

**Université de Sherbrooke**

**La synthèse de la coiffe en tant que nouvelle cible thérapeutique potentielle**

**Par  
Maude Tremblay-Létourneau  
Maîtrise en 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**

**Sherbrooke (Québec) Canada  
Février 2012**

**Membres du jury d'évaluation  
Martin Bisaillon, Département de Biochimie  
Luigi Bouchard, Département de Biochimie  
Luc Gaudreau, Département de 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: 978-0-494-93335-0*

*Our file Notre référence*

*ISBN: 978-0-494-93335-0*

#### 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.

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

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

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

Canada

# Table des matières

<b>Liste des illustrations .....</b>	<b>VI</b>
<b>Liste des sigles, abréviations et symboles .....</b>	<b>VIII</b>
<b>Résumé .....</b>	<b>1</b>
<b>1) Introduction .....</b>	<b>3</b>
1.1. Contexte .....	3
1.2. Stratégie de recherche d'inhibiteur .....	3
1.2.1. Criblage à haut débit .....	5
1.2.2. Criblage virtuel .....	7
1.2.2.1. La sélection des composés .....	8
1.2.2.2. Le calcul des conformations des ligands .....	9
1.2.2.3. L'acquisition des structures protéiques .....	10
1.2.2.4. Le calcul de l'interaction prédite .....	11
1.3. Les différents mécanismes d'action des inhibiteurs.....	13
1.3.1. L'inhibition par des analogues du substrat .....	13
1.3.2. Les interactions protéines-protéines.....	15
1.3.3. L'inhibition allostérique.....	16
1.4. Les cibles pour traitement pharmacologique .....	16
1.4.1. Les antifongiques.....	17
1.4.2. Les protéines de l'hôte .....	17
1.4.3. Les protéines de la maturation des ARNm.....	18

<b>1.5. La maturation des ARNm .....</b>	<b>19</b>
1.5.1. L'ajout d'une structure coiffe en 5' .....	20
1.5.2. Le processus d'édition .....	21
1.5.3. L'épissage .....	23
1.5.4. La polyadénylation .....	25
<b>1.6. Les rôles de la structure coiffe .....</b>	<b>27</b>
1.6.1. La stabilité.....	27
1.6.2. Le transport .....	28
1.6.3. L'épissage .....	29
1.6.4. La traduction.....	30
<b>1.7. Les enzymes responsables de la structure coiffe.....</b>	<b>30</b>
1.7.1. Les ARN triphosphatases .....	32
1.7.2. Les ARN guanylyltransférases.....	33
1.7.3. Les guanine-N7-ARN méthyltransférases.....	34
1.7.4. Les ARN 2'O- méthyltransférases.....	35
<b>1.8. La famille des nucléotidyltransférases .....</b>	<b>36</b>
1.8.1. Les membres des nucléotidyltransférases .....	36
1.8.2. Les ligases à ADN .....	37
1.8.3. Les ligases à ARN .....	38
1.8.4. Les ARN guanylyltransférases.....	39
1.8.5. La structure conservée des nucléotidyltransférases .....	40
1.8.6. Le réarrangement conformationnel lors de la catalyse .....	41
<b>1.9. La caractérisation des ARN guanylyltransférases.....</b>	<b>42</b>
1.9.1. Les ARN guanylyltransférases virales .....	43
1.9.2. Les ARN guanylyltransférases fongiques.....	44
1.9.3. Les ARN guanylyltransférases des métazoaires .....	45
1.9.4. Les inhibiteurs des ARN guanylyltransférases.....	46
<b>1.10. Préambule à l'article.....</b>	<b>48</b>

<b>2) Article .....</b>	<b>50</b>
2.1. Sommaire .....	51
2.2. Implication dans l'article.....	51
<b>3) Discussion .....</b>	<b>85</b>
3.1. Synthèse .....	85
3.2. Criblage virtuel.....	86
3.3. Inhibition par l'acide mycophénolique .....	88
3.4. Spécificité de l'inhibition .....	89
3.5. Inhibition des GTases in vivo.....	91
<b>4) Perspectives .....</b>	<b>95</b>
4.1. Confirmer le site d'interaction entre le MPA et la GTase.....	95
4.1.1. Générer des protéines mutantes .....	95
4.1.1.1. Théorie.....	95
4.1.1.2. Procédure expérimentale.....	96
4.1.2. Validation par interaction en utilisant la microcalorimétrie .....	98
4.1.2.1. Théorie.....	98
4.1.2.2. Procédure expérimentale.....	100
4.2. Développement d'inhibiteur allostérique sélectif .....	101
4.2.1. Utilisation du criblage virtuel orienté vers les enzymes de pathogènes.....	102
4.2.2. Essai en cellules pour trouver un inhibiteur sélectif pour les enzymes de pathogènes .....	103

4.2.2.1. Théorie.....	103
4.2.2.2. Procédure expérimentale .....	104
<b>5) Conclusion .....</b>	<b>104</b>
<b>6) Remerciements.....</b>	<b>106</b>
<b>7) Références.....</b>	<b>107</b>

# Liste des illustrations

Figure 1. Étapes de la recherche et développement .....	4
Figure 2. Schéma récapitulatif des étapes du criblage virtuel .....	10
Figure 3. Schématisation de la structure tertiaire d'une enzyme et des diverses cavités..	14
Figure 4. Inhibition basée sur les interactions protéines-protéines .....	15
Figure 5. Les différentes étapes de la maturation d'un ARN messagers.....	19
Figure 6. Structure de la coiffe .....	21
Figure 7. Mécanisme de l'édition et composantes du complexe de l'éditosome chez <i>Trypanosoma brucei</i> .....	23
Figure 8. Les différentes étapes de l'épissage et de l'association du splicéosome.....	25
Figure 9. Un modèle du complexe multiprotéique du clivage et de la polyadénylation des ARN pré-messagers .....	27
Figure 10. Les différents rôles de la structure coiffe .....	29
Figure 11. Organisation génique des enzymes responsables de la structure coiffe.....	31
Figure 12. Les structures des ARN guanylyltransférases du virus PBCV-1 et de la levure <i>C. albicans</i> .....	33
Figure 13. Structure du site actif de l'ARN guanine-N7-méthyltransférase .....	35
Figure 14. Les différents motifs conservés de la famille des nucléotidyltransférases.....	37
Figure 15. Le changement de conformation de l'ARN guanylyltransférase lors de la catalyse.....	42
Figure 16. Représentation graphique Lineweaver-burk de l'inhibition du MPA sur la GTase de <i>S. cerevisiae</i> . .....	89

Figure 17. La présence de l'inhibiteur MPA diminue la prolifération cellulaire de <i>S. cerevisiae</i> .....	93
Figure 18. L'essai d'extension d'amorce sur des ARNm décoiffés en fonction du temps. ....	94
Figure 19. Les acides aminés de l'IMPDH réagissant avec l'inhibiteur MPA.....	97
Figure 20. . Les différentes composantes de l'ITC .....	99
Figure 21. L'énergie de Gibbs pour les différentes étapes de l'activité enzymatique des GTases.....	101



# Liste des sigles, abréviations et symboles

**ADN** : Acide désoxyribonucléique

**ADP** : Adénosine diphosphate

**AMP** : Adénosine monophosphate

**ARN** : Acide ribonucléique

**ARNm** : ARN messenger

**ARN prém** : ARN pré-messenger

**ATP** : Adénosine triphosphate

**CBC** : Protéines liant la coiffe ( cap binding complex )

**CPSF** : Facteur de spécificité pour le clivage et la polyadénylation

**CTD** : Domaine carboxy-terminal

**FRET** : Transfert d'énergie entre molécules fluorescentes

**GTase** : ARN guanylyltransférase

**HSV** : Virus *Herpes simplex*

**HTS** : Criblage à haut débit ( high-throughput screening )

**IMPDH** : Inosine monophosphate déshydrogénase

**MPA** : Acide mycophénolique

**N7-MTase** : ARN-guanine-N7-méthyltransférase

**NADH** : Nicotinamide adénine dinucléotide

**NTPase** : Nucléosides phosphatases

**NT** : Domaine nucléotidyltransférase

**OB : Domaine de liaison à l'oligonucléotide**

**PBCV-1 : *Paramecium bursaria* Chlorella virus 1**

**PDB : Banque de données protéique ( Protein data bank )**

**PolyA : Polyadénosines**

**RT : Transcriptase inverse**

**RTase : ARN triphosphatase**

**SAM : S-adenosylméthionine**

**VIH : Virus de l'immunodéficience humaine**

**vHTS : Criblage virtuel à haut débit « virtual high-throughput screening »**

# Résumé

Université de Sherbrooke

**La synthèse de la coiffe en tant que nouvelle cible thérapeutique potentielle**

Par

Maude Tremblay-Létourneau  
Maîtrise en biochimie  
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

Sherbrooke (Québec) Canada  
Février 2012

Membres du jury d'évaluation :

Dr. Martin Bisailon, Biochimie  
Dr. Luigi Bouchard, Biochimie  
Dr. Luc Gaudreau, Biologie Faculté des Sciences, Université de Sherbrooke

---

Afin de pallier l'apparition de souches pathogènes résistantes, de nouveaux traitements aux mécanismes d'action inédits doivent être découverts. Une de ces cibles d'intérêt grandissant est la synthèse de la structure coiffe chez les ARN messagers (ARNm) eucaryotes. Cette structure est l'une des modifications co-transcriptionnelles essentielles pour augmenter la demi-vie des ARNm, promouvoir leur export du noyau vers le cytoplasme et augmenter l'efficacité de leur traduction en protéines. La synthèse de la structure coiffe ( $m^7Gppp$ ARN) résulte de trois activités enzymatiques séquentielles qui ajoutent la structure  $m^7Gppp$ . Initialement, une ARN triphosphatase (RTase) hydrolyse le phosphate gamma de l'ARN. Une ARN guanylyltransférase (GTase) transfère ensuite sur cet ARN un groupement GMP à partir d'un GTP. Finalement, une ARN guanine-N7 méthyltransférase (N7-MTase) vient méthyler cette guanine en position N7.

L'activité enzymatique de la GTase est déterminante pour la synthèse de la structure coiffe, ce qui en fait une cible potentiellement puissante afin d'inhiber la synthèse de la structure coiffe. Certaines études ont identifié des analogues de substrat et de produit, la ribavirine et le foscarnet respectivement, comme inhibiteur de GTase. Ces produits agissent sur les différentes GTases puisqu'elles possèdent un site actif très conservés au niveau de la coordination du substrat. Pour augmenter la spécificité des inhibiteurs, il serait avantageux d'identifier des composés ciblant les régions non conservées. Dans la présente étude, la technique du criblage virtuel a permis d'analyser le potentiel d'interaction des molécules à 80 % similaires au substrat naturel avec plusieurs structures de GTases. Parmi les interactions prédites identifiées par le criblage virtuel, une de ces molécules suscitait un intérêt particulier, puisqu'il s'agissait d'un composé utilisé chez l'humain en tant qu'antiviral et immunosuppresseur, l'acide mycophénolique (MPA). L'hypothèse initiale était que ce composé pouvait potentiellement interférer dans l'activité enzymatique des GTases. Pour valider cette prédiction, des essais biochimiques ciblant la réaction complète et les étapes intermédiaires de la GTase de la levure *Saccharomyces cerevisiae* (*S. cerevisiae*) ont été utilisées pour définir l'effet du MPA sur la cinétique de cette GTase. En effet, la réaction de GTase se déroule en deux temps : premièrement, une molécule de GTP est hydrolysée par l'enzyme pour former du GMP, lequel est lié de façon covalente à l'enzyme, et deuxièmement, cette molécule de GMP est transférée sur un ARN diphosphorylé. Ces essais ont confirmé l'interaction prédite par le criblage virtuel. En utilisant cette approche, nous démontrons que le MPA peut inhiber la réaction de GTase en prévenant le transfert catalytique du nucléotide GMP sur l'ARNm accepteur. En ce sens, le MPA représente un nouveau type d'inhibiteur d'ARN guanylyltransférase qui inhibe la deuxième étape de l'activité catalytique. Puisque les résultats laissaient présager une inhibition non compétitive, il fallait confirmer que le MPA n'était pas reconnu par l'enzyme comme substrat. Pour ce faire, la technique de l'électrophorèse par capillarité a montré que l'inhibiteur ne formait pas de lien covalent avec l'enzyme. De plus, nous démontrons qu'une culture de *Saccharomyces cerevisiae* qui croit en présence de MPA dénote une quantité moindre d'ARNm possédant une coiffe. Finalement, des essais biochimiques démontrent que le composé peut également inhiber des enzymes de la même famille que les GTases utilisant des substrats différents : les ligases à ADN. Le MPA inhibe la deuxième étape de la réaction enzymatique d'une ligase, soit le transfert du nucléotide sur l'oligonucléotide. Ce mémoire vise donc à présenter un aperçu du mécanisme d'inhibition de l'ARN guanylyltransférase par un composé allostérique. Nous montrons qu'il est possible de réduire la synthèse de la structure coiffe en affectant les changements conformationnels de cette enzyme.

**Mots clés :** Structure coiffe, ARN messagers, criblage virtuel, ARN guanylyltransférase, nucléotidyltransférase, inhibiteur allostérique

# 1) Introduction

## *1.1. Contexte*

Au début du 20<sup>e</sup> siècle, les infections bactériennes étaient la cause majeure de mortalité (Ellis, 2002). Le développement d'antibiotiques a permis de réduire leur incidence, si bien que vers la deuxième moitié de ce siècle, la menace pour la santé globale a changé pour les infections virales et fongiques (De Clercq, 2004b). L'incidence des infections est influencée par le nombre croissant de patients souffrants d'immunodéficience par le cancer, le vieillissement ou le diabète (Morschhauser, 2010). Afin de pallier à la progression des maladies infectieuses, une première vague de thérapies ciblant ces pathogènes a été générée. Par contre, la médecine fait face constamment à des problèmes de résistance des organismes, autant viraux que fongiques, pour différentes drogues. (Vanden Bossche, 1997). Ce problème de résistance par les souches pathogènes peut être acquis de plusieurs façons; 1) par une diminution de la concentration de la drogue intracellulaire obtenue par la dégradation du composé en sous produit, 2) par une diminution de l'affinité de la drogue pour la cible ou 3) par une modification du métabolisme pour contrecarrer l'effet de la drogue (MacCallum et al., 2010). L'exemple du virus Influenza illustre bien ce problème de résistance puisque la souche H1N1 est résistante à 98 % à l'agent antiviral Oseltamivir tandis que la souche H3N2 résiste à 99 % à l'Amantadine (Das et al., 2010). Dans cette optique de résistance aux médicaments, il est primordial de développer de nouvelles stratégies thérapeutiques aux mécanismes d'actions novateurs qui agissent sur de nouvelles cibles potentielles.

## *1.2. Stratégie de recherche d'inhibiteur*

Le développement de plusieurs antibiotiques dans les années 1950 a permis de contrer l'étalement des maladies infectieuses dans le monde (De Clercq, 2004b). Néanmoins, au

cours des deux dernières décennies, l'apparition de souches pathogènes résistantes aux traitements proposés, soit de nature virale, fongique ou bactérienne, constitue une menace constante pour la santé publique mondiale (Morschhauser, 2010). Par exemple, en 2002, l'Organisation mondiale de la santé a évalué que 95 % des souches de *Staphylococcus aureus* étaient résistantes à la pénicilline, illustrant ainsi la nécessité de développer de nouvelles classes d'action d'antibiotiques.

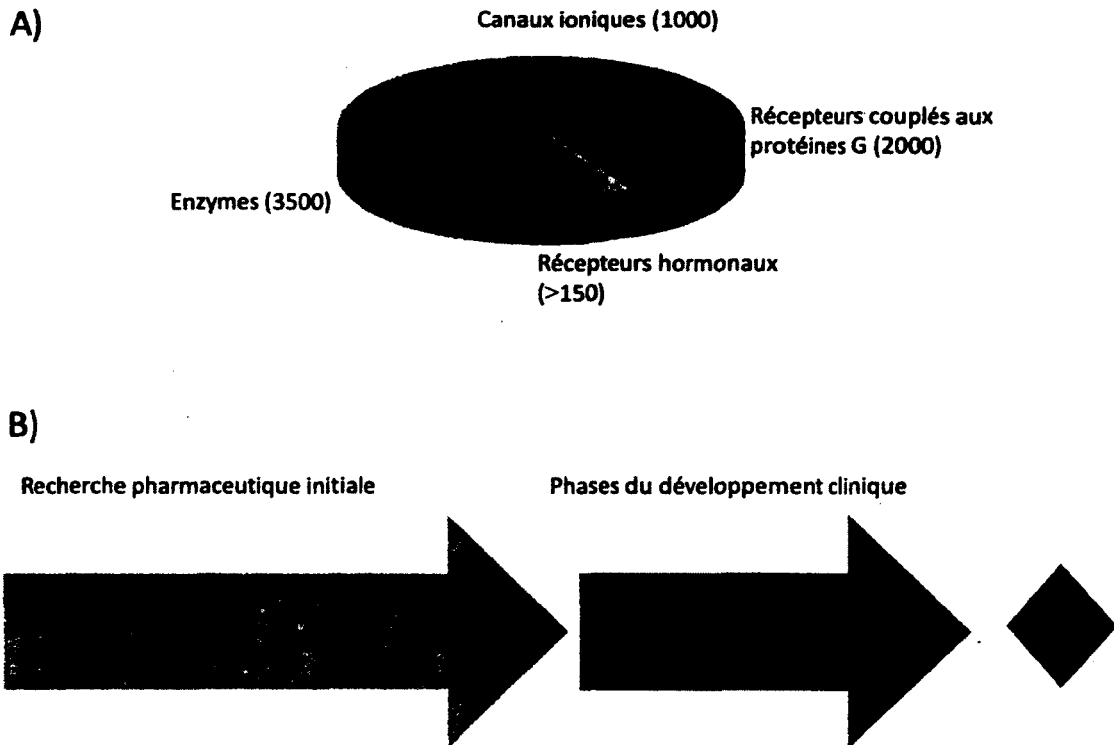


Figure 1. Étapes de la recherche et développement (adapté de Terstappen et Reggiani, 2001). A) Le diagramme circulaire illustre les classes fréquentes de cibles utilisées afin de développer des médicaments potentiels ainsi que le nombre approximatif de composés recensés pour chacune. Ces chiffres sont issus de la consultation de plusieurs banques de données telles que : OMIM (Online Mendelian Inheritance in Man; <http://www3.ncbi.nlm.nih.gov/Omim/>), SWISSPROT (<http://www.expasy.ch/sprot/sprot-top.html>), Incyte LifeSeq Gold Human Gene Database (<http://www.incyte.com/products/lifeseq/lifeseqgold.html>), GeneCards (human genes, proteins and disease; <http://bioinformatics.welzmann.ac.il/cards/>), TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>) et GDB (Genome Database; <http://gdbwww.gdb.org/>). B) Le processus de recherche et développement dans le domaine pharmaceutique peut être fractionné selon deux phases : la recherche initiale permettant les découvertes de la cible choisie et des composés prometteurs, alors que la deuxième phase régit le développement clinique des composés en ce qui a trait aux études pharmacocinétiques (absorption, diffusion, élimination du médicament étudié) ainsi que les études de pharmacodynamiques.

### **1.2.1. Criblage à haut débit**

Les études structurales et mécanistiques des différentes enzymes, autant présentes chez les pathogènes que chez l'humain, permettent de mettre à l'avant-plan les différences inhérentes assujetties à être exploitées dans un but thérapeutique (Bleicher et al., 2003). Ainsi, le nombre grandissant de cibles potentielles jumelé à de larges banques de composés ont poussé l'industrie à développer une technologie très rapide (Hertzberg et Pope, 2000). Afin de trouver des composés spécifiques pour une cible précise, la recherche moderne opte souvent pour le criblage à haut débit, qui permet de tester plusieurs dizaines de milliers de composés par jour (Goddard et Reymond, 2004). Cette technique permet d'évaluer des interactions réelles entre un composé et une cible en limitant les possibilités de faux positifs (Hertzberg et Pope, 2000). Dans le but d'augmenter la rapidité de ce type d'essai, plusieurs avancées technologiques ont permis d'optimiser la méthode (Keseru et Makara, 2006). Tout d'abord, pour minimiser les coûts associés aux composés biologiques et chimiques utilisés par essai, les volumes réactionnels ont passé de l'ordre d'une centaine de microlitres vers une fraction d'un microlitre (Sundberg, 2000). Une autre approche poursuivie dans ce même but a été de remplacer les microplaques 96 puits par des plaques contenant toujours plus de puits, voir jusqu'à plus de 9000 (Hertzberg et Pope, 2000). La deuxième facette d'amélioration du procédé de criblage à haut débit a été de développer des tests spécifiques pour la détection d'une interaction entre un ligand et une protéine. Parmi tous les types d'essais qui sont sujets à être utilisés, deux types d'essais comportant chacun leurs avantages propres seront abordés, soit l'essai biochimique et l'essai cellulaire.

Les essais biochimiques sont fréquemment complétés en employant la scintillation par proximité ou la technique de détection par fluorescence (White, 2000). La scintillation par proximité utilise un composé scintillant attaché sur des microbilles, auxquelles sont immobilisées les cibles d'intérêt (Goddard et Reymond, 2004). Le composé scintillant

possède la capacité de s'exciter et d'émettre un rayonnement lorsqu'il à proximité des rayonnements d'un ligand radioactif qui interagit avec la cible (Hertzberg et Pope, 2000). Les isotopes préférablement utilisés possèdent une faible émission d'énergie, tel le  $^3\text{H}$ . Le rayonnement de cet isotope est rapidement absorbé sur de courtes distances et permet de signaler uniquement les composés qui interagissent directement avec la cible (von Ahsen et Bomer, 2005). D'autre part, la technique basée sur la fluorescence a l'avantage d'être très sensible dû à la rapidité d'excitation de la plupart des fluorophores (Goddard et Reymond, 2004). La méthode la plus courante, le transfert d'énergie entre molécules fluorescentes (FRET), implique le transfert d'énergie du donneur vers une molécule acceptrice. Pour schématiser l'application de cette méthode, prenons la situation où la cible choisie est une protéase. Le substrat de l'enzyme est alors marqué par les deux fluorophores à une distance permettant d'interagir et d'émettre un signal (von Ahsen et Bomer, 2005). La protéase est mise en contact avec une panoplie de composés inhibiteurs potentiels. Les interactions positives sont déterminées par l'absence d'un signal de fluorescence qui se traduit par l'inhibition du clivage du substrat (White, 2000). Bien que ces méthodes permettent de déterminer des interactions réelles pour une cible donnée, des tests supplémentaires doivent déterminer la toxicité du composé, sa spécificité pour la protéine d'intérêt et sa biodisponibilité.

Plusieurs essais cellulaires peuvent être employés, mais uniquement les catégories principales seront abordées : l'activation d'un récepteur à la surface cellulaire, la transcription/la traduction d'un gène rapporteur et la prolifération cellulaire (Goddard et Reymond, 2004). Lorsque le criblage implique un récepteur cellulaire, l'essai cellulaire adéquat repose sur l'utilisation d'un fluorophore réagissant de façon spécifique à l'activation du récepteur en décelant une condition précise. Plusieurs systèmes de détection ont été mis au point afin de reconnaître autant un changement de pH, un changement de calcium intracellulaire que la dépolarisation membranaire (Bleicher et al., 2003). Les deux autres méthodes exigent des étapes d'incubation de plusieurs jours suivies d'un test colorimétrique ou fluorescent. Une méthode fréquemment utilisée repose sur le principe que l'enzyme ciblée, une fois activée, entraîne l'expression du gène rapporteur (White, 2000). Ce type d'essai est beaucoup plus difficile à réaliser, tant au



niveau de l'élaboration d'un test spécifique pour une enzyme précise qu'au niveau de l'optimisation de l'étape de détection (von Ahsen et Bomer, 2005). L'avantage de ce type d'essai est de visualiser l'impact cellulaire du composé, notamment au niveau de la spécificité et de la biodisponibilité. L'efficacité de cette méthode n'est plus à prouver; plusieurs composés pharmacologiques ont été identifiés par le criblage à haut débit (Keseru et Makara, 2006). Toutefois, cette technologie de pointe ne peut être endossée par plusieurs centres de recherches considérant ses coûts exorbitants (Baldi, 2010).

### **1.2.2. Criblage virtuel**

Plusieurs approches ont été développées pour pallier aux facteurs économiques du criblage à haut débit. Parmi toutes les technologies développées, la méthode alternative qui gagne en popularité pour la recherche pharmacologique repose sur l'utilisation des outils informatiques. L'avantage premier est de pouvoir traiter une grande quantité de données rapidement et économiquement (Bajorath, 2002). La modélisation moléculaire, plus précisément l'arrimage moléculaire, consiste en l'évaluation de plusieurs structures chimiques virtuelles interagissant avec un modèle informatique (Baldi, 2010). Ainsi, cette méthode permet de chercher au sein d'une banque de plusieurs composés *in silico* afin de sélectionner un nombre limité de molécules prédites pour interagir avec la cible choisie (Yang et al., 2011). Ces composés prometteurs d'interactions potentielles, calculés selon des algorithmes informatiques, peuvent ensuite être validés *in vitro* et *in vivo*. Les origines d'une telle technique peuvent être retracées au milieu des années 1970 (Taylor et al., 2002). Néanmoins, l'essor du criblage virtuel émane des avancées technologiques en ce qui a trait au positionnement du ligand dans l'espace sous ses multiples conformations ainsi qu'au calcul de l'algorithme pour l'interaction prédite (Terstappen et Reggiani, 2001). Les succès dans ce domaine découlent de la possibilité de combiner les techniques de chimie de synthèse pour tester précisément certains composés non commercialisés (Shoichet, 2004). Effectivement, cette technique permet de rapidement déceler des interactions potentielles avec une cible pour laquelle aucun test biochimique ou cellulaire n'est établi.

### ***1.2.2.1. La sélection des composés***

Il est reconnu qu'en adoptant une méthode axée sur la qualité des composés à cribler plutôt que de tester une grande quantité est un facteur prédominant dans l'identification des composés actifs qui peuvent franchir les différentes phases cliniques (Burford et al., 2011). Une approche par criblage virtuel à haut débit (vHTS) dans la conception et le développement de médicaments est souhaitable pour permettre d'identifier de nouveaux composés comme candidats potentiels (Bajorath, 2002). En effet, cette approche offre plusieurs avantages. Tout d'abord, elle permet d'optimiser un ligand, en lui apportant des modifications subtiles, sans avoir recours à la synthèse chimique de ces composés pour l'analyse par HTS (Bleicher et al., 2003). Ainsi, uniquement les composés possédant une interaction prédite supérieure au composé de base pourront être synthétisés chimiquement pour de futures analyses biochimiques. De plus, plusieurs banques de composés virtuels, telles que PubChem, permettent de filtrer et retenir uniquement les composés qui possèdent un bon potentiel de biodisponibilité (Drews, 2000). Cette caractéristique d'un composé peut être évaluée par la règle de Lipinski qui se base sur l'analyse des propriétés moléculaires d'un composé pharmacocinétique (Kitchen et al., 2004). Cette règle permet approximativement de discerner entre un composé ayant un bon potentiel pharmacologique par opposition à un composé qui serait difficilement absorbable, distribué ou encore métabolisé par l'entremise de quatre critères. Tout d'abord, le poids moléculaire des composés doit être inférieur à la barrière de 500 daltons au-delà de laquelle la perméabilité est grandement compromise. De plus, afin de maintenir la perméabilité des composés au travers de la membrane lipidique, le nombre maximum de 10 accepteurs et 5 donneurs de ponts hydrogènes a été fixée. Finalement, les composés doivent satisfaire une contrainte de rapport de solubilité dans deux solvants, soit l'octanol et l'eau. L'application de cette règle aux composés virtuels permet de réduire la tâche informatique aux composés possédant un bon potentiel pharmaceutique (Klebe, 2006). Malgré l'avantage certain de la règle de Lipinski afin d'alléger les calculs informatiques, il a été démontré que certaines drogues thérapeutiques n'obéissent pas aux contraintes de

la règle de Lipinski. En effet, cette règle n'est pas à toute épreuve et son application peut retirer de l'équation des composés pharmacologiques potentiels.

#### ***1.2.2.2. Le calcul des conformations des ligands***

Les premiers logiciels de criblage virtuel considéraient uniquement les propriétés 2D des composés. Par contre, cette approche s'est avérée rapidement infructueuse compte tenu de la représentation graphique des composés très différente de la réalité (Bajorath, 2002). L'efficacité du criblage virtuel a été nettement améliorée lorsque les calculs informatiques ont intégré les propriétés 3D des différents ligands. Les logiciels sont en mesure de calculer plusieurs conformations potentielles du ligand afin de permettre la flexibilité retrouvée dans un système enzyme-protéine (Ferreira et al., 2010). La méthode traditionnelle repose sur la décortication de la structure du composé en grille, disposé de façon à pouvoir pivoter l'une sur l'autre en différent point de la structure (Kitchen et al., 2004). Cette disposition définit sur le composé des points de flexibilité qui permettent au composé d'adopter différentes conformations. Dans un premier temps, lors d'un essai d'arrimage moléculaire, une première partie du composé est mis de façon rigide sur la protéine au site d'arrimage. Par la suite, les parties restantes du composé sont ajoutées dans les prédictions informatiques par une succession de combinaisons favorables. À chaque ajout de segment de grille du composé, l'énergie est calculée et les molécules partiellement construites sont classées (Klebe, 2006). Cette procédure est répétée jusqu'à la construction complète de la molécule. Le calcul informatique est ensuite reproduit pour chaque composé afin d'obtenir plusieurs essais d'arrimage moléculaire distincts. Il est déterminant pour le succès de l'essai d'arrimage moléculaire que le positionnement de départ soit bien réalisé. Ainsi, les différents logiciels d'arrimage moléculaire, tels qu'AutoDOCK, Dock, FlexX et Gold, n'adoptent pas la même méthodologie de sélection du premier segment du composé à fixer sur la protéine. Malgré ces nuances, tous ces programmes sélectionnent une conformation adoptée par le ligand avec une très basse énergie (Shoichet, 2004).

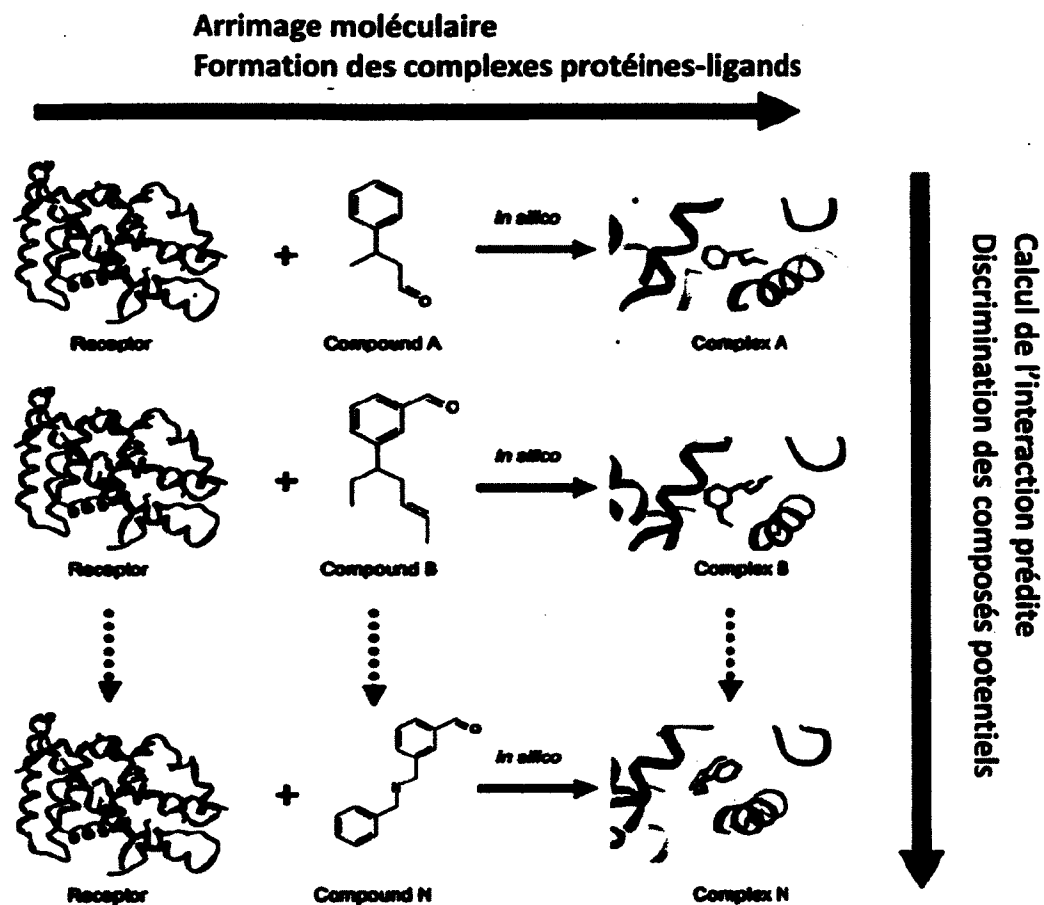


Figure 2. Schéma récapitulatif des étapes du criblage virtuel (adapté de Cavasotto et Orry, 2007). Schéma permettant de discriminer entre les deux étapes du criblage virtuel, soit l'arrimage moléculaire et l'étape de calcul de l'interaction prédite. L'étape d'arrimage moléculaire permet de positionner le ligand dans le site d'interaction de la protéine, et ce, selon l'orientation la plus favorable. Par la suite, le logiciel calcule l'énergie prédite d'interaction ce qui permet de classer les composés selon une affinité croissante.

