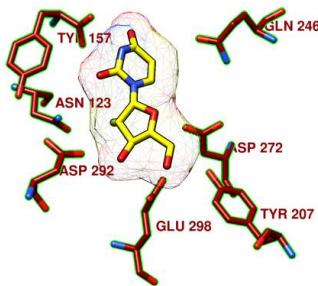




Tesis Doctoral

# SÍNTESIS DE CLOFARABINA MEDIANTE EL USO DE NUCLEÓSIDO 2'-DESOXIRRIBOSILTRANSFERASAS



Memoria de Tesis Doctoral presentada por:

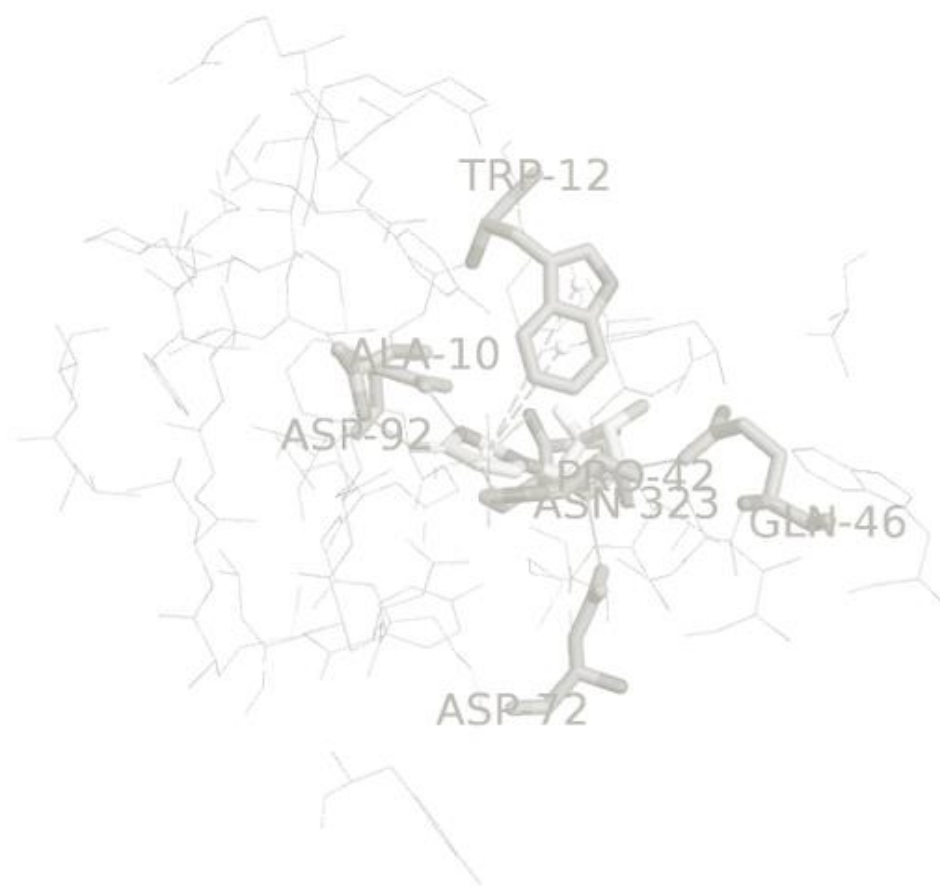
**Víctor Manuel Deroncelé Thomas**

**Director:** Dr. Rafael Montilla Arévalo

**PROGRAMA DE DOCTORADO EN  
QUÍMICA SOSTENIBLE**

**UNIVERSITAT JAUME I**

**2018**





DON **RAFAEL MONTILLA ARÉVALO**, DOCTOR Y DIRECTOR DEL DEPARTAMENTO DE BIOTECNOLOGÍA Y JEFE DE ESTUDIOS DE INKEMIA IUCT GROUP.

CERTIFICA: Que DON. **Víctor Manuel Deroncelé Thomas**, Licenciado en Ciencias Biológicas por la Universidad de Oriente, Cuba, ha realizado bajo su dirección, en el programa de doctorado interuniversitario de Química Sostenible coordinado por la Universitat Jaume I, el trabajo titulado “**Síntesis Clofarabina mediante el uso de Nucleósido 2'-desoxirribosiltransferasas**” y que ésta constituye su memoria de Tesis Doctoral para optar al grado de Doctor en Química Sostenible.

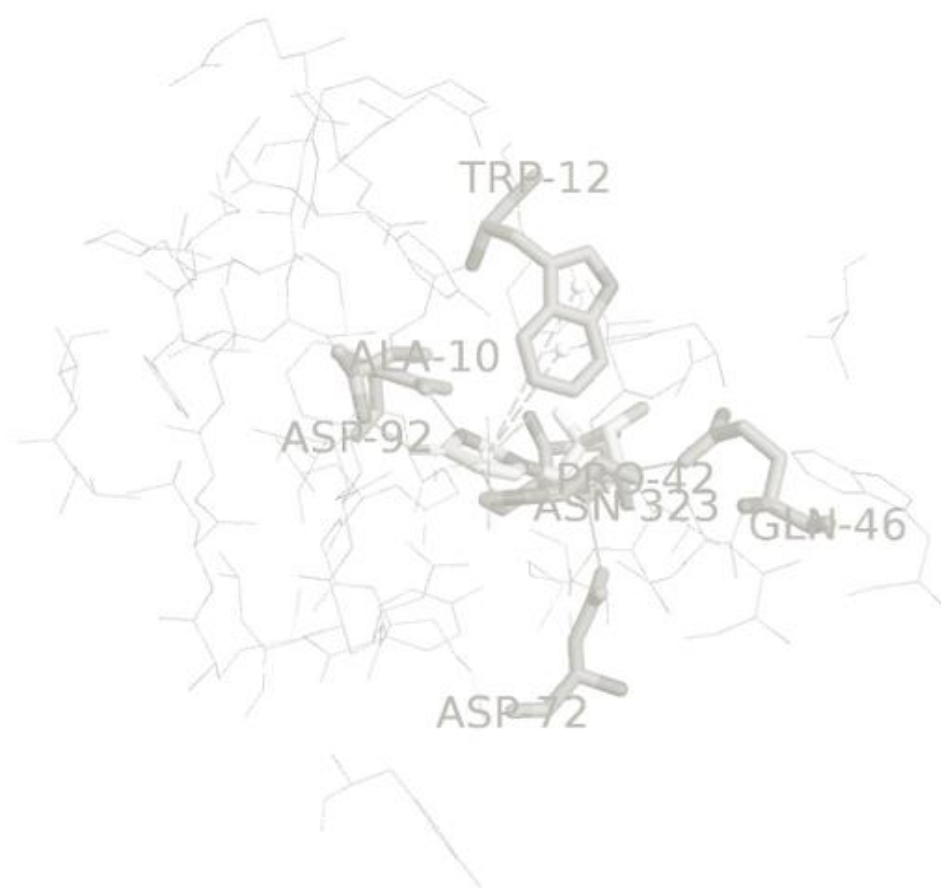
Y para que así conste y en cumplimiento de la legislación vigente, firma el presente certificado en Mollet del Vallés a 20 de Junio de 2018.

**Dr. Rafael Montilla Arévalo**

Jefe de Estudios  
Inkemia IUCT group  
Telf. 935793432  
C. Álvarez de Castro, 63  
08100 Mollet del Vallés

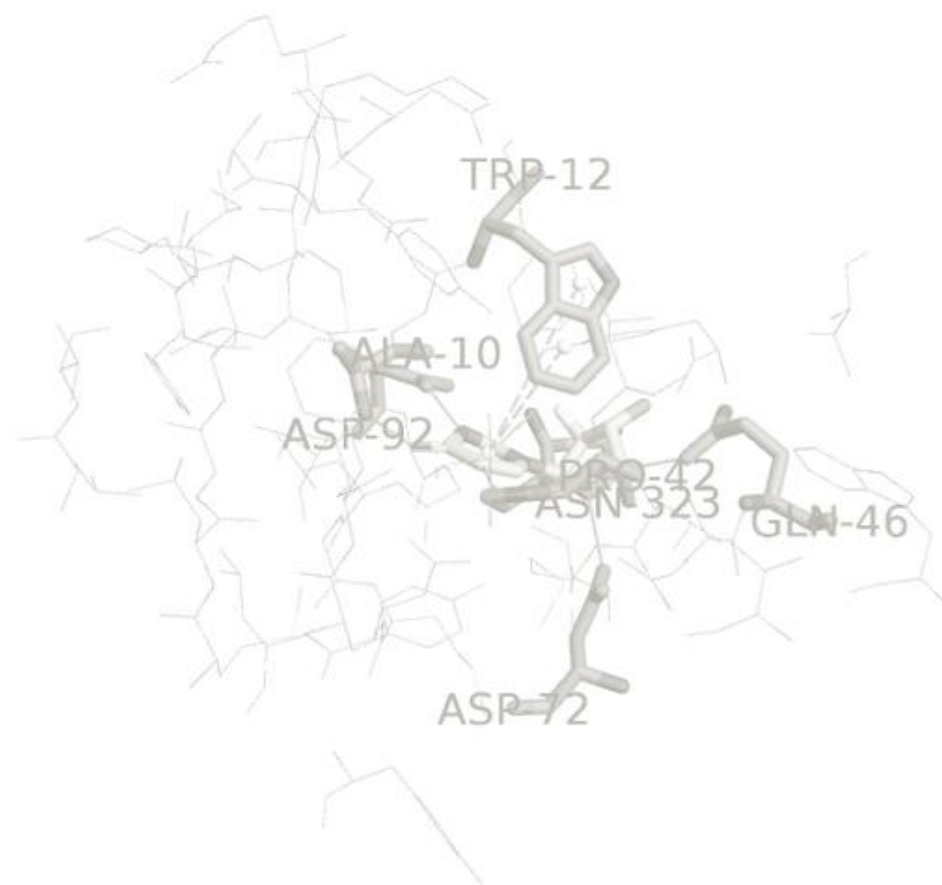


Dedicada a la memoria de mi padre,  
A mi madre y a mi hermana,  
A mi esposa María,  
A mis hijos, Víctor, Dani, Àlex, Leonardo & Elena.



## **Agradecimientos**

*Me gustaría expresar mi más sincero agradecimiento a mis compañeros del grupo Inkemia IUCT Group, este es un trabajo de equipo, de un gran equipo. Al Dr. Rafael Montilla, mi amigo y director de este proyecto. Al Dr. Santiago Luis y a la Dra. Belén Altava, mis catalizadores desde Castellón. Pero sobretodo a mi flaca, siempre apoyándome, incondicionalmente.*



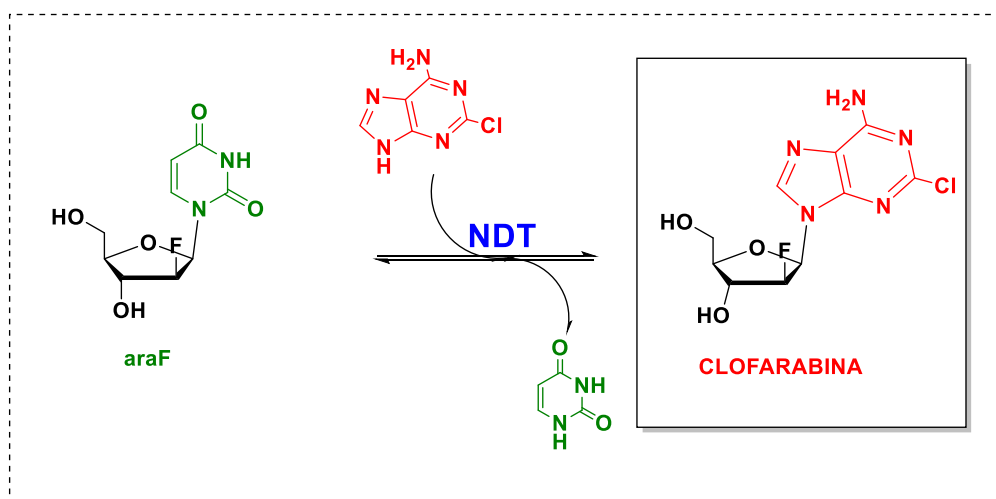


## Resumen

Los 2-desoxi-2-fluoro- $\beta$ -D-arabinofuranósidos de pirimidina y de purina son análogos de nucleósidos con un amplio espectro de actividad biológica, y de entre ellos, la **clofarabina** ha sido aplicada con éxito para el tratamiento de la **leucemia aguda pediátrica**. La síntesis química de este **API** (Ingrediente Farmacéutico Activo) requiere de múltiples pasos, lo que supone procesos costosos y no respetuosos con el medio ambiente. De ahí la propuesta de este trabajo de utilizar y evaluar la síntesis enzimática, como alternativa para simplificar el proceso de producción y reducir el coste económico y medioambiental de su producción actual.

La proteína con actividad desosirribosiltransferasa de *Lactobacillus delbrueckii* subsp. *lactis* fue seleccionada, luego de estudios *in silico* de acoplamiento molecular, como candidata para llevar a cabo la síntesis. En este sentido, la enzima fue clonada, expresada y producida en cantidades suficientes para las pruebas de concepto y los estudios de optimización del proceso biocatalítico.

Para nuestro conocimiento, este trabajo (y la patente que le acompaña), es el primer reporte del uso de una **nucleósido 2'-desosirribosiltransferasa** para la síntesis de **clofarabina**. En el trabajo se hace un análisis de los «cuello de botella» que, en el escenario actual, hacen que la síntesis enzimática no sea una opción, claramente competitiva, en relación a la síntesis química de este API y se hacen propuestas para alcanzar este objetivo.



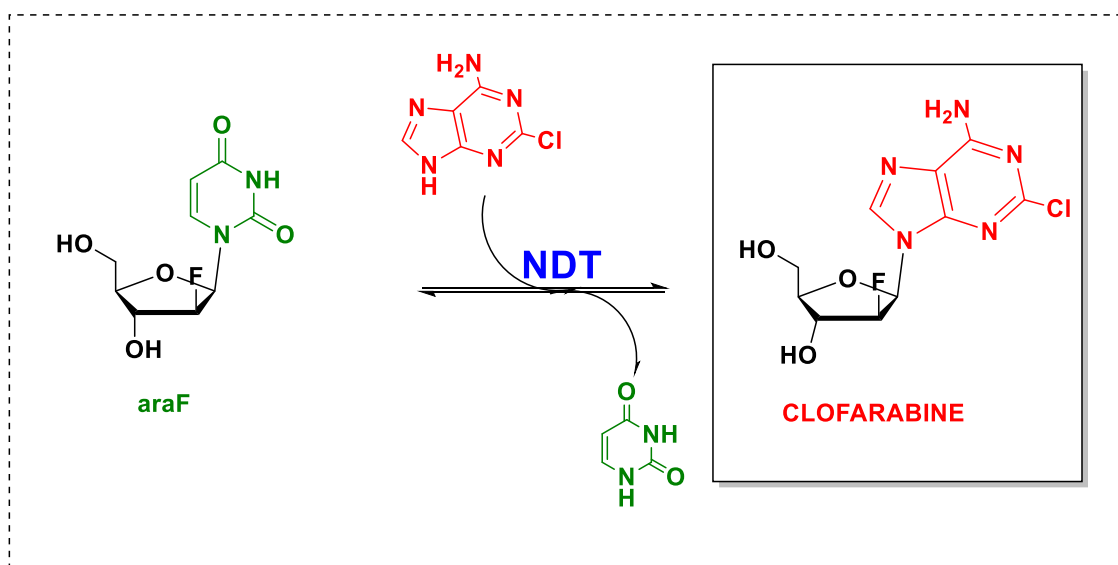


## Abstract.

2-deoxy-2-fluoro- $\beta$ -D-arabinofuranosides of pyrimidine and purine are nucleoside analogues with a broad spectrum of biological activity, and among them, **Clofarabine** has been successfully applied for the treatment of **pediatric acute lymphoblastic leukemia**. The chemical synthesis of this API (Active Pharmaceutical Ingredient) requires multiple steps, which implies costly and non-environmentally friendly processes. Hence the proposal of this study is to use and evaluate the enzymatic synthesis, as an alternative to simplify the production process and reduce the economic and environmental costs of the current chemical process.

The protein with 2'-deoxyribosyltransferase activity from *Lactobacillus delbrueckii* subsp. *lactis* was selected, after *in silico* molecular coupling studies, as a candidate to carry out the synthesis. In this sense, the enzyme was cloned, expressed and produced in sufficient quantities for the proof of concept studies and the process optimization.

To the best of our knowledge, the present study (and the accompanying patent application), is the first report of the use of a nucleoside 2'-deoxyribosyltransferase enzyme for the synthesis of clofarabine. In this study the main bottlenecks that, in the current scenario, make the novel enzymatic synthesis no longer competitive, when is compared to the chemical synthesis of clofarabine, and some proposals are made to reach this objective.





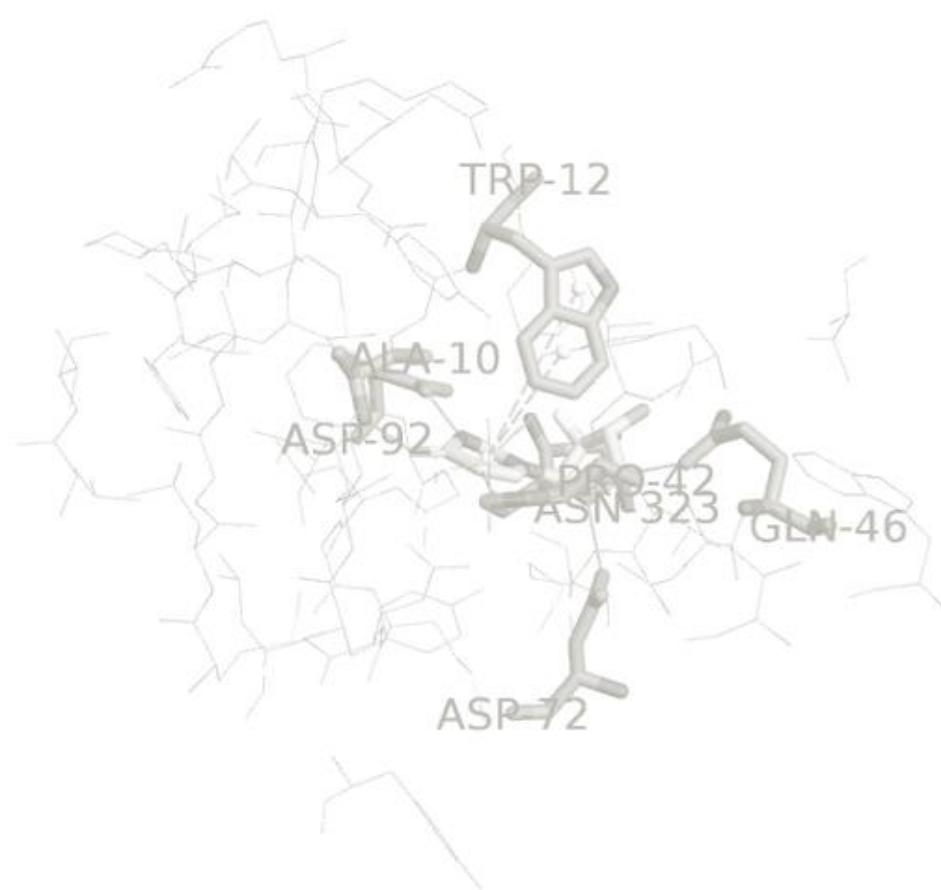
## Prefacio

Esta Tesis se presenta como parte de los requisitos para optar al grado Académico de Doctor en Química Sostenible, en la **Universitat Jaume I**. La misma contiene los resultados obtenidos en investigaciones llevadas a cabo en el ámbito del Departamento *Advanced Biotech* de *Inkemia IUCT group* (<http://www.inkemia.com/>), bajo la supervisión del Dr. Rafael Montilla Arévalo, durante el período comprendido entre octubre del 2015 y julio del 2017.

Los resultados forman parte de una de las tecnologías patentada por PLASMIA Biotech (<http://www.plasmiabiotech.com>), una compañía biotecnológica dedicada a investigar nuevas formas de síntesis, más eficientes, para diferentes tipos de fármacos. Específicamente, se describen estrategias basadas en el **empleo de enzimas que catalizan las reacciones de síntesis de análogos nucleosídicos** en un único paso. La obtención de análogos nucleosídicos, con esta tecnología, supone una **mayor rapidez de síntesis**, una **reducción de costes** y una **menor producción de residuos**, disminuyendo el impacto económico y medioambiental de la producción de estos.

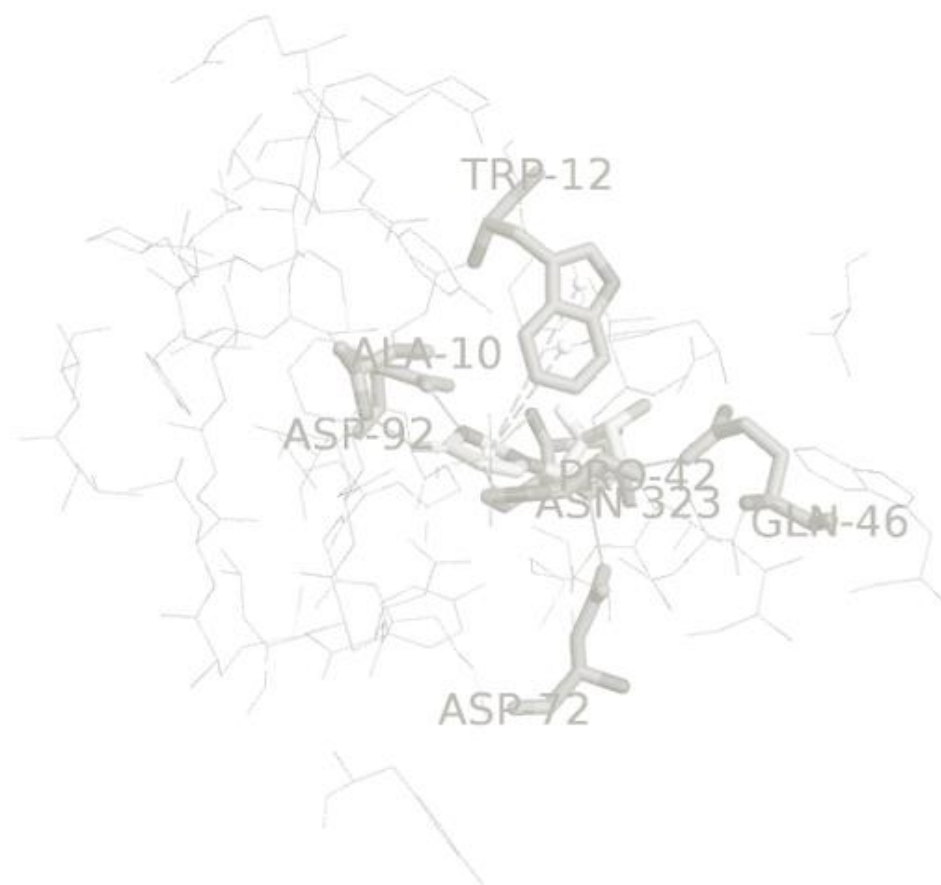
Víctor Manuel Deroncelé Thomas





## Contenido

Resumen .....	ix
Abstract.....	xi
Prefacio .....	xiii
Índice de figuras .....	xvii
Índice de tablas .....	xxi
Abreviaturas.....	xxiii
1. INTRODUCCIÓN.....	1
1.1. Cáncer en pediatría.....	1
1.2. Síntesis de análogos de nucleósidos.....	19
1.3. Química verde y Biocatálisis.....	23
2. HIPÓTESIS Y OBJETIVOS .....	43
2.1. Planteamiento del problema.....	43
2.2. Hipótesis de investigación.....	44
2.3. Objetivos .....	45
3. MATERIALES Y MÉTODOS.....	49
3.1. Cribado virtual de NDTs para la síntesis biocatalítica de clofarabina .....	49
3.2. Síntesis de Clofarabina usando la nucleósido 2'-desoxirribosiltransferasa de <i>Lactobacillus delbrueckii</i> subsp. <i>lactis</i> .....	51
4. RESULTADOS Y DISCUSIÓN.....	72
4.1. Estudios de acoplamiento ( <i>docking</i> ) molecular .....	73
4.2. Síntesis de Clofarabina usando la nucleósido 2'-desoxirribosiltransferasa de <i>Lactobacillus delbrueckii</i> subsp. <i>lactis</i> .....	77
4.3. Escalado de la reacción biocatalítica.....	117
5. CONCLUSIONES Y PERSPECTIVAS.....	145
5.1. Conclusiones.....	145
5.2. Perspectivas futuras.....	147
6. BIBLIOGRAFIA .....	151
ANEXOS.....	165
6.1. Anexo1: Patent Application US 2016/0076070 A1 .....	165





## Índice de figuras

<b>Figura 1:</b> Distribución por grupo diagnóstico de los tumores infantiles en España.....	2
<b>Figura 2:</b> Evolución de una célula sanguínea.....	5
<b>Figura 3:</b> Esquema de tratamiento LAL/SHOP 2005.....	7
<b>Figura 4:</b> Estructura de los análogos nucleosídicos .....	12
<b>Figura 5:</b> Modificaciones químicas para la obtención de análogos de nucleósidos.....	13
<b>Figura 6:</b> Mecanismo general de los análogos nucleosídicos antineoplásicos.....	15
<b>Figura 7:</b> Síntesis de clofarabina. Método Montgomery.....	21
<b>Figura 8:</b> Síntesis de clofarabina. Método ILEX.....	22
<b>Figura 9:</b> Síntesis de clofarabina propuesta por Bauta y col.....	23
<b>Figura 10:</b> Esquema general de la reacción de transglicosilación.....	31
<b>Figura 11:</b> Esquema de la reacción catalizada por NPs .....	31
<b>Figura 12:</b> Esquema general de la reacción de transglicosilación.....	33
<b>Figura 13:</b> Reacción de transglicosilación catalizada por las nucleósido 2'- desoxirribosiltransferasas. B1: purina o pirimidina; B2: purina o pirimidina .....	33
<b>Figura 14:</b> Estructura del centro activo de la NDT de <i>L/NDT</i> empleando 2'- desoxiuridina como sustrato. ....	36
<b>Figura 15:</b> Síntesis quimioenzimática de clofarabina. ....	38
<b>Figura 16:</b> Cascada para la transformación 2-desoxi-2-fluoro- $\alpha$ -D-arabinofuranosa en clofarabina.....	39
<b>Figura 17:</b> Estructura de la 2'-desoxiuridina. ....	50
<b>Figura 18:</b> Estructura del 2-desoxi-2-fluoro- $\alpha$ -D-arabinofuranosa. ....	51
<b>Figura 19:</b> Esquema del vector pGEM-T Easy ( <i>Promega</i> <sup>®</sup> ).....	54
<b>Figura 20:</b> Representación esquemática del vector de expresión pET-22b. ....	58
<b>Figura 21:</b> Esquema de la reacción del ensayo de actividad para <i>LdNDT2</i> .....	69
<b>Figura 22:</b> Síntesis de clofarabina usando nucleósido desoxirribosiltransferasas como biocatalizadores. ....	72
<b>Figura 23:</b> Alineamiento con ClustalW de las secuencias de NDTs utilizadas.. ....	75
<b>Figura 24:</b> Centro activo de <i>L/NDT</i> .....	75
<b>Figura 25:</b> Alineamiento de las secuencias de las NDT de <i>Lactobacillus delbrueckii</i> ( <i>LdNDT2</i> ) y <i>Lactobacillus leichmannii</i> ( <i>L/NDT</i> ).....	78
<b>Figura 26:</b> Subclonaje del gen <i>Ldndt2</i> .....	80

<b>Figura 27:</b> Esquema del plásmido pET22b- <i>Ldndt2</i> .....	81
<b>Figura 28:</b> Análisis de proteínas citosólicas solubles por SDS-PAGE y teñidas con Azul de Coomassie. ....	81
<b>Figura 29:</b> Perfiles de crecimiento del clon recombinante en los medios Terrific broth (TB), Luria broth (LB), medio mínimo (M9) y medio mínimo modificado (MMM)...	82
<b>Figura 30:</b> Análisis de proteínas citosólicas solubles por SDS-PAGE en los cuatro medios ensayados., .....	83
<b>Figura 31:</b> Perfiles de crecimiento y actividad volumétrica del clon recombinante. ....	84
<b>Figura 32:</b> Curva de crecimiento y perfil de inducción de <i>E. coli</i> BL21[pET22ldndt2.	86
<b>Figura 33:</b> Producción de <i>Ld</i> NDT2 en un cultivo discontinuo alimentado. ....	89
<b>Figura 34:</b> Registro gráfico de la evolución de los parametros fermentativos a escala de bioreactor de 5L.....	92
<b>Figura 35:</b> Gel SDS-PAGE de la precipitación fraccionada. ....	93
<b>Figura 36:</b> Primera cromatografía de intercambio iónico de la <i>Ld</i> NDT2 producida por <i>E. coli</i> BL21(pET22- <i>ndt2</i> ).....	94
<b>Figura 37:</b> Visualización del estado de la purificación de la <i>Ld</i> NDT2 en los eluatos positivos, de la primera cromatografía de intercambio iónico. ....	95
<b>Figura 38:</b> Segunda cromatografía de intercambio iónico de la <i>Ld</i> NDT2 producida por <i>E. coli</i> BL21(pET22 <i>Ldndt2</i> ). Cromatografía de intercambio iónico (Q-Sepharose). ....	95
<b>Figura 39:</b> Electroforegrama de segunda cromatografía de intercambio iónico .....	95
<b>Figura 40:</b> Esquema de purificación propuesto.....	96
<b>Figura 41:</b> Esquema detallado del proceso de producción del Biocatalizador.....	98
<b>Figura 42:</b> Esquema de la reacción del ensayo de actividad para <i>Ld</i> NDT2.. ....	106
<b>Figura 43:</b> Prueba de concepto para la síntesis enzimática de clofarabina.. ....	106
<b>Figura 44:</b> Síntesis de clofarabina. Prueba de concepto.....	107
<b>Figura 45:</b> Progreso de la reacción de síntesis de clofarabina mediada por <i>Ld</i> NDT2.	108
<b>Figura 46:</b> Solubilidad de la 2CIA a diferentes pHs.....	109
<b>Figura 47:</b> Efecto del pH sobre la actividad de la <i>Ld</i> NDT2.....	110
<b>Figura 48:</b> Efecto de la temperatura sobre la actividad sintética de la <i>Ld</i> NDT2. ....	110
<b>Figura 49:</b> Efecto de la carga enzimática sobre los rendimientos de conversión de clofarabina.. .....	113
<b>Figura 50:</b> Esquema del diseño Doehlert para el estudio de dos factores.....	114
<b>Figura 51:</b> Compuestos que intervienen en la síntesis de clofarabina.....	119

<b>Figura 52:</b> Perfil HPLC de la composición del crudo final de la reacción de síntesis de clofarabina..	120
<b>Figura 53:</b> Pureza por HPLC de clofarabina recristalizada.....	121
<b>Figura 54:</b> Esquema del proceso de purificación de la clofarabina.....	125
<b>Figura 55:</b> Influencia (simulación) del nivel de expresión de la enzima sobre los costes de producción de la clofarabina.....	129
<b>Figura 56:</b> Influencia (simulación) de los rendimientos de la reacción sobre los costes de producción de la clofarabina.....	130
<b>Figura 57:</b> Síntesis química de clofarabina propuesta por Bauta y col. <sup>37, 38, 125</sup> .....	133
<b>Figura 58:</b> Pentágono radial obtenido a partir del análisis de la síntesis química de la clofarabina protegida, en el proceso propuesto por Bauta y col.....	136
<b>Figura 59:</b> Hexágonos radiales representativos de la síntesis de clofarabina..	136



## Índice de tablas

<b>Tabla 1:</b> Medicamentos antileucémicos utilizados para tratar leucemias linfoblásticas.	8
<b>Tabla 2:</b> Nucleósidos y nucleótidos naturales .....	11
<b>Tabla 3:</b> Análogos de nucleósidos (antimetabolitos) aprobados por la FDA. ....	13
<b>Tabla 4:</b> Requisitos para procesos enzimáticos viables.....	27
<b>Tabla 5:</b> Observaciones importantes sobre biocatálisis en química verde .....	28
<b>Tabla 6:</b> Nucleósido fosforilasas. ....	32
<b>Tabla 7:</b> Estructuras cristalinas utilizadas en este apartado.....	49
<b>Tabla 8:</b> Composición del medio MRS. ....	51
<b>Tabla 9:</b> Secuencias de oligonucleótidos empleados en este apartado.....	52
<b>Tabla 10:</b> Secuencia de la nucleósido 2'-desoxirribosiltransferasa de <i>Lactobacillus delbrueckii</i> subsp. <i>lactis</i> . ....	53
<b>Tabla 11:</b> Plásmidos utilizados.....	53
<b>Tabla 12:</b> Programa de amplificación por PCR.....	54
<b>Tabla 13:</b> Mezcla de reacción de la PCR.....	55
<b>Tabla 14:</b> Mezcla de reacción de PCR para detección de plásmidos con inserto.....	57
<b>Tabla 15:</b> Perfil térmico de PCR para detección de plásmidos con inserto. ....	57
<b>Tabla 16:</b> Composición del medio TB .....	59
<b>Tabla 17:</b> Composición del medio mínimo M9.....	60
<b>Tabla 18:</b> Composición del medio mínimo modificado (MMM).....	61
<b>Tabla 19:</b> Disoluciones sucesivas de proteína patrón para construir la recta de calibración (método de Bradford). ....	68
<b>Tabla 20:</b> Resultados del estudio de acoplamiento ciego. ....	77
<b>Tabla 21:</b> Proteína total y producción de biomasa de los cultivos del clon recombinante <i>E. coli</i> BL21 (DE3)/ [pET22Ldndt2] en diferentes medios de cultivo.....	83
<b>Tabla 22:</b> Resumen de los resultados de las inducciones en medio TB (matraz).....	85
<b>Tabla 23:</b> Resumen de los resultados de las inducciones en medio TB (Biorreactor operado en <i>batch</i> ). ....	87
<b>Tabla 24:</b> Comparación de la producción de LdNDT2 en cultivos en discontinuo (matraz y biorreactor) y discontinuo alimentado (biorreactor).....	91
<b>Tabla 25:</b> Resumen del proceso de purificación de la LdNDT2 recombinante. ....	96

<b>Tabla 26:</b> Coste del medio de fermentación (escala: 1000 litros) para la producción de LdNDT2..	101
<b>Tabla 27:</b> Servicios (0,1€/kWh ).	101
<b>Tabla 28:</b> Coste de los equipos principales.....	103
<b>Tabla 29:</b> Costes operacionales anuales de la producción de la enzima <i>LdNDT2</i> .....	104
<b>Tabla 30:</b> Resumen de los costes anuales para la producción de 294 kg del concentrado con actividad 2'-desoxi-2'-fluororibosil-transferasa. ....	104
<b>Tabla 31:</b> Efecto del uso de cosolventes sobre la actividad de LdNDT2. ....	112
<b>Tabla 32:</b> Resultados experimentales tras aplicar los parámetros del diseño.....	115
<b>Tabla 33:</b> Verificación de los resultados ofrecidos por el modelo obtenido utilizando el algoritmo Simplex. ....	116
<b>Tabla 34:</b> Condiciones de reacción. Escala de 0,18 L.....	117
<b>Tabla 35:</b> Composición del crudo final de la reacción biocatalítica. ....	119
<b>Tabla 36:</b> Solubilidad de los componentes del crudo.....	123
<b>Tabla 37:</b> Coste de los reactivos utilizados en la producción y en la purificación de la clofarabina. ....	124
<b>Tabla 38:</b> Servicios (0,1€/kWh ). ....	124
<b>Tabla 39:</b> Coste de los equipos principales del proceso. ....	125
<b>Tabla 40:</b> Costes operacionales anuales de la producción de la enzima <i>LdNDT2</i> .....	126
<b>Tabla 41:</b> Resumen de los costes de producción del API.....	126
<b>Tabla 42:</b> Resumen de los parámetros identificados como «cuello de botella » del proceso biocatalítico. ....	134
<b>Tabla 43:</b> Puntuación para el cálculo de la Ecoescala.....	137
<b>Tabla 44:</b> Puntuación para el cálculo de la Ecoescala en la síntesis de clofarabina....	141
<b>Tabla 45:</b> Métricas asociadas a la síntesis de clofarabina. ....	142

## Abreviaturas

Abreviaturas	Descripción
$\mu$ max	Velocidad máxima de crecimiento
2'-FdAdo	2'-fluor-2'-desoxiadenosina
2CdA	2-cloro-2'- desoxiadenosina
ALL	<i>Acute lymphocytic leukemia</i> (Leucemia linfocítica aguda)
AML	Leucemia mieloide aguda ( <i>acute myelogenous leukemia</i> )
Ap	Ampicilina
API	Ingrediente Farmacéutico activo ( <i>Active Pharmaceutical Ingredient</i> )
ara-A	9-( $\beta$ -D-arabinofuranosil)adenina; arabinosiladenina
ara-C	Arabinósido de citosina
araF	2'-fluoro- arabinofuranosil-2'-desoxiuridina
ara-U	1-( $\beta$ -D-Arabinofuranosyl)uracil
ARNm	ARN mensajero
ARNr	ARN ribosómico
BPF	Buenas prácticas de fabricación (GMP: <i>Good manufacturing practice</i> )
BSA	Albúmina de suero bovino ( <i>Bovine Serum Albumin</i> )
CADC	Cultivo de Alta Densidad Celular
CAGR	Tasa anual compuesta de crecimiento
CAPEX	<i>Capital Expenditures</i> (gasto de capital)
CIIC	Centro Internacional de Investigaciones sobre el Cáncer
Cl-F-ara-A	Clofarabina (2-cloro-2'-fluoro-2'-desoxiarabinosiladenina)
CMP	Citidin monofosfato
CNT	Transportadores concentradores de nucleosides
CTP	Citidin trifosfato
dAd <sup>2F</sup>	9-(2-Desoxi-2-fluoro- $\beta$ -D-arabinofuranosil)adenina
dAdo	Desoxiadenosina
dAdo <sub>2F</sub>	2'-Desoxi-2'-fluoroadenosina
dCK	Desoxicitidina quinasa
dFdC	2',2'-difluoro 2'-desoxicitidina
dFUrd	2'-desoxi -2 ' Fluorouridina
dGK	Desoxiguanina quinasa

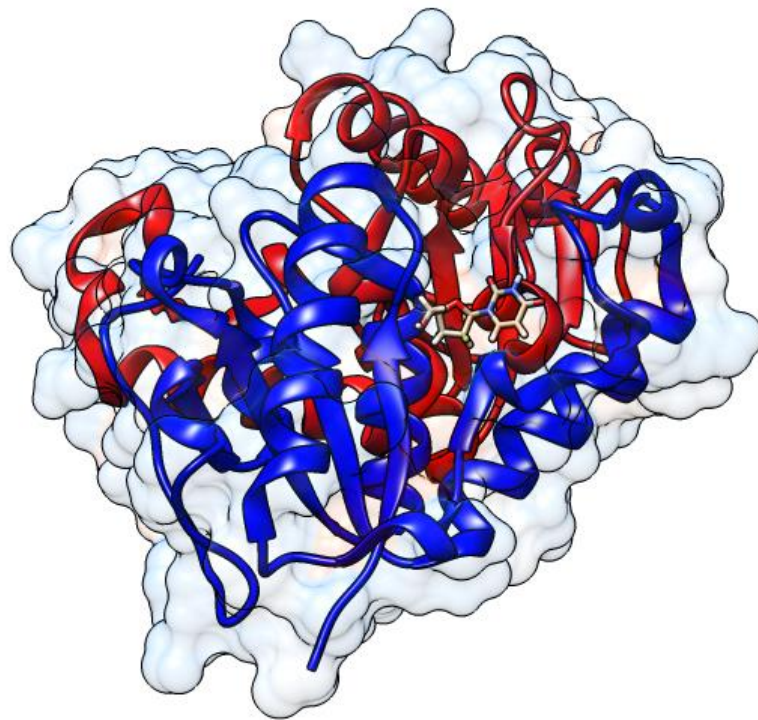
<b>Abreviaturas</b>	<b>Descripción</b>
<b>dGuo</b>	Desoxiguanosina
<b>dIno</b>	Desoxiinosina
<b>DMF</b>	N,N-dimetilformamida anhidra
<b>DMSO</b>	Dimetil sulfóxido
<b>DNA</b>	Ácido desoxiribonucleico ( <i>Deoxyribonucleic acid</i> )
<b>dNTP</b>	Desoxinucleótido trifosfato
<b>DO<sub>600nm</sub></b>	Densidad óptica a 600 nm
<b>DRTasas</b>	Nucleósido 2´desoxirribosiltransferasas
<b>DTT</b>	Dithiothreitol
<b>dUrd</b>	2´-Desoxiuridina
<b>dUrd<sup>2F</sup></b>	1-(2-Deoxy-2-fluoro-β-D-arabinofuranosyl)uracil
<b>dUrd<sub>2F</sub></b>	2´-Deoxy-2´-fluorouridine
<b>EcPNP</b>	Purin nucleósido fosforilasa de <i>E. coli</i>
<b>EcTP</b>	Timidin nucleósido fosforilasa de <i>E. coli</i>
<b>EcUP</b>	Uridin nucleósido fosforilasa de <i>E. coli</i>
<b>EDTA</b>	Ácido etilendiamino tetraacético
<b>E-factor</b>	<i>Environmental factor</i>
<b>EMEA</b>	Agencia europea del medicamento
<b>ENT</b>	Transportadores equilibradores de nucleosides
<b>EPA</b>	<i>Environmental Protection Agency</i> (Agencia de Protección Ambiental)
<b>F-ara-A</b>	2-fluoro-arabinosiladenina, fludarabina
<b>FDA</b>	<i>Food and Drug Administration</i> (Administración de Alimentos y Medicamentos)
<b>G</b>	Tiempo de generación
<b>GMP</b>	Guanosina 5´-monofosfato
<b>GsPyNP</b>	Pirimidin nucleósido fosforilasa de <i>G. stearothermophilus</i>
<b>GtPNP</b>	Purin nucleósido fosforilasa de <i>G. thermoglucosidasius</i>
<b>HPLC</b>	Cromatografía líquida de alta resolución ( <i>High performance liquid chromatography</i> )
<b>IARC</b>	<i>International Agency for Research on Cancer</i>
<b>IMP</b>	Inosina 5´-monofosfato
<b>IPTG</b>	Isopropil-β-D-tiogalactopiranosido
<b>Kan</b>	Kanamicina



<b>Abreviaturas</b>	<b>Descripción</b>
<b>kDa</b>	Kilodaltons
<b>LB</b>	Medio de cultivo Luria Bertani
<b>LdNDT</b>	Nucleósido 2'-desoxirribosiltransferasa tipo II de <i>Lactobacillus delbrueckii</i> subsp. <i>lactis</i>
<b>LINDT</b>	Nucleósido 2'-desoxirribosiltransferasa tipo II de <i>Lactobacillus leichmannii</i>
<b>LrNDT</b>	Nucleósido 2'-desoxirribosiltransferasa tipo II de <i>Lactobacillus reuterii</i>
<b>M9</b>	Medio mínimo M9
<b>MES</b>	Ácido 2-(N-morfolino) etanosulfónico
<b>MMM</b>	Medio mínimo modificado
<b>MTAP</b>	5' -Metiltioadenosina fosforilasa
<b>NCBI</b>	<i>Nacional Center for Biotechnology Information</i>
<b>NdRTs</b>	N-desoxirribosiltransferasas
<b>NDT</b>	Nucleósido 2'-desoxirribosiltransferasa tipo II
<b>NP</b>	Nucleósidos fosforilasas
<b>NRT</b>	N-desoxirribosiltransferasas.
<b>OMS</b>	Organización Mundial de la Salud
<b>OPEX</b>	<i>Operational cost</i> (Costes operacionales)
<b>PAGE</b>	Electroforesis en gel de poliacrilamida
<b>PCR</b>	Reacción en cadena de la polimerasa ( <i>Polimerase Chain Reaction</i> )
<b>PDB</b>	<i>Protein Data Bank</i> (Banco de Datos de Proteínas)
<b>PDT</b>	Nucleósido 2'-desoxirribosiltransferasa tipo I
<b>pET 22b(+)</b>	Plásmido comercial pET 22b(+) (Novagen)
<b>pET 28b(+)</b>	Plásmido comercial pET 28b(+) (Novagen, Reino Unido)
<b>pKa</b>	Constante de disociación ácida ( <i>Acid disssociation constant</i> )
<b>PMI</b>	<i>Process mass intensity</i> (Intensidad de masa en el proceso)
<b>PNP</b>	Purin nucleósido fosforilasa
<b>pO<sub>2</sub></b>	Presión parcial de oxígeno
<b>PPM</b>	Fosfopentomutasa
<b>PyNP</b>	Pirimidin nucleósido fosforilasa
<b>QA</b>	Garantía de calidad ( <i>Quality Assurance</i> )
<b>QC</b>	Control de calidad ( <i>Quality Control</i> )
<b>RK</b>	Riboquinasa

<b>Abreviaturas</b>	<b>Descripción</b>
<b>RMSD</b>	Raíz media de desviación de los cuadrados ( <i>root mean square deviation</i> )
<b>SEHOP</b>	Sociedad Española de Hematología y Oncología Pediátrica
<b>SNC</b>	Sistema nervioso central
<b>TAE</b>	Tampón formado por Tris, borato y EDTA,
<b>TB</b>	<i>Terrific broth</i> (caldo Terrific)
<b>TP</b>	Timidina fosforilasa
<b>TSA</b>	Triptona de Soja Agar
<b>UFC</b>	Unidades Formadoras de Colonias
<b>UI</b>	Unidad internacional
<b>UP</b>	Uridina fosforilasa
<b>X-gal</b>	5-bromo-4-cloro-3-indolil- $\beta$ -D-galactopiranosido

<b>Aminoácidos</b>		
<b>Código 1 letra</b>	<b>Código 3 letras</b>	<b>Aminoácido</b>
A	Ala	Alanina
R	Arg	Arginina
N	Asn	Asparagina
D	Asp	Ácido aspártico
C	Cys	Cisteína
Q	Gln	Glutamina
E	Glu	Ácido glutámico
G	Gly	Glicina
H	His	Histidina
I	Ile	Isoleucina
L	Leu	Leucina
K	Lys	Lisina
M	Met	Metionina
F	Phe	Fenilalanina
P	Pro	Prolina
S	Ser	Serina
T	Thr	Treonina
W	Trp	Triptófano
Y	Tyr	Tirosina
V	Val	Valina

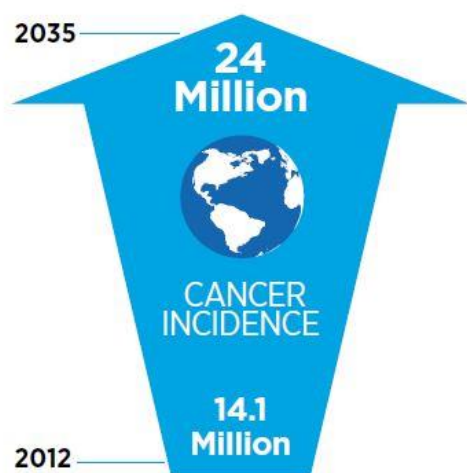


# Introducción.



# 1. INTRODUCCIÓN

En febrero del 2014 el Centro Internacional de Investigaciones sobre el Cáncer (CIIC), que es el organismo de la Organización Mundial de la Salud especializado en la lucha contra dicha enfermedad, publicó un Informe mundial sobre el cáncer (*World Cancer Report 2014*<sup>1</sup>), en el que describían aspectos relacionados con investigación y la lucha contra el cáncer y mostraban las últimas estadísticas sobre la incidencia del cáncer y sus tasas de mortalidad en todo el mundo.



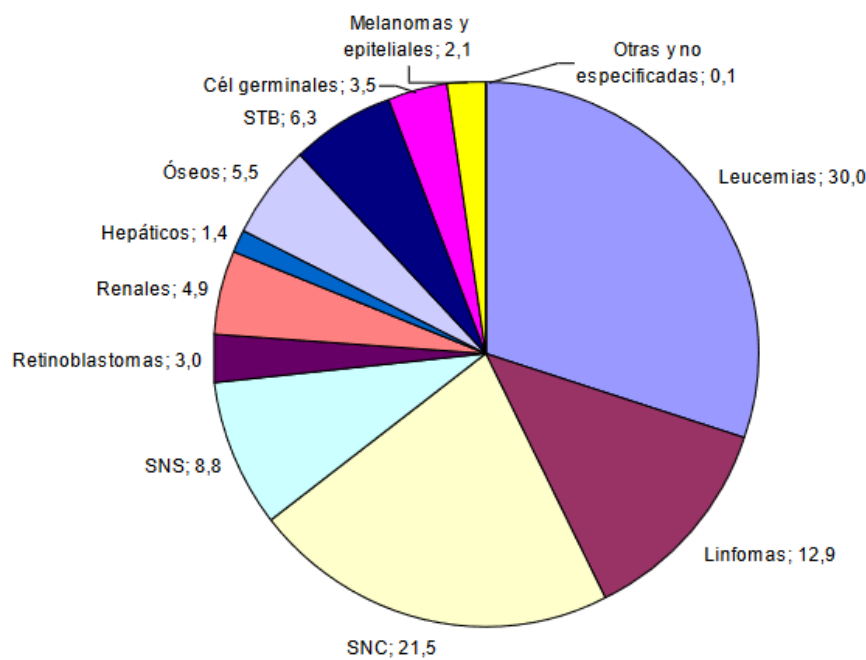
Los avances en la investigación biomédica han provocado un aumento en la esperanza de vida de la población pero, al mismo tiempo, este envejecimiento de la población ha hecho que los casos de cáncer también hayan ido en aumento. Sin ir más lejos, en el año 2012 la carga mundial de cáncer alcanzó, según la OMS, los 14 millones de casos nuevos al año, una cifra que se prevé aumente hasta los 22 millones anuales en los próximos dos

decenios. En el mismo periodo, se prevé que las muertes por cáncer aumenten desde los aproximadamente 8,2 millones de casos anuales hasta los 13 millones de casos anuales<sup>1</sup>. En España, según los datos calculados a partir del proyecto GLOBOCAN 2012<sup>2</sup>, de la *International Agency for Research on Cancer* (IARC), la incidencia anual de cáncer calculada para el año 2015, excluyendo de este cálculo los casos de cáncer de piel no melanoma, será de 227076 personas (135954 varones y 91122 mujeres), siendo el tipo más frecuente el cáncer colorrectal, por delante, en términos globales, del cáncer de pulmón y el cáncer de mama.

## 1.1. Cáncer en pediatría.

El cáncer es la primera causa de muerte por enfermedad entre el primer año de vida y la adolescencia. En España se diagnostica anualmente un nuevo caso de cáncer por cada 6500 niños menores de 15 años. Esto representa 1100 casos nuevos entre los 0 y 14 años y 450 entre los 15 y 19<sup>3</sup>. En menores de 15 años predominan los cánceres

no epiteliales (leucemias, tumores del sistema nervioso central y linfomas) y tumores embrionarios (neuroblastomas, retinoblastomas, nefroblastomas, meduloblastomas, rhabdomyosarcomas embrionarios, tumores de células germinales), mientras que los carcinomas raramente se diagnostican, constituyendo cerca del 1,5% de tumores malignos a esta edad. Según el Registro Nacional de Tumores Infantiles, los tipos de cáncer más frecuentes en España son las leucemias (30%), los tumores del Sistema Nervioso Central (20%) y los linfomas (13%) (Figura 1)<sup>4</sup>.



**Figura 1:** Distribución por grupo diagnóstico de los tumores infantiles en España.

### 1.1.1. *Leucemia.*

La leucemia<sup>♦</sup> o leucosis es un grupo de enfermedades malignas de la médula ósea (cáncer hematológico) que provoca un aumento incontrolado de leucocitos (glóbulos blancos) clonales en la médula ósea, que suelen pasar a la sangre periférica aunque en ocasiones no lo hacen (leucemias aleucémicas). Ciertas proliferaciones malignas de glóbulos rojos se incluyen entre las leucemias (eritroleucemia)<sup>5</sup>.

<sup>♦</sup> Leucemia: Literalmente, significa “*sangre blanca*”; la palabra está formada por dos elementos griegos: *leuc*, una variante de *leuco*= λευκός, "blanco"; y *emia*, αἷμα = "sangre".

## A. Tipos de leucemia en niños.

Con frecuencia la leucemia se describe como **aguda** (que crece rápidamente) o **crónica** (que crece lentamente). Casi todas las leucemias en niños son agudas.

Los dos tipos principales de **leucemia aguda** son:

- *Leucemia linfocítica aguda* (linfoblástica) (*acute lymphocytic leukemia, ALL*): Alrededor de tres de cuatro leucemias en niños son **ALL**. Esta leucemia se origina de formas tempranas de linfocitos en la médula ósea.
- *Leucemia mieloide aguda* (*acute myelogenous leukemia, AML*): este tipo de leucemia, también llamada leucemia mieloide aguda, leucemia mielocítica aguda o leucemia no linfocítica aguda representa la mayoría de los casos remanentes. La **AML** se inicia a partir de las células mieloides que forman los glóbulos blancos (que no son linfocitos), los glóbulos rojos o las plaquetas.
- *Leucemia de linaje híbrido o mixto*: en estas leucemias poco comunes, las células tienen características de la **ALL** y de la **AML**. En niños, son generalmente tratadas como la **ALL** y usualmente responden a este tratamiento como la **ALL**.

Las leucemias crónicas son mucho más comunes en los adultos que en los niños. Suelen crecer más lentamente que las leucemias agudas, aunque también son más difíciles de curar.

Las **leucemias crónicas** se pueden dividir en dos tipos.

- *Leucemia mieloide crónica* (*chronic myelogenous leukemia, CML*): Esta leucemia ocurre rara vez en niños. El tratamiento es similar al empleado en adultos.
- *Leucemia linfocítica crónica* (*chronic lymphocytic leukemia, CLL*): Esta leucemia se presenta muy pocas veces en los niños.

## B. Leucemia Linfoblástica Aguda

La **ALL** es la neoplasia más frecuente en la infancia. Su prevalencia es de un 40% del total de las patologías oncológicas. Se caracteriza por la transformación maligna de una célula progenitora linfoide inmadura que da lugar a una proliferación clonal incontrolada de células bloqueadas en un punto de su diferenciación<sup>6</sup>.

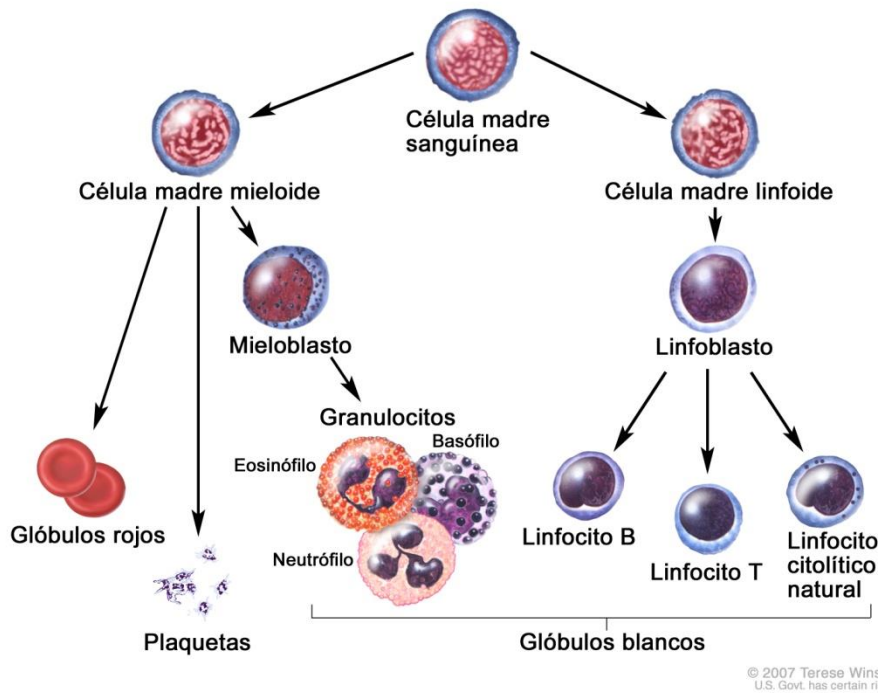
La adquisición por parte de estos precursores linfoides de una serie de anomalías genéticas, hace que se altere su proceso de maduración normal, produciendo una parada madurativa en su proceso de diferenciación y favoreciendo la proliferación de las células transformadas. En consecuencia, se produce una proliferación clonal de estas células inmaduras o blastos que invaden la médula ósea y la sangre periférica, produciendo un fracaso de la función medular e infiltración blástica de órganos y tejidos, principalmente nódulos linfáticos, hígado, bazo y sistema nervioso.

Las leucemias linfoblásticas son aquellas cuya celularidad deriva de la estirpe linfoide, es decir, linfocitos **B** y **T** respectivamente. En otras palabras, las **ALLs** pueden ser de estirpe **B** o **T**. Mientras que las leucemias mieloblásticas derivan de las células de la estirpe mieloide como los glóbulos rojos, neutrófilos, basófilos, eosinófilos y plaquetas (Figura 2). Así se diferencian entre agudas y crónicas de acuerdo con la funcionalidad de las células. Las leucemias agudas se caracterizan por tener una población celular no funcional por estar absolutamente inmaduras a diferencia de las leucemias crónicas en las que las células tienen un mayor grado de maduración. A partir de esta premisa podemos inferir que las leucemias agudas generalmente son más agresivas que las crónicas. El origen de este padecimiento se asocia fuertemente a factores genéticos como el síndrome de Down, síndrome de Li-fraumeni, síndrome de Klinefelter, enfermedad de Wiskott Aldrich, anemia de Fanconi, la exposición a radiaciones ionizantes, a agentes químicos en edades tempranas y a una predisposición citogenética específica como la presencia del cromosoma Filadelfia<sup>7</sup>.

### **C. El tratamiento de la ALL infantil**

El tratamiento dado para el cáncer es muy variable y depende de varios factores incluyendo el tipo, el lugar y la extensión del cáncer, así como en el estado físico del paciente. Los tratamientos son diseñados para matar o extirpar directamente a las células cancerosas o para llevarlas a su muerte por medio de la privación de las señales bioquímicas necesarias para la división celular o para estimular las defensas propias.





**Figura 2:** Evolución de una célula sanguínea. Imagen que muestra las células que derivan la estirpe linfoide y la estirpe mieloide en condiciones de salud\*.

Los tratamientos pueden ser divididos en categorías basadas en su fin y modo de acción♦ e incluyen:

- **Cirugía:** Muchas veces es el primer tratamiento, sobre todo cuando se trata de tumores sólidos. Si el cáncer es detectado en una etapa temprana, la cirugía puede ser suficiente para curar al paciente, al eliminar todas las células cancerosas.
- **Radioterapia:** Tratamiento que usa altas dosis de radiación para destruir células cancerosas y reducir tumores.
- **Quimioterapia:** Tratamiento que usa fármacos para destruir células cancerosas.
- **Inmunoterapia:** Tratamiento que ayuda al sistema inmunitario a combatir el cáncer.
- **Terapia dirigida:** Tratamiento que actúa sobre los cambios que promueven el crecimiento, la división y diseminación de las células cancerosas.
- **Trasplante de células madre:** Procedimiento que restaura las células madre formadoras de sangre que se destruyeron por las dosis elevadas utilizadas en tratamientos del cáncer, como quimioterapia o radioterapia.

\* <https://www.mhealth.org>

♦ <http://www.cancer.gov/>.

El tratamiento principal de los niños con leucemia linfocítica aguda (**ALL**) es quimioterapia. La Tabla 1 muestra un listado de los medicamentos antileucémicos utilizados para tratar leucemias linfoblásticas y las dianas biológicas sobre los cuales actúan.

El tratamiento se divide usualmente en tres fases<sup>♦</sup>

- ***Inducción.***

El objetivo principal de la primera fase del tratamiento o fase de inducción es conseguir la remisión de la enfermedad, es decir, que la médula ósea deje de fabricar células leucémicas o blastos y se instaure la hematopoyesis normal. La duración de esta fase es de 5-6 semanas y los fármacos utilizados son la prednisona (glucocorticoide), vincristina (alcaloide de la vinca), asparraginasa, daunorrubicina (antraciclina) y ciclofosfamida (alquilante).

- ***Consolidación (también llamada intensificación).***

La fase consolidación tiene como objetivo erradicar blastos residuales resistentes al tratamiento recibido. Por lo general, se basa en la administración de altas dosis de metotrexato (análogo del ácido fólico) en infusión de 24 horas junto con rescate con ácido folínico y mercaptopurina oral. El tratamiento se completa con citarabina y triple intratecal.

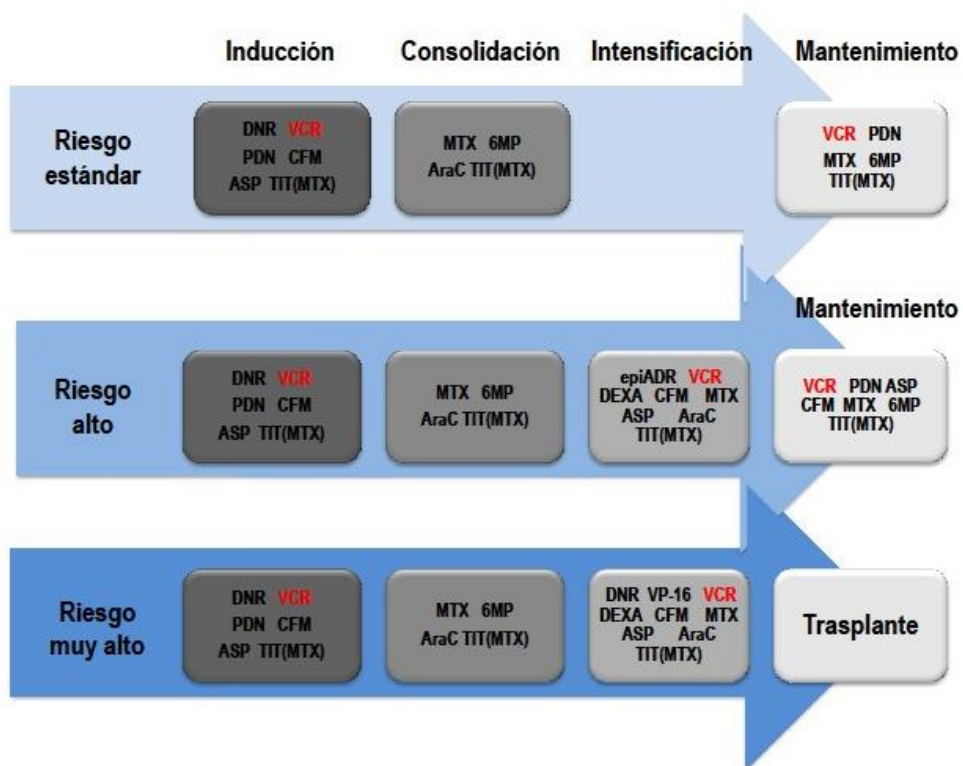
- ***Mantenimiento.***

Si la leucemia sigue en remisión después de la inducción y la consolidación, se puede comenzar la terapia de mantenimiento. La mayoría de los planes de tratamiento usan 6-mercaptopurina diariamente y metotrexato semanalmente, administrados en forma de pastillas, frecuentemente junto con vincristina, que se administra intravenosamente, y un esteroide (prednisona o dexametasona).

---

<sup>♦</sup> [www.cancer.org](http://www.cancer.org) y [www.cancer.gov/espanol](http://www.cancer.gov/espanol)

En España, el protocolo LAL/SHOP aprobado por la Sociedad Española de Hematología y Oncología Pediátrica (SEHOP) ha sido uno de los protocolos más utilizados para el tratamiento de la ALL desde el año 1989. En la Figura 3 se describe esquemáticamente el Protocolo LAL/SHOP 2005, que ha estado vigente en España desde 2005 a 2013. En este esquema los pacientes son asignados inicialmente a un grupo de riesgo, que está determinado por diagnósticos iniciales. En algunos pacientes este riesgo inicial puede ser modificado una vez se dispongan los resultados de los estudios de citogenética/molecular, de respuesta al tratamiento de Inducción y del estudio de la enfermedad residual al final de la Inducción.



**Figura 3:** Esquema de tratamiento LAL/SHOP 2005<sup>8</sup>

La aplicación de protocolos de tratamiento estandarizados ha resultado en una mejoría significativa en pacientes de todos los grupos de riesgo. Sin embargo, uno de los principales problemas que presentan estos protocolos son las diferencias en la respuesta a la quimioterapia entre pacientes.

**Tabla 1:** Medicamentos antileucémicos utilizados para tratar leucemias linfoblásticas<sup>9</sup>.

<b>Clase</b>	<b>Agente</b>	<b>Diana</b>
<b>Análogo de nucleósido purínico</b>		
	<b>Clofarabina</b>	Ribonucleótido reductasa; ADN polimerasa; mitocondria
	Nelarabina	Ribonucleótido reductasa; síntesis
	Forodesina	Purín nucleósido fosforilasa
<b>Alcaloides de la vinca</b>		
	Sulfato de Vincristina	Tubulina
<b>Inhibidores de quinasas</b>		
<b>Inhibidor de la quinasa ABL1</b>		
	Dasatinib; Nilotinib; Imatinib, Ponatinib	Quinasa ABL1; factor B de Crecimiento derivado de plaquetas
<b>Inhibidor Aurora quinasa</b>		
	Alisertib	Aurora quinasa A
<b>Inhibidor Quinasa Janus (JAK)</b>		
	Ruxolitinib; TG101348; CYT387	JAK
<b>Inhibidores de tirosina quinasa</b>		
	Lestaurtinib; Midostaurin; Sorafenib; Quizartinib; Tandutinib; Sunitinib	Tirosina Quinasa 3 Similar a FSM
<b>Otros inhibidores</b>		
<b>Inhibidor de proteasoma</b>		
	Bortezomib	Vía ubiquitina-proteasoma
<b>Diana de rapamicina en mamífero (mTOR)</b>		
	Sirolimus; Temsirolimus; Everolimus	mTOR
<b>Inhibidor de Farnesiltransferasa</b>		
	Tipifarnib; Lonafarnib	Ras, lamin A
<b>Inhibidor de <math>\gamma</math>-secretasa</b>		
	RO4929097	$\gamma$ -secretasa

**Tabla 1:** Medicamentos antileucémicos utilizados para tratar leucemias linfoblásticas (cont.)

Clase	Agente	Diana
<b>Inductor de la apoptosis</b>		
	Obatoclox; Oblimersen	Bcl-2
	Antagonista del receptor de quimiocina (CXCR4) antagonist	
	Plerixafor	Ejes CXCL12 (SDF1)/CXCR4
<b>Terapia Epigenética</b>		
	Inhibidor de la ADN metiltransferasas	
	Azacitidina; Decitabina	ADN metiltransferasas
	Inhibidor de la Histona metiltransferasa	
	EPZ-5676	DOT1L
<b>Inhibidor de Histonas desacetilasa</b>		
	Vorinostat; Panobinostat; Depsipeptide; Ácido valproico	Histona desacetilasa
<b>Terapia Inmune</b>		
	Blinatumomab	CD19 (Presentador CD3 células
	SAR3419	CD19
	DT2219ARL	CD19 y 22
	Rituximab	CD20
	Epratuzumab, Moxetumomab; Inotuzumab ozogamicin	CD22
	Alemtuzumab	CD52
	Terapia Celular Células NK	Receptores tipo inmunoglobulina de las células NK
	Células T con receptor de antígeno quimérico CD19 receptor	CD19

#### **D. El mercado de los medicamentos para el tratamiento de la leucemia linfocítica aguda<sup>10</sup>.**

El mercado de los medicamentos para el tratamiento de la leucemia linfocítica aguda se estima alcance los \$11,3 millones para el año 2020, con una tasa anual compuesta de crecimiento (CAGR) del 4% de 2015 a 2020. Actualmente existen tres regímenes (**Hyper-CVAD, Linker y CALGB 8811**) y tres medicamentos en monoterapia (**Oncaspar, Clolar y Arranon**) disponibles en el mercado. Además, hay tres fármacos: **Graspa** (ERYtech Pharma), **Marqibo** (Talon Therapeutics) e **Inotuzumab ozogamicin** (Pfizer, Inc.) en ensayos clínicos (fase III).

En la actualidad el mercado está siendo impulsado por la creciente tasa de incidencia de los casos, las terapias y esquemas de tratamiento innovadores, encabezados por los tratamientos multi-fármaco. Y además, por el vencimiento del período de validez legal de las patentes de los principales fármacos utilizados en el tratamiento de las ALLs. Así, por ejemplo, la patente del fármaco **Arranon** (nelarabina), expiró en febrero de 2013, mientras que la patente de **Clolar** (clofarabina) expira en 2018. Por otro lado, la patente de **Gleevec**, un fármaco utilizado como terapia combinada en el régimen de **Hyper-CVAD**, ha expirado en 2015.

##### *1.1.2. Análogos de nucleósidos como agentes anticancerígenos.*

El modelo terapéutico clásico para impedir la proliferación de células cancerosas se basa en estrategias que se limitan a evitar que estas células repliquen su ADN. La explotación de nuevos conocimientos sobre mecanismos bioquímicos en los que intervienen desoxinucleótidos ha llevado al desarrollo de grupos de fármacos usados como antimetabolitos, que constituyen una clase importante de compuestos con actividad terapéutica en el tratamiento del cáncer y cuya acción última consiste en interferir la síntesis del ADN<sup>11</sup>.

Los análogos de nucleósidos y de bases nitrogenadas fueron los primeros agentes quimioterapéuticos utilizados para el tratamiento médico del cáncer. Esta familia de compuestos ha crecido para incluir una amplia variedad de nucleósidos derivados de purinas y pirimidinas, con actividad farmacológica sobre tumores sólidos malignos y

trastornos hematológicos. Estos agentes se comportan como antimetabolitos, dado que compiten con los nucleósidos fisiológicos e interactúan con un gran número de dianas intracelulares para inducir citotoxicidad. (La mayoría de los nucleósidos con actividad farmacológica están constituidos, estructuralmente, por tres elementos fundamentales (Figura 4):

El **grupo hidroximetilo**, que es necesario para la fosforilación de la molécula por las quinasas, para conseguir que tenga la actividad biológica deseada.

La **base heterocíclica**, implicada en el proceso de reconocimiento del nucleósido, mediante puentes de hidrógeno específicos.

