Programme: Health and Life Sciences

ASSESSING HUMAN GROWTH HORMONE VARIANTS TO DETERMINE THEIR POTENTIAL RELEVANCE IN ANTI-DOPING AND CLINICAL ANALYSIS.

Memòria presentada per GERARD SUCH SANMARTÍN per a optar al títol de Doctor per la Universitat Pompeu Fabra. Aquesta tesis ha estat realitzada sota la codirecció del Dr. Jordi Segura Noguera i el Dr. Ricardo Gutiérrez Gallego, en el grup d'Investigació en Bioanàlisi i Serveis Analítics, programa de Neuropsicofarmacologia de l'Institut Municipal d'Investigació Mèdica (IMIM-Hospital del Mar). Programa de Doctorat en Ciències de la Salut i de la Vida de la Universitat Pompeu Fabra.

Dr. Jordi Segura Noguera Director de tesis	Dr. Ricardo Gutiérrez Gallego Director de tesis	Gerard Such Sanmartín Doctorand

Barcelona, Novembre 2009

Als meus pares,
al meu germà,
a l'Estela,
(i a les nines).

Imaginar l' "alegria" aliena i complaure's amb ella és el major privilegi dels animals superiors

(Friedrich Nietzsche)

Agraïments / Acknowledgements

Finalment doncs, ja la tenim aquí. La tesi. Sinònim d'anècdotes, d'experiències, d'esforç i sobretot, de constància. I com no podria ser d'una altra manera, acompanyat d'aquell grup de persones que des d'angles distints, han participat en el desenvolupament d'aquest treball. Des d'aquí, i sense proporció amb les línies dedicades, els hi voldria expressar sincerament el meu agraïment més profund.

Gràcies, en primer lloc, al Dr. Jordi Segura, per donar-me la oportunitat de treballar en aquesta institució, i permetre'm descobrir l'apassionant repte que suposa la investigació. I evidentment, gràcies per la seva direcció, certament l'experiència és un grau i en aquest cas, és aclaparadora. Gràcies per l'impuls en la tesi, per la confiança en el meu treball, per les correccions, pels suggeriments, per les crítiques, i sobretot per ajudar-me en el meu creixement com a científic i com a persona. La meva més sincera admiració i respecte.

Gràcies, especialment, al Dr. Ricardo Gutiérrez, per la seva direcció en aquesta tesi. Si he de resumir què he après d'ell, crec que no erro si escric en majúscules la paraula RIGOR. Rigor, que ve a ser com el barret de tot científic. Gràcies perquè n'he après molt, i perquè m'ha ajudat molt. I així de passada, doncs també MOLTES gràcies per les nombroses correccions literàries, i la corresponent paciència. El meu agraïment més sincer.

Voldria també agrair al Dr. David Andreu i al departament de proteòmica de la UPF, als que hi són i als que ja van marxar, la seva ajuda en els meus inicis científics, els seus consells, la seva companyia, i la possibilitat de poder treballar amb els seus equips.

I subratllant-ho, donar les gràcies als companys del departament de farmacologia. Començant per la fauna del despatx, gràcies pels cafès, per les xerrades, per les idees, pels partits, de fet no acabaria mai. GRÀCIES. Entre tots hem deixat moments memorables (que no reproduiré per mantenir un cert estatus). El que si que faré és remarcar la universalitat del despatx. Una argentina (ve a ser com una beguda energètica), un brasiler (*vox populi* que treballa per algun govern), un irlandès (i els seus gossos grocs), una canària (que serem amics si no m'atrapa al SM), un croat (o dos, fent un pilar), i els terratinents de Quart i Sant Pedor (aquests són els perillosos). Amb aquests dos faré una pausa. Gràcies. És evident que aquest trajecte hagués estat molt diferent sense vosaltres. I sortint trobem a un mallorquí. I no em direu que no és molta casualitat, qui podria preveure que el meu company de creixement (hormonal) hagués nascut el

mateix dia del mateix mes que jo? De fet, ja que ell és un parell d'anys més gran, seré jo qui hauré nascut el mateix dia que ell. Gràcies Jaume pel teu suport i per la teva companyia en aquesta tesi. Han estat molts moments i molt diversos. I també han estat molts consells. Gràcies. I de terratinent mallorquí, saltem a terratinent aragonesa. Gràcies per les xerrades, per la guia i pels consells. I *evidentemente* per la companyia ... però no us penseu que podreu guanyar-nos a tennis! Gràcies a l'alacantí (o herculià) per mostrar-me que els meus orígens són d'allà. Bé, per això i per més coses. Un plaer poder coincidir amb tu. Gràcies a Mitous, Civits, Txells i Erikas pels esmorzars de ben bé mitja tesi (de les millors maneres de començar el dia), a la Beth per compartir la seva passió pels pàrquings (i altris), i a la Meritxell per la seva confiança. Al final del passadís, si algú gosa apropar-s'hi, està en JM, flamant tri-campió. És com el chuck norris de la ciència, avisats esteu. I a tots aquells que no anomeno, però amb qui he compartit tot tipus de moments (i no exclusivament de caire científic).

I ja són 10 anys. Es diu ràpid. Quan miro enrere la veritat és que costa veure-hi més enllà. Deu ser la comoditat d'estar a casa. Estela. Podria dir-te gràcies per la paciència, però quan has vingut a les tantes a acompanyar-me mentre feia el western sis-cents, i hem aprofitat per anar a descobrir un restaurant mentre una incubació macerava tranquil·lament, crec que m'avergonyiria si te les donés. Però te les donaré igualment. Gràcies. Gràcies per moltes coses, però en el fons crec que es pot resumir en una única frase. Gràcies per fer que totes aquestes bambolines tinguin sentit.

Finalment, als meus pares i al meu germanet. Gràcies pel caliu. Bé, gràcies per ser família, per estar allà, per aconsellar, per escoltar, per constituir i apuntalar les meves bases. Us la dedico a vosaltres. A vosaltres i a les nines, la nostra particular família no-humana. Gràcies per ser els meus referents.

I would like to thak the world anti-doping agency (WADA) for financial support during all this period.

Gràcies a l'Institut Municipal d'Investigació Mèdica (IMIM) per l'aportació econòmica rebuda per l'enquadernació d'aquesta tesi.

Abbreviations

µl	micro liter
µM	micro molar
2D	bi-dimensional
3D	tri-dimensional
AA	amino acids
Arg	arginine
Asp	aspartic acid
BSA-PBS	bovine Serum Albumin in Phosphate Buffered Saline
CD	circular dichroism
CEIC	clinical Research Ethical Committee
cm	centimetres
CRP	c-reactive protein
Cys	cysteine
DIEA	N,N-diisopropylethylamine
DMF	N,N-Dimetilformamide
DPIV	dipeptidyl peptidase IV
E/S	enzyme / substrate
EDC	1-ethyl-3-(3-dimethylamino-propyl) carbo-diimide
EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme-linked immunosorbent assay
ESI	electrospray ionisation
FA	formic acid
Fmoc	9H-(f)luoren-9-yl(m)eth(o)xy(c)arbonyl
GH	(human) growth hormone
GH – N	gene “N” of human growth hormone
GH – V	gene “V” of human growth hormone
GHBP	growth hormone binding protein
GHD	growth hormone deficiency
GHEA	growth hormone exclusion assay
GHR	growth hormone receptor
Gln	glutamine
Glu	glutamic acid
HBS-EP	buffer saline (10 mM HEPES, 150 mM NaCl, 3.4 mM ethylene-diamine-tetraacetic acid (EDTA) and 0.005% tween20, pH 7.4)
HBTU	2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluoro-phosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	histidine
HOBt	N-hydrozybenzotriazole
HPLC	high-performance liquid chromatography
IGF-I	insulin-like growth factor I
IMAS	institut Municipal d'Assistència Sanitària
IOC	international Olympic Committee
IU	international units
JAK	janus Kinase

kDa	kilodaltons
kV	kilovolts
LP	limited proteolysis
Lys	lysine
m/z	mass / charge
mab / pab	monoclonal / polyclonal antibody
MALDI-TOF	matrix-Assisted Laser Desorption/Ionization Time of Flight
MBL	mannan-binding lectin
mg	milligrams
min	minute
mM	millimolar
mmol	millimol
mQ	milli-Q
MS	mass spectrometry
MW	molecular weight
NHS	N-hydroxy-succinimide
NIBSC	National Institute of Biological Standards and Controls
nm	nanometres
PBST	phosphate Buffered Saline Tween-20
PERL	practical Extraction and Report Language
Phe	phenylalanine
P-III-P	type III pro-collagen
p-IRMA	polyclonal immunoradiometric assay
PMF	peptide mass fingerprint
PRL	(human) prolactin
Pro	proline
PVDF	polyvinylidene fluoride
RIA	radioimmunoassay
Rmax	resonance maximal
Rpm	revolutions per minute
RT	room temperature
RU	resonance units
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
Ser	serine
SPR	surface plasmon resonance
SPS	solid phase synthesis
TAT	1,3,5-triacryloyl-hexahydro-s-triazine
TFA	trifluoroacetic acid
TFE	2,2,2-trifluoroethanol
Thr	threonine
trIFMA	time-resolved immunofluorometric assay
Trp	tryptophan
UV	ultra-visible
V	volts
v/v	volume / volume
WADA	world Anti-Doping Agency
WB	western Blot
WHO	world health organisation PBST

How this thesis is organised.

This manuscript is structured in eight main chapters, each containing several paragraphs. The first chapter includes the *Introduction*, covering the background information on human growth hormone. The second chapter comprises the *Objectives*, including a list of the main targets of the work. The third, fourth, fifth and sixth chapters contain the results obtained and the corresponding discussion. Each one of these four chapters includes in addition a brief introduction followed by materials and methods. The seventh chapter includes the overall *Conclusions* of the thesis, while references are included in the last chapter. Two appendices are included containing the publications derived from the work, and a detailed description of a computer script developed throughout the course of the thesis.

CONTENTS

1. HUMAN GROWTH HORMONE, INTRODUCTION 2

1.1. BIOLOGICAL EFFECTS 3
1.2. REGULATION AND SECRETION 4
1.3. GH VARIANTS AND FRAGMENTS 5
1.4. GH ABUSE 14
1.5. EXISTENT TECHNIQUES 18

2. OBJECTIVES 21

3. HUMAN GROWTH HORMONE VARIANTS 24

4. HUMAN GROWTH HORMONE DETECTION BY SURFACE PLASMON RESONANCE 51

5. HUMAN GROWTH HORMONE PURIFICATION 69

6. HUMAN GROWTH HORMONE BIOLOGICAL MARKERS 84

7. CONCLUSIONS 95

8. REFERENCES 98

9. APPENDIX 118

9.1. MANUSCRIPTS PUBLISHED 118

"Characterisation of the 5 kDa growth hormone isoform" 118

"Generation of 5 and 17 kDa human growth hormone fragments through limited proteolysis" 129

"Growth hormone in sport; beyond beijing 2008" (review) 139

"Surface plasmon resonance immuno assays - A perspective" (review) 150

9.2. SCRIPT: "PERL AUTOMATED FRAGMENTATION" 161

1. INTRODUCTION

1. Human growth hormone, introduction

Human growth hormone (GH) is responsible for a significant and diverse group of biological effects in humans (Segura *et al.*, 2009). Its main function consists in promoting body growth, but it also participates in protein, lipid and carbohydrate metabolism. Primarily produced by the anterior pituitary gland, it is secreted in a pulsatile pattern in response to peptides such as the hypothalamic GH releasing hormone (Goldenberg *et* Barkan, 2007) or the secretagogue ghrelin (Kojima *et* al., 1999). Actually, GH consists in several variants of different sequences, originated either genetically or after post-translational modifications, that in some cases possess specific biological properties. Although a comprehensive quantification is not available, a clear consensus exists towards the 22 kDa variant of 191 amino acids as the predominant GH variant, both in the pituitary gland and in blood circulation (Baumann, 2009).

The recombinant DNA-expression of this protein, reported in 1979 (Goeddel *et al.*, 1979), constitutes the present major source for GH. Utilised to treat GH deficiency pathologies, it substitutes the earlier employed cadaveric GH purified from pituitary glands, that is associated with Creutzfeldt-Jakob disease (Frasier, 1997). Illegal use for doping purposes has been largely suspected, resulting in its inclusion in the lists of banned substances of the International Olympic Committee (IOC) in 1992 (Sonksen, 2009) and by the World Anti-Doping Agency (WADA) in 1999 (Barroso *et al.*, 2008). The development of a method for the detection of exogenous GH is a complex task, mainly due to the lack of differences between the pharmaceutical and the native 22 kDa GH protein (McHugh *et al.*, 2005). Presently, the first method for the detection of GH abuse has been recently approved and already implemented (Barroso *et*

al., 2009), while other approaches are also being developed.

1.1. Biological effects

GH is related to distinct biological actions involving multiple organs and systems, directly associated with postnatal body growth and participating in protein, lipid and carbohydrate metabolism (Ayuk *et* Sheppard, 2006). GH actions on target cells are mediated via the GH receptor (GHR), present at high levels in liver and adipose tissue (Nam *et* Lobie, 2000), while interaction with the prolactin receptor has also been reported (Dannies, 2001). The dimerisation of the GH receptor by one GH molecule initiates the activation of the cytoplasmic tyrosine kinase called Janus Kinase (JAK) 2, leading to a variety of signalling pathways (Forsyth *et* Wallis, 2002). Those implying growth-promoting effects of GH are mostly mediated by the secreted insulin-like growth factor I (IGF-I) (Le Roith *et al.*, 2001; Mauras *et* Haymond, 2005). GH also greatly stimulates lipolysis in adipose tissue (Fain *et al.*, 2008), whereas in carbohydrate metabolism GH is known to induce insulin-resistance, in contrast to the insulin-like effect shown by IGF-I (Mauras *et* Haymond, 2005). Noteworthy, the metabolic effects of GH appear to be stimulated in detriment of the growth-promoting effects during periods of food deprivation. During conditions of energy shortage GH alters fuel consumption from the use of carbohydrates and protein to the use of lipids (Moller *et* Jorgensen, 2009). In this state, the expression of GHR in the liver is reduced (Beauloye *et al.*, 2002). In turn, the decreased concentration of the circulating IGF-I (Wu *et al.*, 2008) mediates the increase of GH concentration by the loss of feedback inhibition (Norrelund, 2005), and apparently not related to the previously presumed promoting role of ghrelin (Kirchner *et al.*, 2009; Smith, 2009).

1.2. Regulation and secretion

GH is mainly synthesised by somatotroph cells in the anterior lobe of the pituitary gland, released to the circulation, and acting through endocrine routes (Goldenberg *et* Barkan, 2007). Classically defined as a peptide hormone, GH is also known to adopt a role of growth or differentiation factor, locally synthesised in developing tissues before the pituitary gland becomes functional, acting by autocrine/paracrine routes (Sanders *et* Harvey, 2008). The regulation of pituitary GH secretion is mainly guided by the hypothalamic factors GH releasing hormone and the inhibitory hormone somatostatin (Gahete *et al.*, 2009), and also by other peripheral factors such as the secretagogue ghrelin produced in the stomach, small intestine and central nervous system (Anderson *et al.*, 2004). The release of the hormone into the blood is pulsatile, consisting in stable low basal levels abruptly interrupted by bursts of secretion (Veldhuis *et* Bowers, 2003). Once in circulation, approximately the 50 % of free GH forms a complex with the GH binding protein (GHBP), that is identical to the extracellular domain of the GHR and though to regulate the levels of free GH (Frystyk *et al.*, 2008). The half-life of the GHBP-bound GH is increased to 27 minutes versus 7 minutes of the free GH (Laursen, 2004). The complex between GH and GHBP is thought to act as a GH repository in blood, mainly during the inter-pulses interval (Baumann, 2002). Since the concentration of GHBP is stable [although depending on the proportion of adipose tissue of the subject (Roelen *et al.*, 1997)], the percentage of the GHBP complex highly varies with dependence of the episodic GH secretion (Frystyk *et al.*, 2008; Laursen, 2004). These GH bursts are stimulated by many factors. Fasting and sleep markedly amplifies the intensity of GH peaks as well as the frequency (Veldhuis *et*

Bowers, 2003). GH secretion also increases after exercise, with a direct relation to the concomitant increase in body temperature (Jorgensen *et al.*, 2003), and highly dependent on the exercise discipline being performed (Saugy *et al.*, 1995; Spolaore *et al.*, 2004). Negatively correlated with age and body composition (body fat percentage), GH secretion also accounts for gender differences, where women have greater daytime GH serum concentrations and greater 24-hour GH secretion (Stokes, 2003). Noteworthy, these differences were not reported in elite athletes in comparison with sedentary subjects (Ubertini *et al.*, 2008). Finally, the GH secretion is continuous in time only during pregnancy, when the placental GH progressively replaces the pituitary GH in maternal circulation within the last trimester (Fuglsang *et* Ovesen, 2006).

Disorders of the GH system include GH hypersecretion, GH deficiency, and GH insensitivity caused due to mutations in GHR. Acromegalia (GH hypersecretion in adults) is characterised by a marked volume expansion, reduced fat mass and higher lean body mass content (Katznelson, 2009). GH deficiency and GH insensitivity are related to increased total and visceral fat, decreased muscle mass and decreased lean body mass (Cummings *et* Merriam, 2003). The administration of GH secretion inhibitors or GH receptors antagonists in the case of GH hypersecretion, GH replacement therapy in the case of GH deficiency, or IGF-I replacement therapy in the case of GH insensitivity have shown to restore a normal state in some cases (Ayuk *et* Sheppard, 2006; Chanson *et* Salenave, 2008; Donangelo *et* Melmed, 2005; Nilsson *et al.*, 2007).

1.3. GH variants and fragments

GH is composed by a numerous group of variants originated both

genetically and through post-translational modifications. At the genetic level, two genes encode GH: the GH-N, expressed in the pituitary gland, and the GH-V, expressed in the placenta. The GH-N gene (five exons and four introns) expresses the main pituitary variant of the GH family, a 22 kDa molecular weight form of 191 amino acids which accounts for more of the 70 % of the pituitary GH (Baumann, 1999; Baumann, 2009). A partial splicing in exon 3 generates the 20 kDa variant, of 176 amino acids lacking the segment AA_{32-46} (Chapman *et al.*, 1981). The total skipping of the exon 3 yields the 17.5 kDa variant, of 151 amino acids lacking the segment AA_{32-71} (Lecomte *et al.*, 1987). The skipping of exons 3 and 4, and 2, 3 and 4 were also reported but their transcription were not observed (Palmetshofer *et al.*, 1995), whereas partial splicing in exons 4 and 5 encodes two 17.8 and 17.0 kDa variants, of 153 and 145 amino acids lacking the segments AA_{85-122} and AA_{91-136} respectively (Zhan *et al.*, 2005). The GH-V gene expresses a 22 kDa placental GH protein of 191 amino acids, 13 of which are different in comparison to the pituitary 22 kDa protein, with an additional glycosylation point and only appearing during pregnancy (Lacroix *et al.*, 2002). Analogously, an 20 kDa placental GH protein has also been described (Boguszewski *et al.*, 1998). Post-translational variations of pituitary-derived GH including acetylations, phosphorylations and deamidations, have been likewise reported (Baumann, 1999), with an additionally identified 23 kDa O-glycosylated variant of the 22 kDa protein (Bustamante *et al.*, 2009; Kohler *et al.*, 2009; Haro *et al.*, 1996a). Similarly, proteolytic fragmentation of the 22 kDa protein allegedly originates cleaved GH variants of 5 and 17 kDa molecular weight (Singh *et al.*, 1983; Sinha *et* Jacobsen, 1994). A list of these variants and their concentrations in blood circulation is provided in table 1, reproduced from Baumann (Baumann, 2009) (with permission).

Table 1. Percentages of the different GH variants in blood circulation. Reproduced with a copyright license obtained from Rightslink®, (Baumann, 2009).

Approximate mean distribution of pituitary GH isoforms in human blood 15–30 min after a secretory pulse.	
Monomeric GH	
22 kDa GH	
Free	22%
Bound to high affinity GHBP	21%
Bound to low affinity GHBP (a_2-macroglobulin)	2%
Total 22 kDa GH	45%
20 kDa GH	
Free	2%
Bound to high affinity GHBP	0.5%
Bound to low affinity GHBP (a_2-macroglobulin)	2%
Total 20 kDa GH	5%
Acidic GH (desamido-, acylated and glycosylated GH)	
Total acidic forms (bound fractions unknown)	5%
Dimeric GH	
22 kDa GH Dimers	
Non-covalent dimers	14%
Disulfide dimers	6%
Total 22 kDa GH dimers (bound fractions unknown)	20%
20 kDa GH Dimers	
Non-covalent dimers	3%
Disulfide dimers	2%
Total 20 kDa GH (bound fractions unknown)	5%
Acidic GH Dimers (desamido-, acylated and glycosylated GH)	
Non-covalent dimers	1.5%
Disulfide dimers	0.5%
Total acidic GH dimers (bound fractions unknown)	2%
Oligomeric GH (trimer-pentamer)	
22 kDa GH Oligomers	
Non-covalent oligomers	7%
Disulfide oligomers	3%
Total 22 kDa GH oligomers (bound fractions unknown)	10%
20 kDa GH Oligomers	
Non-covalent oligomers	1%

Disulfide oligomers	0.5%
Total 20 kDa GH oligomers (bound fractions unknown)	2%
Acidic GH Oligomers (desamido-, acylated and glycosylated GH)	
Non-covalent oligomers	1%
Disulfide oligomers	0.5%
Total acidic GH oligomers (bound fractions unknown)	2%
Fragments (12, 16 and 30 KDa immunoreactive species)	variable

Some of these GH variants might be stored in a pH-reversible aggregated form in the pituitary gland (Dannies, 2002), existing as monomers or homo/hetero-oligomers once in circulation. Equally, some of these variants are known to bind to the GHBP forming a complex of longer half-life time (Baumann, 2002). In serum samples, most concentrated GH variants are the 22 kDa protein and, allegedly the proteolytic 17 and 5 kDa fragments (Sinha *et* Jacobsen, 1994; Lopez-Guajardo *et al.*, 1998), with less amounts of the 20 kDa protein, and phosphorylated and deamidated 22 kDa and 20 kDa variants (Baumann, 2009; Kohler *et al.*, 2009). From the observation of the clearly differentiated biological activities of these proteins, and since the proteolytic 17 and 5 kDa variants are fragments of the 22 kDa GH protein, it was suggested a pro-hormone role for the last (Sinha *et* Jacobsen, 1994).

Growth hormone variants: 22 kDa, 20 kDa, 17 kDa and 5 kDa.

The 22 kDa GH protein has 191 amino acids structured in an anti-parallel four-helix bundle and stabilised by two disulfide bonds (Cys_{53}-Cys_{165} and Cys_{182}-Cys_{189}) (Chantalat *et al.*, 1995; Ultsch *et al.*, 1994). A 3D representation can be found in figure 1. Besides existing as a monomer it also forms both homo and hetero-oligomers, either covalent or non-covalent (Baumann, 1999), including a particularly stable disulfide-linked dimer of 45 kDa (Grigorian *et al.*, 2005). The 22 kDa aggregation is

reversible and guided by the pH of the medium (DeFelippis *et al.*, 1995), promoted by the presence of metals such as Zn^{2+}, Co^{2+} and others (Yang *et al.*, 2000; Dienys *et al.*, 2000). Accounting for monomers and oligomers, the 22 kDa concentration in the pituitary gland is stated to be within the 70-80 % of the total GH, a percentage that is maintained once GH is released into circulation (Baumann, 1991; Baumann, 2009). Both with GHBP and GHR, the 22 kDa protein initially forms a stable 1:1 complex through the so-called site 1 with a high affinity constant, followed by a dimerisation of the receptor forming a 2:1 complex, of a lower affinity constant. This 2:1 complex with the GH receptor, in particular the induced rotation of the receptor (Brown *et al.*, 2005) promotes the JAK2 activation and the posterior GH signal transduction. Approximately a 50 % of free GH circulates bound to GHBP, mostly in a 1:1 complex due to the concentrations ratio, while an additional 5-20 % is reported to be bound to an additional GHBP of a lower affinity constant (low-binding GHBP) (Baumann *et* Shaw, 1990). The regulation of GHBP mediates the biologic activity of GH. GHBP *per se* prolongs GH half-life by increasing its size, too large for glomerular filtration, and also nullifying its GHR-mediated clearance, responsible for part of the GH clearance, by preventing GH from binding to GHR (Baumann, 2002). GHBP in humans originates from a proteolytic shedding at one GHR unit (Zhang *et al.*, 2000). In consequence, the formation of GHBP generates non-functional receptors which in turn further increases the 22 kDa half-life. However, a high concentration of endogenous GHBP also reflects high levels of GHR (Fisker, 2006). Additionally, the 22 kDa protein is also known to interact with the human prolactin (PRL) receptor (Peterson *et* Brooks, 2004) as well as with the correspondent PRL binding protein (Dannies, 2001; Kline *et* Clevenger, 2001). The concentration of the 22 kDa protein under a basal state in healthy individuals lies in the low

picomolar or sub-picomolar range (pg/ml), while after an endogenous pulsatile burst this value eventually reaches the high picomolar range (Leung *et al.*, 2002) or up to (exercise-induced) the low nanomolar range (ng/ml) (Nindl *et al.*, 2001). The mean half-life time of the protein is of ~ 18 minutes (Laursen, 2004).

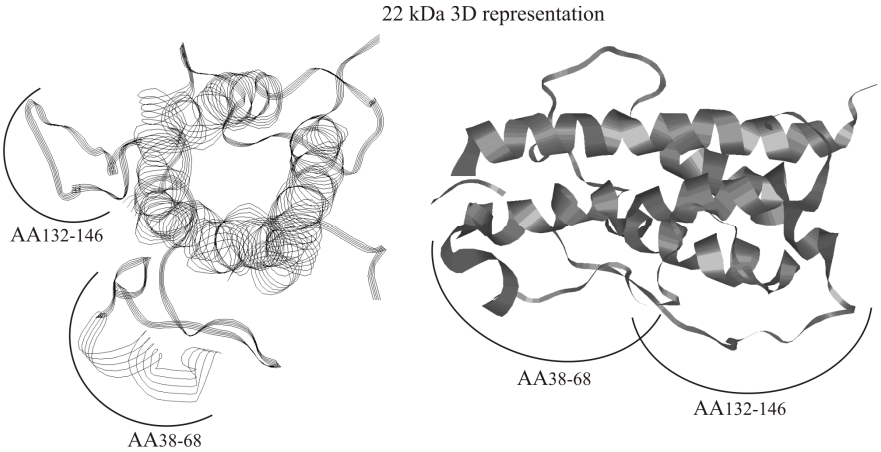


Figure 1. 3D graphical representation of the 22 kDa GH protein structure (1HGU structure from the RCSB protein data bank). The two graphs correspond to the same structure, but to different point of views (left, frontal; right, lateral). The regions $AA_{38\text{-}68}$ and $AA_{132\text{-}146}$, corresponding to two zones of *a priori* high accessibility are indicated.

The 20 kDa GH variant shares the 22 kDa sequence minus the segment $AA_{32\text{-}46}$, derived from an alternative splicing procedure (Chapman *et al.*, 1981). The structure of this protein maintains the same helicoidal conformation than the 22 kDa protein, based on circular dichroism (CD) data (Solomon *et al.*, 2006). The 20 kDa variant binds the GHR or GHBP with high affinity. However, in contrast with the 22 kDa protein that showed a higher affinity towards the bio-inactive 1:1 complex, the 20 kDa protein forms a weak 1:1 complex but a more stable 1:2 complex. Thus, this was suggested to be a more potent agonist than the 22 kDa protein, with a lower percentage bound to GHBP and a higher availability towards

the GHR (Tsunekawa *et al.*, 2000). Noteworthy, the 20 kDa variant was reported to have a longer half-life time (Baumann *et al.*, 1985). In this sense, although the concentration of the 20 kDa variant is maintained nearly constant with respect to the 22 kDa GH protein (~ 7-8 %) (Leung *et al.*, 2002), differences could exist after a GH peak due to these differences.

The 5 kDa and 17 kDa GH fragments allegedly constitute the N-terminal and C-terminal part of the 22 kDa protein, respectively, of sequences AA_{1-43} and AA_{44-191}. The 5 kDa protein was first purified from pituitary glands, identified through Edman sequencing as the first 43 amino acids of the 22 kDa sequence (Singh *et al.*, 1983). The 17 kDa fragment was identified to be produced by a: quote "*not totally specific*" cleavage (Lewis *et al.*, 2000), with the more prominent cleavage point occurring between residues 43 and 44, originating the sequence AA_{44-191}. How and where this cleavage could occur has not been reported. The percentage of both was reported as minor in the pituitary glands (Sinha *et* Jacobsen, 1994), a result that was confirmed for the 5 kDa protein (Shimizu *et al.*, 2004). However, analysis on serum samples showed a high concentration of the 17 kDa fragment, reaching concentrations of the 22 kDa GH protein (Sinha *et* Jacobsen, 1994; Warner *et al.*, 1993), whereas a similar significance was also suggested for the 5 kDa GH fragment (Lopez-Guajardo *et al.*, 1998). These differences suggested an origin of both fragments after the hormone's release from the pituitary gland. Posterior studies were carried out through the recombinant expression of the hypothesised 17 kDa GH sequence (AA_{44-191}), showing no interaction with the GH receptor (Rowlinson *et al.*, 1996). This might be explained by the absence of the first helix, a segment involved with the 22 kDa interaction towards the GH receptor (Behncken *et* Waters, 1999). On the other hand,

it was shown to interact with the human prolactin receptor although with slightly less affinity than the 22 kDa protein (Rowlinson *et al.*, 1996). The key motif identified for this interaction located at the Phe-44 amino acid, with a major mobility in the 17 kDa fragment (N-terminal end of the protein), could reason the shown decreased affinity. In concordance, the 20 kDa variant lacking this amino acid did not show any interaction as well (Peterson *et* Brooks, 1997; Peterson *et* Brooks, 2004). In consequence, part of the 17 kDa fragment biological actions were concluded to be expressed through prolactin receptors (lactogenic effects). Additionally, it is likely that other, unidentified receptors could be involved in sight of the high diabetogenic properties shown: a glucose-intolerance effect higher than the 22 kDa protein (Lewis *et al.*, 1991; Rowlinson *et al.*, 1996; Lewis *et al.*, 2000) although with some reports in conflict (Hettiarachchi *et al.*, 1997). The same conclusions were applied with the 5 kDa fragment, not showing any interaction towards none of the two (GH and prolactin) receptors (Rowlinson *et al.*, 1996; Lewis *et al.*, 2000) but expressing potent insulin-like effects (Salem, 1988; Salem *et* Wolff, 1989; Salem *et al.*, 1989; Frigeri *et al.*, 1988). Nonetheless, the biological actions of the 17 kDa and 5 kDa fragments are opposed. Interestingly, the diabetogenic effects of the 22 kDa protein are observed after the administration of high doses, while the insulin-like effects have been observed with low administered doses and during a short period (Yuen *et al.*, 2002). A graphical scheme of these interactions is shown in figure 2.

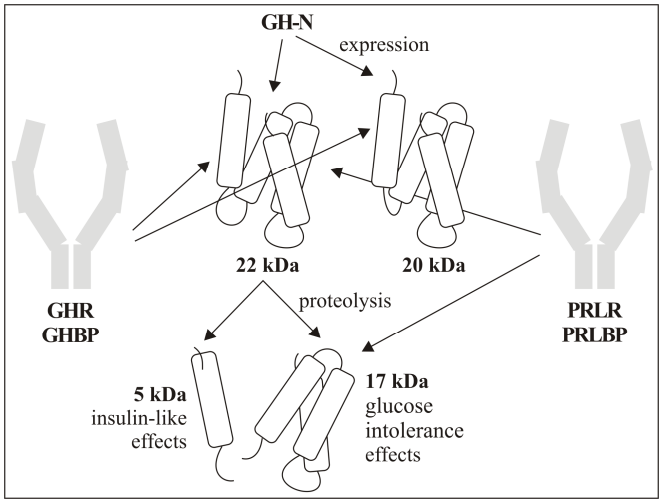


Figure 2. Graphical scheme of the 22 kDa, 20 kDa, 17 kDa and 5 kDa GH variants and their interaction towards the growth hormone (GH) and the prolactin (PRL) receptors (R) and the respective binding proteins (BP).

Recombinant GH

The recombinant 22 kDa GH protein, used to treat pathologies like GH deficiency, has an identical sequence to the native 22 kDa protein. Being a non-glycosylated protein in essence, it is indistinguishable from the endogenous hormone. Specific differences were encountered in some commercial solutions, consisting in specific amino acids exchanges (Met14 – Val14, Met14 – Ile14, Met125 – Ile125) and different degrees of methionine oxidation (Met14, Met 125, and Met 170) or deamidations (Asn12, Asn149, Asn159) (Hepner *et al.*, 2005; Hepner *et al.*, 2006; Jiang *et al.*, 2009). Also isotopic differences between native and pharmaceutical 22 kDa protein were investigated, based on the different isotopic ratio existent in the different nutrients sources used in the protein's *E. Coli* expression (Abramson *et al.*, 1996). In both cases these differences were found to be minimal and/or expected to be corrected once detected. However, following the administration of recombinant 22 kDa, the natural

production of GH is inhibited, this including all the different GH variants. As such, the natural ratio between the 22 kDa and the non-22 kDa GH variants results modified (Wallace *et al.*, 2001a). In this context, basal levels are restored 60 minutes after intravenous administrations, 12-20 hours after intramuscular injections and 16 hours after subcutaneous injections (Laursen, 2004).

