



Rational Design and Synthesis of New Nucleoside Analogues Bearing a Cyclohexane Core

Beatriz Domínguez Pérez

Ph.D. Thesis

Ph.D. in Chemistry

Supervisors:

Dr. Ramon Alibés Arqués

Dr. Félix Busqué Sánchez

Dr. Jean-Didier Márechal

Departament de Química

Facultat de Ciències

2015

Memòria presentada per aspirar al Grau de
Doctor per Beatriz Domínguez Pérez

Beatriz Domínguez Pérez

Vist i plau,

Dr. Ramon Alibés Arqués

Dr. Félix Busqué Sánchez

Dr. Jean-Didier Maréchal

Bellaterra, 13 de maig de 2015

*"When you make the finding yourself-even if you're the last person on Earth to see the
light- you'll never forget it"*

Carl Sagan

A mis padres, a mi hermana

A Jonatan

Table of contents

Abbreviations.....	1
I. General introduction	5
1. Viruses.....	7
1.1. Virus replication.....	8
1.2. Viral diseases in humans.....	9
2. Antiviral drugs.....	10
2.1. Nucleoside analogues.....	11
2.2. Carbocyclic nucleoside analogues	14
2.3. Cyclohexanyl and cyclohexenyl nucleoside analogues.....	17
2.4. Precedents.....	19
3. <i>In silico</i> molecular modelling for drug design	25
3.1. Protein-ligand docking in medicinal chemistry.....	26
3.2. Protein-ligand docking studies on nucleoside analogues.....	26
4. References.....	28
II. Objectives	35
III. Molecular modelling pre-study for the design of anti-HSV novel carbocyclic nucleosides	41
1. Introduction	43
1.1. Molecular modelling of biological macromolecules.....	45
1.1.1. Protein-ligand docking.....	46
1.1.1.1. Limitations	47
1.1.1.2. Protein-ligand docking: inhibition versus activation	49
2. Molecular modelling pre-study of novel carbocyclic nucleoside analogues as anti- HSV	50
2.1. Computational details.....	51
2.2. Activation process.....	53
2.2.1. 1 st phosphorylation step.....	55
2.2.1.1. HSV-1 TK crystallographic structures	55
2.2.1.2. Docking results	57
2.2.2. 2 nd phosphorylation step	64
2.2.2.1. GMPK crystallographic structures	65
2.2.2.2. Docking results	66
2.2.3. 3 rd phosphorylation step.....	71

2.2.3.1. NDPK crystallographic structure.....	72
2.2.3.2. Docking results	74
2.3. Exploring other scaffolds	79
2.3.1. Activation process.....	80
3. Conclusions	83
4. References	84

IV. Synthesis of bicyclo[4.1.0]heptane nucleoside analogues..... 89

1. Introduction	91
2. Synthesis of bicyclo[4.1.0]heptane nucleoside analogues	92
2.1. Synthetic strategy	92
2.2. Diastereoselective synthesis of allylic alcohol 8	93
2.2.1. Preparation of monoketal 11	94
2.2.2. Preparation of enone 12	95
2.2.3. Preparation of allylic alcohol 8	97
2.3. Preparation of 4-hydroxy-2-cyclohexenone, 29	101
2.4. Enantioselective synthesis of both enantiomers of cyclohexenone 29 and their derivatives.....	105
2.5. Asymmetric Transfer Hydrogenation (ATH) Study	110
2.6. Synthesis of the alcohol 16	113
2.6.1. Cyclopropanation reaction	114
2.6.2. Protection of the hydroxyl group of 13	117
2.6.3. Hydrolysis of the ketal	118
2.6.4. Preparation of alcohol 14	119
2.6.5. Preparation of alcohol 16	122
2.7. Introduction of the base moiety	123
2.7.1. Nucleophilic substitution S _N 2	123
2.7.2. Mitsunobu reaction	124
2.7.3. Construction of the base moiety	127
2.7.3.1. Synthesis of purine nucleosides	129
2.7.3.1.1. Microwave radiation in organic chemistry	131
2.7.3.1.2. Synthesis of adenine nucleoside analogue D- 2-A	132
2.7.3.1.3. Synthesis of guanine nucleoside analogue D- 2-G	134
2.7.3.2. Synthesis of pyrimidine nucleosides	135
2.7.3.2.1. Synthesis of uracil nucleoside analogue D- 2-U	136
2.7.3.2.2. Synthesis of thymine nucleoside analogue D- 2-T	137
2.8. Summary	140
3. References	141

V. Study of antiviral activity of prodrug candidates	147
1. Evaluation of the antiviral activity	149
2. Rationalisation of the antiviral activity	152
2.1. Interaction with DNA polymerase	153
2.1.1. Homology modelling.....	154
2.1.2. Generation of a model on an holo conformation of DNA polymerase	155
2.1.3. Computational details	156
2.1.4. Docking results.....	157
3. Conclusions	163
4. References	164
VI. General conclusions	167
VII. Experimental section	173
1. Molecular modelling pre-study for the design of anti-HSV novel carbocyclic nucleosides	175
2. General procedures	178
3. Synthesis of bicyclo[4.1.0]heptane nucleoside analogues	180
4. Study of antiviral activity of prodrug candidates.....	218
5. References	219
VIII. Computational methods	221
1. Molecular modelling	223
1.1 Molecular mechanics.....	223
2. Protein–ligand docking	226
2.1. Search algorithms	226
2.2. Scoring functions	230
3. References	232
Formula index	235
NMR spectra of selected compounds	241

Appendixes — CD

Appendix A. NMR spectra

Appendix B. Figures not included in the manuscript

Abbreviations

A	adenine	DBAD	di- <i>tert</i> -butyl azodicarboxylate
ACN	acetonitrile	DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
AcOH	acetic acid	DCG	D-Cyclohexenyl-G
Ac ₂ O	acetic anhydride	de	diastereomeric excess
ACV	acyclovir, 9-[(2-hydroxyethoxy)methyl]guanine	DEAD	diethyl azodicarboxylate
ACVDP	acyclovir diphosphate	DEPT	Distortionless enhancement by polarisation transfer
ACVMP	acyclovir monophosphate	DIAD	diisopropyl azodicarboxylate
ACVTP	acyclovir triphosphate	DMAP	4-dimethylaminopyridine
ADA	adenosine deaminase	DME	dimethyl ether
ADP	adenosine-5'-diphosphate	DMF	dimethylformamide
AIBN	azobis(isobutyronitrile)	DMSO	dimethylsulfoxide
ATH	Asymmetric Transfer Hydrogenation	DNA	deoxyribonucleic acid
ATP	adenosine-5'-triphosphate	DPPA	diphenylphosphoryl azide
ATR	Attenuated Total Reflectance	dppe	1,2-bis(diphenylphosphino)ethane
AZT	zidovudine, 3'-azido-3'-deoxythymidine	dT	deoxythymidine
9-BBN	9-borabicyclo[3.3.1]nonane	dTDP	deoxythymidine-5'-diphosphate
b.p.	boiling point	dTMP	deoxythymidine-5'-monophosphate
Bz	benzoyl	dTTP	deoxythymidine-5'-triphosphate
BzCl	Benzoyl chloride	EC	Enzyme Commission
C	cytosine	EC50	50% effective concentration
18-C-6	18-crown-6	ee	enantiomeric excess
CAN	ceric ammonium nitrate	EMA	European Medicines Agency
CC50	50% cytotoxic concentration	ESI	Electrospray Ionization
CDI	1,1'-carbonyldiimidazole	Et ₂ O	diethyl ether
CHPLC	chiral high-pressure liquid chromatography	Et ₃ N	triethylamine
CMV	cytomegalovirus	EtOAc	ethyl acetate
COSY	Correlation spectroscopy	EtOH	ethanol
CRIs	Coreceptor inhibitors	Et ₂ Zn	diethylzinc
CTAB	cetyltrimethylammonium bromide	FDA	US Food and Drug Administration
		ff	force field
		FP	final product

Abbreviations

FIs	Fusion inhibitors	2-Me-CBS	1-Methyl-3,3-diphenyltetrahydro-3H-pyrrolo[1,2-c][1,3,2]oxazaborole
G	guanine	MeNH ₂	methylamine
GA	Genetic Algorithm	MeOH	methanol
GDP	guanosine-5'-diphosphate	mGMPK	mouse guanylate kinase
GMP	guanosine-5'-monophosphate	MM	Molecular Mechanics
GMPK	guanylate kinase	MMFF	Merck molecular force field
GTP	guanosine-5'-triphosphate	m.p.	melting point
HBV	Hepatitis B virus	MPO	4-methoxypyridine- <i>N</i> -oxide
HCMV	Human cytomegalovirus	mRNA	messenger ribonucleic acid
HCV	hepatitis C virus	Ms	mesyl
HEL	Human embryonic lung	MsCl	mesyl chloride
HIV	human immunodeficiency virus	MW	microwave heating
HMBC	Heteronuclear multiple bond correlation	NaOMe	sodium methoxide
HRMS	High-Resolution Mass Spectra	NDPK	nucleoside diphosphate kinase
HSQC	Heteronuclear single quantum coherence	NMPK	nucleoside monophosphate kinase
HSV	herpes simplex virus	NMR	Nuclear Magnetic Resonance
IBA	<i>o</i> -iodosobenzoic acid	NNRTIs	Non-Nucleoside Reverse Transcriptase Inhibitors
IBX	<i>o</i> -iodoxybenzoic acid	NOESY	Nuclear Overhauser Effect Spectroscopy
IC	Incremental Construction	NRTIs	Nucleoside Reverse Transcriptase Inhibitors
INIs	integrase inhibitors	P	octanol-water partition ratio
In(OTf) ₃	indium(III) trifluoromethanesulfonate	PDB	Protein Data Bank
<i>i</i> -Pr ₂ NEt	<i>N,N</i> -diisopropylethylamine	Ph	phenyl
IR	Infrared spectroscopy	PIs	protease inhibitors
KHMDS	potassium bis(trimethylsilyl)amide	PLP	Piecewise Linear Potential
LCG	L-Cyclohexenyl-G	PMBCl	<i>p</i> -methoxybenzyl chloride
LDA	lithium diisopropylamide	PPTS	pyridinium <i>p</i> -toluenesulfonate
M	molar	<i>p</i> -TsOH	<i>p</i> -toluenesulfonic acid
MC	Monte Carlo	py	pyridine
<i>m</i> -CPBA	<i>m</i> -chloroperbenzoic acid	QM	quantum mechanics
MCT	methanocarbothymidine	QSAR	quantitative structure-activity relationship
MD	Molecular Dynamics		
Me	Methyl		

RMS	Root mean square
RNA	ribonucleic acid
RT	Reverse Transcriptase
rt	room temperature
(S)-MCT	3'- <i>exo</i> -methanocarbothymidine
SO	Swarm Optimization
T	thymine
TBACl	tetrabutylammonium chloride
TBAF	tetrabutylammonium fluoride
TBDPS	<i>tert</i> -butyldiphenylsilyl
TBDPSCI	<i>tert</i> -butyldiphenylsilyl chloride
TBS	<i>tert</i> -butyldimethylsilyl
TBSCI	<i>tert</i> -butyldimethylsilyl chloride
^t Bu	<i>tert</i> -butyl
^t BuOK	potassium <i>tert</i> -butoxide
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TK	thymidine kinase
TLC	Thin Layer Chromatography
TMSCl	Trimethylsilyl chloride
TMSOTf	trimethylsilyl trifluoromethanesulfonate
U	uracil
VZV	varicella-zoster virus
WDI	World Drug Index

Chapter I: General introduction

1. Viruses

Nowadays nearly everybody knows what a virus is. Etymologically, 'virus' comes from the Latin word *virus* which means poison or toxin. In other words, viruses may be defined as parasites that infect host cells causing, in some cases, diseases to most of living organisms, from archaea and bacteria to animals and plants. Since the discovery of the tobacco mosaic virus by the Dutch microbiologist Martinus Beijerinck in 1898, over 5000 species of viruses have been described.¹

Biologically, viruses have also been defined as submicroscopic obligate intracellular parasites that require the biological machinery of the host cell to replicate and spread.² This dependence on the cell machinery to survive implies that viruses can be found both inside and outside the cells. The virus particle, also known as a virion, is metabolically inert outside the cells. For that reason it is only inside the cells where its replication takes place.³

The general structure of a virus particle contains its genetic material made from either DNA or RNA, but never both, surrounded by a protein coat called capsid (Figure I-1). This capsid, whose role is to protect the genetic material, is made from proteins, known as capsomeres, encoded by the viral genome. The complex formed by the capsid and the virus genetic material is called nucleocapsid. Some virions can also have virus-specific enzymes needed during the infection and replication processes. Furthermore, some viruses possess an envelope of lipids surrounding the nucleocapsid. This viral membrane, which is constructed from the plasma membrane of the infected cell when the virus particle is released, facilitates the entrance into future host cells.⁴

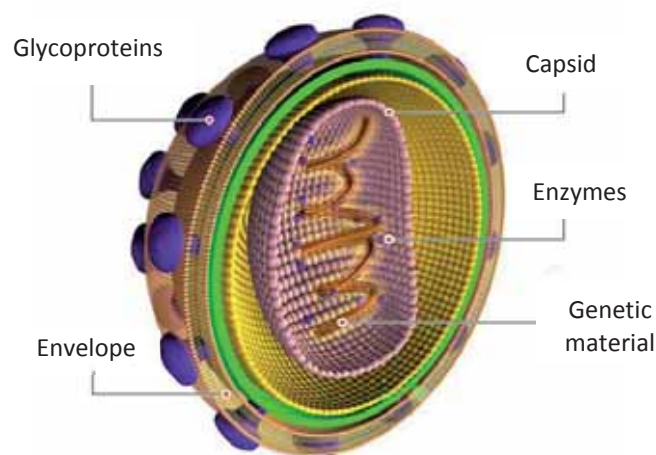


Figure I-1. General structure of viruses.

1.1. Virus replication

As previously mentioned, viruses are dependent upon host cells for their reproduction. During the replication process, a virus induces a living host cell to synthesise the essential components for the synthesis of new viral particles. This process can be divided into six general stages (Figure I-2):

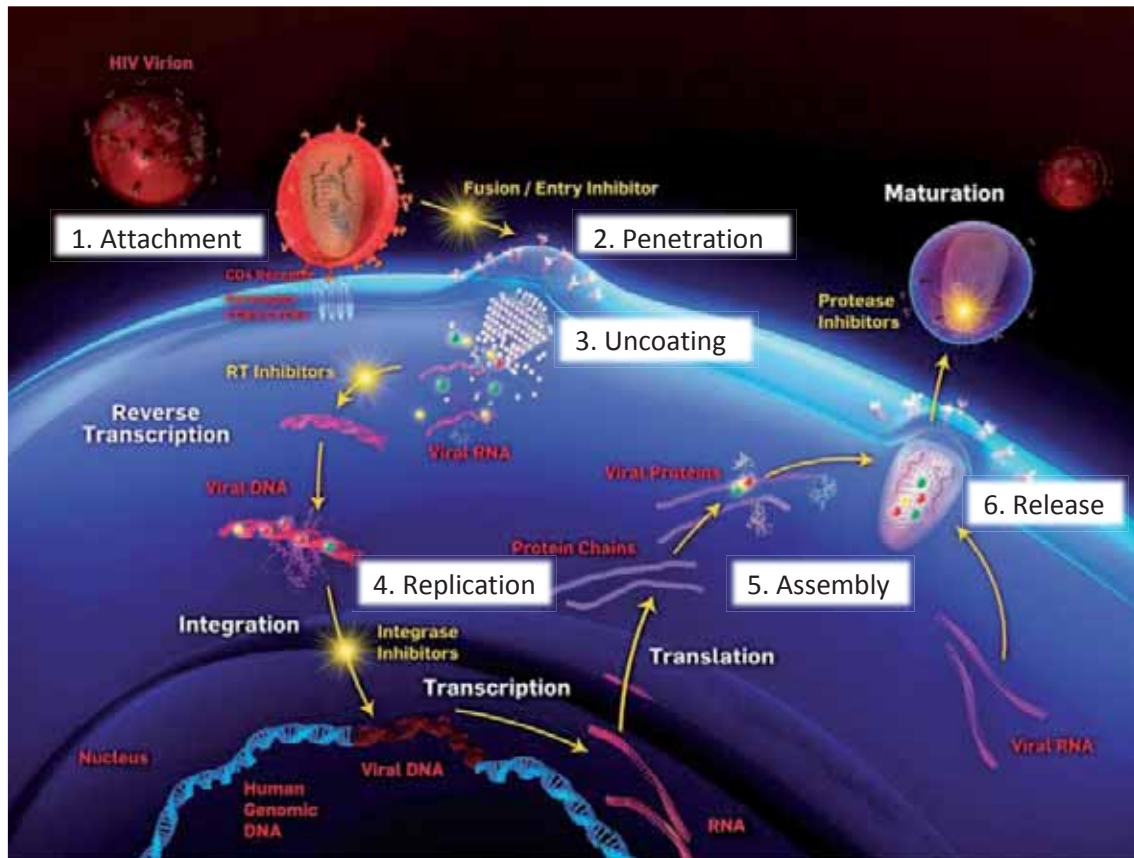


Figure I-2. Viral life cycle.

- 1) *Attachment*. It consists of the recognition of the virion through a binding of a virus-attachment protein to a specific receptor on the host cell membrane. This specificity restricts the virus infectious capabilities to a very limited type of cells or tissues. This step is critical in the viral replication cycle and constitutes a great target for antiviral therapies developed to prevent viral infections. If virus attachment could be blocked, the infection would be prevented.
- 2) *Penetration* of the virion into the target cell. This process may take place by translocation, endocytosis or by fusion of the virus and host cell lipid membranes.
- 3) *Uncoating*. The capsid is completely or partially removed to let the genetic material free for its replication. Unfortunately, this is one of the less studied stages of the replication and therefore is relatively poorly understood.

- 4) *Replication*. This process is developed in three principal steps: the formation of viral mRNAs, the translation into the viral proteins and the replication of the virus genome. Clearly, the nature of the virus genetic material, which can be either RNA or DNA, determines how the viral mRNAs are formed.
- 5) *Assembly*. The new viral genetic material is then assembled into new virions. During assembly, the basic structure of the virus particle is formed. This process may take place either in the cytoplasm or in the nucleus, depending on the site of replication within the cell and the mechanism by which the virus is eventually released from the cell.
- 6) *Release*. The new virions are released from the cell by different ways. Virions without envelope are released by lysis while those viruses with envelope are released by emerging from the cell surface and acquiring their outer lipid shell from the host plasma membrane (a phenomenon known as budding).

The new virions produced are free to infect and replicate in other host cells in the area and start the cycle all over again. It is worth remarking that the complete viral life cycle generally takes between 6 and 8 hours, and as many as 10.000 virions may be released from an infected cell.¹

Undeniably, knowing the virus replication process has been one of the key points in the search for effective antiviral drugs.

1.2. Viral diseases in humans: prevention and therapy

Every year millions of people are infected by different types of viruses, such as the human immunodeficiency virus (HIV), the herpes simplex virus (HSV), the varicella zoster virus (VZV) or the hepatitis virus B and C (HBV and HCV). Not all viruses cause pathogenesis after the infection; pathogenicity is described as the capacity of one organism to produce a disease in another. Although some viruses produce a wide range of diseases in humans, in most cases infections are silent and do not result in any external signs of disease. It is noteworthy that not all the pathogenic symptoms seen in virus infections are caused directly by the virus replication but are side effects of the immune response.

Some viruses, such as herpes viruses, are able to remain latent within the host cells after initial infection and can be reactivated to induce recurrent endemic disease.⁵ Specifically, herpes simplex virus type 1 (HSV-1) is associated with orofacial lesions, whereas genital herpes is frequently induced by herpes simplex virus type 2 (HSV-2). HSV infections are among the

2. Antiviral drugs

most common diseases of humans, with an estimated 60-95% of the adult population being infected by at least one of them.^{6,7}

During the last three decades many efforts have been made on the development of an effective antiviral therapy. Most of the harm caused in cells is usually produced before the first clinical symptoms are detected. For that reason, the prevention of virus infection is still a far better solution than its cure. Vaccines are the most effective way to prevent diseases generated by human viral pathogens and have helped to control some of dreadful viral diseases (e.g. hepatitis B, polio, smallpox).⁸ However, mutations on viruses always rapidly cause immunization against them.

Antiviral drugs are useful in dealing with viral diseases where there is no effective vaccine or when the infection has already taken place. Thus, the treatment in most of the cases relies on the administration of antiviral drugs that inhibit their development.

2. Antiviral drugs

The aim of an effective antiviral drug is to inhibit the virus replication without causing toxic effects on the host cells.

Nowadays, around 60 antiviral drugs have been approved by the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA). Interestingly, almost half of them are against HIV. The others are used in the treatment of herpes virus (e.g. HSV, VZV, cytomegalovirus), HBV and HCV or influenza virus infections.⁹

Almost any stage on the replication process of viruses is susceptible to be a target for antiviral compounds. For example, in the case of HIV, antiviral drugs can be classified in different categories:⁹

- Replicase and transcriptase inhibitors (including Nucleoside Reverse Transcriptase Inhibitors (NRTIs) and the Non-Nucleoside Reverse Transcriptase Inhibitors (NNRTIs))
- Protease inhibitors (PIs)
- Viral entry inhibitors (including coreceptor inhibitors (CRIs) and fusion inhibitors (FIs))
- Integrase inhibitors (INIs)

Amongst the approved antiviral drugs, nucleoside analogues have been the spearhead for the treatment of many widespread diseases caused by viruses.¹⁰⁻¹² Additionally, today several new nucleoside analogues are undergoing preclinical or clinical development.¹³⁻¹⁶

2.1. Nucleoside analogues

The structure of natural nucleosides can be divided into three subunits: a hydroxymethyl group, as a polar group; a sugar moiety, which can be either ribose (RNA) or deoxyribose (DNA); and a purine or pyrimidine base, which can be adenine (A), guanine (G), thymine (T), cytosine (C) or uracil (U) (Figure I-3). Nucleotides are nucleosides that have been phosphorylated one, two or three times, and are building blocks of nucleic acids. Nucleoside analogues are synthetic compounds that have been developed to mimic their physiological counterparts. In order to modulate nucleosides and nucleotides activity, each of their subunits may be modified.

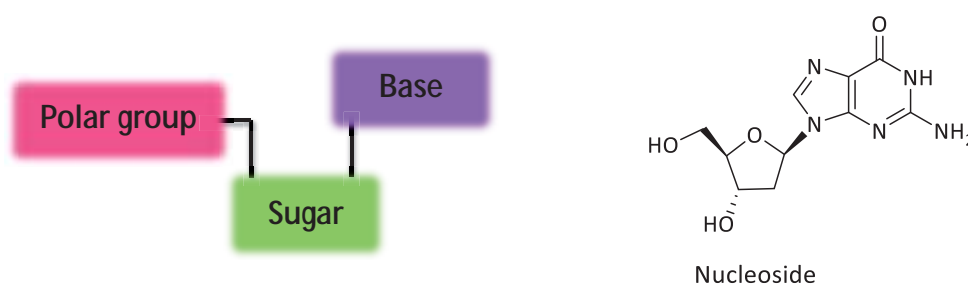


Figure I-3. General structure of a nucleoside and structure of a natural guanosine nucleoside.

The first nucleoside analogue approved as antiviral drug was iododeoxyuridine, IDU, (Figure I-4) which was synthesised by William Prusoff in 1959.¹⁷ Due to its toxicity, it was used only for the treatment of HSV infections of the eyes.¹⁸ A major breakthrough in antiviral therapy was the synthesis of the acyclic guanosine analogue Acyclovir (ACV) (Figure I-4) by Gertrude B. Elion and George H. Hitchings in 1978,¹⁹ which was reported to be active against herpes simplex virus (HSV-1 and HSV-2) and varicella-zoster virus (VZV).¹² Followingly, other acyclic nucleoside analogues were developed against herpes viruses such as Ganciclovir²⁰ (GCV), Penciclovir²¹ (PCV) and their prodrugs Valaciclovir²² (Val-ACV), Valganciclovir²³ (Val-GCV) and Famciclovir²⁴ (FCV), as well as Trifluridine²⁵ (TFT) and Brivudine²⁶ (BVDU) as cyclic nucleoside analogues (Figure I-4). It is worth noting that the nucleotide analogue Cidofovir²⁷ ((S)-HMPC) features a methylphosphonate group instead of a phosphate in order to increase its stability towards hydrolysis.

2. Antiviral drugs

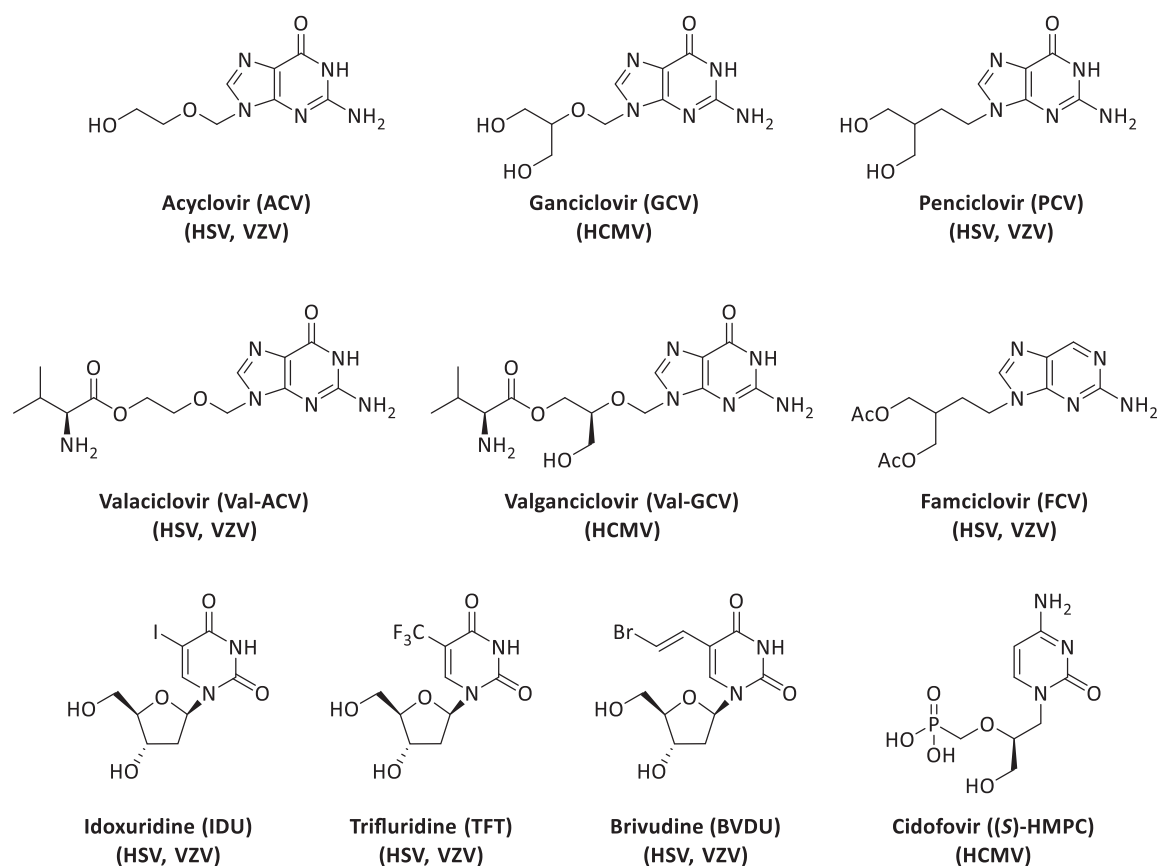


Figure I-4. Antiviral drugs based on nucleoside analogues currently used for the treatment of herpes viruses. The specific herpes virus inhibited by each drug is indicated in parenthesis.

Another breakthrough in antiviral chemotherapy was the discovery of the anti-HIV agent Zidovudine (AZT) in 1985, which contains an 3'-azide group instead of the 3'-hydroxyl group of the natural nucleoside.²⁸ This finding encouraged investigators to search for novel nucleoside analogues with potent anti-HIV activity. To date, seven more nucleoside analogues have been approved by the FDA for the treatment of HIV (Figure I-5).²⁹

Despite these achievements, the development of newer antiviral agents with improved properties is still necessary in order to overcome the main deficiencies of the current drugs such as their toxicity³⁰⁻³³, metabolic instability³³⁻³⁶ and, above all, the emergence of virus drug resistance.^{13,15,31,34-38}

An extensive knowledge of the mechanism of action of nucleoside analogues will contribute to the development of novel antiviral nucleoside analogues to face the resistance issue.

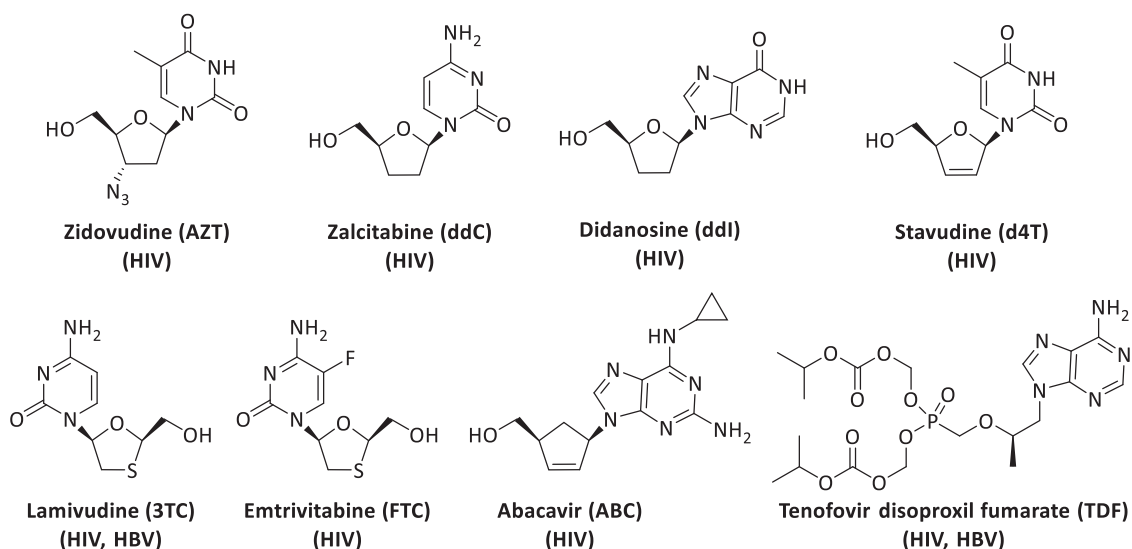


Figure I-5. Antiviral drugs based on nucleoside and nucleotide analogues currently used for the treatment of HIV. The drugs also approved for the treatment of hepatitis B virus (HBV) are indicated in parenthesis.

