

Université de Sherbrooke

**Caractérisation de la principale cible thérapeutique
du Cytomégalovirus humain : l'ADN polymérase UL54**

Par

M. Frédéric Picard-Jean

Département de biochimie

Mémoire présenté à la Faculté de médecine et des sciences de la santé

**En vue de l'obtention du grade de
Maître ès sciences (M.Sc) en biochimie**

24 avril 2008

Évalué par

Dr. Martin Bisailon, Biochimie

Dr. Simon Labbé, Biochimie

Dr. François Malouin, Biologie



Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-49563-6
Our file Notre référence
ISBN: 978-0-494-49563-6

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

Table des matières

<i>Liste des illustrations</i>	<i>V</i>
<i>Liste des sigles, abréviations et symboles</i>	<i>VII</i>
<i>Résumé</i>	<i>i</i>
1) Introduction	1
1.1) Contexte	1
1.2) Problématique	2
2) Propagation et cycle de réplication virale	3
2.1) Infection et transmission	3
2.2) Cellules cibles	4
2.3) Phylogénétique	4
2.4) Le virion	5
2.5) Entrée virale	6
2.6) Cycle de réplication virale	6
2.6.1) Phase immédiate précoce (IP)	6
2.6.2) Phase précoce (P)	7
2.6.3) Phase tardive (T)	8
3) Pharmacopée	9
3.1) Aperçu	9
3.2) Les drogues présentement approuvées pour le traitement du CMV	10
3.2.1) Le Ganciclovir	10
3.2.2) L'Acyclovir	11
3.2.3) Le Cidofovir	12
3.2.4) Le Foscarnet	12
3.2.5) Le Formivirsen	13
3.3) Cytotoxicité	13
3.4) Apparition de résistance au traitement	14
3.5) Problématique	15
3.6) Avenues de recherches et limitations	15
4) Réplication de l'ADN	17
4.1) Cinétique de la réplication	17
4.2) Protéines virales responsables de la réplication du génome	17
5) ADN polymérase de la famille B	20
5.1) Généralités	20
5.2) Importance des ions divalents	22
5.3) Dynamique de la catalyse	23
5.4) Processivité	24
5.5) Fidélité	25
5.6) Conservation	25
6) ADN polymérase UL54	27
7) Forces intermoléculaires	28
7.1) Interactions ioniques	29
7.2) Forces dipôles-dipôles	29
7.3) Ponts hydrogènes	29
7.4) Forces de van der Waals	30
7.5) Empilement aromatique	30
7.6) Forces hydrophobes	31

7.7) Contribution des forces intermoléculaires	31
8) Objectifs	31
Article 1 : Caractérisation de la liaison de l'ADN et des dNTP à l'ADN polymérase du CMV.	33
9) Discussion et conclusion	83
9.1) Résumé	83
9.2) Format	83
10) Résultats supplémentaires	83
10.1) Modèle par homologie	84
10.2) Quantification des interactions électrostatiques de la liaison UL54-ADN	85
11) Discussion complémentaire	86
11.1) Expression de UL54 restreinte aux acides aminés 293-953	86
11.2) Écart entre les Kd et les Km	88
11.3) Forces responsables de la liaisons UL54-ADN	89
11.3.1) Contexte	89
11.3.2) Forces responsables de la liaison GP43-ADN	90
11.3.3) Forces responsables de la liaison UL54-ADN	90
12) Perspectives	91
12.1) Perspectives à court terme	91
12.1.1) Titration calorimétrique isothermique (microcalorimétrie)	91
12.1.1.1) Contexte	91
12.1.1.2) Choix de la microcalorimétrie	92
12.1.1.3) Exigence de la microcalorimétrie	93
12.1.2) Relâche de molécules d'eau au solvant	94
12.1.2.1) Contexte	94
12.1.2.2) Quantification de la relâche de molécules d'eau au solvant	95
12.1.2.3) Procédure expérimentale	96
12.1.3) Discrimination entre l'ADN et l'ARN par UL44	97
12.1.3.1) Contexte	97
12.1.3.2) Choix de la protéine UL44	97
12.1.3.3) État des connaissances sur UL44	98
12.1.3.4) Procédure expérimentale	98
12.2) Perspectives à long terme	99
12.2.1) Caractériser la fidélité de UL54	99
12.2.1.1) Contexte	99
12.2.1.2) Fidélité des ADN polymérases de la famille B	100
12.2.1.3) Objectifs	100
12.2.1.4) UL54-D413A (exo-KO)	101
12.2.1.5) Études précédentes sur la fidélité des polymérases	101
12.2.1.6) Procédure expérimentale : fidélité globale	102
12.2.1.7) Procédure expérimentale : fidélité brute	105
12.2.1.8) Analyse des résultats	105
12.2.1.9) Fidélité de UL54 en mode élongation	106
12.2.1.10) Solution pour augmenter la sensibilité (si nécessaire)	106
12.2.1.10.1) Expression de UL54 ²⁹³⁻¹²⁴²	107
12.2.1.10.2) Expression de UL54 ^{293-953+22C}	107
13) Remerciements	109

14) Bibliographie110

Liste des illustrations

- Figure 1 : Gauche : Rétine atteinte de rétinopathie causée par le cytomégalo virus humain.
Droite : Rétine d'un individu sain (<http://stlukeseye.com>)..... 2
- Figure 2 : Structure du virion du CMV. (Gauche) Schéma présentant le génome, la capside, le tégument, l'enveloppe et les glycoprotéines de surface. (Droite) La capside, qui contient l'ADN, et l'enveloppe des herpesvirus sont des corps denses en électrons, ce qui permet de les visualiser par microscopie électronique. (Modifié de (Taylor, Brockman et al. 2002))..... 5
- Figure 3 : Histologiquement, l'infection au CMV peut être révélée par microscopie par la détection de corps d'inclusions intranucléaires (yeux de hibou)(GREGORY et TAYLOR 2003)..... 8
- Figure 4 : Structure chimique des trois analogues de nucléotides, de leurs deux prodrogues et de l'analogue du pyrophosphate, et qui sont présentement approuvés pour le traitement du CMV. (Mercorelli, Sinigalia et al. 2007)..... 10
- Figure 5 : Activation des drogues anti-cytomégalo virales. Les prodrogues (Valganciclovir et Valaciclovir) sont estérifiées par des estérases intestinales et hépatiques. Une fois dans la cellule, le Ganciclovir et l'Acyclovir sont phosphorylés par la kinase UL97. Les drogues monophosphorylés (-MP) sont ensuite triphosphorylés (-TP) par des kinases cellulaires. Ces drogues inhibent ensuite la polymérase virale UL54. (N.B. le Cidofovir est déjà un analogue monophosphorylé, c'est pourquoi il est annoté Cidofovir-MP) 11
- Figure 6 : Représentation schématique de l'organisation des six protéines essentielles composant le noyau central du large complexe de réplication du CMV 18
- Figure 7 : Structure tridimensionnelle de la polynucléotide polymérase de la famille B. Plus précisément celle du bactériophage RB69, la protéine gp43. Générée à partir du cristal 2P5O (Hogg, Wallace et al. 2004). Les domaines de la paume, des doigts, du pouce, ainsi que le domaine exonucléase sont indiqués. 21
- Figure 8 : Schéma du site actif d'une ADN polymérase. Le brin matrice (noir), servant de support génétique afin de discriminer le bon dNTP entrant (rouge), est sur le point d'être polymérisé sur l'ADN en cours de copie (bleu). À noter : le rôle crucial des deux ions magnésium (vert) lors de la catalyse. Ces cations sont coordonnés par deux acides aspartiques absolument conservés (magenta). (numérotation des asp suivant ceux de la polymérase de RB69) (Modifié de (Steitz 1998))..... 22
- Figure 9 : L'alignement de plusieurs ADN polymérases de la famille a révélé la présence de six domaines hautement conservés I- VI (boîte noire). (Wang, Wong et al. 1989).26
- Figure 10 : (A) Structure tridimensionnelle de la polymérase de la famille B gp43, l'ADN polymérase du bactériophage RB69. Générée à partir du cristal 2P5O (Hogg, Wallace et al. 2004). (B) Modélisation par homologie de la polymérase UL54²⁹³⁻⁹⁵³ du cytomégalo virus (DeepView). (C) Superposition de A et B. 84

- Figure 11 : Constante de dissociation de UL54 avec un ADN de 30nt, en fonction d'une force ionique croissante 85
- Figure 12 : Exemple de résultat attendu lors d'un test visant à quantifier la relâche de molécules d'eau vers le solvant. La liaison entre UL54 et l'ADN serait affectée par la molalité en soluté neutre..... 96
- Figure 13 : Protéine UL44 du cytomégalovirus humain. Électrophorèse sur gel de polyacrylamide de UL44 exprimé en bactérie et purifié grâce à un épitope exahistidine. Coloration au bleu de Coomassie (gauche) et immunobuvardage de type western (droite)..... 99
- Figure 14 : Construction des matrices d'ADN utilisés pour les essais de fidélité de UL54. A) de haut en bas : matrice pour incorporation d'un dATP sur une guanine, d'un dTTP sur une guanine et d'un dTTP sur une tyrosine. B) matrice pour incorporation d'un dATP sur une guanine lorsque UL54 est en élongation..... 102
- Figure 15 : Schéma expérimental visant à déterminer la fidélité de UL54. De haut en bas : l'ADNdb radiomarqué est incubé avec une concentration constante d'enzyme et une concentration croissante de dNTP. La réaction est arrêtée après 30 minutes d'incubation à 37°C. L'élongation du brin amorce radiomarqué est ensuite résolue par électrophorèse sur gel de polyacrylamide dénaturant 104

Liste des sigles, abréviations et symboles

μM	Micromolaire (1×10^{-6} moles par litre)
Å	Ångström, unité de mesure équivalent à 0,1 nanomètre (1×10^{-10} mètre)
ADN	Acide désoxyribonucléique
ADNdb	Acide désoxyribonucléique double brin
ADNsb	Acide désoxyribonucléique simple brin
Arg	Arginine
ARN	Acide ribonucléique
ARNdb	Acide ribonucléique double brin
ARNm	ARN messenger
ARNsb	Acide ribonucléique simple brin
Asp	Acide aspartique
ATP _{γ} ^{32P}	Adénosine triphosphate dont le phosphate gamma est radioactif
a_w	Activité de l'eau
c.-à-d	C'est-à-dire
CD34	Facteur d'adhésion encodé par le gène humain CD34 (utilisé comme marqueur de différenciation cellulaire)
CeHV-6	Cytomégalovirus du singe
CMV	Cytomégalovirus (Herpesvirus humain 5)
D413A	Protéine portant une mutation de l'aspartate 413 en alanine
dNTP	Désoxyribonucléotide triphosphate (soit dATP, dCTP, dGTP ou dTTP)
dNMP	Désoxyribonucléotide monophosphate (soit dAMP, dCMP, dGMP ou dTMP)
dNTP α ^{32P}	Désoxyadénosine triphosphate dont le phosphate alpha est radioactif
exo-KO	Domaine exonucléase knockout
FDA	Food and drug administration
Glu	Acide glutamique
gp43	Polymérase du bactériophage RB69
His	Histidine
IC ₅₀	Concentration requise pour inhiber 50% de la réaction
IE2	Protéine virale immédiate précoce 2
IP	Phase ou gènes immédiats précoces
kbp	Kilopaire de base (1000)
K _{cat}	Constante catalytique
KCl	Chlorure de calcium

K _d	Constante de dissociation (1/K _a)
kDa	Kilodalton
K _m	Constante de Michaelis-Menten
Lys	Lysine
mM	Millimolaire (1x10 ⁻³ moles par litre)
MuHV-1	Cytomégalovirus de le souris
nm	Nanomètre
NMWL	Limite de poids moléculaire nominale
NS3	ARN-triphosphatase du virus du Nil Occidental (elle possède aussi les activités de protéase de l'hélicase)
NS5B	ARN-polymérase-ARN-dépendante du virus de l'hépatite C
<i>ori</i> Lyt	Origine de réplication lytique
P	Phase ou gènes précoces
PCNA	Antigène nucléaire de prolifération cellulaire (facteur de processivité de la polymérase humaine)
Phase S	Phase de synthèse (cycle cellulaire)
pK _a	Constante de dissociation d'un acide
PoHV-4	Cytomégalovirus du chimpanzé
Pol	Polymérase
RB69	Bactériophage RB69
RMN	Résonance magnétique nucléaire
rNTP	Ribonucléotide triphosphate (soit rATP, rCTP, rGTP ou rUTP)
SIDA	Syndrome d'immuno-déficience acquise
T	Phase ou gènes tardif
Tth pol	Polymérase de <i>Thermus thermophilus</i>
UL	Région unique longue du génome viral
UL102	Sous-unité de l'hélicase virale hétérotrimérique
UL105	Sous-unité de l'hélicase virale hétérotrimérique
UL112-113	Protéines virales responsables de l'activation de la transcription
UL114	Uracile ADN glycosylase virale
UL44	Facteur de processivité de la polymérase virale
UL46	Protéine virale de la nucléocapside
UL48	Protéine virale de la nucléocapside
UL49	Protéine virale de la nucléocapside
UL54	Sous-unité catalytique de la polymérase virale
UL54 ²⁹³⁻¹²⁴²	Protéine UL54 comprenant les acides aminés 293 à 1242
UL54 ²⁹³⁻⁹⁵³	Protéine UL54 comprenant les acides aminés 293 à 953
UL54 ^{293-953+22C}	Protéine UL54 comprenant les acides aminés 293 à environ 953, en fusion avec les 22 acides aminés carboxyterminaux de UL54WT (générant ainsi une protéine de 683 a.a.)

UL54-ADN	Complexe formé par la liaison d'une molécule d'ADN à la protéine UL54
UL54WT	Protéine UL54 de type sauvage
UL55	Glycoprotéines virales de surface
UL57	Protéine virale de liaison à l'ADN simple brin
UL70	Sous unité de l'hélicase virale hétérotrimérique
UL80	Protéine virale de la nucléocapside
UL84	Protéine virale de liaison à l'ARN (tige boucle)
UL85	Protéine virale de la nucléocapside
UL86	Protéine virale de la nucléocapside
UL97	Protéine kinase
VHC	Virus de l'hépatite C
VHS	Virus de l'herpès simplex
VNO	Virus du Nil Occidental
zJ	Zepto-Joules (1×10^{-21} joules)
ΔG	Énergie libre de Gibbs
ΔH	Enthalpie
Δn_w	Changement du nombre de molécules d'eau exclus du soluté
ΔS	Entropie
$\Delta \Delta G$	Variation de l'énergie libre de Gibbs

Résumé

Université de Sherbrooke

Caractérisation de la principale cible thérapeutique du Cytomégalovirus humain : l'ADN polymérase UL54

Par

M. Frédéric Picard-Jean

Département de biochimie

Mémoire présenté à la Faculté de médecine et des sciences de la santé

En vue de l'obtention du grade de

Maître ès sciences (M.Sc) en biochimie

24 avril 2008

Évalué par

Dr. Martin Bisailon, Biochimie

Dr. Simon Labbé, Biochimie

Dr. François Malouin, Biologie

Le cytomégalovirus humain (CMV) est un Herpèsvirus causant une infection latente chez 60% à 70% des nord-américains. Sa primo-infection chez les nouveaux-nés et sa réactivation chez les individus immunodéprimés sont associées à de nombreux cas de morbidité et de mortalité. Le traitement des infections au CMV est compliqué par l'absence de vaccin et le nombre restreint de drogues homologuées. Ces dernières

sont malheureusement issues de vieilles générations de composés, et sont associées à une activité modérée, à une importante toxicité, et à l'apparition fréquente de résistance au traitement. Ces composés ciblent tous l'ADN polymérase virale, et inhibent ainsi la réplication du virus. Cependant, comparativement aux autres polymérases virales, cette enzyme est bien peu caractérisée, ce qui limite le développement de composés thérapeutiques de seconde génération plus efficaces, moins toxiques et moins sujets à l'apparition de résistance virale. Ayant entendu ce signal d'alarme, nous avons entrepris de caractériser la principale cible thérapeutique du cytomegalovirus humain, son ADN polymérase. Le gène UL54 du CMV code une protéine de 1242 acides aminés et qui est connue pour être la sous-unité catalytique de l'ADN polymérase virale. Malheureusement, sa grande taille a longtemps limité son expression et sa caractérisation. En nous concentrant sur les régions importantes pour son activité, nous avons généré une protéine un peu plus courte, mais toujours catalytiquement active. Nous avons ainsi pu exprimer, pour la première fois, une grande quantité de protéines UL54, ce qui nous a permis de caractériser en détails sa liaison avec ses deux substrats principaux, soit l'ADN et les dNTP. L'emploi de la spectroscopie à fluorescence nous a permis de suivre la liaison de UL54 à ses substrats, en quantifiant l'impact de la liaison de ces derniers sur la fluorescence intrinsèque du tryptophane de UL54. Nous avons ainsi pu établir que les constantes de dissociation de UL54 pour l'ADN et pour les dNTP étaient respectivement de 48 μ M et de 15 μ M. Notre étude révèle qu'un substrat d'ADN aussi petit que 6nt peut lier la protéine. Nous avons aussi démontré que UL54 lie aussi bien l'ADNsb que l'ADNdb, et que cette liaison semble séquence indépendante. De plus, la protéine est incapable de discriminer entre un substrat d'ADN et un d'ARN. Le profil

thermodynamique de la liaison à l'ADN et aux dNTP a été établi et révèle deux modes de liaison distincts. La liaison de l'ADN est propulsée par l'enthalpie et est fort probablement associée à la relâche de molécules d'eau au solvant. Les interactions hydrophobes et d'empilement représentent les forces majeures stabilisant la liaison, alors que les interactions électrostatiques sont presque négligeables. À l'inverse, la liaison de UL54 au dNTP est propulsée par l'enthalpie, et la stabilisation de la liaison est associée davantage aux interactions électrostatiques, à des ponts hydrogènes et aux forces de van der Waal's. Des essais de dichroïsme circulaire, une seconde technique optique qui peut générer une représentation des structures secondaires et tertiaires d'une protéine, indique que, suite à la liaison de l'ADN, la protéine UL54 opère un changement de conformation, et qu'elle en subit un second suite à la liaison d'un dNTP, ce qui est cohérent avec les changements de conformation observés chez d'autres polymérases mieux caractérisées. Quelques expériences complémentaires sont aussi présentées. Enfin, comme les résistances aux traitements sont un enjeu majeur, et qu'elles sont créées par des erreurs de UL54, l'élaboration d'une démarche expérimentale visant à établir la fidélité de cette protéine est décrite en détail. Nous sommes confiants qu'une telle étude sur la caractérisation biochimique de la principale cible pharmacologique du CMV contribuera au développement rationnel de nouvelles générations de drogues anti-cytomégaloviriques efficaces.

Mots clés : ADN polymérase, Virus à ADN, Cytomégalovirus Humain, liaison de substrat, thermodynamique.

1) Introduction

1.1) Contexte

Au cours du XX^e siècle, l'espérance de vie des Canadiens a bondi de 50 à près de 80 ans. Ces progrès remarquables sont, en grande partie, attribuables à la diminution de la mortalité infantile et au recul des maladies infectieuses (Statistique_Canada 2008). La médecine moderne a réussi à complètement éradiquer le virus de la variole (WorldHealthOrganisation 2008), et elle dispose maintenant d'un arsenal de vaccins contre de nombreux autres, afin de prévenir de nouvelles épidémies. Plus récemment, la découverte de certains antiviraux a permis de mieux traiter plusieurs infections virales. Malgré ces progrès prodigieux, certaines familles de virus représentent encore un défi de taille pour la recherche. C'est notamment le cas des Herpèsvirus, qui sont parmi les pathogènes les mieux adaptés à l'homme (Bego et St-Jeor 2006). L'un des Herpèsvirus les plus répandus à l'échelle planétaire est le Cytomégalo virus (CMV), sa séroprévalence variant de 60% à 100% selon le niveau de développement des pays (Bego et St-Jeor 2006). Chez les sujets immunocompétents, la primo-infection au CVM se traduit par des symptômes modérés s'apparentant à ceux de la mononucléose, comme des maux de tête, de la fatigue ou une angine (laryngite, pharyngite, amygdalite ou staphylite), suite à quoi s'installe une latence virale incurable ponctuée de réactivations généralement asymptomatiques (Gerna, Baldanti et al. 2004). Il y a cependant deux sous-classes de la population à risque de manifestations beaucoup plus graves; ce sont les nouveau-nés, au système

immunitaire encore immature, et les individus aux prises avec une immunodépression. Ces derniers sont susceptibles à une réactivation du virus associée à des symptômes sévères, incluant des pneumonies, des maladies gastro-intestinales, et des rétinopathies (**Figure 1**) (Loregian et Palù 2005) . Dans les pays industrialisés, le CMV représente l'infection materno-fœtale la plus fréquente; jusqu'à un enfant sur 50 peut naître déjà infecté par le CMV, et ce taux peut grimper jusqu'à 40% dans le cas d'une primo-infection maternelle. Parmi les fœtus infectés, 10 à 20% souffriront de séquelles graves pouvant mener à une surdité, une microcéphalie, un retard mental, l'autisme, des difficultés d'apprentissage ou encore une mort fœtale (Malm et Engman 2007).



Figure 1 : Gauche : Rétine atteinte de rétinopathie causée par le cytomégalovirus humain. Droite : Rétine d'un individu sain (<http://stlukeseve.com>).

1.2) Problématique

L'incidence d'infections symptomatiques au CMV augmente dû au nombre grandissant d'individus immunodéprimés en raison de greffe d'organe, de traitement de chimiothérapie ou du SIDA (Loregian et Palù 2005). Dans cette lutte contre le CMV, il n'y a toujours aucun vaccin de disponible, et aucun traitement spécifique n'a encore été développé. Certaines drogues, majoritairement issues de programmes de recherches anticancéreux sur les voies métaboliques des purines et pyrimidines, ont démontré avoir une efficacité antivirale modérée contre le CMV. Ce sont le ganciclovir/valganciclovir (prodrogue), l'acyclovir, le cidofovir et le foscarnet (Biron 2006). Cependant, ces drogues n'offrent qu'une efficacité réduite due à une

pharmacocinétique limitée, des problèmes de toxicité et l'apparition de résistances à la thérapie. C'est pourquoi il y a un criant besoin pour de nouvelles drogues efficaces contre les Herpèsvirus (Loregian et Palù 2005). Mon projet propose de mieux comprendre le cytomégalovirus, plus particulièrement l'enzyme responsable de sa réplication, afin de contribuer au développement de nouveaux traitements efficaces.

2) Propagation et cycle de réplication virale

2.1) Infection et transmission

Le Cytomégalovirus humain est largement présent à la grandeur de la planète, et sa séroprévalence est de 60-70% dans les pays développés. Cependant, celle-ci grimpe à 90-100% chez les hommes homosexuels, les groupes socio-économiques défavorisés et les pays en voie de développement (Gandhi et Khanna 2004; Bego et St-Jeor 2006). Le virus se transmet facilement par contact rapproché entre individus, via un transfert placentaire (de la mère au fœtus) ou par l'entremise de fluides corporels contaminés comme l'urine, la salive, le sperme, les sécrétions vaginales et le lait maternel (Biron 2006; Malm et Engman 2007). Les individus ainsi infectés par le CMV le demeureront à vie due à l'établissement d'une latence virale incurable suivant la primo-infection. C'est ce qui explique que la prévalence de ce virus augmente avec l'âge et est marquée par deux pics d'infections. Le premier, chez les enfants en bas âge, est le résultat de la transmission verticale ainsi que de la contamination par l'urine, les sécrétions respiratoires et la proximité dans les garderies, alors que le second, chez les jeunes adultes, est majoritairement le résultat de la transmission par voie sexuelle (Rowshani, Bemelman et al. 2005).

2.2) Cellules cibles

À l'intérieure de son hôte, le CMV se réplique efficacement dans plusieurs types cellulaires incluant les hépatocytes, les granulocytes, les macrophages, les cellules neuronales, épithéliales, musculaires lisses, myéloïdes, endothéliales et les fibroblastes, les trois derniers représentant son réservoir principal (Schrier, Nelson et al. 1985; Soderberg, Larsson et al. 1993; Sinzger, Grefte et al. 1995; Plachter, Sinzger et al. 1996; Boeckh, Holy et al. 1998; Kahl, Sierge-Axel et al. 2000). De plus, le CMV peut instaurer une latence dans plusieurs types cellulaires, dont les cellules hématopoïétiques progénitrices CD34+, qui lui servent de réservoir renouvelable de virus latents (Goodrum, Jordan et al. 2002). C'est suite à la différenciation finale des cellules hématopoïétiques de la lignée myéloïde en macrophages que le virus peut se réactiver, exposant ainsi un hôte immunodéprimé à une infection aiguë (Taylor-Wiedeman, Sissons et al. 1991; Maciejewski, Bruening et al. 1993; Soderberg, Larsson et al. 1993; Taylor-Wiedeman, Sissons et al. 1994).

2.3) Phylogénétique

Le cytomégalovirus humain (Herpèsvirus Humain 5) est un membre de la famille des Herpèsvirus, qui comprend entre autres les virus de l'herpès simplex, d'Epstein-Barr et de la varicelle-zona. Il appartient à la sous-famille des β -Herpèsvirus et n'infecte que l'homme, bien qu'il y ait des homologues connus pour d'autres mammifères tels le CMV de souris (MuHV-1), le CMV du chimpanzé (PoHV-4) ou encore le CMV du singe rhésus (CeHV-6) (Pellett, Roizman et al. 2007).

2.4) Le virion

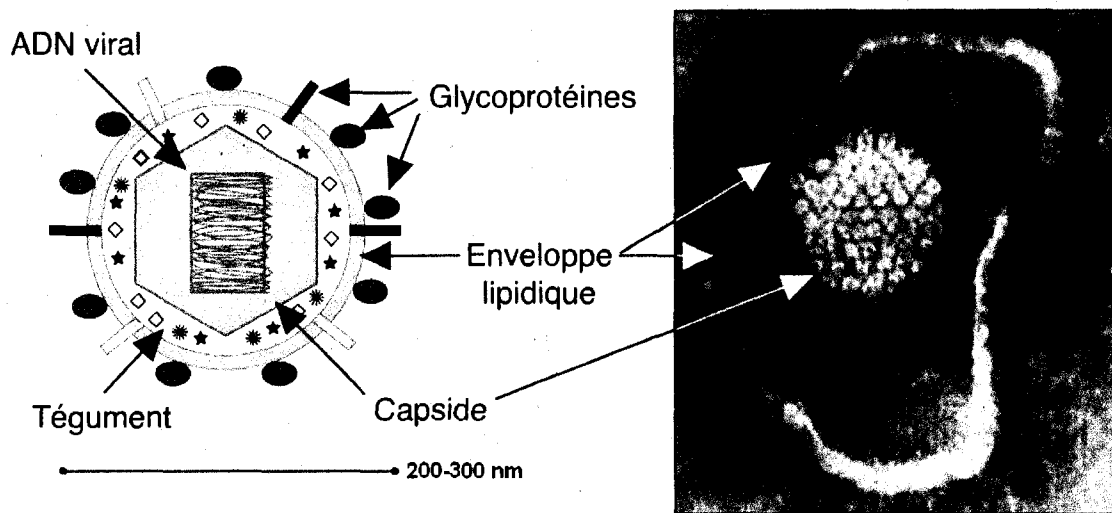


Figure 2 : Structure du virion du CMV. (Gauche) Schéma présentant le génome, la capside, le tégument, l'enveloppe et les glycoprotéines de surface. (Droite) La capside, qui contient l'ADN, et l'enveloppe des herpesvirus sont des corps denses en électrons, ce qui permet de les visualiser par microscopie électronique. (Modifié de (Taylor, Brockman et al. 2002))

Le CMV est un gros virus de 200-300nm de diamètre composé de 71 protéines virales et de 70 protéines cellulaires (**Figure 2**) (Varnum, Streblow et al. 2004). Il est enveloppé d'une bicouche lipidique provenant du réticulum endoplasmique de la cellule hôte et exprime à sa surface au moins vingt glycoprotéines virales dont l'une des plus hautement conservées est UL55 (gB), qui est essentielle à la reconnaissance et l'entrée virale. Elle est aussi une des cibles principales de la réponse immunitaire humorale et une cible potentielle pour le développement d'un éventuel vaccin. Sous la bicouche lipidique se trouve le tégument, une région non structurée représentant 40% de la masse du virion et issue du noyau et du cytoplasme cellulaires. On y retrouve des protéines virales et cellulaires impliquées dans l'assemblage, le désassemblage et la décapsidation du virus. D'autres protéines sont responsables de moduler la réponse de la cellule hôte contre les infections. Le tégument recèle aussi

une nucléocapside icosaédrique de 130nm de diamètre, celle-ci étant formée de 162 capsomères composés de cinq protéines virales encodées par les cadres de lecture UL86, UL85, UL80, UL48-49 et UL46. La nucléocapside contient l'énorme génome viral, un ADN linéaire double brin de 235kbp, ainsi qu'un ARN associé à l'origine de réplication lytique d'ADN (*oriLyt*) (Mocarski, Shenk et al. 2007). Cet imposant génome code pour un grand nombre de protéines, les prédictions variant de 166 à 252 cadres de lectures ouverts (Mocarski, Shenk et al. 2007), mais une estimation de 200 semble faire consensus (Varnum, Streblow et al. 2004; Rowshani, Bemelman et al. 2005; Bego et St-Jeor 2006).