1.4. GH abuse

The fraudulent use of recombinant GH in sport, both by recreational and professional athletes is suspected to exist since the early eighties (Holt *et al.*, 2009b). How and which doses are administrated is uncertain, but a high variability can be expected also considering its use in combination with other drugs (Graham *et al.*, 2009; Holt *et al.*, 2009b) for possible synergistic effects (Ehrnborg *et* Rosen, 2008). GH abuse is based on the potent anabolic properties of the hormone, with also a potential role in fracture healing (Tran *et al.*, 2009) and in their prevention (Doessing *et* Kjaer, 2005). Indeed, athletes testimonials reported assistance in their ability to recover from injuries and fatigue (Holt *et al.*, 2009b). However, the effects observed in GH deficient patients such as an increase of lean body mass and a reduction of fat mass are not clearly reported in healthy subjects (Graham *et al.*, 2009). In acromegalic patients which could be considered “endogenous” cases of doping, muscle weakness is a described symptom. Related, the link between the increased muscle mass obtained with recombinant GH and an incremental strength or power is not clear, and was suggested to be related to fluid retention or an excess of connective tissue (Doessing *et* Kjaer, 2005).

Detection of GH

The lack of differences in the sequence or structure between the exogenous and the endogenous 22 kDa protein, and the pulsatile nature of the hormone excludes the direct measurement of the GH concentration as an evidence of doping. As such, two alternative and complementary approaches have been or are being developed, so-named direct GH detection and indirect GH detection (Segura *et al.*, 2009).

The direct approach relies on the suppression of the endogenous production of GH following the administration of the recombinant protein (Wu *et al.*, 1999). After a GH dose, the 22 kDa GH protein is detected at high levels while the rest of pituitary proteins, grouped in the non-22 kDa variants group, are found at lower levels at least during approximately 24 hours. Following this approach, both the ratio 22 kDa versus non-22 kDa variants (Bidlingmaier *et al.*, 2000; Wallace *et al.*, 2001b; Wallace *et al.*, 2001a; Wu *et al.*, 1999), and versus 20 kDa variant (Hashimoto *et al.*, 1998) have been monitored. However, which components comprise this non-22 kDa group are under continuous debate, with controversy towards the 5 kDa and 17 kDa GH fragments. From the reports of Warner *et al.* and Sinha *et* Jacobsen (Warner *et al.*, 1993; Sinha *et* Jacobsen, 1994) for the 17 kDa fragment, and data provided by Lopez-Guajardo *et al.* (Lopez-Guajardo *et al.*, 1998) for the 5 kDa fragment, a high concentration was to be expected for both; in the case of the 17 kDa fragment detailed at levels even higher than the 22 kDa protein. In consequence, the non-22 kDa group not only would be mainly constituted by these fragments, but its concentration could be even greater than the 22 kDa protein. This assertion was not confirmed in the posterior 22 kDa vs non-22 kDa assays, where the concentration of the non-22 kDa group was always of a much lower magnitude. Actually, latest reviews catalogue the existence of

the 5 kDa and 17 kDa fragments as doubtful, *e.g*. referring to the Sinha *et* Jacobsen manuscript: quote “A report of GH44-191 immunoreactivity in serum is inconclusive as the physical nature of this immunoreactivity has not been positively identified” (Baumann, 2009). Hence, in agreement to this, the non-22 kDa group would be mainly constituted by the 20 kDa GH variant. However, regarding the methods employed to detect the non-22 kDa group, it was very unclear whether these actually recognised either the 17 kDa or the 5 kDa fragments. Sinha *et* Jacobsen evaluated the cross-reactivity of the 17 kDa fragment using a GH radio immunoassay, and reported a very low affinity. Furthermore, Jansson *et al.* (Jansson *et al.*, 1997) proposed a characterisation protocol for any GH immunoassay which included the 5 kDa and 17 kDa GH fragments. This protocol was applied to a series of immunoassays, and equally reported a low or inexistent recognition of the two fragments. Precisely, the method recently implemented as the reference anti-doping method by WADA (Barroso *et al.*, 2009) based on the ratio 22 kDa vs non-22 kDa variants, has only been tested against the 17 kDa fragment in uniquely one of the five employed antibodies (see supplement data of the reference) (Bidlingmaier *et al.*, 2009).

Under these circumstances, the study of the 5 kDa and 17 kDa GH fragments raised a considerable interest. On the one hand, evidences for their existence at the stated concentrations might demand to re-establish the map of the different variants of the GH family, with an unknown but potentially promising utility within the detection of GH abuse and the direct approach. On the other hand, evidences of the contrary could contribute to the clarification of the controversy about these fragments. Independently, both sequences might equally be of interest in the pharmaceutical field in sight of their potent biological properties.

The other approximation to detect GH abuse, the indirect approach, is based on the detection of abnormal concentration values of some proteins (indirect markers) after a prolonged administration of GH. This investigation included the potential candidates IGF-I and type III pro-collagen (P-III-P) (Erotokritou-Mulligan *et al.*, 2008), the insulin-like growth factor-II, insulin-like growth factor binding proteins -2 and -3, C-terminal telopeptide of type I collagen and acid-labile subunit (Dall *et al.*, 2000; Abellan *et al.*, 2005), and also leptin, osteocalcin and type I pro-collagen (Kniess *et al.*, 2003), adiponectin (Giavoli *et al.*, 2004), hemoglobin alpha-chain (Chung *et al.*, 2006), apolipo-proteins A-1 and B, C-reactive protein (CRP) and interleukin-6, intercellular adhesion molecule-1, von Willebrand factor and sCD40L (Bollerslev *et al.*, 2006), surfactant protein D, vitamin D binding protein, haptoglobin, transferrin and mannan-binding lectin (MBL) (Gravholt *et al.*, 2004). Two of these proteins, IGF-I and P-III-P, have been the selected after extensive study, being included in the GH-2000 and GH-2004 group proposed method (Bassett *et* Erotokritou-Mulligan, 2009; Erotokritou-Mulligan *et al.*, 2007; Erotokritou-Mulligan *et al.*, 2009b; Erotokritou-Mulligan *et al.*, 2009a; Powrie *et al.*, 2007), that however, has not yet been officially implemented (Sonksen, 2009; Holt *et al.*, 2009a). Of significance, variability was reported as a function of gender, age and even the sport discipline practised (Abellan *et al.*, 2006; Nelson *et* Ho, 2009), suggesting to consider the inclusion in the test of some of these factors.

As such, the interest for other GH-responsive proteins is maintained. Actually, beyond providing further evidences of GH doping, other markers could also minimise the risk of false accusations. In parallel, data provided by these studies also concerns to the knowledge about GH

treatments and the secondary effects. In comparison with the direct approach, the main advantage of this method is the longer window-of-opportunity where an administration of GH can be detected. Therefore, both approaches can be understood as complementary in regard of the different timeline application.

1.5. Existing techniques

The extremely low concentration of GH in biological samples, within the low or sub-picomolar range in basal states for the 22 kDa GH protein in blood, limits the number of techniques suitable for its direct measurement. The direct detection of GH has been performed using immunoassays such as Immulite (Barth *et* Sibley, 2008), SPR (Heutmekers *et al.*, 2007; Trevino *et al.*, 2009), ELISA [or analogues (Arslan *et al.*, 2008)], or by proteomic approaches (Wisniewski *et al.*, 2009; Wu *et al.*, 2002). The distinct detection of different GH variants have been addressed by ELISA, contrasting the 22 kDa protein either with the 20 kDa variant (Hashimoto *et al.*, 1998) or with the non-22 kDa group of variants (Bidlingmaier *et al.*, 2000; Wallace *et al.*, 2001b; Wallace *et al.*, 2001a; Wu *et al.*, 1999), or following an immunopurification process by 2D Western Blot (Kohler *et al.*, 2008; Kohler *et al.*, 2009; Sakharov *et al.*, 2008). For the detection of GH markers, proteins whose concentration is linked to a GH administration, several commercial kits exist for most of them either by Immulite, RIA or ELISA (Cowan *et* Bartlett, 2009; Abellan *et al.*, 2005), or have been also reported by HPLC-MS (Kay *et al.*, 2009; Bredehoft *et al.*, 2008), too.

Hence, most sensitive methodologies available to target proteins of markedly low concentration, including GH and its variants, are based on

the use of specific antibodies. In the case of GH, the similarity between the different GH variants and the particular specificity of the different antibodies, hinders standardisation between these immunoassays. Crossed interactions might be assumed, resulting in a disparity of results between them that has been largely reported (Strasburger, 1998; Melmed, 1999; Rakover *et al.*, 2000; Wood, 2001; Wieringa *et al.*, 2004; Strasburger, 2004; Popii *et* Baumann, 2004; Strasburger *et* Bidlingmaier, 2005; Bidlingmaier *et* Strasburger, 2007b; Bidlingmaier, 2008; Bidlingmaier *et* Freda, 2009; Bidlingmaier *et* Strasburger, 2007a). In particular and evidencing the role of the different GH variants, the utilisation of monoclonal instead of polyclonal antibodies amplified these differences, reasoned by a more specific detection of different GH variants (Jansson *et al.*, 1997; Bidlingmaier *et* Strasburger, 2007a). The use of monoclonal antibodies provides better sensitivities and a controlled setting of what proteins are recognised, but their use requires a previous characterisation of what proteins do they detect. In this sense, the advent of SPR methodology allows to characterise these interactions investing a minimal amount of time and reagents, providing full information of the binding phenomenon (Gutiérrez-Gallego *et al.*, 2009).

2. OBJECTIVES

2. Objectives

The research described in this thesis was conducted within the context of growth hormone (GH) abuse, focusing the attention on the different GH variants and the effects of a continued GH administration to humans. Among the several GH variants reported in human blood, 22 kDa protein is described as the predominant form. However, two 22 kDa-derived fragments of 5 kDa and 17 kDa molecular weight which could be even more concentrated, remain as doubtful and controversial without evidences to either support or contradict earlier reports. From the sequences homologous with the 22 kDa protein, the hypothetic incidence of the two fragments on different immunoassays for GH is assumed, although isolated well-characterised structures have not been available. Additionally, new approaches for the detection of GH abuse based on these fragments further substantiates the interest of their study.

On the other hand, the effects of a prolonged administration of GH induce changes in the concentration of some proteins, catalogued as GH markers. In this context, preliminary studies have been performed on two proteins in relation to a protocol of GH administration. These results are of interest both within the doping field, in the study of potential novel markers, and within the medical field, in the investigation of the secondary effects of treatments based on the recombinant GH.

As such, the general objectives of this thesis were divided in the following separate parts:

- Study the potential release of the 17 kDa and 5 kDa GH fragments by enzymatic proteolysis of the 22 kDa GH protein, evaluating whether the approach could be representative for an *in-vivo*

situation.

- Generation of the 17 kDa and 5 kDa GH fragments in sufficient amounts for their posterior study and structural characterisation. Exploration of solid phase synthesis for the generation of the 5 kDa fragment.

- Analysis by surface plasmon resonance of the interactions of a number of GH variants, including the 5 kDa and 17 kDa fragments, with monoclonal antibodies employed for the production of GH assays in anti-doping control.

- Assessment of the capabilities of the surface plasmon resonance methodology as a suitable platform for the immunodetection of GH.

- Development of a purification method for GH from blood samples aiming at recovering all GH variants, in particular the 17 kDa GH fragment if present, including a sensitive western blot procedure.

- Study of additional GH abuse markers in an assay of recombinant growth hormone administration to a group of volunteers.

This research project was performed with financial support of the Fundació IMIM, the Generalitat de Catalunya (Consell Català de l'Esport and AGAUR), and of the World Anti-Doping Agency (WADA) within the framework of the following projects:

- Chip technology for the detection of growth hormone abuse.

- Chip technology for the detection of growth hormone abuse – 1 year extension.

- Determination of the specificity of anti-20 kDa antibodies by SPR-search for complementary immunoglobulins.

3 to 6. RESULTS

3. Human growth hormone variants

Human growth hormone (GH) comprises a numerous group of proteins. Once released to blood circulation, there is a clear consensus for the 22 kDa GH protein, of sequence AA_{1-191}, as the most predominant variant representing a 75 % of the total GH (including monomers and oligomers), whereas the rest of variants represent around a 20 % (Baumann, 2009). Two additional GH variants were also detected in blood circulation at high concentrations. These were allegedly originated from a specific cleavage at the 22 kDa protein, releasing fragments of 17 kDa (sequence AA_{44-191}) and 5 kDa (sequence AA_{1-43}) molecular weight (MW). Of importance, for the former there were reported concentrations even higher than the 22 kDa protein (Lopez-Guajardo *et al.*, 1998; Sinha *et* Jacobsen, 1994; Warner *et al.*, 1993). However, the few reports published about these fragments have generated controversy, with uncertainty about their existence. In fact, neither fragment has been further identified in posterior immunoassays. On the contrary, it is also unclear whether the immunoassays used for the detection had sufficient affinity towards them. Based on the results of Jansson *et al.* evaluating different immunoassays (Jansson *et al.*, 1997), in all cases there was a low or null recognition.

Concerning their availability, the 22 kDa, 17 kDa and 5 kDa GH proteins have been successfully expressed in the past by recombinant technology, but only the 22 kDa is accessible commercially as a pharmaceutical. The 5 kDa GH fragment was also reported to be synthesised by solid phase synthesis methodology. The limited length of the 5 kDa sequence makes possible the application of this technique.

Here, the obtention of the 17 kDa and 5 kDa GH fragments has been

addressed by limited proteolysis (LP) of the 22 kDa GH protein, a process that could be comparable to the *in-vivo* proteolysis. Furthermore, the 5 kDa fragment has also been produced through solid phase synthesis (SPS) methodology. Both procedures have generated sufficient amount of material to obtain structural information. The three GH variants have been characterised by mass spectroscopic procedures. Peptide mass fingerprint (PMF) mapping by trypsinolysis under conditions of maximum cleavage have allowed to verify the sequence of the proteins. LP procedures have been used to obtain information the local secondary and tertiary structure of the proteins, based on the evaluation of the cleavages over the sequence under non-reducing conditions. In this case, the generated fragments were a result of a minimum number of cleavages, still linked by disulfide bonds. For the identification of these fragments, a specific script has been developed capable to calculate the corresponding masses.

Material and Methods

Materials. Recombinant 22 kDa GH protein (Genotonorm®) was purchased from Pfizer Laboratories (New York, NY, USA). Thermolysin, pepsin, V8-protease, proteinase K, ß-mercaptoethanol, iodoacetamide, 1,3,5-triacryloyl-hexahydro-s-triazine (TAT), 2,2,2-trifluoroethanol (TFE), sinapinic acid, and α-cyano-4-hydroxy-cinnamic acid were purchased from Sigma-Aldrich (Barcelona, Spain). Sequencing-grade trypsin was purchased from Promega (Madison, WI, USA). Geloader tips were purchased from Eppendorf (Barcelona, Spain). POROS R2 resin (20 µm) and MALDI Sequazyme Peptide Mass Standards Kit were purchased from Applied Biosystems (Barcelona, Spain). Fmoc-protected amino acids were purchased from Senn Chemicals (Dielsdorf, Switzerland). Side chain protecting groups were t-butyloxycarbonyl (Lys, Trp), t-butyl (Asp,

Glu, Ser, Thr, Tyr), 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Arg) and trityl (Asn, Gln, His). 4-(Hydroxymethyl) phenoxymethyl (Wang) resin loaded with the C-terminal Fmoc-Ser(tBu) was from Bachem (Bubendorf, Switzerland). 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluoro-phosphate (HBTU) and N-hydrozybenzotriazole (HOBt) were from Albatross Chem. (Montreal, Canada). CM5 sensor chips, and HBS-EP buffer saline (10 mM HEPES, 150 mM NaCl, 3.4 mM ethylenediaminetetraacetic acid (EDTA) and 0.005% Tween20, pH 7.4) were purchased from Biacore (Izasa, Barcelona, Spain).

Mass spectrometry. Mass spectra were recorded on a Voyager-DE STR Biospectrometry workstation (Applied Biosystems, Barcelona, Spain) using a pulsed nitrogen gas laser (337 nm) in positive linear mode, with accelerating voltage of 20 kV. Sample aliquots (1 µl volume) were mixed with either α-cyano-4-hydroxycinnamic or sinapinic acid matrices diluted in acetonitrile:trifluoroacetic acid (50:0.1 % v/v) saturated aqueous solution in 1:1 (v/v) proportion. Calibration was performed with the Sequazyme Peptide Mass Standards kit and data analysed with the Data Explorer software version 4.5 (Applied Biosystems). Electrospray - Ion trap mass spectrometric analyses were performed on an Esquire3000 (Bruker Daltonics, Barcelona, Spain) equipped with a standard ESI source. The instrument was operated in the positive ion mode, collecting data from *m/z* 50 to 3000. Acquired data were processed using the Data Analysis software v 3.0.

Protein synthesis. Solid phase synthesis was performed on an ABI 433A synthesizer (Applied Biosystems, Foster City, CA) running Fmoc/tBu chemistry programs. The synthesis was initiated at the C-terminus (AA_{43})

with serine-functionalised Wang resin (182 mg equivalent to 0.1 mmol of amino acid): N-terminal Fmoc groups were removed with piperidine [20% in N,N-Dimetilformamide (DMF)], the resin was washed with DMF and the next Fmoc-protected amino acid, activated with HBTU/HOBt in the presence of N,N-diisopropylethylamine (DIEA); amino acid : HBTU : HOBt : DIEA (ratio 1:1:1:2) was coupled in an excess of 10 equivalents within the 20-40 min range adjusted to the nature of each residue, followed by extensive washes with DMF. Coupling of predicted difficult linkage between Glu_{33} and Glu_{32} (Peptide Companion software, CoshiSoft, San Diego, USA) was performed manually and repeated until a negative Kaiser ninhydrin test (Kaiser *et al.*, 1970) was obtained. Manual synthesis included prolonged deprotection times (20% piperidine in DMF, 2 x 5 min) to ensure quantitative Fmoc removal, as well as longer couplings (30 min or more). The progress of the reaction was monitored at positions AA_{33-43}, AA_{32-43} and AA_{18-43} measuring aliquots directly mixed with the matrix and crystallised onto the plate, by MALDI-TOF mass spectrometry. Side chains deprotection and resin cleavage of the full-length protein sequence was done by treatment with trifluoroacetic acid (TFA) in the presence of scavengers (TFA : ethanedithiol : water : triisopropylsilane 94:2.5:2.5:1) for 90 min. The protein material was precipitated from the cleavage filtrate by addition of chilled methyl tert-butyl ether and centrifugation (4000 rpm, 5 ºC, 10 min). The crude protein pellet was solubilised in a minimal volume of acetic acid (10% v/v), freeze-dried and subsequently purified by preparative reversed phase high-performance liquid chromatography (HPLC).

Proteolysis of 22 kDa hGH. Limited proteolysis (LP) of 22 kDa GH (0.1 g/l) with the individual enzymes was performed at room temperature (RT) with an enzyme / substrate (E/S) ratio of 1/250, and at a pH value of

7. Thermolysin and pepsin, the latter exclusively, were also employed at a pH of 4. Sample aliquots were analysed during the first five minutes of the reaction, stopped by addition of trifluoroacetic acid (TFA; for MALDI-TOF) or formic acid (FA; for HPLC-ESI) at a final concentration of 0.2 % (v/v). In order to generate larger amounts of the 5 kDa and 17 kDa fragments, proteolysis was performed at a concentration of 1 g/l, and the reaction was prolonged for 10 min. For peptide mass fingerprinting (PMF) trypsin was employed at an E/S ratio of 1/50 (w/w) for 16 hours at 37 ºC, with prior standard reduction and alkylation (dithiothreitol: 10 mM, 30 min at 56 ºC followed by iodoacetamide: 50 mM, 30 min at RT in the dark), purified through Geloader tips containing Poros R2 resin (Gobom *et al.*, 1999).

Reversed phase HPLC. Separation of proteolytic fragments was performed by reversed phase HPLC-UV (HP1090 series II, Hewlett-Packard, Barcelona, Spain) at a flow rate of 0.8 ml/min, using a C8 column (15x0.46 cm, Teknokroma Tracer Excel, Teknokroma, Barcelona, Spain) and monitoring absorbance at 228 nm. Protein fragments were eluted employing a gradient elution from 30 % to 80 % of solvent A (ACN/H_2O/TFA 95/5/0.085 v/v, solvent B (H_2O/ACN/TFA 95/5/0.1 v/v) over 25 min. The 5 kDa and 17 kDa products were collected manually, quantified by UV absorbance measurements on a Thermo Scientific NanoDrop 1000 Spectrophotometer (Bonsai Technologies, Barcelona, Spain) employing the *protein A280* application of the ND-1000 software v3.3.0., and stored at 4 ºC. For HPLC-ESI (HP1100 series II, Agilent, Barcelona, Spain) assays, FA was used instead of TFA, and a shorter gradient of 16 min was employed. Data were acquired with the HP Chemstation Rev. A.06.03. (HPLC-UV) and EsquireControl v5.0 (HPLC-ESI) software. The synthesised 5 kDa protein, resuspended in H_2O mQ,

was purified by preparative reversed phase HPLC (Shimadzu LC-8A) on a Phenomenex Luna C8 column (10 µm, 2.1 x 25 cm) eluted with a 5-65% acetonitrile gradient (60 min, 25 ml/min, 220 nm UV detection). Fractions estimated to be of sufficient purity by analytical HPLC were pooled and freeze-dried.

Circular dichroism. CD spectra were acquired on a Jasco J-810 spectropolarimeter (Jasco Corporation, Tokyo, Japan) purged with nitrogen (25 ml/min) and using a quartz 0.1 cm path length cell fitted to a Neslab RP-100 thermostated device (Bonsai Technologies, Barcelona, Spain) set at 5 ºC, at protein concentration of 15 µM in 25 mM sodium phosphate buffer, pH 7.4. Alpha-helical secondary structure formation was induced by solvent titration with 10-50 % (v/v) TFE. For each spectrum, two 190-260-nm scans were accumulated, at 10 nm/min, with 0.2 nm resolution, 2s response, and 1.0 nm band width. Data were analyzed by the spectral deconvolution JFIT software (Bernhard Rupp, 1996).

Cross-linking experiments. Intra-molecular cross-linking experiments of the 5 kDa protein used 0.4 g/l protein concentrations in 10 mM calcium chloride solutions, pH 7.0 (minimum incubation of 60 min of the protein in solution at room temperature) with 1 µl of TAT (aqueous saturated solution). The final mixture was incubated at 37 ºC during 1 h and subsequently analysed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE): Mini-Protean 3 cell electrophoresis system (Bio-Rad, Barcelona, Spain), 150 V, 15 % acrylamide, silver staining procedure (Gorg *et al.*, 2000).

Theoretical mass identification of peaks. The obtention of theoretical

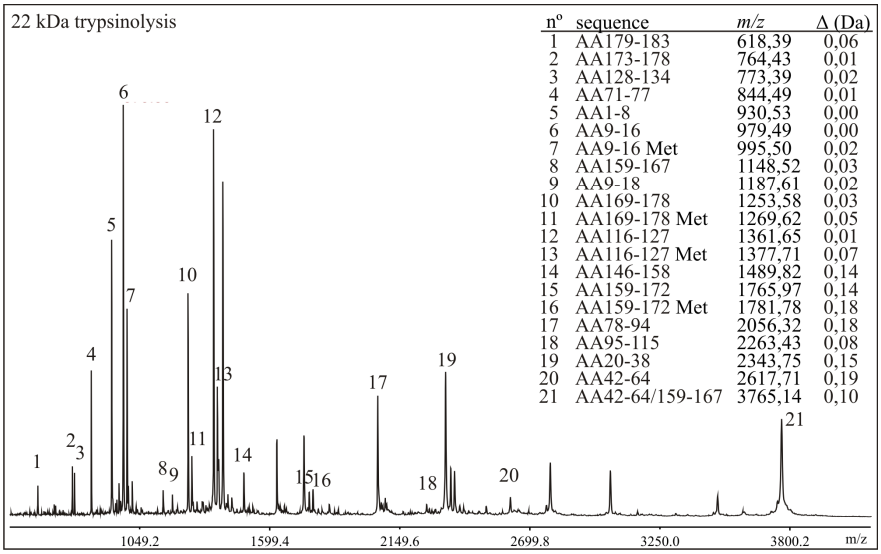
molecular weights of fragments resulting from each proteases activity was performed through the development of a specific script that calculates all possible cleavages for a given sequence and protease. A complete description and the code of the script is included in this thesis as appendix (see chapter 9.2.).

Results

Recombinant 22 kDa GH protein

The sequence of the 22 kDa GH protein was characterised by trypsinolysis (preferred cleavages at the Arg and Lys positions). Most of the observed peaks in the MALDI-TOF spectrum were positively identified. A table with the attributed sequence, the measured *m/z* value and the difference (Da) towards the theoretical mass is included in the figure 3, together with the spectrum. The final sequence coverage corresponded to a 86 %.

The structure of the 22 kDa protein was studied by proteolysis under LP conditions, *i.e.* stopping the reaction during the first stages where the minimum cleavages occurred, using the metallo endopeptidase thermolysin, the endopeptidases pepsin (Fontana *et al.*, 2004), trypsin (Graf *et al.*, 1992), V8-protease (Polverino *et al.*, 1995), and the serine endopeptidase proteinase K. Proteolyses products were analysed by HPLC-ESI technique, evaluating those fragments corresponding to the first cleavages of the enzymes. For each case, the chromatograms and the related mass spectra are included in figure 4. The identification of the observed peaks was done by mass coincidence with the different fragments.



nº	sequence	*m/z*	Δ (Da)
1	AA179-183	618,39	0,06
2	AA173-178	764,43	0,01
3	AA128-134	773,39	0,02
4	AA71-77	844,49	0,01
5	AA1-8	930,53	0,00
6	AA9-16	979,49	0,00
7	AA9-16 Met	995,50	0,02
8	AA159-167	1148,52	0,03
9	AA9-18	1187,61	0,02
10	AA169-178	1253,58	0,03
11	AA169-178 Met	1269,62	0,05
12	AA116-127	1361,65	0,01
13	AA116-127 Met	1377,71	0,07
14	AA146-158	1489,82	0,14
15	AA159-172	1765,97	0,14
16	AA159-172 Met	1781,78	0,18
17	AA78-94	2056,32	0,18
18	AA95-115	2263,43	0,08
19	AA20-38	2343,75	0,15
20	AA42-64	2617,71	0,19
21	AA42-64/159-167	3765,14	0,10

Figure 3. MALDI-TOF mass spectrum corresponding to the trypsinolysis of the recombinant 22 kDa GH protein. Sequences attributed to the obtained peaks, experimental mass values and the differences towards the theoretical values (in Da) are included. Non-marked peaks correspond to unidentified sequences.

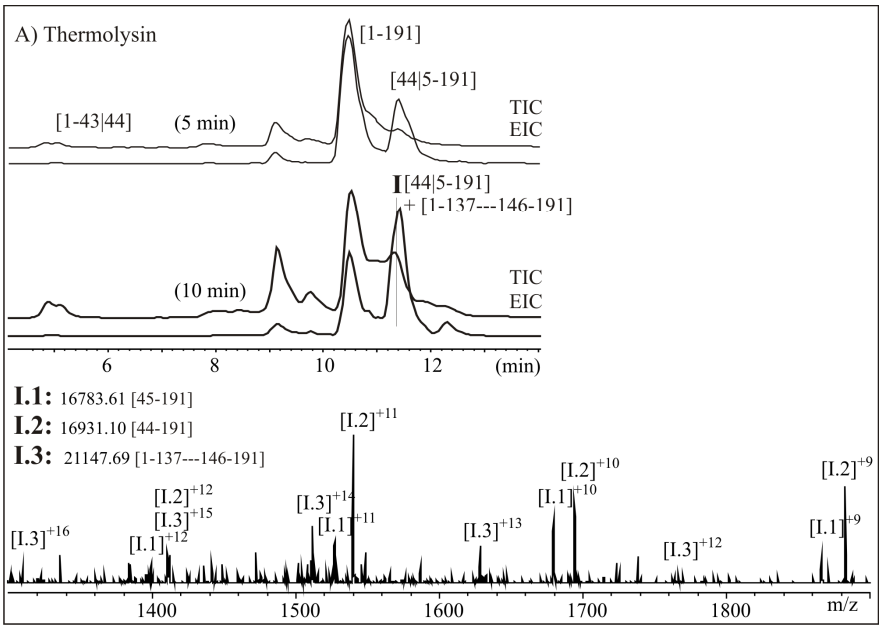
Thermolysinolysis of the 22 kDa protein (figure 4, A) showed a first cleavage at the AA_{43-44} and AA_{44-45} bonds, generating the fragments of sequences AA_{1-43} and AA_{1-44}, and AA_{44-191} and AA_{45-191} (*m/z* 16783.61 and 16931.10; theoretical *m/z* 16780.99 and 16928.17). At longer times of reaction, a third cleavage was observed at positions $AA_{137-138}$ and $AA_{145-146}$, generating the fragments AA_{1-137} and $AA_{146-191}$ that remained linked by the disulfide bond of Cys_{53-165} (*m/z* 21147.69; theoretical *m/z* 21148.92).

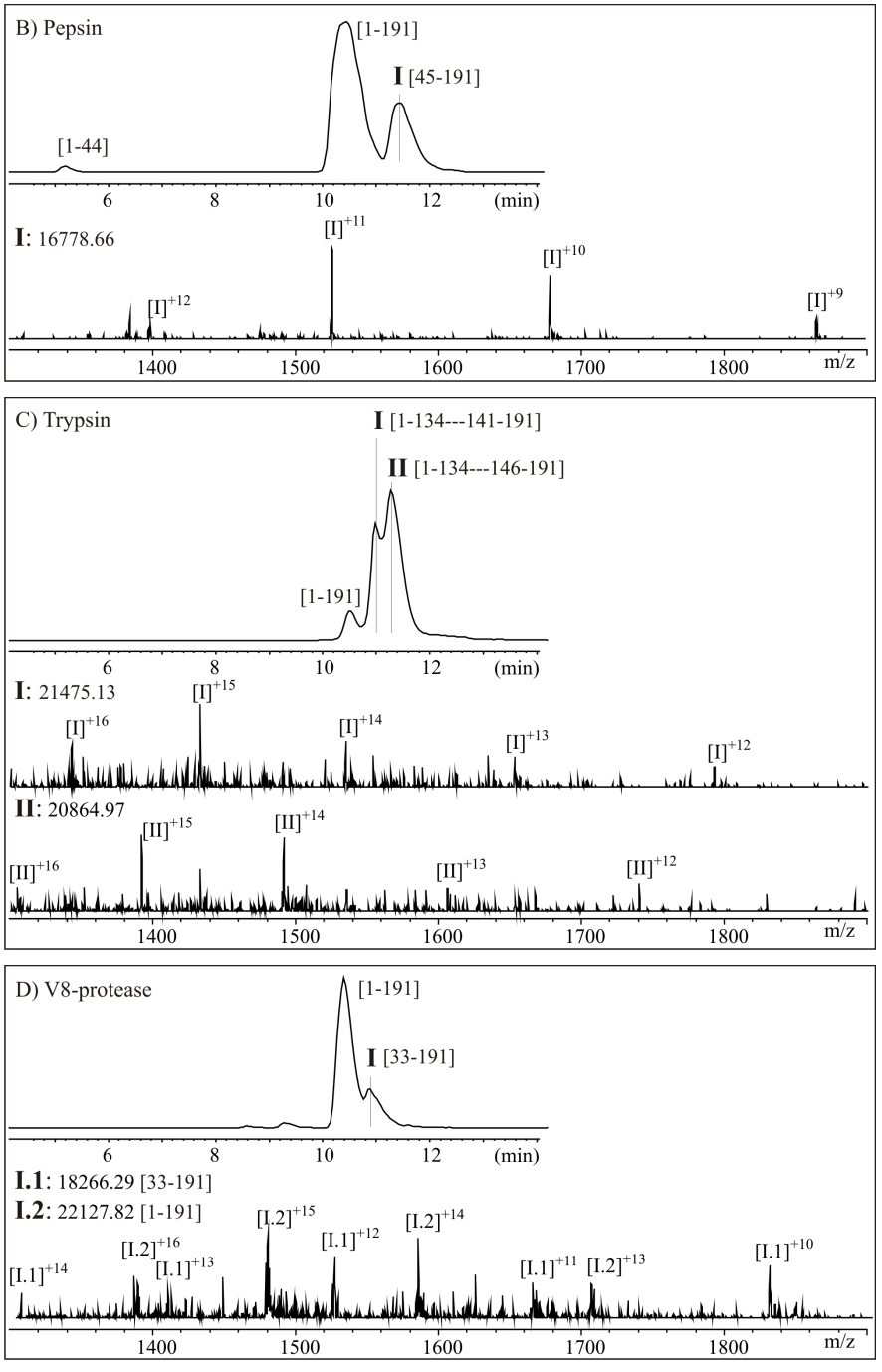
Pepsinolysis (figure 4, B), performed at an acidic pH of 4 where the enzyme is active, showed a simple cleavage at the position AA_{44-45}, generating the fragments AA_{1-44}, and AA_{45-191} (*m/z* 16778.66; theoretical *m/z* 16780.99).

Trypsinolysis (figure 4, C) showed a first cleavage at positions $AA_{134-135}$, $AA_{140-141}$ and $AA_{145-146}$, generating the fragments AA_{1-134} linked with both $AA_{141-191}$ and $AA_{146-191}$ (*m/z* 21475.13 and 20864.97; theoretical *m/z* 21470.30 and 20862.64, respectively).

Proteolysis with V8-protease (figure 4, D; a Glu-specific protease with only 14 possible cleavage sites within the 22 kDa sequence) was shown to target the glutamic acid Glu_{32-33}, generating the fragment AA_{33-191} (*m/z* 18266.29; theoretical *m/z* 18265.44).

Proteinase K proteolysis (figure 4, E) showed a cleavage at positions $AA_{119-120}$ and $AA_{130-131}$, generating the fragment AA_{1-119} linked to the fragment $AA_{131-191}$ (*m/z* 20933.22; theoretical *m/z* 20930.69).