Antiviral nucleoside analogues are prodrug forms of the active compounds that target the viral polymerases, which depending on the virus could be reverse transcriptases, DNA polymerases or RNA polymerases, and must be triphosphorylated after their penetration into the host cell to be active and reach those targets. The triphosphorylated derivatives cannot themselves be used as drugs, because, due to their polarity they are unable to cross the cell membranes. The activation process requires three successive phosphorylation steps carried out by different enzymes called kinases, which whether are from the host or encoded by the virus. The kinases involved in the activation of nucleosides are the following: nucleoside kinase, nucleoside monophosphate kinase (NMPK) and nucleoside diphosphate kinase (NDPK), which catalyse the addition of the first, second, and third phosphoryl groups at 5' position of nucleoside analogues, respectively (Figure I-6).^{39,40} Once nucleoside analogues have been triphosphorylated, they must interact with the viral polymerase.

The inhibition of the replication process by nucleoside analogues may be performed in several ways. If the drug does not contain the 3'-hydroxyl group of the deoxyribose, which is necessary to add further nucleotides into the DNA growing chain, it will act as a chain terminator blocking the elongation process after its incorporation into the primer DNA strand.^{31,41,42} If the nucleoside analogue contains the 3'-hydroxyl group, the replication process may also be interrupted by other mechanisms, such as causing steric hindrance when it has been added to the DNA growing chain. Moreover, nucleoside analogues may stop the replication of the viral polymerase via competitive inhibition with the native triphosphorylated nucleoside (Figure I-6).

2. Antiviral drugs

Nucleoside analogue affinity must be higher for the virus-associated polymerase than for the human one in order to improve its selectivity and reduce its toxicity to the host cells.^{19,40} Remarkably, the low toxicity of some antiviral compounds, such as ACV, is also related to the narrow substrate acceptance of human kinases compared to the viral counterparts.^{43,44} In the case of ACV, the first phosphorylation step is catalysed by the HSV kinase but not by the human kinase thus preventing its activation in non-infected cells, which would have affected normal cell division.¹⁹

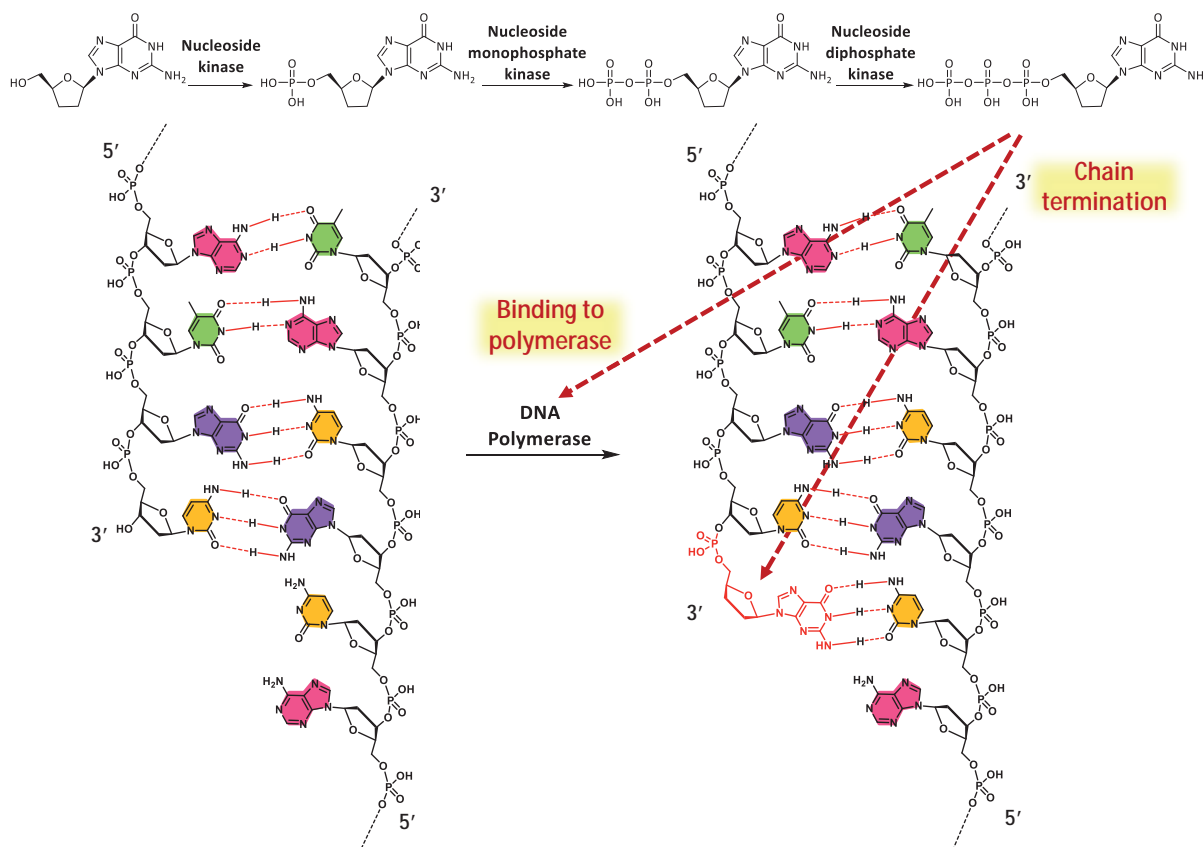


Figure I-6. Mechanism of inhibition of viral replication by 2',3'-dideoxynucleosides.

Even though the synthesis of nucleoside analogues has advanced considerably, more efficient methods are still in demand for the preparation of chiral key intermediates leading to these biologically active compounds.

2.2. Carbocyclic nucleoside analogues

Over the last three decades many efforts have been made on the development of new nucleoside analogues to improve the antiviral efficacy as well as to reduce the toxicity by modifying the structures. In recent years, attention has been focused on carbocyclic nucleosides (also called *carba*-nucleosides) which contain a carbocycle instead of the furanose

ring of the regular nucleosides. The lack of the *N*-glycosidic linkage in these derivatives results in higher resistance to hydrolytic processes and enhance lipophilicity, favouring absorption and penetration through the cell membrane.^{45,46}

The first carbocyclic nucleoside analogues Aristeromycin⁴⁷ and Neplanocin A⁴⁸ (Figure I-7) are natural products with antibiotic and antitumor activity. The discovery of these nucleosides led to the synthesis of novel carbocyclic analogues featuring different ring size.^{49–55}

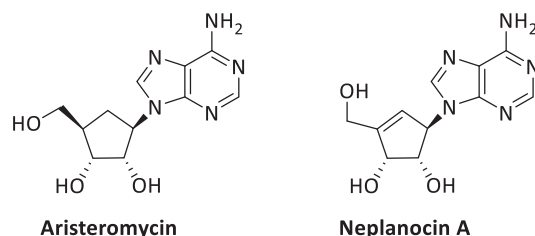


Figure I-7. Structure of the carbocyclic nucleosides Aristeromycin and Neplanocin A.

Since then, a large number of novel carbocyclic nucleoside have been prepared, many of them with interesting biological activities, being the five-membered ring analogues by far the most extensively studied. Among them, Carbovir (CBV), Abacavir (ABC) and Entecavir (ETC) are the most successful examples, showing high activity and low toxicity (Figure I-8).

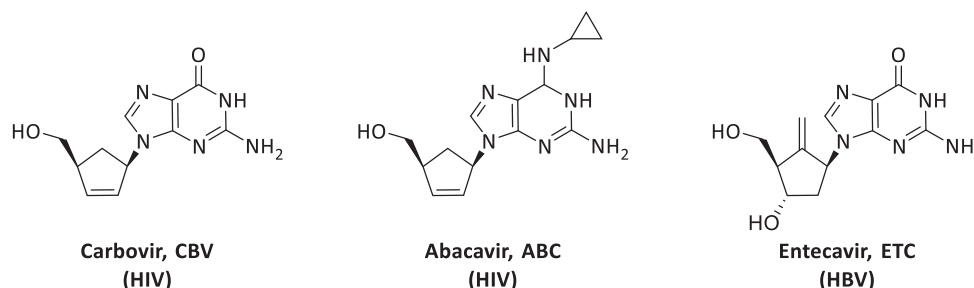


Figure I-8. Structure of the five-membered carbocyclic nucleosides Carbovir, Abacavir and Entecavir. The specific viruses inhibited by each drug are indicated in parenthesis.

Carbovir was synthesised in 1988 and was shown to exhibit potent anti-HIV activity with low toxicity.^{56,57} However, it was limited by its pharmacokinetics and toxicological deficiencies.^{58,59} These problems were solved with the synthesis of the 6-cyclopropylamino derivative, Abacavir, which was approved by the FDA for the HIV treatment in 1998.^{59,60} Another successful nucleoside belonging to the five-membered carbocyclic nucleosides is Entecavir. It was shown to display a potent anti-HBV activity and in 2005 was approved by the FDA for the treatment of chronic HBV.⁶¹

Carbocyclic nucleosides bearing a cyclopropane unit have also been explored. The first asymmetric synthesis of D- and L-cyclopropyl nucleosides, I and II, was achieved by Chu and co-workers (Figure I-9).^{62,63} Unfortunately, none of them exhibited antiviral activity, suggesting that the lack of activity could be associated with the short distance between the 5'-OH group

2. Antiviral drugs

and the nucleobase.⁶⁴ For this reason, the synthesis of cyclopropane analogues has been directed to the preparation of derivatives featuring a methylene spacer between the nucleobase and the carbocyclic ring. One example is compound A-5021, a guanine derivative synthesised by Tsuji and co-workers, which displays antiviral activity against HSV-1, HSV-2 and VZV.⁶⁵ Regarding four-membered carbocyclic analogues, two successful cyclobutane nucleosides examples inspired by the natural compound Oxetanocin-A,⁶⁶ which exhibits anti-HIV activity, are Cyclobut-A and Cyclobut-G, which also display good antiviral potency.⁶⁷⁻⁶⁹ Among six-membered carbocyclic analogues, both enantiomers of Cyclohexenyl G, display potent anti-herpesvirus activity (HSV-1, HSV-2, VZV, CMV).^{70,71}

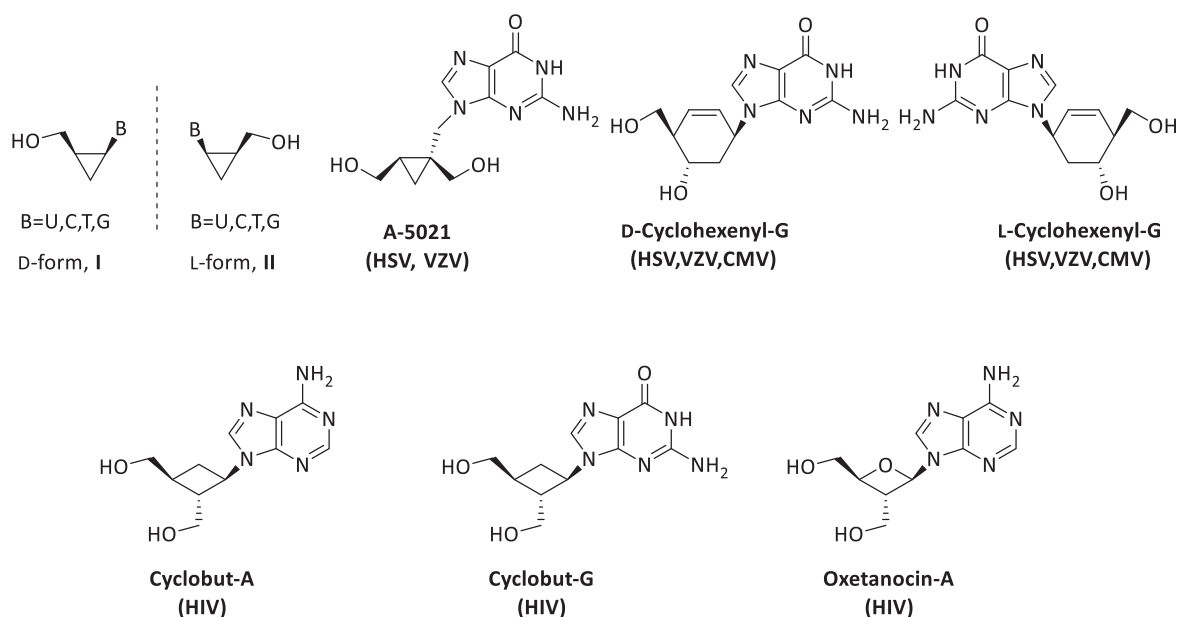


Figure I-9. Three-, six- and four-membered carbocyclic nucleoside analogues that display antiviral activity, and the naturally occurring nucleoside analogue Oxetanocin-A. The specific viruses inhibited by each compound are indicated in parenthesis.

Based on structural characterization of isolated drugs and bound to their targets, it has been postulated that the conformation and puckering of the sugar moiety of nucleosides play a critical role in modulating their biological activity. In recent years, a new series of conformationally locked *carba*-nucleosides has been studied in order to mimic the conformational behaviour of the furanose ring.⁷²⁻⁷⁷ More specifically, conformationally rigid nucleoside analogues built on a bicyclo[3.1.0]hexane system, where the cyclopentane ring was fused to a cyclopropane, have been reported with antiviral activity (Figure I-10).⁷⁸⁻⁸³

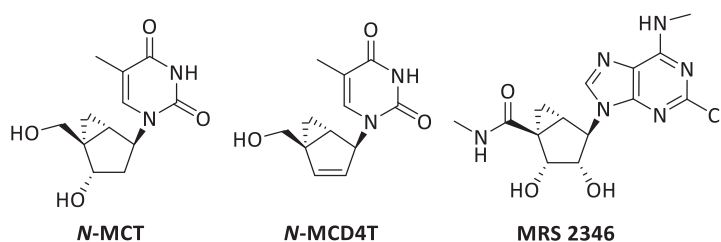


Figure I-10. Structures of some bicyclo[3.1.0]hexane nucleoside analogues with antiviral activity.

The synthesis of new structures and the development of new methodologies to extend the structural diversity of this family of molecules are still very active areas.

2.3. Cyclohexanyl and cyclohexenyl nucleoside analogues

Carbocyclic analogues are molecules that have been widely investigated. However, the number of six-membered analogues is still quite limited now. The first examples of cyclohexanyl nucleosides were dated from the early sixties.^{84,85} During the last decades, different mono- (III and IV),^{86,87} di- (V)^{88,89} and trisubstituted (VI)⁹⁰ cyclohexanyl nucleoside analogues have been synthesised (Figure I-11). Unfortunately, none of them displayed sufficient antiviral activity with only V synthesised in an enantiomerically pure form.

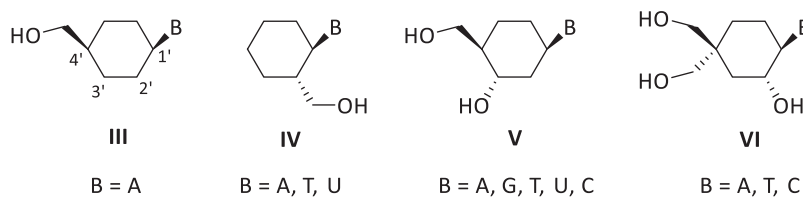


Figure I-11. Structures of cyclohexanyl nucleoside analogues synthesised.

Some conformational studies of the cyclohexene ring suggest that cyclohexenyl nucleoside analogues can be considered as bioisosteres of natural furanose nucleosides.^{70,71,91} Indeed, the conformational behaviour of the cyclohexene ring is similar to that of a furanose ring due to the presence of two sp^2 -hybridised carbon atoms in the cyclohexene ring which reduce its flexibility.⁹²

The natural furanose ring is not planar, so it can exist in different conformations represented in a pseudorotational cycle (Figure I-12). It has two preferential conformations called twist (T) and two other conformations called envelope (E). A cyclohexene ring mainly exists in the half-chair forms, which interconvert via the symmetrical boat form. These half-chair (3_2H (north) and 2_3H (south)) conformations of the cyclohexene ring are structurally equivalent to the twist (3_2T (north) and 2_3T (south)) forms of the furanose ring (Figure I-12).⁷¹

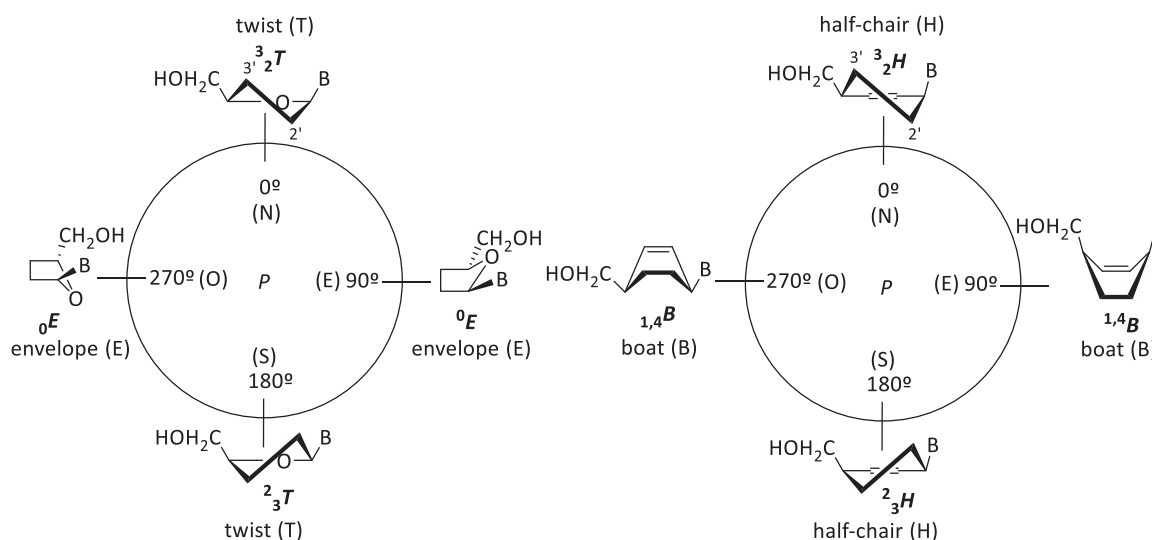


Figure I-12. Twist (T) and envelope (E) conformations of furanose ring and half-chair (H) and boat (B) of cyclohexene ring in a pseudorotational cycle.

The conformation of a nucleoside is determined by competitive steric and stereoelectronic effects. In the case of cyclohexene nucleosides their conformation is controlled by steric effects as well as by the $\pi \rightarrow \sigma^*_{C1'-N}$ interaction between the C5'-C6' double bond and the heterocyclic aglycon.⁷⁰ The $\pi \rightarrow \sigma^*_{C1'-N}$ effect is similar to the anomeric effect in furanose nucleosides and can be explained as an overlap between the antibonding C1'-N and the orbitals of the π bond (Figure I-13). This $\pi \rightarrow \sigma^*_{C1'-N}$ interaction drives the ${}^3_2H \rightleftharpoons {}^2_3H$ equilibrium toward the 3_2H conformation, in which the base moiety is pseudoaxially oriented.

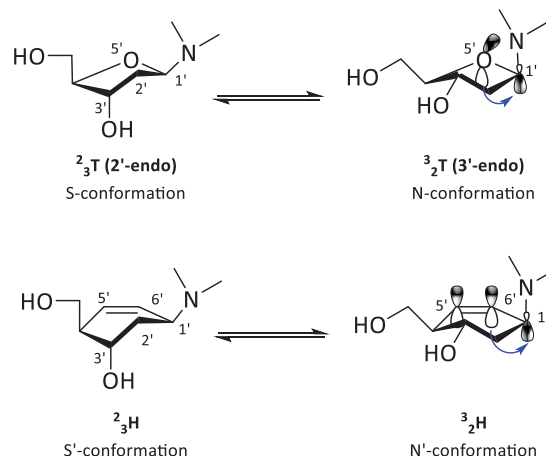


Figure I-13. Comparison between the northern \rightleftharpoons southern conformational equilibrium between a furanose nucleoside and a cyclohexene nucleoside as well as between the anomeric effect in furanose nucleosides and $\pi \rightarrow \sigma^*_{C1'-N}$ effect in cyclohexene nucleosides.

These conformational studies encouraged investigators to search for novel cyclohexenyl nucleosides analogues. To date, six families of cyclohexenyl analogues have been synthesised: a) L-4'-hydroxymethylcyclohexenyl nucleosides, VII;⁹³⁻⁹⁵ b) 3',4'-dihydroxymethylcyclohexenyl

nucleosides, **VIII**;⁹⁶ c) 4'-hydroxycyclohexenyl nucleosides, **IX**;⁹⁷⁻⁹⁹ d) *ara*-cyclohexenyl nucleosides, **X**;^{100,101} e) *ribo*-cyclohexenyl nucleosides, **XI**;¹⁰² and f) 4'-hydroxymethyl-3'-hydroxycyclohexenyl nucleosides, **XII**.^{70,71,103,104} However, most of them did not display any significant biological activity.

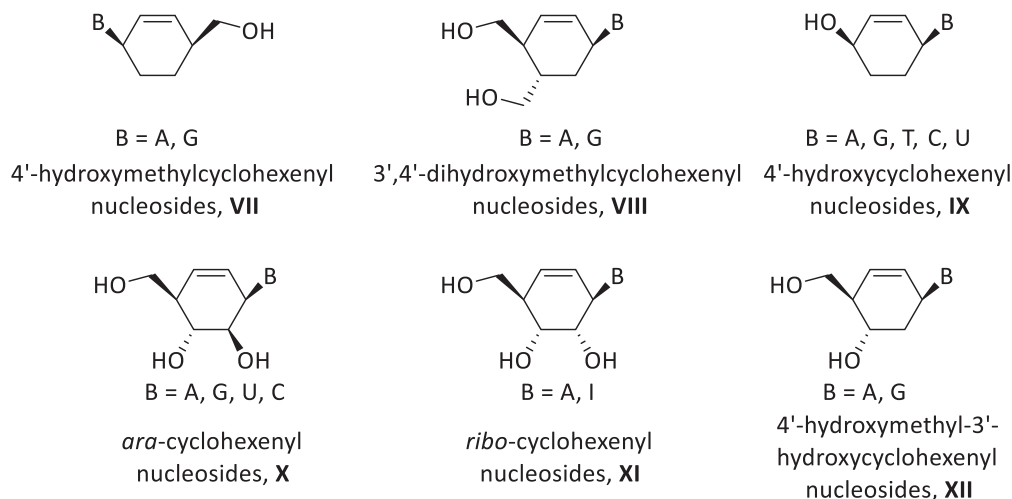


Figure I-14. Structures of some examples of cyclohexenyl nucleoside analogues.

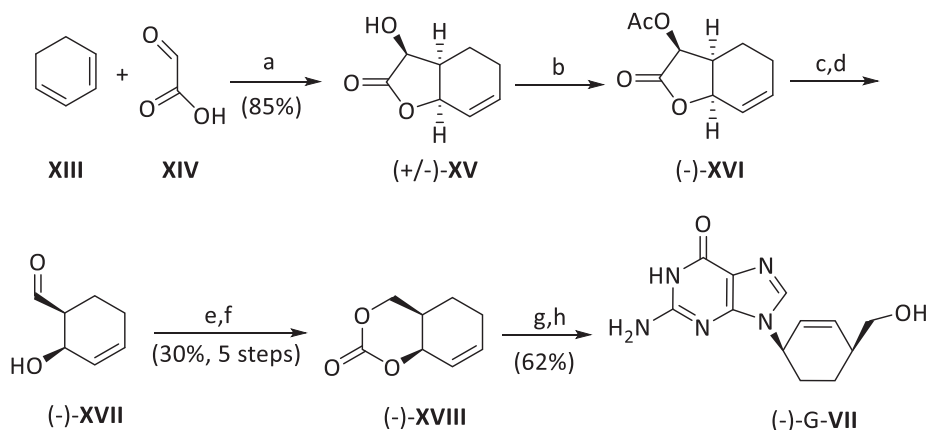
2.4. Precedents

Many methods have been successfully devised to prepare cyclohexenyl analogues, including stepwise base construction from amino alcohols,^{84,85,95,105} Mitsunobu-type base addition^{100,102,106} and Pd(0)-catalysed coupling.^{93,94,96} Herein, a brief overview of some reported examples is given.

Three approaches towards L-homocarbovir **VII** were developed although only one provided optically pure L-homocarbovir **VII** via enzymatic resolution.⁹³⁻⁹⁵ The two first approaches were reported in 1996 by Katagiri and co-workers⁹⁵ and Konkel and Vince.⁹⁴ Both of them started with a hetero Diels-alder reaction to construct the cyclohexene ring and the main difference was the methodology used to introduce the base moiety. In 1998, Olivo and co-workers published the synthesis of enantiomerically pure L-(4'-hydroxymethylcyclohexenyl)-guanine **G-VII** (Scheme I-1).⁹³ The synthesis started with the reaction of 1,3-cyclohexanediene, **XIII**, and glyoxylic acid, **XIV**, which afforded (+/-)-hydroxylactone, **XV**.¹⁰⁷ Kinetic resolution using a *pseudomonas fluorescens* lipase led to the enantiomerically enriched (-)-acetylactone **XVI**. Reduction of the acetyl group and the lactone, followed by an oxidative cleavage provided aldehyde **XVII**. Bicyclocarbonate **XVIII** was obtained by reduction with NaBH₄ to the corresponding alcohol and subsequent addition of triphosgene. Finally, the introduction of the base moiety via Tsuji-Trost reaction and successive

2. Antiviral drugs

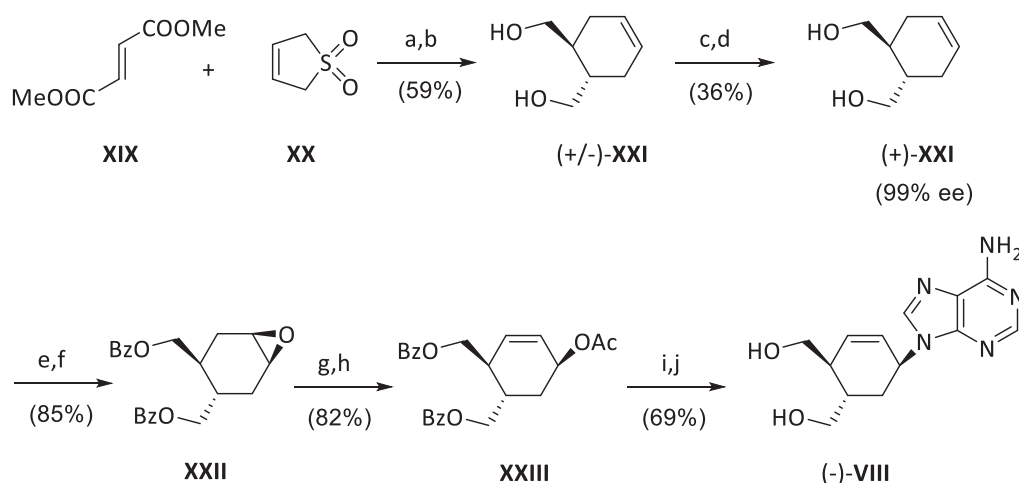
hydrolysis afforded the L-(4'-hydroxymethylcyclohexenyl)-guanine **VII** in 8 steps and 16% overall yield.



Reagents and conditions: (a) H₂O, pH=1; (b) *pseudomonas fluorescens* lipase, vinylacetate; (c) LiAlH₄, THF; (d) NaIO₄, Et₂O-H₂O; (e) NaBH₄, EtOH; (f) triphosgene, Et₃N, CH₂Cl₂; (g) 2-amino-6-chloropurine, Pd(PPh₃)₄, DMSO/THF (1:1); (h) CF₃COOH/H₂O (3:1).

Scheme I-1. Synthesis of enantiomerically pure L-(4'-hydroxymethylcyclohexenyl)guanine, (-)-G-VII.

The first stereoselective synthesis of 3'-4'-dihydroxycyclohexenyl nucleosides analogues **VIII** was carried out by Samuelsson and co-workers in 1996 (Scheme I-2).⁹⁶ The cyclohexene ring was formed by a Diels-Alder reaction of dimethyl fumarate, **XIX**, and 3-sulpholene, **XX**. The next step was a reduction with LiAlH₄ to provide the racemic diol (±)-**XXI**, which was resolved via a lipase-catalysed transesterification process. Protection of the diol (+)-**XXI** with benzoyl chloride followed by the epoxidation using *m*-CPBA gave the epoxide **XXII**. Then, this intermediate was converted to the corresponding allylic alcohol, which was acetylated with acetic anhydride to afford the allylic acetate **XXIII**. The introduction of the base was achieved via a Tsuji-Trost reaction. After hydrolysis, the 3'-4'-dihydroxycyclohexenyl (-)-**VIII** was achieved in 10 steps and 10% overall yield.

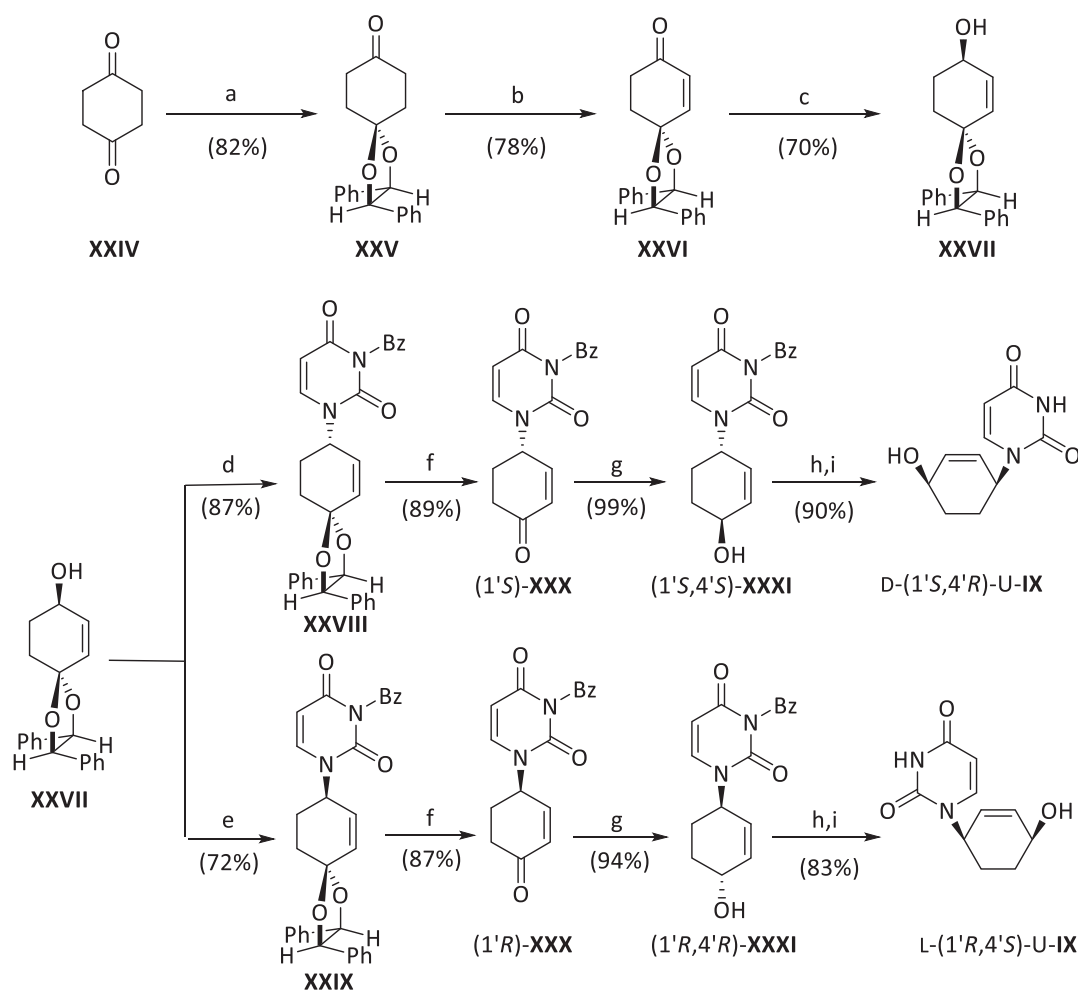


Reagents and conditions: (a) CHCl_3 , reflux; (b) LiAlH_4 , THF; (c) SAM-II, vinylacetate, CHCl_3 ; (d) NaOMe , MeOH; (e) BzCl , py; (f) *m*-CPBA, CH_2Cl_2 ; (g) i) TMSOTf , DBU, toluene, ii) H^+ , MeOH; h) Ac_2O , py; (i) $\text{Pd}(\text{PPh}_3)_4$, NaH, adenine; j) NH_3 .