2.5) Entrée virale

Lors de l'infection d'une cellule, le CMV reconnaît spécifiquement les récepteurs membranaires en liant les sulfates d'héparan. Il s'en suit la fusion de son enveloppe avec la membrane cellulaire, ce qui relâche la nucléocapside dans le cytoplasme. Cette dernière s'associe alors aux microtubules qui la guident jusqu'à un pore nucléaire où la décapsidation délivre le génome viral au noyau (Mocarski, Shenk et al. 2007). Une fois l'ADN du CMV au noyau, la transcription de ses gènes s'opère en trois phases : immédiate précoce (IP), précoce (P) et tardive (T) (Stenberg 1996).

2.6) Cycle de réplication virale

2.6.1) Phase immédiate précoce (IP)

De une à trois heures suivant l'infection, apparaissent les premières protéines non structurales encodées par les gènes IP. Elles ont pour rôle de contrôler l'expression de gènes viraux et cellulaires afin d'optimiser l'environnement pour la production

ultérieure d'ADN viraux (Sinclair et Sissons 2006). C'est notamment lors de cette phase que les protéines impliquées dans la réplication du génome viral, comme la protéine kinase UL97 et l'ADN polymérase UL54, commencent à être exprimées (Rowshani, Bemelman et al. 2005). Ces gènes IP sont cruciaux pour le virus, car l'expression des gènes précoces et tardifs en dépend (Spector 1996). C'est d'ailleurs pourquoi les protéines IP demeurent toujours présentes, même en périodes de latence. Leur expression ne dépend d'aucun gène viral, mais est plutôt influencée par le niveau de différenciation de la cellule; les cellules faiblement différenciées, qui appliquent une plus grande répression sur le promoteur IP, maintiennent le virus latent (Gandhi et Khanna 2004).

2.6.2) Phase précoce (P)

La phase précoce s'étend entre 6 et 24 heures suivant l'infection, au moment où la synthèse de l'ADN virale se prépare. Cette seconde phase coïncide avec l'expression d'au moins 23 gènes cruciaux impliqués dans la réplication génomique, dans la maturation de la capside, et dans l'altération de l'environnement cellulaire (Yu, Silva et al. 2003). À l'expression de ces gènes cruciaux s'ajoute celle de plusieurs gènes importants, notamment impliqués dans la modulation de la réponse antivirale de la cellule hôte. L'expression de UL54 (ADN polymérase) continue à augmenter, et on voit apparaître les protéines impliquées dans l'initiation de la réplication, le tout, dans des cellules où le cycle cellulaire est bloqué en mimant la phase S. En effet, ce blocage permet au virus de bénéficier des fonctions de transcription et de réplication de l'ADN de la cellule hôte (Hertel et Mocarski 2004). Dans ce contexte, la synthèse de l'ADN devient détectable entre 14 et 16 heures post-infection (Browne, Wing et

al. 2001). Rapidement, la grosseur et le nombre de compartiments réplicatifs au sein du noyau augmentent et on voit apparaître une morphologie cellulaire typique au cytomegalovirus, avec un large noyau contenant des corps d'inclusion, aussi appelés yeux de hibou (**Figure 3**) (Anders, Kerry et al. 2006).



Figure 3 : Histologiquement, l'infection au CMV peut être révélée par microscopie par la détection de corps d'inclusions intranucléaires (yeux de hibou)(GREGORY et TAYLOR 2003).

2.6.3) Phase tardive (T)

Encore peu d'informations sont connues sur la régulation des gènes impliqués dans la phase tardive, qui s'amorce 24 heures après le début de l'infection (Mocarski, Shenk et al. 2007). Lors de cette phase, l'expression de beaucoup de gènes précoces se poursuit et la réplication du génome viral continue à s'intensifier. Quelques gènes particuliers commencent aussi à être exprimés avec une cinétique similaire à ceux de la phase précédente. Les protéines tardives sont surtout impliquées dans la maturation de la capsid, l'encapsidation de l'ADN, la maturation des virus et la relâche des virions (Mocarski, Shenk et al. 2007). À partir du moment où les protéines tardives commencent à s'accumuler, l'ADN au noyau commence à être efficacement encapsidé. Son export vers le cytoplasme est encore controversé, bien que de plus en plus de preuves solides supportent l'hypothèse de l'enveloppement/désenveloppement/réenveloppement (Muranyi, Haas et al. 2002; Mettenleiter 2004; Britt, Arvin et al. 2006). Selon cette théorie, la nucléocapside subirait un premier enveloppement à la membrane nucléaire interne, puis se désenvelopperait au cytoplasme, pour ensuite se réenvelopper au réticulum

endoplasmique, celui-ci étant responsable du transport et de l'exocytose des virus (Britt, Arvin et al. 2006). Une fois relâchés, les virions ont une demie-vie de 24 à 48 heures (Riddell, Warren et al. 2000).

3) Pharmacopée

3.1) Aperçu

Cliniquement, l'arsenal thérapeutique dont nous disposons à l'heure actuelle pour traiter les multiples manifestations du CMV humain est très limité. Seulement quatre composés sont présentement homologués pour le traitement systémique du CMV. Le Cidofovir, l'Acyclovir/Valaciclovir et le Gancliclovir/Valganciclovir sont des analogues de nucléotides, et le Foscarnet est un analogue du pyrophosphate (**Figure 4**) (Mercorelli, Sinigalia et al. 2007). Une cinquième thérapie, un ARN antisens commercialisé sous le nom de fomivirsen, est aussi approuvée pour le traitement local des rétinopathies (Grillone et Lanz 2001). À l'exception des deux derniers, ces médicaments n'étaient pas originalement dirigés contre le CMV, mais ils étaient plutôt le fruit de programmes de recherches anti-cancéreuses sur les analogues des purines et pyrimidines durant les années 1970. Le seul traitement spécifiquement développé contre le cytomegalovirus est le fomivirsen, qui fût approuvé par la Food and Drug Administration (FDA) en 1998. Voici quelques informations sur les drogues présentement commercialisées contre le CMV.

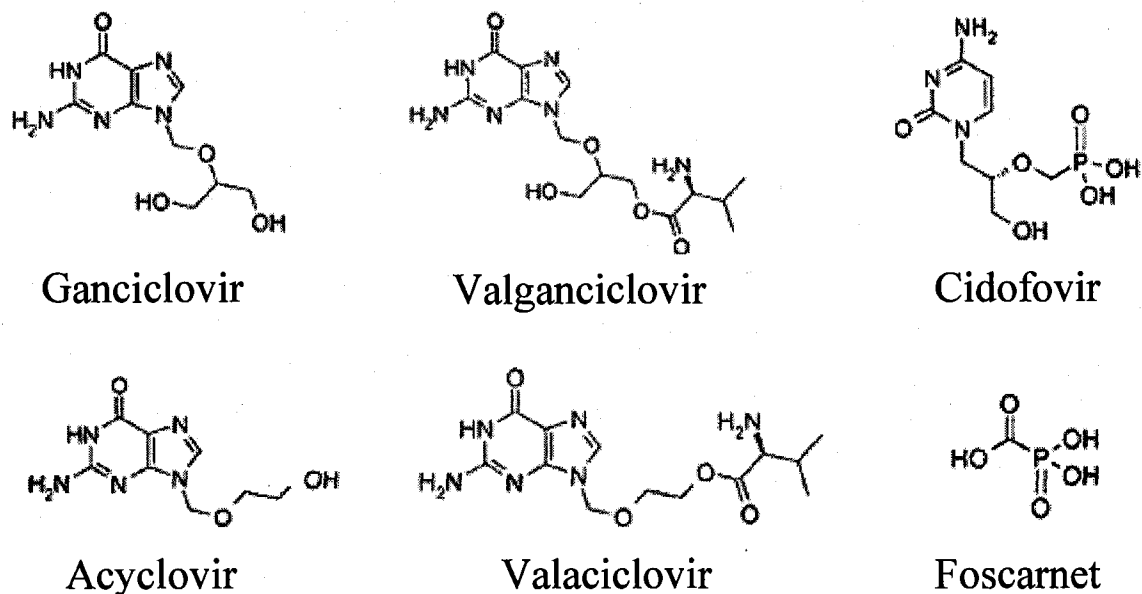


Figure 4 : Structures chimiques des trois analogues de nucléotides, de leurs deux prodrogues et de l'analogue du pyrophosphate, et qui sont présentement approuvés pour le traitement du CMV. (Mercorelli, Sinigalia et al. 2007)

3.2) Les drogues présentement approuvées pour le traitement du CMV

3.2.1) Le Ganciclovir

Le Ganciclovir [9-(1,3-dihydroxy-2-propoxyméthyl)-guanine] est présentement l'outil de choix pour le traitement du cytomégalo virus humain (Markham et Faulds 1994; Mercorelli, Sinigalia et al. 2007). Il s'agit d'un analogue de guanosine non phosphorylé qui subit une première phosphorylation catalysée par la kinase virale UL97, pour être ensuite triphosphorylé par des kinases cellulaires (**Figure 5**) (Littler, Stuart et al. 1992). C'est le Ganciclovir-triphosphate qui interagit avec la polymérase UL54 et inhibe la synthèse d'ADN en compétitionnant avec le nucléoside naturel (dGTP) et en agissant comme terminateur de chaîne. Il inhibe la polymérase virale plus efficacement qu'il peut inhiber la polymérase cellulaire α . Il est présentement

distribué par Roche, sous forme intraveineuse, ainsi que par Baush & Lomb, sous forme d'implants oculaires à libération contrôlée pour le traitement des rétinopathies. Il est classifié comme agent cytotoxique et peut induire une suppression de la moelle osseuse menant à des neutropénies, des leucopénies, des granulocytopenies, de l'anémie, etc. (Razonable et Emery 2004). Des mutations conférant une résistance au Ganciclovir ont été recensées au sein des gènes UL54 et UL97. Une grande partie des mutations cartographiées chez UL54 sont localisées dans son domaine exonucléase (a.a. 300-545) (Mercorelli, Sinigalia et al. 2007). Depuis 2000, il est aussi disponible sous forme d'une prodrogue, le Valganciclovir, qui a une biodisponibilité orale de 60%.

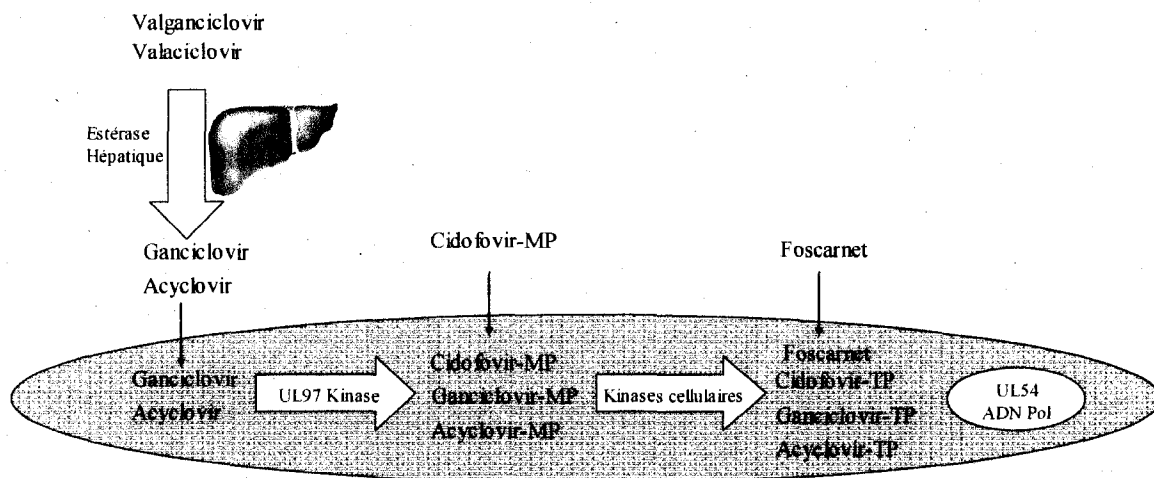


Figure 5 : Activation des drogues anti-cytomégalo-virales. Les prodrogues (Valganciclovir et Valaciclovir) sont estérifiées par des estérases intestinales et hépatiques. Une fois dans la cellule, le Ganciclovir et l'Acyclovir sont phosphorylés par la kinase UL97. Les drogues monophosphorylés (-MP) sont ensuite triphosphorylés (-TP) par des kinases cellulaires. Ces drogues inhibent ensuite la polymérase virale UL54. (N.B. le Cidofovir est déjà un analogue monophosphorylé, c'est pourquoi il est annoté Cidofovir-MP)

3.2.2) L'Acyclovir

L'Acyclovir [9-(2-hydroxyethoxyméthyl)-guanine] est aussi un analogue de guanosine. Il subit la même activation que le Ganciclovir (**Figure 5**), possède le

même mécanisme d'action et les mêmes mécanismes de résistance, à la différence que sa demi-vie est plus courte et qu'il est un substrat moins efficace que le Ganciclovir. Son utilisation principale est dans la prévention de la réactivation virale ou dans le traitement prophylactique chez les greffés.

3.2.3) Le Cidofovir

Le Cidofovir [(S)-1-(3-hydroxy-2-phosphonylmethoxypropyl) cytosine] est, pour sa part, un analogue de dCMP. Il partage les mêmes mécanismes d'activation, d'inhibition et d'administration que le Ganciclovir, à la différence qu'il ne requiert pas d'activation par UL97, car il est déjà monophosphorylé (**Figure 5**). Par conséquent, les mutations conférant une résistance au Cidofovir n'ont été cartographiées que dans le gène UL54. Son administration se fait sous forme intraveineuse et elle comporte plusieurs effets secondaires; cependant, le plus important est sa néphrotoxicité. Il est utilisé comme médicament de seconde ligne.

3.2.4) Le Foscarnet

Le Foscarnet [phosphonoformic acid] est un analogue du pyrophosphate. Il inhibe l'ADN polymérase virale en liant le site de liaison du pyrophosphate et bloque ainsi le clivage du groupement pyrophosphate du dNTP, prévenant de cette façon son incorporation sur le brin d'ADN en élongation. Les mutations conférant une résistance virale au Foscarnet ont aussi été répertoriées dans le gène UL54. Ce médicament est considéré comme un traitement de seconde ligne, mais il est le traitement de choix chez les patients aux prises avec un échec au traitement au Ganciclovir, que ce soit dû à la résistance virale ou à la sévérité de la neutropénie ou

de la leucopénie (Razonable et Emery 2004). Il est cependant néphrotoxique et son administration intraveineuse doit être faite sous surveillance (Naesens et Clercq 2001).

3.2.5) Le Formivirsén

Le Formivirsén fut approuvé par la FDA en 1998. Ce faisant, il devenait le premier ARN antisens commercialisé pour usage thérapeutique (Marwick 1998). Il s'agit d'un ARN de 21 nucléotides, dont la séquence est complémentaire à une région de l'ARNm du gène immédiat précoce 2 (IE2), et qui bloquerait ainsi l'expression de cette dernière. Cette drogue est spécifique contre le CMV, chez qui elle s'est avéré être efficace (Perry et Balfour 1999). Cependant, le Formivirsén n'est disponible que par injection intraoculaire, pour le traitement local des rétinopathies, surtout chez les patients atteints du syndrome d'immunodéficience humaine acquise (Grillone et Lanz 2001).

3.3) Cytotoxicité

De nombreux facteurs peuvent contribuer à la cytotoxicité d'une drogue, les plus fréquents étant le manque de spécificité pour la cible thérapeutique, des métabolites toxiques, des interactions non-spécifiques avec d'autres protéines de l'hôte, ou une bioaccumulation au site d'élimination. Chez les drogues anticytomégaloviriques, c'est le manque de spécificité qui est la principale cause de cytotoxicité (Rubsam, Boucher et al. 1999). En effet, la polymérase UL54 arbore une grande homologie de structure avec les ADN polymérases cellulaires. Bien que ces drogues aient une affinité plus grande pour la polymérase virale que celles cellulaires, elles peuvent parfois être

utilisées comme substrats par la polymérase α , et ainsi générer des altérations génétiques menant à une cytotoxicité. D'ailleurs, de nombreuses études démontrent que ces analogues de nucléotides, dont le Ganciclovir, peuvent être incorporés à l'ADN cellulaire et ainsi induire une phase S différée, ce qui peut mener à la mort cellulaire (Rubsam, Davidson et al. 1998; Boucher, Ostruszka et al. 2000). Ces études sont supportées par d'autres qui démontrent que des cellules possédant des systèmes de réparation de l'ADN défectueux (ADN Polymerase β délété (knockout) ou inhibé) présentent une susceptibilité accrue à la cytotoxicité induite par le Ganciclovir (Tomicic, Thust et al. 2001). En plus de ce manque de spécificité pour leurs cibles, plusieurs de ces médicaments, probablement dû à leur analogie avec les nucléotides, peuvent interagir avec d'autres protéines de l'hôte. Un bon exemple est celui du Cidofovir, qui lie avec une grande affinité un transporteur anionique des tubules proximaux rénaux, ce qui cause une accumulation de la drogue dans le cortex rénal (Ho, Lin et al. 2000). C'est d'ailleurs ce qui est responsable de sa néphrotoxicité. Des composés plus spécifiques pourraient diminuer l'impact sur les polymérases cellulaires, et des composés allostériques, qui ne seraient pas des analogues du substrat d'un très grand nombre d'enzymes cellulaires, pourraient contribuer à diminuer les interactions non-spécifiques avec des protéines de l'hôte.

3.4) Apparition de résistance au traitement

Toutes les thérapies systémiques disponibles à ce jour ciblent UL54 et beaucoup d'entre elles partagent le même mécanisme d'action. Par conséquent, il est fréquent d'observer des résistances croisées entre ces traitements. Ainsi, des mutations chez UL97 peuvent conférer une résistance au Ganciclovir et à l'Acyclovir (Mercorelli,

Sinigalia et al. 2007). Des mutations conférant à UL54 une résistance au Ganciclovir lui confèrent également généralement une résistance à l'Acyclovir, au Cidofovir, et potentiellement aussi au Foscarnet, dépendamment du domaine ayant subi une mutation (Mercorelli, Sinigalia et al. 2007). C'est pourquoi le Foscarnet est souvent un bon médicament de deuxième ligne lors de résistance au Ganciclovir (Mercorelli, Sinigalia et al. 2007). De plus, l'utilisation du Ganciclovir (et de ses homologues) est une lame à deux tranchants, car une bonne partie des mutations de résistance se retrouve dans le domaine exonucléase de la polymérase UL54, et certaines d'entre elles peuvent même abolir cette activité chez UL54. Cela augmente grandement l'infidélité de cette polymérase et peut accélérer l'apparition de résistance à de nouveaux traitements (Chou et Marousek 2007).

3.5) Problématique

Règle générale, les thérapies contre le CMV sont associées à plusieurs effets secondaires et à des contraintes de toxicités qui limitent la dose praticable. L'émergence de résistance au traitement complique d'autant plus le traitement des patients. C'est pourquoi, afin de traiter le cytomegalovirus humain, il y a toujours un grand besoin d'identifier de nouvelles cibles thérapeutiques et de développer de nouveaux composés antiviraux.

3.6) Avenues de recherches et limitations

Encore beaucoup d'efforts devront être investis en recherche avant d'atteindre cet objectif, mais voici quelques unes des percées les plus prometteuses. Certains groupes se concentrent sur l'amélioration d'inhibiteurs nucléosidiques, surtout ceux dérivés du

Cidofovir, comme le composé CMX001 de Chimerix, qui est présentement au stade de phase I (ChimerixInc. 2007). D'autres groupes se sont penchés sur les inhibiteurs non-nucléosidiques, ciblant plusieurs protéines essentielles. Viropharma (filiale de GlaxoSmithKline) mène présentement deux études cliniques de phase III avec leur composé maribavir (CamviaTM), qui est un inhibiteur de la protéine UL97 (ViropharmaInc. 2007). La dimérisation de UL44 étant nécessaire à ses fonctions (Appleton, Loregian et al. 2004), cette interface d'interaction représente une cible thérapeutique potentielle. Le même principe s'applique pour l'interaction UL44/UL54, qui est spécifique et essentielle (Pari et Anders 1993; Loregian, Rigatti et al. 2003). En effet, un peptide, correspondant aux 22 acides aminés du domaine C-terminal de UL54, est capable d'inhiber spécifiquement la liaison de UL44 avec UL54. De plus, le rôle crucial de quelques résidus précis lors de cette liaison (Loregian, Appleton et al. 2004; Loregian, Appleton et al. 2004) ouvre la porte à une inhibition par un petit composé organique. À cet égard, une étude de criblage haut débit sur 50 000 composés a permis l'identification de quelques molécules prometteuses (Loregian et Coen 2006). D'ailleurs, le cristal de UL44 lié à un peptide représentant le domaine C-terminal (a.a. 1221-1242) de UL54 a récemment été résolu, et ces informations pourraient, par l'entremise du dessin rationnel, contribuer à l'amélioration de ces composés (Appleton, Brooks et al. 2006). Enfin, une autre étude de criblage haut débit a permis la découverte d'une nouvelle classe de composés non-nucléosidiques inhibant la polymérase UL54 (Tucker, Clayton et al. 2000). Par exemple, l' Acyloxyanidine 1 arbore un IC₅₀ de 20µM, et aucune

inhibition de la polymérase α n'est détectable à des concentrations de 100 μ M (Tucker, Clayton et al. 2000).

Cependant, l'absence de cristaux et d'informations sur les propriétés biochimiques de UL54 rend l'amélioration et le dessin rationnel de nouvelles générations de composés plus difficiles.

4) Réplication de l'ADN

4.1) Cinétique de la réplication

La réplication de l'ADN génomique du Cytomégalovirus est complexe et n'est pas encore complètement élucidée. Voici un portrait de l'état actuel des connaissances. En moins de 4 heures suivant l'infection, le génome linéaire du CMV se circularise au noyau (LaFemina et Hayward 1983). De nombreuses preuves soutiennent que ce génome circulaire peut servir de matrice à une réplication en cercle roulant (LaFemina et Hayward 1983; Anders, Kacica et al. 1992; McVoy et Alder 1994), durant laquelle le complexe de réplication ne décroche pas de l'ADN et génère un long multimètre d'ADN, qui est par la suite clivé en monomères génomiques linéaires prêts à être encapsidés. C'est environ 16 heures après l'infection que la réplication de l'ADN virale débute, mais elle atteint son maximum entre 60 et 80 heures post-infection.

4.2) Protéines virales responsables de la réplication du génome

La synthèse de l'ADN viral implique un grand nombre de protéines virales et cellulaires et se localise dans des régions sub-nucléaires (Penfold et Mocarski 1997;

Uprichard et Knipe 1997). Le noyau central du large complexe de réplication est constitué de six protéines virales essentielles: UL44, UL54, UL57, UL70, UL102, UL105 (**Figure 6**) (Anders et McCue 1996). La protéine UL54 est la sous-

- UL54 : polymérase
- UL44 : Processivité
- UL70
- UL102
- UL105
- UL57 : lie ADNsb

— Brin mère
 — Brin fille

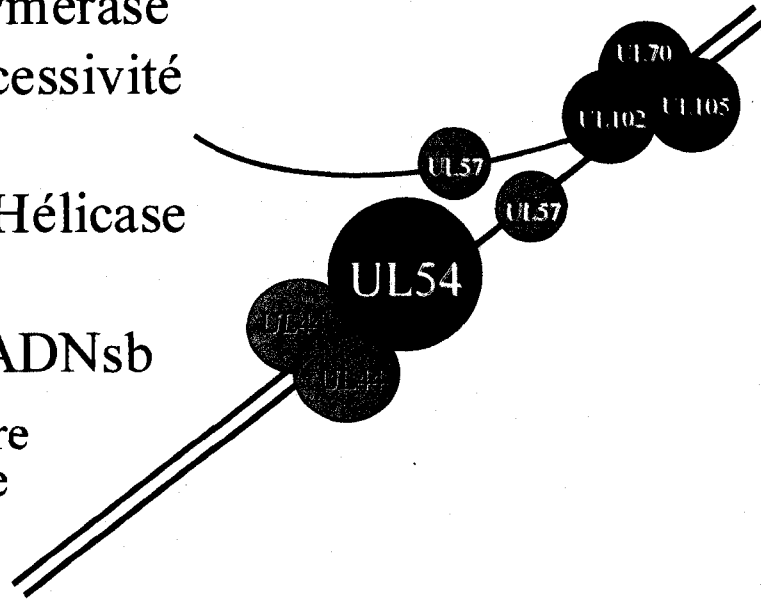


Figure 6 : Représentation schématique de l'organisation des six protéines essentielles composant le noyau central du large complexe de réplication du CMV. unité catalytique de l'ADN polymérase ADN dépendante (Pari et Anders 1993), et elle s'associe avec UL44, qui est son facteur de processivité (Pari, Kacica et al. 1993). De par son domaine de liaison à l'ADN double brin et sa capacité à dimériser autour de l'ADN, UL44 prévient la dissociation précoce de UL54 de son substrat (Ertl et Powell 1992; Wailand, Oien et al. 1994). Le déroulement de l'ADN, en aval de UL54, est assuré par l'hélicase/primase hétérotrimérique, composée de UL70, UL102 et UL105 (Martignetti et Barrell 1991; Pari et Anders 1993). Quant à UL57, c'est une protéine de liaison à l'ADN simple brin qui a pour rôle de prévenir le réappariement de l'ADN suite au passage de l'hélicase/primase (Kemble, McCormick et al. 1987;

Anders et Gibson 1988). À ce noyau de six protéines essentielles s'ajoutent plusieurs protéines nécessaires à une réplication optimale du cytomégalo virus, sans lesquelles on observe des retards de réplication et/ou une diminution, de un à plusieurs ordres de magnitude, de la dose de virus produite. Nommons notamment les protéines virales encodées par les cadres de lectures ouverts suivants : UL84, UL97, UL112-113 et UL114 (Anders et McCue 1996; Prichard, Duke et al. 1996; Sarisky et Hayward 1996; Prichard, Gao et al. 1999; Prichard, Lawlor et al. 2005; Parker, Kim et al. 2006). La protéine UL114 est une uracile ADN glycosylase qui s'associe avec UL44 et contribue à déstabiliser la double hélice d'ADN grâce à sa fonction de balayage (Prichard, Duke et al. 1996; Prichard, Lawlor et al. 2005). Le cadre de lecture ouvert UL112-113 produit, par épissage alternatif, quatre protéines qui contribuent au recrutement de UL44 aux loci sub-nucléaires où la réplication virale prend place (Parker, Kim et al. 2006). Le rôle de la protéine kinase UL97, bien que particulièrement important, n'est pas encore complètement élucidé. Elle a d'abord suscité un grand intérêt de par sa capacité à phosphoryler la prodrogue Ganciclovir (Littler, Stuart et al. 1992; Sullivan, Talarico et al. 1992). On sait qu'elle est aussi capable de s'autophosphoryler et de phosphoryler les histones (Baek, Krosky et al. 2002). Bien que davantage de preuves soient encore nécessaires, il est proposé qu'elle puisse participer ou réguler la réplication en liant et en phosphorylant UL44 (Krosky, Baek et al. 2003; Marschall, Freitag et al. 2003; Baek, Krosky et al. 2004). La reconnaissance spécifique du site d'initiation *oriLyt* est vraisemblablement médiée par la protéine UL84, qui est capable de lier la tige-boucle d'ARN hybridée au promoteur *oriLyt*. Elle pourrait initier la formation du complexe de réplication via ses

interactions directes avec IE2 (protéine immédiate précoce 2) et UL44, et ainsi amorcer la réplication *oriLyt* dépendante (Anders et McCue 1996; Prichard, Duke et al. 1996; Sarisky et Hayward 1996; Colletti, Smallenburg et al. 2007; Gao, Colletti et al. 2008).

Plusieurs protéines sont impliquées dans le complexe de réplication, mais là où les rôles généraux de chacune d'elles commencent à être bien acceptés, les propriétés spécifiques de chacune de ces protéines cruciales sont encore largement inconnues. Le meilleur exemple étant UL54; malgré son importance capitale lors de la réplication et le fait qu'elle soit la cible du faible nombre de thérapies dont nous disposons à ce jour, encore bien peu est connu sur cette ADN polymérase de la famille B.

5) ADN polymérase de la famille B

5.1) Généralités

D'un point de vue évolutionniste, les polynucléotides polymérases sont les éléments centraux du maintien et de la transmission du bagage génétique. De manière surprenante, il n'y a qu'une très faible homologie de séquence primaire entre les différentes polymérases. Cependant, la publication de leurs premières structures cristallisées a révélé de très grandes similitudes au niveau de leur structure tertiaire (Ollis, Brick et al. 1985; Kohlstaedt, Wang et al. 1992; Kim, Eom et al. 1995; Korolev, Nayal et al. 1995; Wang, Sattar et al. 1997). Aujourd'hui, une structure en forme de main droite fait consensus. Ce modèle représente les trois domaines d'une polymérase comme les doigts, la paume et le pouce d'une main droite (**Figure 7**)

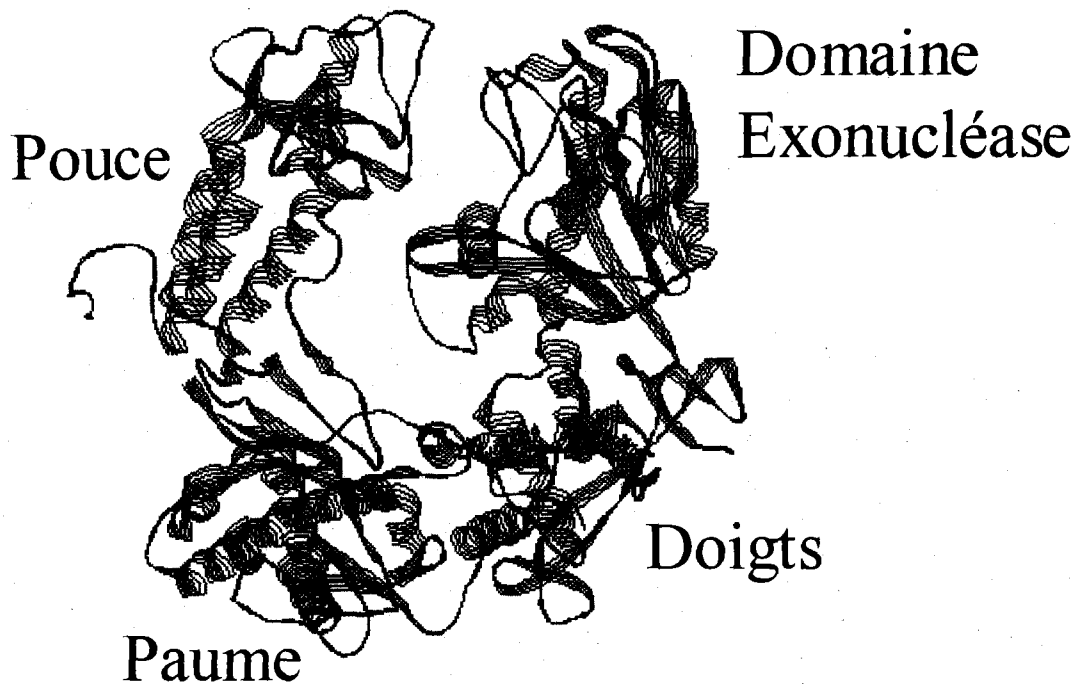


Figure 7 : Structure tridimensionnelle de la polynucléotide polymérase de la famille B. Plus précisément celle du bactériophage RB69, la protéine gp43. Générée à partir du cristal 2P5O (Hogg, Wallace et al. 2004). Les domaines de la paume, des doigts, du pouce, ainsi que le domaine exonucléase sont indiqués.

