

Université du Québec  
Institut National de la Recherche Scientifique  
Institut Armand-Frappier

## **Conservation et modulation de secteurs dynamiques fonctionnels entre ribonucléases homologues**

Par  
Donald Gagné

Thèse présentée pour l'obtention du grade de  
du grade de Philosophiæ Doctor (Ph.D.) en biologie

### **Jury d'évaluation**

Président du jury et examinateur interne	Pr Michel Sylvestre INRS – Institut Armand-Frappier
Examineur externe	Pr Pierre Lavigne Département de biochimie Université de Sherbrooke
Examineur externe	Pr Manon Couture Département de biochimie, de microbiologie et de bio-informatique Université Laval
Directeur de recherche	Pr Nicolas Doucet INRS – Institut Armand-Frappier





## REMERCIEMENTS

Je tiens tout d'abord à remercier mon directeur de recherche, Pr Nicolas Doucet, qui a cru en moi. Il m'a offert l'expertise, les équipements et la liberté d'expression dans la réalisation de mes projets de recherches. Il m'a transmis sa passion pour les sciences et son savoir. Je le remercie pour toutes les occasions où je suis rentré à l'improviste dans son bureau pour lui poser mes nombreuses questions, de même que pour toutes les discussions passionnées que l'on a eues, sans oublier les nombreux congrès locaux et internationaux auxquels j'ai eu la chance de participer, et qui m'ont permis de voyager à travers l'Amérique. C'est aussi grâce à lui que j'ai eu la chance de participer à un projet international qui m'a donné l'occasion de passer 3 mois à Chicago. Ce projet m'a ouvert la porte à la cristallographie, mais m'a aussi donné le goût de poursuivre des études postdoctorales.

J'aimerais sincèrement remercier mes collègues du laboratoire Doucet pour les discussions de groupe, tant scientifiques que philosophiques. Un merci tout particulier à Stéphane Beauchamp, David Bernard, Guillaume Brault Ph.D., Hélène Carlettini, Laurie-Anne Charest, Roger Dubuc, Carlos E. Dulcey, Philippe Egesborg, Benjamin Folch Ph.D., Jacinthe Gagnon Ph.D., Yossef Lopez de los Santos Ph.D., Chitra Narayanan Ph.D., Nhung Nguyen Thi avec lesquels j'ai partagé une partie de vie. Sans oublier les étudiants stagiaires (Glenda Alvarez, Étienne Dubé, Aïssa Muamba, Bruno Ouimet, Jean François Pelletier-Paquette, Félix Pigeon, Carol-Anne Villeneuve, Jordan Volpato Ph.D.), qui malgré leur passage plus bref, ont su me garder jeune. J'aimerais aussi remercier nos collaborateurs (Pr David Chatenet, Pr Jean-François Couture, Pr Marc A. Gauthier, Sophie Gobeil, Andrée-Anne Grosset, Marilyne Labrie, Pr Joelle Pelletier, Pr Yves St-Pierre et Maria C. Vladioiu), avec qui j'ai eu beaucoup de plaisir à travailler et avec qui j'espère avoir l'opportunité de retravailler dans un futur rapproché. J'aimerais saluer tout particulièrement Pr. Miljan Simonović, qui m'a accueilli dans son laboratoire, et Rachel L. French, qui a eu la patience de m'apprendre les rudiments de la cristallographie et m'a donné le goût de continuer. Enfin, un gros merci à tout le monde de l'INRS – Institut Armand-Frappier, professeurs, étudiants et personnel de soutien, que j'ai côtoyés pendant ces 5 merveilleuses années. Ce fut une belle expérience personnelle et professionnelle que je n'oublierai jamais.

Bien sûr, cette thèse n'aurait pu voir le jour sans le soutien des organismes subventionnaires tels que le CRSNG (et la bourse Alexander-Graham-Bell), FRQNT, PROTEO, GRASP et la Fondation Armand-Frappier, qui m'ont permis de poursuivre mes études à temps plein.

Enfin, je tiens à signaler la contribution émotionnelle et psychologique offerte par tous mes amis, parents, et membres de ma famille, mais tout particulièrement celui de ma conjointe, Guylaine Drapeau, et de mes 2 enfants, Christina et Philippe. Ils ont su contribuer, à leur manière, à mon développement personnel et à mon cheminement scientifique.