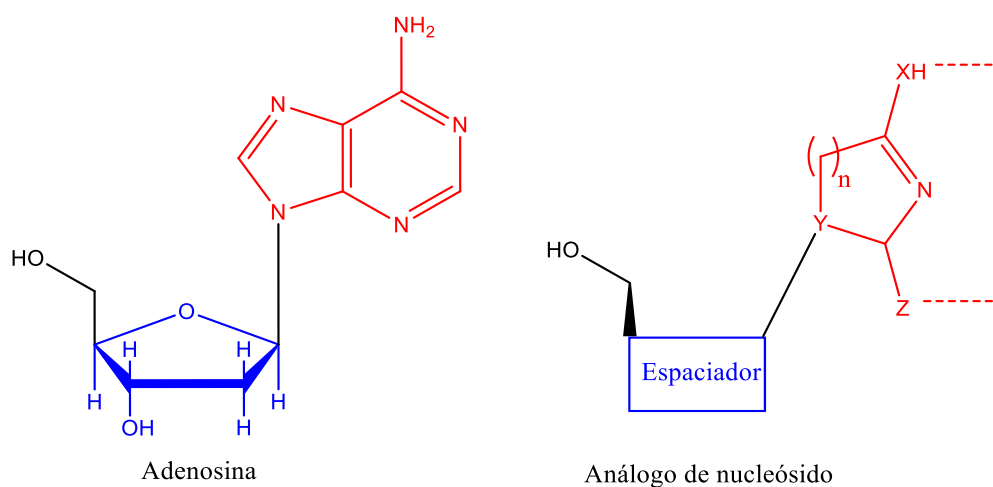
El **anillo furanósico**, que al parecer actúa como una especie de espaciador que permite la correcta orientación del grupo hidroximetilo y de la base (Tabla 2).

**Tabla 2:** Nucleósidos y nucleótidos naturales<sup>♦</sup>

Base	Nucleósido		Nucleótido		
	(=Base + pentosa)		(=nucleósido + fosfato)		
	Ribo-nucleósido	Desoxiribo-nucleósido	NMP	NDP	NTP
			dNMP	dNDP	dNTP
<b>Purinas</b>					
<b>A</b> denina	Adenosina	Desoxiadenosina	AMP	ADP	ATP
			dAMP	dADP	dATP
<b>G</b> uanina	Guanosina	Desoxiguanosina	GMP	GDP	GTP
			dGMP	dGDP	dGTP
<b>Pirimidinas</b>					
<b>C</b> itosina	Citidina	Desoxicitidina	CMP	CDP	CTP
			dCMP	dCDP	dCTP
<b>T</b> imina	Timidina	Desoxitimidina	TMP	TDP	TTP
			dTMP	dTDP	dTTP
<b>U</b> racilo	Uridina	Desoxiuridina	UMP	UDP	UTP
			dUMP	dUDP	dUTP

**NMP** = nucleósido monofosfato; **NDP** = nucleósido difosfato; **NTP** = nucleósido trifosfato; **dNMP** = dexosinucleósido monofosfato; **dNDP** = dexosinucleósido difosfato; **dNTP** = dexosinucleósido trifosfato.

<sup>♦</sup> Tomado de: <http://www.web-books.com/MoBio/Free/Ch3A4.htm>.

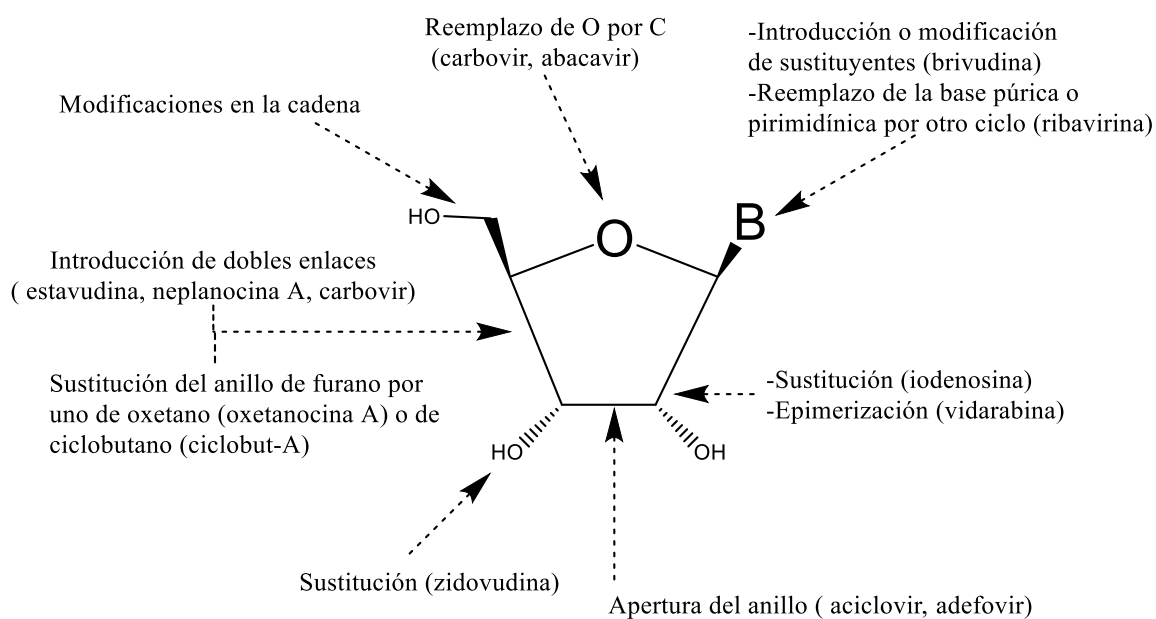


**Figura 4:** Estructura de los análogos nucleosídicos<sup>12</sup>.

La introducción de modificaciones sobre los nucleósidos naturales, les convierte en análogos modificados que mantienen una alta similitud estructural, pero son capaces de bloquear los procesos enzimáticos celulares. Las modificaciones sobre la estructura pueden afectar a la parte del azúcar (o pseudoazúcar), a la base purínica o pirimidínica, o bien a ambas partes. En la Figura 5 se muestran algunos ejemplos de modificaciones comunes para la obtención de análogos de nucleósidos. Pueden, por ejemplo, insertarse heteroátomos (ej., **F**, **Cl**, **I**) en lugar de **-H** o **-OH**; o sustituir **-H** en la posición de **-OH** (desoxiazúcares); inversión de la configuración en un carbono quiral (ej., **C-2'** o **C-3'**); sustitución del azúcar (desoxi) ribofuranosilo por una cadena alifática (ej., aciclovir); L-azúcares, o la sustitución del oxígeno del anillo por **C** o **S**.

Las células humanas tienen la capacidad de recuperar/reciclar las bases purínicas y pirimidínicas necesarias para síntesis de los desoxirribonucleótidos que se utilizan en la síntesis de ADN. Los análogos sintéticos de estos precursores han demostrado ser una clase importante de agentes anticancerígenos y de entre ellos 15 antimetabolitos purínicos y pirimidínicos han sido aprobados por la *Food and Drug Administration* (FDA) para el tratamiento del cáncer (Tabla 3).





**Figura 5:** Modificaciones químicas para la obtención de análogos de nucleósidos<sup>13</sup>

**Tabla 3:** Análogos de nucleósidos (antimetabolitos) aprobados por la FDA<sup>14</sup>.

<b>Análogo de nucleósido</b>	<b>Año de aprobación</b>
Trifluridina/tipiracil	2015
Decitabina	2006
Nelarabina	2005
<b>Clofarabina</b>	<b>2004</b>
Vidaza	2004
Capecitabina	1998
Gemcitabina	1996
Cladribina	1992
Fludarabina	1991
Pentostatina	1991
Floxuridina	1970
Citarabina	1969

<b>Análogo de nucleósido</b>	<b>Año de aprobación</b>
6-tioguanina	1966
5-fluorouracilo	1962
6-mercaptopurina	1953

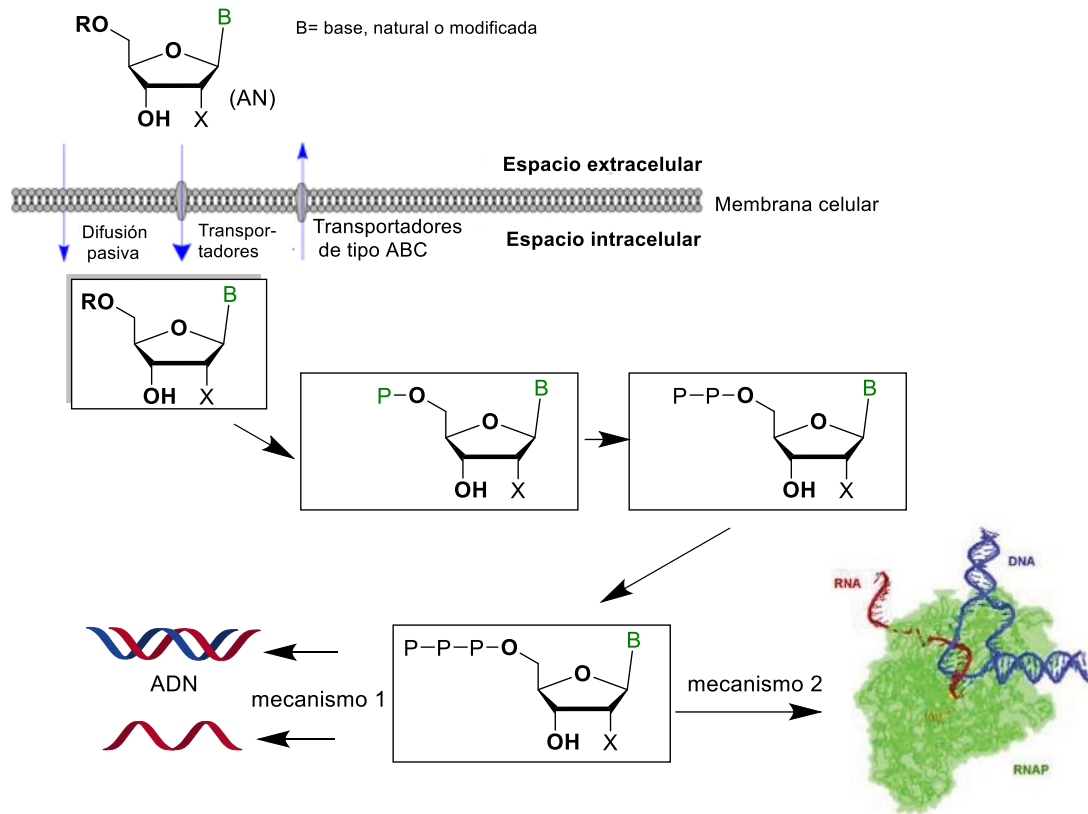
#### 1.1.2.1. Actividad Biológica de los análogos de nucleósidos anticancerígenos.

Las vías principales de administración de los análogos nucleosídicos son la intravenosa y por medio de formulaciones orales. El transporte de los nucleósidos hacia el interior celular puede ser activo o pasivo, por medio de transportadores equilibradores de nucleósidos (**ENT**) o de transportadores concentradores de nucleósidos (**CNT**)<sup>15</sup>. Los transportadores CNT son proteínas de membrana ubicuas que median el transporte de purinas y pirimidinas a favor de gradiente. Los CNTs median el transporte activo (energético-dependiente) hacia el interior celular, aprovechando el gradiente de Na<sup>+</sup>.

Los análogos de nucleósidos son administrados como profármacos por lo cual necesitan que el metabolismo celular los convierta en los correspondientes metabolitos activos. Para ello, enzimas de la rutas de “*salvamento*”<sup>♦</sup> celular, específicamente del tipo fosforilasas, llevan a cabo la fosforilación de los análogos de nucleósidos para generar los nucleósido-5'-monofosfato<sup>16</sup>. Esta fosforilación inicial, a menudo es la etapa limitante en el proceso para generar metabolitos activos (Figura 6)<sup>17</sup> y se han identificado cuatro quinasas para llevar a cabo este proceso: la 2'-desoxicitidina quinasa, la timidina quinasa 1, la timidina quinasa 2 y la desoxiguanidina quinasa.

A menudo la fosforilación inicial se realiza por la 2'-desoxicitidina quinasa, enzima que en las células cancerosas se suele expresar de 3 a 5 veces más que en la mayoría de las células normales.

<sup>♦</sup> Rutas de recuperación, en las que se recuperan las bases para sintetizar los nucleótidos



**Figura 6:** Mecanismo General de los Análogos Nucleosídicos (ANs) antineoplásicos. RNAP: RNA polymerase

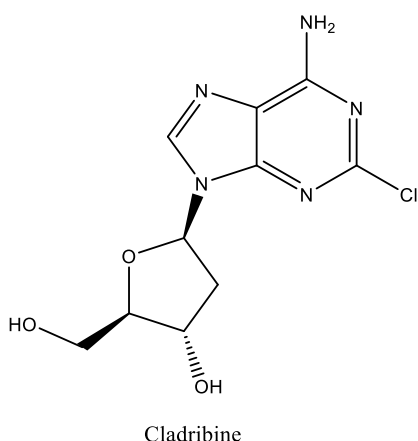
Los nucleósido-5'-monofosfato se convierten en nucleósido-5'-difosfato y nucleósido-5'-trifosfato mediante diversas quinasas celulares, y estos últimos constituyen los sustratos de las ADN polimerasas y pudiendo ser incorporados en el ADN, en los procesos de replicación, síntesis o reparación del ADN, lo que da lugar a horquillas de replicación bloqueadas y terminación anticipada de la cadena. Cuando esto ocurre, se activan varios sensores de daño del ADN, que estimulan la reparación del ADN, detienen la progresión celular y, a menudo, conducen a la apoptosis<sup>18</sup>. Por otro lado, ciertos nucleósido-5'-trifosfato pueden incorporarse en el ARN, dando lugar a la terminación de la transcripción, y provocando inestabilidad en los ARN mensajero (ARNm) y los ARN ribosómicos (ARNr)<sup>19</sup>.

En la Figura 6 se resume el modo en que los análogos de nucleósidos ejercen sus actividades citotóxicas y que se puede resumir planteando que este sigue dos mecanismos generales. El primer mecanismo de acción (**Mecanismo 1**, Figura 6) comienza cuando el análogo entra en la célula y está trifosforilado. Una polimerasa

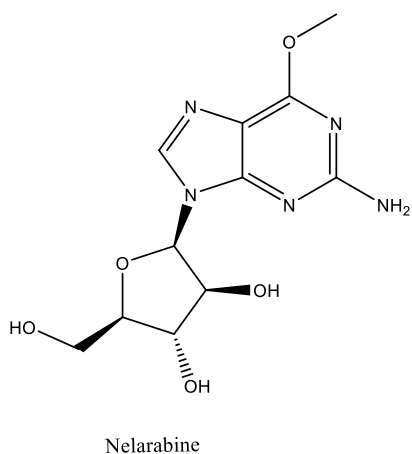
celular utiliza posteriormente el análogo como sustrato e incorpora el mismo en una hebra creciente de ADN o ARN. En el segundo mecanismo de acción (**Mecanismo 2**, Figura 6) el análogo actúa como inhibidor competitivo de enzimas esenciales, por lo general, sobre aquellos involucrados en la síntesis de nucleósidos/nucleótidos. La inhibición de estas enzimas interrumpe la actividad intracelular.

### E. Análogos nucleosídicos antileucémicos.

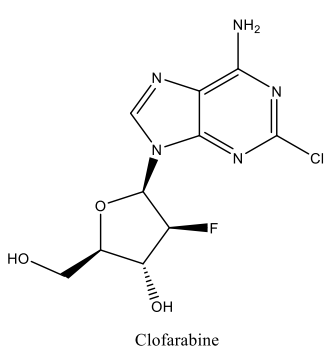
#### - Análogos purinicos:



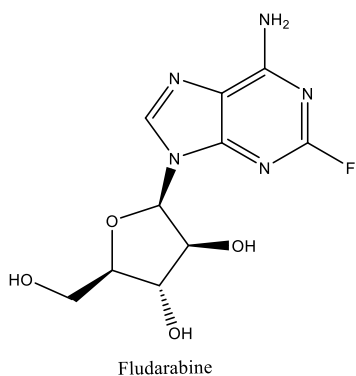
**Cladribina (2-clorodesoxiadenosina, 2-CdA, Leustatin™)** Es un análogo de desoxiadenosina en el que el átomo de hidrógeno, en la posición 2 del anillo de Purina, se ha sustituido por un átomo de cloro<sup>20</sup>. Esta sustitución hace que la cladribina sea resistente a la desaminación por la enzima adenosina desaminasa. La cladribina, es un profármaco, que se convierte en 2-cloro-2'-desoxiadenosina-5'-trifosfato, por la acción de la enzima desoxicitidina quinasa (DCK).



**Nelarabina (9-β-d-arabinofuranosilguanina, Ara-G, Arranon, Atriance).** Es un análogo de guanosina en el que el hidrógeno, en la posición 6 del anillo, está sustituido por un grupo metoxi<sup>21</sup>. Es un profármaco del análogo de desoxiguanosina, ara-G, que es desmetilado y fosforilado intracelularmente hasta ser convertido a la forma activa 5'-trifosfato, ara-GTP.

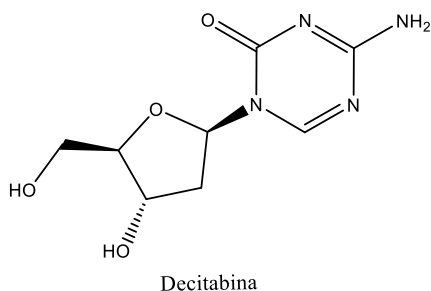


**Clofarabina (9-(2-desoxi-2-fluoro-β-D-arabinofuranosil)-2-cloroadenina, CAFdA, Clolar, Evoltra).** Es el análogo de purina más efectivo en pacientes con leucemia aguda<sup>22, 23</sup>.

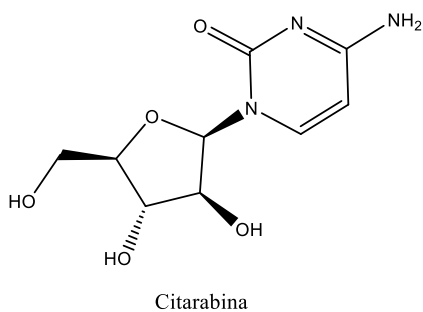


**Fludarabina (9-β-d-arabinosil-2-fluoroadenina, Ara-U, Fludara).** Es un análogo de la deoxiadenosina que se administra a los pacientes como profármaco en forma de fludarabina 5'-monofosfato, que luego entra en las células y se convierte en trifosforilado por desoxicitidina quinasa (dCK) o desoxiguanina quinasa (dGK) para ejercer sus acciones inhibitoras<sup>24</sup>.

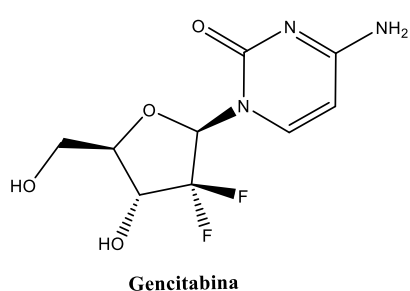
- **Análogos pirimidínicos.**



**Decitabina (5-aza-2'-desoxicitidine, 5-aza-dCyd, Dacogen)** Es un inhibidor de la metiltransferasa. De hecho, en la actualidad, es el inhibidor de la metilación del ADN más utilizado en el tratamiento de células cancerosas<sup>25</sup>.



**Citarabina (1-β-d-arabinofuranosilcitosina, Ara-C, Cytosar-U).** La citarabina es un fármaco antineoplásico análogo de la pirimidina, que inhibe la síntesis de ácido desoxirribonucleico<sup>26</sup>. La acción principal de la citarabina consiste en la inhibición de la síntesis de deoxicitidina, aunque la inhibición de las quinasas citidílicas y la incorporación del compuesto en ácidos nucleicos también podrían desempeñar un papel en sus acciones citostáticas y citocidas.



**Gemcitabina (2',2'-difluoro 2'-desoxicitidina, dFdC, Gemzar)** Es un análogo del arabinósido de citocina (Ara-C), con potente actividad antitumoral. Penetra en la membrana a través de un mecanismo de

transporte de nucleósidos y es trifosforilado (dFdCTP). Este metabolito trifosforilado compete con el desoxicidina trifosfato (dCTP) y se comporta como un inhibidor de la ADN polimerasa. El dFdCDP es además un potente inhibidor de ribonucleósido reductasa, lo cual conduce al agotamiento de grupos de desoxirribonucleótidos necesarios para la síntesis de ADN, potenciando así los efectos de dFdCTP. El dFdCTP se incorpora en el ADN, y luego de de la incorporación de un nucleótido extra, conduce a la terminación de la cadena de ADN<sup>27</sup>.

La historia de estos análogos nucleosídicos nos lleva hasta finales de la década de los 50' cuando la citarabina (**ara-C**, arabinósido de citosina) fue sintetizada. El éxito de ara-C en el tratamiento de leucemias impulsó el desarrollo de los otros análogos de nucleósidos, purínicos y pirimidínicos, entre los que destacaron: **ara-A**, fludarabina (**F-ara-A**, 2-fluoro-arabinosiladenina) y cladribina (**2CdA**, 2-cloro-2'- desoxiadenosina), que a principios de los 80', entraban en la fase de ensayos clínicos.

La presencia de un átomo de cloro en el anillo purínico de la Cladribina mejora la estabilidad del nucleósido frente a la adenosina desaminasa, responsable de la degradación de la adenosina<sup>28</sup>. En la fludarabina, la sustitución por un átomo de flúor, en la misma posición, tiene un efecto similar, pero la presencia de cloro en esa posición es la preferida, dado los severos efectos tóxicos del producto de hidrólisis, 2-fluoroadenina<sup>29</sup>.

Aunque las actividades de **F-ara-A** y de **2CdA** han sido muy prometedoras en la terapia contra el cáncer, existen algunas limitaciones para su uso generalizado. Por ejemplo, sus bajas biodisponibilidades cuando se administran vía oral, debido a la inestabilidad de estos fármacos en solución ácida (tracto gastrointestinal) y sus susceptibilidades a la degradación bacteriana<sup>30</sup>. Además, altas dosis de Cladribina y de Fludarabina resultan a menudo en toxicidad neurológica grave<sup>31</sup>.

Teniendo en cuenta estos resultados, a finales de los años 80' en el Instituto de Investigación *Birmingham* (Alabama) se diseñó un nuevo análogo de 2'-desoxiadenosina, la **Clofarabina** (**Cl-F-ara-A**, 2-cloro-2'-fluoro-2'-desoxiarabinosiladenina)<sup>32</sup>. El nuevo fármaco combinaba las farmacocinéticas más ventajosas de los dos prototipos anteriores y presentaba una actividad antitumoral en

dosis relativamente bajas. Estas características condujeron a la realización de los ensayos clínicos con el fármaco<sup>29</sup>, que culminaron quince años después, en 2004, con la aprobación de la FDA para el tratamiento de leucemia aguda en pacientes pediátricos (1-21 años). El fármaco, licenciado a *Genzyme Corporation* y manufacturado bajo el nombre comercial *Clolar<sup>TM</sup>*, fue aprobado en 2006 en Europa, para el tratamiento de la leucemia linfoblástica aguda infantil.

## 1.2. Síntesis de análogos de nucleósidos.

### 1.2.1. Síntesis química.

Para la síntesis química de nucleósidos, normalmente se sigue un enfoque convergente, lo que significa que la base heterocíclica y el azúcar se preparan de forma independiente y luego son acoplados. Un método comúnmente utilizado es la reacción de Hilbert-Johnson (o Vorbrüggen), en el que una base heterocíclica sililada (nucleófilo) reacciona con el azúcar protegido (electrófilo) en presencia de un ácido de Lewis (normalmente,  $\text{SnCl}_4$  y  $\text{TMSOTf}$ )<sup>33</sup>, aunque existen más posibilidades<sup>34</sup>.

El método Vorbrüggen y otros similares han sido ampliamente aplicados para la síntesis de nucleósidos, pero a menudo han sido cuestionados por los requerimientos de *regio-* y *estéreo-*especificidad necesarios para la actividad biológica de los nucleósidos. En los nucleósidos naturales el azúcar está conectado al átomo **N9** de las bases purínicas o al **N1** de las bases pirimidínicas. Sin embargo, la presencia de múltiples sitios nucleofílicos en las bases heterocíclicas, plantean dificultades para la formación regioselectiva del enlace glicosídico durante la síntesis química. Por otro lado, los nucleósidos naturales se encuentran, exclusivamente, en la configuración anomérica  $\beta$ , en la que la base está orientada por encima del plano del azúcar. Este requisito se puede lograr fácilmente en la síntesis de ribonucleósidos, a través de la formación de un intermediario catiónico cíclico de la ribosa, antes de la formación del enlace glicosídico<sup>35</sup>. Pero, en el caso de la síntesis de desoxirribonucleósidos la situación es más complicada, debido a la ausencia de grupo hidroxilo en posición **2'**.

Como se mostrará en el siguiente apartado, utilizando la síntesis de 2-cloro-2'-fluoro-2'-desoxiarabinosiladenina como ejemplo, la síntesis química de análogos de nucleósidos

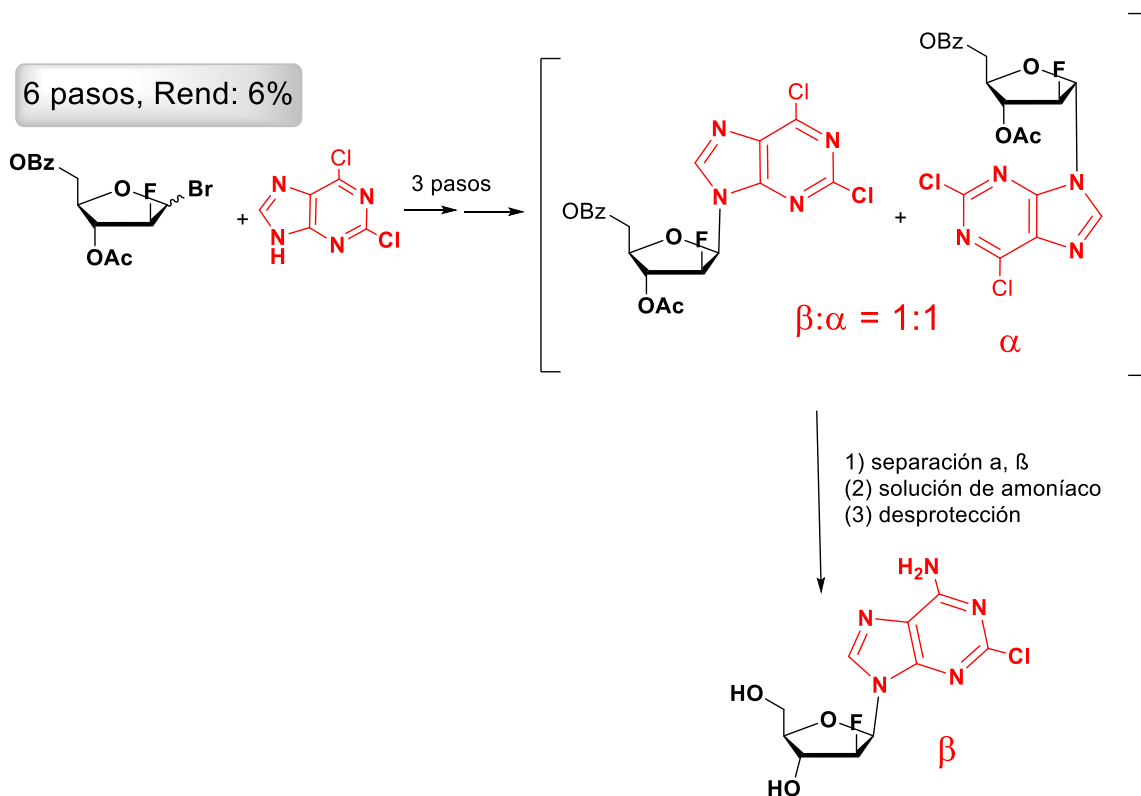
es, por lo general, un procedimiento con varios pasos, que requiere protección y desprotección de grupos funcionales y procedimientos sofisticados para conseguir *regio-* y *estereoselectividad*.

#### 1.2.1.1. Síntesis de clofarabina

Se han establecido varios métodos para la producción de clofarabina y en ellos las rutas de síntesis se basan típicamente en reacciones de acoplamiento entre los derivados de purina y un fragmento arabinofuranosilo. Los dos principales esquemas de síntesis tienen que ver con el método utilizado para introducir el átomo de flúor y comprenden aquellas metodologías en las que: **(i)** el átomo de flúor está previamente en el azúcar (ribofuranosa) y se introduce en la molécula final tras el paso de acoplamiento; **(ii)** el átomo de flúor se introduce luego de la etapa de acoplamiento, mediante transformaciones químicas del nucleósido obtenido previamente.

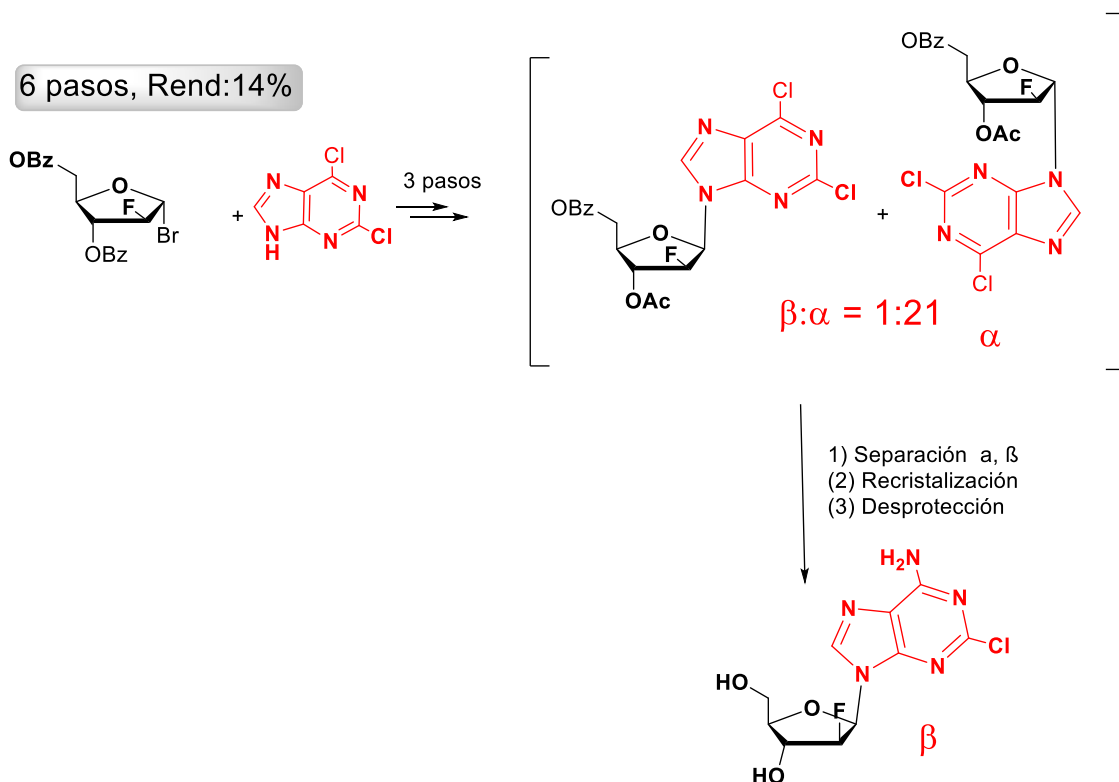
En la patente norteamericana *US Patent 5034518* se describe el esquema del proceso original para la síntesis de clofarabina, descrito por Montgomery y colaboradores<sup>36</sup>. En la ruta de síntesis (Figura 7) se utiliza bromuro de 3-O-acetil-5-O-benzoil-2-desoxi-2-fluoro-D-arabinofuranosilo para el acoplamiento con 2,6-dicloropurina, seguido de una secuencia de aminación/desprotección<sup>32</sup>. Esta etapa de acoplamiento da como resultado una mezcla de nucleósidos del que se obtiene el intermediario anomérico deseado, el 9- $\beta$ , a un rendimiento de tan sólo 32%, seguida de una complicada etapa de separación cromatográfica. El rendimiento global de la reacción (6%) es, de hecho, la mayor desventaja de este método, además de la estabilidad del bromuro de 3-O-acetil-5-O-benzoil-2-desoxi-2-fluoro-D-arabinofuranosilo y de la dificultad de separar la mezcla de isómeros en la etapa de purificación.





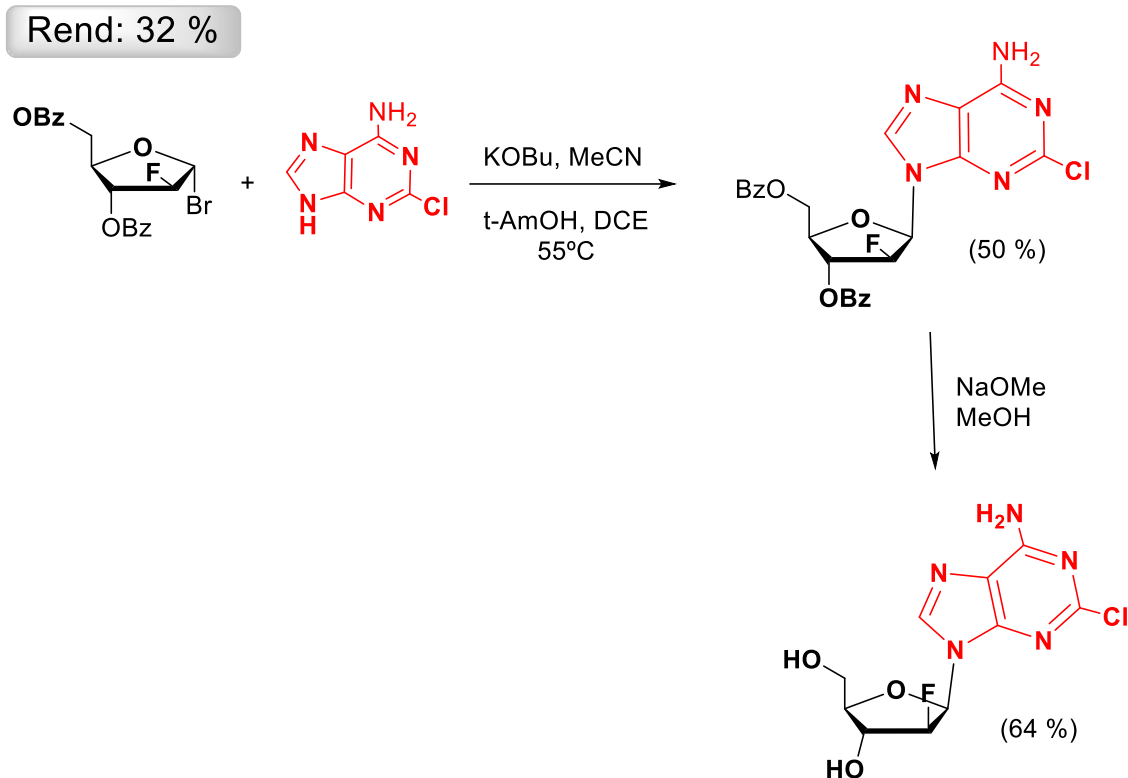
**Figura 7:** Síntesis de clofarabina. Método Montgomery<sup>36</sup>.

El proceso resumido en la Figura 7 se ha venido usando desde entonces, con muy pocas modificaciones, para la síntesis de clofarabina. En el 2004, Bauta y col.<sup>37, 38</sup> propusieron una nueva metodología para la síntesis de clofarabina (Figura 8), el denominado proceso **ILEX** (*ILEX Products Inc.*, San Antonio, TX, USA). El método utiliza como material de partida el anómero ( $\alpha$ -anómero), estereoquímicamente puro, del bromuro de 2-desoxi-2-fluoro-D-arabinofuranosilo, obtenido a partir de la fluorinación química de 1-O-acetil-2,3,5-tri-O-benzoil- $\alpha$ -D-ribofuranosa<sup>39, 40</sup>. El paso de conjugación, para formar el nucleósido, se lleva a cabo a partir de una 2-cloro-6-amino purina pre-formada, con una excelente estereoselectividad ( $\beta/\alpha = 21/1$ ). El proceso de 6 etapas consigue un rendimiento global de alrededor del 14%<sup>37</sup>.



**Figura 8:** Síntesis de clofarabina. Método ILEX<sup>37</sup>.

Aunque el proceso **ILEX** es el proceso utilizado para la producción a escala de la clofarabina, Bauta y col.<sup>37</sup> han propuesto mejoras que consiguen obtener rendimientos globales de entre 28-32%, sin necesidad de etapas de purificaciones mediadas por cromatografía. La Figura 9 resume este proceso, en el cual en la primera etapa se hace reaccionar 2-cloroadenina con el azúcar protegido (1-bromo-2-desoxi-2-fluoro-3,5-di-O-benzoyl- $\alpha$ -D-ribofuranosa) y *tert*-butóxido de potasio, en una mezcla de tres solventes. Luego en una segunda etapa se desprotege y recristaliza la clofarabina obtenida, con menos de un 0,1% del anómero  $\alpha$ .



**Figura 9:** Síntesis de clofarabina propuesta por Bauta y col.<sup>37</sup>. KOBu: *tert*-butóxido de potasio; MeCN: acetonitrilo ; t-AmOH: alcohol *tert*-amilico ; DCE: dicloroetano.

### 1.3. Química verde y Biocatálisis.

Los recursos de la Tierra se están agotando continuamente para satisfacer las necesidades de la creciente población y, por otra parte, las actividades humanas han dado lugar a una enorme carga sobre el medio ambiente, por los desechos tóxicos y las emisiones de gases de efecto invernadero y el consiguiente cambio climático. En este contexto, el establecimiento de una era de **Bioeconomía**<sup>♦</sup>, requiere no sólo un cambio de patrón de producción, sino también una profunda modernización de los procesos de producción donde, por ejemplo, los subproductos puedan ser utilizados como materia prima en la producción de productos básicos como los biocombustibles y otros productos químicos de alto valor añadido.

<sup>♦</sup> La bioeconomía es la ciencia de la gestión de la sustentabilidad. Su finalidad es alcanzar un desarrollo socio-económico sostenible, a través de un uso eficiente de los recursos naturales. (<http://www.fundacionmelior.org/content/tema/que-es-la-bioeconomia-o-economia-ecologica>)

En 1991 Paul Anastas y John Warner<sup>41</sup> formularon el concepto de química verde y un plan para guiar a la industria química hacia prácticas menos peligrosas, que ha sido bien recibido por químicos, industriales y políticos. El creciente número de publicaciones y proyectos de investigación que tratan de desarrollar nuevas vías (más verdes) para la síntesis de productos químicos y la creación de organismos, redes, instituciones, revistas y programas educativos para la consecución de sus objetivos, son una clara indicación de ello.

La iniciativa es una propuesta novedosa para reducir y/o eliminar los problemas ambientales derivados de actividades industriales. Según la *US Environmental Protection Agency* (EPA), la **química verde** es el “uso de la química para la prevención de la contaminación, y el diseño de productos químicos y procesos benévolos para el ambiente”<sup>42</sup>. La **química verde** plantea **12 principios** para conseguir sus objetivos\*.

### **1. Prevención**

Es mejor evitar la formación de residuos que tratarlos o limpiarlos después de que se hayan formado.

### **2. Economía de átomos**

Los métodos de síntesis deberían diseñarse para maximizar la incorporación de todos los materiales utilizados en el producto final.

### **3. Productos químicos intermedios menos tóxicos**

Siempre que sea posible, las metodologías sintéticas deberían diseñarse para usar y generar sustancias con poca o ninguna toxicidad para la salud humana y el medio ambiente.

### **4. Productos finales más seguros**

Los productos químicos deberían diseñarse para mantener la eficacia de su función, a la vez que reducen su toxicidad.

---

\* <https://campushuesca.unizar.es/blog/los-doce-principios-de-la-quimica-sostenible>

## **5. Reducción del uso de sustancias auxiliares**

El empleo de sustancias auxiliares (como disolventes, agentes de separación, etc.) debería evitarse en lo posible, y ser inocuo cuando se empleen.

## **6. Reducción del consumo energético**

El impacto medioambiental y económico de los requerimientos energéticos debe ser reconocido y minimizado. Los métodos sintéticos deberían aplicarse a presión y temperatura ambiente.

## **7. Uso de materias primas renovables**

Deben utilizarse materias primas renovables siempre que sea técnica y económicamente viable.

## **8. Reducción de la derivatización innecesaria**

La derivatización innecesaria (grupos bloqueadores, etapas de protección/desprotección, modificaciones temporales) debe evitarse en la medida de lo posible.

## **9. Uso de catalizadores**

Los reactivos catalíticos (tan selectivos como sea posible) son superiores a los reactivos estequiométricos.

## **10. Diseño para la degradación**

Los productos químicos deben diseñarse de forma que al final de su función no persistan en el entorno y se degraden en productos inocuos.

## **11. Desarrollo de tecnologías analíticas para la monitorización en tiempo real**

Deben desarrollarse metodologías analíticas que permitan el seguimiento y control de procesos en tiempo real, antes de que se formen sustancias peligrosas.

## **12. Minimización del riesgo de accidentes químicos**

Las sustancias químicas y las formas en que se usan dichas sustancias en un proceso químico deben escogerse para minimizar el potencial de accidentes químicos, incluyendo vertidos, explosiones e incendios.

Las reacciones catalizadas por enzimas cumplen con la mayoría de los criterios de la química verde, ofreciendo más de una solución para cada problemática. Así, por ejemplo, los biocatalizadores en lo referente a la generación de residuos, contribuyen a la creación de procesos, más eficientes y con menos consumo energético.

A favor de los biocatalizadores podríamos mencionar numerosas ventajas: pueden actuar como catalizadores no tóxicos; generalmente son muy selectivos, consiguiendo productos con una alta pureza; operan en condiciones de reacción moderadas (ej., a temperatura ambiente), lo que da como resultado un consumo de energía reducido; el medio de reacción es por lo general acuoso, que «*per se*» es considerado como no tóxico; el uso de biocatalizadores evita el uso de metales y disolventes orgánicos; como catalizadores naturales, las enzimas pueden ser consideradas como renovables<sup>43</sup>.

La Biocatálisis puede considerarse como una herramienta atractiva para la síntesis orgánica, sin embargo, persiste cierto recelo de muchos químicos en contra del uso de las enzimas en síntesis química. Muchas veces, estas reticencias tienen que ver con las limitaciones (*percibidas o reales*) de la biocatálisis. Algunos de estos mitos, revisados por Wubbolts y cols.<sup>44</sup>, refieren que:

**L**as enzimas son caras. Es uno de los mitos más extendidos. Sin embargo, Turvfesson y cols.<sup>45, 46</sup> han demostrado que el precio real de un biocatalizador depende de (i) la economía de escala de producción y (ii) la pureza y formulación de la enzima utilizada. Así, en fermentaciones a gran escala, se pueden obtener biocatalizadores a precios por debajo de los **200 euros por kg de biocatalizador**. Hay otras consideraciones que podrían influir en el coste de los biocatalizadores, por ejemplo, si son necesarias preparaciones altamente purificadas o inmovilizadas. En la Tabla 4 se muestran sugerencias, sobre la base del análisis económico de Turvfesson y col.<sup>45</sup>, para estimar la viabilidad económica de un proceso biocatalítico.

**L**as enzimas son ineficientes. Si medimos la "eficiencia" de los biocatalizadores, en términos de eficiencia catalítica, las enzimas son excelentes catalizadores. Pero si medimos la "eficiencia" en función de la calidad y la cantidad de producto, los biocatalizadores ganan en calidad, entre otras razones, por la selectividad intrínseca de muchos enzimas, lo cual permite bioconversiones limpias y altamente selectivas. Pero, el criterio basado en la cantidad de producto, es el "*punto débil*" de la mayoría de los biocatalizadores. Muchas veces esto tiene que ver con la solubilidad de los sustratos, lo

cual hace que estas reacciones biocatalíticas se lleven a cabo usando mezclas muy diluidas. En consecuencia, las cantidades de productos a menudo tienden a ser bajas, lo que requiere un proceso de recuperación/purificación extensivo y un uso ineficiente del espacio del reactor.

**Tabla 4:** Requisitos para procesos enzimáticos viables.

	Precio del producto (Euro/ kg Producto)	Contribución del catalizador al coste (Euro/ kg Producto)
Fármaco	>100	10
Química fina	>15	1,5
Especialidades químicas	5	0,25
Productos químicos a granel	1	0,05

**L**as enzimas son altamente específicas «solo» con sus sustratos naturales. En realidad esto es cierto para los enzimas que participan en rutas anabólicas♦, en los que hay un número muy limitado de sustratos (y muy similares, desde el punto de vista químico), por lo que su uso sobre sustratos o reacciones diferentes a las naturales es muy difícil o inviable. Otras enzimas, implicadas en vías catabólicas o de desintoxicación, pueden ser más versátiles, en lo referido al rango de sustratos («promiscuidad catalítica») manteniendo una alta (*estéreo-*) especificidad, lo que les convierte en catalizadores muy útiles.

**L**a biocatálisis es intrínsecamente "verde". Esto, más que un mito, es una tendencia errónea, en la cual para muchos autores, "biocatálisis" y "química verde" son sinónimos. En esto influyen argumentos como: «condiciones suaves de reacción»; «naturaleza renovable y biodegradable de los catalizadores»; «el uso del agua como disolvente», entre otros. Sin embargo, en muchas ocasiones nos encontramos que muchos procesos biocatalíticos son cuantitativamente menos “verdes” que los procesos (químicos) de referencia en términos, por ejemplo, de residuos generados en la

♦ Las *rutas anabólicas* son divergentes (se forman muchos productos finales diferentes a partir de una pequeña variedad de precursores).

producción de un compuesto dado (**Factor E<sup>♦</sup>**), de disolventes empleados y los auxiliares de producción (por ejemplo, sistemas de filtración, etc.).

En general, las reacciones biocatalíticas tienen un enorme potencial y son una alternativa más ecológica a muchos de los procesos tradicionales de síntesis. Sin embargo, no se puede hacer una declaración ecológica general para la biocatálisis y el impacto ambiental debe evaluarse caso por caso, en lo que respecta a su amabilidad ambiental y viabilidad económica. La Tabla 5 resume algunas de las consideraciones y observaciones importantes sobre el uso de biocatalizadores en química verde.

**Tabla 5:** Observaciones importantes sobre biocatálisis en química verde<sup>47</sup>.

<b>Observación</b>	<b>Análisis crítico</b>
1. Uso de agua como solvente verde	Algunos compuestos orgánicos presentan baja solubilidad en agua; La recuperación de productos a partir de soluciones acuosas, a menudo requiere de la extracción con disolventes orgánicos.
2. Posibilidades de la ingeniería enzimática	Área muy prometedora que, sin embargo, requiere mucho tiempo para su desarrollo desde la idea hasta la implementación a escala industrial.
3. Productividad	Considerando como aceptables, una productividad volumétrica mínima de <b>0,1 gL<sup>-1</sup>h<sup>-1</sup></b> y una concentración final mínima de producto de <b>1 gL<sup>-1</sup></b> , para su implementación a escala industrial, la optimización del proceso en términos de aumento de la concentración del sustrato y el modo de alimentación necesitan de la estabilización de los biocatalizadores
4. Bajas concentraciones de sustratos.	Debido a problemas de inhibición enzimática, se utilizan, por lo general, concentraciones bajas de sustrato dando como resultado reactores sobredimensionados y procesamientos ineficientes de los resultados de la reacción biocatalítica.

<sup>♦</sup> **Factor E:** Relación del peso de los desechos generados en una síntesis respecto al peso del producto final.



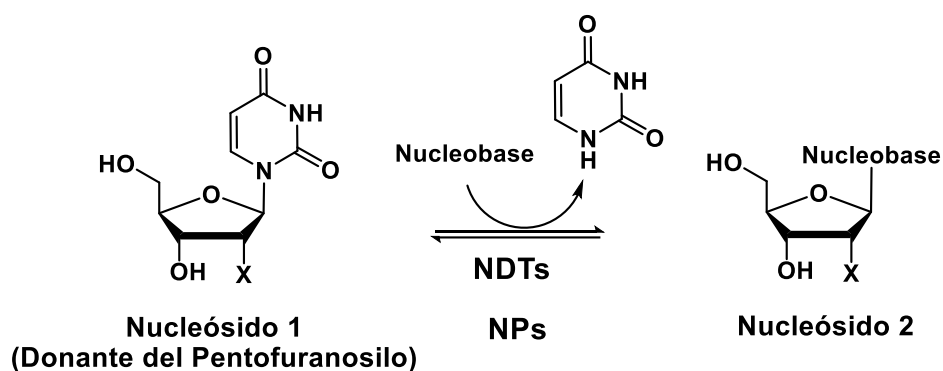
Observación	Análisis crítico
5. Potencial como proceso alternativo de procesamiento	Aunque los procesos biocatalíticos son a menudo más ecológicos que los químicos, para la industria, las razones ecológicas no son los únicos temas a tratar para reemplazar un proceso existente. Por otro lado, a veces no hay alternativas químicas a una vía biotecnológica.
6. Desarrollo farmacéutico	La combinación de pasos químicos y biocatalíticos es el camino más prometedor para algunos productos; Para la obtención de moléculas quirales, si la separación de racematos es compleja y/o no es fiable, deben usarse biotransformaciones enantioselectivas.
7. Precio del catalizador	La contribución de los costes del biocatalizador está muy relacionada al valor de los productos. Pueden variar desde 0,05 € kg <sup>-1</sup> producto (productos químicos a granel) hasta 10 € kg <sup>-1</sup> producto (productos farmacéuticos) <sup>45, 46</sup> .
8. Recuperación/purificación	Las soluciones acuosas, comúnmente usadas en biotransformaciones, requieren una cantidad significativa de disolventes para el aislamiento del producto; Estrategias como la eliminación <i>in situ</i> de productos y la ingeniería de enzimas para mejorar su tolerancia a disolventes podría superar este problema.
9. Uso de líquidos iónicos	Los fluidos funcionales a menudo mejoran la solubilidad del sustrato, pero incurren en gastos adicionales en el procesamiento posterior. Se necesita más información sobre su toxicidad.
10. Espectro de sustratos para biocatálisis	Aunque se dice que la especificidad es una ventaja de las enzimas sobre los catalizadores químicos, algunos biocatalizadores, tales como las lipasas, presentan versatilidad de sustrato y diversas funciones catalíticas.

### 1.3.1. *Síntesis enzimática de análogos de nucleósidos.*

La biocatálisis ofrece la posibilidad de llevar a cabo reacciones *enantio-*, *regio-* y/o *químico-*selectivas, que podrían reducir el número de etapas de reacción, la cantidad de subproductos, etc., y, de este modo, minimizar los costes de síntesis de bioproductos<sup>48</sup>. Aunque la biocatálisis puede utilizarse de muchas maneras, en los procesos de síntesis de análogos de nucleósidos, es en el área de glicosilación o de la transglicosilación en la que se ha aplicado con más éxito.

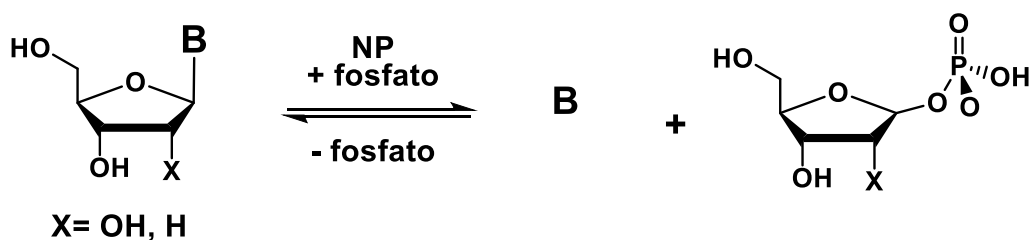
De particular interés en la producción biocatalítica de análogos de nucleósidos son 2 clases de enzimas: **nucleósidos fosforilasas (NP)** y **N-desoxirribosiltransferasas (NdRT)**. Ambas enzimas catalizan la transferencia de un residuo glicosídico a partir de un nucleósido donante a una base heterocíclica. Así, usando diferentes combinaciones de estas enzimas, es posible transferir azúcares de ribosa o desoxirribosa entre bases purínicas y pirimidínicas, así como entre bases de pirimidina o purina. A la vez que es posible transferir azúcares modificados entre bases nucleosídicas o nucleósidos modificados, dependiendo de la especificidad de la enzima utilizada<sup>49</sup>.

La reacción global de ambas enzimas es esencialmente la misma (Figura 10), es decir ambas enzimas catalizan el intercambio de la base de un nucleósido. No obstante, el mecanismo catalítico de la reacción de transglicosilación mediada por **NPs** y **NdRTs** es diferente. Las **NdRTs** (EC 2.4.2.6) catalizan la transferencia directa de un residuo desoxirribofuranosilo, mientras que con las **NPs** (EC 2.4.2.1, EC 2.4.2.2, EC 2.4.2.4 y EC 2.4.2.28) se forma el producto intermediario  $\alpha$ -D-pentofuranosil-1-fosfato, de modo que la reacción catalizada NPs requiere la presencia de fosfato inorgánico. Además, las NPs y NdRTs difieren en sus especificidades de sustrato, ambos ribo- y desoxirribonucleósidos son sustratos naturales de las nucleósido fosforilasas, en cambio, las NdRT son específicas para 2'-desoxirribonucleósidos<sup>50, 51</sup>.



**Figura 10:** Esquema general de la reacción de transglicosilación. En el ejemplo el nucleósido 1 actúa como donante del grupo pentofuranosilo y una nucleobase como aceptor. La reacción puede ser catalizada por una NdRT o por una o dos NPs.

#### 1.3.1.1. Reacciones de transglicosilación catalizadas por nucleósido fosforilasas.



**Figura 11:** Esquema de la reacción catalizada por NPs. **B:** base heterocíclica.

Las reacciones de transglicosilación mediadas por enzimas de la familia de las **NPs** (Tabla 6) tienen lugar en dos etapas consecutivas: en la primera un nucleósido, que actúa como donante del grupo pentofuranosilo, es escindido fosforolíticamente rindiendo *ribosa-1- $\alpha$ -fosfato* o *2'-desoxirribosa-1- $\alpha$ -fosfato*, como intermediario reactivo, y la base heterocíclica correspondiente (Figura 11); y en la segunda el intermediario reactivo se une a otra base heterocíclica que actúa como aceptor de del grupo pentofuranosilo.

**Tabla 6:** Nucleósido fosforilasas.

Nombre de la enzima	Acrónimo	Número EC <sup>♦</sup>
Purin nucleósido fosforilasa	<b>PNP</b>	2.4.2.1
Pirimidin nucleósido fosforilasa	<b>PyNP</b>	2.4.2.2
Uridina fosforilasa	<b>UP</b>	2.4.2.3
Timidina fosforilasa	<b>TP</b>	2.4.2.4
5' -Metiltioadenosina fosforilasa	<b>MTAP</b>	2.4.2.28

Algunos autores han realizado clasificaciones de las **NPs** basadas en diferentes parámetros, de las cuales, la más aceptada es la propuesta por Pugmire y Ealick, elaborada a partir de análisis estructurales<sup>52</sup>:



**NPs tipo I:** Este tipo incluye a las purín-nucleósido fosforilasas (**PNPs**, **EC 2.4.2.1**), específicas de purinas, y a las uridín fosforilasas (**UPs**, **EC 2.4.2.3**), específicas de uridina y 2'-desoxiuridina. Las configuraciones oligoméricas más comunes observadas en estas enzimas son: **a)** PNP's homotriméricas específicas de 6-oxopurinas (guanina e hipoxantina); **b)** PNP's homohexaméricas que aceptan tanto 6-aminopurinas (adenina) como 6-oxopurinas; **c)** UP's homohexaméricas.

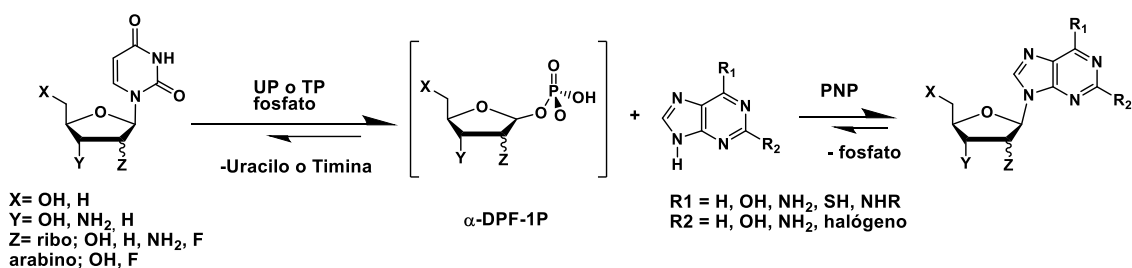


**NPs tipo II:** Todas las **NPs** incluidas en este tipo son específicas de pirimidinas, y se dividen en uridín fosforilasas (**UPs**, **EC 2.4.2.3**), timidín fosforilasas (**TPs** **EC 2.4.2.4**) y pirimidín-nucleósido fosforilasas (**PyNP's**, **EC 2.4.2.2**), que reconocen indistintamente uracilo y timina.

La estrategia de síntesis más explotada, usando **NPs** incluye la transferencia de un resto de pentofuranosa desde nucleósidos de pirimidina a bases de purina (Figura 12).

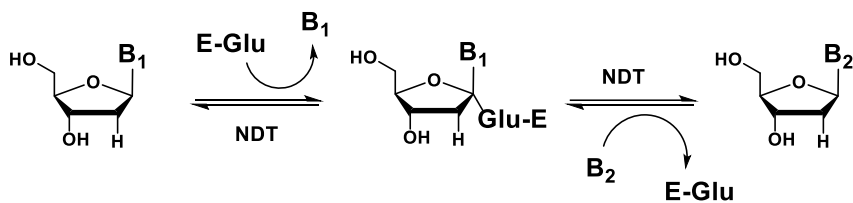
<sup>♦</sup> El nombre de cada enzima puede ser identificado por un código numérico, encabezado por las letras **EC** (*enzyme commission*), seguidas de **cuatro números separados por puntos**. El primer número indica a cual de las seis clases pertenece el enzima, el segundo se refiere a distintas subclases dentro de cada grupo, el tercero y el cuarto se refieren a los grupos químicos específicos que intervienen en la reacción. <http://www.biorom.uma.es>

Desafortunadamente, un enfoque similar no es práctico para la producción de nucleósidos de purinicos. Además, las distinciones en la especificidad por los sustratos de la **TP** y **UP**, supone una mayor diversidad de nucleósidos donantes de pentofurano, para la reacción de transglicosilación. En el esquema de la Figura 11, por ejemplo, si el nucleósido es: 1-(2-desoxi-2-fluoro-β-D-ribofuranosil) uracilo (es decir, **X = Y = OH, Z = ribo F**) o es: 1-(2-desoxi-2-fluoro-arabinofuranosil) timina (es decir, **X = Y = OH, Z = arabino F**) no serían sustratos de **UP**, por lo tanto esta enzima no puede ser empleada como biocatalizador en esta reacción. Sin embargo, ambos nucleósidos, pueden ser sustratos de **TP**, lo cual es posible el uso de **TP** en reacciones de transglicosilación de bases purinicas<sup>53,54</sup>.



**Figura 12:** Esquema general de la reacción de transglicosilación. En el ejemplo el nucleósido. **UP** y **TP**: uridina y timidina fosforilasa, respectivamente; **PNP**: purin nucleósido fosforilasa.

### 1.3.1.2. Reacciones de transglicosilación catalizadas por Nucleósido 2'-desoxirribosiltransferasas.



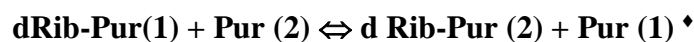
**Figura 13:** Reacción de transglicosilación catalizada por las nucleósido 2'-desoxirribosiltransferasas. **B1**: purina o pirimidina; **B2**: purina o pirimidina.

Las nucleósido-desoxirribosiltransferasas, o **NdRTs** (**EC 2.4.2.6**) son enzimas que catalizan el intercambio de bases entre 2'-desoxirribosilnucleósidos (Figura 13). La reacción se lleva a cabo a través de la formación de un intermedio covalente en el que el hidroxilo glicosídico de la 2'-desoxirribosa es esterificado por el ácido glutámico del sitio activo de las nucleósido 2'-desoxirribosiltransferasas<sup>55</sup>. Esta enzima fue descrita por primera vez por MacNutt en *Lactobacillus helveticus*<sup>56</sup> y posteriormente en otras

especies del mismo género, como *L. leichmanii*<sup>57, 58</sup>, *L. helveticus*<sup>59</sup>, *L. fermentum*<sup>60</sup>, *L. reuteri*<sup>61</sup> y *L. animalis*<sup>62</sup> y están implicadas, sobre todo en bacterias lácticas, en la ruta de salvamento de purinas para la síntesis de ADN, aunque esto no está claro en el caso *Lactococcus lactis*<sup>63</sup>. Se han encontrado **NDTs** también en algunas especies de *Bacillus psychrosaccharolyticus*<sup>55, 64</sup>, y organismos parasitarios unicelulares eucariotas como *Crithidia luciliae*<sup>65</sup> y *Borrelia burgdorferi*<sup>66</sup>.

Las nucleósido 2'-desoxirribosiltransferasas, también denominados trans-N-desoxirribosilasas (**NDT**) exhiben una especificidad estricta hacia 2'-desoxirribosa, regioselectivas (glicosilación a través del **N1** de pirimidinas y **N9** de purinas) y estereoselectivas (se forma únicamente el anómero  $\beta$ )<sup>67</sup>. Aunque son muy específicas, en lo referido a la 2'-desoxirribosa, se han observado diferencias en las especificidades hacia la base nitrogenada, lo que ha permitido su clasificación en dos subclases:

**NDTs tipo I:** son enzimas que catalizan la transferencia de las bases solo entre 2'-desoxirribonucleósidos purínicos y una purina<sup>59</sup> (comúnmente llamadas **NDTs tipo I**, o **PDTs**).



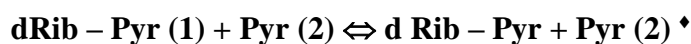
Con el siguiente orden de preferencia por el sustrato donante del grupo 2'-desoxirribosilo<sup>59</sup>: desoxiinosina (*dIno*) > desoxiadenosina (*dAdo*) > desoxiguanosina (*dGuo*).

**NDTs tipo II:** son las enzimas que catalizan la transferencia de bases entre 2'-desoxirribonucleósidos independientemente del tipo de base. Comúnmente llamadas **NDTs tipo II**, o **NDTs**<sup>68</sup>. Las enzimas pertenecientes a este grupo pueden llevar a cabo la transglicosilación entre nucleósidos y bases de citosina, a diferencia de la transglicosilación mediada por nucleósido fosforilasas, que no pueden hacerlo.




---

♦ **dRib** = 2'-desoxi-D-ribosa, **Pur** = base púrica y **Pyr** = base pirimidínica



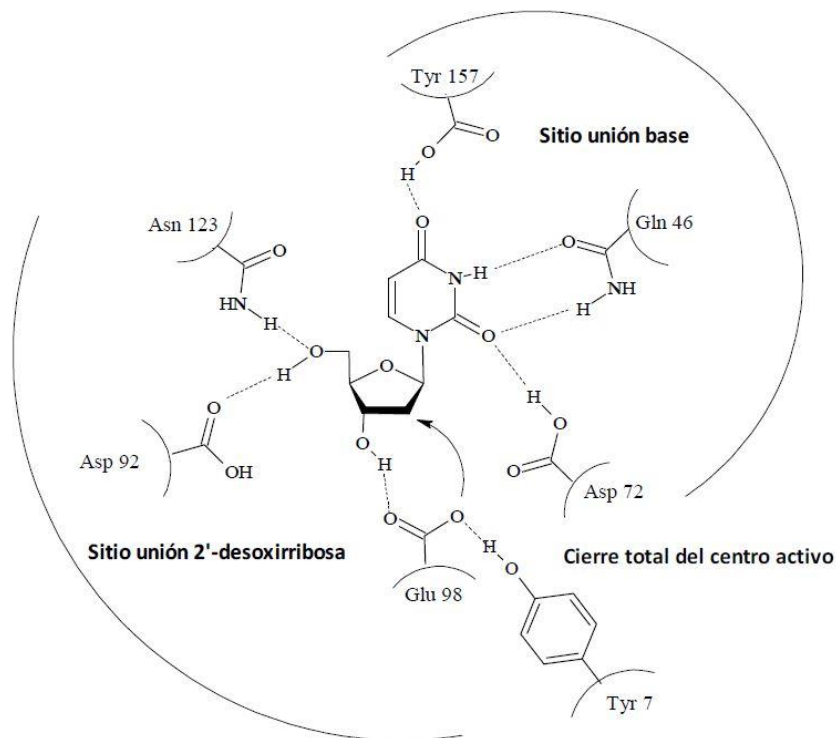
*Esta subclase de NDT tiene una fuerte preferencia por desoxipirimidinas como sustrato donante del grupo 2'-desoxirribosilo*<sup>69</sup>

El mecanismo catalítico descrito para las nucleósido 2'-desoxirribosiltransferasas generalmente es un mecanismo *ping-pong bi-bi*, en la que el producto resultante tiene la misma configuración anomérica del 2'-desoxirribosinucleósido de partida<sup>70, 71</sup>. Analizando la estructura de los centros activos de nucleósido 2'-desoxirribosiltransferasas cristalizadas, ejemplo: NDT de tipo II de *Lactobacillus leichmannii* (**PDB 1F8X**) o la NDT de tipo I de *Lactobacillus helveticus* (**PDB 1S2L**), se ha identificado al ácido glutámico como residuo esencial en la actividad (en el caso de *Lactobacillus leichmannii*, glutámico 98)<sup>69, 72</sup>. Además, estos estudios han revelado que las DRTs presentan dominios  $\alpha/\beta$  y, aunque hay un centro activo en cada subunidad, el sitio activo completo requiere las cadenas laterales de residuos situados en dos subunidades distintas orientadas de forma antiparalela.

Fresco-Taboada y col<sup>73</sup> indicaron que las DRTs, aunque aceptan distintas bases naturales y no naturales, son altamente específicas para la 2'-desoxirribosa, debido a que la interacción del **OH** en posición **2'** con el residuo catalítico de ácido glutámico (Glu) impide la catálisis. En la Figura 14 se muestra el centro activo de la 2'-desoxirribosiltransferasa II de *Lactobacillus leichmannii* (LINDT) en ella puede observarse que el sitio de unión de la 2'-desoxirribosa se alinea con tres residuos ácidos, Asp92, Glu98 y Asn123 de la LINDT, formando un núcleo hidrofílico. El **O5'** de la 2'-desoxirribosa forma puentes de hidrógeno con el carboxilo de la cadena lateral del Asp92 y el grupo amino de la cadena lateral de la Asn123. El **O3'** interacciona con uno de los oxígenos del carboxilato del Glu98, mientras que el otro actúa como nucleófilo que ataca al **C1'** de la 2'-desoxirribosa en una catálisis general ácido-base (Figura 14). Cuando el donador es un ribonucleósido, el oxígeno nucleófilo del Glu catalítico forma un enlace de hidrógeno con el **O2'** de la ribosa y la enzima es inactiva.<sup>55</sup>

---

$\spadesuit$  **dRib** = 2'-desoxi-D-ribosa, **Pur** = base púrica y **Pyr** = base pirimidínica.



**Figura 14:** Estructura del centro activo de la NDT de *LINDT* empleando 2'- desoxiuridina como sustrato<sup>55</sup>. La flecha indica el ataque nucleofílico del carboxilato del *Glu* catalítico al *C1'* de la 2'-desoxirribosa.

#### A. Ventajas de la Biocatálisis usando nucleósido 2'-desoxirribosiltransferasas.

La reacción de transglicosilación catalizada tanto por **NPs** como por **NDTs** es en esencia muy similar (Figura 10), es decir ambas enzimas catalizan la transferencia del azúcar de un nucleósido a una base heterocíclica. Pero, como ya se ha expuesto en apartados anteriores, el mecanismo catalítico de la reacción de transglicosilación mediada por **NPs** y **NdRTs** es diferente. Entonces, **¿qué ventajas ofrece el uso de NdRTs en la síntesis biocatalítica de análogos de nucleósidos?**

Las **NdRTs** catalizan la transferencia directa de un residuo desoxirribofuranosilo, mientras que con las **NPs** se forma el producto intermediario  $\alpha$ -D-pentofuranosil-1-fosfato, de modo que la reacción catalizada **NPs** requiere la presencia de fosfato inorgánico.



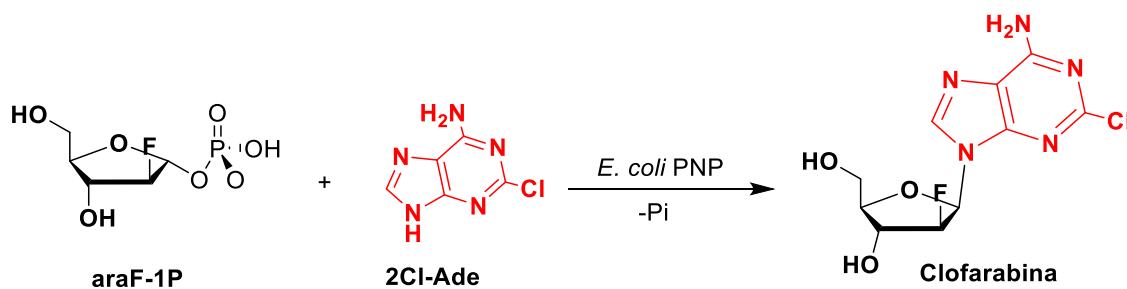
En las **NPs**, el equilibrio termodinámico para **PNP**, aunque no para PyNP, se encuentra desplazado hacia la síntesis del nucleósido, por lo que frecuentemente la reacción de transglicosilación consiste en la combinación de **PyNP-PNP** o **UP-PNP** para sintetizar nucleósidos púricos a partir de nucleósidos pirimidínicos como donadores y bases púricas como aceptores. Esto quiere decir, que son necesarios dos pasos y el empleo de dos enzimas distintas para la síntesis del nucleósido de interés. Por el contrario, la reacción de transglicosilación empleando nucleósido 2'-desoxirribosiltransferasas se produce en un solo paso, empleando una única enzima.

Si se opta por la inmovilización de las enzimas, en el caso de las **NPs** deben co-inmovilizarse dos enzimas, lo que produce una disminución en la efectividad del proceso y la obtención de un biocatalizador heterogéneo; mientras que la inmovilización de una única **DRT** asegura la obtención de un biocatalizador homogéneo<sup>73</sup>.

### 1.3.2. Síntesis quimioenzimática de clofarabina

Los 2-desoxi-2-fluoro- $\beta$ -D-arabinofuranósidos de pirimidina y de purina son análogos de nucleósidos con un amplio espectro de actividad biológica<sup>74</sup> y de entre ellos la clofarabina<sup>74, 75</sup>, (9- (2-desoxi-2-fluoro- $\beta$ -D-arabinofuranosil) -2-cloroadenina, ha sido aplicada con éxito para el tratamiento de la leucemia aguda pediátrica<sup>29, 76</sup>. Para este análogo, varios procesos de síntesis han sido propuestos, muchos de ellos, basados en la condensación de 2,6-dihalopurinas con bromuro de 3-O-acetil-5-O-benzoil-2-desoxi-2-fluoro- $\alpha$ -D-arabinofuranosilo o 1,3,5- tri-O-benzoil-2-desoxi-2-fluoro- $\alpha$ -D-arabinofuranosa<sup>77</sup>. En la mayoría de los casos con bajos rendimientos globales, formación de las mezclas  $\alpha/\beta$  de glucósidos, etc., lo cual supone etapas de purificaciones cromatográficas para separarlos, además del uso de grandes volúmenes de disolventes orgánicos.