(Kohlstaedt, Wang et al. 1992). Les doigts interagissent avec la matrice d'ADN en aval et assure son bon positionnement; ils sont aussi responsables de la liaison du dNTP complémentaire arrivant. La paume recèle les résidus impliqués dans l'activité catalytique, ainsi que ceux responsables de la fidélité de la réplication. Quant au pouce, il fait contact avec le sillon mineur de l'ADN double brin naissant et assure ainsi son bon positionnement. De plus, il a un rôle important dans la processivité, en maintenant l'emprise que la polymérase exerce sur son substrat d'ADN (Brautigam et Steitz 1998; Steitz et Yin 2004).

Les ADN polymérases se subdivisent en quatre familles principales : polymérases I, α , β et les transcriptases inverses. Les polymérases α , aussi appelées famille B,

regroupent essentiellement des polymérase répliquatives comme la polymérase du bactériophage RB69, les polymérase eucaryotes α , δ et ϵ , ainsi que la polymérase virale du cytomégalo virus (Franklin, Wang et al. 2001; Tchessnokov, Gilbert et al. 2006).

5.2) Importance des ions divalents

Pour leur activité, les polymérase requièrent deux cations divalents, dont le rôle a été exhaustivement caractérisé. Ces deux ions sont séparés par 4 Å et sont coordonnés au cœur du site actif par deux acides aspartiques absolument conservés (Brautigam et

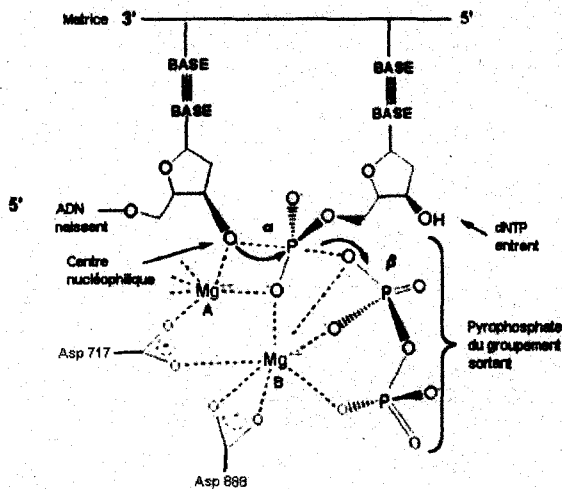


Figure 8 : Schéma du site actif d'une ADN polymérase. Le brin matrice (noir), servant de support génétique afin de discriminer le bon dNTP entrant (rouge), est sur le point d'être polymérisé sur l'ADN en cours de copie (bleu). À noter : le rôle crucial des deux ions magnésium (vert) lors de la catalyse. Ces cations sont coordonnés par deux acides aspartiques absolument conservés (magenta). (numérotation des asp suivant ceux de la polymérase de RB69) (Modifié de (Steitz 1998))

Steitz 1998). Le cation A contacte directement le 3'-OH de l'amorce, ainsi que l'oxygène pro-R du phosphate α du dNTP. Il active le groupement hydroxyle de l'amorce en abaissant son pK_a , ce qui lui permet d'entreprendre une attaque nucléophile

en ligne sur le phosphate α du nucléotide. L'ion B est en contact avec les oxygènes des trois phosphates du nucléotide et est responsable de stabiliser la charge négative induite sur

le phosphate β quittant (Figure 8) (Brautigam et Steitz 1998). De plus, les deux cations sont impliqués dans la stabilisation de l'état de transition trigonal bipyramidal généré par l'attaque nucléophile (Beese et Steitz 1991). Le magnésium, le manganèse

et, dans une moindre mesure, le cobalt, le nickel et le zinc, peuvent être utilisés *in vitro* par les polymérases (Tabor et Richardson 1989; Copeland, Lam et al. 1993; Garcia-Diaza, Bebeneka et al. 2007), mais *in vivo*, cette fonction serait vraisemblablement remplie par le magnésium (Benzaghou, Bougie et al. 2004).

5.3) Dynamique de la catalyse

Au cours d'une catalyse, une polymérase n'est pas rigide; au contraire, les nombreux cristaux de polymérases complexées avec leur ligand à différents stades de la réaction chimique ont révélé que l'enzyme subit une série de changements conformationnels lors d'un cycle catalytique (Brautigam et Steitz 1998; Franklin, Wang et al. 2001). Une apo-polymérase adopte généralement une conformation en forme de main droite entrouverte (conformation ouverte). Suite à la liaison du duplex d'ADN (amorce-matrice), le pouce se referme légèrement sur l'ADN (Franklin, Wang et al. 2001). Un dNTP peut alors se lier au complexe binaire, ce qui induit un changement de conformation au niveau des doigts. Ils se referment de 60° vers le site actif, déplaçant ainsi l'extrémité des doigts de 30 Å vers le pouce et la paume (Franklin, Wang et al. 2001). Cette rotation rapproche le dNTP du site actif et positionne correctement son phosphate α par rapport au 3'-OH de l'amorce et les deux ions divalents, permettant ainsi l'attaque nucléophile. C'est pourquoi ce n'est qu'en position fermée que la protéine est catalytiquement active (Franklin, Wang et al. 2001). Suite à l'ajout d'un nucléotide sur l'amorce (n+1), une molécule de pyrophosphate, correspondant au phosphate β et γ du dNTP, est générée. Chez les polymérases de la famille B, cette molécule est vraisemblablement expulsée vers le solvant via un tunnel formé à l'intersection du pouce, de la paume et des doigts (Capson, JAPeliska et al. 1992;

Franklin, Wang et al. 2001). Suite au départ du pyrophosphate, l'enzyme va transloquer vers l'aval; elle tourne ainsi légèrement autour de l'ADN afin de garder contact avec le sillon mineur (Steitz et Yin 2004). Sa réouverture clôt le cycle catalytique et permet à un nouveau dNTP de se lier et d'amorcer une nouvelle ronde de catalyse. Au cours de la polymérisation d'un brin d'ADN complet, la polymérase se déplace sur l'ADN, un peu comme un écrou sur un boulon.

5.4) Processivité

Les ADN polymérases ont ceci de particulier qu'elles peuvent demeurer associées à leur substrat d'ADN au cours de plusieurs cycles catalytiques consécutifs; par conséquent, elles sont processives. Cependant, les ADN polymérases répliquatives de la famille B requièrent l'aide d'un facteur de processivité afin de rester liées à l'ADN au cours de l'ajout de plusieurs milliers de désoxynucléotides (Brautigam et Steitz 1998). Les facteurs de processivité sont des protéines qui forment généralement des multimères capables de lier fortement l'ADN double brin. Un exemple bien connu est celui de PCNA, le facteur de processivité des polymérases eucaryotes γ et ϵ , qui forme un trimère en forme de beignet autour de l'ADN (Kong, Onrust et al. 1992; Gulbis, Kelman et al. 1996; Wang, Sattar et al. 1997; Appleton, Loregian et al. 2004). Le domaine du pouce participe aussi à la processivité en contribuant à maintenir l'emprise que la polymérase possède sur l'ADN (Brautigam et Steitz 1998), et en interagissant directement avec le facteur de processivité, généralement grâce à des résidus en N-terminal (Wang, Sattar et al. 1997; Appleton, Loregian et al. 2004; Steitz et Yin 2004).

5.5) Fidélité

À chaque nouveau cycle catalytique, le substrat d'une ADN polymérase peut changer. Elle réussit toutefois à incorporer précisément le nucléotide complémentaire à la matrice, étant ainsi fidèle. De par leur fonction répliquative, les ADN polymérases de la famille B ont évolué vers une très grande fidélité. Par exemple, l'ADN polymérase du bactériophage RB69 fait 2×10^{-8} erreurs par paire de bases copiée (Bebenek, Carvera et al. 2002). Cette précision, bien plus grande que la simple différence d'énergie impliquée lors du pairage, est due à de nombreux points de contrôle lors de la polymérisation. D'abord, un mauvais nucléotide a une moins grande affinité pour l'enzyme. Ensuite, si un mauvais nucléotide lie l'enzyme, le changement de conformation sera plus difficile à effectuer, dû à l'encombrement stérique qui subsiste au sein de cet étroit site actif (Steitz 1998; Joyce et Benkovic 2004). Le positionnement d'un mauvais nucléotide est suborbital pour la catalyse (Joyce et Benkovic 2004). Dernièrement, l'incorporation covalente d'un mauvais nucléotide affaiblit l'interaction de l'ADN avec l'enzyme, ce qui ralentit la polymérase et augmente les chances que le brin naissant transloque vers le domaine exonucléase, où le mauvais nucléotide peut être excisé (Brautigam et Steitz 1998; Doublé, Tabor et al. 1998; Franklin, Wang et al. 2001; Steitz et Yin 2004). Le domaine exonucléase, situé au niveau de la paume, est très important pour la fidélité, et il est présent chez la majorité des ADN polymérases de la famille B.

5.6) Conservation

Les polymérases de la famille B sont hautement conservées, surtout au niveau de leur structure tertiaire. En terme de séquence primaire, les domaines des doigts et du

pouce sont les plus variables, et le domaine de la paume, qui abrite le centre catalytique, est le plus conservé. Bien que les polymérase de cette famille aient des tailles variant entre 553 et 1504 acides aminés (Morrison, Christensen et al. 1989; Jung, Leavitt et al. 1990), l'alignement de leurs séquences a permis d'identifier six régions hautement conservées numérotées I-VI, en ordre décroissant de similarité (**Figure 9**) (Wong, Wahl et al. 1988; Ito et Braithwaite 1991; Wang, Sattar et al. 1997; Brautigam et Steitz 1998). C'est au sein de ces régions que l'on retrouve les motifs DxxSLYPS et YGDTDS, qui sont les deux motifs les plus hautement conservés chez les polymérase de la famille B. Ces derniers abritent les deux acides aspartiques obligatoirement conservés (Asp411 et Asp623 chez la Pol α),

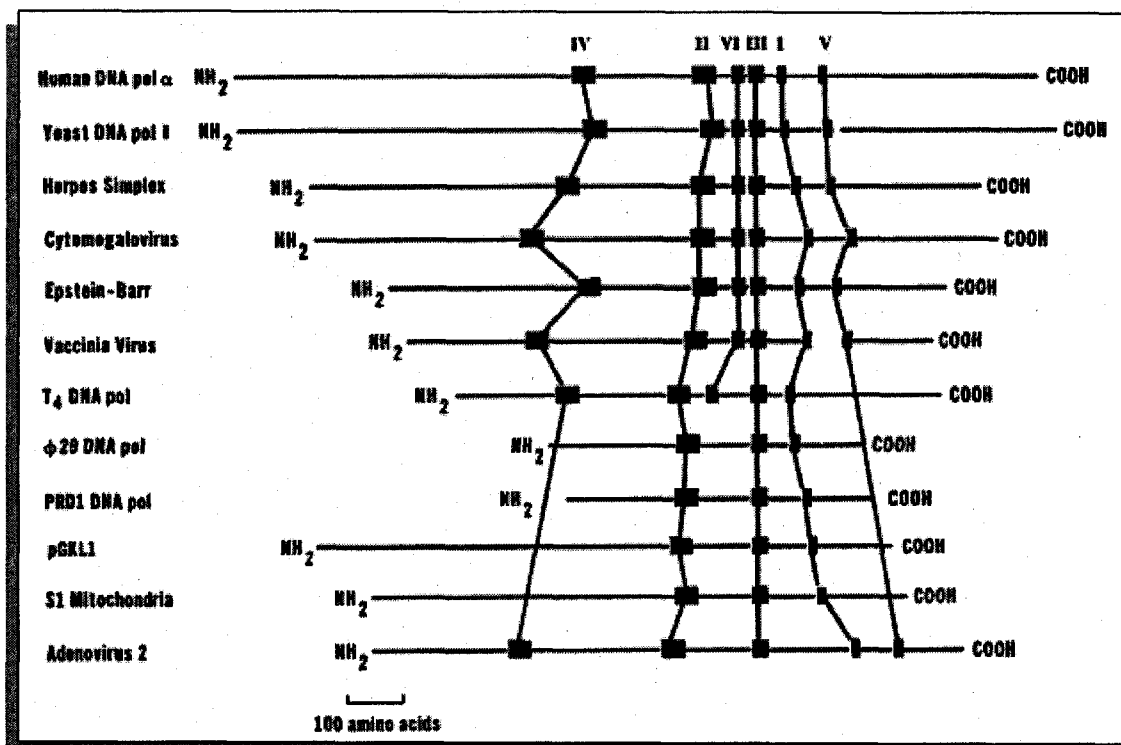


Figure 9 : L'alignement de plusieurs ADN polymérase de la famille α a révélé la présence de six domaines hautement conservés I- VI (boite noire). (Wang, Wong et al. 1989)

et impliqués, de concert avec le triphosphate du dNTP, dans la coordination des deux ions métalliques divalents (Delarue, Poch et al. 1990; Franklin, Wang et al. 2001).

6) ADN polymérase UL54

Le cadre de lecture ouvert 54 de la longue région unique (UL54) du cytomegalovirus humain code pour l'ADN polymérase virale UL54 (Heilbronn, Jahn et al. 1987; Kouzarides, Bankier et al. 1987). Elle fait partie des polymérases de la famille B (pol- α), avec lesquelles elle partage plusieurs caractéristiques, dont les six domaines hautement conservés (I-VI) (Hirai et Watanabe 1976; Ito et Braithwaite 1991). Cette enzyme de 1242 acides aminés requiert aussi l'aide d'un cofacteur de processivité (UL44), qui lui permet de demeurer lié à l'ADN au cours de l'ajout de plusieurs centaines de milliers de désoxynucléotides (Davison, Dolan et al. 2003). Tout comme les polymérases de la Famille B, UL54 est une polymérase répliquative responsable de la copie du génome de 230 kilobases (Davison, Dolan et al. 2003).

En comparaison avec d'autres polymérases virales, les études visant la caractérisation de UL54 ont été très limitées. La majorité des efforts ont été investis lors d'études cliniques sur des souches de virus présentant des résistances à un ou plusieurs médicaments. Ainsi, beaucoup de résidus spécifiquement impliqués dans la liaison ou dans la résistance à certaines drogues contre le cytomegalovirus ont été identifiés (Smith, Cherrington et al. 1997; Cihlar, Fuller et al. 1998; Weinberg, Jabs et al. 2003; Gilbert et Boivin 2005; Tchesnokov, Gilbert et al. 2006). En contrepartie, la majorité des études fondamentales précédentes se sont restreintes à l'utilisation de protéines

issues de synthèse *in vitro* ou de cellules d'insectes infectées avec un baculovirus. Conséquemment, les informations disponibles sur les propriétés biochimiques de UL54 se limitent essentiellement à ses paramètres cinétiques (K_m , K_{cat}) (Crumpacker 1992; Tchesnokov, Gilbert et al. 2006). Il n'existe aucune structure de publiée (cristaux ou RMN) de cette protéine, sa fidélité n'est pas connue, et ses propriétés biochimiques ainsi que la caractérisation de ses interactions avec ses différents ligands restent encore à être élucidées.

7) Forces intermoléculaires

Les macromolécules biologiques comme les protéines et les acides nucléiques interagissent constamment ensemble; c'est la base même de la vie. Ces interactions sont gouvernées par la sommation de plusieurs forces physiques, qui tiennent compte de la nature des macromolécules impliquées. Les liaisons covalentes sont rares au sein des interactions intermoléculaires; elles sont si fortes et stables qu'on les retrouve plutôt au sein des interactions intramoléculaires. Les forces régissant les interactions entre les molécules sont plus faibles et plus nombreuses : électrostatiques (ioniques, dipôles-dipôles et ponts hydrogènes), de Van der Waals, d'empilement aromatique et hydrophobes. Individuellement, ces interactions sont de un à trois ordres de magnitude plus faible qu'un lien covalent, mais dans certaines situations, la sommation de multiples interactions non covalentes additives sur une surface limitée peut être aussi forte qu'une interaction covalente (Leckband et Israelachvili 2001).

7.1) Interactions ioniques

Les interactions ioniques ou charges charges sont relativement fortes (6.3-19 zeptojoules (zJ)), et elles sont attractives lorsque deux charges de signes opposés se rapprochent, alors qu'elles sont répulsives dans le cas de deux charges de même signe (Leckband et Israelachvili 2001). Plusieurs acides aminés chargés (his, lys, arg, glu et asp) peuvent être impliqués dans des interactions ioniques.

7.2) Forces dipôles-dipôles

Les forces dipôles-dipôles fonctionnent comme celles ioniques, à la différence qu'elles impliquent des charges partielles au lieu de charges nettes. Elles sont donc légèrement moins fortes et surviennent entre deux molécules polarisées (Leckband et Israelachvili 2001).

7.3) Ponts hydrogènes

Il y a une interaction dipôle-dipôle particulière et très répandue qui implique un hydrogène : c'est le pont hydrogène. Celui-ci est directionnel et se forme entre un groupement donneur de proton, généralement un hydrogène lié à un groupement fortement électronégatif (l'oxygène ou l'azote, par exemple), et un groupement accepteur, généralement constitué d'un atome électronégatif possédant une densité électronique riche (doublet libre) (Leckband et Israelachvili 2001). Les ponts hydrogènes peuvent être forts (7-50zJ) et sont largement impliqués dans la formation de structures intra et intermoléculaires, comme la double hélice d'ADN, ainsi que les hélices- α et les feuillets- β des protéines.

7.4) Forces de van der Waals

Les forces de van der Waals sont toujours présentes, et elles sont le fruit d'une polarisation momentanée du nuage électronique d'une molécule, ce qui induit une polarisation inverse et temporaire chez une deuxième molécule voisine, générant ainsi une attraction entre deux dipôles induits. Cette force est généralement plus faible (4-8zJ pour de petites molécules) que les interactions dipôles-dipôles, car elle implique des dipôles temporaires, et non des molécules ayant un dipôle permanent. Il est cependant à noter que les forces de van der Waal sont plus importantes chez les grosses molécules, où elles peuvent devenir plus significatives (Leckband et Israelachvili 2001).

Lorsque deux surfaces sont à une faible distance, les forces de van der Waals sont attractives, mais elles deviennent fortement répulsives à des distances encore plus faibles, dû aux chevauchement des orbitales électroniques des atomes; c'est ce qu'on appelle communément l'encombrement stérique.

7.5) Empilement aromatique

L'empilement aromatique se produit lorsque deux cycles aromatiques se rapprochent dans un même plan. Il s'en suit une forte attraction (40-50zJ) entre les deux cycles, due à la conjugaison du système (Delaage 1991). Cette force contribue à la stabilité des acides nucléiques, au moins autant que les ponts hydrogènes entre les bases (Delaage 1991).

7.6) Forces hydrophobes

Les forces hydrophobes apparaissent en condition physiologique où le solvant est l'eau. Elles sont entièrement dues aux changements d'entropie du solvant; c'est en fait l'eau qui force les groupements hydrophobes à se regrouper, afin de minimiser leur impact destructeur sur son réseau de ponts hydrogènes. C'est une force importante représentant environ $17zJ/nm^2$ (Chothra 1984; Chandler 2005).

7.7) Contribution des forces intermoléculaires

Les interactions protéine-protéine et la liaison d'une protéine à son ligand sont le résultat d'une combinaison de toutes ces forces. L'hydrophobicité est généralement le facteur majeur stabilisant une liaison, mais elle est à toute fin pratique non spécifique. À l'inverse, la formation d'interactions électrostatiques et de contacts de van der Waals requiert des surfaces complémentaires. Ces forces apportent souvent une contribution moindre à la stabilisation d'une liaison, mais leur pouvoir de discrétion permet d'en accroître la spécificité (Chothia et Janin 1975).

Un bon exemple serait celui des polymérase α , qui contactent l'ADN via plusieurs résidus évolutivement non-conservés, ce qui indique que les interactions nécessaires à la liaison sont importantes dans leur ensemble, et non individuellement. Dans cet exemple, la protéine reconnaît surtout le phosphosquelette de l'ADN, majoritairement grâce à des ponts hydrogènes et des interactions ioniques (Franklin, Wang et al. 2001).

8) Objectifs

La protéine UL54 demeurant la cible thérapeutique principale chez le CMV, nous avons entrepris une étude détaillée de la liaison de la polymérase du cytomégalovirus avec ses ligands naturels. Nos objectifs principaux sont de produire une grande quantité de protéine UL54 recombinante pure. Nous voulons ensuite caractériser d'un point de vue cinétique, thermodynamique et structurale la liaison de UL54 à ses deux substrats principaux que sont l'ADN et les d'NTP.

**Article 1 : Caractérisation de la liaison de l'ADN et
des dNTP à l'ADN polymérase du CMV.**

**Titre : Characterization of the DNA- and dNTP-binding
activities of the human cytomegalovirus DNA polymerase
catalytic subunit UL54**

Frédéric Picard-Jean, Isabelle Bougie et Martin Bisailon

Biochemical journal (2007) **407**, 331-41

(Le format manuscrit est disponible en annexe)

Implication dans le travail

Ma contribution au travail présenté dans cet article est d'environ 70%,
puisque j'ai produit 15 des 22 figures présentées. Plus précisément, j'ai
été en charge de l'expression de la protéine et d'un des tests d'activité
préliminaire. J'ai aussi été responsable de la caractérisation de la liaison
de UL54 à l'ADN. C'est aussi moi qui ai réalisé les essais de dichroïsme
circulaire ainsi que ceux avec le foscarnet. J'ai aussi participé à la
rédaction du manuscrit.

SOMMAIRE

La sous-unité catalytique de l'ADN polymérase du Cytomégalovirus est cruciale pour la réplication du virus. Les présents travaux exposent l'expression et la purification d'une sous-unité catalytiquement active de l'ADN polymérase du Cytomégalovirus exprimée en bactérie. Comme premier pas en vue d'élucider la nature des interactions entre l'enzyme, l'ADN et les désoxynucléotides (dNTPs), nous avons tiré partie de la fluorescence intrinsèque du tryptophane afin d'évaluer la liaison des ligands à l'enzyme. Cette technique nous a permis de déterminer que le site minimal de liaison à l'enzyme est de 6 nucléotides. Nous présentons aussi la première étude détaillée sur les paramètres thermodynamiques et sur la cinétique de la liaison de l'enzyme à l'ADN et aux dNTPs. Nos analyses thermodynamiques indiquent que la formation du complexe enzyme-ADN est propulsée par une entropie favorable, mais qu'elle est aussi associée à une contribution enthalpique défavorable. À l'inverse, une enthalpie favorable et une entropie défavorable régissent à la liaison des dNTPs avec le complexe enzyme-ADN, ce qui révèle un mode de liaison complètement différent. Afin d'obtenir de plus amples informations sur les modifications structurales qui peuvent survenir lors de la catalyse, nous avons corrélié l'effet de la liaison de l'ADN et des dNTPs sur la structure de la protéine grâce au dichroïsme circulaire. Nos résultats indiquent que l'enzyme subit un premier changement conformationnel lors de la liaison à l'ADN, après quoi, la liaison d'un désoxynucléotide provoque un second changement de structure. Notre étude fournit une compréhension accrue des fondements moléculaires de la reconnaissance de l'ADN et des dNTPs par la sous-unité catalytique de l'ADN polymérase du Cytomégalovirus.

SUMMARY

The catalytic subunit of the human cytomegalovirus DNA polymerase is critical for the replication of the virus. In the present study, we report the expression and purification of a recombinant catalytic subunit of the human cytomegalovirus DNA polymerase expressed in bacteria which retains polymerase activity. As a first step toward elucidating the nature of the interaction between the enzyme, DNA, and deoxynucleotides (dNTPs), we have utilized endogenous tryptophan fluorescence to evaluate the binding of ligands to the enzyme. Using this technique, we demonstrate that the minimal DNA binding site of the enzyme is 6 nucleotides. We also report the first detailed study of the binding kinetics and thermodynamic parameters involved in the interaction between the enzyme, DNA, and dNTPs. Our thermodynamic analyses indicate that the initial formation of the enzyme-DNA binary complex is driven by a favorable entropy change, but is also clearly associated with an unfavorable enthalpic contribution. In contrast, the interaction of dNTPs to the binary complex was shown to depend on a completely different mode of binding that is dominated by a favorable enthalpy change and associated with an unfavorable entropy change. In order to provide additional insights into the structural modifications that occur during catalysis, we correlated the effect of DNA and dNTP binding on protein structure using circular dichroism. Our results indicate that the enzyme undergoes a first conformational change upon the formation of the protein-DNA binary complex which

is followed by a second structural modification upon dNTP binding. Our study provides a better understanding of the molecular basis of DNA and dNTP recognition by the catalytic subunit of the human cytomegalovirus DNA polymerase.

Keywords: Cytomegalovirus, Deoxynucleotides, DNA, Binding kinetics, Thermodynamics, Spectroscopy.

Abbreviations used: ANS, 1-anilino-8-naphthalenesulfonate; CD, circular dichroism; dNTPs, deoxynucleoside triphosphates.

INTRODUCTION

The human cytomegalovirus is a DNA virus that can cause a variety of important diseases in immunocompromised patients and in newborns [1]. These include gastrointestinal diseases, pneumonia, and retinitis, while neonates born with cytomegalic inclusion disease can have symptoms such as microcephaly, hearing loss, liver damage, pneumonitis, and mental retardation [2]. Although the incidence of cytomegalovirus infection has significantly been reduced in recent years due to the use of antiretroviral drugs, cytomegalovirus infections remain difficult to manage [3-6].

Human cytomegalovirus is a member of the *Herpesviridae* family which includes an important number of human pathogens [1]. The DNA polymerase activity of the human cytomegalovirus has been shown to require two distinct polypeptides, a property shared with other viruses from the *Herpesviridae* family. The larger subunit, which is responsible for the catalytic DNA polymerase activity, is encoded by the cytomegalovirus UL54 gene [7-9] which harbors the conserved domains (I-VII) that are also found in other type-B DNA polymerases. The smaller accessory subunit is encoded by the UL44 gene [10] and has been shown to stimulate the polymerase activity of the catalytic subunit by increasing its processivity [10,11]. Moreover, the critical importance of this protein-protein interaction has recently been demonstrated in a recent study using short peptides that disrupt the physical interaction between the two subunits [12]. The presence of these peptides results in the successful inhibition of the polymerase activity, clearly highlighting the importance of the UL54/UL44 interaction. Using an *in vitro* translation system, a previous mutagenesis analysis has also identified a short region located in motif III (amino acids 804 to 807) of the enzyme that is important for the interaction between the enzyme and DNA [9].

As a critical component of the viral replication machinery, the cytomegalovirus DNA polymerase is the specific target for a number of antiviral drugs currently used to inhibit the replication of the virus [13]. These include ganciclovir, cidofovir, and foscarnet that are currently approved for the treatment of cytomegalovirus infections in immunocompromised patients. Both ganciclovir and

cidofovir are nucleoside analogs that act as competitive inhibitors of the DNA polymerase and function as DNA chain terminators [14,15] while foscarnet is an analog of pyrophosphate that acts noncompetitively with respect to the deoxynucleotides [16]. The precise mechanism of action of this drug remains largely speculative although it has been proposed that the pyrophosphate analog could bind close to the active center of the enzyme thereby inhibiting the pyrophosphate exchange reaction [17]. While the use of ganciclovir, cidofovir, and foscarnet can reduce the viral burden, the antiviral therapy has been plagued by the emergence of drug-resistant viral strains [3-6].