## RÉSUMÉ

La flexibilité conformationnelle entre différents ensembles structuraux est une composante essentielle à la fonction des enzymes. Bien qu'il soit connu qu'un vaste éventail de mouvements dynamiques favorise des événements fonctionnels chez les protéines sur plusieurs échelles de temps, il n'est pas encore démontré que des enzymes structurellement et fonctionnellement homologues emploient les mêmes mouvements concertés de résidus pour remplir leur fonction catalytique. Dans la présente étude, nous émettons l'hypothèse que des réseaux de résidus contigus et flexibles dont les mouvements se produisent à l'échelle de temps de la milliseconde (échelle de temps de la catalyse enzymatique) ont évolué dans le but de promouvoir et/ou de préserver une activité catalytique optimale chez des enzymes homologues. Pour ce faire, nous utilisons une combinaison d'expériences de dispersion de relaxation RMN ( $^{15}\text{N}$ -CPMG), d'analyse Model-Free, de titrage RMN, et d'expériences fonctionnelles pour capturer et comparer le rôle de la flexibilité conformationnelle entre des homologues structuraux de la famille des ribonucléases pancréatiques humaines (RNases). En plus de conserver les mêmes résidus catalytiques et la même structure tridimensionnelle, la plupart des membres étudiés montrent une conservation de résidus dynamiques à l'échelle de temps de la  $\mu\text{s}$ -ms, suggérant que la flexibilité conformationnelle peut être conservée parmi des protéines structurellement analogues qui affichent néanmoins une faible identité de séquence. Plutôt que d'afficher un portrait global, nos résultats révèlent que ces résidus dynamiques sont localisés dans des régions contiguës spécifiques à chaque protéine. En utilisant la mutagenèse dirigée, nous montrons qu'il est possible de moduler les régions de résidus dynamiques d'une enzyme et d'en affecter la fonction, soutenant l'hypothèse que la flexibilité conformationnelle est en partie nécessaire pour la fonction catalytique chez des enzymes homologues. En résumé, nous démontrons que la RMN, couplée à la mutagenèse dirigée ainsi qu'à des essais fonctionnels, s'avère un outil de première importance en ingénierie des protéines.

# TABLE DES MATIÈRES

<b>1</b>	<b>Introduction</b> .....	<b>1</b>
1.1	<b>Mise en contexte</b> .....	<b>1</b>
1.2	<b>Les protéines</b> .....	<b>1</b>
1.3	<b>La cinétique enzymatique</b> .....	<b>3</b>
1.4	<b>Les mouvements atomiques</b> .....	<b>5</b>
1.4.1	L'allostérie .....	7
1.4.2	Principales méthodes d'analyse de la dynamique moléculaire .....	8
1.5	<b>La dynamique des protéines</b> .....	<b>10</b>
1.5.1	Principe de la RMN .....	10
1.5.2	La relaxation des spins.....	12
1.5.3	Relaxation longitudinale $R_1$ .....	13
1.5.4	Relaxation transversale $R_2$ .....	14
1.5.5	Effet Overhauser nucléaire <i>NOE</i> .....	17
1.5.6	Les expériences RMN .....	17
1.6	<b>Hypothèses et objectifs</b> .....	<b>26</b>
1.6.1	Hypothèses .....	26
1.6.2	Objectifs.....	27
1.7	<b>Les ribonucléases pancréatiques : un système modèle</b> .....	<b>28</b>
1.7.1	La RNase A .....	28
1.7.2	La famille des RNases pancréatiques humaines .....	33
1.8	<b>Autres techniques</b> .....	<b>37</b>
1.8.1	Production d'enzymes recombinantes .....	37
1.8.2	La mutagenèse dirigée .....	40
1.8.3	Titration calorimétrique.....	41
1.8.4	Analyse spectrale par dichroïsme circulaire .....	43
<b>2</b>	<b>Article No. 1: Importance des fluctuations conformationnelles globales et locales sur la structure et la fonction des membres de la superfamille de la RNase A</b> .....	<b>45</b>
2.1	<b>Présentation de l'article scientifique</b> .....	<b>45</b>
2.2	<b>Contribution des auteurs</b> .....	<b>45</b>