### 1.2.2.3. L'acquisition des structures protéiques

Une fois les ligands sélectionnés et leurs multiples conformations analysées, le vHTS nécessite des données informatiques sur la protéine cible. Depuis 2003, ces informations sont offertes au public via le site Protein Data Bank (PDB) ([www.rcsb.org](http://www.rcsb.org)). Effectivement, ce site recueille plus de 70 000 structures tridimensionnelles de protéines

et d'acides nucléiques obtenus par cristallographie aux rayons X ou par résonance magnétique nucléaire. De plus, ces structures sont souvent accompagnées de ligands qui permettent d'obtenir des informations pertinentes sur le site actif des enzymes. Lorsque la protéine d'intérêt possède un cristal accessible sur PDB, il suffit d'ajouter sa structure au logiciel d'arrimage moléculaire. Par contre, lorsqu'aucun cristal n'est disponible, il est possible de procéder à une modélisation comparative. Cette méthode nécessite toutefois la structure tridimensionnelle d'une protéine partageant une forte homologie à celle d'intérêt (Despins et al., 2010). Le logiciel Swiss-PDB permet d'inférer la structure de la protéine cible en comparant le cristal connu et les deux séquences primaires des peptides. La précision de la structure obtenue dépend principalement du degré d'homologie avec la protéine de référence. Parfois, le degré de similitude n'est pas assez élevé pour inférer une structure prédite ce qui empêche son utilisation pour le vHTS.

#### ***1.2.2.4. Le calcul de l'interaction prédite***

Les calculs réalisés par les programmes d'arrimage moléculaire se situent à deux niveaux, soit le calcul de recherche de conformation possible du ligand et le calcul de la force d'interaction prédite entre le composé et la cible (Pencheva et al., 2010). De façon générale, l'interaction est en fonction de la complémentarité stérique des groupements chimiques qui sont à proximité. Plus précisément, plusieurs types d'interactions intermoléculaires permettent de quantifier l'affinité, tel que les ponts hydrogènes, les liaisons hydrophobes et les liaisons van der Waals (Morris et al., 1998). Les forces van der Waals sont importantes à considérer dans un complexe protéine ligand, puisqu'elles regroupent un ensemble de liaisons faibles, d'attraction ou de répulsion, faisant intervenir des dipôles (Kumar et Nussinov, 2002). Un autre type d'interaction fréquente, mais de plus forte intensité que les liaisons van der Waals sont les liaisons hydrogène. Ces dernières se produisent lorsqu'un ion hydrogène est partagé entre un accepteur et un donneur ce qui forme un pont hydrogène (Kumar et Nussinov, 2002; Schramm, 2005). Un autre type d'interaction est produit par l'attraction électrostatique de deux charges sur des groupements chimiques ioniques. Ce type de liaison moléculaire qui engendre de

fortes interactions peut être représenté par la formation d'un pont salin produit par l'interaction entre des résidus positifs, tels que l'arginine, et des résidus négatifs, tels que l'asparagine (Rousseau et al., 2005). De plus, les logiciels peuvent également calculer des interactions faisant intervenir une attraction par des molécules hydrophobes. Le calcul d'interactions prédites entre un composé et une protéine peut utiliser plusieurs approches telles que : le calcul par complémentarité stérique, le calcul par scores empiriques ou le calcul par scores statistiques (Morris et al., 1998). L'approximation de l'énergie libre de la liaison protéine-ligand par la méthode de complémentarité stérique fait intervenir les données physiques de la thermodynamique afin d'identifier l'emplacement le plus compatible pour le complexe protéine ligand (Taylor et al., 2002). D'autre part, les fonctions d'évaluations empiriques calculent la sommation des interactions intermoléculaires, telles que les liaisons hydrogène, les interactions ioniques et les interactions hydrophobes. Selon cette méthode, utilisée par le logiciel AutoDock, il est donc possible de classer les ligands selon leur probabilité d'interaction avec la cible basée sur l'ensemble des liaisons prédites (Pencheva et al., 2010). Une méthode d'évaluation d'interaction moins prisée par les logiciels de criblage moléculaire repose sur le calcul statistique d'interaction. Ce principe d'évaluation d'affinité se base sur les interactions décrites par des études expérimentales pour chaque ligand (Shoichet, 2004). Ensuite, le logiciel positionne chaque ligand de façon à statistiquement favoriser le plus de liaisons décrites par des études expérimentales. Bien que plusieurs autres fonctions d'évaluation d'interaction soient disponibles, une méthode efficace doit posséder deux qualités : soit un positionnement précis du ligand et une rapidité de calcul d'interaction de l'ordre de quelques minutes par ligand (Klebe, 2006).

Le présent chapitre résume les caractéristiques de deux stratégies de recherche d'inhibiteurs ainsi que leurs avantages et inconvénients. Toutefois, avant de mettre en application l'une de ces stratégies, il est primordial de faire un survol sur les différentes formes d'inhibition. Les diverses facettes selon lesquelles un inhibiteur se lie sur une protéine et influence son activité sont abordées dans la prochaine section.

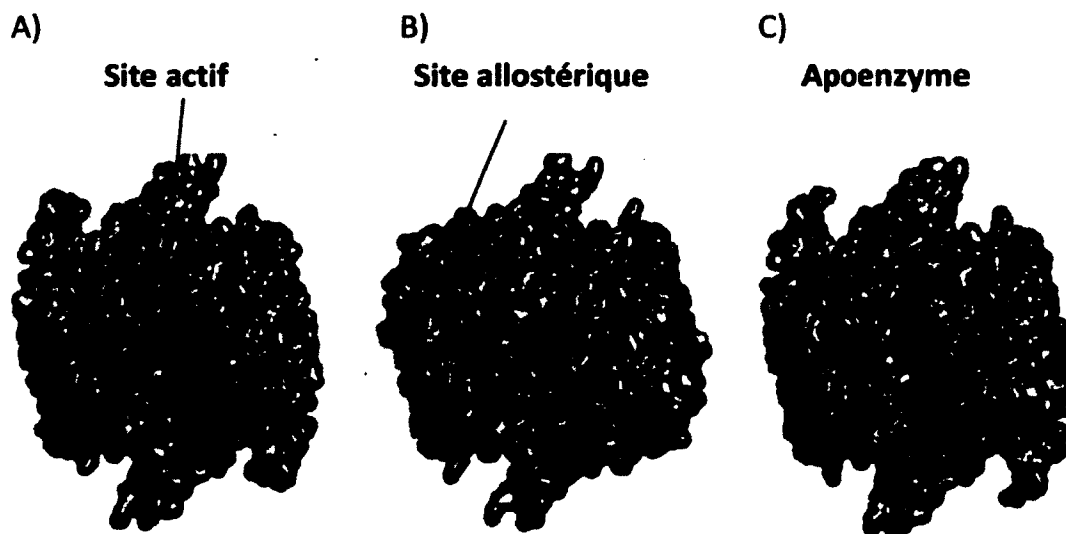
### *1.3. Les différents mécanismes d'action des inhibiteurs*

Une enzyme possède une structure tridimensionnelle au sein de laquelle on retrouve une cavité formant le site actif. La géométrie moléculaire formant le site actif permet la reconnaissance du substrat spécifique de l'enzyme avant d'entamer la réaction enzymatique (Carter, 2000; Issur et al., 2009a). Suite à la liaison du substrat, l'enzyme catalyse la conversion du substrat en produit, lequel est relâché avant de recommencer le cycle. L'inhibition enzymatique se produit selon deux types distincts : soit une inhibition non réversible ou réversible (Carter, 2000). Dans le premier cas, lors d'une inhibition non réversible, l'activité catalytique est compromise par un lien covalent près du site actif ou par la dénaturation de l'enzyme. Lors d'une inhibition réversible, l'inhibition peut-être soit compétitive, par la liaison de l'inhibiteur au site actif ou non compétitif, lorsque l'inhibiteur se lie ailleurs qu'au site actif.

#### **1.3.1. L'inhibition par des analogues du substrat**

Le site actif d'une enzyme est, dans la majorité des cas, au fond d'une cavité qui est délimitée par plusieurs groupements fonctionnels permettant la liaison au substrat (Issur et al., 2009a). Plusieurs études, portant sur des enzymes cellulaires et virales, ont caractérisé les déterminants moléculaires impliqués dans la reconnaissance du substrat. C'est justement en identifiant des composés sélectifs qui ont ouvert la porte afin d'utiliser certains analogues de substrat pour potentiellement limiter la prolifération de pathogènes (Despins et al., 2010). En apportant des modifications chimiques sur les substrats, il est possible d'identifier des composés possédant un bon potentiel inhibiteur d'une enzyme de pathogène. Par exemple, les analogues de nucléotides possèdent plusieurs caractéristiques clé pour leur utilisation à des fins thérapeutiques : une barrière de résistance élevée, une activité antivirale puissante et une bonne biodisponibilité (Fung et al., 2011). Notamment, les premiers inhibiteurs du virus de l'immunodéficience humaine (VIH) étaient des analogues de nucléotides tels que l'Acyclovir et le Famcyclovir (Freeman et Gardiner, 1996). Afin que ces analogues de nucléotides soient reconnus par la transcriptase inverse

lors de la synthèse d'acides nucléiques, ces composés doivent être phosphorylés par des kinases cellulaires (De Clercq, 2004a). Leur sélectivité pour le VIH reposait sur l'inhibition plus prononcée de la transcriptase inverse virale comparativement aux enzymes cellulaires liant les nucléotides. Dans ce même ordre, des études ont mis de l'avant le potentiel inhibiteur du foscarnet, un analogue du pyrophosphate, sur certaines enzymes virales (Freeman et Gardiner, 1996; Yaari et al., 2010). Ce composé possède un potentiel inhibiteur sur les polymérase virales du VIH et les herpèsvirus. Le foscarnet inhibe l'activité polymérase de ces virus en empêchant l'échange du groupement pyrophosphate lors du cycle de polymérisation du virus (Helgstrand et al., 1978; Wagstaff et Bryson, 1994). La comparaison du potentiel inhibiteur de ce composé sur les polymérase virales, ainsi que sur leurs homologues cellulaires, a permis de mettre de l'avant la sélectivité de ce composé sur les enzymes issus des pathogènes. Par contre, leur utilisation demande des études subséquentes afin d'identifier les risques de toxicité qui pourraient être associés à différentes voies métaboliques.

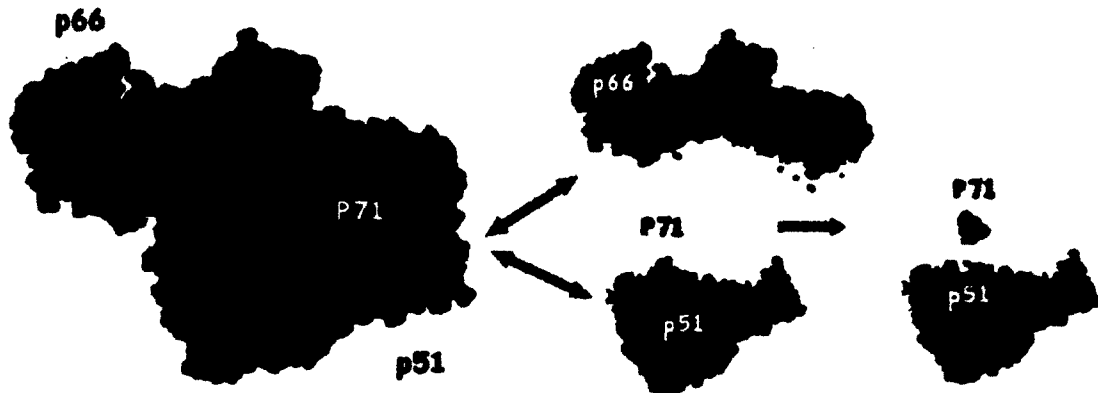


**Figure 3. Schématisation de la structure tertiaire d'une enzyme et des diverses cavités (adapté de Scheer et al., 2006). A) La liaison du substrat sur une enzyme se produit au sein d'un site de liaison, appelé site actif, qui est tapissé par plusieurs acides aminés impliqués dans la fixation du substrat et dans le mécanisme de la catalyse. Ce site spécifique est schématisé par une cavité rose ou le substrat, molécule jaune, est lié. B) L'enzyme peut contenir d'autres sites de liaison, site allostérique, reconnaissant différents métabolites, tels que représenté par un composé blanc. Leur liaison peut induire un changement de conformation au site actif et affecter la réaction enzymatique. C) La protéine seule, exempte de tout substrat porte le nom d'apoenzyme.**



### 1.3.2. Les interactions protéines-protéines

Au sein des voies biologiques et métaboliques essentielles dans une cellule interviennent des interactions protéiques nécessaires pour le bon déroulement de la réaction enzymatique. Une étude récente a identifié que près du tiers des interactions protéiques, dénombrées à plus de 100 000 dans une cellule, peuvent être la cible d'une thérapie (Hajduk et al., 2005). Ce type de thérapie a été validé par l'utilisation d'anticorps ou de petits peptides dont certains sont devenus de puissants inhibiteurs (Whitty et Kumaravel, 2006). Malgré la grande surface partagée par des interactions protéiques, il a été démontré que seulement quelques acides aminés favorisent l'association du complexe protéique et qu'uniquement ces acides aminés peuvent être la cible du composé. Les principaux obstacles d'une telle thérapie reposent sur le principe que le composé utilisé pour entraver l'interaction protéique est très petit comparativement à l'antagoniste. De plus, cette thérapie se confronte au problème que le composé est inefficace lorsque le complexe protéique est déjà formé (Sperandio et al., 2008).



**Figure 4. Inhibition basée sur les interactions protéines-protéines (adapté de Depolier et al., 2005). Des études structurales portées sur la transcriptase inverse (RT) du VIH ont illustré les différentes interactions protéines-protéines cruciales pour assurer la fonction de polymérase. En se basant sur ce principe, il est possible d'utiliser un petit peptide, ci-contre nommé P71, se liant au site d'hétérodimérisation des deux sous-unités de la RT, soit P51 et P66. Le peptide lié permet de déstabiliser la conformation de dimère et de favoriser la dissociation.**

### **1.3.3. L'inhibition allostérique**

Des mécanismes alternatifs d'inhibition, ne ciblant pas directement le site actif, sont une stratégie intéressante pour augmenter la spécificité de l'inhibiteur pour une cible précise (Yang et al., 2011). Effectivement, cette spécificité découle du degré de conservation des enzymes au travers de l'évolution. Afin qu'une enzyme conserve son activité intrinsèque, une pression évolutive est exercée sur certains acides aminés de l'enzyme définissant ainsi les motifs conservés d'une famille d'enzyme. À l'extérieur de ces motifs conservés, les membres d'une même famille d'enzyme sont souvent très peu identiques. Cette variabilité, tant au niveau de la structure tridimensionnelle qu'au niveau de la séquence primaire est propice pour le design d'un inhibiteur spécifique (Tramontano, 2008). Plusieurs exemples de composés se liant ailleurs qu'au site actif ont été identifiés. Lorsque cette liaison est spécifique et puissante, elle permet d'entraver l'activité catalytique d'une enzyme. Par exemple, des études menées sur l'intégrase du VIH ont permis de caractériser un site allostérique qui a ensuite été ciblé afin de développer une thérapie exploitant ce site allostérique (Schames et al., 2004).

### ***1.4. Les cibles pour traitement pharmacologique***

Dans le contexte actuel d'émergence de souches résistantes, le traitement des maladies infectieuses se doit d'envisager des médicaments aux mécanismes d'action différents (MacCallum et al., 2010). Le développement de traitements spécifiques pour un pathogène est ardu en ce sens que plusieurs composés très actifs pour inhiber la réplication du pathogène affectent également les processus de l'hôte (Wutzler et Thust, 2001). De plus, certains organismes pathogènes, tels que les virus à ARN, possèdent une capacité d'évolution très rapide ce qui est directement proportionnel avec l'incorporation de mutation au sein des génotypes (Das et al., 2010). Cette capacité pour une évolution rapide se traduit par la survie du virus en présence d'antiviraux qui inhibaient sa prolifération dans le passé. La clé du succès pour traiter les maladies infectieuses

auxquelles la médecine moderne ne possède pas encore de traitements est d'identifier des mécanismes novateurs (Perea et Patterson, 2002). Parmi la vaste gamme de nouvelles cibles potentielles, uniquement quelques stratégies seront mentionnées dans ce chapitre.

#### **1.4.1. Les antifongiques**

Le développement d'antifongiques se heurte à un obstacle de taille. Une thérapie utilisant des antifongiques se doit de développer une drogue qui cible précisément une caractéristique unique aux levures ciblées. Ceci peut être ardu à identifier considérant que les levures font parties du règne eucaryote et partagent plusieurs chemins réactionnels identiques avec l'humain (Ghannoum et Rice, 1999). Par contre, les gènes qui sont essentiels pour la survie cellulaire et qui diffèrent de l'homologue humain sont potentiellement de bonnes cibles pour le développement de drogue (Qiao et al., 2008). C'est dans cet ordre d'idée que certaines recherches ont mis de l'avant l'importance des interactions entre l'hôte et le pathogène.

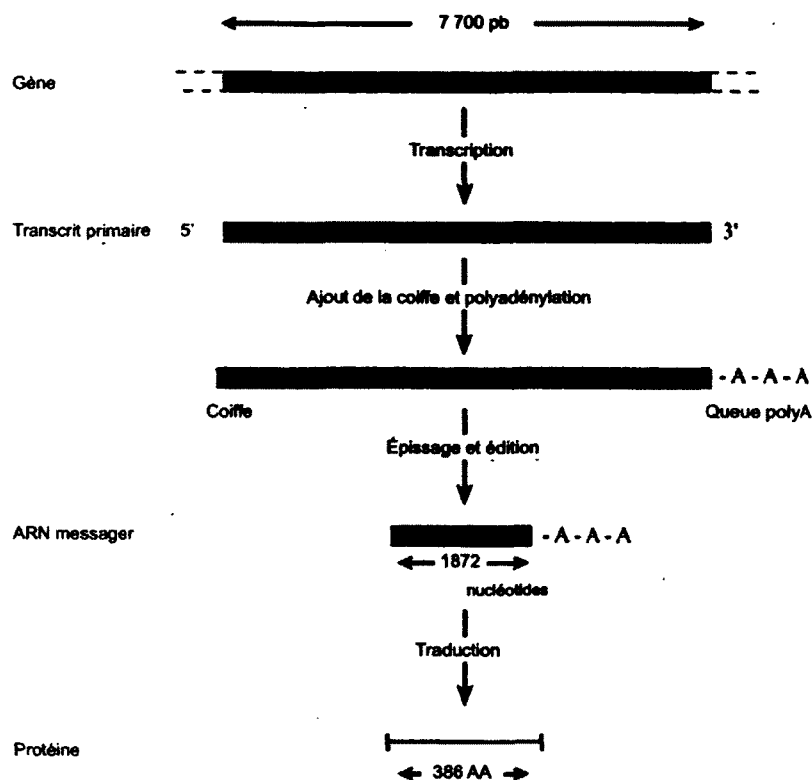
#### **1.4.2. Les protéines de l'hôte**

Le développement d'un traitement est parfois problématique pour certaines infections. C'est notamment le cas du traitement des infections causées par les virus à ARN compte tenu de leur évolution rapide qui permet l'émergence de souches résistantes (Pommier et al., 2005). Les inhibiteurs qui ciblent les protéines virales sont sujets à perdre leur efficacité plus le virus accumule des mutations qui contrecarrent leurs potentiels inhibiteurs. Dans cette optique, des recherches se sont intensifiées afin de cibler des protéines de l'hôte qui réprimeraient la réplication virale (Geller et al., 2007). Étonnamment, cette stratégie mena également à une résistance de la propagation virale. Des études portant sur l'inhibition de la composante cellulaire responsable de la voie de sécrétion nécessaire pour la réplication du virus poliovirus a identifié une molécule prometteuse, le Brefeldin A (Ramachary et Vijayendar Reddy, 2010). Par contre, des

variantes virales du poliovirus résistantes à cet inhibiteur ont été identifiées rapidement. Ce phénomène de résistance a également été reproduit pour le VIH et le HSV. Dans ces cas, les virus peuvent acquérir une résistance en développant de nouveaux modes de sécrétions qui utilisent des facteurs cellulaires alternatifs, non ciblés par les inhibiteurs. Toutefois, une nouvelle cible semble prometteuse afin d'éviter ce mécanisme de résistance; la protéine chaperonne Hsp90. Cette dernière est nécessaire au bon repliement de la structure tertiaire des protéines virales (Sun et al., 2010). En présence d'un inhibiteur de cette chaperonne, la geldanamycin qui bloque l'hydrolyse de l'ATP de Hsp90, le titre intracellulaire de virions est diminuée selon une fonction dose-réponse (Geller et al., 2007). L'inhibition de la protéine chaperonne compromet le repliement de la protéine de la capsid ce qui perturbe la réplication des picornavirus. Jusqu'à ce jour, aucun virus résistant à l'inhibition de Hsp90 n'a pu être isolé (Joshi et Stoddart, 2011). Cette nouvelle cible potentielle démontre que de nouvelles stratégies peuvent contrer l'émergence de souches résistantes.

#### **1.4.3. Les protéines de la maturation des ARNm**

Au sein du dogme central de la biologie moléculaire, c'est-à-dire la transmission de l'information génétique de l'ADN contenu au noyau vers la synthèse protéique au cytoplasme, se retrouve la molécule d'ARN (Shuman, 2002). Ce dogme définit deux mécanismes fondamentaux, soit la transcription et la traduction. La première étape de l'expression de l'information génétique consiste à transcrire un gène en une molécule d'ARN qui sera ensuite traduit en protéine (Coller et Parker, 2004). Ce rôle de médiateur constitue le premier rôle de l'ARN que l'on nomme alors ARN messager. Cet état intermédiaire subit plusieurs modifications essentielles pour permettre leur traduction : l'ajout de la structure coiffe, l'édition, l'épissage et l'ajout de la queue polyadénosines (Gu et Lima, 2005). Toutes ces modifications sont une cible très prometteuse dans un contexte thérapeutique compte tenu de leur importance pour la survie cellulaire (Issur et al., 2011).



**Figure 5.** Les différentes étapes de la maturation d'un ARN messagers (adapté de Lionnet et Croquette, 2005). Pendant la transcription, la molécule d'ARN synthétisée subit plusieurs modifications. Les deux extrémités du transcrit sont modifiées, par l'ajout de la coiffe à l'extrémité 5', et l'ajout de plus de 200 résidus d'adénosines monophosphates à l'extrémité 3'. De plus, le transcrit subit également des modifications qui modulent la séquence interne de nucléotides, soient l'épissage et l'édition. Un gène est constitué de portions codantes, les exons représentés par les tronçons roses, et de portions non codantes, les introns de couleur bleue. La maturation de l'ARNm comporte donc des phases d'excisions des introns qui se déroulent dans le noyau avant d'exporter le transcrit au cytoplasme.

### ***1.5. La maturation des ARNm***

Un point central des processus biologiques est le transfert de l'information génétique encodée par l'ADN en protéine. Pour ce faire, une molécule d'ARN est synthétisée par une ARN polymérase et est transférée au cytoplasme pour être traduite par les ribosomes en protéines. La transcription de cet ARN pré-messager est couplée à plusieurs

modifications essentielles qui permettent de convertir le transcrit en ARNm afin qu'il soit efficacement reconnu par les complexes protéiques responsables, par exemple, de son transport et de sa traduction (Parker et Song, 2004). La maturation d'un ARNm débute par l'addition d'une structure coiffe en 5', suivi du processus d'épissage, d'édition et finalement l'ajout d'une queue polyA en 3' de l'ARNm.

### **1.5.1. L'ajout d'une structure coiffe en 5'**

La maturation d'un ARN pré-messager débute par l'ajout d'une structure coiffe ( $m^7GpppN-$ ) à son extrémité 5' (Schoenberg et Maquat, 2009). Cette structure est obtenue par l'ajout d'une guanosine méthylée en position N7 sur le premier nucléotide d'un ARN nouvellement transcrit par un lien 5'-5' triphosphate (Shuman, 1995). Cette configuration moléculaire représente les déterminants essentiels afin d'accomplir les divers rôles de la structure coiffe. Chez les eucaryotes et certains virus, l'addition de la structure coiffe se produit de façon co-transcriptionnelle par la succession de trois réactions enzymatiques conservées au fil de l'évolution (Ross, 1996). Cette modification est ajoutée sur un transcrit synthétisé par l'ARN polymérase II de 20 à 30 nucléotides de longueur (Cochrane, 2009) (Hirose et Manley, 2000). Dans la première réaction enzymatique, une ARN triphosphatase (RTase) hydrolyse le phosphate terminal à l'extrémité 5'-triphosphate d'un ARNm, formant ainsi une extrémité diphosphorylée (Proudfoot et al., 2002). Cette extrémité sert de substrat pour la deuxième activité enzymatique, soit l'ARN guanylyltransférase (GTase). Cette réaction se produit en deux temps, soit premièrement la formation du complexe covalent enzyme-GMP suivi du transfert de ce nucléotide sur l'ARN diphosphorylé en formant un lien 5'-5' triphosphate (Souliere et al., 2008). Dans la dernière étape, l'ARN guanine N-7-méthyltransférase ajoute un groupement méthyle sur l'azote 7 de la guanine nouvellement ajoutée (Topisirovic et al., 2011). La formation de la structure coiffe est la première modification co-transcriptionnelle d'un ARNm naissant pour permettre sa protection contre les exoribonucléases, son transport au cytoplasme, son épissage ainsi que sa traduction en protéine (Cramer et al., 2001). Pourtant, cette structure de base peut être modifiée par

l'ajout d'un ou plusieurs groupements méthyles. Effectivement, la structure coiffe minimale est appelé type 0 ( $m^7GpppN\dots$ ), alors qu'il est possible d'observer des structures coiffe de type 1 ( $(m^7GpppN_2'OMe pNp\dots)$ ), de type 2 ( $m^7GpppN_2'OMe pN_2'OMe p\dots$ ) selon leur niveau de méthylation sur les nucléotides subséquents (Varani, 1997).

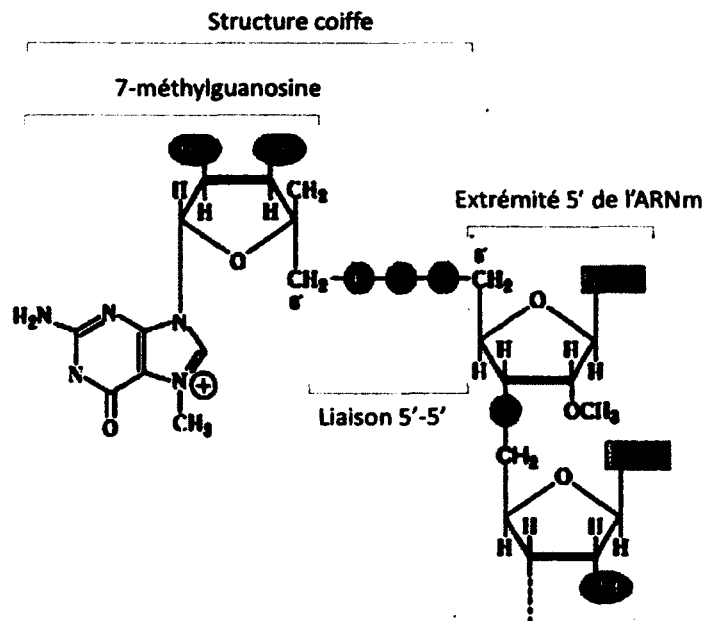


Figure 6. Structure de la coiffe (University of Tokyo). Cette première modification se produit sur les transcrits primaires et consiste en l'hydrolyse de l'extrémité 5' générant une extrémité diphosphorylée avant l'ajout d'une guanosine liée par une liaison triphosphate inhabituelle 5'-5'. La structure coiffe se finalise par l'ajout d'un groupement méthyle sur le cycle imidazole de la guanine terminale sur l'azote 7.

### 1.5.2. Le processus d'édition

Le processus d'édition a été établi pour la première fois en 1986 alors qu'une séquence de plusieurs uridines était ajoutée dans un site spécifique de l'ARNm du cytochrome oxydase 2 des trypanosomes (Feagin et al., 1988). L'investigation approfondie de cette insertion révèle qu'elle reconstitue le bon cadre de lecture et qu'elle est une modification indispensable. Depuis l'identification de cette modification portée sur un ARNm,

différents mécanismes d'altération d'une molécule d'ARN ont été identifiés parmi le règne végétal, animal, fongique de même que chez certains virus (Knoop, 2011). Les phénomènes d'édition englobent tous les processus qui modifient les séquences nucléotidiques de façon spécifique par l'insertion, la délétion ou la conversion de nucléotides. Depuis les deux dernières décennies, l'identification de nouveaux mécanismes d'édition a permis de reconnaître l'édition comme une modification post-transcriptionnelle importante pour la diversité moléculaire (Maas et Rich, 2000). Chez les mammifères, l'édition des ARNm la plus fréquente est la substitution de l'adénosine en inosine (qui fonctionne comme une guanosine) et la conversion de la cytidine en uridine (Pullirsch et Jantsch, 2010). Ces conversions sur un ARNm impliquent un changement d'acide aminé ponctuel sur la protéine ce qui entraîne dans la majorité des cas une altération de la fonction protéique. Ces modifications biochimiques sont l'œuvre d'une enzyme qui reconnaît son substrat soit par la séquence nucléotidique primaire telle que la cytidine désaminase ou soit par une structure tertiaire dans le cas de l'adénosine désaminase. Pour la conversion d'une cytidine en uridine, un dimère de la protéine apobec-1, possédant l'activité catalytique métal-dépendante, doit être assemblé dans le complexe de l'éditosome (Shikanai, 2006). L'éditosome comprend les facteurs qui positionnent le complexe sur la séquence primaire d'amarrage par complémentarité de séquence. L'édition de l'adénosine implique l'activité catalytique d'une adénosine désaminase qui reconnaît le site d'édition avec une tige double brin d'ARN (Shikanai, 2006). Les bases moléculaires permettant ce processus d'édition restent à être identifiées, mais des évidences semblent indiquer qu'une histidine médie le transfert de proton sur l'adénosine avant la catalyse (Stuart et Panigrahi, 2002).



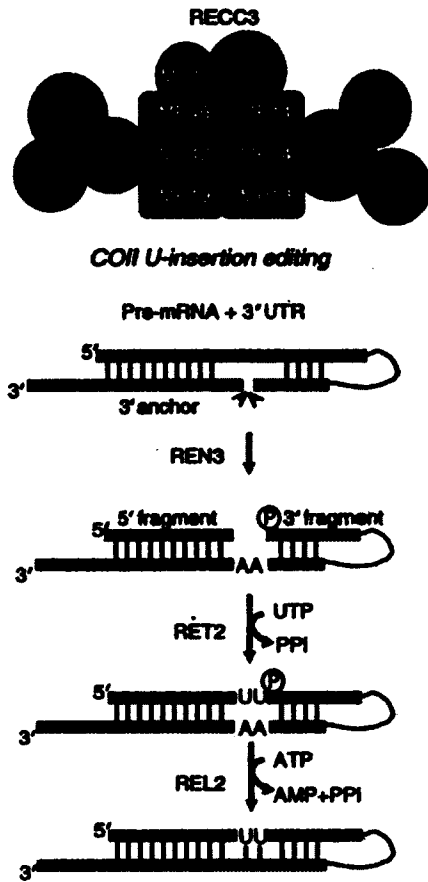
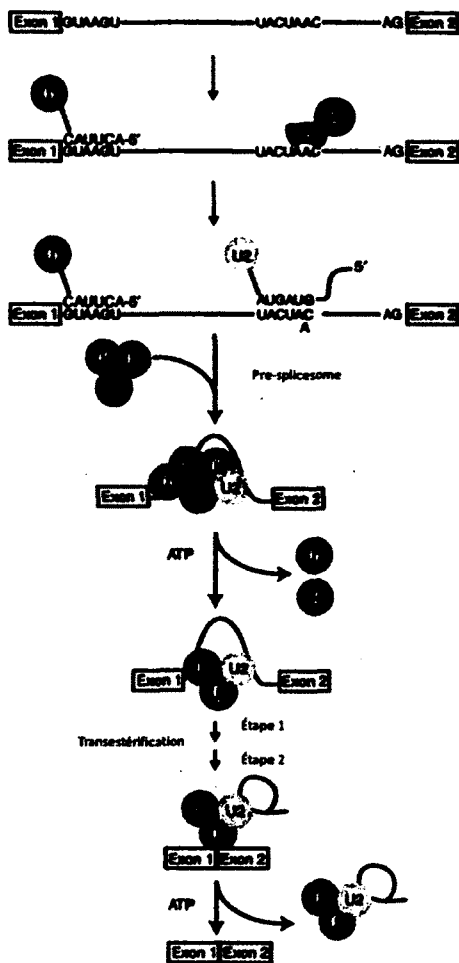


Figure 7. Mécanisme de l'édition et composants du complexe de l'éditosome chez *Trypanosoma brucei* (adapté de Chateigner-Boutin et Small, 2011). Le site de clivage et le nombre d'uridine ajoutée sont spécifiés par de petits ARNs. Ils interagissent avec leur ARNm cible par complémentarité de séquence en formant un duplex d'ARN. La réaction d'édition se déroule par le clivage de l'ARNm par une endonucléase composant le complexe protéique de l'éditosome. Par la suite, l'ajout d'uridine est réalisé par l'uridylyl-transférase et est complété par la ligation de l'ARNm.