B) Pepsin
[1-191]
I [45-191]
[1-44]
6
8
10
12
(min)
I: 16778.66
[I]+11
[I]+10
[I]+12
[I]+9
1400
1500
1600
1700
1800
m/z
C) Trypsin
I [1-134---141-191]
II [1-134---146-191]
[1-191]
6
8
10
12
(min)
I: 21475.13
[I]+15
[I]+16
[I]+14
[I]+13
[I]+12
II: 20864.97
[II]+15
[II]+14
[II]+16
[II]+13
[II]+12
1400
1500
1600
1700
1800
m/z
D) V8-protease
[1-191]
I [33-191]
6
8
10
12
(min)
I.1: 18266.29 [33-191]
I.2: 22127.82 [1-191]
[I.2]+15
[I.2]+16
[I.1]+12
[I.2]+14
[I.1]+14
[I.1]+13
[I.1]+11
[I.2]+13
[I.1]+10
1400
1500
1600
1700
1800
m/z

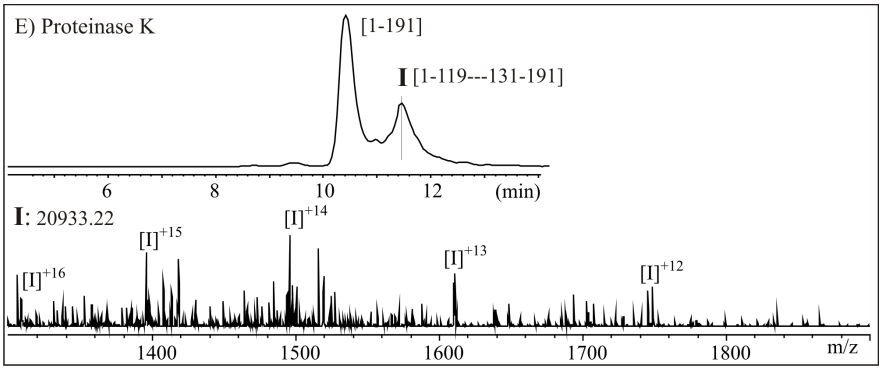


Figure 4. HPLC-ESI chromatograms and mass spectra corresponding to the limited (A) thermolysinolysis, (B) pepsinolysis, (C) trypsinolysis, (D) V8-proteolysis and (E) proteinase K proteolysis of the recombinant 22 kDa GH protein. Sequences attributed to the obtained peaks are indicated on the chromatogram peaks. Each spectrum corresponds to the indicated peak in roman numeral.

Thermolysinolysis and pepsinolysis yielded similar cleavages on the 22 kDa protein, although different pHs were required due to the enzymes' characteristics. The pepsin specificity was hypothesised to be related to a pH-induced modification in the structure of the 22 kDa GH (Spolaore *et al.*, 2004). In order to study whether different conditions could actually modify the structure of the 22 kDa protein, yielding different cleavage patterns with thermolysin, equivalent thermolysinolysis were conducted at pH 4, with TFE (50 %; v/v), and with urea (6 M). TFE is known to strengthen the folding of the protein preferably into an helicoidal form, and urea possesses denaturing properties. The progress of the proteolyses at pH 7 and pH 4 by MALDI-TOF spectrometry is shown in figure 5, including the mass spectra corresponding to different times of reaction. Mass spectra corresponding to proteolyses with TFE and urea are shown in figure 6 (top and bottom, respectively). In all cases the fragments corresponding to the cleavages at positions AA_{43-44} and AA_{44-45} were observed, and eventually also those resulting from a additional cleavages occurring at the positions AA_{41-42} and AA_{42-43}. The generated 17 kDa

fragment was observed to be rapidly proteolysed. The originated pairs maintained a difference coincident to a phenylalanine (Phe 44). With the presence of TFE, the 17 kDa fragment appeared to show a significant major resistance and an accumulation over time. It could be hypothesised to be caused by a higher folding of the fragment induced by TFE. Furthermore, either with TFE, urea or at an acidic pH the global velocity of the proteolysis decreased, reasoned by the lesser thermolysin stability under such conditions. Comparing thermolysinolysis at neutral and at acidic pH, the predominant cleavage at the position $AA_{43\text{-}44}$ observed for the former, changed to the position $AA_{44\text{-}45}$ for the latter, a particularity also shown with TFE. In essence, none of the three conditions were shown to substantially modify the pattern of cleavage observed at neutral pH.

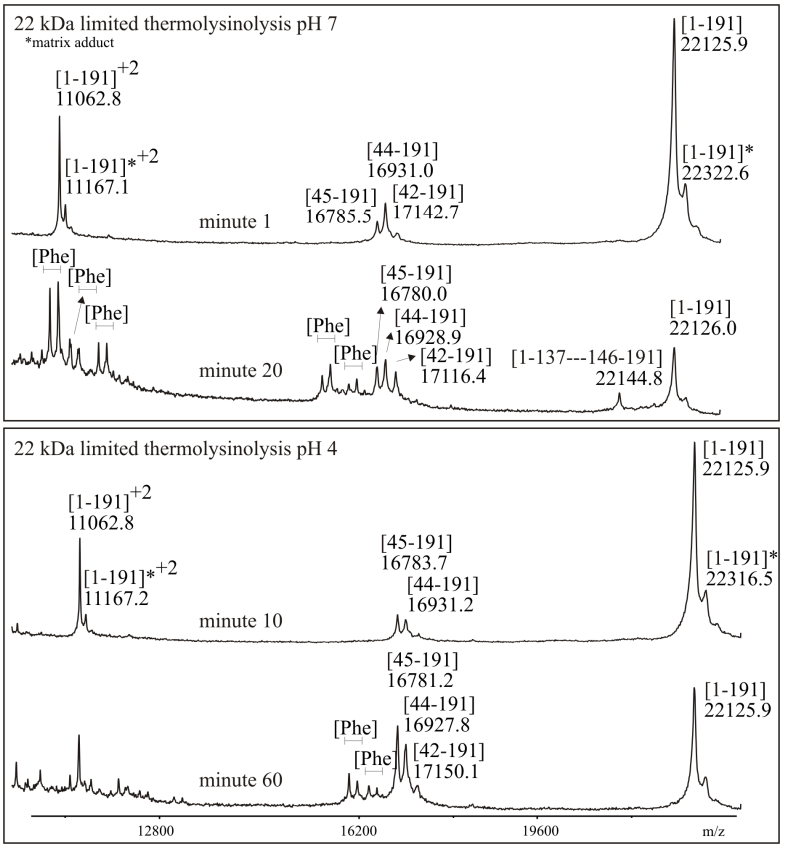


Figure 5. MALDI-TOF mass spectra corresponding to the thermolysinolysis of the recombinant 22 kDa GH protein under a pH value of 7 (top) and 4 (bottom). Sequences attributed to the obtained peaks together with experimental mass values are indicated, where the label [Phe] indicates a difference coincident to the mass of a phenylalanine between the indicated peaks.

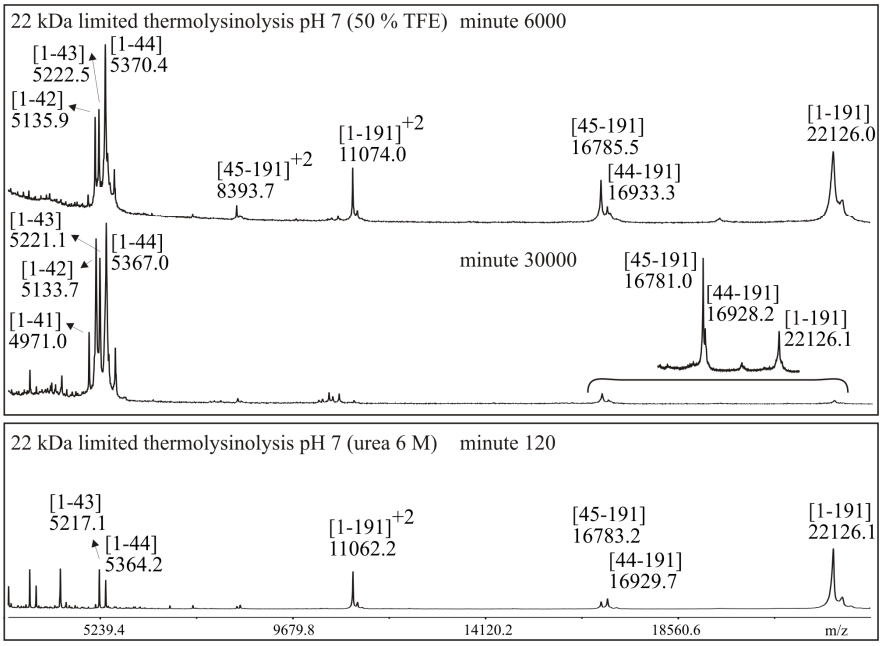


Figure 6. MALDI-TOF mass spectra corresponding to the thermolysinolysis of the recombinant 22 kDa GH protein with the addition of TFE (50 % v/v, top) and urea (6 M, bottom). Sequences attributed to the obtained peaks together with experimental mass values are indicated.

Proteolytic 17 kDa and 5 kDa GH fragments

Thermolysinolysis of the 22 kDa GH protein under LP conditions was conducted at a larger scale in order to purify the 5 kDa and 17 kDa fragments by HPLC. Under optimised conditions approximately 25 % of the initial 22 kDa protein was digested in a single round, resulting in equimolar amounts of the 5 kDa and 17 kDa fragments. Both 5 kDa peaks (subsequently identified as AA_{1-43} and AA_{1-44}) could be separated with

baseline resolution, showing an area ratio of approximately 4:1 for AA_{1-43} and AA_{1-44}, respectively. The complementary 17 kDa sequences (AA_{44-191} and AA_{45-191}) co-eluted and were collected as a mixture. The chromatogram is shown in figure 7. MALDI-TOF mass spectra of each collected peak are included in the figure, indicated through the retention time of the peak.

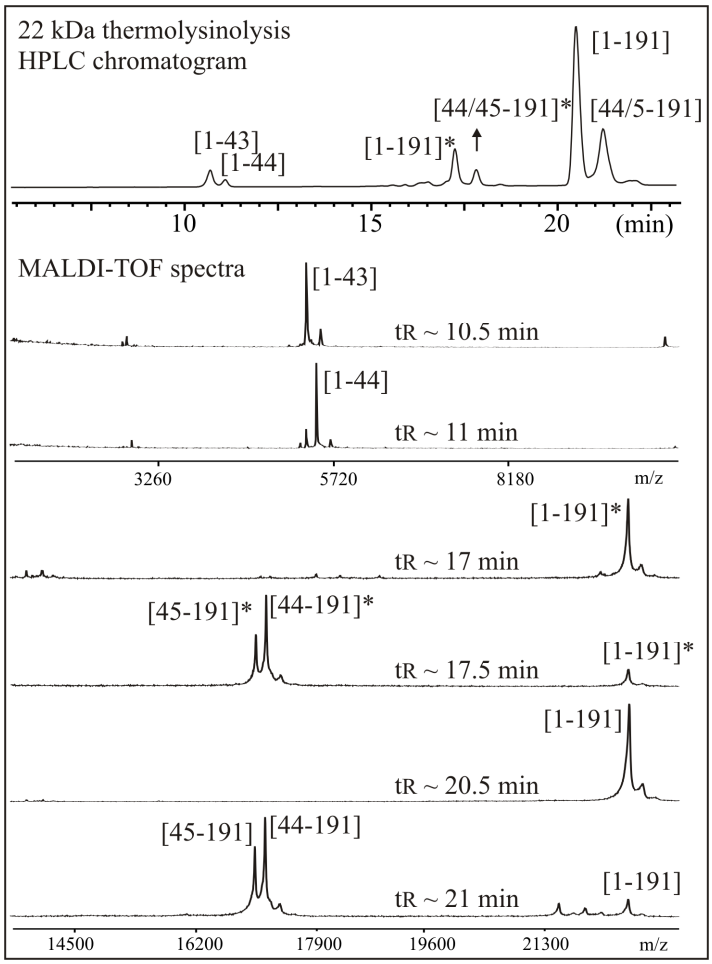


Figure 7. HPLC-UV chromatogram and MALDI-TOF mass spectra of each collected peak of the 22 kDa protein thermolysinolysis stopped at minute 10.

Due to the partial overlapping between the chromatogram peaks, the 5 kDa fragment AA_{1-43} showed in the mass spectrum a minor presence of the sequence AA_{1-44} and vice versa, while the 17 kDa fragments also showed a small percentage of the main 22 kDa sequence AA_{1-191}. Two additional peaks were observed in the chromatogram that eluted just before the 22 kDa GH protein (figure 7, top, peaks marked with an asterisk). The respective mass values and the obtained peptide mass fingerprints coincided with the 22 kDa and 17 kDa sequences. Most probably, a distinct tertiary structures or the formation of aggregates could give rise to the differential chromatographic behaviour.

The collected 17 kDa fragment (sequences AA_{44-191} and AA_{45-191}) was submitted to trypsinolysis in order to verify its sequence identity. MALDI-TOF spectrum is shown in figure 8. Approximately 70 % of the sequence was confirmed. Due to the partial overlapping of the 17 kDa chromatographic peak with the initial 22 kDa peak, a minor contamination of the latter was found in the mixture (shown in figure 7). The peaks AA_{44-191} and AA_{45-191}, in comparison with the 22 kDa derived AA_{42-191} peak, confirmed the respective sequences AA_{44-191} and AA_{45-191}.

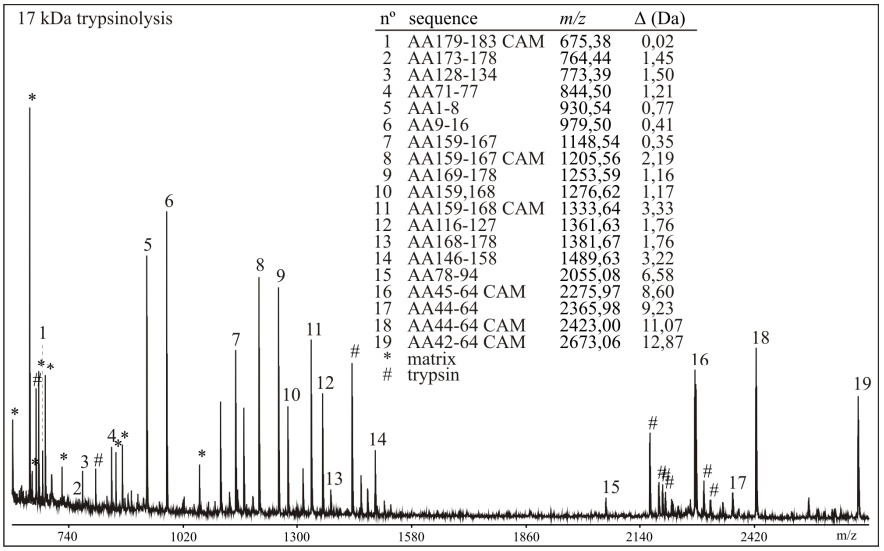


Figure 8. MALDI mass spectra of the trypsinolysis of the 17 kDa collected fragment. Sequences attributed to the obtained peaks, experimental mass values and the differences towards the theoretical values (in Da) are included. Non-marked peaks correspond to unidentified sequences.

The local structure of the purified 17 kDa fragment was studied by thermolysinolysis under LP conditions. The pattern of thermolysin action, shown in the HPLC-ESI chromatogram and mass spectra in figure 9, was similar to the original 22 kDa protein (figure 4, A), resulting in the generation of the sequences $AA_{44/45\text{-}137}$ and $AA_{146\text{-}191}$, linked by a disulfide bond between Cys_{53} and Cys_{165}: *m/z* 15806.63 and 15952.85, theoretical *m/z* 15803.84 and 15951.01, respectively. Based on this data, apparently the tertiary structure of the proteolytic 17 kDa fragment seems to remain essentially unaltered after release from the 22 kDa precursor hormone.

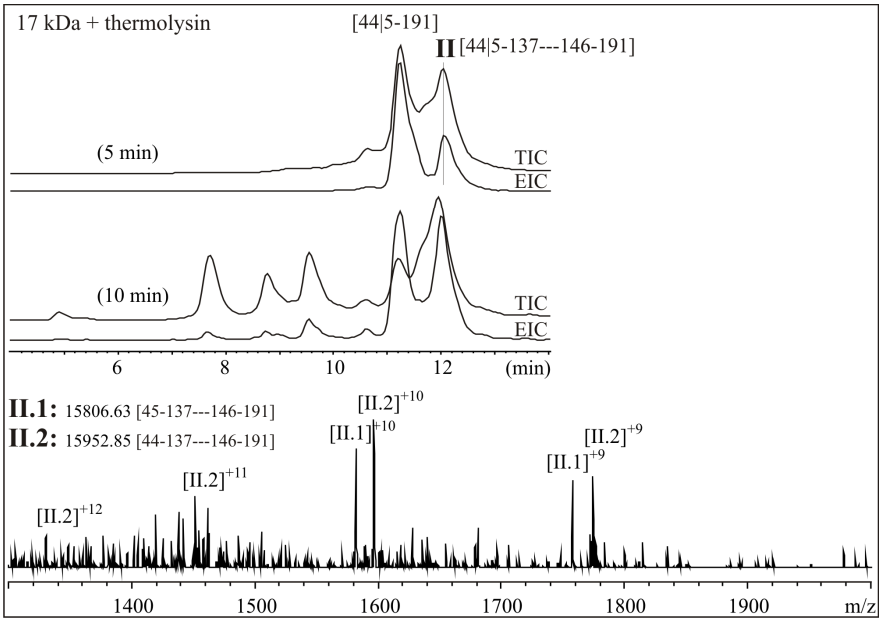


Figure 9. HPLC-ESI chromatograms and mass spectra corresponding to the thermolysinolysis of the purified 17 kDa GH fragment. First chromatogram corresponds to the proteolysis stopped at minute 5, second chromatogram stopped at minute 10. Sequences attributed to the obtained peaks are indicated on the chromatogram peaks. Each spectrum corresponds to the indicated peak in roman numerals.

Synthesis of 5 kDa GH fragment

The solid phase synthesis of the 43 amino acid protein corresponding to the 5 kDa GH fragment was successfully accomplished through Fmoc/t-butyl chemistry (Fields *et* Noble, 1990). The main peak from the reversed-phase HPLC purification of the synthetic material gave a MALDI-TOF mass spectra (figure 10) with a main singly charged ion at *m/z* 5216.93, consistent with the expected average molecular mass of 5215.92 Da, a peak 206 Da higher than the previous one, corresponding to a sinapinic acid photo-adduct (Beavis *et* Chait, 1990), and a third peak at *m/z* 5103.33 Da, *i.e.*, 113.60 Da below the molecular ion, attributed to deletion of either Ile or Leu (113.16 Da of theoretical difference, average mass). This

low-abundance deletion product could not be purified-out by further HPLC steps. A total of 110 mg of the protein were obtained, representing an overall yield of 25 % with respect to the initial amount of resin.

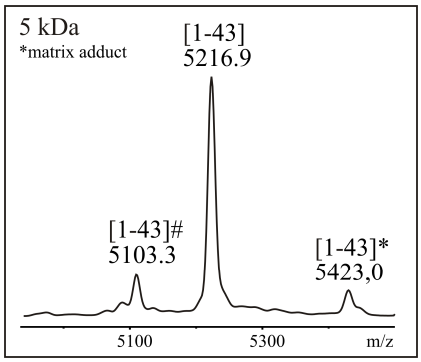


Figure 10. MALDI-TOF mass spectrum corresponding to the purified 5 kDa synthesised protein. Sequences attributed to the obtained peaks together with experimental mass values are indicated.

Verification of the protein sequence was achieved by trypsinolysis. The generated spectrum is shown in figure 11, including the list of attributed sequences. The peak at *m/z* 1734.88, attributed to the sequence $AA_{1\text{-}15}$, corresponded to an uncommon trypsin cleavage (N-terminal side of an Arg). These peaks covered the total sequence of the 5 kDa verifying the correct amino acid sequence of the product. Furthermore, an additional peak was observed at *m/z* 817.45. This peak showed a mass difference of 113.08 Da with respect to the $AA_{1\text{-}8}$ sequence and was consistent with the Ile/Leu deletion observed in the main HPLC fraction, detected here with isotopic resolution. While the mass difference clearly suggested the absence of either Ile_4 or Leu_6, the precise position of the deletion could not be established by electrospray ionisation tandem mass spectrometry experiments. Poor fragmentation was obtained, probably due to the adjacent Pro residue. On the basis of MALDI-TOF mass spectra peak intensities, this deletion-analogue was estimated to be less than 5 % of the

total amount of protein.

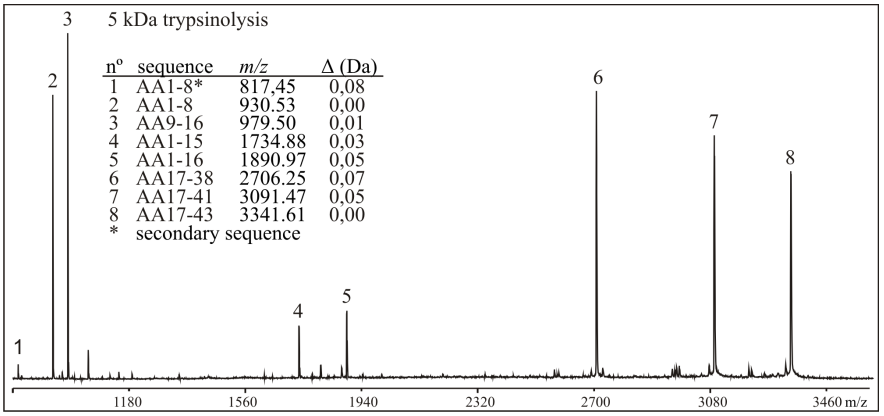


Figure 11. MALDI-TOF mass spectrum corresponding to the purified 5 kDa protein. Sequences attributed to the obtained peaks, experimental mass values and the differences with the theoretical values (in Da) are indicated.

Information about the secondary structure of the 5 kDa synthetic protein was assessed by far-UV circular dichroism. The protein was measured in buffer and after addition of TFE, where a reinforcement of the secondary structure and a disruption of the tertiary structure is expected (Buck, 1998). Figure 12 shows overlaid spectra corresponding to 0 %, 20 % and 50 % (v/v) of TFE (right hand-side graph corresponds to a normalised plot). Spectra were de-convoluted by means of the Yang equation (Yang *et al.*, 1986) to give 40 %, 78 %, and 98 % of α-helical content for 0 %, 20 %, and 50 % TFE concentrations, respectively.

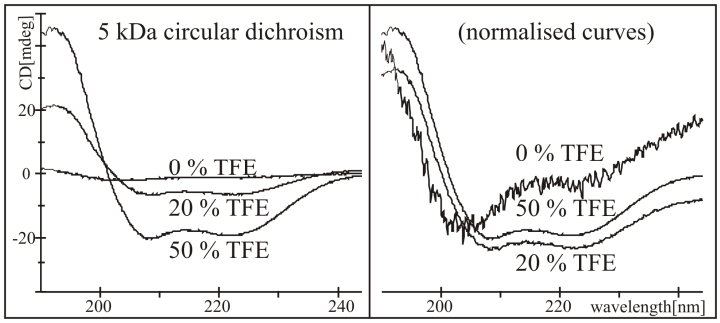


Figure 12. Left panel: spectrum of circular dichroism analysis of the 5 kDa protein in solution and the percentages of tri-fluoroethanol of 0 %, 20 % and 50 %. Right panel: normalised view of the spectrum.

Limited proteolysis of the synthetic 5 kDa fragment provided detailed information about local structure in relation to the amino acids sequence. The particular cleavage patterns obtained are directly related to how exposed and flexible the structure of the protein is (Hubbard, 1998). Digestions were carried out with thermolysin at pH 7 and 4, yielding a cleavage pattern shown in figure 13 (top and bottom, respectively). Fragment assignment of peaks observed at neutral and acidic pH are included in the figure. At acidic pH thermolysin was notably affected and longer times of reactions were required to obtain the first peaks of the digestion. Modifications of thermolysin specificity as function of media pH were discarded with the observation of equal patterns for acidic and neutral digestions with the 22 kDa GH protein (see above).

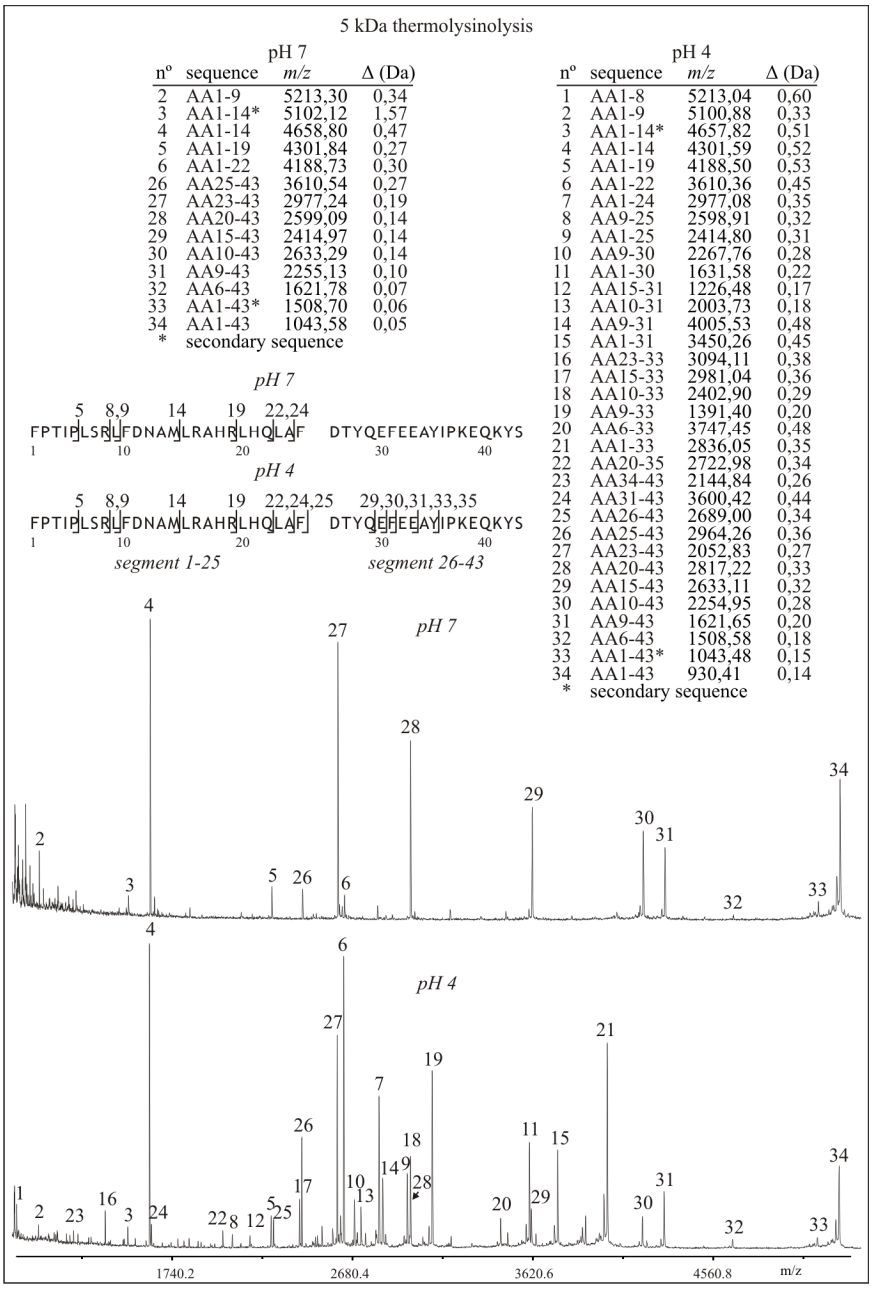


Figure 13. MALDI-TOF mass spectra corresponding to the thermolysinolysis of the 5 kDa protein at a neutral pH (top) and under acidic conditions (bottom). Sequences attributed to the obtained peaks, experimental mass values and the differences towards the theoretical values (in Da) are indicated in the left part (top and bottom parts correspond to the respective mass spectra).

Finally, the oligomerisation state of the 5 kDa protein was determined through covalent intra-molecular cross-linking reactions of the oligomers in solution and posterior analysis with SDS-PAGE electrophoresis. Whereas covalent oligomers could be directly viewed in the gel, non-covalent oligomers required to previously covalently fix them using the TAT reagent (Dienys *et al.*, 1998). Shown in figure 14, non-covalent dimers were observed in protein analysis (lane 1), but multiple non-covalent oligomers were showed under the same buffer conditions used in limited proteolysis (lane 3). This non-covalent oligomerisation state was disrupted with the addition of urea (8 M) as expected (Bennion *et* Daggett, 2003), observing exclusively the monomer in such conditions (lane 4). To underline, LP of the protein with the addition of urea generated comparable mass spectra (not shown), thus discarding any relation between the protein's oligomerisation and the LP cleavage pattern.

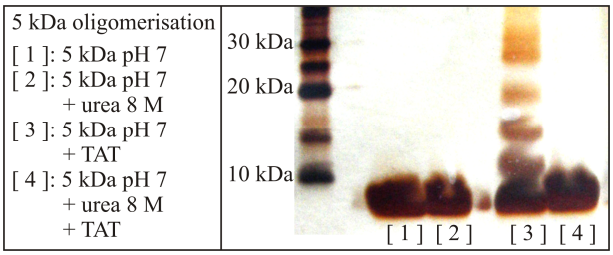


Figure 14. SDS-PAGE of the 5 kDa protein analysed before and after intra-molecular cross-linking reaction with the 1,3,5-triacryloyl-hexahydro-s-triazine (TAT) reagent.

Discussion

Thermolysinolysis of the recombinant 22 kDa GH protein resulted in the 17 kDa and 5 kDa fragments, allegedly reported to exist *in vivo* at high concentrations in blood circulation. The release of both fragments occurs through a specific cleavage at the $AA_{43\text{-}44}$ bond, with additional cleavages

at the adjacent bonds. Based on LP precepts, first cleavages of an enzyme seem indicative of a lack of structure, high mobility and exposition for the targeted region (Fontana *et al.*, 2004; Hubbard, 1998; Hubbard *et al.*, 1998). Hence, the thermolysin activity reported here suggests a lack of structure for the AA_{38-47} region in the free state of the protein in solution. This is in agreement to the existent data on the free state of the protein, reporting an undefined (Kasimova *et al.*, 2002) or irregular (Chantalat *et al.*, 1995) structure for the region, in comparison with the GH receptor – bound protein, where it is shown to be structured in a mini-helix (de Vos *et al.*, 1992; Duda *et* Brooks, 2003).

The same thermolysin-like activity was also observed for pepsin at acidic pH (Spolaore *et al.*, 2004), results also reproduced here. In this case, it was hypothesised a relation between the pH of the media and the structure of the AA_{38-47} region. The fact that thermolysin performs comparable cleavages at neutral pH suggests that this activity is not a function of the pH. Indeed, the equivalent activity of thermolysin shown either at neutral or acidic pH, with TFE or with urea also appears to indicate a resistant structure of the protein. This is in agreement to previous reports stating a highly stable folding state of the protein, adopting at low pHs a "molten globule" state characterised by a higher mobility of the unstructured regions (Kasimova *et al.*, 2002; Bam *et al.*, 1996). Here, the slight change in the specificity regarding the preferred cleavage at positions AA_{43-44} or AA_{44-45} shown at neutral and acidic pH / with TFE, respectively, could be probably attributed to a minor conformational accessibility.

The next cleavages shown by thermolysin occurred in the segment $AA_{134-150}$, that constitutes the largest loop of the protein. In concordance, most of proteases previously evaluated were reported to perform a first cleavage in

the same region (Wroblewski *et al.*, 1993; Alam *et al.*, 1998; Komatsu *et al.*, 2007). Thus both regions could constitute the most unstructured, exposed and flexible parts of the protein. In fact, the first thermolysin cleavage in the AA_{38-47} segment appears not to modify the resulting structure of the 17 kDa fragment, as shown in the respective LP analysis, allowing a second cleavage at the $AA_{134-150}$ region. This situation was already reported in the past for 22 kDa proteolysis in rat thyroid glands extracts (Wroblewski *et al.*, 1993). The fact that the opposite situation has not been observed, suggests that a first activity in the $AA_{134-150}$ loop could strongly perturb the rest of the protein structure. Nevertheless, the reasons that determine which of the two regions are cleaved first are unclear. For some proteases, such as V8-protease and proteinase K, the first regions to be cleaved were located within structured parts of the protein, at the first (AA_{9-34}) and at the third ($AA_{106-128}$) α-helices, respectively. These results are in apparent conflict with LP basis, where a proteolysis over unstructured regions is expected. As such, it underlines the necessity to include a sufficiently wide group of proteases in the design of any LP experiment, in order to yield a proper interpretation of the results.

The activity observed for thermolysin allowed to develop a purification protocol to produce the 17 kDa and 5 kDa GH fragments. This process could be considered as homologous to the hypothetical *in vivo* generation, thus preserving possible structural particularities of the fragments considering that these derive from the 22 kDa protein. Notwithstanding, this process has not yielded pure proteins. The final collected mixtures could be entirely representative of the hypothesised real GH variants, stated to be a mixture of fragments mostly derived from a main cleavage at the AA_{43-44} bond (Lewis *et al.*, 2000). For the 5 kDa fragment, the possibility of a chemical synthesis by SPS technique was feasible given

the limited number of amino acids. This procedure has been accomplished here, obtaining more than 100 mg of essentially pure protein. In this case, the final structure is derived from the inherent tendency of the sequence once deprotected and released from the resin used for the synthesis. In consequence, this could differ from the native structure of the protein once released from the 22 kDa protein. The characterisation through CD and LP measurements have provided some insight of its structural properties.