2.3	<b>Résumé</b> .....	<b>46</b>
2.4	<b>Article scientifique</b> .....	<b>47</b>
2.5	<b>Discussion</b> .....	<b>59</b>
2.6	<b>Résultats supplémentaires</b> .....	<b>60</b>
<b>3</b>	<b>Article No. 2: Altération de l'activité enzymation suite à la perturbation de la dynamique conformationnelle d'une boucle du site actif</b> .....	<b>61</b>
3.1	<b>Présentation de l'article scientifique</b> .....	<b>61</b>
3.2	<b>Contribution des auteurs</b> .....	<b>62</b>
3.3	<b>Résumé</b> .....	<b>63</b>
3.4	<b>Article scientifique</b> .....	<b>64</b>
3.5	<b>Discussion</b> .....	<b>76</b>
3.6	<b>Résultats supplémentaires</b> .....	<b>77</b>
3.6.1	Comparaison des structures native et mutée .....	77
3.6.2	Analyse structurale du mutant RNase A <sub>A109G</sub> .....	78
<b>4</b>	<b>Article No. 3: Conservation de régions flexibles entre homologues structuraux et fonctionnels</b> .....	<b>81</b>
4.1	<b>Présentation de l'article scientifique</b> .....	<b>81</b>
4.2	<b>Contribution des auteurs</b> .....	<b>82</b>
4.3	<b>Résumé</b> .....	<b>82</b>
4.4	<b>Article scientifique</b> .....	<b>84</b>
4.5	<b>Discussion</b> .....	<b>104</b>
4.6	<b>Résultats supplémentaires</b> .....	<b>105</b>
4.6.1	Caractérisation biophysique de la RNase 4 humaine .....	105
4.6.2	Comparaison biophysique d'EDN et d'ECP.....	113
<b>5</b>	<b>Article No. 4: Réseaux de corrélations à longue distance de déplacements chimiques concertés suite à la liaison de ligands à l'angiogénine humaine ...</b>	<b>123</b>
5.1	<b>Présentation de l'article scientifique</b> .....	<b>123</b>
5.2	<b>Contribution des auteurs</b> .....	<b>124</b>
5.3	<b>Résumé</b> .....	<b>124</b>
5.4	<b>Article scientifique</b> .....	<b>126</b>
5.5	<b>Discussion</b> .....	<b>135</b>
<b>6</b>	<b>Conclusions et Perspectives</b> .....	<b>137</b>

<b>7</b>	<b>Références .....</b>	<b>139</b>
----------	-------------------------	------------

## LISTE DES TABLEAUX

Tableau 1.1	Interactions entre les sous-sites de la RNase A et l'ARN	31
Tableau 1.2	Résumé des paramètres cinétiques pour la coupure d'ARN sur l'enzyme native de la RNase A.	32