### 1.5.3. L'épissage

La plupart des ARN pré-messagers synthétisés possèdent des séquences codantes (exons) intercalées par des séquences internes non codantes (introns). Ces dernières doivent donc être retirées durant le processus d'épissage pour permettre la jonction des exons

(Proudfoot et al., 2002). Le processus d'épissage se déroule dans le noyau au sein du complexe nucléoprotéique appelé splicéosome (Parmley et al., 2007). Le mécanisme permettant de cliver les introns et de joindre les deux exons adjacents est obtenu par deux trans-estérifications se déroulant dans le splicéosome. Les composantes protéiques du complexe splicéosome sont classées sous deux types de facteurs d'épissage, soit les hnRNPs et les snRNPs (U1, U2, U4, U5 et U6) (Gu et Lima, 2005). Ces derniers interagissent avec des éléments consensus situés soit au site d'épissage 5', c'est-à-dire à la jonction exon-intron (AG/GURAGU (R : purine, Y : pyrimidine), au site d'épissage 3', c'est-à-dire à la jonction intron-exon (YAG/RNNN), ou soit au site de formation du lariat (CURAY) situé à plus de cent nucléotides en amont de fin de l'intron. L'évènement d'épissage est initié avec l'association de U1 au site d'épissage 5', qui permet le recrutement des autres composantes du splicéosome, et l'interaction de U2 avec le site consensus de formation du lariat (Parmley et al., 2007). Cet appariement permet de mettre en évidence l'adénosine, n'interagissant pas avec U2, et ainsi favoriser l'attaque nucléophile qu'elle exerce au site d'épissage 5'. Une première trans-estérification du 2'OH de l'adénosine sur la jonction exon-intron engendre la séparation de l'exon en 5' du lariat formé (Proudfoot et al., 2002). La protéine charpente U4 achemine U5 au site d'épissage 3' et dirige U6 pour remplacer U1 au site d'épissage 5'. Enfin, la deuxième réaction de trans-estérification se déroule lorsque l'extrémité 3'OH libre de l'exon 5' attaque l'exon situé en aval. Cette attaque permet de joindre les deux exons et de relâcher le lariat résiduel (Hirose et al., 1999). Un nombre considérable d'évidences indiquent que l'activité catalytique du splicéosome est basée sur l'ARN, alors que les protéines qui y sont associées influencent la stabilité et l'assemblage (Hocine et al., 2010).

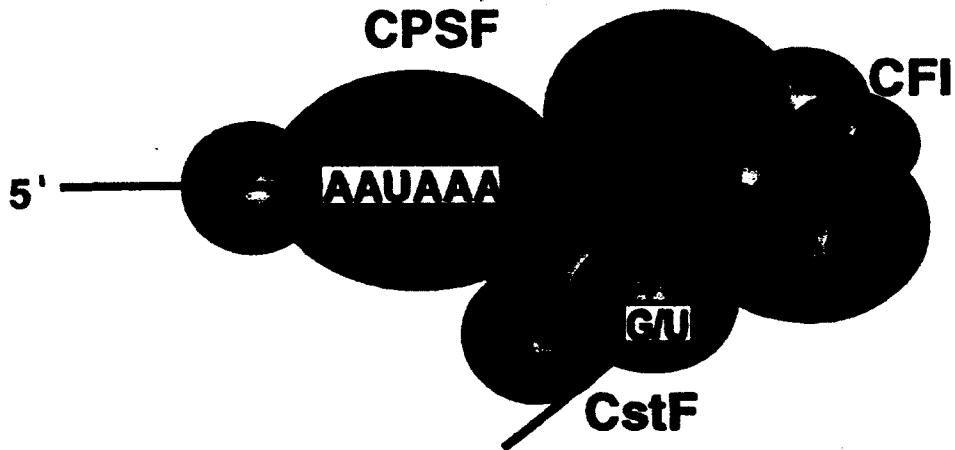


**Figure 8. Les différentes étapes de l'épissage et de l'association du splicéosome (adapté de Quin et al., 2007). L'épissage est le processus selon lequel les séquences introniques sont retirées de l'ARNm. L'assemblage du splicéosome débute avec l'alignement de bases du petit ARN U1 au site d'épissage 5', suivi du recrutement de plusieurs facteurs protéiques d'épissage et de facteurs auxiliaires.**

#### 1.5.4. La polyadénylation

Presque tous les ARNm des eucaryotes, à l'exception des transcrits d'histones, sont maturés à leur extrémité 3' par l'ajout de près de 200 molécules d'adénosines qui ne sont pas encodées par la séquence génique (Proudfoot et al., 2002). Avant de pouvoir synthétiser la queue de polyadénosines (PolyA), une séquence signal induit la coupure du

transcrit afin de le décrocher de l'ARN polymérase II (Kuhn et Wahle, 2004). Le signal de coupure sur l'ARN pré-messager se situe entre la séquence hautement conservée AAUAAA et une séquence en aval (DES : downstream sequence element) riche en U ou G/U (Colgan et Manley, 1997). Le complexe du clivage endonucléotidique requiert l'association des facteurs protéiques CPSF (cleavage/polyadenylation specific factor) et CstF au site AAUAAA et au DES respectivement pour recruter tous les autres facteurs impliqués dans la coupure du transcrit (Proudfoot et O'Sullivan, 2002). Ensuite, les facteurs protéiques CFI et CFII sont recrutés sur le transcrit afin de diriger adéquatement la coupure au site de clivage. Le complexe multimérique de la machinerie de polyadénylation est complet lorsque les deux dernières composantes s'associent, soit le CTD et la poly-(A)-polymérase (PAP), qui interagît directement avec CPSF, CFI et CFII (Bienroth et al., 1993). La PAP catalyse la coupure endonucléotidique au site de clivage, composé du dinucléotide CA, ce qui entraîne la libération des facteurs CstF, CFI et CFII. Suite à cette coupure, la polyadénylation par PAP nécessite la présence du facteur protéique CPSF, toujours associé à l'ARN pré-messager, afin de synthétiser l'ajout du substrat spécifique, l'adénosine (Flaherty et al., 1997). La processivité de cette enzyme se voit augmentée lorsque PAB II (polyA binding protein) se lie systématiquement à chaque dizaine d'adénosines (Gray et al., 2000). Une fois la longueur maximale de la queue poly(A) atteinte, l'activité de PAP s'estompe complètement avec la dissociation du complexe PAP-CPSF (Ross, 1996). Cette modification permet de terminer la synthèse du transcrit, d'épisser l'intron terminal adéquatement, de stabiliser l'ARNm tout en permettant d'augmenter sa traduction (Kuhn et Wahle, 2004).



**Figure 9.** Un modèle du complexe multiprotéique du clivage et de la polyadénylation des ARN pré-messagers (Ryan et al., 2004). Le facteur de spécificité pour le clivage et la polyadénylation (CPSF) se fixe au signal AAUAAA en amont du site du clivage. D'autres facteurs protéiques se fixent au site riche en GU ou en U en aval et induisent la formation d'une boucle de l'extrémité 3' de l'ARN pré-messager. Le recrutement de la poly(A) polymérase (PAP) stimule le clivage, qui permet de libérer les facteurs de clivage et de permettre la polymérisation.

## ***1.6. Les rôles de la structure coiffe***

### **1.6.1. La stabilité**

La dégradation des ARNm est importante notamment pour le recyclage des nucléotides ainsi que pour permettre à la cellule de répondre rapidement aux divers stimuli externes (Levy et Miller, 1998). Malgré que la régulation de la dégradation des ARNm varie énormément beaucoup d'un transcrit à l'autre, chaque ARNm doit être stabilisé de manière à pouvoir éviter la dégradation hâtive par rapport à leur cycle cellulaire. En effet, l'ajout de la structure coiffe permet de stabiliser les ARNm contre les exoribonucléases (Gu et Lima, 2005). La nature de ce pont 5'-5' triphosphate dissimule les phosphates de l'extrémité 5', ce qui les isole de l'activité catalytique des exoribonucléases 5'-3' (Garneau et al., 2007). De plus, cette structure particulière est reconnue par plusieurs complexes protéiques qui s'agglomèrent sur l'extrémité 5' des ARNm et préviennent leur accessibilité par les exoribonucléases (Coller et Parker, 2004).

### **1.6.2. Le transport**

Puisqu'un ARNm est transcrit au noyau et est traduit en protéine au cytoplasme, un processus de translocation est primordial pour permettre le transport des molécules au travers de la membrane nucléaire (Lewis et Izaurralde, 1997). Deux protéines nucléaires sont responsables de la liaison de la structure coiffée au noyau, soit CBP20 et CBP 80 (cap binding protein). La protéine CBP80 possède une séquence de localisation nucléaire en amino-terminal qui peut être reconnue par l'importine  $\alpha$  au noyau. Le complexe formé de CBP20, CBP80, l'importine  $\alpha$  et l'ARNm interagissent avec le complexe du pore nucléaire et transite au cytoplasme (McKendrick et al., 2001). Pour favoriser la dissociation du complexe CBP et permettre l'association d'un complexe important pour la traduction de l'ARNm, l'importine  $\beta$  se lie à l'importine  $\alpha$  ce qui libère ce dernier. Lorsque l'importine  $\alpha$  n'est plus liée à CBP80, le site de localisation nucléaire exposé permet au complexe CBP d'interagir avec le complexe du pore nucléaire (Dias et al., 2010).

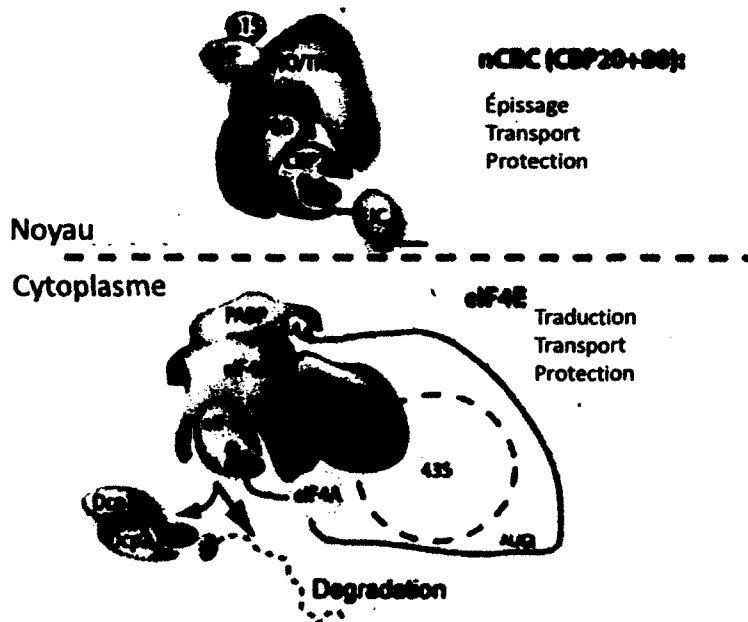


Figure 10. Les différents rôles de la structure coiffe (adapté de Topisirovic et al., 2010). La structure coiffe sert d'entité protectrice à l'extrémité 5' de l'ARNm. Elle est également nécessaire à l'exportation de l'ARNm vers le cytoplasme et à la liaison de ce dernier avec la petite sous-unité du ribosome lors de l'étape d'initiation de la traduction. La structure coiffe est reconnue par plusieurs protéines, notamment le facteur d'initiation de la traduction, eIF4e et les protéines du complexe CBC qui permettent ces différentes fonctions.

### 1.6.3. L'épissage

Plusieurs études, autant *in vitro* qu'*in vivo*, ont démontré la nécessité de la structure coiffe pour permettre un processus d'épissage efficace (Shuman, 1995). Par différentes études utilisant des ARN coiffés avec des analogues de m7-guanine combinés avec des essais de compétition avec des ARN coiffés, certaines conditions augmentaient le niveau d'ARN pré-messager contenant toujours leurs introns (Lewis et Izaurralde, 1997). À partir de ces études, une corrélation marquée lie les déterminants moléculaires de la coiffe des ARNm qui induisent une diminution de l'épissage de l'ARN avec l'incapacité à lier le complexe CBC, responsable de l'export des ARNm (Schoenberg et Maquat, 2009). Lors du

processus d'épissage, des complexes protéiques sont assemblés de part et d'autre de l'intron à exciser. Il a été démontré que l'absence du complexe CBC nuit au recrutement de la protéine snRNP U1 au site d'épissage 5' et par le fait même abolit l'épissage de l'intron juxtaposé à la structure coiffe. Par contre, la présence de la coiffe n'interfère pas avec l'épissage des introns subséquent sur l'ARN pré-messager (Parker et Song, 2004).

#### **1.6.4. La traduction**

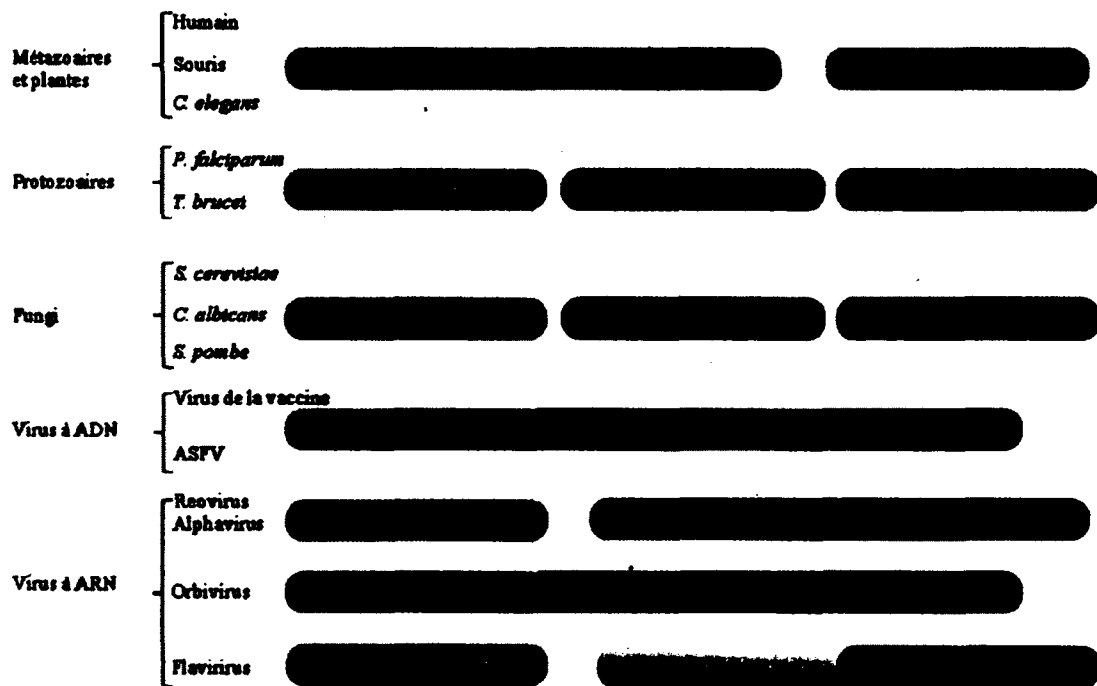
La protéine qui interagit avec la structure coiffe la plus exhaustivement étudiée est le facteur d'initiation de la traduction, eIF4E (Fischer, 2009). Cette protéine appartient au complexe protéique eIF4 qui est composé d'une hélicase à ARN (eIF4A), une protéine liant l'ARN (eIF4B), une protéine activatrice (eIF4H), la protéine charpente assemblant le complexe (eIF4G) et eIF4E (Sonenberg, 2008). Il est bien établi qu'eIF4E interagit directement avec la structure coiffe et permet de recruter l'ARNm au complexe d'initiation de la traduction. De par sa liaison aux ribosomes, via l'interaction avec le facteur eIF3, eIF4G recrute l'ARNm à la petite sous-unité du ribosome (Rhoads, 2009). Ce dernier défile sur l'ARNm à la recherche du codon d'initiation avant de recruter la grande sous-unité, l'unité catalytique, pour former le complexe permettant le cycle d'élongation (Baron-Benhamou et al., 2003). Les déterminants moléculaires de la spécificité d'eIF4E pour la structure coiffe est le résultat d'interaction de la structure coiffe avec deux tryptophanes et un acide glutamique.

#### ***1.7. Les enzymes responsables de la structure coiffe***

Cette structure coiffe est présente chez tous les eucaryotes, ainsi que chez plusieurs pathogènes tels que les virus à ADN et les virus à ARN (Shuman, 2002). Les enzymes responsables de synthétiser cette structure sont essentielles pour le développement de l'organisme, tel que démontré par le phénotype de létalité dans les souches déficientes



pour l'un des trois gènes responsables de cette synthèse (Issur et al., 2011). En effet, leur nécessité pour la survie cellulaire fut premièrement démontrée avec la GTase de *S. cerevisiae* qui, lorsqu'elle possède des mutations ponctuelles abolissant son activité, inhibe la croissance cellulaire. Par la suite, l'identification subséquente des gènes responsables des activités MTase et RTase a également permis de démontrer que ces activités sont essentielles *in vivo* pour la survie des levures. Ces travaux se sont répandus à différents organismes pour toujours constater les preuves directes de la nécessité de la structure coiffe *in vivo*. De plus, le présent chapitre permet de mettre à l'avant-plan les similarités et les différences entre les enzymes responsables de la structure coiffe chez les divers organismes.



Issur et al., 2010

**Figure 11. Organisation génique des enzymes responsables de la structure coiffe (Issur et al., 2010). Les trois enzymes responsables de la synthèse de la structure coiffe peuvent être distribuées sur un ou plusieurs peptides, dépendamment de l'organisme. Le code de couleur de cette figure illustre les enzymes utilisant des mécanismes catalytiques similaires.**

### 1.7.1. Les ARN triphosphatases

Suite à la polymérisation par l'ARN polymérase II, les transcrits d'ARN pré-messagers possèdent une extrémité 5' triphosphate (Takagi et al., 1997). Pour initier la synthèse de la structure coiffée, la première réaction enzymatique est l'hydrolyse du phosphate gamma de l'ARN par la RTase, générant une extrémité 5' diphosphate. Les enzymes des RTases ne sont pas conservées dans le règne eucaryote et la caractérisation de plusieurs d'entre elles a permis d'identifier deux familles d'enzymes : les triphosphatases métallo-dépendantes et les triphosphatases métallo-indépendantes (Tsukamoto et al., 1997).

Les RTases métallo-dépendantes sont présentes chez les eucaryotes inférieurs et les virus et sont dérivées de la famille d'enzymes des nucléosides phosphatases (NTPase). Les RTases métallo-dépendantes ont été caractérisées par des études extensives sur l'enzyme de *Saccharomyces cerevisiae*, Cet1. Des études structurales ont élucidé la structure dimérique formée de deux tunnels parallèles, composés chacun de huit feuillets antiparallèles (Ho et al., 1998a). Les cristaux ont démontré la présence de chaînes latérales polaires et hydrophiles interagissant avec les ions manganèse. Les RTases métallo-dépendantes peuvent être différenciées en différentes sous-familles structurales telles que le groupe des triphosphatases metallo-dépendantes en tunnel (TTM) et le repliement RTase/hélicase (Shuman, 2001). Comme son nom l'indique, le premier groupe est caractérisé par une structure en tunnel délimitée par huit feuillets bêta antiparallèles. Les RTases de la levure *S. cerevisiae*, du virus de la vaccine et du virus PBCV-1 appartiennent à ce groupe (Bisailon et Shuman, 2001). L'étude approfondie d'enzymes virales a permis d'identifier une classe de RTase qui se caractérise par une fusion avec la fonction hélicase, retrouvée notamment chez les flavivirus.

À l'inverse, les RTases métallo-indépendantes présentes chez les eucaryotes supérieurs, tels que les métazoaires et les plantes, ont évolué de la superfamille des cystéines phosphatases (Takagi et al., 1997). Les détails structuraux des RTases métallo-

indépendantes ont été établis avec l'enzyme de la souris, MceI. Le site actif de ces enzymes est caractérisé par le motif conservé HCxxxxR(S/T). La structure de cette enzyme a pu mettre à l'avant-plan les différences mécanistiques et structurales distinguant l'enzyme humaine et celles de plusieurs pathogènes (Changela et al., 2001). Compte tenu des différences marquées au niveau de la structure tridimensionnelle de ces deux familles d'enzymes, mais également au niveau de la spécificité du substrat et du mécanisme catalytique, les RTases pathogènes sont une bonne cible potentielle.

### 1.7.2. Les ARN guanylyltransférases

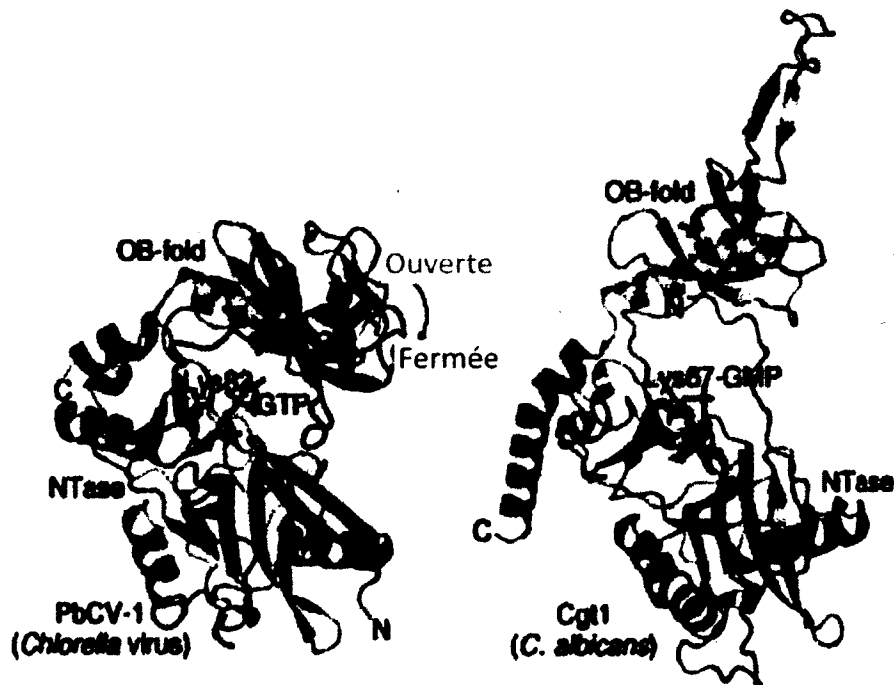


Figure 12. Les structures des ARN guanylyltransférases du virus PBCV-1 et de la levure *C. albicans* (adapté de Gosh et Lima, 2010). La figure illustre la disposition des deux domaines de l'enzyme ainsi que la position de l'acide aminé catalytique au fond du site actif.

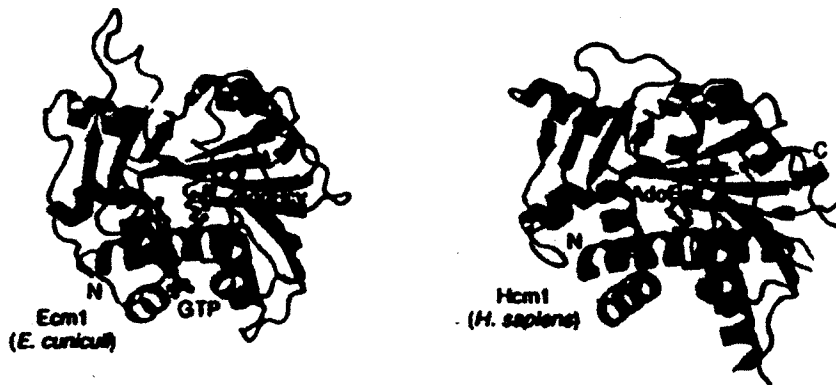
L'extrémité 5' de l'ARNm après l'activité enzymatique de la RTase est reconnue comme substrat par la deuxième activité enzymatique de la formation de la structure coiffe; la

GTase. L'activité chimique de la GTase est réalisée lorsque le résidu lysine catalytique et conservée attaque le phosphate alpha d'une molécule de GTP permettant la dissociation du pyrophosphate et la liaison covalente enzyme-GMP (Proudfoot et al., 2002). Un ARNm se lie ensuite à la protéine ce qui permet le transfert du GMP sur l'extrémité 5' diphosphorylée de l'ARN pour compléter la réaction. Les GTases sont structurellement et mécaniquement appareillées à la superfamille des nucléotidyltransférases (Schoenberg et Maquat, 2009). Les membres de cette famille partagent un mécanisme catalytique similaire qui implique la formation d'un état intermédiaire de l'enzyme liée de façon covalente à un nucléotide monophosphate. Les études cristallographiques ont permis d'identifier chez tous les membres de cette famille une architecture similaire pointant un site actif et des domaines semblables (Fresco et Buratowski, 1994). Les GTases sont les enzymes les moins bien caractérisées de toutes les enzymes responsables de la synthèse de la coiffe. Par contre, de nouvelles évidences illustrent que chez la plupart des virus à ARN, la GTase encodée diffère en fonction de la nature du substrat qu'elle reconnaît (Swift et McCammon, 2008).

### **1.7.3. Les guanine-N7-ARN méthyltransférases**

La dernière étape de la synthèse de la coiffe est réalisée par la N7-MTase qui transfère un groupement méthyle sur l'ARN nouvellement coiffé. Le cofacteur donneur de groupement méthyle est issu de la déméthylation du S-adénosylméthionine (SAM) en S-adénosylhomocystéine (SAH) (Parker et Song, 2004). Les études cristallographiques menées sur la N7-MTase (Ecm1) du parasite *Encephalitozoon cuniculi* en présence de trois différents ligands, le cofacteur SAM, le produit SAH et un analogue de la coiffe, ont dessiné le mécanisme catalytique de l'enzyme (Hausmann et al., 2002). Les structures ont révélé qu'Ecm1 possède une grande cavité au sein de laquelle se trouve deux sites de liaisons espacés : l'un pour le cofacteur donneur de méthyle SAM et l'autre pour l'accepteur de méthyle (Schwer et al., 2001). L'analyse de ces structures supporte le modèle selon lequel Ecm1 catalyse le transfert du groupement méthyle en positionnant les deux substrats dans une configuration optimale pour une attaque nucléophile

(Fabrega et al., 2004). Les résultats cristallographiques ont démontré qu'Ecml n'intervient pas au niveau de la stabilisation d'un état de transition, de l'activation du nucléophile et du relâchement des produits (Schwer et al., 2001). Des essais de mutagenèse dirigée envers plusieurs N7-MTase, dont l'enzyme humaine, ont identifié les acides aminés conservés qui permettent la coordination des divers substrats (Hausmann et al., 2002). L'ensemble des données expérimentales connues à ce jour suggère que les N7-MTases sont les enzymes de la synthèse de la structure coiffe possédant le mode catalytique le plus conservé. Les N7-MTases possèdent des structures et des motifs conservés caractéristiques de la classe I de la grande famille des méthyltransférases dépendantes au cofacteur SAM.



**Figure 13.** Structure du site actif de l'ARN guanine-N7-méthyltransférase (Gosh et Lima, 2010). Les structures tertiaires des N7-MTases de *E. cuniculi* et de l'homme sont très conservées. Le positionnement du produit de la réaction, le composé SAH ainsi qu'un nucléotide représentant le substrat accepteur de groupement méthyle permettent de discerner le site actif.

#### 1.7.4. Les ARN 2'O- méthyltransférases

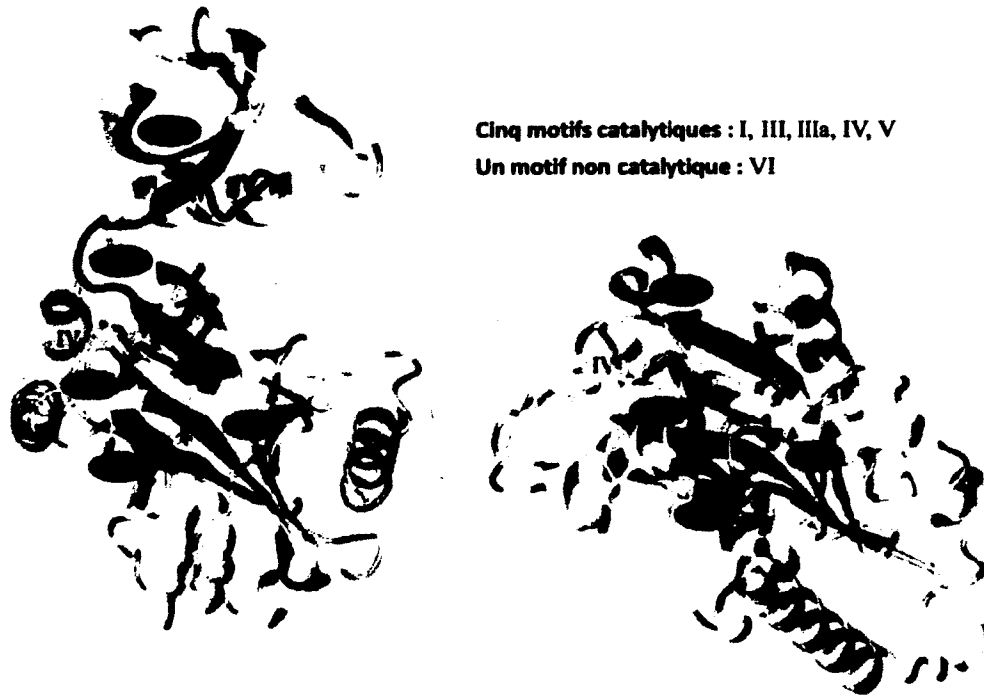
La structure canonique de la coiffe peut être agrémentée de groupement méthyle sur les nucléotides subséquents. Les différentes structures coiffe résultant de ces modifications sont appelées coiffe 0, coiffe 1, coiffe 2, en fonction du nombre de méthylations présentes sur les sucres des acides nucléiques. La proportion des différentes coiffes varie d'un organisme à l'autre, mais les eucaryotes inférieurs tendent à posséder majoritairement la

coiffe de type 0 alors qu'il s'agit de la coiffe 2 chez les eucaryotes supérieurs (Bisaillon et Lemay, 1997). Malgré que ces types de structure coiffe ont été découvert il y a 35 ans, leur fonction n'est pas complètement élucidée. Toutefois, des études récentes suggèrent qu'elles sont impliquées dans la distinction moléculaire des ARN endogènes et exogènes (Daffis et al., 2010). Jusqu'à ce jour, plusieurs 2'O-MTases ont été identifiées, autant des enzymes cellulaires que virales. Des études approfondies des 2'O-MTases du virus de la vaccine et du coronavirus SARS ont permis de décrire l'activité catalytique de ces enzymes (Chen et al., 2011; Furuichi et Shatkin, 2000; Lockless et al., 1998). De façon similaire au N7-MTase, leur site actif permet de lier le donneur de groupement méthyle SAM et de reconnaître spécifiquement un ARN possédant la structure coiffe canonique (Fechter et Brownlee, 2005).

## ***1.8. La famille des nucléotidyltransférases***

### **1.8.1. Les membres des nucléotidyltransférases**

La superfamille nucléotidyltransférase est composée des ligases à ADN, des ligases à ARN et des GTases (Shuman et Lima, 2004). Chacune de ces enzymes partage des caractéristiques chimiques et structurales qui les distinguent des autres familles de protéine. Chimiquement, chaque membre de cette superfamille possède un mécanisme catalytique similaire impliquant la formation d'un intermédiaire de réaction qui est proprement dit une liaison covalente entre l'enzyme et un nucléotide (Shuman et Schwer, 1995). Par contre, les étapes subséquentes varient selon leur nature exacte dépendamment si l'enzyme est une ligase ou une GTase.



**Figure 14.** Les différents motifs conservés de la famille des nucléotidyltransférases. Le positionnement des différents motifs conservés sur une GTase, à gauche de la figure, et une ligase à ADN, à droite de la figure.

### 1.8.2. Les ligases à ADN

Les ligases sont des protéines ubiquitaires dans une cellule puisqu'elles accomplissent plusieurs fonctions essentielles telles que la réplication, la réparation et la recombinaison de l'ADN (Doherty et Suh, 2000). Peu importe la fonction accomplie par les ligases, la catalyse se déroule toujours par la formation d'un pont phosphodiester sur une cassure simple brin entre les extrémités 3'hydroxyle et 5'phosphate (Pascal, 2008). Afin de réaliser la réaction enzymatique, ces enzymes dépendent de la présence d'un cofacteur, soit l'ATP ou le NAD. Le cofacteur ATP est préférentiellement utilisé par les différentes

ligases dans le règne du vivant alors que les ligases dépendantes au cofacteur NAD<sup>+</sup> sont retrouvées uniquement chez les eubactéries (Sriskanda et Shuman, 1998). Malgré les nuances au niveau du cofacteur, la synthèse du pont phosphodiester se produit toujours par trois réactions catalytiques distinctes. Initialement, l'enzyme forme un lien covalent avec une molécule d'AMP via l'interaction avec la chaîne latérale de la lysine catalytique. Le lien phosphoramidate est généré après l'attaque de la lysine sur l'ATP suivie du relâchement de pyrophosphate. Chez les ligases dépendantes au NAD<sup>+</sup>, ce lien est obtenu par le clivage du NADH suivi de la libération de nicotinamide mononucléotide (Shuman et Ru, 1995). Par la suite, l'AMP est transférée sur l'extrémité 5' phosphate du site de cassure monocaténaire pour former une terminaison diphosphorylée. Pour compléter la réaction de ligation, l'extrémité 3'hydroxyle attaque le phosphate bêta de l'intermédiaire ADN-adénylate formé lors de la deuxième réaction (Odell et al., 2000). L'attaque permet de joindre les deux polynucléotides tout en libérant l'AMP.

### **1.8.3. Les ligases à ARN**

Dans un système biologique, différents processus surviennent lorsque des molécules d'ARN sont altérées et doivent être réparées (Wood et al., 2004). Parmi ces mécanismes biologiques, les ligases sont au cœur même des réactions enzymatiques permettant de relier des molécules d'ARN l'une à l'autre. Les études menant à comprendre la suite catalytique des ligases à ARN ont d'abord été axées sur la ligase présente chez le bactériophage T4 (Wang et Shuman, 2005). Le focus des recherches a été premièrement porté sur la T4 ligase à ARN 1 (Rnl1), mais rapidement il a été montré que cette enzyme était une mauvaise candidate pour l'étude du mécanisme global des ligases. En effet, Rnl1 possède uniquement 3 des 6 motifs conservés de la famille des nucléotidyltransférases (Bonfond et al., 2005). Les recherches ont été orientées vers une autre ligase du bactériophage T4, soit la ligase à ARN 2 (Rnl2) qui possède les motifs typiques I, III, IIIa, IV et V de cette famille. La caractérisation de cette enzyme a permis de dresser un portrait global des ligases à ARN qui sont présentes dans les différents



règles phylogénétiques. Les études cristallographiques de Rnl2 ont permis d'obtenir les évidences du mécanisme catalytique de ces ligases (Swift et al., 2009). Initialement, une molécule d'ATP est attaquée par la lysine catalytique de telle manière à initier un lien covalent entre l'enzyme et une molécule d'AMP, tout en relâchant du pyrophosphate (Pascal, 2008). Cette molécule d'AMP est ensuite transférée sur un ARN accepteur portant une extrémité monophosphorylée par une liaison phosphoanhydre. La ligation est complétée lorsque l'extrémité 3'-OH attaque le lien phosphoanhydre, libérant une molécule d'AMP, tout en formant le lien phosphodiester au niveau de la cassure monocaténaire (El Omari et al., 2006).