By comparison with other viral DNA polymerases, characterization studies of the cytomegalovirus enzyme have been very limited. Based on clinical studies of drug-resistant isolates, most studies have focused on the identification of specific residues involved in the binding to the anti-cytomegalovirus drugs [4,6,19-22]. Moreover, biochemical studies have been limited to the identification of the kinetic parameters (K_M , k_{cat}) of the polymerase reaction [13,22]. This lack of biochemical information on the enzymatic activity of the protein stems from the fact that previous studies have mostly used proteins synthesized by *in vitro* translation systems or insect cells infected with baculoviruses. Our understanding of the DNA synthesis by the cytomegalovirus DNA polymerase remains largely incomplete. One area that requires additional characterization concerns how the enzyme interacts with DNA and dNTPs. In the present study, we report the expression and purification of a recombinant catalytic subunit (UL54) of the human cytomegalovirus protein expressed in bacteria

which retains polymerase activity. Using fluorescence spectroscopy and circular dichroism studies, we describe a detailed thermodynamic study on the formation of the enzyme-DNA binary complex which provides insights into the relative enthalpic and entropic contributions to the binding energy, as well as on the effects of DNA binding on both the structure and stability of the protein. Moreover, we have also characterized the binding of DNA molecules of various lengths and various sequences to determine both the minimal length for stable binding and the specificity of the DNA-enzyme interaction. Finally, we also monitored the binding of dNTPs to the enzyme-DNA complex, and completed a structural analysis of the protein in order to gain insights into the structural modifications that occur during catalysis. Such analyses are crucial in order to understand the molecular basis of the interaction between the catalytic subunit of the cytomegalovirus DNA polymerase and its ligands.

MATERIALS AND METHODS

Protein expression and purification

A plasmid for the expression of the catalytic subunit of the cytomegalovirus DNA polymerase was generated by inserting a fragment of the cytomegalovirus UL54 gene

in the pET-28a expression plasmid (Novagen). This fragment encodes for amino acids 293 to 953 of the protein, and was cloned between the *Nde*I and *Hind*III cloning sites of the pET-28a plasmid. In this context, the protein is fused in frame with an N-terminal peptide containing 6 tandem histidine residues, and expression of the His-tagged protein is driven by a T7 RNA polymerase promoter. The resulting recombinant plasmid (pET-UL54) was transformed into *Escherichia coli* BL21(DE3) and a 1000-ml culture of *E. coli* BL21(DE3)/pET-UL54 was grown at 37°C in Luria-Bertani medium containing 30 µg/ml kanamycin until the A_{600} reached 0.5. The culture was adjusted to 0.4 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and 2% ethanol, and the incubation continued at 37°C for 3 h. The cells were then harvested by centrifugation, and the pellet stored at -80°C. All subsequent procedures were performed at 4°C. Thawed bacteria pellets were resuspended in 50 ml of lysis buffer A (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, and 10% sucrose) and cell lysis achieved by the addition of lysozyme and Triton X-100 to final concentrations of 50 µg/ml and 0.1%, respectively. The lysates were sonicated to reduce viscosity, and any insoluble material removed by centrifugation at 13,000 rpm for 45 min. The soluble extract was applied to a 5-ml column of Ni-nitrilotriacetic acid-agarose (Qiagen) that had been equilibrated with buffer A containing 0.1% Triton X-100. The column was washed with the same buffer, and then eluted stepwise with buffer B (50 mM Tris-HCl, pH 8.0, 0.1 M NaCl, and 10% glycerol) containing 50, 100, 200, 500, and 1000 mM imidazole. The polypeptide composition of the column fractions was monitored by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE). The recombinant protein was retained on the column and recovered in the 200 and

500 mM imidazole fractions. Following dialysis against buffer C, buffer C (50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 2 mM dithiothreitol, 10% glycerol, and 0.05% Triton X-100) the purified preparation was stored at -80°C. Protein concentration was determined by the Bio-Rad dye binding method using bovine serum albumin as the standard.

Polymerization assay

A DNA substrate (P20/T60) consisting of heteropolymeric oligonucleotides was used to monitor the polymerase activity of the enzyme. The substrate consisted of a primer (5'-TAATACGACTCACTATAGGG-3') and a template (5'-TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCCGACCCGCCCTATAGTGAGTCGTA-3'). Equimolar concentrations of the oligonucleotides were heated at 95 °C for 5 minutes and cooled at room temperature. Standard reaction mixtures (50 µl) containing 25 mM Tris-HCl, pH 8.0, 10 mM MgCl₂, 50 mM NaCl, 0.2 µM [α -³²P]ATP, 500 mM of each dNTP, 5 µM of DNA template, and purified DNA polymerase (0.6 µM) were incubated for 30 min at 37 °C. The reaction products were analyzed by electrophoresis through a 10% polyacrylamide gel containing 8 M urea. The radiolabeled DNA products were located by autoradiography.

Fluorescence measurements

Fluorescence was measured using a Hitachi F-2500 fluorescence spectrophotometer. Excitation was performed at a wavelength of 290 nm. Background emission was eliminated by subtracting the signal from either buffer alone or buffer containing the appropriate quantity of substrate.

The extent to which ligands (DNA or dNTPs) bind to the enzyme was determined by monitoring the fluorescence emission of a fixed concentration of proteins and titrating with a given ligand. The binding can be described by Equation 1,

$$K_D = \frac{[\text{UL54}] [\text{ligand}]}{[\text{UL54} \cdot \text{ligand}]}$$

where K_D is the apparent dissociation constant, $[\text{UL54}]$ is the concentration of the protein, $[\text{UL54} \cdot \text{ligand}]$ is the concentration of complexed protein, and $[\text{ligand}]$ is the concentration of unbound ligand.

The proportion of ligand-bound protein as described by Equation 1 is related to measured fluorescence emission intensity by Equation 2,

$$\Delta F / \Delta F_{max} = [\text{UL54} \cdot \text{ligand}] / [\text{UL54}]_{tot}$$

where ΔF is the magnitude of the difference between the observed fluorescence intensity at a given concentration of ligand and the fluorescence intensity in the

absence of ligand, ΔF_{max} is the difference at infinite [ligand], and $[\text{UL54}]_{\text{tot}}$ is the total protein concentration.

If the total ligand concentration, $[\text{ligand}]_{\text{tot}}$, is in large molar excess relative to $[\text{UL54}]_{\text{tot}}$, then it can be assumed that [ligand] is approximately equal to $[\text{ligand}]_{\text{tot}}$. Equations 1 and 2 can then be combined to give Equation 3,

$$\Delta F/\Delta F_{max} = [\text{ligand}]_{\text{tot}} / (K_D + [\text{ligand}]_{\text{tot}})$$

The K_D values were determined from a nonlinear least square regression analysis of titration data by using Equation 3.

Circular dichroism spectroscopy measurements

Circular dichroism (CD) measurements were performed with a Jasco J-810 spectropolarimeter. The samples were analyzed in quartz cells with pathlengths of 1 mm. Far-UV and near-UV wavelength scans were recorded from 200 to 250 nm and from 250 to 340 nm, respectively. All the CD spectra were corrected by subtraction of the background for the spectrum obtained with either buffer alone or buffer containing the ligand. The ellipticity results were expressed as mean residue ellipticity, $[\theta]$, in degrees·cm²·dmol⁻¹.

Equilibrium unfolding experiments

Thermal transitions were monitored by following the change in CD ellipticity of the protein (1 μM) at 222 nm. The samples were heated from 10 $^{\circ}\text{C}$ to 95 $^{\circ}\text{C}$, at a heating rate of 1 $^{\circ}\text{C}/\text{min}$. The ellipticity results were expressed as mean residue ellipticity, $[\theta]$, in $\text{degrees}\cdot\text{cm}^2\cdot\text{dmol}^{-1}$. The fraction of unfolded protein at each temperature was determined by calculating the ratio $[\theta_{222}]/[\theta_{222}]_d$, where $[\theta_{222}]_d$ is the molar ellipticity for the completely unfolded enzyme.

ANS binding measurements

Binding of ANS (1-anilino-8-naphtalenesulfonate) was evaluated by measuring the fluorescence enhancement of ANS (50 μM) upon excitation at a wavelength of 380 nm. The emission spectra were integrated from 400 to 600 nm.

RESULTS

Expression, purification and DNA polymerase activity of the protein

In order to investigate the initial binding of deoxynucleotides to the catalytic subunit of the human cytomegalovirus DNA polymerase (UL54), we intended to purify a large quantity of the enzyme. Previous biochemical studies performed with the enzyme have either used a coupled *in vitro* transcription/translation system [22,23] or insect cells infected with recombinant baculoviruses [18,24] to generate the full-length protein (1242 amino acids). Since a limited amount of protein is obtained from these systems, an expression plasmid was created to express the recombinant protein in *E. coli*. Because of the large size of the enzyme, we focused on the region encompassing amino acids 293 to 953 of the protein. This domain contains the consensus sequences of the DNA polymerase type-B family [25] as well as most of the residues [16, 17] that have been associated with drug resistance in UL54 (Fig. 1). The truncated enzyme was therefore expressed in *E. coli* with an N-terminal His-tagged epitope, and purified from soluble bacterial extracts by nickel-agarose chromatography. The purified protein was located mainly in the 200 and 500 mM imidazole eluates (Fig. 2A). All subsequent experiments were performed with the protein found in the 500 mM fraction, as it displayed the highest level of purity. The concentration of protein in this fraction was estimated to be 130 ng/mL. Immunoblotting analysis using a monospecific anti-His antibody also confirmed the identity of the protein (Fig. 2B).

The ability of the purified enzyme to catalyze the synthesis of DNA was initially analyzed in a gel-based assay using a defined heteropolymeric primer/template (P20/T60) substrate. As shown in figure 2C, the purified enzyme had the ability to perform multiple deoxynucleotide incorporation events. However, as reported previously with the native enzyme [22], the recombinant protein was not highly processive as evidenced by the numerous shorter elongation products that were also synthesized. Additionally, we also used a coupled in vitro transcription-translation system to generate the full-length protein and compare its activity with the recombinant truncated enzyme. As shown in figure 2C, both proteins displayed similar polymerization patterns upon incubation with dNTPs.

A second polymerization assay was used to quantify the specific activity of the recombinant enzyme. In this assay, the purified protein was added to the same primer/template combination, but the incorporation of radiolabeled dATP was monitored by liquid scintillation counting following the precipitation of the DNA reaction products. As can be seen in figure 2D, the extent of dATP incorporation during a 30 min reaction was proportional to the amount of input UL54 protein. A specific activity of 120 nmoles/min/ μ g was evaluated from this experiment. Foscarnet, an antiviral used for the treatment of human cytomegalovirus infections and a potent inhibitor of the cytomegalovirus DNA polymerase activity [3,13,19], was shown to inhibit the polymerization activity of the purified protein. As shown in

figure 2E, the addition of 2.7 μM of foscarnet resulted in 50 % of inhibition of the polymerase reaction.

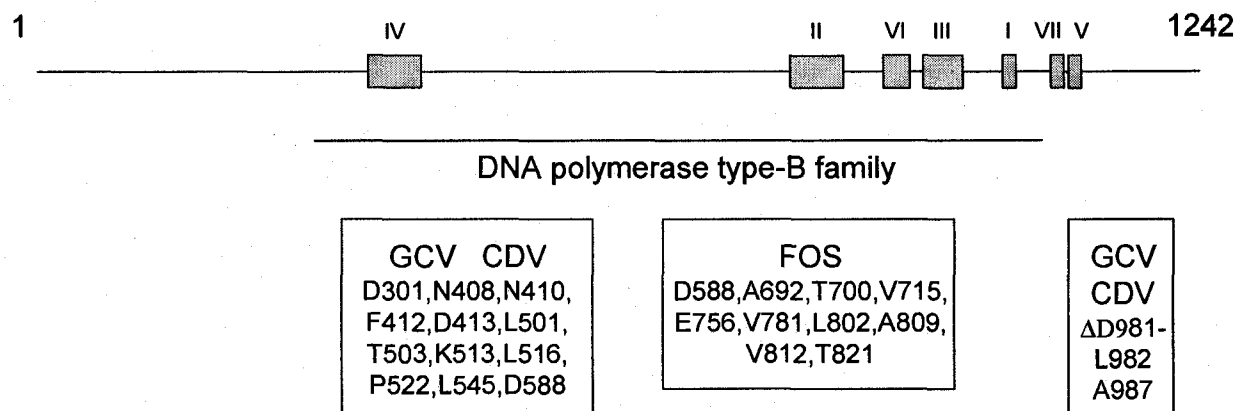


Figure 1 Functional domains of the cytomegalovirus DNA polymerase

The conserved domains (I-VII) found in other type-B DNA polymerases are shown. These domains, as initially determined by Wang *et al.*, have been defined as follows according to the amino acid numbering of the enzyme: IV (379-421), II (696-742), VI (771-790), III (805-845), I (905-919), VII (962-970), and V (978-988). The consensus region of the DNA polymerase type-B family is also indicated (293-953). Additionally, the regions that have been shown to be critical for drug resistance are also shown along with the specific amino acids that have been shown to be associated with drug resistance.

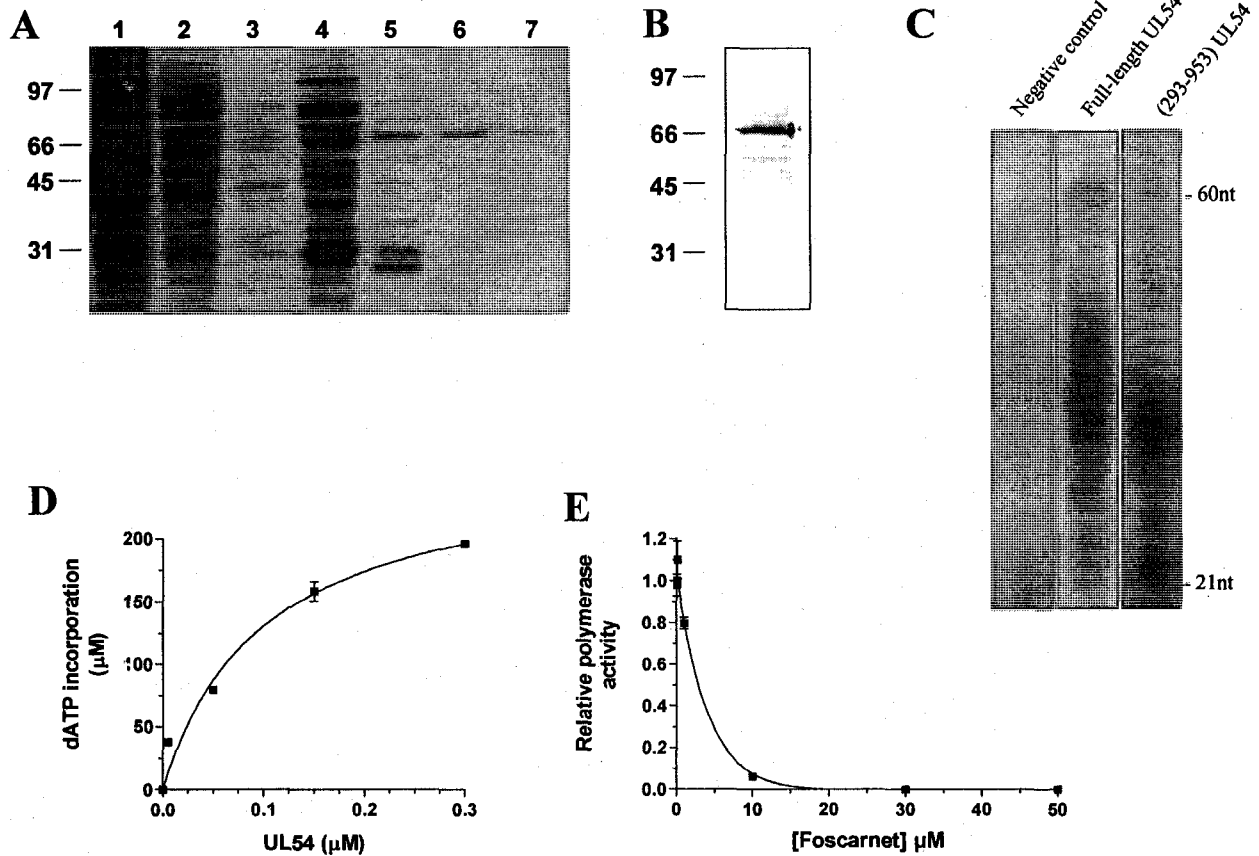


Figure 2 Purification and enzymatic activity of the purified cytomegalovirus DNA polymerase

(A) The peptide compositions of various purification fractions were analyzed by SDS-PAGE. The gels were fixed and stained with Coomassie blue dye. Lane 1, nickel-agarose flow-through; lane 2, nickel-agarose wash; lane 3, 50 mM imidazole eluate; lane 4, 100 mM imidazole eluate; lane 5, 200 mM imidazole eluate; lane 6, 500 mM imidazole eluate; lane 7, 1000 mM imidazole eluate. The positions and sizes (in kDa) of the size markers are indicated on the left. (B) The purified cytomegalovirus protein (500 mM imidazole fraction) was analyzed by

immunoblotting using a monospecific anti-His antibody. **(C)** Polymerization assays performed with a heteropolymeric primer/template DNA substrate (P20/T60). The reaction mixtures (50 μ l) containing 20 mM Tris-HCl, pH 7.5, 10 mM MgCl₂, 50 mM NaCl, 5 μ M of primer/template DNA substrate, 0.2 mM [α -³²P]ATP, 500 mM of dGTP, dCTP, and dTTP were incubated for 30 min at 37 °C in the presence of 0.6 μ M of either the full-length protein expressed in a coupled transcription/translation system (lane 2) or recombinant enzyme (lane 3). A negative control (vector alone) synthesized with the coupled transcription/translation system was also used in the assay (lane 1). The reaction products were analyzed by electrophoresis through a 10% polyacrylamide gel containing 8 M urea. An autoradiogram of the gel is shown, and the positions of the size markers are also indicated. **(D)** Increasing amounts of enzyme were added to the heteropolymeric primer/template DNA substrate (P20/T60). The reaction mixtures containing 5 μ M of primer/template DNA substrate, 0.6 μ M of enzyme, 1 mM [α -³²P]ATP, 1 mM of dGTP, dCTP, and dTTP were incubated for 30 min at 37 °C. The reaction products were analyzed by liquid scintillation counting following the precipitation of the DNA reaction products. **(E)** Polymerization assays were performed in the presence of increasing concentrations of foscarnet, and the incorporation of radiolabeled ATP was quantified by liquid scintillation counting.

Intrinsic fluorescence properties of the cytomegalovirus DNA polymerase

We and others have previously shown that the binding of DNA to free enzymes can result in a significant decrease in protein emission fluorescence [26-32]. We therefore

intended to use fluorescence spectroscopy to characterize the binding of DNA to the cytomegalovirus DNA polymerase. The fluorescence emission spectrum of the purified enzyme in standard buffer at 22°C is shown in figure 3A. The purified protein used in our assays harbors a single tryptophan residue (Trp-780) which provided a significant fluorescence signal upon excitation at 290 nm. The emission maximum of the enzyme (338 nm) is blue-shifted relative to that of free L-tryptophan, which under the same conditions is observed to be at 350 nm. The λ_{max} of tryptophan is highly sensitive to the polarity of the microenvironment in which its indole sidechain is located. Blue shifts of protein emission spectra have been ascribed to shielding of the tryptophan residues from the aqueous phase [33]. This shielding is the result of the protein's three-dimensional structure. Accordingly, denaturation of the enzyme with 8M urea results in a red shift of λ_{max} towards 350 nm (Fig. 3A). The molar intensity of the fluorescence emission spectrum of the purified cytomegalovirus DNA polymerase was also evaluated to determine if significant protein aggregation, or if the loss of protein from solution through adhesion, could influence the binding parameters. As shown in figure 3A, an increase in fluorescence is observed with increasing concentrations of the enzyme. A linear change of 0.03 fluorescence intensity units/nM of protein was observed over the range examined. This relatively small change can be attributed to minor losses of proteins from solution through adhesion. All subsequent binding experiments were therefore performed at a protein concentration of 1 μM , with the assumption that the DNA binding assays were not complicated by the presence of an aggregation equilibrium.

DNA binding activity

In order to monitor the formation of the enzyme-DNA binary complex, increasing concentrations of DNA were added to the purified cytomegalovirus DNA polymerase, and the fluorescence intensity was monitored. We initially monitored the binding of a DNA substrate of 30 nt encompassing the cytomegalovirus immediate-early promoter sequence. We observed that the binding of DNA to a fixed concentration of the enzyme resulted in a modification of the intensity of the intrinsic fluorescence of the protein. Typical emission spectra are shown in figure 3B. The addition of increasing amounts of DNA produced a decrease in the fluorescence intensity. About 80% of the intrinsic protein fluorescence was accessible to the quencher DNA substrate (Fig. 3C). A double-reciprocal plot of the saturation isotherm, generated by plotting the change of fluorescence intensity as a function of added DNA, indicated a K_D value of 47.5 μM (Fig. 3C).

Fluorescence spectroscopy was also used to evaluate the Gibbs free energy of binding (ΔG), as well as both the enthalpy (ΔH) and the entropy changes (ΔS) associated with the binding of DNA to the enzyme. Evaluation of these thermodynamic parameters yields important insight into the nature of the DNA binding reaction. Our binding studies indicate that the Gibbs free energy of binding for the interaction of DNA with the enzyme was -24.8 kJ/mol. The precise enthalpic and entropic contributions to the free energy of binding were then determined by measuring the initial binding of DNA to the enzyme as a function of temperature.

Analysis of a van't Hoff plot for the interaction between DNA and the enzyme (Fig. 3D) revealed that binding of the DNA substrate is entropically favored ($T\Delta S = 32.7$ kJ/mol), with a minor contribution coming from an unfavorable enthalpy change ($\Delta H = 7.9$ kJ/mol). The implications of these thermodynamic findings will be further discussed in the Discussion section.

The use of fluorescence spectroscopy also allowed us to monitor the kinetics of the DNA binding to the cytomegalovirus DNA polymerase (Fig. 3E). The progress of the binding reaction was followed for 30 seconds upon addition of a saturating concentration of DNA. The results show that there is a rapid exponential decrease in fluorescence intensity following the addition of DNA substrate. An apparent association rate of $3.6 \mu\text{M}^{-1}\cdot\text{sec}^{-1}$ was estimated from the data. Half-maximal quenching was observed at approximately 0.1 sec, while maximal quenching was achieved after 0.3 sec of incubation with DNA, and remained constant thereafter. The exponential decrease in fluorescence observed following the addition of DNA was not due to photobleaching, as similar results were obtained when the protein was incubated away from the light source (data not shown).

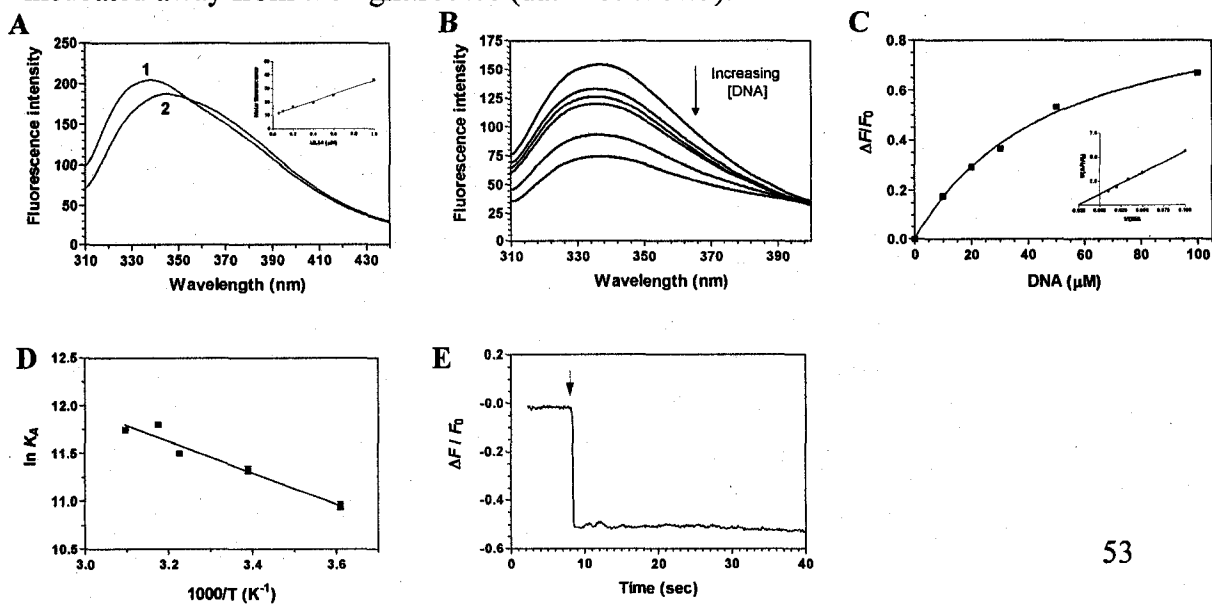


Figure 3 Titration of the cytomegalovirus DNA polymerase with DNA

(A) Background corrected fluorescence emission spectra of the enzyme. 1, Purified protein in 50 mM Tris-HCl, and 50 mM KOAc, pH 7.5; 2, Purified protein after a 2-hour exposure to an 8 M solution of urea at 25°C. Fluorescence spectra were recorded at an excitation wavelength of 290 nm. The molar fluorescence of the protein is also displayed in the right-hand corner. Various concentrations of the purified protein were assayed in 50 mM Tris-HCl, and 50 mM KOAc, pH 7.5. Excitation was performed at 290 nm, and emission was monitored at 338 nm. **(B)** Increasing concentrations of a DNA substrate of 30 nt encompassing the cytomegalovirus immediate-early promoter sequence were added to a 1 μ M solution of the enzyme in binding buffer (20 mM Tris-HCl, pH 7.5, 50 mM NaCl, and 10 mM MgCl₂) and the emission spectrum was scanned from 310 to 440 nm. **(C)** A saturation isotherm can be generated from these data by plotting the change in fluorescence intensity at 338 nm as a function of added DNA. A double-reciprocal plot of the saturation isotherm is shown in the right-hand corner. **(D)** Thermodynamic parameters of the interaction between DNA and the cytomegalovirus DNA polymerase. Binding reactions were performed at various temperatures, and the respective association constants were evaluated. A van't Hoff plot for the interaction between DNA and the protein is shown. The effect of temperature on the association constant was evaluated at pH 7.5. **(E)** Kinetic analysis of real-time binding of DNA to the protein. A 1 μ M solution of

the enzyme was incubated with 100 μ M DNA. Emission was monitored for 40 sec at 338 nm, and excitation was performed at 290 nm.

Minimal DNA length for stable binding

We next sought to determine the minimal length of DNA required for stable binding of the cytomegalovirus DNA polymerase. Fluorescence spectroscopy assays were therefore performed with DNA substrates ranging from 30 to 5 nt (Fig. 4A). Our data indicate that the enzyme could bind efficiently to DNA substrates of 20, 15, 12, 10, 9, 8, 7, and 6 nt in length (Table 1). However, we could not detect the binding of the enzyme to the DNA substrate of 5 nt, and no accurate binding constants could be efficiently and repeatedly determined for this DNA substrate. Based on these data, we conclude that the minimal DNA binding site size of the cytomegalovirus is 6 nt.

Assuming that each nucleotide within the minimal DNA binding site makes the same number of contacts with the enzyme, the apparent K_D should in theory display a small hyperbolic decrease, since each additional nucleotide can provide a new binding position. This is precisely what was observed (Fig. 4B), as the free energy of binding for substrates ranging from 10 to 20 nt changes linearly with a slope of 157 J/mol/nt. Beyond that length, the change in free energy of binding was smaller. Moreover, we observed that the specific constant of binding (B_{max}/K_d), which takes into account both the strength and the maximal level of binding, shows that the

binding efficiency increases dramatically up to a length of 20 nt, and leveled off thereafter (Fig. 4C).

The specificity of the protein-DNA interaction was also monitored by using various nucleic acids substrates. Binding of the enzyme to the DNA substrate is not sequence specific as similar dissociation constants were obtained for DNA substrates (30 nt) of various sequences (Fig. 4D and Table 1). Moreover, similar binding parameters were determined when a double-stranded DNA substrate was used in the binding assays. Surprisingly, similar binding parameters were also obtained for an RNA substrate of 30 nt, indicating that the enzyme does not efficiently discriminate between ribo- and deoxyribonucleic acids (Fig. 4D and Table 1).