Scheme I-2. Synthesis of adenine 3',4'-hydroxymethylcyclohexenyl nucleoside.

The first family of 4'-hydroxycyclohexenyl nucleoside analogues IX was synthesised by Arango and co-workers but in a racemic form.⁹⁹ Taking into account the importance of preparing enantiomerically pure nucleosides, our research group developed an enantioselective approach toward the synthesis of both D- and L-enantiomers of 4'-hydroxycyclohexenyl nucleoside analogues IX (Scheme I-3).¹⁰⁶ This approach relied on the use of (*R,R*)-hydrobenzoin as chiral auxiliary to induce enantioselectivity. Thus, the synthesis started from the commercially available 1,4-cyclohexanedione, XXIV, which was monoprotected with (*R,R*)-hydrobenzoin to give monoketal XXV. This monoketal was oxidised to enone XXVI, which was then stereoselective reduced using catecholborane and (*R*)-2-Me-CBS to afford the key allylic alcohol XXVII. Two different methodologies were used to introduce the base moiety: a Mitsunobu-type reaction was applied to achieve the D-series of cyclohexenyl nucleosides, while palladium-catalysed coupling led to the L-isomers. Then, removal of the ketal, followed by the reduction with catecholborane led to the *trans* isomers (1'*S*,4'*S*)-XXXI and (1'*R*,4'*R*)-XXXI. The inversion of the hydroxyl group was carried out via Mitsunobu-type reaction using *p*-nitrobenzoic as a nucleophile. A final ammonolysis removed the nitrobenzoates and leading to uracil analogues D-(1'*S*,4'*R*)-U-IX and L-(1'*R*,4'*S*)-U-IX in 8 steps and 31% and 22% overall yield, respectively. The synthesis of adenine analogues following the same strategy was also reported.

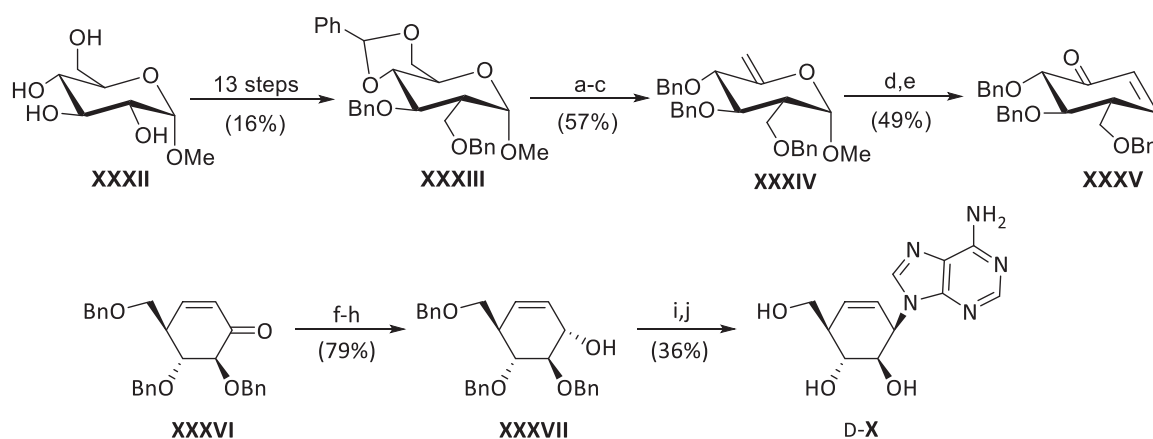
2. Antiviral drugs



Reagents and conditions: (a) (*R,R*)-hydrobenzoin, *p*-T₂SOH, benzene; (b) (i) Br₂, diethyl ether; (ii) DBU, dioxane; (c) cathecolborane, (*S*)-2-Me-CBS, CH₂Cl₂; d) *N*²-benzoyluracil, DBAD, Ph₃P, THF; e) (i) ClCO₂Et, py, DMAP, CH₂Cl₂; (ii) *N*³-benzoyluracil, [(η³-C₃H₅)PdCl]₂, dppe, DMF; (f) CF₃COOH/H₂O (14:1); (g) catecholborane, (*S*)-2-Me-CBS for (1'*S*)-XXX or (*R*)-2-Me-CBS for (1'*R*)-XXX, CH₂Cl₂; (h) *p*-NO₂BzOH, DBAD, Ph₃P, THF; (i) MeNH₂, EtOH.

Scheme I-3. Synthesis of (4'-hydroxycyclohexenyl)uracil D-(1'*S*,4'*R*)-U-IX and its enantiomer L-(1'*R*,4'*S*)-U-IX.

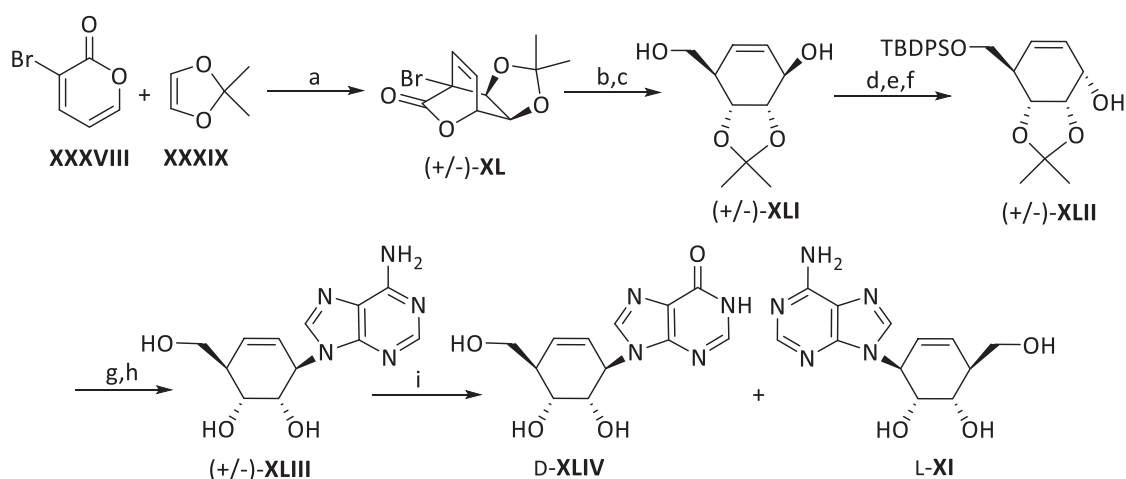
In 2007, Hederwijn and co-workers reported the enantioselective synthesis of *ara*-cyclohexenyl nucleosides **X** in enantiomerically pure form (Scheme I-4).¹⁰⁰ The sequence is started from commercially available and inexpensive methyl- α -D-glucopyranose, **XXXII**, which was converted to the corresponding derivative **XXXIV** in 16 steps. Then, a Ferrier rearrangement followed by an elimination reaction provided enone **XXXV**. Selective reduction and introduction of the base via Mitsunobu reaction furnished the *D-ara*-cyclohexenyl nucleoside **D-X** in 23 steps and 0.7% overall yield.



Reagents and conditions: (a) $\text{Cu}(\text{OTf})_2$, $\text{BH}_3 \cdot \text{THF}$, THF; (b) PPh_3 , I_2 , imidazole, toluene; (c) DBU, THF; (d) HgCl_2 , acetone/ H_2O (4:1); (e) MsCl , DMAP, py; (f) NaBH_4 , $\text{CeCl}_3 \cdot 7\text{H}_2\text{O}$, EtOH/THF (1:1); (g) BzCl , py; (h) K_2CO_3 , MeOH; (i) adenine, DIAD, PPh_3 , dioxane; (j) BCl_3 , CH_2Cl_2 .

Scheme I-4. Synthesis of *ara*-cyclohexenyladenine, D-X.

In 2005, the same research group reported the synthesis of *ribo*-cyclohexenyl nucleosides XI (Scheme I-5).¹⁰² The synthetic route started with a Diels-Alder reaction between 3-bromo-2*H*-pyran-2-one, XXXVIII, and 2,2-dimethyl-1,3-dioxolane, XXXIX, to generate the bicyclic intermediate (+/-)-XL. Reduction of the bromine and subsequent reduction of the lactone provided diol (+/-)-XLI. Protection of the primary hydroxyl group and the inversion of the configuration of the second hydroxyl afforded the corresponding alcohol (+/-)-XLII. Then, introduction of the base moiety via Mitsunobu reaction delivered the *ribo*-cyclohexenyl adenine XI in a racemic form in 6 steps and 19% overall yield. Finally, an enzymatic kinetic resolution using adenosine deaminase (ADA), which selectively converts the D-like enantiomer into an inosine analogue, led to the inosine nucleoside D-XLIV and the adenine nucleoside L-XI.

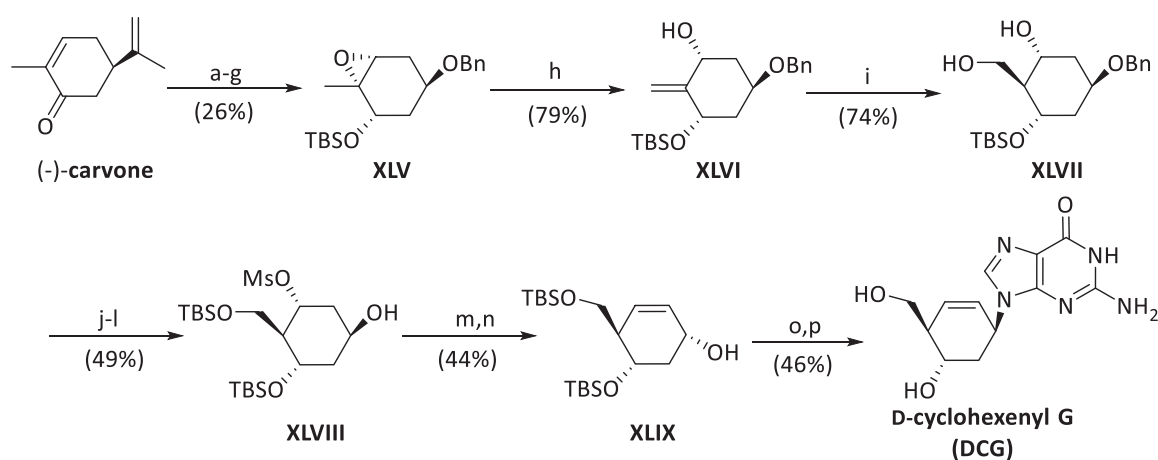


Reagents and conditions: (a) CH_2Cl_2 , 90 °C; (b) *n*- Bu_3SnH , AIBN, toluene; (c) LiAlH_4 , THF; (d) TBSCl, imidazole, DMF; (e) MnO_2 , CH_2Cl_2 ; (f) NaBH_4 , $\text{CeCl}_3 \cdot 7\text{H}_2\text{O}$, MeOH; (g) PPh_3 , DIAD, adenine, dioxane; (h) $\text{CF}_3\text{COOH}/\text{H}_2\text{O}$ (3:1); (i) ADA, H_2O .

Scheme I-5. Synthesis of *L-ribo*-cyclohexenyl adenine, L-XI.

2. Antiviral drugs

Disappointingly, none of the compounds previously described showed any significant antiviral activity. On the other hand, Hederwijn and co-workers reported in 1999 the synthesis of both enantiomers of Cyclohexenyl G (DCG and LCG), which displayed antiviral activity against herpesvirus (Scheme I-6).^{70,71} The starting material was (*R*)-carvone which was converted into epoxide **XLV** in 7 steps. The regioselective aperture of the epoxide followed by the hydroboration with 9-BBN of the exo double bond led to diol **XLVII**, which was further elaborated to obtain alcohol **XLVIII** which was oxidised to enone and then reduced using NaBH₄ to afford allylic alcohol **XLIX**. Finally, introduction of the base moiety via Mitsunobu reaction and successive deprotection with TFA delivered D-cyclohexenyl G (DCG) in 12 steps and 1.5% overall yield.

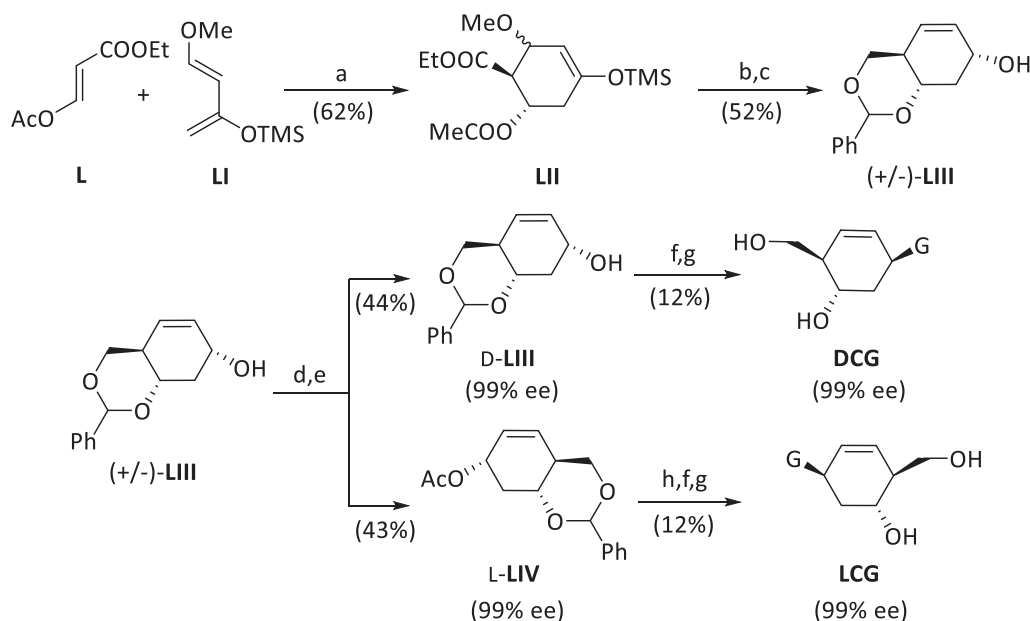


Reagents and conditions: (a) H₂O₂, NaOH, MeOH; (b) L-Selectride, THF; (c) TBSCl, imidazole, DMF; (d) OsO₄, KIO₄, THF/H₂O; (e) *m*-CPBA, CHCl₃, pH=8; (f) K₂CO₃, MeOH; (g) NaH, BnBr, TBAI, THF; (h) LiTMP, Et₂AlCl, toluene; (i) i) 9-BBN, THF; ii) H₂O₂, NaOH, H₂O; (j) TBSCl, imidazole, DMF; (k) MsCl, Et₃N, CH₂Cl₂; (l) Pd-C, HCOONH₄, MeOH; (m) MnO₂, CH₂Cl₂; (n) NaBH₄, CeCl₃·7H₂O, MeOH; (o) 2-amino-6-chloropurine, DEAD, PPh₃, 1,4-dioxane; (p) CF₃COOH/H₂O (3:1).

Scheme I-6. Synthesis of D-cyclohexenyl G (DCG).

However, the synthesis was long and time-consuming, and it was not suited for the preparation of large amounts of the final products. As a consequence, in 2004 the same research group developed a new synthetic approach (Scheme I-7).^{103,104} The new route started from a Diels-Alder reaction of ethyl (2*E*)-3-acetyloxy-2-propenoate, **L**, with Danishefsky's diene **LI** to construct the six-membered ring skeleton **LII**, which was reduced using LiAlH₄ to afford the corresponding alcohol which was then protected as the benzylidene acetal, with concomitant formation of the allylic alcohol moiety, to yield racemic intermediate (±)-**LIII**. The separation of both enantiomers was carried out via an enzymatic kinetic resolution using *Candida antarctica* lipase B, Novozyme® 435, which only acetylates alcohols with *S* configuration. Once both enantiomers of **LIII** were separated, the next step was the

introduction of the base moiety via a Mitsunobu reaction. After removal of the protecting group using TFA, both enantiomers of cyclohexenyl G, DCG and LCG, were obtained in 7 and 8 steps, respectively, and 2% overall yield in both cases.



Reagents and conditions: (a) Hydroquinone, 180 °C; (b) LiAlH₄, THF; (c) PhCH(OMe)₂, *p*-TsOH, dioxane; (d) Novozyme® 435, isopropenyl acetate, CH₂Cl₂; (e) recrystallization twice in EtOAc/n-Hexane 50%; (f) PPh₃, DEAD, 2-amino-6-chloropurine, dioxane; (g) TFA/H₂O (3:1); (h) NH₃, MeOH.

Scheme I-7. Synthesis of both enantiomers DCG and LCG.

In summary, in the last decades only few synthetic strategies towards the synthesis of enantiomerically pure cyclohexenyl nucleosides have been reported. Therefore, the development of enantiomerically pure six-membered carbocyclic nucleosides is still an important challenge in antiviral research.

On the other hand, a way to save time and cost in drug design is the use of molecular modelling, which has recently emerged as a powerful tool, being able to predict the affinity of a substrate before synthesising it.

3. *In silico* molecular modelling for drug design

The design of new drugs is a time-consuming and multi-step process. Most of the reported syntheses of drugs relied heavily on random variations of lead compounds on a trial and error basis. Advances in molecular biology and genetics have provided a detailed understanding of drug targets. This, combined with the advances in computer hardware and software for the investigation of biological processes, has been a revolutionary change in medicinal chemistry.¹⁰⁸

In the last years, *in silico* molecular modelling studies have arisen as a powerful key tool for drug discovery. Molecular modelling encompasses all theoretical methods and computational techniques used to model or mimic the behaviour of molecules and molecular systems.¹⁰⁹

One of the most widely used molecular modelling techniques in computer-aided drug design is protein-ligand docking, which tries to predict the structure of an intermolecular complex between different constituent molecules. This is based on the search for a ligand that is able to fit both geometrically and energetically the binding site of a protein. Pioneered during the early 1980s,¹¹⁰ it remains an active area of research as it has demonstrated to be a valuable tool for drug discovery programs. A more detailed description of all these techniques is given in chapter III.

3.1 Protein-ligand docking in medicinal chemistry

Protein-ligand docking has a wide variety of uses and applications in medicinal chemistry, including structure-activity studies, lead optimization, finding potential leads by virtual screening, providing binding hypothesis to facilitate predictions for mutagenesis studies, chemical mechanism studies and combinatorial library design. For instance, virtual screening is commonly used to generate hits against drugs targets for which the structure is known, and docking is also heavily used in structure-based design projects to prioritise medicinal chemistry efforts.

Nowadays, molecular dockings are broadly applied, being capable of predicting known ligand binding modes with average accuracies of about 1.5-2.0 Å and success rates in the range of 70-80%. Hence, molecular modelling has become a useful tool in the research of novel antiviral compounds.^{75,81}

3.2 Protein-ligand docking studies on nucleoside analogues

HSV infections are still among the most frequent human diseases despite many nucleoside analogues are currently used for their treatment.⁶ Despite their safety and efficacy, many nucleoside analogues have limited oral bioavailability and can become ineffective due to the development of drug resistance. Thus, there is still a need for the development of new anti-HSV agents, which requires a better knowledge of their mechanism of action.

In recent years, protein–ligand docking calculations have been used in the development of new nucleoside analogues for the treatment of HSV. In the next paragraphs, some examples of docking studies on nucleoside analogues as anti-HSV agents are reviewed.

In 2001, Hederwijn and co-workers reported the synthesis and the antiviral activity of both D- and L- enantiomer of Cyclohexenyl G, including docking calculations on HSV-1 TK (Figure I-15). As both of them displayed similar antiviral activity against different herpes viruses, the authors carried out molecular modelling studies on both enantiomers in the HSV-1 TK binding site to understand how two enantiomers can be bound to the same enzyme. Analysing the predicted orientations, they concluded that the same amino acids were involved in binding both enantiomers.

In 2007, Marquez and co-workers reported the racemic synthesis of a new bicyclo[3.1.0]hexane nucleoside analogue *iso*-MCT (Figure I-15).⁷⁵ This compound was envisaged after the encouraging results of an earlier investigation in which the authors prepared other bicyclo[3.1.0]hexane nucleosides which were proved to be substrates of HSV-1 TK. The crystallographic structures of these compounds in HSV-1 TK binding site was used for the docking studies on *iso*-MCT prior to its synthesis. According to those docking experiments, only one enantiomer was expected to be recognized as a substrate by HSV-1 TK. This work was extended in 2008 by the publication of the stereoselective synthesis and biological evaluation of both enantiomers.¹¹¹ This biological evaluation confirmed that the D-enantiomer was the biologically active enantiomer, as it had been predicted by molecular modelling.

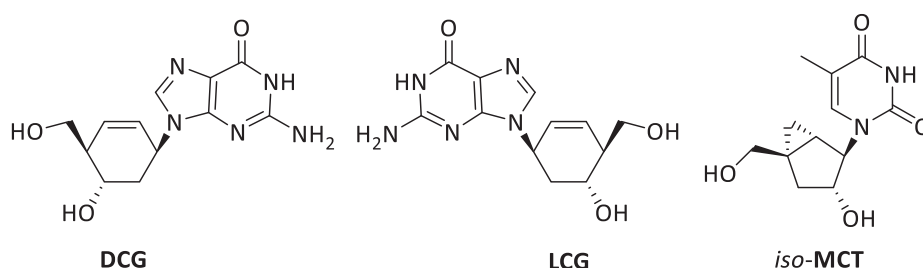


Figure I-15. Structures of the nucleoside analogues investigated by means of protein–ligand dockings as anti-HSV agents.

In recent years, our group have combined drug design research with *in silico* tools. Specifically, we have used a protein-ligand docking study to rationalise the lack of activity against HSV as well as HIV of some nucleoside analogues that had been previously synthesised (Figure I-16).^{112,113} Molecular docking studies of these compounds were performed on the whole activation process for HSV, and on the activation process as well as on the interaction with the viral polymerase for HIV. These studies revealed that most of the studied compounds cannot be activated to the required triphosphorylated form. However, according to these

4. References

results, nucleoside analogues **LIII** and **LVII** can be phosphorylated, thus the lack of activity must be related to their introduction into the viral DNA strand. It is worth highlighting that it was the first study considering the simulation of the entire activation process to rationalise antiviral activities.

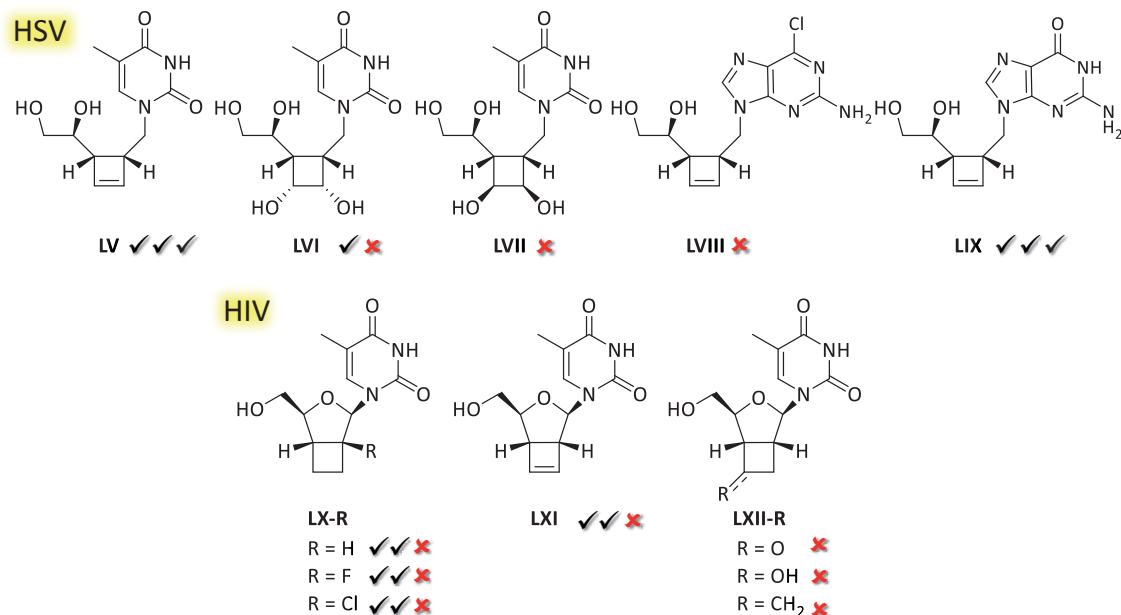


Figure I-16. Protein-ligand docking results of the first, second and third activation steps of cyclobutane and cyclobutene nucleoside analogues carried out in our research group.

4. References

- (1) Collier, L. H.; Oxford, J. S. *Human Virology*; 3rd ed.; Oxford University Press, 2006.
- (2) Wagner, K. E.; Hewlett, M. J. *Basic Virology*; 2nd ed.; Blackwell Publishing, 2004.
- (3) Cann, A. J. *Principles of molecular virology*; 5th ed.; Academic Press, 2011.
- (4) Madigan, M. T.; Martinko, J. M.; Stahl, D. A.; Clark, D. P. *Brock Biology of Microorganisms*; 13th ed.; Pearson Education, 2011.
- (5) Preston, C. M. Herpesviruses : Latency. *Encyclopedia of Virology*, 2006, 2, 436–442.
- (6) Brady, R. C.; Bernstein, D. I. *Antiviral Res.* 2004, 61, 73–81.
- (7) Bulletin of the World Health Organization <http://www.who.int/bulletin/volumes/86/10/07-046128/en/> (accessed Nov 21, 2014).
- (8) Arvin, A. M.; Greenberg, H. B. *Virology* 2006, 344, 240–249.
- (9) De Clercq, E. *Future Virol.* 2008, 3, 393–405.

-
- (10) De Clercq, E. *Antiviral Res.* **2010**, *85*, 19–24.
- (11) Martin, J. C.; Hitchcock, M. J. M.; De Clercq, E.; Prusoff, W. H. *Antiviral Res.* **2010**, *85*, 34–38.
- (12) Razonable, R. R. *Mayo Clin. Proc.* **2011**, *86*, 1009–1026.
- (13) Esté, J. A.; Cihlar, T. *Antiviral Res.* **2010**, *85*, 25–33.
- (14) Cihlar, T.; Ray, A. S. *Antiviral Res.* **2010**, *85*, 39–58.
- (15) Asahchop, E. L.; Wainberg, M. A.; Sloan, R. D.; Tremblay, C. L. *Antimicrob. Agents Chemother.* **2012**, *56*, 5000–5008.
- (16) Price, N. B.; Prichard, M. N. *Curr. Opin. Virol.* **2011**, *1*, 548–554.
- (17) Prusoff, W. H. *Biochim. Biophys. Acta* **1959**, *32*, 295–296.
- (18) Field, H. J. Antiviral Agents. *Encyclopedia of Virology*, 2005, 25–37.
- (19) Elion, G. B.; Furman, P. A.; Fyfe, J. A.; De Miranda, P.; Beauchamp, L.; Schaeffert, H. J. *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 5716–5720.
- (20) Smith, K. O.; Galloway, K. S.; Kennell, W. L.; Ogilvie, K. K.; Radatus, B. K. *Antimicrob. Agents Chemother.* **1982**, *22*, 55–61.
- (21) Harnden, M. R.; Jarvest, R. L. *Tetrahedron Lett.* **1985**, *26*, 4265–4268.
- (22) Colla, L.; De Clercq, E.; Busson, R.; Vanderhaeghe, H. *J. Med. Chem.* **1983**, *697*, 602–604.
- (23) Pescovitz, M. D.; Rabkin, J.; Merion, R. M.; Paya, C. V.; Pirsch, J.; Freeman, R. B.; O’Grady, J.; Robinson, C.; To, Z.; Wren, K.; Banken, L.; Buhles, W.; Brown, F. *Antimicrob. Agents Chemother.* **2000**, *44*, 2811–2815.
- (24) Gill, K. S.; Wood, M. J. *Clin. Pharmacokinet.* **1996**, *31*, 1–8.
- (25) Heidelberger, C.; Parsons, D.; Remy, D. C. *J. Am. Chem. Soc.* **1962**, *84*, 3597–3598.
- (26) De Clercq, E.; Descamps, J.; De Somer, P.; Barr, P. J.; Jones, A. S.; Walker, R. T. *Proc. Natl. Acad. Sci. USA* **1979**, *76*, 2947–2951.
- (27) Snoeck, R.; Sakuma, T.; De Clercq, E.; Rosenberg, I.; Holy, A. *Antimicrob. Agents Chemother.* **1988**, *32*, 1839–1844.
- (28) Broder, S. *Antiviral Res.* **2010**, *85*, 1–18.
- (29) Jordheim, L. P.; Durantel, D.; Zoulim, F.; Dumontet, C. *Nat. Rev. Drug Discov.* **2013**, *12*, 447–464.
- (30) Calmy, A.; Hirschel, B.; Cooper, D. A.; Carr, A. *Antivir. Ther.* **2009**, *14*, 165–179.