#### **1.8.4. Les ARN guanylyltransférases**

La cascade enzymatique des GTases est initiée par la liaison du GTP et d'un ion divalent au site actif de l'enzyme (Fresco et Buratowski, 1994). La liaison des substrats est suivie d'un changement conformationnel de l'enzyme qui est décrit par le rapprochement des deux domaines de l'enzyme de manière à se resserrer autour du site actif (Bougie et al., 2004). Sous cette conformation fermée, les deux domaines interagissent avec le nucléotide pour favoriser son positionnement adéquat dans le site actif (Ho et al., 1998b). La lysine catalytique de l'enzyme produit une attaque nucléophile sur le nucléotide ce qui engendre une liaison covalente entre ces deux parties tout en relâchant une molécule de pyrophosphate (Souliere et al., 2008). Une fois l'hydrolyse du nucléotide, les interactions intermoléculaires entre le nucléotide et l'enzyme induisent un changement de conformation pour restituer la conformation ouverte (Cong et Shuman, 1995). Cette ouverture du site actif permet l'entrée de l'accepteur d'ARN à être coiffé. La réaction est complétée lorsqu'il y a un transfert du nucléotide monophosphate à l'extrémité 5' des ARNm (Souliere et al., 2008).

### **1.8.5. La structure conservée des nucléotidyltransférases**

Malgré les nuances au niveau de leur substrat, ces enzymes ont été regroupées en une famille unique compte tenu de leur similitude en ce qui a trait à leur mécanique enzymatique (Issur et al., 2011). Ces enzymes possèdent toutes un domaine catalytique à l'extrémité amino-terminale, le domaine nucléotidyltransférase (NT), et un domaine non catalytique à l'extrémité carboxyl-terminale, le domaine de liaison à l'oligonucléotide (OB) (Cong et Shuman, 1993; Shuman et al., 1994). Le domaine NT, qui coordonne le site d'appariement de liaison du nucléotide tout en possédant l'acide aminé responsable de la catalyse, est défini par la présence de quatre motifs conservés (I, III, IIIa et IV) toujours disposés dans cet ordre précis (Sriskanda et Shuman, 2002b). Des études de mutagenèse ont montré l'importance de plusieurs acides aminés à l'intérieur de ces motifs conservés. S'étendant à l'extérieur du feuillet bêta du motif conservé IIIa, la chaîne latérale d'une phénylalanine, conservée chez tous les membres de cette famille, permet une interaction d'empilement aromatique avec le cycle purine du nucléotide (Sriskanda et Shuman, 1998). À l'opposé de cette phénylalanine, se retrouve le motif IV qui est composé majoritairement de résidus hydrophobes (Sriskanda et Shuman, 2002a). Un de ces résidus, variable selon les divers membres de la famille des nucléotidyltransférases, est responsable d'interagir selon des forces hydrophobes avec le groupement purine du nucléotide. La force intermoléculaire générée par ces deux résidus hydrophobes, la phénylalanine du motif IIIa et le résidu du motif IV, permet de coincer l'hétérocycle du nucléotide selon un empilement en sandwich (Shuman et Ru, 1995). Les études cristallographiques ont démontré l'importance de ces interactions qui permettent de diminuer les fluctuations latérales du nucléotide et ainsi favoriser un positionnement proche des groupements réactifs nucléophiles et électrophiles. En stabilisant le nucléotide, le temps de réaction est nettement amélioré (Cong et Shuman, 1995). À l'extrémité carboxyl-terminale de ces enzymes, le domaine OB est structuré sous forme de tunnel composé de cinq feuillets bêta antiparallèles et contient le sixième motif conservé (VI) de cette famille d'enzymes (Shuman et Lima, 2004). Le motif contient la séquence interne KxDK dont le rôle fonctionnel diffère selon qu'il est situé sur une ligase ou une GTase. Cette séquence permet de lier le substrat d'ADN et de favoriser le transfert

du nucléotide sur ce dernier, alors que des évidences suggèrent un rôle de coordination du nucléotide.

Malgré la grande similarité de ces enzymes, la spécificité des GTases et des ligases pour leur nucléotide respectif, soit la guanosine et l'adénosine, relève de la nature d'un nombre restreint d'acides aminés (Sriskanda et Shuman, 1998). Les déterminants moléculaires expliquant la spécificité des GTases pour la guanosine résident dans l'interaction de la lysine du motif IV, hautement conservée chez les GTases et absente chez les ligases à ADN, avec le groupement carboxyle du carbone 6 sur le cycle purine (Hakansson et al., 1997). Contrairement aux GTases où l'interaction entre le nucléotide et l'enzyme est médiée par un acide aminé unique, la spécificité des ligases pour l'adénosine découle d'un acide aminé variable chez les ligases (Shuman et Lima, 2004). Par exemple, la ligase à ADN ATP-dépendante interagit avec le groupement amine du carbone 6 du groupement aromatique de l'adénosine via la chaîne latérale d'un glutamate. Toutefois, chez le virus eucaryote PBCV-1 cette interaction, responsable de la spécificité du substrat, est obtenue par l'entremise de la chaîne latérale d'une thréonine.

#### **1.8.6. Le réarrangement conformationnel lors de la catalyse**

L'analyse des résultats cristallographiques suggère un mécanisme de changement conformationnel majeur conservé autant chez les GTases que chez les ligases (Shuman et Lima, 2004). La GTase du virus PBCV-1 a été cristallisée sous deux formes distinctes, soit ouverte et fermée (Hakansson et Wigley, 1998). Le cristal de la forme ouverte montre que le domaine OB est éloigné du domaine NT, et ce, avec une distance de 30,3 Å entre les deux centres de masse de ces domaines (Swift et McCammon, 2008). Sous cette forme, le nucléotide est lié à l'enzyme sous une forme non réactive, disposé perpendiculairement par rapport à la lysine nucléophile (Fabrega et al., 2003). Par contre, la forme fermée de l'enzyme témoigne d'un rapprochement du domaine OB

autour du site actif et en se rapprochant du domaine NT de 4,8 Å (Swift et McCammon, 2008). Ce changement de conformation induit un repliement de la chaîne triphosphate du GTP, le positionnant favorablement pour l'attaque catalytique du nucléophile.

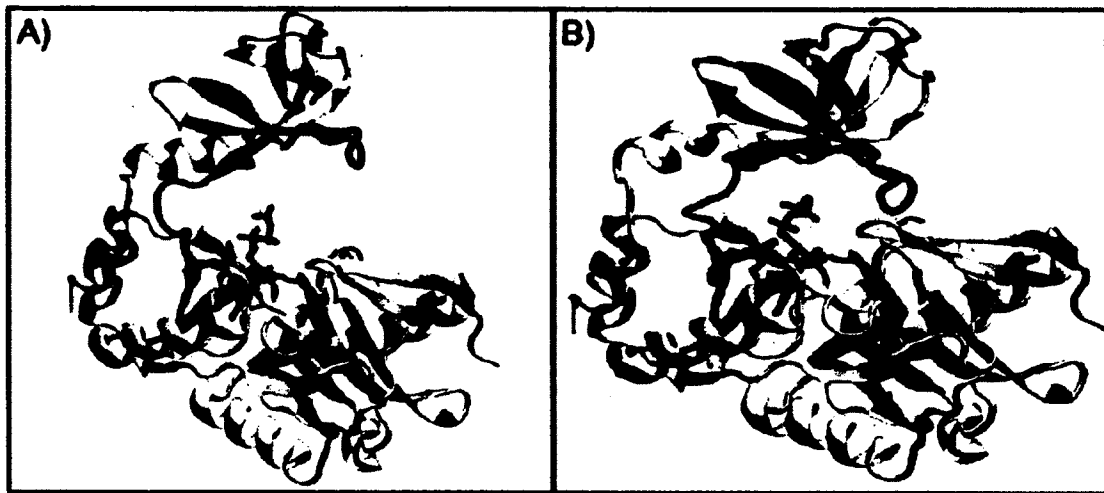


Figure 15. Le changement de conformation de l'ARN guanylyltransférase lors de la catalyse (Swift et McCammon, 2008). La représentation sous forme de ruban montrant les deux conformations : A) ouverte ainsi que B) la forme fermée.

### ***1.9. La caractérisation des ARN guanylyltransférases***

La structure coiffe étant présente uniquement dans les cellules eucaryotes, les GTases sont absentes des protéomes bactériens et archaebactériens. Par contre, plusieurs virus infectant des cellules eucaryotes encodent pour leurs propres enzymes de synthèse de la coiffe. Peu importe leur provenance, les GTases possèdent un mécanisme catalytique et une liaison au substrat similaire alors que leur structure diffère. Conséquemment, la GTase présente chez *PBCV-1* (344 acides aminés) représente la plus petite GTase répertoriée, alors que les GTases cellulaires sont nettement plus structurées et complexes. Par exemple, la GTase de *S. cerevisiae* (459 acides aminés) est une enzyme monofonctionnelle qui possède plusieurs éléments additionnels sur le domaine OB

permettant des interactions protéiques (Sawaya et Shuman, 2003). Chez les eucaryotes supérieurs, la structure des GTases s'est complexifiée davantage. Par exemple, la GTase de *Mus musculus* (597 acides aminés) se situe sur un polypeptide bifonctionnelle en fusion avec la RTase (Shuman, 1995). Ce chapitre permettra donc de faire un survol des différentes enzymes GTases présentes chez une variété d'organismes.

### **1.9.1. Les ARN guanylyltransférases virales**

Les enzymes virales ont joué un rôle crucial dans l'élucidation du mécanisme de la formation de la structure coiffe. Cette structure a été premièrement identifiée chez les réovirus et le virus de la vaccine. Chez le virus de la vaccine, les trois activités enzymatiques se retrouvent sur un hétérodimère contenant le polypeptide D1 (95 kDa) et D12 (33 kDa) (Shuman, 1995). L'activité enzymatique de la RTase et de la GTase est située conjointement sur le peptide D1 en portion amino-terminale. Ces deux activités enzymatiques ne possèdent pas de domaines distincts en ce sens que leurs motifs régissant le site actif sont dispersés au sein du peptide. La N7-MTase se retrouve fragmentée sur deux peptides dont la première portion se retrouve à l'extrémité carboxyl-terminale de D1 et la deuxième moitié sur la protéine D12 (Gu et Lima, 2005). De plus, ce complexe sert également de facteur d'initiation lors de la transcription des gènes intermédiaires du virus. La GTase de PBCV-1 a également permis d'étendre nos connaissances sur le mécanisme catalytique des GTases. En effet, les nombreuses études cristallographiques en présence de plusieurs ligands et sous différentes conformations ont permis d'identifier le réseau d'interaction entre les acides aminés et le substrat (Hakansson et al., 1997). Ces études ont permis de visualiser le changement de conformation qui se produit lors de la catalyse et d'identifier l'amplitude de ce changement. Malgré le mécanisme hautement conservé chez les GTases, il a été possible d'identifier des nuances chez les protéines de certains organismes. Parmi les GTases au mécanisme catalytique alternatif se retrouve la protéine p90 des *birnavirus* (Spies et Muller, 1990). En effet, le lien phosphoamine de la protéine déroge de l'acide aminé catalytique habituel pour relier le groupement GMP à un résidu sérine ou thréonine

(Ogino et Banerjee, 2011). De plus, la GTase des *flavivirus* se retrouve sur la protéine d'envergure Ns5, qui remplit également les activités enzymatiques de l'ARN polymérase dépendante à l'ARN et de la N7-MTase (Issur et al., 2009b). Malgré les évidences biochimiques démontrant que ce peptide peut accomplir l'activité GTase, l'absence du motif catalytique consensus laisse présager que cette protéine réalise l'activité enzymatique via un acide aminé catalytique différent.

### 1.9.2. Les ARN guanylyltransférases fongiques

Les GTases de *S. cerevisiae*, de *C. albicans* et de *S. pombe* ont toutes été identifiées sur un peptide monofonctionnel appelé respectivement, Ceg1, Cgt1 et Pce1. Ces GTases fongiques possèdent 38 % de similarité de leur séquence d'acides aminés (Ho et al., 1998b). De plus, il a été montré que Cgt1 et Pce1 peuvent compléter une souche de *S. cerevisiae* lorsque Ceg1 possède des mutations létales. Les diverses interactions protéiques des GTases ont été démontrées autant *in vivo* qu'*in vitro*. En effet, cette protéine interagit avec le domaine de l'extrémité carboxyl-terminale (CTD) de l'ARN polymérase II au niveau de l'heptapeptide YSPTSPS (Tyr-Ser-Pro-Thr-Ser-Pro-Ser) répété 26 fois chez *S. cerevisiae*. L'heptapeptide possède un patron de phosphorylation sur les résidus sérine dépendamment des étapes de la transcription. En effet, avant la phase d'élongation du complexe de transcription, l'heptapeptide possède des phosphorylations sur la serine à la position 5 de l'heptamère et permet de recruter Ceg1 (Gu et Lima, 2005). Il est maintenant clairement établi que Ceg1 interagit avec la RTase pour permettre son recrutement près de l'ARNm nouvellement transcrit.

La GTase de *S. cerevisiae* a servi de modèle afin d'étudier exhaustivement le mécanisme d'action de cette enzyme. Des études ont combiné l'impact des substitutions d'acides aminés pour une alanine, des substitutions conservatives et des données obtenues par cristallographie pour dresser un dessin rationnel des divers rôles des acides aminés autour du site actif. Ainsi, ces études ont permis de déterminer l'importance de certaines chaînes latérales délimitant le site de reconnaissance du substrat. Initialement, la Lys70, plus

précisément le groupement  $\epsilon$ -amino, est à proximité du phosphate bêta du GTP (Wang et al., 1997). Les essais de substitution de la lysine pour une arginine ou une histidine n'ont pas permis de contrer le phénotype de létalité, qui s'explique par une pression de sélection très forte pour l'acide aminé catalytique. Le site catalytique est complètement formé lorsque le changement de conformation de la forme ouverte à fermée se produit ce qui ressert le motif IV autour du site actif permettant certains résidus d'induire plusieurs interactions clés avec le substrat. En effet, l'Arg369 du motif IV permet, d'une part, d'interagir avec le phosphate  $\beta$  du GTP et, d'autre part, avec l'Asp372. La formation de deux ponts hydrogènes de l'Arg369, qui possède une chaîne latérale bidentée, permet d'expliquer pourquoi la substitution par une lysine devient létale (Wang et al., 1997). De plus, deux lysines présentes sur le motif V permettent d'interagir avec le phosphate  $\alpha$  du GTP. Sous la forme ouverte, la Lys247 régit ces liaisons, alors que sous la forme fermée il s'agit de la Lys249. Ces résidus basiques permettront d'accélérer la catalyse en stabilisant les états de transition. Finalement, l'Asp225 du motif IV permet de favoriser deux réactions chimiques (Fabrega et al., 2003). Tout d'abord, lorsque l'enzyme est sous la forme fermée, le caractère acide de ce résidu permet d'interagir avec un hydrogène du groupement amine de la lysine catalytique ce qui permet de favoriser la formation du lien phosphoamide entre la lysine et le GTP. Par la suite, lorsque l'enzyme réadopte un changement de conformation vers la forme ouverte, l'aspartate libère le proton du groupement amide, déstabilisant le lien phosphoamide, et facilitant l'attaque de l'extrémité 5' de l'ARN sur la molécule de GMP (Swift et McCammon, 2008).

### **1.9.3. Les ARN guanylyltransférases des métazoaires**

Les GTases des eucaryotes supérieurs ont été identifiées et caractérisées chez plusieurs organismes, tels que chez *C. elegans* et *M. musculus*. Ces GTases ont été identifiées sur un polypeptide bi-fonctionnelle qui possède un domaine amino-terminale avec l'activité RTase et un domaine carboxyl-terminale avec l'activité GTase (Sawaya et Shuman, 2003). Malgré que les deux activités soient situées sur le même peptide, les deux enzymes sont fonctionnelles en l'absence de l'autre. Ainsi, pour Mce1, l'enzyme

bifonctionnelle chez la souris, il est possible de séparer la protéine de 597 acides aminés en une entité RTase (1-210) et un domaine GTase (210-597) (Ho et al., 1998b). Il est possible de faire des recoupements entre les enzymes de métazoaires et de levures. De façon analogue aux enzymes de levures, uniquement le domaine de l'activité GTase permet d'interagir avec le domaine CTD de l'ARN polymérase II. De plus, Mce1 peut compléter une souche de *S. cerevisiae* lorsque la GTase endogène possède des mutations létales. En utilisant ce système *in vivo*, ils ont pu déterminer 16 acides aminés essentiels, dont la lysine catalytique (Lys294), pour la survie cellulaire. Un très fort niveau de concordance s'observe entre les acides aminés de Mce1 et ceux de Ceg1 (Ho et al., 1998b). En effet, des 16 mutants létaux identifiés, 11 impliquent le même phénotype chez la levure. Des études exhaustives sur les mutations compensatoires ont permis de mettre à l'avant-plan plusieurs nuances au niveau de certaines interactions clés pour la coordination du substrat à l'enzyme

#### **1.9.4. Les inhibiteurs des ARN guanylyltransférases**

Jusqu'à présent, les inhibiteurs des GTases ont été identifiés de par leur homologie de structure avec les substrats et produits naturels de l'enzyme. En effet, des études ont analysé l'impact du composé antiviral foscarnet, qui cible l'ADN polymérase du cytomegalovirus, sur la GTase du virus PBCV-1. Il a été montré que le foscarnet est un inhibiteur puissant de l'activité guanylyltransférase. Des évidences suggèrent que ce composé interfère avec la liaison du substrat au site actif. En effet, le foscarnet est un analogue de pyrophosphate (PP<sub>i</sub>) lequel est autant le produit de l'activité catalytique des GTases que des ADN polymérases. Le foscarnet est recruté au site actif des GTases et cause, par encombrement stérique, une inhibition de la liaison du substrat. Toutefois, ce composé est peu sélectif pour les enzymes pathogènes et inhibe ces enzymes avec un potentiel inhibiteur de 2 mM.

L'analogue de nucléotide la ribavirine est un antiviral qui limite la propagation de plusieurs virus, tel l'hépatite C, en bloquant l'ARN polymérase de ces virus. La ribavirine



se retrouve sous une forme triphosphorylée en cellule ce qui permet son intégration dans l'ARN lors de la transcription et induit la terminaison de la transcription. Toutefois, ce composé possède un mode d'action pléiotropique et des études ont montré des évidences biochimiques que ce composé peut interagir avec l'enzyme GTase du virus de la vaccine. En approfondissant cette inhibition, il a été montré que l'enzyme peut lier la ribavirine et former un complexe covalent tout à fait similaire avec le substrat naturel. De plus, la ribavirine peut également être transférée sur un ARN et donc générer un ARNm coiffé alternativement. Cette coiffe ne peut être méthylée par les N7-MTase ce qui se traduit par un ARNm stable, mais non efficacement traduit. Le cas de la ribavirine démontre bien l'étendue de la possibilité d'utiliser la GTase en tant que cible potentielle. Par contre, plusieurs études subséquentes devront s'attarder à augmenter la spécificité des composés.

Compte tenu de la grande conservation des GTases au niveau du site actif, il est difficile d'arriver à obtenir une inhibition spécifique en utilisant des composés similaires au substrat et au produit. C'est pourquoi, en utilisant une inhibition allostérique, distante du site actif où la pression de sélection favorise la conservation au travers de l'évolution, il est possible d'augmenter la spécificité pour une GTase spécifique. L'objectif de la présente étude est de parvenir à identifier des composés potentiels par la méthode du criblage virtuel et de confirmer cette interaction autant *in vitro* qu'avec des essais *in vivo*. Dans l'optique de pouvoir rapidement transférer les essais *in vitro* aux essais cellulaires, l'enzyme GTase de *S. cerevisiae* est le modèle de cette étude. Le génome haploïde de cet organisme est l'outil idéal pour des études génétiques. En effet, par génie génétique, il est possible de retirer un gène d'intérêt, d'exprimer une protéine exogène et d'identifier rapidement des mutations compensatoires. Ceg1 est très bien caractérisée autant au niveau des acides aminés essentiels que ceux provoquant un phénotype de thermosensibilité.

### ***1.10. Préambule à l'article***

La structure coiffe est d'une importance primordiale au métabolisme des ARNm. Sa synthèse est issue de réactions enzymatiques exécutées par des protéines présentes autant dans le règne eucaryote que chez certains de ces pathogènes. De plus, ces enzymes sont essentielles pour la survie cellulaire comme il est démontré par la létalité des souches de *S. cerevisiae* délétées pour l'un de ces trois gènes. Pour ces raisons, les enzymes responsables de la synthèse de la structure coiffe sont de bonnes cibles potentielles pour le développement d'antiviraux et d'antifongiques. Parmi ces trois enzymes, la GTase est une cible de choix pour inhiber la structure coiffe puisqu'elle effectue l'étape limitante qui conditionne la vitesse de réaction de tout le processus d'ajout de la coiffe. Malgré qu'elle représente une cible robuste, un obstacle s'oppose à l'application d'une telle thérapie. Le mécanisme catalytique de cette enzyme est hautement conservé de même que les séquences d'acides aminés coordonnant la liaison au substrat. Toutefois, à l'extérieur des motifs conservés, les différentes protéines ne possèdent pas un niveau élevé d'homologie. Ceci permet d'utiliser cette divergence à profit pour le développement d'inhibiteurs spécifiques.

La pierre angulaire de la stratégie thérapeutique envers plusieurs virus et levures pathogènes repose sur l'utilisation de nucléosides ou d'analogues de nucléotides. Toutefois, cette stratégie s'accompagne inmanquablement d'effets secondaires causés par une toxicité en affectant d'autres voies métaboliques. En effet, plusieurs processus biologiques impliquent l'interaction entre les nucléotides et une panoplie d'enzymes. En utilisant une banque de composés variés, il est possible d'identifier un composé beaucoup plus sélectif que des analogues de nucléotides.



## **2) Article**

**Le criblage virtuel à haut débit identifie l'acide mycophénolique comme un nouvel inhibiteur de la synthèse de la structure coiffe.**

**Titre : Virtual High-Throughput Screening Identifies Mycophenolic Acid as a Novel RNA Capping Inhibitor**

**Maude Tremblay-Létourneau, Simon Despins, Isabelle Bougie, Martin Bisailon\***  
RNA Group, Département de Biochimie, Faculté de Médecine et des Sciences de la Santé  
, Université de Sherbrooke, Sherbrooke, Québec, Canada

**Publié**

**Tremblay-Létourneau M, Despins S, Bougie I, Bisailon M, 2011 Virtual High-Throughput Screening Identifies Mycophenolic Acid as a Novel RNA Capping Inhibitor. PLoS ONE 6(9): e24806. doi:10.1371/journal.pone.0024806**

## **2.1. Sommaire**

L'ARN guanylyltransférase (GTases) est impliquée dans la synthèse de la structure coiffe (m<sup>7</sup>Gppp-ARN) présente à l'extrémité 5' des ARN messagers (ARNm) eucaryotes. Les GTases font partie de la superfamille des nucléotidyltransférases, au sein de laquelle on retrouve également les ligases d'ARN et d'ADN. Les GTases catalysent une réaction en deux étapes qui débute par la liaison au substrat GTP suivi d'une hydrolyse de ce dernier formant ainsi l'intermédiaire de réaction enzyme-GMP lié de façon covalente. La molécule de GMP est ensuite transférée, lors de la deuxième étape de la réaction, à l'extrémité diphosphorylée d'un ARNm nouvellement transcrit pour former la structure Gppp-ARN. Dans la présente étude, une combinaison de criblage virtuel de banque de composé, de modélisation moléculaire par homologie et plusieurs essais biochimiques a été utilisée afin d'identifier un nouvel inhibiteur de GTase. En utilisant cette approche, nous démontrons que l'acide mycophénolique (MPA) peut inhiber la réaction GTase en prévenant le transfert catalytique du GMP sur un accepteur d'ARN. En ce sens, MPA représente un nouveau type d'inhibiteur dirigé contre GTase qui inhibe la deuxième étape de la réaction catalytique. De plus, nous démontrons que l'addition de MPA dans une culture de *S. cerevisiae* induit une réduction des ARNm coiffés. Finalement, des essais biochimiques démontrent que MPA peut inhiber les ligases à ADN par l'inhibition de la deuxième étape de la réaction enzymatique. L'implication biologique de l'inhibition médiée par MPA sur les membres de la famille des nucléotidyltransférase est discutée.

## **2.2. Implication dans l'article**

Mon implication dans le manuscrit est présente autant au niveau de la réalisation des expériences, de l'interprétation des résultats et de l'écriture de l'article. J'ai réalisé presque la totalité des figures du présent article à l'exception de l'essai de liaison de l'ARN à la protéine en présence de MPA. J'ai participé activement à l'interprétation des résultats afin de mettre au point les essais pour valider les hypothèses. De plus, j'ai

participé à l'écriture de l'article de façon conjointe avec mon directeur de recherche ainsi qu'à la révision du manuscrit en entier.

Pour ce qui est de la contribution des autres auteurs présents sur cet article, Simon Despins a réalisé une figure dans ce manuscrit. Pour sa part, Isabelle Bougie a réalisé les essais préliminaires du projet et les techniques de clonage afin de générer des bactéries qui permettent d'exprimer de grande quantité de GTase de *S. cerevisiae* recombinante.

# Virtual High-Throughput Screening Identifies Mycophenolic Acid as a Novel RNA Capping Inhibitor

**Maude Tremblay-Létourneau, Simon Despins, Isabelle Bougie, Martin Bisaillon\***  
RNA Group, Département de Biochimie, Faculté de Médecine et des Sciences de la Santé  
, Université de Sherbrooke, Sherbrooke, Québec, Canada

Address correspondence to :

Martin Bisaillon, 3001, 12 e avenue, Sherbrooke, Qc, Canada J1H 5N5

Fax: 819-564-5340

E-mail: [Martin.Bisaillon@usherbrooke.ca](mailto:Martin.Bisaillon@usherbrooke.ca)

# Abstract

The RNA guanylyltransferase (GTase) is involved in the synthesis of the m<sup>7</sup>Gppp-RNA cap structure found at the 5' end of eukaryotic mRNAs. GTases are members of the covalent nucleotidyl transferase superfamily, which also includes DNA and RNA ligases. GTases catalyze a two-step reaction in which they initially utilize GTP as a substrate to form a covalent enzyme-GMP intermediate. The GMP moiety is then transferred to the diphosphate end of the RNA transcript in the second step of the reaction to form the Gppp-RNA structure. In the current study, we used a combination of virtual database screening, homology modeling, and biochemical assays to search for novel GTase inhibitors. Using this approach, we demonstrate that mycophenolic acid (MPA) can inhibit the GTase reaction by preventing the catalytic transfer of the GMP moiety onto an acceptor RNA. As such, MPA represents a novel type of inhibitor against RNA guanylyltransferases that inhibits the second step of the catalytic reaction. Moreover, we show that the addition of MPA to *S. cerevisiae* cells leads to a reduction of capped mRNAs. Finally, biochemical assays also demonstrate that MPA can inhibit DNA ligases through inhibition of the second step of the reaction. The biological implications of these findings for the MPA-mediated inhibition of members of the covalent nucleotidyl superfamily are discussed.



# Introduction

The synthesis and maturation of eukaryotic mRNAs are crucial events for gene expression. During mRNA synthesis, eukaryotic mRNAs undergo a series of essential modifications before being exported to the cytoplasm where they are translated into proteins [1]. These processing events include the addition of a cap structure at the 5' terminus, the splicing out of introns, the editing of specific nucleotides, and the acquisition of a poly(A) tail at the 3' terminus. The RNA cap structure found at the 5' end of mRNAs is critical for the splicing of the cap-proximal intron, the transport of mRNAs from the nucleus to the cytoplasm, and for both the stability and translation of mRNAs [2,3]. The cap is synthesized by a series of three enzymatic reactions [4]. The first step involves the hydrolysis of the RNA 5'-triphosphate end of the nascent RNA by an RNA triphosphatase to form a diphosphate end. An RNA guanylyltransferase then catalyzes a two-step reaction in which it initially utilizes GTP as a substrate to form a covalent enzyme-GMP intermediate. The GMP moiety is then transferred to the diphosphate end of the RNA transcript in the second step of the reaction to form the GpppN structure. The guanosine residue is finally methylated by an RNA (guanine-N7)-methyltransferase to form the typical m<sup>7</sup>GpppN cap structure.

A number of different microbial pathogens code for their own enzymes involved in the synthesis of a cap structure [5,6,7,8,9,10]. Although the RNA cap structures originating from human and microbial enzymes are often identical, the physical organization of the genes, subunit composition, structure and catalytic mechanisms of the microbial-encoded enzymes involved in the synthesis of the RNA cap structure are often significantly different from those of host cells [2]. As a consequence these pathogenic capforming enzymes are potential targets for anti-microbial drugs.

During the past few years, both the RNA triphosphatase and the RNA (guanine-N7) methyltransferase (N7-MTase) components of the RNA capping machinery have been major targets for the development of drugs directed against RNA cap synthesis [11,12,13,14,15,16,17,18,19,20]. Of all the enzymes involved in RNA capping, the RNA guanylyltransferase (GTase) has traditionally been considered a poor candidate as an anti-microbial target because of the high mechanistic and structural conservation of this enzyme across species [21]. Based on various crystal structures of GTases, a general mechanism for phosphoryltransfer has previously been elucidated which involves conformational changes between an open and closed form of the enzyme [22,23]. In the first step of the reaction, GTP binds to the open form of the enzyme which promotes closure of the Nterminal nucleotidyl transferase (NT) domain and the C-terminal oligomer-binding (OB) fold domain. This closure is stabilized by interactions between the residues of the NT domain, the bound nucleotide, and residues on the OB fold domain. Domain closure is then followed by hydrolysis of the GTP substrate to produce the enzyme-GMP covalent intermediate. Hydrolysis of GTP disrupts the interactions between the bound guanylate and the C-terminal OB fold domain, thus destabilizing the closed form of the enzyme, which opens up with the concomitant release of pyrophosphate. This exposes the RNA-binding site of the enzyme, thereby allowing the subsequent transfer of the GMP moiety onto the acceptor RNA. Figure 1 summarizes the mechanistic and structural pathway used by GTases.

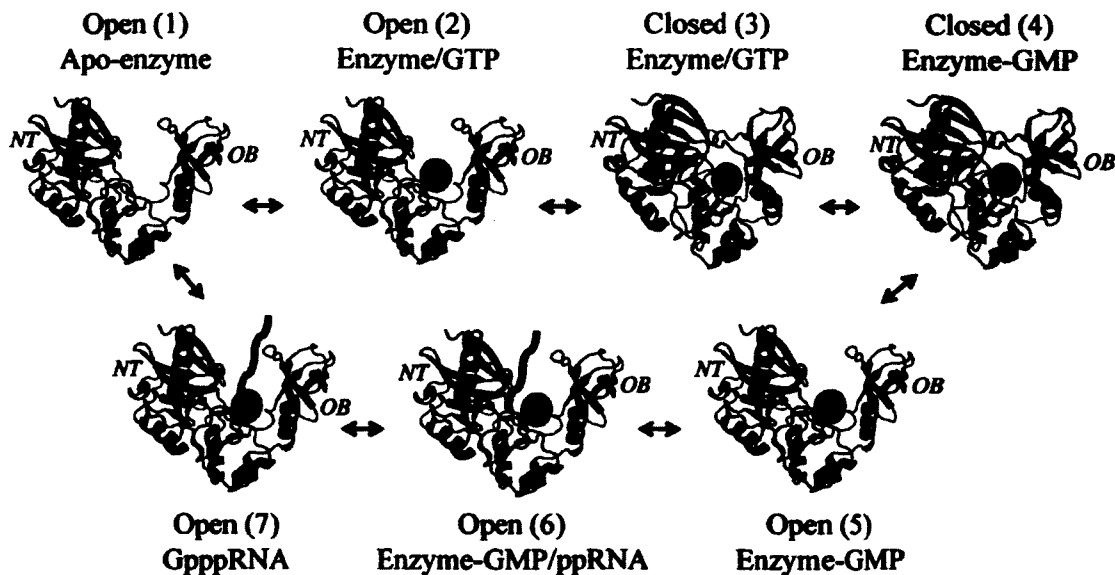


Figure 1. Structural and mechanistic pathway used by RNA guanylyltransferases. The mechanism for phosphoryltransfer involves conformational changes between an open and closed form of the enzyme [22], [23]. GTP (grey sphere) initially binds to the apo-enzyme (structure 1) which promotes closure of the N-terminal nucleotidyl transferase (NT) domain and the C-terminal oligomer-binding (OB) fold domain (structure 3). This is followed by hydrolysis of the GTP substrate to produce the enzyme-GMP covalent intermediate (structure 4). Hydrolysis of GTP disrupts the interactions between the bound guanylate and the C-terminal OB fold domain, thus destabilizing the closed form of the enzyme, which opens up with the concomitant release of pyrophosphate (structure 5). This exposes the RNA-binding site of the enzyme (exact location unknown), thereby allowing the subsequent transfer of the GMP moiety onto the acceptor RNA (structure 7). The capped RNA is then released and the enzyme can reinitiate the pathway.

Recent *in vitro* studies have shown that foscarnet is a potent inhibitor of the GTase reaction [24]. Its mechanism of action is purported to occur through substrate binding inhibition on account of its analogous nature to pyrophosphate (PPi), a product of the GTase reaction. Ribavirin, a broad-spectrum nucleoside analogue approved for the treatment of various viral infections, is another inhibitor of the GTase activity [25]. Biochemical studies have shown that ribavirin triphosphate can actually be used as a substrate by the vaccinia virus GTase to form a covalent enzyme-ribavirin monophosphate intermediate reminiscent of the covalent enzyme-GMP intermediate [25]. Furthermore, ribavirin monophosphate can be transferred to the diphosphate end of an RNA transcript to form the unusual RpppN structure [25]. However, RNA transcripts

blocked with ribavirin are not recognized efficiently by the cap-binding protein eIF4E, and are not translated into proteins [26]. The use of guanine-N7 methylation-inert cap donor molecules could potentially prove to be an interesting line of research for the development of antimicrobial drugs. However, on account of the possibility of off-targets, the risk of major side effects upon treatment with GTase substrate/product analogs remains. Several issues related to the specificity problem faced with these inhibitors can likely be partially resolved by the development of non-nucleoside inhibitors.