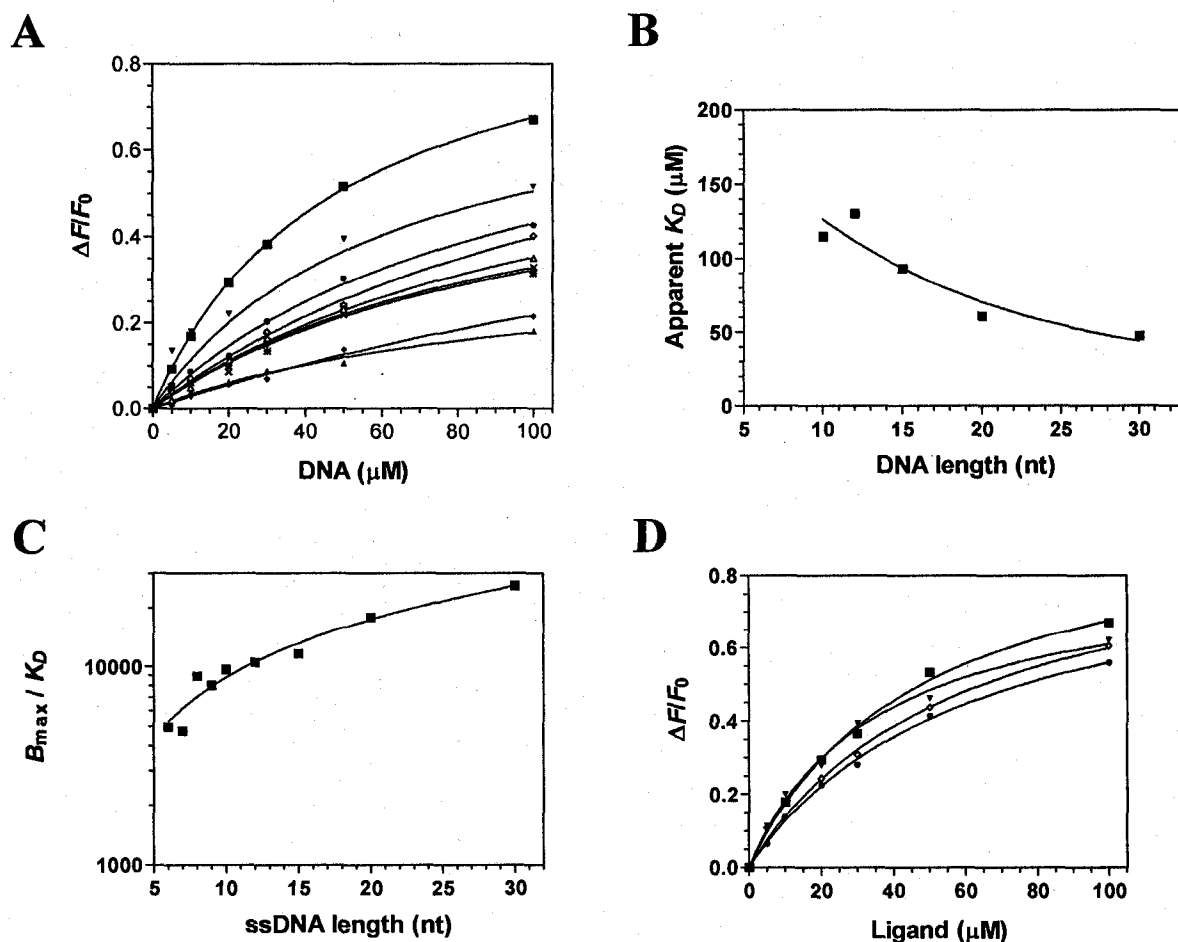


Figure 4 Characterization of the interaction between DNA and the catalytic subunit of the cytomegalovirus DNA polymerase

(A) Affinity of the enzyme for DNA of 30 (■), 20 (▼), 15 (●), 12 (◇), 10 (Δ), 9 (X), 8 (*), 7 (◆), and 6 nucleotides (▲). Increasing concentrations of DNA were added to a 1 μM solution of the enzyme and the emission spectrum was scanned from 310 to 440 nm upon excitation at 290 nm. Saturation isotherms can be generated from these data by plotting the change in fluorescence intensity at 338 nm as a function of added DNA. (B) Interaction of the enzyme with DNA of various lengths. The apparent K_D values were plotted as a function of substrate length. (C) The specificity of binding was also monitored as a function of the length of the various DNA molecules used in the binding assays. (D) The binding of various substrates was monitored by fluorescence spectroscopy. Increasing concentrations of a DNA substrate of 30 nt encompassing the cytomegalovirus immediate-early promoter sequence (■), a DNA substrate corresponding to the complement of the cytomegalovirus immediate-early promoter sequence (▼), a DNA substrate of 30 nt of random sequence (●), and an RNA substrate of 30 nt corresponding to cytomegalovirus immediate-early promoter sequence (◇) were used in the binding assays.

Table 1 Dissociation constants (K_D), association constants (K_A), and binding free energies (ΔG°) for the interaction of UL54 with DNA.

Ligand	Length (nt)	K_D (μM)	K_A (M^{-1})	ΔG° (J/mol)	Ligand structure
Promoter-30	30	48	21 050	-24 660	5' GGGAGCTCGTTTAGTGAACCGTCAGATCTC 3'
Promoter-20	20	61	16 420	-24 050	5' GGGAGCTCGTTTAGTGAACC 3'
Promoter-15	15	93	10 720	-22 990	5' GGGAGCTCGTTTAGT 3'
Promoter-12	12	130	7681	-22 170	5' GGGAGCTCGTTT 3'
Promoter-10	10	115	8719	-22 480	5' GGGAGCTCGT 3'
Promoter-9	9	97	10 320	-22 900	5' GGGAGCTCG 3'
Promoter-8	8	99	10 070	-22 840	5' GGGAGCTC 3'
Promoter-7	7	95	10 500	-22 940	5' GGGAGCT 3'
Promoter-6	6	94	10 680	-22 980	5' GGGAGC 3'
Promoter-5	5	No binding detected			5' GGGAG 3'
Rv-Promoter-30	30	49	20 450	-24 590	5'GAGATCTGACGGTTCATAAACGAGCTCCC 3'
ds-Promoter-30	30	42	23 580	-24 950	5' GGGAGCTCGTTTAGTGAACCGTCAGATCTC 3' 3' CCCTCGAGCAAATCACTTGGCAGTCTAGAG 5'
Random-30	30	61	16 390	-24 040	5'TCAGTAACTTAGCCTTGCTTGAATCATCC 3'
RNA-30	30	59	17 010	-24 130	5'GGGAGCUCGUUUAGUGAACCGUCAGAUUC 3'

Binding of dNTPs

Fluorescence spectroscopy has been shown to be a powerful technique to monitor the binding of nucleotides/deoxynucleotides to enzymes [6,31,32]]. Therefore, we intended to use fluorescence spectroscopy to characterize the binding of deoxynucleotides to the binary complex of the cytomegalovirus DNA polymerase bound to DNA. In order to monitor the binding of deoxynucleotides to the enzyme-DNA complex, increasing concentrations of dATP were added to the purified

cytomegalovirus DNA polymerase bound to DNA, and the fluorescence intensity was monitored. We observed that the binding of dATP to a fixed concentration of the enzyme-DNA complex resulted in a modification of the intensity of the intrinsic fluorescence of the protein. Typical emission spectra are shown in figure 5A. The addition of increasing amounts of dATP produced a decrease in the fluorescence intensity. About 30% of the intrinsic protein fluorescence was accessible to the quencher dATP substrate. A double-reciprocal plot of the saturation isotherm, generated by plotting the change of fluorescence intensity as a function of added dATP, indicated a K_D value of 15 μM (Fig. 5A). The binding of the three other deoxynucleotides was also evaluated. As shown in Table 2, the protein displayed similar binding activity for all four dNTPs, although dTTP appeared as a slightly better substrate as judged by the Gibbs free energy of binding. It should be noted that the dNTP binding activity was absolutely dependent on the presence of a divalent cation. No binding could be detected in the absence of metal ions (data not shown).

Fluorescence spectroscopy was also used to evaluate a number of different parameters involved in the interaction between dNTPs and the enzyme-DNA complex. A critical component of the energetics of protein-ligand binding is the interaction of protein residues with the electrostatic field of the ligand. In order to evaluate the contribution of electrostatic interactions to the dNTP binding activity, binding assays were performed in the presence of increasing ionic strength. The contribution of specific electrostatic interactions to the binding of dATP to the cytomegalovirus DNA polymerase-DNA complex was investigated by monitoring the

salt-dependence of the binding process. As shown in figure 5B, the dATP-enzyme interaction was attenuated at elevated KCl concentrations. The experimentally observed equilibrium constant was clearly dependent on the ionic strength of the solution. The equilibrium binding experiments showed that the apparent K_D at 500 mM KCl was 63 μ M, more than 4 times higher than at 10 mM KCl. This change in binding affinity corresponds to a $\Delta\Delta G$ of 3.6 kJ/mol. Evaluation of the Gibbs energy due to electrostatic interactions ($\Delta\Delta G_{ES}$) by extrapolation of the binding reaction to an ionic strength of 1.0 M, the standard state where electrostatic interactions are effectively eliminated, revealed that 12.9% of the binding energy is derived from electrostatic interactions. These findings indicate that hydrogen bonding and/or electrostatic interactions with the triphosphate moiety contribute to the binding of dNTPs to the enzyme-DNA complex.

Fluorescence spectroscopy was also used to evaluate the Gibbs free energy of binding (ΔG), as well as both the enthalpy (ΔH) and the entropy changes (ΔS) associated with the binding of dATP to the enzyme. Evaluation of these thermodynamic parameters yields important insight into the nature of the dATP binding reaction. Our binding studies indicate that the Gibbs free energy of binding for the interaction of dATP with the enzyme was -28.6 kJ/mol. The precise enthalpic and entropic contributions to the free energy of binding were then determined by measuring the initial binding of dATP to the enzyme as a function of temperature. The binding reaction was shown to be connected with a high enthalpy of association, $\Delta H = -56.5$ kJ/mol. Analysis of a van't Hoff plot for the interaction between dATP

and the enzyme (Fig. 5C) revealed that the $T\Delta S$ value for the binding reaction was -27.9 kJ/mol, clearly indicating that the initial binding step is primarily driven by enthalpy, with an unfavorable entropic contribution. Moreover, these data indicate that the reaction is spontaneous at low temperatures but tends to reverse at higher temperatures. Accordingly, the binding of dATP could not be observed repeatedly at temperatures higher than 70°C (data not shown). This can most probably be attributed to the denaturation of the protein that occurs at higher temperatures, and the concomitant destabilization of the architecture of the enzyme. The implications of these thermodynamic findings will be further discussed in the Discussion section.

Fluorescence spectroscopy was also used to monitor the kinetics involved in the binding of deoxynucleotides to the enzyme-DNA binary complex (Fig. 5D). The binding reaction was followed for 30 seconds upon addition of a saturating concentration of dATP. Our data indicate that there is a rapid exponential decrease in fluorescence intensity following the addition of dATP corresponding to an apparent association rate of $0.4 \mu\text{M}^{-1}\cdot\text{sec}^{-1}$. Half of the binding reaction was observed at approximately 0.5 sec, while maximal binding was attained after 1.5 sec of incubation with dATP, and remained constant thereafter. As was the case with the DNA binding assays, the exponential decrease in fluorescence observed following the addition of dATP was not due to photobleaching, as similar results were obtained when the protein was incubated away from the light source (data not shown).

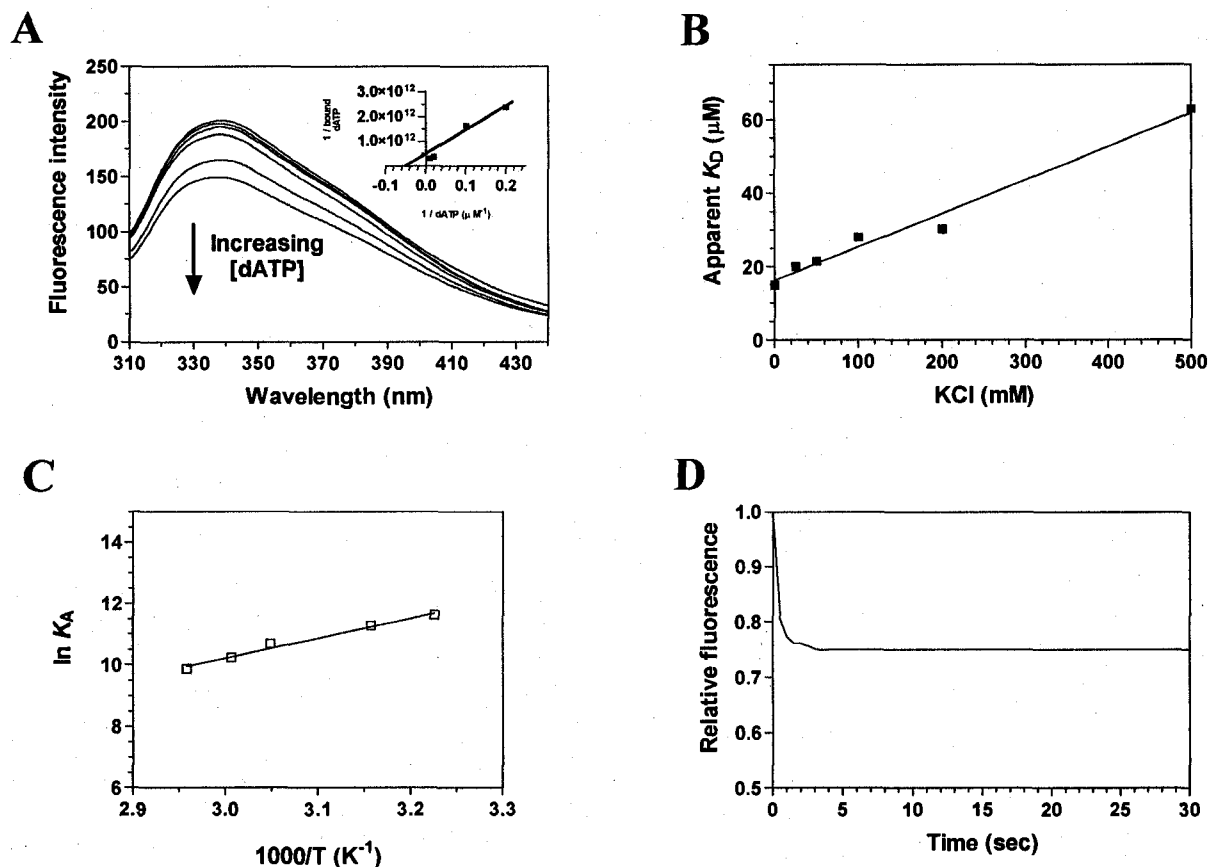


Figure 5 Titration of the cytomegalovirus DNA polymerase with dATP

(A) Increasing concentrations of dATP were added to a 1 μM solution of the enzyme in binding buffer (20 mM Tris-HCl, pH 7.5, 50 mM NaCl, and 10 mM MgCl_2) and the emission spectrum was scanned from 310 to 440 nm. A double-reciprocal plot of the saturation isotherm is shown in the right-hand corner. (B) The effect of increasing ionic strength on the apparent dissociation constant of the enzyme for dATP was investigated. Increasing concentrations of KCl were added to the binding reactions to generate the desired ionic strengths. (C) Thermodynamic parameters of the interaction between dATP and the cytomegalovirus DNA polymerase. Binding

reactions were performed at various temperatures, and the respective association constants were evaluated. A van't Hoff plot for the interaction between dATP and the protein is shown. The effect of temperature on the association constant was evaluated at pH 7.5. **(D)** Kinetic analysis of real-time binding of dATP to the protein. A 1 μ M solution of the enzyme was incubated with 1 mM dATP. Emission was monitored for 30 sec at 338 nm, and excitation was performed at 290 nm.

Table 2 Dissociation constants (K_D), association constants (K_A), and binding free energies (ΔG°) for the interaction of UL54 with deoxynucleoside triphosphates

Ligand	K_D	K_A	ΔG°
dATP	15 μ M	66 670 M ⁻¹	-27 520 J/mol
dGTP	14 μ M	71 430 M ⁻¹	-27 690 J/mol
dCTP	16 μ M	62 500 M ⁻¹	-27 360 J/mol
dTTP	7 μ M	142 900 M ⁻¹	-29 410 J/mol

Effect of DNA binding on the structure of the enzyme

Limited information is available on the structure of the human cytomegalovirus DNA polymerase. Structural studies have been hindered by the absence of a defined crystal structure. Therefore, structural information has been obtained through the use of modelling studies [34] based on other solved polymerases for which crystallographic

information is available. As a consequence, we next sought to investigate whether the binding of DNA to the cytomegalovirus DNA polymerase results in a modification of the protein architecture. Near-UV circular dichroism (CD) spectra can provide useful information on the structural features of a protein, reflecting the environments of the aromatic amino acid side chains and giving information about the tertiary structure of a protein. In an effort to determine if the binding of DNA results in the modification of the cytomegalovirus DNA polymerase structure, near-UV CD spectra were recorded both in the presence and the absence of DNA. Analysis of the near-UV CD spectra of the protein in both the absence and presence of DNA was performed from 250 to 400 nm. As can be seen in figure 6A, a significant reduction of the amplitude of the signal is observed over the 250-300 nm region when the protein is incubated with DNA. Overall, the CD spectra suggest that the protein undergoes conformational changes upon the formation of the enzyme-DNA binary complex.

In order to gain additional insights into the structural modifications that occur upon DNA binding, we investigated the binding of a structural fluorescent reporter to the enzyme. The exposure of the hydrophobic area of the cytomegalovirus DNA polymerase was evaluated by measuring the binding of ANS (1-anilino-8-naphtalenesulfonate) to the protein. ANS is a reporter of exposed hydrophobic surfaces on proteins that binds with high affinity to hydrophobic patches, which results in an enhancement of ANS intrinsic fluorescence [35]. Our data revealed that the unliganded protein binds very weakly to ANS, probably reflecting limited hydrophobic regions at the surface of the protein (Fig. 6B). A very modest modification of the ANS fluorescence is observed when the protein is incubated in

the presence of saturating concentrations of DNA (Fig. 6B). Increasing the urea concentrations did not modify the emission intensity. Overall, these data indicate that the conformational change that is observed upon DNA binding does not involve significant hydrophobic exposure on the surface of the enzyme.

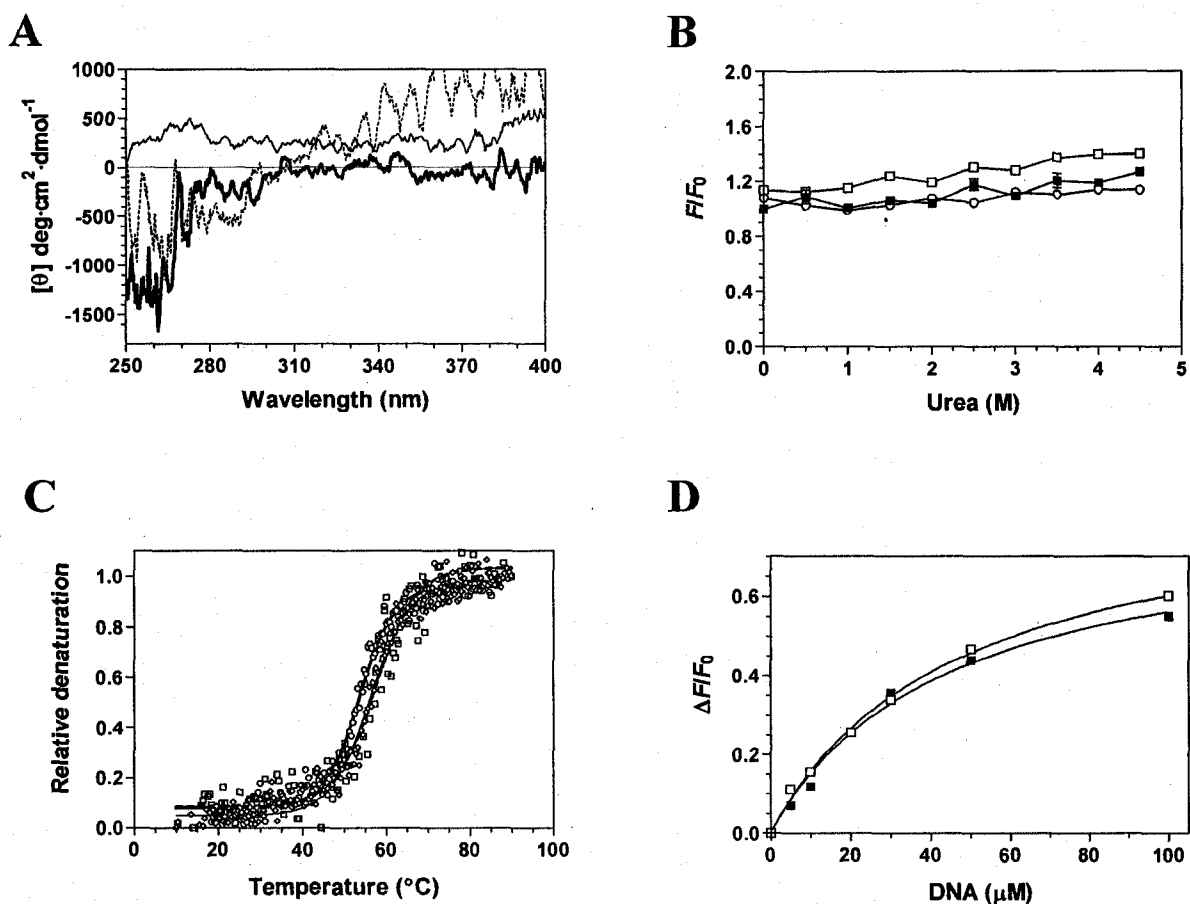


Figure 6 Structural consequences of ligand binding to the cytomegalovirus DNA polymerase

(A) Near-UV CD spectra were recorded for the unliganded protein (*black line*), the protein-DNA binary complex (*thick line*), and the transition state of the enzyme (protein-DNA-dATPαS) (*dotted line*). Formation of the protein-DNA binary complex was induced by incubating the enzyme (1 μ M) with 100 μ M DNA (primer template P20/T60). The transition state was generated by adding 1 mM of dATPαS to the binary complex. The spectra were recorded from 250 to 400 nm, and the average of 3

wavelength scans is presented. **(B)** Binding of ANS to the cytomegalovirus DNA polymerase during urea denaturation. The unliganded protein (1 μM) (■), binary complex (□), and transition state (○) were unfolded with various concentrations of urea at 22 °C for 1 hour. Fluorescence emission was monitored after ANS addition (50 μM) at an excitation wavelength of 380 nm. The integrated fluorescence area between 400 and 600 nm was evaluated. **(C)** Thermal denaturation of the enzyme. Thermal denaturation was recorded for the unliganded protein (◇), the protein-DNA binary complex (□) or the transition state (○). Circular dichroism spectra were recorded at a constant wavelength of 222 nm from 10 to 90 °C at a protein concentration of 1 μM . **(D)** The affinity of the enzyme for the substrate of 30 nt encompassing the cytomegalovirus immediate-early promoter sequence was also monitored in the absence (■) or presence of 1 mM foscarnet (□).

Table 3 Thermodynamic unfolding parameters measured by thermal denaturation

Protein	T_m (°C)	ΔH_{VH} (kcal / mol)	ΔS (kcal / mol / K)
UL54	56.2	-82.0	-0.25
UL54·DNA	56.8	-56.4	-0.17
UL54·DNA·dATP α S·Mg ²⁺	54.6	-86.3	-0.26

The thermodynamic parameters of unfolding ΔH_{VH} (van't Hoff enthalpy of denaturation) and ΔS (entropy of denaturation) were determined by thermal denaturation of the unliganded cytomegalovirus DNA polymerase and the protein bound to various ligands. Temperature-induced denaturations were monitored by circular dichroism spectroscopy at 222 nm.

Structural and Mechanistic Implications

Because our circular dichroism assays indicated conformational changes upon DNA binding to the cytomegalovirus DNA polymerase, this raises the possibility that the formation of the catalytic active site of the enzyme involves additional structural rearrangements. We therefore intended to use spectroscopic approaches to gain insights into the structural modifications that occur during catalysis. An heteropolymeric DNA substrate (P20/T60) and a non-hydrolyzable dATP analog, dATP α S, were used to mimic the transition state (enzyme-DNA-dATP-Mg²⁺). The non-hydrolyzable dATP α S analog was used since the presence of magnesium, DNA, and the classic dATP substrate would result in catalysis and incorporation of nucleotides to the DNA chain. Prior control experiments demonstrated that dATP α S is not efficiently used by the cytomegalovirus DNA polymerase in the presence of magnesium ions (data not shown). As can be seen in 6A, analysis of the near-UV CD spectra revealed that structural changes are occurring when both dATP and magnesium ions were added to the enzyme-DNA complex, indicating that the catalytic center of the enzyme undergoes structural rearrangements upon formation of the transition state.

In an effort to gain insights into the reaction chemistry, the stability of the enzyme bound to various ligands was assessed by circular dichroism spectroscopy. Thermal denaturation assays were performed both in the presence and absence of

ligands, and unfolding of the enzyme was evaluated by monitoring the changes in the α -helix content of the protein (222 nm). Thermal denaturation of the unliganded protein initially revealed a midpoint of thermal transition (T_m) of 56.2 °C (Fig. 6C). Formation of the enzyme-DNA binary complex had no significant effect on the midpoint of thermal transition (Fig. 6C and Table 3), and evaluation of the thermodynamic parameters of unfolding revealed similar values for the unliganded protein or the protein bound to DNA (Table 3). However, denaturation studies performed on the transition state (enzyme-DNA-dATP-Mg²⁺) indicated a reduced stability, reflecting the transient nature of the intermediate (Fig. 6C and Table 3). Note that all of the denaturation assays revealed monophasic unfolding curves, suggestive of a two-state unfolding model. No intermediate state could be detected during the unfolding process.

Finally, we sought to investigate whether the DNA binding activity of the UL54 could be influenced by the presence of foscarnet. Foscarnet is an analog of pyrophosphate which is used for the treatment of cytomegalovirus infections. This antiviral acts noncompetitively with respect to dNTPs but the precise mechanism of action of this drug remains elusive. However, it has previously been proposed that the pyrophosphate analog could bind close to the active center of the enzyme thereby inhibiting the pyrophosphate exchange reaction. We therefore intended to monitor if the binding of foscarnet could influence the binding of DNA to the UL54 protein. Fluorescence spectroscopy was used to measure the binding of DNA in the presence of foscarnet because initial experiments revealed that foscarnet had no influence on

the intrinsic fluorescence properties of the UL54 protein (data not shown). The enzyme was therefore incubated with increasing concentrations of DNA in the presence or absence of foscarnet (1 mM). As shown in figure 6D, the presence of foscarnet had no influence on the DNA binding activity of the enzyme, as similar K_D values were determined in the absence (47.5 μM) or presence of foscarnet (45.8 μM). This result clearly indicates that the binding of the enzyme to the DNA substrate is not influenced by the presence of foscarnet. Note that a similar conclusion was obtained using electrophoretic mobility shift assays performed with a radiolabeled DNA substrate (data not shown).

DISCUSSION

The human cytomegalovirus DNA polymerase is a critical enzyme for the replication of the virus [7-9,13]. As a specific target for antiviral chemotherapy, it plays a key role in the treatment of infected individuals [14-17]. However, characterization of the cytomegalovirus DNA polymerase has been very limited, and a number of critical questions about the reaction chemistry remain to be addressed. These include 1) what are the thermodynamic parameters that are involved in the formation of the enzyme-DNA binary complex of DNA? 2) what are the conformational changes that occur during the binding of DNA to the enzyme? 3) what is the minimal length of the DNA substrate required for stable binding? 4) is the binding sequence-specific? and 5) what are the thermodynamic and structural implications of dNTP binding? In this regard, the multidimensional properties of fluorescence spectroscopy and circular dichroism

provide accurate sensitivity to monitor numerous aspects of protein-ligand interactions.

In the present study, spectroscopic approaches were used to allow, for the first time, a precise quantitation of the kinetic and thermodynamic parameters associated with the binding of DNA to the cytomegalovirus DNA polymerase. While these binding parameters do not provide a complete picture of the binding activity, they can suggest which features are likely to be important for the interaction, and provide a framework to construct an accurate model of the complex. For instance, changes in both the magnitude and the sign of ΔH and ΔS , in conjunction with structural data generated by spectroscopy, can provide crucial information about the structural alterations that accompany ligand binding in terms of 1) changes in solvation state; 2) interactions between ligand and protein, such as electrostatic and hydrophobic interactions; and 3) changes in conformation/dynamics induced by ligand binding. Our thermodynamic analyses indicate that the formation of the protein-DNA binary complex is an entropy-driven process ($T\Delta S = 32.7$ kJ/mol). Such entropic contributions, which have been observed in other enzymes involved in DNA binding, are generally associated with conformational changes and/or from the release of water molecules to the bulk solvent upon DNA binding [36]. Our circular dichroism assays clearly demonstrated the existence of conformational changes in the cytomegalovirus DNA polymerase following the binding of DNA. Interestingly, hydrophobic interactions are also associated with a relatively small ΔH (as compared to ΔG) and a positive ΔS . Since the parameters for the UL54-DNA interaction fit these characteristics, it is tempting to speculate that hydrophobic interactions are also an

important determinant for the formation of the complex. In accordance with this hypothesis, examination of the ionic strength dependence for the interaction between DNA and the protein reveals that the affinity of the enzyme for DNA is only slightly affected by an increased ionic strength (data not shown). Taken together with the thermodynamic parameters, these data suggest that hydrophobic and stacking interactions probably play a significant role in the formation of the enzyme-DNA complex than ionic contacts, which would be expected to be completely disrupted upon an increase in ionic strength of the solvent.

In comparison to the DNA binding analyses, our thermodynamic studies of the interaction between the enzyme and dNTPs revealed a completely different mode of binding. Binding of dNTPs to the enzyme-DNA binary complex was dominated by a favorable enthalpy change, but is also clearly associated with an unfavorable entropy change. Favorable negative enthalpy changes are generally associated with contributions from hydrogen bonds, van der Waal's interactions, or ionic interactions [36], while the unfavorable negative entropy changes are associated with the exposure of hydrophobic surfaces to the surface of the protein, and/or to a decrease in conformational flexibility [37]. Although our salt-dependence assays clearly established the importance of electrostatic interactions for the binding of dATP to the protein-DNA binary complex, both our fluorescence spectroscopy and circular dichroism studies indicate that aromatic residues move to a more hydrophobic environment upon binding, suggesting that a decrease in conformational flexibility is probably encountered upon dATP binding.