4. References

- (31) El Safadi, Y.; Vivet-Boudou, V.; Marquet, R. *Appl. Microbiol. Biotechnol.* **2007**, *75*, 723–737.
- (32) Chen, C.-H.; Vazquez-Padua, M.; Cheng, Y.-C. *Mol. Pharmacol.* **1991**, *39*, 625–628.
- (33) Gallant, J. E.; Staszemski, S.; Pozniak, A. L.; DeJesus, E.; Suleiman, J. M. A. H.; Miller, M. D.; Coakley, D. F.; Lu, B.; Toole, J. J.; Cheng, A. K. *J. Am. Med. Assoc.* **2004**, *292*, 191–201.
- (34) Kuritzkes, D. R. *Curr. Opin. Virol.* **2011**, *1*, 582–589.
- (35) Larder, B. A.; Darby, G.; Richman, D. D. *Science* **1989**, *243*, 1731–1734.
- (36) Morfin, F.; Thou, D. *J. Clin. Virol.* **2003**, *26*, 29–37.
- (37) Little, S. J.; Holte, S.; Routy, J.-P.; Daar, E. S.; Markowitz, M.; Collier, A. C.; Koup, R. A.; Mellors, J. W.; Connick, E.; Conway, B.; Kilby, M.; Wang, L.; Withcomb, J. M.; Hellmann, N. S.; Richman, D. D. *N. Engl. J. Med.* **2002**, *347*, 385–394.
- (38) Paredes, R.; Clotet, B. *Antiviral Res.* **2010**, *85*, 245–265.
- (39) Deville-Bonne, D.; El Amri, C.; Meyer, P.; Chen, Y.; Agrofoglio, L. A.; Janin, J. *Antiviral Res.* **2010**, *86*, 101–120.
- (40) Furman, P. A.; Fyfe, J. A.; St. Clair, M. H.; Weinhold, K.; Rideout, J. L.; Freeman, G. A.; Lehrman, S. N.; Bolognesi, D. P.; Broder, S.; Mitsuya, H.; Barry, D. W. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 8333–8337.
- (41) Arts, E. J.; Wainberg, M. A. *Antimicrob. Agents Chemother.* **1996**, *40*, 527–540.
- (42) Elion, G. B. *J. Antimicrob. Chemother.* **1983**, *12*, 9–17.
- (43) Prichard, M. N.; Keith, K. A.; Johnson, M. P.; Harden, E. A.; McBrayer, A.; Luo, M.; Qiu, S.; Chattopadhyay, D.; Fan, X.; Torrence, P. F.; Kern, E. R. *Antimicrob. Agents Chemother.* **2007**, *51*, 1795–1803.
- (44) Eriksson, S.; Munch-Petersen, B.; Johansson, K.; Eklund, H. *Cell. Mol. Life Sci.* **2002**, *59*, 1327–1346.
- (45) Shealy, Y. F.; Clayton, J. D. *J. Am. Chem. Soc.* **1966**, *88*, 3885–3887.
- (46) Shealy, Y. F.; Clayton, J. D. *J. Am. Chem. Soc.* **1969**, *91*, 3075–3083.
- (47) Kusaka, T.; Yamamoto, H.; Shibata, M.; Muroi, M.; Kishi, T.; Mizuno, K. *J. Antibiot.* **1968**, *21*, 255–263.
- (48) Yaginuma, S.; Muto, N.; Tsujino, M.; Sudate, Y.; Hayashi, M.; Otani, M. *J. Antibiot.* **1981**, *34*, 359–366.
- (49) Hury, D. M.; Okabe, M. *Chem. Rev.* **1992**, *92*, 1745–1768.
- (50) Jeong, L. S.; Lee, J. A. *Antivir. Chem. Chemother.* **2004**, *15*, 235–250.

- (51) Ferrero, M.; Gotor, V. *Chem. Rev.* **2000**, *100*, 4319–4347.
- (52) Wang, J.; Rawal, R. K.; Chu, C. K. *Medicinal Chemistry of Nucleic Acids*; Zhang, L. H.; Xi, Z.; Chattopadhyaya, J., Eds.; 1st ed. Ch.; John Wiley & Sons, 2011; pp. 1–100.
- (53) Marquez, V. E. In *Advances in Antiviral Drug Design*; Elsevier Ltd, 1996; Vol. 2, pp. 89–146.
- (54) Kulikowski, T. *Pharm. World Sci.* **1994**, *16*, 127–138.
- (55) Boutureira, O.; Matheu, M. I.; Díaz, Y.; Castellón, S. *Chem. Soc. Rev.* **2013**, *42*, 5056–5072.
- (56) Vince, R.; Hua, M.; Brownell, J.; Daluge, S.; Lee, F.; Shannon, W. M.; Lavelle, G. C.; Qualls, J.; Weislow, O. S.; Kiser, R.; Canonico, P. G.; Schultz, R. H.; Narayanan, V. L.; Mayo, J. G.; Shoemaker, R. H.; Boyd, M. R. *Biochem. Biophys. Res. Commun.* **1988**, *156*, 1046–1053.
- (57) Vince, R.; Brownell, J. *Biochem. Biophys. Res. Commun.* **1990**, *168*, 912–916.
- (58) Tan, X.; Chu, C. K.; Boudinot, F. D. *Adv. Drug Deliv. Rev.* **1999**, *39*, 117–151.
- (59) Daluge, S. M.; Good, S. S.; Faletto, M. B.; Miller, W. H.; Clair, M. H. S. T.; Boone, L. R.; Tisdale, M.; Parry, N. R.; Reardon, J. E.; Dornsife, R. E.; Averett, D. R.; Krenitsky, T. A. *Antimicrob. Agents Chemother.* **1997**, *41*, 1082–1093.
- (60) Crimmins, M. T.; King, B. W. *J. Org. Chem.* **1996**, *61*, 4192–4193.
- (61) Zoulim, F. *J. Clin. Virol.* **2006**, *36*, 8–12.
- (62) Lee, M.; Lee, D.; Zhao, Y.; Newton, M. G.; Chun, M. W.; Chu, C. K. *Tetrahedron Lett.* **1995**, *36*, 3499–3502.
- (63) Zhao, Y.; Yang, T.; Lee, M.; Lee, D.; Newton, M. G.; Chu, C. K. *J. Org. Chem.* **1995**, *60*, 5236–5242.
- (64) Pierra, C.; Olgen, S.; Cavalcanti, S. C.; Cheng, Y. C.; Schinazi, R. F.; Chu, C. K. *Nucleosides Nucleotides* **2000**, *19*, 253–268.
- (65) Sekiyama, T.; Hatsuya, S.; Tanaka, Y.; Uchiyama, M.; Ono, N.; Iwayama, S.; Oikawa, M.; Suzuki, K.; Okunishi, M.; Tsuji, T. *J. Med. Chem.* **1998**, *41*, 1284–1298.
- (66) Hoshino, H.; Shimizu, N.; Shimada, N.; Takita, T.; Takeuchi, T. *J. Antibiot.* **1987**, *40*, 1077–1078.
- (67) Honjo, M.; Maruyama, T.; Sato, Y.; Horii, T. *Chem. Pharm. Bull.* **1989**, *37*, 1413–1415.
- (68) Norbeck, D. W.; Kern, E.; Hayashi, S.; Rosenbrook, W.; Sham, H.; Herrin, T.; Plattner, J. J.; Erickson, J.; Clement, J.; Swanson, R.; Shipkowitz, N.; Hardy, D.; Marsh, K.; Arnett, G.; Shannon, W.; Broder, S.; Mitsuya, H. *J. Med. Chem.* **1990**, *33*, 1281–1285.

4. References

- (69) Ichikawa, Y.; Narasakab, K. *J. Chem. Soc., Chem. Commun.* **1989**, 1919–1921.
- (70) Wang, J.; Herdewijn, P. *J. Org. Chem.* **1999**, *64*, 7820–7827.
- (71) Wang, J.; Froeyen, M.; Hendrix, C.; Andrei, G.; Snoeck, R.; De Clercq, E.; Herdewijn, P. *J. Med. Chem.* **2000**, *43*, 736–745.
- (72) Marquez, V. E.; Siddiqui, M. A.; Ezzitouni, A.; Russ, P.; Wang, J.; Wagner, R. W.; Matteucci, M. D. *J. Med. Chem.* **1996**, *39*, 3739–3747.
- (73) Ezzitouni, A.; Barchi, J. J.; Marquez, V. E. *J. Chem. Soc., Chem. Commun.* **1995**, 1345–1346.
- (74) Choi, Y.; Sun, G.; George, C.; Nicklaus, M. C.; Kelley, J. A.; Marquez, V. E. *Nucleosides Nucleotides* **2003**, *22*, 2077–2091.
- (75) Comin, M. J.; Agbaria, R.; Ben-Kasus, T.; Huleihel, M.; Liao, C.; Sun, G.; Nicklaus, M. C.; Deschamps, J. R.; Parrish, D. A.; Marquez, V. E. *J. Am. Chem. Soc.* **2007**, *129*, 6216–6222.
- (76) Melman, A.; Wang, B.; Joshi, B. V.; Gao, Z.-G.; De Castro, S.; Heller, C. L.; Kim, S.-K.; Jeong, L. S.; Jacobson, K. A. *Bioorg. Med. Chem.* **2008**, *16*, 8546–8556.
- (77) Marquez, V. E.; Schroeder, G. K.; Ludek, O. R.; Siddiqui, M. A.; Ezzitouni, A.; Wolfenden, R. *Nucleosides Nucleotides* **2009**, *28*, 614–632.
- (78) Marquez, V. E.; Ezzitouni, A.; Russ, P.; Siddiqui, M. A.; Ford, H.; Feldman, R. J.; Mitsuya, H.; George, C.; Barchi, J. J. *J. Am. Chem. Soc.* **1998**, *120*, 2780–2789.
- (79) Marquez, V. E.; Choi, Y.; Comin, M. J.; Russ, P.; George, C.; Huleihel, M.; Ben-Kasus, T.; Agbaria, R. *J. Am. Chem. Soc.* **2005**, *127*, 15145–15150.
- (80) Joshi, B. V.; Moon, H. R.; Fettingner, J. C.; Marquez, V. E.; Jacobson, K. A. *J. Org. Chem.* **2005**, *70*, 439–447.
- (81) Comin, M. J.; Vu, B. C.; Boyer, P. L.; Liao, C.; Hughes, S. H.; Marquez, V. E. *ChemMedChem* **2008**, *3*, 1129–1134.
- (82) Russ, P. L.; Gonzalez-Moa, M. J.; Vu, B. C.; Sigano, D. M.; Kelley, J. A.; Lai, C. C.; Deschamps, J. R.; Hughes, S. H.; Marquez, V. E. *ChemMedChem* **2009**, *4*, 1354–1363.
- (83) Choi, Y.; George, C.; Comin, M. J.; Barchi, J. J.; Kim, H. S.; Jacobson, K. A.; Balzarini, J.; Mitsuya, H.; Boyer, P. L.; Hughes, S. H.; Marquez, V. E. *J. Med. Chem.* **2003**, *46*, 3292–3299.
- (84) Schaeffer, H. J.; Marathe, S.; Alks, V. *J. Pharm. Sci.* **1964**, *53*, 1368–1370.
- (85) Schaeffert, H. J.; Godse, D. D.; Liu, G. *J. Pharm. Sci.* **1964**, *53*, 1510–1515.
- (86) Viña, D.; Santana, L.; Uriarte, E. *Nucleosides Nucleotides* **2001**, *20*, 1363–1365.

- (87) Halazy, S.; Kenny, M.; Dulworth, J.; Eggenspiller, A. *Nucleosides Nucleotides* **1992**, *11*, 1595–1606.
- (88) Maurinsh, Y.; Rosemeyer, H.; Esnouf, R.; Medvedovici, A.; Wang, J.; Ceulemans, G.; Lescrinier, E.; Hendrix, C.; Busson, R.; Sandra, P.; Seela, F.; Van Aerschot, A.; Herdewijn, P. *Chem. Eur. J.* **1999**, *5*, 2139–2150.
- (89) Maurinsh, Y.; Schraml, J.; De Winter, H.; Blaton, N.; Peeters, O.; Lescrinier, E.; Rozenski, J.; Van Aerschot, A.; De Clercq, E.; Busson, R. *J. Org. Chem.* **1997**, *62*, 2861–2871.
- (90) Mikhailov, S. N.; Blaton, N.; Rozenski, J.; Balzarini, J.; De Clercq, E.; Herdewijn, P. *Nucleosides Nucleotides* **1996**, *15*, 869–878.
- (91) Wang, J.; Froeyen, M.; Herdewijn, P. *Adv. Antivir. Drug Des.* **2004**, *4*, 119–145.
- (92) Herdewijn, P.; De Clercq, E. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1591–1597.
- (93) Olivo, H. F.; Yu, J. *J. Chem. Soc. Perkin Trans. 1* **1998**, *3*, 391–392.
- (94) Konkel, M. J.; Vince, R. *Tetrahedron* **1996**, *52*, 799–808.
- (95) Katagiri, N.; Ito, Y.; Shiraishi, T.; Maruyama, T.; Sato, Y.; Kaneko, C. *Nucleosides Nucleotides* **1996**, *15*, 631–647.
- (96) Rosenquist, A.; Kvarnström, I.; Classon, B.; Samuelsson, B. *J. Org. Chem.* **1996**, *61*, 6282–6288.
- (97) Ramesh, K.; Wolfe, M. S.; Lee, Y.; Vander Velde, D.; Borchardt, R. T. *J. Org. Chem.* **1992**, *57*, 5861–5868.
- (98) Perez-Perez, M.-J.; Rozenski, J.; Busson, R.; Herdewijn, P. *J. Org. Chem.* **1995**, *60*, 1531–1537.
- (99) Arango, J. H.; Geer, A.; Rodríguez, J.; Young, P. E.; Scheiner, P. *Nucleosides Nucleotides* **1993**, *12*, 773–784.
- (100) Horváth, A.; Ruttens, B.; Herdewijn, P. *Tetrahedron Lett.* **2007**, *48*, 3621–3623.
- (101) Wang, J.; Viña, D.; Busson, R.; Herdewijn, P. *J. Org. Chem.* **2003**, *68*, 4499–4505.
- (102) Vijgen, S.; Nauwelaerts, K.; Wang, J.; Van Aerschot, A.; Lagoja, I.; Herdewijn, P. *J. Org. Chem.* **2005**, *70*, 4591–4597.
- (103) Gu, P.; Griebel, C.; Van Aerschot, A.; Rozenski, J.; Busson, R.; Gais, H.-J.; Herdewijn, P. *Tetrahedron* **2004**, *60*, 2111–2123.
- (104) Wang, J.; Morral, J.; Hendrix, C.; Herdewijn, P. *J. Org. Chem.* **2001**, *66*, 8478–8482.
- (105) Terán, C.; Santana, L.; Uriarte, E.; Viña, D.; De Clercq, E. *Nucleosides Nucleotides* **2003**, *22*, 787–789.

4. References

- (106) Ferrer, E.; Alibés, R.; Busqué, F.; Figueredo, M.; Font, J.; De March, P. *J. Org. Chem.* **2009**, *74*, 2425–2432.
- (107) Lubineau, A.; Augé, J.; Lubin, N. *Tetrahedron Lett.* **1991**, *32*, 7529–7530.
- (108) Patrick, G. L. *An introduction to medicinal chemistry*; 5th ed.; Oxford University Press, 2013.
- (109) Leach, A. R. *Molecular modelling: principles and applications*; 2nd ed.; Pearson Education, 2001.
- (110) Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. *J. Mol. Biol.* **1982**, *161*, 269–288.
- (111) Comin, M. J.; Vu, B. C.; Boyer, P. L.; Liao, C.; Hughes, S. H.; Marquez, V. E. *ChemMedChem* **2008**, *3*, 1129–1134.
- (112) Figueras, A.; Miralles-Llumà, R.; Flores, R.; Rustullet, A.; Busqué, F.; Figueredo, M.; Font, J.; Alibés, R.; Maréchal, J.-D. *ChemMedChem* **2012**, *7*, 1044–1056.
- (113) Miralles-Llumà, R.; Figueras, A.; Busqué, F.; Alvarez-Larena, A.; Balzarini, J.; Figueredo, M.; Font, J.; Alibés, R.; Maréchal, J.-D. *Eur. J. Org. Chem.* **2013**, 7761–7775.

Chapter II: Objectives

The development of novel chemotherapeutic agents based on carbocyclic nucleoside analogues continues to be of great importance in drug design. In particular, cyclohexenyl and conformationally locked nucleoside analogues, which mimic the conformational behaviour of their endogenous counterparts, seem to be potential as antiviral agents. Some examples, such as both enantiomers of Cyclohexenyl G (DCG and LCG) or *D-iso*-MCT, have been reported to display anti-HSV activity (Figure II-1). In addition, the synthesis of new six-membered carbocyclic nucleosides is still a challenge due to the complexity of their stereoselective synthesis.

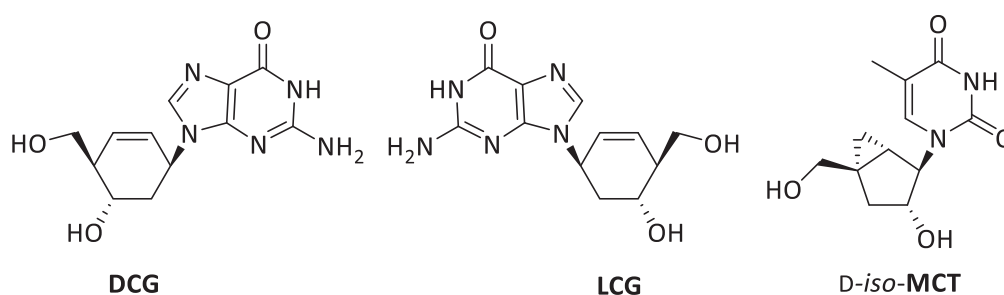


Figure II-1. Nucleoside analogues with potent antiviral activity.

In recent years, protein-ligand dockings have arisen as a powerful tool for the rational design of biologically active compounds, being able to predict the binding affinity of drug candidates to their targets before the synthesis.

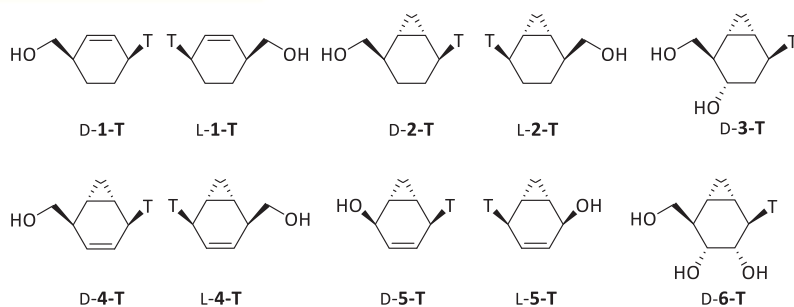
The main objective of this dissertation is to perform the rational design by means of molecular docking and the synthesis of a series of six-membered carbocyclic nucleosides based on the skeleton of D- and L-Cyclohexenyl G as anti-herpes virus agents. This objective can be divided into three different parts that are briefly outlined below.

- **OBJECTIVE 1: Molecular modelling pre-study for the design of anti-HSV novel carbocyclic nucleosides**

Our research group studied the whole activation process of previously synthesised nucleoside analogues by means of protein-ligand docking in order to rationalise their antiviral activity against different viruses, such as HSV. Taking advantage of the knowledge gathered in this study about the mechanism of action of HSV, a first class of cyclohexenyl nucleoside derivatives are envisaged to study their activation process through molecular modelling and select those with higher potential. More specifically, different cyclohexenyl and bicyclo[4.1.0]heptane nucleosides are planned to be investigated as anti-HSV agents (Figure II-2). Cyclopropane-fused derivatives are also selected with the aim of studying the effect of

replacing the double bond of the cyclohexene with a fused cyclopropane, which could confer more flexibility to the carbocycle while still mimicking cyclohexene conformations. Both enantiomers of these nucleosides are planned to be studied considering the precedent of D- and L-Cyclohexenyl G, which were both active. Both pyrimidine and purine candidates are considered, while known antiviral compounds dT, ACV, DCG and LCG will be used as reference compounds to validate the results (benchmarks).

Pyrimidine compounds



Purine compounds

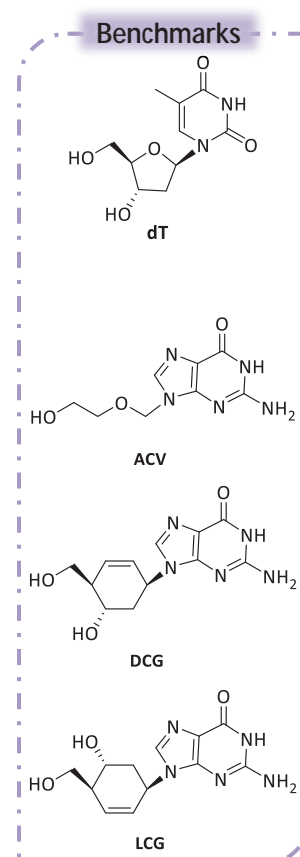
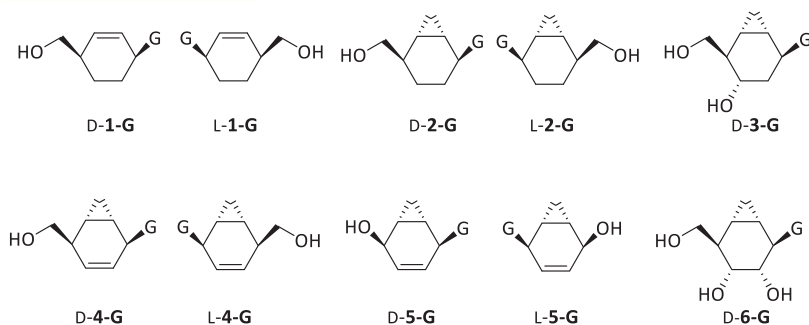


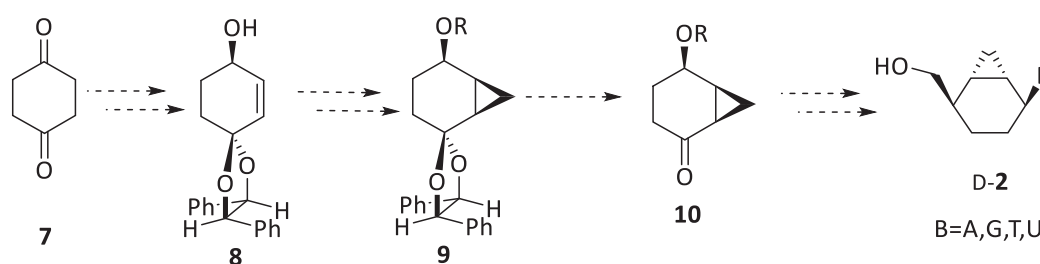
Figure II-2. Nucleoside analogues to be studied as candidates against HSV-1.

With this purpose, we aim to evaluate their drug-likeness and analyse the activation process by which nucleosides are converted into the triphosphorylated derivatives.

The computational study would lead to a list of potential prodrug candidates and after analysing the results, we would select those candidates more prone to be triphosphorylated with the aim of synthesising them.

- **OBJECTIVE 2: Synthesis of bicyclo[4.1.0]heptane nucleoside analogues**

The second objective is devoted to the enantioselective synthesis of bicyclo[4.1.0]heptane nucleoside analogue **2**, following a similar strategy reported by our research group (Scheme II-1). The synthesis would start with a monoprotection of the 1,4-cyclohexanedione **7**, followed by an oxidation to enone and successive asymmetric reduction of the carbonyl to afford allylic alcohol **8**. Then, a cyclopropanation followed by a protection of the alcohol would furnish **9**, which would be converted into the key intermediate **10** after the hydrolysis of the ketal. Subsequent modifications of **10** would lead to bicyclo[4.1.0]heptane nucleoside analogue D-2.



Scheme II-1. Synthetic pathway foreseen to prepare bicyclo[4.1.0]heptane nucleoside analogues.

- **OBJECTIVE 3: Study of antiviral activity of prodrug candidates**

Finally, these synthesised nucleoside analogues would be screened for antiviral activity against different viruses, such as HSV.

In order to rationalise their antiviral activity additional modelling is required to investigate the interaction with the target. In particular, a protein-ligand docking study of the triphosphorylated candidates into the target would be performed.

**Chapter III: Molecular
modelling pre-study for the
design of anti-HSV novel
carbocyclic nucleosides**

1. Introduction

Medicinal chemistry pursues the design and synthesis of novel and effective pharmaceutical agents with a desired biological effect on the human body or some other living system. Drugs have been defined as “compounds which interact with a biological system to produce a biological response”.¹ This biological response is related to several molecular variables such as its absorption, metabolism and interaction with its target. The major drug targets are normally large molecules (macromolecules), such as lipids, carbohydrates, nucleic acids and proteins. These macromolecules have a binding site into which the drug fits (Figure III-1). The study of this interaction and the pharmacological effect that is produced is known as pharmacodynamics.¹

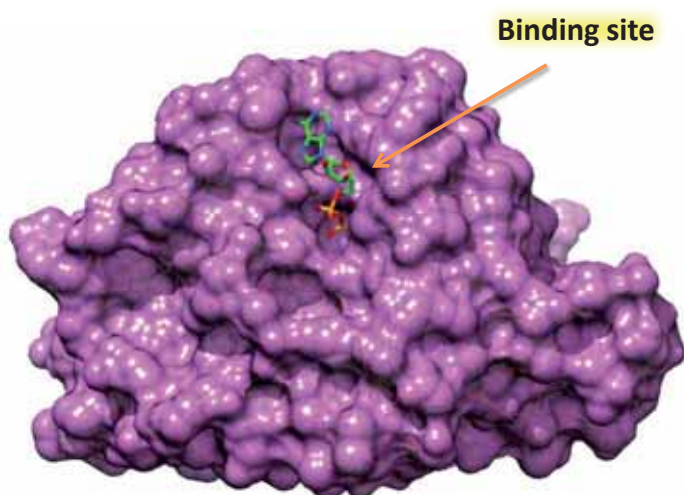


Figure III-1. Representation of the surface of a protein (in purple) with the crystallized ligand (in green; PDB code: 1VTK).

Nucleoside analogues as those that are objective of this work are prodrugs. This means that, they must be converted into their active form in order to interact with their physiological target. This process of activation consists of three phosphorylation steps catalysed by different kinases.² Once nucleoside analogues have been triphosphorylated, successful candidates interact with the viral polymerase to gain antiviral activity.

As stated in Chapter I, the development of new antiviral agents is time consuming. However, the application of computational techniques in medicinal chemistry during the last decades has allowed the study of biological molecules and their biological processes at the atomic level. Molecular modelling techniques have become a useful tool in drug discovery and development.³

2. Molecular modelling pre-study for the design of anti-HSV novel carbocyclic nucleosides

In general, computational approaches for the study of the biological activity of chemical compounds can be divided into two branches: ligand-based and structure-based.⁴ Ligand-based approaches, such as quantitative structure-activity relationship (QSAR) models, rely on the knowledge of a set of chemicals for which certain activities towards a given target are known (binding, inhibition values, etc.). Based on physico-chemical descriptors related to these ligands, one could dress predictive models without structural knowledge of the target. On the other hand, structure-based approaches, such as protein-ligand docking, may be used when the three-dimensional structure of the biological target is available, and the substrate activity can be extrapolated from the interactions of the ligand in the binding site. Those modelling provide with a real size model of the interaction and consider both chemical and biological partners. However, their viability is extremely dependent on the amount of experimental structures of the targets available for the project.

The Protein Data Bank (PDB, <http://www.rcsb.org/pdb/>) is a worldwide archive of structural data of large biological molecules, including proteins and nucleic acids, established at Brookhaven National Laboratories (BNL) in 1971 as an archive for biological macromolecular crystal structures. Each data entry contains the atomic coordinates of the structure as well as pertinent information of the structure such as species from which the molecule has been obtained, experimental details or related literature citations. The Protein Data Bank has become a milestone in molecular modelling, since all the crystallographic protein–ligand structures used in the molecular modelling studies have been retrieved from this repository.

The activity of a drug relies on its bioavailability, which is defined as the amount of drug that is actually absorbed from a given dose. In order to avoid unnecessary work synthesising inactive molecules, Lipinski *et al.* proposed a set of four rules that would predict whether a molecule was likely to be orally bioavailable.⁵ These rules were derived from an analysis of 2245 compounds from the World Drug Index (WDI) aimed at identifying features that were important in making a drug orally active. The Lipinski's rules state that a good candidate compound should present:

- a molecular mass less than 500
- a calculated value of logP less than 5 (P is the octanol-water partition ratio used to represent molecular lipophilicity)
- less than 10 hydrogen bond acceptor groups
- less than 5 hydrogen bond donor groups

These four physico-chemical parameter ranges were named the “rule of 5” because the cut-off points for each of the four parameters were all multiple of 5. This ligand-based approach has been widely used in the synthesis of new drugs as a useful tool, but it does not always work. Some orally active therapeutic compounds, such as some antibiotics, antifungals, vitamins and cardiac glycosides, do not obey the “Lipinski’s rule of five”.

Once the drug absorption is predicted, the interaction with its biological target may be analysed by means of molecular modelling. The next section outlines a few general principles about molecular modelling techniques aimed at modelling biological macromolecules and their interaction with small molecules. A more extensive description of these techniques is given in the computational methods.

1.1. Molecular modelling of biological macromolecules

Molecular modelling is used to calculate the structure, the energy and any related properties of molecules. Methods used in this field can be mainly split into two categories: molecular mechanics (MM) and quantum mechanics (QM). Molecular mechanics use equations based on classical physics to calculate force fields and apply them to nuclei without considering the electrons. Molecules are treated as charged spheres (the atoms) joined together by springs (the bonds). By contrast, quantum mechanics use equations based on quantum physics to calculate the properties of a molecule by considering the interactions between electrons and nuclei. Unlike molecular mechanics, atoms are not treated as solid spheres and electrons are included in the calculations. As a result, MM is faster and less intensive on computer time than QM but it cannot treat the reactivity of the system.

The method of calculation chosen depends on what calculation needs to be done as well as the size of the system to be studied. Thus, molecular mechanic methods are normally used for the study of macromolecules containing thousands of atoms such as proteins, whereas quantum mechanics are normally restricted to small systems. In the context of macromolecules, due to the high degrees of freedom of the system, some approximations should be considered in order to reduce the computational cost. In these cases, a common solution consists in using simplified force fields centered on non-covalent terms (scoring functions) and reducing the number of degree explored during the conformational search. Molecular docking, which is based on these premises, is one of the most widely used techniques on this field.⁶⁻⁹

Molecular docking can be defined as the prediction of the structure of a complex of two or more molecules. Nowadays, molecular docking has a wide variety of applications in drug discovery, such as structure-activity studies, lead optimization and the discovery of potential leads by virtual screening, since it offers a relatively fast and economic alternative to standard experimental techniques, allowing the *in silico* prediction of the binding modes and affinities for molecular recognition.

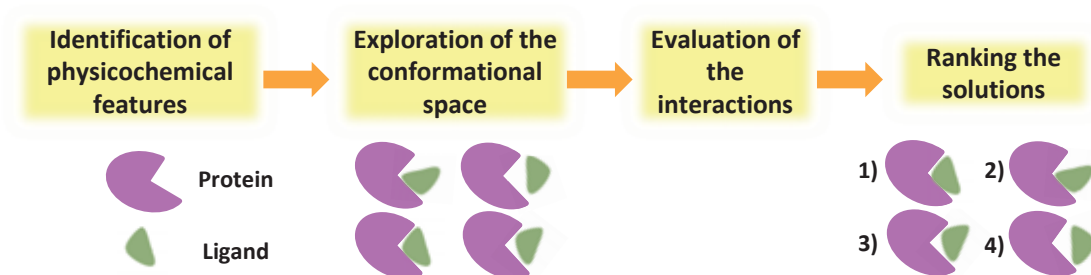
Normally, molecular docking is performed between a small molecule (the ligand) and a target macromolecule (the protein, also called receptor), which is known as a protein-ligand docking. Pioneered during the early 1980s,¹⁰ it remains an active area of research as it has demonstrated to be a valuable tool for drug discovery programs.¹¹⁻¹³

The main goal of the present chapter is to perform a rational design of novel nucleoside analogues as antiviral agents based on a series of computational approaches. As most of this work will be carried out using protein-ligand docking, in the next subsections, a brief overview of docking techniques is presented, including their limitations.