In the current study, we used a combination of virtual database screening, homology modeling, and biochemical assays to search for novel GTase inhibitors. We demonstrate that mycophenolic acid, a compound which is currently used both in cancer and immunosuppressive chemotherapy, is an inhibitor of the GTase reaction. The biological implications of these findings for the MPA-mediated inhibition of RNA capping are discussed.

## **Experimental procedures**

Structure-based virtual screening to identify potential candidate compounds that can bind to GTases, the crystal structures of various GTases were retrieved from the Protein Data Bank (*Chlorella* virus: 1CKN and 1CKO; *Candida albicans*: 1P16; *S. cerevisiae*: 3KYH). The chemical databases used in our virtual screening included the Sigma-Aldrich, Ambinter, ASINEX, IS Chemical technology, MolPort, and Vitas M Laboratory catalogs. Collectively, these 6 databases offered a collection of 13,800,000 small-molecule compounds. The molecular docking program DOCK (Version 4.0) was used to perform the virtual screening [27]. Compounds displaying at least 80% of structural similarity to GTP or ATP (25,000 compounds) were screened for the binding to the four GTases. The levels of similarities were measured using the Tanimoto equation [28] and the PubChem dictionary-based binary fingerprint (<http://pubchem.ncbi.nlm.nih.gov/>). Based on the

binding models of these compounds predicted by DOCK, the X-SCORE program (Version 1.1) was applied to obtain an estimate of the binding affinities of these compounds [29]. The compounds were then ranked according to their binding affinities as estimated by X-SCORE.

### *Expression and purification of proteins*

The RNA guanylyltransferases from *S. cerevisiae* (Ceg1), vaccinia virus (D1R), Chlorella virus (A103R), and human (HCE) were expressed and purified as described before [24,25,30,31]. The ATP-dependent ligase DNA ligase from Chlorella virus (ChVLig) was also expressed and purified as described previously [32].

### *Assay for enzyme-GMP complex formation*

The assay was performed by incubating the enzyme (0.1 mM) with 10 mM [ $\alpha$ - $^{32}$ P]GTP in a buffer containing 50 mM Tris-HCl, pH 8, 5 mM DTT, and 5 mM MgCl<sub>2</sub> for 5 min at 30°C. The reactions were stopped by the addition of EDTA to 10 mM and SDS to 1%. The reactions were analyzed by electrophoresis through a 12.5% polyacrylamide gel containing 0.1% SDS. The radiolabeled proteins were visualized by autoradiography of the gel. The extent of covalent complex formation was quantitated by scanning the gel with a PhosphorImager (Amersham Biosciences).

### *Preparation of RNA substrates*

An RNA substrate of 81 nucleotides was synthesized with the MAXIscript kit (Ambion) using T7 RNA polymerase. The RNA transcript was synthesized from the pBS-KSII+ plasmid (Stratagene) that had been linearized with HindIII. The RNA substrate was purified on a denaturing 20% polyacrylamide gel and visualized by ultraviolet shadowing. The corresponding band was excised and then eluted from the gel by an overnight incubation in 0.1% SDS and 0.5 M ammonium acetate. The RNA was then

precipitated with ethanol and quantitated by spectrophotometry. Alternatively, radiolabeled RNA substrates were also synthesized by adding [ $\alpha$ - $^{32}\text{P}$ ]ATP or [ $\alpha$ - $^{32}\text{P}$ ]GTP to the transcription reaction. The purified 5'-triphosphorylated RNA was further processed to obtain a diphosphorylated 5' end using the *S. cerevisiae* RNA 5'-triphosphatase (Cet1) which was expressed and purified as described before [33]. The diphosphorylated RNA (ppRNA) was precipitated with ethanol, resuspended, quantitated by spectrophotometry, and stored at  $-20^{\circ}\text{C}$ .

### *Molecular docking*

Docking calculations were carried out using the Docking Serversoftware and the Dreiding force field was used for energy minimization of MPA using built-in Chemaxon tools in Docking Server [34]. PM6 semi-empirical charges calculated by MOPAC2007 were added to the ligand atoms. Non-polar hydrogen atoms were merged and rotatable bonds were defined [35]. Docking calculations were carried out using the Chlorella virus and *Candida albicans* RNA guanylyltransferase crystal structures (Protein Data Bank 1CKN and 1P16). Essential hydrogen atoms, Kollman united atom type charges and solvation parameters were added with the aid of AutoDock tools [36]. Affinity (grid) maps of  $20620620 \text{ \AA}^{\circ}$  grid points and  $0.375 \text{ \AA}^{\circ}$  spacing were generated using the Autogrid program [36]. AutoDock parameter set- and distance-dependent dielectric functions were used in the calculation of the van der Waals and the electrostatic terms, respectively. Docking simulations were performed using the Lamarckian genetic algorithm and the Solis and Wets local search method [37]. Initial position, orientation, and torsions of the ligand molecules were set randomly. Each docking experiment was derived from two different runs that were set to terminate after a maximum of 2,500,000 energy evaluations. The population size was set to 150. During the search, a translational step of  $0.2 \text{ \AA}^{\circ}$ , and quaternion and torsion steps of  $5 \text{ \AA}^{\circ}$  were applied.

### *UV-crosslinking experiments*

UV-crosslinking between the internally <sup>32</sup>P-labeled RNA of 81 nt and the yeast GTase was performed in a crosslink buffer (50 mM Tris pH 7.5, 5 mM DTT, 5 mM MgCl<sub>2</sub>). The protein (12 mM) was incubated with GTP (1 mM) and different concentrations of MPA for 10 minutes at 30°C. The radiolabeled RNA (3 mM) was added to the reaction mixture and incubated for 5 minutes at 30°C. The reaction mixture was exposed to UV light (254 nm, 20 Joules/cm<sup>2</sup>) for 5 min at 30°C using a Stratalinker 2400 UV Crosslinker (Stratagene). The crosslink mixture was denatured (50 mM Tris-HCl, pH 7.0, 5% sucrose, 5% bmercaptoethanol, 2% SDS) and separated by electrophoresis on a 12% SDS-PAGE. The gel was analyzed by phosphorimaging.

#### *Ligation assay*

The ligation reaction was performed as described previously [32]. Briefly, a 36 bp DNA duplex harboring a centrally placed nick was used as a substrate. The 18-mer constituting the 5'- phosphate-terminated strand 5'-d(ATTCCGATAGTGACTACA)-3' was 5'-radiolabeled and gel purified as described before. This labeled 18-mer was then annealed to a complementary 36-mer in the presence of a 3'-OH 18-mer strand 5'-d(CATATCCGTGTCGCCCTT)-3' [38,39]. Ligation reaction mixtures containing 50 mM Tris-HCl (pH 7.5), 5 mM DTT, 10 mM MgCl<sub>2</sub>, 1 mM ATP, nicked duplex substrate and the *Chlorella* virus DNA ligase were incubated at 22°C for 15 min. The reactions were stopped by the addition of 1 ml 0.5 M EDTA and 5 ml formamide. The samples were heated at 95°C for 5 min and then analyzed by electrophoresis through a 17% polyacrylamide gel containing 7 M urea. The extent of ligation was determined by scanning the gel with a PhosphorImager (Amersham Biosciences).

#### *Primer extension analysis of 5' ends*

Primer extension reactions were performed as described previously [40] using a 5' <sup>32</sup>P-labeled 18-mer DNA oligonucleotide complementary to the 5' region of the SSA1 mRNA (positions +1 to +19). Total RNA was extracted from *S. cerevisiae* cells that were grown in the presence or absence of 500 mg/ml MPA for 3 h at 30°C. The primer

extension reactions were analyzed by electrophoresis through a 8% polyacrylamide gel containing 7 M urea in TBE and visualized by autoradiography.

## Results and discussions

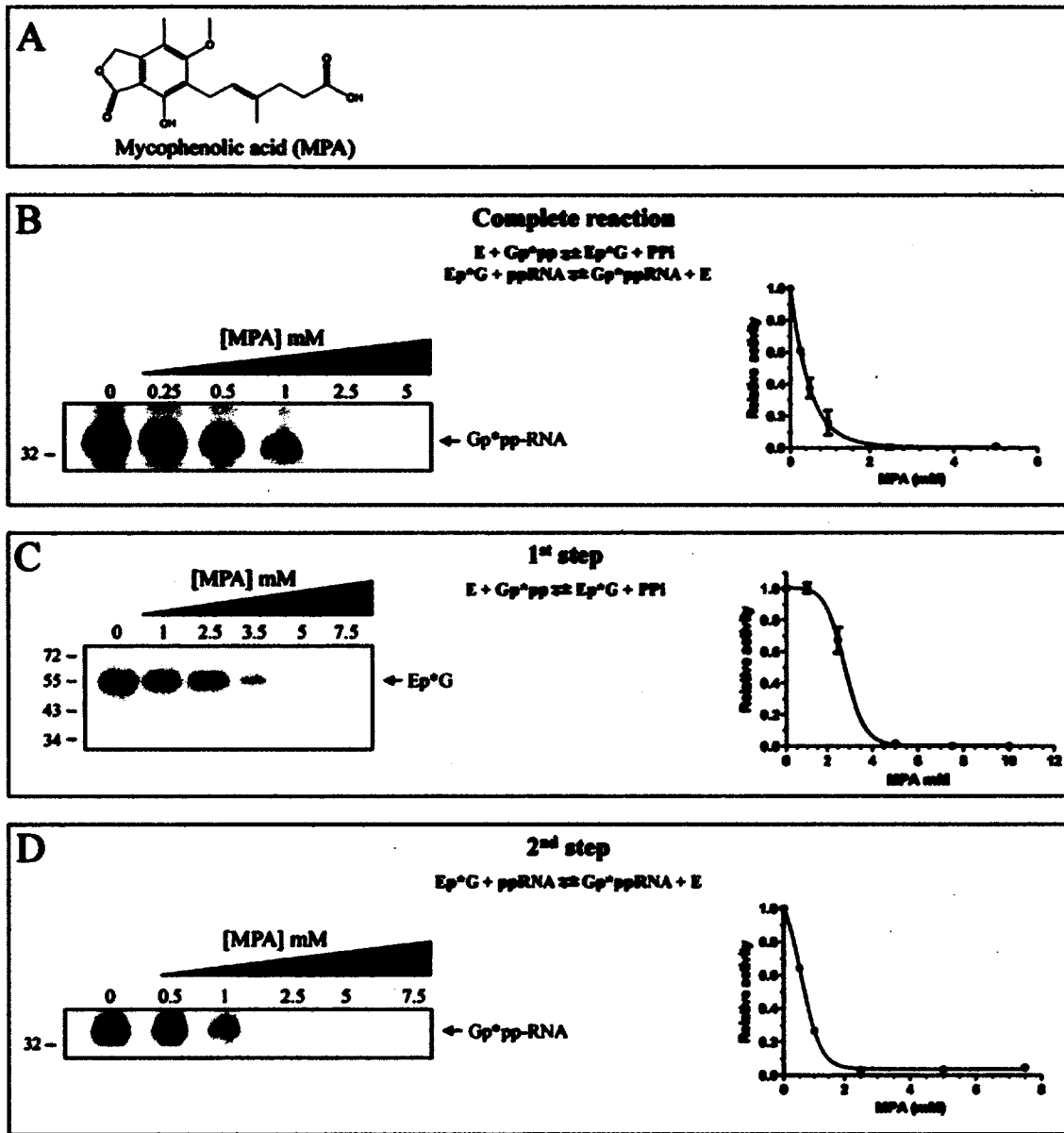
### *Identification of mycophenolic acid through virtual database screening*

We initially performed a virtual screen of more than 25,000 purine-related compounds for their ability to bind to the GTases of Chlorella virus (open and closed forms), *S. cerevisiae* (open form), and *C. albicans* (open form). The ligands were docked and ranked according to their respective docking scores. Our initial screen indicated that very few compounds bound to the open forms of these three different GTases. In all cases, the highest scoring compounds were not predicted to bind with very strong affinities to the enzymes (predicted  $K_d$  2 mM). However, one compound (mycophenolic acid) docked efficiently in the cavity of the closed form of the Chlorella virus GTase (predicted  $K_d$  of 280 nM). Mycophenolic acid (MPA, Fig. 2A) is a well-known inhibitor of inosine monophosphate dehydrogenase (IMPDH), a key cellular enzyme required for biosynthesis of guanine nucleotides [41]. Because T- and B-lymphocytes are critically dependent for their proliferation on de novo synthesis of purines, whereas other cell types can utilize salvage pathways, IMPDH inhibitors have potent cytostatic effects on lymphocytes [42,43]. Accordingly, MPA has been used both in cancer and immunosuppressive chemotherapy as well as in antiviral and antifungal therapy [44,45,46,47,48,49,50,51,52]. Because GTP is required for the transcription and replication of cellular and microbial genomes, it has traditionally been assumed that the decrease in the cytosolic concentration of GTP could affect both cell growth and the multiplication of fungal and viral pathogens.



### *MPA inhibits the RNA guanylyltransferase activity*

The ability of MPA to inhibit the GTase of the model organism *S. cerevisiae* (Ceg1 protein) was then investigated. The GTase activity is actually a two-step ping-pong reaction in which the enzyme first reacts with GTP to produce the enzyme-GMP (EpG) covalent intermediate with the concomitant release of pyrophosphate [2]. In the second step of the reaction, the GMP moiety is then transferred to a 5'-diphosphate RNA. The ability of MPA to inhibit the complete GTase reaction (both steps) was monitored using a standard GTase assay in which the purified enzyme was incubated with [ $\alpha$ -<sup>32</sup>P]GTP and a 5'-diphosphate acceptor RNA. The reaction products were then analyzed on a denaturing polyacrylamide gel. As can be seen in figure 2B, the presence of increasing concentrations of MPA significantly decreased the transfer of the radiolabeled GMP moiety onto the acceptor RNA. MPA was shown to inhibit the overall GTase activity by 50% at 360 mM, and by 80% at 900 mM. In order to characterize the inhibition of the GTase activity by MPA, we next set out to investigate which of the two catalytic steps of the reaction is inhibited by MPA. The first step of the reaction, i.e. the formation of the enzyme-GMP covalent intermediate, was monitored by incubating the purified enzyme in the presence of [ $\alpha$ -<sup>32</sup>P]GTP. The radiolabeled covalent enzyme-GMP complex was then visualized by autoradiography following electrophoresis on a denaturant polyacrylamide gel. Our results indicate that very high concentrations of MPA are required to inhibit the formation of the enzyme-GMP covalent complex (Fig. 2C). We determined that a concentration of 3 mM of MPA is required to inhibit the first step of the reaction by 50%, a concentration that corresponds to 8-times the amount required to inhibit the overall GTase reaction (both steps). The effect of MPA on the second step of the RNA reaction (i.e. the transfer of the GMP moiety onto an acceptor RNA) was next evaluated by pre-incubating the enzyme with GTP to ensure formation of the covalent enzyme-GMP complex, followed by the addition of the acceptor 5'-diphosphate RNA in the presence of MPA. Using such an approach, it was determined that a concentration of 640 mM of MPA is sufficient to inhibit 50% of the second step of the GTase reaction (Fig. 2D). Taken together, these results indicate that MPA inhibits the GTase reaction mainly through inhibition of the catalytic transfer of the GMP moiety onto an acceptor RNA.



**Figure 2.** Mycophenolic acid inhibits the RNA guanylyltransferase activity.

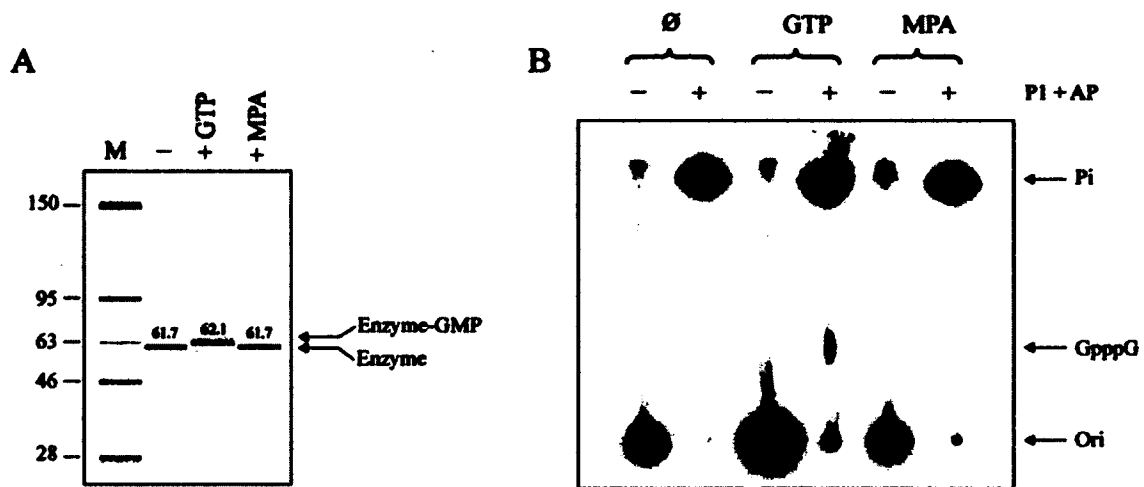
(A) Molecular structure of mycophenolic acid (MPA). (B) Increasing concentrations of MPA inhibit the complete RNA guanylyltransferase reaction. A standard GTase assay in which the purified enzyme (1  $\mu$ M) was incubated with both [ $\alpha$ -<sup>32</sup>P]GTP and a 5'-diphosphate acceptor RNA was performed in the presence of increasing concentrations of MPA. The reaction products were analyzed on a denaturing polyacrylamide gel and quantified (*right side of the panel*). (C) MPA is not a potent inhibitor of the first step of the GTase reaction. The formation of the enzyme-GMP covalent intermediate was monitored by incubating the purified enzyme (1  $\mu$ M) in the presence of [ $\alpha$ -<sup>32</sup>P]GTP and increasing concentrations of MPA. The radiolabeled covalent enzyme-GMP complex was then visualized by autoradiography following electrophoresis on a denaturing 12.5% polyacrylamide gel. The radiolabeled enzyme-GMP complex was quantified by phosphorimaging (*right side of the panel*). (D) The second step of the GTase reaction is inhibited by MPA. The transfer of the GMP moiety onto an acceptor RNA was evaluated by pre-incubating the enzyme (1  $\mu$ M) with [ $\alpha$ -<sup>32</sup>P]GTP (10 mM) to ensure formation of the radiolabeled covalent

enzyme-GMP complex, followed by the addition of the acceptor 5'-diphosphate RNA (3  $\mu$ M) in the presence of MPA. Formation of the radiolabeled capped GpppRNA was monitored following electrophoresis on a denaturant polyacrylamide gel. The radiolabeled GpppRNA was then quantified by phosphorimaging (*right side of the panel*).

### *MPA is not a substrate for the RNA guanylyltransferase*

In the typical GTase reaction, the nucleophilic attack on the aphosphate of GTP by the enzyme during the first step of the reaction results in the formation of a covalent intermediate in which GMP is linked via a phosphoamide bond to a lysine residue of the enzyme [2]. Interestingly, it has previously been demonstrated that the nucleotide analog ribavirin triphosphate can be used as a substrate by the GTase to form an enzyme-RMP covalent intermediate [25]. Although MPA is not a nucleoside analog per se, we were still interested to monitor its ability to potentially be used as a substrate by the enzyme. The appearance of a slower migrating protein species is traditionally observed upon electrophoresis through a polyacrylamide gel when an RNA capping enzyme is incubated with GTP [25]. This slower migrating species corresponds to the enzyme with covalently bound GMP. We therefore incubated the enzyme in the presence of GTP or MPA and the polypeptide was analyzed by capillary electrophoresis. The appearance of a slower migrating protein species was observed repeatedly when the protein was incubated with GTP (Fig. 3A). However, this was not observed when the enzyme was incubated with MPA (Fig. 3A). We therefore conclude that MPA is not a substrate for the GTase. Accordingly, the transfer of MPA to an acceptor RNA could not be detected when the enzyme was incubated in the presence of MPA (Fig. 3B). This was tested by incubating the enzyme with an RNA substrate (81 nt) synthesized in the presence of [ $\alpha$ -<sup>32</sup>P]GTP. This RNA substrate harbored a radiolabelled diphosphate 5'-end (5' **ppG**-RNA 3', where the boldface indicates the radiolabelled moiety) and 27 internally labelled guanosine residues. The RNA was then incubated with the GTase in the presence of GTP or MPA. The products of the reaction were extracted with phenol/chloroform and the RNA acceptor molecules were recovered by ethanol precipitation. Aliquots of the RNA samples were then digested with nuclease P1 and alkaline phosphatase and analyzed by polyethyleneiminecellulose thin layer chromatography (Fig. 3B). The transfer of

radiolabeled GMP to RNA was confirmed by demonstrating the release of a GpppG structure following digestion of the RNA samples with both nuclease P1 and alkaline phosphatase. However, the transfer of MPA to the acceptor RNA could not be detected. Overall, these results demonstrate that MPA is not a substrate for the GTase, and that it cannot be transferred to an acceptor RNA.



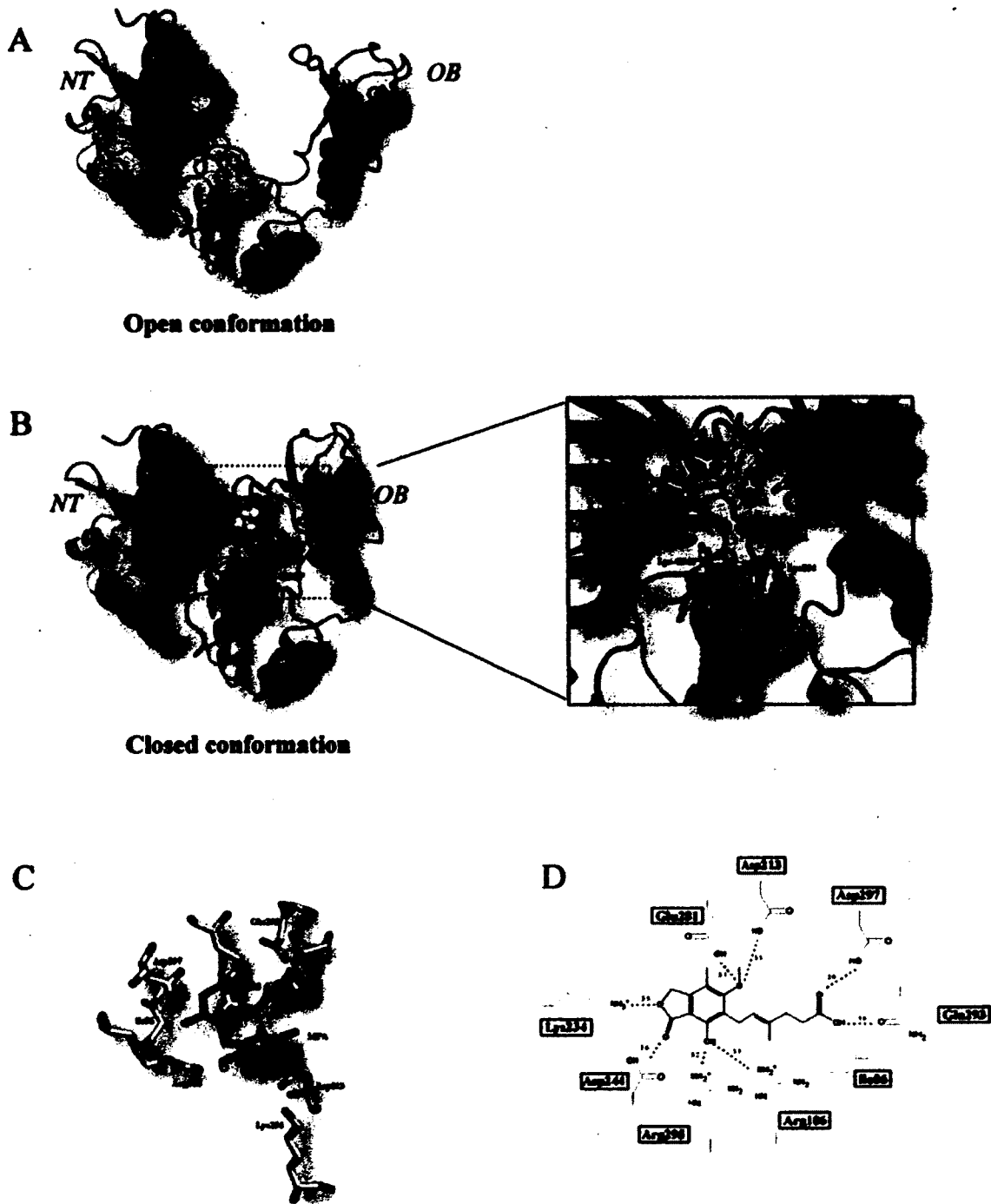
**Figure 3.** MPA is not a substrate for the RNA guanylyltransferase.

(A) Capillary electrophoresis analysis of the RNA capping reaction. The GTase reactions were performed in the presence of the purified enzyme (1  $\mu$ M) and GTP (1 mM) or MPA (1 mM), and the reaction products were analyzed by capillary electrophoresis. An untreated protein was also used as a control (-). The positions and sizes (in kDa) of the size markers (M) are indicated on the left. Masses are shown above the corresponding bands. (B) RNA capping reaction. The reaction mixtures contained 1  $\mu$ g of purified enzyme, 23 pmol of radiolabelled 5' diphosphate-terminated RNA (5' ppG-RNA 3', where the boldface indicates the radiolabelled moiety), and either 1 mM GTP or 1 mM MPA. An untreated control was also used in these assays ( $\emptyset$ ). The reactions were incubated at 30°C for 30 min, and unincorporated nucleotides were removed by multiple rounds of ethanol precipitation. The RNAs were extracted with phenol/chloroform and recovered by ethanol precipitation. Aliquots of the RNA samples were adjusted to 50 mM NaOAc, pH 5.2, and digested with nuclease P1 (5  $\mu$ g) for 60 min at 37°C. The reaction was then adjusted to 50 mM Tris-HCl, pH 8.0, and digested with alkaline phosphatase (1 unit) for 60 min at 37°C (P1+AP). The reaction products were analyzed by thin layer chromatography on a polyethyleneimine-cellulose plate developed with 0.5 M LiCl and 1 M formic acid. An autoradiogram of the plate is shown. The positions of the chromatographic origin (*ori*), inorganic phosphate (*Pi*), and GpppG are indicated.

### *Mechanistic/Structural implications*

In order to better understand the mechanism by which MPA inhibits the transfer of GMP to RNA, we set out to use the power of molecular docking to provide information on the interaction between MPA and both the open and closed conformers of the RNA guanylyltransferase (Fig. 4A and B). Using extensive computational docking and structure optimizations, we generated models of MPA bound to the closed form enzyme-GMP complex of the *Chlorella* virus GTase. More than 2,000,000 energy evaluations were performed in order to provide an accurate of steepest descent energy minimization and did not contain energetically unfavorable bonds, angles or torsions. Analysis of the molecular docking model predicts that MPA binds in the cleft created by the closure of the N-terminal and C-terminal domains. The space-filling model suggests that the molecular structure of MPA is sterically complementary to the cleft between the domains (Fig. 4B). The molecular docking model provides instructive findings on the interaction between specific residues of the enzyme and MPA (Fig. 4B). Arg106 and Asp213 of the N-terminal NT domain and Lys234, Asp244, Gln293, Arg295, Asp297, and Glu281 of the C-terminal OB fold domain would be involved in the coordination of the MPA through hydrogenbonding (Fig. 4C). Ile86 (N-terminal) is predicted to make a hydrophobic interaction with MPA, which is located 4.0 Å away from the bound GMP (Fig. 4D). As observed in our initial virtual screening, the molecular docking analyses suggest that MPA only binds to the closed form of the enzyme. No significant binding of MPA was detected when the molecular docking experiments were performed on the open forms of the *C. albicans*, *S. cerevisiae* and *Chlorella* virus RNA guanylyltransferases bound with GMP. It is tempting to speculate that the opening of the active site of the enzyme that is normally observed following the hydrolysis of GTP to produce the enzyme-GMP covalent intermediate [23] is inhibited by MPA. Following the formation of the GMP adduct, the enzyme must open up to provide access for the incoming mRNA substrate, since this site is blocked off in the closed form of the enzyme [23]. The presence of multiple interactions between MPA and residues of both the N-terminal and C-terminal domains of the enzyme likely inhibits this critical conformational change that is required for the binding of RNA. In order to validate this hypothesis, cross-linking assays were used to monitor the binding of radiolabelled RNA to the enzyme-GMP

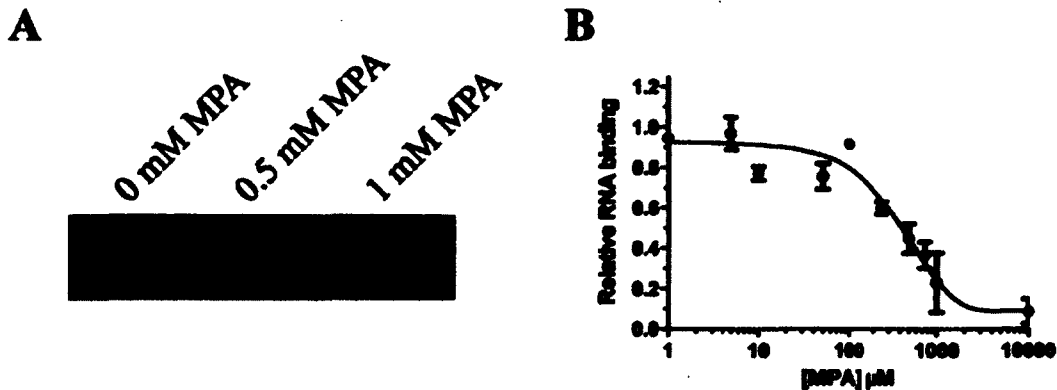
complex with or without MPA. An RNA harboring a 5'-triphosphate end was used as a ligand instead of the classical RNA with a 5'-diphosphate since the latter would lead to the transfer of the GMP moiety onto the RNA. Using this approach, an apparent  $K_d$  of 30 mM could be estimated for the binding of the RNA to the enzyme-GMP complex (data not shown). However, the binding of RNA to the enzyme-GMP/MPA complex was severely decreased in the presence of MPA (Fig. 5A). An  $IC_{50}$  value of 410 mM could be estimated from the inhibition assays (Fig. 5B). These results demonstrate that the presence of MPA inhibits the binding of the enzyme-GMP complex to RNA.



**Figure 4.** Molecular docking model for the binding of MPA to the closed form of the *Chlorella* virus GTase bound to GMP.

(A) Ribbon diagram of the open form of the enzyme (PDB: 1CKO). The positions of the N-terminal nucleotidyl transferase (*NT*) and the C-terminal oligomer-binding (*OB*) fold domains are indicated. (B) Ribbon diagrams looking at the interaction of MPA with the GTase (PDB: 1CKN) bound to GMP. (C) Close-up view of the MPA binding pocket with emphasis on the residues interacting with MPA. (D) The

side chains of amino acids that are predicted to interact with MPA are shown, and the distances are indicated (*in angströms*).



**Figure 5.** Binding of MPA to the enzyme-GMP complex inhibits the interaction with RNA.

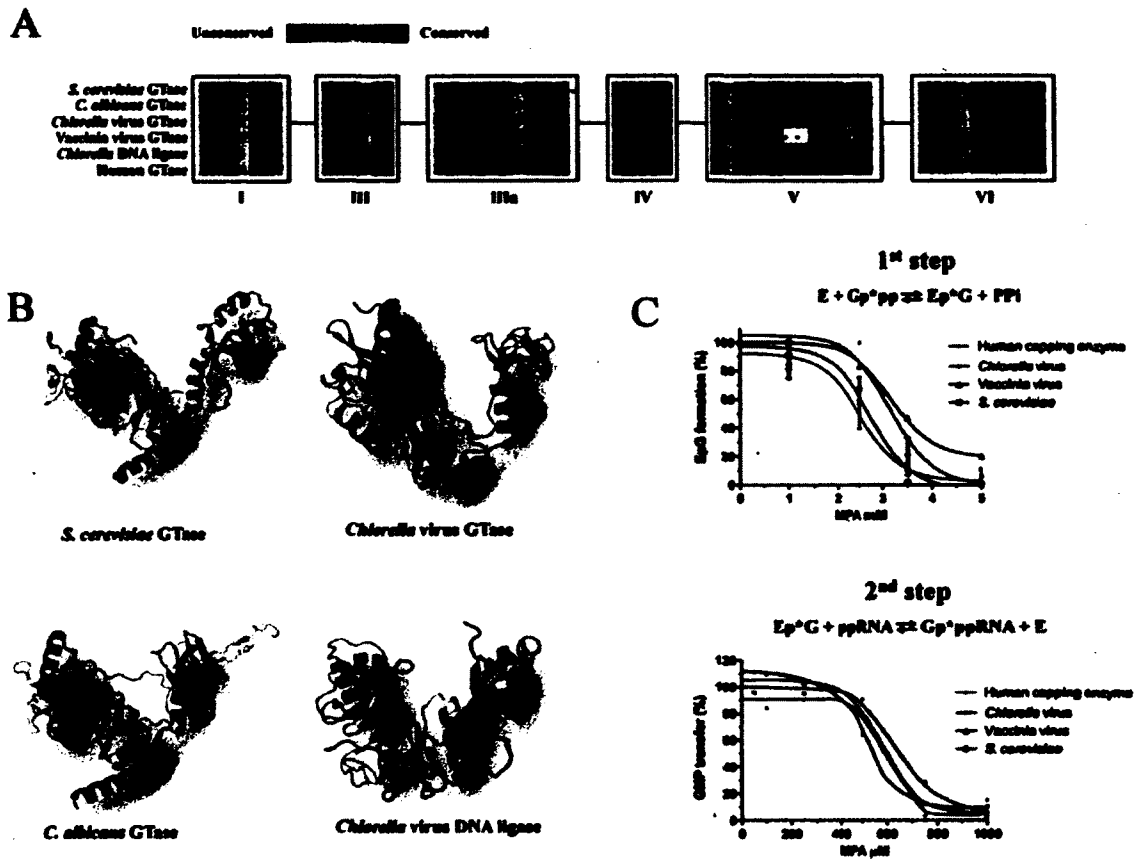
(A) The enzyme-GMP complex was incubated with a radiolabeled RNA substrate (3  $\mu\text{M}$ ) of 81 nucleotides in the presence of increasing concentrations of MPA. UV-cross-linking assays were performed to monitor the binding of radiolabelled RNA to the enzyme-GMP complex and visualized by SDS-PAGE analysis and autoradiography. (B) The reaction products were quantified by phosphorimaging.