The absence of a crystal structure of the cytomegalovirus DNA polymerase makes it difficult to precisely locate the conformational changes that occur upon DNA and dNTP binding. However, based on the crystal structure of the closely related bacteriophage RB69 DNA polymerase [38], a three-dimensional modeling of the cytomegalovirus DNA polymerase has recently been proposed which can provide useful information on the location of specific residues of the cytomegalovirus DNA polymerase [34]. Interestingly, crystallographic studies of many DNA polymerases have revealed similar structural features that are similar to a partially opened right-hand with thumb, palm, and fingers domains [39-42]. The thumb domain associates with the DNA template, the palm harbors the essential carboxylate residues that are involved in catalysis, and the fingers domain is involved in the interaction with the incoming nucleotide [39-42]. Moreover, the crystal structures of various conformations of the RB69 DNA polymerase have been elucidated, and have provided crucial information on the structural rearrangements that occur during catalysis. For instance, a comparison of the RB69 DNA polymerase crystals with and without DNA indicated structural modifications upon DNA binding. These DNA-induced changes are mainly confined to the thumb domain of the enzyme. Upon DNA binding, the entire thumb domain is translocated relative to its position in the unliganded enzyme, closing down around the DNA, and conferring a more 'open' conformation to the protein. Analysis of the RB69 polymerase crystal structure also indicated that the thumb domain contacts predominantly the phosphodiester backbone of the DNA. Based on our nucleic acid binding assays, we demonstrated that the

cytomegalovirus UL54 protein does not discriminate between DNA or RNA, consistent with the binding of the enzyme to the phosphodiester backbone of nucleic acids. It is therefore tempting to speculate that the UL44 accessory subunit, which associates with UL54, is critical for the specificity of the DNA polymerase for the DNA substrate/template. Such an interaction between UL54 and UL44 might also be crucial for sequence-specific binding, since the binding of the UL54 protein to DNA is clearly not sequence-specific.

Interestingly, the K_D values that were determined in the present study for the various dNTPs (Table 1) are significantly higher than the K_M values that were previously evaluated [14,22,34] from polymerization assays (ranging from 0.16 to 0.8 μM). Similar K_M values were also obtained with the purified protein used in the current study (data not shown). A likely explanation for the observation that the K_M values are smaller than the respective K_D values is that the uptake of the occupied binding site for dNTPs is probably fast relative to the reappearance of the unoccupied binding site during catalysis. Moreover it is interesting to note that substrate-independent V_{max} has previously been observed for the UL54 subunit [22]. This suggests that a substrate-independent step in catalysis is rate-limiting. We envision that this is probably the reappearance of the unoccupied binding site for dNTPs during catalysis, which is consistent with a small K_M/K_D ratio.

Crystal structures have now been elucidated for a number of DNA polymerases bound to various ligands (DNA, dNTPs, metal ions). A general model for DNA

polymerization has been generated by comparison of these different structures [39-42]. The initial step involves the binding of the DNA polymerase to the DNA substrate, which causes the thumb domain of the enzyme to close around the DNA, leading to a more 'open' conformation to the protein. The next step involves the binding of the dNTP to this complex. This causes a conformational change in the fingers domain of the enzyme, which rotate toward the catalytic center of the enzyme, conferring a more 'closed' catalytically competent conformation to the protein [39-42]. It is interesting to notice that such conformational changes have now been observed for the UL54 catalytic subunit of the cytomegalovirus DNA polymerase. Using spectroscopic approaches, the current study demonstrated that the protein undergoes structural modifications upon formation of the enzyme-DNA complex which is followed by another conformational change upon dNTP binding. Although the complete understanding of the mechanisms underlying the replication of cytomegalovirus is still largely incomplete, characterization of the individual biochemical steps involved in catalysis should provide the basis for further studies in this direction. We believe that structural measurements performed on purified proteins should enable us to correlate conformation and kinetic parameters of ligand recognition, and add another dimension to investigate the individual biochemical steps involved in the replication of cytomegalovirus.

ACKNOWLEDGEMENTS

We thank Dr. Jay A. Nelson (Oregon Health and Sciences University) for generously providing the initial plasmid harboring the UL54 gene. We are also very grateful to Audrey Dubé, Dominique Lévesque, and François Bachand for technical assistance. This work was supported by a grant from the Canadian Institutes for Health Research. M.B. is a New Investigator Scholar from the Canadian Institutes for Health Research.

REFERENCES

1. Britt, W. J. and Alford, C. A. (1996). In *Fields Virology*, Lippincott-Raven Publishers, Philadelphia, PA
2. Trincado, D. E., Scott, G. M., White, P. A., Hunt, C., Rasmussen, L. and Rawlinson, W. D. (2000). Human cytomegalovirus strains associated with congenital and perinatal infections. *J. Med. Virol.* **61**, 481-487
3. Chou, S. W. 2001. Cytomegalovirus drug resistance and clinical implications. *Transpl. Infect. Dis.* **3**, 20-24
4. Gilbert, C. and Boivin, G. (2005). Human cytomegalovirus resistance to antiviral drugs. *Antimicrob. Agents Chemother.* **49**, 873-883
5. Springer, K. L. and Weinberg, A. (2004). Cytomegalovirus infection in the era of HAART: fewer reactivations and more immunity. *J. Antimicrob. Chemother.* **54**, 582-586
6. Weinberg, A., Jabs, D. A., Chou, S., Martin, B. K., Lurain, N. S., Forman, M. S. and Crumpacker, C. (2003). Mutations conferring foscarnet resistance in a cohort of patients with acquired immunodeficiency syndrome and cytomegalovirus retinitis. *J. Infect. Dis.* **187**, 777-784
7. Ertl, P. F. and Powell, K. L. (1992). Physical and functional interaction of human cytomegalovirus DNA polymerase and its accessory protein (ICP36) expressed in insect cells. *J. Virol.* **66**, 4126-4133
8. Heilbronn, R., Jahn, G., Burkle, A., Freese, U. K., Fleckenstein, B. and zur Hausen, H. (1987). Genomic localization, sequence analysis, and transcription of the putative human cytomegalovirus DNA polymerase gene. *J. Virol.* **61**, 119-124

9. Ye, L. B. and Huang E. S. (1993). In vitro expression of the human cytomegalovirus DNA polymerase gene: effects of sequence alterations on enzyme activity. *J. Virol.* **67**, 6339-6347
10. Mar, E. C., Patel, P. C. and Huang, E. S. (1981). Human cytomegalovirus-associated DNA polymerase and protein kinase activities. *J. Gen. Virol.* **57**, 149-156
11. Weiland, K. L., Oien, N. L., Homa, F. and Wathen, M. W. (1994). Functional analysis of human cytomegalovirus polymerase accessory protein. *Virus Res.* **34**, 191-206
12. Loregian, A., Rigatti, R., Murphy, M., Schievano, E., Palu, G. and Marsden, H. S. (2003). Inhibition of human cytomegalovirus DNA polymerase by C-terminal peptides from the UL54 subunit. *J. Virol.* **77**, 8336-8344
13. Crumpacker, C. S. (1992). Mechanism of action of foscarnet against viral polymerases. *Am. J. Med.* **92**, 3S-7S
14. Xiong, X., Smith, J. L., Kim, C., Huang, E. S. and Chen, M. S. (1996). Kinetic analysis of the interaction of cidofovir diphosphate with human cytomegalovirus DNA polymerase. *Biochem. Pharmacol.* **51**, 1563-1567
15. Xiong, X., Smith, J. L. and Chen, M. S. (1997). Effect of incorporation of cidofovir into DNA by human cytomegalovirus DNA polymerase on DNA elongation. *Antimicrob. Agents Chemother.* **41**, 594-599
16. Eriksson, B., Oberg, B. and Wahren, B. (1982). Pyrophosphate analogues as inhibitors of DNA polymerases of cytomegalovirus, herpes simplex virus and cellular origin. *Biochim. Biophys. Acta* **696**, 115-123

17. Derse, D., Bastow, K. F. and Cheng, Y. (1982). Characterization of the DNA polymerases induced by a group of herpes simplex virus type I variants selected for growth in the presence of phosphonoformic acid. *J. Biol. Chem.* **257**, 10251-10260
18. Griffiths, P. D. (2001). Cytomegalovirus therapy: current constraints and future opportunities. *Curr. Opin. Infect. Dis.* **14**, 765-768
19. Chou, S., Lurain, N. S., Thompson, K. D., Miner, R. C. and Drew, W. L. (2003). Viral DNA polymerase mutations associated with drug resistance in human cytomegalovirus. *J. Infect. Dis.* **188**, 32-39
20. Cihlar, T., Fuller, M. D. and Cherrington, J. M. (1998). Characterization of drug resistance-associated mutations in the human cytomegalovirus DNA polymerase gene by using recombinant mutant viruses generated from overlapping DNA fragments. *J. Virol.* **72**, 5927-5936
21. Smith, I. L., Cherrington, J. M., Jiles, R. E., Fuller, M. D., Freeman, W. R. and Spector, S. A. (1997). High-level resistance of cytomegalovirus to ganciclovir is associated with alterations in both the UL97 and DNA polymerase genes. *J. Infect. Dis.* **176**, 69-77
22. Tchesnokov, E. P., Gilbert, C., Boivin, G. and Gotte, M. (2006). Role of helix P of the human cytomegalovirus DNA polymerase in resistance and hypersusceptibility to the antiviral drug foscarnet. *J. Virol.* **80**, 1440-1450
23. Cihlar, T., Fuller, M. D. and Cherrington, J. M. (1997). Expression of the catalytic subunit UL54 and the accessory protein UL44 of human cytomegalovirus DNA

- polymerase in a coupled in vitro transcription/translation system. *Prot. Exp. Purif.* **11**, 209-218
24. Loregian, A., Appleton, B. A., Hogle, J. M. and Coen, D. M. (2004). Residues of human cytomegalovirus DNA polymerase catalytic subunit UL54 that are necessary and sufficient for interaction with the accessory protein UL44. *J. Virol.* **78**, 158-167
25. Ito, J. and Brathwaite, D. K. (1991). Compilation and alignment of DNA polymerase sequences. *Nucleic Acids Res.* **19**, 4045-4057
26. Bougie, I. and Bisailon, M. (2003). Initial binding of the broad spectrum antiviral nucleoside ribavirin to the hepatitis C virus RNA polymerase. *J. Biol. Chem.* **278**, 52471-52478
27. Bougie, I., Charpentier, S. and Bisailon, M. (2003). Characterization of the metal ion binding properties of the hepatitis C virus RNA polymerase. *J. Biol. Chem.* **278**, 3868-3875
28. Bougie, I., Parent, A. and Bisailon, M. (2004). Thermodynamics of ligand binding by the yeast mRNA-capping enzyme reveals different modes of binding. *Biochem. J.* **384**, 411-420
29. Flowers, S., Biswas, E. E. and Biswas, S. B. (2003). Conformational dynamics of DnaB helicase upon DNA and nucleotide binding: analysis by intrinsic tryptophan fluorescence quenching. *Biochemistry* **42**, 1910-1921
30. Painter, G. R., Wright, L. L., Hopkins, S. and Furman, P. A. (1991). Initial binding of 2'-deoxynucleoside 5'-triphosphates to human immunodeficiency virus type 1 reverse transcriptase. *J. Biol. Chem.* **266**, 19362-19368

31. Pan, J. Y., Sanford, J. C. and Wessling-Resnick, M. (1995). Effect of guanine nucleotide binding on the intrinsic tryptophan fluorescence properties of Rab5. *J. Biol. Chem.* **270**, 24204-24208
32. Zhou, T. and Rosen, B. P. (1997). Tryptophan fluorescence reports nucleotide-induced conformational changes in a domain of the ArsA ATPase. *J. Biol. Chem.* **272**, 19731-19737
33. Eftink, M. R. and Ghiron, C. A. (1976). Exposure of tryptophanyl residues in proteins. Quantitative determination by fluorescence quenching studies. *Biochemistry* **15**, 672-680
34. Shi, R., Azzi, A., Gilbert, C., Boivin, G. and Lin, S. X. (2006). Three-dimensional modeling of cytomegalovirus DNA polymerase and preliminary analysis of drug resistance. *Proteins* **64**, 301-307
35. Labowicz, J. R. (1999). In *Principles of Fluorescence Spectroscopy*, 2nd Ed., Kluwer/Plenum, New York, NY
36. Beaudette, N. V. and Langerman, N. (1980). The thermodynamics of nucleotide binding to proteins. *Crit. Rev. Biochem.* **9**, 145-169
37. Tame, J. R. H., O'Brien, R. and Ladbury, J. E. (1998). In *Biocalorimetry: Applications of calorimetry in biological sciences*, John Wiley and Sons, London
39. Brautigam, C. A. and Steitz, T. A. (1998). Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes. *Curr. Opin. Struct. Biol.* **8**, 54-63
40. Delarue, M., Poch, O., Tordo, N., Moras, D. and Argos, P. (1990). An attempt to unify the structure of polymerases. *Protein Eng.* **6**, 461-467

41. Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G. and Steitz, T. A. (1985).
Structure of large fragment of Escherichia coli DNA polymerase I complexed
with dTMP. *Nature* **313**, 762-766
42. Savino, C., Federici, L., Johnson, K. A., Vallone, B., Nastopoulos, V., Rossi, M.,
Pisani, F. M. and Tsernoglou, D. (2004). Insights into DNA replication: the
crystal structure of DNA polymerase B1 from the archaeon *Sulfolobus*
solfataricus. *Structure* **12**, 2001-2008
- 43 Wang T.S., Wong S.W. and Korn D. (1989) Human DNA polymerase alpha:
predicted functional domains and relationships with viral DNA polymerases.
FASEB . **3** 14-21

9) Discussion et conclusion

9.1) Résumé

Pour la première fois, la polymérase UL54 du cytomégalovirus a été exprimée et purifiée à partir d'un système hétérologue bactérien. La grande quantité de protéines ainsi générée a permis la première étude visant la caractérisation exhaustive de l'activité de liaison de l'ADN et des désoxynucléotides à la protéine UL54, la sous-unité catalytique de l'ADN polymérase du Cytomégalovirus. Les vitesses d'association, les affinités, les profils thermodynamiques, les forces conductrices, les conséquences structurales et les répercussions sur la stabilité ont été résolus, dressant ainsi un bien meilleur portrait d'ensemble de la principale cible pharmacologique de ce virus.

9.2) Format

L'espace alloué pour cette discussion ne sera pas utilisé afin de simplement réitérer les conclusions que vous venez de lire dans l'article 1. Il sera plutôt utilisé afin d'ajouter quelques résultats complémentaires, ainsi que de pousser davantage la réflexion sur certains points. Enfin, les perspectives proposées en vue de compléter ce projet seront exposées.

10) Résultats supplémentaires

10.1) Modèle par homologie

La structure tertiaire de UL54 n'a pas encore été résolue, ni par cristallographie, ni par résonance magnétique nucléaire, et ni par microscopie cryogénique électronique. Cependant, au cours de la dernière décennie, plusieurs autres polymérases de la famille B ont vu leur structure tertiaire être déterminée avec une haute résolution. À ce sujet, l'une des polymérases les plus étudiées est l'ADN polymérase gp43, du bactériophage RB69. Les structures de seize de ses cristaux sont connues, et la résolution de ces derniers va jusqu'à 2.21 Å (Wang, Sattar et al. 1997; Franklin, Wang et al. 2001; Freisinger, Grollman et al. 2004; Hogg, Wallace et al. 2004; Zahn, Belrhali et al. 2007). En nous basant sur la structure tertiaire de gp43 (2P5O (Hogg, Wallace et al. 2004)) (**Figure 10 A**), nous avons employé le programme DeepView afin de générer un modèle par homologie de l'enzyme UL54²⁹³⁻⁹⁵³ (**Figure 10 B**). La superposition des deux structures révèle que, malgré une très faible homologie de séquence primaire, les deux protéines partagent une très forte homologie de structure tertiaire (**Figure 10 C**). De plus, ce modèle par homologie confirme que le segment

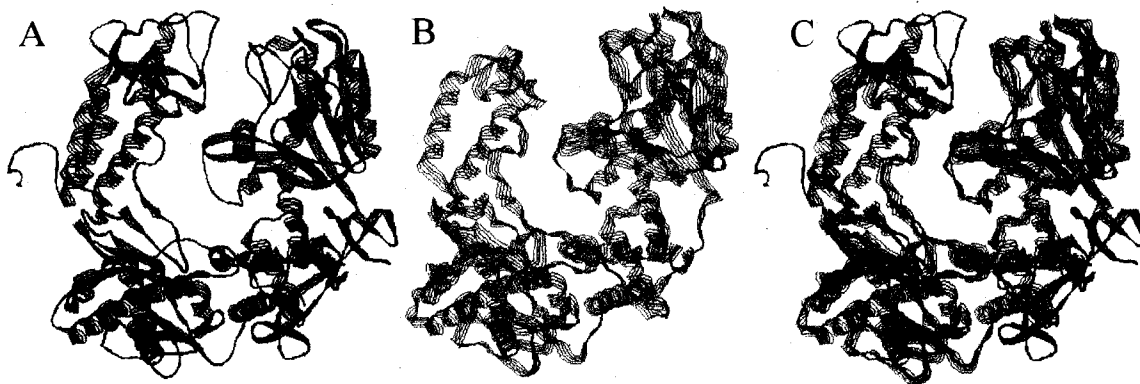


Figure 10 : (A) Structure tridimensionnelle de la polymérase de la famille B gp43, l'ADN polymérase du bactériophage RB69. Générée à partir du cristal 2P5O (Hogg, Wallace et al. 2004). (B) Modélisation par homologie de la polymérase UL54²⁹³⁻⁹⁵³ du cytomegalovirus (DeepView). (C) Superposition de A et B.

293-953 de la protéine UL54 que nous avons exprimée au laboratoire comporte les séquences nécessaires à l'adoption de la structure tertiaire en forme de main droite caractéristique des polymérase.

10.2) Quantification des interactions électrostatiques de la liaison

UL54-ADN

Dans l'article 1, le profil thermodynamique de la liaison de UL54 à l'ADN a été établi. Ce profil prédit une nette dominance des forces hydrophobes, des interactions d'empilements et de la relâche de molécules d'eau au solvant, alors que les interactions électrostatiques et la formation de ponts hydrogènes ne seraient que marginales. Cependant, la quantification de ces interactions électrostatiques n'a pas été présentée dans l'article 1.

Cette quantification s'effectue en observant l'impact, sur la constante de dissociation UL54-ADN, d'une augmentation de la force ionique du milieu (**Figure 11**).

L'interaction UL54-ADN n'est que très faiblement atténuée en présence de fortes

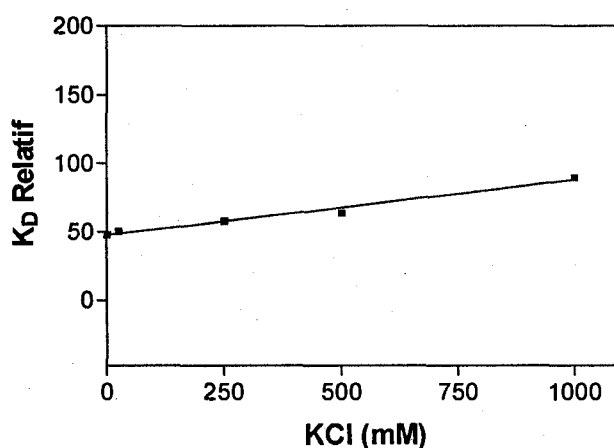


Figure 11 : Constante de dissociation de UL54 avec un ADN de 30nt, en fonction d'une force ionique croissante.

concentrations de chlorure de potassium (KCl). Cet essai démontre que dans des conditions où les interactions électrostatiques sont éliminées (1M KCl), le K_d est de 90 μ M. Cette différence d'affinité correspond à une variation d'énergie libre de Gibbs ($\Delta\Delta G$) de 1.5 kJ/mol,

ce qui indique qu'uniquement 6% de l'énergie totale de liaison est dérivé d'interactions électrostatiques. La nature des forces impliquées dans la stabilisation de la liaison UL54-ADN sera discutée plus loin.

Ce résultat supporte les conclusions de l'étude thermodynamique de la liaison UL54-ADN. De plus, cette propriété qu'a UL54 de ne pas être affecté par la force ionique quand vient le temps de lier son substrat principal, l'ADN, pourrait, à tout le moins en partie, expliquer une autre de ses propriétés. En effet, les premières études cinétiques sur cette enzyme ont révélé que cette polymérase était active même en présence de hautes concentrations de sel (HUANG 1975). Même si, à lui seul, ce résultat ne permet pas d'expliquer cette caractéristique, il permet de confirmer que la polymérase peut toujours lier l'ADN à ces concentrations de sel.

11) Discussion complémentaire

11.1) Expression de UL54 restreinte aux acides aminés 293-953

L'ADN polymérase UL54 étant l'une des principales cibles thérapeutiques chez le CMV, plusieurs groupes en ont fait leur sujet d'étude. Ils ont d'ailleurs réussi à exprimer une protéine pleine longueur grâce à des systèmes de transcription/traduction *in vitro*, ou encore à l'aide de cellules d'insectes. Bien que ces systèmes offrent beaucoup d'avantages, ils ont l'inconvénient de ne produire qu'une faible quantité de protéine. Or, les techniques comme la spectroscopie à fluorescence et le dichroïsme circulaire, qui nous offrent la possibilité de pousser l'étude biochimique de cette protéine beaucoup plus loin, requièrent une grande

quantité de protéine. C'est pourquoi nous nous sommes tournés vers le système d'expression particulièrement performant qu'est l'expression de protéines recombinantes chez *E. coli* sous contrôle d'un promoteur T7 inductible. Cependant, l'une des limitations de ce système est la taille des protéines qui peuvent être efficacement exprimées. La protéine UL54, avec 1242 acides aminés et un poids moléculaire de 137.3 kDa, excède cette limite. Une analyse plus détaillée de la protéine UL54 nous a permis de conclure que l'expression d'une protéine recombinante contenant les acides aminés de 293 à 953 avait de très bonnes chances de générer une enzyme catalytiquement active. En effet, cette protéine conserve les domaines du pouce, de la paume, des doigts ainsi que le domaine exonucléase. C'est au sein de ces domaines que l'on retrouve les sites de liaison pour les substrats de la polymérase : ADN, dNTPs et ions. La superposition d'un modèle par homologie de UL54²⁹³⁻⁹⁵³ (que nous avons généré avec le cristal de la polymérase gb43 du bactériophage RB69) démontre bien que UL54²⁹³⁻⁹⁵³ adopte bien une structure tridimensionnelle en forme de main droite (figure 10 C). Les domaines N-terminal et C-terminal qui n'ont pas été retenus sont, pour leur part, impliqués dans les interactions protéine-protéine et sont, par conséquent, responsables de la localisation, de la régulation et du recrutement de cofacteurs et d'autres protéines, mais ils ne sont pas impliqués directement dans l'activité enzymatique. Le fragment UL54²⁹³⁻⁹⁵³ est d'ailleurs catalytiquement fonctionnel et présente une activité comparable à UL54 pleine longueur, une protéine contrôlée exprimée par transcription/traduction *in vitro* (Figure 2C de l'article 1 page 49). Comme notre étude porte principalement sur la caractérisation de la protéine UL54 avec ses ligands principaux (ADN, dNTPs), nous

sommes très confiants que les paramètres déterminés à partir de la région catalytique UL54²⁹³⁻⁹⁵³ sont valables et peuvent être extrapolés à la protéine pleine longueur. D'ailleurs, pratiquement toutes les mutations conférant une résistance aux drogues anticytomégalovirique recensées chez UL54 se situent dans la région 293-953. Une limitation potentielle de l'utilisation de la protéine tronquée UL54²⁹³⁻⁹⁵³, serait lors de l'étude ou du développement d'inhibiteurs allostériques, ce qui n'est pas le but de la présente étude.

11.2) Écart entre les K_D et les K_M

De manière intéressante, les K_D entre UL54 et les dNTPs (7-15 μ M) sont significativement plus élevés que les K_M retrouvés dans la littérature (0.16-0.8 μ M) (Wang, Wong et al. 1989; Xiong, Smith et al. 1996; Tchesnokov, Gilbert et al. 2006). Nous avons d'ailleurs obtenu des valeurs K_M similaires à ceux établis dans la littérature, avec la protéine purifiée utilisée au cours de cette étude. Cette différence entre les K_D et les K_M peut être expliquée par plusieurs facteurs. D'abord, un K_D est une valeur d'équilibre et ne renseigne en rien sur les vitesses d'association et de dissociation elles-mêmes, mais plutôt sur leur rapport relatif. Dans le cas présent, l'association et la dissociation sont probablement très rapides, ce qui voudrait dire que les dNTPs peuvent lier très rapidement UL54, mais peuvent aussi s'en dissocier très rapidement s'ils ne sont pas utilisés, augmentant ainsi la valeur du K_D . Chez les polymérases, et fort probablement aussi chez UL54, ce n'est pas la liaison du substrat qui est l'étape limitante (Joyce et Benkovic 2004), ce qui semble indiquer que l'utilisation du substrat est probablement rapide par rapport à la réapparition d'un site

de liaison non occupé. Chez plusieurs polymérases, l'affinité pour les dNMPs est moins grande que celle pour les dNTPs, ce qui favorise un K_M plus faible que le K_D (Doronin, Lavrik et al. 1987). Deux expériences rapides pourraient vérifier ces hypothèses. Dans un premier temps, nous pourrions utiliser la même procédure expérimentale que celle employée pour déterminer les K_D entre UL54 et les dNTPs, mais cette fois avec chacun des dNMPs ainsi que du pyrophosphate. Ainsi, nous pourrions comparer l'affinité du produit et du substrat. Dans un deuxième temps, nous pourrions tirer profit de la possibilité d'analyse en temps réel que nous offre la fluorescence à spectroscopie afin de mesurer le taux d'association des dNTPs à UL54 comme nous l'avons fait pour l'association initiale de UL54 à l'ADN. Ainsi, nous pourrions avoir une meilleure idée de la vitesse d'association entre UL54 et les dNTPs.

11.3) Forces responsables de la liaisons UL54-ADN

11.3.1) Contexte

Est-ce que toutes les polymérases de la famille B partagent le même mécanisme de reconnaissance pour l'ADN ? Sachant que chez cette famille de polymérases, la majorité des résidus impliqués dans la liaison de l'ADN, autant au niveau du pouce et de la paume que des doigts, ne sont pas conservés (Franklin, Wang et al. 2001), la réponse est potentiellement non. Le faible niveau de conservation des résidus impliqués dans une tâche si importante suggère que leur interaction avec l'ADN n'est pas importante individuellement, mais bien collectivement (Franklin, Wang et al.

2001). D'ailleurs, leur structure secondaire et tertiaire sont beaucoup plus conservées que leur séquence primaire (Brautigam et Steitz 1998).

11.3.2) Forces responsables de la liaison GP43-ADN

Chez beaucoup de polymérase de la famille B, comme celle du bactériophage RB69 (GP43), qui a été largement étudiée, l'enzyme contacte le phospho-squelette du sillon mineur de l'ADN grâce à des ponts hydrogènes, avec ses chaînes principales et latérales (Wang, Sattar et al. 1997; Franklin, Wang et al. 2001). Ces interactions sont aussi consolidées par plusieurs interactions charges-charges, surtout avec le brin amorce (Franklin, Wang et al. 2001).

11.3.3) Forces responsables de la liaison UL54-ADN

La liaison de l'ADN par UL54 semble légèrement différente. En effet, la contribution des interactions électrostatiques (ponts H, interactions charges-charges) à la liaison se chiffre à aussi peu que 6% de l'énergie totale de liaison, ce qui est beaucoup moins important que pour la majorité des polymérase de cette famille. Cette faible contribution des interactions électrostatiques est compensée par de nombreuses interactions d'empilements aromatiques, d'importantes forces hydrophobes et la relâche, vers le solvant, de nombreuses molécules d'eau qui étaient coordonnées au site de liaison de UL54 ou à l'intérieur du sillon mineur de l'ADN. Les empilements aromatiques peuvent être énergétiquement très favorables (~40-50zJ). Là où ils contribuent au moins autant que les ponts hydrogènes à la stabilisation de la double hélice d'ADN, ils peuvent aussi fortement stabiliser la liaison de cette dernière à une protéine (Delaage 1991). Les forces hydrophobes favorables sont le résultat du

transfert des groupements aromatiques des nucléosides du solvant vers le site de liaison de UL54. Ce patron de forces, couplé au profil thermodynamique entropique, est typique de la liaison d'un acide nucléique à une protéine (Ladbury et Doyle 2004). On voit donc se dessiner un mode de reconnaissance de l'ADN légèrement différent entre UL54 et RB69, ce qui n'est peut-être pas si étonnant, étant donné la quasi non-conservation des acides impliqués dans cette liaison. Cependant, malgré un mode de reconnaissance apparemment distinct, la liaison de l'ADN à UL54 et RB69 est très similaire. Les deux lient l'ADN avec une constante d'affinité du même ordre, elles sont non spécifiques à la séquence d'acides nucléiques, et elles lient toutes deux l'ADN via le sillon mineur, trois caractéristiques importantes pour leur activité.

12) Perspectives

Les perspectives de ce travail sont exposées en deux volets, le premier regroupant les perspectives à court terme et contenant une série d'expériences visant à compléter et à confirmer les résultats du présent projet. Le second volet porte sur les perspectives à plus long terme et expose des voies qu'il serait intéressant d'explorer, ainsi que les approches que nous privilégierions.

12.1) Perspectives à court terme

12.1.1 Titration calorimétrique isothermique (microcalorimétrie)

12.1.1.1) Contexte

Beaucoup de résultats présentés dans l'article 1 sont issus de techniques fiables, éprouvées et reproductibles (fluorescence à spectroscopie, dichroïsme circulaire, ...). Ce sont surtout des techniques indirectes, qui ont l'incroyable avantage d'être « spectatrice » de la liaison, sans la biaiser en interagissant directement avec elle. Il serait cependant intéressant de contre-vérifier quelques résultats importants à l'aide d'une seconde technique indirecte. Une concordance des résultats ne ferait qu'en renforcer leurs poids.