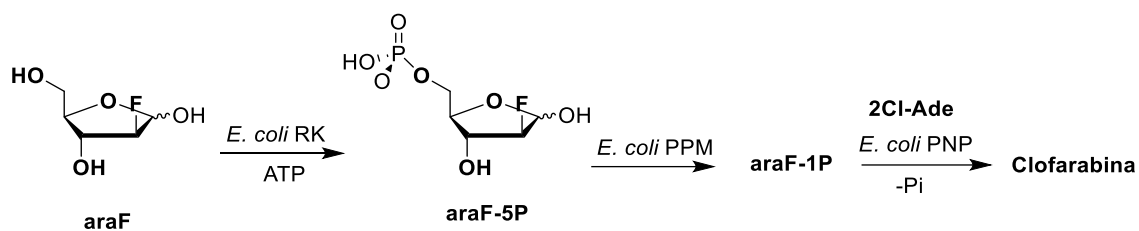
La búsqueda de alternativas “verdes”, más eficientes, ha conducido a proponer a las enzimas para llevar a cabo la transferencia de un residuo de pentofuranosa 2'-fluorinado obtenido a partir de un nucleósido pirimidina (donante) a una base purínica (aceptora). En este contexto, en la patente de Krenitsky y col.<sup>78</sup> (*Wellcome Res. Labs*) se describe la primera síntesis enzimática de 2-desoxi-2-fluoro- D-arabinofuranosil nucleósidos. En este estudio se transfiere un residuo de pentofuranosa, a partir de 1- (2-desoxi-2-fluoro-β-D-arabinofuranosil) timina, usando la timidina fosforilasa (**TP**) de *E. coli* y una purín nucleósido fosforilasa (**PNP**) de *E. coli*. Este proceso tenía dos “cuello de botella”: **i**) baja actividad de la **TP**, lo que suponía el uso de grandes cantidades de las enzimas: **TP** (160.000 unidades♦) y **PNP** (290.000 unidades♦) y, **ii**) largos tiempos de reacción. Recientemente, Fateev y col.<sup>79, 80</sup> han publicado dos estrategias para la síntesis de 2-cloro-9-(2-desoxi-2-fluoro-β-D-arabinofuranosil) adenina. La primera consistía en la síntesis química de 2-desoxi-2-fluoro-α-D-arabinofuranosa-1-fosfato y su condensación con 2-cloroadenina, en una reacción catalizada por una PNP recombinante de *E. coli* (Figura 15), con un 67 % de rendimiento.



**Figura 15:** Síntesis quimioenzimática de clofarabina. **araF-1P:** 2-desoxi-2-fluoro-α-D-arabinofuranosa-1-fosfato; **2Cl-Ade:** 2-cloroadenina; **PNP:** purín nucleósido fosforilasa.

Para la segunda estrategia proponían una cascada de reacciones catalizadas por varias enzimas relacionadas con la síntesis de nucleósidos, con un 45% de rendimiento. En la propuesta se parte de 2-desoxi-2-fluoro-D-arabinosa que es convertida al derivado fosfatado (2-desoxi-2-fluoro-D-arabinofuranosa-5-fosfato) usando la enzima riboquinasa (**RK**), luego una fosfopentomutasa (**PPM**), que convierte el derivado 5-fosfato en 2-desoxi-2-fluoro-α-D-arabinofuranosa-1-fosfato, que finalmente es condensado con 2-cloroadenina usando una **PNP** (Figura 16).

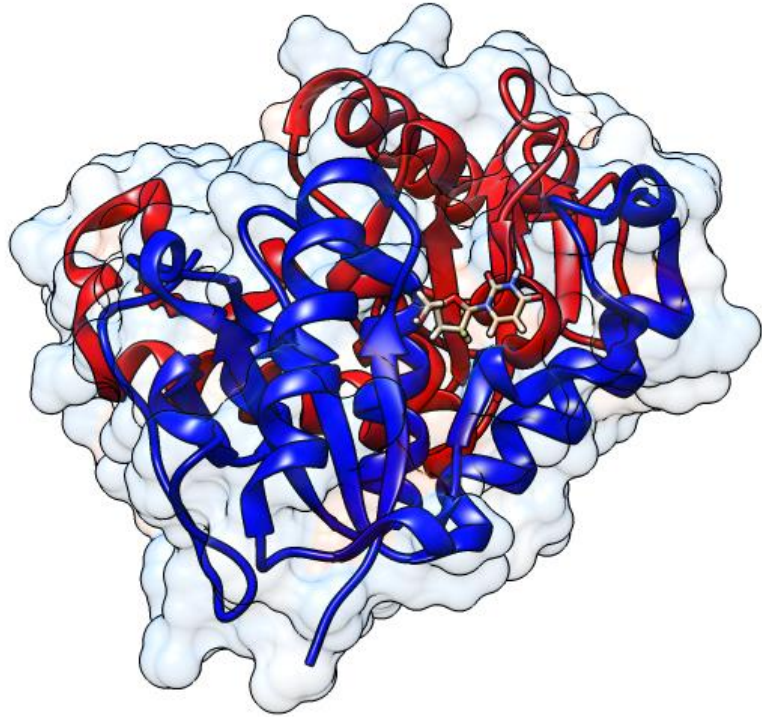
♦ Una unidad se define como la cantidad de enzima que fosforila un nucleósido para rendir 1 μmol de la base correspondiente por minuto.



**Figura 16:** Cascada para la transformación 2-desoxi-2-fluoro- $\alpha$ -D-arabinofuranosa en clofarabina. **araF**: 2-desoxi-2-fluoro- $\alpha$ -D-arabinofuranosa; **araF-5P**: 2-desoxi-2-fluoro- $\alpha$ -D-arabinofuranosa-5-fosfato; **araF-1P**: 2-desoxi-2-fluoro- $\alpha$ -D-arabinofuranosa-1-fosfato; **2Cl-Ade**: 2-cloroadenina; **PNP**: purín nucleósido fosforilasa; **PPM**: fosfopentomutasa; **RK**: riboquinasa.

En la publicación de Fernández-Lucas y col.<sup>81</sup> se describe por primera vez el reconocimiento de arabinonucleósidos por enzimas de tipo **NdRTs**. En el mismo, utilizando una nucleósido 2'-desoxirribosiltransferasa de tipo II de *Lactobacillus reuteri* (*Lr*NDT), lograron transferir un residuo 2'-desoxi-2'-fluoro ribofuranosilo, usando como donante, 2'-fluoro-2'-desoxiuridina y como base aceptora, adenina. Este resultado, en principio, no era esperado, las **NDTs** son muy específicas para 2'-desoxirribosa debido a que la interacción del 3'-OH del azúcar con la *Glu* del centro activo es vital para una orientación correcta del azúcar y por tanto para la actividad (Figura 14). De acuerdo con los experimentos de Anand y col.<sup>82</sup>, cuando se utilizan ribonucleósidos como donantes, el átomo nucleofílico del oxígeno de la *Glu98* del centro catalítico forma puentes de hidrógeno con el O-2' de la ribosa y la enzima se inactiva. Al parecer, en algunas NDTs, como ocurre con la *Lr*NDT, este puente de hidrógeno no ocurre cuando se utilizan como sustratos arabinonucleósidos y 2'-fluoro-2'-desoxirribonucleósidos y *Lr*NDT puede reconocerlos como sustratos y transferir el azúcar, aunque con tiempos de reacción muy largos, en comparación a los sustratos naturales de la enzima.





# HIPÓTESIS Y OBJETIVOS.



## 2. HIPÓTESIS Y OBJETIVOS

### 2.1. Planteamiento del problema

La catálisis enzimática es, hoy por hoy, una de las tecnologías más verdes para la síntesis de moléculas de alto valor añadido. En procesos multietapas, el uso de biocatalizadores altamente activos y específicos, evita etapas de activación/desactivación, protección/desprotección de grupos funcionales, lo cual se traduce en menos subproductos y/o desechos, menor consumo energético, etc.<sup>83</sup> En particular, la síntesis de *ingredientes activos farmacéuticos* (API), está asociada generalmente al uso y a la generación de grandes cantidades de sustancias peligrosas, por lo que están considerados entre los peores, en términos de **Factor E** (relación de masa de residuos respecto al producto deseado)<sup>84</sup>.

Para la síntesis de *clofarabina*, un análogo de nucleósido de purina de segunda generación autorizado por la EMEA♦ para su uso en el tratamiento de la leucemia linfoblástica aguda en pacientes pediátricos, se han establecido varios métodos de síntesis. Muchos de ellos, basados en la condensación de 2-cloroadenina con una 2-desoxi-2-fluoro-D-arabinofuranosa protegida<sup>77</sup>. Estos métodos de síntesis, en general, implican rendimientos globales bajos, la formación de las mezclas  $\alpha/\beta$  de glucósidos, etc., lo cual supone etapas de purificaciones cromatográficas para separarlos, además del uso de grandes volúmenes de disolventes orgánicos. Estos inconvenientes afectan el rendimiento del proceso, en términos de pureza y costes del producto final, así como de impacto ambiental.

La producción de clofarabina, usando reacciones de transglicosilación enzimática, ha sido abordada usando nucleósido fosforilasas (NPs; EC: 2.4.2.n)<sup>53, 80, 85-87</sup>, sin embargo, no existe una metodología (publicada) en la que se describa una reacción biocatalítica, comercialmente viable, para la producción de *clofarabina*. Entre otras razones, cuando se usan NPs no es factible utilizar como materia prima, un nucleósido como donante del pentofuranosilo activado (2-desoxi-2-fluoro- $\alpha$ -D-arabinofuranosa-1-fosfato),

---

♦ La **Agencia Europea de Medicamentos, EMA** o **EMEA** (por sus siglas en inglés), es una agencia de la Unión Europea que se encarga de la evaluación de las solicitudes de autorización de comercialización de medicamentos en la Unión Europea y su supervisión.

necesario para la reacción de transglicosilación. En este sentido, la estrategia usando **NPs**, pasa por suministrar el pentofuranosilo activado, ya sea usando un paso previo de síntesis química<sup>88, 89</sup>, o usando una cascada de reacciones catalizadas por varias enzimas relacionadas con la síntesis de nucleósidos. En ambos casos, con muy bajos rendimientos, hecho al que también contribuye la extremadamente baja solubilidad de la base aceptora, la 2-cloroadenina.

En la literatura consultada encontramos que una nucleósido 2'-desoxirribosiltransferasa de tipo II, como lo es la NDT de *Lactobacillus reuteri*, es capaz de reconocer arabinonucleósidos como sustratos<sup>81</sup>. Esto hecho fue considerado por estos autores como algo “inusual”, dado que a este tipo de enzimas se les suponía una estricta especificidad por nucleósidos 2'-desoxirribosílicos. En esta misma publicación se demostró que esta enzima es capaz de transferir un residuo 2'-desoxi-2'-fluoro ribofuranosilo a adenina, aunque con muy bajos rendimientos (aproximadamente:  $3,8 \times 10^{-4} \% \diamond$ ).

## 2.2. Hipótesis de investigación

¿Podrían las enzimas de tipo **NdRTs** llevar a cabo la síntesis biocatalítica de **clofarabina** (9-(2-desoxi-2-fluoro- $\beta$ -D-arabinofuranosil)-2-cloroadenina), en condiciones ambientalmente sostenibles, además de viables desde el punto de vista de la economía del bioproceso?





---

$\diamond$  Calculado a partir de los datos de Fernández-Lucas y col.<sup>81</sup>, en el que 1 mM del sustrato se transforma en 24 horas, con la *Lr*NDT ( $0,0085 \text{ mg.mL}^{-1}$ ), lo cual resulta en una actividad específica, para la síntesis de 2'-desoxi-2'-fluoroadenosina de  $0,3 \times 10^{-3} \text{ U.mg}^{-1}$ , que según nuestros cálculos sería un rendimiento de conversión de  $3,75 \times 10^{-4} \%$ .

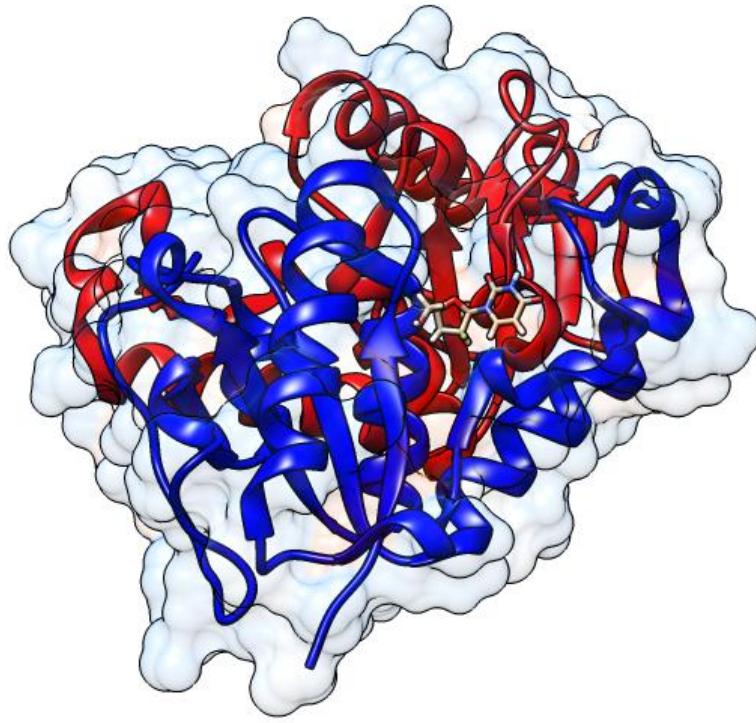


## 2.3. Objetivos

Para probar la hipótesis indicada, en el presente trabajo se plantearon los siguientes objetivos de investigación:

-  Estudios de acoplamiento (*docking*) molecular para estimar la posibilidad de que enzimas del tipo **NdRTs** puedan llevar a cabo la reacción de síntesis biocatalítica de *clofarabina*.
-  Clonaje del gen de la **NdRT** candidata en un microorganismo heterólogo con el fin de obtener niveles de expresión que permitan la aplicación industrial de la misma.
-  Aplicación de la **NdRT** a en la síntesis de *clofarabina*.
-  Evaluar la viabilidad técnica, ambiental y económica de la síntesis biocatalítica de *clofarabina* usando **NdRT**.





# **MATERIALES Y MÉTODOS.**



### 3. MATERIALES Y MÉTODOS

#### 3.1. Cribado virtual de NDTs para la síntesis biocatalítica de clofarabina

##### 3.1.1. Selección de las NDTs candidatas.

En este apartado se utilizaron tres estructuras cristalinas (Tabla 7) de proteínas con actividad nucleósido desoxirribosiltransferasa (EC 2.4.2.6), tomadas de la base de datos *Protein Data Bank* (PDB)<sup>90</sup>. Específicamente, se escogieron las anotadas como *Nucleósido 2-desoxirribosiltransferasa*, es decir, aquellas que presuntamente catalizan la transferencia de desoxirribosa entre purinas o pirimidinas (**Pur ↔ Pur, Pur ↔ Pyr, o Pyr ↔ Pyr**), aunque con una gran preferencia por desoxipirimidinas como sustrato donante<sup>69</sup>. La visualización de las moléculas en tres dimensiones se llevó a cabo usando el programa *Chimera*<sup>91</sup>▲.

**Tabla 7:** Estructuras cristalinas utilizadas en este apartado.

PDB ID	UNIPROT <sup>◇</sup>	Descripción
2A0K	Q57VC7	Estructura cristalina de la Nucleósido 2-desoxirribosiltransferasa de <i>Trypanosoma brucei</i>
1F8X	Q9R5V5	Estructura cristalina de la Nucleósido 2-desoxirribosiltransferasa de <i>Lactobacillus leichmannii</i>
3EHD	Q831Y8	Estructura cristalina de la proteína PF05014 de <i>Enterococcus faecalis</i> V583

##### 3.1.1.1. Estudios de acoplamiento (docking) molecular.

Los procedimientos de *docking* se llevaron a cabo usando la tecnología **KIN**, de la plataforma SAAS\* (*Mind the Byte, S.L.*, Parc Científic de Barcelona). Para determinar el sitio activo de la enzima se empleó el método de *docking* ciego o acoplamiento molecular ciego<sup>92</sup>, que consistió en realizar un *docking* en una cuadrícula (*grid*) que cubriera toda la superficie de la proteína para que los sustratos buscaran los

▲ <http://www.cgl.ucsf.edu/chimera/>

◇ UniProt ([www.uniprot.org/](http://www.uniprot.org/)) es una base de datos de secuencias de proteínas anotadas

\* <https://saas.mindthebyte.com>

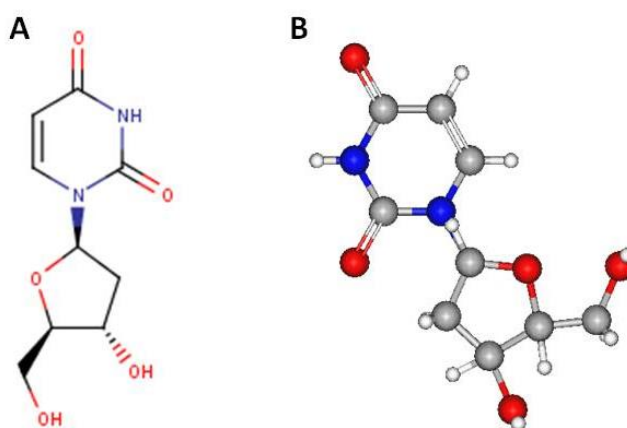
mejores sitios para su unión. Posteriormente, tomando como base el análisis de la energía libre de *Gibbs* de unión ( $\Delta G$  unión), representativo de la afinidad por dicho sitio, y criterios de geometría (lo suficientemente cerca del nucleófilo *Glu* en el centro activo de la enzima, se determinó el lugar más apropiado (de entre las *poses*<sup>Δ</sup> generadas) para el acoplamiento.

### A. Preparación de las proteínas.

Las moléculas de agua, los ligandos, metales e inhibidores fueron eliminados de las estructuras cristalinas de las proteínas evaluadas (Tabla 7), utilizando los programas *Pymol* y *Chimera*<sup>91</sup>. Con estos programas se generan ficheros *\*.pdb* modificados de estas proteínas.

### B. Preparación de los ligandos.

Las estructuras 2D y 3D de **desoxiuridina** (1-(2-desoxi-β-D-ribofuranosil)uracil) (Figura 17) y de **Ara F** (1-(2-desoxi-2-fluoro-β-D-arabinofuranosil)-uracilo) (Figura 18), fueron obtenidas de la base de datos *ChemSpider*<sup>°</sup>. Estas estructuras luego fueron descargadas en formato *MOL* y convertidas a ficheros: *.sk2*, *.mol2* y *.pdb*, con la ayuda del programa *MarvinSketch*<sup>×</sup> (*ChemAxon Ltd.*).

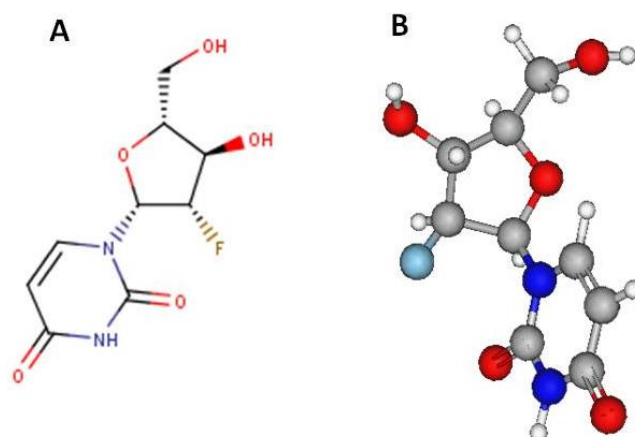


**Figura 17:** Estructura de la 2'-desoxiuridina. A: Estructura bidimensional (2D) de la **dUrd**. B: Estructura tridimensional (3D) de la **dUrd**.

<sup>Δ</sup> **Pose** – Un candidato al modo de enlace.

<sup>°</sup> <http://www.chemspider.com>

<sup>×</sup> (<http://www.chemaxon.com/products/marvin/marvinsketch/>)



**Figura 18:** Estructura del 2-desoxi-2-fluoro- $\alpha$ -D-arabinofuranosa **A:** Estructura bidimensional (2D) de AraF. **B:** Estructura tridimensional (3D) de AraF.

### 3.2. Síntesis de Clofarabina usando la nucleósido 2'-desoxirribosiltransferasa de *Lactobacillus delbrueckii* subsp. *lactis*.

#### 3.2.1. Obtención del vector de expresión.

##### 3.2.1.1. Aislamiento del gen de interés

##### A. Crecimiento del microorganismo.

*Lactobacillus delbrueckii* subsp. *lactis* DSM 20072 se cultivó en 10 mL de caldo MRS (*Biokar, Bioser, Barcelona, España*) durante 48 h a 32 °C. La composición del medio de cultivo MRS se muestra en la Tabla 8.

**Tabla 8:** Composición del medio MRS.

Ingrediente	Cantidad	Unidad
Peptona de soja	18	g/L
Extracto de levadura	4	g/L
Glucosa	20	g/L
Tween 80	1	mL/L
K <sub>2</sub> HPO <sub>4</sub>	2	g/L
Citrato triamónico	2	g/L
Acetato de sodio	3	g/L
Sulfato de magnesio 7H <sub>2</sub> O	0,2	g/L

## B. Extracción de ADN genómico.

Para aislar el ADN genómico se utilizó el kit comercial *ATP™ Genomic DNA Mini Kit* (*ATP biotech Inc.*), siguiendo las instrucciones proporcionadas por el fabricante para bacterias Gram positivas. Al final del procedimiento, la concentración e integridad del ADN extraído se determinaron usando el equipo y los reactivos del *Qubit® DNA Assay Kit*, *Qubit® 2.0 Fluorometer* (*Life Technologies, CA, USA*) y mediante electroforesis en geles de agarosa.

## C. Plásmidos y oligonucleótidos

Los oligonucleótidos utilizados como cebadores en las reacciones en cadena de la polimerasa (PCR) fueron sintetizados por la empresa **Biolegio BV†** y se especifican en la Tabla 9. El diseño de los oligonucleótidos para la amplificación y posterior clonaje de los genes de las enzimas se basó en la secuencia del gen de la nucleósido 2'-desoxirribosiltransferasa de *Lactobacillus delbrueckii* subsp. *lactis*, depositada en *GenBank‡*, con el número de acceso *WP\_003616023.1* (ver secuencia en la Tabla 10). A estas secuencias se les adicionaron secuencias de corte de las enzimas de restricción seleccionadas para poder clonar la construcción genética (vector de expresión):

- *Ldndt2<sub>Fw</sub>*, que corresponde al extremo 5' del gen *ndt2* y que presenta un sitio de restricción *NdeI*.
- *Ldndt2<sub>Rv</sub>*, que corresponde al extremo 3' del gen *ndt2* y que presenta sitio de restricción *SalI*.

**Tabla 9:** Secuencias de oligonucleótidos empleados en este apartado.

Nombre	Secuencia (5' a 3')	Descripción
<i>Ldndt2<sub>fw</sub></i>	5' - <u>catatg</u> ccaaaaaagacgatctacttc -3'	Empleado en la clonación de <i>Ldndt2</i>
<i>Ldndt2<sub>rv</sub></i>	5' - <u>gtcgact</u> tagtatacggcacc-3'	Empleado en la clonación de <i>Ldndt2</i>
<i>T7</i>	5' - taatacgaactcactataggg-3'	Secuenciación de construcciones en <i>pET22b</i> y <i>pGEM</i>
<i>T7-term</i>	5' - ctagttattgctcagcggg -3'	Secuenciación de construcciones en <i>pET22b</i>
<i>SP6</i>	5' - catacgatttaggtgac actatag -3'	Secuenciación de construcciones en <i>pGEM</i>

Los nucleótidos subrayados se corresponden a las dianas de restricción añadidas para el clonaje

† [www.biolegio.com](http://www.biolegio.com)

‡ [www.ncbi.nlm.nih.gov/genbank](http://www.ncbi.nlm.nih.gov/genbank)



**Tabla 10:** Secuencia de la nucleósido 2'-desoxirribosiltransferasa de *Lactobacillus delbrueckii* subsp. *lactis*.

---

> Nucleósido desoxirribosiltransferasa de *Lactobacillus delbrueckii* subsp. *lactis* DSM 20072.  
 ATGACGGCAACAGCACGTATTTATTTAGGTTTCATCTTTTCACCGTGACCACCATCAAGCC  
 CGGGCCAAGCAGGTCCGCGAATACCTGGCGGCCAACCCGACGGTTGAACATGTCCACTTT  
 CCCTTTGACCAGCAGGTAACCGACCCGGCTGAAGACCCGAATGTCGGCTTTGGCGACGAA  
 CGCAGCCTGCTTTGGAAGCAGCAGACCTACCAAAACGATATCAACGGCCTGTCCATGGCC  
 AACTGCGGGGTCTTCTCTACGACATGGATGAACTGGATGACGGCTGCGCCTTTGAAATC  
 GGCTTTTTTGCGGGCGCAGAACAAGCCGGTCATCATCGTGCCTTTCACTAACCAGCCGGAA  
 AAGAAGCCAATCATGAACTTGATGCTGTCTGAAGGGGCAACTGCTTTTATCAAGGGTGAC  
 CAGGAAGTGGAACTCTTAAAGACTTACGACTTTAAACAAGTGGCCATGCCGGCTGCTGACT  
 GGCTACGGCACTATTTAA

---

Los plásmidos empleados en este apartado y sus características más relevantes se detallan en la Tabla 11.

Tabla 11: **Plásmidos utilizados.**

Plásmido	Genotipo relevante o propiedades	Fuente o referencia
<i>pGEM-T- Easy</i>	<i>Ap<sup>r</sup></i> , vector para clonar fragmentos de ADN	Promega
<i>pET-22b(+)</i>	<i>Ap<sup>r</sup></i> , vector de clonación e hiperexpresión, <i>oriColE1</i>	Novagen
<i>pGEMT - Ldndt2</i>	Fusión entre <i>pGEM-T- Easy</i> y el gen <i>Ldndt2</i> de <i>L. delbrueckii</i>	Este trabajo
<i>pET-22b- Ldndt2</i>	Fusión entre <i>pET-22b(+)</i> y el gen <i>Ldndt2</i> de <i>L. delbrueckii</i>	Este trabajo

#### D. Subclonaje del gen *ndt2* de *Lactobacillus delbrueckii*

El vector plasmídico empleado como vector de clonación de productos de PCR es *pGEM®-T easy* (Promega<sup>♦</sup>) (Figura 19). El vector se suministra cortado y con extremos 3'T protuberantes, de modo que actúan como extremos cohesivos con los extremos 3' de los amplificadores por PCR que contienen adeninas. Los subclonajes con este plásmido solo exigen la ligación del inserto mediante la enzima *T4 ligasa*, no requieren de la presencia de enzimas de restricción y permiten la identificación de las colonias positivas mediante color.

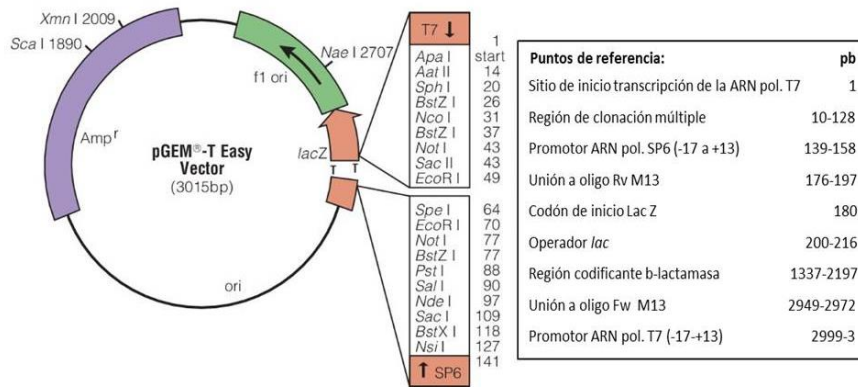
---

♦ <https://www.promega.es/resources/protocols/technical-manuals/0/pgem-t-and-pgem-t-easy-vector-systems-protocol/>

### 3.2.2. Clonaje del gen *ndt2* en vector “pGEM-T Easy” (pGEM-Ldndt2).

#### A. Amplificación del gen *ndt2* mediante PCR.

El gen codificante de la enzima *N*-desoxirribosiltransferasa de *Lactobacillus delbrueckii* subsp. *lactis* fue amplificado, mediante PCR, empleando los oligonucleótidos *Ldndt2<sub>fw</sub>* y *Ldndt2<sub>rv</sub>* (Tabla 9), utilizando el ADN genómico de *L. delbrueckii* como molde, en las condiciones descritas en la Tabla 12.



**Figura 19:** Esquema del vector pGEM-T Easy (*Promega*<sup>®</sup>). En la tabla de la derecha se describen los principales elementos de este vector.

**Tabla 12:** Programa de amplificación por PCR.

Paso	Temp. (°C)	Duración
1	94°C	2 min
2	46°C	5 min
3	72 °C	2 min
4	95°C	1 min
5	46°C	2 min
6	72 °C	2 min
7	Repetición pasos 4-6	30 ciclos
8	72 °C	10 min

La amplificación del DNA mediante PCR se realizó en un termociclador PTC-100 (*MJ Research*), utilizando la mezcla de reacción descrita en la Tabla 13.

**Tabla 13:** Mezcla de reacción de la PCR.

Componente	Volumen
ADN genómico de (20 ng/ $\mu$ L)	5 $\mu$ L
Tampón de la <i>Taq</i> polimerasa (10X)	5 $\mu$ L
<i>Ldndt2<sub>fw</sub></i> (10 $\mu$ M)	1 $\mu$ L
<i>Ldndt2<sub>rv</sub></i> (10 $\mu$ M)	1 $\mu$ L
dNTP mix (10 mM dNTPs)	1 $\mu$ L
<i>Taq</i> polimerasa (5 U/ $\mu$ L)	0,25 $\mu$ L
H <sub>2</sub> O libre de nucleasas	hasta 50 $\mu$ L

### B. Extracción del producto de PCR.

Para extraer las secuencias amplificadas tras la PCR, la mezcla de reacción se somete a una electroforesis en geles de agarosa 1% (p/v), Electroforesis de DNA en tampón *TAE* utilizando la cubeta de electroforesis horizontal *Enduro 7.10 (Labnet International)*, a 120V durante aproximadamente 40 minutos. La tinción de los geles llevó a cabo con *Sybr safe<sup>TM</sup> (Invitrogen, Life Technologies)* bajo luz UV, utilizando para ello el sistema *Gel-doc<sup>TM</sup> (Bio-Rad Laboratories)*. Una vez teñido el gel, el corte de las bandas se llevó a cabo sobre un transiluminador de luz ultravioleta que permite su visualización. La purificación de los fragmentos de ADN a partir de geles de agarosa se llevo a cabo con el *kit ATP<sup>TM</sup> Gel/PCR extraction kit (ATP biotech Inc.)*.

### C. Clonaje producto de PCR.

Con el producto de PCR purificado, se llevó a cabo la clonación en el plásmido pGEM<sup>®</sup>-T (*Promega*), usando para la reacción de ligación la *ligasa* del fago *T4*, siguiendo el protocolo recomendado por Promega. En general, las reacciones de ligación realizadas en este trabajo, se hicieron en volúmenes de 10  $\mu$ L, en relaciones

molares vector:inserto 1:3 y en tiempos de incubación de 12 h a 4°C. La construcción obtenida se denominó *pGEM-Ldndt2*.

#### **D. Transformación de *E. coli* competentes y selección de clones.**

La transformación de las células competentes de *Escherichia coli* **TOP 10** (Invitrogen) se realizó mediante choque térmico, utilizando 5µL del producto de ligación (*pGEM-Ldndt2*). El choque térmico se llevo a cabo mediante la incubación de la suspensión en hielo, durante 20 min, seguido de la incubación a 42 °C durante 45 seg. e incubación final en hielo durante 2 min. A continuación se añadieron 600 µL de medio **LB** (triptona, 10 g L<sup>-1</sup>; extracto de levadura, 5 g L<sup>-1</sup>; NaCl, 10 g L<sup>-1</sup>, pH 7,0) a la suspensión de células transformadas y se completó con la incubación a 37 °C durante 1 h en agitación. Finalmente las bacterias transformadas se sembraron en placas de **LB** agarizado suplementado con ampicilina (100 µg/mL), IPTG (0,5 mM) y **X-Gal** (80 µg/mL). Las colonias que incorporaron el vector intacto fueron de color azul, mientras que las que incorporaron el plásmido recombinante (*pGEM-Ldndt2*) fueron blancas; esto debido a la interrupción del gen de la β-galactosidasa.

#### **E. Selección de colonias con inserto mediante PCR de colonia (“Colony PCR”)**

En la mayoría de los casos los clones que contienen productos de *PCR* producen colonias blancas. En este trabajo, la presencia del inserto y su orientación también fueron comprobadas mediante **PCR**, con los oligonucleótidos **T7** y **SP6**, que flanquean el sitio de inserción, con los *oligos* del “amplicón”, o una combinación de *oligos* del amplicón y del vector<sup>§</sup>.

Para la *PCR* de colonia se utilizó como molde 5 µL de suspensiones bacterianas (una colonia en 20 µL de agua destilada) en reacciones de PCR convencionales, con los oligonucleótidos seleccionados. En la

Tabla 14 se muestra la mezcla de reacción utilizada, usando los dos *oligos* del vector y en la Tabla 15, el perfil térmico utilizado. Para las amplificaciones usando los *oligos* del “amplicón” se utiliza el perfil térmico descrito en la Tabla 12.

---

<sup>§</sup> Dependiendo de la pareja de oligos utilizada, el perfil térmico utilizado fue modificado convenientemente.

**Tabla 14:** Mezcla de reacción de PCR para detección de plásmidos con inserto.

<b>Reactivo</b>	<b>Volumen por reacción (µL)</b>
Agua purificada	9,87
Tampón de la <b>Taq</b> polimerasa (10X)	2,5
Mezcla <b>dNTP</b> (10 mM)	0,5
Oligo <b>T7</b> (10 µM)	1
Oligo <b>SP6</b> (10 µM)	1
<i>Taq</i> polimerasa (5 U/µL)	0,125

**Tabla 15:** Perfil térmico de PCR para detección de plásmidos con inserto.

<b>Paso</b>	<b>Temp (°C)</b>	<b>Duración</b>
1	95°C	5 min
2	95°C	1 min
3	48°C	1 min
4	72 °C	1 min
5	Repetición pasos 2-4	30 ciclos
6	72 °C	10 min

Los resultados de la amplificación se visualizaron en geles de agarosa al 1%. Solo aquellas colonias que habían incorporado el plásmido con el inserto dieron resultado positivo de amplificación. Las colonias que presentaron un inserto del tamaño esperado fueron recuperadas y sus plasmidos se aislaron para su secuenciación.

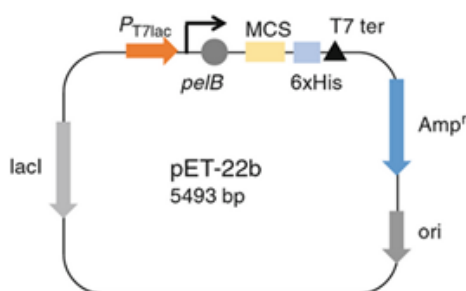
#### **F. Minipreparaciones de DNA plasmídico y secuenciación de los productos.**

Las colonias positivas, seleccionadas mediante PCR de colonia, se hicieron crecer en medio **LB-amp**, para aislar el DNA plasmídico, mediante purificación por mini-preparaciones o “mini-preps”. En este trabajo se empleó el kit *ATP™ Plasmid DNA Mini Kit* (*ATP Biotech*), para la obtención de “mini-preps” plasmídicas, siguiendo la metodología propuesta por el fabricante.

Para la secuenciación se utilizaron los oligonucleótidos universales: T7 y SP6. La secuenciación de ADN se realizó en el **Servei de Genòmica i Bioinformàtica de la Universitat Autònoma de Barcelona**.

### 3.2.3. Clonaje del gen *ndt2* en el vector de expresión.

Como vector de expresión en este capítulo se eligió el plásmido pET22b(+) (Figura 20). El vector *pET-22b(+)* posee la secuencia de una señal *pelB* en el extremo N-terminal para una localización de la proteína en el espacio periplasmático. El vector permite opcionalmente obtener la proteína de interés fusionada con una etiqueta de histidinas en el extremo C-terminal. El vector pET-22b(+) puede ser secuenciado con los primers **T7** y *T7 terminator* y confiere resistencia a ampicilina.



**Figura 20:** Representación esquemática del vector de expresión pET-22b<sup>93</sup>. La transcripción es iniciada por el promotor T7, regulada por el operador lac/lacI y terminada por T7 ter.

Para obtener las construcciones recombinantes en el vector de expresión se llevaron a cabo dobles digestiones de las construcciones en **pGEM-T** y de los vectores de expresión, con endonucleasas de restricción (*New England Biolabs*), empleando los tampones y las condiciones de digestión recomendados por el fabricante. Los productos de digestión se purificaron y se ligaron en proporción inserto:vector 3:1 utilizando la ligasa del fago T4 en las mismas condiciones. Los productos de ligación se transformaron en células competentes de *E. coli* Top10 (*Invitrogen*) y se seleccionaron en placas de agar LB con ampicilina (100 µg/mL). Los plásmidos recombinantes extraídos a partir de los clones positivos (Ap<sup>r</sup>), se estudiaron mediante análisis de restricción, PCR y secuenciación, para seleccionar aquellos que hayan incorporado la construcción con el gen *ndt2*. El plásmido purificado a partir de estas

células se usó para transformar las células BL21(DE3), en las cuales se llevó a cabo la expresión de proteína recombinante.

### 3.2.3.1. Preparación y validación del banco celular de la cepa recombinante *E. coli* BL21 (pET22-Ldndt2).

#### A. Crecimiento y crioconservación del clon recombinante

Una colonia de la cepa recombinante crecida en medio LB con ampicilina (100 µg/mL), se inoculó en 50 mL de medio *TB* líquido (Tabla 16). El medio líquido inoculado se incubó a 37 °C, hasta que se alcanzó una  $DO_{600nm} = 1,0$ . A continuación, el caldo de cultivo se mezcló asepticamente con 50 mL de una solución de glicerol esteril al 25% y alícuotas de 1 mL de esta mezcla se transfirieron a crioviales (*Cryo.s*<sup>TM</sup>, *Greiner bio-one*) estériles y se conservaron a -80°C.

**Tabla 16:** Composición del medio TB (*Terrific Broth*).

Ingrediente	Cantidad	Unidad
Peptona de soja	12	g/L
Extracto de levadura	24	g/L
KH <sub>2</sub> PO <sub>4</sub>	2,32	g/L
K <sub>2</sub> HPO <sub>4</sub>	12,54	g/L
Glicerol	5,04	g/L
Ampicilina	100	µg/mL

#### B. Viabilidad del cultivo

Para la determinación de la viabilidad se seleccionaron aleatoriamente tres crioviales del banco celular y se inocularon diluciones (diluciones seriadas) de las mismas en placas con medio LB agar suplementadas con ampicilina. La pureza del cultivo se determinó tras el análisis micro y macroscópico de colonias previamente crecidas en TSA (agar triptona de soja, *Scharlau*).

#### C. Curvas de crecimiento y determinación de la velocidad de crecimiento.

Para evaluar la producción de la proteína recombinante *Ld*NDT2 se ensayaron medios que se usan comúnmente para cultivar *E. coli* recombinantes. Específicamente, los cultivos se llevaron a cabo en cuatro medios: caldo de Luria (**LB**), caldo *Terrific* (**TB**), medio mínimo M9 (**M9**, Tabla 17) y medio mínimo modificado (**MMM**, Tabla 18). Todos los medios fueron formulados como se especifica en el *Handbook of Microbiological Media*<sup>94</sup>. Se añadió inoculo (1 mL) a 50 mL de cada uno de los cuatro medios en matraces de 250 mL, incubados con agitación, a 180 rpm y 37°C. Se tomaron muestras cada hora y graficando el log de los valores de *DO*<sub>600nm</sub> versus tiempo, se pudo determinar la velocidad de crecimiento (Ecuación 1) y el tiempo de generación (Ecuación 2).

**Ecuación 1** 
$$\mu = \frac{d \ln DO_{600nm}}{dt}$$

**Ecuación 2** 
$$G = \frac{\ln 2}{\mu}$$

**Tabla 17:** Composición del medio mínimo M9.

<b>Ingrediente</b>	<b>Cantidad</b>	<b>Unidad</b>
Na <sub>2</sub> HPO <sub>2</sub>	12,8	g/L
KH <sub>2</sub> PO <sub>4</sub>	3	g/L
NaCl	0,5	g/L
NH <sub>4</sub> Cl	2	g/L
Glucosa	20	g/L
CaCl <sub>2</sub>	0,01	g/L
MgSO <sub>4</sub>	0,12	g/L
FeCl <sub>3</sub>	0,002	g/L
Antibiótico (Ampicilina)	100	µg/mL



**Tabla 18:** Composición del medio mínimo modificado (MMM).

<b>Ingrediente</b>	<b>Cantidad</b>	<b>Unidad</b>
Glucosa	30	g/L
MgSO <sub>4</sub> × 7H <sub>2</sub> O	1,2	g/L
KH <sub>2</sub> PO <sub>4</sub>	13,3	g/L
(NH <sub>4</sub> ) <sub>2</sub> HPO <sub>4</sub>	4	g/L
Ácido cítrico	1,7	g/L
Sol. Elementos traza	10	g/L
Tiamina HCl	100 µL	mL/L
Antibiótico (Ampicilina)	100	µg/mL

#### **D. Perfiles de inducción y estabilidad de los plásmidos.**

Se inocularon dos matraces Erlenmeyer con 50 mL de medio, con crioviales de la cepa recombinante (ver apartado 3.2.3.1). Los matraces fueron incubados, toda la noche, a 37°C, en agitación (200 rpm). Estos crecimientos sirvieron para inocular matraces Erlenmeyer con 700 mL de medio seleccionado. Las cepas fueron incubadas a 37°C, con agitación y el crecimiento fue monitorizado, cada hora, a DO<sub>600nm</sub>. Adicionalmente se tomaron alícuotas de 1 mL, para determinar la concentración de proteínas (ver apartado 3.2.8) y la actividad enzimática (ver apartado 3.2.9). Para ello, las muestras fueron lisadas y el *debris* celular eliminado mediante centrifugación (13000 rpm, 5 min). La concentración de proteínas en el sobrenadante fue analizada usando el método de Bradford, siguiendo las instrucciones del fabricante (ver apartado 3.2.8) y la actividad nucleósido desoxirribosiltransferasa se determinó de acuerdo al método descrito en el apartado 3.2.9.

Basándonos en los resultados de las curvas de crecimiento, los cultivos fueron inducidos con 1,0 mM *IPTG*, en la fase exponencial de crecimiento. El crecimiento y la toma de muestras se prolongaron 4 horas tras la inducción.

Para la determinación de la estabilidad del plásmido, se tomaron muestras del cultivo a intervalos regulares de tiempo y diluciones de las mismas fueron sembradas en placas de agar LB (sin ampicilina) y en placas de LB selectivas (con ampicilina). Las placas fueron incubadas a 37°C, durante 16 horas. La relación entre el número de *UFC*

(unidades formadoras de colonias) en el medio selectivo y las UFC en el medio no selectivo fue usada para determinar el porcentaje de células portadoras del plásmido, y como indicador de la estabilidad plasmídica<sup>95</sup>.

#### 3.2.4. Producción de la nucleósido 2'-desoxirribosiltransferasa (*LdNDT2*) de *Lactobacillus delbrueckii subsp. lactis* DSM 20072.

Para llevar a cabo la producción de NDT de *L. delbrueckii*, se utilizaron los medios LB y TB suplementados con ampicilina (100 µg/mL). Se hizo crecer el microorganismo recombinante *E. coli* BL21 (pET22-*Ldndt2*) en 50 mL de medio de cultivo (suplementado con ampicilina), en un matraz de 250 mL a 37° C, con una agitación de 200 rpm en un incubador termostatzado con agitación orbital incubador orbital **WY200** (*IvyMen System*), durante 18 horas. Este crecimiento sirvió para inocular, en una relación 1/100, matraces Erlenmeyer de un litro que contenían 450 mL de medios **LB** o **TB** suplementados con ampicilina (100 µg/mL). Los medios se incubaron en las mismas condiciones hasta observar una densidad óptica (DO<sub>600nm</sub>) de entre 0,6 y 0,8; momento en el cual se adicionó **IPTG** a una concentración final de 1 mM y se mantuvo creciendo, en estas condiciones, 4 horas adicionales a 37°C. Tras la inducción, se tomaron muestras (1mL del caldo fermentativo) cada hora, estas se centrifugaron (12000 rpm, 2min) y el sedimento se conservó a -20°C, hasta el momento de ser analizadas mediante SDS- PAGE (apartado 3.2.8). Al final de la inducción, el caldo de cultivo se centrifugó a 4000×g durante 15 minutos a 4 °C y, tras separar el sobrenadante de las células, se procedió a la lisis del sedimento celular (ver apartado 3.2.6).

Se usaron matraces de agitación para probar el efecto del tiempo de inducción (fase exponencial temprana y fase exponencial media) sobre la producción de *LdNDT2*. Se inocularon (por duplicado) matraces de 250 mL que contenían 50 mL de **TB** suplementados con ampicilina. Estos cultivos fueron inducidos con 1 mM de **IPTG**. Todos los cultivos se incubaron a 180 rpm y 37°C. La proteína total, la proteína recombinante y la densidad celular final se midieron tras de 8 h de fermentación.

### 3.2.5. Producción de la nucleósido 2'-desoxirribosiltransferasa (*LdNDT2*) de *L. delbrueckii* en cultivo discontinuo (*batch*) y discontinuo alimentado (*fed-batch*).

El escalado de la producción de la nucleósido 2'-desoxirribosiltransferasa (*LdNDT2*) de *L. delbrueckii* se llevó a cabo en sistemas de cultivo discontinuos y discontinuos alimentados. Para estas fermentaciones se inoculó el contenido de un criovial del banco celular en un matraz Erlenmeyer de 250 mL con 50 mL de medio LB suplementado con ampicilina. Este inóculo se hizo crecer en agitación (200 rpm) a 37°C entre 12-16 h y 2 mL de este cultivo sirvieron para inocular un matraz de 2L que contenía 400 mL de medio **TB** suplementado con antibiótico. La cepa recombinante en medio **TB** se hizo crecer, toda la noche, a 180 rpm (37°C).

Para las fermentaciones en *batch*, 125 mL del crecimiento *overnight* se inocularon asépticamente en un biorreactor de 5L (*BioFlo®*, *New Brunswick Scientific*) con 2,5 L de medio **TB** estéril. El programa *Biocommand Bioprocessing* (*New Brunswick Scientific*) se usó para controlar las fermentaciones y registrar los datos de las fermentaciones. Durante las fermentaciones discontinuas el oxígeno disuelto se mantuvo al 30% de saturación y fue controlado mediante la agitación (hasta 500 rpm) y/o suministrando oxígeno puro. La agitación, el pH, la temperatura y el flujo de aire fueron establecidas al inicio de las fermentaciones: 200 rpm; 7,0; 37 °C y 4 L/min, respectivamente. El pH del cultivo durante la fermentación se controló entre 6,8 -7 usando una solución de hidróxido de amonio al 25%. La formación de espuma se controló usando agentes antiespumantes (*Antifoam 289*, *Sigma*). En las fermentaciones en *batch*, los cultivos fueron crecidos entre 4-6 horas después de la inoculación e inducidos a una concentración final de **IPTG** (1mM). Cuatro horas después de la inducción, la biomasa producida fue recuperada mediante centrifugación a 4000 x g durante 30 min. y lisada utilizando homogenizador de alta presión (***Panda Plus 2.000***, ***GEA-Niro Soavi***).

En el caso de las fermentaciones discontinuas alimentadas, 200 mL del crecimiento *overnight* se inocularon asépticamente a un biorreactor de 5L (*BioFlo®*, *New Brunswick Scientific*), con 2,5 mL de medio TB. Los cultivos se iniciaron con los siguientes parámetros: aireación, 3-7 L/min; agitación, 200-800 rpm; pH, 6,8-7 y a 37 °C. El

oxígeno disuelto se controla a una saturación de ~30 %. Al principio, el cultivo se operó en modo *batch*, a 37°C, hasta que prácticamente se ha consumido todo el glicerol del medio (esto también se pone de manifiesto, por un fuerte incremento del oxígeno disuelto). En este punto (~ 5 h después de la inoculación) se comenzó a alimentar con un “medio de alimentación”, que contenía: extracto de levadura, glicerol, sulfato de magnesio y una solución de elementos traza. La alimentación se realizó durante un periodo de aproximadamente 9 horas. Si el oxígeno disuelto caía por debajo del 30 % de saturación, se detenía esta alimentación, hasta que recuperaran los niveles de saturación establecidos. Cuando el cultivo alcanzó niveles de DO<sub>600nm</sub> cercanos a 30 (aprox. 10 horas), se indujo la expresión mediante la adición de ***IPTG*** (a una concentración final de 1,0 mM) y el cultivo se mantuvo a 37°C durante 4-8 horas. El medio de alimentación contenía (por litro) 120 g de extracto de levadura, 5 g de (NH<sub>4</sub>)SO<sub>4</sub>, 5 g de MgSO<sub>4</sub>·7H<sub>2</sub>O y 500 g de glicerol. Y éste se alimentó usando la estrategia **DO-stat** bajo el control automático del programa BioCommand (*New Brunswick Scientific*). La biomasa producida fue recuperada mediante centrifugación a 4000 x g durante 30 min. y lisada utilizando un homogeneizador de alta presión (*Panda Plus 2000, GEA-Niro Soavi*).

### 3.2.6. Preparación y tratamiento de la biomasa.

#### A. Disrupción celular a pequeña escala.

Para analizar los niveles de expresión mediante ***SDS-PAGE***, las muestras, post-inducción (1mL), tomadas durante la expresión de la proteína recombinante (apartado 3.2.3.1) fueron lisadas (sonicadas) a pequeña escala usando un sonicador Branson Digital Sonifier 450 (*Branson Ultrasonics Corp.*). Para ello, los sedimentos se resuspendieron en 100 µl de tampón de lisis (Tris-HCl, 20 mM, pH 7,2; 5 mM de EDTA y 1 mM de ***DTT***) y se lisaron usando 3 pulsos de sonicación de 5 segundos con una amplitud del 30%. El resultado de la lisis se centrifugó a 12,000 rpm durante 2 min., el sobrenadante se analizó mediante ***SDS PAGE*** y el contenido de proteína total se obtuvo mediante el método de *Bradford* (apartado 3.2.8).

Para la preparación del extracto celular, el sedimento resultante de la expresión de la *NDT* (apartado 3.2.6) se resuspendió en 20 mL del tampón de lisis<sup>♦</sup> (Tris-HCl, 20 mM; pH 7,2; EDTA, 5 mM, DTT, 1 mM) y se enfrió, en hielo, durante 20 min. La suspensión fue entonces sonicada según las siguientes condiciones: 10 ciclos de 10 segundos con una amplitud del 30 %, a intervalos de 10 segundos. Tras lisar totalmente las células, se centrifugó la suspensión durante 20 min a 17.000 x g a 4 °C y se procedió a analizar mediante electroforesis en geles de poliacrilamida, tanto el extracto (sobrenadante) como los sedimentos obtenidos después de centrifugar el lisado, con el fin de determinar si la enzima se expresaba de forma soluble o en forma de cuerpos de inclusión. La cantidad de proteína total en el extracto se cuantificó mediante el método de Bradford (ver apartado 3.2.8). La eficacia de la lisis celular se analizó mediante *SDS-PAGE* (ver apartado 3.2.8).

### **B. Disrupción celular a gran escala.**

Para la preparación del extracto celular, el sedimento resultante de la expresión de la *NDT* en el biorreactor, usando el sistema discontinuo alimentado, (apartado 3.2.4) se resuspendió en tampón de lisis (Tris-HCl, 20 mM, pH 7,2; EDTA, 5 mM; DTT, 1 mM) y se enfrió, en hielo, durante 30 min. La suspensión se disruptó usando un homogeneizador de alta presión a escala de laboratorio (*Panda Plus 2000, GEA-Niro Soavi*), a 1000 bar (4 pases).

### **C. Determinación de Biomasa**

Con el fin de determinar la equivalencia de la absorbancia de una muestra en términos de peso seco se construyó una curva de calibración, tomándose la absorbancia de las diluciones de una solución madre de biomasa de absorbancia 1 medida a una longitud de onda de 600 nm. Posteriormente se filtraron suspensiones (diluciones) usando filtros de membrana (tamaño de poro de 0,22 µm) previamente pesados. La biomasa filtrada se seca, hasta peso constante, y se pesa. El valor de peso seco de la biomasa es el resultante de sustraer entre el peso final de la membrana (con biomasa) el peso inicial de la membrana (sin biomasa). Graficando el peso seco vs. la

---

<sup>♦</sup> Por lo general se resuspende el sedimento celular para obtener una concentración final de aproximadamente 4DO<sub>600nm</sub> por mL.

absorbancia de la muestra, se obtiene la ecuación que relaciona la absorbancia de una muestra y peso seco de la biomasa.

### 3.2.7. Purificación de la nucleósido 2'-desoxirribosiltransferasa (*LdNDT2*) de *L. delbrueckii*.

La fracción soluble obtenida tras centrifugar los lisados celulares se sometió a un proceso de precipitación fraccionada usando sulfato de amonio, como primer paso para la purificación de la enzima *NDT* de *L. delbrueckii*. Para ello se utilizaron dos concentraciones de sulfato amónico para precipitar las proteínas de forma fraccionada. En el primer paso (20 % de saturación), este fue añadido lentamente a la solución enzimática, en condiciones de agitación suave y a 4 °C. Una vez disuelto completamente se dejaron transcurrir 20 min y se centrifugó a 15,000 rpm a 4 °C durante 20 min. El precipitado se resuspendió en tampón Tris-HCl, 20 mM, pH 7,2, y se continuó añadiendo sulfato amónico para alcanzar el punto de 70% de saturación. Una vez centrifugado el sedimento, se resuspende en tampón Tris 20 mM, pH 7,2. Ambas fracciones fueron diafiltradas utilizando filtros Amicon™ (*Millipore*), con un tamaño de poro de 10 kDa.

Como segundo paso para la purificación de la enzima, después de la precipitación fraccionada en sulfato amónico, se realizaron dos purificaciones sucesivas usando un intercambiador aniónico HiTrap Q Sepharose FF (*GE healthcare life sciences*). La fracción proveniente de la precipitación con sulfato de amonio, una vez dializada, fue centrifugada a 17,000 g, durante 20 min y ajustada a una concentración de 1 mg/mL con tampón A<sup>Q-seph</sup> (Tris-HCl, 20 mM, pH 7,2). La solución se pasó por la columna de Q-sepharose (5 mL, a una velocidad de flujo 1mL/min), previamente equilibrada con 5 volúmenes de columna (VC) de tampón B<sup>Q-seph</sup> (igual al tampón A<sup>Q-seph</sup>, pero con una concentración final de NaCl de 1M) y 10 VC de tampón A<sup>Q-seph</sup>. Los aluatos resultantes de la fase de carga fueron colectados y la columna fue lavada con 5VC de tampón A<sup>Q-seph</sup> y los lavados fueron colectados como fracciones de 5 mL. La proteína fue entonces eluida con un gradiente escalonado de NaCl (0,2; 0,4; 0,6; 0,8 y 1 M.), a una velocidad de 2 mL/min. Los eluatos fueron colectados, esta vez, en fracciones de 2,5 mL y analizados por *SDS-PAGE*, para detectar la presencia de la *LdNDT2*. La segunda cromatografía se realizó con las fracciones positivas de la

primera y se realizó con las siguientes concentraciones de cloruro sódico (gradiente): 0,2; 0,25; 0,3; 0,35 y 0,4M.

### 3.2.8. Análisis de Proteínas.

#### A. Electroforesis en geles de poliacrilamida (SDS-PAGE).

Las fracciones solubles e insolubles y los eluatos de los resultados de las etapas de purificación fueron analizados mediante SDS-PAGE. Para este propósito se usaron geles preformados (*NuPAGE™ Novex™ 4-12 % Bis-Tris Protein Gels*). Las muestras provenientes de los cultivos fueron normalizadas sobre la base de sus densidades ópticas y preparadas para el análisis electroforético siguiendo protocolos estándar<sup>96</sup>. Brevemente, las fracciones solubles y de proteínas totales proteicas se combinaron con 1 volumen de tampón de carga 2x (Tris·Cl, 100 mM pH 6,8; DTT, 200 mM; SDS, 4 %; bromofenol azul, 0,2 % y glicerol, 20 %). Luego las mezclas se calentaron a 95°C, durante 5 min, se dejaron enfriar a temperatura ambiente. Las muestras fueron procesadas electroforéticamente (aprox. 90 min a 150 V) y las proteínas separadas se detectaron utilizando el protocolo propuesto en el kit *GelCode Blue Safe Protein Stain (Thermo scientific)*.

#### B. Determinación de la concentración de proteínas.

La concentración de proteína presente en las disoluciones se determinó según el método de *Bradford*<sup>97</sup>, que se basa en la variación en el espectro de absorción visible del colorante azul brillante de Coomassie G-250 en presencia de una proteína. La adsorción de proteína provoca que el valor del máximo de absorción se desplace de 465 nm a 595 nm. La preparación del reactivo de Bradford se llevó a cabo partiendo de un concentrado comercial de Bio-Rad (*Dye reagent*) siguiendo las instrucciones del fabricante (dilución 1:5 para el método estándar y sin dilución para el método micro). Como patrón de proteína se utilizó una solución de albúmina de suero bovina (BSA) de concentración 1 mg/mL en agua purificada (*Grifols*) o en el tampón de muestra, según fuera el caso. Se realizaron disoluciones sucesivas de la proteína patrón como se muestra en la Tabla 19, completándose con agua purificada (o tampón) hasta un volumen total de 0,1 mL. A continuación se añadieron 5 mL de reactivo de Bradford y se homogenizó la mezcla. La adsorción de la proteína al colorante se produce instantáneamente siendo máxima y estable entre los 5 y 20 primeros minutos. Todos los

ensayos se realizaron a temperatura ambiente y por duplicado según la tabla de dilución utilizada (Tabla 19). Los resultados de absorbancia en función de la concentración de proteína, en  $\mu\text{g/mL}$  para el volumen final de 5,1 mL, se ajustaron por regresión lineal al modelo de una recta de primer grado, cuya linealidad se valoró según el valor de  $R^2$ .

**Tabla 19:** Disoluciones sucesivas de proteína patrón para construir la recta de calibración (método de Bradford).

Nº	$\mu\text{L}$ de Proteína (1 mg/mL)	$\mu\text{L}$ de Agua purificada o tampón	mL de Bradford	BSA ( $\mu\text{g/mL}$ )
0	0	100	5,00	0
1	20	80	5,00	2,94
2	30	70	5,00	4,41
3	40	60	5,00	5,88
4	50	50	5,00	7,35
5	60	40	5,00	8,82
6	70	30	5,00	10,29
7	80	20	5,00	11,76
8	90	10	5,00	13,23
9	100	0	5,00	14,71

### C. Determinaciones densitométricas de la concentración de proteínas.

Los valores de área e intensidad de las bandas de interés obtenidas en los geles de poliacrilamida, se analizaron densitométricamente, usando un el sistema de documentación (*Gel Doc™ EZ System, BIO-RAD*). La concentración de proteínas en las bandas se determinó por interpolación de los valores de intensidad en una curva estándar realizada con varias concentraciones, en mg/mL, de BSA.

#### 3.2.9. Ensayo de actividad nucleósido 2'-desoxirribosiltransferasa

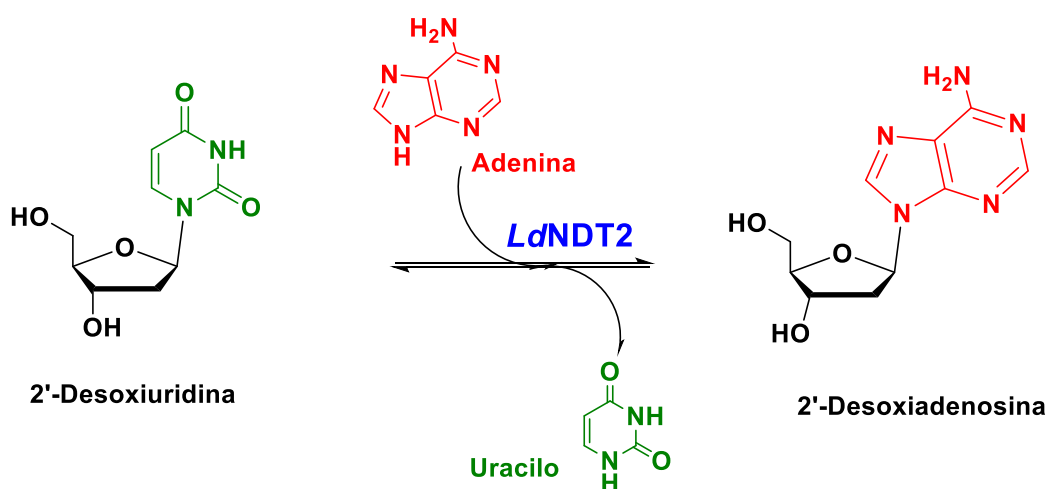
El método para la determinación de la actividad nucleósido 2'-desoxirribosiltransferasa empleado en este trabajo se basó en la cuantificación mediante HPLC de los 2'-desoxirribonucleósidos formados a partir de un 2'-desoxirribonucleósido y una base de partida<sup>61</sup>. Específicamente, en este apartado se utilizó como reacción estándar la obtención de 2'-desoxiadenosina, a partir de 2'-desoxiuridina como sustrato donador de



2'-desoxirribosa y adenina como base aceptora (Figura 21). La actividad enzimática se cuantificó mediante el empleo de la unidad internacional (UI), definida como la cantidad de enzima que produce 1  $\mu\text{mol}$  de 2'-desoxirribonucleósido por minuto en las condiciones de ensayo.

El ensayo para la determinación de la actividad enzimática se llevó a cabo a 40°C, en tampón MES 50 mM (*morpholineethanesulfonic acid*), pH 6,5. Una solución de adenina 10 mM y desoxiuridina a 10 mM se incubó a la temperatura de reacción, durante 10 min. La reacción se inició tras adicionar la enzima purificada o los extractos (30  $\mu\text{L}$ ; 0,068  $\mu\text{g}$  proteína total/mL) para obtener un volumen final de reacción de 240  $\mu\text{L}$ . Después de 5 min de reacción, se detuvo la reacción adicionando 240  $\mu\text{L}$  de metanol frío y la mezcla se calentó a 95°C durante 5 min. Finalmente la mezcla se centrifugó durante dos minutos a 9700 rpm. El sobrenadante de esta centrifugación se diluyó 1/3 en agua de calidad HPLC y se filtró por un tamaño de poro de 10 KDa para poder ser analizada posteriormente por HPLC.

La mezcla de reacción filtrada se analizó por HPLC en una columna C18 Tracer Excel ODSA, 5  $\mu\text{m}$ , 250 x 46 mm (*Teknokroma*). Las condiciones de elución fueron agua/metanol 96/4 como fase móvil inicial y posterior gradiente aumentando el porcentaje de metanol, flujo de 1 mL/min, volumen de inyección de 10  $\mu\text{L}$  y detección a 254 nm.



**Figura 21:** Esquema de la reacción del ensayo de actividad para *LdNDT2*.

### 3.2.10. *Reacciones de Síntesis de clofarabina usando la nucleósido 2'-desoxirribosiltransferasa de Lactobacillus delbrueckii subsp. lactis.*

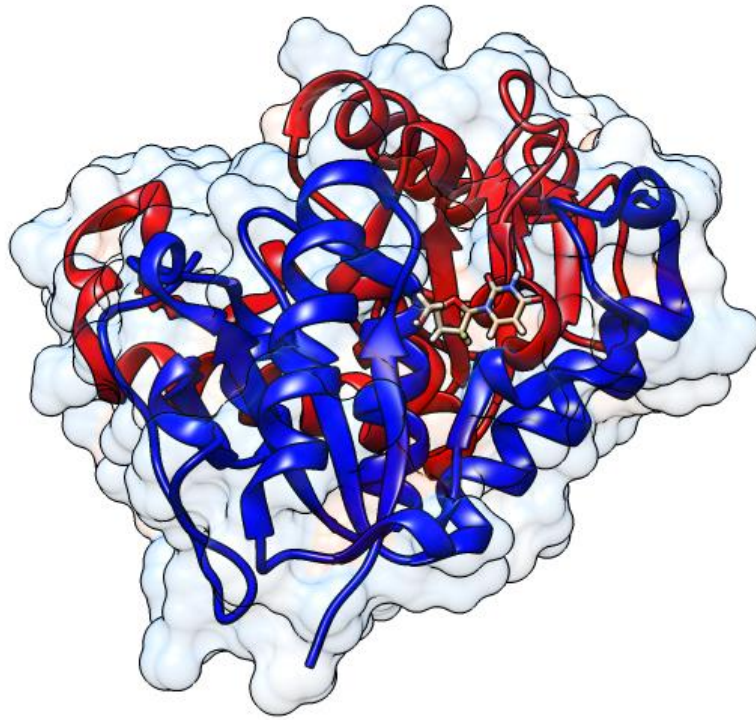
En este apartado se sintetizó el análogo de nucleósido clofarabina, en ambientes acuosos y usando cosolventes, para mejorar la solubilidad de las materias primas y empleando extractos crudos con actividad *NDT* o *LdNDT2* purificada. Las reacciones a pequeña escala (200-1000  $\mu$ L) se llevaron a cabo en tubos Eppendorf incubados en incubador orbital, a 300 rpm. A escalas superiores las reacciones se realizaron en reactores de vidrio de fondo redondo, equipados con un puerto para toma de muestras, condensador de reflujo y un termómetro y agitación magnética. Los reactores se colocaron en un baño de aceite mineral, con un controlador capaz de mantener la temperatura en el intervalo de  $\pm 1^\circ\text{C}$  con respecto al punto de consigna (*set point*).

#### **A. Producción de clofarabina en medio acuoso.**

Para la síntesis enzimática de clofarabina se termostataron ( $50^\circ\text{C}$ ) suspensiones de 2-cloroadenina, 10-50 mM y 2'-fluoro-arabinofuranosil-2'-desoxiuridina (**araF**), 20-50 mM en Tampón MES, pH 6,5, durante 30 min. Para iniciar la reacción de síntesis, la enzima o los extractos enzimáticos fueron adicionados (0,9-20 U/  $\mu\text{mol}_{\text{araF}}$ ) y la reacción se mantuvo a  $50^\circ\text{C}$ . La duración de las reacciones osciló entre 2-72 horas.

#### **B. Producción de clofarabina en medio acuoso y cosolventes orgánicos.**

Para la síntesis enzimática de clofarabina se termostataron ( $50^\circ\text{C}$ ) suspensiones de 2-cloroadenina, 20-50 mM y 2'-fluoro-arabinofuranosil-2'-desoxiuridina (**araF**), 20-50 mM en Tampón MES, pH 6,5, suplementado con un cosolvente orgánico, durante 30 min. Para iniciar la reacción de síntesis, la enzima o los extractos enzimáticos fueron adicionados (0,9-20 U/  $\mu\text{mol}_{\text{araF}}$ ) y la reacción se mantuvo a  $50^\circ\text{C}$ . La duración de las reacciones osciló entre 2-72 horas.

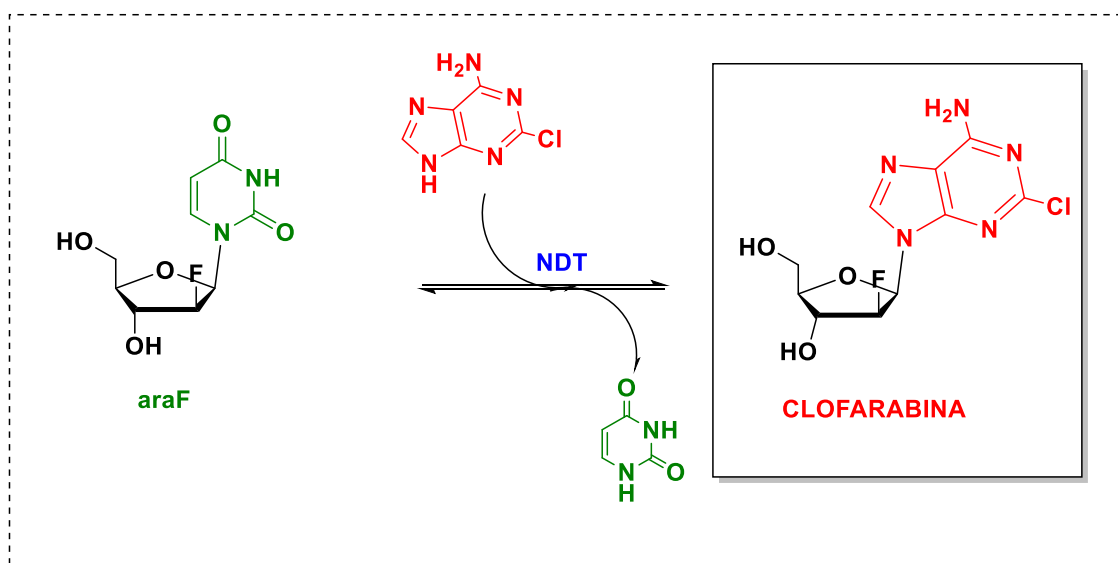


# **RESULTADOS Y DISCUSIÓN.**

## 4. RESULTADOS Y DISCUSIÓN.

La clofarabina es un análogo de nucleósido de purina de segunda generación, autorizado por la **EMEA** para su uso en el tratamiento de la leucemia linfoblástica aguda en pacientes pediátricos. Su síntesis química ha sido desarrollada y optimizada<sup>37, 40</sup>, aunque su producción a gran escala está ligada al uso de grandes volúmenes de disolventes orgánicos y a etapas de purificación para obtener el anómero deseado, con el consiguiente efecto sobre el rendimiento global del proceso.

En este sentido, en este trabajo nos propusimos la búsqueda de alternativas, "verdes", para la obtención de este análogo de nucleósido, usando nucleósido desoxirribosiltransferasas como biocatalizadores (Figura 22).



**Figura 22:** Síntesis de clofarabina usando nucleósido desoxirribosiltransferasas como biocatalizadores.

Los primeros estudios sobre N-desoxirribosiltransferasas revelaron que estas enzimas tienen una amplia tolerancia hacia las modificaciones de la purina o la pirimidina<sup>68, 98, 99</sup> y, aunque son muy específicas para 2-desoxirribosa, Kaminski y col.<sup>100</sup> han demostrado que una simple mutación en el centro activo de esta enzima puede ampliar esta especificidad, por ejemplo, para utilizar 2,3-dideoxirribosa y 2,3-didehidro-2,3-dideoxirribosa. También se ha descrito el reconocimiento de 2'-desoxirribonucleósidos

con flúor en la posición 2'C<sup>101</sup>, de esta manera usando una *LINDT* recombinante se ha transferido la 2'-desoxi-2'-fluoro- $\alpha$ -D-ribofuranosa, a partir de 2'-desoxi-2'-Fluorouridina (**dFUrd**), a 1,2,4-triazol-3-carboxamida para de esta forma obtener 2'-desoxivirazol, aunque con bajos rendimientos, en una reacción con una duración de 14 días. Y, finalmente, los resultados obtenidos por Fernández Lucas y col.<sup>61</sup> con la NDT de *L. ruteri*, dan argumentos para pensar en la plasticidad de estas enzimas, toda vez que la *Lr*NDT puede reconocer arabinonucleósidos y 2'-fluorodesoxirribonucleósidos como sustratos.