CD data showed a partial helicoidal structure in agreement with the conformation shown on the AA_{1-43} sequence within the 22 kDa structure (Kasimova *et al.*, 2002), where the central part (AA_{9-34}) of the 5 kDa sequence is described to form an a-helix. The addition of TFE to the media resulted in a fast increment of the a-helix content revealing a low resistance of the rest of the protein to adopt a folded state, synonym of a notable degree of flexibility within the chain. LP data showed a major enzymatic activity over the AA_{1-25} N-terminal part of the sequence in neutral conditions, thus suggesting that the rest of sequence, the C-terminal segment (approximated to the AA_{26-43} segment) was then structured. These results differ from the a-helix at the AA_{9-34} region present in the 22 kDa sequence. In addition, a global cleavage was observed at acidic pH, suggesting a major unfolding of the protein. Furthermore, the 5 kDa fragment was observed to oligomerise at neutral pH, although the enzyme's activity was shown to be independent of the oligomerisation state. This structure suggested for the 5 kDa fragment, different from the AA_{1-43} part of the 22 kDa sequence, could be related to the lack of affinity shown by the anti-GH antibodies towards this fragment, and to the lack of affinity shown by the anti-5 kDa antibodies towards the 22 kDa protein. This situation, evidenced and discussed in chapter 4., was similar both for the synthetic and for the proteolytic 5 kDa

fragments, the latter obtained by LP of the 22 kDa protein. Thus, similar structured between both 5 kDa fragments were suggested.

The enzyme allegedly processing *in vivo* the 22 kDa GH in humans and releasing the 17 kDa and 5 kDa products remains unidentified. However, the activity shown by thermolysin indicates that this process is possible, and that GH has the structural requirements for being hydrolysed between AA_{43} and AA_{44} by appropriate enzymes. Where this production could occur is also unknown, albeit that the low concentration found in the pituitary and the high concentration allegedly found in blood, suggests that it might occur after its release from the gland (Sinha *et* Jacobsen, 1994). The biological relevance of both fragments consists in a potent insulin-like effect for the 5 kDa fragment inducing a glucose clearance (anti-diabetogenic effect) (Frigeri *et al.*, 1988; Mondon *et al.*, 1988), and opposite glucose intolerance properties for the 17 kDa fragment (diabetogenic effect) (Hettiarachchi *et al.*, 1997). Specific and unidentified receptors mediating these actions are presumed, in sight of the low (17 kDa) or null (5 kDa) interaction of both fragments towards the GH and human prolactin receptors (Haro *et al.*, 1996b). Finally, both fragments could have a potential use as pharmaceutical drugs. The endogenous 17 kDa fragment was speculated to prevent hypoglycemia in post-exercise states (Wallace *et al.*, 2001b), suggesting a possible pharmacological use in this sense. Similarly, the potent insulin-like activity of the 5 kDa fragment could be used in treatments for type 2 diabetes mellitus. Actually, it could possibly be a better candidate than the 22 kDa protein which was already proposed (Ahn *et al.*, 2006), due to the possible lack of secondary effects. Furthermore, the 5 kDa fragment has been identified as a novel substrate for the dipeptidyl peptidase IV (DPIV), responsible for the inactivation of an incretin protein which elicit

insulin secretion from the islets (Zhu *et al.*, 2003). Used for treatments of the same disease, the 5 kDa fragment could be also revealed as a potential DPIV inhibitor (Boyle *et* Freeman, 2007).

4. Human growth hormone detection by surface plasmon resonance

The measurement of human growth hormone (GH) in biological samples through immunoassays yield results that display a pronounced disparity (see chapter 1.5.). One of the main causes for such variability has been stated to reside in the heterogeneous composition of growth hormone and the differential detection by the antibodies used. Two causes contribute to this situation. First, the lack of a characterisation protocol in current immunoassays. Second, the uncertainty of which exactly are the GH variants in blood circulation. A first protocol of characterisation was proposed by Jansson *et al.* (Jansson *et al.*, 1997), that included the 22 kDa, 20 kDa, 17 kDa and 5 kDa GH proteins in the process. The inclusion of these four proteins was supported by the literature. There is a clear consensus with the 22 kDa and 20 kDa variants as the most predominant GH proteins in circulation (Baumann, 2009). Concerning the 17 kDa and 5 kDa fragments, allegedly even more concentrated than the 22 kDa protein (Sinha *et* Jacobsen, 1994), the controversy is still there. Between all of them, a high cross-reactivity can be assumed given the coincident sequences: the $AA_{1\text{-}191}$ sequence of the 22 kDa protein includes both the 5 kDa (although anti-22 kDa antibodies did not cross-react to this variant; see below) and 17 kDa sequences, $AA_{1\text{-}43}$ and $AA_{44\text{-}191}$ respectively, and the 20 kDa sequence is constituted by the $AA_{1\text{-}191}$ sequence lacking the segment $AA_{32\text{-}46}$.

In this work, the same four GH variants have been employed individually in a characterisation protocol directly evaluating a number of anti-GH antibodies rather than the GH immunoassays themselves. The group of analysed (monoclonal) antibodies is composed by two anti-GH antibodies (A36, A1549), the former reported to recognise uniquely the 22 kDa

variant (Tsushima *et al.*, 1999), two other antibodies directed towards the 20 kDa GH variant [D05 (Hashimoto *et al.*, 1998), hGH33 (Mellado *et al.*, 1996)], and one monoclonal antibody (A5K), generated in collaboration with a Chinese laboratory and directed towards the synthetic 5 kDa GH fragment previously obtained (see chapter 3.).

Cross-reactivity studies were performed by means of surface plasmon resonance (SPR). Briefly, based on an optical phenomenon, SPR detectors monitor changes in refractive index near a sensor surface, within the first ~ 300 nm of a small flow cell, through which an aqueous solution (running buffer) passes under continuous flow (1-100 µl/min). As such, the accumulation of material over the surface results in a change of the refractive index, corresponding to a signal which is plotted as response or resonance units (RUs) versus time (a sensorgram). Monitoring of specific interactions between antibodies and antigens requires a previous immobilisation of one of the two onto the surface, whereas the other is injected in aqueous solution. The sensorgram obtained comprises a first association and a posterior dissociation phase (see figure 15), whereas one RU is proportional to the binding of ~ 1 pg of protein per mm^2 in the BIAcore 3000 instrument design (surface and volume of a flow cell of 1.2 mm^2 and 20 nl, respectively). Typically more than 10-30 RU are needed to provide a signal clearly differentiated from background noise (Jason-Moller *et al.*, 2006).

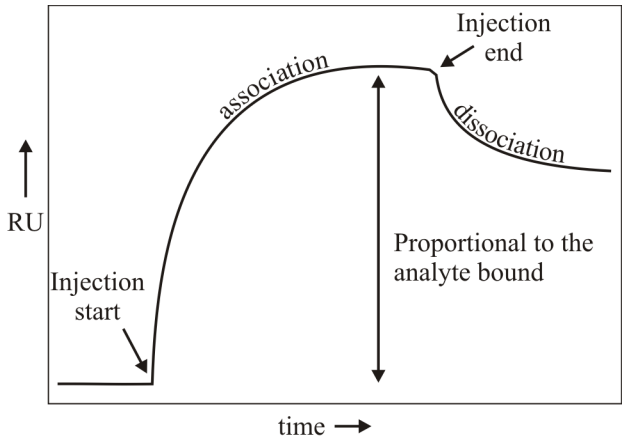


Figure 15. Typical sensorgram obtained from a SPR measurement. Association and dissociation phases are indicated in the graph.

This SPR-based characterisation protocol has allowed to characterise each antibody towards the different GH variants, obtaining for each case kinetic data (Hahnefeld *et al.,* 2004). Furthermore, the capabilities of the SPR for the direct detection of GH have also been evaluated. In order to improve the sensitivity of the assay, a sandwich approach has been performed to amplify the SPR signal, taking profit of the rise of the signal proportional to the mass accumulated over the surface.

Material and methods

Materials. Recombinant 22 kDa protein (WHO, IS 98/574) was purchased from National Institute of Biological Standards and Controls (NIBSC, Hertfordshire, UK). Recombinant 20 kDa protein was kindly provided by Mitsui & Co. LTD (Tokyo, Japan). Proteolytic 17 kDa GH fragment and synthetic 5 kDa GH fragment were obtained as described in chapter 3. Monoclonal anti-22 kDa antibody (mAb) (A36020047P, hereafter referred as A36) was purchased from Biospacific Corporate (Emeryville, CA, USA). Anti-GH mAb (A1549) was purchased from Bioclone

(Marrickville, Australia). Antibodies hGH-33 and D05 were provided by Dr. Mellado (CSIC Madrid, Spain) and Dr. Ueki (Mitsui Pharmaceuticals, Japan), respectively. 1-Ethyl-3-(3-dimethylamino-propyl) carbo-diimide (EDC), *N*-hydroxy-succinimide (NHS), ethanolamine, sensor chips CM5, and HBS-EP buffer (10 mM hepes, 150 mM NaCl, 3.4 mM ethylene diamine tetraacetic acid (EDTA) and 0.005% tween20, pH 7.4) were purchased from Biacore (GE Healthcare, Barcelona, Spain). Pepsin, papain, and dithiothreitol (DTT) were purchased at Sigma-Aldrich (Barcelona, Spain). Centricon YM centrifugal filter devices of 30 kDa molecular weight cut-off were purchased to Millipore Corporation (Billerica, Mass., USA).

Generation of anti-5 kDa monoclonal antibody (A5K). Two six weeks old female Balb/c mice were immunised three times by subcutaneous injection. The first immunisations were with 50 µg of synthetic AA_{1-43} in Freund's complete adjuvant, followed by a second and third immunisations with 50 µg of synthetic AA_{1-43} in Freund's incomplete adjuvant every two weeks, respectively. One month following the last immunisation, serum was obtained from each mouse and assayed for the presence of antibodies specific to synthetic AA_{1-43}, with the screening enzyme-linked immunosorbent assay (ELISA) procedure, described below. The mouse with the highest antibody titre received a boost of 50 µg of synthetic AA_{1-43} in Freund's incomplete adjuvant three days prior to removal of the spleen. The splenocytes were fused to SP2/0 myeloma cells using the procedure described by Kennett (Kennett, 1979). Hybridomas were selected in hypoxanthine-aminopterin-hymidine medium. Hybridomas producing anti-synthetic AA_{1-43} were screened through the antibody capture ELISA. A limited number of hybridomas corresponding to a single clone were re-assayed for anti-5 kDa activity

through the supernatants, selecting the clone of highest sensitivity.

ELISA procedure: High binding ELISA plates (Corning 3922, Corning, New York, USA) were coated with 100 µl/well of a 1 µg/ml synthetic AA_{1-43} solution and incubated overnight at 4 °C. The wells were blocked with 200 µl/well of 1% BSA-PBS (8 mM sodium phosphate, 2 mM potassium phosphate, 140 mM sodium chloride, 10 mM potassium chloride, 1% w/v bovine serum albumin (BSA), pH 7.4) for 1 h at 37 °C. A total of 100 µl / well of hybridoma culture supernatant or two-fold serial dilutions of monoclonal antibody in 1% BSA-PBS were added to the plate for 1 h at 37 °C. The wells were washed three times with PBS-Tween 20 (8 mM sodium phosphate, 2 mM potassium phosphate, 140 mM sodium chloride, 10 mM potassium chloride, 0.1% v/v Tween 20, pH 7.4). 100 µl of peroxidase-coupled goat anti-mouse IgG (ImmunoHunt, Beijing, China) at a dilution of 1:10000 were then added to the wells and the binding reaction was carried out for 1 h at 37 °C. The wells were then washed five times with PBS-Tween 20. A total of 100 µl of 3,3´,5,5´-tetramethylbenzidine substrate solution was then added to each well. After 15 min incubation, the reaction was stopped with the addition of 50 µl / well of 2 M H_2SO_4. The optical density (450 nm) of each well was subsequently measured with a Tecan Sunrise Elisa plate reader. Hybridomas of interest were further subcloned by limiting dilution. Supernatants from wells containing a single clone were re-screened to verify the binding to synthetic AA_{1-43}.

Surface Plasmon Resonance. SPR experiments were performed using a BIAcore™ 3000 (GE Healthcare, Barcelona, Spain) at 25.0 ºC, employing CM5 sensor chips. The immobilisation of the antibodies to singular surface channels was achieved by standard amine coupling indicated by

the manufacturer. These were injected at antibodies concentrations of 0.06 g/l (60 µl), 0.8 g/l for the A5K antibody, in a buffered solution at pH 5.0 (10 mM sodium acetate). A reference channel was generated following the same procedure without antibody injection in order to account for non-specific binding events. Sensorgrams were recorded at a flow rate of 5 µl/min and analyte (typically a sample volume for a 3-min association phase) was allowed to interact with the surface-bound material. In the sandwich approach, the secondary antibody A1549 was injected subsequently at a volume and concentration of 30 µl and 65 nM. The regeneration of the surfaces for the next round of experiments was achieved with a 10 µl pulse of a 10 mM Gly solution buffered with formic acid to a pH of 2.0. All analysed samples were diluted in HBS-EP buffer. Kinetic experiments were performed at a higher flow rate of 20 µl/min and covering a concentration range between 1- and 20 nM in 6 different solutions (3-min association phase). Analysis of the SPR curves was performed through a mathematical fit consisting in a 1:1 interaction employing the BIAcore Control evaluation software 4.0.1.

Antibody proteolysis. Papain and pepsin proteolyses were carried out in sodium phosphate 10 mM (pH 7, papain) or sodium acetate 10 mM (pH 4, pepsin), using the enzymes immobilised to agarose in E/S ratios 30/1 (w:w), with an initial amount of antibody of 10 µg. The final mixture was filtered in Centricon devices and the retentate was collected in the appropriate buffer for the SPR immobilisation. Papain enzyme was previously activated by incubation with DTT 60 mM (30 min, 37 ºC), removed by centrifugation and discarding the supernatant.

SDS-PAGE. SDS-PAGE was carried out on Mini-Protean 3 Cell Electrophoresis System of Bio-Rad Laboratories (Hercules, CA, USA),

under conditions of 120 V constant and 7.5% acrylamide gels.

Immulite. Measurements of GH levels were performed using a solid-phase, two-site chemiluminescent immunometric assay from Diagnostic Products Corporation (Immulite-1000, Siemens-DPC, LA, USA).

Results

The antibodies anti-GH (A36 and A1549), anti-20 kDa (D05 and hGH33) and anti-5 kDa (A5K) were immobilised separatedly onto carboxymethylated dextran surfaces, generating covalent links with the amine groups located at the antibody's surface. After the immobilisation step, each antibody showed a certain percentage of active antibodies, values shown in table 2, that mostly determine the sensitivity of the surface (Bonroy *et al.*, 2006). These values entirely depend on the random distribution of their amine groups and thus, cannot be predicted.

Table 2. For each antibody immobilised, it is indicated the (1) immobilisation value (resonance units, RU), the (2) calculated Rmax* (RU), the (3) value corresponding to a saturating concentration of the antigen (RU), and the (4) surface activity's percentage [#]. * *Rmax = immobilisation value · stoichiometry · MW analyte / MW antibody.* [#] Obtained from the quotient between the experimental and theoretical maximum response of the surface.

	(1)	(2)	(3)	(4)
A36	13093	3840	2189	57 %
A1549	12430	3646	1083	30 %
D05	9702	2587	60	2 %
hGH33	8153	2174	172	8%
A5K	11504	767	593	76 %

Once immobilised, the characterisation of the different affinities of the antibodies was performed by assaying the group of GH variants composed by the 22 kDa, 20 kDa, 17 kDa and 5 kDa GH proteins. Data

corresponding to cross-reactivity percentages is shown in table 3, also including the hypothesised segment of sequence that could contain the epitope. The sensorgrams for each antibody surface are shown in figure 16 (A-E).

Table 3. Relation of crossed activities for the five antibodies evaluated and the 22, 20, 17 and 5 kDa GH variants, expressed in percentages relative to the highest value. Bold numbers refer to the preferential GH variant recognised. The sequence segment hypothesised to contain the epitope is indicated in each case.

	22 kDa	20 kDa	17 kDa	5 kDa	segment
A36	**100**	< 1.3	< 1.1	< 0.3	AA_{32-43}
A1549	**100**	76.9	31.1	< 0.9	AA_{44-191}
D05	< 2.7	**100**	< 2.7	< 2.7	AA_{31}-AA_{47}
hGH33	< 2.8	42.9	**100**	< 0.7	AA_{46-50}
A5K	< 3.7	< 3.7	< 3.7	**100**	AA_{1-43}

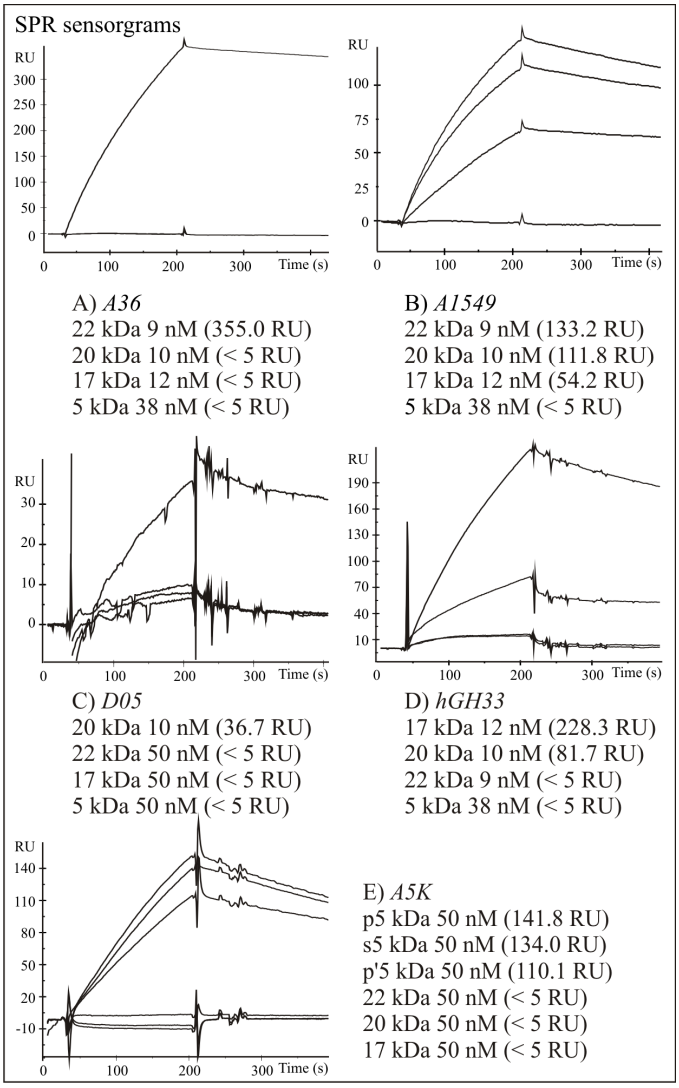


Figure 16. SPR Sensorgrams (reference curve subtracted) corresponding to the interactions of the four GH variants: 5 kDa, 17 kDa, 20 kDa and 22 kDa, and the antibodies A36, A1549, D05, hGH33 and A5K. Values expressed in resonance units (RU) are indicated between brackets. * p5 kDa = proteolytic 5 kDa of sequence $AA_{1\text{-}43}$; p'5 kDa = proteolytic 5 kDa of sequence $AA_{1\text{-}44}$; s5 kDa = synthetic 5 kDa of sequence $AA_{1\text{-}43}$.

Subsequently, for each pair of antibody and responding GH variant, a series of concentrations within the range of 1-80 nM were analysed. The fit of a mathematical model to the obtained curves, accounting for a 1:1

stoichiometric reaction with a correction for mass transport effects was carried out. Association (k_a) and dissociation (k_d) rate constants, and the equilibrium dissociation constant (k_D) were obtained, shown in table 4.

Table 4. Relation of association (k_a) and dissociation (k_d) rate constants, and the resulting dissociation equilibrium (k_D) constant for each pair of antibody and antigen, expressed in 1/Ms, 1/s and M units, respectively.

		A36	A1549	D05	hGH33	A5K
22 kDa	k_a	$8.92 \cdot 10^{5}$	$1.68 \cdot 10^{6}$			
	k_d	$1.51 \cdot 10^{-4}$	$7.79 \cdot 10^{-4}$			
	k_D	$1.7 \cdot 10^{-10}$	$4.6 \cdot 10^{-10}$			
20 kDa	k_a		$1.56 \cdot 10^{6}$	$8.5 \cdot 10^{5}$	$1.69 \cdot 10^{4}$	
	k_d		$6.36 \cdot 10^{-4}$	$8.5 \cdot 10^{-4}$	$1.59 \cdot 10^{-4}$	
	k_D		$4.1 \cdot 10^{-10}$	$1.6 \cdot 10^{-8}$	$9.4 \cdot 10^{-9}$	
17 kDa	k_a		$6.61 \cdot 10^{5}$		$1.23 \cdot 10^{5}$	
	k_d		$4.6 \cdot 10^{-4}$		$1.49 \cdot 10^{-3}$	
	k_D		$6.9 \cdot 10^{-10}$		$1.2 \cdot 10^{-8}$	
5 kDa	k_a					$4.55 \cdot 10^{4}$
	k_d					$2.52 \cdot 10^{-3}$
	k_D					$5.5 \cdot 10^{-8}$

Additionally, this batch of GH variants was also analysed through the commercial kit immunoassay based on the Immulite system which employs a pair of monoclonal and polyclonal antibodies. Under equimolar concentration settings similar cross reactivity (namely 100 %) was obtained for the 22 kDa and 20 kDa GH variants, a 5 % for the 17 kDa fragment, and less than 0.25 % for the 5 kDa fragment.

From all the antibodies evaluated, the D05 showed a significant low percentage of active antibodies after the immobilisation step (table 2). This value, indicative of a favoured coupling through the Fab domains of the antibody, severely penalised the final sensitivity of the surface. A fragmentation of the antibody through papainolysis and pepsinolysis prior to the immobilisation was addressed in an attempt to modify the resulting

antibody orientation. The SDS-PAGE analysis (figure 17) of the D05 papainolysis generated Fab'$_2$ (~ 100 kDa) and Fc' (~ 20 kDa) fragments (lane 1), and for equivalent pepsinolysis (lane 2) similar Fab'$_2$ fragments observed at a slightly lower MW (~ < 100 kDa), with an additional band at ~ 65 kDa that was attributed to a Fab'$_2$ fragment partially digested (Inouye *et* Ohnaka, 2001). In this case, no Fc' fragments were observed, in concordance with the reported digestion of the Fc dominion of the antibodies by pepsin (Boushaba *et al.*, 2003) (and posterior filtration with the 30 kDa filter device). The immobilisation of the pepsinolysis mixture yielded a slight improvement in the resulting activity percentages (6.3 % for pepsinolysis, 1.5 % for papainolysis, 2 %, for the entire antibody), although this was much lower than the expected.

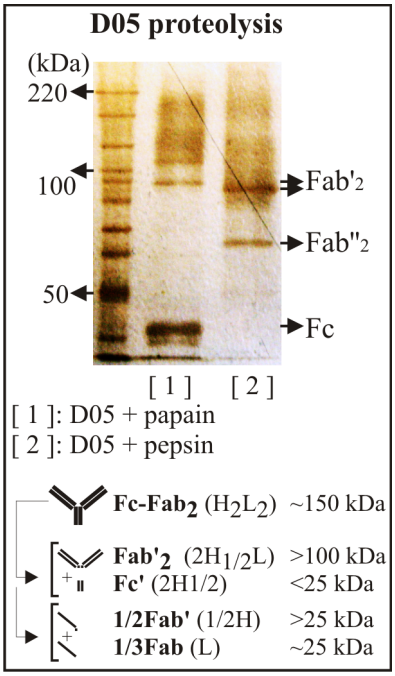
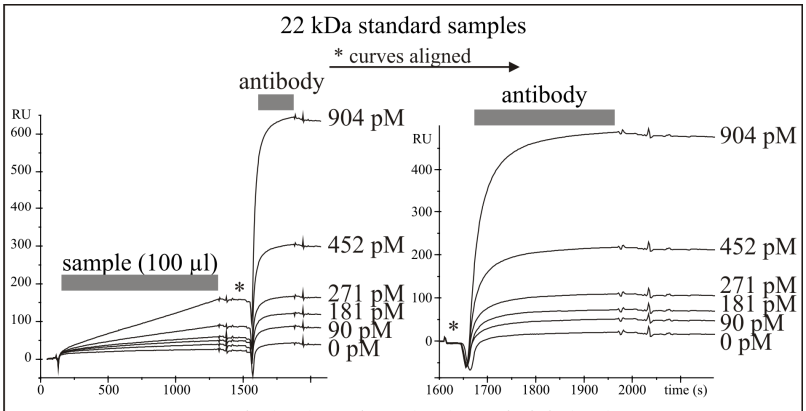


Figure 17. SDS-PAGE of the D05 antibody analysed after proteolysis with papain and pepsin. A scheme of the different antibody fragments obtained after proteolysis, with the correspondent molecular weights (kDa), is included. H and L refer to heavy and light chains of the antibody.

Among the antibodies evaluated, the A36 antibody had a high percentage of activity after the immobilisation step. Together with its low k_D towards the 22 kDa protein, it constituted the best candidate to develop a SPR-based method for the detection of GH. In this case, highly specific towards the 22 kDa GH variant. Furthermore, a signal enhancement was achieved by a sandwich approach, employing the A1549 antibody which showed a comparable k_D. The signal increment would be proportional to the mass accumulated over the surface, *i.e.* ~ 150 kDa for an antibody. Assays were conducted at a low flow rate (5 µl/min), in order to observe the maximum yield in the binding reaction. A first assay consisted in the analysis of 22 kDa samples at the concentration range of 90 – 904 pM injecting 100 µl of volume (+ 40 µl of dead volume), followed by the injection of 30 µl of A1549 at a concentration of 65 nM. Significantly lower amplifications were obtained when using lower A1549 concentrations. The curves obtained in the respective sensorgrams, included in figure 18, showed a proper amplification of the signal with the secondary antibody, providing a clear response at concentrations as low as 181 pM (4 ng/ml) with a final value of ~ 115 RU.

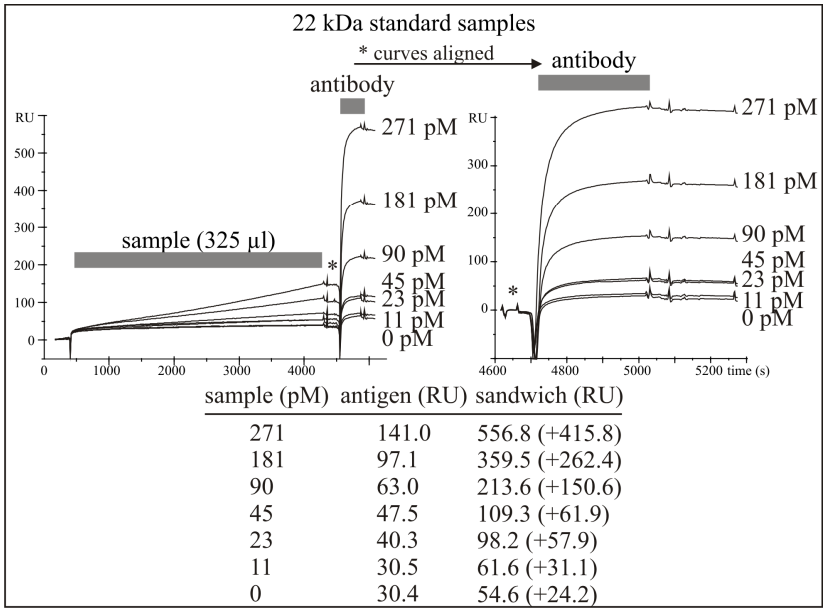


sample (pM)	antigen (RU)	sandwich (RU)
904	153.3	631.6 (+478.3)
452	81.5	295.0 (+213.5)
271	52.1	159.3 (+107.2)
181	43.6	114.8 (+71.2)
90	31.4	79.5 (+48.1)
0	9.5	36.0 (+16.5)

Figure 18. Values corresponding to the injection of 100 µl of a series of 22 kDa samples (pM) over a surface with the A36 antibody coated, followed by the injection of the A1549 antibody. Left sensorgrams: entire SPR sensorgrams. Right sensorgrams: part of the same SPR sensorgrams, aligned, corresponding to the A1549 amplification. Bottom: values indicating response (RU) for the 22 kDa increment (column 2) and for the total increment (column 3), the part corresponding to the antibody between brackets.

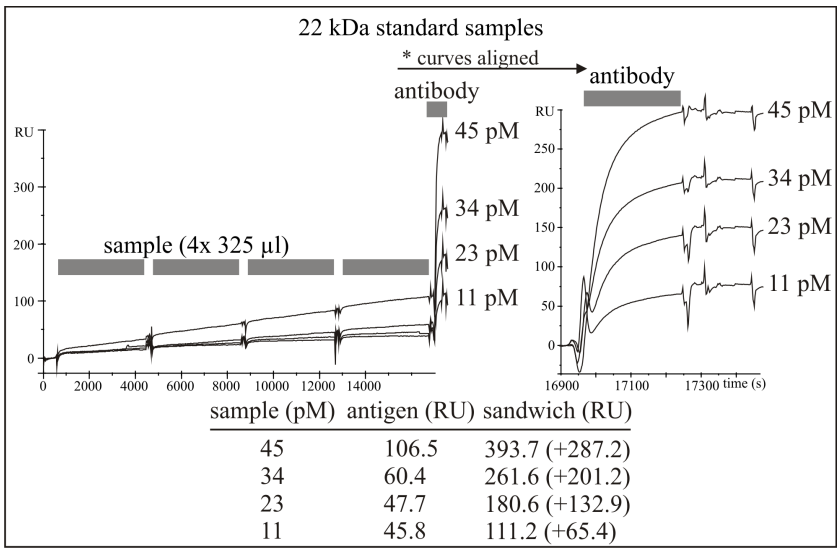
The sustained binding of the 22 kDa protein at the surface, observed from the sensorgram's slope (Rich *et* Myszka, 2008), suggested a higher amplification if employing higher sample volumes. In contrast, the slope for A1549 indicated a state near to equilibrium. Thus, 22 kDa samples were analysed at the maximum of 325 µl of sample volume per injection (limited by the instrument design; + 40 µl of dead volume), with one injection (22 kDa range between 11-271 pM) and four consecutive injections (range between 11-45 pM). For the former (sensorgrams included in figure 19), a clear response was obtained at concentrations as low as 45 pM (1 ng/ml) with a final value of ~ 109 RU. For the latter (sensorgrams included in figure 20), ~ 180 RU were obtained for the analysis of a concentration of 23 pM (0.5 ng/ml). For the three

experimental designs, SPR values (taken 2.5 minutes into the dissociation phase) versus 22 kDa concentrations are represented in figure 21, including the corresponding adjusted equations.



sample (pM)	antigen (RU)	sandwich (RU)
271	141.0	556.8 (+415.8)
181	97.1	359.5 (+262.4)
90	63.0	213.6 (+150.6)
45	47.5	109.3 (+61.9)
23	40.3	98.2 (+57.9)
11	30.5	61.6 (+31.1)
0	30.4	54.6 (+24.2)

Figure 19. Values corresponding to the injection of 325 µl of sample volume. See legend of figure 18.



sample (pM)	antigen (RU)	sandwich (RU)
45	106.5	393.7 (+287.2)
34	60.4	261.6 (+201.2)
23	47.7	180.6 (+132.9)
11	45.8	111.2 (+65.4)

Figure 20. Values corresponding to the injection of 1.3 ml of sample volume. See legend of figure 18.

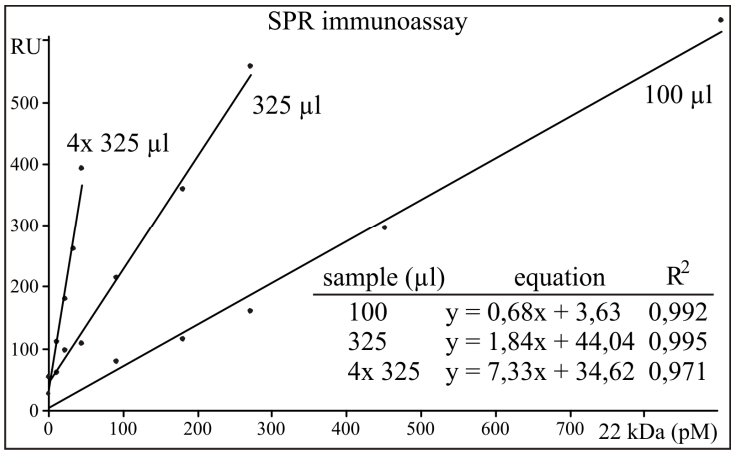


Figure 21. Graphical representation of 22 kDa concentration versus the SPR response, of the three assays characterised by the samples volumes of 100 µl, 325 µl and 4x325 µl. Curves equations are included.

Discussion

Characterisation of anti-GH antibodies

The antibodies characterised here, allegedly responsive to total GH, or specific to the 20 kDa and the 5 kDa variants, have revealed a wide variability with respect to the different GH variants that are recognised. Strong cross reactions were observed for the anti-GH A1549 and the anti-20 kDa hGH33 antibodies, in contrast with the highly specific anti-GH A36 and anti-20 kDa D05 antibodies. Thus, differences could be expected when comparing hypothetical immunoassays based on these antibodies. Furthermore, the variability observed within the group of 22 kDa, 20 kDa, 17 kDa and 5 kDa GH proteins is probable to exist as well with other GH variants and members of the GH family (*e.g.* prolactin or placental lactogen). In fact, a group of these variants has already been employed in the characterisation of other anti-GH antibodies in our lab, and the results

obtained support this hypothesis. This different specificity has been stated to cause at least partially the present disparity of results between GH immunoassays (see chapter 1.5.). As such, it is strongly suggested to implement a characterisation protocol, similar to the one here reported, for the characterisation of the current anti-GH immunoassays.