12.1.1.2) Choix de la microcalorimétrie

Vu le patron d'expériences que nous avons exécuté et surtout, le profil des informations que nous en avons tiré, une nouvelle technique semble toute désignée pour confirmer nos résultats; il s'agit de la titration calorimétrique isothermique. Cette technique permet de suivre la liaison entre une protéine et un acide nucléique, en quantifiant la variation de chaleur qui en résulte. Pour ce faire, elle fait appel à un microcalorimètre, un appareil capable de détecter des variations de température de l'ordre du millionième de degré Celsius. Expérimentalement, l'appareil contient deux cellules. Une cellule de référence, qui sert de masse thermique et qui ne contient que du tampon, et une cellule test, qui contient la protéine d'intérêt. Cette dernière cellule est titrée avec des concentrations croissantes de ligand, jusqu'à ce que l'équilibre soit atteint et qu'il n'y ait plus de variation de température. L'analyse de la variation différentielle de la chaleur émise ou absorbée en fonction de la concentration de ligand nous permettra de rapidement confirmer la constante d'association UL54-ADN. L'intégration de l'aire sous la courbe de chaque injection nous permettra de résoudre et de confirmer les paramètres thermodynamiques de la liaison : énergie

libre de Gibbs (ΔG), enthalpie (ΔH) et entropie (ΔS). Enfin, l'analyse des résultats nous permettra aussi de résoudre la stœchiométrie de l'association. L'analyse de cette dernière information pourrait aussi supporter nos résultats d'affinité, car si on peut prouver que le ratio de liaison UL54:ADN est bien 1:1, on confirmerait du même coup que la liaison est spécifique et que l'ADN ne se lie pas de manière non spécifique n'importe où sur UL54.

12.1.1.3) Exigence de la microcalorimétrie

La microcalorimétrie requiert une quantité relativement grande de protéines, ce dont nous disposons, mais cependant, à des concentrations légèrement supérieures à celle présentement à notre disposition. Nous avons en inventaire plus que la quantité de protéines totale nécessaire, mais les concentrations présentement à notre disposition varient entre $2\mu\text{M}$ et $5\mu\text{M}$, alors que la microcalorimétrie requiert des lots de protéines concentrés entre $8\mu\text{M}$ et $20\mu\text{M}$ (Meier-Andrejszki, Bjeli et al. 2007). Nous devons donc légèrement reconcentrer UL54 par un faible facteur de 2 à 5. Pour ce faire, nous suggérons deux approches potentielles. La première serait d'utiliser la lyophilisation, qui tire partie de la sublimation du solvant d'un échantillon gelé dans une chambre à basse pression. Nous avons d'ailleurs à notre disposition un lyophilisateur (Labconco, FreeZone 1 Liter Benchtop Freeze Dry Systems 7740020). Cependant, un des problèmes potentiels avec la lyophilisation est que, parfois, la structure des protéines peut changer lors de la sublimation de l'eau. Un essai préalable de dichroïsme circulaire pourrait nous assurer de l'intégrité de la structure tertiaire de notre protéine. Une seconde approche serait d'utiliser des colonnes concentratrices par centrifugation, comme les Amicon Ultra-4 50,000 NMWL. Leur

principe d'utilisation est aussi fort simple : lors de la centrifugation, le filtre retient les protéines de poids moléculaire supérieur à 50 kDa, alors qu'une partie du solvant le traverse, ce qui concentre les protéines. Pour ce qui est des concentrations d'ADN requises, nous disposons d'ADN en quantité et en concentration bien plus que suffisantes pour satisfaire aux expériences proposées.

12.1.2) Relâche de molécules d'eau au solvant

12.1.2.1) Contexte

La liaison de UL54 à l'ADN est régie par une forte entropie très favorable, ainsi qu'une enthalpie légèrement défavorable. Le patron de liaison généralement associé à un tel profil énergétique est généralement régi par 1) la relâche de molécules d'eau et de contre-ions vers le solvant; 2) l'importance des interactions hydrophobes et d'empilements aromatiques et 3) une très faible contribution des interactions électrostatiques à la liaison. Une contribution entropique peut aussi être associée à un changement de conformation. Nos résultats de dichroïsme circulaire semblent indiquer qu'un changement de conformation survient chez UL54 suite à sa liaison avec l'ADN. Un tel changement de conformation a d'ailleurs été répertorié chez de nombreuses polymérases de la famille B, où le domaine du pouce se referme légèrement sur l'ADN suivant sa liaison. Nous avons aussi mis en évidence que la force ionique n'affectait à peu près pas l'équilibre UL54-ADN, ce qui supporte la prémisse d'une très faible contribution des interactions électrostatiques à la liaison quantitativement; elles ne représentent que 6% de l'énergie de liaison. Afin d'étoffer notre modèle, il serait particulièrement intéressant de confirmer et même de quantifier

la relâche prédite de molécules d'eau vers le solvant lors de la liaison de la protéine avec son substrat. L'approche que je propose serait la suivante : déterminer la dépendance des paramètres thermodynamiques pour la molalité d'un soluté neutre (par exemple du glycérol) lors de la liaison de UL54 à L'ADN.

12.1.2.2) Quantification de la relâche de molécules d'eau au solvant

Cette expérience est basée sur le principe, clairement démontré dans la littérature, que l'énergie libre d'une liaison, impliquant la relâche d'eau au solvant, présente une dépendance linéaire à l'égard de la concentration osmolaire en soluté neutre (Mitchell et Litman 1999; Swaminathan, Nandi et al. 1999). Cette dépendance est due à l'exclusion de ce soluté, l'eau entourant la surface d'interaction protéine-ADN. La magnitude de l'effet de ce stress osmotique dépend à la fois de l'osmolalité du soluté, qui est équivalent au potentiel chimique de l'eau brute, et de la différence entre le nombre de molécules d'eau exclues du soluté qui sont liées au complexe protéine-ADN et celles liées à la protéine libre et à l'ADN libre. L'impact de l'influence de l'osmolalité sur l'équilibre entre l'association et la dissociation du complexe protéine-ADN (K_A) est donné par l'équation A (Mitchell et Litman 1999), où la variation du logarithme naturel de l'affinité en fonction de la variation de l'osmolalité est état de $\Delta n_w / 55.56$, soit le changement dans le nombre de molécules d'eau exclues du soluté lors de la liaison (l'activité de l'eau a_w) par kilogramme d'eau (1kg d'eau = 55.56 moles).

$$d \ln(K_A) / d [\text{osmolalité}] = - \Delta n_w / 55.56 \quad \text{Équation A}$$

En portant graphiquement le logarithme naturel de la constante d'association UL54-ADN en fonction de la molalité en soluté neutre de la solution, on s'attend à obtenir une fonction linéaire.

Une pente positive nous indiquerait qu'il y a bel et

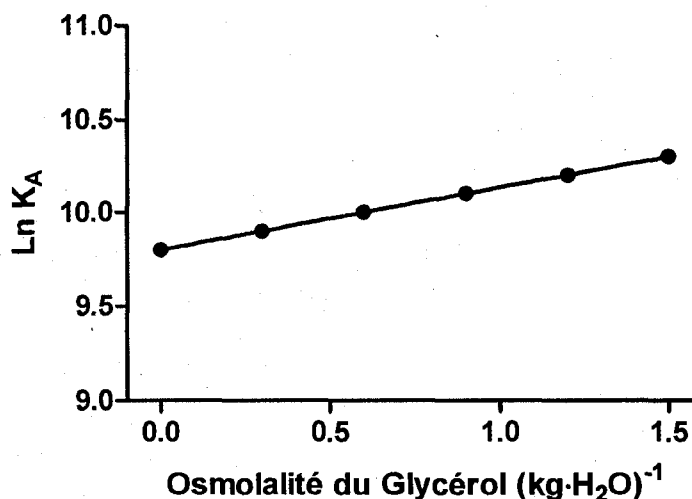


Figure 12 : Exemple de résultat attendu lors d'un test visant à quantifier la relâche de molécules d'eau vers le solvant. La liaison entre UL54 et l'ADN serait affectée par la molalité en soluté neutre.

bien relâche de molécules d'eau vers le solvant lors de la liaison. L'amplitude de cette pente serait proportionnelle au nombre de molécules d'eau relâchées, ce qui nous permettrait ainsi de confirmer et de quantifier le nombre de molécules d'eau relâchées au solvant lors de la liaison de UL54 à l'ADN. Un exemple de résultat attendu est fourni (Figure 12).

12.1.2.3) Procédure expérimentale

Je suggère l'emploi du glycérol et de l'éthylène glycol comme solutés neutres, car ce sont les deux composés les plus cités dans la littérature pour remplir cette fonction. De plus, l'emploi de deux solutés neutres dans le cadre de deux expériences indépendantes nous fournira un duplicata, augmentant ainsi la crédibilité de la technique. La détermination de la constante d'association UL54-ADN pour chacune des molalités de soluté neutre pourra s'effectuer à l'aide de la spectroscopie à fluorescence.

12.1.3) Discrimination entre l'ADN et l'ARN par UL44

12.1.3.1) Contexte

Nous avons démontré dans l'article 1 que la protéine UL54 n'était pas capable de discriminer entre un substrat d'ADN et un d'ARN. Cette propriété peut sembler étonnante, mais bien qu'elle ne soit pas la norme, elle n'est pas unique non plus. La nature nous fournit plusieurs exemples d'ADN-polymérase-ADN-dépendantes capables de lier l'ARN. En fait, certaines, comme la Tth pol (la polymérase de *Thermus thermophilus*), sont même catalytiquement actives avec un ARN comme substrat (Myers et Gelfand 1991). L'inverse étant aussi vrai, certaines ARN-polymérase-ARN-dépendantes peuvent démontrer une activité en présence d'un substrat d'ADN, la polymérase NS5B du VHC étant un bon exemple (Behrens, Tomei et al. 1996). Les hypothèses possibles visant à expliquer cette faible spécificité sont nombreuses. Une première stipule que si les pressions sélectives visant la conservation de cette spécificité sont très faibles ou absentes, la protéine perdra sa spécificité. Une réplication dans un sous-environnement très contrôlé, où la protéine n'est en présence que d'ADN ou d'ARN, pourrait être un exemple d'une pression sélective faible. Une seconde possibilité serait que cette activité de discrimination soit assumée par un cofacteur.

12.1.3.2) Choix de la protéine UL44

Nous proposons de vérifier si le cofacteur de processivité UL44 peut supporter une discrimination entre l'ADN et l'ARN. Le choix de cette protéine, parmi celles du complexe de réplication, s'explique par les faits qu'elle soit connue pour être une

protéine de liaison à l'ADN double brin, qu'elle interagisse directement avec UL54, et surtout que sa fonction principale soit d'ancrer UL54 sur son substrat. Serait-elle responsable d'ancrer UL54 sur son bon substrat, l'ADN?

12.1.3.3) État des connaissances sur UL44

Cette protéine a déjà été partiellement caractérisée. Elle lie l'ADNdb avec une affinité de l'ordre du nanomolaire Cette liaison est séquence indépendante, mais elle arbore une préférence pour l'ADN double brin versus l'ADN simple brin (Loregian, Sinigalia et al. 2007). Sa cristallisation révèle qu'elle se dimérise pour adopter une structure quaternaire en forme de fer à cheval, au centre de laquelle elle lie l'ADNdb (Appleton, Loregian et al. 2004), ce qui est consistant avec sa stochiométrie de liaison UL44:ADN de 2:1 (Loregian, Sinigalia et al. 2007). Cette liaison est associée à une enthalpie moyennement défavorable et est, par conséquent, propulsée par l'entropie. Cette forte entropie est le fruit de la relâche de cations vers le solvant lors de la liaison (Loregian, Sinigalia et al. 2007). Ce modèle est supporté par les données structurales, qui semblent indiquer que de nombreux résidus aspartates de UL44 contactent les phosphates du phosphosquelette de l'ADNdb via des interactions charges-charges (Appleton, Loregian et al. 2004), ce qui est consistant avec la forte influence négative de la force ionique du milieu sur la liaison UL44-ADNdb (Loregian, Sinigalia et al. 2007). Cependant, aucun groupe de recherche ne s'est attardé à comparer l'affinité de UL44 pour l'ADN et l'ARN.

12.1.3.4) Procédure expérimentale

Nous avons déjà exprimé la protéine recombinante UL44 en bactérie, et nous l'avons purifié par l'ajout d'un épitope hexahistidine (FIGURE 13). Comme la protéine possède un tryptophane (Davison, Dolan et al. 2003), et que nous en disposons d'une grande quantité hautement purifiée, l'emploi de la spectroscopie à fluorescence semble tout indiqué pour quantifier l'affinité de UL44 à l'ARN et à l'ADN. La quantification de la variation de fluorescence d'une concentration fixe de protéines titrées, avec une concentration croissante d'ADNsb, d'ADNdb, d'ARNsb et d'ARNdb, nous permettrait de calculer

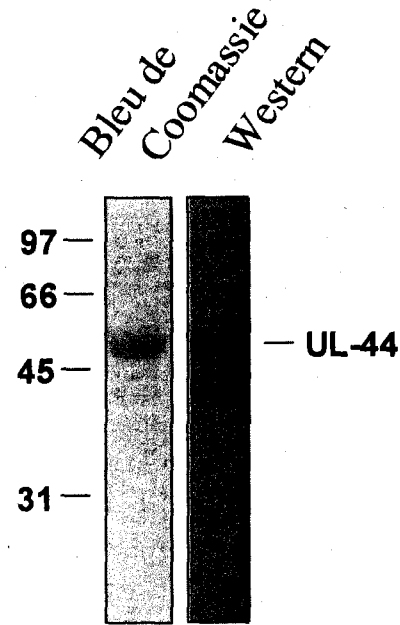


Figure 13 : Protéine UL44 du cytomégalo virus humain. Électrophorèse sur gel de polyacrylamide de UL44 exprimé en bactérie et purifié grâce à un épitope exahistidine. Coloration au bleu de Coomassie (gauche) et immunobuvardage de type western (droite)

l'affinité de UL44 pour chacun de ces substrats. Pour chacun des substrats, nous utiliserions la même séquence de bases. Un test de bio-informatique préalable nous permettrait de sélectionner des séquences qui n'adoptent pas de structure secondaire. Cette technique serait rapide et efficace afin de savoir si UL44 est capable de discriminer l'ADN de l'ARN.

12.2) Perspectives à long terme

12.2.1) Caractériser la fidélité de UL54

12.2.1.1) Contexte

La protéine UL54, bien qu'étant l'enzyme clé de la réplication du CMV, est relativement peu caractérisée pour une polymérase virale. La présente étude palie en partie à cette lacune, cependant, il n'en demeure pas moins que de nombreuses propriétés de cette enzyme demeurent encore inconnues. Une des caractéristiques qu'il serait particulièrement intéressant d'étudier est sa fidélité. Paradoxalement, alors que tant d'efforts sont consacrés sur les conséquences de son infidélité, c'est-à-dire l'apparition de mutations conférant des résistances au traitement, très peu est connu sur sa fidélité.

12.2.1.2) Fidélité des ADN polymérases de la famille B

Brièvement, la fidélité est un mécanisme particulièrement complexe qui permet à l'enzyme de pratiquement toujours incorporer le bon substrat sur la matrice d'ADN, alors que ce substrat change potentiellement à chaque ronde de catalyse. Cette précision est le fruit d'une sommation de points de contrôle. Ainsi, un mauvais nucléotide d'une plus faible affinité entrave les changements de conformation, son positionnement n'est pas optimal pour la catalyse, et son incorporation ralentit la polymérase. Advenant une incorporation erronée, les polymérases de la famille B possèdent aussi un domaine exonucléase 3'-5' leur permettant d'exciser le mauvais nucléotide.

12.2.1.3) Objectifs

Lors de cette étude, je propose la caractérisation de la fidélité sous plusieurs angles. D'une part, l'étude des mauvaises incorporations, et d'autre part, l'activité

exonucléase de UL54. Il serait aussi intéressant de regarder si la fidélité de la polymérase diffère lors de l'initiation par rapport à l'élongation.

12.2.1.4) UL54-D413A (exo-KO)

Lors d'études précédentes sur le domaine exonucléase du virus de l'herpès simplex (VHS), il a été démontré que les activités polymérase et exonucléases étaient indépendantes. En effet, une ou plusieurs mutations précises au sein du domaine exonucléase permettaient de l'inactiver, tout en conservant de 80% à 99% de l'activité polymérase intacte. Cette activité exonucléase s'est d'ailleurs avérée importante, mais non létale, car *in vivo*, une polymérase exonucléase déficiente peut supporter une réplication virale affaiblie (Hall, Orth et al. 1995; Hwang, Liu et al. 1997). Chez UL54, un cas clinique particulier, arborant une résistance au Gancicovir et au Cidofovir, mais pas au foscarnet, présentait une mutation dans le domaine II du domaine exonucléase de UL54, plus particulièrement une mutation de l'aspartate 413 en alanine (D413A). Cette mutation inactive l'activité exonucléase, mais ne semble pas affecter l'activité polymérase. De plus, les virus portant cette mutation ne présentent pas une réplication gravement compromise (Chou et Marousek 2007). Par conséquent, la protéine UL54-D413A représente un outil particulièrement intéressant pour l'étude de la fidélité, car elle permet de dissocier les mauvaises incorporations de la correction (proofreading).

12.2.1.5) Études précédentes sur la fidélité des polymérase

Afin de produire une caractérisation détaillée de la fidélité de UL54, nous proposons de se baser sur une technique éprouvée qui a fait ses preuves lors de la caractérisation

de nombreuses polymérase, dont celle des mitochondries humaines, et la polymérase η de *Saccharomyces cerevisiae*. Il s'agit de l'analyse cinétique d'incorporation de désoxynucléotides *in vitro* en condition d'excès de substrat (Washington, Johnson et al. 1999; Lee et Johnson 2006). Cette approche permet de déterminer le K_D , le k_{cat} et la spécificité de la polymérase pour chacune des 16 combinaisons de paire de bases (dATP:dT, dTTP:dT, dATP:dG ...). Cela nous renseigne non seulement sur le taux de mutations, mais aussi sur celles qui sont le plus probables de survenir.

12.2.1.6) Procédure expérimentale : fidélité globale

Expérimentalement, il s'agit d'incuber la protéine recombinante dans un tampon supportant l'activité polymérase (75mM Tris-HCl pH 8.0, 6.5mM MgCl₂, 0.5mM DTT, 90mM NaCl (Cihlar, Fuller et al. 1997; Loregian, Rigatti et al. 2003)), et en présence d'un excès d'ADN double brin comportant une extrémité franche et une extrémité 5' saillante. Ces matrices d'ADN sont conçues de manière à ce que le premier nucléotide

simple brin soit celui sur lequel on veut tester l'incorporation (dATP:dG). Il est suivi de 60 nucléotides complémentaires à celui que l'on désire incorporer (dATP:dG) (FIGURE 14 A). On ajoute ensuite une concentration croissante du désoxynucléotide

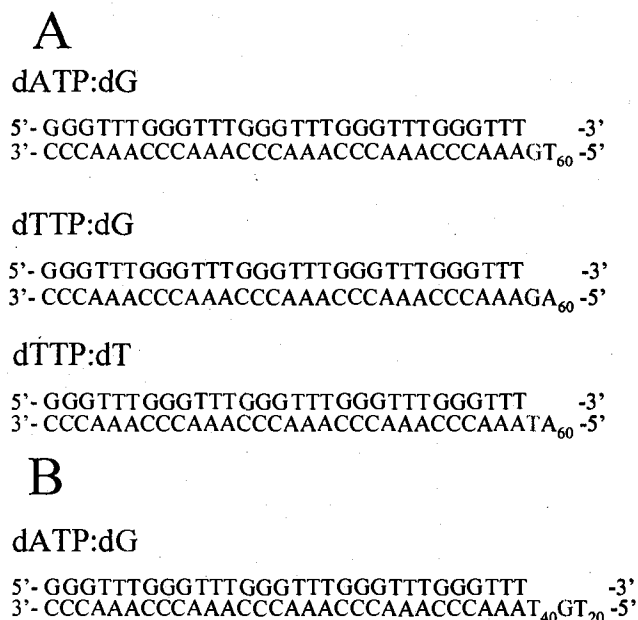


Figure 14 : Construction des matrices d'ADN utilisés pour les essai de fidelité de UL54. A) de haut en bas : matrice pour incorporation d'un dATP sur une guanine, d'un dTTP sur une guanine et d'un dTTP sur une tyrosine. B) matrice pour incorporation d'un dATP sur une guanine lorsque UL54 est en élongation.

dont on veut tester l'incorporation, que l'on incube à 37⁰C pour une période de temps fixe. Ainsi, pour un bon nucléotide, il y aura polymérisation simple, alors que pour un mauvais nucléotide, l'incorporation étant beaucoup moins favorisée, les concentrations requises attendues seront beaucoup plus grandes. Cependant, suite à cette mauvaise incorporation (ex. dATP sur G), la matrice est conçue de telle sorte que les 60 incorporations suivantes seront favorisées (ex. dATP sur T). Si une mauvaise incorporation survient, et qu'elle n'est pas immédiatement excisée, elle sera suivie de 60 bonnes incorporations, ce qui protégera cette mauvaise incorporation d'une exonucléation non couplée à la transcription.

Pour la détection quantitative, nous proposons un marquage préalable en 5' du brin amorce par phosphorylation à l'ATP_γ³²P. Les produits de polymérisation pourront ensuite être résolus par électrophorèse sur gel de polyacrylamide dénaturant. La **Figure 15** présente un résumé du schéma expérimental.

Advenant une résolution insatisfaisante, un protocole alternatif de détection pourrait faire appel à une précipitation suivie d'un compte à scintillation, et/ou à un marquage radioactif par incorporation de dNTP α ³²P.

La protéine UL54 sauvage (UL54WT) possède un domaine exonucléase actif, et les probabilités qu'une mauvaise incorporation soit immédiatement enlevée sont grandes. Ne fournissant qu'un seul nucléotide à la polymérase à cette position, elle n'a d'autre choix que de faire une erreur et de continuer ou de s'arrêter. Pour que la polymérisation continue, il faut donc 1) que la polymérase incorpore le mauvais nucléotide; 2) que le domaine exonucléase ne l'excise pas immédiatement; 3) que la

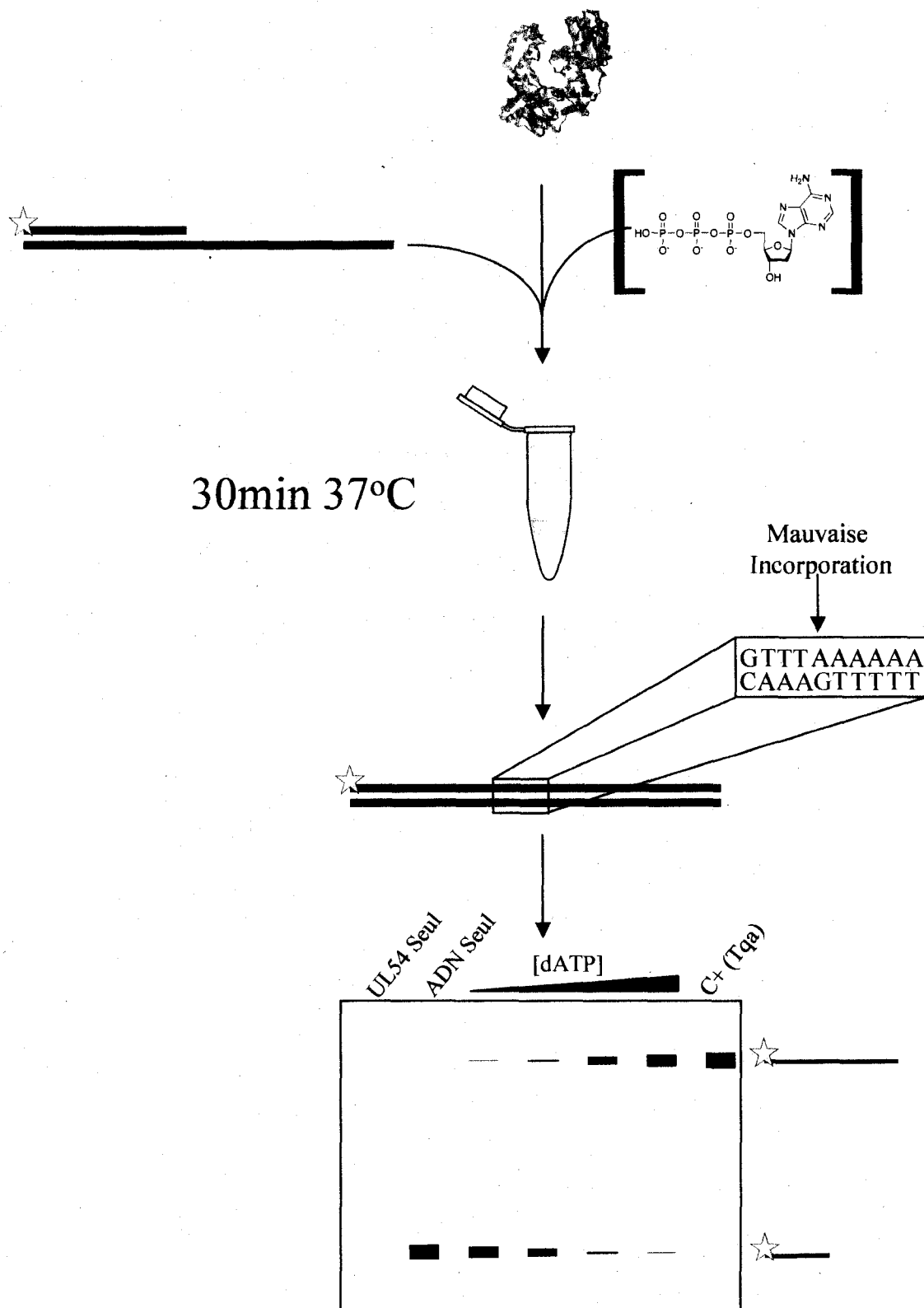


Figure 15 : Schéma expérimental visant à déterminer la fidélité de UL54. De haut en bas : l'ADNdb radiomarqué est incubé avec une concentration constante d'enzyme et une concentration croissante de dNTP. La réaction est arrêtée après 30 minutes d'incubation à 37°C. L'élongation du brin amorce radiomarqué est ensuite résolue par électrophorèse sur gel de polyacrylamide dénaturant.

polymérisation continue en incorporant le bon nucléotide suivant, après quoi toutes les incorporations subséquentes seront favorables. Les résultats de cette expérience dépeindront la fidélité d'ensemble de UL54, c'est-à-dire le taux final de mutations que la polymérase fait (mauvaise incorporation moins correction).

12.2.1.7) Procédure expérimentale : fidélité brute

Afin de départager l'importance relative des mauvaises incorporations par rapport aux corrections, nous proposons l'utilisation de UL54 comportant une mutation de l'aspartate 413 en alanine (UL54D413A). Cette mutation est connue pour abolir l'activité exonucléase, tout en conservant l'activité polymérase (Chou et Marousek 2007). Nous proposons de refaire les mêmes séries d'expériences avec UL54D413A. Cette seconde série de résultats nous indiquera le taux d'erreurs brut de UL54, car ces erreurs ne pourront plus être enlevées.

12.2.1.8) Analyse des résultats

Après avoir mit ces résultats en perspective, nous serons en mesure de dresser un portrait complet de la fidélité de UL54, sans biais de sélection naturelle. Nous pourrons quantifier les K_D et les k_{cat} pour chacune des 16 combinaisons de paire de bases. En les relativisant par rapport à la bonne incorporation, nous allons pouvoir déterminer la spécificité de la polymérase pour chacune. Ainsi, nous serons en mesure de prédire la probabilité d'apparition de chacune des douze mutations possibles. Les résultats générés avec UL54D413A nous indiqueront quelles erreurs d'incorporation sont le plus souvent commises par la polymérase. De plus, en comparant les résultats de UL54WT avec ceux de UL54D413A, nous aurons une idée des mauvaises

incorporations les plus souvent réparées, et celles les plus souvent ignorées. L'analyse de ces résultats nous fournira des informations importantes sur la nature de la fidélité de UL54. Nous serons à même de quantifier les différences de fidélité entre les purines et les pyrimidines, par exemple, ce qui pourrait épauler les recherches de composés de deuxième génération basés sur le Ganclivir (analogue de purine) ou sur le Cidofovir (analogue de pyrimidine).

12.2.1.9) Fidélité de UL54 en mode élongation

Cette conception expérimentale représente une adaptation du test de fidélité classique. Il serait cependant intéressant de pousser le concept un peu plus loin, en regardant si le stade (initiation et élongation) où la polymérase se trouve a une influence sur sa fidélité. Pour ce faire, nous proposons d'utiliser le même principe expérimental, avec une matrice d'ADN légèrement différente, où la base sujette à une mauvaise incorporation n'est plus en première position, mais en quarantième position (**FIGURE 14 B**). Nous pourrions ainsi voir si en phase d'élongation, à vitesse maximale, UL54 présente un patron de fidélité différent.

Il serait aussi facile d'adapter la technique pour tester la cinétique d'incorporation d'analogues de nucléotides déjà connus, ainsi que l'importance relative de leur endonucléation.

12.2.1.10) Solution pour augmenter la sensibilité (si nécessaire)

Nos essais préliminaires indiquent que la sous-unité catalytique seule (UL54²⁹³⁻⁹⁵³) démontre une activité polymérase. Cette activité devrait être assez importante pour les tests proposés. Cependant, advenant que la sensibilité de notre méthode requiert une

activité catalytique plus forte, il serait possible d'utiliser son facteur de processivité (UL44) afin de stimuler son activité. Nous disposons déjà de la protéine recombinante UL44 exprimée en bactérie et purifiée par l'ajout d'un épitope hexahistidine (**FIGURE 13**). Le domaine d'interaction de UL54 pour UL44 est cartographié et relativement bien caractérisé. Les 22 derniers acides aminés carboxi-terminaux (1220-1242) de UL54 ont été démontrés comme étant essentiels et suffisants pour l'interaction avec UL44 (Appleton, Loregian et al. 2004; Loregian, Appleton et al. 2004; Appleton, Brooks et al. 2006; Loregian, Sinigalia et al. 2007).