1.1.1 Protein-ligand docking

The aim of protein–ligand docking is to predict the most favoured ligand conformations and orientations (commonly referred as poses, binding modes or, simply, solutions) when it is bound to a given target protein and estimate the binding affinity. This way, we can obtain a model of the ligand-protein complex structure, which can be used for the prediction of the biological activity of the system.

A docking simulation can be divided into four different processes: (1) identification of the physicochemical features of the binding site and the ligand; (2) exploration of the conformational space available to the ligand in the binding site and subsequent generation of protein–ligand complexes; (3) calculation of the binding energies between the ligand and the protein; and (4) ranking the solutions using a scoring function (Scheme III-1). Steps (2) and (3) are closely related, since most docking programs calculate the binding energies of the orientations while they are being generated, to guide the search towards the best poses.



Scheme III-1. Schematic representation of a docking procedure.

Docking protocols can be described as a combination of two components: a search algorithm and a scoring function. The search algorithm is related to the examination of the conformational space in the receptor and the generation of a number of ligand orientations in the protein binding site whereas the scoring function is designed to evaluate and rank the binding modes predicted on the basis of the search algorithm. Then these poses have to be analysed to evaluate if they are in good agreement with the known requirements of the target protein. The mathematical details and principles of these algorithms are described in more detail in the computational methods.

1.1.1.1. Limitations

Despite protein–ligand dockings have become a useful tool in drug discovery, they present several limitations. Even though dockings approaches have achieved several improvements in both search algorithms and scoring functions, one of the major limiting factors are weak accuracies in the scoring function. This limitation is due to the numerous simplifications that are assumed to allow the computational evaluation of the ligand affinity of an extremely large number of poses. For instance, in most scoring functions entropy, solvent effects and electrostatic interactions are completely neglected or at least not fully accounted for.^{8,9} It is true that in some scoring functions rotational entropy^{14,15} and solvation^{16–18} contributions have been included in order to tackle these limitations, but these terms are only partial descriptions of the real entropic and solvation effects occurring on protein–ligand binding.

In addition, there are also some limitations when water molecules and metals are important in protein–ligand interactions. Despite the fact that water molecules and metals can play an essential role in ligand–protein binding, they are normally not taken into account when performing docking studies.^{8,11} The most successful docking algorithms are nowadays able to introduce explicit water molecules in their runs,^{19–21} and many programs also have optimized metal ion parameters to take them into consideration during the calculations.^{22–24} However,

dealing with metals and water in protein binding sites is still one of the major challenges in this field.

The flexibility of the protein receptor is another relevant limitation of current molecular docking programs.^{6,8,25–28} An ideal docking should consider both the protein and the ligand flexible, which means that a large number of degrees of freedom are needed: translation and rotation of one molecule relative to another, which involves 6 degrees of freedom, as well as the internal conformational degrees of freedom of both the ligand and the receptor. This is impractical due to the size of the search space, so several docking methods follow the assumption that protein structures are rigid entities and thus there are no structural changes during the binding process, which is based on the “lock and key” theory postulated by Emil Fischer in 1894 (Figure III-2).²⁹ In this model, Emil suggested that both the enzyme and the substrate have specific geometric shapes that fit exactly into each other. However, although this model explains enzyme specificity, it does not explain the stabilization of the corresponding transition state that enzymes achieve. In 1958, Daniel Koshland suggested a slight modification to this theory and postulated the “induced fit” theory (Figure III-2).³⁰ He proposed that the binding of the substrate induces conformational changes in the active site of the protein, due to multiple weak interactions with the substrate. Nowadays, more modern theories based on the “conformational selection” model are well established, in which the protein is described as an ensemble of differently populated conformations in equilibrium, and postulating that the ligand selects the most favoured conformation.^{27,31–33} It is worth noting that this specific conformation does not have to be the most populated.

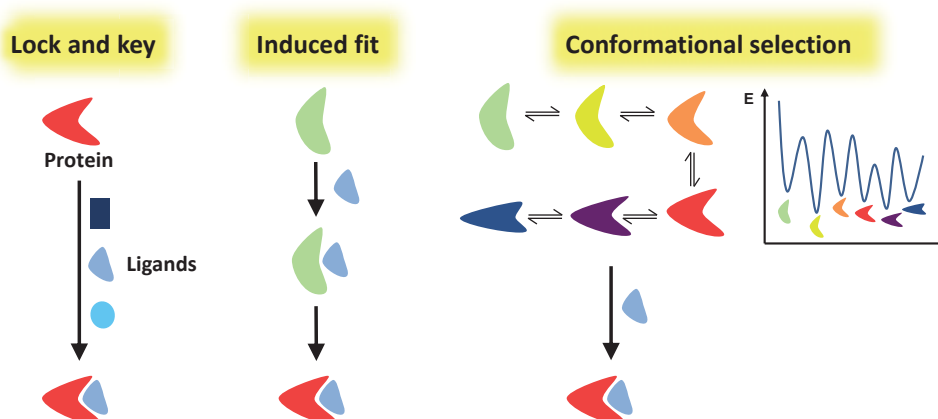


Figure III-2. Simplified representation of the three main models for protein–ligand binding.

In order to model protein flexibility several approximations have been applied,^{6,25–27} which can be divided in two main groups: models in which only the binding site of the protein is set flexible (local) and models which simulate the flexibility as a whole (global). The soft-

docking method was the first one able to include local protein flexibility by decreasing the van der Waals repulsion energy term between the atoms in the binding site and those in the ligand which allows some overlap between the ligand and the protein (Figure III-3).³⁴ However, this method can account for only small conformational changes. Another approximation to include partial protein flexibility consists in the use of a library of discrete rotameric states for each type of side chain (Figure III-3).³⁵

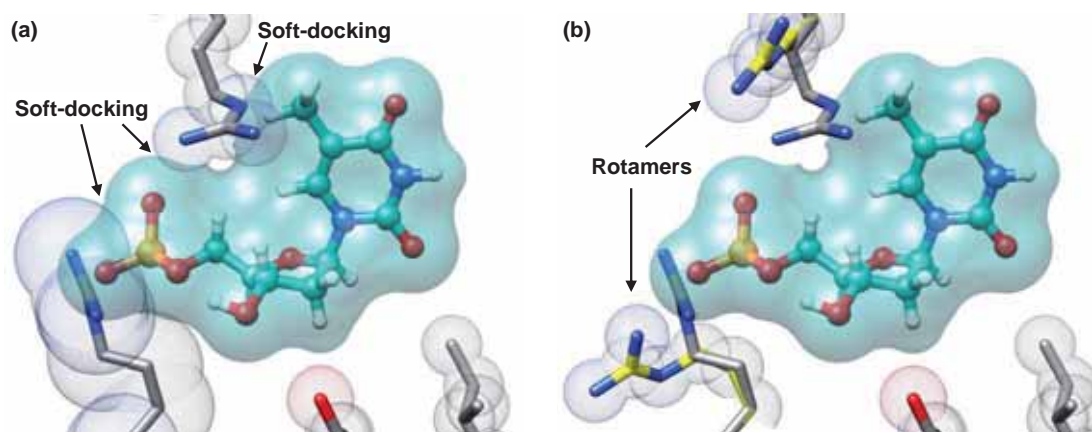


Figure III-3. Methods developed to model local protein flexibility: (a) soft-docking and (b) use of rotamers (original residues shown in grey and rotamers in yellow).

Regarding full protein flexibility, there are several approaches to deal with that problem. One of the most widely used approaches consists in a generation of an ensemble of rigid protein structures to represent different possible conformational changes of the protein.⁶ These structures can be obtained from multiple crystal structures, NMR studies or be generated by computational techniques, such as molecular dynamic simulations and Monte Carlo simulations.²⁶

Despite these limitations, molecular dockings are nowadays widely used in medicinal chemistry because they are capable of predicting known ligand binding modes with average accuracies of about 1.5-2.0 Å and success rates in the range of 70-80%.⁸

1.1.1.2. Protein-ligand docking: inhibition versus activation

Protein–ligand docking is a widely-used computational tool in drug discovery for the design of novel potential drug candidates as inhibitors of enzymes.³⁶ In this field, dockings are normally applied in two different ways: to study the interactions of a given compound with its biological target and thus design new drug candidates based on these interactions, or to evaluate the interactions between a large set of compounds and a biological target, which is used to select only a small number of them as possible leads.^{37–39}

2. Molecular modelling pre-study for the design of anti-HSV novel carbocyclic nucleosides

Although dockings are widely used to study binding interactions, they are not optimised for catalytic processes since they are based on the use force fields to calculate binding affinities and thus cannot study reactivity (i.e. no transition states can be determined). Another limitation is when metal ions are involved in catalysis. It is really complex to consider metal ions in catalytic docking procedures and they are only taken into account structurally.

However, dockings can be used to catch pre-catalytic structures if enough structural knowledge is available on the catalytic mechanism and assuming into the computation that: 1) pre-reactive structures should not be far for the main binding mode (one of the lowest energy solutions could react) and 2) the reactive groups are well located (residues, phosphates, etc.) so that the reaction occurs. The use of docking into the prediction of catalytically competent orientations is extremely increasing but still remains a complex exercise.

An additional complexity in antiviral nucleoside prodrugs is their activation through successive catalytic processes, since they required an analysis of the binding affinity of the prodrugs to the enzymes as well as the catalytic features of the whole activation process. In other words, not only the binding affinity of the prodrugs to the enzyme is important for their activation but also the pre-catalytic orientation of the ligand into the binding site for several enzymes one step at the time. Nowadays, protein-ligand dockings are also used in multi-target approaches, in which one or more compounds are docked into different receptors.³⁶ The activation process of prodrugs such as nucleoside analogues may be studied as a cascade docking, which means all the activation steps could be successively studied considering multi-receptor approaches.

2. Molecular modelling pre-study for the design of anti-HSV novel carbocyclic nucleosides

Herpes simplex virus (HSV) infections are still among the most frequent human diseases despite many nucleoside analogues are currently used for their treatment.⁶ In this field, cyclohexenyl nucleosides are a promising class of antiviral compounds, wherein replacement of the oxygen atom of the furanose ring by a double bond induces annular flexibility, similar to that of the regular nucleoside.⁴³ As previously mentioned, it has been reported that both enantiomers of Cyclohexenyl G (DCG and LCG) display antiviral activity against some herpes viruses (HSV-1, HSV-2, VZV, CMV) (Figure III-4).⁴⁴

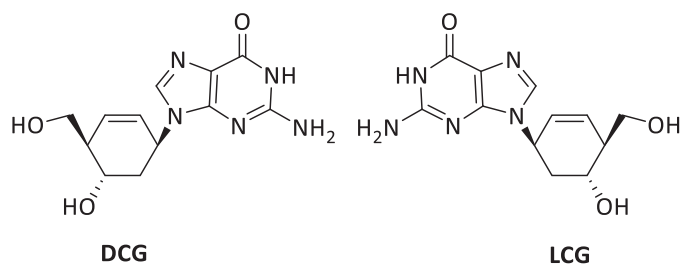


Figure III-4. Structures of D- and L-Cyclohexenyl G.

Taking advantage of a preliminary computational study for the rational design of carbocyclic nucleoside analogues as anti-HSV agents carried out in our research group,⁴⁵ and the knowledge that the conformational behaviour of the cyclohexene ring is similar to that of the natural furanose, we decided to perform a rational drug design of a series of *carba*-nucleosides based on the skeleton of a cyclohexene moiety (Figure III-5) and study the effect of different modifications, such as the fusion of a cyclopropane ring to the cyclohexene, which could confer more flexibility to the carbocycle while still mimicking cyclohexene conformations (compounds type 3, Figure III-5). Other structural features to be evaluated would be the simultaneous presence of an insaturation and a cyclopropane in the cyclohexane ring (compounds type 4 and 5, Figure III-5) and the direct attachment of the hydroxyl moiety to the carbocycle (compound type 5, Figure III-5) discarding the methylene linker unit present in Cyclohexenyl G. Regarding the precedent of Cyclohexenyl G, both enantiomers of these nucleosides were also suggested to be studied.

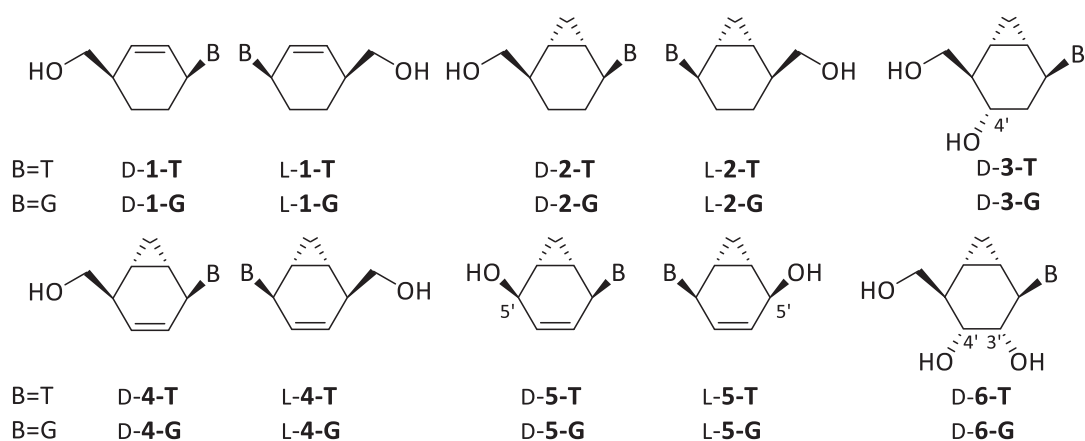


Figure III-5. Nucleoside analogues for the rational design.

2.1. Computational details

Protein-ligand dockings were used to predict all binding modes and energies. Those calculations were performed with the docking program GOLD (version 5.2.2),⁴⁶ and the scoring

2. Molecular modelling pre-study for the design of anti-HSV novel carbocyclic nucleosides

function used was *ChemPLP*.⁴⁷ Molecular graphics and visualization of docking results were performed with the *UCSF Chimera* package.⁴⁸

The three-dimensional structures of the ligands were initially optimized using the *Marvin* work package⁴⁹ and the Merck molecular force field (MMFF) minimization.⁵⁰ Some modifications on the available PDB structures of the proteins must be performed with *UCSF Chimera* before carrying out the docking calculations. When the structures presented more than one identical subunit, duplicate parts were removed. All crystallographic waters, ions and ligands were also deleted from each enzyme. Hydrogen atoms were added, the protonation state of the histidine residues was checked manually and atom charges were assigned for the Amber force field using the Antechamber plugin of UCSF Chimera.⁵¹

The centre of the binding site in the X-ray structures was used as the central point of the docking cavity, the radius of which varies for each enzyme (Table III-1). A series of residues in the binding pocket were set as flexible (Table III-1) by using the rotameric rotation scheme implemented in GOLD program.⁵² Ligand flexibility was also considered in all the cases.

Table III-1. PDB structures and computational details of the docking calculations.

Enzyme	PDB code	Chain	Crystallized ligand	Centre of the cavity (radius)	Flexible residues
HSV-1 TK	1KIM	A	dT	Tyr-172, CD2 (20 Å)	His-58, Lys-62, Glu-83, Arg-222
HSV-1 TK	2KI5	A	ACV	Tyr-172, CD2 (20 Å)	His-58, Lys-62, Glu-83, Arg-222
HSV-1 TK	1VTK	A	dTMP (+ ADP)	Tyr-172, CD2 (20 Å)	His-58, Lys-62, Glu-83, Arg-222
Mouse GMPK	1LVG	A	GMP (+ ADP)	Asp-101, CG (10 Å)	Ser-37, Arg-41, Tyr-53
<i>Dictyostelium discoideum</i> NDPK	1NDC	A	dTDP	Lys-16, NZ (11 Å)	-
Human NDPK	1NUE	AB	GDP	Lys-12, NZ (13 Å)	-

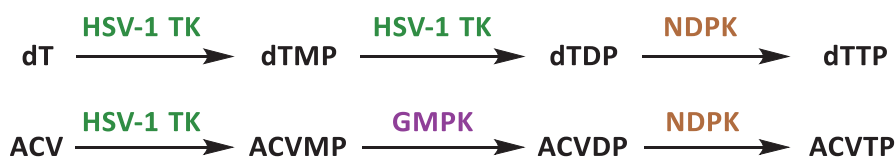
Finally, each nucleoside was docked into each enzyme and 20 predicted orientations were obtained and ranked (50 predicted poses for the third phosphorylation step for both pyrimidine and purine compounds). The results were analysed both in structural and energetic terms, checking how many binding modes (from the 20 poses predicted for each compound)

were consistent with a catalytic orientation, and if their corresponding binding energies were similar or even lower than those of the reference compounds (benchmarks). We therefore consider that those orientations are consistent with the catalysis.

2.2. Activation process

As mentioned in the introduction of the present dissertation, nucleoside analogues are prodrugs that must be converted into their triphosphorylated derivatives by three different kinases before interacting with their biological target.

In the case of HSV, the first phosphorylation step is carried out by a nucleoside kinase coded by the virus itself, which enables the selectively recognition of nucleoside analogues by the viral enzyme but not by its human counterpart.^{53,54} The whole activation process in HSV-1 infected cells is catalysed by the following kinases: HSV-1 thymidine kinase (HSV-1 TK) for the first phosphorylation step; HSV-1 TK and human guanylate kinase (GMPK) for the second phosphorylation of pyrimidine and purine derivatives, respectively, and human nucleoside diphosphate kinase (NDPK) for the third phosphorylation (Scheme III-2).



Scheme III-2. Kinases involved in the activation of pyrimidine (represented by dT) and purine analogues (represented by ACV) in HSV-1 infected cells.

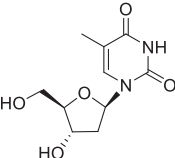
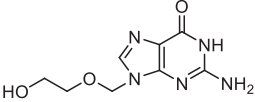
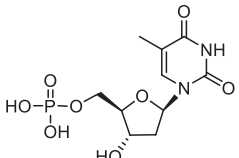
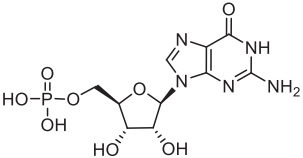
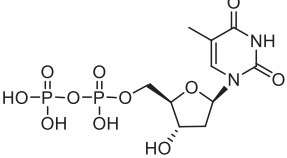
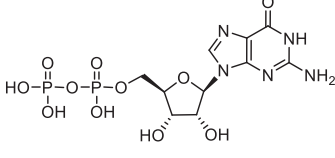
Once nucleosides are triphosphorylated, they have to interact with HSV-1 DNA polymerase, which is the enzyme responsible for the DNA replication of HSV-1, to block the viral replication.⁵⁵

Before studying the activation process of compounds **1-6** proposed, their drug-likeness should be evaluated. The evaluation of compounds **1-6** was carried out according to the “Lipinski’s rule of 5”. The standard molecular descriptors of all these compounds as well as dT, ACV and both enantiomers of Cyclohexenyl G, were calculated with the ChemBioOffice package.⁵⁶ The full series of compounds responds positively to the criteria, which supports their good drug-likeness (Table VII-1, Chapter VII).

Once their drug-likeness was studied, protein–ligand docking calculations on HSV-1 TK, GMPK and NDPK were performed using crystallographic structures available with natural ligands or approved antiviral drugs (Table III-2).

2. Molecular modelling pre-study for the design of anti-HSV novel carbocyclic nucleosides

Table III-2. PDB structures used to study the activation process of novel nucleoside analogues.

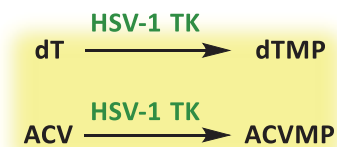
Enzyme	Organism	PDB code	Crystallized ligand	Docking calculations ^[a]
TK	HSV-1	1KIM	 <p>dT</p>	D-1-T, L-1-T, D-2-T, L-2-T, D-3-T, D-4-T, L-4-T, D-5-T, L-5-T and D-6-T
TK	HSV-1	2KI5	 <p>ACV</p>	D-1-G, L-1-G, D-2-G, L-2-G, D-3-G, D-4-G, L-4-G, D-5-G, L-5-G and D-6-G
TK	HSV-1	1VTK	 <p>dTMP (+ ADP)</p>	D-1-TMP, L-1-TMP, D-2-TMP, L-2-TMP, D-3-TMP, D-4-TMP, L-4-TMP, D-5-TMP, L-5-TMP and D-6-TMP
GMPK	<i>Mus musculus</i>	1LVG	 <p>GMP (+ ADP)</p>	D-1-GMP, L-1-GMP, D-2-GMP, L-2-GMP, D-3-GMP, D-4-GMP, L-4-GMP, D-5-GMP, L-5-GMP and D-6-GMP
NDPK	<i>Dictyostelium discoideum</i>	1NDC	 <p>dTDP</p>	D-1-TDP, L-1-TDP, D-2-TDP, L-2-TDP, D-3-TDP, D-4-TDP, L-4-TDP, D-5-TDP, L-5-TDP and D-6-TDP
NDPK	<i>Homo sapiens</i>	1NUE	 <p>GDP</p>	D-1-GDP, L-1-GDP, D-2-GDP, L-2-GDP, D-3-GDP, D-4-GDP, L-4-GDP, D-5-GDP, L-5-GDP and D-6-GDP

^[a] MP and DP stand for monophosphorylated and diphosphorylated compounds, respectively.

The series of nucleoside analogues **1-6** were docked into the active site of each kinase. Calculations with dT, ACV, DCG and LCG were also performed to provide with structural and energetic benchmarks. Docking results were analysed both in structural and energetic terms, checking how many binding modes (from the 20 poses predicted for each compound) were

consistent with a catalytic orientation, and if their corresponding binding energies were similar or even lower than those of the reference compounds (benchmarks).

2.2.1. 1st phosphorylation step



The first phosphorylation step is catalysed by herpes simplex virus type 1 thymidine kinase. HSV-1 TK (EC number: 2.7.1.21) is a key enzyme in the HSV metabolism, being responsible for promoting the transfer of the γ -phosphate group from ATP to the 5'-OH group of thymidine. In contrast to other thymidine kinases that are highly specific, HSV-1 TK acts as phosphorylating agent toward a wide variety of nucleoside analogues,^{53,54,57,58} is able to phosphorylate not only pyrimidine nucleosides⁵⁹⁻⁶¹ but also purine analogues,^{53,54,62} to accept a large diversity of sugar moieties^{42,53,54,61-63} as well as catalyses the introduction of the first and second phosphate to pyrimidine analogues.^{2,59,64} ATP is the common phosphate donor of this enzyme, but it also shows high affinity for others, such as cytidine, uridine and guanosine triphosphate and their deoxy analogues.⁶⁵

2.2.1.1. HSV-1 TK crystallographic structures

To date, there are almost 40 crystallographic structures of HSV-1 TK available: in its apo and holo forms^a as well as for wild type and mutagenic forms. Those holo forms containing dT and ACV were selected for the study of pyrimidine and purine analogues, respectively (Table III-1). These crystallographic structures containing a substrate represent a good approximation to take into account the required structural pre-organization of the enzyme for the catalysis.

HSV-1 TK is a homodimeric enzyme with 376 residues per monomer.^{57,66} Each subunit consists of 13 α -helices, two 3_{10} -helices and seven β -sheets (Figure III-6). The five-stranded parallel β -sheet forms part of the core of the enzyme, which contains the active site.^{57,67} This active site consists of a nucleoside binding region and an ATP.⁶⁸ It is worth noting that while in most of the kinases the donor and the acceptor normally bind in a non-ordered manner, in HSV-1 TK dT is first bound and so the first and second phosphorylation steps can occur successively with the phosphate acceptor remaining at its position.^{65,66}

^a The apo form is the structure of the protein without any ligand bound whereas the holo form is the structure of the protein in complex with the ligand or a cofactor.

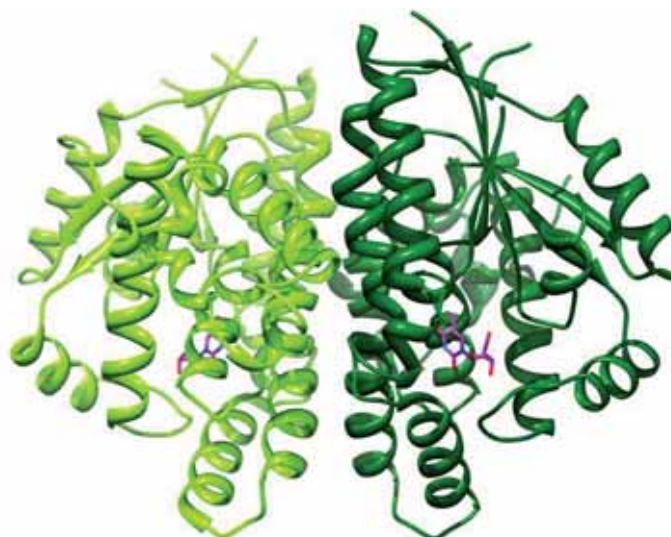


Figure III-6. HSV-1 TK crystallized with dT (pink) in its binding site (PDB code: 1KIM).

The interactions between dT and HSV-1 TK in the binding site (PDB code: 1KIM)⁵⁷ are detailed in Figure III-7. The nucleobase moiety is sandwiched between Met-128 and Tyr-172,⁶⁸ and it is stabilized by pairwise hydrogen bond interactions with Gln-125 as well as by two water-mediated hydrogen bonds with the side chain of Arg-176. The 5-methyl group of the pyrimidine ring is placed in a hydrophobic environment formed by Tyr-132, Ala-167 and Ala-168. The sugar moiety interacts with the protein via its hydroxyl groups. The 3'-OH makes hydrogen bond interactions with Tyr-101 and Glu-225, whereas its 5'-OH is hydrogen bonded to Arg-163 and Glu-83.

It is worth highlighting that Glu-83 is the responsible for the activation of the 5'-OH by deprotonating it to make it a better nucleophile,^{58,68,69} so that phosphorylation can take place.

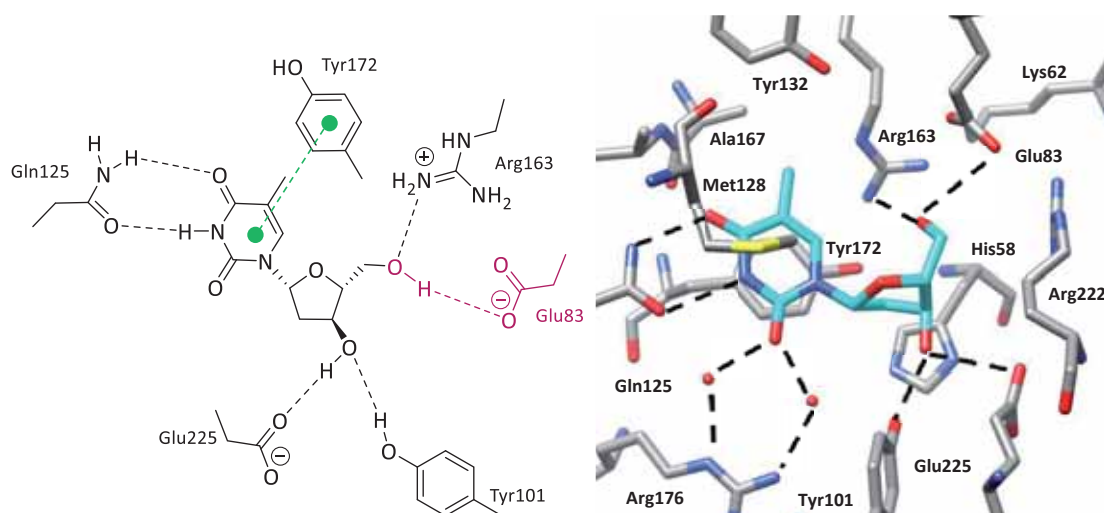


Figure III-7. Representation of the main interactions of dT in HSV-1 TK binding site (PDB code: 1KIM). Hydrogen bonds between dT and residues are depicted as dotted lines and the catalytic residue Glu-83 is depicted in maroon. Images generated using ChemBioOffice⁵⁶ software (left) and UCSF Chimera⁴⁸ package (right).

ACV was chosen as a reference for the study of purine analogues. The binding mode of ACV (PDB code: 2KI5)⁷⁰ is similar to that of dT (Figure III-8). The main difference is that Gln-125 has rotated 180° its amide to interact via hydrogen bond with the purine ring. According to the X-ray structure, the acyclic moiety of ACV can assume two distinct orientations, each of them matching to the previously described poses of the 3'-OH and 5'-OH of dT.⁷⁰ But knowing that Glu-83 activates the hydroxyl, only the conformation in which the OH is pointing towards Glu-83 is considered a phosphorylation orientation.

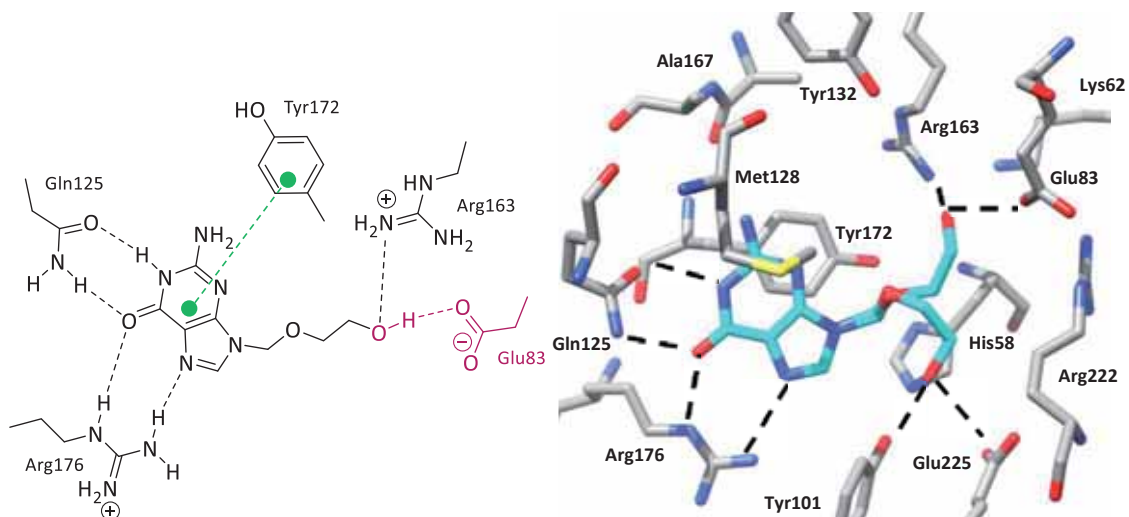


Figure III-8. Representation of the main interactions of ACV in HSV-1 TK binding site (PDB code: 2KI5). Hydrogen bonds between ACV and residues are depicted as dotted lines and the catalytic residue Glu-83 is depicted in maroon. Images generated using ChemBioOffice⁵⁶ software (left) using UCSF Chimera⁴⁸ package (right).