#### 4.1. Estudios de acoplamiento (*docking*) molecular

Con el objetivo de valorar la potencialidad de enzimas del tipo NDT para llevar a cabo la reacción de síntesis de clofarabina, se realizó un cribado, mediante estudios de acoplamiento (*docking*) molecular, de estructuras cristalinas anotadas en la base de datos **Protein Data Bank** como: “*Nucleoside 2-deoxyribosyltransferase*”, usando como ligandos: *i*) un sustrato natural de estas enzimas, la 2'-desoxiuridina (**dUrd**) y *ii*) el sustrato propuesto para la síntesis de clofarabina, 2'-fluoro-arabinofuranosil-2'-desoxiuridina (**araF**).

En la selección de las estructuras cristalinas, se ha tenido en cuenta su pertenencia a las DRTasas de la clase II (NDTs), es decir, aquellas que catalizan la transferencia de desoxirribosa entre purinas o pirimidinas (**Pur**  $\leftrightarrow$  **Pur**, **Pur**  $\leftrightarrow$  **Pyr**, o **Pyr**  $\leftrightarrow$  **Pyr**), aunque con una gran preferencia por desoxipirimidinas como sustrato donante<sup>69</sup>. En este sentido, se escogieron tres estructuras (Tabla 7) que, a pesar de no tener una elevada identidad entre ellas, sí que muestran una gran similitud estructural.

En la Figura 23 se muestra un alineamiento de estas tres secuencias y analizando en detalle las mismas, observamos la conservación de los residuos involucrados en la actividad biocatalítica de las 2'-desoxirribosiltransferasas<sup>59, 69, 72, 102</sup>, es decir, los que intervienen en el reconocimiento de la 2'-desoxirribosa en *Lactobacillus leichmannii*, que son *Tyr7*, *Asp72*, *Asp92*, *Glu98* y *Asn123*. De éstos, en el alineamiento de las secuencias, se pueden identificar, en la NDT de *Enterococcus faecalis*, que los residuos *Tyr5*, *Asp80*, *Glu86* y *Asn130* podrían estar involucrados en el reconocimiento de la 2'-

desoxirribosa; y en la NDT de *Trypanosoma brucei*, se pueden identificar los residuos *Tyr5*, *Asp76*, *Glu82* y *Asn126*. Aunque las principales diferencias se observan en la conservación de los residuos que participan en la el reconocimiento de la base (residuos *Glu 46*, *Asp72* y *Tyr157* en *Lactobacillus leichmannii*). Aquí se sustituye la **Glu** (de la *LINDT*) por *Ala42*, en *3EHD* y *Glu42* en *2A0K*, en tanto que los otros dos están ausentes en estas secuencias.

A partir de los resultados del alineamiento no se puede asegurar que **3EHD** y **2A0K** sean NDTs de tipo II, toda vez que en ambas faltan 2 de los 3 residuos polares y aromáticos implicados en la correcta orientación y estabilización de la base, tal y como ocurre en la proteína *LINDT*, **Gln** y **Tyr**. Fresco-Taboada y col.<sup>64</sup> indicaron que la **Gln** en las NDT se localiza en un bucle que actúa como tapa del centro activo y que asegura el ajuste del sustrato en el mismo. En el caso de las **NDTs**, la presencia de la **Gln** fuerza a que ese bucle se cierre sobre el sitio activo y lo aísla del medio; mientras que en las **PDTs**, en las que el residuo **Gln** está ausente, el bucle se encuentra en una conformación más abierta y sólo se produce ese ajuste cuando la base es voluminosa, una purina, ya que las pirimidinas tienen un tamaño menor. El otro residuo ausente, la *Tyr* terminal, es según trabajos de mutagénesis dirigida realizados por Short y col.<sup>72</sup> más crítica en el caso de las **PDTs**, ya que es este residuo el que forma los enlaces de hidrógeno con la purina, mientras que en las **NDTs** estos puentes los forma la **Gln**, ausente en las **PDTs**.

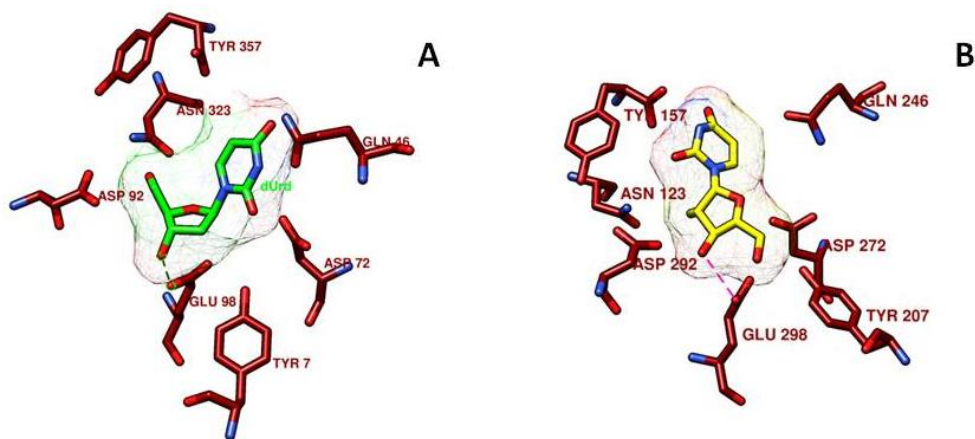
A la luz de estos resultados, las nucleósido 2-desoxirribosiltransferasas de *Enterococcus faecalis* y *Trypanosoma brucei* podrían tener mayor preferencia por las purinas. De todas maneras, otros aminoácidos podrían interaccionar con la base y suplir la ausencia de los residuos determinantes de la actividad en *LINDT*. Así, por ejemplo, al comparar la disposición del centro activo de *BpNDT*<sup>♦</sup> con el de otras **NDTs**, Fresco-Taboada y col.<sup>64</sup> concluyeron que *BpNDT* podría tener mayor preferencia por las purinas, pero que puede llevar a cabo la reacción de transferencia con aceptores pirimidínicos dado que el carboxilato de la *Lys142*, en esta enzima, interviene en el mecanismo de reacción.

---

♦ 2'-desoxirribosiltransferasa de *Bacillus psychrosaccharolyticus*.

	1	11	21	31	41	
Conservation	* * ** : . .	. *	. * * :	:: : . .	: * : *	
1F8X	1 MPKKTLYFGA	GW--FTDRQN	KAYKEAMEAL	KE-NPTIDLE	NSYVPLDNQY	47
3EHD	1 MTK--LYFAG	PLFSQADLRY	NAY--LVEQI	RQLDKTIDL-	--YLPQENAJA	43
2A0K	1 MRK--LYLAG	PAVENPDMGA	SYYNKVRELL	KK-ENVMP--	--LIPTDNEJA	43
	51	61	71	81	91	
Conservation	:	. *	: :	* :: .	* *	
1F8X	48 KGIRVDEHPE	YLHDKVWATA	TYNNDLNGIK	TNDIML---G	VYIPDEEDVVG	94
3EHD	44 ----INDKSA	YADSKMIALA	D---TENVL	ASDLLV---A	LLDGPTIDAG	82
2A0K	44 ----TEALDI	R--QK-----	N---IQMIK	DCDAVIADLS	PFRGHEEPDCG	78
	101	111	121	131	141	
Conservation	. * : * * :	. * : . .	*		: **	
1F8X	95 LGMELGYALS	QGKYVLLVIP	DEDY-----	-----	---GKPINLM	125
3EHD	83 VASELGVAYA	KGIPVVALYT	DSRQQGADNH	QKLDALNEIA	ENQFHYLNLY	132
2A0K	79 TAEVGCAAA	LNMVLTFTS	DRRNMRKYG	SGVDKDNLRV	EGFGLPFNLM	128
	151	161	171	181		
Conservation	. .	* : :	:			
1F8X	126 SWGV-SDNVI	KMSQLKDFNF	NKPRF-DFYE	GAVY	157	
3EHD	133 TVGLIKLNGR	VVSSEEDLLE	EIKORLS---	---	159	
2A0K	129 LYDGEVEFDS	FESAFKYFLA	NFPSK-----	---	153	

**Figura 23:** Alineamiento con ClustalW de las secuencias de NDTs utilizadas. **1F8X:** Nucleósido 2'-desoxirribosiltransferasa tipo II de *Lactobacillus leichmannii*; **3EHD:** proteína PF05014 de *Enterococcus faecalis* V583; **2A0K:** Nucleósido 2-desoxirribosiltransferasa de *Trypanosoma brucei*. El grado de conservación se indica con los símbolos ClustalW. Las regiones señaladas con el cuadro blanco, indican los sitios de reconocimiento del sustrato pirimidinico.



**Figura 24:** Centro activo de *LINDT*. **A:** Complejo (*in silico*) de **dUrd** unida a *LINDT*; **B:** Complejo *LINDT* con **AraF**. Se muestra un mapa de la densidad electrónica de los sustratos en el interior del centro activo (modelo semitransparente).

En este sentido, se realizaron estudios de acoplamiento molecular con la **dUrd** y con **AraF** y se estimaron las posiciones de estas moléculas en el centro activo de las NDTs seleccionadas. Los criterios de selección de la enzima para llevar a cabo la reacción propuesta fueron: i): el de la cercanía geométrica a los residuos involucrados

en el centro catalítico y ii): criterios energéticos (energía favorable) para finalmente establecer un enlace covalente entre el residuo **Glu** nucleofílico y el pentofuranosilo. La Tabla 20 muestra los resultados del estudio de acoplamiento ciego de las mejores “*poses*” y sus principales interacciones con los residuos de los centros activos de las tres enzimas seleccionadas. Analizando los datos de la energía libre de unión de los dos ligandos utilizados en este estudio se puede apreciar que **dUrd** (-7,5 Kcal/mol) es capaz de formar un complejo 0,6 Kcal/mol más estable que la **AraF** (-6,9 Kcal/mol) cuando el receptor es la *L*INDT. Este resultado puede estar relacionado al hecho de que **dUrd** es capaz de establecer un mayor número de interacciones con los residuos involucrados en el centro activo.

Sin embargo, esto no ocurre así cuando se analizan los resultados del acoplamiento con los otros dos receptores. El centro activo de estas dos enzimas es más volumétrico dado que sus sustratos, los nucleósidos purínicos, son más voluminosos. Esto hace que la ubicación de los sustratos en estudio (**dUrd** y **AraF**), encuentre posibilidades de establecer puentes de hidrógeno con residuos diferentes a los esperados para el caso de una NDT de tipo II. En la Tabla 20 se puede apreciar que las “*poses*” que se acercan a los criterios energéticos y de cercanía al **Glu** catalítico, no se corresponden con los residuos típicos del centro catalítico de una dRTasa de tipo II.

En la Figura 24 se puede apreciar que ambos sustratos (**dUrd** y **AraF**) se posicionan en el centro activo del enzima *L*INDT, de manera que pueden ser reconocidos por las cadenas laterales de los residuos *Asp72*, *Asn323*, *Glu98* (residuo catalítico, posicionado por *Tyr7*), *Asn92*, *Glu46* y *Tyr7*, que pueden formar una densa red de puentes de hidrógeno con este sustrato natural (Figura 24 **A**). **AraF**, Figura 24 **B**, también se ubica de manera que, espacialmente, puede ser reconocido por los residuos antes mencionados, aunque la presencia del flúor impone torsiones adicionales en **AraF** que hace que determinados residuos en el centro activo no puedan establecer puentes de hidrógeno (Tabla 20).



**Tabla 20:** Resultados del estudio de acoplamiento ciego. **dUrd:** desoxiuridina; **AraF:** 2'-fluoro- arabinofuranosil-2'-desoxiuridina. **A:** y **B:** cadena en la cual se localiza el residuo.

PDB ID	Ligando	Energía de unión (kcal/mol)	Residuos involucrados
1F8X	dUrd	-7,5	Gln46A, Glu98A, Asp72A, Asp92A, Asn323B, Tyr357B
	AraF	-6.9	Ala10A, Glu98A, Gln46A
2A0K	dUrd	-6,9	Gly16A, Gln50A, Asn61A, Thr87A, Glu90A, Asn134B
	AraF	-7,5	Gln50A, Glu90A, Asn134B
3EHD	dUrd	-6.5	Gly8B, Asp60B, Asn130B, Tyr132B
	AraF	-6,8	Gly8B, Asp60B, Asn130B, Tyr132B

A la luz de estos resultados, en este trabajo se estima el uso de la NDT de *Lactobacillus leichmannii* (**1F8X**), para la síntesis biocatalítica de clofarabina.

## 4.2. Síntesis de Clofarabina usando la nucleósido 2'-desoxirribosiltransferasa de *Lactobacillus delbrueckii subsp. lactis*

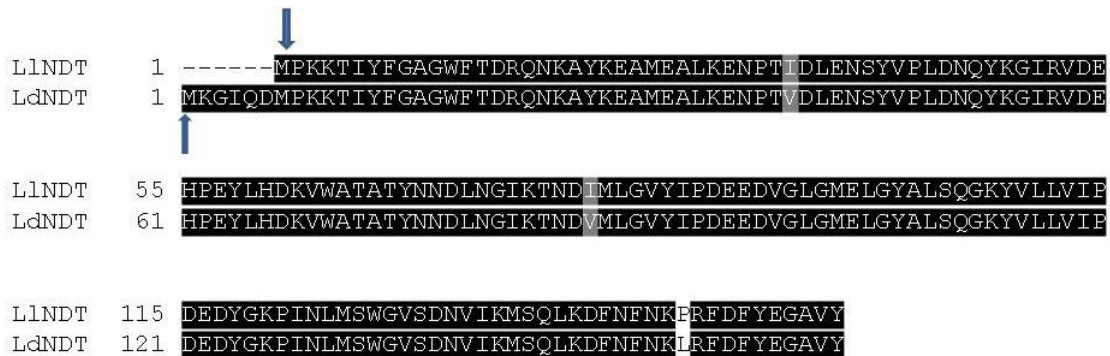
### 4.2.1. Obtención de la nucleósido 2'-desoxirribosiltransferasa recombinante de *Lactobacillus delbrueckii subsp. lactis*.

#### 4.2.1.1. Amplificación, clonación y secuenciación del gen *Ldndt2*

Una vez seleccionada, en los estudios *in silico*, la enzima de *Lactobacillus leichmannii*, en este trabajo se ha clonado y expresado esta enzima (específicamente, la de *L. delbrueckii*), con el objetivo de disponer del biocatalizador para las reacciones de transglicosilación. En la base de datos **Genbank**, el gen *ndt2* (número de acceso: EGD27012.1) codificaría para una proteína de 163 residuos aminoacídicos, con una masa molecular de 18,741 kDa. Observando detalladamente el producto del gen *ndt2* (proteína NDT2 de *L. delbrueckii*), encontramos que contiene dos metioninas, muy

cercanas (*Met1* y *Met7*) en la región N-terminal (Figura 25), es decir dos posibles codones de inicio (*ATG*), en este gen. No obstante, en el alineamiento de ambas secuencias (las secuencias de *L. delbrueckii* y de *L. leichmannii*; (Figura 25), a partir de la segunda metionina, las secuencias son prácticamente idénticas\*. En este sentido se diseñaron oligos para amplificar la secuencia evitando los 7 primeros residuos, incluyendo la *Met7*, pues los trabajos de Kaminsky y col.<sup>59</sup>, Porter y col.<sup>69</sup> y Short y col.<sup>104</sup> indican que este residuo no está presente en la proteína de *L. leichmannii* madura, una vez expresada en *E. coli*.

Para la amplificación del gen *ndt2* por PCR se añadieron los sitios de restricción *NdeI* y *Sall* en los extremos del gen. La amplificación por PCR permitió la obtención de un fragmento de 483 pb, que concuerda con el tamaño del gen *ndt2* que codifica para la síntesis de la NDT de *Lactobacillus delbrueckii subsp. lactis*. El producto génico se purificó y se subclonó en el vector pGEM-T-Easy y se obtuvo la construcción recombinante, pGEM-*Ldndt2* (Figura 26).



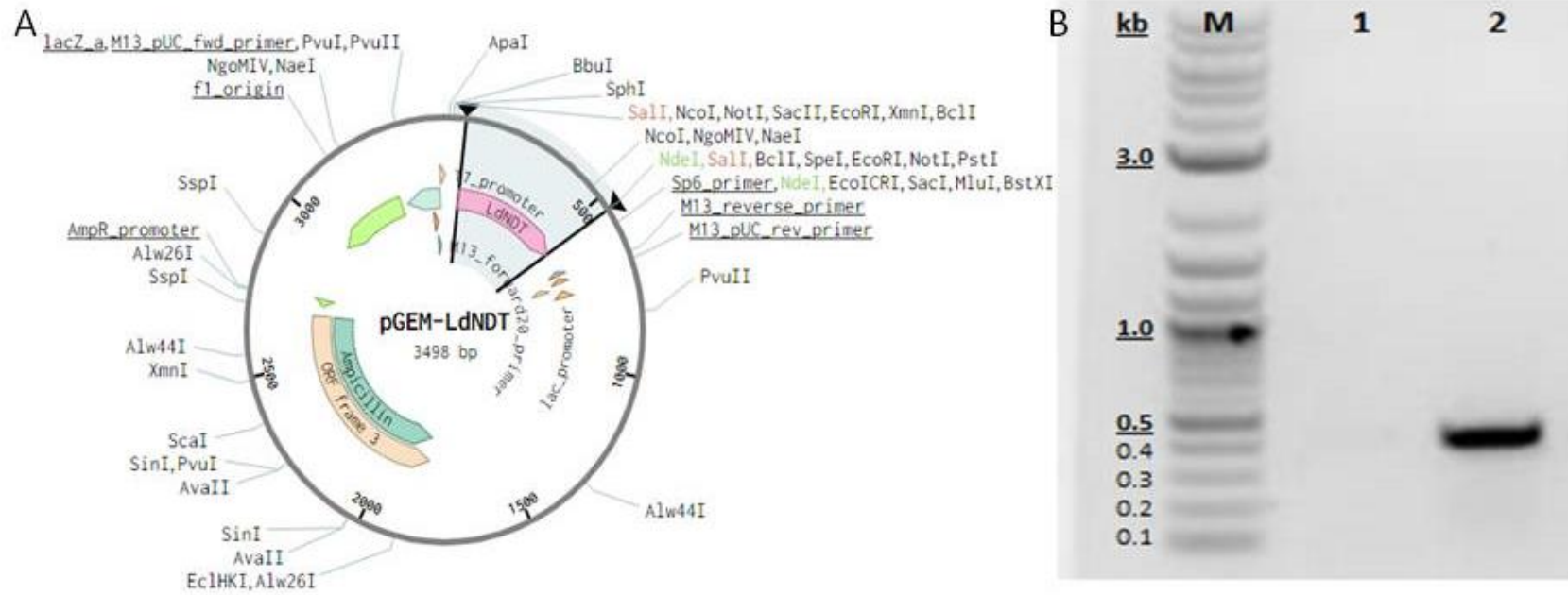
**Figura 25:** Alineamiento de las secuencias de las NDT de *Lactobacillus delbrueckii* (*LdNDT2*) y *Lactobacillus leichmannii* (*LINDT*). Se muestran los residuos (sombreado claro) diferentes en estas secuencias.

Para la expresión recombinante se utilizó el sistema *E. coli* BL21(DE3)/pET22b(+). En este caso, a partir del vector recombinante pGEM-*Ldndt2*, se obtuvo, tal y como se describe en el apartado 3.2.1.1, el fragmento 483 pb. Dicho fragmento se purificó y se clonó utilizando como vector el plásmido pET22b(+) para obtener el vector recombinante pET22b-*Ldndt2* (Figura 27).

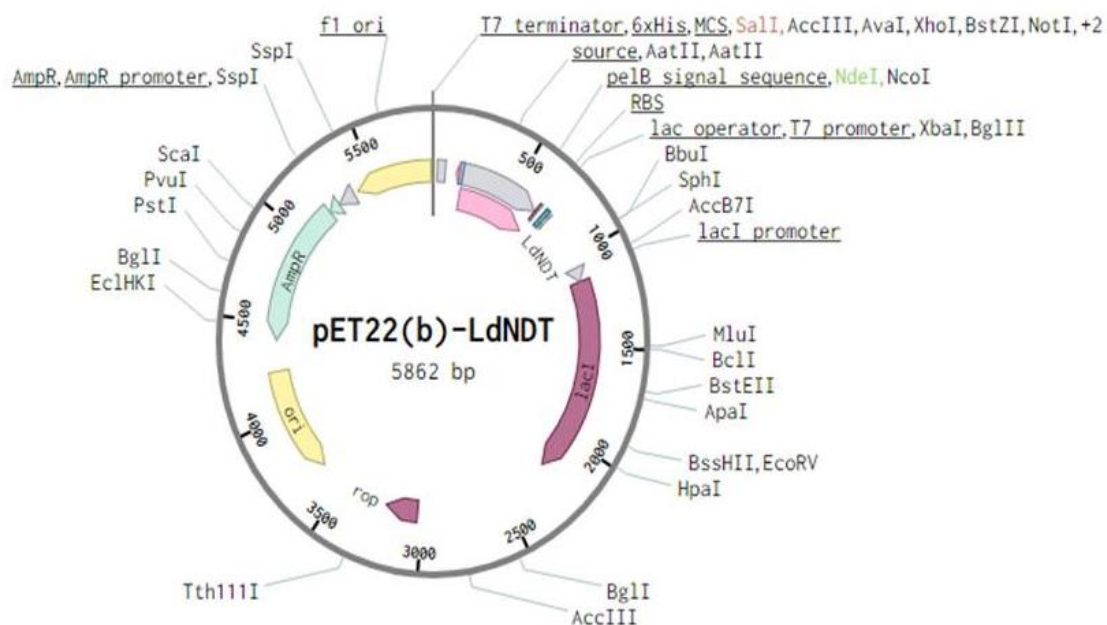
\* Las cepas tipo, *Lactobacillus delbrueckii*, *Lactobacillus lactis*, *Lactobacillus leichmannii* y *Lactobacillus bulgaricus* exhiben homología de ADN-ADN de 90-100% entre sí. Estas cepas se consideran así pertenecientes a una misma especie que, según las reglas de prioridad, conservarían el nombre de *Lactobacillus delbrueckii*<sup>103</sup>.

Con la a nueva construcción, con el gen *ndt2*, se transformaron células competentes de *E. coli* BL21(DE3) y el producto de transformación se sembró en placas de Petri con medio LB con ampicilina. Los clones positivos se analizaron para comprobar la incorporación del inserto y para estudiar su capacidad de expresar la proteína recombinante *LdNDT2* en medio LB según los protocolos del apartado 3.2.4.

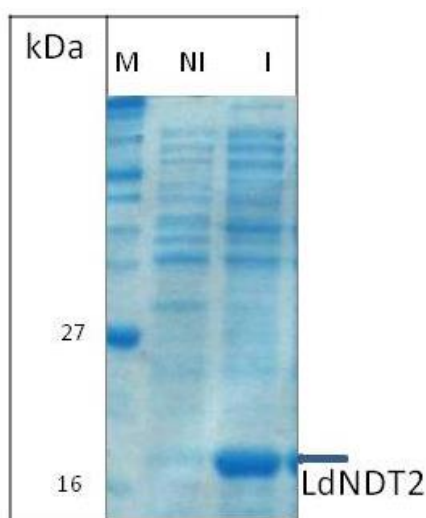
La Figura 28 muestra una banda muy significativa (carril I, Figura 28) de aproximadamente 18 kDa, que corresponde con el peso molecular esperado de la proteína de NDT2, la cual apenas se observa en el control no inducido (carril NI, Figura 28). Por otro lado, el estudio densitométrico de los geles de proteínas teñidos con Coomassie para determinar el porcentaje de proteína recombinante soluble demostró que aproximadamente el 50 %, de las proteínas solubles en el extracto, corresponden a la proteína recombinante NDT2, en comparación con un 4 % de las células no inducidas). Esto indica que el mecanismo de regulación de la expresión del gen clonado, en el sistema de expresión *E. coli* BL21(DE3)/pET22b(+) funciona de manera adecuada y que la proteína se está expresando en forma soluble en el citoplasma de las células.



**Figura 26:** Subclonaje del gen *ldndt2*. A: Plásmido recombinante pGEM-Ldndt2. B: PCR-screening de clon recombinante con la construcción pGEM-Ldndt2. Kb: marcador; Carril 2: Se observan la banda de 483pb, que indica la presencia del inserto.



**Figura 27:** Esquema del plásmido pET22b-Ldndt2

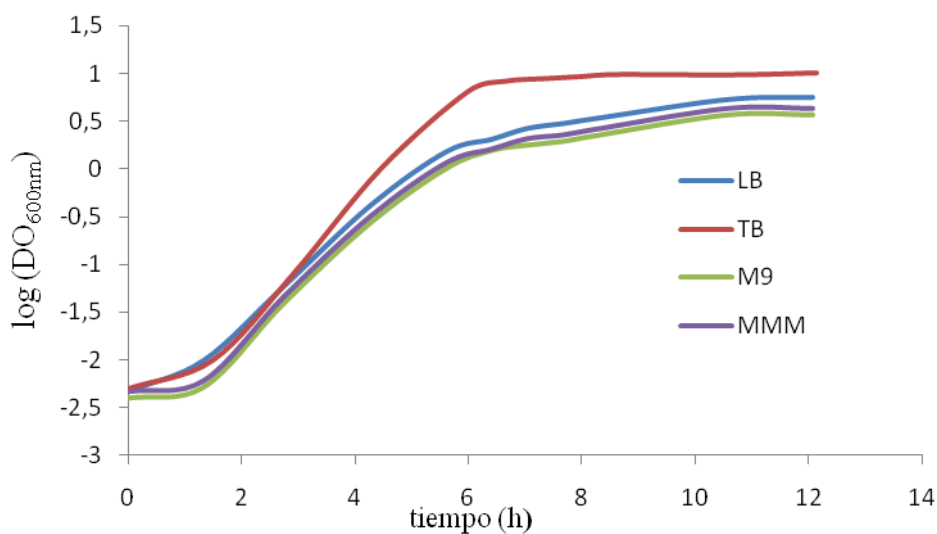


**Figura 28:** Análisis de proteínas citosólicas solubles por SDS-PAGE y teñidas con Azul de Coomassie. M: Marcadores de peso molecular; NI: No inducido I: Inducido.

#### 4.2.2. Producción de la proteína recombinante. Fermentaciones.

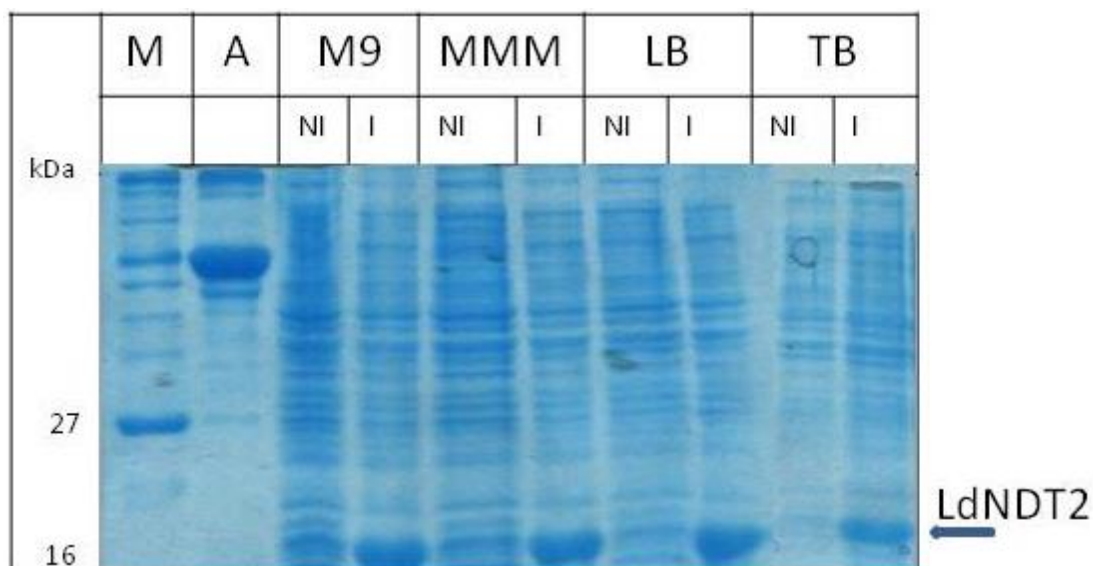
La cepa recombinante *E. coli* BL21 [pET22-Ldndt2] fue crecida, sin inducción, para determinar sus perfiles de crecimiento, en fermentaciones sumergidas, y obtener parámetros relacionados con el crecimiento, a partir de las curvas de crecimiento

(Figura 29). En los 4 medios utilizados (LB, TB, M9 y MMM), el clon recombinante entró en la fase de crecimiento logarítmico 2h después de la inoculación y en fase estacionaria después de 6 horas. La velocidad máxima de crecimiento ( $\mu_{max}$ ) del cultivo en medio LB fue de 0,4995 y de 0,6438 en medio TB y el tiempo de duplicación de 1,387 h en medio LB y de 1,076 h en medio TB. En cuanto a los medios sintéticos, la  $\mu_{max}$  en M9 fue de 0,4847 y en el medio mínimo modificado (MMM) de 0,4821; con tiempos de duplicación de 1,430 h y de 1,437 h, respectivamente. Finalmente, en los cuatro medios, la estabilidad del plásmido estuvo por encima del 90% tras 16 horas de fermentación.



**Figura 29:** Perfiles de crecimiento del clon recombinante en los medios Terrific broth (TB), Luria broth (LB), medio mínimo (M9) y medio mínimo modificado (MMM).

Para determinar el efecto de la composición del medio de cultivo sobre el crecimiento del clon recombinante *E. coli* BL21 [pET22Ldndt2], se analizó además la producción de biomasa, la producción de proteína total y la expresión de la enzima. La Tabla 21 muestra una comparativa de los cuatro medios ensayados en lo referido a densidad celular y nivel de expresión proteica. En los medios complejos (LB y TB), se observó un incremento de la densidad celular, siendo el medio TB el medio con mejores resultados. En cambio, en los medios minerales (MMM y M9) se obtuvieron menores rendimientos de biomasa. De la Tabla 21 se puede además concluir que se obtienen concentraciones superiores de proteína total (por gramo de biomasa celular) en los medios complejos, en comparación con los medios sintéticos. Sin embargo, la relación proteína total/biomasa para todos los medios es muy similar.



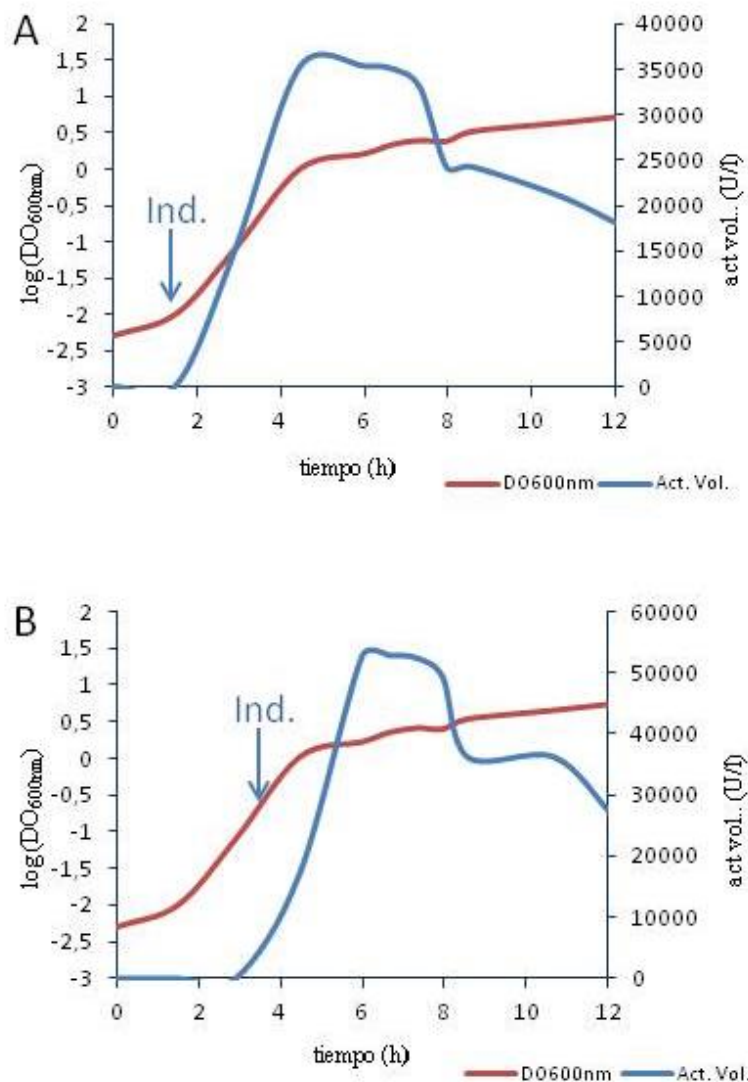
**Figura 30:** Análisis de proteínas citosólicas solubles por SDS-PAGE en los cuatro medios ensayados. **TB:** *Terrific broth*; **LB:** *Luria-Bertani*; **M9:** medio mínimo; **MMM:** medio mínimo modificado; **M:** Marcadores de peso molecular; **NI:** No inducido **I:** Inducido. **A:** *BSA*,

**Tabla 21:** Proteína total y producción de biomasa de los cultivos del clon recombinante *E. coli* BL21 (DE3)/ [pET22*Ldndt2*] en diferentes medios de cultivo.

Medio	Proteína total (g/L)	Biomasa (g/L)	Relación Proteína total/Biomasa
<b>LB</b>	1,85 ± 0,21	9,69 ± 0,22	0,19
<b>TB</b>	2,62 ± 0,24	14,20 ± 0,65	0,18
<b>M9</b>	1,42 ± 0,14	7,48 ± 0,50	0,19
<b>MMM</b>	1,47 ± 0,13	7,35 ± 0,38	0,20

Teniendo en cuenta los resultados del análisis de los perfiles de crecimiento, se decide continuar los estudios utilizando medio TB. La alta producción de biomasa en este medio tiene que ver con la presencia, entre sus componentes, de nutrientes de fácil digestión y ricos en péptidos y oligoelementos y de glicerol como fuente de carbono. De hecho, el medio TB contiene un poco más de Bacto® Tryptone°, pero casi 5 veces más de extracto de levadura, en comparación con el medio LB, lo cual se traduce en un mayor suministro de factores de crecimiento y de precursores de ácidos nucleicos. Por otro lado, contiene tampón fosfato para mantener un pH fisiológico durante el crecimiento exponencial; y el glicerol (que al metabolizarse no genera ácidos), a una concentración que no limita la velocidad de crecimiento.

° Es un digerido pancreático de caseína, se utiliza como fuente de nitrógeno en los medios de cultivos



**Figura 31:** Perfiles de crecimiento y actividad volumétrica del clon recombinante *E. coli* BL21 (DE3)/ [pET22*Ldndt2*] inducido en fase exponencial temprana (A) y en fase exponencial media (B). La flecha señala el momento aproximado de la inducción (Ind.).

Teniendo en cuenta el perfil de crecimiento, en medio **TB**, se valoró el mejor momento para la inducción. Los cultivos fueron entonces inducidos en fase exponencial temprana (2h) y en fase exponencial media (3,5 h) (Figura 31), con IPTG (1mM). En la Tabla 22 se resumen los resultados obtenidos. La comparación denota que, induciendo en estos dos momentos de la fase exponencial, la productividad de biomasa (gBS/L/h) es muy similar, lo que podría indicar que la expresión de la proteína no afecta la velocidad de crecimiento. No obstante, inducir los cultivos en fase exponencial media, parece ser más efectiva, a juzgar por la actividad volumétrica (U.L<sup>-1</sup>), 54480 unidades de actividad



enzimática por litro de caldo fermentativo, que es un 54,5% superior a la obtenida cuando se induce al inicio de fase exponencial.

**Tabla 22:** Resumen de los resultados de las inducciones en medio TB (matraz). **BS:** Biomasa seca; **DO:** densidad óptica.

<b>Parámetro*</b>	<b>Fase log temprana</b>	<b>Fase log media</b>
Máxima DO (600nm)	8,31	9,35
Biomasa (g BS.L <sup>-1</sup> ) ♦	3,24	3,64
Productividad (g BS.L <sup>-1</sup> .h <sup>-1</sup> )	0,40**	0,45**
Actividad volumétrica (kU.L <sup>-1</sup> )	35,25	54,48
Productividad enzimática (kU.L <sup>-1</sup> .h <sup>-1</sup> )	4,40	6,81

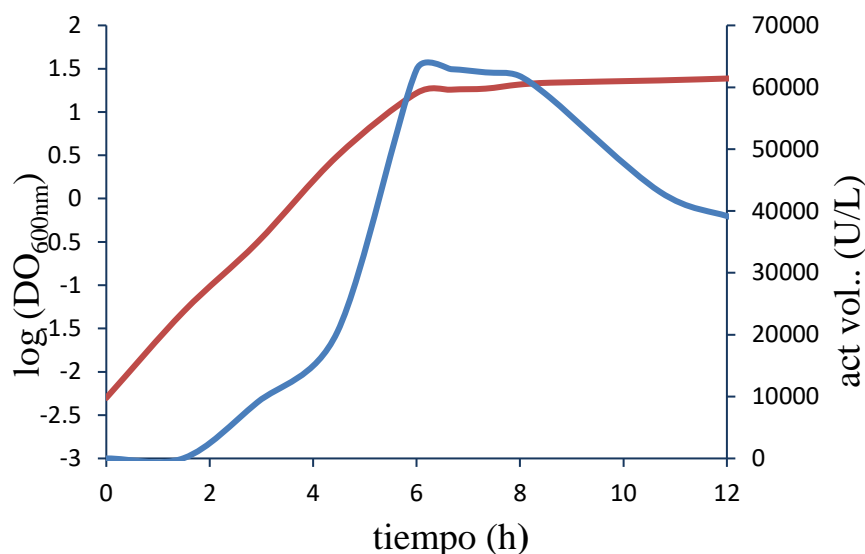
\*4 horas post-inducción.

\*\*8 horas de fermentación.

#### 4.2.3. *Ensayos preliminares de producción en biorreactor. Operación en discontinuo.*

El objetivo de estos ensayos preliminares es extrapolar los resultados obtenidos a nivel de matraz a un biorreactor de 7L con un volumen de trabajo de 5L para, por una parte tener una idea de los costes de producción y además, disponer de suficiente cantidad de la enzima para los ensayos de purificación y síntesis de clofarabina. Para ello, se utilizó el medio TB, en condiciones similares a las empleadas en los ensayos a pequeña escala, controlando la concentración de O<sub>2</sub> disuelto, fijándola en 30 % (agitación en cascada) y el pH, fijándose en 7,0 mediante la adición de NaOH (2N) o HCl (2N). Por otro lado, teniendo en cuenta los resultados obtenidos a nivel de matraz, las inducciones se realizaron cuando el cultivo alcanzó la fase exponencial media, esto es a una densidad óptica de DO<sub>600nm</sub> = 1,0 (Figura 32), añadiendo 1 mM de IPTG al biorreactor.

♦ El método empleado de rutina para medir la biomasa de los cultivos se basa en la determinación espectrofotométrica a 600 nm de la turbidez de la solución del medio de cultivo. Luego, se puede convertir este valor a concentración de biomasa (gBSL<sup>-1</sup>) multiplicando el valor de DO<sub>600 nm</sub> por 0,39 g/L (un factor obtenido a partir de una correlación entre turbidez y peso seco celular).



**Figura 32:** Curva de crecimiento y perfil de inducción de *E. coli* BL21 [pET22ldndt2], en medio TB en biorreactor de 7 L (volumen operacional: 5 L). La inducción se realizó con 1 mM de IPTG. En rojo, la actividad volumétrica y en azul la biomasa. Act vol: actividad volumétrica.

A esta escala se reproducen los perfiles de crecimiento y de expresión obtenidos a nivel de matraz, obteniéndose una velocidad máxima de crecimiento ( $\mu_{max}$ ) del cultivo en medio TB de  $0,6137 \text{ h}^{-1}$  y un tiempo de duplicación de  $1,129 \text{ h}^{-1}$ . Por otra parte, la actividad volumétrica muestra un incremento constante, desde el momento de la inducción (4,3-5h) hasta 4 horas post-inducción. A partir de este punto se observó una rápida disminución (degradación) de la actividad volumétrica, lo que podría indicar deficiencia de nutrientes, presumiblemente, de uno o más aminoácidos esenciales. El agotamiento de ciertos aminoácidos puede conducir a cambios en los flujos metabólicos, en la síntesis proteica y, posiblemente, también conduce a determinadas respuestas, a corto plazo, antes de que la célula induzca el conjunto de enzimas para la síntesis de este aminoácido esencial. Éstas podrían incluir, la degradación de la proteína recombinante, para proporcionar los aminoácidos necesarios para la síntesis de proteínas esenciales. El nivel de producción de *LdNDT2*, en estas condiciones alcanzó **163 mg/L<sup>♠</sup>**, en el punto más alto de producción. Por lo tanto, las células deben ser cosechadas antes de que se produzca la deficiencia de nutrientes). En la Tabla 23 se resumen los resultados obtenidos a nivel de biorreactor operado en *batch*, sin alimentación.

<sup>♠</sup> Calculado a partir del análisis densitométrico de la fracción soluble, en geles de poliacrilamida. Aproximadamente, la banda correspondiente a la proteína *LdNDT2* representaba, en este caso, el 31-32 % del total de las proteínas solubles.

**Tabla 23:** Resumen de los resultados de las inducciones en medio TB (Biorreactor operado en *batch*). **BS:** Biomasa seca; **DO:** densidad óptica.

<b>Parámetro</b>	
Máxima DO (600nm)	21,8
Biomasa (g BS.L <sup>-1</sup> )	8,50
Productividad (g BS.L <sup>-1</sup> .h <sup>-1</sup> )	1,06
Actividad volumétrica (kU.L <sup>-1</sup> )	62,50
Productividad enzimática (kU.L <sup>-1</sup> .h <sup>-1</sup> )	7,81

#### 4.2.4. Ensayos preliminares de producción en biorreactor.

##### *Operación en discontinuo alimentado*

El rendimiento de la producción de proteínas recombinantes depende tanto de la concentración de la biomasa como de la productividad celular específica. En este contexto, para lograr una alta concentración de biomasa en el cultivo, el sistema de cultivo discontinuo alimentado es la opción preferida para obtener cultivos de alta densidad<sup>105</sup>. Las estrategias de alimentación incluyen métodos de retroalimentación indirecta (tales como **pH-stat\*** y **DO-stat\***), métodos de alimentación predeterminados (tales como alimentación exponencial) y alimentación de acuerdo con la absorción o demanda del sustrato<sup>106</sup>.

Durante la fermentación, a medida que el cultivo bacteriano crece, la pO<sub>2</sub> disminuye debido al aumento en la captación de O<sub>2</sub> por el cultivo. En la Figura 34 se muestra una representación de la evolución de este (pO<sub>2</sub>) y otros parámetros en el biorreactor operado en discontinuo. En la misma se puede apreciar que el cultivo comenzó con una saturación de oxígeno cercana al 80%, y con el avance de la fermentación, la pO<sub>2</sub> fue disminuyendo hasta estabilizarse, entre las 4-5 horas (punto **A** de la Figura 34), al 30% de saturación. En este punto, la velocidad de agitación comenzó a aumentar exponencialmente, y 4-5 horas después, se estabiliza (punto **B** de la Figura 34), coincidiendo con una disminución brusca del consumo de oxígeno.

\* Alimentación de nutriente cuando el pH empieza a aumentar, debido a el agotamiento del sustrato.

\* Alimentación de nutriente cuando el oxígeno disuelto (OD) empieza a aumentar, debido a el agotamiento de la fuente de carbono.

En base a los perfiles de consumo de oxígeno observados, se consideró iniciar los ensayos de producción del enzima *LdNDT2*, en discontinuo alimentado, controlando la alimentación en base al nivel de oxígeno disuelto, el denominado **DO-stat**. Bajo esta estrategia de alimentación, las células de *E. coli* se cultivaron durante 20 h, desde 0,02 g.L<sup>-1</sup> hasta **35,23 g.L<sup>-1</sup>**, y se obtuvieron unas **90 unidades de Densidad Óptica**.

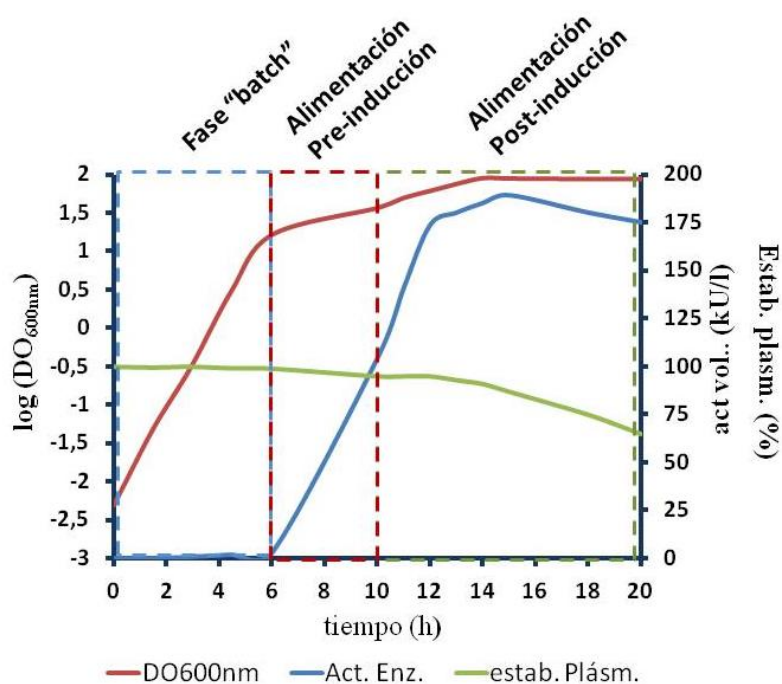
El crecimiento y la alimentación en este cultivo se dividieron en tres fases (Figura 33). Una fase de crecimiento no limitante, en “*batch*”, que duró hasta que se alcanzaron alrededor de 6 gramos de biomasa seca por litro de cultivo (BS.L<sup>-1</sup>), a una velocidad de crecimiento de 0,60 h<sup>-1</sup>, momento en el cual comenzó la fase de **alimentación “pre-inducción”**, con el medio de alimentación concentrado. Este punto coincidió con el mencionado **punto A** de la Figura 34 y la tercera fase, la de **alimentación “post-inducción”**, comenzó aproximadamente a las 10 horas de cultivo, cuando la concentración de biomasa alcanzó los **14 gramos de peso seco por litro** de medio de cultivo, momento en el que se indujo el cultivo con IPTG.

El cultivo inducido se mantuvo creciendo exponencialmente durante cuatro horas, aunque con una tasa de crecimiento menor ( $\mu = 0,1 \text{ h}^{-1}$ ). Al final de esta fase se obtuvieron aproximadamente 35 g de BS.L<sup>-1</sup>, unas 90 unidades de densidad óptica, es decir se cuadruplicó el rendimiento (en biomasa) respecto a la fermentación discontinua, no alimentada (Tabla 24). Por otro lado, la actividad enzimática volumétrica fue máxima a las cuatro horas post inducción (Figura 33), a partir de este momento la actividad volumétrica dejó de aumentar, a pesar de que la biomasa aun continuó aumentando. En este caso, la explicación del cultivo no alimentado, referida a deficiencia de nutrientes, no es factible, toda vez que el cultivo ha estado alimentado, con soluciones concentradas de nutrientes (extracto de levadura, (NH<sub>4</sub>)SO<sub>4</sub>, MgSO<sub>4</sub>·7H<sub>2</sub>O y glicerol), que consiguen mantener el crecimiento exponencial del cultivo.

Por otro lado, la explotación comercial de cepas recombinantes depende, en gran medida, de la estabilidad de los plásmidos y de los niveles de expresión. La estabilidad de las construcciones recombinantes está determinada por factores como: el número de copias del plásmido, los patrones de replicación, el tipo de sustrato, la composición del medio, la temperatura del cultivo y toxicidad de la proteína expresada. En la Figura 33 junto a los perfiles de crecimiento, se ha graficado la estabilidad (retención) del plásmido recombinante, durante el cultivo. Cuando el cultivo en matraz y en

bioreactor fue operado en modo discontinuo (*batch*), la retención de los plásmidos al final de los ciclos fermentativos estuvo por encima del 90%, pero en las fermentaciones semicontínuas (*batch*-alimentado), a las 4-5 horas post-inducción, la inestabilidad plasmídica comenzó a hacerse significativa, al punto de que una tercera parte de las células podría no tener plásmidos y por tanto, no eran productivas (Figura 33).

Una de las posibles explicaciones para la inestabilidad del plásmido esta en la existencia de un diferencial, significativo, de velocidad de crecimiento entre las células que contienen y las que no contienen plásmidos. Las células que retienen plásmidos están obligadas a asignar una parte de sus recursos para apoyar actividades relacionadas con plásmidos que incluyen mantenimiento y replicación del ADN plasmídico y expresión de genes plasmídicos. De cualquier manera, el cómo estabilizar la construcción recombinante parece ser una de las vías para incrementar la productividad (actividad volumétrica) de la *LdNDT2* producida en modo semicontínuo.



**Figura 33:** Producción de *LdNDT2* en un cultivo discontinuo alimentado. En cultivo se alimenta antes y después de la inducción con IPTG. **Act. Enz.:** Actividad volumétrica (kU/L); **Estab. Plasm.:** Estabilidad del plásmido (%).

#### 4.2.5. Comparación de la producción de LdNDT2 en cultivos en discontinuo y discontinuo alimentado.

La Tabla 24 presenta los resultados obtenidos en relación con la productividad de LdNDT2 en las tres estrategias utilizadas, es decir, usando fermentaciones discontinuas (en matraz y en biorreactor) y en fermentaciones discontinuas alimentadas (en biorreactor). En cultivos semicontinuos se alcanzaron 90 unidades de densidad óptica, unos 35 g BS.L<sup>-1</sup>, unas 4 veces superior al rendimiento obtenido sin alimentación. Aunque los parámetros de productividad (de biomasa y de enzimas) no mejoraron al mismo nivel, lo cual significa que estas fermentaciones preliminares, aun no son óptimas. Para futuras fermentaciones se podrían resolver cuestiones como: **¿Cuándo se debe inducir el cultivo semicontínuo?**

Cuando analizábamos los rendimientos obtenidos en los cultivos discontinuos, observamos que inducir a diferentes tiempos, modificaba los perfiles de crecimiento y de expresión. En el semicontinuo, una vez adicionado el IPTG, también se observó una disminución de la velocidad de crecimiento. Se conoce que el uso de un promotor fuerte dentro del sistema de expresión, a menudo conduce a una síntesis brusca de proteína recombinante. Esto tiene como resultado la inhibición del crecimiento celular producto de la síntesis de una gran cantidad de proteína recombinante<sup>107</sup>. Luego, observamos que en la fermentación semicontinua, a las 4 horas después de la inducción, se alcanzó la máxima actividad volumétrica y el cultivo aún continuaba creciendo. En este sentido, estudiar la posibilidad de inducir 4 horas antes de alcanzar el máximo de biomasa teórico, podría suponer una alta producción volumétrica de la enzima.

El nivel de producción de LdNDT2, en estas condiciones alcanzó 606 mg de la proteína por litro de caldo fermentativo<sup>♠</sup>, en el punto más alto de producción, unas 189 kU/L, lo que supone un incremento significativo (casi 4 veces) en comparación con la proteína obtenida en el cultivo en discontinuo.

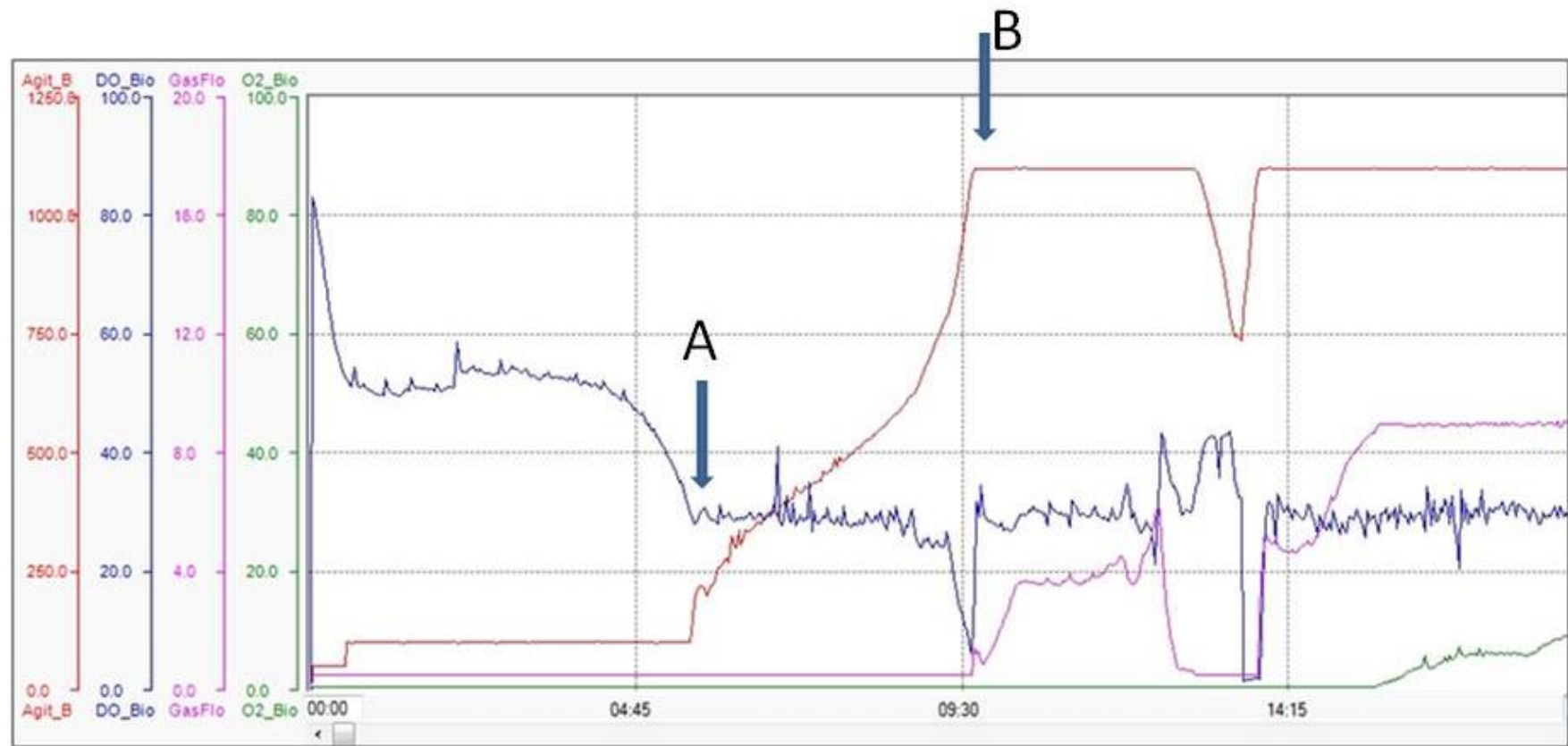
---

<sup>♠</sup> Calculado a partir del análisis densitométrico de la fracción soluble, en geles de poliacrilamida. Aproximadamente, la banda correspondiente a la proteína LdNDT2 representaba, en este caso, el 32 % del total de las proteínas solubles.

**Tabla 24:** Comparación de la producción de LdNDT2 en cultivos en discontinuo (matraz y biorreactor) y discontinuo alimentado (biorreactor).

<b>Parámetro*</b>	<b>Discontinuo (matraz)</b>	<b>Discontinuo Biorreactor</b>	<b>Discontinuo alimentado Biorreactor*</b>
DO 600nm Máxima	9,35	21,8	90,34
Biomasa (g BS.L <sup>-1</sup> )	3,64	8,50	35,23
Productividad (g BS.L <sup>-1</sup> .h <sup>-1</sup> )	0,45	1,06	2,20
Actividad volumétrica (kU.L <sup>-1</sup> )	54,48	62,50	189,02
Productividad enzimática (kU.L <sup>-1</sup> .h <sup>-1</sup> )	6,81	7,81	11,81

\*A las 16 horas de fermentación.



**Figura 34:** Registro gráfico de la evolución de los parámetros fermentativos a escala de bioreactor de 5L. A: 30% de saturación de O<sub>2</sub>, inicio de la alimentación pre-inducción; B: Alimentación post-inducción.

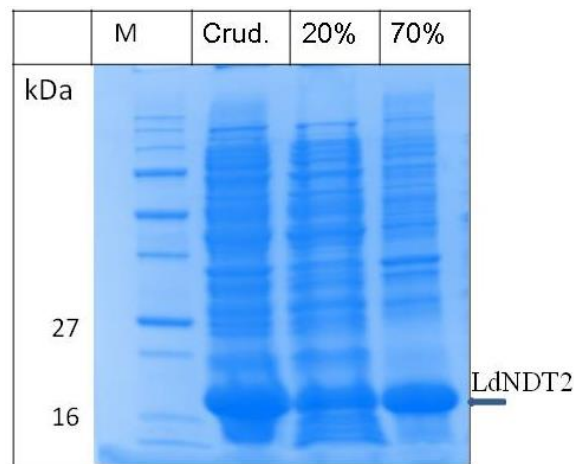


#### 4.2.6. Purificación de la *LdNDT2* recombinante producida en *E. coli*

El crudo (lisado celular clarificado) contiene, además de la proteína *LdNDT2* otras proteínas acompañantes que podrían de alguna manera afectar la actividad de la enzima, por ejemplo, compitiendo por los sustratos. Se ha documentado que la promiscuidad de las NPs, puede hacer que estas enzimas acepten un amplio rango de sustratos (nucleósidos) con modificaciones en la porción del azúcar, ej. 2'- o 3'-NH<sub>2</sub>, 2'-F (ribo- o arabino-) y 2'-OH (arabino)<sup>108</sup>. De hecho, nucleósidos antileucémicos como Cladribina, Fludarabina, Clofarabina y Nelarabina han sido sintetizados usando enzimas del tipo NPs<sup>109, 110</sup>. De modo que la presencia de estas enzimas en el crudo podría afectar el rendimiento de la reacción. En este trabajo para demostrar, por un lado, que la actividad *2'-desoxi-2'-fluororibosil-transferasa* de interés se corresponde con la acción de la *LdNDT2* y además, valorar aspectos relacionados con las cinéticas y las especificidades de esta, se ha procedió a la purificación de la enzima.

##### A. Precipitación fraccionada con sulfato amónico.

Uno de los objetivos en la purificación de proteínas es la obtención del máximo rendimiento de proteína pura en el menor número de pasos, lo cual no solo favorece la estabilidad de la proteína, sino también disminuye los costes de producción. La mayoría de los procesos de purificación requieren de más de un paso para conseguir el grado de pureza deseado, lo que puede provocar pérdida del producto. En este caso, se han empleado dos pasos cromatográficos para purificar la *NDT2* de *Lactobacillus delbrueckii*, previa precipitación fraccionada con sulfato de amonio.



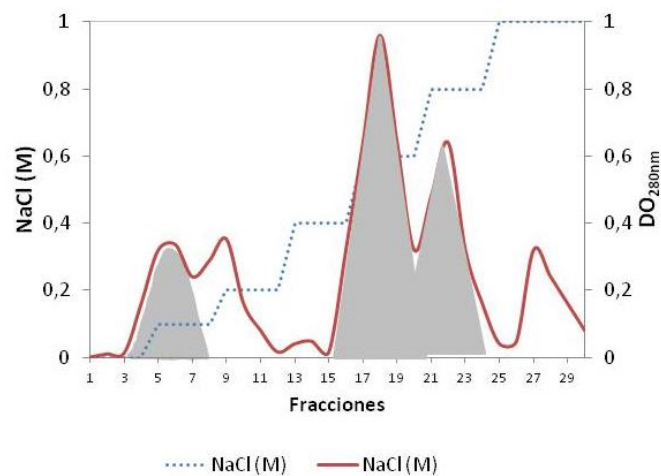
**Figura 35:** Gel SDS-PAGE de la precipitación fraccionada usando dos concentraciones de sulfato amónico (20 y el 70% de saturación). **M:** Marcadores de peso molecular; **Crud:** Extracto crudo (fracción soluble).

El primer paso de purificación consistió en la precipitación usando dos concentraciones de sulfato amónico (20 y el 70% de saturación de sal) y en ambas fracciones se localizó la enzima *Ld*NDT2 (Figura 35). Esto se debe a que ésta es una técnica de purificación muy poco precisa, utilizada sobre todo para concentrar y estabilizar proteínas. A la luz de estos resultados, y en el sentido de recuperar la mayor cantidad posible de actividad enzimática, se escogió la precipitación, en un solo paso, usando un 70% de saturación de sal. Los rendimientos de esta etapa de purificación se muestran en la tabla resumen de la purificación de la enzima (Tabla 25).

### B. Purificación mediante cromatografía de intercambio iónico.

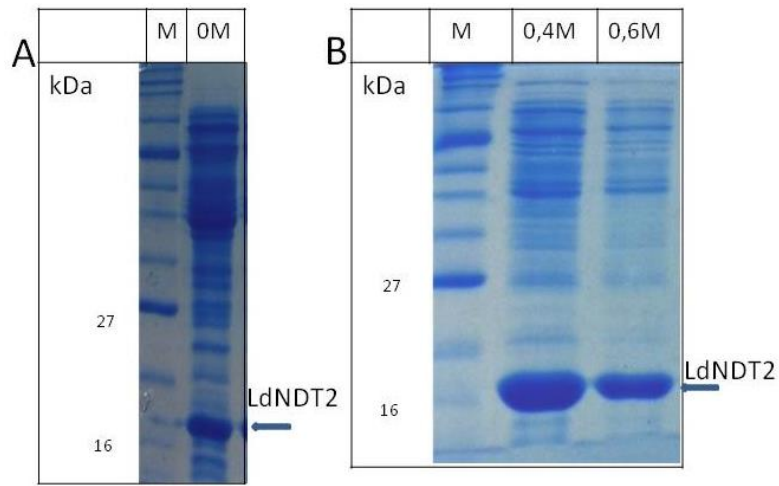
Tras la precipitación fraccionada con sulfato amónico se llevaron a cabo dos cromatografías, consecutivas, de intercambio iónico. Esta cromatografía basa su principio de separación en la relación que existe entre la carga neta de las proteínas y el pH de la fase móvil, ya que ésta relación es única para cada proteína. La interacción entre las moléculas cargadas y la matriz es controlada para favorecer la unión o elución de moléculas específicas y lograr así la separación de la molécula de interés.

A la vista de los datos proporcionados por el servidor ExPASy\* (*Expert Protein Analysis System*) acerca de la estructura y que indica que el punto isoeléctrico (pI) de la proteína es 4,65, se llevaron a cabo las cromatografía de intercambio iónico, utilizando un intercambiador aniónico y un gradiente escalonado de NaCl (0M-1M), en la primera cromatografía y de 0,3-0,5 M en la segunda cromatografía. En la Figura 36 se muestran los perfiles de elución de la primera cromatografía y el electroforegrama de las fracciones positivas se muestra en la Figura 37.

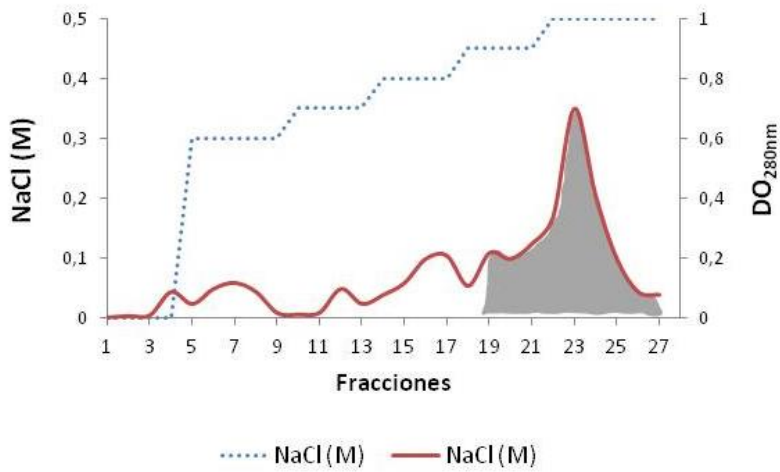


**Figura 36:** Primera cromatografía de intercambio iónico de la *Ld*NDT2 producida por *E. coli* BL21(pET22-*ndt2*). Cromatografía de intercambio iónico (Q-Sepharose).

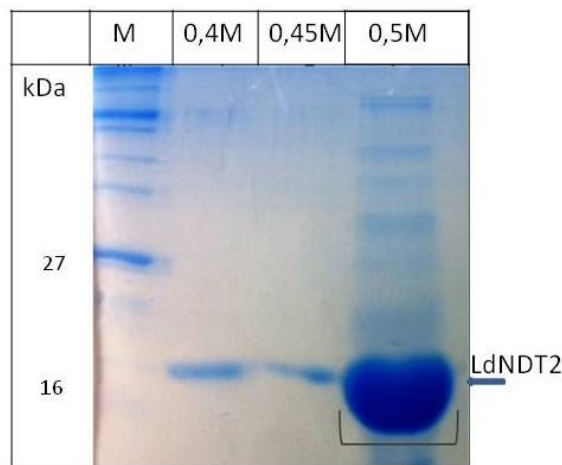
\* <http://web.expasy.org/protparam/>



**Figura 37:** Visualización del estado de la purificación de la LdNDT2 en los eluatos positivos, de la primera cromatografía de intercambio iónico. **M:** Marcadores de peso molecular; **A:** Fracción no retenida; **B:** Perfil de las fracciones eluidas con 0,4 - 0,6M de NaCl.

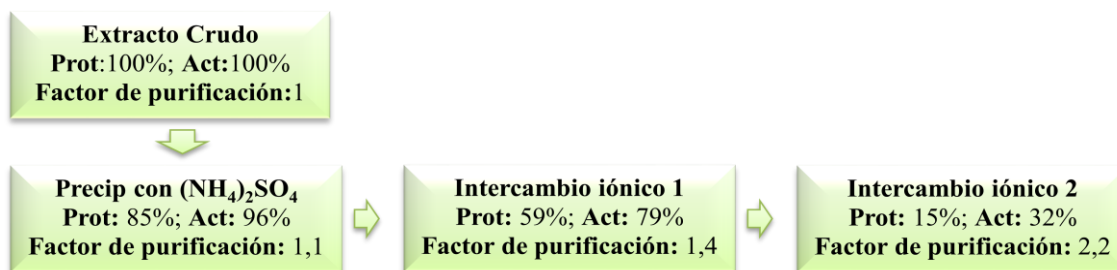


**Figura 38:** Segunda cromatografía de intercambio iónico de la LdNDT2 producida por *E. coli* BL21(pET22Ldndt2). Cromatografía de intercambio iónico (Q-Sepharose).



**Figura 39:** Electroforegrama de segunda cromatografía de intercambio iónico

La purificación de la enzima fue llevada a cabo con éxito como se demuestra en el SDS-PAGE de la Figura 39 donde la LdNDT aparece en la última calle (NaCl; 0,5M), prácticamente pura (más del 95%) y concentrada. El esquema final de purificación quedó como sigue (Figura 40):



**Figura 40:** Esquema de purificación propuesto. **Prot:** Cantidad de proteína; **Act:** Actividad enzimática total.

Durante la precipitación con  $(\text{NH}_4)_2\text{SO}_4$  y los dos pasos cromatográficos posteriores se fueron sumando datos con el objetivo de obtener una tabla resumen (Tabla 25) de todo el proceso de purificación seguido mediante la medición de la actividad nucleósido desoxirribosiltransferasa, concentración de proteínas y la medición densitométrica de la pureza de la proteína en geles de poliacrilamida.

**Tabla 25:** Resumen del proceso de purificación de la LdNDT2 recombinante $\diamond$ .

<b>Fracción/etapa</b>	<b>Vol. (mL)</b>	<b>Prot (mg)</b>	<b>Act. Total (kU)</b>	<b>Act. Esp. (kU/mg)</b>	<b>Factor de purif. (veces)</b>	<b>Rend. (%)</b>
Extracto Crudo*	250	255,1	51,45	0,20	1,00	100
Precipitación en $(\text{NH}_4)_2\text{SO}_4$ **	50	216,8	49,56	0,22	1,10	96,3
Intercambio iónico 1	25	150,3	40,59	0,27	1,35	78,9
Intercambio iónico 2	8	38,25	16,23	0,43	2,15	31,5

\* Sobrenadante del disruptado celular (proteínas solubles)

\*\* Resultado de la precipitación utilizando un 70% de saturación.

$\diamond$  En la Tabla 25

**Tabla 25:**

**Vol:** Volumen (mL) se refiere al volumen total de solución medido en la etapa de purificación.

**Prot:** Proteína total (mg).

**Act. total:** Actividad total (unidades)

**Act. Esp.:** Actividad específica (unidades/mg).

**Factor de Purif.:** Grado de purificación (veces) - En este caso se refiere al aumento de la actividad específica, es decir, el grado de purificación se obtiene dividiendo la actividad específica en cualquier etapa por la actividad específica del extracto crudo inicial.

**Rend.:** Rendimiento (%) - el rendimiento se basa en la recuperación de la actividad después de cada paso. La actividad del crudo original se fija arbitrariamente al 100%. El rendimiento (%) se calcula a partir de la actividad total (unidades) en cada paso dividido por la actividad total (unidades) en el crudo, multiplicado por 100.

Con estos resultados a partir de la producción en un biorreactor de 5 litros, operado en modo semicontinuo, es posible obtener **aproximadamente 1,42 g** de proteína purificada (más del 95% de pureza, según análisis densitométrico de los geles de SDS PAGE), es decir el 15% de la proteína total del lisado celular clarificado.