The A36, the D05 and the A5K antibodies showed in all cases exclusive affinities towards the 22 kDa, 20 kDa and 5 kDa proteins, respectively. In these cases, those sequence segments not present in the non-recognised variants should contain the epitope, *i.e.* AA_{32-43} for the A36, the bond AA_{31}-AA_{47} for the D05, and the AA_{1-43} for the A5K. For the case of A36, the lack of affinity for the 5 kDa fragment might be explained by the apparent structural differences of the fragment (see chapter 3.). Equally, the lack of affinity of the A5K antibody to the 22 kDa protein might have the same explanation. In agreement, other anti-5 kDa antibodies were also reported to have a very low cross-reactivity towards the 22 kDa protein (Lopez-Guajardo *et al.*, 1998; Shimizu *et al.*, 2004). The wide specificity observed for the A1549 antibody, could correspond to an epitope located at *a priori* any part of the 17 kDa sequence (AA_{44-191}). The hGH33 antibody recognised the 20 kDa and 17 kDa GH proteins, but not the 22 kDa protein. In this case, the global structure of the 17 kDa fragment was apparently not modified once released from the 22 kDa protein (see chapter 3.). However, the part of the 22 kDa sequence not present in both variants (17 kDa: AA_{1-43}; 20 kDa: AA_{32-46}) could induce particular structural differences in the N-terminus of the 17 kDa fragment and the equivalent sequence in the 20 kDa protein, that could therefore contain the epitope. In the case of the Immulite immunoassay, the epitope recognised should be located in the AA_{1-32} region, not showing interaction towards the 5 kDa fragment for the same reasons above described. The little cross-

reaction towards the 17 kDa fragment could correspond to the presence of the 22 kDa protein in the final 17 kDa purification.

SPR immunoassay

Based on a SPR sandwich protocol employing the A36 (immobilised) and A1549 (secondary) antibodies, the recombinant 22 kDa GH protein has been detected in a non-biological matrix at the concentration of 23 pM (0.5 ng/ml). In this design, the final sensitivity of the assay is entirely dependent of the sample volume employed. Unfortunately, considering the impossibility to analyse undiluted biological samples in this specific instrument (BIAcore 3000), these values do not allow the direct detection of GH in biological fluids. Notwithstanding, this method is still suitable for the analysis of purified samples, an option that has been explored in the next chapter.

Additionally, as a proof-of-principle, this design could be employed with other antibodies, specific for other GH variants. For the surface antibody, it should show a sufficient activity after the immobilisation process. The amine chemistry employed here provides, in our experience, the highest values of density of the surface (resulting in the highest sensitivity). With this chemistry, the final orientation of the antibody depends on the antibody primary sequence. As such, in those cases with poor values of immobilisation, the prior proteolysis of the antibody could yield a better orientation in the surface. This process is however also dependent on the characteristics of the antibody. In the case here reported, this procedure yielded minor improvements.

Finally, comparing the WB and SPR-based immunoassays, both approaches offer a different equilibrium between sample consumption,

sensitivity and the information obtained. WB requires pre-purification of the highest sample volume, but provides a protein identification by molecular weight. In comparison, ELISA and analogues immunoassays offer the minimum sample consumption, and provides a unique numerical value. Complementarily, SPR lowers the volume requirements and generates a curve corresponding to the association and dissociation phase of the binding reaction in real time.

5. Human growth hormone purification

The human GH family is comprised by several variants of different sequence composition and with distinct post-translational modifications (Baumann, 2009). One of these allegedly consists in a rather specific proteolytic cleavage at the AA_{43-44} bond, originating the so-named N-terminal 5 kDa and C-terminal 17 kDa GH fragments (Lewis *et al.*, 2000). These were detected at high concentrations in blood samples (Lopez-Guajardo *et al.*, 1998; Sinha *et* Jacobsen, 1994), the 17 kDa fragment at even higher concentrations than the 22 kDa GH protein, employing both a (polyclonal) antibody detecting most of GH variants, and a specific (monoclonal) antibody for the 17 kDa variant (Sinha *et* Jacobsen, 1994). However, these results were not reproduced in other posterior assays, generating a controversy that still persists. Conversely, concerning the GH immunoassays of general use, it remains unclear whether these actually recognise any of the two fragments. In this regard, a thorough characterisation of some GH immunoassays was performed by Jansson *et al.* (Jansson *et al.*, 1997), reporting a lack or very low immunoreactivity towards both fragments. The structural particularities reported for the 5 kDa fragment (see chapter 3.) could explain the low or inexistent affinity with the evaluated anti-GH antibodies.

Here, a protocol for the immunopurification of blood and pituitary human GH has been developed based on the A1549 antibody. This antibody was shown to recognise the 22 kDa, 20 kDa and 17 kDa variants (see chapter 4.). Purified samples from healthy non-professional athletes, included basal and post-exercise states, have been submitted to Western Blot (WB) detection and additionally, to SPR analysis following the reported (chapter 4.) sandwich approach. The objective was to shed light on the conflicting

results reported for the presence or absence of some of the GH variants.

Materials and methods

Materials. Recombinant 22 kDa GH protein (WHO, IS 98/574) was purchased from National Institute of Biological Standards and Controls (NIBSC, Hertfordshire, UK). 20 kDa protein was kindly provided by Mitsui & Co. LTD (Tokyo, Japan). 17 kDa GH fragment was obtained as described previously (see chapter 3.), provided in a solution 6:4 (ACN:H_2O). Monoclonal anti-22 kDa antibody (mAb) (A36020047P, hereafter referred to as A36) was purchased from Biospacific Corporate (Emeryville, CA, USA). Anti-GH mAb (A1549) was purchased from Bioclone (Marrickville, Australia). Rabbit polyclonal anti-GH antibody GTX74405 corresponding to batches nº 18312 and 21491 was purchased from GeneTex (AntibodyBCN, Barcelona, Spain). Anti-Rabbit peroxidase antibody A9169 was purchased from Sigma-Aldrich (Barcelona, Spain). Anti-Mouse peroxidase antibody P0260 was purchased from Dako (Barcelona, Spain). 96-Well high-binding microtiter plates Corning ref #3590 were purchased at Sigma-Aldrich (Barcelona, Spain). Immobilon-P PVDF 0.45µm 7x8.4cm sheet IPVH07850, and Amicon Ultra-15, PLGC Ultracel-PL Membrane, 10 kDa, UFC901024 were purchased from Millipore (Barcelona, Spain). Amersham ECL™ Advance Western Blotting Detection Kit, RPN2135, was purchased from GE Healthcare (Barcelona, Spain). PerfectWestern small (B104) and medium (B101), clear containers were purchased from GenHunter (CliniSciences, France). 1-Ethyl-3-(3-dimethylamino-propyl) carbodiimide (EDC), N-hydroxy-succinimide (NHS), ethanolamine, sensor chips CM5, and HBS-EP buffer (10 mM hepes, 150 mM NaCl, 3.4 mM ethylene diamine tetraacetic acid (EDTA) and 0.005% tween20, pH 7.4) were purchased from Biacore (GE

Healthcare, Barcelona, Spain).

Serum samples. Blood samples were taken from volunteers that gave their consent to participate in the study and ethical approval was obtained from the local ethical committee. Samples from 4 volunteers (age 26-38 years, males) were obtained before and immediately after a 45 min running exercise (outdoor, ~ 30 ºC) under maximum effort. Blood samples were collected into vacuette serum tubes (Greiner bio-one, Madrid, Spain), and centrifuged for 15 min at 2700×g. Serum was stored at 4 ºC until analysis. Basal serum sample from another non-exercising volunteer was obtained from Hospital Del Mar (Barcelona, Spain).

Pituitary samples. Three pituitary glands were provided by the Hospital Del Mar (Barcelona, Spain). These were washed thrice (buffer solution, 50 mM Tris-HCl, 2 mM EDTA, at pH 7.4)) and homogenised in 1.5 mL of buffer with the addition of a protease inhibitor cocktail (Complete, Roche, 1 tablet for 50 ml of buffer) through a homogeniser Heidolph (Heidolph Instruments GmbH & Co. KG, Schwabach Germany) for the initial pieces of glands (~ 5 min), and then a homogenizer Ultra Turrax (IKA Works, Inc, Wilmington, NC, USA). After centrifugation (10 min, 500×g), the supernatant was transferred to a new tube and frozen at -80°C until analysis.

Solid phase immunopurification. 100 µl of A1549 antibody (5 µg/ml in PBS solution) per well were incubated in a microtiter plates support (Butler, 2004) at 37 ºC (1 h) and washed 2x with PBST (0.1 % Tween-20) (1 min). 100 µl of plasma sample were incubated in each well at 4 ºC (overnight), 1x rinsed, 1x washed (2 min), and incubated with formic acid (2 %) (5 min). Eluents were recovered from each well, mixed with

equivalent volume of H_2O mQ and filtered over Amicon filters (20 min, 4000 g) to pre-concentrate and obtain a salt-free solution. The resulting volume (~ 200 µl) was evaporated via speed-vac, reconstituted with 5 µl of H_2O mQ (no shaking, 15 min room temperature) and mixed with 5 µl SDS sample loader buffer 4x (non-reducing). Samples were loaded onto the gel. Standards samples followed the same pre-concentration and reconstitution steps, from a initial dilution in HBS-EP buffer. For SPR analyses, Amicon filtrations were repeated twice after the addition of 3 ml of H_2O mQ, to ensure a total wash of the acid formic from the eluted. The final purification once evaporated to dryness was re-suspended into a final volume of 390 µl (in HBS-EP buffer) and filtered through a 0.22 µm filter.

Western Blot. SDS-PAGE was carried out on Mini-Protean 3 Cell Electrophoresis System of Bio-Rad Laboratories (Hercules, CA, USA), under conditions of 120 V constant and 15% acrylamide gels, 1.5 mm thick. Semy-dry protein transfer to the PVDF membrane (TE77PWR, GE Healthcare) was performed at 45 mA, 45 min, in buffer. The region of the membrane comprising the ~ 150 kDa molecular weight was excised, in order to eliminate interferences of the residual antibodies present in the purified solution. The membrane was blocked in blocking solution for 1 h. The primary antibody was diluted 1:300 (GTX74405) or 1:75 (A1549) in new blocking solution, and after washing three times for 5 min with PBST (PBS 0.1 % Tween 20) the membrane was incubated with secondary antibody for 1 h (1:7500). Detection was performed using ECL™ Advance chemiluminescent reagent.

Surface plasmon resonance. Surface plasmon resonance (SPR) experiments were performed using a BIAcore™ 3000 (GE Healthcare,

Barcelona, Spain) at 25.0 ºC, employing CM5 sensor chips. The immobilisation of the A36 antibody to the surface channel was done by standard amine coupling as indicated by the manufacturer. A reference channel was generated following the same procedure without antibody injection in order to account for non-specific binding events. Both 22 kDa GH protein and secondary antibody were diluted in HBS-EP buffer, and employed at the concentrations and volumes indicated below. Sensorgrams were recorded at a flow rate of 5 µl/min. Surface regeneration was achieved with a 5 µl pulse of a glycine 10 mM solution adjusted to a pH of 2.0 with formic acid for the next round of experiments.

Immulite. Measurements of GH concentrations were performed using a solid-phase, two-site chemiluminescent immuno-metric assay from Diagnostic Products Corporation (Immulite-1000, Siemens-DPC, LA, USA).

Results

Western blot analysis of blood samples.

Pre-exercise (basal) samples of five volunteers, and post-exercise samples of four were quantified by Immulite (the method detects the 22 kDa and 20 kDa GH variants). Included in table 5, post-exercise concentrations showed a notable increase with respect to basal values, likely induced by the design of the running routine, of sufficient duration, high environmental temperature, and maximum effort (Ftaiti *et al.*, 2008; Jorgensen *et al.*, 2003; Wheldon *et al.*, 2006).

Table 5. GH values obtained through Immulite measurements corresponding to five volunteers.

Volunteers	Pre-exercise ng/ml	Post-exercise ng/ml
A	0.14	16.10
B	1.91	23.45
C	0.11	47.34
D	1.85	59.90
E	2.08	--

Purifications employing the A1549 antibody were performed on post-exercise samples of volunteers A, C and D, and on basal sample of volunteer E, and evaluated by WB utilising the A1549 antibody. Figure 22 shows the final WB image. Required volumes were determined by the sensitivity of the assay, consisting in 2.4 ml for post-exercise samples and in 9.6 ml for the basal sample.

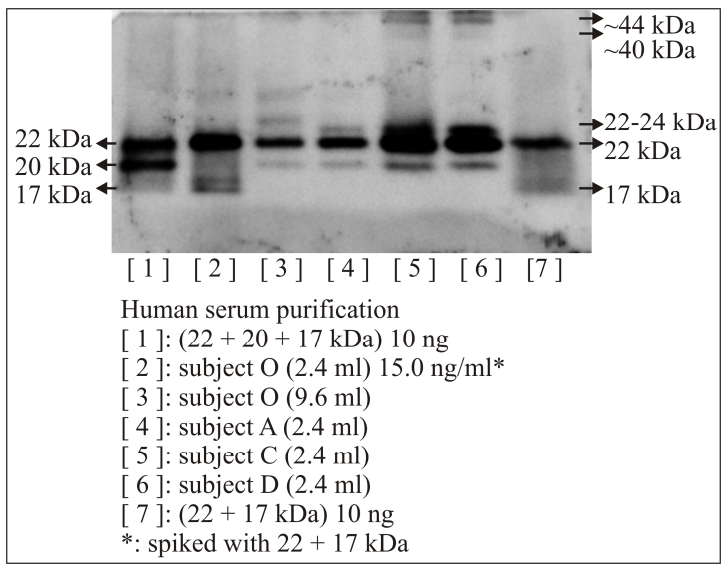


Figure 22. Western Blot (A1549 antibody) of purified serum samples. Lane 1,7: GH standards. Lane 2: 2.4 ml sample of volunteer E, spiked with 22 kDa and 17 kDa GH standards. Lane 3: 9.6 ml sample of volunteer E. Lane 4-6: 2.4 ml samples of volunteers A, C and D post-exercise. Further details are included in

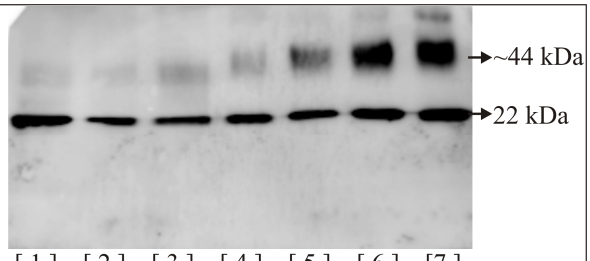
the left panel.

The purification of 15 ng/ml of the 22 kDa and 17 kDa protein, spiked in 2.4 ml of human serum corresponding to the basal state of the volunteer E (lane 2), resulted in the observation of two bands, attributed to the 22 kDa and 17 kDa proteins, in coincidence with the molecular weights of the standards (lanes 1 and 7). The minor intensity of the 17 kDa band in the purified sample (lane 2) was reasoned by the lower affinity of the A1549 antibody towards the 17 kDa fragment (chapter 4., dissociation constants for the 22 kDa, 20 kDa and 17 kDa proteins of 0.46, 0.41 and 0.69 nM, respectively). Differences between the intensities could also be observed in the standard samples (lanes 1 and 7). In this case, proteins were diluted in HBS-EP buffer and submitted to the same pre-concentration and reconstitution steps. The high aggregation tendency shown for the 17 kDa protein, in consonance with previous reports (Sinha *et* Jacobsen, 1994), together with the different proteic content in comparison with the purifications solutions could explain the reduced intensity observed for the band.

Purifications of serum from volunteer E and from volunteers A, C and D (corresponding to post-exercise samples) resulted in the observation of two bands, attributed to the 22 kDa and 20 kDa GH variants from the observation of their molecular weights (lanes 3-6). The ratio between the two proteins could also roughly coincide with the reported 55 % and 6 % for the monomeric 22 kDa and 20 kDa proteins, respectively, of the total GH (Baumann, 2009). It was not observed any band in the region of the 17 kDa GH fragment in any of the samples.

On the other hand, a third band was observed in all post-exercise samples located between approximately 22 and 24 kDa, at higher concentration

than the 20 kDa protein. In the (basal) volunteer E, this band could be barely observed. In contrast, two slight additional bands could be shown at slightly higher molecular weight regions, between approximately 24 and 30 kDa. The identity of these proteins is unknown. Additional bands at much higher molecular weights were observed in the four volunteers, especially for volunteers C and D. These could be tentatively be attributed to dimers of the 22 kDa and 20 kDa proteins, as the same intensity ratio between the two was apparently maintained. Presumably, this dimerisation could also be observed in additional purifications of serum from volunteers A, B, C and D, analysed by WB with a polyclonal anti-GH antibody (GTX74405). Alternatively, in the WB of figure 23, the antibody utilised was exclusive for the 22 kDa protein. In all samples (standards included), diffuse bands of compounds with higher molecular weights than the 22 kDa attributed band were shown. The high specificity of the polyclonal antibody (batch nº 21491) for the 22 kDa protein was unexpected, and contrasted with another batch (batch nº 18312) of the same antibody that had a lower and comparable specificity to the A1549 antibody.



Human serum purification
[1]: (22 + 20 + 17 kDa) 10 ng
[2]: (22 + 20 + 17 kDa) 5 ng
[3]: (22 + 20 + 17 kDa) 3 ng
[4]: subject A (2.4 ml)
[5]: subject B (2.4 ml)
[6]: subject C (2.4 ml)
[7]: subject D (2.4 ml)

Figure 23. Western Blot of the microtiter immunopurification of 2.4 ml of post-exercise serum of volunteers A, B, C and D, and GH standards, using the polyclonal antibody GTX74405 batch nº 21491.

Western blot analysis of pituitary samples.

Analysis of three different dilutions of an homogenate from three human pituitary glands was performed by the same WB procedure with the A1549 antibody. The high GH concentration (from Immulite measurements), made unnecessary any purification protocol. The final WB image is shown in figure 24.

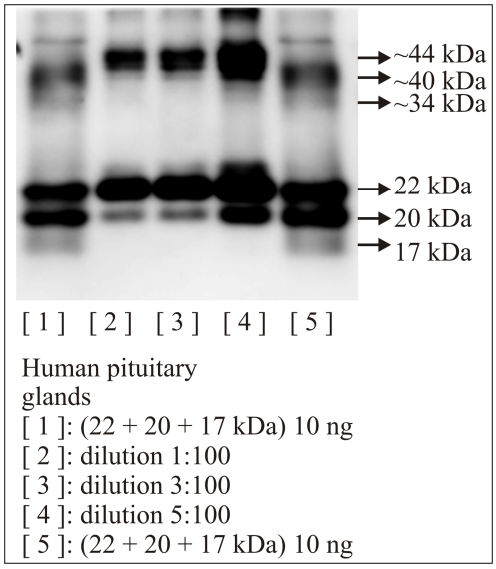


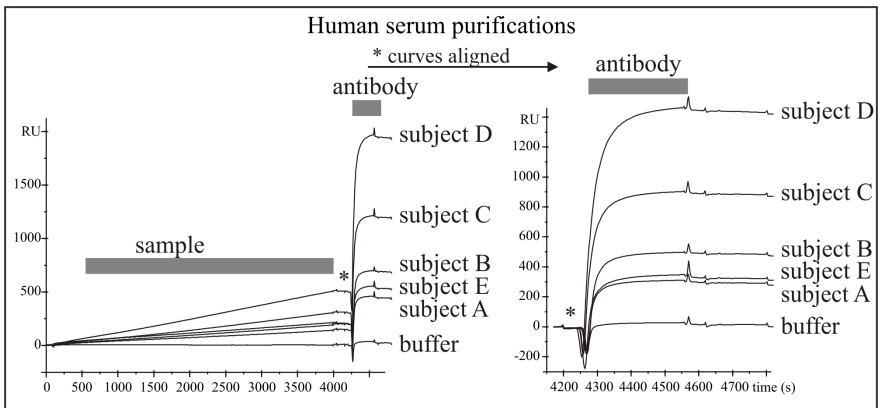
Figure 24. Western Blot of the analysis of distinct dilutions of an homogenate of three human pituitary gland.

Two bands were observed at the molecular weight corresponding to the 22 kDa and 20 kDa GH variants (lanes 2-4), in comparison with the standards (lanes 1,5). As in the case of blood samples, neither band appeared at the region of 17 kDa. Again, a third band between 22 and 24 kDa was observed at a slightly lower concentration than the 20 kDa

protein (lanes 2.4). Multiple bands could also be observed at higher molecular weights, initially attributed to dimers of the 22 kDa and 20 kDa variants.

Surface plasmon resonance analysis of blood samples.

Purifications of the blood samples from volunteers A, B, C, D and E were furthermore evaluated by SPR, following the previously reported method (see chapter 4.). In this case, the higher sensitivity of the SPR approach allowed to reduce the volume of serum to be purified. For post-exercise and basal samples, 400 µl and 3 ml of volume were required, respectively. The SPR-based method was totally specific towards the 22 kDa GH protein. Sensorgrams obtained are shown in figure 25.



sample	antigen (RU)	sandwich (RU)
subject A	140.6	434.9 (+294.3)
subject B	193.3	681.5 (+488.2)
subject C	306.2	1192.4 (+886.2)
subject D	503.7	1942.8 (+1439.1)
subject E	201.9	525.9 (+324.0)
buffer	2.8	18.7 (+15.9)

Figure 25. SPR sensorgrams of post-exercise (A, B, C, D) and basal (E) samples after purification through antibody A1549, over a surface with the immobilised antibody A36, and with the posterior signal amplification through the same antibody A1549.

In all volunteers the sensitivity of the assay was appropriate for an

accurate measurement. The concentrations for each volunteer sample, using the calibration equation of the assay (y = 1,84x+44,04; R^2 = 0.995), and taking into account the analysed volume (325 µl) and the initial purified volume, yielded values of 3.81, 6.22, 11.2, 18.52 and 0.63 ng/ml for volunteers A, B, C, D and E, that corresponded to a recovery percentage of (mean ± s.d.) 27.0 (± 3.4) % for the purification protocol, in comparison with values from Immulite measurements.

Discussion

17 kDa GH fragment

The 17 kDa GH fragment has not been positively identified neither in pituitary nor in blood samples. This has also been reported by several other authors, however the employed immunological in those assays might have displayed a null or low cross-reactivity towards this protein. This was evidenced in different immunoassays characterised by Jansson *et al.* (Jansson *et al.*, 1997): the polyclonal immunoradiometric assay (p-IRMA from Pharmacia & Upjohn), employed in the so-called growth hormone exclusion assay (GHEA) described by Boguszewski *et al.* (Boguszewski *et al.*, 1996), the commercial DELFIA® monoclonal time-resolved immunofluorometric assay (trIFMA from Wallac), and the monoclonal-based immunoradiometric assay (m-IRMA from Pharmacia & Upjohn). Comparable results were reported for a GH radioimmunoassay, characterised by Sinha *et* Jacobsen (Sinha *et* Jacobsen, 1994). In other assays, data regarding the 17 kDa fragment were not provided. This is the case for the six assays characterised by Wu *et al.* and Bidlingmaier *et al.*, two (*Pit-GH* and *rhGH*) (Wu *et al.*, 1999) and four (*recA*, *recB*, *pitA* and *pitB*) (Bidlingmaier *et al.*, 2009) sandwich-type assays. Information of cross-reactivity with the 17 kDa fragment was only

provided for one primary antibody of the pitB assay, that in fact showed a high affinity (see supplemental table of the second reference). Equally, the immunoassay provided by Kohler *et al.* was not evaluated with the 17 kDa fragment (Kohler *et al.*, 2008). Regarding the commercial GH kit based on Immulite, this does not recognise the fragment. In this manuscript, the method employed to both purify and analyse human blood samples does recognise the 17 kDa GH fragment as well as several other GH proteins. Consequently, our results are conclusive concerning the lack of null-identification of the fragment, although penalised by the limited number of samples analysed. In turn, these data are in conflict with the previous report from Sinha *et* Jacobsen (Sinha *et* Jacobsen, 1994), where the 17 kDa GH fragment was detected at concentrations even higher than the 22 kDa protein.

Other non-22 kDa GH variants in blood circulation

A clear consensus exists towards the 22 kDa protein as the most concentrated GH variant. It is stated to represent about a ~ 55 % of the total GH in blood circulation accounting only for the monomer, whereas oligomers (of all variants) represent an additional ~ 25 % (Baumann, 2009). However, a novel third variant has recently been detected showing an increased concentration during an endogenous GH peak, and being rather undetectable at basal levels (Kohler *et al.*, 2008). The molecular weight and concentration of this reported variant coincides with the novel band detected here, between 22 and 24 kDa, at elevated concentrations in post-exercise samples. Although still minor in comparison with the 22 kDa protein, it is more concentrated than the 20 kDa variant at non-basal states. The identity of this protein is still unknown. Its molecular weight is similar to the protein also detected (above) in pituitary samples, attributed to the O-glycosylated pituitary variant of the 22 kDa GH protein

previously identified (Haro *et al.*, 1996a), and recently fully characterised (Bustamante *et al.*, 2009; Kohler *et al.*, 2009). The variable proportion of this protein in blood in comparison with the 22 kDa and 20 kDa variants suggests a specific pituitary secretion for this variant, or conversely a non-pituitary origin. In this sense, glycosylation of proteins have been described in inflammatory diseases (Gornik *et* Lauc, 2008), a fact that could be related with the exercise-induced GH secretion, and the exercise-derived muscle stress and damage (Clarkson *et* Hubal, 2002). Under this hypothesis, other GH variants could equally suffer similar glycosylations, including the recombinant 22 kDa protein in doping cases. On the contrary, this 22-24 kDa serum variant has not been detected here when utilising the polyclonal antibody, that showed a total specificity towards the 22 kDa protein. Given the fact that the epitope of this antibody appears to be located at the 22 kDa segment AA_{32-43}, thus not responsive to the 20 kDa and 17 kDa proteins, this might be in conflict with the unique difference between the glycosylated and the non-glycosylated 22 kDa protein, consisting in the oligosaccharide attachment at threonine 60 described in the pituitary O-glycosylated 22 kDa variant (Bustamante *et al.*, 2009; Kohler *et al.*, 2009).

As such, the identity of this novel variant still remains to be determined. Nonetheless, it is evidenced that the GH protein distribution in blood is not constant in time. In this case, it is apparently linked to the GH release into blood, at least to an exercise-induced secretion. In agreement to this concept, other unidentified bands between 24 and 30 kDa could be also observed uniquely in the basal volunteer, even though with a very low intensity. Whether these represent additional GH variants or artefacts has not been determined, but these were not observed in any post-exercise volunteers. This notion of a variable proportion of GH proteins that

dependent on the state of the individual was also previously described. After acute exercise, the non-22 kDa group was reported to increase (8.2 ± 0.4 versus 1.0 ± 1.0 in pre-exercise state, mean ± sem, percentages of total GH) (Wallace *et al.*, 2001b; Wallace *et al.*, 2001a). When recombinant GH was administered, the non-22 kDa group was higher but did not change with exercise (3.3 ± 0.2 versus 3.1 ± 0.1) (Wallace *et al.*, 2001a). As an hypothesis, this increase could be related to the 22-24 kDa variant, although it is unknown whether this variant is recognised or not. On the contrary, the administration of GH in GH deficient patients has been shown not to result in the identification of the 22-24 kDa variant (Kohler *et al.*, 2009). In any case, the precise determination of which GH variants of variable concentration, as a function of the state, and the consequences within either the medical or the doping field needs further investigation. In this sense, an abnormal non-22 kDa high concentration was detected in short children, indicating a possible mechanism for growth failure (Boguszewski *et al.*, 1997). Furthermore, the possibility that some of these non-22 kDa variants could be generated from the recombinant 22 kDa protein after its administration cannot be discarded. In this case, the employed immunoassays would require a complete characterisation to establish which variants are detected and which are not.

Methodological considerations

In GH analyses, the presence of several variants increases the complexity of the assays, based on the specificity of antibodies when employing immunoassays. Most sensitive approaches, based on ELISA or analogous procedures provide a single numerical value corresponding to the measurement. This number is representative of all variants recognised, each one with its particular affinity. In contrast, WB measurements allow a visual differentiation of all the recognised variants. Between the two

methods, the main difference resides in the much higher sensitivity of the former. In this context, the SPR approach reported here represents an intermediate point between these two methodologies. In comparison to the WB analysis, the major sensitivity has allowed to use a ~ 5 times less sample in SPR. Additionally, the minimal part of the GH sample that is consumed allows further analyses with the remainder of the sample, if required.

6. Human growth hormone biological markers

The suspected administration of GH within the athletes community is persecuted for medical and ethical reasons since 1992 (Sonksen, 2009). Aimed at detecting abuse, the indirect (or marker) approach monitors different proteins that increase (or decrease) in concentration as a result of a prolonged administration of GH (Segura *et al.*, 2009). The group of studied proteins is large (see a brief list in chapter 1.4.), but uniquely two proteins (IGF-I and P-III-P, based on the studies of groups GH-2000 and GH-2004) have been included in the proposed method (Holt *et al.*, 2009a). Belonging to the GH/IGF-I axis (IGF-I) and bone markers (P-III-P), both are described to increase their concentration after a prolonged use of GH, reaching maximums after 2-4 and 6-8 weeks, respectively, and with a reported higher response in men (Nelson *et al.*, 2008). Conversely, slight variation of these markers have been reported under other circumstances. P-III-P concentration was shown to increase after injury, but did not lead to false positive findings (Erotokritou-Mulligan *et al.*, 2008). Also salivary IGF-I concentration was shown to increase after physical exercise, in contrast to serum IGF-I (Antonelli *et al.*, 2009). Ethnical differences could exist, at least during puberty between African and European Americans for either IGF-I and its binding proteins (Casazza *et al.*, 2008), these also being affected by dietary factors (Colangelo *et al.*, 2005). Noteworthy, these differences were stated not to invalidate the method (Erotokritou-Mulligan *et al.*, 2009a). In this regard, the combination of different markers aims at providing an absolute certainty about the results obtained. None withstanding these data, the study of novel and/or additional markers still constitutes a continuous focus of attention, within the doping field for additional evidences of GH abuse, but also within the medical field, for the study of possible secondary

effects derived from the treatment with recombinant GH.

Here, the mannan binding lectin (MBL) and C-reactive protein (CRP) have been studied as potential markers of a GH abuse. MBL is a plasma protein of the innate immune system, part of the inflammatory complement system that initiates the complement cascade and activates inflammation after binding to carbohydrate structures on microbial surfaces (Dean *et al.*, 2005). Endogenous concentrations display a genetically determined, high inter-individuals variability (Presanis *et al.*, 2003), that is supposed to be stable over time and also in relation to gender, age and physical activity (Ytting *et al.*, 2007). However, in Chinese adults, it was shown to decrease beyond 40-50 years of age (Ip *et al.*, 2004). Furthermore, MBL was reported to markedly increase in heavy smokers (Maffei *et al.*, 2005) and in response to bacterial infections (Aittoniemi *et al.*, 1997). Under GH treatment, its concentration was found to duplicate, not mediated through IGF-I. This rise in MBL concentration was hypothesised to be somehow related with the increased mortality observed during GH therapy in critically ill patients (Hansen *et al.*, 2001; Gravholt *et al.*, 2004). On the other hand, CRP is an inflammatory marker (Van Dyke *et* Kornman, 2008) that shows a rise in concentration in response to inflammatory processes, also after tissue injury or infection suggesting a contribution to host defence (Black *et al.*, 2004; Ho *et* Lipman, 2009), and as well reflecting non-inflammatory states such as obesity or chronic fatigue (Kushner, 2001). With long-term intensive regular exercise its concentration has been shown to decrease (Mattusch *et al.*, 2000; Balducci *et al.*, 2009), while after acute physical exercise, its concentration has been reported not to change (Edge *et al.*, 2009; Czarkowska-Paczek *et al.*, 2005), to show a moderate increase (Semple *et al.*, 2006), or a high increase (Scharhag *et al.*, 2005; Murakami

et al., 2009; Kim *et al.*, 2009), apparently as a function of the sports discipline and intensity. CRP concentration shows a high inter-individual and low intra-individual variability (Macy *et al.*, 1997). Under administration of GH, CRP concentration was reported to decrease (Graham *et al.*, 2007; Sesmilo *et al.*, 2000), or not to change (Albert *et al.*, 2007) when a minor dose was employed. A decrease was also observed in patients with pituitary GH-secreting adenomas suggested a protective effect of GH (Ikeda, 2009).

Material and Methods

Materials. Recombinant 22 kDa GH protein (Genotonorm®) was purchased from Pfizer Laboratories (New York, NY, USA). Commercial MBL ELISA (HK323) was purchased at Hycult Biotech (Uden, Netherlands).

Clinical trials. The studies were single blind and conducted according to the corresponding international ethics guidelines, in a randomized and controlled fashion. Revision of the protocol was performed by and permission obtained from the Ethical Committee from Hospital del Mar in Barcelona (CEIC-IMAS, Spain) and volunteers gave written informed consent. The volunteers were recruited applying a number of selection criteria that included a history of normal growth and development and normal analytical values as well as exploratory complementary information performed prior to enrolling in the trial. Recombinant 22 kDa protein was administered subcutaneously at 08.00 h in the morning, with the exception of control subjects. A limited number of administrations, once daily during 3 and 7 days respectively, with 6 IU of GH were carried out. Venous blood samples were obtained from the antecubital vein and

collected in plastic Vacutainer® tubes with gel separator. After an interval for clotting, samples were centrifuged at 4 ºC, and then the serum was aliquoted (0.5 ml) and frozen at – 20 ºC until assayed. Particularities of the two trials were:

A assay. Three (male) healthy volunteers (A1-3), of (mean ± standard deviation, s.d.) age 21.7 ± 0.6 (yr), weight 76.7 ± 3.7 (kg), and height 179.3 ± 1.6 (cm).

B assay. Nine (male) healthy volunteers (B1-9), of age 24.2 ± 2.2 (yr), weight 76.1 ± 6.1 (kg), and height 176.6 ± 6.7 (cm). Controls (B4 and B7) did not follow the administration protocol.

Immulite. Measurements of CRP levels were performed using a solid-phase, two-site chemiluminescent immunometric assay from Diagnostic Products Corporation, (Immulite-1000, Siemens-DPC, LA, USA).