### *Specificity*

GTases are members of the RNA/DNA nucleotidyltransferase superfamily that share six conserved sequence motifs (Fig. 6A) and a similar three-dimensional architecture consisting of an N-terminal NT domain and a C-terminal OB fold domain [21]. As observed previously, the NT domain (aa 1–243) of the *C. albicans* GTase aligns to the *Chlorella virus* GTase enzyme with 1.9  $\text{\AA}$  rmsd over 210 amino acids (26% side chain identity) [53]. In addition, the *C. albicans* GTase OB domain (aa 244–390) aligns to the *Chlorella virus* GTase OB domain with 1.8  $\text{\AA}$  rmsd over 72 amino acids (28% identity) [53]. Because of the high level of structural conservation between members of this family (Fig. 6B), we hypothesized that MPA should inhibit the GTases of various organisms. We therefore expressed and purified the RNA guanylyltransferases from vaccinia virus



and *Chlorella* virus, as well as the corresponding human enzyme. Our results demonstrate that MPA inhibits all these GTases to the same extent than the *S. cerevisiae* homolog (Fig. 6C). IC50 values similar to the ones obtained for both the first (0.3 mM) and second (0.620 mM) of the GTase reaction were observed when MPA was added to the reactions catalyzed by the vaccinia, *Chlorella*, and human enzymes. Overall, these results indicate that MPA inhibits the activity of different GTases through a similar mechanism that mainly prevents the catalytic transfer of the GMP moiety onto an acceptor RNA.



**Figure 6.** MPA inhibits the GTases of various origins.

(A) Sequence conservation in members of the RNA/DNA nucleotidyltransferase superfamily. Members of this family share six conserved motifs (I, III, IIIa, IV, V, VI). An amino acid alignment of the GTases from *S. cerevisiae*, *C. albicans*, *Chlorella* virus, vaccinia virus, human, and the DNA ligase of *Chlorella* virus is presented. (B) Members of the RNA/DNA nucleotidyltransferase superfamily harbor a similar three-dimensional architecture consisting of an N-terminal NT domain and a C-terminal OB fold domain. The structures of the *S. cerevisiae* GTase (PDB: 3KYH), *C. albicans* GTase (PDB: 1P16), *Chlorella* virus GTase (PDB: 1CKN), and *Chlorella* virus DNA ligase (PDB: 1P8L) are shown. (C) The effect of MPA on the first and second step of the GTase reaction was monitored on the GTases from *S. cerevisiae*, *Chlorella*

virus, vaccinia virus, and human. Both reactions were performed in the presence of increasing concentrations of MPA as described in the legend of figure 1.

### *Inhibition of DNA ligase*

GTases are members of the covalent nucleotidyl transferase superfamily, which also includes DNA and RNA ligases [21]. The crystal structures of various family members revealed a common tertiary structure consisting of an N-terminal NT domain and a C-terminal OB-fold domain [22,23,53,54,55]. Based on both the sequence and structural similarities between GTases and DNA/RNA ligases (Fig. 6A and B), it is tempting to speculate that MPA might also be an inhibitor of DNA/RNA ligases. DNA/RNA ligation entails three sequential nucleotidyl transfer steps, and the first two steps are mechanistically related to the GTase reaction [56]. In the first step of the ligation reaction, nucleophilic attack on the  $\alpha$ -phosphorus of ATP by ligase results in the formation of a covalent ligase-adenylate intermediate with the concomitant release of pyrophosphate. In the second step, AMP is transferred to the 5'-end of the 5'-phosphate-terminated DNA strand to form a DNA-adenylate complex. In the last step of the reaction, the polynucleotides are joined with the concomitant release of AMP [56]. To verify the potential ability of MPA to inhibit DNA/RNA ligases, we have monitored the effect of MPA on the purified DNA ligase encoded by Chlorella virus. The ability of MPA to inhibit the complete ligation reaction was monitored by incubating the purified enzyme with ATP and a DNA duplex containing a centrally located nick. The reaction products were then analyzed on a denaturing polyacrylamide gel. As can be seen in Fig. 7A, the presence of increasing concentrations of MPA significantly decreased strand joining by the Chlorella virus DNA ligase. MPA was shown to inhibit the strand-joining activity by 50% at 700 mM. We next set out to investigate which of the first two catalytic steps of the ligation reaction is inhibited by MPA. The first step of the reaction, i.e. the formation of the enzyme-AMP covalent intermediate, was monitored by incubating the purified enzyme in the presence of [ $\alpha$ - $^{32}$ P]ATP. The radiolabeled covalent enzyme-AMP complex was then visualized by autoradiography following electrophoresis on a denaturant polyacrylamide gel. Similarly to what was observed in the case of the GTase

activity, our results indicate that very high concentrations of MPA are required to inhibit the first step of the reaction, i.e. the formation of the enzyme-AMP covalent complex (Fig. 7B). We determined that a concentration of, 3.0 mM of MPA is required to inhibit the first step of the ligation reaction by 50%, a concentration that corresponds to 4-times the amount required to inhibit the overall reaction. The effect of MPA on the second step of the ligation reaction (i.e. the transfer of the AMP from ligase-adenylate to a 5'-phosphate terminus acceptor DNA) was next evaluated. In a typical ligation reaction, the adenylate-DNA intermediate would not be detected since the enzyme is highly efficient in ligating the two DNA strands [57]. However, as reported previously [58], the adenylate-DNA intermediate can accumulate to high levels when the enzyme acts on a substrate that contains a 1-nt gap between the reactive 3'-OH and 5'-phosphate strands. Using such an approach, the incubation of the enzyme with ATP and a gapped DNA substrate resulted in the conversion of the 5'-radiolabeled 18-mer strand into an adenylated species (AppDNA) that migrated 1-nt slower than the input 18-mer during polyacrylamide gel electrophoresis (Fig. 7C). The ability of MPA to inhibit this second step of the ligase reaction was investigated by pre-incubating the enzyme with ATP to ensure formation of the covalent enzyme-AMP complex, and by incubating the complex in the presence of the ligation substrate that contains a 1-nt gap between the reactive 3'-OH and 5'-phosphate strands. Using this strategy, it was determined that a concentration of 600 mM of MPA is sufficient to inhibit 50% of the second step of the ligation reaction (Fig. 7C). Taken together, these results indicate that MPA inhibits the ligation reaction mainly through inhibition of the second step of the reaction, i.e. the catalytic transfer of the AMP moiety onto the 5'-phosphate terminus of the nicked DNA substrate.

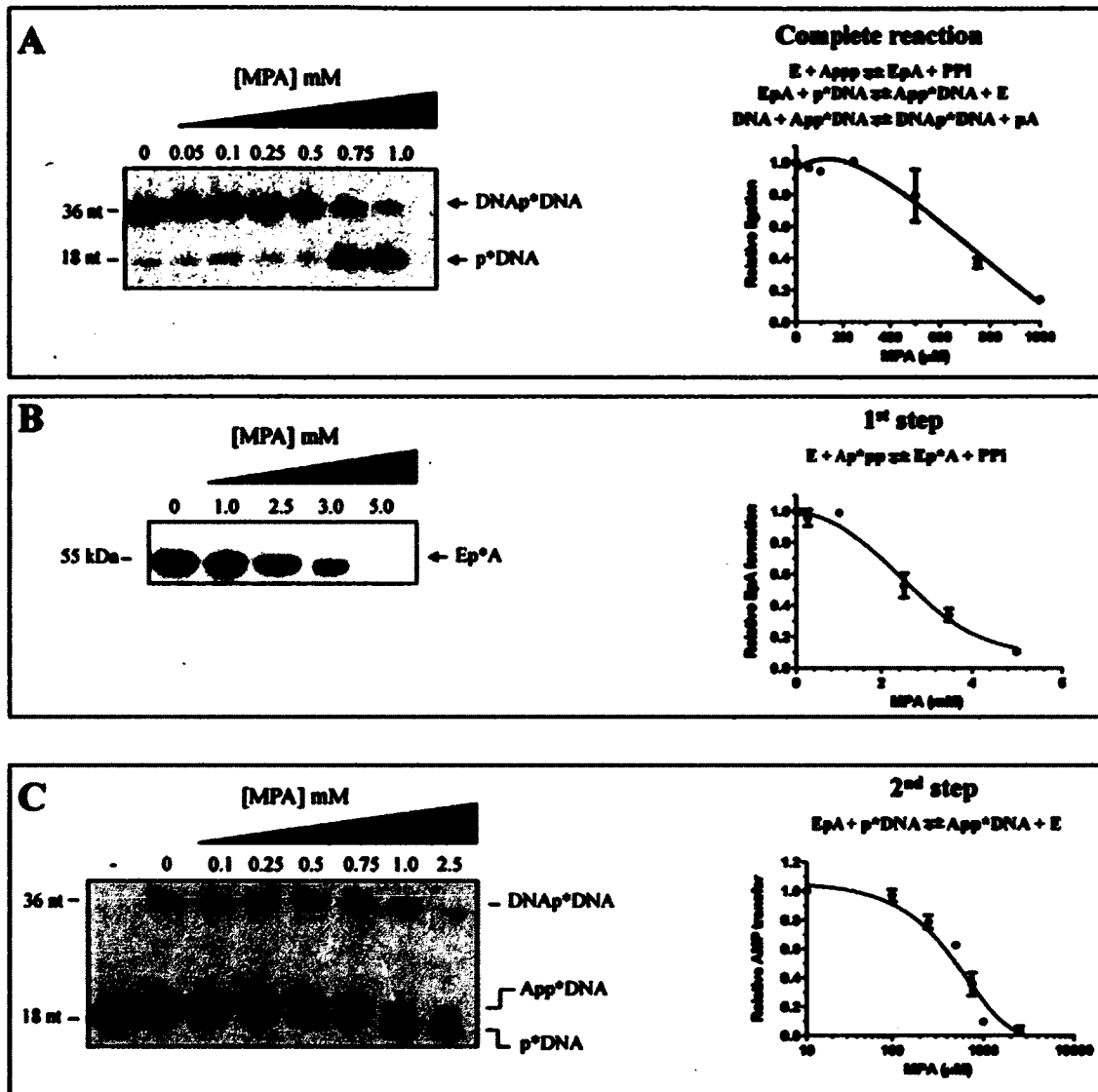


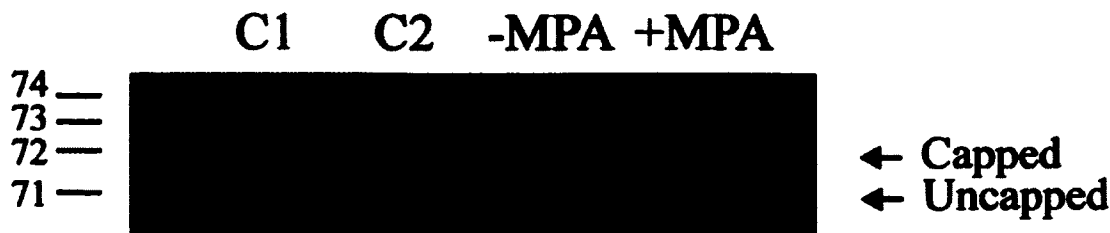
Figure 7. MPA inhibits the DNA ligase activity.

(A) MPA inhibits the strand-joining activity of DNA ligase. Ligation reactions (50 mM Tris-HCl pH 7.5, 5 mM DTT, 10 mM MgCl<sub>2</sub>, 1 mM ATP, 500 fmol 1-nt gap DNA substrate and 6 pmol of the *Chlorella* virus DNA ligase) were performed at 22°C for 15 min in the presence of increasing concentrations of MPA. The samples were analyzed by electrophoresis through a 17% polyacrylamide gel containing 7 M urea. An autoradiogram of the gel is shown. The positions of the input 5'-monophosphate 18-mer strand (pDNA) and the 36-mer ligation product are indicated. The radiolabeled ligated product was then quantified by phosphorimaging (right side of the panel). (B) High concentrations of MPA are required to inhibit the first step of the ligase reaction. The formation of the enzyme-AMP covalent intermediate was monitored by incubating the purified enzyme in the presence of [ $\alpha$ -<sup>32</sup>P]ATP and increasing concentrations of MPA. The radiolabeled covalent enzyme-AMP complex was then visualized by autoradiography following electrophoresis on a denaturant polyacrylamide gel. The radiolabeled enzyme-AMP complex was then quantified by phosphorimaging (right side of the panel). (C) The second step of the ligase reaction is inhibited by MPA. The transfer of the AMP moiety onto a radiolabeled 5'-monophosphate 18-mer strand (pDNA) was evaluated by pre-incubating the enzyme with ATP to ensure formation of the radiolabeled

covalent enzyme-AMP complex, followed by the addition of a 1-nt gapped substrate in the presence of increasing concentrations of MPA. Conversion of the 5'-<sup>32</sup>P-labeled 18-mer strand into an adenylated species (AppDNA) was monitored by electrophoresis on a denaturant polyacrylamide gel. Lane 1: reaction performed in the absence of protein (-). The formation of the radiolabeled AppDNA was quantified by phosphorimaging (*right side of the panel*).

### *Inhibition of RNA cap formation in cells*

The effect of MPA on the capping of mRNAs was then assessed *in vivo*. A primer extension assay was used to monitor the effect of MPA on the formation of the RNA cap structure in *S. cerevisiae*. Previous studies have shown that during cDNA synthesis by reverse transcriptase, the presence of a cap structure results in the synthesis of products that harbor an extra 3' nucleotide [40,59]. We therefore performed primer extension assays with a 5' P<sup>32</sup>-labeled 18-mer oligonucleotide complementary to the 5' region of the SSA1 mRNA. The oligonucleotide was annealed to total mRNAs extracted from cells that were grown in the presence or absence of 500 mg/ml MPA for 3 h, and extended with reverse transcriptase. Our results indicate that cDNA products with apparent chain lengths of 71 and 72 nucleotides, corresponding to uncapped and capped mRNAs, were synthesized using mRNAs extracted from untreated cells (Fig. 8). However, treatment with MPA resulted in a marked reduction of the 72 nucleotides species, indicating that cap formation was inhibited in the presence of MPA. Quantitative analysis indicate that cap formation was reduced by 50% in the presence of 500 mg/ml MPA (Fig. 8). We conclude that the addition of MPA to *S. cerevisiae* cells leads to an inhibition of RNA cap synthesis.



**Figure 8.** MPA inhibits RNA capping in *S. cerevisiae* cells.

Primer extension assays were performed with a 5' P<sup>32</sup>-labeled 18-mer oligonucleotide complementary to the 5' region of the *SSA1* mRNA. The oligonucleotide was annealed to total mRNAs extracted from cells that were grown in the absence (-MPA) or presence (+MPA) of 500 µg/ml MPA for 3 h, and extended with reverse transcriptase. The primer extension reactions were analyzed by electrophoresis through a 8% polyacrylamide gel containing 7 M urea in TBE and visualized by autoradiography. Control P<sup>32</sup>-labeled RNA transcripts of 74 (C1) and 73 (C2) nucleotides were run in parallel. The positions and sizes (in nt) of the size markers are indicated on the left.

## Conclusion

The current study provides the first biochemical evidence that MPA can directly interact with an RNA guanylyltransferase and inhibit its activity. We demonstrated that MPA inhibits the RNA guanylyltransferase reaction by preventing the catalytic transfer of the GMP moiety onto an acceptor RNA. Moreover, our RNA binding studies demonstrated that the binding of the enzyme-GMP intermediate to RNA is inhibited in the presence of MPA. As such, MPA represents a novel type of inhibitor against RNA guanylyltransferases that inhibits the second step of the catalytic reaction. Several inhibitors of the RNA guanylyltransferase activity have previously been identified [60]. However, these inhibitors all target the formation of the enzyme-GMP complex. For instance, foscarnet inhibits the formation of the enzyme-GMP intermediate on account of its analogous nature to pyrophosphate (PPi), a product of the RNA guanylyltransferase reaction [24]. Ribavirin, a broad-spectrum nucleoside analogue used as an antiviral for severe respiratory syncytial virus, Hepatitis C and other viral infections, is another example of an RNA guanylyltransferase inhibitor that prevents the formation of the

enzyme-GMP intermediate [25]. Non-nucleoside competitive inhibitors have also been generated against the RNA guanylyltransferase of respiratory syncytial virus [61]. MPA, with its ability to inhibit the second step of the RNA guanylyltransferase reaction, has the potential to serve as a template for the development of more potent inhibitors. In fact, a number of MPA derivatives have been developed in recent years [62,63]. Analysis of the interaction between these derivatives and RNA guanylyltransferases could shed light on the chemistry of the RNA capping reaction and lead to the development of more efficient anti-proliferative/antimicrobial drugs.

What is the biological relevance of the present finding? MPA has been shown to cause a reduction of the cellular GTP pools through the inhibition of IMPDH which is required for the de novo biosynthesis of GTP [41,42]. A decrease in GTP concentrations could potentially have a negative effect on the capping of mRNAs. Evidence for this mechanism comes from studies performed with both the Sindbis virus and the Borna disease virus that showed that the viruses cannot replicate in cultured cells treated with ribavirin because the level of GTP falls too low to permit the capping of viral RNAs [48,64,65]. However, mounting evidence indicates that the antiproliferative/antimicrobial effect of MPA is not mediated entirely through the reduction of the intracellular GTP pool. For instance, MPA is not always a potent viral inhibitor [66]. The second mechanism by which MPA could inhibit the capping of mRNAs is by directly inhibiting the activity of GTases, as we demonstrated in the current study. Moreover, we do not exclude the possibility that MPA may have an increased frequency of utilization in vivo when the levels of GTP are lowered through the inhibition of IMPDH. In the present study, we demonstrated that MPA can inhibit RNA guanylyltransferases and ligases. MPA therefore appears as a pleiotropic agent that may function through multiple targets (IMPDH, RNA guanylyltransferase, ligases), as recently suggested through proteomic analysis [67]. As such, MPA shares many properties with ribavirin. Both compounds lead to a reduction of the de novo synthesis of GTP through IMPDH inhibition, and both molecules can inhibit RNA capping, albeit through different mechanisms of inhibition.

## **Références**

1. Furuichi Y, Shatkin AJ (2000) Viral and cellular mRNA capping: past and prospects. *Adv Virus Res* 55: 135–184.
2. Shuman S (2001) Structure, mechanism, and evolution of the mRNA capping apparatus. *Prog Nucleic Acid Res Mol Biol* 66: 1–40.
3. Bisaillon M, Lemay G (1997) Viral and cellular enzymes involved in synthesis of mRNA cap structure. *Virology* 236: 1–7.
4. Ghosh A, Lima CD (2010) Enzymology of RNA cap synthesis. *WIREs RNA* 1:152–172.
5. Issur M, Geiss BJ, Bougie I, Picard-Jean F, Despins S, et al. (2009) The flavivirus NS5 protein is a true RNA guanylyltransferase that catalyzes a two-step reaction to form the RNA cap structure. *RNA* 15: 2340–2350.
6. Bisaillon M, Lemay G (1997) Characterization of the reovirus lambda1 protein RNA 59-triphosphatase activity. *J Biol Chem* 272: 29954–29957.
7. Bouvet M, Debarnot C, Imbert I, Selisko B, Snijder EJ, et al. (2010) In vitro reconstitution of SARS-coronavirus mRNA cap methylation. *PLoS Pathog* 6: e1000863.
8. Myette JR, Niles EG (1996) Characterization of the vaccinia virus RNA 5'-triphosphatase and nucleotide triphosphate phosphohydrolase activities. Demonstrate that both activities are carried out at the same active site. *J Biol Chem* 271: 11945–11952.
9. Shuman S, Hurwitz J (1981) Mechanism of mRNA capping by vaccinia virus guanylyltransferase: characterization of an enzyme-guanylate intermediate. *Proc Natl Acad Sci U S A* 78: 187–191.



10. Ho CK, Van Etten JL, Shuman S (1996) Expression and characterization of an RNA capping enzyme encoded by Chlorella virus PBCV-1. *J Virol* 70:6658–6664.
11. Pugh CS, Borchardt RT, Stone HO (1978) Sinefungin, a potent inhibitor of virion mRNA(guanine-7-)-methyltransferase, mRNA(nucleoside-29-)-methyltransferase, and viral multiplication. *J Biol Chem* 253: 4075–4077.
12. Pugh CS, Borchardt RT (1982) Effects of S-adenosylhomocysteine analogues on vaccinia viral messenger ribonucleic acid synthesis and methylation. *Biochemistry* 21: 1535–1541.
13. Zheng S, Hausmann S, Liu Q, Ghosh A, Schwer B, et al. (2006) Mutational analysis of Encephalitozoon cuniculi mRNA cap (guanine-N7) methyltransferase, structure of the enzyme bound to sinefungin, and evidence that cap methyltransferase is the target of sinefungin's antifungal activity. *J Biol Chem* 281: 35904–35913.
14. Paolantonacci P, Lawrence F, Robert-Gero M (1985) Differential effect of sinefungin and its analogs on the multiplication of three Leishmania species. *Antimicrob Agents Chemother* 28: 528–531.
15. Benghiat E, Crooks PA, Goodwin R, Rottman F (1986) Inhibition of vaccinia RNA guanine 7-methyltransferase by compounds designed as multisubstrate adducts. *J Pharm Sci* 75: 142–145.
16. Luzhkov VB, Selisko B, Nordqvist A, Peyrane F, Decroly E, et al. (2007) Virtual screening and bioassay study of novel inhibitors for dengue virus mRNA cap (nucleoside-29O)-methyltransferase. *Bioorg Med Chem* 15: 7795–7802.
17. Podvinec M, Lim SP, Schmidt T, Scarsi M, Wen D, et al. (2010) Novel inhibitors of dengue virus methyltransferase: discovery by in vitro-driven virtual screening on a desktop computer grid. *J Med Chem* 53: 1483–1495.
18. Lima CD, Wang LK, Shuman S (1999) Structure and mechanism of yeast RNA triphosphatase: an essential component of the mRNA capping apparatus. *Cell* 99: 533–543.

19. Gu M, Lima CD (2005) Processing the message: structural insights into capping and decapping mRNA. *Curr Opin Struct Biol* 15: 99–106.
20. Issur M, Despins S, Bougie I, Bisailon M (2009) Nucleotide analogs and molecular modeling studies reveal key interactions involved in substrate recognition by the yeast RNA triphosphatase. *Nucleic Acids Res* 37: 3714–3722.
21. Shuman S, Liu Y, Schwer B (1994) Covalent catalysis in nucleotidyl transfer reactions: essential motifs in *Saccharomyces cerevisiae* RNA capping enzyme are conserved in *Schizosaccharomyces pombe* and viral capping enzymes and among polynucleotide ligases. *Proc Natl Acad Sci U S A* 91: 12046–12050.
22. Doherty AJ, Hakansson K, Ho CK, Shuman S, Wigley DB (1997) Crystallization of the RNA guanylyltransferase of *Chlorella virus* PBCV-1. *Acta Crystallogr D Biol Crystallogr* 53: 482–484.
23. Hakansson K, Doherty AJ, Shuman S, Wigley DB (1997) X-ray crystallography reveals a large conformational change during guanyl transfer by mRNA capping enzymes. *Cell* 89: 545–553.
24. Souliere MF, Perreault JP, Bisailon M (2008) Kinetic and thermodynamic characterization of the RNA guanylyltransferase reaction. *Biochemistry* 47: 3863–3874.
25. Bougie I, Bisailon M (2004) The broad spectrum antiviral nucleoside ribavirin as a substrate for a viral RNA capping enzyme. *J Biol Chem* 279: 22124–22130.
26. Yan Y, Svitkin Y, Lee JM, Bisailon M, Pelletier J (2005) Ribavirin is not a functional mimic of the 7-methyl guanosine mRNA cap. *RNA* 11: 1238–1244.
27. Ewing TJ, Makino S, Skillman AG, Kuntz ID (2001) DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. *J Comput Aided Mol Des* 15: 411–428.

28. Fontaine F, Bolton E, Borodina Y, Bryant SH (2007) Fast 3D shape screening of large chemical databases through alignment-recycling. *Chemistry Central Journal* 1: 1–14.
29. Wang R, Lai L, Wang S (2002) Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J Comput Aided Mol Des* 16: 11–26.
30. Bougie I, Parent A, Bisaillon M (2004) Thermodynamics of ligand binding by the yeast mRNA-capping enzyme reveals different modes of binding. *Biochem J* 384: 411–420.
31. Wen Y, Yue Z, Shatkin AJ (1998) Mammalian capping enzyme binds RNA and uses protein tyrosine phosphatase mechanism. *Proc Natl Acad Sci U S A* 95: 12226–12231.
32. Sriskanda V, Shuman S (1998) Mutational analysis of Chlorella virus DNA ligase: catalytic roles of domain I and motif VI. *Nucleic Acids Res* 26: 4618–4625.
33. Bisaillon M, Bougie I (2003) Investigating the role of metal ions in the catalytic mechanism of the yeast RNA triphosphatase. *J Biol Chem* 278: 33963–33971.
34. Mayo SL (1990) DREIDING: a generic force field for molecular simulations. *J Phys Chem* 94: 8897–8909.
35. Stewart JJ (2007) Optimization of parameters for semiempirical methods V: modification of NDDO approximations and application to 70 elements. *J Mol Model* 13: 1173–1213.
36. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, et al. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem* 19: 1639–1661.
37. Solis FJ, Wets RJB (1981) Minimization by random search techniques. *Math Oper Res* 6.

38. Shuman S (1995) Vaccinia virus DNA ligase: specificity, fidelity, and inhibition. *Biochemistry* 34: 16138–16147.
39. Sekiguchi J, Shuman S (1997) Nick sensing by vaccinia virus DNA ligase requires a 5' phosphate at the nick and occupancy of the adenylate binding site on the enzyme. *J Virol* 71: 9679–9684.
40. Schwer B, Mao X, Shuman S (1998) Accelerated mRNA decay in conditional mutants of yeast mRNA capping enzyme. *Nucleic Acids Res* 26: 2050–2057.
41. Qiu Y, Fairbanks LD, Ruckermann K, Hawrlowicz CM, Richards DF, et al. (2000) Mycophenolic acid-induced GTP depletion also affects ATP and pyrimidine synthesis in mitogen-stimulated primary human T-lymphocytes. *Transplantation* 69: 890–897.
42. Ichikawa Y, Ihara H, Takahara S, Takada K, Shrestha GR, et al. (1984) The immunosuppressive mode of action of mizoribine. *Transplantation* 38: 262–267.
43. Allison AC, Hovi T, Watts RW, Webster AD (1975) Immunological observations on patients with Lesch-Nyhan syndrome, and on the role of denovo purine synthesis in lymphocyte transformation. *Lancet* 2: 1179–1183.
55. Lee JY, Chang C, Song HK, Moon J, Yang JK, et al. (2000) Crystal structure of NAD(+)-dependent DNA ligase: modular architecture and functional implications. *EMBO J* 19: 1119–1129.
56. Tomkinson AE, Levin DS (1997) Mammalian DNA ligases. *Bioessays* 19: 893–901.
57. Ho CK, Van Etten JL, Shuman S (1997) Characterization of an ATP-dependent DNA ligase encoded by Chlorella virus PBCV-1. *J Virol* 71: 1931–1937.
58. Sriskanda V, Shuman S (1998) Chlorella virus DNA ligase: nick recognition and mutational analysis. *Nucleic Acids Res* 26: 525–531.
59. Davison AJ, Moss B (1989) Structure of vaccinia virus early promoters. *J Mol Biol* 210: 749–769.

60. Issur M, Picard-Jean F, Bisailon M (2011) The RNA capping machinery as an anti-infective target. *WIREs RNA* 2: 184–192.
61. Liuzzi M, Mason SW, Cartier M, Lawetz C, McCollum RS, et al. (2005) Inhibitors of respiratory syncytial virus replication target cotranscriptional mRNA guanylation by viral RNA-dependent RNA polymerase. *J Virol* 79: 13105–13115.
62. Elbarbry FA, Shoker AS (2007) Therapeutic drug measurement of mycophenolic acid derivatives in transplant patients. *Clin Biochem* 40: 752–764.
63. Habib E, Leon F, Bauer JD, Hill RA, Carvalho P, et al. (2008) Mycophenolic Derivatives from *Eupenicillium parvum*. *J Nat Prod* 71: 1915–1918.
64. Scheidel LM, Durbin RK, Stollar V (1987) Sindbis virus mutants resistant to mycophenolic acid and ribavirin. *Virology* 158: 1–7.
65. Jordan I, Briese T, Averett DR, Lipkin WI (1999) Inhibition of Borna disease virus replication by ribavirin. *J Virol* 73: 7903–7906.
66. Planterose DN (1969) Antiviral and cytotoxic effects of mycophenolic acid. *J Gen Virol* 4: 629–630.
67. Escobar-Henriques M, Balguerie A, Monribot C, Boucherie H, Daignan-Fornier B (2001) Proteome analysis and morphological studies reveal multiple effects of the immunosuppressive drug mycophenolic acid specifically resulting from guanylic nucleotide depletion. *J Biol Chem* 276: 46237–46242.

## **Acknowledgment**

We thank Dr. Aaron Shatkin for the generous gift of the cDNA encoding the human capping enzyme. We also thank Dr. Sherif Abou Elela for the generous gift of the yeast strain knock out for Xrn1.



# 3) Discussion

## 3.1. Synthèse

Les résultats issus de la présente étude décrivent le premier exemple d'inhibition allostérique qui cible l'enzyme GTase de la cascade enzymatique de la synthèse de la coiffe. Le protocole utilisé pour ce projet de recherche permet de valider nos résultats autant par des méthodes *in silico*, *in vitro* ainsi qu'*in vivo*. Par ces différentes méthodes expérimentales, le composé MPA a été montré pour interférer avec la mécanique des GTases et d'obstruer le changement conformationnel de ces enzymes. Puisque l'activité catalytique de la GTase est la succession de deux réactions enzymatiques, des essais d'inhibition *in vitro* sur ces étapes ont permis d'identifier que le MPA interfère majoritairement avec la dernière étape enzymatique, soit le transfert du nucléotide sur l'ARN. L'utilisation de l'arrimage moléculaire a permis de prédire le site d'interaction du ligand sur l'enzyme, soit dans la cavité entre les deux domaines, au-dessus du site actif. Finalement, cette étude fait la première démonstration de l'inhibition de l'activité d'une GTase *in vivo*.

Cette étude décrit le mécanisme d'inhibition du MPA sur l'étape limitante de la synthèse de la structure coiffe. La portée de cette étude repose sur le principe que plusieurs pathogènes humains, d'origines virales ou fongiques, encodent pour leurs propres enzymes qui résultent de l'ajout de cette la coiffe sur l'extrémité 5' de leurs ARNm. Puisque les enzymes encodées par les pathogènes possèdent des structures ou des mécanismes catalytiques distincts des homologues humains et quelles soient essentielles à la propagation des pathogènes, elles représentent une excellente avenue thérapeutique. Dans cette optique, des études antérieures ont permis de décrire les déterminants moléculaires des sites actifs d'enzymes pathogènes en utilisant une série d'analogues de nucléotides. Cette approche est très intéressante afin d'identifier des composés ayant un

bon potentiel inhibiteur de la RTase. En effet, la RTase humaine possède un site actif qui reconnaît spécifiquement l'ARN comme substrat alors que les homologues présents chez certains pathogènes possèdent un site actif liant également les nucléotides libres. Il est donc possible de cibler spécifiquement les RTases pathogènes en utilisant la thérapie basée sur des analogues de nucléotides. Par contre, cette approche peut difficilement être appliquée aux GTases puisque plusieurs pathogènes possèdent un site actif qui coordonne tous le même nucléotide avec une très haute conservation. Toutefois, à l'extérieur de ces motifs conservés, les différentes GTases possèdent une très faible homologie. Par cette perspective, cette étude approfondit les connaissances afin de mettre à l'avant-plan une inhibition spécifique pour une GTase.

Avant qu'une thérapie utilisant des composés dirigés vers les GTases soit possible, des études supplémentaires doivent inévitablement se pencher sur plusieurs aspects. En effet, ces différents points seront abordés dans le prochain chapitre, mais au préalable, ce chapitre apporte des explications supplémentaires ne faisant pas partie intégrante de l'article présenté au chapitre précédent.

### ***3.2. Criblage virtuel***

Mes travaux de recherches se sont basés sur les prédictions informatiques obtenues lors du criblage virtuel. Afin d'obtenir des résultats valables, plusieurs facteurs devaient être pris en considération. Tout d'abord, il fallait obtenir des données informatiques quant à la structure tertiaire de la protéine. Puisque notre modèle d'étude utilise la GTase de *S.cerevisiae* et qu'aucun cristal n'est disponible pour cette enzyme, la structure tertiaire de cette enzyme a été prédite par un modèle d'homologie. Par l'entremise de « Protein Data bank », la GTase de *Candida albicans* a permis de servir de modèle pour la prédiction de la structure tridimensionnelle de son homologue chez *S. cerevisiae*. L'utilisation d'un modèle par homologie est une technique qui a fait ses preuves à maintes reprises. Le facteur déterminant afin de prédire une structure adéquate relève d'une bonne homologie entre la protéine cristallisée et celle dont la structure est



inconnue. Les GTases de *C. ablicans* et de *S. cerevisiae* possèdent une homologie à plus de 60 %. De plus, afin de pallier la lacune de la grande majorité des logiciels d'arrimage moléculaire en ce qui à trait à la flexibilité de la protéine, les deux conformations de la GTase ont été intégrées à l'algorithme. Pour ce faire, les cristaux de la GTase du virus de la vaccine, sous la forme ouverte et fermée, ont également été utilisés dans les étapes du criblage virtuel. Les GTases possèdent une structure tridimensionnelle caractérisée par un site actif situé au creux d'une cavité profonde délimitée par deux domaines sur la protéine. Pour les essais d'arrimage moléculaire, un rayon de 20 Angströms autour du site actif permet d'agir en tant que site potentiel d'interaction pour les différents ligands.