#### 4.2.7. Costes de producción del biocatalizador.

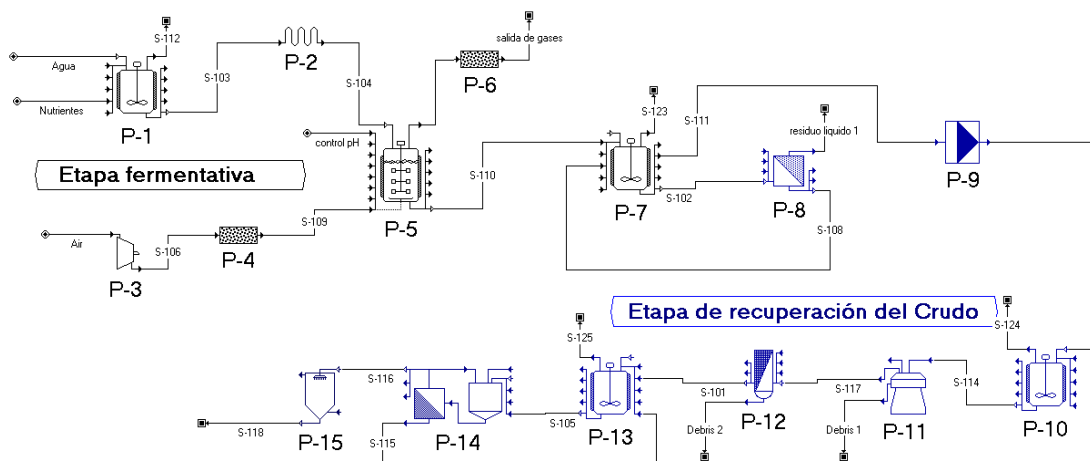
Para analizar los costes de producción del biocatalizador, el primer aspecto a tener en cuenta es la formulación final del mismo, es decir, seleccionar en qué forma será usado. Como se verá más adelante (Figura 53), en este trabajo se obtuvo el producto de interés, con un aceptable nivel de pureza (< 3% de impurezas), usando tanto los crudos (lisados clarificados), como los enzimas purificados. En este sentido, la formulación de elección, en este trabajo, fue la enzima aislada (lisado clarificado y concentrado), pero no purificada, sometida a tratamientos que garanticen altos niveles de actividad y estabilidad. Esta elección es muy común entre los bioprocesos, dado que minimiza los costes de producción<sup>111</sup>, y de acuerdo al análisis realizado por Straathof y col.<sup>112</sup> cerca del 40% de las reacciones biocatalíticas industriales utiliza los biocatalizadores en este formato.

Para los cálculos se tuvieron en cuenta las dos etapas principales: **fermentación** y **aislamiento**. La Figura 41 resume el proceso general de producción y recuperación de la proteína recombinante *LdNDT2*, en fermentación semicontinua. Para los cálculos se han tenido en cuenta los resultados obtenidos a nivel de biorreactor de **5 litros**, asumiendo que se obtienen **35 g de BS/L** y de éste aproximadamente el **32%** es la *LdNDT2* soluble, “sin purificar”, en el lisado clarificado. Por otro lado, también se asumió que solo se realiza un ciclo fermentativo por semana (72 horas por *lote*) y que el sistema sería operado por dos técnicos a tiempo completo. No se incluyeron en los cálculos los aspectos relacionados con las Buenas Prácticas de Fabricación (BPF).

En este trabajo se estimaron los costes de de producción de la *LdNDT2* recombinante, en un proceso que se resume de la siguiente manera:

- 1- El medio de fermentación se prepara en un tanque (mezclador) y es esterilizado a 121°C.
- 2- El fermentador se inocula y es operado en modo semicontinuo durante 38 horas, a 37°C.
- 3- La expresión de la *LdNDT2* se induce usando IPTG.

- 4- La aireación se lleva a cabo usando un compresor a razón de 1vvm (volumen por volumen de reactor por minuto).
  - 5- La potencia de mezclado será de 2,5 W/L.
  - 6- Métricas de la fermentación:
    - a. Densidad celular; 35 g BS/L
    - b. La proteína LdNDT2 representa el 1,72 % de la biomasa seca (32% de las proteínas presentes en el extracto crudo clarificado).
  - 7- Duración del proceso, incluyendo la preparación, la fermentación y la limpieza: 3 días (72 horas).
  - 8- Al final de la fermentación se recupera la biomasa usando microfiltración y las células se resuspenden en tampón. La suspensión celular se lisa usando un homogenizador. El debris celular se separa de la fracción soluble usando una centrifuga de discos y la fracción soluble se filtra para asegurar la eliminación de los restos celulares y se concentra (y dializa) mediante ultrafiltración.
- Nota: En este punto se prescinde de la concentración con sulfato de amonio, toda vez que ese efecto se consigue mediante el proceso de ultrafiltración. Se ha de enfatizar que el equipo de filtración es el mismo en todos los procesos, solo se requiere un cambio de la membrana filtrante (microfiltración, ultrafiltración).*
- 9- El concentrado diafiltrado se seca usando un secador de bandejas y se acondiciona convenientemente.



**Figura 41:** Esquema detallado del proceso de producción del Biocatalizador. **P-1, P-7, P-10 y P-13:** mezcladores; **P-2:** Esterilización continua; **P-3:** bomba de aire; **P-4 y P6:** Filtros de aire; **P-5:** Biorreactor; **P-8:** Sistema de filtración; **P-9:** Homogenizador; **P-11:** centrifuga; **P-14** sistema de diafiltración, **P-15:** Sistema de secado.

## A. Estimación de los costes de producción de la enzima.

### CAPEX

- Adquisición de Equipos
- Adq. medios transporte
- Adq. de Maquinaria
- Adq. de Instalaciones

Para la estimación de los costes de producción del biocatalizador se tuvieron en cuenta las recomendaciones de Tufvesson y col.<sup>45, 46</sup> y Lima-Ramos y col.<sup>113</sup>, y que ponen el énfasis en que éstas estimaciones deben hacerse al inicio de cada proyecto, incluso cuando aún no se disponen de todas las especificaciones ( u otros datos).

En el caso que nos ocupa, se asumió que la proteína iba a ser producida, en una planta piloto, con toda la capacidad y los servicios instalados, y que ésta estaría ubicada en territorio de la Unión Europea. Pero, **¿cuánto nos costará escalar la producción de la LdNDT2?**. Pues, para ello se han de tener en cuenta, las inversiones de capital (CAPEX\*) y las operaciones (OPEX\*) necesarias para producir los lotes necesarios del biocatalizador.

### OPEX

- Mantenimiento equipos y maquinaria
- Consumos
- Alquileres

Se utilizaron las métricas del proceso a escala de biorreactor de 5 L para estimar los costes de producción de **294 kg** de un crudo enzimático concentrado (liofilizado) con actividad *2'-desoxi-2'-fluororibosil-transferasa*, en 35 campañas (al año), en un biorreactor de 1,26 m<sup>3</sup> (80% de volumen de trabajo) operado en modo semicontinuo.

## B. Consideraciones generales.

Las consideraciones generales de la fermentación fueron:

- 🌍 La planta piloto, con todas las instalaciones necesarias estaría en Europa Occidental.
- 🌍 La producción estimada del concentrado con actividad 2'-desoxi-2'-fluororibosil-transferasa sería de **294 kg/año**.
- 🌍 La planta sería operada, en **35 campañas**, con una ocupación de la capacidad instalada de **2520 horas** al año.

---

\* El gasto de capital (**Capex**), es el coste de desarrollo o el suministro de componentes no consumibles para el producto o sistema.

\* **OPEX**, del inglés "Operating expense", es un coste permanente para el funcionamiento de un producto, negocio o sistema. Puede traducirse como gasto de funcionamiento, gastos operativos, o gastos operacionales.

Los costes y precios, en euros (€).

### C. El proceso.

- Capacidad del biorreactor: 1,26 m<sup>3</sup>, con un 80 % de volumen de trabajo.
- El fermentador se inocularía y se operaría en modo semicontinuo durante 36 horas.
- El tiempo total de una campaña, incluida la preparación, la fermentación, la recuperación y la limpieza sería de 3 días de trabajo.
- Se utilizaría un tanque para la preparación del medio de cultivo y de aquí el medio pasaría al biorreactor, previa esterilización (121°C) usando un intercambiador de calor.
- La expresión de la *LdNDT2* se induciría mediante la adición de IPTG.
- La potencia de mezclado sería de 3,75 kW/m<sup>3</sup>.
- El contenido de agua en la biomasa, tras la centrifugación, sería del 10% del peso seco celular.
- Las métricas de la fermentación:
  - Densidad celular: 35 gBS/L
  - El contenido de proteína sería de 0,05 g de proteína total/gBS
  - El nivel de expresión de la proteína recombinante sería de 0,017 g de *LdNDT2*/gBS.

### D. Servicios y costes operacionales para la producción *LdNDT2* production (35 batches)<sup>♦</sup>

#### Materia prima.

Para estimar el coste de las materias primas utilizadas en el bioproceso, la información fue obtenida a partir de la revisión de trabajos similares en revistas especializadas<sup>45, 113, 115</sup>, y las estimaciones de la materia prima consumida se obtuvo directamente de los balances de masa del proceso. Algunos precios han sido obtenidos a través de contacto personal con proveedores o de colegas de empresas química o biotecnológicas. En la Tabla 26 se muestra el precio de las materias primas utilizadas para la producción de la enzima recombinante.

---

<sup>♦</sup> El coste de operación es la suma de todos los gastos relacionados con materias primas, mano de obra, servicios, eliminación de residuos, gastos generales, etc..



**Tabla 26:** Coste del medio de fermentación (escala: 1000 litros) para la producción de LdNDT2. Se estimó el coste total para la producción de 35 lotes al año.

<b>Materia prima</b>	<b>Coste unitario (€/kg)</b>	<b>Cantidad (kg)</b>	<b>Coste (€)</b>
<b>Sales del medio</b>	2,96	527	1559,92
<b>Peptona</b>	3	546	1638
<b>Extracto de levadura glicerol</b>	1,8	1092	1965,6
<b>Agua</b>	0,01	35000	350
<b>Ampicilina</b>	38	3,5	133
<b>IPTG</b>	154	0,83	127,82
<b>Total</b>			<b>5901,74*</b>

**\* 168,62 €/lote.**

### Servicios

Los servicios incluyen los costes de la energía necesaria para el calentamiento o la refrigeración, la energía necesaria para la agitación, aireación y otros equipos del bioproceso, y que pueden estimarse a partir de los balances de masa y energía. En la Tabla 27 se resumen los costes de los servicios anuales para la producción del enzima, se incluyen además los costes del tratamiento y/o eliminación de residuos<sup>45</sup>.

**Tabla 27:** Servicios (0,1€/kWh\* ).

Operación	capacidad	MWh
Esterilización (121°C)	3,15 J/kg	0,65
Aeración (1vvm)	3,75 kW/m <sup>3</sup>	1,35
Agitación (500 rpm)	3,75 kW/m <sup>3</sup>	1,35
Tratamiento de residuos <sup>^</sup>	~12,35 m <sup>3</sup>	

Coste de los servicios: **359,7 €/lote**

**Coste total de los servicios (35 lotes): € 12589,5.**

\* European Energy Portal, <http://www.energy.eu>

<sup>^</sup> Por lo general el tratamiento de las aguas residuales suele costar entre 0,5–2 €/m<sup>3</sup>, y el tratamiento de residuos sólidos (no peligrosos) alrededor de 25 €/ton.<sup>116</sup>

### Otros costes operacionales.

A partir de los costes directos de la mano de obra se han inferido otros costes de producción o hacer estimaciones anuales de los costes de inversión. Así, por ejemplo, se pudieron estimar los costes de supervisión (costes operativos directos) y los costes indirectos (nóminas, control de calidad, etc.), que representan entre el 80-115% de los costes directos de la planta. Por otro lado, el mantenimiento anual (mano de obra y material para el mantenimiento), representa el 6-10% de la inversión de capital fijo. Luego están los costes operativos fijos (depreciación, impuestos, seguros, etc.) que suponen el 12-17% de la inversión anual.

### E. Gastos de Capital.

CAPEX (*capital expenditures*) o gastos de capital es la cantidad que se gasta para adquirir o mejorar los activos productivos (tales como edificios, maquinaria y equipos, vehículos) con el fin de aumentar la capacidad o eficiencia de una empresa.

#### Coste de los equipos

En la Tabla 28 se muestra el coste de los equipos principales de la planta de producción y esta información se utilizó, aplicando “ciertos factores”, para calcular la inversión de capital. El precio de los mismos se obtuvo a partir de la información solicitada a los proveedores, en programas de bioprocesos, como *Superpro Designer*, o consultando el sitio web *MatChe Inc.*<sup>▲</sup> Por otro lado, cuando no se dispuso del precio, para un equipo de determinada capacidad, pero se disponía de datos a otra escala, se hicieron aproximaciones usando la “regla de los 6 décimos”. La relación puede expresarse como se muestra en la Ecuación 3.

#### Ecuación 3:

$$\text{Coste del equipo B} = \text{Coste del equipo A} \left( \frac{\text{Capacidad del equipo B}}{\text{Capacidad del equipo A}} \right)^{0,6}$$

El exponente (0,6) de la ecuación se conoce como el factor coste-capacidad y como promedio, éste tiende a 0,6 y es por ello que esta relación se conoce también como la regla de los seis décimos.

A partir de la estimación de la inversión en equipos y aplicando ciertos multiplicadores se pudieron inferir, los costes anuales de mantenimiento (costes operativos directos), que representan del 6% al 10% de la inversión de capital fijo.

---

<sup>▲</sup> [www.matche.com](http://www.matche.com)

## Mano de obra.

Con los costes de capital instalado, como se ha descrito en la Tabla 28, se pudieron estimar los pagos anuales totales de capital y a partir de este inferir (teniendo en cuenta, los periodos de ocupación/inactividad ofrecidos por el programa *SuperPro Designer*), los costes de la mano de obra. En este trabajo los costes de la mano de obra se estimaron como una fracción de los pagos anuales totales del capital instalado (15%)<sup>117</sup> y los costes relacionados con Garantía de calidad (QA) / control de calidad (QC) se estimaron alrededor un 15% del coste total de la mano de obra. Así, el coste de la mano de obra se estimó en **12178,86 €/35 lotes**, y el Coste total de QC/QA/35 lotes: **1826,829 €**.

**Tabla 28:** Coste de los equipos principales.

Equipos	Unid.	Capacidad	Precio <sup>◇</sup>	TIC <sup>*</sup>	Anualidad (€/año) <sup>♣</sup>	Mantenimiento (€/año) <sup>♥</sup>	Otros (€/año) <sup>♣</sup>	Coste /lote
Intercambiador de calor (esterilización)	1	0,30m <sup>3</sup> /h	12500	62500	8898,8	889,9	1334,8	317,8
Tanque de mezcla y almacenamiento.	2	1,25 m <sup>3</sup>	3500	17500	2491,7	249,2	373,7	89,0
Compresor centrífugo	1	5,24 kW	13500	67500	9610,7	961,1	1441,6	343,2
Homogenizador	1	0,26 m <sup>3</sup> /h	2600	13000	1850,9	185,1	277,6	66,1
Filtro sin fin	1	0,63m <sup>2</sup>	1050	5250	747,5	74,7	112,1	26,7
Diafiltración	1	17,9 m <sup>2</sup>	6500	32500	4627,4	462,7	694,1	165,3
Microfiltración	1	6,43 m <sup>2</sup>	5500	27500	3915,5	391,5	587,3	139,8
Fermentador	1	1,26 m <sup>3</sup>	43900	219500	31252,4	3125,2	4687,9	1116,2
Centrífuga de disco	1	0,27 m <sup>3</sup> /h	3000	15000	2135,7	213,6	320,4	76,3
Liofilizador <sup>▽</sup>	1	0,09 m <sup>2</sup>	22000	110000	15661,8	1566,2	2349,3	559,4
<b>Total</b>								<b>2899,7</b>

<sup>◇</sup> Aplicando la Ecuación 3 para estimar los precios a partir de los datos obtenidos en SuperPro designer y de la literatura consultada<sup>45, 46, 113, 118, 119</sup>.

<sup>\*</sup> TIC – Costes totales de instalación, Lang Factor = 5; <sup>♣</sup>Tasa de interés anual: 7% y a 10 años. <sup>♥</sup> Los costes de mantenimiento representan un 10% del TIC.

<sup>♣</sup> Otras tasas incluidos los seguros, etc, representan el 15% del TIC;

<sup>▽</sup> 75 kg de agua/ciclo.

La Tabla 29 resume los principales costes operativos que incluyen el coste estimado de las materias primas, los servicios, los asociados a la mano de obra y los costes dependientes de las instalaciones (depreciación, mantenimiento, etc.).

**Tabla 29:** Costes operacionales anuales de la producción de la enzima *Ld*NDT2 (35 lotes).

<b>Índice de coste</b>	<b>€/35 lotes</b>
Materia prima	5901,74
Mano de obra	12178,86
Costes relacionados con la capacidad instalada (equipos) <sup>□</sup>	101489,5
Análisis/QC/QA	1826,83
Servicios	12589,5
<b>Total</b>	<b>133986,43</b>

**Tabla 30:** Resumen de los costes anuales para la producción de 294 kg del concentrado con actividad 2'-desoxi-2'-fluororibosil-transferasa.

<b>Índice de Coste</b>	<b>€/35 lotes</b>
<b>Producción anual (kg)</b>	<b>294,00</b>
<b>Costes de operación (€/35 batches)</b>	<b>133986,43</b>
<b>Costes de Producción (€/kg)</b>	<b>455,73</b>

En este apartado se han estimado los costes de producción de la enzima *Ld*NDT2, tomando como base para los cálculos la producción una fermentación en un biorreactor de **1000 L** (volumen de fermentación), operado en modo semicontínuo. En base a nuestros cálculos, con 35 campañas fermentativas es posible obtener **294 kg** de un concentrado con actividad 2'-desoxi-2'-fluororibosil-transferasa y la producción de un kilogramo del catalizador costaría **455,73** euros, es decir, aproximadamente **0,45 €/g** de biocatalizador.

### ¿Qué nos dice este resultado?

El análisis realizado puede, en primer lugar, proporcionar directrices sobre qué objetivos se ha de trabajar pensando en los rendimientos de producción (kg clofarabina / kg de concentrado enzimático). En segundo lugar, puede identificar los «*cuello de botella*» del proceso (la etapa fermentativa o la etapa de recuperación), identificando

<sup>□</sup> Multiplicando por 35 el coste total para un lote (Tabla 28)

los aspectos del bioproceso (rendimientos de la fermentación y rendimientos de la recuperación), variables operativas y del diseño del bioproceso en general. Finalmente, la combinación de estos análisis podría conducir a nuevas estrategias para el desarrollo del proceso. Por ejemplo, nuevas estrategias de ingeniería de la *LdNDT2*, mejoras en la expresión o en las etapas de recuperación, etc..

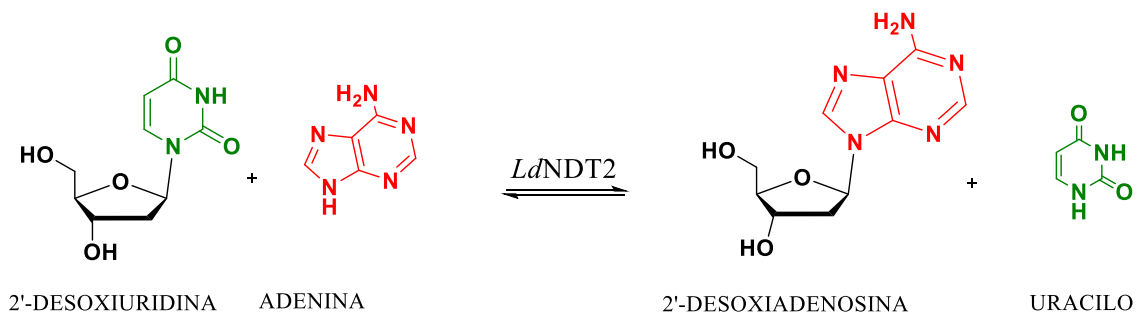
Si echamos un vistazo al precio de las materias primas del proceso (Tabla 26 y más adelante la Tabla 37), el precio del biocatalizador, es de los más altos. Pero, como veremos más adelante, el precio del catalizador no significa mucho en sí mismo, sino cómo contribuye al valor del producto a obtener, en términos, por ejemplo, de condiciones de reacción más suaves, rendimientos más altos, mayor pureza del producto final, disminución de las etapas de purificación, reducción de emisiones o condiciones “únicas” de manufactura<sup>120</sup>. Ya en la Tabla 4 mostraban las sugerencias, sobre la base del análisis económico de Tuvfesson y col.<sup>45</sup>, de cómo contribuyen los biocatalizadores a la viabilidad económica de un proceso biocatalítico. Enfatizando en que diferentes tipos de productos químicos suelen tener diferentes requisitos, en lo referido a la contribución del biocatalizador. Por ejemplo, la producción “a granel”, de un producto químico básico, precisa que esta contribución esté en el rango de 1 € / kg. Para casos como el que nos ocupa, un API, cuyos precios de venta son considerablemente más altos (más de 650 € / kg), por regla general, se plantea que esta contribución podría ser de alrededor del 10%, en este caso, puede estar entre de **60-70** €/ por kg de clofarabina<sup>46</sup>. Cuando analicemos la producción de clofarabina (apartado 4.3.2), tendremos más datos sobre la contribución del concentrado de *LdNDT2* al coste de producción de la clofarabina.

#### 4.2.8. Síntesis enzimática de clofarabina.

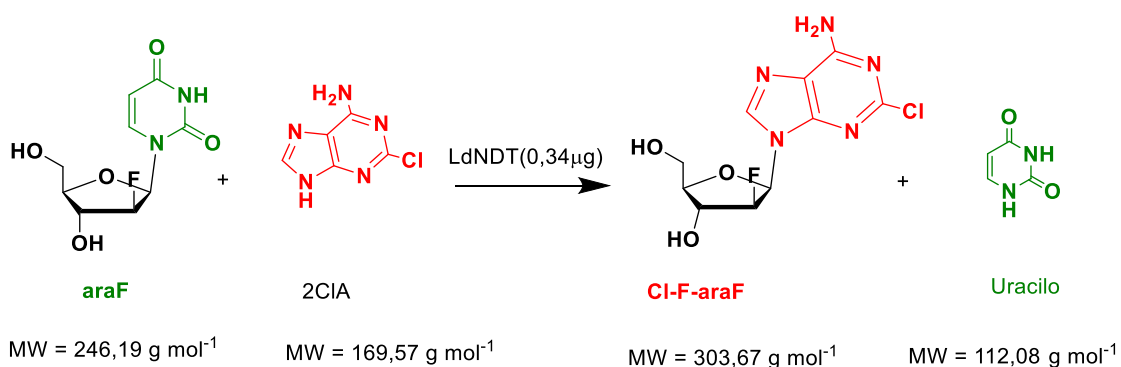
La producción de clofarabina por transglicosilación usando nucleósido fosforilasas (NPs; EC: 2.4.2.n) ya ha sido explorada por varios autores<sup>53, 80, 85-87</sup>, sin embargo, no existe una metodología (publicada) en la que se describa una reacción biocatalítica, comercialmente viable, para la producción de clofarabina. Entre otras razones, si se usan NPs, no es factible usar como materia prima un nucleósido como donante del pentofuranosilo activado (2-desoxi-2-fluoro- $\alpha$ -D-arabinofuranosa-1-fosfato), necesario para la reacción de transglicosilación. En este sentido, la estrategia usando NPs, pasa por suministrar el pentofuranosilo activado, ya sea usando un paso previo de síntesis

química<sup>88, 89</sup> o usando una cascada de reacciones catalizadas por varias enzimas relacionadas con la síntesis de nucleósidos.

En este trabajo se explora la posibilidad de llevar a cabo la síntesis de clofarabina usando una NDT de tipo II, la *L. delbrueckii*, toda vez que los estudios de acoplamiento molecular, indicaron que era factible la síntesis. El reconocimiento de arabinonucleósidos como sustratos no es común entre este grupo de enzimas, por lo cual se realizaron pruebas de concepto, utilizando, en primer lugar condiciones de reacción similares a las empleadas por Fernández-Lucas y col.<sup>61</sup>, para la síntesis de 2'-fluor-2'-desoxiadenosina (**2'-FdAdo**), es decir, la concentración de los sustratos se fijó en 1 mM, la enzima a una concentración de 0,0085 mg/ml y la temperatura de reacción, 40°C (Figura 42). Para la síntesis de clofarabina, se sustituye el sustrato natural del enzima, desoxiuridina, por 1-(2-desoxi-2-fluoro-β-D-arabinofuranosil)uracilo (**AraF**) y se utiliza 2-cloroadenina, como base aceptora (Figura 43). Se usó el extracto concentrado (seco) con actividad *Ld*NDT2 obtenido en este trabajo, cuya actividad específica es aproximadamente 150 U/mg proteína<sup>♦</sup>.



**Figura 42:** Esquema de la reacción del ensayo de actividad para *Ld*NDT2. Condiciones de reacción: Tampón 50 mM MES (ácido morfolíneetansulfónico), pH 6,5; 40°C; 2'-desoxiuridina (10 mM); adenina (10 mM); *Ld*NDT (0,0085 mg/mL).



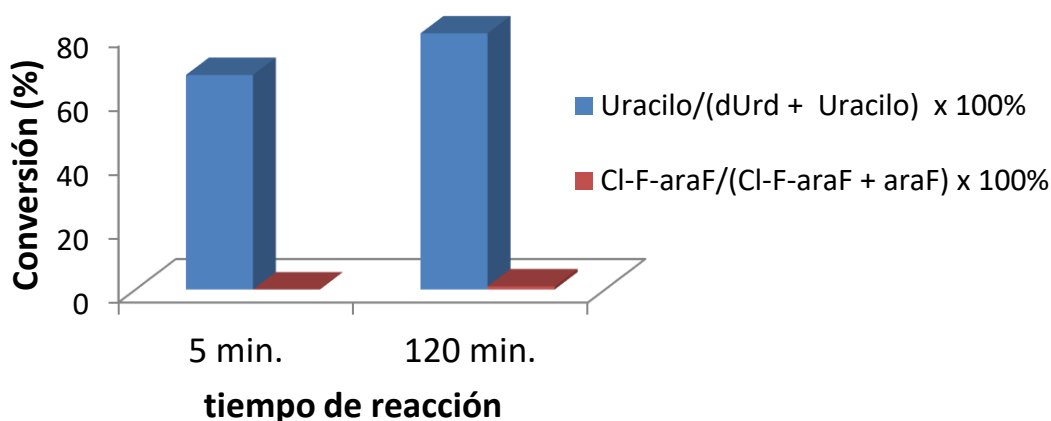
**Figura 43:** Prueba de concepto para la síntesis enzimática de clofarabina. **AraF**: 2'-desoxi-2'-fluorouridina; **2CIA**: 2-Cloroadenina; **Cl-F-araF**: Clofarabina. Condiciones de reacción: Tampón MES 50 mM; pH 6,5; 40°C; *Ld*NDT (0,0085 mg/mL).

<sup>♦</sup> Esta actividad se basa en la determinación descrita en el apartado 3.2.9 usando la reacción de la Figura 42.

Dada la baja solubilidad de la 2-cloroadenina, la reacción biocatalítica propuesta, es un sistema en suspensión. En estas condiciones se investigó la conversión de los sustratos de partida en clofarabina y en paralelo, se ha ensayó una reacción, con sustratos naturales de esta enzima. En esta primera aproximación se tomaron muestras a 5 y 120 minutos de reacción y la formación de los productos fue confirmada por comparación de los tiempos de retención y el espectro UV de los picos, con patrones auténticos. A partir de las áreas de los picos obtenidos en por HPLC, se estimaron las conversiones según la Ecuación 4. Como en la reacción de síntesis de clofarabina la base es insoluble, en los cálculos se emplearon los datos de desaparición del sustrato, **AraF**, en lugar de la base.

**Ecuación 4**      **Conversión (%) =  $\frac{ClFaraF}{ClFaraF+araF} \times 100$**

**Ecuación 5**      **Conversión (%) =  $\frac{Uracilo}{dUrd +Uracilo} \times 100$**



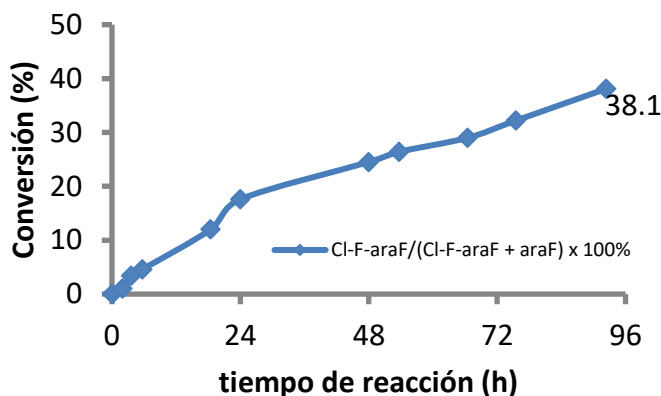
**Figura 44:** Síntesis de clofarabina. Prueba de concepto.

Como se esperaba, dado los resultados del estudio de acoplamiento (*docking*) molecular, el reconocimiento de **araF** como sustrato de la *Ld*NDT2 fue, en comparación con el sustrato natural, desoxiuridina, mucho menos eficiente. De hecho a los 5 minutos de reacción, no se detectó la formación de clofarabina, en tanto que en ese mismo tiempo se obtuvo una conversión de aproximadamente un 70%, del sustrato natural. No obstante, a las 2 horas de reacción se detectó la formación de clofarabina, aunque a un nivel de conversión significativamente bajo, cerca de un 1%.

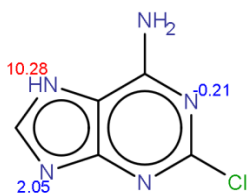
La síntesis de clofarabina, no había sido abordada antes, usando NDTs, por lo que a falta de estos datos para valorar la dimensión de estos niveles de conversión se ha hecho una comparación con los rendimientos obtenidos por Fernández-Lucas y col<sup>61</sup>.

Como se ha comentado anteriormente, estos autores utilizan la **NdRT** de *L. reuteri* para la síntesis de 2'-fluoro-2'-desoxiadenosina (**dAdo<sub>2F</sub>**), usando 2'-fluoro-2'-desoxiuridina (**dUrd<sub>2F</sub>**), como donante y adenina como base aceptora. Según los datos aportados por la publicación, la reacción se lleva a cabo con una concentración de 1mM de cada sustrato y añadiendo 0,34μg de enzima en 40 μL de reacción. Esto significa que se han empleado 0,0085 mg de enzima por mL de reacción y tras de 24 horas de reacción plantean una actividad específica de síntesis de 2'-fluoro-2'-desoxiadenosina de 0,3 mU/mg. Si transformamos estos datos en porcentajes de conversión, respecto a la base (adenina), serían aproximadamente  $3,7 \times 10^{-4} \%$ , es decir, la conversión obtenida usando la *Ld*NDT2 (en 2 horas), aunque baja, es significativamente superior a la obtenida usando *Lr*NDT, con un sustrato similar, en 24 horas.

A raíz de estos resultados se propuso el estudio del progreso de la reacción a intervalos de tiempo más largos. En la Figura 45 se muestran los resultados obtenidos tras extender a 4 días la reacción biocatalítica. La conversión mostró un incremento sostenido (prácticamente lineal) hasta alcanzar un valor de 38,1%.



**Figura 45:** Progreso de la reacción de síntesis de clofarabina mediada por *Ld*NDT2. Se utilizan las condiciones de reacción descritas en la Figura 43.

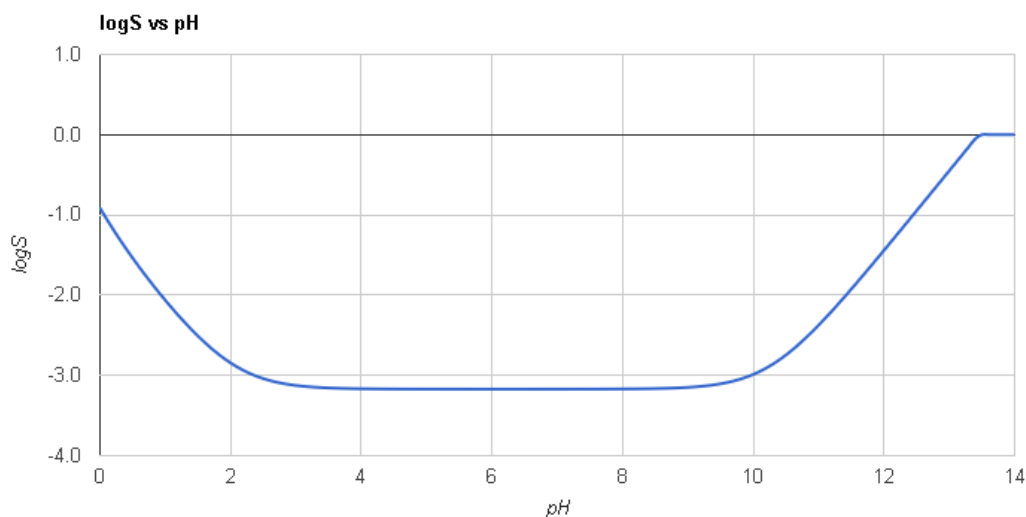


En aras de mejorar estos parámetros se propusieron una serie de experimentos conducentes a definir de los límites de actuación de la *Ld*NDT2. En primer lugar, tratar el tema de la solubilidad de los sustratos, en especial de la 2-cloroadenina. En la reacción ésta actúa como base aceptora y dada su extremadamente baja solubilidad en agua (aprox. 70 mg/L, a 50°C), constituye uno de los “*cuello de botella*” en el escalado de la reacción. La Figura 46 muestra una representación de la solubilidad de **2CIA**, a diferentes valores de pH. Este nucleósido es muy insoluble a valores neutros de pH, con valores de pKa muy altos (por ejemplo, fuertemente ácido: pKa=10,28), de modo que podría esperarse que su solubilidad pudiera aumentar significativamente a través de la



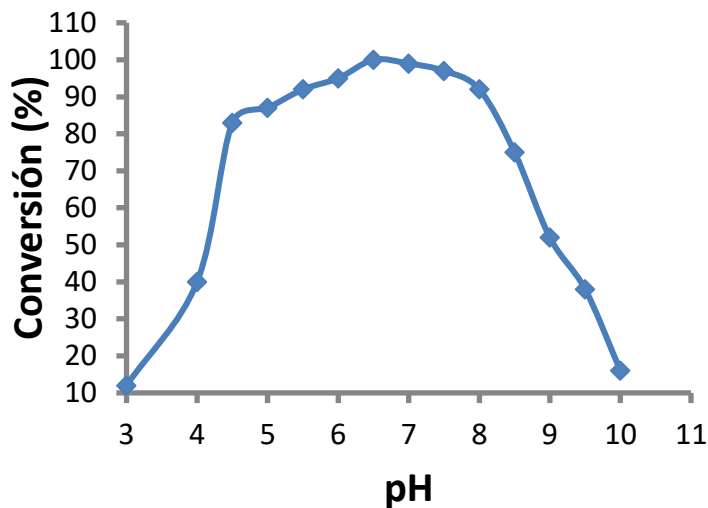
desprotonación, a medida que el pH del medio de reacción se incrementase de condiciones ácidas a básicas (Figura 46). pHs por encima de 10, o por debajo de 3, podrían cambiar el modo de unión de un sustrato cargado al centro activo de la enzima, o afectar la estabilidad de la enzima. En las reacciones catalizadas por enzimas, los cambios de pH podrían afectar el estado de ionización de los grupos funcionales en el centro activo y por tanto la velocidad de las reacciones.

De cualquier manera, enzimas con mecanismos catalíticos similares, han sido usadas a pH10, para la síntesis de 2'-desoxiguanosina, por el grupo del Dr. Terreni<sup>121</sup>, aunque las enzimas utilizadas, la **UP** de *E. coli* y la **PNP** de *Bacillus subtilis*, estaban estabilizadas mediante inmovilización, para llevar a cabo la síntesis en estas condiciones. En todo caso, se ha de establecer un compromiso entre la solubilidad de los sustratos y la actividad/ estabilidad del enzima.



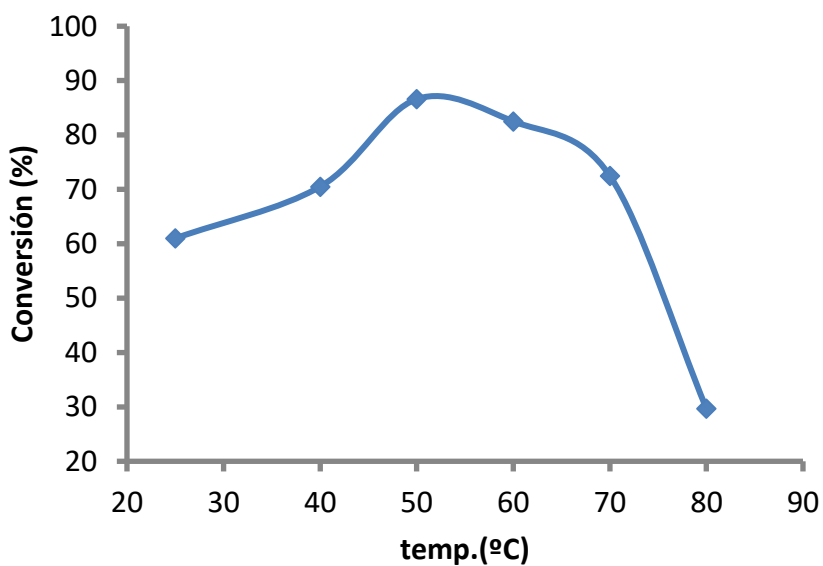
**Figura 46:** Solubilidad de la 2CIA a diferentes pHs.

Para definir los «*limites de actuación*» de la *LdNDT2* se investigaron en primer lugar parámetros que podrían afectar a la solubilidad de los sustratos, entre ellos, el pH, la temperatura, y el uso de cosolventes, con sustratos naturales del enzima. La Figura 47 muestra los porcentajes de conversión obtenidos utilizando medios de reacción con pHs que van desde 3 hasta 10. *LdNDT2*, tiene un pH óptimo de actuación entre 6,5 y 7,5 y retiene más del 60% de su actividad máxima a valores de pH de entre 4,5 y 8,5. Fuera de estos límites la actividad cae en picado, lo cual supone una limitación a la hora de intentar la solubilización de los sustratos a pHs extremos.



**Figura 47:** Efecto del pH sobre la actividad de la LdNDT2. Los porcentajes de conversión están referidos al sustrato natural de la enzima (reacción de la Figura 42).

Como se muestra en la Figura 48, la *LdNDT2* tiene una temperatura óptima de actuación alrededor de los 50°C, y mantiene su actividad por encima del 60%, desde los 25°C hasta los 70°C. La temperatura, podría favorecer la solubilidad de los sustratos, pero temperaturas por encima de los 60°C, podrían comprometer la estabilidad de la 2'-fluoro-2'-desoxiuridina<sup>122</sup>.



**Figura 48:** Efecto de la temperatura sobre la actividad sintética de la LdNDT2.

A la luz de los resultados obtenidos, usar altas temperaturas o valores de pH “no fisiológicos” podrían no ser aplicables en el mejoramiento de la productividad, usando *Ld*NDT2. Alternativamente, en este trabajo se ha considerado el uso de cosolventes, con especial atención a aquellos considerados como “disolventes verdes”, como por ejemplo, derivados de la dimetilamida (N,N-dimetildecanamida, Agnique AMD-10™) o del glicerol (ej. glicerol formal). En la Tabla 31 se muestran las condiciones de reacción ensayadas y para cada entrada se describe el tipo y concentración del disolvente empleado.

Como se puede apreciar en la Tabla 31 los resultados mostraron que las actividades, en medio convencional (tampón MES 50 mM, pH 6,5), son por lo general superiores al de los medios con cosolventes, salvo aquellas en la que se ha usado AMD-10™ al 5% (entradas 15 y 16), aunque usar concentraciones superiores supone, al parecer, la pérdida de actividad de la enzima. En estos ensayos se emplearon dos concentraciones de sustratos iniciales, la intención inicial era averiguar, si el hecho de adicionar cosolventes, suponía un aumento de la disponibilidad de los sustratos y si todo ello redundaría en un aumento de la actividad enzimática. Pero los resultados obtenidos (Tabla 31) indicaron que, aumentar de 1 a 10 mM, la concentración inicial de sustratos, conduce, “presumiblemente”, a un aumento de la actividad enzimática, independientemente del uso de cosolventes.

Estos experimentos preliminares, usando la *Ld*NDT2, con sus sustratos naturales, permitieron fijar las condiciones iniciales para la reacción de síntesis de clofarabina. Es decir, usar concentraciones superiores de sustratos (>1mM), a pH de entre 6,5-8, a 50°C y de momento, no usar cosolventes. Lo cual condujo a un escenario en el que, dada la escasa solubilidad de los sustratos de la reacción, se establece un sistema en suspensión (heterogéneo), con las dificultades técnicas que esto conlleva.

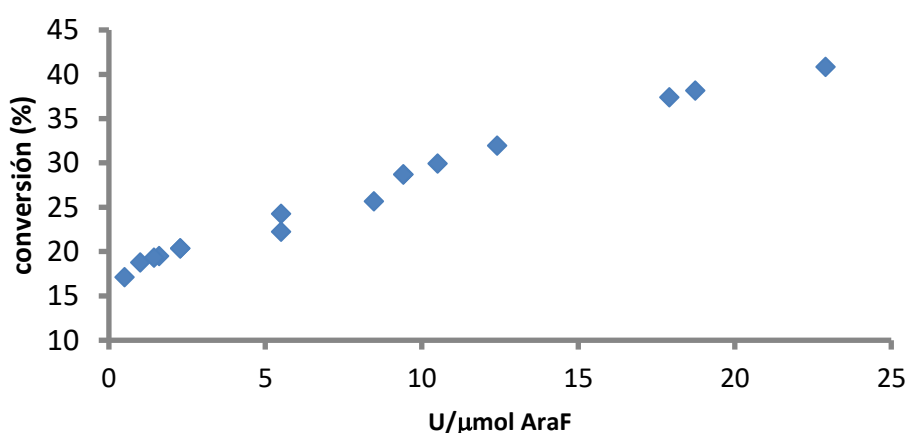
**Tabla 31:** Efecto del uso de cosolventes sobre la actividad de LdNDT2. Condiciones de reacción: Tampón MES 50 mM (ácido morfolineetansulfónico), pH 6,5; 50°C; 2'-desoxiuridina (**2dUrd**); adenina (**Ade**); LdNDT2 (0,0085 mg/mL). Se ha tomado como 100%, la actividad de la entrada 1.

<b>Entrada</b>	<b>dUrd (mM)</b>	<b>Ade (mM)</b>	<b>Solvente<sup>√</sup> (%)</b>	<b>Actividad (%)</b>
<b>1</b>	<b>1</b>	<b>1</b>	—	<b>100</b>
2	10	10	—	134
3	1	1	THF 5%	59,23
4	10	10	THF 5%	75,6
5	1	1	THF 10%	52,27
6	10	10	THF 10%	67,37
7	1	1	THF 20%	48,98
8	10	10	THF 20%	50,81
9	1	1	GF 5%	59,75
10	10	10	GF 5%	75,99
11	1	1	GF 10%	50,33
12	10	10	GF10%	59,33
13	1	1	GF 20%	47,25
14	10	10	GF 20%	46,61
<b>15</b>	<b>1</b>	<b>1</b>	<b>AMD 5%</b>	<b>121,12</b>
<b>16</b>	<b>10</b>	<b>10</b>	<b>AMD 5%</b>	<b>125,87</b>
17	1	1	AMD 10%	65,56
18	10	10	AMD 10%	70,48
19	1	1	AMD 20%	49,66
20	10	10	AMD 20%	53,15

<sup>√</sup> THF: tetrahidrofurano; GF: glicerol formal; AMD: N,N-dimetildecánamida

A partir de estos resultados se realizaron experimentos, esta vez con los sustratos conducentes a la formación de clofarabina (Figura 43) y en las condiciones que maximizan y/o estabilizan la actividad de la *Ld*NDT2. En primer lugar, se abordaron aspectos relacionados con los rendimientos de conversión de clofarabina. La síntesis de este API precisa de largos tiempos de reacción (Figura 45), probablemente, por efecto inhibitorio de los sustratos y/o los productos de la reacción. Ésta ha sido una de las limitaciones comunes de la mayoría de las aproximaciones enzimáticas, en las cuales se han usado nucleósidos fluoro sustituidos en posición 2', es decir 2'-desoxi-2'-fluoro –ribonucleosidos o –arabinonucleósidos<sup>53, 81, 88, 108</sup>.

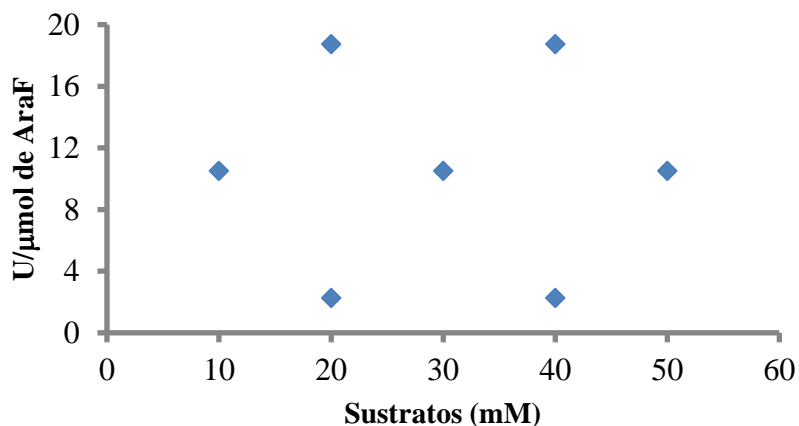
Una de las soluciones propuestas ha consistió en aumentar la cantidad de unidades de actividad desoxirribosiltransferasa en las reacciones, es decir, el efecto de la carga enzimática. En este trabajo se han usado cargas enzimáticas en el rango de 1-20 U<sup>▼</sup> de *Ld*NDT2/ $\mu$ mol de **AraF**, se ha encontrado que existe una correlación lineal ( $R^2 > 0,98$ ) entre el aumento de los niveles de conversión y la carga enzimática, en otras palabras, los niveles de conversión mejoran con la cantidad de enzima adicionada (Figura 49). Todo ello, apunta a que un aumento de la carga limita el efecto inhibitorio de los sustratos utilizados en la reacción biocatalítica.



**Figura 49:** Efecto de la carga enzimática sobre los rendimientos de conversión de clofarabina. Condiciones de reacción: Sustratos (1mM); Tampón MES (ácido morfolineetansulfónico), 50 mM; pH 6,5; 50°C.

▼ Una unidad de actividad desoxirribosiltransferasa se define como la cantidad de enzima que convierte 1  $\mu$ mol del sustrato natural de la enzima (dUrd) por min, a 40°C.

Desde el punto de vista práctico, es deseable usar los sustratos de la biorreacción a concentraciones tan altas como sea posible (siempre y cuando los fenómenos de inhibición no sean obvios) en aras de conseguir altos rendimientos volumétricos y emplear la menor cantidad “relativa” de biocatalizador posible, para reducir la contribución de éste al coste final del API. Como se ha anticipado, con los sustratos naturales (Tabla 31), la concentración inicial de sustrato influía sobre la actividad del enzima, y en este caso, elevar la concentración de éstos de 1 a 10 mM, suponía, como promedio un incremento de un 25% en la actividad del enzima.



**Figura 50:** Esquema del diseño Doehlert para el estudio de dos factores: Carga de la enzima (U/μmol de AraF) concentración inicial de los sustratos (Los dos sustratos a la misma concentración).

Se ha aplicado un diseño experimental de tipo Doehlert<sup>√</sup> considerando como variables **la concentración inicial de los sustratos y la carga de la enzima en la reacción** (unidades de enzima por μmol de AraF), para valorar el efecto de la carga inicial de los sustratos y su correlación con la carga enzimática. La Figura 50 muestra la distribución los siete puntos utilizados en el diseño Doehlert para los dos factores estudiados, el primero de los factores se estudió a cinco niveles y el segundo a tres. La Tabla 32 muestra los resultados obtenidos al utilizar los parámetros del diseño.

<sup>√</sup> Entre los diferentes diseños de superficie de respuesta posibles se eligió el diseño de Doehlert dado que ofrece una distribución uniforme de puntos en el espacio de respuesta experimental. Dicha uniformidad de la distribución describe una figura romboidal (en el caso de dos variables se produce un hexágono, Figura 50), en la que se asignan cinco niveles a la variable de mayor peso o importancia y tres a la otra variable.

**Tabla 32:** Resultados experimentales tras aplicar los parámetros del diseño.

<b>Entrada</b>	<b>AraF (mM)</b>	<b>2CIA (mM)</b>	<b>LdNDT (U/<math>\mu</math>mol araF)</b>	<b>Conversión (%)</b>
1	30	30	10,5	36,5
2	50	50	10,5	19,51
3	40	40	18,73	32,82
4	10	10	10,5	48,27
5	20	20	2,27	39,97
6	20	20	18,73	47,7
7	40	40	2,27	25,51
<b>Óptimo simplex</b>	<b>17,71</b>	<b>17,71</b>	<b>19,91</b>	<b>49,71</b>

Debido a la baja solubilidad de los componentes de reacción, el medio de reacción biocatalítico es un sistema en suspensión, por lo que es muy probable que determinadas combinaciones sufran condiciones limitadas de transferencia de masa sólido-líquido. En el análisis, la variable de proceso más importante fue la mejora de la del porcentaje de conversión y este fue utilizado para escoger la combinación de estos dos parámetros que permita maximizarlo. Para ello, se empleó el algoritmo *Simplex*<sup>123</sup>, con el cual el óptimo se alcanza a través de un procedimiento de búsqueda secuencial. Los resultados de este análisis estadístico apuntan a que el óptimo (máxima conversión), teniendo en cuenta que se han analizado concentraciones de sustrato de entre **10-50 mM\*** y cargas de enzimas de entre **1-20 U/ $\mu$ mol de AraF**, que la combinación óptima sería la de **23,42 mM** de los sustratos y una carga enzimática de **19,95 U/ $\mu$ mol de AraF**. Con esta combinación, se alcanzaría teóricamente un **49,71 %** de conversión (Tabla 32).

En este punto se ha de aclarar que una solución "óptima" (ofrecida por el algoritmo) es óptima nada más en lo que se refiere al modelo específico que se está usando para representar el problema real (ecuación) , y tal solución no se convierte en una guía confiable hasta que se verifica su comportamiento. En este sentido, si observamos los resultados de las conversiones, usando concentraciones de sustratos a 10 mM y la carga enzimática a 10,5 U/ $\mu$ mol de AraF (entrada 4, Tabla 32), se aprecia que la conversión planteada por el modelo (Ecuación 6) es "relativamente" similar, en términos de

\* Para la optimización se han escogido concentraciones iniciales de sustrato, claramente saturantes, pensando en los procesos de aislamiento y purificación de los productos de reacción.

conversión, a la “óptima” ofrecida por el diseño. Pero, observando con un poco más de detalle vemos que el rendimiento de la entrada 4 (Tabla 32) se obtiene usando las concentraciones más bajas de los sustratos, lo cual podría significar que los sustratos ejercen un efecto inhibitorio sobre la actividad del enzima.

**Ecuación 6:**

$$Conv (\%) = A1 + A2[araF] + A3[2CIA] + A4[araF]^2 + A5[2CIA]^2 + A6[araF][2CIA]$$

donde:

$$A1 = 4,82 \times 10^1; A2 = -3,18 \times 10^{-1}; A3 = 2,92 \times 10^{-1}; A4 = -6,52 \times 10^{-3}; A5 = 9,63 \times 10^{-3}; A6 = -1,27 \times 10^{-3}$$

La combinación sugerida por el modelo fue probada experimentalmente, junto a otras que permitieron valorar los resultados del experimento de optimización. La Tabla 33 muestra los experimentos y sus resultados indican que, efectivamente, el modelo es creíble, pues experimentalmente los resultados son muy similares al ofrecido por el modelo. Por otro lado, se ha constatado que las concentraciones iniciales más bajas favorecen los rendimientos de conversión, como demuestra el usar la misma carga enzimática (cerca de 20 U/μmol **araF**), para la menor concentración de sustrato (10 mM). En todo caso, con la vista puesta en el aislamiento de los productos de reacción y sobre todo en la economía del bioproceso se puede sopesar, en función de los rendimientos (concentración final de clofarabina en el crudo de reacción), el usar concentraciones de sustratos y carga de biocatalizadores superiores al óptimo encontrado.

**Tabla 33:** Verificación de los resultados ofrecidos por el modelo (Ecuación 6) obtenido utilizando el algoritmo Simplex<sup>√</sup>.

<b>Entrada</b>	<b>AraF (mM)</b>	<b>2CIA (mM)</b>	<b>LdNDT2 (U/μmol araF)</b>	<b>Conversión. (%)</b>
1	10	10	10,5	48,27
2	17,71	17,71	19,5	48,53
3	17,71	17,71	10,1	43,75
4	10	10	19,4	52,56

<sup>√</sup> El método secuencial Simplex es una alternativa al cálculo de un modelo matemático para alcanzar las condiciones óptimas. En este método, cada futuro experimento se decide en función de los resultados obtenidos en los experimentos anteriores



### 4.3. Escalado de la reacción biocatalítica

Los resultados obtenidos a pequeña escala mostraron que es posible alcanzar rendimientos “moderadamente” satisfactorios, cuando se utilizan concentraciones saturantes de los sustratos. El modelo obtenido (Ecuación 6) indica que, a pesar de que las concentraciones más altas de sustratos pueden inhibir los rendimientos de conversión, puede compensarse esta pérdida en los rendimientos con una mayor carga enzimática.

Por otro lado, el sistema, a pesar de ser un sistema en suspensión reproduce los resultados y las expectativas teóricas, de modo que es posible predecir los rendimientos con combinaciones que permitan la viabilidad de los procesos de recuperación del producto de la reacción.

En este sentido, el escalado de la producción a nivel de reactor de 0,18 L, se planteó usando las condiciones similares a la de los experimentos a pequeña escala (Tabla 34). En esta escala se han obtenido rendimientos de conversión de un **17,2%**, que es menor de lo esperado, según modelo ( $\approx$  **24%**, según la Ecuación 6), probablemente por el efecto del cambio de escala (agitación, transferencia de masa, etc.). La reacción de transglicosilación en esta escala tuvo una productividad de **0,026 g.L<sup>-1</sup>.h<sup>-1</sup>**, muy baja si nos fijamos en el criterio de Straathof y col.<sup>112</sup>, que plantean una productividad promedio (para un API) de cerca de 15,5 g.L<sup>-1</sup>.h<sup>-1</sup>, para un proceso viable, desde el punto de vista económico.

**Tabla 34:** Condiciones de reacción. Escala de 0,18 L

<b>Parámetro</b>	
pH	6,5
Temperatura (°C)	50
<b>AraF</b> (mM)	50
<b>2CIA</b> (mM)	50
<i>Ld</i> NDT2 (U/ $\mu$ mol AraF)	20
Agitación (r.p.m.)	150
Tiempo de reacción (h)	96

Pero, considerando que no es una reacción optimizada, estos rendimientos son claramente mejorables, entre otras posibilidades, podría pensarse, por ejemplo, en mejorar la solubilidad de la **2CIA**, en la inmovilización del enzima, la evolución del enzima, etc. Por otro, la concentración final del producto fue de **2,71 g/L**, que es baja respecto a la concentración promedio de la producción comercial de nucleótidos (65 g/L)<sup>124</sup>, resultó adecuada para acometer con garantías el proceso de aislamiento y purificación de la clofarabina a partir del crudo de reacción ( apartado 4.3.1).

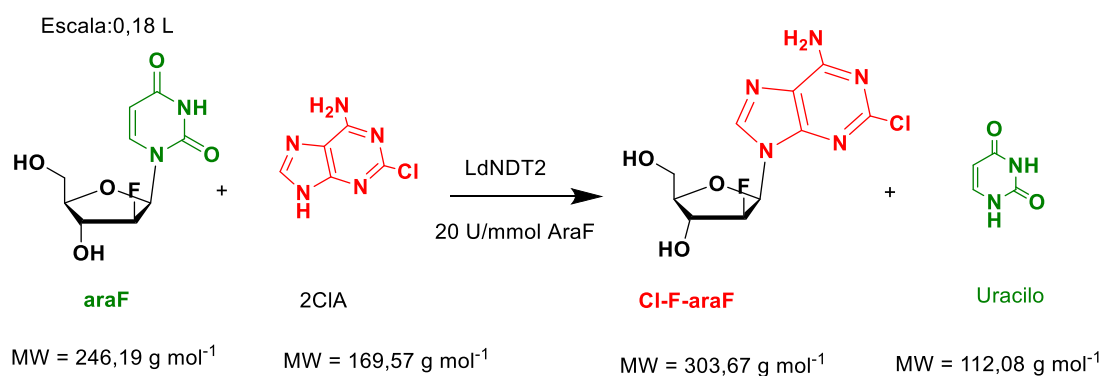
#### 4.3.1. *Purificación de la clofarabina*

En el escalado de la producción se utilizó una concentración inicial de sustratos de 50 mM y al final del proceso una conversión del 17,2 %. El crudo obtenido, al no ser una conversión completa, contiene además de los sustratos iniciales, los dos productos de la reacción (es decir, 2-cloroadenina, uracilo, AraF y Clofarabina), disueltos en tampón MES. En la Figura 51 se muestran las cantidades iniciales y finales de la reacción biocatalítica (Tabla 35) y en la Figura 52 se muestra un cromatograma con el perfil típico de un crudo de reacción, al final de la reacción.

No hay muchas publicaciones en la que se describa un proceso factible de aislamiento y purificación de clofarabina. Desde su descubrimiento, los métodos de purificación están muy relacionados con la separación del anómero correcto y de impureza propias del proceso de síntesis. Los métodos más efectivos incluyen el uso de columnas cromatográficas, HPLC preparativos, entre otras<sup>37, 39, 40, 125, 126</sup>, que en su mayoría son procesos multipasos y/o con limitados niveles de recuperación, lo que dificulta su aplicación a escala comercial. También se puede constatar que muchos de estos procesos de aislamiento y/o purificación incluyen una o más etapas de recristalización para alcanzar la pureza final del producto.

En el proceso propuesto en este trabajo, la estereoespecificidad de la reacción enzimática (transglicosilación), supone que el único producto formado sea el anómero  $\beta$ , y esto facilita el proceso de purificación. Sin embargo, en el crudo de reacción persisten parte de los reactivos de partida y uracilo (subproducto), aunque sus solubilidades permiten proponer la cristalización, para la recuperación de la clofarabina y del resto de los componentes del crudo de reacción (Figura 52).

La Tabla 35 muestra la composición final del crudo de la reacción de la Figura 51. La mezcla, al final de reacción, continúa siendo un sistema “*en suspensión*”, puesto que uno de los reactivos iniciales, la **2CIA**, es muy insoluble a concentraciones por encima de 0,77 mM. Precisamente esta insolubilidad se aprovechó para separarla, mediante filtración “*en caliente*”. El filtrado resultante, se concentró y se enfrió hasta la aparición de un sólido, que fue separado por filtración. En este filtrado se separa el **AraF**, en tanto que, en el sólido se concentra la clofarabina y el subproducto de la reacción, el uracilo. Este sólido se lavó y resuspendió en agua y al final se obtiene la clofarabina (recristalizada), con una pureza<sup>♦</sup> de **99,47%** (Figura 53) y un rendimiento del proceso de purificación es del **68,8 %**.

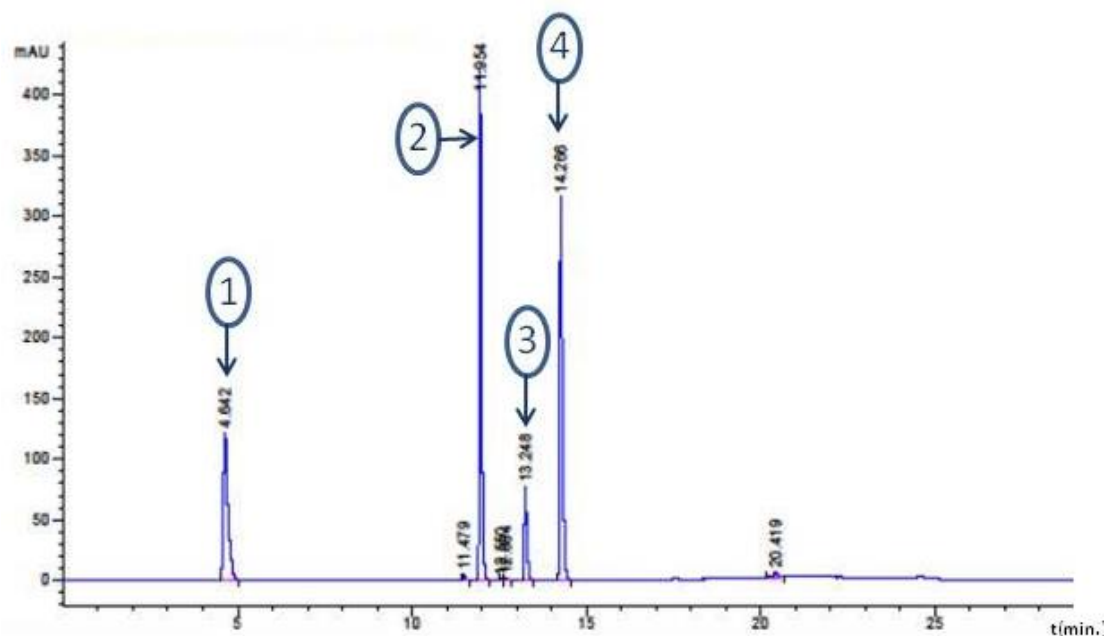


**Figura 51:** Compuestos que intervienen en la síntesis de clofarabina

Tabla 35: Composición del crudo final de la reacción biocatalítica.

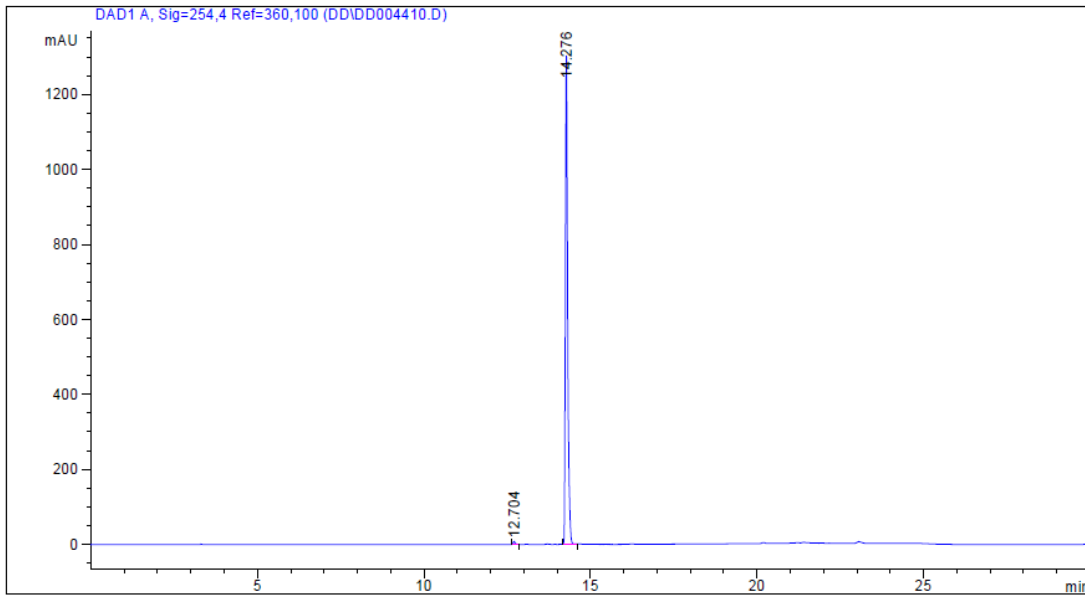
	<b>AraF</b>	<b>2CIA</b>	<b>C-F-AraF</b>	<b>Uracilo</b>
t=0h	9,34 mmol (2,3 g)	9,37 mmol (1,584 g)		
Conv. 100%			9,34 mmol (2,83 g)	9,34 mmol (1,04 g)
t= 96 h conv. 17,2%	7,73 mmol (1,904 g)	7,76 mmol (1,312 g)	3,29 mmol (0,488 g)	1,60 mmol (0,180g)

<sup>♦</sup> La pureza química de una sustancia activa, o ingrediente en general, hace referencia a la mayor o menor contaminación con especies químicamente similares pero no iguales. Aunque suele usarse el término con cierta ambigüedad, la pureza debe calcularse únicamente en base a aquellas **especies químicas relacionadas** con el principio activo y que enmascaran y disminuyen la riqueza final de éste. Es el caso, por ejemplo, de los productos secundarios originados durante los procesos de obtención o síntesis (productos isómeros, homólogos, etc.). <http://formulacionfarmaceutica.blogspot.com.es/2015/09/pureza-y-elementos-extranos-de-los.html>



**Figura 52:** Perfil HPLC de la composición del crudo final de la reacción de síntesis de clofarabina. Pico1: uracilo (4,6 min.); Pico 2: **AraF** (11,9 min.) ; Pico 3: **2CIA** (13,2 min.) y Pico 4: Clofarabina (14,3 min.).

A pesar del “potencial” avance tecnológico que representa la producción biocatalítica de este importante análogo nucleosídico, los rendimientos globales, calculados como el producto de los rendimientos de las etapas de biocatálisis y de recuperación, son muy bajos, **11,8%**, si se comparan con el proceso actual de obtención de clofarabina<sup>37, 38</sup>. Las bajas productividades, los bajos rendimientos, la solubilidad y la concentración de los sustratos, unido al posible efecto inhibitorio de los sustratos y/o los productos, son aspectos a tener en cuenta a la hora de hacer competitiva esta propuesta.



```

=====
                          Area Percent Report
=====

Sorted By      :      Signal
Multiplier     :      1.0000
Dilution      :      1.0000
Use Multiplier & Dilution Factor with ISTDs

Signal 1: DAD1 A, Sig=254,4 Ref=360,100

Peak RetTime Type Width Area Height Area
# [min] [min] [min] [mAU*s] [mAU] %
-----|-----|-----|-----|-----|-----|
  1  12.704 BB  0.0593  32.85449  8.28538  0.5276
  2  14.276 BV  0.0725 6194.83496 1301.58752 99.4724

Totals :                      6227.68945 1309.87290

=====
*** End of Report ***

```

**Figura 53:** Pureza por HPLC de clofarabina recristalizada.

### 4.3.2. Integración del proceso de síntesis biocatalítica de clofarabina

Para determinar los costes de producción biocatalítica de la clofarabina, en los cálculos se tuvo en cuenta las dos etapas principales: **Biorreacción** y **purificación**. La Figura 54 resume el diagrama de flujo del proceso de producción y recuperación del API. El mismo, ha sido elaborado usando el programa *Superpro designer* y muestra las principales operaciones unitarias (etapas de proceso) y con colores diferenciados, las secciones del proceso. Para los cálculos se tuvieron en cuenta los resultados obtenidos a nivel de reactor de **0,18 litros**, asumiendo que se obtienen **1,87 g de clofarabina/L**, con una pureza de **99,7%** en un proceso con un rendimiento global del **11,83 %**. Por otro lado, también se asumió que sólo se realizaba un ciclo biocatalítico por semana (144 horas por campaña y que el sistema sería operado por un técnico a tiempo completo. No se han incluido en los cálculos los aspectos relacionados con las normas de correcta fabricación.

En el trabajo se han estimado los costes de un proceso integrado de síntesis biocatalítica de clofarabina, usando la enzima LdNDT2, en un proceso que podemos resumir de la siguiente manera:

- 1- La reacción enzimática tiene lugar en el reactor **P1**. Los sustratos y la enzima son alimentados siguiendo las proporciones de la Tabla 34, en un proceso con una duración de 72 horas.
- 2- Al final del proceso, se disminuye (de manera controlada) la temperatura de la reacción, en el mismo reactor **P1** y al final del proceso de enfriamiento se transfiere el crudo al filtro **P2**. En esta operación se recupera la **2CIA** que no ha reaccionado durante la reacción biocatalítica.
- 3- El filtrado, que contiene, la **clofarabina**, junto al subproducto (**uracilo**) y **AraF** se transfiere al tanque de condensación, donde es parcialmente concentrado en el tanque de reacción **P3** y enfriado.
- 4- El crudo resultante es filtrado usando el filtro **P4** del que se recupera **AraF** en el filtrado y el sólido recuperado, que contiene exclusivamente **clofarabina** y **uracilo**.
- 5- El sólido se redissuelve y se pasa a un reactor de enfriamiento (**P5**) y el precipitado formado es filtrado usando el filtro **P6**. El filtrado se corresponde con el **uracilo**.
- 6- La **clofarabina** (sólido), se redissuelve y cristaliza en el cristalizador **P7**.
- 7- Al final del proceso, se recupera la clofarabina, mediante filtración al vacío (filtro **P8**).

*Nota:*

- *En el diagrama de la Figura 54 algunos iconos, en ciertas unidades de proceso, indican que estos equipos se utilizan más de una vez. Es decir, el diagrama del proceso representado es, esencialmente, una representación gráfica de la «receta» del proceso y muestra la secuencia de ejecución.*
- *La duración de todo el proceso post-biorreacción es de 72 horas.*
- *Las precipitaciones de los productos se basan en la diferencia de solubilidad de los 4 productos, a diferentes temperaturas (Tabla 36).*

**Tabla 36 :** Solubilidad de los componentes del crudo.

Temperatura (°C)	Solubilidad (mM)			
	AraF	2-Cloroadenina	Clofarabina	Uracilo
12	< 52,8	< 0,77	< 7,57	< 7,14
25	< 255,9	< 0,77	< 7,57	< 18,28
40	> 420,8	< 0,77	< 13,5	< 46,4
50	> 420,8	< 0,77	< 20,08	< 67,8

#### 4.3.3. *Estimación de los costes de operación.*

Se han utilizado las métricas del proceso a escala de reactor de 0,18 L para estimar los costes de producción de clofarabina, en 35 lotes (al año), en un reactor de 200 litros operado en modo discontinuo (*batch*). El coste de operación es la suma de todos los gastos relacionados con materias primas, mano de obra, servicios, eliminación de residuos, gastos generales, etc.

##### **A. Escenario: Producción de clofarabina (Biorreacción y purificación).**

- 1 reactor (Tanque agitado).
- Volumen total: **200 L**

En este escenario se propone la producción de aproximadamente **360 gramos** de clofarabina pura por lote. Se han tenido en cuenta las mismas aproximaciones y «*reglas de oro*» hechas para el cálculo de los gastos y los precios de las materias primas del bioproceso para la producción del biocatalizador (apartado 4.2.7).

## Materia prima.

**Tabla 37:** Coste de los reactivos utilizados en la producción y en la purificación de la clofarabina. Se estima el coste total para la producción de 35 lotes al año.

Materia prima <sup>♦</sup>	Coste unitario (€/kg)	Cantidad (kg)	Coste (€)
Agua	0,01	9,625	96,25
2CIA	290,88	59,24 <sup>a</sup>	17231,7
AraF	488,90	86,16 <sup>a</sup>	42123,6
<b>Biocatalizador (<i>Ld</i>NDT2)</b>	455,73	18,48	8421,89
<b>Total</b>			<b>67873,44</b>

<sup>a</sup>Teniendo en cuenta que se recuperan, lote a lote, los sustratos que no han reaccionado y se incorporan al siguiente lote<sup>§</sup>.