Both HSV-1 TK structures described do not contain ATP in the binding site but a sulphate anion, which is located in the putative position of the β -phosphate of ATP, close to the 5'-OH function of dT.⁷¹

Summarising, it is postulated that the interaction energy between the substrate and HSV-1 TK derives mainly from the interaction with six residues: Glu-83, Tyr-101, Gln-125, Met-128, Tyr-172 and Glu-225.^{58,67,72} Thus, the predicted orientations must present most of these interactions with the substrate to be compatible with the catalysis, including a close enough position of the 5'-OH to Glu-83.

2.2.1.2. Docking results

The docking protocol was validated by carrying out docking calculations of the crystallized ligands dT and ACV into the corresponding HSV-1 TK X-ray structures. The lowest energy poses were perfectly overlapped to the ligand poses in the crystallographic structures, and thus dT and ACV were used as benchmarks for pyrimidine and purine analogues, respectively. The corresponding binding energies are -70 score units for dT and -63 score units

for ACV. The higher binding energy values of ACV compared to dT is due to the low affinity to HSV-TK.^{57,62,70}

▪ Pyrimidine nucleoside analogues

Most of the predicted binding modes of D-1-T and L-1-T are compatible with the catalysis (Figure III-9). The lack of interactions with Glu-225 and Tyr-101 due to the absence of a 3'-OH is not considered restrictive for the catalysis, since the same behaviour is observed for ACV.⁷⁰ In the case of L-1-T, the majority of the binding modes overlapped with the natural ligand dT, whereas for D-1-T, most of the poses showed a slight displacement of the base due to the flexibility of some residues of the binding site, mainly Glu-83, which actually forms a hydrogen bond with the ligand and is the catalytic residue. Similar binding energies are calculated for both D- and L- enantiomers of 1-T with respect to dT, being -70 and -67 score units, respectively.

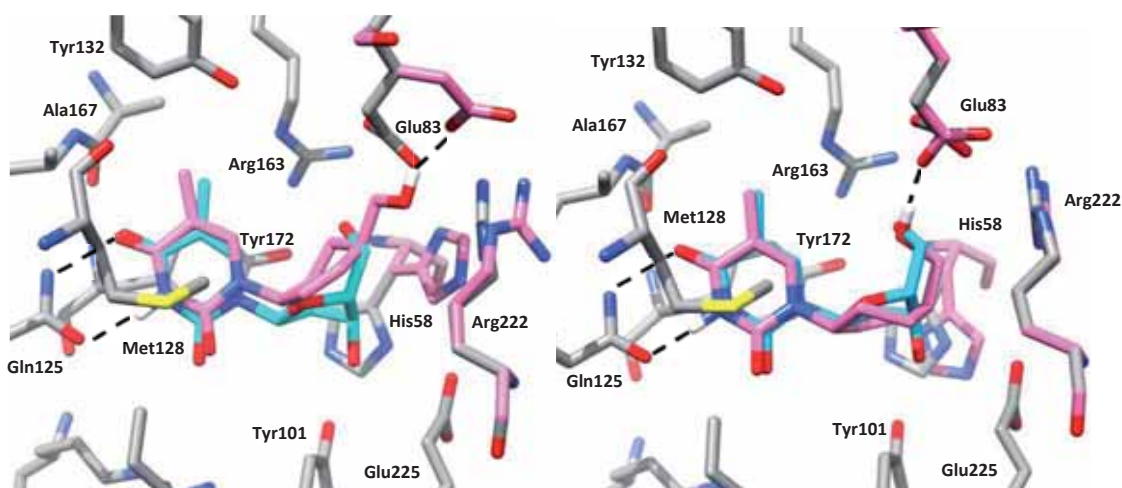


Figure III-9. D-1-T (left) and L-1-T (right) in pink superimposed to crystallographic dT (blue) in HSV-1 TK (PDB:1KIM, X-ray residues shown in grey). Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are not shown and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

Similarly, most of the predicted complexes of D-2-T and L-2-T are properly posed for the catalytic activity of the enzyme (Figure III-10), and their binding energies are similar or even lower to those of dT, being -73 and -70 score units, respectively. Again, there is a slight displacement of the base and the carbocycle but the hydroxymethyl group is still orientated toward Glu-83, which is responsible for its activation. These results showed that the presence of a fused cyclopropane ring does not alter the binding mode of the nucleosides in this enzyme.

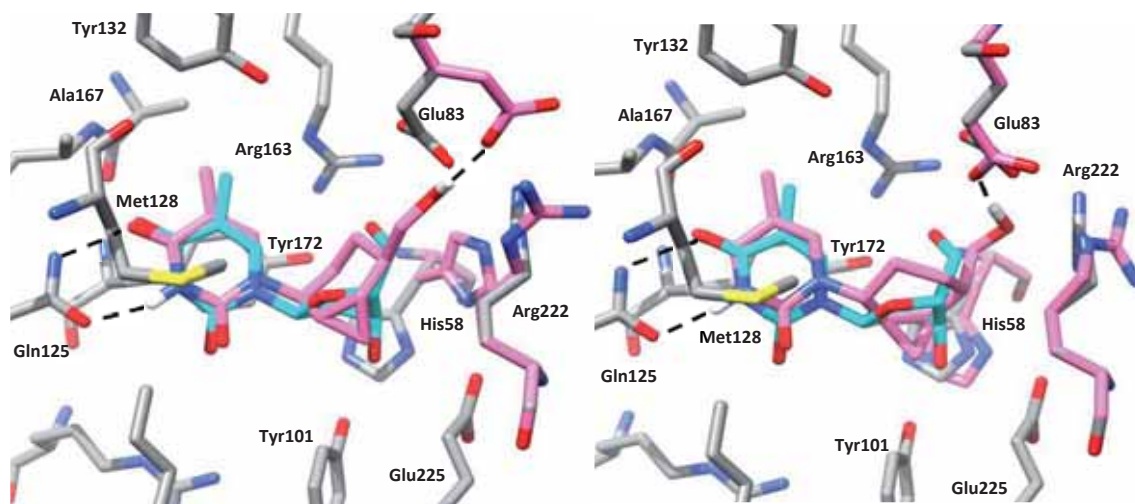


Figure III-10. D-2-T (left) and L-2-T (right) in pink superimposed to crystallographic **dT** (blue) in HSV-1 TK (PDB:1KIM, X-ray residues shown in grey). Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are not shown and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

Docking calculations on D-3-T showed that it is likely to be phosphorylated and the corresponding binding energy is -73 score units, which is lower than the binding energy of the benchmark. Although this compound presents a 4'-OH which resembles the 3'-OH of the natural furanose, there is no hydrogen bond interaction with Tyr-101 and Glu-225 but with Arg-163.

Concerning D-4-T and L-4-T, the binding energies are similar to those of dT, being -68.9 score units in both cases, and their predicted complexes are consistent with the catalysis (Figure III-11).

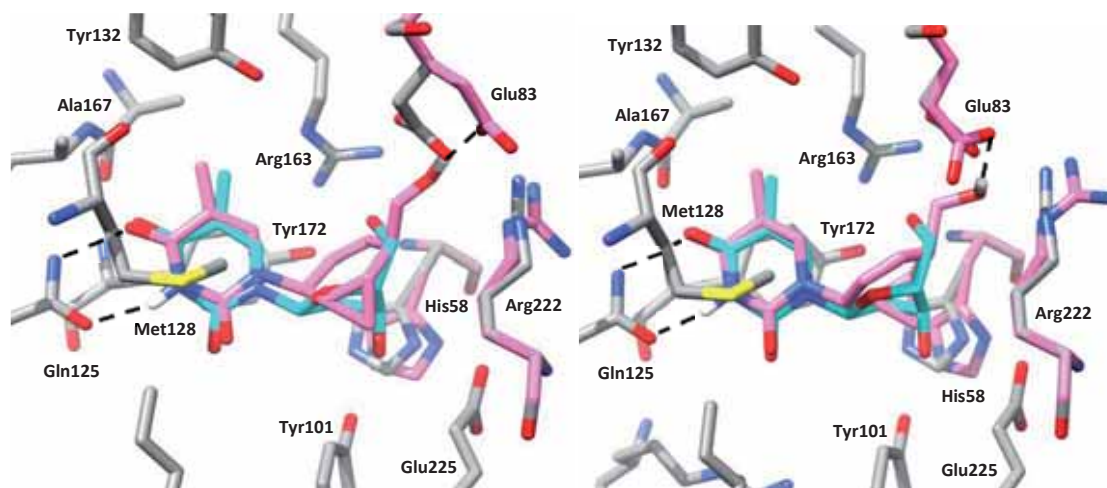


Figure III-11. D-4-T (left) and L-4-T (right) in pink superimposed to crystallographic **dT** (blue) in HSV-1 TK (PDB:1KIM, X-ray residues shown in grey). Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are not shown and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

Regarding D-5-T and L-5-T, in which the hydroxymethyl group of the natural nucleoside is replaced by a hydroxyl, most of the low energy binding modes are in agreement with the

catalysis. However, in the case of D-5-T the corresponding binding energies are higher with respect to dT, being between -64 and -59 score units. Therefore, D-5-T is less prone to be phosphorylated than L-5-T.

Finally, docking calculations on D-6-T showed that most of the predicted poses were both structurally and energetically compatible with the catalysis, being the phosphoryl transfer plausible to occur.

In most of the cases, there is a slight displacement of the base and sugar moiety due to the flexibility of the Glu-83 residue. Despite the lack of interaction with Tyr-101 and Glu-225 as a consequence of the absence of the 3'-OH, most predicted orientations overlapped with those of the benchmark and in all cases, the interaction with the catalytic residue Glu-83 is observed. Therefore, our docking analysis suggests that all the pyrimidine derivatives considered in this work at the first instance are likely to be converted into their monophosphorylated derivatives by HSV-1 TK.

▪ Purine nucleoside analogues

Regarding purine analogues, docking calculations on both enantiomers of Cyclohexenyl G (DCG and LCG) were also carried out and similar poses to ACV were obtained (Figure III-12), validating the use of DCG and LCG together with ACV as benchmarks. Their binding energies are -66 and -65 score units, respectively.

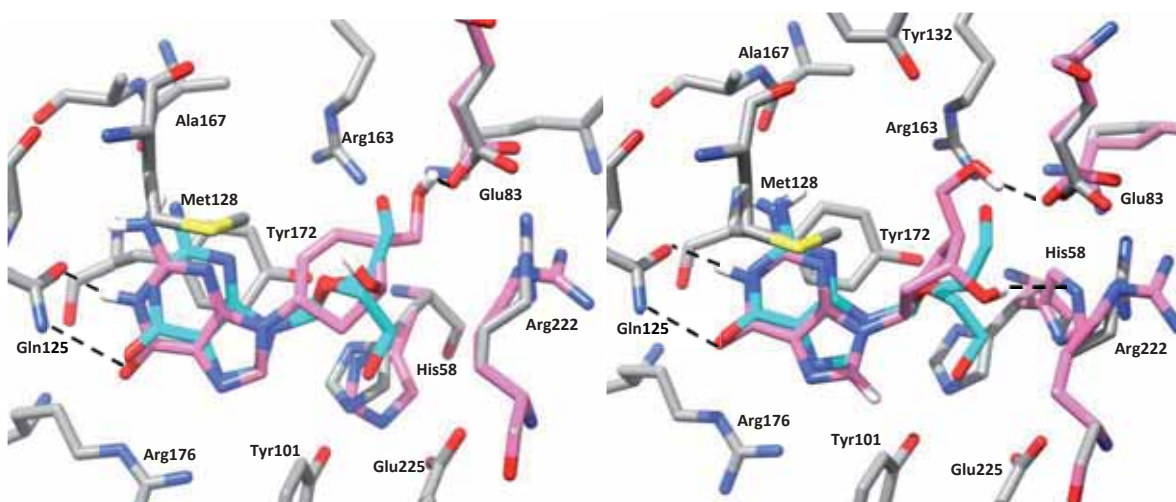


Figure III-12. DCG (left) and LCG (right) in pink superimposed to crystallographic ACV (blue) in HSV-1 TK (PDB: 2KI5, X-ray residues shown in grey). Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are not shown and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

Docking calculations on purine derivatives led to two different behaviours: D- and L-**1-G**, D- and L-**2-G** and D-**3-G** are more prone to be phosphorylated while both enantiomers of **4-G** and **5-G**, and D-**6-G** are not.

Concerning both enantiomers of **1-G**, most of the predicted binding modes of D-**1-G** are compatible with the catalysis with binding energies between -65 and -56 score units, while a few binding modes of compound L-**1-G** in HSV-1TK binding site are in agreement with the catalysis (Figure III-13). A displacement of the base and the carbocycle is observed, but there are interactions between the base and Gln-125 and also between the hydroxymethyl group and Glu-83, which is responsible for the deprotonation. Thus, the phosphorylation of L-**1-G** is less likely to happen than for D-**1-G** but based on the approximation used in this protocol one cannot discard that it could actually occur.

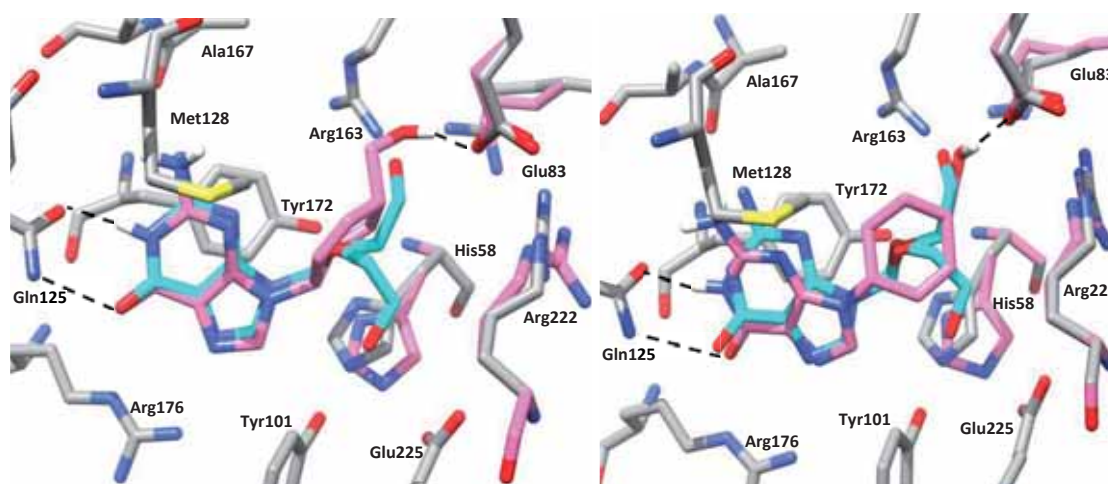


Figure III-13. D-**1-G** (left) and L-**1-G** (right) in pink superimposed to crystallographic ACV (blue) in HSV-1 TK (PDB: 2KI5, X-ray residues shown in grey). Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are not shown and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

Similarly, the predicted complexes of both enantiomers of cyclopropane-fused derivatives **2-G** are consistent with the catalytic activity of the enzyme (Figure III-14), and the corresponding energies are lower than those of the ACV, being -70 score units for D-**2-G** and -65 score units for L-**2-G**. In both cases, the base is properly posed and the hydroxymethyl group is interacting with Glu-83 via a hydrogen bond.

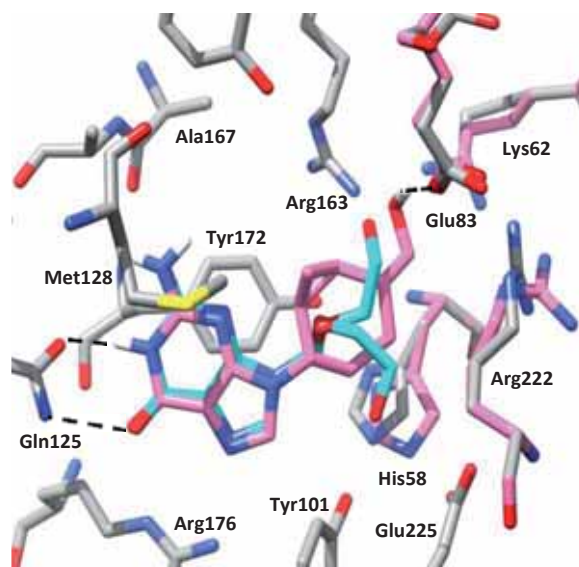


Figure III-14. D-2-G (pink) superimposed to crystallographic ACV (blue) in HSV-1 TK (PDB: 2KI5, X-ray residues shown in grey). Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are not shown and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

By contrast, docking calculations on D-3-G led to two different orientations: one in which the base and the carbocycle are displaced and leads to hydrogen bonds with Gln-125 far weaker, and also the hydroxymethyl group and 3'-OH are pointing towards Glu-83; an another one in which the nucleobase is properly posed but the cyclohexanyl conformation has changed to pseudo-axial and then both 3'-OH and the hydroxymethyl group are hydrogen bonded to Glu-225 instead of Glu-83 (Figure III-15). Thus, the phosphorylation is less plausible to happen but it is possible.

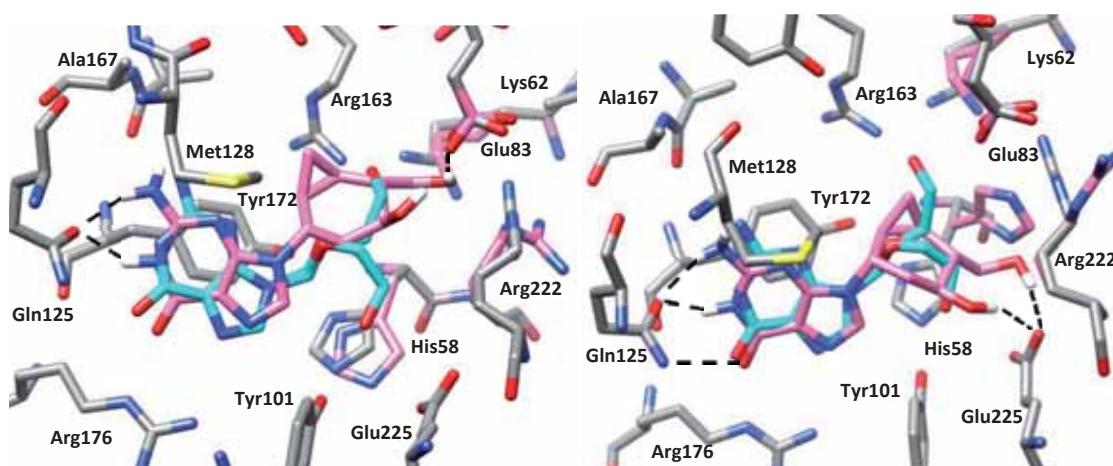


Figure III-15. D-3-G orientations consistent with the catalysis (left) and D-3-G orientations not consistent with the catalysis (right) in pink superimposed to crystallographic ACV (blue) in HSV-1 TK (PDB: 2KI5, X-ray residues shown in grey). Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are not shown and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

Regarding both enantiomers of **4-G**, none of the predicted poses were in agreement with the catalysis and so they are unlikely to be phosphorylated by HSV-1 TK (Figure III-16). The presence of a double bond in the cyclopropane-fused cyclohexane makes it planar and rigid and so the hydroxymethyl cannot be oriented towards Glu-83. Instead of that, the hydroxymethyl is interacting with Tyr-132.

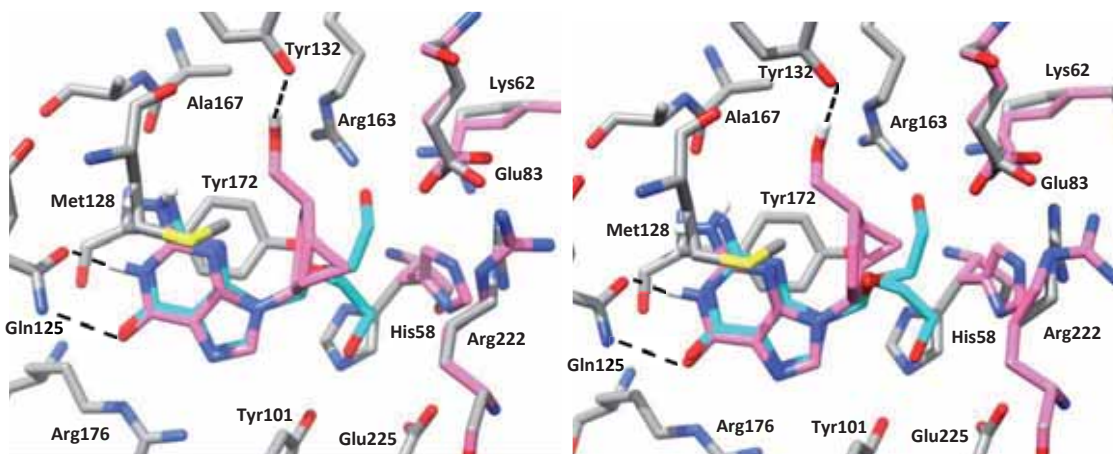


Figure III-16. D-**4-G** (left) and L-**4-G** (right) in pink superimposed to crystallographic ACV (blue) in HSV-1 TK (PDB: 2KI5, X-ray residues shown in grey). Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are not shown and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

The same behaviour is observed for D-**5-G** and L-**5-G** in which the hydroxymethyl group was replaced by a hydroxyl. Although the base is properly posed, the cyclohexane moiety is too rigid that 5'-OH is far away from Glu-83 and cannot interact with it. As a result, in the case of D-**5-G** most of the predicted orientations present an internal hydrogen bond between 4'-OH and the nucleobase, whereas in the case of L-**5-G** the carbocycle has switched and the hydroxyl is interacting with Glu-225 via a hydrogen bond (Figure III-17). Therefore, none of them are expected to be phosphorylated.

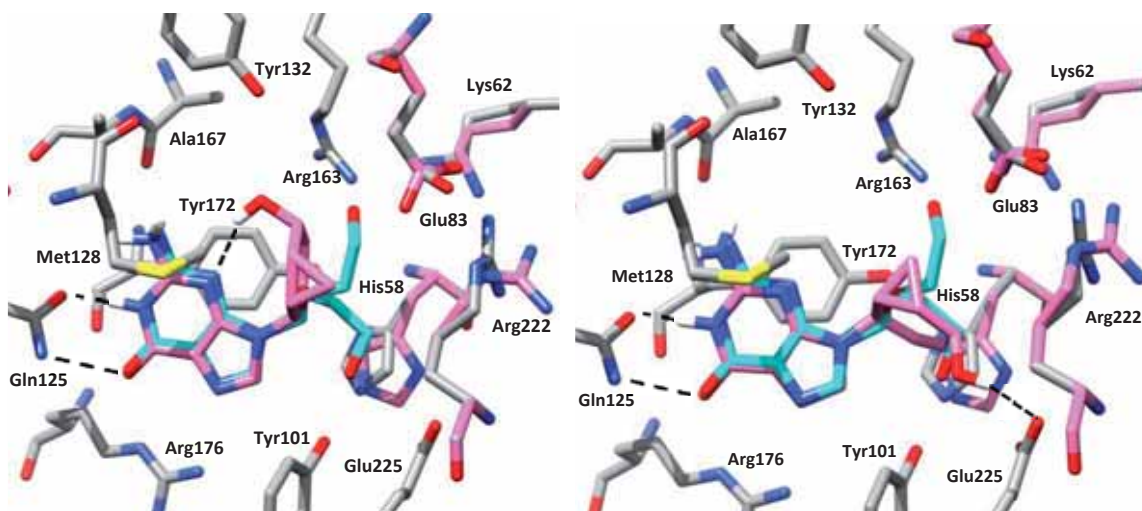
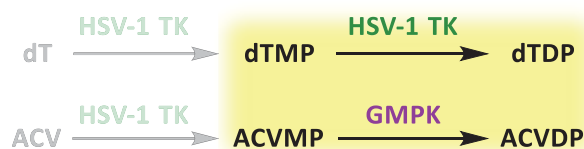


Figure III-17. D-5-G (left) and L-5-G (right) in pink superimposed to crystallographic ACV (blue) in HSV-1 TK (PDB: 2KI5, X-ray residues shown in grey). Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are not shown and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

Concerning D-6-G, docking calculations led to two different orientations: one in which there is an internal hydrogen bond between the hydroxymethyl and the nucleobase and the 3'-OH is pointed towards Glu-83, and also another one in which the hydroxymethyl group is hydrogen bonded to Tyr-132 and Glu-83 is interacting with 4'-OH instead of 3'-OH. However, the phosphoryl transfer for this compound does not seem of the most reliable because the binding energy is far higher than the reference one (-53 score units for D-6-G and -63 and -66 score units for ACV and DCG, respectively). Therefore the synthesis of this compound does not appear an interesting bet.

In summary, all pyrimidine derivatives **1-6** are likely to be phosphorylated and so are both enantiomers of **1-G** and **2-G** and D-**3-G**, whereas the rest of purine compounds are not.

2.2.2. 2nd phosphorylation step



The second phosphorylation step is catalyzed by two different kinases: HSV-1 TK for pyrimidine derivatives and the human GMPK for purine derivatives. GMPK (EC number: 2.7.4.8) is a monophosphate kinase, which belongs to the nucleoside monophosphate (NMP) kinase family, responsible for the phosphoryl transfer from ATP to GMP.

2.2.2.1. GMPK crystallographic structures

There is no crystallographic structure of human GMPK available, but there is one structure of the mouse GMPK in complex with GMP and ADP obtained by Sekulic and co-workers.⁷³ It is worth highlighting that there is a high sequence similarity between human and mouse structures (88% identity), which assures that the information of the mouse enzyme is directly transferable to its human counterpart. So we decided to use the mouse GMPK (mGMPK) structure (PDB code: 1LVG)⁷³ to carry out the docking calculations of the second phosphorylation of purine analogues.

mGMPK is a monomer of 198 residues. It consists of eight α -helices and two β -sheets that form three structurally distinct regions, namely the core, LID and NMP-binding region (Figure III-18). It is known that this kinase changes its conformation when the substrate and ADP are binding, so the structure presenting both the substrate and ADP in its active site is the best approximation to consider these conformational changes.

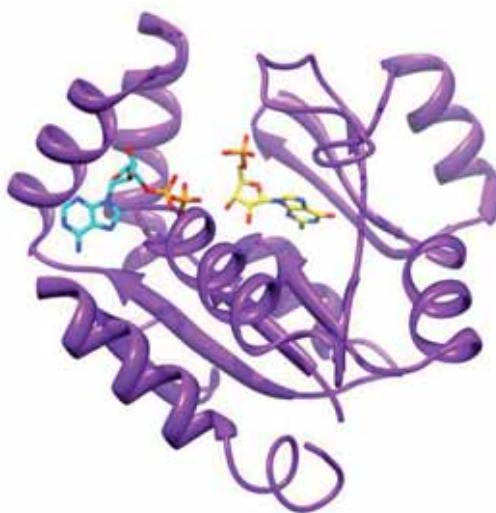


Figure III-18. mGMPK crystallized with GMP (yellow) and ADP (blue) in its binding site (PDB code: 1LVG).

The interaction between GMP and the enzyme mGMPK in the binding site (PDB: 1LVG) is detailed in Figure III-19. The binding mode is different from the HSV-1 TK. The purine ring of GMP is hydrogen bonded to Thr-83, Ser-37, Glu-72 and Asp-103. The ribose moiety is stabilized via a hydrogen bond between the 2'-OH and Asp-101. The α -phosphate of the substrate is hydrogen bonded to Tyr-53, Tyr-81, Arg-41, Arg-44 and Arg-148. It is worth noting that Arg-44 and Arg-148 are suggested to have a catalytic role in the phosphoryl transfer since they also interact with the γ -phosphate of ATP.⁷³ Regarding ADP, its base is interacting via hydrogen bonds with Asp-172 and Asn-171, as well as via a stacking interaction with Arg-133. The α - and β -phosphates of ADP are hydrogen bonded to Gly-16, Lys-17, Ser-18, Thr-19 and Arg-137.

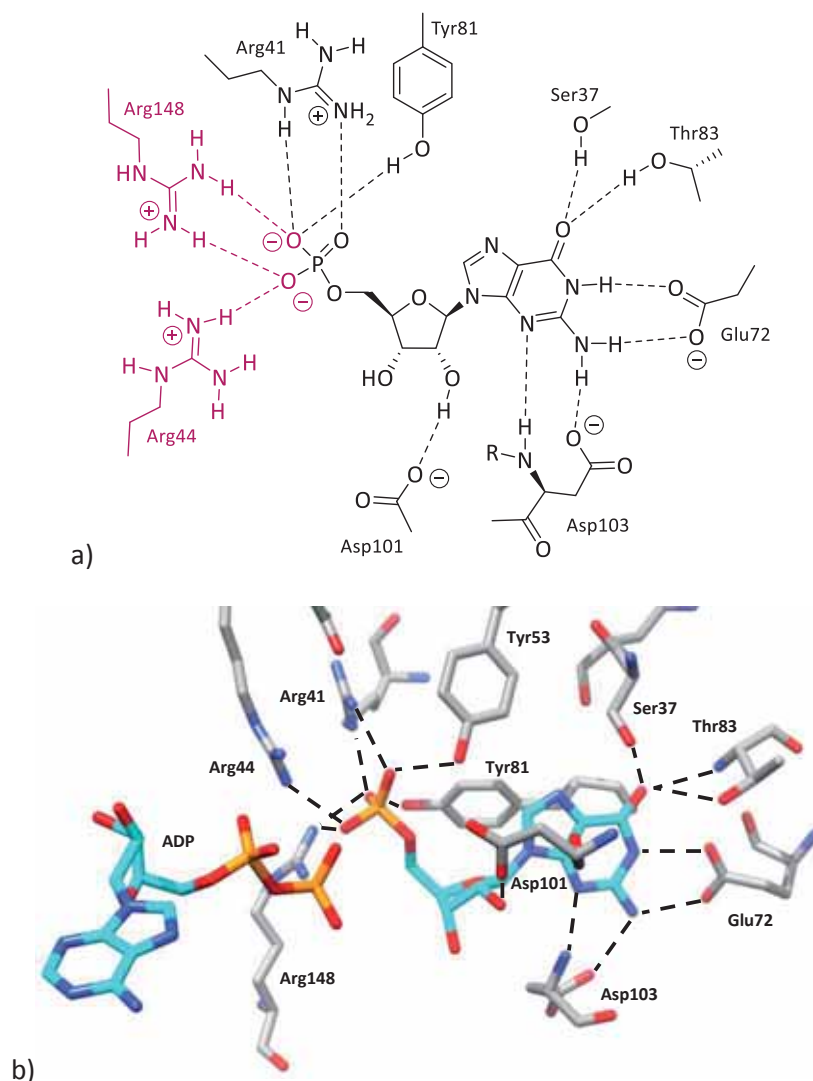


Figure III-19. Representation of the main interactions of GMP in mGMPK binding site (PDB code: 1LVG). Hydrogen bonds between GMP and residues are depicted as dotted lines and the catalytic residues Arg-44 and Arg-148 and their interactions are depicted in maroon. Crystallographic waters and hydrogen atoms are not shown for clarity. Images generated using (a) ChemBioOffice⁵⁶ software and (b) USCF Chimera⁴⁸ package.