Dans un autre ordre d'idée, les ligands ont été sélectionnés en fonction de plusieurs filtres. Tout d'abord, la banque chimiothèque virtuelle « PubChem » qui recoupe plus de 25 millions de composés a servi de point de départ. La taille de la librairie a été réduite de moitié en sélectionnant uniquement les composés qui peuvent être obtenus chez un fournisseur scientifique, dans le but de valider les interactions prédites de façon expérimentales. Puisque la chimiothèque comptait plusieurs millions de composés, il était impératif de réduire considérablement son contenu. Pour identifier des composés susceptibles d'interagir avec la GTase, uniquement les composés dont la structure tridimensionnelle ressemblait à 80 % près de la structure d'une purine, soit l'ATP ou le GTP, ont été sélectionnés. Finalement, la règle de Lipinski a été employée afin de discriminer uniquement les composés qui possèdent un bon potentiel de biodisponibilité. Cette règle, également nommée règle des cinq, regroupe quatre contraintes au niveau de la chimie des composés afin de sélectionner ceux avec un potentiel pharmacologique plus important. Suite à l'application des règles de Lipinski, une banque approximative de mille composés a été utilisée pour le criblage virtuel.

Les résultats du criblage virtuel ont été classés selon leur interaction prédite du complexe enzyme-ligand. En analysant les différents composés prédits pour interagir sur une des deux conformations des GTases, plusieurs composés, des analogues du substrat naturel, se lient au site actif de l'enzyme. Par contre, un composé, l'acide mycophénolique

(MPA), possède un site d'interaction prédit au-dessus du site actif et fait interagir des acides aminés différents de ceux responsables de la liaison au nucléotide. En s'attardant aux détails issus du criblage virtuel, on observe une interaction dix fois plus puissante sur la forme fermée de l'enzyme comparativement à la conformation ouverte. Lorsque l'on s'attarde à la littérature scientifique, le MPA est un composé utilisé à plusieurs fins thérapeutiques soit à titre d'immunosuppresseur, d'antiviral et d'anti-tumorale. La principale cible de l'acide mycophénolique est l'IMPDH, qui est l'étape limitante de la synthèse *de novo* de guanosine intracellulaire. Des études ont mis de l'avant des évidences selon lesquelles ce composé pourrait également interférer avec d'autres voies métaboliques cellulaires. En ce basant sur les résultats du criblage virtuel jumelé à la littérature scientifique, l'interaction entre ce composé et la GTase de *S. cerevisiae* a été analysée.

### ***3.3. Inhibition par l'acide mycophénolique***

Dans l'article figurant dans le chapitre précédant, nous avons montré par deux méthodes différentes que le MPA n'est pas reconnu par la GTase en tant que substrat. Afin de caractériser l'inhibition observée par le MPA, des essais de cinétique d'inhibition seront présentés à titre de résultats supplémentaires. Nos résultats précédents tendent à démontrer que l'interaction du MPA sur la GTase n'est pas au site actif de l'enzyme, mais bien dans un site allostérique. De manière à confirmer cette hypothèse, les différentes réactions de GTase ont été réalisées en utilisant l'enzyme GTase de *S. cerevisiae* purifiée comme mentionnée ultérieurement en présence de différentes concentrations de MPA (0,5 mM à 10 mM). Les réactions enzymatiques ont été réalisées en maintenant constant tous les paramètres à l'exception de la concentration d'inhibiteur. Ce faisant, il était possible de convertir les données sous forme graphique de type Michaelis-Menten afin de déterminer la vitesse de la réaction en fonction de la concentration de substrat et d'inhibiteur. Dans l'optique de déterminer les paramètres cinétiques de l'inhibition de la réaction enzymatique, on peut faire la linéarisation du

graphique pour obtenir la représentation Lineweaver-Burk. De façon congruente avec une inhibition compétitive mixte, l'addition de MPA résulte en une diminution de la vitesse maximale ( $V_{max}$ ) et de la constante d'affinité pour le GTP. Ces résultats démontrent que le MPA ne se lie pas au même site que le substrat naturel de l'enzyme. Une inhibition compétitive mixte est caractérisée par une liaison à la conformation exempte de substrat ainsi qu'une enzyme complexée avec le substrat. Le profil d'inhibition du MPA sur les deux réactions individuelles de la GTase illustre bien que le composé peut bel et bien lier l'enzyme libre et complexée au substrat.

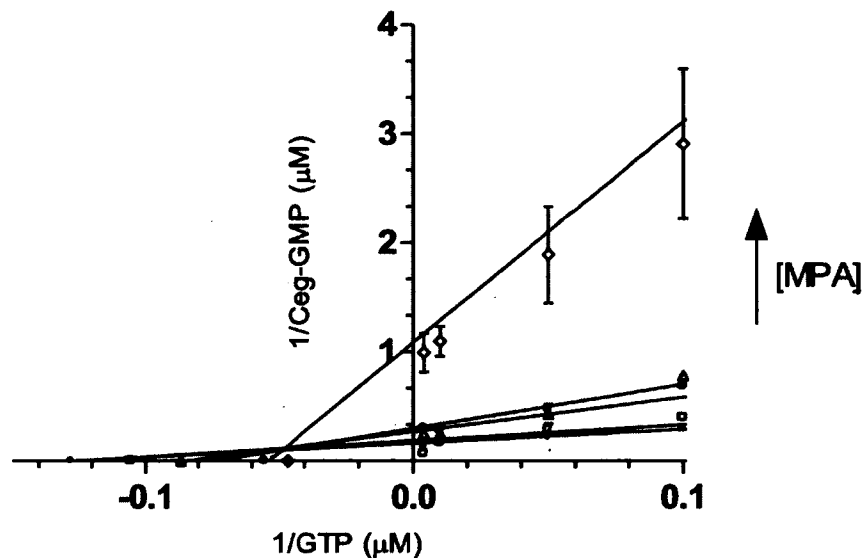


Figure 16. Représentation graphique Lineweaver-burk de l'inhibition du MPA sur la GTase de *S. cerevisiae*. L'activité GTase a été évaluée en absence d'inhibiteur (rond plein) ou en présence de 2,5 mM (carré vide), 5 mM (triangle vide), 7,5 mM (rond vide) et 10 mM (losange vide) de MPA.

### 3.4. Spécificité de l'inhibition

Le composé MPA, que nous avons identifié comme inhibiteur de la GTase de *S. cerevisiae*, s'est avéré être un inhibiteur agissant sur plusieurs GTases, notamment les enzymes testées lors de l'étude présente soit l'enzyme humaine ainsi que deux GTases virales. De plus, nos résultats suggèrent que ce composé se lie sur la protéine dans un site adjacent au site actif. Son mécanisme d'inhibition s'opère en reconnaissant la forme fermée de l'enzyme et en entravant la suite de la cascade enzymatique de l'enzyme. Selon cette hypothèse, le MPA ne serait pas spécifique pour la liaison d'un substrat précis, soit le GTP ou l'ARN. Afin de confirmer cette indépendance quant au substrat naturel des GTases, nous voulions évaluer son potentiel inhibiteur sur une enzyme réalisant une cascade enzymatique similaire au changement conformationnel des GTases. En ce sens, l'enzyme sélectionnée devait reconnaître des ligands différents des GTases afin de retirer de l'équation les paramètres associés au substrat d'ARN, l'oligonucléotide accepteur. Pour ce faire, au sein de la grande famille des nucléotidyltransférases se retrouve les ligases à ADN qui reconnaissent spécifiquement l'ADN et l'ATP. En utilisant l'enzyme A544R de PBCV-1, nous avons identifié que le composé possède le même patron d'inhibition sur cette enzyme. Ce résultat démontre que l'impact du MPA n'est pas dû par une compétition avec un substrat spécifique, mais bien par une liaison avec un site allostérique sur l'enzyme.

Le potentiel inhibiteur du MPA varie légèrement d'une enzyme à l'autre mais il est autant présent chez les GTases que chez les ligases à ADN. Considérant que le composé interagit sur un site allostérique distinct du site actif et que les membres de la famille des nucléotidyltransférases ne possèdent pas un fort taux d'homologie de ces membres à l'extérieur de six motifs délimitant le site actif, il faut déterminer comment le composé peut parvenir à inhiber les enzymes sélectionnées pour cette étude. Ceci s'explique par la conservation des acides aminés qui forme le site d'interaction du MPA pour les enzymes. En effet, tel que prédit par les essais d'arrimage moléculaire, le MPA interagit avec des acides aminés situés sur les deux domaines de l'enzyme. Lorsqu'on s'attarde au taux de conservation des acides aminés impliqués dans la coordination du MPA, on remarque que huit des acides aminés responsables d'interagir avec le composé sont présents de façon

constitutive chez les nucléotidyltransférases. En effet, des cinq acides aminés situés sur le domaine amino-terminal, quatre font parties intégrantes des motifs conservés tels que l'isoleucine situé sur le motif I, l'acide aspartique du motif IV, ainsi qu'une lysine et une acide aspartique située sur le motif V. D'autre part, le MPA interagit avec une arginine, un acide aspartique et une glutamine sur le domaine carboxyl-terminal qui font toutes parties du motif IV et qui sont conservés autant chez les ligases à ADN et les GTases.

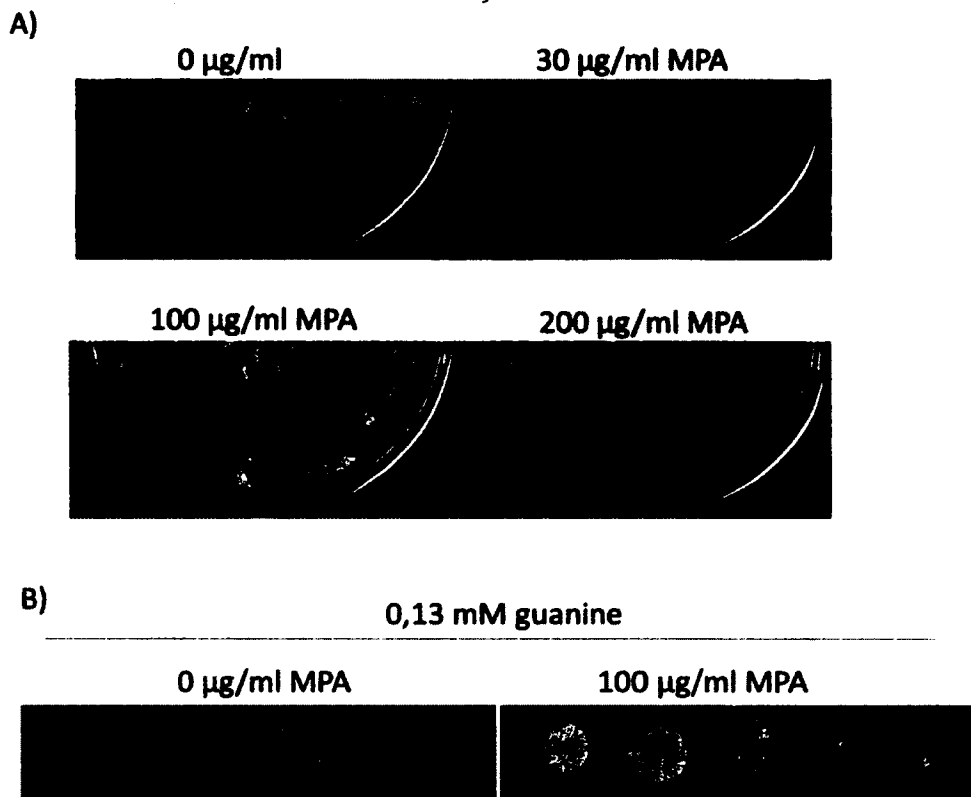
### ***3.5. Inhibition des GTases in vivo***

Afin de faire suite à la démonstration de l'inhibition des GTases par le MPA, et plus précisément par l'inhibition du transfert du nucléotide sur l'accepteur d'oligonucléotide, il fallait prouver cette inhibition dans un système *in vivo*. Toutefois, l'ajout de cette structure coiffe est essentielle pour la survie cellulaire ce qui complique l'inhibition de cette structure particulière à l'extrémité 5' des ARNm sans entraver la viabilité cellulaire. Néanmoins, une étude a démontré qu'il était possible de cartographier l'extrémité 5' par un essai d'extension d'amorce. En effet, le principe de cette expérience est lorsque l'ADN complémentaire d'un ARNm est synthétisé, la transcriptase inverse intègre la guanine de la coiffe au sein de l'ADN complémentaire. Cette expérience a été validée par plusieurs laboratoires tout en utilisant des mutants thermosensibles pour l'enzyme guanylyltransférase.

L'un des rôles essentiels de la structure coiffe est la protection des ARNm envers les exoribonucléases 5'-3'. Puisque notre objectif est de quantifier les ARNm coiffés versus ceux qui ne possèdent pas de structure coiffe, nous devons trouver un moyen de limiter la dégradation des ARN non coiffés. Pour ce faire, nous avons utilisé une souche de *S. cerevisiae* dont le gène *Xrn1* (exoribonucléase cytoplasmique 5'-3') a été délété tout en ne montrant aucun phénotype de déficit de croissance (Winzeler et al., 1999). Par contre, puisque l'homologue nucléaire de *Xrn1*, *Rat1* possède un déficit de croissance une fois

déléte, les essais cellulaires ont été réalisés avec une souche, donnée généreusement par le Dr Sherif Abou Elela, où le gène *Xrn1* uniquement était enlevé .

Préalablement à l'utilisation de l'extension d'amorce, nous devons vérifier plusieurs paramètres. Tout d'abord, il fallait identifier l'impact du MPA sur la croissance d'une population de *S. cerevisiae*. Les concentrations employées pour analyser l'impact du MPA sur l'activité guanylyltransférase *in vivo* se sont basées sur les études qui mesurent l'effet du MPA sur sa cible principale, soit l'inosine monophosphate déshydrogénase (IMPDH). Le MPA est un inhibiteur très puissant de l'IMPDH et c'est dans cette optique que nous avons utilisé des concentrations cinq fois plus élevées, soit de 100 mg/L à 500 mg/L pour vérifier son impact sur la GTase. Lorsque qu'une colonie de *S. cerevisiae* pousse en présence de MPA, la viabilité cellulaire est affectée partiellement par l'inhibition de l'IMPDH. Pour renverser le phénotype létal dû à l'inhibition de l'IMPDH, le milieu de culture peut être supplémenté en guanine. Selon plusieurs études, une concentration de 0.13 mM de guanine permet de rétablir la concentration intracellulaire de ce nucléotide nécessaire pour la croissance, tel que montré par la figure 17. Selon certaines études, l'ajout de guanine ne fait que partiellement restaurer l'inhibition cellulaire du MPA et laisse envisager que ce composé peut également cibler d'autres composantes cellulaires.



**Figure 17. La présence de l'inhibiteur MPA diminue la prolifération cellulaire de *S. cerevisiae*. A) L'effet d'addition de différentes concentrations de MPA (30, 100 et 300 µg/ml) sur la croissance cellulaire de *S. cerevisiae* BY4741 placé à 30°C. B) Les dilutions en série de culture de *S. cerevisiae* en présence de 0,13 mM de guanine mis en culture en absence (gauche) ou en présence (droite) de MPA.**

Avant de procéder aux essais cellulaires en présence d'inhibiteur de l'activité GTase, nous devons d'abord confirmer que l'essai d'extension d'amorce permettait bien de distinguer entre un ARN non coiffé et un ARN coiffé. Pour ce faire, l'extraction d'ARNm d'une culture de *S. cerevisiae*, a été mise en présence avec une enzyme recombinante responsable d'hydrolyser la structure coiffé avant de réaliser l'essai d'extension d'amorce. L'enzyme en question est la protéine virale D10 du virus de la vaccine puisque plusieurs études ont caractérisé son activité cinétique et de ses cofacteurs

*in vitro*. Une concentration constante d'ARN a été mise en présence de 1  $\mu$ M de D10 selon différents temps d'incubation. Les résultats obtenus démontrent clairement une diminution du ratio de l'ARN coiffé par rapport à l'ARN non coiffé. Après 60 minutes d'incubation avec la protéine de décoiffage, l'extension d'amorce du gène rapporteur SSA1 montre une diminution de l'ordre de 20%. Dans les essais de décoiffage, une très grande quantité d'ARNm est mise en incubation avec l'enzyme D10, ce qui peut expliquer que le ratio ne soit pas diminué davantage. Une fois ces travaux préliminaires réalisés, il est possible de poursuivre avec les essais cellulaires sur *S. cerevisiae*.

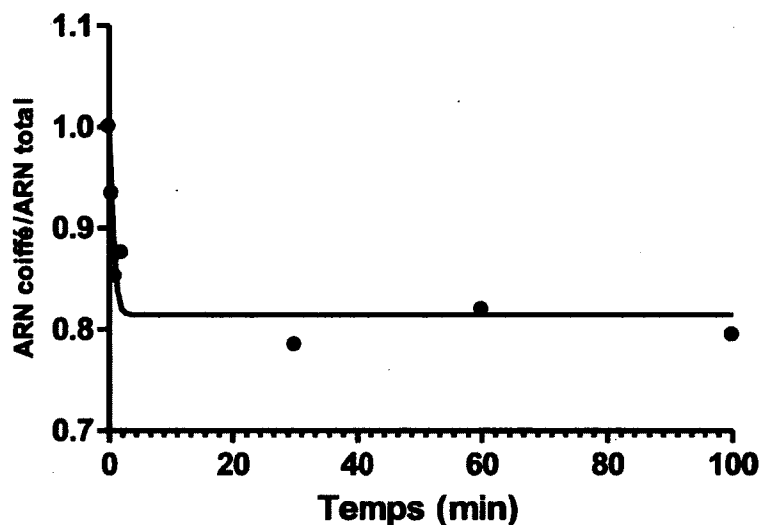


Figure 18. L'essai d'extension d'amorce sur des ARNm décoiffés en fonction du temps. Les ARNm extraits des cellules de *S. cerevisiae* ont été décoiffés en utilisant l'enzyme D10 du virus de la vaccine selon des temps définis. Ensuite, un oligonucléotide complémentaire à l'ARNm SSA1 a été hybridé à l'extraction des ARNm avant l'extension utilisant la transcriptase inverse. Le graphique illustre le ratio des ARN coiffés sur les ARN totaux.



## **4) Perspectives**

### **4.1. Confirmer le site d'interaction entre le MPA et la GTase**

L'utilisation de l'arrimage moléculaire dans cette étude nous a permis de prédire le site d'interaction du MPA sur la GTase. Toutefois, il sera intéressant de confirmer par une méthode expérimentale le site d'appariement ainsi que les acides aminés interagissant avec la molécule de MPA. Dans l'optique d'identifier la constante d'affinité du MPA pour l'enzyme, nous proposons d'utiliser plusieurs méthodes, soit l'utilisation de protéines mutantes et l'utilisation du microcalorimètre.

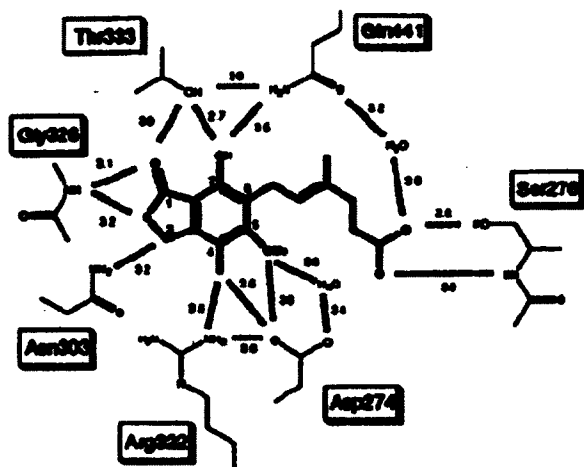
#### **4.1.1. Générer des protéines mutantes**

##### **4.1.1.1. Théorie**

La reconnaissance du substrat spécifique d'une protéine pour son ligand est souvent médiée par l'interaction entre quelques acides aminés bien précis. Les chaînes latérales des acides aminés délimitant le site actif permettent d'interagir avec différentes portions du ligand pour coordonner sa liaison avec la protéine. Plusieurs études ont pu confirmer les acides aminés essentiels de certaines interactions protéine-ligand en générant des protéines modifiées remplaçant de façon ponctuelle ces acides aminés. En effet, en ciblant les acides aminés non catalytiques qui possèdent une chaîne latérale préconisant un type d'interaction bien précis et en les modifiant par des acides aminés neutres, tels que l'alanine, il est possible d'abolir cette interaction et potentiellement de déstabiliser le complexe protéine-ligand. En générant des protéines recombinantes possédant des modifications des acides aminés entourant le site de liaison, il est possible de déterminer l'impact de chacune des interactions pour la stabilisation de l'inhibiteur.

#### **4.1.1.2. Procédure expérimentale**

Plusieurs études de caractérisation, notamment des études cristallographiques portant sur la liaison, du MPA à l'IMPDH ont permis d'obtenir des informations sur son site de liaison. Afin de confirmer le caractère essentiel de certains acides aminés pour la liaison du MPA, des études de mutagenèse dirigée ont ciblé la thréonine 333 pour la remplacer par une isoleucine et la glutamine 441 pour une alanine. Ces études ont démontré qu'elles diminuent le potentiel inhibiteur du MPA de 300 et 25 fois respectivement (Sintchak et al., 1996). Le cas de la mutation de la thréonine 333 a également été montré dans une étude chez un modèle murin qui développe un phénotype de résistance au MPA par un facteur de 10 000 (Lightfoot et Snyder, 1994). Les études cristallographiques démontrent plusieurs liaisons hydrogènes entre la thréonine et la glutamine en question avec le MPA. Plus précisément, l'atome d'oxygène sur le groupement aromatique du MPA est situé à proximité de la chaîne latérale de la thréonine. L'extrémité de la chaîne acide du MPA permet de former deux liaisons hydrogènes avec une sérine. Si cette dernière est remplacée par une alanine, l'inhibition du MPA sur l'IMPDH est diminuée par un facteur de sept selon l'inhibition trouvée sur l'enzyme de type sauvage (Sintchak et al., 1996). Par opposition, les essais de mutagenèse de la sérine 275, la serine 327 et la glutamine 368 pour une alanine ont très peu d'impact sur le potentiel inhibiteur par le MPA (Sintchak et al., 1996). En se portant sur les études cristallographiques, il était possible de proposer que le site d'interaction ne faisait pas intervenir ces acides aminés.



**Figure 19. Les acides aminés de l'IMPDH réagissant avec l'inhibiteur MPA. Cette illustration schématise les résultats obtenus par cristallographie de la protéine IMPDH (PDB: 1JR1). Les résidus importants sont illustrés sous la représentation de tube et les liaisons hydrogènes sont illustrées par des lignes foncées entre les résidus et les atomes de MPA. Les distances sont reportées en angströms (Sintchak et al., 1996).**

Nos résultats d'arrimage moléculaire nous ont permis d'identifier des interactions potentielles entre 9 acides aminés de la GTase et le MPA. Il reste toutefois à confirmer ces interactions de façon expérimentale. Pour ce faire, des protéines recombinantes possédant une ou plusieurs mutations ponctuelles seraient générées. La majorité des acides aminés interagissant avec le MPA sont des acides aminés chargés. De plus, six acides aminés prédits pour coordonner la liaison du MPA interagissent avec le groupement aromatique, qui a été démontré dans les études du complexe MPA-IMPDH, comme un groupement déterminant pour les interactions avec une protéine. Puisque les acides aminés prédits pour se lier avec le MPA interagissent avec ce dernier via des chaînes latérales chargées, des protéines au sein desquelles ces acides aminés ont été remplacés par une alanine, un acide aminé neutre, pourraient être synthétisées. Ainsi, les acides aminés qui sont susceptibles de contribuer le plus pour l'interaction avec le MPA sont l'asparagine 297, l'arginine 295, la lysine 234 et l'asparagine 244. Ces derniers sont à proximité, approximativement 3 Å, des groupements réactifs de l'inhibiteur et semblent interagir avec ceux-ci par des liens favorables. En générant des protéines modifiées qui

portent une ou plusieurs des modifications mentionnées plus haut, une diminution du potentiel inhibiteur du MPA devrait être perçue sur l'activité enzymatique de l'activité GTase.

#### **4.1.2. Validation par interaction en utilisant la microcalorimétrie**

##### **4.1.2.1. Théorie**

Le premier principe de la thermodynamique stipulant que l'énergie totale d'un système isolé reste constante peut être appliqué à l'étude des réactions biochimiques. Celui-ci est basé sur l'étude de la chaleur produite ou absorbée par une réaction biochimique. En fait, la réaction chimique forme un système thermodynamique pouvant donner lieu à un échange d'énergie avec le milieu extérieur. Le microcalorimètre (ITC pour « isothermal calorimetry ») est l'instrument qui permet de mesurer cet échange d'énergie sous forme de chaleur produite ou absorbée par une réaction. Cette méthode permet de caractériser la sommation des interactions moléculaires de façon très précise. Dans cet environnement contrôlé, la cellule d'échantillonnage contenant la protéine d'intérêt est comparée à la cellule de référence. L'appareil ajoute le ligand à la cellule d'échantillonnage en le titrant de façon contrôlée. Le microcalorimètre permet de déceler le changement de température induit par l'interaction du ligand avec la protéine.

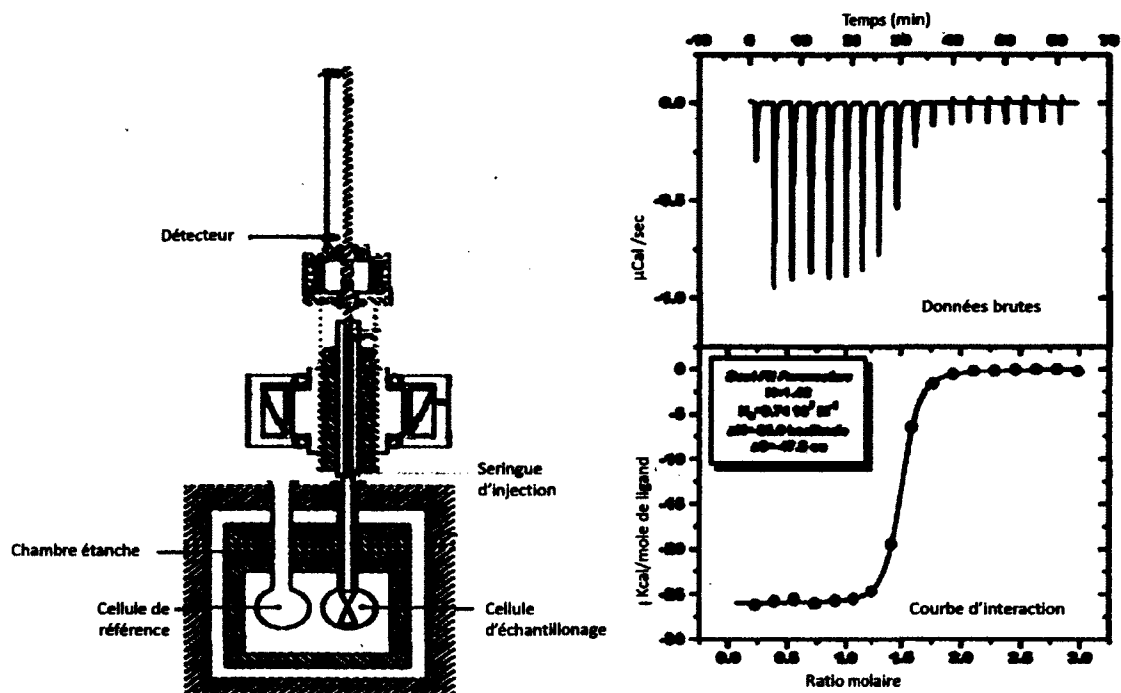


Figure 20. . Les différentes composantes de l'ITC. Dans la portion de gauche, la figure schématise un ITC en indiquant les éléments principaux, soit la cellule de référence et d'échantillonnage contenues dans la chambre isolée. Dans la portion de droite, la figure illustre les résultats attendus pour un essai d'interaction d'un ligand pour une protéine.

Un essai d'ITC permet d'obtenir un profil thermodynamique complet de l'interaction entre une protéine et son ligand, tout en calculant une constante d'affinité, le nombre de sites d'interaction avec le ligand, l'enthalpie, l'entropie et l'énergie libre. Pour ce faire, plusieurs répétitions d'un volume constant du ligand sont ajoutées à une protéine en solution. Les résultats générés sont représentés sous forme de thermogramme appelé association isothermique, où la chaleur totale (kcal par mole de ligand) est fonction de l'association du ligand et de la protéine. La forme de la courbe contient toutes les informations nécessaires pour caractériser l'association du ligand sur la protéine. Initialement, chaque molécule du ligand ajoutée à la solution se lie avec la protéine, mais au fur et à mesure que le ligand est ajouté, l'association diminue puisque les sites

d'interaction pour le ligand sont partiellement chargés par le ligand. Vers la fin de la réaction, l'ajout de ligand n'induit aucune modification à l'énergie du système puisque tous les sites d'interaction de la protéine sont déjà saturés.

#### **4.1.2.2. Procédure expérimentale**

Des études exhaustives de thermodynamiques sur la protéine GTase de PBCV-1 ont permis de déterminer les différents états intermédiaires lors de la cascade enzymatique. Erreur ! Source du renvoi introuvable.. Bien que l'ITC n'a pas été la méthode employée pour obtenir le profil de l'énergie libre de Gibbs, l'étendue des variations de ces valeurs énergétiques laisse envisager que ces réactions pourraient être perceptibles avec la sensibilité de l'appareil ITC. La qualité des résultats obtenus par cette méthode dépend énormément d'une concentration protéique très grande, voir autour de 20  $\mu\text{M}$ . Cette contrainte peut facilement être écarté en utilisant un système d'expression de la GTase en bactérie. En effet, cette protéine est particulièrement bien exprimée dans ce système puisqu'il est possible d'obtenir un volume de l'ordre du millilitre à 40  $\mu\text{M}$  de protéine. Pour réaliser cet essai, la solution contenant la protéine GTase de *S cerevisiae* sera ajoutée dans l'ITC dans la chambre adjacente à la cellule de référence. Le MPA est introduit dans l'injecteur, lequel sera ajouté à la solution protéique selon un apport constant de ligand selon le temps. S'il y a une interaction entre ces deux molécules et que cette dernière induit une réaction, soit accompagnée de libération ou d'absorption de chaleur, il est possible de dessiner une courbe de liaison dose-réponse. Ce faisant, il sera possible d'obtenir autant la constante de liaison, mais également des données au niveau de la thermodynamique du complexe enzyme-MPA.

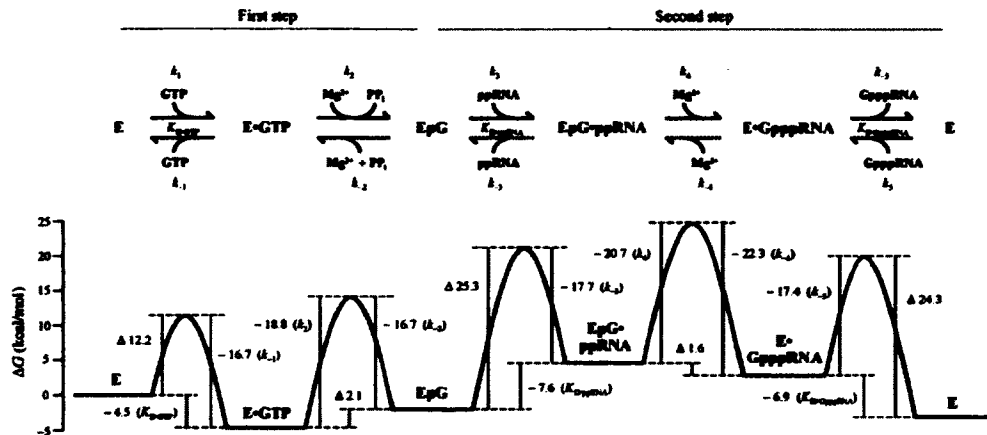


Figure 21. L'énergie de Gibbs pour les différentes étapes de l'activité enzymatique des GTases (Soulière et al., 2008). L'activité enzymatique de la GTase se décompose en réactions élémentaires de vitesse différente qui possèdent un niveau d'énergie bien précis.

## 4.2. Développement d'inhibiteur allostérique sélectif

Les résultats de la présente étude permettent de mettre à l'avant-plan un exemple allostérique de l'enzyme guanylyltransférase. Toutefois, tel que discuté, l'inhibition observée par le MPA s'observe autant chez les GTases virales, du virus PBCV-1 et du virus de la vaccine, chez la GTase de levure *S. cerevisiae* et chez la GTase humaine. La spécificité de l'inhibition est le point central de tout éventuel développement d'un inhibiteur de GTase. Différentes approches peuvent être exploitées pour parvenir à sélectionner des composés ciblant préférentiellement les enzymes pathogènes versus l'enzyme humaine. Tout d'abord, l'approche bio-informatique permettrait de discriminer entre un composé possédant un potentiel pharmacologique. De plus, en utilisant un système *in vivo* chez la levure *S. cerevisiae*, on pourrait confirmer les prédictions informatiques.

#### **4.2.1. Utilisation du criblage virtuel orienté vers les enzymes de pathogènes**

Tout criblage virtuel se décompose en trois étapes d'égale importance, soit la mise au point de la chimiothèque de départ, le criblage proprement dit et la sélection d'une liste de touches virtuelles. Plusieurs aspects peuvent être apportés à ces trois étapes afin d'augmenter le potentiel inhibiteur et la sélectivité des composés.