## Servicios y costes operacionales para la producción biocatalítica de clofarabina (35 lotes)

**Tabla 38:** Servicios (0,1€/kWh\* ).

Operación	Capacidad	MWh
Bombas	5 kW/m <sup>3</sup>	0,45
Agitación y calentamiento	5 kW/m <sup>3</sup>	1,45
Tratamiento de residuos <sup>^</sup>	~4 m <sup>3</sup>	

Coste de los servicios: **190,00 €/batch**

**Coste total de los servicios (35 lotes): € 6650,00.**

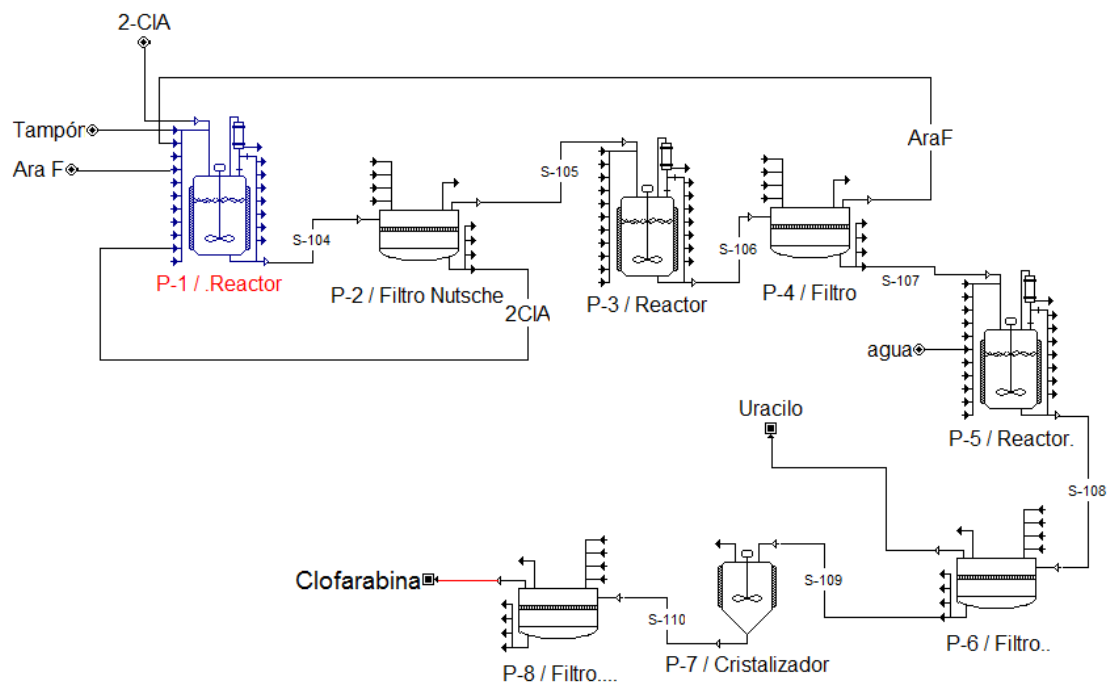
<sup>♦</sup> Dada la dificultad para obtener los precios de materias primas, como la **2CIA** y **AraF**, los precios se estimaron a partir de del precio a escala de laboratorio (<http://www.carbosynth.com>), dividiendo el precio original por 30. La incertidumbre de este enfoque es alta, pero es un buen punto de partida para las estimaciones de costes<sup>46</sup>.

<sup>§</sup> Lote a lote se recuperan como promedio: 63% de AraF y >99% de 2CIA y el material recuperado se incorpora a la siguiente reacción.

\* European Energy Portal, <http://www.energy.eu>

<sup>^</sup> Por lo general el tratamiento de las aguas residuales suele costar entre 0,5–2 €/m<sup>3</sup>, y el tratamiento de residuos sólidos (no peligrosos) alrededor de 25 €/ton.<sup>116</sup>.





**Figura 54:** Esquema del proceso de purificación de la clofarabina. Se muestran todas las unidades de proceso y los principales productos derivados del mismo. **2CIA:** 2-cloroadenina, **AraF:** 2'-fluoro- arabinofuranosil-2'-desoxiuridina.

**Tabla 39:** Coste de los equipos principales del proceso.

Equipos	unidades	Capacidad	Precio de compra (€) <sup>a</sup>	TIC <sup>*</sup>	Anualidad (€/año) <sup>♦</sup>	Mantenimiento (€/año) <sup>♥</sup>	Otros (€/año) <sup>▲</sup>	Coste /batch
Reactor (termostatizado)	1	0,20m <sup>3</sup>	27600	138000	19648,4	1964,8	2947,3	701,7
Filtro <i>Nutsche</i>	1	0,35 m <sup>2</sup>	19000	95000	13526,1	1352,6	2028,9	483,1
Reactor (termostatizado)	1	0,05m <sup>3</sup>	12500	62500	8898,8	889,9	1334,8	317,8
Cristalizador	1	0,10 m <sup>3</sup>	35600	178000	25343,6	2534,4	3801,5	905,1
<b>Total</b>								<b>2407,7</b>

<sup>a</sup><http://matche.com/equipcost/>

<sup>\*</sup> TIC – *Total Installed cost* (costes totales de instalación) - Factor Lang = 5

<sup>♦</sup> Tasa de interés: 7% y 10 años de vida útil.

<sup>♥</sup> El mantenimiento es el 10% del TIC

<sup>▲</sup> Otros, incluidos los impuestos, los seguros, etc., representan el 15% del TIC.

## Mano de obra

A partir de los costes de capital instalado (Tabla 39), se estimaron los costes de la mano de obra, como una fracción de los pagos anuales totales del capital instalado (15%)<sup>117</sup>, y los costes relacionados con Garantía de calidad (QA) / control de calidad (QC) , como una fracción (15%) del coste total de la mano de obra. Así, el coste de la mano de obra se estimó en **10112,53 €/35 lotes**, y el Coste total de QC/QA/35 lotes en **1516,88 €**.

La Tabla 40 resume los principales costes operativos que incluyen el coste estimado de las materias primas, los servicios, los asociados a la mano de obra y los costes dependientes de las instalaciones (depreciación, mantenimiento, etc.).

**Tabla 40:** Costes operacionales anuales de la producción de la enzima *Ld*NDT2 (35 lotes).

Indicador de coste	€/35 batches
Materia prima	67873,44
Mano de obra	10112,53
Análisis/QC/QA	1516,88
Servicios	6650,00
Costes relacionados con la capacidad instalada	84271,16
<b>Total</b>	<b>170424,01</b>

**Tabla 41:** Resumen de los costes de producción del API.

Indicador de coste	€/35 batches
Producción anual (kg)	<b>12,6</b>
Costes de operación (€) (35 campañas)	<b>170424,01</b>
Costes de Producción anual (€/kg)	<b>13525,71</b>

En este trabajo se han estimado los costes de producción del API, la clofarabina, tomando como base para los cálculos la producción en un reactor de **200 L**. En base a nuestros cálculos, con tres campañas sería posible obtener **12,6 kg** de clofarabina pura y su producción costaría **13525,71 euros/kg**, es decir, aproximadamente **13,52 €** por gramo de clofarabina pura.

Los costes de producción son muy altos, en comparación con otros APIs, teniendo en cuenta que el precio de venta de este API, según *Pharmacompass*<sup>\*</sup>, es de 785,016 dólares/kg (637,18 €/kg)<sup>♦</sup>. Probablemente, estos costes pueden ser ajustados si se introducen mejoras en el esquema global del bioproceso. En el siguiente apartado (4.3.4) se hacen algunas consideraciones en este sentido.

Por otro lado, en el apartado 4.2.7, se dejó una pregunta abierta, en relación a la contribución del coste del biocatalizador al coste del API. Entonces se mencionaba el hecho de que, para casos como el que nos ocupa, un API, cuyos precios de venta son considerablemente altos (637,18 €/kg), por regla general, esta contribución debería ser de alrededor del 10%, es decir, debería rondar los **60-70 €/** por kg de clofarabina<sup>46</sup>. En este caso, el escenario propuesto planteó la utilización de 18,48 kg del extracto enzimático (unos 8421,89 €), para producir 12,6 kg de clofarabina, lo cual supone cerca del 5% del coste total de producción, razonable, desde el punto de vista del porcentaje, pero inviable desde el punto de vista comercial, 668,40 €/ kg de clofarabina. Es en este punto donde cobra una importancia capital la posibilidad de reutilizar el enzima, por medio de la inmovilización del biocatalizador. Así, por ejemplo, con sólo reutilizar 12 veces el enzima, y asumiendo una pérdida de un 20% en la actividad de la enzima (y teniendo en cuenta los costes de la inmovilización), se puede llevar esta contribución a la mitad. Fresco-Taboada y col. han observado que estas enzimas pueden ser recicladas hasta 30 veces consecutivas, con sólo un 25% de pérdida de actividad al final del proceso<sup>127-130</sup>.

---

<sup>\*</sup> <https://www.pharmacompass.com/pharma/ingredients/clofarabine> (abril/2018)

<sup>♦</sup> En abril de 2018, EUR/USD = 0,81159

#### 4.3.4. Análisis de los «cuello de botella» detectados en el proceso de producción biocatalítica de clofarabina

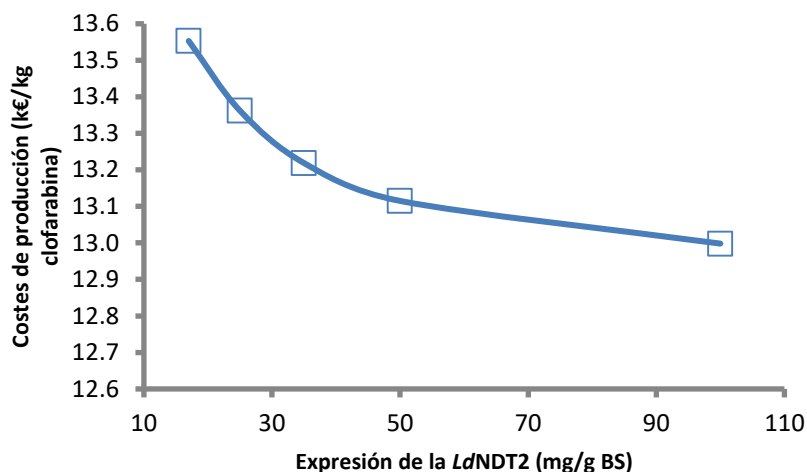
Los biocatalizadores, en especial aquellos enzimas producidos por microorganismos, representan una alternativa muy atractiva y sostenible para la síntesis de análogos de nucleósidos dado que pueden utilizarse en condiciones de reacción «suaves». Sin embargo, a la luz de los resultados obtenidos en este trabajo (y en buena parte de la literatura consultada), a menudo los procesos biocatalíticos necesitan ser adaptados a las condiciones químicas existentes en la producción a gran escala, al ambiente de las reacciones a las cuales se pretende su aplicación y en general cumplir con determinados criterios tecno-económicos, que les hagan competitivos frente a los procesos químicos de síntesis. En el caso específico de la síntesis biocatalítica de clofarabina, estas adaptaciones podrían incluir:

- 1- El rendimiento de producción del biocatalizador**, es decir, mejoras en la obtención (rendimiento volumétrico) del catalizador, lo cual redundaría en el precio del biocatalizador y al final en el de la reacción biocatalítica. En este trabajo, el coste del biocatalizador está dentro de los límites aceptables para biocatalizadores usados en la producción de APIs de alto valor añadido. Se plantea que es «coste ideal», para una enzima no purificada (o semipurificada) usada en procesos a escala industrial, debería estar entre 250–1000 €/kg<sup>45</sup>. El coste de producción del concentrado de *LdNDT2* se ha estimado, en este trabajo en **455,73 €/kg**.

A pesar de ello, este parámetro podría ser mejorado, por ejemplo, incidiendo en los niveles de expresión. En el caso de la *LdNDT2*, se han alcanzado niveles de expresión de **0,017 g de *LdNDT2*/gBS**, valores definitivamente mejorables, con la vista puesta en reducir los costes de producción (los rendimientos promedio están alrededor de 0,05-0,1g de enzima recombinante/gBS<sup>45</sup>). Por otro lado, se han alcanzado densidades celulares de **35g/BS por litro** de caldo fermentativo, cuando el promedio es 50g BS/L.

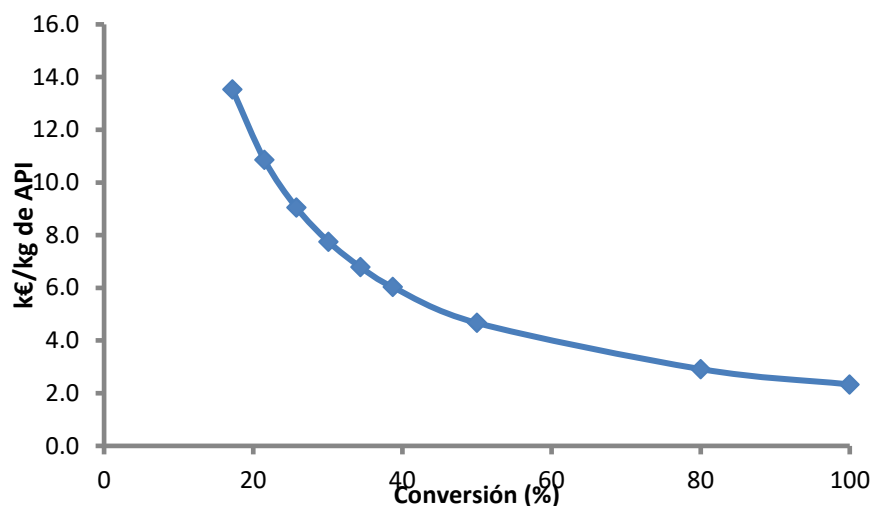
De cualquier manera las mejoras en la expresión tienen un límite (biológico), hasta un 10-12% de la biomasa en el mejor de los casos<sup>131</sup>, y esto hace que el efecto de mejorar no sea, quizá, el más importante. En la Figura 55 se ha

simulado el efecto de diferentes niveles de expresión sobre la economía del proceso de producción de clofarabina y los resultados muestran que variando los niveles de expresión, desde 17 mg/gBS (situación actual) hasta 100 mg/gBS, la mejora es de un 4,1%, lo que indica que no cambia significativamente el coste de producción de la clofarabina.



**Figura 55:** Influencia (simulación) del nivel de expresión de la enzima sobre los costes de producción de la clofarabina. k€: miles de Euros.

**2- Rendimiento de la reacción.** Puesto que operar con altos rendimientos influye de manera significativa en la disminución de los costes de producción, especialmente aquellos asociados a las materia prima, además de simplificar los procesos de recuperación de la clofarabina. El escenario aquí presentado parte de un rendimiento de conversión muy bajo, **17,2%**. En la Figura 56 se representa como cambian los costes totales del proceso de producción del API, si se incrementan los rendimientos de conversión. Tomando como ejemplo una mejora del 13,4%, lo cual es factible toda vez que este rendimiento de conversión (19,51%) ya ha sido obtenido a pequeña escala, en las mismas condiciones de reacción (Tabla 32), el coste de producción biocatalítica de la clofarabina rondaría los 11966 €/kg, es decir, 1568 € menos por kg de API. De cualquier manera, sigue estando por encima del precio de venta actual del API (637,18 €/kg). En las condiciones actuales de reacción, usar la vía del aumento de rendimiento de conversión para alcanzar costes por debajo del precio actual de venta, no se avizora como solución «única», pues ni con un 100% de conversión se consigue este objetivo (Figura 56).



**Figura 56:** Influencia (simulación) de los rendimientos de la reacción sobre los costes de producción de la clofarabina. **k€:** miles de Euros

No obstante, mejorar este parámetro es imprescindible, más adelante se discute brevemente el tema de la concentración inicial de los sustratos pero, enfocados en incrementar los rendimientos de conversión, una de las posibles soluciones podría estar en la ingenierización de esta enzima, para mejorar por ejemplo la tolerancia hacia los sustratos de la reacción. Las *N*-desoxirribosiltransferasas al parecer son enzimas con una gran « plasticidad » y en este sentido, por ejemplo, se han conseguido mutantes capaces de reconocer 2,3-didesoxirribosas y a 2,3-didehidro-2,3-didesoxirribosas como sustratos<sup>60</sup>.

La ingenierización del enzima debería ir acompañada de la ingenierización del bioproceso, toda vez que en aras de hacer viable este bioproceso, se necesitan altas concentraciones de sustratos y, en nuestro caso, una de las consecuencias es la escasa solubilidad de los mismos, y derivado de esto la disponibilidad de éstos para la reacción. En este sentido, se ha sopesado abordar el problema de la solubilidad desde varias aristas: i) añadiendo una segunda fase al medio de reacción, usando un disolvente orgánico o una resina (soporte sólido); ii) hacer funcionar el reactor en modo discontinuo alimentado, entre otras.

- 3- La productividad volumétrica** (*kg de clofarabina/unidad de volumen/h*). Este aspecto está muy relacionado con el anterior y afecta a los gastos en CAPEX, servicios, mano de obra, etc. Supone que un menor tiempo de reacción (**eficiencia catalítica**) sería lo deseable. Aquí, aspectos relacionados con la

ocupación de los equipos (en el escenario actual, 6-7 días), la concentración del producto, la reutilización del enzima son, entre otros, aspectos a tener en cuenta. Es muy probable que las velocidades de la reacción para la producción de clofarabina, puedan incrementarse por medio de una mayor dosis del biocatalizador, pero esto supone un incremento de los costes permisibles del biocatalizador, además de aumentar los costes de la recuperación del producto final. Por lo tanto, se ha de identificar una solución de compromiso para garantizar buenos rendimientos espaciales y temporales.

**4- Coste de la materia prima del proceso.** Aunque buena parte de los sustratos de la reacción se recuperan y reúsan en los procesos de biotransformación, el coste de la materia prima para la producción de clofarabina representa el 40% de todo el bioproceso. Los costes de producción están a menudo dominados por el coste de las materias primas y a menudo representan entre el 40% y el 90% de los costes de la biotransformación<sup>113</sup>, de ahí la necesidad de rendimientos de reacción muy elevados, en nuestro caso, altas conversiones, para bajar los costes de capital, los requerimientos energéticos (agitación, calentamiento, enfriamiento) y disminuir las emisiones, el impacto ambiental y los costes relacionados con la mano de obra.

**5- Solubilidad de los sustratos. Uso de solventes.** Ya se ha enfatizado en que, probablemente, los sustratos de esta biorreacción sean los que más influyen en los costes y en las cinéticas de producción de clofarabina. Por ello se considera un aspecto crítico, el aumentar la dosificación de éstos a concentraciones similares a otros procesos biocatalíticos exitosos, desde el punto de vista industrial que según Pollard y Woodley<sup>132</sup> están entre 50-100 g/L. En el caso de la producción de este API, la extremadamente baja solubilidad de **2CIA** en agua hace que, en el escenario actual, se trabaje con muy bajas concentraciones iniciales de sustratos, lo cual supone bajas productividades volumétricas y, como se ha indicado antes, altos costes de ocupación de los equipos y de procesado para la recuperación del API.

En este caso, estrategias de alimentación controlada de los sustratos, en especial de la **2CIA**, podría ser una de las soluciones. Esta estrategia podría resolver

problemas relacionados con la baja solubilidad, las velocidades de disolución, los problemas de inhibición/toxicidad y de recuperación del producto final.

También, y a la luz de experimentos preliminares usando cosolventes, se ha pensado en ellos para aumentar la solubilidad de los sustratos en la fase acuosa. Usando por ejemplo cosolventes miscibles en agua o cosolventes inmiscibles en agua, estos últimos actuando como reservorio para dosificar los sustratos hacia la fase acuosa.

**6- Concentración final del producto** (clofarabina). En el proceso aquí propuesto se han logrado concentraciones finales de clofarabina de alrededor de 0,5 g/L. Este posiblemente sea de los retos más importantes a superar para escalar la producción de este API, ya que, dependiendo del sector del mercado, por lo general, lo recomendado es conseguir concentraciones cercanas a los 50 g/L<sup>132</sup>. En condiciones naturales, las enzimas operan a concentraciones de sustratos y de productos que rondan las partes por millón (ppm), de aquí que forzarlas a trabajar en las condiciones que imponen los bioprocesos en el escalado constituye un reto de ingeniería, tanto proteica, como de procesos.

La Tabla 42 resume los «*cuello de botella*» detectados en el proceso de producción biocatalítica de clofarabina y sus las soluciones propuestas para los mismos.

#### 4.3.5. *Comparación entre la síntesis química y la síntesis enzimática.*

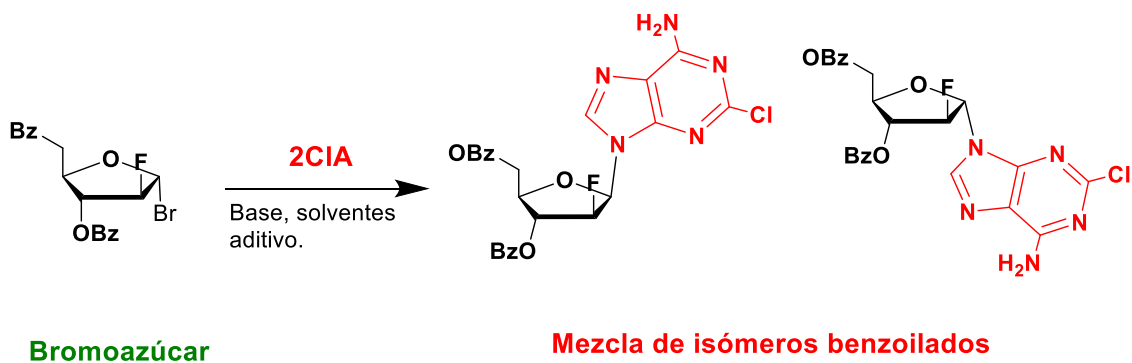
El proceso de síntesis propuesto en este trabajo, aún está lejos de ser considerado como “viable”, desde el punto de vista de la economía del bioproceso, vistos los costes de fabricación y los rendimientos obtenidos. De cualquier manera, en este punto, y para tener más argumentos, a favor o en contra del proceso biocatalítico, se ha considerado una comparativa con el proceso de síntesis más extendido para la síntesis química de clofarabina, el propuesto por Bauta y col.<sup>37</sup>.

El proceso de síntesis química es un proceso eficiente de producción en el cual se obtiene al final clofarabina pura (menos de 0,1% de contaminación con el anómero  $\alpha$ ), sin utilizar separaciones cromatográficas. Para ello, hacen reaccionar la 2-cloroadenina



con 1-bromo-2-desoxi-2-fluoro-3,5-di-*O*-benzoil- $\beta$ -D-ribofuranosa y *tert*-butóxido de potasio, todo en una mezcla de tres solventes (acetonitrilo/ alcohol *tert*-amílico / 1,2-dicloroetano), para obtener en 19 horas, a 50°C, una mezcla de isómeros benzoilados (Figura 57), con una relación anomérica de 164:1 ( $\beta/\alpha$ ). El rendimiento de esta etapa es del 50%.

Luego se somete la clofarabina protegida a una serie de pasos que brevemente se pueden resumir de la siguiente manera: A una suspensión obtenida disolviendo la clofarabina protegida con metanol, se le adiciona una solución de metóxido de sodio (30% p/p., en metanol) y se calienta, unas 7 horas. Cuando se completa la reacción, la solución resultante es neutralizada con ácido acético glacial y la mezcla enfriada a -10°C. El sólido resultante se filtra y se lava con metanol frío y se vuelve a resuspender en metanol. La suspensión es entonces calentada (63°C) con agitación hasta la disolución del sólido y entonces es enfriada entre -8 y -12°C para obtener los cristales de clofarabina mediante filtrado al vacío. El rendimiento de esta etapa de purificación es del 64% y el global de todo el proceso del 32%.



**Figura 57:** Síntesis química de clofarabina propuesta por Bauta y col.<sup>37, 38, 125</sup>

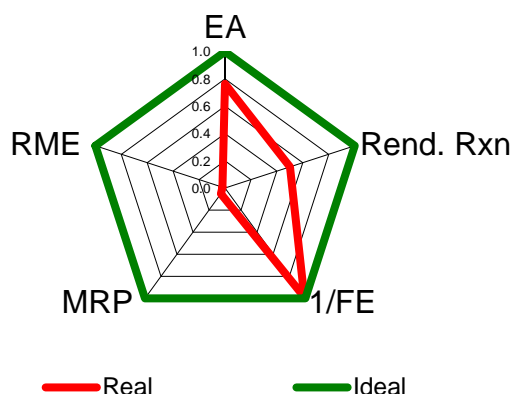
**Tabla 42: Resumen de los parámetros identificados como «cuello de botella» del proceso biocatalítico.**

Parámetro	Posibles soluciones
Relacionados con el biocatalizador	Densidad celular
	(gBS/L)
	Actividad específica Expresión de la <i>Ld</i> NDT
	(U/g BSL <sup>-1</sup> )
	Actividad enzimática
Estabilidad	Ingeniería de proteínas, mutagénesis
	Inmovilización
	Ingeniería de Proteínas
	Cambios en el modo de operación
Relacionado con los sustratos	Actividad NDT (h)
	Cambios en el medio de reacción
	Solubilidad (g/L)
Parámetros de proceso.	Toxicidad (g/L),
	Inhibición (g/L)
	Uso de solventes miscibles o inmiscibles
	Alimentación controlada de los sustratos, o uso de una segunda fase
	Inmovilización del biocatalizador
	Configuración del reactor
	Adición de surfactantes/cosolventes

Comparar el proceso antes descrito con el proceso biocatalítico, es una tarea difícil, toda vez que hay que tener en cuenta muchos parámetros y datos que no siempre están disponibles. Aquí, para los cálculos nos hemos apoyado en herramientas como el programa *Ecoscale*<sup>133</sup> y los algoritmos desarrollados por Andraos<sup>134, 135</sup>. Estos últimos, miden cuán “verdes” son los procesos de síntesis a partir de coeficientes estequiométricos y de los resultados de los procedimientos experimentales. Al final se produce un diagrama, radial, en forma de pentágono, cuyos vértices son: la Economía de átomos (**AE**), la Eficiencia másica de la reacción (**RME**), el parámetro de recuperación de materiales (**MRP**), el rendimiento de la reacción (**Rend. Rxn**) y el inverso del Factor Estequiométrico (**FE**). Estos cinco parámetros pueden ser resumidos en un parámetro llamado vector de magnitud (**VMR**), definido como la suma cuadrática de los cinco parámetros anteriores. Todos estos parámetros están normalizados entre cero y uno.

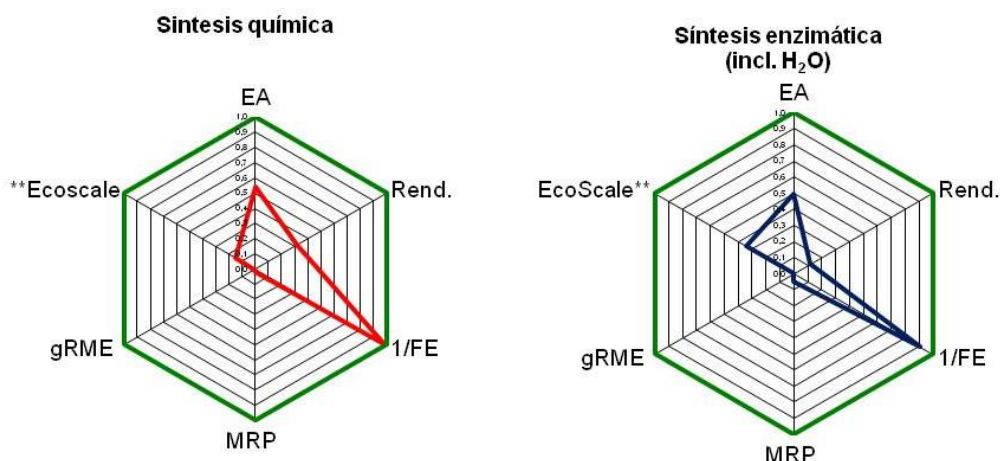
En la Figura 58 se muestra el resultado de este análisis para la reacción de síntesis propuesta por Bauta y col.<sup>37</sup>, sólo hasta la obtención de la clofarabina protegida (Figura 57). En la Figura 58, en el pentágono, 1 (en verde) representa el escenario ideal con respecto a cada una de las cinco métricas aplicadas. Hasta la obtención de la clofarabina protegida (Figura 57), el proceso sintético se caracteriza por coeficientes **FE**<sup>-1</sup> cercanos al ideal, rendimientos moderados, economía atómica relativamente alta, pero muy baja eficiencia másica de la reacción. En la Tabla 45 se muestran los valores de estos parámetros.

En diagrama de la Figura 58 no tiene en cuenta, o no calcula, el impacto ambiental, ni la naturaleza de las sustancias que intervienen en el proceso. Para evaluar la calidad de estas reacciones, en base a parámetros económicos y ecológicos, se ha usado la herramienta semicuantitativa desarrollada por Van Aken y col.<sup>133</sup>. *Ecoscale* utiliza una escala del 0 al 100, de forma que el 0 representa una reacción que ha resultado completamente fallida (rendimiento = 0%), mientras que el 100 representa una transformación ideal en la que un sustrato se transforma, con un rendimiento del 100%, a partir de reactivos de bajo precio, a temperatura ambiente, con un riesgo mínimo para el operador y un impacto medioambiental despreciable<sup>136</sup>. La Tabla 43 muestra como penalizan los 6 parámetros (Rendimiento; Precio de reactivos; catalizadores y disolventes; Seguridad; Manipulación; Temperatura/tiempo; Procesado y purificación) evaluados por *Ecoscale*.



**Figura 58:** Pentágono radial obtenido a partir del análisis de la síntesis química de la clofarabina protegida, en el proceso propuesto por Bauta y col.<sup>36</sup> **EA:** Economía atómica; **Rend. Rxn:** Rendimiento; **RME:** Eficiencia másica de la reacción; **FE:** Factor estequiométrico.

En este trabajo, la comparativa de los procesos de síntesis de clofarabina, los diagramas de la Figura 59 no son pentagonales, sino hexagonales, dado que uno de los vértices incluye el resultado del análisis del impacto medioambiental que aporta *Ecoscale*. Al igual que en la representación del pentágono de la Figura 58, los vértices del polígono corresponden al proceso ideal, en tanto que el polígono distorsionado, en el interior, representa el procedimiento real. Estas distorsiones nos ayudaran en las comparativas, pero sobre todo, a identificar que parámetros han de ser mejorados en aras de una síntesis más eficiente y respetuosa con el medio ambiente.



**Figura 59:** Hexágonos radiales representativos de la síntesis de clofarabina. ; **Rend. Rxn:** Rendimiento; **RME:** Eficiencia másica de la reacción; **FE:** Factor estequiométrico, \*\*:Se excluyen las penalizaciones correspondientes a la seguridad.

Tabla 43: **Puntuación para el cálculo de la Ecoescala.**

<b>Parámetro</b>	<b>Penalización</b>
<b>1- Rendimiento</b>	$(100 - \text{Rendimiento})/2$
<b>2- Precio de los reactivos</b> (para obtener 10 mmol de producto )	
Barato (< 10 US\$)	0
Caro (entre 10 y 50 US\$)	3
Muy caro (> 50 US\$)	5
<b>3- Seguridad</b> (Basado en los códigos de peligrosidad)	
GHS09 (Peligroso para el medio ambiente)	5
GHS06 (tóxico)	5
GHS02 (inflamable)	5
GHS01 (explosivo)	10
GHS07, GHS08 (extremadamente toxico)	10
<b>4- Manipulación</b>	
Manipulación convencional	0
Instrumentos para la adición de reactivos	1
Técnicas no convencionales de activación	2
Equipos de presión > 1 atm	3
Material de vidrio, especial	1
Atmósfera inerte	1
Recinto cerrado	3
<b>5- Temperatura/Tiempo</b>	
Temperatura ambiente, < 1 h	0
Temperatura ambiente, < 24 h	1
Calefacción < 1 h	2
Calefacción > 1 h	3
Enfriamiento a 0 °C	4
Enfriamiento a < 0 °C	5
<b>6- Procesado/Purificación</b>	
Ninguno	0
Enfriamiento a Temperatura ambiente	0
Adición de disolventes	0
Filtración	0
Eliminación de disolventes p.e. < 150 °C	0
Cristalización y filtración	1
Eliminación de disolventes p.e. > 150 °C	2
Extracción en fase sólida	2
Destilación	3
Sublimación	3
Extracción Líquido-líquido	3
Separación cromatográfica	10

### A. Economía atómica (EA).

El concepto de economía atómica se define como “la masa molecular de los productos dividida por la suma de las masas moleculares de todos los reactantes”, expresando el resultado en % (Ecuación 7):

$$\text{Ecuación 7} \quad \text{EA} = \left( \frac{\text{MW del Producto}}{\sum \text{MW reactantes}} \right) \times 100$$

En el caso de la obtención de la clofarabina, EA de la síntesis química, es relativamente más alta (**54%**) que la EA de la síntesis enzimática (**49,7%**). En cualquier caso, esto significa que en procesos ideales, las dos vías de síntesis tienen como residuos de la reacción prácticamente la mitad de los reactivos de partida (46% en el proceso de síntesis química y 50,3% la biocatalítica). En el caso de la síntesis enzimática la explicación podría estar en la toxicidad (para la enzima) de los reactivos de partida. Aquí cobra importancia lo discutido anteriormente, en lo referido a la posibilidad de dosificar de manera controlada los reactivos de partida y de este modo limitar la acumulación de los productos de partida.

### B. Rendimiento (R).

Posiblemente uno de los parámetros más críticos a la hora de valorar la eficiencia de un proceso de síntesis, pues mide la cantidad efectiva de productos y residuos generados a partir de los reactantes. El rendimiento (%) puede ser definido según la Ecuación 8. En ambos casos el rendimiento global de la reacción es bajo, **11,8%** para la síntesis enzimática y **32%** para la síntesis química.

**Ecuación 8:**

$$R = \left( \frac{\text{número de moles obtenidos del producto}}{\text{máximo número de moles teóricos, que pueden obtenerse a partir del reactivo limitante}} \right) \times 100$$

Para comparar ambos procesos, en función de los dos parámetros analizados hasta el momento, se ha empleado un parámetro que los relaciona y tiene en cuenta para los cálculos, la ecuación estequiométrica, el exceso de reactivos y el rendimiento. El referido parámetro, **eficiencia atómica**, fue introducido por K. M. Doxsee<sup>137</sup> y puede calcularse usando la Ecuación 9. Según este parámetro, la síntesis química es más sostenible (**17,3%** vs. **5,86**) que el proceso enzimático.

**Ecuación 9:** 
$$EFA = \frac{\text{Rendimiento (\%)} \times \text{Economía Atómica(\%)}}{100}$$

### C. Eficiencia másica de reacción (RME<sub>Curzons</sub>)

En el 2001, Curzons y *col.* definieron el concepto de “Eficiencia másica de reacción” (RME<sub>Curzons</sub>) como el porcentaje de la masa de los reactivos que permanece en el producto final. En su cálculo están incluidos de forma indirecta (Ecuación 10), la estequiometría, el rendimiento y la economía atómica. Los valores obtenidos para este parámetro son muy similares, **11,4%** para la síntesis enzimática y **10,1%** para la síntesis química, indicativo del alto impacto ambiental de ambos procesos.

**Ecuación 10:** 
$$RME_{Curzons} = \frac{\text{Masa del producto}}{\sum \text{Masa de todos los reactivos}} \times 100$$

### D. Factor estequiométrico (FE)

El factor estequiométrico (FE) permite una comparación entre un proceso experimental y un proceso similar en el que todos los reactivos deben utilizarse estequiométricamente, es decir refleja el grado en que se utiliza un exceso de reactivo(s) . Se define el factor estequiométrico según la Ecuación 11:

**Ecuación 11:**

$$FE = 1 + \frac{\sum \text{Masa de todos los reactivos} - \sum \text{Masa de todos los reactivos en un proceso estequiométrico}}{\sum \text{Masa de todos los reactivos en un proceso estequiométrico}}$$

Los factores estequiométricos calculados, para la síntesis química y enzimática fueron 1,102 y 1,012, respectivamente<sup>138</sup>. Lo que significa que en el proceso propuesto por Bauta y *col.*<sup>37</sup>, un 10% de los reactantes es usado en exceso y el enzimático un 1%.

### E. Eficiencia másica global de las reacciones de síntesis y Rendimiento másico efectivo.

Si consideramos en las comparativas a todos los reactivos y productos, incluyendo los solventes o cualquier otra sustancia que tome parte de las reacciones de síntesis de la clofarabina, se puede utilizar el parámetro métrico gRME<sup>139, 140</sup> (Eficiencia másica global de la reacción), es decir el balance de masas de todo el proceso, incluidos los solventes. La Ecuación 12 permite el cálculo de este parámetro, y resulta que en el caso de la síntesis química, la eficiencia global de la reacción es de **0,4 %**, teniendo en cuenta que el proceso intervienen además de los reactivos de partida (bromoazúcar y 2CIA), solventes de la reacción (acetonitrilo, alcohol tert-amílico y DCE), los catalizadores y los reactivos utilizados para la recuperación y/o purificación de la clofarabina (acetato de butilo, heptano, metanol, ácido acético y DCE). Es decir, más del 99% de la masa incorporada al proceso de síntesis química, no forma parte del producto final (clofarabina). En el caso de la síntesis enzimática, el gRME fue del **0,3%**, teniendo en cuenta que en el cálculo también se tiene en cuenta el único solvente de la reacción, el agua.

Ecuación 12: 
$$\text{gRME} = \frac{\text{masa del producto}}{\sum \text{Masa de todos los reactivos del proceso}}$$

### F. MRP: Parámetro de recuperación de masa.

El parámetro de recuperación de masa (MRP) considera el uso de otros materiales como solventes u otros materiales usados en la recuperación, purificación, etc. y puede ser calculado según Ecuación 13. Cuando todos los catalizadores, solventes y otros materiales post-reacción son recuperados o eliminados, el MRP=1. En las estrategias de síntesis comparadas se recupera un porcentaje muy bajo del material no usado, un **2,23%** en la síntesis química y **5,05%** en la enzimática.

Ecuación 13 
$$\text{MRP} = \frac{\text{RME} * \text{FE}}{\text{Rend.} * \text{EA}}$$



## G. La Ecoescala.

En los hexágonos radiales de la Figura 59 uno de los vértices se corresponde con un parámetro que va más allá de los balances de masa para comparar los dos procesos de síntesis de clofarabina, la **Ecoescala**, una herramienta desarrollada por Van Aken y col.<sup>133</sup>, para evaluar la calidad de las reacciones. En la Tabla 44 se muestran los puntos de penalización asociados a los dos procesos de síntesis comparados en este trabajo y la puntuación final de la reacción (o Ecoescala), calculada según la Ecuación 14

$$\text{Ecuación 14} \quad \text{Ecoescala} = 100 - \sum \text{puntos de penalización}$$

En la Tabla 44 se muestran los resultados de la Ecoescala de los procesos de síntesis propuestos por Bauta y col.<sup>37</sup> y su comparación con el proceso de síntesis enzimática propuesto en este trabajo. Dado que los valores obtenidos están muy por debajo de 50, ambos procesos de síntesis de clofarabina son “inaceptables”, desde el punto de vista de esta herramienta. De los dos procesos, el de síntesis química tiene mayor cantidad de penalizaciones (Tabla 44) fundamentalmente, por el precio de los reactivos de partida, el uso de solventes (y su toxicidad) y su eliminación al final de la reacción. En el caso de la síntesis enzimática, penalizan, en gran medida, los rendimientos de la reacción y los reactivos de partida (precios y toxicidad).

**Tabla 44: Puntuación para el cálculo de la Ecoescala en la síntesis de clofarabina**

Parámetros	Penalizaciones por vía	
	Química	Enzimática
Rendimiento <sup>a</sup>	34 (32%)	37.5 (11,8%)
Seguridad	30	10
Reactivos <sup>b</sup>	30	13
Manipulación	1	0
Temperatura/Tiempo	12	7
Procesado/Purificación	8	4
Ecoescala (excluyendo seguridad)	15	34,6

<sup>a</sup> Entre paréntesis el rendimiento global de la síntesis.

<sup>b</sup> Se incluyen todos los reactivos de la síntesis, solventes, reactivos auxiliares y catalizadores.

## H. Intensidad másica del proceso y Factor E

Aunque no han sido incluidos en la representación radial de la Figura 59, en muchos procesos de síntesis se utilizan las métricas de Intensidad másica (**PMI**) y de **Factor E**, para estimar cuan “verdes” son los procesos sintéticos. La intensidad másica del proceso puede ser calculada como el inverso de la Eficiencia másica de reacción (Ecuación 15), o como la suma de toda la masa usada en un proceso (global o parcial) dividido por la masa del producto final<sup>141</sup>( Ecuación 16). La intensidad másica calculada para el proceso de síntesis química propuesto por Bauta y col.<sup>37</sup> fue de 260,47 y para el proceso enzimático propuesto en este trabajo fue de 371,6. Esto significa que para obtener 1 kg de clofarabina, se introducen 260,47 kg y 371,6 kg, en los procesos químico y enzimático, respectivamente.

**Ecuación 15** 
$$\text{PMI} = \frac{1}{\text{gRME}}$$

**Ecuación 16** 
$$\text{PMI} = \frac{\sum \text{masas introducidas en el proceso}}{\text{masa del producto final}}$$

Por su parte el **Factor E**<sup>142</sup> se calcula dividiendo la masa total de residuos producida en la preparación de un compuesto, por la masa total de producto sintetizado (Ecuación 17). El **factor E** y la **PMI** se pueden relacionar mediante la Ecuación 18, y en los procesos comparados suponen que se generan 259,47 kg de residuos, en el proceso de síntesis química, y 370,6 kg en el enzimático, para producir un kg de clofarabina.

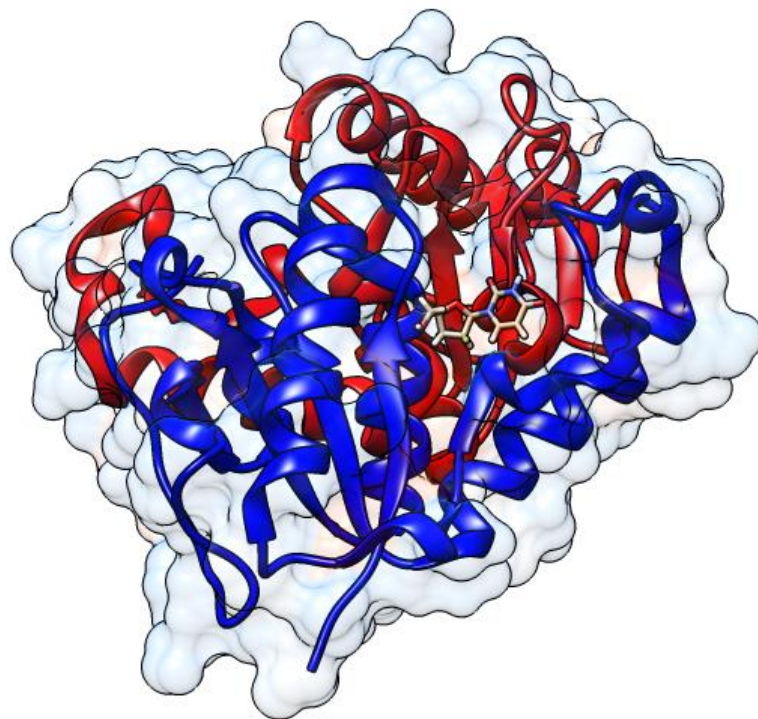
**Ecuación 17** 
$$\text{factor E} = \frac{\text{masa total de residuos}}{\text{masa de producto}}$$

**Ecuación 18** 
$$\text{PMI} = \text{Factor E} + 1$$

**Tabla 45: Métricas asociadas a los escenarios comparados para la síntesis de clofarabina.**

	EA (%)	Rend. (%)	RME <sub>Curzons</sub> (%)	MRP (%)	CE	gRME (%)	VMR	EcoEscala
Síntesis química	54	32	10,1	2,23	1,102	4	52,4	15
Síntesis enzimática	49,7	17,2	11,4	5,04	1,012	3	46,5	34,6

La Tabla 45 muestra un resumen de las métricas aplicadas para comparar, cuantitativamente, la «sostenibilidad» del nuevo proceso biocatalítico de síntesis propuesto en este trabajo con el proceso de síntesis química usado en la actualidad. A primera vista, el proceso biocatalítico no mejora el proceso de obtención química del API, sobre todo en términos de rendimientos, porque el resto de los parámetros evaluados son “relativamente” similares. Sin embargo, al hacer un análisis más cuidadoso de estas dos rutas de síntesis, nuestra visión de los mismos podría cambiar si pensamos que en el proceso biotecnológico, los cálculos tienen en cuenta al agua. Pero, si la excluimos de los cálculos, esta evaluación, no sería del todo correcta. El agua es por lo general considerada el solvente «verde» pero, esta afirmación es válida solo para el agua pura. Aguas contaminadas con microorganismos o sustancias químicas, aunque sea en cantidades muy pequeñas, constituyen un problema medioambiental y necesitan ser tratadas y/o purificadas antes de reutilizarlas o liberarlas.



# **CONCLUSIONES Y PERSPECTIVAS.**

## 5. CONCLUSIONES Y PERSPECTIVAS.

### 5.1. Conclusiones.

Los 2-desoxi-2-fluoro- $\beta$ -D-arabinofuranósidos de pirimidina y de purina son análogos de nucleósidos con un amplio espectro de actividad biológica<sup>74</sup> y de entre ellos la clofarabina<sup>74, 75</sup>, 9- (2-desoxi-2-fluoro- $\beta$ -D-arabinofuranosil) -2-cloroadenina, ha sido aplicada con éxito para el tratamiento de la leucemia aguda pediátrica<sup>29, 76</sup>. Se han establecido varios métodos de síntesis química para su producción, la mayoría, con bajos rendimientos globales, formación de las mezclas  $\alpha/\beta$  de glucósidos, etc., lo cual supone etapas de purificaciones cromatográficas para separarlos, además del uso de grandes volúmenes de disolventes orgánicos.

La búsqueda de alternativas “verdes”, más eficientes, ha conducido a proponer a las enzimas para llevar a cabo la síntesis de este importante API y en este contexto, en la actualidad no existen alternativas viables, desde el punto de vista comercial, para la síntesis de este análogo. Las aproximaciones biocatalíticas llevadas a cabo hasta el momento para la síntesis de clofarabina, se han hecho usando enzimas del tipo nucleósido fosforilasas. Pero estas estrategias pasan por el uso de dos enzimas y por la necesidad de aportar el desoxi-2-fluoro- $\alpha$ -D-arabinofuranosa-1-fosfato, sea por síntesis química (Figura 15), o usando una cascada de reacciones catalizadas por varias enzimas relacionadas con la síntesis de nucleósidos (riboquinasas y fosfopentomutasas (Figura 16).

El presente estudio fue iniciado bajo la hipótesis de utilizar enzimas de tipo NDT, capaces de llevar, al igual que las NPs, reacciones de transglicosilación pero, a diferencia de éstas, no precisan del intermediario  $\alpha$ -D-pentofuranosil-1-fosfato. Específicamente, la hipótesis de trabajo se basó en demostrar, primero, la posibilidad de obtener clofarabina usando NDTs y luego sentar las bases para la síntesis de este API, de manera sostenible, desde el punto de vista económico y ambiental. En este apartado se resumen las principales conclusiones y perspectivas de este trabajo.

**A. La nucleósido 2'-desoxirribosiltransferasa de *Lactobacillus leichmannii* ha sido seleccionada para llevar a cabo la reacción de síntesis de clofarabina.**

En la base de datos **Protein Data Bank** ([www.rcsb.org/pdb](http://www.rcsb.org/pdb)), tres proteínas cristalinas aparecen anotadas como “*Nucleoside 2-deoxyribosyltransferase*”. Las mismas fueron utilizadas como receptores en los estudios de acoplamiento (*docking*) molecular, usando como ligandos: *i*) un sustrato natural de estas enzimas, la 2'-desoxiuridina (**dUrd**) y *ii*) el sustrato propuesto para la síntesis de clofarabina, 2'-fluoro-arabinofuranosil-2'-desoxiuridina (**araF**). Del estudio se pudo concluir que dada la cercanía geométrica del 2'-fluoro-arabinofuranosil-2'-desoxiuridina, a los residuos involucrados en el centro catalítico, la orientación espacial y los valores energéticos (energía favorable) para finalmente establecer un enlace covalente entre el residuo **Glu** nucleofílico y el pentofuranosilo, la nucleósido 2'-desoxirribosiltransferasa de *Lactobacillus leichmannii*, ha sido seleccionada para llevar a cabo la reacción de síntesis de clofarabina.

**B. Se ha clonado el gen *ndt2* de *L. delbrueckii* y se ha producido la enzima en cultivos de alta densidad celular.**

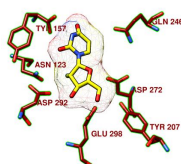
Con unas pequeñas variaciones en la secuencia original del gen ***ndt2*** de *L. delbrueckii*, que incluyen la eliminación de los 7 primeros residuos y el cambio de tres residuos en el resto de la secuencia, la enzima con actividad 2'-desoxi-2'-fluororibosil-transferasa ha sido clonada y expresada exitosamente. En este estudio además, se ha establecido una fermentación semicontinua, con la que se consiguen 0,017g de la enzima por gramo de biomasa seca y el coste estimado de la producción de un kilogramo de un extracto crudo (liofilizado, con actividad 2'-desoxi-2'-fluororibosil-transferasa) ronda los 456 euros.

**C. Se ha conseguido la síntesis de clofarabina usando la NdRT de *L. delbrueckii*.**

La síntesis enzimática de clofarabina ya había sido abordada con anterioridad, usando nucleósido fosforilasas. En este trabajo, por primera vez se logra la síntesis de este análogo nucleosídico, usando una **NdRT**. En el escenario planteado, se consiguen rendimientos de conversión de aproximadamente 11,8% y unos rendimientos globales del proceso de obtención de este API purificado de un 17,2%.

**D. La síntesis biocatalítica de clofarabina usando NdRT es una alternativa prometedora, que ha de mejorar determinadas condiciones y costes de operación para ser competitiva frente al proceso de síntesis química.**

La aplicación industrial del proceso propuesto debe superar la barrera del “pobre” rendimiento global obtenido y de la productividad volumétrica, en aras de ser competitiva con el proceso de síntesis química. No obstante, el análisis de la sostenibilidad del proceso, indica que el proceso enzimático es una alternativa, muy prometedora para la síntesis de clofarabina.

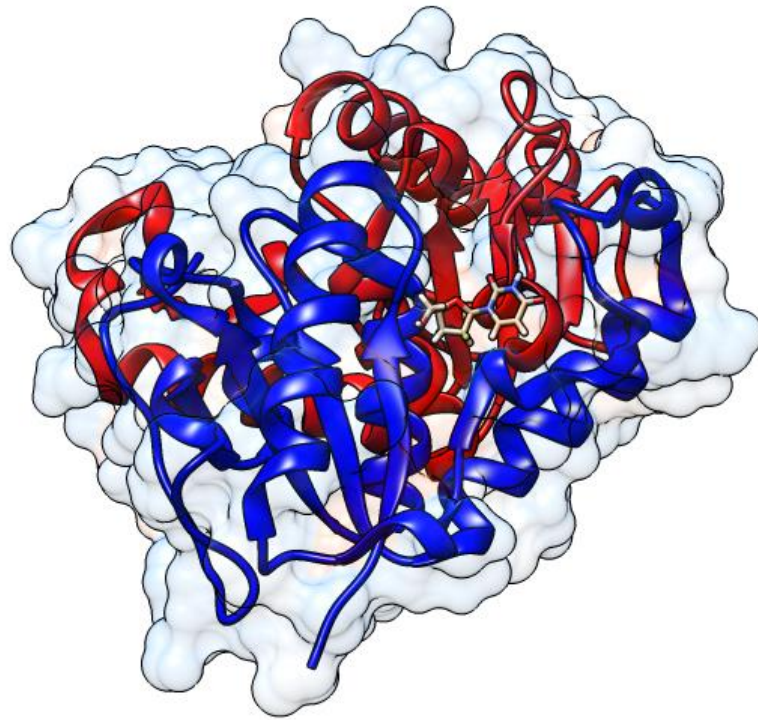


## **5.2. Perspectivas futuras**

En resumen, los datos y discusiones presentados en esta **Tesis Doctoral**, demuestran el potencial y la plasticidad de las enzimas Nucleósido 2'-desoxirribosiltransferasas, para la síntesis de 2'-desoxifluoroarabinosil nucleósidos y en especial de clofarabina. Las perspectivas de trabajos futuros se orientan en dos direcciones: i) la optimización de la reacción enzimática y ii) aumentar el repertorio de **NdRTs**. La primera consistiría en los trabajos destinados a optimizar la reacción de transglicosilación presentada en este estudio. En concreto, se podrían mejorar aspectos relacionados con la actividad catalítica de esta enzima así, por ejemplo, a partir del conocimiento del centro activo del enzima modificar mediante ingeniería de proteínas o mutagénesis la eficiencia catalítica. También sería interesante cambiar el modo de operación de las bioreacciones para, por ejemplo, limitar los efectos de inhibición por sustratos/productos o estudiar la utilización de solventes orgánicos, dada la baja solubilidad de los sustratos en agua. Otro punto interesante sería el estudio de nuevos sustratos para la reacción enzimática, pues sus precios hacen prácticamente inviable su utilización, además de aumentar la estabilidad del biocatalizador a través de la inmovilización.

Luego, la segunda línea de actuación, supondría la búsqueda de nuevos biocatalizadores, más versátiles y robustos, con mayor eficiencia catalítica y estabilidades mejoradas bajo las condiciones de reacción que impone la síntesis de 2'-desoxifluoroarabinosil nucleósidos.





# **BIBLIOGRAFÍA.**



## 6. BIBLIOGRAFIA

1. Stewart, B. W.; Wild, C. P., World cancer report 2014. **2014**.
2. Ferlay, J.; Soerjomataram, I.; Dikshit, R.; Eser, S.; Mathers, C.; Rebelo, M.; Parkin, D. M.; Forman, D.; Bray, F., Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer* **2015**, *136* (5), E359-E386.
3. Bautista, F.; Gallego, S.; Canete, A.; Mora, J.; Heredia, C. D.; Cruz, O.; Fernandez, J. M.; Rives, S.; Madero, L.; Castel, V.; Cela, M. E.; Ramirez, G.; Sabado, C.; Acha, T.; Astigarraga, I.; Sastre, A.; Munoz, A.; Guibelalde, M.; Moreno, L., Landscape of early clinical trials for childhood and adolescence cancer in Spain. *Clinical & Translational Oncology* **2016**, *18* (7), 708-713.
4. Infantiles, R. E. d. T. Cifras básicas del cáncer infantil en España. <http://www.uv.es/rnti/cifrasCancer.html>.
5. Vachani, C. Leucemia: Los Fundamentos. <http://es.oncolink.org/tipos-de-cancer/leucemia/leucemia-los-fundamentos>.
6. JF, M.; KR, R.; CP, S.; DG, P., *Acute Lymphoblastic Leukemia*. 6 th ed ed.; Pizzo A, Poplack DG, eds: **2011**; p 518-565.
7. Hitzler, J. K.; Zipursky, A., Origins of leukaemia in children with Down syndrome. *Nature Reviews Cancer* **2005**, *5* (1), 11-20.
8. Echebarria Barona, A. B. Implicación de lospolimorfismos genéticos en la neurotoxicidad por Vincristina en pacientes pediátricos diagnosticados de Leucemia Linfoblástica Aguda. **Tesis doctoral**, Universidad del País Vasco, Facultad de Medicina y Odontología, **2016**.
9. Inaba, H.; Greaves, M.; Mullighan, C. G., Acute lymphoblastic leukaemia. *Lancet* **2013**, *381* (9881), 1943-1955.
10. Markets, M. , Acute lymphocytic leukemia therapeutics market:2011–2020. **2016**.
11. Abeijón martínez, P. Heterocarbonucleósidos: Síntesis y Evaluación biológica de análogos tiofenocondensados del abacavir. Universidade de Santiago . Faculdade de farmacia, **2008**.

12. Romeo, G.; Chiacchio, U.; Corsaro, A.; Merino, P., Chemical Synthesis of Heterocyclic-Sugar Nucleoside Analogues. *Chemical Reviews* **2010**, *110* (6), 3337-3370.
13. Avendaño Lopez, M. D. C., *Introducción a la Química Farmacéutica*. 2<sup>a</sup> edición ed.; **2001**.
14. Shelton, J.; Lu, X.; Hollenbaugh, J. A.; Cho, J. H.; Amblard, F.; Schinazi, R. F., Metabolism, Biochemical Actions, and Chemical Synthesis of Anticancer Nucleosides, Nucleotides, and Base Analogs. *Chemical Reviews* **2016**, *116* (23), 14379-14455.
15. Zhang, J.; Visser, F.; King, K. M.; Baldwin, S. A.; Young, J. D.; Cass, C. E., The role of nucleoside transporters in cancer chemotherapy with nucleoside drugs. *Cancer and Metastasis Reviews* **2007**, *26* (1), 85-110.
16. Van Rompay, A. R.; Johansson, M.; Karlsson, A., Phosphorylation of nucleosides and nucleoside analogs by mammalian nucleoside monophosphate kinases. *Pharmacology & Therapeutics* **2000**, *87* (2-3), 189-198.
17. Smal, C.; Vertommen, D.; Bertrand, L.; Ntamashimikiro, S.; Rider, M. H.; Van Den Neste, E.; Bontemps, F., Identification of in vivo phosphorylation sites on human deoxycytidine kinase - Role of Ser-74 in the control of enzyme activity. *Journal of Biological Chemistry* **2006**, *281* (8), 4887-4893.
18. Achanta, G.; Pelicano, H.; Feng, L.; Plunkett, W.; Huang, P., Interaction of p53 and DNA-PK in response to nucleoside analogues: Potential role as a sensor complex for DNA damage. *Cancer Research* **2001**, *61* (24), 8723-8729.
19. Sahasranaman, S.; Howard, D.; Roy, S., Clinical pharmacology and pharmacogenetics of thiopurines. *European Journal of Clinical Pharmacology* **2008**, *64* (8), 753-767.
20. Satishkumar, S.; Vuram, P. K.; Relangi, S. S.; Gurram, V.; Zhou, H.; Kreitman, R. J.; Montemayor, M. M. M.; Yang, L. J.; Kaliyaperumal, M.; Sharma, S.; Pottabathini, N.; Lakshman, M. K., Cladribine Analogues via O-6-(Benzotriazolyl) Derivatives of Guanine Nucleosides. *Molecules* **2015**, *20* (10), 18437-18463.
21. Buie, L. W.; Epstein, S. S.; Lindley, C. M., Nelarabine: A novel purine antimetabolite antineoplastic agent. *Clinical Therapeutics* **2007**, *29* (9), 1887-1899.
22. Robak, P.; Robak, T., Older and new purine nucleoside analogs for patients with acute leukemias. *Cancer Treat Rev* **2013**, *39* (8), 851-61.
23. Robak, T.; Robak, P., Purine nucleoside analogs in the treatment of rarer chronic lymphoid leukemias. *Curr Pharm Des* **2012**, *18* (23), 3373-88.

24. Cividini, F.; Pesi, R.; Chaloin, L.; Allegrini, S.; Camici, M.; Cros-Perrial, E.; Dumontet, C.; Jordheim, L. P.; Tozzi, M. G., The purine analog fludarabine acts as a cytosolic 5'-nucleotidase II inhibitor. *Biochemical Pharmacology* **2015**, *94* (2), 63-68.
25. Navada, S. C.; Steinmann, J.; Lubbert, M.; Silverman, L. R., Clinical development of demethylating agents in hematology. *Journal of Clinical Investigation* **2014**, *124* (1), 40-46.
26. Chhikara, B. S.; Parang, K., Development of cytarabine prodrugs and delivery systems for leukemia treatment. *Expert Opinion on Drug Delivery* **2010**, *7* (12), 1399-1414.
27. Kim, T. M.; Kim, S.; Ahn, Y.-O.; Lee, S.-H.; Kim, D.-W.; Heo, D. S., Anti-cancer activity of gemcitabine against natural killer cell leukemia/lymphoma. *Leukemia & Lymphoma* **2014**, *55* (4), 940-943.
28. Carson, D. A.; Wasson, D. B.; Kaye, J.; Ullman, B.; Martin, D. W.; Robins, R. K.; Montgomery, J. A., DEoxycytidine kinase-mediated toxicity of deoxyadenosine analogs toward malignant human-lymphoblasts invitro and toward murine 11210 leukemia invivo. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **1980**, *77* (11), 6865-6869.
29. Bonate, P.; Arthaud, L.; Cantrell, W.; Stephenson, K.; Secrist, J.; Weitman, S., Discovery and development of clofarabine: a nucleoside analogue for treating cancer. *Nature Reviews Drug Discovery* **2006**, *5* (10), 855-U2.
30. Reichelova, V.; Liliemark, J.; Albertioni, F., Liquid-chromatographic study of acid stability of 2-chloro-2'-arabino-fluoro-2'-deoxyadenosine, 2-chloro-2'-deoxyadenosine and related analogs. *Journal of Pharmaceutical and Biomedical Analysis* **1995**, *13* (4-5), 711-714.
31. Galmarini, C. M.; Popowycz, F.; Joseph, B., Cytotoxic nucleoside analogues: Different strategies to improve their clinical efficacy. *Current Medicinal Chemistry* **2008**, *15* (11), 1072-1082.
32. Montgomery, J. A.; Shortnacyfowler, A. T.; Clayton, S. D.; Riordan, J. M.; Secrist, J. A., Synthesis and biologic activity of 2'-fluoro-2-halo derivatives of 9-beta-d-arabinofuranosyladenine. *Journal of Medicinal Chemistry* **1992**, *35* (2), 397-401.
33. Vorbruggen, H.; Lagoja, I. M.; Herdewijn, P., Synthesis of ribonucleosides by condensation using trimethylsilyl triflate. *Current protocols in nucleic acid chemistry / edited by Serge L. Beaucage ... [et al.]* **2007**, *Chapter 1*, Unit 1.13-Unit 1.13.

34. Zhang, Q.; Sun, J.; Zhu, Y.; Zhang, F.; Yu, B., An Efficient Approach to the Synthesis of Nucleosides: Gold(I)-Catalyzed N-Glycosylation of Pyrimidines and Purines with Glycosyl ortho-Alkynyl Benzoates. *Angewandte Chemie-International Edition* **2011**, *50* (21), 4933-4936.
35. Liu, P.; Sharon, A.; Chu, C. K., Fluorinated nucleosides: Synthesis and biological implication. *Journal of Fluorine Chemistry* **2008**, *129* (9), 743-766.
36. Montgomery, J. A.; Secrist III, J. A., 2-fluoro-9-(2-deoxy-2-fluoro- $\beta$ -D-arabinofuranosyl) adenine nucleosides. Google Patents: **1991**.
37. Bauta, W. E.; Schulmeier, B. E.; Burke, B.; Puente, J. F.; Cantrell, W. R.; Lovett, D.; Goebel, J.; Anderson, B.; Ionescu, D.; Guo, R. C., A new process for antineoplastic agent clofarabine. *Organic Process Research & Development* **2004**, *8* (6), 889-896.
38. Bauta, W. E.; Burke, B. D.; Schulmeier, B. E.; Cantrell Jr, W. R.; Lovett, D. P.; Puente, J., Process for preparing purine nucleosides. Google Patents: **2004**.
39. Bauta, W. E.; Schulmeier, B. E.; Cantrell, W. R.; Lovett, D.; Puente, J.; Burke, B. D.; Lovett, D. P. Preparation of purine nucleosides useful as antitumor and antiviral agents, comprising coupling adenine derivative with blocked arabinofuranosyl. WO2003011877-A2; US2003114663-A1; US6680382-B2; EP1412369-A2; AU2002330955-A1; JP2005504036-W; US2005033043-A1; EP1412369-B1; DE60226447-E; US7528247-B2; US2009286971-A1; JP4593917-B2; US7947824-B2; US2011282045-A1; WO2003011877-A3; US2013035306-A1.
40. Anderson, B. G.; Bauta, W. E.; Cantrell, W. R., Jr.; Engles, T.; Lovett, D. P., Isolation, Synthesis, and Characterization of Impurities and Degradants from the Clofarabine Process. *Organic Process Research & Development* **2008**, *12* (6), 1229-1237.
41. Anastas, P., Twenty years of green chemistry. *Chemical & Engineering News* **2011**, *89* (26), 62-63.
42. Manley, J. B.; Anastas, P. T.; Cue, B. W., Frontiers in Green Chemistry: meeting the grand challenges for sustainability in R&D and manufacturing. *Journal of Cleaner Production* **2008**, *16* (6), 743-750.
43. Wohlgemuth, R., Biocatalysis - key to sustainable industrial chemistry. *Current Opinion in Biotechnology* **2010**, *21* (6), 713-724.
44. Schoemaker, H. E.; Mink, D.; Wubbolts, M. G., Dispelling the myths - Biocatalysis in industrial synthesis. *Science* **2003**, *299* (5613), 1694-1697.

45. Tufvesson, P.; Lima-Ramos, J.; Nordblad, M.; Woodley, J. M., Guidelines and Cost Analysis for Catalyst Production in Biocatalytic Processes. *Organic Process Research & Development* **2011**, *15* (1), 266-274.
46. Tufvesson, P.; Fu, W. J.; Jensen, J. S.; Woodley, J. M., Process considerations for the scale-up and implementation of biocatalysis. *Food and Bioproducts Processing* **2010**, *88* (C1), 3-11.
47. Wenda, S.; Illner, S.; Mell, A.; Kragl, U., Industrial biotechnology-the future of green chemistry? *Green Chemistry* **2011**, *13* (11), 3007-3047.
48. Araki, T.; Ikeda, I.; Matoishi, K.; Abe, R.; Oikawa, T.; Matsuba, Y.; Ishibashi, H.; Nagahara, K.; Fukui, Y. Method for producing cytosine nucleoside compounds. **2003**.
49. Hanrahan, J. R.; Hutchinson, D. W., The enzymatic-synthesis of antiviral agents. *Journal of Biotechnology* **1992**, *23* (2), 193-210.
50. Lewkowicz, E. S.; Iribarren, A. M., Nucleoside Phosphorylases. *Current Organic Chemistry* **2006**, *10* (11), 1197-1215.
51. Fernandez-Lucas, J., Multienzymatic synthesis of nucleic acid derivatives: a general perspective. *Applied Microbiology and Biotechnology* **2015**, *99* (11), 4615-4627.
52. Pugmire, M. J.; Ealick, S. E., Structural analyses reveal two distinct families of nucleoside phosphorylases. *Biochemical Journal* **2002**, *361*, 1-25.
53. Tuttle, J. V.; Tisdale, M.; Krenitsky, T. A., Purine 2'-deoxy-2'-fluororibosides as antiinfluenza virus agents. *Journal of Medicinal Chemistry* **1993**, *36* (1), 119-125.
54. Mikhailopulo, I. A.; Miroshnikov, A. I., New Trends in Nucleoside Biotechnology. *Acta Naturae* **2010**, *2* (2), 36-58.
55. Fresco-Taboada, A.; de la Mata, I.; Arroyo, M.; Fernandez-Lucas, J., New insights on nucleoside 2'-deoxyribosyltransferases: a versatile Biocatalyst for one-pot one-step synthesis of nucleoside analogs. *Applied Microbiology and Biotechnology* **2013**, *97* (9), 3773-3785.
56. Kalckar, H. M.; Macnutt, W. S.; Hoffjorgensen, E., Trans-n-glycosidase studied with radioactive adenine. *Biochemical Journal* **1952**, *50* (3), 397-400.
57. Beck, W.; Levin, M., Purification, kinetics, and repression control of bacterial trans-n-deoxyribosylase. *Journal of Biological Chemistry* **1963**, *238* (2), 702.

58. Carson, D. A.; Wasson, D. B., Synthesis of 2',3'-dideoxynucleosides by enzymatic trans-glycosylation. *Biochemical and Biophysical Research Communications* **1988**, *155* (2), 829-834.
59. Kaminski, P. A., Functional cloning, heterologous expression, and purification of two different N-deoxyribosyltransferases from *Lactobacillus helveticus*. *Journal of Biological Chemistry* **2002**, *277* (17), 14400-14407.
60. Kaminski, P. A.; Dugué, L.; Pochet, S., Expanding substrate specificity of nucleoside 2'-deoxyribosyltransferase. *Nucleic Acids Symp Ser (Oxf)* **2008**, (52), 495-6.
61. Fernández-Lucas, J.; Acebal, C.; Sinisterra, J. V.; Arroyo, M.; de la Mata, I., *Lactobacillus reuteri* 2'-deoxyribosyltransferase, a novel biocatalyst for tailoring of nucleosides. *Appl Environ Microbiol* **2010**, *76* (5), 1462-70.
62. Britos, C. N.; Jose Lapponi, M.; Cappa, V. A.; Rivero, C. W.; Trelles, J. A., Biotransformation of halogenated nucleosides by immobilized *Lactobacillus animalis* 2'-N-deoxyribosyltransferase. *Journal of Fluorine Chemistry* **2016**, *186*, 91-96.
63. Miyamoto, Y.; Masaki, T.; Chohnan, S., Characterization of N-deoxyribosyltransferase from *Lactococcus lactis* subsp *lactis*. *Biochimica Et Biophysica Acta-Proteins and Proteomics* **2007**, *1774* (10), 1323-1330.
64. Fresco-Taboada, A.; Serra, I.; Fernandez-Lucas, J.; Acebal, C.; Arroyo, M.; Terreni, M.; de la Mata, I., Nucleoside 2'-Deoxyribosyltransferase from Psychrophilic Bacterium *Bacillus psychrosaccharolyticus* - Preparation of an Immobilized Biocatalyst for the Enzymatic Synthesis of Therapeutic Nucleosides. *Molecules* **2014**, *19* (8), 11231-11249.
65. Steenkamp, D. J.; Halbich, T. J. F., Substrate-specificity of the purine-2'-deoxyribonucleosidase of crithidia-luciliae. *Biochemical Journal* **1992**, *287*, 125-129.
66. Lawrence, K. A.; Jewett, M. W.; Rosa, P. A.; Gherardini, F. C., *Borrelia burgdorferi* bb0426 encodes a 2'-deoxyribosyltransferase that plays a central role in purine salvage. *Molecular Microbiology* **2009**, *72* (6), 1517-1529.
67. Huang, M. C.; Montgomery, J. A.; Thorpe, M. C.; Stewart, E. L.; Secrist, J. A.; Blakley, R. L., Formation of 3-(2'-deoxyribofuranosyl) and 9-(2'-deoxyribofuranosyl) nucleosides of 8-substituted purines by nucleoside deoxyribosyltransferase. *Archives of Biochemistry and Biophysics* **1983**, *222* (1), 133-144.
68. Holguin, J.; Cardinaud, R., Trans-n-deoxyribosylase - purification by affinity chromatography and characterization. *European Journal of Biochemistry* **1975**, *54* (2), 505-514.