2.2.2.2. Docking results

The second phosphorylation step is carried out by two different kinases: HSV-1 TK for pyrimidine compounds and GMPK for purine compounds. Therefore, the analysis of pyrimidine and purine derivatives will be presented separately.

▪ Pyrimidine nucleoside analogues

The docking study of pyrimidine nucleosides was carried out with a crystallographic structure of HSV-1 TK which contains dTMP and also ADP (PDB code: 1VTK)⁶⁶ because it represents a better description of HSV-1 TK state before the second catalytic reaction. The docking protocol was validated by carrying out calculations on the crystallized ligand, which

confirmed that the lowest energy orientations closely match the crystallographic structure and thus validate our docking tool for these compounds.

It is worth mentioning that the second phosphoryl transfer by HSV-1 TK is more restrictive about the presence of a 3'-OH during the catalysis.⁷⁴ In other words, compounds without an alcohol that mimics the 3'-OH and therefore without interactions with Tyr-101 and Glu-225 will be less likely to be phosphorylated. In our case, most of the compounds do not present any alcohol at a similar position and so what is important to consider in predicted poses compatible with the catalysis is that the base is puckered in the right position and the phosphate is oriented towards Arg-44 and Arg-148, which are the catalytic residues, and ADP.

Docking calculations on both enantiomers of **1-TMP** and **2-TMP** showed that the second phosphorylation is in agreement with the catalysis, both structurally and energetically (Figure III-20). In all these cases, the base is correctly oriented and the phosphate does not perfectly overlap with the crystallographic structure due to the flexibility of the side chains, but it is still oriented towards ADP. The corresponding binding energies are -72 and -69 score units for D-**1-TMP** and L-**1-TMP**, and -80 and -67 score units for D-**2-TMP** and L-**2-TMP**, which are similar or even lower than the reference one (-73 score units).

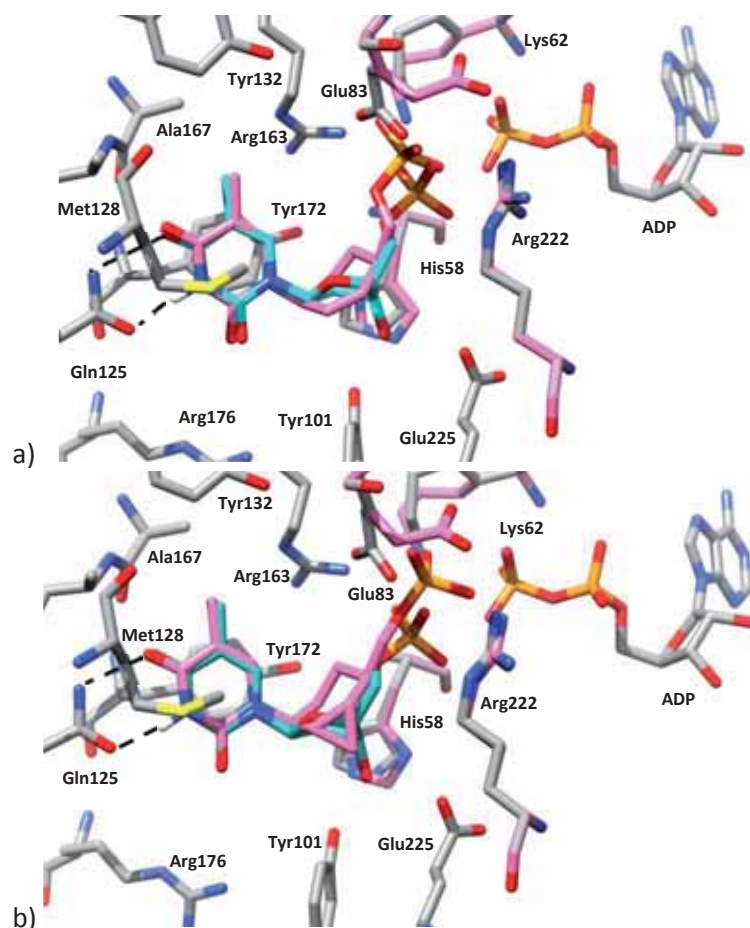


Figure III-20. a) D-1-TMP (pink) superimposed to crystallographic dTMP (blue) in HSV-1 TK (PDB: 1VTK, X-ray residues shown in grey). b) D-2-TMP (pink) superimposed to crystallographic dTMP (blue) in HSV-1 TK (PDB: 1VTK, X-ray residues shown in grey). Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are not shown and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

Docking calculations on D-3-TMP showed two different orientations: one in which the base and the phosphate are properly oriented and the 4'-OH is hydrogen bonded to Glu-225 and the cyclohexane moiety presents a pseudo-equatorial conformation, and another in which the base and the phosphate are also properly posed, the 3'-OH is hydrogen bonded to Tyr-172 and the cyclohexane moiety is in pseudo-axial conformation. In both predicted orientations the second phosphoryl transfer is expected to happen.

Concerning both enantiomers of 4-TMP, the predicted binding modes are compatible with the catalytic reactivity of the enzyme. In most poses of both enantiomers, thymine is stacked against Tyr-172 and the phosphate is pointed towards ADP, and so the phosphorylation is likely to happen.

A similar behaviour was predicted for D-5-TMP and L-5-TMP, in which several predicted orientations were consistent with the catalysis. However, the corresponding binding energies (-68 score units for D-5-TMP and -63 score units for L-5-TMP) were higher than the reference

one, -73 score units. These results suggest that the phosphoryl transfer is less likely to occur but it is still possible.

Regarding **D-6-TMP**, docking calculations showed that most of the binding modes are not perfectly posed, being the phosphorylation less plausible to occur (Figure III-21). The nucleobase does not perfectly overlap the crystallographic structure, the phosphate is pointing up and none of the hydroxyl groups are interacting with Tyr-101 and Glu-225. Moreover, the binding energy is -66 score units which is higher than the reference one, -73 score units. Therefore, the second phosphorylation step is not expected to happen.

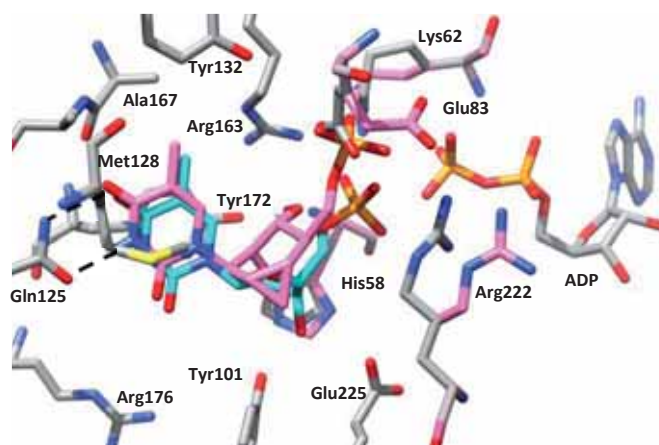


Figure III-21. D-6-TMP (pink) superimposed to crystallographic dTMP (blue) in HSV-1 TK (PDB: 1VTK, X-ray residues shown in grey). Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are not shown and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

▪ Purine nucleoside analogues

For purine derivatives, the docking study was performed with the X-ray structure of mGMPK previously described (PDB code: 1LVG). The docking protocol was validated by carrying out docking calculations of the crystallized ligand, GMP, which confirmed that the lowest energy orientations closely match the crystallographic structure.

Additionally, docking calculations on ACVMP as well as both enantiomers DCGMP and LCGMP were also carried out and similar poses to GMP were obtained, validating the use of them as benchmarks. The corresponding binding energies are -87, -98 and -93 score units for ACVMP, DCGMP and LCGMP, respectively.

Docking calculations on both enantiomers of **1-GMP** and **2-GMP** showed that the phosphorylation is plausible to occur, since most of the predicted binding modes are consistent with the catalysis and the corresponding binding energies are -92, -95, -96 and -98 score units for D- and L-**1-GMP** and D- and L-**2-GMP**, respectively, being similar or even lower than those of the references.

Similarly, the predicted orientations of **D-3-GMP** are in agreement with the catalysis, both structurally and energetically (Figure III-22). The base is perfectly overlapped and the phosphate is pointing towards Arg-44 and Arg-148, which are suggested to have a catalytic role in the phosphoryl transfer.

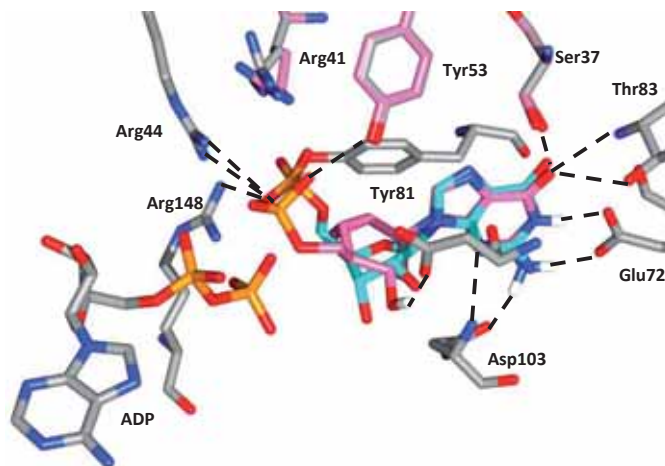


Figure III-22. D-3-GMP (pink) superimposed to crystallographic GMP (blue) in mGMPK (PDB: 1LVG, X-ray residues shown in grey). Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are not shown and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

Regarding both enantiomers of **4-GMP** and **5-GMP**, most of the predicted binding modes are consistent with the catalysis, and therefore they are likely to be phosphorylated by GMPK. The binding energies are -96 and -99 score units for D-**4-GMP** and L-**4-GMP**, and -92 and -99 score units for D-**5-GMP** and L-**5-GMP**, respectively, which are similar than those of the references.

Finally, docking calculations on D-**6-GMP** led to two opposite orientations: one in which the base is properly posed and the phosphate is pointing towards ADP, and another one in which a rotation of the purine base is observed (Figure III-23). This orientation is stabilized by hydrogen bonds between the base and Ser-37 and Tyr-53 as well as the interactions between the phosphate and Arg-44. However, the high number of orientations incompatible with the catalysis indicated that the phosphoryl transfer is less plausible to happen.

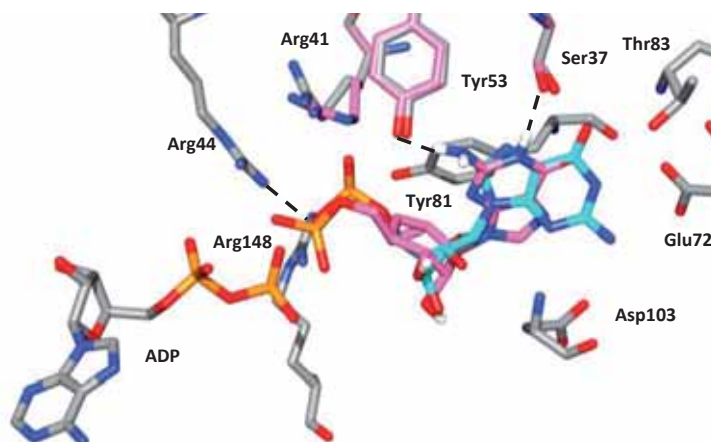
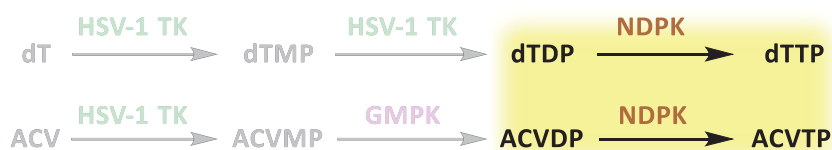


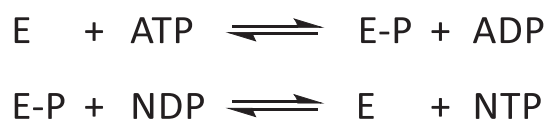
Figure III-23. D-6-GMP (pink) superimposed to crystallographic GMP (blue) in mGMPK binding site (PDB: 1LVG, X-ray residues shown in grey). Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are not shown and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

To sum up, a second phosphorylation is likely to occur for almost all the studied compounds except for D-6-TMP and D-6-GMP compounds.

2.2.3. 3rd phosphorylation step



The last phosphorylation step is catalysed by the human NDPK (EC number 2.7.4.6.) for both pyrimidine and purine derivatives. It is responsible for the third phosphoryl transfer from a nucleoside triphosphate (usually ATP or GTP) to a diphosphate acceptor (NDP) via a ping-pong mechanism (Scheme III-3). Most kinases bind the donor and the acceptor at two different sites and the phosphoryl is directly transferred from the donor to the acceptor. But in the case of NDPK, the binding site of the donor is the same as the acceptor, which means that the first product has to be released from the enzyme before the second substrate comes in.^{75,76} In other words, NDPK binds first the donor, phosphorylates a catalytic histidine residue, then dissociates the donor and binds the acceptor and, finally, transfers the phosphoryl group to the diphosphorylated substrate.



Scheme III-3. Ping-pong mechanism for the activation of diphosphorylated derivatives by NDPK using ATP as the phosphoryl donor. E: enzyme; E-P: enzyme-phosphate; NDP: nucleoside diphosphate; NTP: nucleoside triphosphate.

As the enzyme accepts all common nucleotides and derivatives as substrates, it is a major source of precursors for DNA and RNA synthesis in all cells.

2.2.3.1. NDPK crystallographic structures

Human NDPK is responsible for the third phosphorylation of both pyrimidine and purine nucleosides. However, human NDPK crystallized structures are only available for purine nucleosides. As a result, a slime mold *Dictyostelium discoideum* NDPK X-ray structure was selected for the docking calculations on pyrimidine nucleosides, since the active site and the residues involved in nucleotide binding are analogous to its human counterpart with similar *in vitro* properties.⁷⁷ Thus, docking calculations were performed on a complex of slime mold *Dictyostelium discoideum* NDPK with dTDP (PDB code: 1NDC) for pyrimidine derivatives,⁷⁸ whereas for purine derivatives a human NDPK crystallized with GDP was used (PDB code: 1NUE).⁷⁷

NDPK is a homo-hexamer with around 150 residues each subunit (Figure III-24).⁷⁷ Except for minor differences due to crystal packing, all subunits are identical.

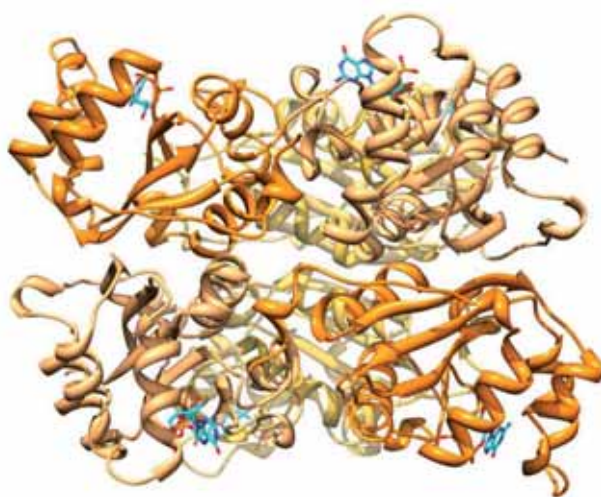


Figure III-24. NDPK crystallized with GDP (blue) in its binding site (PDB code: 1NUE).

The interaction between GDP and human NDPK in the binding site (PDB code: 1NUE) is detailed in Figure III-25. The purine base is sandwiched between Phe-60 and Val-112 and hydrogen bonded to Glu-152, which comes from an adjacent subunit. The ribose moiety is

III. Molecular modelling pre-study for the design of anti-HSV novel carbocyclic nucleosides

totally buried in the protein, and the 2'-OH and 3'-OH are interacting via hydrogen bond with Lys-12 and Asn-115, respectively. There is also an internal hydrogen bond interaction between the 3'-OH and the β -phosphate. The α -phosphate remains accessible to the solvent and it is interacting with a magnesium ion, and the β -phosphate is hydrogen bonded to Arg-88 and Thr-94 and also water-mediated to His-118.

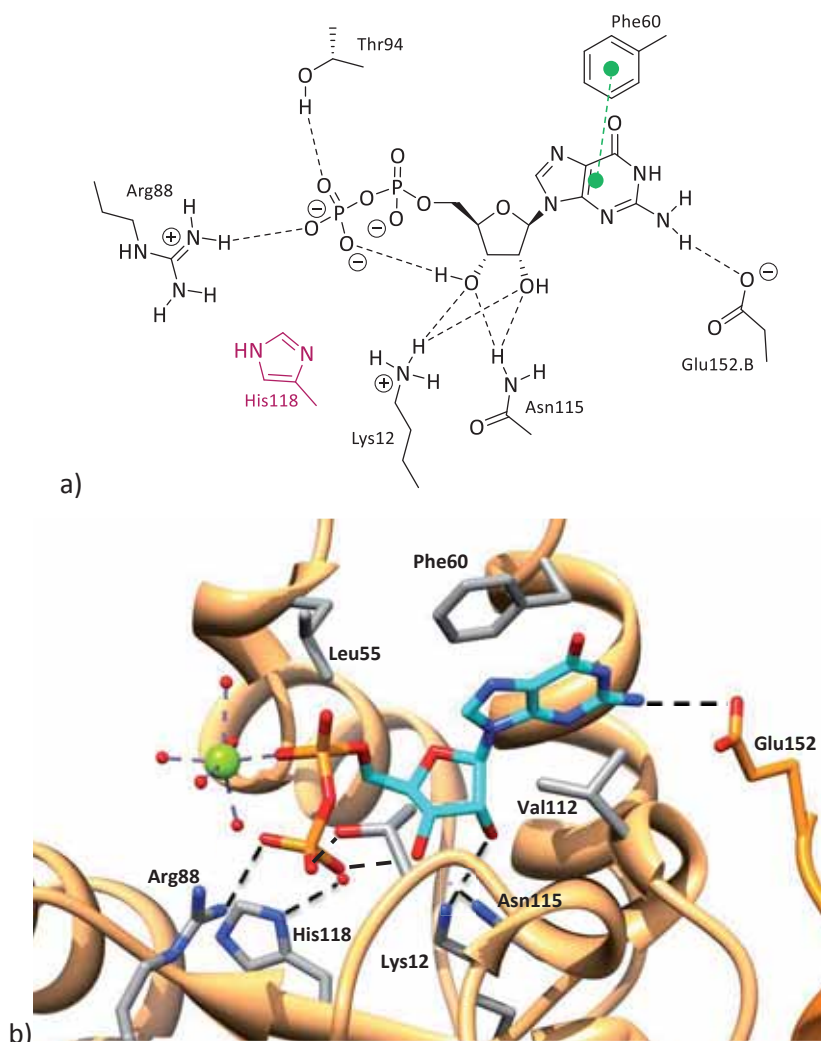


Figure III-25. Representation of the main interactions of GDP in NDPK binding site (PDB code: 1NUE). Hydrogen bonds between GDP and residues are depicted as dotted lines and the catalytic residue His-118 is depicted in maroon. The crystallographic ion Mg^{2+} is shown, and Glu-152 of another subunit is shown in orange. Hydrogen atoms are not shown for clarity. Images generated using (a) ChemBioOffice⁵⁶ software and (b) USCF Chimera⁴⁸ package.

The binding mode of dTDP in *Dictyostelium discoideum* NDPK (1NDC) is the same as that of GDP in human NDPK (Figure III-26). The base is also cleft between Phe-64 and Val-116, and it is stabilised by two water-mediated hydrogen bonds with the side chain of Glu-115 from an adjacent subunit. There is a magnesium ion bridging the α - and β -phosphates and β -phosphate is hydrogen bonded to 3'-OH, Thr-98, Arg-92 and Arg-109 and water-mediated to the catalytic His-122.

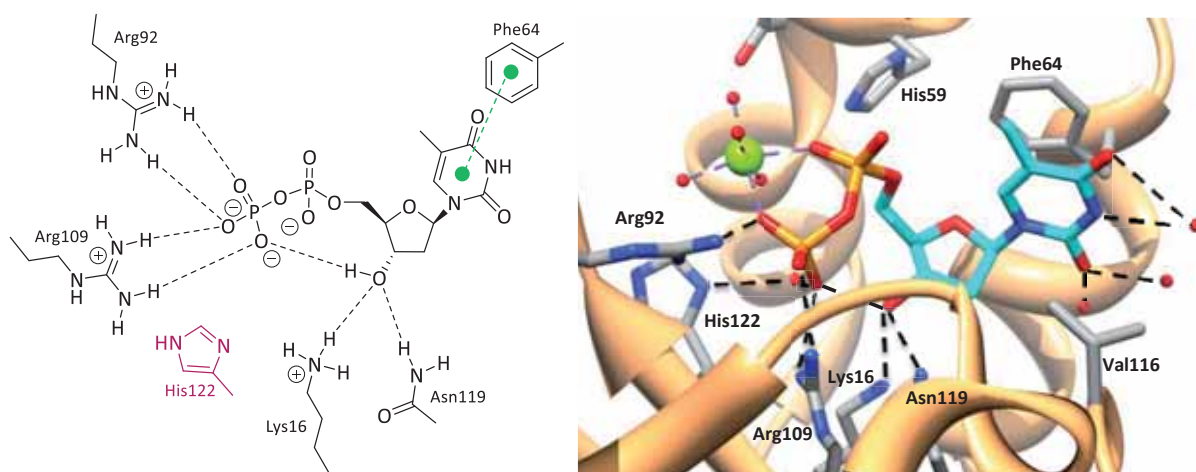


Figure III-26. Representation of the main interactions of dTDP in NDPK binding site (PDB code: 1NDC). Hydrogen bonds between dTDP and residues are depicted as dotted lines and the catalytic residue His-122 is depicted in maroon. The crystallographic ion Mg^{2+} is shown, and the residue Thr-98 as well as hydrogen atoms are not shown for clarity. Images generated using ChemBioOffice⁵⁶ software (left) and USCF Chimera⁴⁸ package (right).

2.2.3.2. Docking results

The last phosphoryl transfer is carried out by the same kinase for both pyrimidine and purine nucleosides. However, the analysis of these derivatives will be presented separately, since different crystallographic structures were used for pyrimidine and purine nucleosides.

▪ Pyrimidine nucleoside analogues

The docking protocol was validated by docking dTDP into the corresponding crystallographic structure (PDB code: 1NDC). The first calculations were carried out defining the binding site from a 9 Å cavity generated around Lys-16 and without any flexible residue. Disappointingly, most of the predicted binding modes did not match the crystallographic structure, probably because of docking limitations in dealing with large solvent exposed binding sites. In order to overcome this limitation, some explicit waters were added and the calculations were repeated. However, no better predicted orientations were obtained. Then, some residues were set as flexible, the scoring function was changed as well as the size of the binding site but there was any improvement neither. Fortunately, properly predicted orientations were obtained when all the residues were kept fixed, the binding site was set at 11 Å and three explicit waters were taken into account. It is worth highlighting that the reproduction of the crystallographic structure was not as good as in the previous activation steps and therefore predictions on this last phosphoryl transfer should be read with caution.

Considering this and the fact that the NDPK catalysis consists of an initial phosphoryl transfer to His-122 and subsequent phosphorylation of the substrate, docking results were analysed focusing mainly in the orientation of the β -phosphate towards the catalytic His-122.

Regarding both enantiomers of **1-TDP** and **2-TDP**, several predicted binding modes are consistent with the catalysis (Figure III-27), being the corresponding binding energies -95 and -94 score units for D- and L-**1-TDP**, and -93 and -94 score units for D- and L-**2-TDP**, which are similar than the reference (-94 score units). Thus, all of them are expected to be phosphorylated.

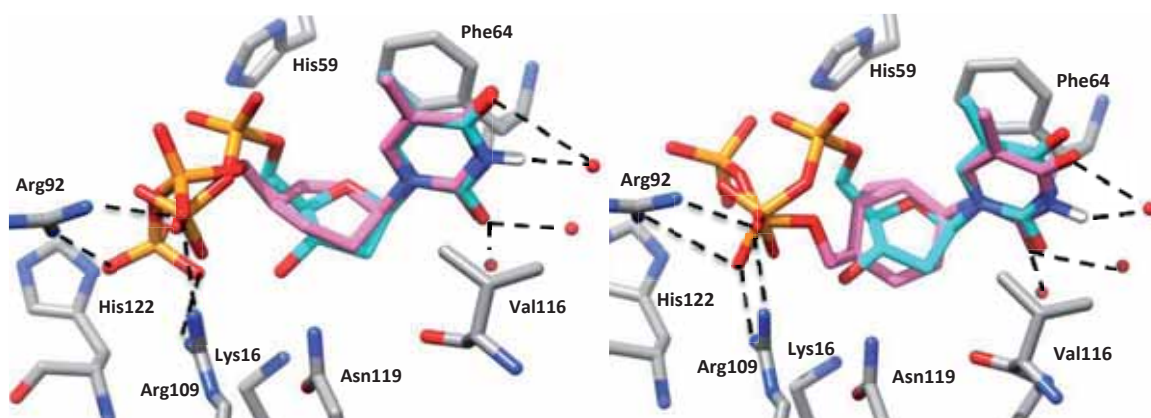


Figure III-27. D-**1-TDP** (left) and D-**2-TDP** (right) in pink superimposed to crystallographic dTDP (blue) in NDPK (PDB: 1NDC, X-ray residues shown in grey). Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are not shown and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

Docking calculations on D-**3-TDP** showed that a large number of predicted poses are correctly oriented, in which the thymine is properly stacked and the β -phosphate is pointing towards His-122. The binding energy is -98 score units, which is lower than the reference one.

Several of the binding modes of both enantiomer of **4-TDP** and **5-TDP** are in agreement with the catalytic activity of the enzyme, in which the β -phosphate was pointing towards His-122 and so the phosphoryl transfer is likely to occur for all these nucleosides.

By contrast, docking calculation on D-**6-TDP** led to a reduced number of predicted orientation properly posed for the catalysis. In fact, most of the binding modes showed the β -phosphate pointing upwards or the thymine rotated outside its pocket (Figure III-28). Therefore, the second phosphorylation is not expected to happen.

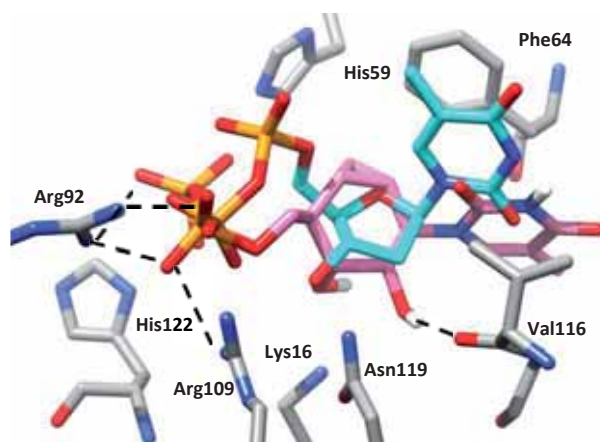


Figure III-28. D-6-TDP (pink) superimposed to crystallographic dTDP (blue) in NDPK (PDB: 1NDC, X-ray residues shown in grey). Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are not shown and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

▪ Purine nucleoside analogues

Regarding purine derivatives, the docking protocol was validated by docking GDP into the human NDPK (PDB code: 1NUE). The first calculations were carried out defining the same parameters than for pyrimidine derivatives: the binding site from a 9 Å cavity generated around Lys-12 and without any flexible residue. And again, most binding modes did not match the crystallographic structure maybe due to the docking limitations dealing with widely solvent exposed binding sites. Some modifications in the parameters were tested in order to overcome this problem, such as setting some residues as flexible, changing the scoring function or even changing the size of the binding site, but without any success. Finally, several predicted binding modes were properly posed when all the residues were kept fixed, another adjacent subunit that contained Glu-152 was added as well as three explicit waters and the binding site was set at 13 Å. Despite these good results, the reproduction of the crystallographic structure was not as good as the previous steps, and so the analysis of the docking results were mainly focused in the proper orientation of the β -phosphate towards the catalytic His-122.

Docking calculations on ACVDP, DCGDP and LCGDP were performed to provide with structural and energetic benchmarks. As most of the predicted poses were properly overlapped to the crystallographic structure, they were taken as references (Figure III-29). Their corresponding binding energies are -79 score units for ACVDP, and -91 and -85 score units for DCGDP and LCGDP, respectively.

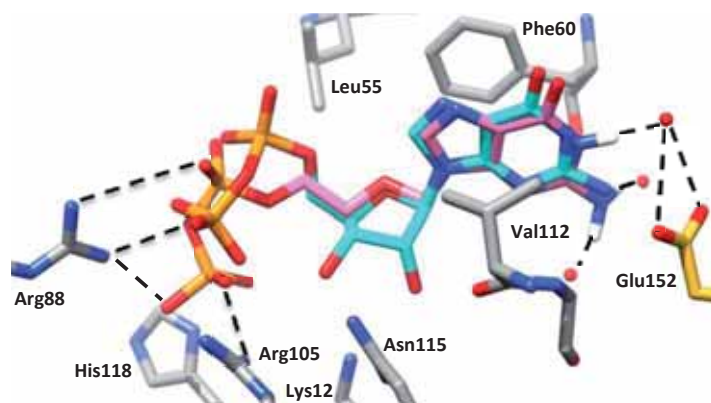


Figure III-29. ACVDP (pink) superimposed to crystallographic GDP (blue) in NDPK (PDB: 1NUE, X-ray residues shown in grey). Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are only shown when interacted with the ligand and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

Several of the binding modes predicted for both enantiomers of **1-GDP** and **2-GDP** were consistent with the catalysis (Figure III-30), so all these nucleosides are expected to be phosphorylated by NDPK. The corresponding binding energies are -82 and -84 score units for D- and L-**1-GDP**, and -80 and -84 score units for D- and L-**2-GDP**, respectively, which are similar or even lower than those of the benchmarks.

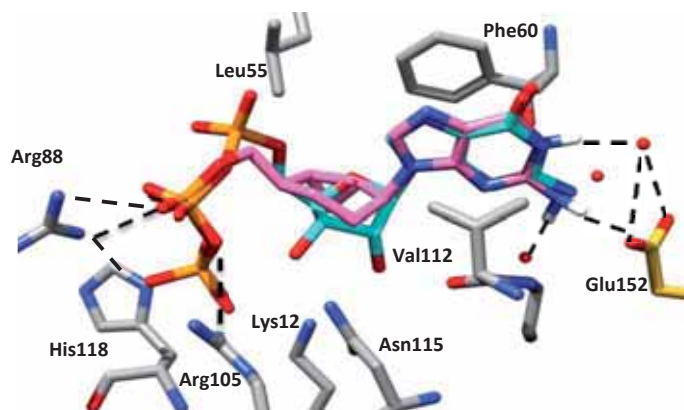


Figure III-30. D-**2-GDP** (pink) superimposed to crystallographic GDP (blue) in NDPK (PDB: 1NUE, X-ray residues shown in grey). Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are only shown when interacted with the ligand and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

Docking calculations on D-**3-GDP** showed that the phosphoryl transfer is plausible since several of the lowest predicted binding modes showed the β -phosphate oriented towards His-118, the nucleobase and the carbocycle properly posed and the 3'-OH hydrogen bonded to β -phosphate (Figure III-31).