Results

Individuals from both A and B assays where administered with daily doses of recombinant 22 kDa GH protein during three and seven days respectively, except for the two controls in B assay (B4 and B7). MBL analyses were performed at hours 0, 48, 56, 72 (A assay) and 0, 72, 168, 192, 312, 480 (B assay). CRP analyses were performed at hours 0, 4, 48, 52, 56, 72 (A assay) and 0, 72, 148, 168, 192, 312, 480 (B assay). The *during-assay* period comprises the interval 0-72 h (A assay) and 0-168 h (B assay), including the last 24 hours after the final GH dose. The *post-assay* period comprises the interval 192-480 h (B assay).

MBL analysis

MBL concentration increased with GH administration, although a high

inter-individual variability existed in the magnitude of this response. One volunteer (B9) showed a very low concentration (below the assay's sensitivity) and a null response towards GH, and was excluded for data analysis. For the other 9 individuals following the GH administration, 24 hours after the third dose (72 hours) the MBL concentration rose nearly 3 times (versus the initial concentration): 259 (± 183) (mean ± s.d.) %, while for the 2 controls the value was 74 (± 36) %. During these first 72 hours, the increase was shown to follow from the first dose, with values for the 3 volunteers (A assay) at 48, 56 and 72 hours of 196 (± 93), 286 (± 187) and 346 (± 306) %, respectively. A prolonged administration of GH resulted in a higher increase in the MBL concentration in all volunteers except for on case (volunteer B5). Values at 168 hours (7 GH administrations) for the 6 volunteers (B assay) were of 309 (± 169) %. Values from the B assay are depicted in figure 26 (left). Specific data of each individual is included in table 6. Data regarding the wash-out step were also monitored (B assay). MBL values were slightly decreased 2 days after the last GH dose, restored in most cases 7 days after, and in all cases 14 days after. With regard to the absolute MBL values, basal MBL concentrations of the 12 individuals were representative of the stated high inter-individual variability, of 544 (± 508) ng/ml.

CRP analysis

An effect of a GH administration on CRP could not be demonstrated convincingly. Absolute CRP concentrations in basal state showed a high inter-individual variability, of 559 (± 805) ng/ml for the 12 individuals. One volunteer (B6) showed very low concentrations (below the assay's sensitivity) and was excluded for data analysis. Noteworthy, sudden increases were observed for volunteers B5 (312 hours), B9 (148 hours), and B7 (control, 312 hours), with an irregular pattern of concentrations.

During the first 72 hours of administration (A assay), CRP showed a notable decrease of the concentration (figure 26; right-top). Apparently, the maximum decrease was already reached after 48 hours, with a percentage of 66 (± 22) %, similar to the value obtained at 72 hours, of 65 (± 26) %. Prolonged administration showed a decrease of the concentration at 72 hours less pronounced (B assay), of 79 (± 63) %, but of 55 (± 28) % if volunteer B5 was excluded (taking into account its irregular pattern). Beyond 72 hours, volunteers B1, B2, B3 and B8 continued to experiment an incremented decrease, of 45 (± 24) % at 168 hours. Values of the two control subjects (B4 and B7) were of 77 (± 37) % at 72 hours, and of 76 (± 17) % at 168 hours, *i.e.* experimenting a comparable decrease in CRP concentration even though GH was not administered. Data regarding the wash-out step was also monitored (B assay). Values were restored in most cases after 7 days, with the exception of volunteers B5, B7 and B9 that remained at high levels. Figure 26, right-bottom, depicts the mean value of volunteers B1, B2, B3 and B8, and of the two control subjects (B7 peak point excluded). Specific data for each individual is included in table 6.

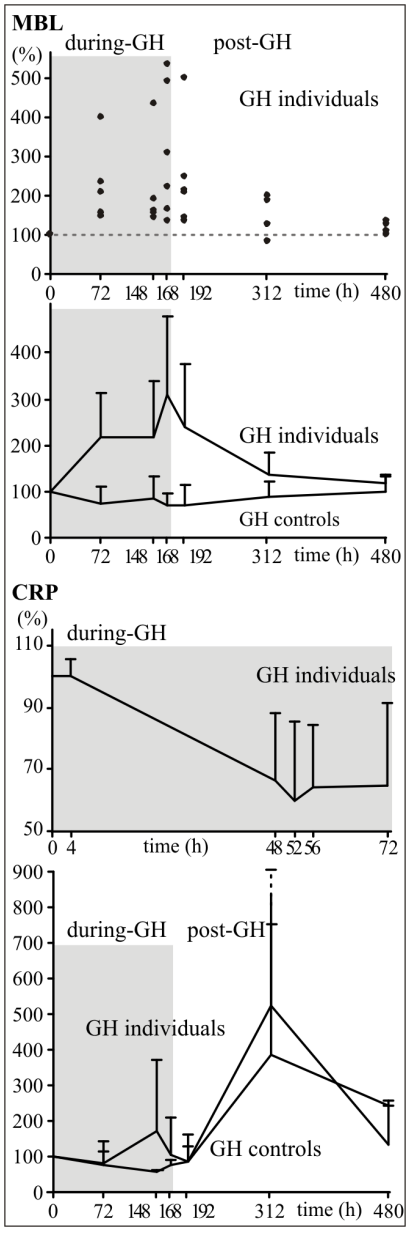


Figure 26. MBL (left) and CRP (right) values expressed as percentages of their initial value, in response of an administration protocol of recombinant GH. During-GH stage (grey area) refers up to 24 hours after the last dose. Post-GH stage refers to the wash-out period (posterior to 24 hours after the last dose). B-assay (0-480 h). A assay (0-72 h). *Left-top*: Scatter graph of the volunteers receiving GH. *Rest*: Concentrations (mean ± s.d.) of the volunteers and controls.

Table 6. MBL and CRP concentration values for each individual corresponding to the clinical assays A (hours 0-72) and B (hours 0-480). Values correspond to the median of two values.

	time (h)	A1	A2	A3	B1	B2	B3	B4	B5	B6	B7	B8	B9
MBL (ng/ml)	0	550	<4	1300	1100	600	180	1480	300	600	400	12	<4
	48	900	12	1600									
	56	1150	20	1950									
	72	1000	28	2050	4400	1.250	280	720	450	900	400	28	<4
	148				--*	1150	280	770	484	850	480	52	<4
	168				5400	1.850	400	800	400	1.000	350	64	<4
	192				2300	1.500	380	600	400	850	400	60	<4
	312				1400	750	340	1000	250	500	450	24	<4
	480				1200	650	240	1080	300	750	500	16	<4
CRP (ng/ml)	0	115	426	565	332	418	140	3050	170	<100	424	663	303
	4	122	409	553									
	48	104	208	339									
	52	100	154	322									
	56	100	204	323									
	72	109	190	314	248	174	129	1560	336	<100	436	157	129
	148				317	<100	127	1570	407	<100	253	213	1650
	168				199	<100	<100	1940	295	<100	371	173	858
	192				164	<100	111	1720	308	<100	490	101	527
	312				382	<100	166	3830	4410	<100	2730	1120	376
	480				217	138	153	7000	572	198	1070	671	490

* value omitted due to incoherent values of replicates

Discussion

The concentration of MBL had been reported to increase with the administration of recombinant GH, showing a sustained rise during 6 days of administration (Hansen *et al.*, 2001). For doses (once daily) of 15 IU (n = 9), 10.4 IU (n = 16) and 7.5 IU (n = 10), mean values (± s.d.) of 201 (± 26) %, 209 (± 32) % [approximated from the figure 1 included in (Hansen *et al.*, 2001)] and 198 (± 16) % were given, respectively. Here, employing a dose of 6 IU once daily during 3 and 7 days, values obtained at 72 hours (n = 9) and at 168 hours (n = 6) were of 259 (± 183) % and 309 (± 169) %, respectively. Although the limited number of individuals is underlined, here it has been observed a much higher inter-individual variability with

respect to the effects caused by the GH administration, with one case with a total lack of MBL (both basal or GH induced), representative of the stated percentage of individuals with a genetically impaired MBL production (Presanis *et al.*, 2003). However, the rest of individuals displayed an increase of the MBL concentration. Given the fairly low doses administered, the possibility of observing stronger increments employing higher GH doses specifically in those individuals of lower responses should not be discarded.

Similarly, the response of CRP after an administration of GH was reported to decrease in doses of 4.4 IU once daily during 6 days (Graham *et al.*, 2007), and also in long-term treatments over GHD patients (Sesmilo *et al.*, 2000), but not with doses of 0.5 IU (Albert *et al.*, 2007). Here, a higher dose (6 IU) provided distinct responses, revealing a high variability between the subjects. Of 12 cases, 1 individual had a too low CRP concentration to be evaluated, and 3 individuals (including 1 control) experimented sudden peak of concentration. Furthermore, the behaviour of the controls was comparable to the individuals of the study, with a similar decrease in CRP concentration without an identified reason. The unexpected CRP increments observed later actually occurred despite the administration of recombinant GH. This is in contrast with the protective effects reported in pituitary adenomas (Ikeda, 2009), although in that case there were higher concentrations of GH.

Concerning the applicability of the two proteins as potential candidates for markers of GH abuse, MBL has shown to be GH-responsive but with elevated differences between individuals. The highly different values of reference between individuals discard the applicability of a universal criterium, requiring a MBL record for every athlete. For the CRP protein,

the stated effect of GH has not been clearly observed in our measurements. Although in four individuals a decrease of CRP is shown, the similar behaviour observed for the two control subjects does not allow to establish any relationship. As a potential marker, the same precepts are applicable, *i.e.* the necessity of reference values of the athlete requiring thus continuous and longitudinal analyses.

In summary, our preliminary studies pinpoint MBL as a potential biomarker for the detection of GH abuse, although the response of MBL versus GH administration is heterogeneous within individuals. Its monitoring would be possible under the application of the biological passport concept. The potential contribution of MBL within the detection of GH abuse merits further investigation, with particular interest directed towards the behaviour of this protein in athletes.

7. CONCLUSIONS

7. Conclusions

The following conclusions can be extracted from the research, conducted within the framework of this thesis:

1- The production of the 5 kDa and 17 kDa GH fragments through a specific enzymatic proteolysis of the 22 kDa GH protein has been successfully performed under a physiological pH. This result supports that a similar procedure could be feasible to occur *in-vivo*.

2- The 17 kDa and 5 kDa GH fragments were obtained by either proteolysis or solid phase synthesis (5 kDa only) in sufficient amounts to allow structural approximations. Based on limited proteolysis data, the 17 kDa protein was shown to preserve the same conformation once released from the 22 kDa protein. In contrast, the 5 kDa folding state appeared to be modified.

3- Cross-reactivity studies between antibodies and individual GH variants, *i.e.* 22, 20, 17 and 5 kDa, suggests that current GH immunoassays do not recognise the same group of variants, and when so, with different affinities. The detailed characterisation of different GH immunoassays is thus strongly recommended.

4- Analyses of GH based on surface plasmon resonance technique are suitable for the evaluation and characterisation of standards and purified GH samples. Sensitivity requirements for the direct analysis of GH in biological samples are at the limit of the capabilities of the SPR instrumentation currently available.

5- Analysis of serum samples of male subjects either under basal or post-exercise states, have not resulted in a positive identification

of the 17 kDa GH fragment. This is in conflict with a previously published alleged specific assay for the 17 kDa protein where similar concentrations to the 22 kDa protein were reported. However, the finding is in keeping with the absence of such a variant claimed by other authors.

6- From the results obtained, new parameters of interest are revealed. A third GH variant with an approximately molecular weight between 22 and 24 kDa appears as the second most predominant GH form after an exercise-induced GH secretion. The potential presence of this exercise-related unidentified GH protein should be taken into account when interpreting direct GH variants in the context of anti-doping tests. On the other hand, among the indirect markers of GH abuse, the mannan-binding lectin may constitute a promising candidate for longitudinal follow up in athletes, showing a considerable response after GH administration.

8. REFERENCES

8. References

Abellan, R., Ventura, R., Pichini, S., Di Giovannandrea, R., Bellver, M., Olive, R., Pacifici, R., Pascual, J. A., Zuccaro, P., *et* Segura, J. (2006). Effect of physical fitness and endurance exercise on indirect biomarkers of recombinant growth hormone misuse: insulin-like growth factor I and procollagen type III peptide. *Int.J.Sports Med.* 27, 976-983.

Abellan, R., Ventura, R., Pichini, S., Pascual, J. A., Pacifici, R., Di Carlo, S., Bacosi, A., Segura, J., *et* Zuccaro, P. (2005). Evaluation of immunoassays for the measurement of insulin-like growth factor-I and procollagen type III peptide, indirect biomarkers of recombinant human growth hormone misuse in sport. *Clin.Chem.Lab Med.* 43, 75-85.

Abramson, F. P., Osborn, B. L., *et* Teffera, Y. (1996). Isotopic differences in human growth hormone preparations. *Anal.Chem.* 68, 1971-1972.

Ahn, C. W., Kim, C. S., Nam, J. H., Kim, H. J., Nam, J. S., Park, J. S., Kang, E. S., Cha, B. S., Lim, S. K., Kim, K. R., Lee, H. C., *et* Huh, K. B. (2006). Effects of growth hormone on insulin resistance and atherosclerotic risk factors in obese type 2 diabetic patients with poor glycaemic control. *Clin.Endocrinol.(Oxf)* 64, 444-449.

Aittoniemi, J., Rintala, E., Miettinen, A., *et* Soppi, E. (1997). Serum mannan-binding lectin (MBL) in patients with infection: clinical and laboratory correlates. *APMIS* 105, 617-622.

Alam, K. S., Morimoto, M., Yoshizato, H., Fujikawa, T., Furukawa, K., Tanaka, M., *et* Nakashima, K. (1998). Expression and purification of a mutant human growth hormone that is resistant to proteolytic cleavage by thrombin, plasmin and human plasma in vitro. *J.Biotechnol.* 65, 183-190.

Albert, S. G., Haas, M. J., *et* Mooradian, A. D. (2007). The effects of recombinant human growth hormone (rhGH) supplementation on adipokines and C-reactive protein in obese subjects. *Growth Horm.IGF.Res.* 17, 19-25.

Anderson, L. L., Jeftinija, S., *et* Scanes, C. G. (2004). Growth hormone secretion: molecular and cellular mechanisms and in vivo approaches. *Exp.Biol.Med.(Maywood.)* 229, 291-302.

Antonelli, G., Gatti, R., Prearo, M., *et* De Palo, E. F. (2009). Salivary free insulin-like growth factor-i levels: effects of an acute physical exercise in athletes. *J.Endocrinol.Invest* 32, 1-5.

Arslan, A. A., Gu, Y., Zeleniuch-Jacquotte, A., Koenig, K. L., Liu, M., Velikokhatnaya, L., Shore, R. E., Toniolo, P., Linkov, F., *et* Lokshin, A. E.

(2008). Reproducibility of serum pituitary hormones in women. *Cancer Epidemiol.Biomarkers Prev.* 17, 1880-1883.

Ayuk, J. *et* Sheppard, M. C. (2006). Growth hormone and its disorders. *Postgrad.Med.J.* 82, 24-30.

Balducci, S., Zanuso, S., Nicolucci, A., Fernando, F., Cavallo, S., Cardelli, P., Fallucca, S., Alessi, E., Letizia, C., Jimenez, A., Fallucca, F., *et* Pugliese, G. (2009). Anti-inflammatory effect of exercise training in subjects with type 2 diabetes and the metabolic syndrome is dependent on exercise modalities and independent of weight loss. *Nutr.Metab Cardiovasc.Dis. In press.*

Bam, N. B., Cleland, J. L., *et* Randolph, T. W. (1996). Molten globule intermediate of recombinant human growth hormone: stabilization with surfactants. *Biotechnol.Prog.* 12, 801-809.

Barroso, O., Mazzoni, I., *et* Rabin, O. (2008). Hormone abuse in sports: the antidoping perspective. *Asian J.Androl* 10, 391-402.

Barroso, O., Schamasch, P., *et* Rabin, O. (2009). Detection of GH abuse in sport: Past, present and future. *Growth Horm.IGF.Res.* 19, 369-374.

Barth, J. H. *et* Sibley, P. E. (2008). Standardization of the IMMULITE systems growth hormone assay with the recombinant IS 98/574. *Ann.Clin.Biochem.* 45, 598-600.

Bassett, E. E. *et* Erotokritou-Mulligan, I. (2009). Statistical issues in implementing the marker method. *Growth Horm.IGF.Res.* 19, 361-365.

Baumann, G. (1991). Growth hormone heterogeneity: genes, isohormones, variants, and binding proteins. *Endocr.Rev.* 12, 424-449.

Baumann, G. (1999). Growth hormone heterogeneity in human pituitary and plasma. *Horm.Res.* 51 Suppl 1, 2-6.

Baumann, G. (2002). Growth hormone binding protein. The soluble growth hormone receptor. *Minerva Endocrinol.* 27, 265-276.

Baumann, G. *et* Shaw, M. A. (1990). A second, lower affinity growth hormone-binding protein in human plasma. *J.Clin.Endocrinol.Metab* 70, 680-686.

Baumann, G., Stolar, M. W., *et* Buchanan, T. A. (1985). Slow metabolic clearance rate of the 20,000-dalton variant of human growth hormone: implications for biological activity. *Endocrinology* 117, 1309-1313.

Baumann, G. P. (2009). Growth hormone isoforms. *Growth Horm.IGF.Res.* 19, 333-340.

Beauloye, V., Willems, B., de, C., V, Frank, S. J., Edery, M., *et* Thissen, J. P. (2002). Impairment of liver GH receptor signaling by fasting. *Endocrinology* 143, 792-800.

Beavis, R. C. *et* Chait, B. T. (1990). Rapid, sensitive analysis of protein mixtures by mass spectrometry. *Proc.Natl.Acad.Sci.U.S.A* 87, 6873-6877.

Behncken, S. N. *et* Waters, M. J. (1999). Molecular recognition events involved in the activation of the growth hormone receptor by growth hormone. *J.Mol.Recognit.* 12, 355-362.

Bennion, B. J. *et* Daggett, V. (2003). The molecular basis for the chemical denaturation of proteins by urea. *Proc.Natl.Acad.Sci.U.S.A* 100, 5142-5147.

Bernhard Rupp (1996). JFIT. *In* http://www.ruppweb.org/cd/cdtutorial.htm.

Bidlingmaier, M. (2008). Problems with GH assays and strategies toward standardization. *Eur.J.Endocrinol.* 159 Suppl 1, S41-S44.

Bidlingmaier, M. *et* Freda, P. U. (2009). Measurement of human growth hormone by immunoassays: Current status, unsolved problems and clinical consequences. *Growth Horm.IGF.Res. In press.*

Bidlingmaier, M. *et* Strasburger, C. J. (2007a). Growth hormone assays: current methodologies and their limitations. *Pituitary.* 10, 115-119.

Bidlingmaier, M. *et* Strasburger, C. J. (2007b). What endocrinologists should know about growth hormone measurements. *Endocrinol.Metab Clin.North Am.* 36, 101-108.

Bidlingmaier, M., Suhr, J., Ernst, A., Wu, Z., Keller, A., Strasburger, C. J., *et* Bergmann, A. (2009). High-Sensitivity Chemiluminescence Immunoassays for Detection of Growth Hormone Doping in Sports. *Clin.Chem.* 55, 445-453.

Bidlingmaier, M., Wu, Z., *et* Strasburger, C. J. (2000). Test method: GH. *Baillieres Best.Pract.Res.Clin.Endocrinol.Metab* 14, 99-109.

Black, S., Kushner, I., *et* Samols, D. (2004). C-reactive Protein. *J.Biol.Chem.* 279, 48487-48490.

Boguszewski, C. L., Hynsjo, L., Johannsson, G., Bengtsson, B. A., *et* Carlsson, L. M. (1996). 22-kD growth hormone exclusion assay: a new approach to measurement of non-22-kD growth hormone isoforms in human blood. *Eur.J.Endocrinol.* 135, 573-582.

Boguszewski, C. L., Jansson, C., Boguszewski, M. C., Rosberg, S., Carlsson, B., Albertsson-Wikland, K., *et* Carlsson, L. M. (1997). Increased proportion of

circulating non-22-kilodalton growth hormone isoforms in short children: a possible mechanism for growth failure. *J.Clin.Endocrinol.Metab* 82, 2944-2949.

Boguszewski, C. L., Svensson, P. A., Jansson, T., Clark, R., Carlsson, L. M., *et* Carlsson, B. (1998). Cloning of two novel growth hormone transcripts expressed in human placenta. *J.Clin.Endocrinol.Metab* 83, 2878-2885.

Bollerslev, J., Ueland, T., Jorgensen, A. P., Fougner, K. J., Wergeland, R., Schreiner, T., *et* Burman, P. (2006). Positive effects of a physiological dose of GH on markers of atherogenesis: a placebo-controlled study in patients with adult-onset GH deficiency. *Eur.J.Endocrinol.* 154, 537-543.

Bonroy, K., Frederix, F., Reekmans, G., Dewolf, E., De Palma, R., Borghs, G., Declerck, P., *et* Goddeeris, B. (2006). Comparison of random and oriented immobilisation of antibody fragments on mixed self-assembled monolayers. *J.Immunol.Methods* 312, 167-181.

Boushaba, R., Kumpalume, P., *et* Slater, N. K. (2003). Kinetics of whole serum and prepurified IgG digestion by pepsin for F(ab')2 manufacture. *Biotechnol.Prog.* 19, 1176-1182.

Boyle, P. J. *et* Freeman, J. S. (2007). Application of incretin mimetics and dipeptidyl peptidase IV inhibitors in managing type 2 diabetes mellitus. *J.Am.Osteopath.Assoc.* 107 Suppl, S10-S16.

Bredehoft, M., Schanzer, W., *et* Thevis, M. (2008). Quantification of human insulin-like growth factor-1 and qualitative detection of its analogues in plasma using liquid chromatography/electrospray ionisation tandem mass spectrometry. *Rapid Commun.Mass Spectrom.* 22, 477-485.

Brown, R. J., Adams, J. J., Pelekanos, R. A., Wan, Y., McKinstry, W. J., Palethorpe, K., Seeber, R. M., Monks, T. A., Eidne, K. A., Parker, M. W., *et* Waters, M. J. (2005). Model for growth hormone receptor activation based on subunit rotation within a receptor dimer. *Nat.Struct.Mol.Biol.* 12, 814-821.

Buck, M. (1998). Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins. *Q.Rev.Biophys.* 31, 297-355.

Bustamante, J. J., Gonzalez, L., Carroll, C. A., Weintraub, S. T., Aguilar, R. M., Munoz, J., Martinez, A. O., *et* Haro, L. S. (2009). O-Glycosylated 24 kDa human growth hormone has a mucin-like biantennary disialylated tetrasaccharide attached at Thr-60. *Proteomics.* 9, 3474-3488.

Butler, J. E. (2004). Solid supports in enzyme-linked immunosorbent assay and other solid-phase immunoassays. *Methods Mol.Med.* 94, 333-372.

Casazza, K., Higgins, P. B., Fernandez, J. R., Goran, M. I., *et* Gower, B. A. (2008). Longitudinal analysis of the insulin-like growth factor system in African-American and European American children and adolescents. *J.Clin.Endocrinol.Metab* 93, 4917-4923.

Chanson, P. *et* Salenave, S. (2008). Acromegaly. *Orphanet.J.Rare.Dis.* 3, 17.

Chantalat, L., Jones, N., Korber, F., Navaza, J., *et* Pavlovsky, A. G. (1995). The Crystal Structure of Wild-Type Growth Hormone at 2.5 A Resolution. *Protein and Peptide Letters* 2, 333-340.

Chapman, G. E., Rogers, K. M., Brittain, T., Bradshaw, R. A., Bates, O. J., Turner, C., Cary, P. D., *et* Crane-Robinson, C. (1981). The 20,000 molecular weight variant of human growth hormone. Preparation and some physical and chemical properties. *J.Biol.Chem.* 256, 2395-2401.

Chung, L., Clifford, D., Buckley, M., *et* Baxter, R. C. (2006). Novel biomarkers of human growth hormone action from serum proteomic profiling using protein chip mass spectrometry. *J.Clin.Endocrinol.Metab* 91, 671-677.

Clarkson, P. M. *et* Hubal, M. J. (2002). Exercise-induced muscle damage in humans. *Am.J.Phys.Med.Rehabil.* 81, S52-S69.

Colangelo, L. A., Chiu, B. C., Liu, K., Kopp, P. A., Gann, P. H., *et* Gapstur, S. M. (2005). IGF-1, IGFBP-3, and nutritional factors in young black and white men: the CARDIA Male Hormone Study. *Nutr.Cancer* 53, 57-64.

Cowan, D. A. *et* Bartlett, C. (2009). Laboratory issues in the implementation of the marker method. *Growth Horm.IGF.Res.* 19, 357-360.

Cummings, D. E. *et* Merriam, G. R. (2003). Growth hormone therapy in adults. *Annu.Rev.Med.* 54, 513-533.

Czarkowska-Paczek, B., Bartlomiejczyk, I., Gabrys, T., Przybylski, J., Nowak, M., *et* Paczek, L. (2005). Lack of relationship between interleukin-6 and CRP levels in healthy male athletes. *Immunol.Lett.* 99, 136-140.

Dall, R., Longobardi, S., Ehrnborg, C., Keay, N., Rosen, T., Jorgensen, J. O., Cuneo, R. C., Boroujerdi, M. A., Cittadini, A., Napoli, R., Christiansen, J. S., Bengtsson, B. A., Sacca, L., Baxter, R. C., Basset, E. E., *et* Sonksen, P. H. (2000). The effect of four weeks of supraphysiological growth hormone administration on the insulin-like growth factor axis in women and men. GH-2000 Study Group. *J.Clin.Endocrinol.Metab* 85, 4193-4200.

Dannies, P. S. (2001). A serum prolactin-binding protein: implications for growth hormone. *Trends Endocrinol.Metab* 12, 427-428.

Dannies, P. S. (2002). Mechanisms for storage of prolactin and growth hormone in secretory granules. *Mol.Genet.Metab* 76, 6-13.

de Vos, A. M., Ultsch, M., *et* Kossiakoff, A. A. (1992). Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science* 255, 306-312.

Dean, M. M., Minchinton, R. M., Heatley, S., *et* Eisen, D. P. (2005). Mannose binding lectin acute phase activity in patients with severe infection. *J.Clin.Immunol.* 25, 346-352.

DeFelippis, M. R., Kilcomons, M. A., Lents, M. P., Youngman, K. M., *et* Havel, H. A. (1995). Acid stabilization of human growth hormone equilibrium folding intermediates. *Biochim.Biophys.Acta* 1247, 35-45.

Dienys, G., Sereikaite, J., Gavenas, G., Kvederas, R., *et* Bumelis, V. A. (1998). Cross-linking of protein subunits by 1,3, 5-triacryloyl-hexahydro-s-triazine. *Bioconjug.Chem.* 9, 744-748.

Dienys, G., Sereikaite, J., Luksa, V., Jarutiene, O., Mistiniene, E., *et* Bumelis, V. A. (2000). Dimerization of human growth hormone in the presence of metal ions. *Bioconjug.Chem.* 11, 646-651.

Doessing, S. *et* Kjaer, M. (2005). Growth hormone and connective tissue in exercise. *Scand.J.Med.Sci.Sports* 15, 202-210.

Donangelo, I. *et* Melmed, S. (2005). Treatment of acromegaly: future. *Endocrine.* 28, 123-128.

Duda, K. M. *et* Brooks, C. L. (2003). Identification of residues outside the two binding sites that are critical for activation of the lactogenic activity of human growth hormone. *J.Biol.Chem.* 278, 22734-22739.

Edge, J., Mundel, T., Weir, K., *et* Cochrane, D. J. (2009). The effects of acute whole body vibration as a recovery modality following high-intensity interval training in well-trained, middle-aged runners. *Eur.J.Appl.Physiol* 105, 421-428.

Ehrnborg, C. *et* Rosen, T. (2008). Physiological and pharmacological basis for the ergogenic effects of growth hormone in elite sports. *Asian J.Androl* 10, 373-383.

Erotokritou-Mulligan, I., Bassett, E. E., Bartlett, C., Cowan, D., McHugh, C., Seah, R., Curtis, B., Wells, V., Harrison, K., Sonksen, P. H., *et* Holt, R. I. (2008). The effect of sports injury on insulin-like growth factor-I and type 3 procollagen: implications for detection of growth hormone abuse in athletes. *J.Clin.Endocrinol.Metab* 93, 2760-2763.

Erotokritou-Mulligan, I., Bassett, E. E., Cowan, D. A., Bartlett, C., McHugh, C., Sonksen, P. H., *et* Holt, R. I. (2009a). Influence of ethnicity on IGF-I and procollagen III peptide (P-III-P) in elite athletes and its effect on the ability to detect GH abuse. *Clin.Endocrinol.(Oxf)* 70, 161-168.

Erotokritou-Mulligan, I., Bassett, E. E., Kniess, A., Sonksen, P. H., *et* Holt, R. I. (2007). Validation of the growth hormone (GH)-dependent marker method of detecting GH abuse in sport through the use of independent data sets. *Growth Horm.IGF.Res.* 17, 416-423.

Erotokritou-Mulligan, I., Eryl, B. E., Cowan, D., Bartlett, C., Milward, P., Sartorio, A., Sonksen, P. H., *et* Holt, R. I. (2009b). The use of growth hormone (GH)-dependent markers in the detection of GH abuse in sport: Physiological intra-individual variation of IGF-I, type 3 pro-collagen (P-III-P) and the GH-2000 detection score. *Clin.Endocrinol.(Oxf). In press.*

Fain, J. N., Cheema, P., Tichansky, D. S., *et* Madan, A. K. (2008). Stimulation of human omental adipose tissue lipolysis by growth hormone plus dexamethasone. *Mol.Cell Endocrinol.* 295, 101-105.

Fields, G. B. *et* Noble, R. L. (1990). Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *Int.J.Pept.Protein Res.* 35, 161-214.

Fisker, S. (2006). Physiology and pathophysiology of growth hormone-binding protein: methodological and clinical aspects. *Growth Horm.IGF.Res.* 16, 1-28.

Fontana, A., de Laureto, P. P., Spolaore, B., Frare, E., Picotti, P., *et* Zambonin, M. (2004). Probing protein structure by limited proteolysis. *Acta Biochim.Pol.* 51, 299-321.

Forsyth, I. A. *et* Wallis, M. (2002). Growth hormone and prolactin--molecular and functional evolution. *J.Mammary.Gland.Biol.Neoplasia.* 7, 291-312.

Frasier, S. D. (1997). The not-so-good old days: working with pituitary growth hormone in North America, 1956 to 1985. *J.Pediatr.* 131, S1-S4.

Frigeri, L. G., Teguh, C., Ling, N., Wolff, G. L., *et* Lewis, U. J. (1988). Increased sensitivity of adipose tissue to insulin after in vivo treatment of yellow Avy/A obese mice with amino-terminal peptides of human growth hormone. *Endocrinology* 122, 2940-2945.

Frystyk, J., Andreasen, C. M., *et* Fisker, S. (2008). Determination of free growth hormone. *J.Clin.Endocrinol.Metab* 93, 3008-3014.

Ftaiti, F., Jemni, M., Kacem, A., Zaouali, M. A., Tabka, Z., Zbidi, A., *et* Grelot, L. (2008). Effect of hyperthermia and physical activity on circulating growth hormone. *Appl.Physiol Nutr.Metab* 33, 880-887.

Fuglsang, J. *et* Ovesen, P. (2006). Aspects of placental growth hormone physiology. *Growth Horm.IGF.Res.* 16, 67-85.

Gahete, M. D., Duran-Prado, M., Luque, R. M., Martinez-Fuentes, A. J., Quintero, A., Gutiérrez-Pascual, E., Cordoba-Chacon, J., Malagon, M. M., Gracia-Navarro, F., *et* Castano, J. P. (2009). Understanding the multifactorial control of growth hormone release by somatotropes: lessons from comparative endocrinology. *Ann.N.Y.Acad.Sci.* 1163, 137-153.

Giavoli, C., Cappiello, V., Corbetta, S., Ronchi, C. L., Morpurgo, P. S., Ferrante, E., Beck-Peccoz, P., *et* Spada, A. (2004). Different effects of short- and long-term recombinant hGH administration on ghrelin and adiponectin levels in GH-deficient adults. *Clin.Endocrinol.(Oxf)* 61, 81-87.

Gobom, J., Nordhoff, E., Mirgorodskaya, E., Ekman, R., *et* Roepstorff, P. (1999). Sample purification and preparation technique based on nano-scale reversed-phase columns for the sensitive analysis of complex peptide mixtures by matrix-assisted laser desorption/ionization mass spectrometry. *J.Mass Spectrom.* 34, 105-116.

Goeddel, D. V., Heyneker, H. L., Hozumi, T., Arentzen, R., Itakura, K., Yansura, D. G., Ross, M. J., Miozzari, G., Crea, R., *et* Seeburg, P. H. (1979). Direct expression in Escherichia coli of a DNA sequence coding for human growth hormone. *Nature* 281, 544-548.

Goldenberg, N. *et* Barkan, A. (2007). Factors regulating growth hormone secretion in humans. *Endocrinol.Metab Clin.North Am.* 36, 37-55.

Gorg, A., Obermaier, C., Boguth, G., Harder, A., Scheibe, B., Wildgruber, R., *et* Weiss, W. (2000). The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 21, 1037-1053.

Gornik, O. *et* Lauc, G. (2008). Glycosylation of serum proteins in inflammatory diseases. *Dis.Markers* 25, 267-278.

Graf, L., Li, C. H., *et* Jibson, M. D. (1992). Human growth hormone molecule with no loss of the complex. *Science* 255, 306-312.

Graham, M. R., Baker, J. S., Evans, P., Hullin, D., Thomas, N. E., *et* Davies, B. (2009). Potential benefits of recombinant human growth hormone (rhGH) to athletes. *Growth Horm.IGF.Res.* 19, 300-307.