69. Porter, D. J. T.; Merrill, B. M.; Short, S. A., Identification of the active-site nucleophile in nucleoside 2-deoxyribosyltransferase as glutamic-acid-98. *Journal of Biological Chemistry* **1995**, *270* (26), 15551-15556.
70. Danzin, C.; Cardinaud, R., Deoxyribosyl transfer catalysis with trans-n-deoxyribosylase - kinetic study of purine (pyrimidine) to pyrimidine (purine) trans-n-deoxyribosylase. *European Journal of Biochemistry* **1976**, *62* (2), 365-372.
71. Danzin, C.; Cardinaud, R., Deoxyribosyl transfer catalysis with trans-n-deoxyribosylase - kinetic studies of purine to purine trans-n-deoxyribosylase. *European Journal of Biochemistry* **1974**, *48* (1), 255-262.
72. Short, S. A.; Armstrong, S. R.; Ealick, S. E.; Porter, D. J. T., Active site amino acids that participate in the catalytic mechanism of nucleoside 2'-deoxyribosyltransferase. *Journal of Biological Chemistry* **1996**, *271* (9), 4978-4987.
73. Fresco-Taboada, A.; de la Mata, I.; Arroyo, M.; Fernández-Lucas, J., New insights on nucleoside 2'-deoxyribosyltransferases: a versatile biocatalyst for one-pot one-step synthesis of nucleoside analogs. *Appl Microbiol Biotechnol* **2013**, *97* (9), 3773-85.
74. Pankiewicz, K. W., Fluorinated nucleosides. *Carbohydrate Research* **2000**, *327* (1-2), 87-105.
75. Smee, D. F.; Sidwell, R. W., Anti-cowpox virus activities of certain adenosine analogs, arabinofuranosyl nucleosides, and 2'-fluoro-arabinofuranosyl nucleosides. *Nucleosides Nucleotides & Nucleic Acids* **2004**, *23* (1-2), 375-383.
76. Ghanem, H.; Jabbour, E.; Faderl, S.; Ghandhi, V.; Plunkett, W.; Kantarjian, H., Clofarabine in leukemia. *Expert Review of Hematology* **2010**, *3* (1), 15-22.
77. Tennila, T.; Azhayeva, E.; Vepsalainen, J.; Laatikainen, R.; Azhayev, A.; Mikhailopulo, I. A., Oligonucleotides containing 9-(2-deoxy-2-fluoro-beta-D-arabinofuranosyl)-adenine and -guanine: Synthesis, hybridization and antisense properties. *Nucleosides Nucleotides & Nucleic Acids* **2000**, *19* (10-12), 1861-1884.
78. Tuttle, J. V.; Krenitsky, T. A. Therapeutic nucleosides. **1992**.
79. Fateev, I. V.; Antonov, K. V.; Konstantinova, I. D.; Muravyova, T. I.; Seela, F.; Esipov, R. S.; Miroshnikov, A. I.; Mikhailopulo, I. A., The chemoenzymatic synthesis of clofarabine and related 2'-deoxyfluoroarabinosyl nucleosides: the electronic and stereochemical factors determining substrate recognition by E. coli nucleoside phosphorylases. *Beilstein J Org Chem* **2014**, *10*, 1657-69.

80. Fateev, I. V.; Antonov, K. V.; Konstantinova, I. D.; Kharitonova, M. I.; Seela, F.; Esipov, R. S.; Miroshnikov, A. I.; Mikhailopulo, I. A., The chemo-enzymatic synthesis of clofarabine and related nucleosides: the role of electronic and stereochemical factors of substrates in reactions catalyzed by E. coli nucleoside phosphorylases. *Febs Journal* **2013**, *280*, 363-363.
81. Fernandez-Lucas, J.; Acebal, C.; Sinisterra, J. V.; Arroyo, M.; de la Mata, I., Lactobacillus reuteri 2'-Deoxyribosyltransferase, a Novel Biocatalyst for Tailoring of Nucleosides. *Applied and Environmental Microbiology* **2010**, *76* (5), 1462-1470.
82. Anand, R.; Kaminski, P. A.; Ealick, S. E., Structures of purine 2'-deoxyribosyltransferase, substrate complexes, and the ribosylated enzyme intermediate at 2.0 Å resolution. *Biochemistry* **2004**, *43* (9), 2384-2393.
83. Sheldon, R. A., Fundamentals of green chemistry: efficiency in reaction design. *Chemical Society Reviews* **2012**, *41* (4), 1437-1451.
84. Sheldon, R. A., E factors, green chemistry and catalysis: an odyssey. *Chemical Communications* **2008**, (29), 3352-3365.
85. Fateev, I. V.; Kharitonova, M. I.; Antonov, K. V.; Konstantinova, I. D.; Stepanenko, V. N.; Esipov, R. S.; Seela, F.; Temburnikar, K. W.; Seley-Radtke, K. L.; Stepchenko, V. A.; Sokolov, Y. A.; Miroshnikov, A. I.; Mikhailopulo, I. A., Recognition of Artificial Nucleobases by E-coli Purine Nucleoside Phosphorylase versus its Ser90Ala Mutant in the Synthesis of Base-Modified Nucleosides. *Chemistry-a European Journal* **2015**, *21* (38), 13401-13419.
86. Fateev, I. V.; Antonov, K. V.; Konstantinova, I. D.; Muravyova, T. I.; Seela, F.; Esipov, R. S.; Miroshnikov, A. I.; Mikhailopulo, I. A., The chemoenzymatic synthesis of clofarabine and related 2'-deoxyfluoroarabinosyl nucleosides: the electronic and stereochemical factors determining substrate recognition by E-coli nucleoside phosphorylases. *Beilstein Journal of Organic Chemistry* **2014**, *10*, 1657-1669.
87. Krenitsky, T. A.; Koszalka, G. W.; Tuttle, J. V.; Rideout, J. L.; Elion, G. B., AN Enzymic-synthesis of purine d-arabinonucleosides. *Carbohydrate Research* **1981**, *97* (1), 139-146.
88. Yamada, K.; Matsumoto, N.; Hayakawa, H., Stereoselective synthesis of 2-deoxy-2-fluoroarabinofuranosyl- $\alpha$ -1-phosphate and its application to the synthesis of 2'-deoxy-2'-fluoroarabinofuranosyl purine nucleosides by a chemo-enzymatic method. *Nucleosides Nucleotides & Nucleic Acids* **2009**, *28* (11-12), 1117-1130.

89. Krzeminski, J.; Nawrot, B.; Pankiewicz, K. W.; Watanabe, K. A., Nucleosides .157. Synthesis of 9-(2-deoxy-2-fluoro-beta-d-arabinofuranosyl)hypoxanthine - the 1st direct introduction of a 2'-beta-fluoro substituent in preformed purine nucleosides - studies directed toward the synthesis of 2'-deoxy-2'-substituted arabinonucleosides .8. *Nucleosides & Nucleotides* **1991**, *10* (4), 781-798.
90. Berman, H. M.; Kleywegt, G. J.; Nakamura, H.; Markley, J. L., The Protein Data Bank archive as an open data resource. *Journal of Computer-Aided Molecular Design* **2014**, *28* (10), 1009-1014.
91. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **2004**, *25* (13), 1605-1612.
92. Meng, X. Y.; Zhang, H. X.; Mezei, M.; Cui, M., Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Current Computer-Aided Drug Design* **2011**, *7* (2), 146-157.
93. Gupta, V.; Sengupta, M.; Prakash, J.; Charan Tripathy, B., *Basic and Applied Aspects of Biotechnology*. Springer Singapore: **2017**.
94. Atlas, R. M., *Handbook of Microbiological Media*. Fourth Edition ed.; **2010**.
95. Zhang, Y. K.; Li, T. M.; Liu, J. J., Low temperature and glucose enhanced T7 RNA polymerase-based plasmid stability for increasing expression of glucagon-like peptide-2 in Escherichia coli. *Protein Expression and Purification* **2003**, *29* (1), 132-139.
96. Chong, L., Molecular cloning - A laboratory manual, 3rd edition. *Science* **2001**, *292* (5516), 446-446.
97. Bradford, M. M., Rapid and sensitive method for quantitation of microgram quantities of protein utilizing principle of protein-dye binding. *Analytical Biochemistry* **1976**, *72* (1-2), 248-254.
98. Cardinaud, R.; Holguin, J., Nucleoside deoxyribosyltransferase-ii from lactobacillus-helveticus - substrate-specificity studies - pyrimidine-bases as acceptors. *Biochimica Et Biophysica Acta* **1979**, *568* (2), 339-347.
99. Huang, M. C.; Hatfield, K.; Roetker, A. W.; Montgomery, J. A.; Blakley, R. L., Analogs of 2'-deoxyadenosine - facile enzymatic preparation and growth inhibitory effects on human cell-lines. *Biochemical Pharmacology* **1981**, *30* (19), 2663-2671.

100. Kaminski, P. A.; Dugue, L.; Pochet, S., Expanding substrate specificity of nucleoside 2'-deoxyribosyltransferase. *Nucleic acids symposium series (2004)* **2008**, (52), 495-6.
101. Slater, M. J.; Gowrie, C.; Freeman, G. A.; Short, S. A., Enzymatic synthesis and antiviral activity of 2'-deoxy-2'-fluoro-ribavirin. *Bioorganic & Medicinal Chemistry Letters* **1996**, 6 (22), 2787-2790.
102. Porter, D. J. T.; Short, S. A., Nucleoside 2-deoxyribosyltransferase - pre-steady-state kinetic-analysis of native enzyme and mutant enzyme with an alanyl residue replacing glu-98. *Journal of Biological Chemistry* **1995**, 270 (26), 15557-15562.
103. Weiss, N.; Schillinger, U.; Kandler, O., *Lactobacillus-lactis*, *Lactobacillus-leichmannii* and *Lactobacillus-bulgaricus*, subjective synonyms of *Lactobacillus-delbrueckii*, and description of *Lactobacillus-delbrueckii* subsp *lactis* comb nov and *Lactobacillus-delbrueckii* subsp *bulgaricus* comb no. *Systematic and Applied Microbiology* **1983**, 4 (4), 552-557.
104. Armstrong, S. R.; Cook, W. J.; Short, S. A.; Ealick, S. E., Crystal structures of nucleoside 2-deoxyribosyltransferase in native and ligand-bound forms reveal architecture of the active site. *Structure* **1996**, 4 (1), 97-107.
105. Babaeipour, V.; Shojaosadati, S. A.; Maghsoudi, N., Maximizing Production of Human Interferon-gamma in HCDC of Recombinant E. coli. *Iranian Journal of Pharmaceutical Research* **2013**, 12 (3), 563-572.
106. Kim, B. S.; Lee, S. C.; Lee, S. Y.; Chang, Y. K.; Chang, H. N., High cell density fed-batch cultivation of Escherichia coli using exponential feeding combined with pH-stat (vol 26, pg 147, 2004. *Bioprocess and Biosystems Engineering* **2004**, 26 (4), 283-283.
107. Hoffmann, F.; Weber, J.; Rinas, U., Metabolic adaptation of Escherichia coli during temperature-induced recombinant protein production: 1. Readjustment of metabolic enzyme synthesis. *Biotechnol Bioeng* **2002**, 80 (3), 313-9.
108. Zhou, X. R.; Szeker, K.; Jiao, L. Y.; Oestreich, M.; Mikhailopulo, I. A.; Neubauer, P., Synthesis of 2,6-Dihalogenated Purine Nucleosides by Thermostable Nucleoside Phosphorylases. *Advanced Synthesis & Catalysis* **2015**, 357 (6), 1237-1244.
109. Mikhailopulo, I.; Fateev, I.; Esipov, R.; Antonov, K.; Konstantinova, I.; Seela, F.; Stepchenko, V.; Miroshnikov, A., The role of the Ser90 residue of the catalytic site of E. coli PNP in the synthesis of purine and 8-aza-7-deazapurine nucleosides. *Febs Journal* **2013**, 280, 363-363.

110. Stepchenko, V. A.; Seela, F.; Esipov, R. S.; Miroshnikov, A. I.; Sokolov, Y. A.; Mikhailopulo, I. A., Enzymatic Synthesis of 2'-Deoxy-beta-D-ribonucleosides of 8-Azapurines and 8-Aza-7-deazapurines. *Synlett* **2012**, (10), 1541-1545.
111. Lange, J.-P., - Lignocellulose conversion: an introduction to chemistry, process and economics. **2007**, *1* (1), 48.
112. Straathof, A. J. J.; Panke, S.; Schmid, A., The production of fine chemicals by biotransformations. *Current Opinion in Biotechnology* **2002**, *13* (6), 548-556.
113. Lima-Ramos, J.; Tufvesson, P.; Woodley, J. M., Application of environmental and economic metrics to guide the development of biocatalytic processes. *Green Processing and Synthesis* **2014**, *3* (3), 195-213.
114. Flickinger, M. C.; Drew, S. W., *Encyclopedia of Bioprocess Technology*. Wiley Online Library: **1999**.
115. Sharma, S. S.; Chong, S. R.; Harcum, S. W., Simulation of large-scale production of a soluble recombinant protein expressed in Escherichia coli using an intein-mediated purification system. *Applied Biochemistry and Biotechnology* **2005**, *126* (2), 93-117.
116. Heinzle, E.; Biver, A. P.; Cooney, C. L., *Development of sustainable bioprocesses : modeling and assessment*. John Wiley & Sons: Chichester, **2007**; p XXI, 294 p.
117. Peters, M.; Timmerhaus, K.; West, R., *Plant Design and Economics for Chemical Engineers*. McGraw-Hill Education: **2003**.
118. Tufvesson, P.; Ekman, A.; Sardari, R. R. R.; Engdahl, K.; Tufvesson, L., Economic and environmental assessment of propionic acid production by fermentation using different renewable raw materials. *Bioresource Technology* **2013**, *149*, 556-564.
119. Tufvesson, L. M.; Tufvesson, P.; Woodley, J. M.; Borjesson, P., Life cycle assessment in green chemistry: overview of key parameters and methodological concerns. *International Journal of Life Cycle Assessment* **2013**, *18* (2), 431-444.
120. Butters, M.; Catterick, D.; Craig, A.; Curzons, A.; Dale, D.; Gillmore, A.; Green, S. P.; Marziano, I.; Sherlock, J. P.; White, W., Critical assessment of pharmaceutical processes--A rationale for changing the synthetic route. *Chem Rev* **2006**, *106* (7), 3002-27.
121. Ubiali, D.; Rocchietti, S.; Scaramozzino, F.; Terreni, M.; Albertini, A. M.; Fernández-Lafuente, R.; Guisán, J. M.; Pregnotato, M., Synthesis of 2'-Deoxynucleosides by Transglycosylation with New Immobilized and Stabilized Uridine

Phosphorylase and Purine Nucleoside Phosphorylase. *Advanced Synthesis & Catalysis* **2004**, *346* (11), 1361-1366.

122. Zhou, X.; Szeker, K.; Janocha, B.; Böhme, T.; Albrecht, D.; Mikhailopulo, I. A.; Neubauer, P., Recombinant purine nucleoside phosphorylases from thermophiles: preparation, properties and activity towards purine and pyrimidine nucleosides. *FEBS J* **2013**, *280* (6), 1475-90.

123. Lundstedt, T.; Seifert, E.; Abramo, L.; Thelin, B.; Nystrom, A.; Pettersen, J.; Bergman, R., Experimental design and optimization. *Chemometrics and Intelligent Laboratory Systems* **1998**, *42* (1-2), 3-40.

124. Visser, D. F.; Rashamuse, K. J.; Hennessy, F.; Gordon, G. E. R.; Van Zyl, P. J.; Mathiba, K.; Bode, M. L.; Brady, D., High-yielding cascade enzymatic synthesis of 5-methyluridine using a novel combination of nucleoside phosphorylases. *Biocatalysis and Biotransformation* **2010**, *28* (4), 245-253.

125. Bauta, W. E.; Burke, B. D.; Schulmeier, B. E.; Cantrell, W. R.; Lovett, D. P.; Puente, J., Process for preparing purine nucleosides. Google Patents: **2011**.

126. Zabudkin, A.; Matvienko, V.; Matviienko, i.; Sypchenko, I. Method for the synthesis of clofarabine. **2016**.

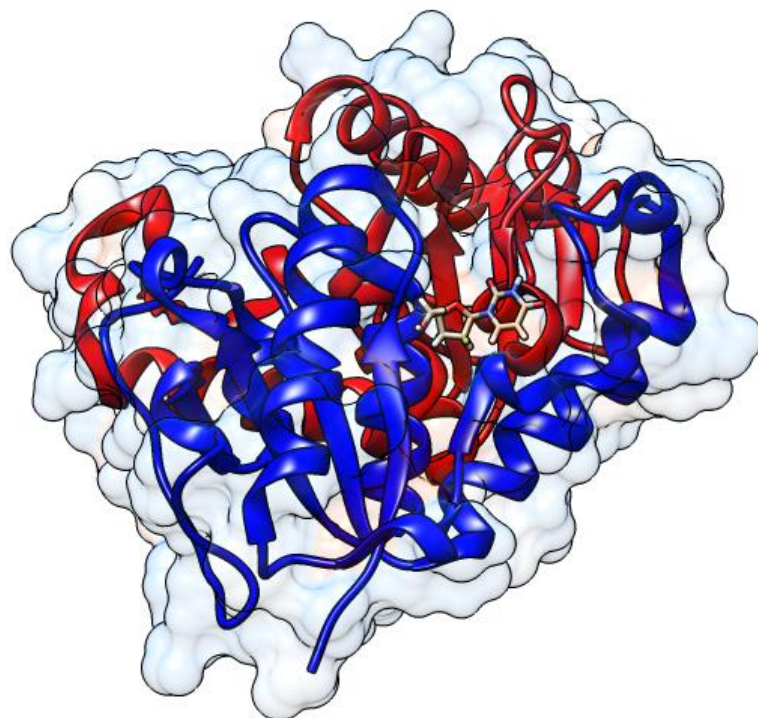
127. Fernandez-Lucas, J.; Fresco-Taboada, A.; Acebal, C.; de la Mata, I.; Arroyo, M., Enzymatic synthesis of nucleoside analogues using immobilized 2'-deoxyribosyltransferase from *Lactobacillus reuteri*. *Applied Microbiology and Biotechnology* **2011**, *91* (2), 317-327.

128. Fernández-Lucas, J.; Fresco-Taboada, A.; de la Mata, I.; Arroyo, M., One-step enzymatic synthesis of nucleosides from low water-soluble purine bases in non-conventional media. *Bioresour Technol* **2012**, *115*, 63-9.

129. Fresco-Taboada, A.; Serra, I.; Fernández-Lucas, J.; Acebal, C.; Arroyo, M.; Terreni, M.; de la Mata, I., Nucleoside 2'-deoxyribosyltransferase from psychrophilic bacterium *Bacillus psychrosaccharolyticus*--preparation of an immobilized biocatalyst for the enzymatic synthesis of therapeutic nucleosides. *Molecules* **2014**, *19* (8), 11231-49.

130. Fresco-Taboada, A.; Serra, I.; Arroyo, M.; Fernandez-Lucas, J.; de la Mata, I.; Terreni, M., Development of an immobilized biocatalyst based on *Bacillus psychrosaccharolyticus* NDT for the preparative synthesis of trifluridine and decytabine. *Catalysis Today* **2016**, *259*, 197-204.

131. Rosano, G. L.; Ceccarelli, E. A., Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers in Microbiology* **2014**, *5*.
132. Pollard, D. J.; Woodley, J. M., Biocatalysis for pharmaceutical intermediates: the future is now. *Trends in Biotechnology* **2007**, *25* (2), 66-73.
133. Van Aken, K.; Streckowski, L.; Patiny, L., EcoScale, a semi-quantitative tool to select an organic preparation based on economical and ecological parameters. *Beilstein Journal of Organic Chemistry* **2006**, *2*.
134. Andraos, J., Critical Evaluation of Published Algorithms for Determining Material Efficiency Green Metrics of Chemical Reactions and Synthesis Plans. *Acs Sustainable Chemistry & Engineering* **2016**, *4* (4), 1917-1933.
135. Andraos, J.; Sayed, M., On the use of "green" metrics in the undergraduate organic chemistry lecture and lab to assess the mass efficiency of organic reactions. *Journal of Chemical Education* **2007**, *84* (6), 1004-1010.
136. Garcia Calvo-Flores, F. In *Parámetros para el análisis de las reacciones en Química Sostenible*, Anales de la Real Sociedad Española de Química, **2009**; pp 42-49.
137. Doxsee, K.; Hutchison, J., Green Organic Chemistry: Strategies. *Tools, and Laboratory Experiments* **2004**, 155-158.
138. Andraos, J., Unification of reaction metrics for green chemistry: applications to reaction analysis. *Organic process research & development* **2005**, *9* (2), 149-163.
139. Andraos, J.; Hent, A., Simplified Application of Material Efficiency Green Metrics to Synthesis Plans: Pedagogical Case Studies Selected from Organic Syntheses. *Journal of Chemical Education* **2015**, *92* (11), 1820-1830.
140. Constable, D. J.; Curzons, A. D.; Cunningham, V. L., Metrics to 'green' chemistry—which are the best? *Green Chemistry* **2002**, *4* (6), 521-527.
141. Curzons, A. D.; Constable, D. J.; Mortimer, D. N.; Cunningham, V. L., So you think your process is green, how do you know?—Using principles of sustainability to determine what is green—a corporate perspective. *Green Chemistry* **2001**, *3* (1), 1-6.
142. Sheldon, R. A., The E factor: fifteen years on. *Green Chemistry* **2007**, *9* (12), 1273-1283.



**ANEXOS.**



## ANEXOS

### 6.1. Anexo1: Patent Application US 2016/0076070 A1

Biocatalytic Production Of Nucleoside Analogues As Active Pharmaceutical Ingredients

#### Applicants

- Plasmia Biotech S L

#### Inventors

- Pascual Gilabert Marta
- **Deroncelé Thomas Victor Manuel**
- Montilla Arévalo Rafael



US 20160076070A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2016/0076070 A1**

**PASCUAL GILABERT et al.**

(43) **Pub. Date: Mar. 17, 2016**

(54) **BIOCATALYTIC PRODUCTION OF NUCLEOSIDE ANALOGUES AS ACTIVE PHARMACEUTICAL INGREDIENTS**

**Publication Classification**

(51) **Int. Cl.**  
*C12P 19/38* (2006.01)  
*C12P 19/40* (2006.01)  
(52) **U.S. Cl.**  
CPC ..... *C12P 19/385* (2013.01); *C12P 19/40* (2013.01)

(71) Applicant: **PLASMIA BIOTECH, S.L.**, Barcelona (ES)

(72) Inventors: **Marta PASCUAL GILABERT**, Barcelona (ES); **Victor Manuel DERONCELÉ THOMAS**, Barcelona (ES); **Rafael MONTILLA ARÉVALO**, Barcelona (ES)

(57) **ABSTRACT**  
A biocatalytic process for producing active pharmaceutical ingredients (APIs) or intermediates thereof, wherein those APIs or their intermediates are nucleoside analogues (NAs) of formula I

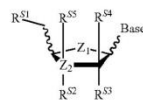
(21) Appl. No.: **14/787,678**

(22) PCT Filed: **Apr. 29, 2014**

(86) PCT No.: **PCT/EP2014/058761**

§ 371 (c)(1),  
(2) Date: **Oct. 28, 2015**

Formula I



**Related U.S. Application Data**

(60) Provisional application No. 61/817,064, filed on Apr. 29, 2013.

and wherein said NAs are active as pharmaceutically relevant antivirals and anticancer medicaments, intermediates or pro-drugs thereof.

**BIOCATALYTIC PRODUCTION OF  
NUCLEOSIDE ANALOGUES AS ACTIVE  
PHARMACEUTICAL INGREDIENTS**

FIELD OF THE INVENTION

**[0001]** The present invention relates to novel enzymatic processes useful for the production of Nucleoside Analogues (NAs) active as pharmaceutically relevant antivirals and anticancer medicaments, intermediates or prodrugs thereof.

BACKGROUND OF THE INVENTION

**[0002]** Nucleoside analogues (NAs) are synthetic compounds structurally related to natural nucleoside. In terms of their structure, nucleosides are constituted by three key elements: (i) the hydroxymethyl group, (ii) the heterocyclic nitrogenous base moiety, and (iii) the furanose ring, which in several instances seems to act as a spacer presenting the hydroxymethyl group and the base in the correct orientation.

**[0003]** NAs are extensively used as antiviral and antitumor agents. These molecules have been traditionally synthesized by different chemical methods which often require time-consuming multistep processes including protection-deprotection reactions on the heterocyclic base and/or the pentose moiety to allow the modification of naturally occurring nucleosides (Boryski J. 2008. Reactions of transglycosylation in the nucleoside chemistry. *Curr Org Chem* 12:309-325). This time consuming multistep processes often lead to low yields and increased costs. Indeed, chemical methods usually increase the difficulty of obtaining products with correct stereo- and regioselectivity, generating secondary products (Condezo, L. A., et al. 2007. Enzymatic synthesis of modified nucleosides, p. 401-423. *Biocatalysis in the pharmaceutical and biotechnology industries*. CRC Press, Boca Raton, Fla., Mikhailopulo, I. A. 2007). Moreover, the chemical methods include the use of chemical reagents and organic solvents that are expensive and environmentally harmful.

**[0004]** Since several non-natural nucleosides acting as antiviral or anticancer agents have modifications on their sugar moiety, it is interesting to explore the possibility of developing a novel and effective industrial biocatalyst to catalyze the enzymatic synthesis of nucleosides analogues i.e. to develop a synthesis of active pharmaceutical ingredients (APIs) or their intermediates which can be applied on an industrial scale.

**[0005]** Surprisingly, it was found that the drawbacks of previous cited chemical synthesis routes can be avoided and NAs can be obtained with a conversion higher than 50% and/or an anomeric purity higher than 95%. That is possible based on the use of Nucleoside Desoxyriboseyl Transferase (NDT or NdT) enzymes that we claim in the present invention.

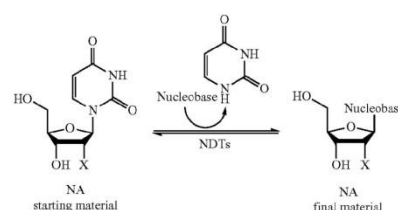
**[0006]** The advantages of the NDT bioenzymatic synthesis are:

- [0007]** (i) One-pot synthesis,
- [0008]** (ii) Reduced number of steps,
- [0009]** (iii) Higher conversions and yields,
- [0010]** (iv) Avoidance of organic solvents in the enzymatic step,
- [0011]** (v) No protection/deprotection strategies, e.g. for the hydroxyl groups in the sugar, are needed,
- [0012]** (vi) Mild reaction conditions: environmentally-friendly technology (water or aqueous medium, neutral pH).

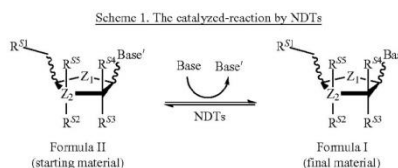
- [0013]** (vii) Extremely good selectivity: stereoselectivity—enantioselectivity, chemo-regioselectivity,
- [0014]** (viii) Fewer or no side reactions: impurity profile (reduced by-products content),
- [0015]** (ix) Reduction in overall waste generation,
- [0016]** (x) Process productivity,
- [0017]** (xi) Overall lower cost of production

DESCRIPTION OF THE INVENTION

**[0018]** The process described herein, using NDT enzymes is outlined as follows:



**[0019]** more precisely, as follows (where  $Z_1$ ,  $Z_2$ ,  $R^{S1}$ ,  $R^{S2}$ ,  $R^{S3}$ ,  $R^{S4}$ ,  $R^{S5}$  are as defined herein):



**[0020]** For the purposes of present description, the following terms are further defined as follows.

**[0021]** The term “nucleoside” refers to all compounds in which a heterocyclic base is covalently coupled to a sugar, and especially preferred coupling of the nucleoside to the sugar includes a C1’-(glycosidic) bond of a carbon atom in a sugar to a carbon- or heteroatom (typically nitrogen) in the heterocyclic base. Therefore, in the present context the term “nucleoside” means the glycoside of a heterocyclic base.

**[0022]** As further used herein, the term “sugar” refers to all carbohydrates and derivatives thereof, wherein particularly contemplated derivatives include deletion, substitution or addition of a chemical group or atom in the sugar. For example, especially contemplated deletions include 2’-deoxy and/or 3’-deoxy sugars. Especially contemplated substitutions include replacement of the ring-oxygen with sulphur or methylene, or replacement of a hydroxyl group with a halogen, an amino-, sulphydryl-, or methyl group, and especially contemplated additions include methylene phosphonate groups. Further contemplated sugars also include sugar analogues (i.e., not naturally occurring sugars), and particularly carbocyclic ring systems. The term “carbocyclic ring system” as used herein refers to any molecule in which a plurality of

carbon atoms form a ring, and in especially contemplated carbocyclic ring systems the ring is formed from 3, 4, 5, or 6 carbon atoms.

**[0023]** The term “nucleoside” may be used broadly as to include non-naturally occurring nucleosides, naturally occurring nucleosides as well as other nucleoside analogues. Illustrative examples of nucleosides are ribonucleosides comprising a ribose moiety as well as deoxyribonucleosides comprising a deoxyribose moiety. With respect to the bases of such nucleosides, it should be understood that this may be any of the naturally occurring bases, e.g. adenine, guanine, cytosine, thymine, and uracil, as well as any modified variants thereof or any possible unnatural bases.

**[0024]** The term “nucleoside analogue”, “NA” or “NAs” as used herein refers to all nucleosides, isomers or enantiomers thereof as well as racemates or enantio-enriched mixtures thereof, which are oxy- or deoxy-analogues of the naturally-occurring DNA and RNA nucleosides deoxycytidine, deoxyuridine, deoxyadenosine, deoxyguanosine and thymidine. Preferably, said NAs are nucleosides which comprise a sugar moiety and/or a base moiety which is different from the corresponding sugar moiety and/or base moiety of each of deoxycytidine, deoxyuridine, deoxyadenosine, deoxyguanosine or thymidine. More preferably, said NAs are non-naturally occurring nucleosides comprised of a sugar moiety and base moiety wherein at least one of said sugar moiety and said base moiety is not found in naturally-occurring DNA or RNA, more preferably in naturally-occurring polynucleotides. Alternatively, said NAs are nucleosides in which the sugar is not a ribofuranose and/or in which the heterocyclic base is not a naturally occurring base (e.g., A, G, C, T, I, etc.). Similarly, the term “nucleotide” refers to a nucleoside, either D-nucleoside or L-nucleoside, to which a phosphate group is coupled to the sugar.

**[0025]** The term “one-step/one-pot”, “one-step one-pot” or “one step-one pot” refers to a method of synthesis of chemical compounds through a single step in which the materials used are mixed together in a single vessel and allowed to react, rather than conducting the reaction in a sequence of separate stages. This strategy is used to improve reaction efficiency, increase yield and save time and resources.

**[0026]** The terms “heterocyclic ring” or “heterocyclic base” are used interchangeably herein and refer to any compound in which plurality of atoms form a ring via a plurality of covalent bonds, wherein the ring includes at least one atom other than a carbon atom. Particularly contemplated heterocyclic bases include 5- and 6-membered rings containing at least 1 to 4 heteroatoms each independently selected from nitrogen, oxygen and sulphur as the non-carbon atom (e.g., imidazole, pyrrole, triazole, dihydropyrimidine). Further contemplated heterocycles may be fused (i.e., covalently bound) to another ring or heterocycle, and are thus termed “fused heterocycle” or “fused heterocyclic base” as used herein. Especially contemplated fused heterocycles include a 5-membered ring fused to a 6-membered ring (e.g., purine, pyrrolo[2,3-d]pyrimidine), and a 6-membered ring fused to another 6-membered or higher ring (e.g., pyrido[4,5-d]pyrimidine, benzodiazepine). Examples of these and further preferred heterocyclic bases are given below. Still further contemplated heterocyclic bases may be aromatic, or may include one or more double or triple bonds. Moreover, contemplated heterocyclic bases and fused heterocycles may further be substituted in one or more positions. And any one of the rings being optionally substituted with one, two or three

substituents each independently selected from the group consisting of halogen, hydroxy, nitro, cyano, carboxyl, C1-6alkyl, C1-6alkoxy, C1-6alkoxyC1-6alkyl, C1-6alkyl-carbonyl, amino, mono- or diC1-6 alkyl amino, azido, mercapto, polyhaloC1-6alkyl, polyhaloC1-6alkoxy, and C3-7-cycloalkyl.

**[0027]** The terms “nucleobase” covers naturally occurring nucleobases as well as non-naturally occurring nucleobases. It should be clear to the person skilled in the art that various nucleobases which previously have been considered “non-naturally occurring” have subsequently found in nature. Thus, “nucleobase” includes not only the known purine and pyrimidine heterocycles, but also heterocyclic analogues (such as N-substituted heterocycles) and tautomers thereof. See formulas A to H. Illustrative examples of nucleobases are adenine, guanine, thymine, cytosine, uracil, purine, xanthine, 2-chloroadenine, 2-fluoroadenine, pentyl(5-fluoro-2-oxo-1,2, dihydropyrimidin-4-yl)carbamate, cytosine N-alkyl carbamates, cytosine N-alkylesters, 5-azacytosine, 5-bromovinyluracil, 5-fluorouracil, 5-trifluoromethyluracil, 6-methoxy-9H-purin-2-amine and (R)-3,6,7,8-tetrahydroimidazo[4,5-d][1,3]diazepin-8-ol. The term “nucleobase” is intended to cover every and all of these examples as well as analogues and tautomers, and regioisomers thereof.

**[0028]** The term “tautomer” or “tautomeric form” refers to structural isomer of different energies which are interconvertible via a low energy barrier. For example, proton tautomers (also known as prototropic tautomers) include interconversion via migration of a proton, such as keto-enol and imine-enamine isomerizations. Valence tautomers include interconversions by reorganization of some of the bonding electrons.

**[0029]** The term “regioisomer” refers to structural isomer, or constitutional isomer in the sense that refers to molecules with the same molecular formula that whose atoms are bonded in different order of connectivity.

**[0030]** The term “conversion” refers to is the percentage of starting material that is transformed into products, either the expected final product, byproducts, or even into products of degradation.

**[0031]** The term “anomeric purity” refers to the amount of a particular anomer of a compound divided by the total amount of all anomers of that compound present in the mixture multiplied by 100%.

**[0032]** The term “2'-fluoro-arabino nucleoside-type” refers to any nucleoside analogue in which the carbon atom at the 2' position of the sugar moiety is substituted by a fluorine atom in a 2'-arabino configuration, independently of the nucleobase linked to the before mentioned sugar.

**[0033]** The term “intermediate” or “intermediates” refer to any nucleoside analogue type compounds which may be transformed into an active pharmaceutical ingredient (API) of nucleosidic structure by means of suitable additional chemical reactions. Therefore, intermediates are molecules that may be considered as API precursors. Any nucleoside analogue type compounds which are not used as APIs or their intermediates are disclaimed in the present invention.

**[0034]** Moreover, naturally-occurring DNA and RNA nucleosides and NAs which might have been disclosed in the prior art prepared by biocatalytic processes but which are unable to be produced on an industrial-scale at g/L using the method of the present invention (i.e. are only produced on a microlab demonstration scale or even at milligram quantities at lab scale not useful for production purposes) are disclaimed

in the present invention since such NAs either cannot serve as APIs or be suitable for production on a commercial-scale, respectively.

**[0035]** Therefore summing up, by one hand, the chemical processes available in the prior art for industrial production of NAs are much more complex than the biocatalytic process of present invention. By the other hand, biocatalytic processes disclosed in the state of the art for the synthesis of nucleosides are limited to lab scale and to naturally occurring nucleosides, not to NAs, nor particularly at industrial scale, neither to APIs.

**[0036]** The terms “alkyl” and “unsubstituted alkyl” are used interchangeably herein and refer to any C<sub>1</sub>-C<sub>25</sub> linear, branched, or cyclic hydrocarbon in which all carbon-carbon bonds are single bonds.

**[0037]** The terms “alkenyl” and “unsubstituted alkenyl” are used interchangeably herein and refer to any C<sub>1</sub>-C<sub>25</sub> linear, branched, or cyclic alkyl with at least one carbon-carbon double bond.

**[0038]** Furthermore, the terms “alkynyl” and “unsubstituted alkynyl” are used interchangeably herein and refer to any C<sub>1</sub>-C<sub>25</sub> linear, branched, or cyclic alkyl or alkenyl with at least one carbon-carbon triple bond. The terms “aryl” and “unsubstituted aryl” are used interchangeably herein and refer to any aromatic cyclic alkenyl or alkynyl, being as a group or part of a group is phenyl or naphthalenyl, each optionally substituted with one, two or three substituents selected from halo, hydroxy, nitro, cyano, carboxyl, C<sub>1-6</sub>alkyl, C<sub>1-6</sub>alkoxy, C<sub>1-6</sub>alkoxyC<sub>1-6</sub>alkyl, C<sub>1-6</sub>alkylcarbonyl, amino, mono- or diC<sub>1-6</sub>alkylamino, azido, mercapto, polyhaloC<sub>1-6</sub>alkyl, polyhaloC<sub>1-6</sub>alkoxy, and polyhaloC<sub>1-6</sub>alkoxy. The term “alkaryl” is employed where an aryl is covalently bound to an alkyl, alkenyl, or alkynyl.

**[0039]** The term “substituted” as used herein refers to a replacement of an atom or chemical group (e.g., H, NH<sub>2</sub>, or OH) with a functional group, and particularly contemplated functional groups include nucleophilic groups (e.g., —NH<sub>2</sub>, —OH, —SH, —NC, etc.), electrophilic groups (e.g., C(O)OR, C(X)OH, etc.), polar groups (e.g., —OH), non-polar groups (e.g., aryl, alkyl, alkenyl, alkynyl, etc.), ionic groups (e.g., —NH<sub>3</sub><sup>+</sup>), and halogens (e.g., —F, —Cl), and all chemically reasonable combinations thereof. Thus, the term “functional group” and the term “substituent” are used interchangeably herein and refer to nucleophilic groups (e.g., —NH<sub>2</sub>, —OH, —SH, —NC, —CN, etc.), electrophilic groups (e.g., C(O)OR, C(X)OH, C(Halogen)OR, etc.), polar groups (e.g., —OH), non-polar groups (e.g., aryl, alkyl, alkenyl, alkynyl, etc.), ionic groups (e.g., —NH<sub>3</sub><sup>+</sup>), and halogens.

**[0040]** As also used herein, the terms “heterocycle” and “heterocyclic base” are used interchangeably herein and refer to any compound in which plurality of atoms form a ring via a plurality of covalent bonds, wherein the ring includes at least one atom other than a carbon atom. Particularly contemplated heterocyclic bases include 5- and 6-membered rings containing at least 1 to 4 heteroatoms each independently selected from nitrogen, oxygen and sulphur as the non-carbon atom (e.g., imidazole, pyrrole, triazole, dihydropyrimidine). Further contemplated heterocycles may be fused (i.e., covalently bound) to another ring or heterocycle, and are thus termed “fused heterocycle” or “fused heterocyclic base” as used herein. Especially contemplated fused heterocycles include a 5-membered ring fused to a 6-membered ring (e.g., purine, pyrrolo[2,3-d]pyrimidine), and a 6-membered ring

fused to another 6-membered or higher ring (e.g., pyrido[4,5-d]pyrimidine, benzodiazepine). Examples of these and further preferred heterocyclic bases are given below. Still further contemplated heterocyclic bases may be aromatic, or may include one or more double or triple bonds. Moreover, contemplated heterocyclic bases and fused heterocycles may further be substituted in one or more positions. And any one of the rings being optionally substituted with one, two or three substituents each independently selected from the group consisting of halogen, hydroxy, nitro, cyano, carboxyl, C<sub>1-6</sub>alkyl, C<sub>1-6</sub>alkoxy, C<sub>1-6</sub>alkoxyC<sub>1-6</sub>alkyl, C<sub>1-6</sub>alkylcarbonyl, amino, mono- or diC<sub>1-6</sub>alkylamino, azido, mercapto, polyhaloC<sub>1-6</sub>alkyl, polyhaloC<sub>1-6</sub>alkoxy, and C<sub>3-7</sub>cycloalkyl.

**[0041]** The nucleic acid molecule encoding an enzyme having N-deoxyribosyltransferase (NDT) activity according to present invention comprises:

**[0042]** a) a nucleotide sequence as shown in SEQ ID NO:1, SEQ ID NO:3 or SEQ ID NO:5; or

**[0043]** b) a nucleotide sequence which is the complement of SEQ ID NO:1, SEQ ID NO:3 or SEQ ID NO:5; or

**[0044]** c) a nucleotide sequence which is degenerate with SEQ ID NO:1, SEQ ID NO:3 or SEQ ID NO:5; or

**[0045]** d) a nucleotide sequence hybridizing under conditions of high stringency to SEQ ID NO:1, SEQ ID NO:3 or SEQ ID NO:5, to the complements of SEQ ID NO:1, SEQ ID NO:3 or SEQ ID NO:5, or to a hybridization probe derived from SEQ ID NO:1, SEQ ID NO:3 or SEQ ID NO:5, or their complement thereof; or

**[0046]** e) a nucleotide sequence having at least 80% sequence identity with SEQ ID NO:1, SEQ ID NO:3 or SEQ ID NO:5; or

**[0047]** f) a nucleotide sequence having at least 65% sequence identity with SEQ ID NO:1, SEQ ID NO:3 or SEQ ID NO:5, wherein said sequence preferably encodes or is complementary to a sequence encoding at least a NDT enzyme or a functional part thereof.

**[0048]** g) a nucleotide sequence encoding for an amino acid sequence selected from: SEQ ID NO:2, SEQ ID NO:4 or SEQ ID NO:6.

**[0049]** Conditions of stringency hybridization in the sense of the present invention are defined as those described by Sambrook et al., *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory Press (1989), 1.1011.104. According to this, hybridization under stringent conditions means that a positive hybridization signal is still observed after washing for one hour with 1×SSC buffer and 0.1% SDS at 55° C., preferably at 62° C. and most preferred at 68° C., in particular, for one hour in 0.2×SSC buffer and 0.1% SDS at 55° C., preferably at 62° C. and most preferred at 68° C.

**[0050]** Moreover, in the sense of present description, the invention also covers nucleotide or amino acid sequences which, on nucleotide or amino acid levels, respectively, have an identity of at least 70%, particularly preferred at least 80% and most preferred at least 90% to the nucleotide or amino acid sequence shown in SEQ ID NO. 1, 3 or 5 (nucleotide) or SEQ ID NO. 2, 4 or 6 (amino acid). Percent identity is determined according to the following equation:

$$I = (n/L) \times 100$$

wherein I are percent identity, L is the length of the basic sequence and n is the number of nucleotide or amino acid difference of a sequence to the basic sequence.

**[0051]** Still another subject matter of the present invention is a recombinant vector comprising at least one copy of the nucleic acid molecule as defined above, operatively linked with an expression control sequence. The vector may be any prokaryotic or eukaryotic vector. Examples of prokaryotic vectors are chromosomal vectors such as bacteriophages (e.g. bacteriophage Lambda) and extrachromosomal vectors such as plasmids (see, for example, Sambrook et al., supra, Chapter 1-4). The vector may also be a eukaryotic vector, e.g. a yeast vector or a vector suitable for higher cells, e.g. a plasmid vector, viral vector or plant vector. Suitable eukaryotic vectors are described, for example, by Sambrook et al., supra, Chapter 16. The invention moreover relates to a recombinant cell transformed with the nucleic acid or the recombinant vector as described above. The cell may be any cell, e.g. a prokaryotic or eukaryotic cell. Prokaryotic cells, in particular, *E. coli* cells, are especially preferred.

**[0052]** For the purpose of present description, the invention also covers variants, or precursors, or orthologues, or combinations of SEQ ID Nos: 1-6.

**[0053]** The term "variant" as used throughout the specification is to be understood to mean a nucleotide sequence of a nucleic acid or amino acid sequence of a protein or polypeptide that is altered by one or more nucleotides or amino acids, respectively. The variant may have "conservative" changes, wherein a substituted nucleotide or amino acid has similar structural or chemical properties to the replaced nucleotide or amino acid. A variant may also have "non-conservative" changes or a deletion and/or insertion of one or more nucleotides or amino acids. The term also includes within its scope any insertions/deletions of nucleotides or amino acids for a particular nucleic acid or protein or polypeptide. A "functional variant" will be understood to mean a variant that retains the functional capacity of a reference nucleotide sequence or a protein or polypeptide.

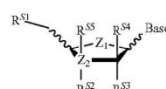
**[0054]** The term "complement" or "complementary" as used herein may mean that each strand of 20 double-stranded nucleic acids such as, DNA and RNA, is complementary to the other in that the base pairs between them are non-covalently connected via two or three hydrogen bonds. For DNA, adenine (A) bases complement thymine (T) bases and vice versa; guanine (G) bases complement cytosine (C) bases and vice versa. With RNA, it is the same except that adenine (A) bases complement uracil (U) bases instead of thymine (T) bases. Since there is only one 25 complementary base for each of the bases found in DNA and in RNA, one can reconstruct a complementary strand for any single strand.

**[0055]** The term "orthologue" as used throughout the specification is to be understood as homologous gene or miRNAs sequences found in different species.

**[0056]** For the purposes of present description the term "NDT" means a biocatalyst pertaining to the family of the enzymes having nucleoside deoxyribosyltransferase or N-deoxyribosyl transferase activity. Said NDT may be a nucleoside deoxyribosyltransferase enzyme or N-deoxyribosyl transferase enzyme or may be another enzyme which has nucleoside deoxyribosyltransferase activity under specific reaction conditions and/or when said enzyme is modified to give nucleoside deoxyribosyltransferase activity, such as a glycosidase that is modified to catalyze the condensation reaction rather than the hydrolysis reaction (i.e. to carry out the glycosidase reaction in reverse), thereby behaving as a nucleoside deoxyribosyltransferase.

**[0057]** Also for the purposes of present description the term "Active Pharmaceutical Ingredient" or "API" means a compound, basically a NA, which shows therapeutic activity in human beings or animals. All the APIs disclosed in the present description are referred to under either any isomeric form, or as a racemic or enantiomerically-enriched mixture of isomers.

**[0058]** Present description discloses a biocatalytic process for producing active pharmaceutical ingredients (APIs) or intermediates thereof, being those APIs or their intermediates, nucleoside analogues (NAs), D isomers of formula I



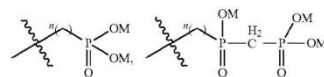
Formula I

**[0059]** wherein,

**[0060]**  $Z_1$  being O,  $CH_2$ , S, NH;

**[0061]**  $Z_2$  being O, C, S, N, independently of  $Z_1$ ;

**[0062]**  $R^{S1}$  being hydrogen, methyl, OH, ether or ester thereof selected from:



**[0063]** being n is 0 or 1 and M is hydrogen or a pharmaceutically acceptable counter-ion such as sodium, potassium, ammonium or alkylammonium;

**[0064]**  $R^{S2}$  being hydrogen, halogen, preferably F, OH or an ether or ester residue thereof, CN,  $NH_2$ , SH,  $C=CH$ ,  $N_3$ ;

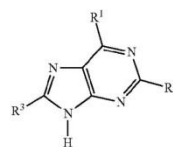
**[0065]**  $R^{S3}$  being hydrogen, in case of NA derived from deoxyribonucleosides or being selected from: OH,  $NH_2$ , halogen (preferably F),  $OCH_3$ , when the NA is derived from ribonucleosides;

**[0066]**  $R^{S4}$  being hydrogen, OH or an ether or ester residue thereof,  $NH_2$ , halogen, preferably F,

**[0067]** providing  $R^{S1}$  and  $R^{S4}$  are different when both were ethers or esters of OH residues;

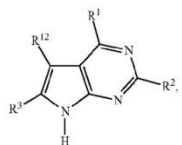
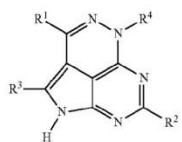
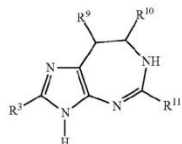
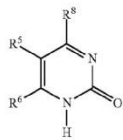
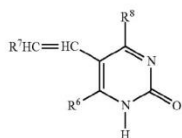
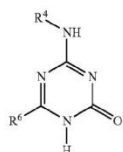
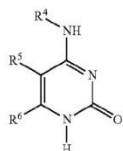
**[0068]**  $R^{S5}$  being hydrogen, OH or an ether or ester residue thereof,  $NH_2$  or halogen, preferably F;

**[0069]** Base bound to the ribose moiety being selected from those of formula A-H or their tautomers and regioisomers thereof:

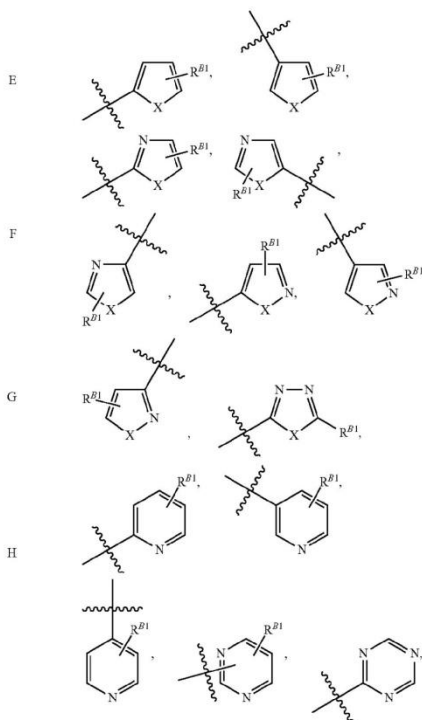


A

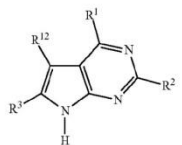
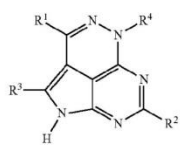
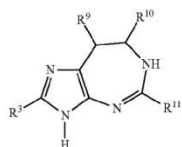
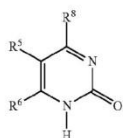
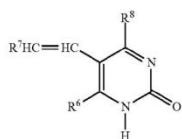
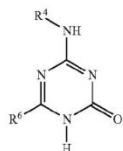
-continued



[0070] wherein,

B [0071] R<sup>1</sup> being hydrogen, methyl, optionally substituted alkyl chain, halogen, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, SR<sup>13</sup>;[0072] R<sup>2</sup> being hydrogen, methyl, optionally substituted alkyl chain, halogen, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, SR<sup>13</sup>;[0073] R<sup>3</sup> being hydrogen, methyl, optionally substituted alkyl chain;C [0074] R<sup>4</sup> being hydrogen, methyl, optionally substituted alkyl chain, COR<sup>13</sup>, CONR<sup>13</sup>R<sup>14</sup>, CO<sub>2</sub>R<sup>13</sup>, C(S)OR<sup>13</sup>, CH<sub>2</sub>-heterocyclic ring;[0075] R<sup>5</sup> being hydrogen, methyl, an optionally substituted alkyl chain, an optionally substituted alkenyl chain, an optionally substituted alkynyl chain, halogen, trihaloalkyl, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, CN, COR<sup>13</sup>, CONR<sup>13</sup>R<sup>14</sup>, CO<sub>2</sub>R<sup>13</sup>, C(S)OR<sup>13</sup>, OCONR<sup>13</sup>R<sup>14</sup>, OCO<sub>2</sub>R<sup>13</sup>, OC(S)OR<sup>13</sup>, NHCONR<sup>13</sup>R<sup>14</sup>, NHCO<sub>2</sub>R<sup>13</sup>, NHC(S)OR<sup>13</sup>, SO<sub>2</sub>NR<sup>13</sup>R<sup>14</sup>;D [0076] and any optionally substituted heterocycle or optionally substituted aryl of, independently, R<sup>1</sup>, R<sup>2</sup>, R<sup>3</sup>, R<sup>4</sup> or R<sup>5</sup>, selected from:

-continued



[0098] wherein.

[0099] R<sup>1</sup> being hydrogen, methyl, optionally substituted alkyl chain, halogen, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, SR<sup>13</sup>;[0100] R<sup>2</sup> being hydrogen, methyl, optionally substituted alkyl chain, halogen, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, SR<sup>13</sup>;[0101] R<sup>3</sup> being hydrogen, methyl, optionally substituted alkyl chain;[0102] R<sup>4</sup> being hydrogen, methyl, optionally substituted alkyl chain, COR<sup>13</sup>, CONR<sup>13</sup>R<sup>14</sup>, CO<sub>2</sub>R<sup>13</sup>, (S)OR<sup>13</sup>, CH<sub>2</sub>-heterocyclic ring;C' R<sup>5</sup> being hydrogen, methyl, an optionally substituted alkyl chain, an optionally substituted alkenyl chain, an optionally substituted alkynyl chain, halogen, trihaloalkyl, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, CN, COR<sup>13</sup>, CONR<sup>13</sup>R<sup>14</sup>, CO<sub>2</sub>R<sup>13</sup>, C(S)OR<sup>13</sup>, OCONR<sup>13</sup>R<sup>14</sup>, OCO<sub>2</sub>R<sup>13</sup>, OC(S)OR<sup>13</sup>, NHCONR<sup>13</sup>R<sup>14</sup>, NHCO<sub>2</sub>R<sup>13</sup>, NHC(S)OR<sup>13</sup>, SO<sub>2</sub>NR<sup>13</sup>R<sup>14</sup>;[0104] and any optionally substituted heterocycle or optionally substituted aryl of, independently, R<sup>1</sup>, R<sup>2</sup>, R<sup>3</sup>, R<sup>4</sup> or R<sup>5</sup>, selected from:

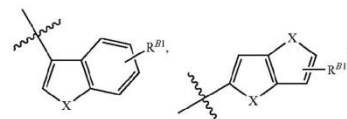
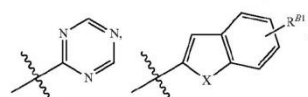
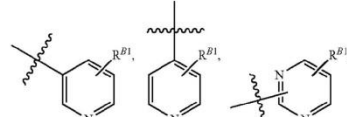
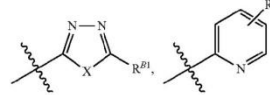
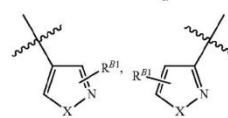
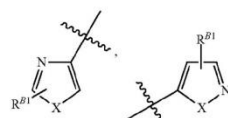
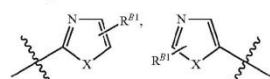
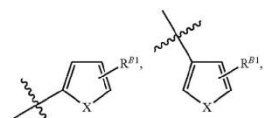
D'

E'

F'

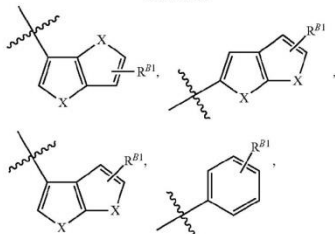
G'

H'





-continued



[0105] wherein

[0106] X being O, S, N—R<sup>B2</sup>, Se; R<sup>B1</sup> being H, straight or branched C<sub>1-10</sub> alkyl, F, Cl, Br, I, X—R<sup>B2</sup>, —C≡C—R<sup>B2</sup>, CO<sub>2</sub>R<sup>B2</sup>; R<sup>B2</sup> being H, straight or branched C<sub>1-5</sub> alkyl, phenyl;

[0107] R<sup>6</sup> being hydrogen, optionally substituted alkyl chain;

[0108] R<sup>7</sup> being hydrogen, halogen, trihaloalkyl, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, CN, COR<sup>13</sup>, CONR<sup>13</sup>R<sup>14</sup>, CO<sub>2</sub>R<sup>13</sup>, C(S)OR<sup>13</sup>;

[0109] R<sup>8</sup> being hydroxyl or amino, OR<sup>13</sup>, OSO<sub>2</sub>R<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, CN, COR<sup>13</sup>, CONR<sup>13</sup>R<sup>14</sup>, CO<sub>2</sub>R<sup>13</sup>, C(S)OR<sup>13</sup>;

[0110] R<sup>9</sup> being hydrogen, methyl, optionally substituted alkyl chain, halogen, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, COR<sup>13</sup>, CONR<sup>13</sup>R<sup>14</sup>, CN, CO<sub>2</sub>R<sup>13</sup>, C(S)OR<sup>13</sup>, OCONR<sup>13</sup>R<sup>14</sup>, OCO<sub>2</sub>R<sup>13</sup>, OC(S)OR<sup>13</sup>, NHCONR<sup>13</sup>R<sup>14</sup>, NHCO<sub>2</sub>R<sup>13</sup>, NHC(S)OR<sup>13</sup>;

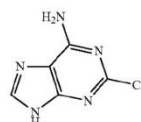
[0111] R<sup>10</sup> and R<sup>11</sup> are independently of each other hydrogen, methyl, optionally substituted alkyl chain;

[0112] R<sup>12</sup> being hydrogen, methyl, optionally substituted alkyl chain, halogen, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, CN, COR<sup>13</sup>, CONR<sup>13</sup>R<sup>14</sup>, CO<sub>2</sub>R<sup>13</sup>, C(S)OR<sup>13</sup>, OCONR<sup>13</sup>R<sup>14</sup>, OCO<sub>2</sub>R<sup>13</sup>, OC(S)OR<sup>13</sup>, NHCONR<sup>13</sup>R<sup>14</sup>, NHCO<sub>2</sub>R<sup>13</sup>, NHC(S)OR<sup>13</sup>;

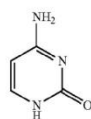
[0113] R<sup>13</sup> and R<sup>14</sup> are independently of each other hydrogen, substituted or unsubstituted alkyl chain, substituted or unsubstituted alkenyl chain, substituted or unsubstituted alkenyl chain, substituted or unsubstituted heterocyclic or phenyl rings.

[0114] The heterocyclic ring constituting the base in formula II of the starting materials may be selected from: uracil, adenine, cytosine, guanine, thymine, hypoxanthine, xanthine, thiouracil, thioguanine, 9-H-purine-2-amine, 7-methylguanine, 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5,6-dihydrouracil, 5-methylcytosine and 5-hydroxymethylcytosine, pteridone, and any substituted derivative thereof. Preferably the heterocyclic ring constituting the base' in formula II of the starting materials is selected from: uracil, adenine, cytosine, guanine, thymine, hypoxanthine, xanthine, thiouracil, thioguanine, 9-H-purine-2-amine, 7-methylguanine, 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5,6-dihydrouracil, 5-methylcytosine and 5-hydroxymethylcytosine and any substituted derivatives thereof.

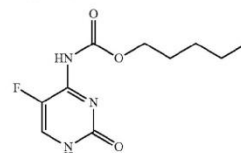
[0115] Moreover, the free nucleobase to be transferred by the NDT enzyme is, preferably, selected from:



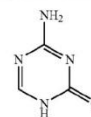
2-chloro-9H-purin-6-amine  
(2-chloroadenine)



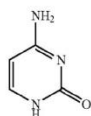
4-aminopyrimidin-2(1H)-one  
(cytosine)



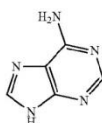
pentyl (5-fluoro-2-oxo-1,2-dihydropyrimidin-4-yl)carbamate



4-amino-1,3,5-triazin-2(1H)-one  
(5-azacytosine)

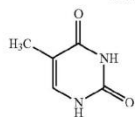
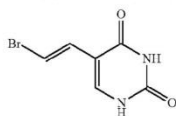
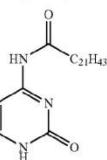


4-aminopyrimidin-2(1H)-one  
(cytosine)

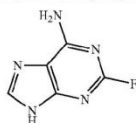
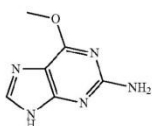


9H-purin-6-amine  
(adenine)

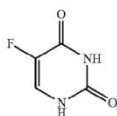
-continued

5-methylpyrimidine-2,4(1H,3H)-dione  
(thymine; 5-methyluracil)(E)-5-(2-bromovinyl)pyrimidine-2,4(1H,3H)-dione  
(E)-5-bromovinyluracil)

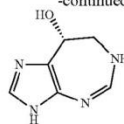
N-(2-oxo-1,2-dihydropyrimidin-4-yl)docosanamide

2-fluoro-9H-purin-6-amine  
(2-fluoroadenine)

6-methoxy-9H-purin-2-amine

5-fluoropyrimidine-2,4(1H,3H)-dione  
(5-fluorouracil)

-continued



(R)-3,6,7,8-tetrahydroimidazo[4,5-d][1,3]diazepin-8-ol

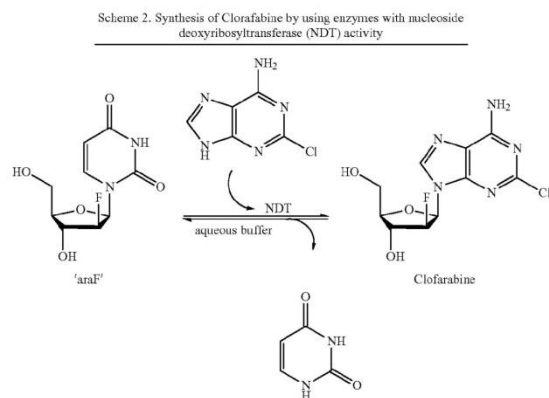
[0116] In a preferred embodiment of the aforementioned process,  $Z_2$  is C. With the process described herein, the APIs or intermediates thereof produced are selected from: Clofarabine (Cl-F-araA), Decitabine (aza-dCyd), Cytarabine (ara-C), Vidarabine, Brivudine, Encitabine (BH-AC), Zalcitabine (ddC), Cladribine (Cl-dAdo), Fludarabine (F-araA), Nelarabine (MAY), Zidovudine, Floxuridine (FUDR),  $\beta$ -Thymidine, idoxuridine (IdU), trifluridine (TFT), acedurid (EdU), ribavirin, didanosine (ddl) and Pentostatin.

[0117] More preferably, the APIs or intermediates thereof produced are selected from: Clofarabine, Cytarabine, Vidarabine, Brivudine, Encitabine, Zalcitabine, Cladribine, Fludarabine, Nelarabine, Zidovudine, Floxuridine and Pentostatin; yet more preferably the APIs or intermediates thereof produced are selected from Clofarabine, Cytarabine and Zidovudine (AZT); and still more preferably the described process is particularly intended to the industrial production of Clofarabine, Cytarabine and an intermediate of Zidovudine; furthermore preferably Clofarabine or intermediates thereof, yet even more preferably Clofarabine.

[0118] However, for the purpose of limiting the scope of present invention, the following NAs are specifically disclaimed: CAS No. 2627-62-5, CAS No. 7481-89-2, CAS No. 50-91-9, CAS No. 50-90-8, CAS No. 54-42-2, CAS No. 10356-76-0, CAS No. 2239-64-7, CAS No. 4546-70-7, CAS No. 15176-29-1, CAS No. 70-00-8, CAS No. 838-07-3, CAS No. 10212-20-1, but only with regard to a biocatalytic synthesis mediated by mesophilic NDT, wherein the pilot production of the aforementioned list of NAs did not render sufficient yields to scale-up that production to industry efficiently.

[0119] For the purposes of present description organisms from which the enzymes having NDT activity originate used may be mesophiles, thermophiles or hyperthermophiles. Mesophiles or mesophilic are those able to work or to carry out a nucleoside deoxyribosyltransferase activity, at temperatures ranging from 18 to up 60° C., with an optimal temperature range of 40-55° C. Organisms or NDT enzymes, thermophiles or thermophilic are those able to work or to carry out a nucleoside deoxyribosyltransferase activity, at temperatures ranging over 60° C. and up to 80° C. Organisms or NDT enzymes, hyperthermophiles or hyperthermophilic, are those able to work or to carry out a nucleoside deoxyribosyltransferase activity, at temperatures over 80° C. and up to 100° C., with an optimal temperature range of 80-95° C.

[0120] Still in one more preferred embodiment of practicing the process describe hereto, the API, or intermediates thereof, produced is Clofarabine (Scheme 2).



**[0121]** The process of invention applied to the production of clorafabine has several advantages over the industrial process used in the prior art for the chemical synthesis of this NA, the so called ILEX procedure (Bauta et al, Org. Proc. Res. Dev. 2004, 8, 889-896):

- [0122]** (i) One-pot synthesis,
- [0123]** (ii) Reduced number of steps,
- [0124]** (iii) Higher conversions and yields,
- [0125]** (iv) Avoidance of organic solvents in the enzymatic step,
- [0126]** (v) No protection/deprotection strategies are needed, e.g. for the hydroxyl groups in the sugar,
- [0127]** (vi) Mild reaction conditions: environmentally-friendly technology (water or aqueous medium, neutral pH),
- [0128]** (vii) Extremely good selectivity: stereoselectivity—enantioselectivity, chemo-regioselectivity,
- [0129]** (viii) Fewer or no side reactions: impurity profile (reduced by-products content),
- [0130]** (ix) Reduction in overall waste generation,
- [0131]** (x) Process productivity
- [0132]** (xi) Overall lower cost of production

**[0133]** Moreover, there are other additional advantages of the biocatalytic process of invention over the chemical process used in the prior art. Particularly relevant is the absence in the process of the invention of the organic solvents used in the chemical process of Bauta et al., as dichloromethane (DCM), acetonitrile (ACN), 1,2-dichloroethane (DCE), methanol, heptane, etc. . . . All those organic solvents must be removed prior to any process waste discharge to the environment. An supplementary inconvenient of ILEX process with regard to the procedure of invention is its complexity, as far as ILEX process comprises at least 3 procedure steps, against the single step-one pot process of invention.

**[0134]** The source of native or recombinant enzyme used in the process of the invention are Bacteria or Archaea organisms. Enzymes having nucleoside deoxyribosyltransferase activity used in the process of invention can be isolated from a microorganism selected from, as a way of example, Bacte-

ria, particularly from *Lactobacillus* or *Lactococcus* species or from Archaea particularly from any of the following genus: *Thermofilum*; *Haloarcula*; *Natronococcus*; *Natrialba*; *Halobiforma*; *Methanosarcina*; *Methanomethylovorans*; *Methanoculleus*; *Methanosphaerula*; *Methanohalobium*; *Methanosalsum* or *Methanosaceta*.

**[0135]** Still another embodiment of present invention relates to the use in the process disclosed of a recombinant NDT enzyme comprising an amino acid sequence encoded by a nucleic acid sequence, or fragments thereof, isolated from Archaea.

**[0136]** As described all along present description one of the preferred embodiments for practicing the process of invention is based on the use as biocatalyst of NDT enzymes for performing the base transfer one step-one pot reaction. Preferably NDT enzymes are of recombinant type, comprising gene sequences of fragments thereof, encoding for NDT enzymatic activities able to carry out nucleobase transfer one step-one pot reactions, at temperatures ranging over 60° C. and up to 100° C. and, more preferably, at the suitable reaction conditions described herein.

**[0137]** Suitable conditions for carrying out the different embodiments of the process of invention comprise:

**[0138]** a) Temperature ranging 18-100° C., preferably 20-100° C., and more preferably ranging 40-100° C., furthermore preferably 50-100° C.

**[0139]** b) Reaction time ranging 1-600 h

**[0140]** c) Concentration of starting material ranging 1-500 mM

**[0141]** d) Stoichiometry nucleoside starting material: nucleobase ranges from 1:5 to 5:1.

**[0142]** e) Amount of enzyme having NDT activity or NDT enzyme ranging 0.001-100 mg/ml, preferably 0.001-10 mg/mL.

**[0143]** f) Free nucleobase added to the reaction medium, optionally dissolved in an organic solvent

**[0144]** g) Aqueous reaction medium optionally also containing up to 40% of a suitable organic solvent. Preferably up to 20% and more preferably up to 5%.

[0145] Preferred organic solvents to be added to the reaction medium or to be used for dissolving previously the free nucleobases are polar aprotic solvents, preferably selected from: Tetrahydrofuran, acetonitrile, acetone or Dimethylformamide (DMF).

[0146] The process according to present invention also includes an isolation and/or purification steps of the NA produced by standard operation means selected from: chromatography, precipitation, filtration, concentration or crystallization.

[0147] For the purposes also of present description, the term recombinant enzyme or recombinant type enzyme should be understood as a protein or enzyme that is derived from recombinant DNA. The term "recombinant DNA" is a form of DNA that does not exist naturally, which is created by combining DNA sequences that would not normally occur together.

[0148] The process of the invention, specifically covers the one wherein the API produced is Clofarabine (Scheme 2), the 2'-deoxyribonucleoside used as starting material is 2'-fluoro-arabinofuranosyl-2'-deoxyuridine, the nucleobase also used as starting material to be transferred by the NDT enzyme, is 2-chloroadenine and the NDT is a natural occurring NDT enzyme, isolated from *Lactobacillus delbrueckii* (formerly called *Lactobacillus leichmannii*).

[0149] One of the embodiments of the process of invention covers the one wherein the API produced is, Clofarabine (Scheme 2), the 2'-deoxyarabinonucleoside used as starting material, is any suitable 2'-fluoro-arabino-type, preferably 2'-fluoro-arabinofuranosyl-2'-deoxyuridine, the nucleobase also used as starting material to be transferred by the NDT enzyme, is 2-chloroadenine.

[0150] In another embodiment of the process of invention, the API produced is Cytarabine, the 2'-deoxyarabinonucleoside used as starting material, is arabinofuranosyl-2'-deoxyuridine, the nucleobase also used as starting material to be transferred by the NDT enzyme is cytosine.

[0151] In another embodiment of the process of invention, the API produced is an intermediate in the synthesis of Zidovudine, the nucleoside used as starting material is 3'-amino-2',3'-dideoxyadenosine, the nucleobase also used as starting material to be transferred by the NDT enzyme is thymine. Said intermediate of Zidovudine, thus obtained, comprises a primary amine moiety which may be readily transformed into zidovudine using a further azidation step using, for example inorganic azides or azo-transfer compounds such as trifluoromethanesulfonyl azide or imidazole-1-sulfonyl azide.

[0152] Forming part of the same inventive concept, present description also discloses recombinant nucleoside deoxyriboseyltransferase enzymes (NDTs), able to carry out the process of the invention described above. The said native or recombinant NDT enzymes can be isolated from, bacteria preferably selected from *Lactobacillus delbrueckii* (formerly *Lactobacillus leichmannii*) or *Lactococcus lactis*; or from Archaea, preferably selected from: *Thermofilum pendens*; *Haloarcula sinaiensis*; *Natronococcus amylolyticus*; *Natrialba hulunbeirensis*; *Halobiforma nitratireducens*; *Methanosarcina maezi*; *Methanomethylovorans hollandica*; *Methanococcus bourgensis*; *Methanosphaerula palustres*; *Methanohalobium vestigatum*; *Methanosalsum zhilinae* or *Methanoseta harundinacea*.

[0153] In one preferred embodiment of the invention, the NDT enzyme is obtained from

[0154] a) *Lactobacillus delbrueckii* (formerly called *Lactobacillus leichmannii*) nucleotide encoding sequence shown in SEQ ID NO 1; or

[0155] b) a nucleotide sequence which is the complement of SEQ ID. NO: 1; or

[0156] c) a nucleotide sequence which is degenerate with SEQ ID. NO: 1; or

[0157] d) a nucleotide sequence hybridizing under conditions of high stringency to SEQ ID. NO: 1, to the complement of SEQ ID. NO: 1, or to a hybridization probe derived from SEQ ID. NO: 1, or their complement thereof; or

[0158] e) a nucleotide sequence having at least 80% sequence identity with SEQ ID. NO: 1; or

[0159] f) a nucleotide sequence having at least 65% sequence identity with SEQ ID. NO: 1, wherein said sequence preferably encodes or is complementary to a sequence encoding at least a NDT enzyme or a functional part thereof.

[0160] g) a nucleotide sequence encoding for a polypeptide having NDT activity, the amino acid sequence of which is at least 80% identical to the amino acid sequence shown in SEQ ID. NO: 2, further preferably a nucleotide sequence encoding for the amino acid sequence shown in SEQ ID. NO: 2.