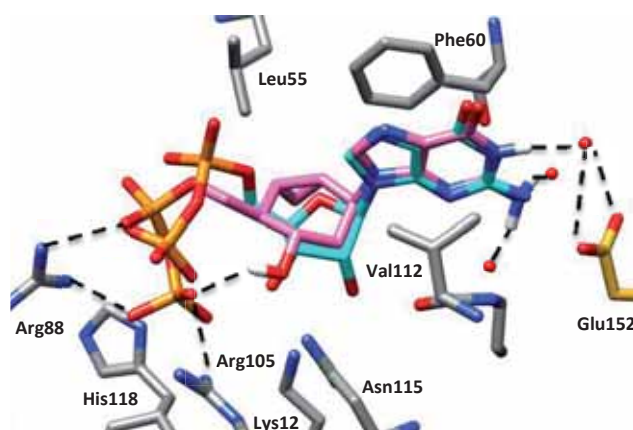


Figure III-31. D-3-GDP (pink) superimposed to crystallographic GDP (blue) in NDPK (PDB: 1NUE, X-ray residues shown in grey). Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are only shown when interacted with the ligand and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

Regarding both enantiomers of **4-GDP** docking calculations in all the cases showed to that the third phosphoryl transfer is plausible, since several lowest binding modes were properly overlapped with the crystallographic structure.

On the other hand, docking calculations on both enantiomers of **5-GDP** led to two different behaviours: D-**5-GDP** lowest binding modes were consistent with the catalysis whereas L-**5-GDP** showed a reduced number of predicted orientation properly posed for the catalysis. In fact, most of the binding modes of L-**5-GDP** showed the β -phosphate pointing upwards. Therefore, the second phosphorylation more likely to happen for D-**5-GDP** than for L-**5-GDP**.

And finally, a high number of predicted binding modes of D-**6-GDP** were properly overlapped with the crystallographic structure (Figure III-32), showing the base correctly stacked and stabilized by two hydrogen bonds with water molecules and a water-mediated hydrogen bond with the side chain of Glu-152 from an adjacent subunit, as well as the β -phosphate pointing towards His-118.

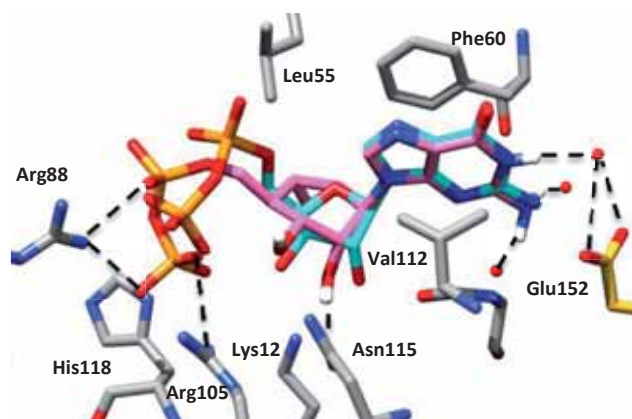


Figure III-32. D-6-GDP (pink) superimposed to crystallographic GDP (blue) in NDPK (PDB: 1NUE, X-ray residues shown in grey). Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are only shown when interacted with the ligand and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

In summary, the third phosphorylation step is plausible to occur for all these nucleosides except for D-6-TDP and L-5-GDP. It is worth recalling that there are some limitations on docking calculations with NDPK due to its widely solvent exposed binding site, and therefore these docking results are not as good as those of the previous steps.

2.3. Exploring other scaffolds

The *in silico* molecular modelling study performed on cyclohexenyl nucleoside analogues **1-6** (Figure III-33) showed that most of them were expected to be activated and therefore they seem good candidates as anti-HSV agents.

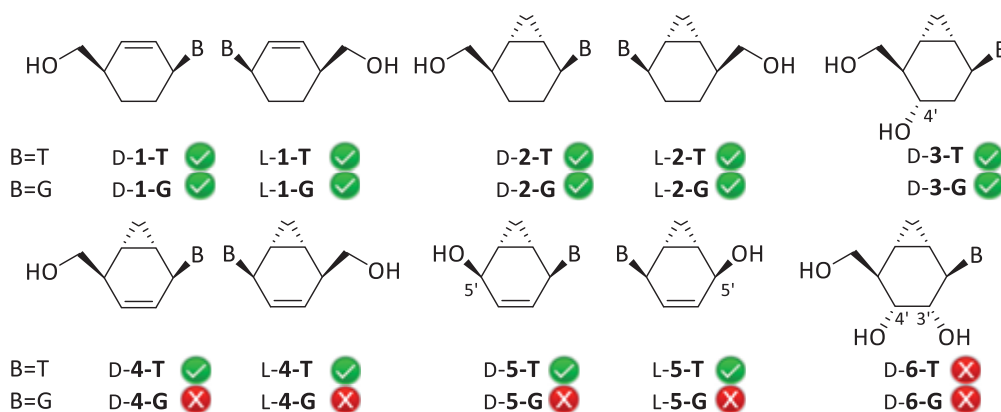


Figure III-33. Molecular modelling study results.

Based on those compounds more prone to be triphosphorylated of the original series, a new family of nucleoside analogues built on a bicyclo[4.1.0]heptane scaffold were envisaged to be synthesised. Concretely, we focused on bicycle[4.1.0]heptane nucleoside **2** with different nucleobases: thymine (T), guanine (G), adenine (A) and uracil (U). As the whole activation

2. Molecular modelling pre-study for the design of anti-HSV novel carbocyclic nucleosides

process of nucleoside analogues D-2-G and D-2-T has already been studied, only docking calculations on D-2-A and D-2-U must be carried out.

2.3.1. Activation process

Protein-ligand docking studies on the kinases involved in the activation process (HSV-1, GMPK and NDPK) were performed on the crystallographic structures previously used (Table III-3).

Table III-3. PDB structures used to study the activation process of D-2-U and D-2-A.

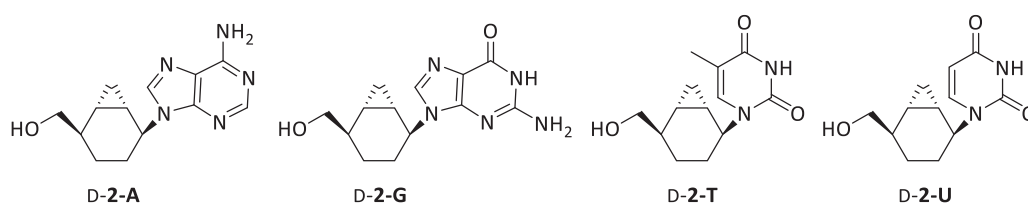
Enzyme	Organism	PDB code	Crystallized ligand	Docking calculations ^a
TK	HSV-1	1KIM	dT	D-2-U
TK	HSV-1	2KI5	ACV	D-2-A
TK	HSV-1	1VTK	dTMP (+ ADP)	D-2-UMP
GMPK	<i>Mus musculus</i>	1LVG	GMP (+ ADP)	D-2-AMP
NDPK	<i>Dictyostelium discoideum</i>	1NDC	dTDP	D-2-UDP
NDPK	<i>Homo sapiens</i>	1NUE	GDP	D-2-ADP

^a MP and DP stand for monophosphorylated and diphosphorylated compounds, respectively.

Nucleoside analogues D-2-A and D-2-U were docked separately into the active site of each kinase following the same criteria as in the preceding steps. Docking results were analysed both in structural and energetic terms, checking how many binding modes were consistent with a catalytic orientation, and if their corresponding binding energies were similar or even lower than those of the benchmarks (dT and ACV for pyrimidine and purine derivatives, respectively).

The results obtained from docking calculations on D-2-A and D-2-U as well as the previous studies on D-2-G and D-2-T are summarized in Table III-4. As it is shown on the table, all the nucleoside analogues are expected to be triphosphorylated except the uracil derivative D-2-U, which fails in the last phosphorylation step. Docking results on each phosphorylation step are explained in detail in the next paragraphs.

Table III-4. Protein-ligand docking results of the whole activation process of nucleoside analogues synthesised.



Requirement		D-2-A	D-2-G	D-2-T	D-2-U
Activation process	OH → OMP	✓	✓	✓	✓
	OMP → ODP	✓	✓	✓	✓
	ODP → ODP	✓	✓	✓	✗

Docking calculations of the first phosphorylation step showed that both **D-2-A** and **D-2-U** are likely to be phosphorylated, since most of the predicted binding modes are compatible with the catalytic activity of the enzyme and their binding energies are similar or even lower than those of the benchmarks (-64 and -67 score units for **D-2-A** and **D-2-U**; -63 and -70 score units for ACV and dT, respectively) (Figure III-34). In the case of **D-2-U**, most of the predicted binding modes showed the uracil ring perfectly overlapped with the crystallographic structure and the hydroxyl group pointing towards Glu-83, which is the catalytic residue. In the case of **D-2-A**, although the adenine ring was not correctly overlapped with the crystallographic structure, in most of the predicted poses the nucleobase was cleft between Tyr-172 and Met-128 and hydrogen bonded to Gln-125. Moreover, the 5'-OH was oriented towards Glu-83.

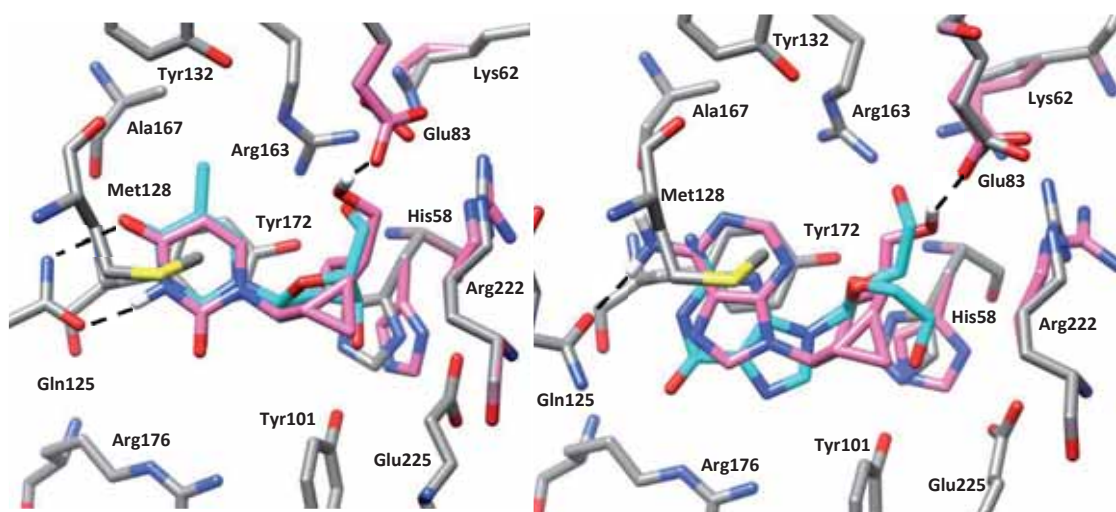


Figure III-34. **D-2-U** in pink superimposed to crystallographic dT (blue) in HSV-1 TK (PDB:1KIM, X-ray residues shown in grey) (left); **D-2-A** in pink superimposed to crystallographic ACV (blue) in HSV-1 TK (PDB:2K15, X-ray residues shown in grey) (right);. Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are not shown and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

2. Molecular modelling pre-study for the design of anti-HSV novel carbocyclic nucleosides

Protein-ligand dockings of the second phosphorylation step revealed that both **D-2-U** and **D-2-A** are also expected to be converted into their diphosphorylated derivatives (Figure III-35). In both cases, the base is properly posed and the phosphate does not perfectly overlap with the crystallographic structure due to the flexibility of the side chains, but it is still oriented towards ADP. The corresponding binding energies are -74 score units for **D-2-UMP** and -88 score units for **D-2-AMP**, which are lower than the reference one (-73 and -79 score units for dTMP and ACVMP, respectively), and therefore suggest that the affinity of the prodrugs to the activating enzymes are excellent for poses catalytically consistent.

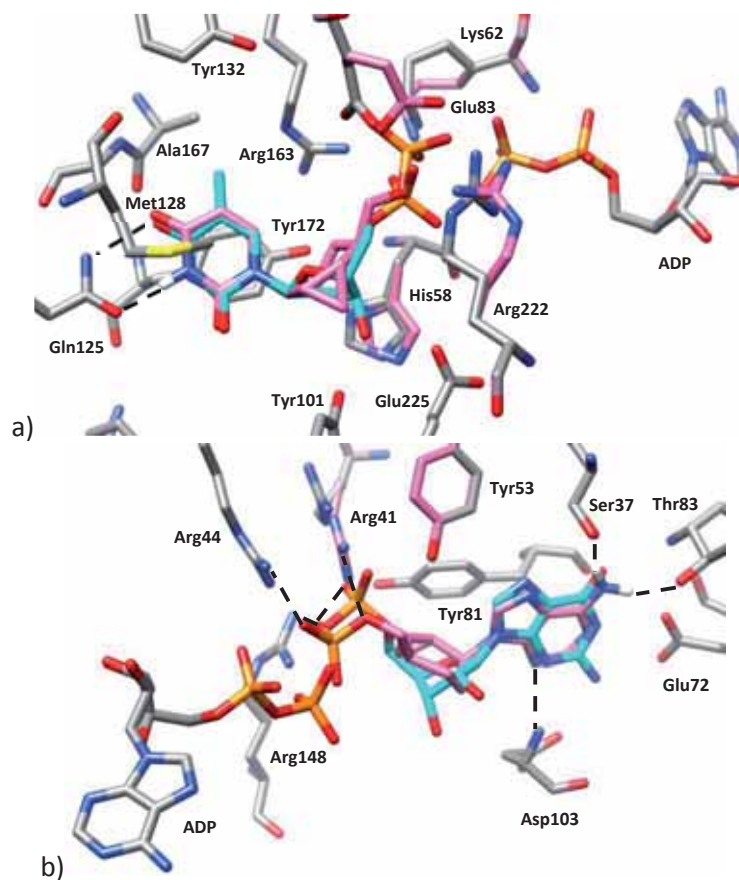


Figure III-35. a) **D-2-UDP** in pink superimposed to crystallographic dTDP (blue) in HSV-1 TK (PDB:1VTK, X-ray residues shown in grey); b) **D-2-ADP** in pink superimposed to crystallographic GDP (blue) in *Mus musculus* GMPK (PDB:1LVG, X-ray residues shown in grey). Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are not shown and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

Regarding the last activation step, due to the docking limitations in dealing with widely solvent exposed binding sites, docking results were analysed focusing mainly in the orientation of the β -phosphate towards the catalytic histidine.

Calculation on the third activation step revealed that most of the predicted binding modes of **D-2-ADP** are in agreement with the catalysis, and thus it is expected to be triphosphorylated (Figure III-36). Despite the adenine ring is not correctly posed, it is still

sandwiched between Phe-60 and Val-112 and interacting with the Glu-152 of an adjacent subunit. In the case of D-2-UDP, the predicted poses are structurally consistent with the catalysis, since the uracil ring is properly posed and the β -phosphate is pointing towards His-122. However, this phosphoryl transfer is not energetically favourable, since the binding energies of D-2-UDP (-80 score units) were higher than those of the benchmark, dTDP (-94 score units). Therefore, the last phosphorylation of D-2-UDP is unlikely to happen.

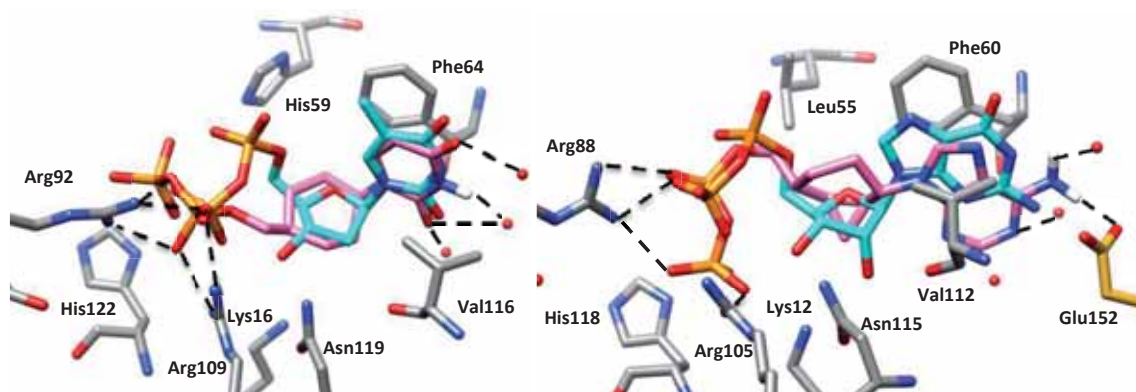


Figure III-36. D-2-UDP in pink superimposed to crystallographic dTDP (blue) in *Dictyostelium discoideum* NDPK (PDB:1NDC, X-ray residues shown in grey) (left); b) D-2-ADP in pink superimposed to crystallographic GDP (blue) in human NDPK (PDB:1NUE, X-ray residues shown in grey) (right). Hydrogen bonds are depicted as dotted lines. For the sake of clarity, crystallographic waters are not shown and hydrogen atoms are only shown when bound to a heteroatom of the ligand.

3. Conclusions

A molecular modelling study was performed as a first stage of the rational design of new series of cyclohexenyl nucleoside analogues **1-6** as anti-HSV agents. To do that, their drug-likeness and their activation process by which they are converted into the triphosphorylated derivatives were investigated.

The evaluation of their drug-likeness showed that these nucleoside candidates fulfil the “Lipinski’s rule of five”, supporting their good drug-likeness.

Once their drug-likeness was evaluated, the activation process of these nucleosides was analysed by means of protein-ligand docking. The theoretical calculations revealed that only nucleosides D-6-T, D- and L-4-G, D- and L-5-G and D-6-G are not likely to be converted in their active forms. In particular, all these purine derivatives fail in the first and third phosphorylation steps whereas pyrimidine derivative D-6-T fails in the second and third phosphoryl transfer.

Furthermore, calculations revealed which skeletons are more favoured in each activation step. In general terms, dockings demonstrated that the replacement of the double bond by a fused cyclopropane ring does not work against the phosphoryl transfer. By contrast,

the flexibility of the carbocycle seems to play an important role on the activation process, since the additional double bond to the fused cyclohexane of purine derivatives **4-G** and **5-G**, which forces the carbocycle to be planar and rigid, is detrimental to the criterion we used in our protocol. Additionally, the presence of hydroxyl groups in the carbocycle does not enhance the phosphoryl transfer.

In summary, the present study indicated that among the studied compounds, pyrimidine nucleosides are more likely to be converted into their triphosphorylated derivatives than purine nucleosides. This study also proved that the same interactions in the HSV-1 binding site were observed for both D- and L-enantiomers and that the flexibility of the carbocycle is important for the activation process.

To complete this rational design the interaction between the triphosphorylated derivatives of these nucleosides and HSV-1 DNA polymerase should be studied. However, no crystallographic structure of this enzyme with DNA was available and thus the interaction could not be investigated.

This computational study prompted us to synthesise a new family of nucleoside analogues built on a bicyclo[4.1.0]heptane scaffold, **2**, with different nucleobases (A,G,T,U).

4. References

- (1) Patrick, G. L. *An introduction to medicinal chemistry*; 5th ed.; Oxford University Press, 2013.
- (2) Deville-Bonne, D.; El Amri, C.; Meyer, P.; Chen, Y.; Agrofoglio, L. A.; Janin, J. *Antiviral Res.* **2010**, *86*, 101–120.
- (3) Leach, A. R. *Molecular modelling: principles and applications*; 2nd ed.; Pearson Education, 2001.
- (4) Kirchmair, J.; Distinto, S.; Roman Liedl, K.; Markt, P.; Maria Rollinger, J.; Schuster, D.; Maria Spitzer, G.; Wolber, G. *Infect. Disord. - Drug Targets* **2011**, *11*, 64–93.
- (5) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* **1997**, *23*, 3–25.
- (6) Huang, S.-Y.; Zou, X. *Int. J. Mol. Sci.* **2010**, *11*, 3016–3034.
- (7) Moitessier, N.; Englebienne, P.; Lee, D.; Lawandi, J.; Corbeil, C. R. *Br. J. Pharmacol.* **2008**, *153 Suppl*, S7–26.
- (8) Sousa, S. F.; Fernandes, P. A.; Ramon, M. J. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 15–26.
- (9) Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. *J. Comput. Aided. Mol. Des.* **2002**, *16*, 151–166.

- (10) Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. *J. Mol. Biol.* **1982**, *161*, 269–288.
- (11) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. *Nat. Rev. Drug Discov.* **2004**, *3*, 935–949.
- (12) Jorgensen, W. L. *Science* **2004**, *303*, 1813–1818.
- (13) Clark, D. E. *Expert Opin. Drug Discov.* **2006**, *1*, 103–110.
- (14) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J.; Al, M. E. T. *J. Comput. Chem.* **1998**, *19*, 1639–1662.
- (15) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. *J. Comput. Aided Mol. Des.* **1997**, *11*, 425–445.
- (16) Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. *J. Comput. Chem.* **2003**, *24*, 1637–1656.
- (17) Sitkoff, D.; Sharp, K. A.; Honig, B. *J. Phys. Chem.* **1994**, *98*, 1978–1988.
- (18) Still, W. C.; Tempczyk, A.; Hawley, R. C. *J. Am. Chem. Soc.* **1990**, *112*, 6127–6129.
- (19) Birch, L.; Murray, C. W.; Hartshorn, M. J.; Tickle, I. J.; Verdonk, M. L. *J. Comput. Aided Mol. Des.* **2002**, *16*, 855–869.
- (20) De Graaf, C.; Pospisil, P.; Pos, W.; Folkers, G.; Vermeulen, N. P. E. *J. Med. Chem.* **2005**, *48*, 2308–2318.
- (21) Roberts, B. C.; Mancera, R. L. *J. Chem. Inf. Model.* **2008**, *48*, 397–408.
- (22) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. *J. Mol. Biol.* **1996**, *261*, 470–489.
- (23) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997**, *267*, 727–748.
- (24) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. *J. Med. Chem.* **2004**, *47*, 1739–1749.
- (25) Carlson, H. A.; McCammon, J. A. *Mol. Pharmacol.* **2000**, *57*, 213–218.
- (26) Teodoro, M. L.; Phillips, G. N.; Kavral, L. E. In *IEEE International Conference on Robotics & Automation*; 2001; pp. 960–965.
- (27) Carlson, H. A. *Curr. Opin. Chem. Biol.* **2002**, *6*, 447–452.
- (28) Tuffery, P.; Derreumaux, P. *J. R. Soc. Interface* **2012**, *9*, 20–33.
- (29) Fischer, E. *Ber. Dtsch. Chem. Ges.* **1894**, *27*, 2985–2993.
- (30) Koshland, D. E. *Proc. Natl. Acad. Sci.* **1958**, *44*, 98–104.
- (31) Boehr, D. D.; Nussinov, R.; Wright, P. E. *Nat. Chem. Biol.* **2009**, *5*, 789–796.
- (32) Lange, O. F.; Lakomek, N.; Farès, C.; Schröder, G. F.; Walter, K. F. A.; Becker, S.; Meiler, J.; Grubmüller, H.; Griesinger, C.; Groot, B. L. De. *Science* **2008**, *320*, 1471–1475.
- (33) Mittermaier, A.; Kay, L. E. *Science* **2006**, *312*, 224–228.
- (34) Jiang, F.; Kim, S.-H. *J. Mol. Biol.* **1991**, *219*, 79–102.

- (35) Leach, A. R. *J. Mol. Biol.* **1994**, *235*, 345–356.
- (36) Jenwitheesuk, E.; Horst, J. A.; Rivas, K. L.; Van Voorhis, W. C.; Samudrala, R. *Trends Pharmacol. Sci.* **2008**, *29*, 62–71.
- (37) Mooij, W. T. M.; Verdonk, M. L. *Proteins* **2005**, *61*, 272–287.
- (38) Jones, G.; Willett, P.; Glen, R. C. *J. Comput. Aided. Mol. Des.* **1995**, *9*, 532–549.
- (39) Jones, G.; Willett, P.; Glen, R. C. *J. Mol. Biol.* **1995**, *245*, 43–53.
- (40) Brady, R. C.; Bernstein, D. I. *Antiviral Res.* **2004**, *61*, 73–81.
- (41) Comin, M. J.; Agbaria, R.; Ben-Kasus, T.; Huleihel, M.; Liao, C.; Sun, G.; Nicklaus, M. C.; Deschamps, J. R.; Parrish, D. A.; Marquez, V. E. *J. Am. Chem. Soc.* **2007**, *129*, 6216–6222.
- (42) Comin, M. J.; Vu, B. C.; Boyer, P. L.; Liao, C.; Hughes, S. H.; Marquez, V. E. *ChemMedChem* **2008**, *3*, 1129–1134.
- (43) Herdewijn, P.; De Clercq, E. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1591–1597.
- (44) Wang, J.; Froeyen, M.; Hendrix, C.; Andrei, G.; Snoeck, R.; De Clercq, E.; Herdewijn, P. *J. Med. Chem.* **2000**, *43*, 736–745.
- (45) Miralles Llumà, R. Synthesis and computational studies of nucleoside analogues as antiviral agents, Doctoral Thesis, UAB, 2013.
- (46) GOLD 5.2.2, 2011, Cambridge Crystallographic Data Centre (<http://www.ccdc.cam.ac.uk>).
- (47) Korb, O.; Stu, T.; Exner, T. E. *J. Chem. Inf. Model.* **2009**, *49*, 84–96.
- (48) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605–1612.
- (49) Marvin 5.4, 2010, ChemAxon (<http://www.chemaxon.com>).
- (50) Molecular, M.; Field, F.; Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490–519.
- (51) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graph. Model.* **2006**, *25*, 247–260.
- (52) Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. *Proteins: Struct., Funct., Bioinf.* **2003**, *623*, 609–623.
- (53) Elion, G. B.; Furman, P. A.; Fyfe, J. A.; De Miranda, P.; Beauchamp, L.; Schaeffert, H. J. *Proc. Natl. Acad. Sci. U. S. A.* **1977**, *74*, 5716–5720.
- (54) Fyfe, J. A.; Keller, P. M.; Furman, P. A.; Miller, R. L.; Elion, G. B. *J. Biol. Chem.* **1978**, *253*, 8721–8727.
- (55) Harnden, M. R.; Parkin, A.; Parratt, M. J.; Perkins, R. M. *J. Med. Chem.* **1993**, *36*, 1343–1355.
- (56) ChemBioOffice Ultra 13, 2013, PerkinElmer (<http://www.cambridgesoft.com/>).
- (57) Champness, J. N.; Bennett, M. S.; Wien, F.; Visse, R.; Summers, W. C.; Herdewijn, P.; De Clercq, E.; Ostrowski, T.; Jarvest, R. L.; Sanderson, M. R. *Proteins: Struct., Funct., Bioinf.* **1998**, *32*, 350–361.

- (58) Pilger, B. D.; Perozzo, R.; Alber, F.; Wurth, C.; Folkers, G.; Scapozza, L. *J. Biol. Chem.* **1999**, *274*, 31967–31973.
- (59) Chen, M. S.; Prusoff, W. H. *J. Biol. Chem.* **1978**, *253*, 1325–1327.
- (60) Bohman, C.; Balzarini, J.; Wigerincks, P.; Aerschots, A. Van; Herdewijn, P. *J. Biol. Chem.* **1994**, *269*, 8036–8043.
- (61) Martic, M.; Pernot, L.; Westermaier, Y.; Perozzo, R.; Kraljevic, T. G.; Kristafor, S.; Raic-Malic, S.; Scapozza, L.; Ametamey, S. *Nucleosides Nucleotides* **2011**, *30*, 293–315.
- (62) Balzarini, J.; Bohman, C.; De Clercq, E. *J. Biol. Chem.* **1993**, *268*, 6332–6337.
- (63) Schelling, P.; Claus, M. T.; Johner, R.; Marquez, V. E.; Schulz, G. E.; Scapozza, L. *J. Biol. Chem.* **2004**, *279*, 32832–32838.
- (64) Chen, M. S.; Walker, J.; Prusoff, W. H. *J. Biol. Chem.* **1979**, *254*, 10747–10753.
- (65) Perozzo, R.; Jelesarov, I.; Bosshard, H. R.; Folkers, G.; Scapozza, L. *J. Biol. Chem.* **2000**, *275*, 16139–16145.
- (66) Wild, K.; Bohner, T.; Folkers, G.; Schulz, G. E. *Protein Sci.* **1997**, *6*, 2097–2106.
- (67) De Winter, H.; Herdewijn, P. *J. Med. Chem.* **1996**, *39*, 4727–4737.
- (68) Alber, F.; Kuonen, O.; Scapozza, L.; Folkers, G.; Carloni, P. *Proteins: Struct., Funct., Genet.* **1998**, *31*, 453–459.
- (69) Sulpizi, M.; Schelling, P.; Folkers, G.; Carloni, P.; Scapozza, L. *J. Biol. Chem.* **2001**, *276*, 21692–21697.
- (70) Bennett, M. S.; Wien, F.; Champness, J. N.; Batuwangala, T.; Rutherford, T.; Summers, W. C.; Sun, H.; Wright, G.; Sanderson, M. R. *FEBS Lett.* **1999**, *443*, 121–125.
- (71) Brown, D. G.; Visse, R.; Sandhu, G.; Davies, A.; Rizkallah, P. J.; Melitz, C.; Summers, W. C.; Sanderson, M. R. *Nat. Struct. Biol.* **1995**, *2*, 876–881.
- (72) Luyten, I.; De Winter, H.; Busson, R.; Lescrinier, T.; Creuven, I.; Durant, F.; Balzarini, J.; De Clercq, E.; Herdewijn, P. *Helv. Chim. Acta* **1996**, *79*, 1462–1474.
- (73) Sekulic, N.; Shuvalova, L.; Spangenberg, O.; Konrad, M.; Lavie, A. *J. Biol. Chem.* **2002**, *277*, 30236–30243.
- (74) Marquez, V. E.; Choi, Y.; Comin, M. J.; Russ, P.; George, C.; Huleihel, M.; Ben-Kasus, T.; Agbaria, R. *J. Am. Chem. Soc.* **2005**, *127*, 15145–15150.
- (75) Lascau, I.; Gonin, P. *J. Bioenerg. Biomembr.* **2000**, *32*, 237–246.
- (76) Garces, E.; Cleland, W. W. *Biochemistry* **1969**, *8*, 633–640.
- (77) Morera, S.; Lacombe, M.; Yingwu, X.; Lebras, G.; Janin, J. *Structure* **1995**, *3*, 1307–1314.
- (78) Cherfils, J.; Moréra, S.; Lascau, I.; Véron, M.; Janin, J. *Biochemistry* **1994**, *33*, 9062–9069.