Graham, M. R., Baker, J. S., Evans, P., Kicman, A., Cowan, D., Hullin, D., *et* Davies, B. (2007). Evidence for a decrease in cardiovascular risk factors following recombinant growth hormone administration in abstinent anabolic-androgenic steroid users. *Growth Horm.IGF.Res.* 17, 201-209.

Gravholt, C. H., Leth-Larsen, R., Lauridsen, A. L., Thiel, S., Hansen, T. K., Holmskov, U., Naeraa, R. W., *et* Christiansen, J. S. (2004). The effects of GH and hormone replacement therapy on serum concentrations of mannan-binding lectin, surfactant protein D and vitamin D binding protein in Turner syndrome. *Eur.J.Endocrinol.* 150, 355-362.

Grigorian, A. L., Bustamante, J. J., Hernandez, P., Martinez, A. O., *et* Haro, L. S. (2005). Extraordinarily stable disulfide-linked homodimer of human growth hormone. *Protein Sci.* 14, 902-913.

Gutiérrez-Gallego, R., Bosch, J., Such-Sanmartín, G., *et* Segura, J. (2009). Surface plasmon resonance immuno assays - A perspective. *Growth Horm.IGF.Res.* 19, 388-398.

Hahnefeld, C., Drewianka, S., *et* Herberg, F. W. (2004). Determination of kinetic data using surface plasmon resonance biosensors. *Methods Mol.Med.* 94, 299-320.

Hansen, T. K., Thiel, S., Dall, R., Rosenfalck, A. M., Trainer, P., Flyvbjerg, A., Jorgensen, J. O., *et* Christiansen, J. S. (2001). GH strongly affects serum concentrations of mannan-binding lectin: evidence for a new IGF-I independent immunomodulatory effect of GH. *J.Clin.Endocrinol.Metab* 86, 5383-5388.

Haro, L. S., Lewis, U. J., Garcia, M., Bustamante, J., Martinez, A. O., *et* Ling, N. C. (1996a). Glycosylated human growth hormone (hGH): a novel 24 kDa hGH-N variant. *Biochem.Biophys.Res.Commun.* 228, 549-556.

Haro, L. S., Singh, R. N., Lewis, U. J., Martinez, A. O., Galosy, S. S., Staten, N. R., *et* Krivi, G. G. (1996b). Human growth hormone deletion mutant (hGH44-191) binds with high affinity to lactogenic receptors but not to somatogenic receptors. *Biochem.Biophys.Res.Commun.* 222, 421-426.

Hashimoto, Y., Ikeda, I., Ikeda, M., Takahashi, Y., Hosaka, M., Uchida, H., Kono, N., Fukui, H., Makino, T., *et* Honjo, M. (1998). Construction of a specific and sensitive sandwich enzyme immunoassay for 20 kDa human growth hormone. *J.Immunol.Methods* 221, 77-85.

Hepner, F., Csaszar, E., Roitinger, E., Pollak, A., *et* Lubec, G. (2006). Massspectrometrical analysis of recombinant human growth hormone Norditropin reveals amino acid exchange at M14_V14 rhGH. *Proteomics.* 6, 775-784.

Hepner, F., Cszasar, E., Roitinger, E., *et* Lubec, G. (2005). Mass spectrometrical analysis of recombinant human growth hormone (Genotropin(R)) reveals amino acid substitutions in 2% of the expressed protein. *Proteome.Sci.* 3, 1.

Hettiarachchi, M., Watkinson, A., Leung, K. C., Sinha, Y. N., Ho, K. K., *et* Kraegen, E. W. (1997). Human growth hormone fragment (hGH44-91) produces insulin resistance and hyperinsulinemia but is less potent than 22 kDa hGH in the rat. *Endocrine.* 6, 47-52.

Heutmekers, T. H., Bremer, M. G., Haasnoot, W., *et* Nielen, M. W. (2007). A rapid surface plasmon resonance (SPR) biosensor immunoassay for screening of somatotropins in injection preparations. *Anal.Chim.Acta* 586, 239-245.

Ho, K. M. *et* Lipman, J. (2009). An update on C-reactive protein for intensivists. *Anaesth.Intensive Care* 37, 234-241.

Holt, R. I., Bassett, E. E., Erotokritou-Mulligan, I., McHugh, C., Cowan, D., Bartlett, C., *et* Sonksen, P. H. (2009a). Moving one step closer to catching the GH cheats: The GH-2004 experience. *Growth Horm.IGF.Res.* 19, 346-351.

Holt, R. I., Erotokritou-Mulligan, I., *et* Sonksen, P. H. (2009b). The history of doping and growth hormone abuse in sport. *Growth Horm.IGF.Res.* 19, 320-326.

Hubbard, S. J. (1998). The structural aspects of limited proteolysis of native proteins. *Biochim.Biophys.Acta* 1382, 191-206.

Hubbard, S. J., Beynon, R. J., *et* Thornton, J. M. (1998). Assessment of conformational parameters as predictors of limited proteolytic sites in native protein structures. *Protein Eng* 11, 349-359.

Ikeda, H. (2009). Specificity of systemic inflammatory response syndrome during the peri-operative period in patients with GH-secreting adenoma. *Cytokine* 46, 92-95.

Inouye, K. *et* Ohnaka, S. (2001). Pepsin digestion of a mouse monoclonal antibody of IgG1 class formed F(ab')(2) fragments in which the light chains as well as the heavy chains were truncated. *J.Biochem.Biophys.Methods* 48, 23-32.

Ip, W. K., To, Y. F., Cheng, S. K., *et* Lau, Y. L. (2004). Serum mannose-binding lectin levels and mbl2 gene polymorphisms in different age and gender groups of southern Chinese adults. *Scand.J.Immunol.* 59, 310-314.

Jansson, C., Boguszewski, C., Rosberg, S., Carlsson, L., *et* Albertsson-Wikland, K. (1997). Growth hormone (GH) assays: influence of standard preparations, GH isoforms, assay characteristics, and GH-binding protein. *Clin.Chem.* 43, 950-956.

Jason-Moller, L., Murphy, M., *et* Bruno, J. (2006). Overview of Biacore systems and their applications. *Curr.Protoc.Protein Sci.* Chapter 19, 13.

Jiang, H., Wu, S. L., Karger, B. L., *et* Hancock, W. S. (2009). Mass spectrometric analysis of innovator, counterfeit, and follow-on recombinant human growth hormone. *Biotechnol.Prog.* 25, 207-218.

Jorgensen, J. O., Krag, M., Kanaley, J., Moller, J., Hansen, T. K., Moller, N., Christiansen, J. S., *et* ORskov, H. (2003). Exercise, hormones, and body temperature. regulation and action of GH during exercise. *J.Endocrinol.Invest* 26, 838-842.

Kaiser, E., Colescott, R. L., Bossinger, C. D., *et* Cook, P. I. (1970). Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal.Biochem.* 34, 595-598.

Kasimova, M. R., Kristensen, S. M., Howe, P. W., Christensen, T., Matthiesen, F., Petersen, J., Sorensen, H. H., *et* Led, J. J. (2002). NMR studies of the backbone flexibility and structure of human growth hormone: a comparison of high and low pH conformations. *J.Mol.Biol.* 318, 679-695.

Katznelson, L. (2009). Alterations in body composition in acromegaly. *Pituitary.* 12, 136-142.

Kay, R. G., Barton, C., Velloso, C. P., Brown, P. R., Bartlett, C., Blazevich, A. J., Godfrey, R. J., Goldspink, G., Rees, R., Ball, G. R., Cowan, D. A., Harridge, S. D., Roberts, J., Teale, P., *et* Creaser, C. S. (2009). High-throughput ultra-high-performance liquid chromatography/tandem mass spectrometry quantitation of insulin-like growth factor-I and leucine-rich alpha-2-glycoprotein in serum as biomarkers of recombinant human growth hormone administration. *Rapid Commun.Mass Spectrom.* 23, 3173-3182.

Kennett, R. G. (1979). Monoclonal antibodies. Hybrid myelomas--a revolution in serology and immunogenetics. *Am.J.Hum.Genet.* 31, 539-547.

Kim, H. J., Lee, Y. H., *et* Kim, C. K. (2009). Changes in serum cartilage oligomeric matrix protein (COMP), plasma CPK and plasma hs-CRP in relation to running distance in a marathon (42.195 km) and an ultra-marathon (200 km) race. *Eur.J.Appl.Physiol* 105, 765-770.

Kirchner, H., Gutiérrez, J. A., Solenberg, P. J., Pfluger, P. T., Czyzyk, T. A., Willency, J. A., Schurmann, A., Joost, H. G., Jandacek, R. J., Hale, J. E., Heiman, M. L., *et* Tschop, M. H. (2009). GOAT links dietary lipids with the endocrine control of energy balance. *Nat.Med.* 15, 741-745.

Kline, J. B. *et* Clevenger, C. V. (2001). Identification and characterization of the prolactin-binding protein in human serum and milk. *J.Biol.Chem.* 276, 24760-24766.

Kniess, A., Ziegler, E., Kratzsch, J., Thieme, D., *et* Muller, R. K. (2003). Potential parameters for the detection of hGH doping. *Anal.Bioanal.Chem.* 376, 696-700.

Kohler, M., Puschel, K., Sakharov, D., Tonevitskiy, A., Schanzer, W., *et* Thevis, M. (2008). Detection of recombinant growth hormone in human plasma by a 2-D PAGE method. *Electrophoresis* 29, 4495-4502.

Kohler, M., Thomas, A., Puschel, K., Schanzer, W., *et* Thevis, M. (2009). Identification of Human Pituitary Growth Hormone Variants by Mass Spectrometry. *J.Proteome.Res.* 8, 1071-1076.

Kojima, M., Hosoda, H., Date, Y., Nakazato, M., Matsuo, H., *et* Kangawa, K. (1999). Ghrelin is a growth-hormone-releasing acylated peptide from stomach. *Nature* 402, 656-660.

Komatsu, N., Saijoh, K., Otsuki, N., Kishi, T., Micheal, I. P., Obiezu, C. V., Borgono, C. A., Takehara, K., Jayakumar, A., Wu, H. K., Clayman, G. L., *et* Diamandis, E. P. (2007). Proteolytic processing of human growth hormone by multiple tissue kallikreins and regulation by the serine protease inhibitor Kazal-Type5 (SPINK5) protein. *Clin.Chim.Acta* 377, 228-236.

Kushner, I. (2001). C-reactive protein elevation can be caused by conditions other than inflammation and may reflect biologic aging. *Cleve.Clin.J.Med.* 68, 535-537.

Lacroix, M. C., Guibourdenche, J., Frendo, J. L., Muller, F., *et* Evain-Brion, D. (2002). Human placental growth hormone--a review. *Placenta* 23 Suppl A, S87-S94.

Laursen, T. (2004). Clinical pharmacological aspects of growth hormone administration. *Growth Horm.IGF.Res.* 14, 16-44.

Le Roith, D., Scavo, L., *et* Butler, A. (2001). What is the role of circulating IGF-I? *Trends Endocrinol.Metab* 12, 48-52.

Lecomte, C. M., Renard, A., *et* Martial, J. A. (1987). A new natural hGH variant--17.5 kd--produced by alternative splicing. An additional consensus sequence which might play a role in branchpoint selection. *Nucleic Acids Res.* 15, 6331-6348.

Leung, K. C., Howe, C., Gui, L. Y., Trout, G., Veldhuis, J. D., *et* Ho, K. K. (2002). Physiological and pharmacological regulation of 20-kDa growth hormone. *Am.J.Physiol Endocrinol.Metab* 283, E836-E843.

Lewis, U. J., Lewis, L. J., Salem, M. A., Staten, N. R., Galosy, S. S., *et* Krivi, G. G. (1991). A recombinant-DNA-derived modification of human growth hormone

(hGH44-191) with enhanced diabetogenic activity. *Mol.Cell Endocrinol.* 78, 45-54.

Lewis, U. J., Sinha, Y. N., *et* Lewis, G. P. (2000). Structure and properties of members of the hGH family: a review. *Endocr.J.* 47 Suppl, S1-S8.

Lopez-Guajardo, C. C., Armstrong, L. S., Jordan, L., Staten, N. R., Krivi, G. G., Martinez, A. O., *et* Haro, L. S. (1998). Generation, characterization and utilization of anti-human growth hormone 1-43, (hGH1-43), monoclonal antibodies in an ELISA. *J.Immunol.Methods* 215, 179-185.

Macy, E. M., Hayes, T. E., *et* Tracy, R. P. (1997). Variability in the measurement of C-reactive protein in healthy subjects: implications for reference intervals and epidemiological applications. *Clin.Chem.* 43, 52-58.

Maffei, G., Brouwer, N., Dolman, K. M., van, d., V, Roos, D., *et* Loos, B. G. (2005). Plasma levels of mannan-binding lectin in relation to periodontitis and smoking. *J.Periodontol.* 76, 1881-1889.

Mattusch, F., Dufaux, B., Heine, O., Mertens, I., *et* Rost, R. (2000). Reduction of the plasma concentration of C-reactive protein following nine months of endurance training. *Int.J.Sports Med.* 21, 21-24.

Mauras, N. *et* Haymond, M. W. (2005). Are the metabolic effects of GH and IGF-I separable? *Growth Horm.IGF.Res.* 15, 19-27.

McHugh, C. M., Park, R. T., Sonksen, P. H., *et* Holt, R. I. (2005). Challenges in detecting the abuse of growth hormone in sport. *Clin.Chem.* 51, 1587-1593.

Mellado, M., Rodriguez-Frade, J. M., Kremer, L., *et* Martinez-Alonso, C. (1996). Characterization of monoclonal antibodies specific for the human growth hormone 22K and 20K isoforms. *J.Clin.Endocrinol.Metab* 81, 1613-1618.

Melmed, S. (1999). Confusion in clinical laboratory GH and IGF-I reports. *Pituitary.* 2, 171-172.

Moller, N. *et* Jorgensen, J. O. (2009). Effects of growth hormone on glucose, lipid, and protein metabolism in human subjects. *Endocr.Rev.* 30, 152-177.

Mondon, C. E., Reaven, G. M., Ling, N., Lewis, U. J., *et* Frigeri, L. G. (1988). Amino-terminal peptide of growth hormone enhances insulin action in normal rats. *Endocrinology* 123, 827-833.

Murakami, S., Kurihara, S., Koikawa, N., Nakamura, A., Aoki, K., Yosigi, H., Sawaki, K., *et* Ohtani, M. (2009). Effects of oral supplementation with cystine and theanine on the immune function of athletes in endurance exercise:

randomized, double-blind, placebo-controlled trial. *Biosci.Biotechnol.Biochem.* 73, 817-821.

Nam, S. Y. *et* Lobie, P. E. (2000). The mechanism of effect of growth hormone on preadipocyte and adipocyte function. *Obes.Rev.* 1, 73-86.

Nelson, A. E. *et* Ho, K. K. (2009). Demographic factors influencing the GH system: Implications for the detection of GH doping in sport. *Growth Horm.IGF.Res.* 19, 327-332.

Nelson, A. E., Meinhardt, U., Hansen, J. L., Walker, I. H., Stone, G., Howe, C. J., Leung, K. C., Seibel, M. J., Baxter, R. C., Handelsman, D. J., Kazlauskas, R., *et* Ho, K. K. (2008). Pharmacodynamics of growth hormone abuse biomarkers and the influence of gender and testosterone: a randomized double-blind placebo-controlled study in young recreational athletes. *J.Clin.Endocrinol.Metab* 93, 2213-2222.

Nilsson, A. G., Svensson, J., *et* Johannsson, G. (2007). Management of growth hormone deficiency in adults. *Growth Horm.IGF.Res.* 17, 441-462.

Nindl, B. C., Hymer, W. C., Deaver, D. R., *et* Kraemer, W. J. (2001). Growth hormone pulsatility profile characteristics following acute heavy resistance exercise. *J.Appl.Physiol* 91, 163-172.

Norrelund, H. (2005). The metabolic role of growth hormone in humans with particular reference to fasting. *Growth Horm.IGF.Res.* 15, 95-122.

Palmetshofer, A., Zechner, D., Luger, T. A., *et* Barta, A. (1995). Splicing variants of the human growth hormone mRNA: detection in pituitary, mononuclear cells and dermal fibroblasts. *Mol.Cell Endocrinol.* 113, 225-234.

Peterson, F. C. *et* Brooks, C. L. (1997). Identification of a motif associated with the lactogenic actions of human growth hormone. *J.Biol.Chem.* 272, 21444-21448.

Peterson, F. C. *et* Brooks, C. L. (2004). Different elements of mini-helix 1 are required for human growth hormone or prolactin action via the prolactin receptor. *Protein Eng Des Sel* 17, 417-424.

Polverino, d. L., Toma, S., Tonon, G., *et* Fontana, A. (1995). Probing the structure of human growth hormone by limited proteolysis. *Int.J.Pept.Protein Res.* 45, 200-208.

Popii, V. *et* Baumann, G. (2004). Laboratory measurement of growth hormone. *Clin.Chim.Acta* 350, 1-16.

Powrie, J. K., Bassett, E. E., Rosen, T., Jorgensen, J. O., Napoli, R., Sacca, L., Christiansen, J. S., Bengtsson, B. A., *et* Sonksen, P. H. (2007). Detection of growth hormone abuse in sport. *Growth Horm.IGF.Res.* 17, 220-226.

Presanis, J. S., Kojima, M., *et* Sim, R. B. (2003). Biochemistry and genetics of mannan-binding lectin (MBL). *Biochem.Soc.Trans.* 31, 748-752.

Rakover, Y., Lavi, I., Masalah, R., Issam, T., Weiner, E., *et* Ben Shlomo, I. (2000). Comparison between four immunoassays for growth hormone (GH) measurement as guides to clinical decisions following GH provocative tests. *J.Pediatr.Endocrinol.Metab* 13, 637-643.

Rich, R. L. *et* Myszka, D. G. (2008). Survey of the year 2007 commercial optical biosensor literature. *J.Mol.Recognit.* 21, 355-400.

Roelen, C. A., Koppeschaar, H. P., de Vries, W. R., Snel, Y. E., Doerga, M. E., Zelissen, P. M., Thijssen, J. H., *et* Blankenstein, M. A. (1997). Visceral adipose tissue is associated with circulating high affinity growth hormone-binding protein. *J.Clin.Endocrinol.Metab* 82, 760-764.

Rowlinson, S. W., Waters, M. J., Lewis, U. J., *et* Barnard, R. (1996). Human growth hormone fragments 1-43 and 44-191: in vitro somatogenic activity and receptor binding characteristics in human and nonprimate systems. *Endocrinology* 137, 90-95.

Sakharov, D. A., Thevis, M., *et* Tonevitsky, A. G. (2008). Analysis of major isoforms of human growth hormone before and after intensive physical exercise. *Bull.Exp.Biol.Med.* 146, 466-469.

Salem, M. A. (1988). Effects of the amino-terminal portion of human growth hormone on glucose clearance and metabolism in normal, diabetic, hypophysectomized, and diabetic-hypophysectomized rats. *Endocrinology* 123, 1565-1576.

Salem, M. A., Lewis, U. J., Haro, L. S., Kishi, K., McAllister, D. L., Seavey, B. K., Bee, G., *et* Wolff, G. L. (1989). Effects of hypophysectomy and the insulin-like and anti-insulin pituitary peptides on carbohydrate metabolism in yellow Avy/A (BALB/c x VY)F1 hybrid mice. *Proc.Soc.Exp.Biol.Med.* 191, 408-419.

Salem, M. A. *et* Wolff, G. L. (1989). Potentiation of response to insulin and anti-insulin action by two human pituitary peptides in lean agouti A/a, obese yellow Avy/A, and C57BL/6J-ob/ob mice. *Proc.Soc.Exp.Biol.Med.* 191, 113-123.

Sanders, E. J. *et* Harvey, S. (2008). Peptide hormones as developmental growth and differentiation factors. *Dev.Dyn.* 237, 1537-1552.

Saugy,M., Cardis,C., Rivier,L., Brissom,G., Ayotte,C., Hemmersbach,P., Haug,E., *et* Segura,J. Urinary Human Growth Hormone (U-HGH): II. Measurements in Competition Tests: More the Expression of the Renal Function than the Endocrinological Status. M.Donike, H. Geyer A. Gotzmann U. Mareck-Engelke eds. 2, 223-233. 1995. Recent advances in doping analysis. Sport und Buch Strauß, Köln.

Scharhag, J., Meyer, T., Gabriel, H. H., Schlick, B., Faude, O., *et* Kindermann, W. (2005). Does prolonged cycling of moderate intensity affect immune cell function? *Br.J.Sports Med.* 39, 171-177.

Segura, J., Gutiérrez-Gallego, R., Ventura, R., Pascual, J. A., Bosch, J., Such-Sanmartín, G., Nikolovski, Z., Pinyot, A., *et* Pichini, S. (2009). Growth hormone in sport: beyond Beijing 2008. *Ther.Drug Monit.* 31, 3-13.

Semple, S. J., Smith, L. L., McKune, A. J., Hoyos, J., Mokgethwa, B., San Juan, A. F., Lucia, A., *et* Wadee, A. A. (2006). Serum concentrations of C reactive protein, alpha1 antitrypsin, and complement (C3, C4, C1 esterase inhibitor) before and during the Vuelta a Espana. *Br.J.Sports Med.* 40, 124-127.

Sesmilo, G., Biller, B. M., Llevadot, J., Hayden, D., Hanson, G., Rifai, N., *et* Klibanski, A. (2000). Effects of growth hormone administration on inflammatory and other cardiovascular risk markers in men with growth hormone deficiency. A randomized, controlled clinical trial. *Ann.Intern.Med.* 133, 111-122.

Shimizu, T., Murakami, Y., Nishiki, M., Koshimura, K., Oki, Y., *et* Kato, Y. (2004). Contents and release of human growth hormone1-43 (hGH1-43) in human pituitary glands and hGH-producing pituitary adenoma cells. *Biomedical Research* 25, 141-146.

Singh, R. N. P., Seavey, B. K., Lewis, L. J., *et* Lewis, U. J. (1983). Human growth hormone peptide 1-43: isolation from pituitary glands. *J.Protein Chem* 2, 425-436.

Sinha, Y. N. *et* Jacobsen, B. P. (1994). Human growth hormone (hGH)-(44-191), a reportedly diabetogenic fragment of hGH, circulates in human blood: measurement by radioimmunoassay. *J.Clin.Endocrinol.Metab* 78, 1411-1418.

Smith, R. G. (2009). From GH to Billy Ghrelin. *Cell Metab* 10, 82-83.

Solomon, G., Reicher, S., Gussakovsky, E. E., Jomain, J. B., *et* Gertler, A. (2006). Large-scale preparation and in vitro characterization of biologically active human placental (20 and 22K) and pituitary (20K) growth hormones: placental growth hormones have no lactogenic activity in humans. *Growth Horm.IGF.Res.* 16, 297-307.

Sonksen, P. (2009). The International Olympic Committee (IOC) and GH-2000. *Growth Horm.IGF.Res.* 19, 341-345.

Spolaore, B., Polverino, d. L., Zambonin, M., *et* Fontana, A. (2004). Limited proteolysis of human growth hormone at low pH: isolation, characterization, and complementation of the two biologically relevant fragments 1-44 and 45-191. *Biochemistry* 43, 6576-6586.

Stokes, K. (2003). Growth hormone responses to sub-maximal and sprint exercise. *Growth Horm.IGF.Res.* 13, 225-238.

Strasburger, C. J. (1998). Laboratory assessment of GH. *Growth Horm.IGF.Res.* 8 Suppl A, 41-46.

Strasburger, C. J. (2004). Taking one step at a time. *Clin.Endocrinol.(Oxf)* 60, 540.

Strasburger, C. J. *et* Bidlingmaier, M. (2005). How robust are laboratory measures of growth hormone status? *Horm.Res.* 64 Suppl 2, 1-5.

Tran, G. T., Pagkalos, J., Tsiridis, E., Narvani, A. A., Heliotis, M., Mantalaris, A., *et* Tsiridis, E. (2009). Growth hormone: does it have a therapeutic role in fracture healing? *Expert.Opin.Investig.Drugs* 18, 887-911.

Trevino, J., Calle, A., Rodriguez-Frade, J. M., Mellado, M., *et* Lechuga, L. M. (2009). Surface plasmon resonance immunoassay analysis of pituitary hormones in urine and serum samples. *Clin.Chim.Acta* 403, 56-62.

Tsunekawa, B., Wada, M., Ikeda, M., Banba, S., Kamachi, H., Tanaka, E., *et* Honjo, M. (2000). The binding between the stem regions of human growth hormone (GH) receptor compensates for the weaker site 1 binding of 20-kDa human GH (hGH) than that of 22-kDa hGH. *J.Biol.Chem.* 275, 15652-15656.

Tsushima, T., Katoh, Y., Miyachi, Y., Chihara, K., Teramoto, A., Irie, M., *et* Hashimoto, Y. (1999). Serum concentration of 20K human growth hormone (20K hGH) measured by a specific enzyme-linked immunosorbent assay. Study Group of 20K hGH. *J.Clin.Endocrinol.Metab* 84, 317-322.

Ubertini, G., Grossi, A., Colabianchi, D., Fiori, R., Brufani, C., Bizzarri, C., Giannone, G., Rigamonti, A. E., Sartorio, A., Muller, E. E., *et* Cappa, M. (2008). Young elite athletes of different sport disciplines present with an increase in pulsatile secretion of growth hormone compared with non-elite athletes and sedentary subjects. *J.Endocrinol.Invest* 31, 138-145.

Ultsch, M. H., Somers, W., Kossiakoff, A. A., *et* de Vos, A. M. (1994). The crystal structure of affinity-matured human growth hormone at 2 A resolution. *J.Mol.Biol.* 236, 286-299.

Van Dyke, T. E. *et* Kornman, K. S. (2008). Inflammation and factors that may regulate inflammatory response. *J.Periodontol.* 79, 1503-1507.

Veldhuis, J. D. *et* Bowers, C. Y. (2003). Human GH pulsatility: an ensemble property regulated by age and gender. *J.Endocrinol.Invest* 26, 799-813.

Wallace, J. D., Cuneo, R. C., Bidlingmaier, M., Lundberg, P. A., Carlsson, L., Boguszewski, C. L., Hay, J., Boroujerdi, M., Cittadini, A., Dall, R., Rosen, T., *et* Strasburger, C. J. (2001a). Changes in non-22-kilodalton (kDa) isoforms of growth hormone (GH) after administration of 22-kDa recombinant human GH in trained adult males. *J.Clin.Endocrinol.Metab* 86, 1731-1737.

Wallace, J. D., Cuneo, R. C., Bidlingmaier, M., Lundberg, P. A., Carlsson, L., Boguszewski, C. L., Hay, J., Healy, M. L., Napoli, R., Dall, R., Rosen, T., *et* Strasburger, C. J. (2001b). The response of molecular isoforms of growth hormone to acute exercise in trained adult males. *J.Clin.Endocrinol.Metab* 86, 200-206.

Warner, M. D., Sinha, Y. N., *et* Peabody, C. A. (1993). Growth hormone and prolactin variants in normal subjects. Relative proportions in morning and afternoon samples. *Horm.Metab Res.* 25, 425-429.

Wheldon, A., Savine, R. L., Sonksen, P. H., *et* Holt, R. I. (2006). Exercising in the cold inhibits growth hormone secretion by reducing the rise in core body temperature. *Growth Horm.IGF.Res.* 16, 125-131.

Wieringa, G. E., Barth, J. H., *et* Trainer, P. J. (2004). Growth hormone assay standardization: a biased view? *Clin.Endocrinol.(Oxf)* 60, 538-539.

Wisniewski, E. S., Rees, D. K., *et* Chege, E. W. (2009). Proteolytic-based method for the identification of human growth hormone. *J.Forensic Sci.* 54, 122-127.

Wood, P. (2001). Growth hormone: its measurement and the need for assay harmonization. *Ann.Clin.Biochem.* 38, 471-482.

Wroblewski, V. J., Kaiser, R. E., *et* Becker, G. W. (1993). Proteolysis of human growth hormone by rat thyroid gland in vitro: application of electrospray mass spectrometry and N-terminal sequencing to elucidate a metabolic pathway. *Pharm.Res.* 10, 1106-1114.

Wu, M., Wang, A., Bernard, G. C., Hall, J. B., Beal, W. E., Michael, A. R., Boisclair, Y. R., *et* Jiang, H. (2008). Increased degradation of insulin-like growth factor-I in serum from feed-deprived steers. *Domest.Anim Endocrinol.* 35, 343-351.

Wu, S. L., Amato, H., Biringer, R., Choudhary, G., Shieh, P., *et* Hancock, W. S. (2002). Targeted proteomics of low-level proteins in human plasma by LC/MSn: using human growth hormone as a model system. *J.Proteome.Res.* 1, 459-465.

Wu, Z., Bidlingmaier, M., Dall, R., *et* Strasburger, C. J. (1999). Detection of doping with human growth hormone. *Lancet* 353, 895.

Yang, J. T., Wu, C. S., *et* Martinez, H. M. (1986). Calculation of protein conformation from circular dichroism. *Methods Enzymol.* 130, 208-269.

Yang, T. H., Cleland, J. L., Lam, X., Meyer, J. D., Jones, L. S., Randolph, T. W., Manning, M. C., *et* Carpenter, J. F. (2000). Effect of zinc binding and precipitation on structures of recombinant human growth hormone and nerve growth factor. *J.Pharm.Sci.* 89, 1480-1485.

Ytting, H., Christensen, I. J., Thiel, S., Jensenius, J. C., Svendsen, M. N., Nielsen, L., Lottenburger, T., *et* Nielsen, H. J. (2007). Biological variation in circulating levels of mannan-binding lectin (MBL) and MBL-associated serine protease-2 and the influence of age, gender and physical exercise. *Scand.J.Immunol.* 66, 458-464.

Yuen, K., Ong, K., Husbands, S., Chatelain, P., Fryklund, L., Gluckman, P., Ranke, M., Cook, D., Rosenfeld, R., Wass, J., *et* Dunger, D. (2002). The effects of short-term administration of two low doses versus the standard GH replacement dose on insulin sensitivity and fasting glucose levels in young healthy adults. *J.Clin.Endocrinol.Metab* 87, 1989-1995.

Zhan, X., Giorgianni, F., *et* Desiderio, D. M. (2005). Proteomics analysis of growth hormone isoforms in the human pituitary. *Proteomics.* 5, 1228-1241.

Zhang, Y., Jiang, J., Black, R. A., Baumann, G., *et* Frank, S. J. (2000). Tumor necrosis factor-alpha converting enzyme (TACE) is a growth hormone binding protein (GHBP) sheddase: the metalloprotease TACE/ADAM-17 is critical for (PMA-induced) GH receptor proteolysis and GHBP generation. *Endocrinology* 141, 4342-4348.

Zhu, L., Tamvakopoulos, C., Xie, D., Dragovic, J., Shen, X., Fenyk-Melody, J. E., Schmidt, K., Bagchi, A., Griffin, P. R., Thornberry, N. A., *et* Sinha, R. R. (2003). The role of dipeptidyl peptidase IV in the cleavage of glucagon family peptides: in vivo metabolism of pituitary adenylate cyclase activating polypeptide-(1-38). *J.Biol.Chem.* 278, 22418-22423.

9. APPENDIX

Such-Sanmartín G, Bosch J, Segura J, Wu M, Du H, Chen G, Wang S, Vila-Perelló M, Andreu D, Gutiérrez-Gallego R.
Characterisation of the 5 kDa growth hormone isoform.
Growth Factors. 2008 Jun;26(3):152-62.

Such-Sanmartín G, Bosch J, Segura J, Gutiérrez-Gallego R.
Generation of 5 and 17 kDa human growth hormone fragments through limited proteolysis.
Growth Factors. 2009 Oct;27(5):255-64.

Segura J, Gutiérrez-Gallego R, Ventura R, Pascual JA, Bosch J, Such-Sanmartín G, Nikolovski Z, Pinyot A, Pichini S.
Growth hormone in sport: beyond Beijing 2008.
Ther Drug Monit. 2009 Feb;31(1):3-13.

Gutiérrez-Gallego R, Bosch J, Such-Sanmartín G, Segura J.
Surface Plasmon resonance immuno assays - A perspective.
Growth Horm IGF Res. 2009Aug;19(4):388-98.

9.2. Appendix: script “Perl automated fragmentation”

In limited proteolysis experiments, the resulting fragments correspond to those released with the first cleavages of the enzyme. For their mass identification and the attribution of a protein sequence, the use of scripts or programs is common. However, in all of the options evaluated during this research (fundamentally those provided by the expasy.org domain) these were exclusively oriented to proteolysis generating a maximum cleavage of the protein, and thus predicting fragments that could not be further cleaved. This situation is inversed in limited proteolysis experiments, where the released fragments derive from a minimum number of cleavages. In this context, the manual calculation of the theoretical fragments is feasible when the number of possible combinations is reduced. However, in sequences that contain disulfide bonds, the number of combinations can be strongly increased due to the multiple possibilities, considering that many fragments, once released, will still be linked through the disulfide bonds. This is the case of the thermolysinolysis of the 22 kDa protein. Shown in figure A.1. and organised in ranges of 1 kDa, for a 22 kDa sequence without considering the disulfide bonds it is shown a moderate number of possible combinations (< 400). In contrast, for the native sequence with the two disulfide bonds, this number increases drastically (up to 50000).