[0161] In another preferred embodiment of the invention, the NDT enzyme is obtained from

[0162] a) *Lactococcus lactis* nucleotide encoding sequence shown in SEQ ID NO. 3 or

[0163] b) a nucleotide sequence which is the complement of SEQ ID. NO:3; or

[0164] c) a nucleotide sequence which is degenerate with SEQ ID. NO:3; or

[0165] d) a nucleotide sequence hybridizing under conditions of high stringency to SEQ ID. NO:3, to the complement of SEQ ID. NO:3, or to a hybridization probe derived from SEQ ID. NO:3, or their complement thereof; or

[0166] e) a nucleotide sequence having at least 80% sequence identity with SEQ ID. NO:3; or

[0167] f) a nucleotide sequence having at least 65% sequence identity with SEQ ID. NO:3, wherein said sequence preferably encodes or is complementary to a sequence encoding at least a NDT enzyme or a functional part thereof.

[0168] g) a nucleotide sequence encoding for a polypeptide having NDT activity, the amino acid sequence of which is at least 80% identical to the amino acid sequence shown in SEQ ID. NO:4, further preferably a nucleotide sequence encoding for the amino acid sequence shown in SEQ ID. NO:4.

[0169] In still other preferred embodiment of present invention the NDT enzyme is obtained from *Thermofilum pendens* nucleotide encoding sequence shown in SEQ ID NO. 5 or

[0170] a) a nucleotide sequence which is the complement of SEQ ID. NO:5; or

[0171] b) a nucleotide sequence which is degenerate with SEQ ID. NO:5; or

[0172] c) a nucleotide sequence hybridizing under conditions of high stringency to SEQ ID. NO:5, to the

complement of SEQ ID. NO:5, or to a hybridization probe derived from SEQ ID. NO:5, or their complement thereof; or

**[0173]** d) a nucleotide sequence having at least 80% sequence identity with SEQ ID. NO:5; or

**[0174]** e) a nucleotide sequence having at least 65% sequence identity with SEQ ID. NO:5, wherein said sequence preferably encodes or is complementary to a sequence encoding at least a NDT enzyme or a functional part thereof.

**[0175]** f) a nucleotide sequence encoding for a polypeptide having NDT activity, the amino acid sequence of which is at least 80% identical to the amino acid sequence shown in SEQ ID. NO:6, further preferably a nucleotide sequence encoding for the amino acid sequence shown in SEQ ID. NO:6.

**[0176]** Also in the same line of integrating a single inventive linked concept, present description also discloses the use of a nucleoside deoxyribosyltransferase (NDT) in the production of APIs or intermediates thereof, being those APIs or their intermediates, nucleoside analogues (NAs) useful as anti-cancer or anti-viral medicaments. More preferably, the previously mentioned use is achieved by the production process and variants thereof, also previously detailed, of NAs as APIs or intermediates thereof. Particularly, related to that previously mentioned use, recombinant nucleoside deoxyribosyltransferases (NDTs) are preferred, in the production of APIs, being those APIs or their intermediates, nucleoside analogues (NAs) particularly useful as anti-cancer or anti-viral medicaments.

**[0177]** Among the APIs or intermediates thereof produced by such uses the following may be found: Clofarabine (Cl-F-araA), Decitabine (aza-dCyd), Cytarabine (ara-C), Vidarabine, Brivudine, Encitabine (BH-AC), Zalcitabine (ddC), Cladribine (Cl-dAdo), Fludarabine (F-araA), Nelarabine (MAY), Zidovudine, Floxuridine (FUDR),  $\beta$ -Thymidine, idoxuridine (IdU), trifluridine (TFT), acedurid (EdU), ribavirin, didanosine (ddl) and Pentostatin.

**[0178]** More preferably, the APIs or intermediates thereof produced by the use of the enzymes disclosed herein, are selected from: Clofarabine, Cytarabine, Vidarabine, Brivudine, Encitabine, Zalcitabine, Cladribine, Fludarabine, Nelarabine, Zidovudine, Floxuridine and Pentostatin; yet more preferably the APIs or intermediates thereof produced are selected from Clofarabine, Cytarabine and Zidovudine or intermediates thereof; and still more preferably the described uses are particularly dedicated to the industrial production of Clofarabine, Cytarabine and an intermediate of Zidovudine; furthermore preferably Clofarabine or intermediates thereof, yet more preferably Clofarabine.

**[0179]** Enzymes having N-deoxyribosyltransferase (NDT) activity to be used according to present invention include recombinant NDT enzymes encoded by a nucleic acid sequence selected from: SEQ ID No. 1, SEQ ID No. 3 or SEQ ID No. 5; or

**[0180]** a) a nucleotide sequence which is the complement of SEQ ID NO:1, SEQ. ID NO:3 or SEQ ID. NO:5; or

**[0181]** b) a nucleotide sequence which is degenerate with SEQ ID NO:1, SEQ. ID NO:3 or SEQ ID. NO:5; or

**[0182]** c) a nucleotide sequence hybridizing under conditions of high stringency to SEQ ID NO: 1, SEQ. ID NO:3 or SEQ ID. NO:5, to the complements of SEQ ID NO:1, SEQ. ID NO:3 or SEQ ID. NO:5, or to a hybrid-

ization probe derived from SEQ ID NO:1, SEQ. ID NO:3 or SEQ ID. NO:5, or their complement thereof; or

**[0183]** d) a nucleotide sequence having at least 80% sequence identity with SEQ ID NO:1, SEQ. ID NO:3 or SEQ ID. NO:5; or

**[0184]** e) a nucleotide sequence having at least 65% sequence identity with SEQ ID NO:1, SEQ. ID NO:3 or SEQ ID. NO:5, wherein said sequence preferably encodes or is complementary to a sequence encoding at least a NDT enzyme or a functional part thereof.

**[0185]** f) a nucleotide sequence encoding for an amino acid sequence selected from: SEQ ID NO:2, SEQ. ID NO:4 or SEQ ID NO:6.

**[0186]** Also included in present invention are recombinant expression vectors comprising sequences encoding a nucleoside deoxyribosyltransferase (NDT) operably linked to one or more control sequences that direct the expression or overexpression of said deoxyribosyltransferase in a suitable host. A preferred recombinant expression vector according to present invention is any carrying and expressing or overexpressing genes encoding said NDT enzymatic activities.

**[0187]** Preferred recombinant expression vectors according to present invention are carrying and expressing or overexpressing nucleic acid sequence selected from: SEQ ID No. 1, SEQ ID No. 3 or SEQ ID No. 5; or

**[0188]** a) a nucleotide sequence which is the complement of SEQ ID NO:1, SEQ. ID NO:3 or SEQ ID NO:5; or

**[0189]** b) a nucleotide sequence which is degenerate with SEQ ID NO:1, SEQ. ID NO:3 or SEQ ID NO:5; or

**[0190]** c) a nucleotide sequence hybridizing under conditions of high stringency to SEQ ID NO: 1, SEQ. ID NO:3 or SEQ ID NO:5, to the complements of SEQ ID NO:1, SEQ. ID NO:3 or SEQ ID NO:5 or to a hybridization probe derived from SEQ ID NO:1, SEQ. ID NO:3 or SEQ ID NO:5, or their complement thereof; or

**[0191]** d) a nucleotide sequence having at least 80% sequence identity with SEQ ID NO: 1, SEQ. ID NO:3, or SEQ ID NO:5; or

**[0192]** e) a nucleotide sequence having at least 65% sequence identity with SEQ ID NO:1, SEQ. ID NO:3 or SEQ ID NO:5, wherein said sequence preferably encodes or is complementary to a sequence encoding at least a NDT enzyme or a functional part thereof.

**[0193]** a nucleotide sequence encoding for an amino acid sequence selected from: SEQ ID NO:2, SEQ. ID NO:4 or SEQ ID NO:6.

**[0194]** The invention also covers the use of the recombinant expression vectors previously said, for the production either of recombinant NDT enzymes or for the production of active pharmaceutical ingredients (APIs) or intermediates thereof, being those APIs or their intermediates, nucleoside analogues (NAs) particularly useful as anti-cancer or anti-viral medicaments. Particularly APIs or intermediates thereof produced by the aforesaid use are selected from: Clofarabine (Cl-F-araA), Decitabine (aza-dCyd), Cytarabine (ara-C), Vidarabine, Brivudine, Encitabine (BH-AC), Zalcitabine (ddC), Cladribine (Cl-dAdo), Fludarabine (F-araA), Nelarabine (MAY), Zidovudine, Floxuridine (FUDR),  $\beta$ -Thymidine, idoxuridine (IdU), trifluridine (TFT), acedurid (EdU), ribavirin, didanosine (ddl) and Pentostatin.

**[0195]** More preferably, the APIs or intermediates thereof produced are selected from: Clofarabine, Cytarabine, Vidara-

bine, Brivudine, Enocitabine, Zalcitabine, Cladribine, Fludarabine, Nelerabine, Zidovudine, Floxuridine and Pentostatin; yet more preferably the APIs or intermediates thereof produced are selected from Clofarabine, Cytarabine and Zidovudine; and still more preferably the expression vectors described herein are particularly suitable for industrial production of Clofarabine, Cytarabine and an intermediate of Zidovudine; furthermore preferably Clofarabine or intermediates thereof, yet even more preferably Clofarabine.

[0196] More preferably, the previously mentioned use for the production of APIs or intermediates thereof is achieved by the production process and variants thereof, also previously detailed.

[0197] The invention also covers host cells comprising the recombinant expression vectors, previously described, particularly when said host cell is *Escherichia coli*. Included as part of the same inventive single linked concept, the use of host cells comprising the recombinant expression vectors as previously described, for the production of recombinant NDTs, is also contemplated. Analogously, the use of host cells comprising the recombinant expression vectors as previously described, for the production of active pharmaceutical ingredients (APIs) or intermediates thereof, being those APIs or their intermediates, nucleoside analogues (NAs) useful as anti-cancer or anti-viral medicaments, is also part of present invention. Particularly if those host cells comprising the recombinant expression vectors as previously described, are used for producing APIs or intermediates thereof selected from: Clofarabine (Cl-F-araA), Decitabine (*aza-dCyd*), Cytarabine (ara-C), Vidarabine (araA), Brivudine, Enocitabine (BH-AC), Zalcitabine (ddC), Cladribine (Cl-dAdo), Fludarabine (F-araA), Nelerabine (MAY), Zidovudine, Floxuridine (FUDR),  $\beta$ -Thymidine, idoxuridine (IdU), trifluridine (TFT), acedurid (EdU), ribavirin, didanosine (ddl) and Pentostatin.

[0198] More preferably, the APIs or intermediates thereof produced using the host cells of present invention, transformed with the recombinant expression vectors, previously described, are selected from: Clofarabine, Cytarabine, Vidarabine, Brivudine, Enocitabine, Zalcitabine, Cladribine, Fludarabine, Nelerabine, Zidovudine, Floxuridine and Pentostatin; yet more preferably the APIs or intermediates thereof produced are selected from Clofarabine, Cytarabine and Zidovudine; and still more preferably the host transformed cells described herein are particularly suitable for industrial production of Clofarabine, Cytarabine and an intermediate of Zidovudine; furthermore preferably Clofarabine or intermediates thereof, yet more preferably Clofarabine.

[0199] More preferably, the previously mentioned use is achieved by the production process and variants thereof, also previously detailed, of NAs as APIs or intermediates thereof.

#### Example 1

##### Cloning and Expression NDTs from *Lactobacillus delbrueckii* Subsp. *lactis* DSM 20072

[0200] Cloning.

[0201] The gene encoding N-deoxyribosyltransferase from *Lactobacillus delbrueckii* subsp. *lactis* (in GenBank accession number EGD27012.1) was amplified by polymer chain reaction (PCR) from *Lactobacillus delbrueckii* subsp. *lactis* DSM 20072 genomic DNA. The gene was cloned into the polylinker region of the expression vectors: pET22 b(+), using the restriction sites NdeI and Sall

[0202] The primers used were designed with proper modifications at 5'-ends (see sequences):

```
Ldndt2 fw: (SEQ. ID No. 19)
5'-CATATGCCAAAAAGACGATCTACTTC-3'

Ldndt2 rv: (SEQ. ID No. 20)
5'-GTCGACTTAGTATACGGCACC-3'
```

[0203] The corresponding gene for nucleoside 2-deoxyribosyltransferase sequence was amplified by PCR, using the Platinum Taq enzyme (Invitrogen), with the corresponding forward and reverse oligonucleotides (SEQ. ID. From No. 19 and No. 20). The amplified 0.-kb product was inserted into a pGEM-T vector. The cloned region was completely sequenced and it was found to be the same to the data bank sequence. The insert DNA was digested with NdeI and Sall and then ligated into an expression vector of pET-22b(+) digested with the same restriction enzymes.

[0204] The ligation product was transformed into chemically competent *E. coli* DH5 $\alpha$  cells. Positive plasmids were selected and subsequently transformed into BL21(DE3) chemically competent cells.

[0205] Expression and Purification of NDT from *Lactobacillus delbrueckii* Subsp. *lactis*

[0206] *E. coli* strain bearing the recombinant plasmid were grown aerobically in LB medium supplemented with ampicillin (100  $\mu\text{g ml}^{-1}$ ) until OD<sub>600 nm</sub> of 0.5 was reached. Over-expression of the protein was achieved by inducing the *E. coli* culture with 0.5 mM isopropyl-1-thio- $\beta$ -D-galactopyranoside (IPTG) at 37° C. for 4 h. The bacterial cells were harvested by centrifugation at 4° C. at 4,000xg for 20 minutes. Harvested cells were resuspended in 20 ml of Phosphate Buffered Saline (PBS Buffer) for cells collected from one liter cell culture. The cell suspensions were lysed using sonication (Branson Sonifier 450). Cell lysate (supernatant) was obtained by centrifugation at 13000xg for 45 min to obtain an enzyme corresponding to the amino acid sequence SEQ. ID. NO. 2 encoded by nucleotide sequence SEQ. ID. NO. 1 and having NDT activity.

#### Example 2

##### Cloning and Expression of NDTs from *L. lactis*

[0207] Cloning.

[0208] The gene encoding N-deoxyribosyltransferase from *Lactococcus lactis* subsp. *lactis* (in GenBank accession number AE006284) was amplified by polymer chain reaction (PCR) from *Lactococcus lactis* subsp. *lactis* genomic DNA. The gene was cloned into the polylinker region of the expression vectors: pET22 b(+), using the restriction sites NdeI and EcoRI

[0209] The primers used were designed with proper modifications at 5'-ends (see sequences):

```
Llac fw: (SEQ. ID No. 17)
5'-GCCATATGAACCAAGTGTGTTAATCAAG-3'

Llac rv: (SEQ. ID No. 18)
5'-GCGAATTCTACTGGTATTTCCACTATA-3'
```

**[0210]** The corresponding gene for nucleoside 2-deoxyribosyltransferase (NDT) sequence was amplified by PCR, using the Platinum Taq enzyme (Invitrogen), with the corresponding forward and reverse oligonucleotides (SEQ. ID. No. 17 and No. 18). The amplified 0.5-kb product was inserted into a pGEM-T vector. The cloned region was completely sequenced and it was found to be the same as the data bank sequence. The insert DNA was digested with NdeI and EcoRI and then ligated into an expression vector of pET-22b (+) digested with the same restriction enzyme.

**[0211]** The ligation product was transformed into chemically competent *E. coli* DH5 $\alpha$  cells. Positive plasmids were selected and subsequently transformed into BL21(DE3) chemically competent cells.

**[0212]** Expression and Purification of NDT from *Lactococcus lactis* Subsp. *lactis*

**[0213]** *E. coli* strain bearing the recombinant plasmid were grown aerobically in LB medium supplemented with ampicillin (100  $\mu$ g ml<sup>-1</sup>) until OD<sub>600 nm</sub> of 0.5 was reached. Overexpression of the protein was achieved by inducing the *E. coli* culture with 0.5 mM isopropyl-1-thio- $\beta$ -D-galactopyranoside (IPTG) at 30° C. for 4 h. The bacterial cells were harvested by centrifugation at 4° C. at 4,000 $\times$ g for 20 minutes. Harvested cells were resuspended in 20 ml of Phosphate Buffered Saline (PBS Buffer) for cells collected from one liter cell culture. The cell suspensions were lysed using sonication (Branson Sonifier 450). Cell lysate (supernatant) was obtained by centrifugation at 13000 $\times$ g for 45 min.

#### Example 3

##### Cloning and Expression of a Nucleoside 2-Deoxyribosyltransferase Homologue Enzyme from *T. pendens*

**[0214]** Construction and Transformation of the *Escherichia coli* NDT

**[0215]** Cloning

**[0216]** The sequence (SEQ ID No. 5) was retrieved from GenBank (complement strain) and it comes from *Thermophilum pendens* Hrk 5 chromosome. The Tpen\_0017 gene locus (NC\_008698, Version: NC\_008698.1, Region: 14900 . . . 15331) was amplified by polymer chain reaction (PCR) from *Thermophilum pendens* Hrk 5 (DSM 2475) genomic DNA. Full-length 0017 gene was cloned without tags and optionally was cloned as either a C-His6-tagged and N-terminal His6-tagged fusion protein.

**[0217]** The gene was cloned into the polylinker region of the expression vectors: pET102/D-TOPO or pET22 b(+), using the restriction sites NdeI and EcoRI or NdeI and XhoI.

**[0218]** The primers used were designed with properly modifications at 5'-ends (see sequences):

```
Tpen_0017_For: (SEQ. ID No. 7)
5'-catatgaaggctacctggcg-3'

Tpen_0017_4Rev: (SEQ. ID No. 8)
5'-gaattatgcatgtcaacgtacc-3'

Tpen_0017_For: (SEQ. ID No. 9)
5'-caccatgaaggctacctggcg-3'
```

-continued

```
Tpen_0017_RevI: (SEQ. ID No. 10)
5'-tcattgcatgtcaacgtacc-3'

Tpen_0017_RevII: (SEQ. ID No. 11)
5'-ttgcatgtcaacgtacctg-3'

Tpen_0017_Rev: (SEQ. ID No. 12)
5'-ctcgagtcattgcatgtcaacg-3'

Tpen_0017_2Rev: (SEQ. ID No. 13)
5'-ctcgagttgcatgtcaacgtca-3'

Tpen_0017_3Rev: (SEQ. ID No. 14)
5'-gaatttcattgcatgtcaacgtc-3'

Tpen_0017_4Rev: (SEQ. ID No. 15)
5'-gaattatgcatgtcaacgtacc-3'

Tpen_0017_5Rev: (SEQ. ID No. 16)
5'-ggaattccgcttgcatgtcaacgtacc-3'
```

**[0219]** The forward primer was designed in order to adapt the start signal to the *E. coli* usage (GTG to ATG). The corresponding gene for Tpen\_0017 sequence was amplified by PCR, using the Platinum Taq enzyme (Invitrogen), with the corresponding forward and reverse oligonucleotides (SEQ. ID. From No. 7 to No. 16). The amplified fragment was subcloned in pGEM-T Easy and then digested and cloned into the polylinker region of the pET22 b(+) vector which carries the ampicillin resistance gene. The cloned region was completely sequenced and it was found to be the same to the data bank sequence. The ligation product was transformed into chemically competent cells *E. coli* TOP 10 cells (Invitrogen). Positive plasmids were selected and subsequently transformed into BL21(DE3) chemically competent cells.

**[0220]** Expression and Purification of NDI from *T. pendens*

**[0221]** *E. coli* strain bearing the recombinant plasmid was grown in LB medium supplemented with ampicillin (100  $\mu$ g ml<sup>-1</sup>) until OD<sub>600 nm</sub> of 0.5-0.8 was reached. Overexpression of the protein was achieved by inducing the *E. coli* culture with 1 mM isopropyl-1-thio- $\beta$ -D-galactopyranoside (IPTG) at 37° C. for 4 h. The bacterial cells were harvested by centrifugation at 4 OC at 4,000 $\times$ g for 20 minutes. Harvested cells were resuspended in 20 mL of Phosphate Buffered Saline (PBS Buffer) for cells collected from one liter cell culture. The cell suspension was lysed using sonication (Branson Sonifier 450). Cell lysate (supernatant) was obtained by centrifugation at 13000 $\times$ g for 20 min. Extracts (supernatants) were stored at -80° C.

#### Example 4

##### N-Deoxyribosyltransferase Assays

**[0222]** The activity assay was performed at 40° C., in 50 mM MES (morpholineethanesulfonic acid) buffer (pH 6.5). Under these conditions, a solution of 10 mM adenine and 10 mM 2'-deoxyuridine was incubated during 10 min. Then, the enzyme was added (2.04  $\mu$ g of purified enzyme or 30  $\mu$ l of cell extract) up to a final volume of 240  $\mu$ l, and the reaction was left to proceed during 5 min. The reaction was quenched through the addition of 240  $\mu$ l of cold MeOH in an ice bath.

and further heated at 95° C. for 5 min. Sample was centrifuged at 9700 rpm for 2 min, and the supernatant was diluted ½ with water. The diluted aliquot was filtered through 0.45 µm hydrophilic filter and then through a 10 KDa Ultrafiltration membrane. The sample was analyzed using a high performance liquid chromatograph (HPLC) with Tracer Excel C18 (Teknokroma, 5 µm, 25×0.46 cm) at 40° C., under MeOH/H<sub>2</sub>O gradient elution conditions, 1 ml/min flow. Detection was performed at least at 254 nm.

[0223] The quantification was performed through calculations on the released base, according to a standard solution of 2'-deoxyadenosine eluted in the same conditions. The enzymatic activity is expressed as, either:

[0224] (1) Specific activity (Sp act) as units-mg protein<sup>-1</sup> (µmoles of 2'-deoxyadenosine min<sup>-1</sup>·mg<sup>-1</sup>)

[0225] (2) Enzymatic activity as units·ml<sup>-1</sup> (µmoles of 2'-deoxyadenosine·min<sup>-1</sup>·ml<sup>-1</sup>)

wherein one unit of enzyme activity was defined as the amount of enzyme required to produce 1 µmol of product per min under standard conditions described herein.

#### Example 5

##### Clofarabine Production in Aqueous Media Using NDT Enzyme at 4.5 U/µmol<sub>araF</sub>

[0226] A suspension of 20 mM 2-chloroadenine (0.142 g, 0.84 mmol) and 20 mM 2'-fluoro-arabinofuranosyl-2'-deoxyuridine (araF) (0.206 g, 0.84 mmol) in aqueous buffer at pH 6.5 (42 ml) was thermostated at 50° C. during 30 min. Then, NDT enzyme (SEQ ID NO: 2) was added (4.5 U/µmol<sub>araF</sub>, 0.422 mg/ml) dropwise and the reaction was stirred at 50° C. during at least 10 days at the same conditions. Then, the suspension was hot filtered, and the solid was washed and dried. The aqueous filtrate was partially concentrated, cooled down and filtered. The recovered solid was allowed to partially precipitate under basic pH conditions and concentration. Once filtered and washed, the solid may be optionally crystallized or recrystallized using a suitable solvent, such as a polar protic or polar aprotic solvent, in combination with water or a suitable apolar solvent.

#### Example 6

##### Clofarabine Production in Aqueous Media Using NDT Enzyme at 1.78 U/µmol<sub>araF</sub>

[0227] A suspension of 50 mM 2-chloroadenine (0.356 g, 2.1 mmol) and 50 mM 2'-fluoro-arabinofuranosyl-2'-deoxyuridine (araF) (0.517 g, 2.1 mmol) in aqueous buffer at pH 6.5 (42 ml) was thermostated at 50° C. during 30 min. Then, NDT enzyme (SEQ ID NO: 2) was added (1.78 U/µmol<sub>araF</sub>, 0.422 mg/ml) dropwise and the reaction was stirred at 50° C. during at least 10 days at the same conditions. Then, the suspension was hot filtered, and the solid was washed and dried. The aqueous filtrate was partially concentrated, cooled down and filtered. The recovered solid was allowed to partially precipitate under basic pH conditions and concentration. Once filtered and washed, the solid may be optionally crystallized or recrystallized using a suitable solvent, such as a polar protic or polar aprotic solvent, in combination with water or a suitable apolar solvent.

#### Example 8

##### Clofarabine Production in Aqueous Buffer Containing Organic Solvent Using NDT Enzyme at 0.9 U/µmol<sub>araF</sub>

[0228] A suspension of 20 mM 2-chloroadenine (10.3 mg, 0.06 mmol) and 20 mM 2'-fluoro-arabinofuranosyl-2'-deoxyuridine (araF) (14.8 mg, 0.06 mmol) in a mixture of aqueous buffer at pH 6.5 and 5% THF (reaction volume 3 ml) was thermostated at 50° C. during 30 min. Then, NDT enzyme (SEQ ID NO: 2) was added (0.9 U/µmol<sub>araF</sub>, 0.088 mg/ml) dropwise and the reaction was stirred at 50° C. during at least 5 days at the same conditions. Then, the suspension was hot filtered, and the solid was washed and dried. The aqueous filtrate was partially concentrated, cooled down and filtered. The recovered solid was allowed to partially precipitate under basic pH conditions and concentration. Once filtered and washed, the solid may be optionally crystallized or recrystallized using a suitable solvent, such as a polar protic or polar aprotic solvent, in combination with water or a suitable apolar solvent.

#### Example 9

##### Clofarabine Production in Aqueous Media Using NDT Enzyme at 0.9 U/µmol<sub>araF</sub>

[0229] A suspension of 100 mM 2-chloroadenine (0.712 g, 4.2 mmol) and 100 mM 2'-fluoro-arabinofuranosyl-2'-deoxyuridine (araF) (1.034 g, 4.2 mmol) in aqueous buffer at pH 6.5 (42 ml) was thermostated at 50° C. during 30 min. Then, NDT enzyme (SEQ ID NO: 2) was added (0.9 U/µmol<sub>araF</sub>, 0.422 mg/ml) dropwise and the reaction was stirred at 50° C. during at least 10 days at the same conditions. Then, the suspension was hot filtered, and the solid was washed and dried. The aqueous filtrate was partially concentrated, cooled down and filtered. The recovered solid was allowed to partially precipitate under basic pH conditions and concentration. Once filtered and washed, the solid may be optionally crystallized or recrystallized using a suitable solvent, such as a polar protic or polar aprotic solvent, in combination with water or a suitable apolar solvent.

#### Example 10

##### Clofarabine Production Using NDT Enzyme at 12.8 U/µmol<sub>araF</sub>

[0230] A suspension of 10 mM 2-chloroadenine (3.47 g, 20.1 mmol) and 10 mM 2'-fluoro-arabinofuranosyl-2'-deoxyuridine (araF) (5.0 g, 20.1 mmol) in aqueous buffer at pH 6.5 (2 L) was thermostated at 50° C. during 30 min. Then, NDT enzyme (SEQ ID NO: 2) was added (12.8 U/µmol<sub>araF</sub>, 0.642 mg/ml) dropwise and the reaction was stirred at 50° C. during at least 3 days at the same conditions. Then, the suspension was hot filtered, and the solid was washed and dried. The aqueous filtrate was partially concentrated, cooled down and filtered. The recovered solid was allowed to partially precipitate and concentrate. Once filtered and washed, the solid was crystallized using a suitable solvent, such as a polar protic or polar aprotic solvent, in combination with water or a suitable apolar solvent.



## Example 11

Clofarabine Production Using NDT Enzyme at 1.2 U/ $\mu\text{mol}_{\text{araF}}$ 

[0231] A suspension of 50 mM 2-chloroadenine (3.03 g, 17.9 mmol) and 50 mM 2'-fluoro-arabinofuranosyl-2'-deoxyuridine (araF) (4.4 g, 17.9 mmol) in aqueous buffer at pH 6.5 (0.35 L) was thermostated at 50° C. during 30 min. Then, NDT enzyme (SEQ ID NO: 2) was added (1.2 U/ $\mu\text{mol}_{\text{araF}}$ , 0.37 mg/ml) dropwise and the reaction was stirred at 50° C. during at least 4 days at the same conditions. Then, the suspension was hot filtered, and the solid was washed and dried. The aqueous filtrate was partially concentrated, cooled down and filtered. The recovered solid was allowed to partially precipitate and concentrate. Once filtered and washed, the solid was crystallized using a suitable solvent, such as a polar protic or polar aprotic solvent, in combination with water or a suitable apolar solvent.

## Example 12

Clofarabine Production Using NDT Enzyme at 9 U/ $\mu\text{mol}_{\text{araF}}$ 

[0232] A suspension of 10 mM 2-chloroadenine (0.06 g, 0.35 mmol) and 10 mM 2'-fluoro-arabinofuranosyl-2'-deoxyuridine (araF) (0.09 g, 0.35 mmol) in aqueous buffer at pH 6.5 (42 ml) was thermostated at 50° C. during 30 min. Then, NDT enzyme (SEQ ID NO: 2) was added (9 U/ $\mu\text{mol}_{\text{araF}}$ ) dropwise and the reaction was stirred at 50° C. during at least 1 days at the same conditions. Then, the suspension was hot filtered, and the solid was washed and dried. The aqueous filtrate was partially concentrated, cooled down and filtered. The recovered solid was allowed to partially precipitate and concentrate. Once filtered and washed, the solid was crystallized using a suitable solvent, such as a polar protic or polar aprotic solvent, in combination with water or a suitable apolar solvent.

## Example 13

Clofarabine production using NDT enzyme at 1.1 U/ $\mu\text{mol}_{\text{araF}}$ 

[0233] A suspension of 5 mM 2-chloroadenine and 5 mM 2'-fluoro-arabinofuranosyl-2'-deoxyuridine (araF) in ammonium containing aqueous Tris-HCl buffer was thermostated at 25° C. during 25 min. Then, NDT enzyme (SEQ ID NO: 4) was added (1.1 U/ $\mu\text{mol}_{\text{araF}}$ ) dropwise and the reaction was stirred at 25° C. during at least 4 days at the same conditions. Then, the suspension was hot filtered, and the solid was washed and dried. The aqueous filtrate was partially concentrated, cooled down and filtered. The recovered solid was allowed to partially precipitate and concentrate. Once filtered and washed, the solid was crystallized using a suitable solvent, such as a polar protic or polar aprotic solvent, in combination with water or a suitable apolar solvent.

## Example 14

Production of Zidovudine Intermediate Using NDT Enzyme at 0.06 U/ $\mu\text{mol}_{\text{araF}}$ 

[0234] A suspension of 1 mM thymine and 1 mM 3'-amino-2',3'-dideoxyadenosine in aqueous 50 mM MES buffer was thermostated at 50° C. during 20 min. Then, NDT enzyme

(SEQ ID NO: 2) was added (0.06 mg/ $\mu\text{mol}_{\text{base}}$ ) dropwise and the reaction was stirred at 50° C. during at least 1 day under the same conditions. Then, the suspension was hot filtered, and the solid was washed and dried. The aqueous filtrate was partially concentrated, cooled down and filtered. The recovered solid was allowed to partially precipitate and concentrate. Once filtered and washed, the solid was crystallized using a suitable solvent, such as a polar protic or polar aprotic solvent, in combination with water or a suitable apolar solvent.

## Example 15

Production of Zidovudine Intermediate Using NDT Enzyme at 0.14 U/ $\mu\text{mol}_{\text{araF}}$  and 60 mM Thymine

[0235] A suspension of 60 mM thymine and 60 mM 3'-amino-2',3'-dideoxyadenosine in aqueous 50 mM MES buffer was thermostated at 50° C. during 20 min. Then, NDT enzyme (SEQ ID NO: 2) was added (0.14 mg/ $\mu\text{mol}_{\text{base}}$ ) dropwise and the reaction was stirred at 50° C. during at least 1 day under the same conditions. Then, the suspension was hot filtered, and the solid was washed and dried. The aqueous filtrate was partially concentrated, cooled down and filtered. The recovered solid was allowed to partially precipitate and concentrate. Once filtered and washed, the solid was crystallized using a suitable solvent, such as a polar protic or polar aprotic solvent, in combination with water or a suitable apolar solvent.

## Example 16

Production of Zidovudine Intermediate Using NDT Enzyme at 0.07 U/ $\mu\text{mol}_{\text{araF}}$ 

[0236] A suspension of 100 mM thymine and 100 mM 3'-amino-2',3'-dideoxyadenosine in aqueous 50 mM MES buffer was thermostated at 50° C. during 20 min. Then, NDT enzyme (SEQ ID NO: 2) was added (0.07 mg/ $\mu\text{mol}_{\text{base}}$ ) dropwise and the reaction was stirred at 50° C. during at least 1 day under the same conditions. Then, the suspension was hot filtered, and the solid was washed and dried. The aqueous filtrate was partially concentrated, cooled down and filtered. The recovered solid was allowed to partially precipitate and concentrate. Once filtered and washed, the solid was crystallized using a suitable solvent, such as a polar protic or polar aprotic solvent, in combination with water or a suitable apolar solvent.

## Example 17

Production of Zidovudine Intermediate Using NDT Enzyme at 0.14 U/ $\mu\text{mol}_{\text{araF}}$  and 80 mM thymine

[0237] A suspension of 80 mM thymine and 80 mM 3'-amino-2',3'-dideoxyadenosine in aqueous 50 mM MES buffer was thermostated at 50° C. during 20 min. Then, NDT enzyme (SEQ ID NO: 2) was added (0.14 mg/ $\mu\text{mol}_{\text{base}}$ ) dropwise and the reaction was stirred at 50° C. during at least 1 day under the same conditions. Then, the suspension was hot filtered, and the solid was washed and dried. The aqueous filtrate was partially concentrated, cooled down and filtered. The recovered solid was allowed to partially precipitate and concentrate. Once filtered and washed, the solid was crystal-

lized using a suitable solvent, such as a polar protic or polar aprotic solvent, in combination with water or a suitable apolar solvent.

#### Example 18

Cytarabine Production Using NDT Enzyme at 0.085 U/ $\mu$ mol<sub>araF</sub>

[0238] A suspension of 1 mM cytosine and 1 mM 1-(B-D-arabinofuranosyl)uracil in aqueous 50 mM MES buffer was

thermostated at 50° C. during 30 min. Then, NDT enzyme (SEQ ID NO: 2) was added (0.085 mg/ml) dropwise and the reaction was stirred at 50° C. during at least 1 day under the same conditions. Then, the suspension was hot filtered, and the solid was washed and dried. The aqueous filtrate was partially concentrated, cooled down and filtered. The recovered solid was allowed to partially precipitate and concentrate. Once filtered and washed, the solid was crystallized using a suitable solvent, such as a polar protic or polar aprotic solvent, in combination with water or a suitable apolar solvent.

#### SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 20

<210> SEQ ID NO 1

<211> LENGTH: 474

<212> TYPE: DNA

<213> ORGANISM: *Lactobacillus delbrueckii*

<400> SEQUENCE: 1

```
atgccaaaa agacgatcta cttcggtgcc ggctggttca ctgaccgcca aaacaaagcc    60
tacaaggaag ccatggaagc cctcaaggaa aacccaacga ttgacctgga aaacagctac    120
gttccoctgg acaaccagta caagggtatc cgggttgatg aacacccgga atacctgeat    180
gacaagggtt gggctacggc cacctacaac aacgacttga acgggatcaa gaccaacgac    240
atcatgctgg gtgtctacat cctgacgaa gaagacgtcg gcctgggcat ggaactgggt    300
tacgccttga gccaaaggcaa gtacgtcett ttggtcatcc cggacaaga ctacggcaag    360
ccgatcaacc tcatgagctg gggcgtcagc gacaacgtga tcaagatgag ccagctgaag    420
gacttcaact tcaacaagcc ggccttcgac ttctacgaag gtgccgtata ctaa        474
```

<210> SEQ ID NO 2

<211> LENGTH: 157

<212> TYPE: PRT

<213> ORGANISM: *Lactobacillus delbrueckii*

<400> SEQUENCE: 2

```
Met Pro Lys Lys Thr Ile Tyr Phe Gly Ala Gly Trp Phe Thr Asp Arg
 1          5          10          15
Gln Asn Lys Ala Tyr Lys Glu Ala Met Glu Ala Leu Lys Glu Asn Pro
 20          25          30
Thr Ile Asp Leu Glu Asn Ser Tyr Val Pro Leu Asp Asn Gln Tyr Lys
 35          40          45
Gly Ile Arg Val Asp Glu His Pro Glu Tyr Leu His Asp Lys Val Trp
 50          55          60
Ala Thr Ala Thr Tyr Asn Asn Asp Leu Asn Gly Ile Lys Thr Asn Asp
 65          70          75          80
Ile Met Leu Gly Val Tyr Ile Pro Asp Glu Glu Asp Val Gly Leu Gly
 85          90          95
Met Glu Leu Gly Tyr Ala Leu Ser Gln Gly Lys Tyr Val Leu Leu Val
100          105          110
Ile Pro Asp Glu Asp Tyr Gly Lys Pro Ile Asn Leu Met Ser Trp Gly
115          120          125
Val Ser Asp Asn Val Ile Lys Met Ser Gln Leu Lys Asp Phe Asn Phe
130          135          140
```

-continued

Asn	Lys	Pro	Arg	Phe	Asp	Phe	Tyr	Glu	Gly	Ala	Val	Tyr
145					150					155		

<210> SEQ ID NO 3  
 <211> LENGTH: 480  
 <212> TYPE: DNA  
 <213> ORGANISM: *Lactobacillus lactis* subsp. *lactis*  
 <400> SEQUENCE: 3

ttgaacaagt	tgtttaatca	agcagttaac	gtttaccttg	cgcaccatt	ttttagtga	60
agtcaataa	aaaagttga	acttttagaa	aatgcacttt	caaaaaataa	aacagtagca	120
aaotTTTTTA	gccaatgag	atgtcaacat	cctgaatctt	taccacaaga	agttgaagct	180
tttacccttg	aatgggcaa	agcgacaatg	gaaatgatg	taaagaggt	aaataaagca	240
gatatcatg	ttgcaattgt	tgatttcgat	catcaagata	ctgattctgg	aacagcttgg	300
gagcttgct	acgccatcgc	tttagaaaa	ccaacctatc	ttattcgttt	tgaagatact	360
attccagcaa	atataatgct	cactgagcga	aatagagctt	tcttcacca	gattgaacaa	420
gtggaagaat	atgatttttt	agagtctaaa	ctaatcccat	atagtggaaa	ataccagtag	480

<210> SEQ ID NO 4  
 <211> LENGTH: 159  
 <212> TYPE: PRT  
 <213> ORGANISM: *Lactobacillus lactis* subsp. *lactis*  
 <400> SEQUENCE: 4

Met	Asn	Lys	Leu	Phe	Asn	Gln	Ala	Val	Asn	Val	Tyr	Leu	Ala	Ala	Pro
1				5					10					15	
Phe	Phe	Ser	Glu	Ser	Gln	Ile	Lys	Lys	Val	Glu	Leu	Leu	Glu	Asn	Ala
			20				25						30		
Leu	Ser	Lys	Asn	Lys	Thr	Val	Ala	Asn	Phe	Phe	Ser	Pro	Met	Arg	Cys
			35			40						45			
Gln	His	Pro	Glu	Ser	Leu	Pro	Gln	Glu	Val	Glu	Ala	Phe	Thr	Pro	Glu
	50					55					60				
Trp	Ala	Lys	Ala	Thr	Met	Glu	Asn	Asp	Val	Asn	Glu	Val	Asn	Lys	Ala
	65				70				75					80	
Asp	Ile	Ile	Val	Ala	Ile	Val	Asp	Phe	Asp	His	Gln	Asp	Thr	Asp	Ser
			85					90						95	
Gly	Thr	Ala	Trp	Glu	Leu	Gly	Tyr	Ala	Ile	Ala	Leu	Glu	Lys	Pro	Thr
			100				105						110		
Tyr	Leu	Ile	Arg	Phe	Glu	Asp	Thr	Ile	Pro	Ala	Asn	Ile	Met	Leu	Thr
			115				120						125		
Glu	Arg	Asn	Arg	Ala	Phe	Phe	Thr	Gln	Ile	Glu	Gln	Val	Glu	Glu	Tyr
			130				135					140			
Asp	Phe	Leu	Glu	Ser	Lys	Leu	Ile	Pro	Tyr	Ser	Gly	Lys	Tyr	Gln	
			145		150				155						

<210> SEQ ID NO 5  
 <211> LENGTH: 432  
 <212> TYPE: DNA  
 <213> ORGANISM: *Thermofilum pendens* Hrk 5  
 <400> SEQUENCE: 5

gtgaaggtct	acctggcggc	tccgatgagg	ggtgaccgta	gcgcgctggc	aaacgtgaag	60
aagctgttgc	aagccttggg	ggagaggggg	tacgtcgtgt	tgaacgaagca	cgtagcggac	120

-continued

---

```

gacgtgctcg acgtggagaa gggatgacg cctagagagg tcttcgagag ggatataagg 180
ttgctggaag aggcggacgt cctgggtggcg gaggtatcgt acccgagcct cggcgtgggc 240
ttcgagatag cgtactttct gctgaggggg aagccgggtga tagccctggc cttgcccag 300
aggctggaat cggatccgc gatgataagg ggtataacgt gggagaactt caggctggta 360
gcctactcgg acgtcgacga ggcaatagaa aaattagaca gcatgttgcc aggtagcgtt 420
gacatgcaat ga 432

```

```

<210> SEQ ID NO 6
<211> LENGTH: 143
<212> TYPE: PRT
<213> ORGANISM: Thermofilum pendens Hrk 5

```

```

<400> SEQUENCE: 6

```

```

Met Lys Val Tyr Leu Ala Ala Pro Met Arg Gly Asp Arg Ser Ala Leu
1          5          10          15
Ala Asn Val Lys Lys Leu Leu Gln Ala Leu Glu Glu Arg Gly Tyr Val
20          25          30
Val Leu Thr Lys His Val Ala Asp Asp Val Leu Asp Val Glu Lys Gly
35          40          45
Met Thr Pro Arg Glu Val Phe Glu Arg Asp Ile Arg Leu Leu Glu Glu
50          55          60
Ala Asp Val Leu Val Ala Glu Val Ser Tyr Pro Ser Leu Gly Val Gly
65          70          75          80
Phe Glu Ile Ala Tyr Phe Leu Leu Arg Gly Lys Pro Val Ile Ala Leu
85          90          95
Ala Leu Arg Glu Arg Leu Glu Ser Val Ser Ala Met Ile Arg Gly Ile
100         105         110
Thr Trp Glu Asn Phe Arg Leu Val Ala Tyr Ser Asp Val Asp Glu Ala
115         120         125
Ile Glu Lys Leu Asp Ser Met Leu Pro Gly Ser Val Asp Met Gln
130         135         140

```

```

<210> SEQ ID NO 7
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

```

```

<400> SEQUENCE: 7

```

```

catatgaagg tctacctggc g 21

```

```

<210> SEQ ID NO 8
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

```

```

<400> SEQUENCE: 8

```

```

gaattcttgc atgtoaacgc tacc 24

```

```

<210> SEQ ID NO 9
<211> LENGTH: 23
<212> TYPE: DNA

```

-continued

---

```

<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 9
caccatgaag gtctacctgg cgg                23

<210> SEQ ID NO 10
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 10
tcattgcatg tcaacgctac c                21

<210> SEQ ID NO 11
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 11
ttgcatgtca acgctacctg                20

<210> SEQ ID NO 12
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 12
ctcagatcat tgcattgcaa cg                22

<210> SEQ ID NO 13
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 13
ctcagattgc atgtcaacct ca                22

<210> SEQ ID NO 14
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 14
gaattctcat tgcattgcaa cgtc                24

<210> SEQ ID NO 15
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

```

---

-continued

---

<400> SEQUENCE: 15  
gaattcttgc atgtcaacgt cacc 24

<210> SEQ ID NO 16  
<211> LENGTH: 28  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 16  
ggaattccgc ttgcatgtca acgtacc 28

<210> SEQ ID NO 17  
<211> LENGTH: 27  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 17  
gccatgatgaa caagttgttt aatcaag 27

<210> SEQ ID NO 18  
<211> LENGTH: 28  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 18  
gcgaattcta ctggtatattt ccactata 28

<210> SEQ ID NO 19  
<211> LENGTH: 27  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Construct

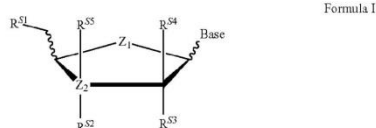
<400> SEQUENCE: 19  
catatgccaa aaaagacgat ctaettc 27

<210> SEQ ID NO 20  
<211> LENGTH: 21  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 20  
gtcgacttag tatacggcac c 21

---

1. A biocatalytic process for producing active pharmaceutical ingredients (APIs) or intermediates thereof, being those APIs or their intermediates, nucleoside analogues (NAs), D isomers, of formula I

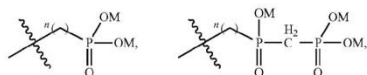


wherein,

$Z_1$  being O,  $\text{CH}_2$ , S, NH;

$Z_2(\text{R}^{S2})(\text{R}^{S5})$  being O, C( $\text{R}^{S2})(\text{R}^{S5})$ , S( $\text{R}^{S2})(\text{R}^{S5})$ , S( $\text{R}^{S5})$ , SO,  $\text{SO}_2$ , N( $\text{R}^{S2})(\text{R}^{S5})$ , N( $\text{R}^{S2}$ ), N( $\text{R}^{S5}$ ), independently of  $Z_1$ ;

$\text{R}^{S1}$  being hydrogen, methyl, OH, ether or ester thereof selected from:



being n is 0 or 1 and M is hydrogen or a pharmaceutically acceptable counter-ion such as sodium, potassium, ammonium or alkylammonium;

$\text{R}^{S2}$  being hydrogen, halogen, preferably F, OH or an ether or ester residue thereof, CN,  $\text{NH}_2$ , SH,  $\text{C}\equiv\text{CH}$ ,  $\text{N}_3$ ;

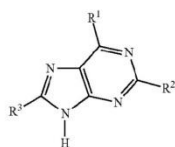
$\text{R}^{S3}$  being hydrogen, in case of NA derived from deoxyribonucleosides or being selected from: OH,  $\text{NH}_2$ , halogen (preferably F),  $\text{OCH}_3$ , when the NA is derived from ribonucleosides;

$\text{R}^{S4}$  being hydrogen, OH or an ether or ester residue thereof,  $\text{NH}_2$  or halogen, preferably F,

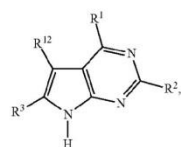
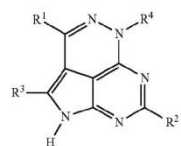
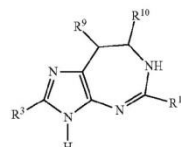
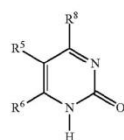
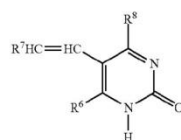
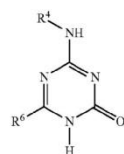
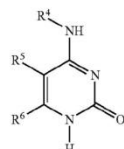
providing  $\text{R}^{S3}$  and  $\text{R}^{S4}$  are different when both were ethers or esters of OH;

$\text{R}^{S5}$  being hydrogen, OH or an ether or ester residue thereof,  $\text{NH}_2$  or halogen, preferably F;

Base bound to the ribose moiety being selected from those of formula A-H or their tautomers and regioisomers thereof:



-continued



wherein,

R<sup>1</sup> being hydrogen, methyl, optionally substituted alkyl chain, halogen, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, SR<sup>13</sup>;

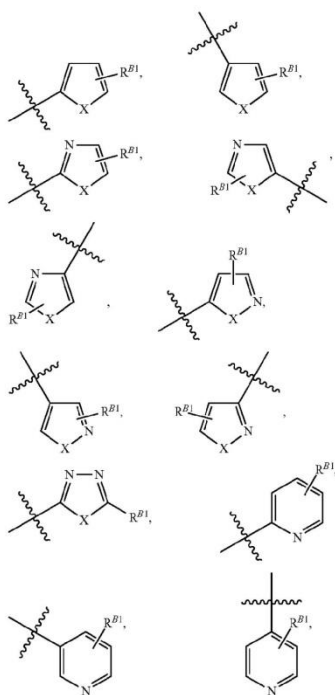
R<sup>2</sup> being hydrogen, methyl, optionally substituted alkyl chain, halogen, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, SR<sup>13</sup>;

R<sup>3</sup> being hydrogen, methyl, optionally substituted alkyl chain, halogen, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, SR<sup>13</sup>;

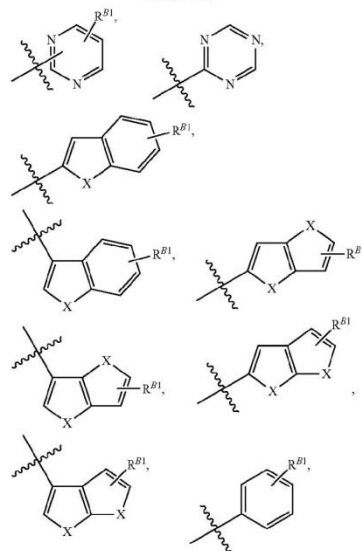
R<sup>4</sup> being hydrogen, methyl, optionally substituted alkyl chain, COR<sup>13</sup>, CONR<sup>13</sup>R<sup>14</sup>, CO<sub>2</sub>R<sup>13</sup>, C(S)OR<sup>13</sup>, CH<sub>2</sub>-heterocyclic ring;

R<sup>5</sup> being hydrogen, methyl, an optionally substituted alkyl chain, an optionally substituted alkenyl chain, an optionally substituted alkynyl chain, halogen, trihaloalkyl, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, CN, COR<sup>13</sup>, CONR<sup>13</sup>R<sup>14</sup>, CO<sub>2</sub>R<sup>13</sup>, C(S)OR<sup>13</sup>, OCONR<sup>13</sup>R<sup>14</sup>, OCO<sub>2</sub>R<sup>13</sup>, OC(S)OR<sup>13</sup>, NHCONR<sup>13</sup>R<sup>14</sup>, NHCO<sub>2</sub>R<sup>13</sup>, NHC(S)OR<sup>13</sup>, SO<sub>2</sub>NR<sup>13</sup>R<sup>14</sup>;

and any optionally substituted heterocycle or optionally substituted aryl of, independently, R<sup>1</sup>, R<sup>2</sup>, R<sup>3</sup>, R<sup>4</sup> or R<sup>5</sup>, selected from:



-continued



wherein

X being O, S, N—R<sup>B2</sup>, Se; R<sup>B1</sup> being H, straight or branched C<sub>1-10</sub> alkyl, F, Cl, Br, I, X—R<sup>B2</sup>, —C≡C—R<sup>B2</sup>, CO<sub>2</sub>R<sup>B2</sup>; R<sup>B2</sup> being H, straight or branched C<sub>1-5</sub> alkyl, phenyl;

R<sup>6</sup> being hydrogen, optionally substituted alkyl chain;

R<sup>7</sup> being hydrogen, halogen, trihaloalkyl, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, CN, COR<sup>13</sup>, CONR<sup>13</sup>R<sup>14</sup>, CO<sub>2</sub>R<sup>13</sup>, C(S)OR<sup>13</sup>;

R<sup>8</sup> being hydroxyl or amino, OR<sup>13</sup>, OSO<sub>2</sub>R<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, CN, COR<sup>13</sup>, CONR<sup>13</sup>R<sup>14</sup>, CO<sub>2</sub>R<sup>13</sup>, C(S)OR<sup>13</sup>;

R<sup>9</sup> being hydrogen, methyl, optionally substituted alkyl chain, halogen, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, COR<sup>13</sup>, CONR<sup>13</sup>R<sup>14</sup>, CN, CO<sub>2</sub>R<sup>13</sup>, C(S)OR<sup>13</sup>, OCONR<sup>13</sup>R<sup>14</sup>, OCO<sub>2</sub>R<sup>13</sup>, OC(S)OR<sup>13</sup>, NHCONR<sup>13</sup>R<sup>14</sup>, NHCO<sub>2</sub>R<sup>13</sup>, NHC(S)OR<sup>13</sup>, OR<sup>13</sup>;

R<sup>10</sup> and R<sup>11</sup> are independently of each other hydrogen, methyl, optionally substituted alkyl chain;

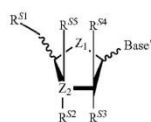
R<sup>12</sup> being hydrogen, methyl, optionally substituted alkyl chain, halogen, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, CN, COR<sup>13</sup>, CONR<sup>13</sup>R<sup>14</sup>, CO<sub>2</sub>R<sup>13</sup>, C(S)OR<sup>13</sup>, OCONR<sup>13</sup>R<sup>14</sup>, OCO<sub>2</sub>R<sup>13</sup>, OC(S)OR<sup>13</sup>, NHCONR<sup>13</sup>R<sup>14</sup>, NHCO<sub>2</sub>R<sup>13</sup>, NHC(S)OR<sup>13</sup>;

R<sup>13</sup> and R<sup>14</sup> are independently of each other hydrogen, substituted or unsubstituted alkyl chain, substituted or unsubstituted alkenyl chain, substituted or unsubstituted alkynyl chain, substituted or unsubstituted phenyl ring, substituted or unsubstituted heterocyclic ring;

the process being carried out, preferably, in a one-step/one-pot reaction, wherein said reaction comprises the addition, in a suitable reaction aqueous medium and under suitable reaction conditions, of an enzyme having



nucleoside deoxyribosyltransferase (NDT) activity, either native or recombinant to a mixture of starting materials comprising at least a nucleoside, D isomer, of formula II



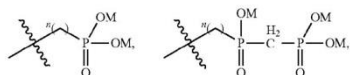
Formula II

wherein,

Z<sub>1</sub> being O, CH<sub>2</sub>, S, NH;

Z<sub>2</sub>(R<sup>S2</sup>)(R<sup>S5</sup>) being O, C(R<sup>S2</sup>)(R<sup>S5</sup>), S(R<sup>S2</sup>)(R<sup>S5</sup>), S(R<sup>S2</sup>), S(R<sup>S5</sup>), SO, SO<sub>2</sub>, N(R<sup>S2</sup>)(R<sup>S5</sup>), N(R<sup>S2</sup>), N(R<sup>S5</sup>), independently of Z<sub>1</sub>;

R<sup>S1</sup> being hydrogen, methyl, OH, ether or ester thereof selected from:



being n is 0 or 1 and M is hydrogen or a pharmaceutically acceptable counter-ion such as sodium, potassium, ammonium or alkylammonium;

R<sup>S2</sup> being hydrogen, halogen, preferably F, OH or an ether or ester residue thereof, CN, NH<sub>2</sub>, SH, C=CH, N<sub>3</sub>;

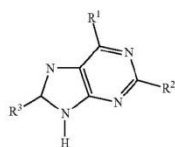
R<sup>S3</sup> being hydrogen, in case of NA derived from desoxyribonucleosides or being selected from: OH, NH<sub>2</sub>, halogen (preferably F), OCH<sub>3</sub>, when the NA is derived from ribonucleosides;

R<sup>S4</sup> being hydrogen, OH or an ether or ester residue thereof, NH<sub>2</sub> or halogen, preferably F, providing R<sup>S1</sup> and R<sup>S4</sup> are different when both were ethers or esters of OH residues;

R<sup>S5</sup> being hydrogen, OH or an ether or ester residue thereof, NH<sub>2</sub> or halogen, preferably F;

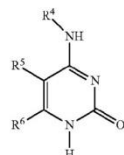
Base' bound to the ribose moiety, being selected from any heterocyclic ring

and at least a free nucleobase, to be transferred by the enzyme having NDT activity, being selected from those of formula A'-H' or their tautomers and regioisomers thereof:

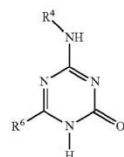


A'

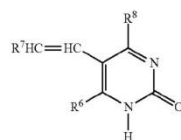
-continued



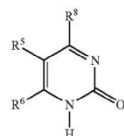
B'



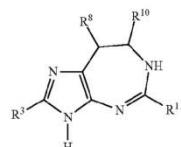
C'



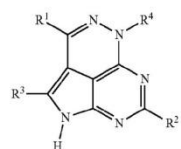
D'



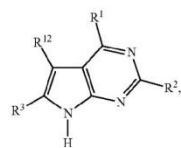
E'



F'



G'



H'

wherein,

R<sup>1</sup> being hydrogen, methyl, optionally substituted alkyl chain, halogen, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, SR<sup>13</sup>;

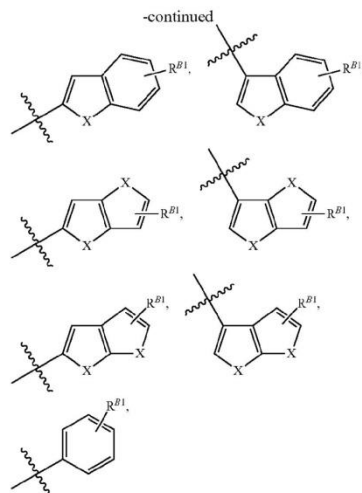
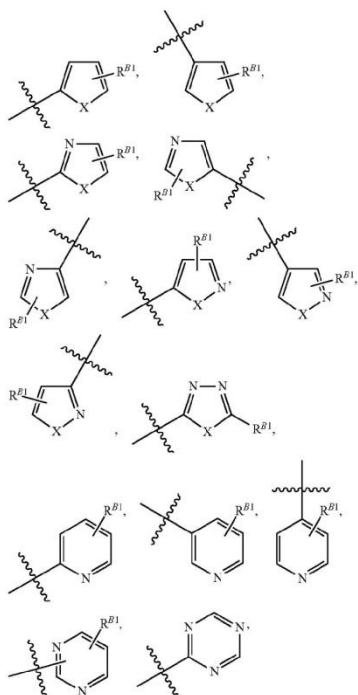
R<sup>2</sup> being hydrogen, methyl, optionally substituted alkyl chain, halogen, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, SR<sup>13</sup>;

R<sup>3</sup> being hydrogen, methyl, optionally substituted alkyl chain, halogen, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, SR<sup>13</sup>;

R<sup>4</sup> being hydrogen, methyl, optionally substituted alkyl chain, COR<sup>13</sup>, CONR<sup>13</sup>R<sup>14</sup>, CO<sub>2</sub>R<sup>13</sup>, C(S)OR<sup>13</sup>, CH<sub>2</sub>-heterocyclic ring;

R<sup>5</sup> being hydrogen, methyl, an optionally substituted alkyl chain, an optionally substituted alkenyl chain, an optionally substituted alkynyl chain, halogen, trihaloalkyl, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, CN, COR<sup>13</sup>, CONR<sup>13</sup>R<sup>14</sup>, CO<sub>2</sub>R<sup>13</sup>, C(S)OR<sup>13</sup>, OCONR<sup>13</sup>R<sup>14</sup>, OCO<sub>2</sub>R<sup>13</sup>, OC(S)OR<sup>13</sup>, NHCONR<sup>13</sup>R<sup>14</sup>, NHCO<sub>2</sub>R<sup>13</sup>, NHC(S)OR<sup>13</sup>, SO<sub>2</sub>NR<sup>13</sup>R<sup>14</sup>;

and any optionally substituted heterocycle or optionally substituted aryl of, independently, R<sup>1</sup>, R<sup>2</sup>, R<sup>3</sup>, R<sup>4</sup> or R<sup>5</sup>, selected from:



wherein

X being O, S, N—R<sup>B2</sup>, Se; R<sup>B1</sup> being H, straight or branched C<sub>1-10</sub> alkyl, F, Cl, Br, I, X—R<sup>B2</sup>, —C≡C—R<sup>B2</sup>, CO<sub>2</sub>R<sup>B2</sup>, R<sup>B2</sup> being H, straight or branched C<sub>1-5</sub> alkyl, phenyl;

R<sup>6</sup> being hydrogen, optionally substituted alkyl chain;

R<sup>7</sup> being hydrogen, halogen, trihaloalkyl, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, CN, COR<sup>13</sup>, CONR<sup>13</sup>R<sup>14</sup>, CO<sub>2</sub>R<sup>13</sup>, C(S)OR<sup>13</sup>;

R<sup>8</sup> being hydroxyl or amino, OR<sup>13</sup>, OSO<sub>2</sub>R<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, CN, COR<sup>13</sup>, CONR<sup>13</sup>R<sup>14</sup>, CO<sub>2</sub>R<sup>13</sup>, C(S)OR<sup>13</sup>;

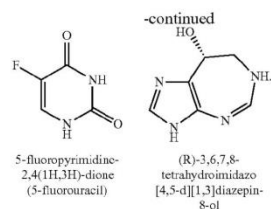
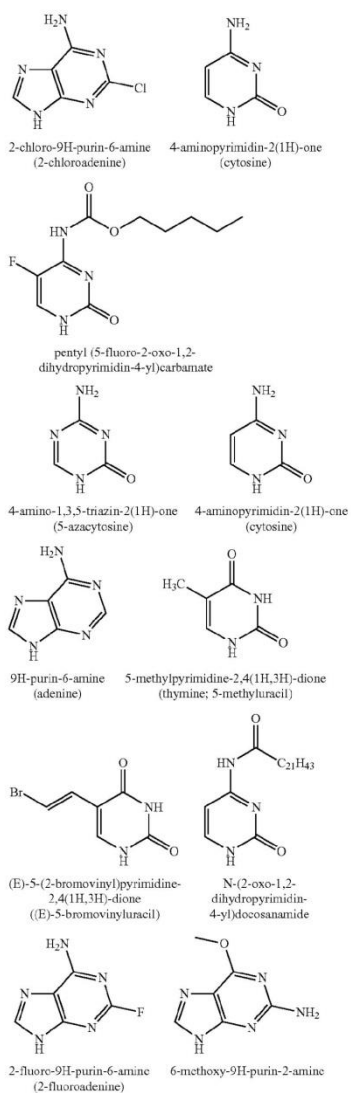
R<sup>9</sup> being hydrogen, methyl, optionally substituted alkyl chain, halogen, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, COR<sup>13</sup>, CONR<sup>13</sup>R<sup>14</sup>, CN, CO<sub>2</sub>R<sup>13</sup>, C(S)OR<sup>13</sup>, OCONR<sup>13</sup>R<sup>14</sup>, OCO<sub>2</sub>R<sup>13</sup>, OC(S)OR<sup>13</sup>, NHCONR<sup>13</sup>R<sup>14</sup>, NHCO<sub>2</sub>R<sup>13</sup>, NHC(S)OR<sup>13</sup>;

R<sup>10</sup> and R<sup>11</sup> are independently of each other hydrogen, methyl, optionally substituted alkyl chain;

R<sup>12</sup> being hydrogen, methyl, optionally substituted alkyl chain, halogen, OR<sup>13</sup>, NR<sup>13</sup>R<sup>14</sup>, CN, COR<sup>13</sup>, CONR<sup>13</sup>R<sup>14</sup>, CO<sub>2</sub>R<sup>13</sup>, C(S)OR<sup>13</sup>, OCONR<sup>13</sup>R<sup>14</sup>, OCO<sub>2</sub>R<sup>13</sup>, OC(S)OR<sup>13</sup>, NHCONR<sup>13</sup>R<sup>14</sup>, NHCO<sub>2</sub>R<sup>13</sup>, NHC(S)OR<sup>13</sup>;

R<sup>13</sup> and R<sup>14</sup> are independently of each other hydrogen, substituted or unsubstituted alkyl chain, substituted or unsubstituted alkenyl chain, substituted or unsubstituted alkynyl chain, substituted or unsubstituted phenyl ring, substituted or unsubstituted heterocyclic ring.

2. A process according to claim 1, wherein the free nucleobase to be transferred by the enzyme having NDT activity is selected from:



3. The process according to claim 1 wherein the APIs or intermediates thereof produced are selected from: Clofarabine (Cl-F-araA), Decitabine (aza-dCyd), Cytarabine (ara-C), Vidarabine, Brivudine, Enocitabine (BH-AC), Zalcitabine (ddC), Cladribine (Cl-dAdo), Fludarabine (F-araA), Nelarabine (MAY), Zidovudine, Floxuridine (FUDR),  $\beta$ -Thymidine, idoxuridine (IdU), trifluridine (TFT), acedurid (EdU), ribavirin, didanosine (ddl) and Pentostatin.

4. The process according to claim 3 wherein the APIs or intermediates thereof produced are selected from: Clofarabine, Cytarabine, Vidarabine, Brivudine, Enocitabine, Zalcitabine, Cladribine, Fludarabine, Nelarabine, Zidovudine, Floxuridine and Pentostatin.

5. The process according to claim 3 wherein the APIs or intermediates thereof produced are selected from Clofarabine, Cytarabine and Zidovudine.

6. The process according to claim 1 wherein the API, or intermediates thereof, produced is Clofarabine.

7. The process according to any of claims 1 to 6 wherein the enzyme having NDT activity is a NDT isolated from *Lactobacillus delbrueckii* (formerly called *Lactobacillus leichmannii*).

8. The process according to any of claims 1 to 6 wherein the enzyme having NDT activity is a NDT isolated from *Lactococcus lactis*.

9. The process according to claim 7 wherein the NDT is obtained from

- Lactobacillus delbrueckii* (formerly called *Lactobacillus leichmannii*) nucleotide sequence shown in SEQ ID NO. 1; or
- a nucleotide sequence which is the complement of SEQ ID. NO.1; or
- a nucleotide sequence which is degenerate with SEQ ID. NO.1; or
- a nucleotide sequence hybridizing under conditions of high stringency to SEQ ID. NO. 1, to the complement of SEQ ID. NO.1, or to a hybridization probe derived from SEQ ID. NO.1, or their complement thereof; or
- a nucleotide sequence having at least 80% sequence identity with SEQ ID. NO. 1; or
- a nucleotide sequence having at least 65% sequence identity with SEQ ID. NO. 1, wherein said sequence preferably encodes or is complementary to a sequence encoding at least a NDT enzyme or a functional part thereof;
- a nucleotide sequence encoding for a polypeptide having NDT activity, the amino acid sequence of which is at least 80% identical to the amino acid sequence shown in SEQ ID. NO.2.

10. The process according to claim 8 wherein a NDT is obtained from

- a) *Lactococcus lactis* nucleotide sequence shown in SEQ ID NO. 3 or
- b) a nucleotide sequence which is the complement of SEQ ID NO.3; or
- c) a nucleotide sequence which is degenerate with SEQ ID NO.3; or
- d) a nucleotide sequence hybridizing under conditions of high stringency to SEQ ID NO.3, to the complement of SEQ ID NO.3, or to a hybridization probe derived from SEQ ID NO.3, or their complement thereof; or
- e) a nucleotide sequence having at least 80% sequence identity with SEQ ID NO.3; or
- f) a nucleotide sequence having at least 65% sequence identity with SEQ ID NO.3, wherein said sequence preferably encodes or is complementary to a sequence encoding at least a NDT enzyme or a functional part thereof;
- g) a nucleotide sequence encoding for a polypeptide having NDT activity, the amino acid sequence of which is at least 80% identical to the amino acid sequence shown in SEQ ID NO.4.

**11.** Use of a recombinant enzyme having nucleoside deoxyribosyltransferase activity in the production of APIs or intermediates thereof according to the process of any of claims **1** to **10**, being those APIs or their intermediates, nucleoside analogues (NAs) particularly useful as anti-cancer or anti-viral medicaments, wherein said recombinant enzyme is characterized by being encoded by a nucleic acid sequence comprising a sequence selected from: SEQ ID No. 1 or SEQ ID No. 3;

- a) a nucleotide sequence which is the complement of SEQ ID NO:1 or SEQ. ID NO:3; or
- b) a nucleotide sequence which is degenerate with SEQ ID NO:1 or SEQ. ID NO:3; or
- c) a nucleotide sequence hybridizing under conditions of high stringency to SEQ ID NO:1 or SEQ. ID NO:3, to the complements of SEQ ID NO:1 or SEQ. ID NO:3, or to a hybridization probe derived from SEQ ID NO:1 or SEQ. ID NO:3, or their complement thereof; or
- d) a nucleotide sequence having at least 80% sequence identity with SEQ ID NO:1 or SEQ. ID NO:3; or
- e) a nucleotide sequence having at least 65% sequence identity with SEQ ID NO:1 or SEQ. ID NO:3, wherein

said sequence preferably encodes or is complementary to a sequence encoding at least a NDT enzyme or a functional part thereof;

- f) a nucleotide sequence encoding for an amino acid sequence selected from: SEQ ID NO:2 or SEQ. ID NO:4.

**12.** A recombinant expression vector comprising the sequence encoding a natural or recombinant nucleoside deoxyribosyltransferase (NDT) operably linked to one or more control sequences that direct the expression or overexpression of said deoxyribosyltransferase in a suitable host, wherein said recombinant expression vector, carries and expresses or overexpresses a nucleic acid sequence selected from: SEQ ID No. 1 or SEQ ID No. 3;

- a) a nucleotide sequence which is the complement of SEQ ID NO:1 or SEQ. ID NO:3; or
- b) a nucleotide sequence which is degenerate with SEQ ID NO:1 or SEQ. ID NO:3; or
- c) a nucleotide sequence hybridizing under conditions of high stringency to SEQ ID NO:1 or SEQ. ID NO:3, to the complements of SEQ ID NO:1 or SEQ. ID NO:3, or to a hybridization probe derived from SEQ ID NO:1 or SEQ. ID NO:3, or their complement thereof; or
- d) a nucleotide sequence having at least 80% sequence identity with SEQ ID NO:1 or SEQ. ID NO:3; or
- e) a nucleotide sequence having at least 65% sequence identity with SEQ ID NO:1 or SEQ. ID NO:3, wherein said sequence preferably encodes or is complementary to a sequence encoding at least a NDT enzyme or a functional part thereof;
- f) a nucleotide sequence encoding for an amino acid sequence selected from: SEQ ID NO:2 or SEQ. ID NO:4.

**13.** Use of a recombinant expression vector, according to claim **12**, for the production of active pharmaceutical ingredients (APIs) or intermediates thereof, being those APIs or their intermediates, nucleoside analogues (NAs) particularly useful as anti-cancer or anti-viral medicaments.

**14.** A host cell comprising the recombinant expression vector of claim **12**.

**15.** Use of a host cell according to claim **14** for the production of recombinant enzymes having NDT activity.

\* \* \* \* \*



US 20160076070A1

(19) **United States**

(12) **Patent Application Publication**  
**PASCUAL GILABERT et al.**

(10) **Pub. No.: US 2016/0076070 A1**  
(43) **Pub. Date: Mar. 17, 2016**

(54) **BIOCATALYTIC PRODUCTION OF NUCLEOSIDE ANALOGUES AS ACTIVE PHARMACEUTICAL INGREDIENTS**

(71) Applicant: **PLASMIA BIOTECH, S.L.**, Barcelona (ES)

(72) Inventors: **Marta PASCUAL GILABERT**, Barcelona (ES); **Victor Manuel DERONCELÉ THOMAS**, Barcelona (ES); **Rafael MONTILLA ARÉVALO**, Barcelona (ES)

(21) Appl. No.: **14/787,678**

(22) PCT Filed: **Apr. 29, 2014**

(86) PCT No.: **PCT/EP2014/058761**

§ 371 (c)(1),

(2) Date: **Oct. 28, 2015**

**Related U.S. Application Data**

(60) Provisional application No. 61/817,064, filed on Apr. 29, 2013.

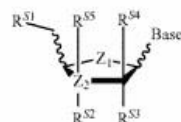
**Publication Classification**

(51) **Int. Cl.**  
**C12P 19/38** (2006.01)  
**C12P 19/40** (2006.01)

(52) **U.S. Cl.**  
CPC ..... **C12P 19/385** (2013.01); **C12P 19/40** (2013.01)

(57) **ABSTRACT**

A biocatalytic process for producing active pharmaceutical ingredients (APIs) or intermediates thereof, wherein those APIs or their intermediates are nucleoside analogues (NAs) of formula I



Formula I

and wherein said NAs are active as pharmaceutically relevant antivirals and anticancer medicaments, intermediates or pro-drugs thereof.