Comme le démontre cette étude, il est possible de synthétiser des inhibiteurs allostériques des GTases en entravant leur changement de conformation nécessaire pour leur activité enzymatique. Dans ce même ordre idée d'inhibiteur allostérique, toujours en se basant sur les composés de la chimiothèque PubChem, on pourrait retirer le filtre de sélection sélectionnant les composés qui possèdent un niveau de similarité à 80% avec le GTP. Étant donné que nous avons démontré qu'il est possible d'interagir avec les GTases de façon allostérique, le paramètre de similarité avec le GTP reconnu au site actif n'est pas pertinent pour identifier un inhibiteur allostérique. Ainsi, la taille de la chimiothèque initiale ne serait plus d'environ mille composés, mais plutôt de l'ordre des dizaines de milliers. De plus, la chimiothèque permettrait de recouper des composés variés tant au niveau de leur structure et de leur groupement réactif. La taille accrue de la banque de composés virtuels ouvre la porte à identifier des composés liant les GTases de façon beaucoup plus puissante.

Toujours afin d'optimiser la sélectivité des composés identifiés envers une GTase particulière, une attention particulière doit être apportée au site d'arrimage moléculaire. En effet, tel que discuté dans le chapitre précédent illustrant les résultats, le site d'essai d'arrimage moléculaire se situe à quelques angströms (Å) du site actif. En analysant le site d'interaction au niveau moléculaire du MPA pour la protéine, 8 des 9 acides aminés responsables de l'interaction sont hautement conservés chez les membres de la famille des nucléotidyltransférases. Pour éviter que les composés identifiés par le criblage virtuel soient reconnus de plusieurs GTases, le site sélectionné pour les essais d'arrimage ne devra pas inclure les acides aminés contenus dans les motifs conservés. Pour ce faire, la boîte de simulation potentielle pourrait être située à 12 Å au-dessus du site actif, au point



de rapprochement des domaines OB et du domaine NT. Ce site de simulation offre l'avantage d'être variable d'une GTase à l'autre.

Pour améliorer la sélectivité pour les enzymes GTases pathogènes par rapport à l'enzyme humaine, on pourrait appliquer le criblage virtuel autant pour la GTase cible que pour l'enzyme humaine. Ce faisant, il serait possible de comparer les résultats obtenus pour chacun des tableaux d'interactions prédits et d'identifier les composés qui discriminent les différentes enzymes. L'étape du criblage pourrait également être effectuée par plusieurs logiciels afin d'identifier des recouvrements entre les cibles prédites et d'éviter les faux positifs.

#### **4.2.2. Essai en cellules pour trouver un inhibiteur sélectif pour les enzymes de pathogènes**

##### **4.2.2.1. Théorie**

Bien que les essais d'arrimage moléculaire permettent de mettre à l'avant-plan des interactions potentielles, il est primordial de confirmer par un modèle *in vivo* robuste la spécificité des molécules potentielles. Un tel système a déjà été mis au point et validé en permettant d'identifier des composés inhibiteurs pour la MTase des levures. Pour ce faire, des souches haploïdes de *S. cerevisiae* sont modifiées par délétion des trois gènes essentiels responsables de la synthèse de la coiffe. Ces souches sont ensuite complémentées par des vecteurs permettant d'exprimer les enzymes de *S. cerevisiae*, du virus PBCV-1, du virus de la vaccine, et de l'humain. Les différentes souches générées sont dites isogéniques puisqu'uniquement les protéines synthétisant la structure coiffe diffèrent. Ainsi, en mettant ces souches en présence de différents composés, on peut identifier les composés qui entraînent un déficit de croissance des souches exprimant les enzymes pathogènes tout en ne causant pas de phénotype pour la souche avec l'enzyme humaine. Cet outil permet autant d'évaluer la biodisponibilité des composés que l'identification *in vivo* des composés affectant les enzymes de la coiffe.

#### **4.2.2.2. Procédure expérimentale**

Cette méthode procure une évaluation sommaire de la spécificité des composés et de leur toxicité cellulaire ce qui représente son avantage certain. Toutefois, pour réaliser cet essai en cellules, il est nécessaire d'obtenir les différentes souches isogéniques. Pour ce faire, il faut empêcher l'expression de la GTase endogène chez une souche de *S. cerevisiae*. Le gène peut être totalement inhibé par invalidation génique (knock-out) ou par diminution de son expression par l'interférence à l'ARN. Tout en réprimant l'expression du gène d'intérêt, il faut induire l'expression du gène de GTase, qui permet de récupérer l'enzyme manquante pour la synthèse de la coiffe, de l'organisme voulu. Ainsi, différentes souches isogéniques de *S. cerevisiae* exprimeraient spécifiquement la GTase humaine et certaines GTases de pathogènes. En débutant avec la molécule de MPA tel qu'inhibiteur de GTase, d'autres composés similaires seraient sélectionnés pour être testés sur les différentes souches isogéniques. Pour déceler la sensibilité des souches pour un composé, une méthode automatisée très sensible, utilisée dans le laboratoire du professeur Martin Toussaint à Sherbrooke, permettrait de récolter les données quant à la croissance des souches tout en maintenant en agitation à une température constante les souches (Toussaint et Conconi, 2006).

## **5) Conclusion**

Les différents résultats rapportés dans ce mémoire montrent que l'utilisation du criblage virtuel peut être utilisée pour identifier des interactions entre une protéine et un ligand. Cette étude fait la démonstration que le composé MPA peut inhiber l'activité enzymatique de la GTase de *S. cerevisiae* en interférant avec la deuxième étape de sa réaction, soit le transfert du nucléotide sur l'ARN. De plus, par l'utilisation de deux

méthodes distinctes, nous démontrons que l'action du composé sur l'enzyme n'est pas réalisée en compétition avec le substrat de l'enzyme, mais relève plutôt d'une inhibition allostérique. Le site de liaison du MPA pour la GTase de *S. cerevisiae* est prédit par arrimage moléculaire pour se situer entre les deux domaines de l'enzyme au dessus du site actif. Cette prédiction concorde avec le résultat selon lequel le MPA bloque la liaison de l'ARN sur l'enzyme. Par la suite, pour valider que l'inhibition observée avec le MPA ne soit pas due aux substrats des GTases, les mêmes essais d'inhibition sont portés sur des enzymes de la même famille que les GTases : les ligases à ADN. Les résultats montrent un même patron d'inhibition, soit une inhibition du transfert du nucléotide sur l'oligonucléotide. Finalement, des essais *in vivo* de *S. cerevisiae* résultent en une diminution des ARNm possédant la structure coiffe à leur extrémité 5'.

Ce mémoire présente la première démonstration d'une inhibition allostérique sur une GTase. De plus, il fait la preuve qu'il est possible de visualiser l'inhibition d'une enzyme de la synthèse de la coiffe, dans un système *in vivo*, générant des ARNm sans structure coiffe. Toutefois, avant d'envisager utiliser la GTase en tant que cible potentielle, plusieurs recherches devront augmenter la spécificité des composés vers une enzyme pathogène précise.

## **6) Remerciements**

Tout d'abord, je tiens à remercier mon directeur de recherche Martin Bisaillon, sans lequel ce travail n'aurait pas été possible. Il a su donner son appui tout au long de mes études graduées pour la poursuite de mes essais expérimentaux, apporter de judicieux conseils pour l'écriture d'article scientifique et lors de mes présentations scientifiques à un public spécialisé. Il a toujours porté une attention particulière à la formation aux étudiants ce qui s'est traduit par une qualité croissante de l'enseignement.

Par la suite, je souhaite remercier la Faculté de médecine et des sciences de la santé pour l'ensemble des activités organisées pour favoriser la participation des étudiants. Que ce soit à titre scientifique, tel que la Journée scientifique de la FMSS, ou que ce soit à titre académique pour la remise de bourse institutionnelle, la Faculté place l'étudiant au cœur de son programme.

Je tiens également à remercier les membres du groupe de recherche dans lequel ce travail s'est déroulé, soit Isabelle Bougie, Simon Despins, Nishal Issur, Marie Soulière, Frédéric Picard-Jean, Vincent Dutilly, Angélique Lanthier. Merci à tous pour toutes les discussions que nous avons eues, votre soutien et ces bons moments passés ensemble.

Merci Jean-Sébastien pour tes encouragements et ton soutien tout au long de mes études graduées. Je remercie énormément mes parents pour m'avoir toujours motivée dans la poursuite de mes études.

## 7) Références

Bajorath, J. (2002). Integration of virtual and high-throughput screening. *Nat Rev Drug Discov* 1, 882-894.

Baldi, A. (2010). Computational approaches for drug design and discovery: An overview, Vol 1.

Baron-Benhamou, J., Fortes, P., Inada, T., Preiss, T., et Hentze, M.W. (2003). The interaction of the cap-binding complex (CBC) with eIF4G is dispensable for translation in yeast. *RNA* 9, 654-662.

Bienroth, S., Keller, W., et Wahle, E. (1993). Assembly of a processive messenger RNA polyadenylation complex. *EMBO J* 12, 585-594.

Bisaillon, M., et Lemay, G. (1997). Viral and cellular enzymes involved in synthesis of mRNA cap structure. *Virology* 236, 1-7.

Bisaillon, M., et Shuman, S. (2001). Structure-function analysis of the active site tunnel of yeast RNA triphosphatase. *J Biol Chem* 276, 17261-17266.

Bleicher, K.H., Bohm, H.J., Muller, K., et Alanine, A.I. (2003). Hit and lead generation: beyond high-throughput screening. *Nat Rev Drug Discov* 2, 369-378.

Bonnefond, L., Giege, R., et Rudinger-Thirion, J. (2005). Evolution of the tRNA(Tyr)/TyrRS aminoacylation systems. *Biochimie* 87, 873-883.

Bougie, I., Parent, A., et Bisaillon, M. (2004). Thermodynamics of ligand binding by the yeast mRNA-capping enzyme reveals different modes of binding. *Biochem J* 384, 411-420.

Burford, N.T., Watson, J., Bertekap, R., et Alt, A. (2011). Strategies for the identification of allosteric modulators of G-protein-coupled receptors. *Biochem Pharmacol* 81, 691-702.

Carter, S.K. (2000). Clinical strategy for the development of angiogenesis inhibitors. *Oncologist 5 Suppl 1*, 51-54.

Changela, A., Ho, C.K., Martins, A., Shuman, S., et Mondragon, A. (2001). Structure and mechanism of the RNA triphosphatase component of mammalian mRNA capping enzyme. *EMBO J 20*, 2575-2586.

Chen, Y., Su, C., Ke, M., Jin, X., Xu, L., Zhang, Z., Wu, A., Sun, Y., Yang, Z., Tien, P., et al. (2011). Biochemical and Structural Insights into the Mechanisms of SARS Coronavirus RNA Ribose 2'-O-Methylation by nsp16/nsp10 Protein Complex. *PLoS Pathog 7*, e1002294.

Cochrane, A. (2009). How does the journey affect the message(RNA)? *RNA Biol 6*, 169-170.

Colgan, D.F., et Manley, J.L. (1997). Mechanism and regulation of mRNA polyadenylation. *Genes Dev 11*, 2755-2766.

Coller, J., et Parker, R. (2004). Eukaryotic mRNA decapping. *Annu Rev Biochem 73*, 861-890.

Cong, P., et Shuman, S. (1993). Covalent catalysis in nucleotidyl transfer. A KTDG motif essential for enzyme-GMP complex formation by mRNA capping enzyme is conserved at the active sites of RNA and DNA ligases. *J Biol Chem 268*, 7256-7260.

Cong, P., et Shuman, S. (1995). Mutational analysis of mRNA capping enzyme identifies amino acids involved in GTP binding, enzyme-guanylate formation, and GMP transfer to RNA. *Mol Cell Biol 15*, 6222-6231.

Cramer, P., Srebrow, A., Kadener, S., Werbajh, S., de la Mata, M., Melen, G., Nogues, G., et Kornblihtt, A.R. (2001). Coordination between transcription and pre-mRNA processing. *FEBS Lett 498*, 179-182.

Daffis, S., Szretter, K.J., Schriewer, J., Li, J., Youn, S., Errett, J., Lin, T.Y., Schneller, S., Zust, R., Dong, H., et al. (2010). 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature 468*, 452-456.

Das, K., Aramini, J.M., Ma, L.C., Krug, R.M., et Arnold, E. (2010). Structures of influenza A proteins and insights into antiviral drug targets. *Nat Struct Mol Biol 17*, 530-538.

De Clercq, E. (2004a). Antiviral drugs in current clinical use. *J Clin Virol 30*, 115-133.

- De Clercq, E. (2004b). Antivirals and antiviral strategies. *Nat Rev Microbiol* 2, 704-720.
- Despins, S., Issur, M., Bougie, I., et Bisaillon, M. (2010). Deciphering the molecular basis for nucleotide selection by the West Nile virus RNA helicase. *Nucleic Acids Res* 38, 5493-5506.
- Dias, S.M., Cerione, R.A., et Wilson, K.F. (2010). Unloading RNAs in the cytoplasm: an "importin" task. *Nucleus* 1, 139-143.
- Doherty, A.J., et Suh, S.W. (2000). Structural and mechanistic conservation in DNA ligases. *Nucleic Acids Res* 28, 4051-4058.
- Drews, J. (2000). Drug discovery: a historical perspective. *Science* 287, 1960-1964.
- El Omari, K., Ren, J., Bird, L.E., Bona, M.K., Klarmann, G., LeGrice, S.F., et Stammers, D.K. (2006). Molecular architecture and ligand recognition determinants for T4 RNA ligase. *J Biol Chem* 281, 1573-1579.
- Ellis, M. (2002). Invasive fungal infections: evolving challenges for diagnosis and therapeutics. *Mol Immunol* 38, 947-957.
- Fabrega, C., Hausmann, S., Shen, V., Shuman, S., et Lima, C.D. (2004). Structure and mechanism of mRNA cap (guanine-N7) methyltransferase. *Mol Cell* 13, 77-89.
- Fabrega, C., Shen, V., Shuman, S., et Lima, C.D. (2003). Structure of an mRNA capping enzyme bound to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Mol Cell* 11, 1549-1561.
- Feagin, J.E., Shaw, J.M., Simpson, L., et Stuart, K. (1988). Creation of AUG initiation codons by addition of uridines within cytochrome b transcripts of kinetoplastids. *Proc Natl Acad Sci U S A* 85, 539-543.
- Fechter, P., et Brownlee, G.G. (2005). Recognition of mRNA cap structures by viral and cellular proteins. *J Gen Virol* 86, 1239-1249.
- Ferreira, R.S., Simeonov, A., Jadhav, A., Eidam, O., Mott, B.T., Keiser, M.J., McKerrow, J.H., Maloney, D.J., Irwin, J.J., et Shoichet, B.K. (2010). Complementarity between a docking and a high-throughput screen in discovering new cruzain inhibitors. *J Med Chem* 53, 4891-4905.
- Fischer, P.M. (2009). Cap in hand: targeting eIF4E. *Cell Cycle* 8, 2535-2541.

Flaherty, S.M., Fortes, P., Izaurralde, E., Mattaj, I.W., et Gilmartin, G.M. (1997). Participation of the nuclear cap binding complex in pre-mRNA 3' processing. *Proc Natl Acad Sci U S A* *94*, 11893-11898.

Freeman, S., et Gardiner, J.M. (1996). Acyclic nucleosides as antiviral compounds. *Mol Biotechnol* *5*, 125-137.

Fresco, L.D., et Buratowski, S. (1994). Active site of the mRNA-capping enzyme guanylyltransferase from *Saccharomyces cerevisiae*: similarity to the nucleotidyl attachment motif of DNA and RNA ligases. *Proc Natl Acad Sci U S A* *91*, 6624-6628.

Fung, J., Lai, C.L., Seto, W.K., et Yuen, M.F. (2011). Nucleoside/nucleotide analogues in the treatment of chronic hepatitis B. *J Antimicrob Chemother.*

Furuichi, Y., et Shatkin, A.J. (2000). Viral and cellular mRNA capping: past and prospects. *Adv Virus Res* *55*, 135-184.

Garneau, N.L., Wilusz, J., et Wilusz, C.J. (2007). The highways and byways of mRNA decay. *Nat Rev Mol Cell Biol* *8*, 113-126.

Geller, R., Vignuzzi, M., Andino, R., et Frydman, J. (2007). Evolutionary constraints on chaperone-mediated folding provide an antiviral approach refractory to development of drug resistance. *Genes Dev* *21*, 195-205.

Ghannoum, M.A., et Rice, L.B. (1999). Antifungal agents: mode of action, mechanisms of resistance, and correlation of these mechanisms with bacterial resistance. *Clin Microbiol Rev* *12*, 501-517.

Goddard, J.P., et Reymond, J.L. (2004). Enzyme assays for high-throughput screening. *Curr Opin Biotechnol* *15*, 314-322.

Gray, N.K., Collier, J.M., Dickson, K.S., et Wickens, M. (2000). Multiple portions of poly(A)-binding protein stimulate translation in vivo. *EMBO J* *19*, 4723-4733.

Gu, M., et Lima, C.D. (2005). Processing the message: structural insights into capping and decapping mRNA. *Curr Opin Struct Biol* *15*, 99-106.

Hajduk, P.J., Huth, J.R., et Tse, C. (2005). Predicting protein druggability. *Drug Discov Today* *10*, 1675-1682.

Hakansson, K., Doherty, A.J., Shuman, S., et Wigley, D.B. (1997). X-ray crystallography reveals a large conformational change during guanyl transfer by mRNA capping enzymes. *Cell* *89*, 545-553.



Hakansson, K., et Wigley, D.B. (1998). Structure of a complex between a cap analogue and mRNA guanylyl transferase demonstrates the structural chemistry of RNA capping. *Proc Natl Acad Sci U S A* 95, 1505-1510.

Hausmann, S., Vivares, C.P., et Shuman, S. (2002). Characterization of the mRNA capping apparatus of the microsporidian parasite *Encephalitozoon cuniculi*. *J Biol Chem* 277, 96-103.

Helgstrand, E., Eriksson, B., Johansson, N.G., Lannero, B., Larsson, A., Misiorny, A., Noren, J.O., Sjoberg, B., Stenberg, K., Stening, G., et al. (1978). Trisodium phosphonoformate, a new antiviral compound. *Science* 201, 819-821.

Hertzberg, R.P., et Pope, A.J. (2000). High-throughput screening: new technology for the 21st century. *Curr Opin Chem Biol* 4, 445-451.

Hirose, Y., et Manley, J.L. (2000). RNA polymerase II and the integration of nuclear events. *Genes Dev* 14, 1415-1429.

Hirose, Y., Tacke, R., et Manley, J.L. (1999). Phosphorylated RNA polymerase II stimulates pre-mRNA splicing. *Genes Dev* 13, 1234-1239.

Ho, C.K., Pei, Y., et Shuman, S. (1998a). Yeast and viral RNA 5' triphosphatases comprise a new nucleoside triphosphatase family. *J Biol Chem* 273, 34151-34156.

Ho, C.K., Sriskanda, V., McCracken, S., Bentley, D., Schwer, B., et Shuman, S. (1998b). The guanylyltransferase domain of mammalian mRNA capping enzyme binds to the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J Biol Chem* 273, 9577-9585.

Hocine, S., Singer, R.H., et Grunwald, D. (2010). RNA processing and export. *Cold Spring Harb Perspect Biol* 2, a000752.

Issur, M., Despins, S., Bougie, I., et Bisaillon, M. (2009a). Nucleotide analogs and molecular modeling studies reveal key interactions involved in substrate recognition by the yeast RNA triphosphatase. *Nucleic Acids Res* 37, 3714-3722.

Issur, M., Geiss, B.J., Bougie, I., Picard-Jean, F., Despins, S., Mayette, J., Hobdey, S.E., et Bisaillon, M. (2009b). The flavivirus NS5 protein is a true RNA guanylyltransferase that catalyzes a two-step reaction to form the RNA cap structure. *RNA* 15, 2340-2350.

Issur, M., Picard-Jean, F., et Bisaillon, M. (2011). The RNA capping machinery as an anti-infective target. *Wiley Interdiscip Rev RNA* 2, 184-192.

- Joshi, P., et Stoddart, C.A. (2011). Impaired infectivity of ritonavir-resistant HIV is rescued by heat shock protein 90AB1. *J Biol Chem* 286, 24581-24592.
- Keseru, G.M., et Makara, G.M. (2006). Hit discovery and hit-to-lead approaches. *Drug Discov Today* 11, 741-748.
- Kitchen, D.B., Decornez, H., Furr, J.R., et Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov* 3, 935-949.
- Klebe, G. (2006). Virtual ligand screening: strategies, perspectives and limitations. *Drug Discov Today* 11, 580-594.
- Knoop, V. (2011). When you can't trust the DNA: RNA editing changes transcript sequences. *Cell Mol Life Sci* 68, 567-586.
- Kuhn, U., et Wahle, E. (2004). Structure and function of poly(A) binding proteins. *Biochim Biophys Acta* 1678, 67-84.
- Kumar, S., et Nussinov, R. (2002). Close-range electrostatic interactions in proteins. *ChemBiochem* 3, 604-617.
- Levy, M., et Miller, S.L. (1998). The stability of the RNA bases: implications for the origin of life. *Proc Natl Acad Sci U S A* 95, 7933-7938.
- Lewis, J.D., et Izaurralde, E. (1997). The role of the cap structure in RNA processing and nuclear export. *Eur J Biochem* 247, 461-469.
- Lightfoot, T., et Snyder, F.F. (1994). Gene amplification and dual point mutations of mouse IMP dehydrogenase associated with cellular resistance to mycophenolic acid. *Biochim Biophys Acta* 1217, 156-162.
- Lockless, S.W., Cheng, H.T., Hodel, A.E., Quioco, F.A., et Gershon, P.D. (1998). Recognition of capped RNA substrates by VP39, the vaccinia virus-encoded mRNA cap-specific 2'-O-methyltransferase. *Biochemistry* 37, 8564-8574.
- Maas, S., et Rich, A. (2000). Changing genetic information through RNA editing. *Bioessays* 22, 790-802.
- MacCallum, D.M., Coste, A., Ischer, F., Jacobsen, M.D., Odds, F.C., et Sanglard, D. (2010). Genetic dissection of azole resistance mechanisms in *Candida albicans* and their validation in a mouse model of disseminated infection. *Antimicrob Agents Chemother* 54, 1476-1483.

McKendrick, L., Thompson, E., Ferreira, J., Morley, S.J., et Lewis, J.D. (2001). Interaction of eukaryotic translation initiation factor 4G with the nuclear cap-binding complex provides a link between nuclear and cytoplasmic functions of the m(7) guanosine cap. *Mol Cell Biol* 21, 3632-3641.

Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., et Olson, A.J. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry* 19, 1639-1662.

Morschhauser, J. (2010). Regulation of multidrug resistance in pathogenic fungi. *Fungal Genet Biol* 47, 94-106.

Odell, M., Sriskanda, V., Shuman, S., et Nikolov, D.B. (2000). Crystal structure of eukaryotic DNA ligase-adenylate illuminates the mechanism of nick sensing and strand joining. *Mol Cell* 6, 1183-1193.

Ogino, T., et Banerjee, A.K. (2011). An unconventional pathway of mRNA cap formation by vesiculoviruses. *Virus Res* 162, 100-109.

Parker, R., et Song, H. (2004). The enzymes and control of eukaryotic mRNA turnover. *Nat Struct Mol Biol* 11, 121-127.

Parmley, J.L., Urrutia, A.O., Potrzebowski, L., Kaessmann, H., et Hurst, L.D. (2007). Splicing and the evolution of proteins in mammals. *PLoS Biol* 5, e14.

Pascal, J.M. (2008). DNA and RNA ligases: structural variations and shared mechanisms. *Curr Opin Struct Biol* 18, 96-105.

Pencheva, T., Soumana, O.S., Pajeva, I., et Miteva, M.A. (2010). Post-docking virtual screening of diverse binding pockets: comparative study using DOCK, AMMOS, X-Score and FRED scoring functions. *Eur J Med Chem* 45, 2622-2628.

Perea, S., et Patterson, T.F. (2002). Antifungal resistance in pathogenic fungi. *Clin Infect Dis* 35, 1073-1080.

Pommier, Y., Johnson, A.A., et Marchand, C. (2005). Integrase inhibitors to treat HIV/AIDS. *Nat Rev Drug Discov* 4, 236-248.

Proudfoot, N., et O'Sullivan, J. (2002). Polyadenylation: a tail of two complexes. *Curr Biol* 12, R855-857.

Proudfoot, N.J., Furger, A., et Dye, M.J. (2002). Integrating mRNA processing with transcription. *Cell* 108, 501-512.

Pullirsch, D., et Jantsch, M.F. (2010). Proteome diversification by adenosine to inosine RNA editing. *RNA Biol* 7, 205-212.

Qiao, J., Liu, W., et Li, R. (2008). Antifungal resistance mechanisms of *Aspergillus*. *Nihon Ishinkin Gakkai Zasshi* 49, 157-163.

Ramachary, D.B., et Vijayendar Reddy, Y. (2010). A general approach to chiral building blocks via direct amino acid-catalyzed cascade three-component reductive alkylations: formal total synthesis of HIV-1 protease inhibitors, antibiotic agglomerins, brefeldin A, and (R)-gamma-hexanolide. *J Org Chem* 75, 74-85.

Rhoads, R.E. (2009). eIF4E: new family members, new binding partners, new roles. *J Biol Chem* 284, 16711-16715.

Ross, J. (1996). Control of messenger RNA stability in higher eukaryotes. *Trends Genet* 12, 171-175.

Rousseau, D.L., Li, D., Couture, M., et Yeh, S.R. (2005). Ligand-protein interactions in nitric oxide synthase. *J Inorg Biochem* 99, 306-323.

Sawaya, R., et Shuman, S. (2003). Mutational analysis of the guanylyltransferase component of Mammalian mRNA capping enzyme. *Biochemistry* 42, 8240-8249.

Schames, J.R., Henchman, R.H., Siegel, J.S., Sotriffer, C.A., Ni, H., et McCammon, J.A. (2004). Discovery of a novel binding trench in HIV integrase. *J Med Chem* 47, 1879-1881.

Schoenberg, D.R., et Maquat, L.E. (2009). Re-capping the message. *Trends Biochem Sci* 34, 435-442.

Schramm, V.L. (2005). Enzymatic transition states: thermodynamics, dynamics and analogue design. *Arch Biochem Biophys* 433, 13-26.

Schwer, B., Lehman, K., Saha, N., et Shuman, S. (2001). Characterization of the mRNA capping apparatus of *Candida albicans*. *J Biol Chem* 276, 1857-1864.

Shikanai, T. (2006). RNA editing in plant organelles: machinery, physiological function and evolution. *Cell Mol Life Sci* 63, 698-708.

Shoichet, B.K. (2004). Virtual screening of chemical libraries. *Nature* 432, 862-865.

Shuman, S. (1995). Capping enzyme in eukaryotic mRNA synthesis. *Prog Nucleic Acid Res Mol Biol* 50, 101-129.

Shuman, S. (2001). The mRNA capping apparatus as drug target and guide to eukaryotic phylogeny. *Cold Spring Harb Symp Quant Biol* 66, 301-312.

Shuman, S. (2002). What messenger RNA capping tells us about eukaryotic evolution. *Nat Rev Mol Cell Biol* 3, 619-625.

Shuman, S., et Lima, C.D. (2004). The polynucleotide ligase and RNA capping enzyme superfamily of covalent nucleotidyltransferases. *Curr Opin Struct Biol* 14, 757-764.

Shuman, S., Liu, Y., et Schwer, B. (1994). Covalent catalysis in nucleotidyl transfer reactions: essential motifs in *Saccharomyces cerevisiae* RNA capping enzyme are conserved in *Schizosaccharomyces pombe* and viral capping enzymes and among polynucleotide ligases. *Proc Natl Acad Sci U S A* 91, 12046-12050.

Shuman, S., et Ru, X.M. (1995). Mutational analysis of vaccinia DNA ligase defines residues essential for covalent catalysis. *Virology* 211, 73-83.

Shuman, S., et Schwer, B. (1995). RNA capping enzyme and DNA ligase: a superfamily of covalent nucleotidyl transferases. *Mol Microbiol* 17, 405-410.

Sintchak, M.D., Fleming, M.A., Futer, O., Raybuck, S.A., Chambers, S.P., Caron, P.R., Murcko, M.A., et Wilson, K.P. (1996). Structure and mechanism of inosine monophosphate dehydrogenase in complex with the immunosuppressant mycophenolic acid. *Cell* 85, 921-930.

Sonenberg, N. (2008). eIF4E, the mRNA cap-binding protein: from basic discovery to translational research. *Biochem Cell Biol* 86, 178-183.

Souliere, M.F., Perreault, J.P., et Bisailon, M. (2008). Kinetic and thermodynamic characterization of the RNA guanylyltransferase reaction. *Biochemistry* 47, 3863-3874.

Sperandio, O., Miteva, M.A., Segers, K., Nicolaes, G.A., et Villoutreix, B.O. (2008). Screening Outside the Catalytic Site: Inhibition of Macromolecular Inter-actions Through Structure-Based Virtual Ligand Screening Experiments. *Open Biochem J* 2, 29-37.

Spies, U., et Muller, H. (1990). Demonstration of enzyme activities required for cap structure formation in infectious bursal disease virus, a member of the birnavirus group. *J Gen Virol* 71 (Pt 4), 977-981.

Sriskanda, V., et Shuman, S. (1998). Mutational analysis of Chlorella virus DNA ligase: catalytic roles of domain I and motif VI. *Nucleic Acids Res* 26, 4618-4625.

Sriskanda, V., et Shuman, S. (2002a). Role of nucleotidyl transferase motif V in strand joining by chlorella virus DNA ligase. *J Biol Chem* 277, 9661-9667.

Sriskanda, V., et Shuman, S. (2002b). Role of nucleotidyltransferase motifs I, III and IV in the catalysis of phosphodiester bond formation by Chlorella virus DNA ligase. *Nucleic Acids Res* 30, 903-911.

Stuart, K., et Panigrahi, A.K. (2002). RNA editing: complexity and complications. *Mol Microbiol* 45, 591-596.

Sun, X., Barlow, E.A., Ma, S., Hagemeyer, S.R., Duellman, S.J., Burgess, R.R., Tellam, J., Khanna, R., et Kenney, S.C. (2010). Hsp90 inhibitors block outgrowth of EBV-infected malignant cells in vitro and in vivo through an EBNA1-dependent mechanism. *Proc Natl Acad Sci U S A* 107, 3146-3151.

Sundberg, S.A. (2000). High-throughput and ultra-high-throughput screening: solution- and cell-based approaches. *Curr Opin Biotechnol* 11, 47-53.

Swift, R.V., Durrant, J., Amaro, R.E., et McCammon, J.A. (2009). Toward understanding the conformational dynamics of RNA ligation. *Biochemistry* 48, 709-719.

Swift, R.V., et McCammon, J.A. (2008). Catalytically requisite conformational dynamics in the mRNA-capping enzyme probed by targeted molecular dynamics. *Biochemistry* 47, 4102-4111.

Takagi, T., Moore, C.R., Diehn, F., et Buratowski, S. (1997). An RNA 5'-triphosphatase related to the protein tyrosine phosphatases. *Cell* 89, 867-873.

Taylor, R.D., Jewsbury, P.J., et Essex, J.W. (2002). A review of protein-small molecule docking methods. *J Comput Aided Mol Des* 16, 151-166.

Terstappen, G.C., et Reggiani, A. (2001). In silico research in drug discovery. *Trends Pharmacol Sci* 22, 23-26.

Topisirovic, I., Svitkin, Y.V., Sonenberg, N., et Shatkin, A.J. (2011). Cap and cap-binding proteins in the control of gene expression. *Wiley Interdiscip Rev RNA* 2, 277-298.

Toussaint, M., et Conconi, A. (2006). High-throughput and sensitive assay to measure yeast cell growth: a bench protocol for testing genotoxic agents. *Nat Protoc* 1, 1922-1928.

- Tramontano, E. (2008). The exploding field of the HCV polymerase non-nucleoside inhibitors: summary of a first generation compounds. *Mini Rev Med Chem* 8, 1298-1310.
- Vanden Bossche, H. (1997). Mechanisms of antifungal resistance. *Rev Iberoam Micol* 14, 44-49.
- Varani, G. (1997). A cap for all occasions. *Structure* 5, 855-858.
- von Ahsen, O., et Bomer, U. (2005). High-throughput screening for kinase inhibitors. *ChemBiochem* 6, 481-490.
- Wagstaff, A.J., et Bryson, H.M. (1994). Foscarnet. A reappraisal of its antiviral activity, pharmacokinetic properties and therapeutic use in immunocompromised patients with viral infections. *Drugs* 48, 199-226.
- Wang, L.K., et Shuman, S. (2005). Structure-function analysis of yeast tRNA ligase. *RNA* 11, 966-975.
- Wang, S.P., Deng, L., Ho, C.K., et Shuman, S. (1997). Phylogeny of mRNA capping enzymes. *Proc Natl Acad Sci U S A* 94, 9573-9578.
- White, R.E. (2000). High-throughput screening in drug metabolism and pharmacokinetic support of drug discovery. *Annu Rev Pharmacol Toxicol* 40, 133-157.
- Whitty, A., et Kumaravel, G. (2006). Between a rock and a hard place? *Nat Chem Biol* 2, 112-118.
- Winzeler, E.A., Shoemaker, D.D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J.D., Bussey, H., *et al.* (1999). Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* 285, 901-906.
- Wood, Z.A., Sabatini, R.S., et Hajduk, S.L. (2004). RNA ligase; picking up the pieces. *Mol Cell* 13, 455-456.
- Wutzler, P., et Thust, R. (2001). Genetic risks of antiviral nucleoside analogues--a survey. *Antiviral Res* 49, 55-74.
- Yaari, S., Koslowsky, B., Wolf, D., Chajek-Shaul, T., et Hershcovici, T. (2010). CMV-related thrombocytopenia treated with foscarnet: a case series and review of the literature. *Platelets* 21, 490-495.

Yang, S.H., Khadka, D.B., Cho, S.H., Ju, H.K., Lee, K.Y., Han, H.J., Lee, K.T., et Cho, W.J. (2011). Virtual screening and synthesis of quinazolines as novel JAK2 inhibitors. *Bioorg Med Chem* 19, 968-977.