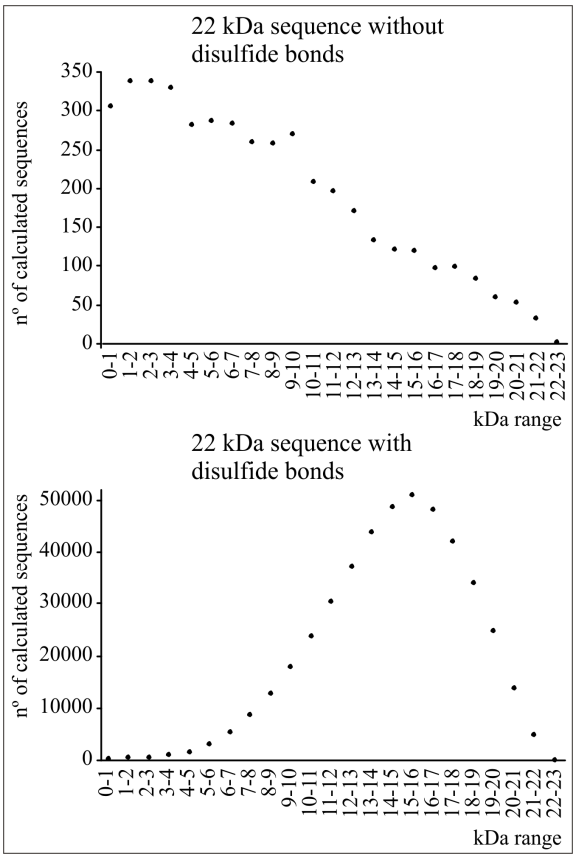


Figure A.1. Graph showing the number of theoretical fragments released after thermolysinolysis of the 22 kDa GH protein, for a particular range of masses indicated in intervals of 1 kDa. Data obtained from the number of sequences generated through the PAF script.

In order to calculate the theoretical fragmentation of a protein under limited proteolysis conditions, a specific script [(P)erl (A)utomated (F)ragmentation – PAF] has been developed to calculate all the possible sequences with the corresponding masses, for a given protein and a particular enzyme. The PAF script has been developed using the script language Practical Extraction and Report Language (Perl v5.10.0), following an imperative, structured, object oriented paradigm (*i.e.* a fundamental style of computer programming based on the use of

"objects", data structures consisting in datafields and methods together). The fundamental instructions sequence is schematised in the logical flow diagram shown in figure A.2. Below, the perl code itself (self-commented) is reproduced.

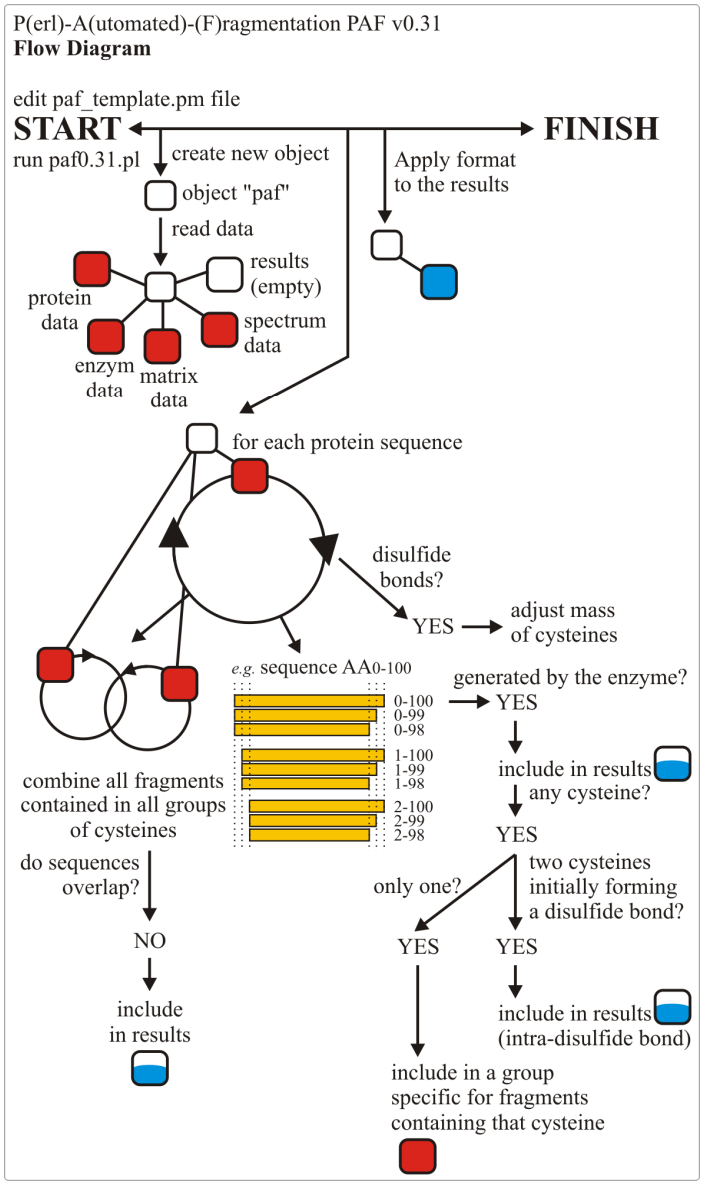


Figure A.2. Logic flow diagram of the PAF (v 3.1) Perl script.

Initially the user is required to manually edit the template file adding the information of the proteins that will be processed (sequence, disulfide bonds, …), of the enzyme responsible of the proteolysis (amino acids cleaved, preferred side of the bond, …), of the matrix employed, and of the final spectrum (isotopic or average, …). Once the script is executed, this information is stored in the program, depicted in the diagram flow with the red boxes symbolising containers of information, linked to the main object. Subsequently, for each protein sequence introduced, all possible fragments are calculated, storing only those compatible with the enzyme's specificity. From all these fragments, those that contain a cysteine are included in separate groups, except when the two cysteines that already form a disulfide bond are present. Then, all fragments contained in these groups are combined between them, discarding those sequences that overlap themselves, and stored. Finally, all these data is sorted, formatted and printed to a file.

Perl code - paf0.31.pl

```
#!c:/Perl/bin/perl
# P(erl)-A(utomated)-F(ragmentation) PAF v0.31, 2009.09.08

package PAF;
$VERSION = 0.31;
# ------------------------------------------------------
use strict;
use PAF_TEMPLATE 0.31;
my $paf;
$$ paf = undef;
bless $ paf, "PAF";
my $debug = 0;
$| = 1;
# ------------------------------------------------------
$ paf-> new();



# ------------------------------------------------------
sub new {
  my ($sec,$min,$hour) = localtime(time);
  my $init = $sec+$min*60+$hour*60*60;
  my $file = 'debug.txt' if ($debug);
  open DEBUG, ">>$file" if ($debug);



  print '---------  ---------  ---------  ---------'."\n";
  print '---------  ---------  ---------  ---------'."\n";
  print '---------  ---------  ---------  ---------'."\n";
  print '---------  ---------  ---------  ---------'."\n";
  print '---------  ---------  ---------  ---------'."\n";
  print '                      '."\n";
  print 'PAF v0.31 starts  '."\n";
  print '                      '."\n";
  print '---------  ---------  ---------  ---------'."\n";
  print '---------  ---------  ---------  ---------'."\n";
  print '---------  ---------  ---------  ---------'."\n";
  print '---------  ---------  ---------  ---------'."\n";
  print '---------  ---------  ---------  ---------'."\n";



  my $sp = {}; # data source
  $sp->{protein}= []; # proteines data source
  $sp->{enzym}= []; # enzym data source
  $sp->{matrix}= []; # matrix data source
  $sp->{spectrum}= {}; # spectrum data source
  my @exProteins; # expasy proteins
  my $rs = {}; # responses
  PAF_TEMPLATE -> new($sp,\@exProteins); # read info
  $ paf -> get_aminoAcids($sp); # get aminoacids array


  # if sequences are in swissprot format, we download them.
  if ($sp->{spectrum}->{expasy}) {
    use Bio::Perl; # must be installed
    my $temp;
    foreach my $r(@exProteins) {
      $temp = get_sequence('swissprot', $r);
      $sp->{protein}->[@{$sp->{protein}}]->{name}= $r;
      $sp->{protein}->[$#{$sp->{protein}}]->{sequence}= $temp->seq();
      print "\n".'- Sequence of '.$r.' downloaded';
      }
    }


  # start analysis of each protein sequence introduced
  LOOP: foreach (@{$sp->{protein}}) {
    if ($_->{name}) {
      print "\n".$_->{name};
      $ paf -> get_aminoAcids2($sp,\$_->{dislufide}); # set mass of cysteines
      print "\n".'building linear fragments';
      $ paf -> analysis($_,$sp,$rs); # cleaves the sequence and treat cysteines
      print "\n".'building fragments linked through disulfide bonds';
      $paf -> reconstruct($_,$sp,$rs) if ($_->{dislufide}); # join fragments with cysteine
      print "\n".'sorting and writting results';
      $paf -> results($_,$sp,$rs); # construct the results
```

```
    }else {
      last LOOP;
      }
    }

  ($sec,$min,$hour) = localtime(time);
  $init = $sec+$min*60+$hour*60*60 - $init;
  print "\n".'results sorted and written, PAF has employed '.$init.' seconds, <press enter to end>';
  <>;
  close DEBUG if ($debug);
  }


sub get_aminoAcids {
  my (undef,$sp) = @_;
  if ($sp->{spectrum}->{analysis}) {# average
    $sp->{aa}->{Hy}= 1.00797;
    $sp->{aa}->{H2O}= 18.015345;
    $sp->{aa}->{A}= 71.0788;
    $sp->{aa}->{R}= 156.1876;
    $sp->{aa}->{N}= 114.1039;
    $sp->{aa}->{D}= 115.0886;
    $sp->{aa}->{E}= 129.1155;
    $sp->{aa}->{Q}= 128.1308;
    $sp->{aa}->{G}= 57.0520;
    $sp->{aa}->{H}= 137.1412;
    $sp->{aa}->{I}= 113.1595;
    $sp->{aa}->{L}= 113.1595;
    $sp->{aa}->{K}= 128.1742;
    $sp->{aa}->{M}= 131.1986;
    $sp->{aa}->{F}= 147.1766;
    $sp->{aa}->{P}= 97.1167;
    $sp->{aa}->{S}= 87.0782;
    $sp->{aa}->{T}= 101.1051;
    $sp->{aa}->{W}= 186.2133;
    $sp->{aa}->{Y}= 163.1760;
    $sp->{aa}->{V}= 99.1326;
  }else {# isotopic
    $sp->{aa}->{Hy}= 1.007825;
    $sp->{aa}->{H2O}= 18,010566;
    $sp->{aa}->{Sod}= 22.989;
    $sp->{aa}->{Pot}= 38.9630;
    $sp->{aa}->{A}= 71.03711;
    $sp->{aa}->{R}= 156.10111;
    $sp->{aa}->{N}= 114.04293;
    $sp->{aa}->{D}= 115.02694;
    $sp->{aa}->{E}= 129.04259;
    $sp->{aa}->{Q}= 128.05858;
    $sp->{aa}->{G}= 57.02146;
    $sp->{aa}->{H}= 137.05891;
    $sp->{aa}->{I}= 113.08406;
    $sp->{aa}->{L}= 113.08406;
    $sp->{aa}->{K}= 128.09496;
    $sp->{aa}->{M}= 131.04049;
    $sp->{aa}->{F}= 147.06841;
    $sp->{aa}->{P}= 97.05276;
    $sp->{aa}->{S}= 87.03203;
    $sp->{aa}->{T}= 101.04768;
    $sp->{aa}->{W}= 186.07931;
    $sp->{aa}->{Y}= 163.06333;
    $sp->{aa}->{V}= 99.06841;
    }
  $sp->{Naa}->{A}= 'Alanine';
  $sp->{Naa}->{R}= 'Arginine';
  $sp->{Naa}->{N}= 'Asparagine';
  $sp->{Naa}->{D}= 'Aspartic Acid';
  $sp->{Naa}->{C}= 'Cysteine';
  $sp->{Naa}->{E}= 'Glutamic Acid';
  $sp->{Naa}->{Q}= 'Glutamine';
  $sp->{Naa}->{G}= 'Glycine';
  $sp->{Naa}->{H}= 'Histidine';
  $sp->{Naa}->{I}= 'Isoleucine';
  $sp->{Naa}->{L}= 'Leucine';
  $sp->{Naa}->{K}= 'Lysine';
  $sp->{Naa}->{M}= 'Methionine';
  $sp->{Naa}->{F}= 'Phenylalanine';
  $sp->{Naa}->{P}= 'Proline';
  $sp->{Naa}->{S}= 'Serine';
  $sp->{Naa}->{T}= 'Threonine';
  $sp->{Naa}->{W}= 'Tryptophan';
```

```
  $sp->{Naa}->{Y}= 'Tyrosine';
  $sp->{Naa}->{V}= 'Valine';

  $sp->{sNaa}->{A}= 'Ala';
  $sp->{sNaa}->{R}= 'Arg';
  $sp->{sNaa}->{N}= 'Asn';
  $sp->{sNaa}->{D}= 'Asp';
  $sp->{sNaa}->{C}= 'Cys';
  $sp->{sNaa}->{E}= 'Glu';
  $sp->{sNaa}->{Q}= 'Gln';
  $sp->{sNaa}->{G}= 'Gly';
  $sp->{sNaa}->{H}= 'His';
  $sp->{sNaa}->{I}= 'Ile';
  $sp->{sNaa}->{L}= 'Leu';
  $sp->{sNaa}->{K}= 'Lys';
  $sp->{sNaa}->{M}= 'Met';
  $sp->{sNaa}->{F}= 'Phe';
  $sp->{sNaa}->{P}= 'Pro';
  $sp->{sNaa}->{S}= 'Ser';
  $sp->{sNaa}->{T}= 'Thr';
  $sp->{sNaa}->{W}= 'Trp';
  $sp->{sNaa}->{Y}= 'Tyr';
  $sp->{sNaa}->{V}= 'Val';
 }
sub get_aminoAcids2 {# take into account the hydrogen mass with disulfide bonds
  my (undef,$sp,$disulphide) = @_;
  if ($sp->{spectrum}->{analysis}) {# average
    $sp->{aa}->{C}= 103.1448;
    $sp->{aa}->{C}-= $sp->{aa}->{Hy}if $$disulphide;
  }else {# isotopic
    $sp->{aa}->{C}= 103.00919;
    $sp->{aa}->{C}-= $sp->{aa}->{Hy}if $$disulphide;
    }
  }
sub analysis {
  my ($self,$prot,$sp,$rs) = @_;
  my ($ei,$ej); #enzym cleavage positions
  my ($ei2,$ej2); #if cleavage is in both positions
  my ($ri,$rj); #human numbers (+1)
  my $enzymLeft; #variables to known if cleavage is valid or not
  my $enzymatic;
  my $etxt = '';
  my $value;

  my $ccc = 0;
  for (my $i=0; $i<length($prot->{sequence}); $i++) {
    $ccc++;if ($ccc == 100) {print '.';$ccc = 0;}
    for (my $j=length($prot->{sequence})-1; $j>$i; $j--) {

      # now each cycle is a portion of the sequence: $i and $j, $ri and $rj in human format.
      $ri = $i+1+$prot->{sum};
      $rj = $j+1+$prot->{sum};

      # setting the exact positions of enzyme cleavage (-1=left,0=right,1=both;)
      $enzymLeft = 0;
      $enzymatic = 0;
      $etxt = '';

      if ($sp->{enzym}->[$prot->{enzym}]->{cutting}== 1) {
        $ei = $i - 1;
        $ej = $j;
        $ei2 = $i;
        $ej2 = $j + 1;
      }else {
        $ei = $i + $sp->{enzym}->[$prot->{enzym}]->{cutting};
        $ej = $j + 1 + $sp->{enzym}->[$prot->{enzym}]->{cutting};
        }

      # can the fragment be produced by the enzyme?
      # establishing if left side of evaluated fragment is enzymatically possible
      if ($i == 0) {$enzymLeft = 1;}
      else {for (my $k=0; $k<@{$sp->{enzym}->[$prot->{enzym}]->{aa}}; $k++) {
        $enzymLeft = 1 if (substr($prot->{sequence},$ei,1) eq $sp->{enzym}->[$prot->{enzym}]->{aa}->[$k]);
        $enzymLeft = 1 if ($ei2 and substr($prot->{sequence},$ei2,1) eq $sp->{enzym}->[$prot->{enzym}]->{aa}->[$k]);
        }}
      # and if left side was enzymatic, now let's see right side
      if ($enzymLeft) {
      if ($j == length($prot->{sequence})-1) {$enzymatic = 1;}
      else {for (my $l=0; $l<@{$sp->{enzym}->[$prot->{enzym}]->{aa}}; $l++) {
        $enzymatic = 1 if (substr($prot->{sequence},$ej,1) eq $sp->{enzym}->[$prot->{enzym}]->{aa}->[$l]);
```

```
      $enzymatic = 1 if ($ei2 and substr($prot->{sequence},$ej2,1) eq $sp->{enzym}->[$prot->{enzym}]->{aa}->[$l]);
      }}}
    $etxt = 'and enzymatic ' if $enzymatic;

    if ($enzymatic) {# too numbers to calculate if not

    # see if weight value is of our interest (again, too numbers to calculate if all are included)
    $value = $self->get_weight($sp,\$prot->{sequence},$i,$j);
    if (!$sp->{spectrum}->{limDown}or !$sp->{spectrum}->{limUp}
    or ($sp->{spectrum}->{limDown}and $value > $sp->{spectrum}->{limDown})
    or ($sp->{spectrum}->{limUp}and $value < $sp->{spectrum}->{limUp})
    or ($sp->{spectrum}->{limDown}and $sp->{spectrum}->{limUp}
      and $value > $sp->{spectrum}->{limDown}and $value < $sp->{spectrum}->{limUp})) {

      my $bonded = 0;
      my $numOfCys = 0; # number of cysteines NOT REDUCED in the fragment
      my $fName = $prot->{name}.'FINAL';

      # we have one fragment, we have to include in the responses array
      # but if it has cysteines, it has to be included to generate other linked fragments as well
      $rs->{$fName}->[@{$rs->{$fName}}]->{sequence}= $ri.'-'.$rj;
      $rs->{$fName}->[$#{$rs->{$fName}}]->{weight}= $value;
      $rs->{$fName}->[$#{$rs->{$fName}}]->{definition}= 'lineal, ';

      # let's see all pairs of dilufide bonds, if it has disulfides of course
      if ($prot->{dislufide}) {

        foreach my $t(0..$#{$prot->{united_cys}}) {
          # if first cysteine is inside the fragment
          if ($ri < $prot->{united_cys}->[$t]->[0]
          and $rj > $prot->{united_cys}->[$t]->[0]) {
            $bonded += 1;
            $numOfCys++;
            }
          # if second cysteine is inside the fragment
          if ($ri < $prot->{united_cys}->[$t]->[1]
          and $rj > $prot->{united_cys}->[$t]->[1]) {
            $bonded += 2;
            $numOfCys++;
            }

          # if there is only the first cysteine inside the fragment then it can be linked to another fragment as well
          # and also, that cysteine has to be reduced
          if ($bonded == 1) {
            $rs->{$prot->{name}}->[$t]->[0]->[@{$rs->{$prot->{name}}->[$t]->[0]}]->{sequence}= $ri.'-'.$rj;
            $rs->{$prot->{name}}->[$t]->[0]->[$#{$rs->{$prot->{name}}->[$t]->[0]}]->{weight}= $value;
            $rs->{$prot->{name}}->[$t]->[0]->[$#{$rs->{$prot->{name}}->[$t]->[0]}]->{right}= $rj;
            $rs->{$fName}->[$#{$rs->{$fName}}]->{weight}+= $sp->{aa}->{Hy};
            $rs->{$fName}->[$#{$rs->{$fName}}]->{definition}.= 'one cysteine that has to be reduced, ';
            $numOfCys--;
          # if there is only the second cysteine inside the fragment then it can be linked to another fragment as well
          # and also, that cysteine has to be reduced
          }elsif ($bonded == 2) {
            $rs->{$prot->{name}}->[$t]->[1]->[@{$rs->{$prot->{name}}->[$t]->[1]}]->{sequence}= $ri.'-'.$rj;
            $rs->{$prot->{name}}->[$t]->[1]->[$#{$rs->{$prot->{name}}->[$t]->[1]}]->{weight}= $value;
            $rs->{$prot->{name}}->[$t]->[1]->[$#{$rs->{$prot->{name}}->[$t]->[1]}]->{left}= $ri;
            $rs->{$fName}->[$#{$rs->{$fName}}]->{weight}+= $sp->{aa}->{Hy};
            $rs->{$fName}->[$#{$rs->{$fName}}]->{definition}.= 'one cysteine that has to be reduced, ';
            $numOfCys--;
          # zero cysteines are inside the fragment, do nothing
          }elsif ($bonded == 0) {
          # the two cysteines are inside the fragment, let's change the definition
          }elsif ($bonded == 3) {
            $rs->{$fName}->[$#{$rs->{$fName}}]->{definition}.= 'with disulfide bond inside, ';
            }
          $bonded = 0;
          }
        # include versions of cysteines reduced, when there are disulfide bonds BOTH can become reduced
        for (my $t=2; $t<=$numOfCys; $t+=2) {
          $rs->{$fName}->[@{$rs->{$fName}}]->{sequence}= $ri.'-'.$rj;
          $rs->{$fName}->[$#{$rs->{$fName}}]->{weight}= $value + $t*$sp->{aa}->{Hy};
          $rs->{$fName}->[$#{$rs->{$fName}}]->{definition}= 'lineal, intra-disulfide bond('.$t
          .' cys) reduced '.$prot->{name};
          }
        }
      }
    }}
  }
}
sub get_weight {
```

```
  my(undef,$sp,$seq,$start,$end) = @_;
  $start = 0 if (!$start);
  $end = length($$seq)-1 if (!$end);
  my $value = 0;
  for (my $i=$start; $i<=$end; $i++) {
    $value += $sp->{aa}->{substr($$seq,$i,1)};
    }
  $value += $sp->{aa}->{H2O};
  $value += $sp->{aa}->{Hy}*$sp->{spectrum}->{reading};
  $value;
  }
sub reconstruct {
  my ($self,$prot,$sp,$rs) = @_;
  my $fName = $prot->{name}.'FINAL';
  # each cycle has two groups with fragments with the first and the second cysteine.
  # we'll do all combinations between both
  foreach my $t(0..$#{$rs->{$prot->{name}}}) {
    my $ccc = 0;
    # through fragments with the first cysteine of the given pair
    foreach my $tt(@{$rs->{$prot->{name}}->[$t]->[0]}) {
      $ccc++;if ($ccc == 500) {print '.';$ccc = 0;}
      # through fragments with the second cysteine of the given pair
      foreach my $ttt(@{$rs->{$prot->{name}}->[$t]->[1]}) {
        if ($tt->{right}<$ttt->{left}) {
          $rs->{$fName}->[@{$rs->{$fName}}]->{sequence}= $tt->{sequence}.' / '.$ttt->{sequence}
          .' (cys'.$prot->{united_cys}->[$t]->[0].'-cys'.$prot->{united_cys}->[$t]->[1].')';
          $rs->{$fName}->[$#{$rs->{$fName}}]->{weight}= $tt->{weight}+ $ttt->{weight};
          $rs->{$fName}->[$#{$rs->{$fName}}]->{definition}= 'disulphide linked, '.$prot->{name};
          }
        }
      }
    }
  # different disulfide bonds are not contemplated, only the original ones
  }
sub results {
  my($self,$prot,$sp,$rs,$final) = @_;
  my $fName = $prot->{name}.'FINAL';

  my(undef,$min,$hora,$dia,$mes,$any) = localtime(time);
  my $fileName = substr($any,1).'.'.$mes.'.'.$dia.' - '.$prot->{name}.'.txt';

  $prot->{Weight}= $paf->get_weight($sp,\$prot->{sequence});

  my $output = <<EOF;
P(erl)-A(utomated)-F(ragmentation) PAF v0.31, 2009.09.08
Autor: Gerard Such
Last Date: 08 September 2009
-----------------------------------------------------------
fileName: $fileName

- protein name: $prot->{name}
- protein weight: $prot->{Weight}
- enzym name: $sp->{enzym}->[$prot->{enzym}]->{name}
- matrix name: $sp->{matrix}->[$prot->{matrix}]->{name}
- matrix weight: $sp->{matrix}->[$prot->{matrix}]->{weight}
- average mass(1) or isotopic mass(0): $sp->{spectrum}->{analysis}
- positive(1) or negative(-1) read mode: $sp->{spectrum}->{reading}


  PROTEIN SEQUENCE
EOF
  my $rows = 9;
  my $ii;
  my $ij;
  $output .= '1 ';
  for (my $i=0; $i<length($prot->{sequence}); $i++) {
    $output .= $sp->{sNaa}->{substr($prot->{sequence},$i,1)}.'-';
    if ($i == $rows) {
      $ii = $i+1;
      $ij = $i+2;
      $output .= $ii."\n".$ij.'-';
      $rows += 10;
      }
    }

  $output .= <<EOF;

-----------------------------------------------------------

EOF
```

```
  # sort data, a long-time operation
  @{$rs->{$fName}}= sort {$a->{weight}<=> $b->{weight}}@{$rs->{$fName}};
  $output .= substr($_->{weight},0,9).' , '.$_->{sequence}.' , '.$_->{definition}."\n" foreach (@{$rs->{$fName}});

  open OUTPUT, ">$fileName";
  print OUTPUT $output;
  close OUTPUT;
  }
sub debug {
  #$self->debug('message');
  my(undef,$message) = @_;
  print DEBUG $message."\n";
  }
```

Perl code – paf_template.pm

```
# Template PAF 0.31 --- 09.09.08

# ------------------------------------------------------------------------
# NOTES
# DISULFIDE BONDS MUST BE SEQUENTIAL, OTHERWISE IT IS NOT IMPLEMENTED
# ------------------------------------------------------------------------

package PAF_TEMPLATE;
$VERSION = 0.31;
use strict;


sub new {
my(undef,$sp,$exproteins) = @_;
my $template;
$$template = undef;
bless $template, " PAF_TEMPLATE";



# Spectrum
# ------------------------------------------------------------------------
$sp->{spectrum}->{analysis}= 1; #average=1,isotopic=0
$sp->{spectrum}->{reading}= 1; #positive=1,negative=-1

# (Comment the lines if you don't want limites. It is for speed issues.)
#$sp->{spectrum}->{limDown}= 800; # Establish the lowest mass limit to be shown, format issue
#$sp->{spectrum}->{limUp}= 3000; # Establish the highest mass limit to be shown, format issue

# Aminoacids sequence
# ------------------------------------------------------------------------
$sp->{spectrum}->{expasy}= 0; #expasy names (e.g. Q9BYQ7)=1, comment if otherwise
$sp->{protein}->[0]->{name}=  '22K';
$sp->{protein}->[0]->{sum}= 0; # to add in the numeration, e.g. =43, now 1-10 will be 44-53.
$sp->{protein}->[0]->{sequence}= 'FPTIPLSRLF'.  'DNAMLRAHRL'.'HQLAFDTYQE'.'FEEAYIPKEQ'
                .'KYSFLQNPQT'. 'SLCFSESIPT'.  'PSNREETQQK'.'SNLELLRISL'
                .'LLIQSWLEPV'.  'QFLRSVFANS'.  'LVYGASDSNV'.'YDLLKDLEEG'
                .'IQTLMGRLED'.  'GSPRTGQIFK'.  'QTYSKFDTNS'.'HNDDALLKNY'
                .'GLLYCFRKDM'.'DKVETFLRIV'.  'QCRSVEGSCG'.'F';
$sp->{protein}->[0]->{dislufide}= 1; # comment if cisteines are reduced (or there aren't) and sequence is lineal
# insert the cysteines unions (sequence starts at 1, not at 0)
$sp->{protein}->[0]->{united_cys}->[0]->[0] =  53;
$sp->{protein}->[0]->{united_cys}->[0]->[1] =  165;
$sp->{protein}->[0]->{united_cys}->[1]->[0] =  182;
$sp->{protein}->[0]->{united_cys}->[1]->[1] =  189;
# matrix and enzyme information
$sp->{protein}->[0]->{matrix}= 0; #0=sinapinic,1=alphaciano,2=benzoic,3=picolinic
$sp->{protein}->[0]->{enzyme}= 0; #0=thermolysin,1=pepsin,2=elastase,3=trypsin,4=proteinase K
# --------------------
$sp->{protein}->[1]->{name}=  '17K';
$sp->{protein}->[1]->{sum}= 43; # to add in the numeration, e.g. =43, now 1-10 will be 44-53.
$sp->{protein}->[1]->{sequence}= 'FLQNPQT'. 'SLCFSESIPT'.'PSNREETQQK'.'SNLELLRISL'
                .'LLIQSWLEPV'.  'QFLRSVFANS'.  'LVYGASDSNV'.'YDLLKDLEEG'
                .'IQTLMGRLED'.  'GSPRTGQIFK'.  'QTYSKFDTNS'.'HNDDALLKNY'
                .'GLLYCFRKDM'.'DKVETFLRIV'.  'QCRSVEGSCG'.'F';
$sp->{protein}->[1]->{dislufide}= 1; # comment if cisteines are reduced (or there aren't) and sequence is lineal
# insert the cysteines unions (sequence starts at 1, not at 0)
$sp->{protein}->[1]->{united_cys}->[0]->[0] =  53;
$sp->{protein}->[1]->{united_cys}->[0]->[1] =  165;
$sp->{protein}->[1]->{united_cys}->[1]->[0] =  182;
$sp->{protein}->[1]->{united_cys}->[1]->[1] =  189;
# matrix and enzyme information
$sp->{protein}->[1]->{matrix}= 0; #0=sinapinic,1=alphaciano,2=benzoic,3=picolinic
$sp->{protein}->[1]->{enzyme}= 0; #0=thermolysin,1=pepsin,2=elastase,3=trypsin,4=proteinase K
## --------------------
$sp->{protein}->[2]->{name}=  '5K';
$sp->{protein}->[2]->{sum}= 0; # to add in the numeration, e.g. =43, now 1-10 will be 44-53.
$sp->{protein}->[2]->{sequence}= 'FPTIPLSRLF'.  'DNAMLRAHRL'.'HQLAFDTYQE'.'FEEAYIPKEQ'
                .'KYS';
$sp->{protein}->[2]->{dislufide}= 0; # comment if cisteines are reduced (or there aren't) and sequence is lineal
# matrix and enzyme information
$sp->{protein}->[2]->{matrix}= 0; #0=sinapinic,1=alphaciano,2=benzoic,3=picolinic
$sp->{protein}->[2]->{enzyme}= 0; #0=thermolysin,1=pepsin,2=elastase,3=trypsin,4=proteinase K
# --------------------
```

```
# --------------------

$sp->{enzym}->[0]->{name}= 'thermolysin';
$sp->{enzym}->[0]->{cutting}= 0; #-1=left,0=right,1=both;
$sp->{enzym}->[0]->{aa}->[0] = 'L'; #0.943 hydrophobicity
$sp->{enzym}->[0]->{aa}->[1] = 'F'; #1.00
$sp->{enzym}->[0]->{aa}->[2] = 'I'; #0.943
$sp->{enzym}->[0]->{aa}->[3] = 'A'; #0.616
$sp->{enzym}->[0]->{aa}->[4] = 'M'; #0.738
$sp->{enzym}->[0]->{aa}->[5] = 'V'; #0.825
$sp->{enzym}->[0]->{aa}->[6] = 'W'; #0.878
$sp->{enzym}->[0]->{aa}->[7] = 'Y'; #0.880
$sp->{enzym}->[0]->{aa}->[8] = 'C'; #0.680
$sp->{enzym}->[0]->{aa}->[9] = 'G'; #0.501
$sp->{enzym}->[0]->{aa}->[10] = 'P'; #0.711

$sp->{enzym}->[1]->{name}= 'pepsin';
$sp->{enzym}->[1]->{cutting}= 1; #-1=left,0=right,1=both;
$sp->{enzym}->[1]->{aa}->[0] = 'L';
$sp->{enzym}->[1]->{aa}->[1] = 'F';
$sp->{enzym}->[1]->{aa}->[2] = 'E';

$sp->{enzym}->[2]->{name}= 'elastase';
$sp->{enzym}->[2]->{cutting}= -1; #-1=left,0=right,1=both;
$sp->{enzym}->[2]->{aa}->[0] = 'A';
$sp->{enzym}->[2]->{aa}->[1] = 'V';
$sp->{enzym}->[2]->{aa}->[2] = 'L';
$sp->{enzym}->[2]->{aa}->[3] = 'I';
$sp->{enzym}->[2]->{aa}->[4] = 'G';
$sp->{enzym}->[2]->{aa}->[5] = 'S';

$sp->{enzym}->[3]->{name}= 'trypsin';
$sp->{enzym}->[3]->{cutting}= -1; #-1=left,0=right,1=both;
$sp->{enzym}->[3]->{aa}->[0] = 'R';
$sp->{enzym}->[3]->{aa}->[1] = 'K';
$sp->{enzym}->[3]->{aa}->[2] = 'H';

$sp->{enzym}->[4]->{name}= 'proteinase K';
$sp->{enzym}->[4]->{cutting}= -1; #-1=left,0=right,1=both;
$sp->{enzym}->[4]->{aa}->[0] = 'L'; #0.943 hydrophobicity
$sp->{enzym}->[4]->{aa}->[1] = 'F'; #1.00
$sp->{enzym}->[4]->{aa}->[2] = 'I'; #0.943
$sp->{enzym}->[4]->{aa}->[3] = 'A'; #0.616
$sp->{enzym}->[4]->{aa}->[4] = 'M'; #0.738
$sp->{enzym}->[4]->{aa}->[5] = 'V'; #0.825
$sp->{enzym}->[4]->{aa}->[6] = 'W'; #0.878
$sp->{enzym}->[4]->{aa}->[7] = 'Y'; #0.880
$sp->{enzym}->[4]->{aa}->[8] = 'C'; #0.680
$sp->{enzym}->[4]->{aa}->[9] = 'G'; #0.501
$sp->{enzym}->[4]->{aa}->[10] = 'P'; #0.711

$sp->{matrix}->[0]->{name}= 'sinapinic';
$sp->{matrix}->[0]->{weight}= 224;
$sp->{matrix}->[1]->{name}= 'alphaciano';
$sp->{matrix}->[1]->{weight}= 189;
$sp->{matrix}->[2]->{name}= 'benzoic';
$sp->{matrix}->[2]->{weight}= 154;
$sp->{matrix}->[3]->{name}= 'picolinic';
$sp->{matrix}->[3]->{weight}= 139;

}
```