12.2.1.10.1) Expression de UL54²⁹³⁻¹²⁴²

La protéine UL54²⁹³⁻⁹⁵³ dont nous disposons ne possède pas ce domaine. Nous pourrions cependant cloner, exprimer et purifier la protéine UL54²⁹³⁻¹²⁴², pour générer une protéine de 106 kDa, ce qui est dans les limites possibles de notre système d'expression. D'ailleurs, nous avons déjà exprimé des protéines d'une telle taille au laboratoire (données non publiées). Il serait aussi important de positionner l'épitope histidine en N-terminal (comme pour UL54²⁹³⁻⁹⁵³), afin de s'assurer qu'il n'interfère pas avec l'interface de dimérisation.

12.2.1.10.2) Expression de UL54^{293-953+22C}

Une seconde approche plus audacieuse pourrait aussi être envisagée. En sachant qu'un peptide de 22 acides aminés comportant la séquence C-terminal de UL54 est capable de lier fortement UL44, et que le rôle de facteur de processivité de UL44 est médié par sa fonction d'encre sur l'ADN, il serait possible, en se fiant sur le modèle par homologie pour les positionner dans la même région, de fusionner, un peu comme un

épitope, les 22 acides aminés carboxiterminaux aux acides aminés 293 à ≈953 de UL54 (UL54^{293-953+22C}). On générerait ainsi une protéine plus courte, mais somme toute capable de lier UL44 et de s'ancrer à l'ADN pour être plus processive.

13) Remerciements

J'aimerais remercier mon directeur de recherche, Dr. Martin Bisaillon, pour m'avoir offert la possibilité d'accomplir ma maîtrise dans son laboratoire ainsi que pour sa disponibilité, sa compréhension et son génie. Je remercie aussi notre assistante de recherche dévouée, Isabelle Bougie, pour son soutien technique, pour ses réponses à mes nombreuses questions et pour tout le matériel qu'elle m'a commandé. Merci aussi aux étudiants avec qui j'ai partagé ces dernières années. Merci à Marie Soulière pour les conseils scientifiques et pour toutes les gâteries qu'elle nous a généreusement partagées. Merci à Nishal pour les conseils scientifiques et les notions de culture générale qui a su transmettre. Enfin, un gros merci à toute ma famille pour le soutien inconditionnel qu'ils m'ont apporté et pour m'avoir encourager et m'avoir valorisé dans ce que j'aimais faire.

14) Bibliographie

- Anders, D. et W. Gibson (1988). Location, transcription analysis, and partial nucleotide sequence of the cytomegalovirus gene encoding an early DNA-binding protein with similarities to ICP8 of herpes simplex virus type 1. *J Virol* **62**: 1364-72.
- Anders, D., M. Kacica, et al. (1992). Boundaries and Structure of Human Cytomegalovirus oriLyt, a Complex Origin for Lytic-Phase DNA Replication. *J Virol* **66**: 3373-84.
- Anders, D., J. Kerry, et al. (2006). DNA synthesis and late viral gene expression. *Human Herpesviruses: Biology, Therapy and Immunoprophylaxis* **1**: 292-307.
- Anders, D. et L. McCue (1996). The human cytomegalovirus genes and proteins required for DNA synthesis. *Intervirology* **39**: 378-88.
- Appleton, B., J. Brooks, et al. (2006). Crystal Structure of the Cytomegalovirus DNA polymerase Subunit UL44 in Complex with the C-Terminus from the Catalytic Subunit: differences in structure and function relative to unliganded UL44. *J Biol* **281**: 5224-32.
- Appleton, B., A. Loregian, et al. (2004). The cytomegalovirus DNA polymerase subunit UL44 forms a C clamp-shaped dimer. *Mol Cell* **15**: 233-44.
- Baek, M., P. Krosky, et al. (2004). Phosphorylation of the RNA polymerase II carboxyl-terminal domain in human cytomegalovirus-infected cells and in vitro by the viral UL97 protein kinase. *Virology* **324**: 184-93.
- Baek, M.-C., P. Krosky, et al. (2002). Specific phosphorylation of exogenous protein and peptide substrates by the human cytomegalovirus UL97 protein kinase: importance of the P+5 position. *J Biol Chem* **277**: 29593-9.
- Bebenek, A., G. T. Carvera, et al. (2002). Dissecting the Fidelity of Bacteriophage RB69 DNA Polymerase: Site-Specific Modulation of Fidelity by Polymerase Accessory Proteins. *Genetics* **162**: 1003-8.
- Beese, L. et T. Steitz (1991). Structural basis for the 3'-5' exonuclease activity of *Escherichia coli* DNA polymerase I: a two metal ion mechanism. *EMBO J* **10**: 25-33.
- Bego, M. et S. St-Jeor (2006). Human cytomegalovirus infection of cells of hematopoietic origin: HCMV-induced immunosuppression, immune evasion, and latency. *Experimental hematology* **34**: 555-70.
- Behrens, S. E., L. Tomei, et al. (1996). Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. *EMBO J* **15**: 12-22.

- Benzaghoul, I., I. Bougie, et al. (2004). Effect of Metal Ion Binding on the Structural Stability of the Hepatitis C Virus RNA Polymerase. *J Biol Chem* **48**: 49755-61.
- Biron, K. (2006). Antiviral drugs for cytomegalovirus diseases. *Antiviral Research* **71**: 154-63.
- Boeckh, M., C. Holy, et al. (1998). Occult Cytomegalovirus infection of marrow stroma. *Clin Infect Dis* **26**: 209-10.
- Boucher, P. D., L. J. Ostruszka, et al. (2000). Synergistic Enhancement of Herpes Simplex Virus Thymidine Kinase/Ganciclovir-mediated Cytotoxicity by Hydroxyurea. *Cancer Research* **60**: 1631-6.
- Brautigam, C. et T. Steitz (1998). Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes. *Curr Opin Struct Biol* **8**: 54-63.
- Brautigam, C. et T. Steitz (1998). Structural basis for the 3'-5' exonuclease activity of Escherichia coli DNA polymerase I: a two metal ion mechanism. *J Mol Biol* **27**: 363-77.
- Britt, W., A. Arvin, et al. (2006). Maturation and egress. *Human Herpesviruses: Biology, Therapy and Immunoprophylaxis* **1**: 308-20.
- Browne, E. P., B. Wing, et al. (2001). Altered Cellular mRNA Levels in Human Cytomegalovirus-Infected Fibroblasts: Viral Block to the Accumulation of Antiviral mRNAs. *J Virol* **75**: 12319-12330.
- Capson, T., JAPeliska, et al. (1992). Kinetic characterization of the polymerase and exonuclease activities of the gene 43 protein of bacteriophage T4. *Biochemistry* **31**: 10984-94.
- Chandler, D. (2005). Interfaces and the driving forces of hydrophobic assembly. *Nature* **437**: 640-7.
- ChimerixInc. (2007). <http://www.chimerix-inc.com>.
- Chothia, C. et J. Janin (1975). Principles of protein-protein recognition. *Nature* **256**: 705-8.
- Chothra, C. (1984). Principles that Determine the Structure of Proteins. *Ann. Rev. Biochem* **53**: 537-72.
- Chou, S. et G. I. Marousek (2007). Accelerated Evolution of Maribavir Resistance in a Cytomegalovirus Exonuclease Domain II Mutant. *J Virol* **82**: 246-53.

- Cihlar, T., M. Fuller, et al. (1998). Characterization of drug resistance-associated mutations in the human cytomegalovirus DNA polymerase gene by using recombinant mutant viruses generated from overlapping DNA fragments. *J Virol* **72**: 5927-36.
- Cihlar, T., M. D. Fuller, et al. (1997). Expression of the Catalytic Subunit (UL54) and the Accessory Protein (UL44) of Human Cytomegalovirus DNA Polymerase in a Coupled *In Vitro* Transcription/Translation System. *Protein Expression and Purification* **11**: 209-218.
- Colletti, K., K. Smallenburg, et al. (2007). Human cytomegalovirus UL84 interacts with an RNA stem-loop sequence found within the RNA/DNA hybrid region of oriLyt. *J Virol* **81**: 7077-85.
- Copeland, W., N. Lam, et al. (1993). Fidelity studies of the human DNA polymerase alpha. The most conserved region among alpha-like DNA polymerases is responsible for metal-induced infidelity in DNA synthesis. *J Biol Chem* **268**: 11041-9.
- Crumpacker, C. (1992). Mechanism of action of foscarnet against viral polymerases. *Am. J. Med* **92**: 3S-7S.
- Davison, A., A. Dolan, et al. (2003). The human cytomegalovirus genome revisited: comparison with the chimpanzee cytomegalovirus genome. *J Gen Virol* **84**: 17-28.
- Delaage, M. (1991). Physico-Chemical Aspects of Molecular Recognition. *Molecular Recognition Mechanisms* **1**: 1-13.
- Delarue, M., O. Poch, et al. (1990). An attempt to unify the structure of polymerases. *Protein Eng* **3**: 461-7.
- Doronin, S., O. Lavrik, et al. (1987). The efficiency of dNTP complex formation with human placenta DNA polymerase alpha as demonstrated by affinity modification. *FEBS Lett.* **216**: 221-4.
- Doublie, S., S. Tabor, et al. (1998). Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature* **391**: 251-8.
- Ertl, P. et K. Powell (1992). Physical and functional interaction of human cytomegalovirus DNA polymerase and its accessory protein. *J Virol* **66**: 4126-33.
- Franklin, M., J. Wang, et al. (2001). Structure of the Replicating Complex of a pol a Family DNA Polymerase. *Cell* **105**: 657-67.
- Freisinger, E., A. Grollman, et al. (2004). Lesion (in)tolerance reveals insights into DNA replication fidelity. *EMBO J* **23**: 1494-505.
- Gandhi, M. et R. Khanna (2004). Human cytomegalovirus: clinical aspects, immune regulation, and emerging treatments. *Infectious Diseases* **4**: 725-38.

Gao, Y., K. Colletti, et al. (2008). Identification of human cytomegalovirus UL84 virus- and cell-encoded binding partners by using proteomics analysis. *J Virol* **82**: 96-104.

Garcia-Diaza, M., K. Bebeneka, et al. (2007). Role of the catalytic metal during polymerization by DNA polymerase lambda. *DNA Repair* **6**: 1333-40.

Gerna, G., F. Baldanti, et al. (2004). Pathogenesis of Human Cytomegalovirus Infection and Cellular Targets. *Human Immunology* **65**: 381-6.

Gilbert, C. et G. Boivin (2005). Human cytomegalovirus resistance to antiviral drugs. *Antimicrob. Agents Chemother* **49**: 873-83.

Goodrum, F. D., C. T. Jordan, et al. (2002). Human cytomegalovirus gene expression during infection of primary hematopoietic progenitor cells: a model for latency. *Proc Natl Acad Sci U S A* **99**: 16255-60.

GREGORY, H. et M. D. TAYLOR (2003). Cytomegalovirus. *American Family Physician* **67**: 519-24.

Grillone, L. et R. Lanz (2001). Fomivirsen. *Drugs Today (Barc)* **37**: 245-55.

Gulbis, J., Z. Kelman, et al. (1996). Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA. *Cell* **18**: 297-306.

Hall, J. D., K. L. Orth, et al. (1995). Mutations within conserved motifs in the 3'-5' exonuclease domain of herpes simplex virus DNA polymerase. *Journal of General Virology* **76**: 2999-3008.

Heilbronn, R., G. Jahn, et al. (1987). Genomic localization, sequence analysis, and transcription of the putative human cytomegalovirus DNA polymerase gene. *J Virol* **61**: 119-24.

Hertel, L. et E. S. Mocarski (2004). Global Analysis of Host Cell Gene Expression Late during Cytomegalovirus Infection Reveals Extensive Dysregulation of Cell Cycle Gene Expression and Induction of Pseudomitosis Independent of US28 Function. *J Virol* **78**: 11988-12011.

Hirai, K. et Y. Watanabe (1976). Induction of alpha type DNA polymerases in human cytomegalovirus-infected WI-38 cells. *Biochim Biophys Acta* **447**: 328-39.

Ho, E. S., D. C. Lin, et al. (2000). Mendel and T. Cihlar, Cytotoxicity of antiviral nucleotides adefovir and cidofovir is induced by the expression of human renal organic anion transporter. *J. Am. Soc. Nephrol* **11**: 383-93.

- Hogg, M., S. Wallace, et al. (2004). Crystallographic snapshots of a replicative DNA polymerase encountering an abasic site. *EMBO J* **23**: 1483-93.
- HUANG, E.-S. (1975). Human Cytomegalovirus III. Virus-Induced DNA Polymerase. *J Virol* **165**: 298-310.
- Hwang, Y. T., B. Y. Liu, et al. (1997). Effects of mutations in the Exo III motif of the herpes simplex virus DNA polymerase gene on enzyme activities, viral replication, and replication fidelity. *J Virol* **71**: 7791-8.
- Ito, J. et D. Braithwaite (1991). Compilation and alignment of DNA polymerase sequences. *Nucleic Acids Res* **19**: 4045-57.
- Joyce, C. M. et S. J. Benkovic (2004). DNA Polymerase Fidelity: Kinetics, Structure, and Checkpoints. *Biochemistry* **43**: 14317-24.
- Jung, G., M. Leavitt, et al. (1990). Site-specific mutagenesis of PRD1 DNA polymerase: mutations in highly conserved regions of the family B DNA polymerase. *Biochem Biophys Res Commun* **170**: 1294-300.
- Kahl, M., D. Sierge-Axel, et al. (2000). Efficient lytic infection of Human arterial Endothelial Cells by Human Cytomegalovirus strains. *J Virol* **74**: 7628-35.
- Kemble, G., A. McCormick, et al. (1987). A cytomegalovirus protein with properties of herpes simplex virus ICP8: partial purification of the polypeptide and map position of the gene. *J Virol* **61**: 3143-51.
- Kim, Y., S. Eom, et al. (1995). Crystal structure of *Thermus aquaticus* DNA polymerase. *Nature* **376**: 612-6.
- Kohlstaedt, L., J. Wang, et al. (1992). Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* **256**: 1783-90.
- Kong, X., R. Onrust, et al. (1992). Three-dimensional structure of the beta subunit of *E. coli* DNA polymerase III holoenzyme: a sliding DNA clamp. *Cell* **69**: 425-37.
- Korolev, S., M. Nayal, et al. (1995). Crystal structure of the large fragment of *Thermus aquaticus* DNA polymerase I at 2.5 Å resolution: structural basis for thermostability. *Proc Natl Acad Sci U S A* **92**: 9264-8.
- Kouzarides, T., A. T. Bankier, et al. (1987). Sequence and transcription analysis of the human cytomegalovirus DNA polymerase gene. *J Virol* **61**: 125-33.
- Krosky, P. M., M.-C. Baek, et al. (2003). The Human Cytomegalovirus UL44 Protein Is a Substrate for the UL97 Protein Kinase. *J Virol* **77**: 7720-7.

- Ladbury, J. et M. Doyle (2004). *Biocalometry 2, Application of calorimetry in the biological sciences*. Chichester, Wiley.
- LaFemina, R. et G. Hayward (1983). Replicative forms of human cytomegalovirus DNA with joined termini are found in permissively infected human cells but not in non-permissive Balb/c-3T3 mouse cells. *J Gen Virol* **64**: 373-89.
- Leckband, D. et J. Israelachvili (2001). Intermolecular forces in biology. *Quarterly Reviews of Biophysics* **34**: 105-267.
- Lee, H. R. et K. A. Johnson (2006). Fidelity of the Human Mitochondrial DNA Polymerase. *J Biol Chem* **281**: 36236-40.
- Littler, E., A. Stuart, et al. (1992). Human cytomegalovirus UL97 open reading frame encodes a protein that phosphorylates the antiviral nucleoside analogue ganciclovir. *Nature* **358**: 160-2.
- Loregian, A., B. Appleton, et al. (2004). Residues of human cytomegalovirus DNA polymerase catalytic subunit UL54 that are necessary and sufficient for interaction with the accessory protein UL44. *J Virol* **78**: 158-67.
- Loregian, A., B. Appleton, et al. (2004). Specific residues in the connector loop of the human cytomegalovirus DNA polymerase accessory protein UL44 are crucial for interaction with the UL54 catalytic subunit. *J Virol* **78**: 9084-92.
- Loregian, A. et D. Coen (2006). Selective anti-cytomegalovirus compounds discovered by screening for inhibitors of subunit interactions of the viral polymerase. *Chem Biol* **13**: 191-200.
- Loregian, A. et G. Palù (2005). Disruption of the interaction between the subunits of herpesvirus DNA polymerases as a novel antiviral strategy. *Clin Microbiol Infect* **11**: 437-46.
- Loregian, A., R. Rigatti, et al. (2003). Inhibition of human cytomegalovirus DNA polymerase by C-terminal peptides from the UL54 subunit. *J Virol* **77**: 8226-44.
- Loregian, A., R. Rigatti, et al. (2003). Inhibition of Human Cytomegalovirus DNA Polymerase by C-Terminal Peptides from the UL54 Subunit. *J Virol* **77**: 8336-44.
- Loregian, A., E. Sinigalia, et al. (2007). Binding parameters and thermodynamics of the interaction of the human cytomegalovirus DNA polymerase accessory protein, UL44, with DNA: implications for the processivity mechanism. *Nucleic Acids Res* **35**: 4779-91.
- Maciejewski, J. P., E. E. Bruening, et al. (1993). Infection of mononucleated phagocytes with human cytomegalovirus. *J Virol* **67**: 327-36.

- Malm, G. et M. Engman (2007). Congenital cytomegalovirus infections. *Semin Fetal Neonatal Med* **12**(3): 154-9.
- Markham, A. et D. Faulds (1994). Ganciclovir. An update of its therapeutic use in cytomegalovirus infection. *Drugs* **48**: 455-84.
- Marschall, M., M. Freitag, et al. (2003). The protein kinase pUL97 of human cytomegalovirus interacts with and phosphorylates the DNA polymerase processivity factor pUL44. *Virology* **311**: 60-71.
- Martignetti, J. et B. Barrell (1991). Sequence of the HindIII T fragment of human cytomegalovirus, which encodes a DNA helicase. *J Gen Virol* **72**: 1113-21.
- Marwick, C. (1998). First "antisense" drug will treat CMV retinitis. *JAMA* **280**: 871.
- McVoy, M. et S. Alder (1994). Human Cytomegalovirus DNA replicates after early circularization by concatemer formation, and inversion occurs within the concatemer. *J Virol* **68**: 1040-51.
- Meier-Andrejszki, L., S. Bjeli, et al. (2007). Thermodynamics of b-HLH-LZ Protein Binding to DNA: The Energetic Importance of Protein-DNA Contacts in Site-Specific E-Box Recognition by the Complete Gene Product of the Max p21 Transcription Factor. *Biochemistry* **46**: 12427-40.
- Mercorelli, B., E. Sinigalia, et al. (2007). Human cytomegalovirus DNA replication: antiviral targets and drugs. *Rev Med Virol* **Sous presse**.
- Mettenleiter, T. (2004). Budding events in herpesvirus morphogenesis. *Virus Res* **106**: 167-80.
- Mitchell, D. C. et B. J. Litman (1999). Effect of Protein Hydration on Receptor Conformation: Decreased Levels of Bound Water Promote Metarhodopsin II Formation. *Biochemistry* **38**: 7617-23.
- Mocarski, E., T. Shenk, et al. (2007). Cytomegaloviruses. *Fields Virology* **5**: 2701-72.
- Morrison, A., R. Christensen, et al. (1989). REV3, a *Saccharomyces cerevisiae* gene whose function is required for induced mutagenesis, is predicted to encode a nonessential DNA polymerase. *J Bacteriol* **171**: 5659-67.
- Muranyi, W., J. Haas, et al. (2002). Cytomegalovirus Recruitment of Cellular Kinases to Dissolve the Nuclear Lamina. *Science* **297**: 854-57.
- Myers, T. et D. Gelfand (1991). Reverse transcription and DNA amplification by a *Thermus thermophilus* DNA polymerase. *Biochemistry* **30**: 7661-6.

- Naesens, L. et E. D. Clercq (2001). Recent developments in herpesvirus therapy. *Herpes* **8**: 12-6.
- Ollis, D., P. Brick, et al. (1985). Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. *Nature* **313**: 762-6.
- Pari, G. et D. Anders (1993). Eleven loci encoding trans-acting factors are required for transient complementation of human cytomegalovirus oriLyt-dependent DNA replication. *J Virol* **67**: 6979-88.
- Pari, G., M. Kacica, et al. (1993). Open reading frames UL44, IRS1/TRS1, and UL36-38 are required for transient complementation of human cytomegalovirus oriLyt-dependent DNA synthesis. *J Virol* **67**: 2575-82.
- Parker, M., Y. Kim, et al. (2006). Interactions among four proteins encoded by the human cytomegalovirus UL112-113 region regulate their intranuclear targeting and the recruitment of UL44 to prereplication foci. *J Virol* **80**: 2718-27.
- Pellett, P., B. Roizman, et al. (2007). The Family Herpesviridae: A Brief Introduction. *Fields Virology* **2**: 2479-99.
- Penfold, M. et E. Mocarski (1997). Formation of cytomegalovirus DNA replication compartments defined by localization of viral proteins and DNA synthesis. *Virology* **239**: 46-61.
- Perry, C. et J. Balfour (1999). Fomivirsen. *Drugs* **57**: 375-80.
- Plachter, B., C. Sinzger, et al. (1996). Cell types involved in replication and distribution of human cytomegalovirus. *Adv Virus Res* **46**: 195-261.
- Prichard, M., D. Duke, et al. (1996). Human cytomegalovirus uracil DNA glycosylase is required for the normal temporal regulation of both DNA synthesis and viral replication. *J Virol* **70**: 3018-25.
- Prichard, M., N. Gao, et al. (1999). A recombinant human cytomegalovirus with a large deletion in UL97 has a severe replication deficiency. *J Virol* **73**: 5663-70.
- Prichard, M., H. Lawlor, et al. (2005). Human cytomegalovirus uracil DNA glycosylase associates with ppUL44 and accelerates the accumulation of viral DNA. *Virol J* **15**: 1-14.
- Razonable, R. et V. Emery (2004). Management of CMV infection and disease in transplant patients. *Herpes* **11**: 77-86.
- Riddell, S., E. Warren, et al. (2000). Application of T cell immunotherapy for human viral and malignant diseases. *Ernst Schering Res Found Workshop* **30**: 53-73.

- Rowshani, A., F. Bemelman, et al. (2005). Clinical and Immunologic Aspect of Cytomegalovirus Infection in solid Organ transplant Recipients. *Transplantation* **79**(4): 381-6.
- Rubsam, L., B. Davidson, et al. (1998). Superior cytotoxicity with ganciclovir compared with acyclovir and 1-*-D*-arabinofuranosylthymine in herpes simplex virus-thymidine kinase-expressing cells: a novel paradigm for cell killing. *Cancer Research* **58**: 3873-82.
- Rubsam, L. Z., P. D. Boucher, et al. (1999). Cytotoxicity and Accumulation of Ganciclovir Triphosphate in Bystander Cells Cocultured with Herpes Simplex Virus Type 1 Thymidine Kinase-expressing Human Glioblastoma Cells. *Cancer Research* **59**: 669-75.
- Sarisky, R. et G. Hayward (1996). Evidence that the UL84 gene product of human cytomegalovirus is essential for promoting oriLyt-dependent DNA replication and formation of replication compartments in cotransfection assays. *J Virol* **70**: 7398-413.
- Schrier, R., J. Nelson, et al. (1985). Detection of human Cytomegalovirus in peripheral blood Lymphocytes in a natural infection. *Science* **230**: 1048-51.
- Sinclair, J. et P. Sissons (2006). Latency and reactivation of human cytomegalovirus. *J of Gen Virol* **87**: 1763-79.
- Sinzger, C., A. Grefte, et al. (1995). Fibroblasts, epithelial cells, endothelial cells and smooth muscle cells are major targets of human cytomegalovirus infection in lung and gastrointestinal tissues. *J Gen Virol* **76**: 741-50.
- Smith, I., J. Cherrington, et al. (1997). High-level resistance of cytomegalovirus to ganciclovir is associated with alterations in both the UL97 and DNA polymerase genes. *J. Infect. Dis* **176**: 69-77.
- Soderberg, C., S. Larsson, et al. (1993). Definition of a subset of human peripheral blood mononuclear cells that are permissive to human cytomegalovirus infection. *J Virol* **67**: 3166-75.
- Spector, H. (1996). Activation and regulation of human cytomegalovirus early genes. *Intervirolgy* **39**: 316-77.
- Statistique_Canada (2008). L'espérance de vie s'est allongée de façon constante au cours du 20e siècle. <http://www.statcan.ca/francais/freepub/91-003-XIF/2007001/com-fr.htm>.
- Steitz, T. (1998). A mechanism for all polymerases. *Nature* **391**: 231-2.
- Steitz, T. et Y. Yin (2004). Accuracy, lesion bypass, strand displacement and translocation by DNA polymerases. *Philos Trans R Soc Lond B Biol Sci* **29**: 17-23.

- Stenberg, R. (1996). The human cytomegalovirus major immediate-early gene. *Intervirology* **39**(5-6): 343-9.
- Sullivan, V., C. L. Talarico, et al. (1992). A protein kinase homologue controls phosphorylation of ganciclovir in human cytomegalovirus-infected cells. *Nature* **358**: 162-4.
- Swaminathan, C. P., A. Nandi, et al. (1999). Thermodynamic Analyses Reveal Role of Water Release in Epitope Recognition by a Monoclonal Antibody against the Human Guanylyl Cyclase C Receptor. *J Biol Chem* **274**: 3127-8.
- Tabor, S. et C. C. Richardson (1989). Effect of manganese ions on the incorporation of dideoxynucleotides by bacteriophage T7 DNA polymerase and Escherichia coli DNA polymerase I. *Proc Natl Acad Sci U S A* **86**: 4076-80.
- Taylor, T., M. Brockman, et al. (2002). Herpes simplex virus. *Front Biosci* **7**: 752-64.
- Taylor-Wiedeman, J., J. G. Sissons, et al. (1991). Monocytes are a major site of persistence of human cytomegalovirus in peripheral blood mononuclear cells. *J Gen Virol* **72**: 2059-64.
- Taylor-Wiedeman, J., P. Sissons, et al. (1994). Induction of endogenous human cytomegalovirus gene expression after differentiation of monocytes from healthy carriers. *J Virol* **68**: 1597-604.
- Tchesnokov, E., C. Gilbert, et al. (2006). Role of helix P of the human Cytomegalovirus DNA polymerase in resistance and hypersusceptibility to the antiviral drug foscarnet. *J Virol* **80**: 1440-50.
- Tchesnokov, E. P., C. Gilbert, et al. (2006). Role of Helix P of the Human Cytomegalovirus DNA Polymerase in Resistance and Hypersusceptibility to the Antiviral Drug Foscarnet. *J Virol* **80**: 1440-1450.
- Tomicic, M. T., R. Thust, et al. (2001). DNA Polymerase β Mediates Protection of Mammalian Cells against Ganciclovir-induced Cytotoxicity and DNA Breakage. *Cancer Research* **61**: 7399-403.
- Tucker, J., T. Clayton, et al. (2000). Structure-activity relationships of acycloxyamidine cytomegalovirus DNA polymerase inhibitors. *Bioorg Med Chem* **8**: 601-615.
- Uprichard, S. et D. Knipe (1997). Assembly of herpes simplex virus replication proteins at two distinct intranuclear sites. *Virology* **229**: 113-125.
- Varnum, S., D. Streblow, et al. (2004). Identification of Proteins in Human Cytomegalovirus (HCMV) Particles: the HCMV Proteome. *J Virol* **78**(20): 10960-6.

ViropharmaInc. (2007). <http://www.viropharma.com>.

Wailand, K., N. Oien, et al. (1994). Functional analysis of human cytomegalovirus polymerase accessory protein. *Virus Res* **34**: 191-206.

Wang, J., A. Sattar, et al. (1997). Crystal structure of a pol alpha family replication DNA polymerase from bacteriophage RB69. *Cell* **27**: 1087-99.

Wang, T., S. Wong, et al. (1989). Human DNA polymerase alpha: predicted functional domains and relationships with viral DNA polymerases. *FASEB J* **1**: 14-21.

Washington, M. T., R. E. Johnson, et al. (1999). Fidelity and Processivity of *Saccharomyces cerevisiae* DNA Polymerase eta. *J Biol Chem* **274**: 36835-8.

Weinberg, A., D. Jabs, et al. (2003). Mutations conferring foscarnet resistance in a cohort of patients with acquired immunodeficiency syndrome and cytomegalovirus retinitis. *J. Infect. Dis* **187**: 777-84.

Wong, S., A. Wahl, et al. (1988). Human DNA polymerase alpha gene expression is cell proliferation dependent and its primary structure is similar to both prokaryotic and eukaryotic replicative DNA polymerases. *EMBO J* **7**: 37-47.

WorldHealthOrganisation (2008). Smallpox.
<http://www.who.int/mediacentre/factsheets/smallpox/en/>.

Xiong, X., J. L. Smith, et al. (1996). Kinetic analysis of the interaction of cidofovir diphosphate with human cytomegalovirus DNA polymerase. *Biochem. Pharmacol* **51**: 1563-7.

Yu, D., M. Silva, et al. (2003). Functional map of human cytomegalovirus AD169 defined by global mutational analysis. *Proc Natl Acad Sci U S A* **100**: 12396-401.

Zahn, K., H. Belrhali, et al. (2007). Caught bending the a-rule: crystal structures of translesion DNA synthesis with a non-natural nucleotide. *Biochemistry* **46**: 10551-61.