## LISTE DES FIGURES

Figure 1.1	Structure d'une protéine.	2
Figure 1.2	Effet de la concentration du substrat sur la vitesse de la réaction catalytique	4
Figure 1.3	Cycle thermodynamique du mécanisme de reconnaissance moléculaire	6
Figure 1.4	Échelles de temps d'évènements dynamiques	7
Figure 1.5	Interconnectivité entre la structure de la protéine, sa dynamique et sa fonction	8
Figure 1.6	Techniques d'analyse d'évènements dynamiques en fonction de l'échelle de temps et du type de mouvements	10
Figure 1.7	Orientation du moment magnétique d'un atome avec $I = \frac{1}{2}$ dans un champ magnétique $\mathbf{B}_0$	11
Figure 1.8	Description du signal RMN	13
Figure 1.9	Approche expérimentale pour l'enregistrement du $R_1$ (inversion recovery) et du $R_2$ (spin echo) dans un champ magnétique $\mathbf{B}$	14
Figure 1.10	Échange conformationnel caractérisé par Carr-Purcell-Meiboom-Gill ( $^{15}\text{N}$ -CPMG)	16
Figure 1.11	Représentation schématique d'une expérience de dispersion de relaxation $^{15}\text{N}$ -CPMG	17
Figure 1.12	Expérience de dispersion de relaxation de Carr-Purcell-Meiboom-Gill (CPMG)	18
Figure 1.13	Tenseur de diffusion de symétrie axiale	20
Figure 1.14	Attribution des résidus de la chaîne principale en $^1\text{H}$ et $^{15}\text{N}$ de la protéine HIP2	22
Figure 1.15	Spectres RMN permettant l'attribution des résonances d'une protéine	23
Figure 1.16	Exemple d'une expérience de titrage RMN, avec la protéine PDCD5(20-104) marquée au $^{15}\text{N}$ et le ligand non marqué p53(15-61)	24
Figure 1.17	Analyse de projection du déplacement chimique (CHESPA)	26
Figure 1.18	Représentation schématique de la structure tertiaire de la RNase A	28
Figure 1.19	Mécanisme catalytique de la famille des RNases	29
Figure 1.20	Alignement de séquences des huit RNases humaines	33
Figure 1.21	Arbre phylogénétique montrant l'évolution des gènes et pseudogènes des RNases chez les vertébrés	34
Figure 1.22	Régulation de l'opéron <i>lac</i>	39
Figure 1.23	Description des étapes de la méthode de mutagenèse dirigée QuikChange™	41
Figure 1.24	Représentation graphique d'un ITC	42
Figure 1.25	Analyse spectrale de dichroïsme circulaire	44
Figure 3.1	Cytotoxicité des RNases	62
Figure 3.2	Comparaison de la structure tridimensionnelle de la RNase A et du mutant RNase A <sub>A109G</sub>	78
Figure 3.3	Stabilité des RNase A et RNase A <sub>A109G</sub>	79
Figure 4.1	Conservation évolutive de la RNase 4 humaine	106
Figure 4.2	Spectre $^1\text{H}$ - $^{15}\text{N}$ -HSQC à deux dimensions de la RNase 4 humaine	107
Figure 4.3	Représentation schématique d'analogues substrats et produits de la réaction d'une ribonucléase utilisés pour les expériences de	108



	spectroscopie RMN	
Figure 4.4	Variation des déplacements chimiques ( $\Delta\delta$ ) en $^1\text{H}$ et $^{15}\text{N}$ de la RNase 4 humaine en présence de ligands	110
Figure 4.5	Échange conformationnel à l'échelle de temps de la $\mu\text{s}$ - $\text{ms}$ chez la RNase 4 humaine	111
Figure 4.6	Homologie structurale entre EDN (Eosinophil-Derived Neurotoxine; RNase 2) et ECP (Eosinophil Cationic Protein; RNase 3)	114
Figure 4.7	Spectre $^1\text{H}$ - $^{15}\text{N}$ -HSQC à deux dimensions d'EDN	116
Figure 4.8	Titration RMN d'EDN en présence de ligands	117
Figure 4.9	Titration RMN d'ECP en présence de ligands	118
Figure 4.10	Comparaison de l'échange conformationnel entre EDN et ECP dans leur forme libre	119
Figure 4.11	Comparaison de l'échange conformationnel entre EDN et ECP en présence des produits de la réaction 5'-AMP et 3'-UMP, et de l'oligonucléotide dACACA	120

## LISTE DES ABRÉVIATIONS

### Ribonucléases

RNase	Ribonucléase
RNase A	RNase A de bovin
HPR1	Human Pancreatic Ribonuclease 1; RNase 1 humaine
EDN	Eosinophil-Derived Neurotoxin; RNase 3 humaine
ECP	Eosinophil Cationic Protein; RNase 2 humaine
Angiogénine	Angiogenin; ANG; RNase 5 humaine
RNase k6	RNase 6 humaine

### Abréviations

$\Delta\delta$	Variation du déplacement chimique (exprimée en ppm)
$\Delta G, \Delta H, \Delta S$	Variation d'énergie libre, d'enthalpie et d'entropie
ADN	Acide désoxyribonucléique
CD	Circular Dichroism (Dichroïsme circulaire)
$^{15}\text{N}$ -CPMG	Carr-Purcell-Meiboom-Gill
CSA	Chemical Shift Anisotropy (Anisotropie du déplacement chimique)
Da	Dalton (1 Da) = 1 g/mol
EDTA	Ethylenediaminetetraacetic acid
FDA	Food and Drug Administration
HSQC	Heteronuclear Single-Quantum Coherence
IPTG	Isopropyl- $\beta$ -D-thiogalactopyranoside
ITC	Isothermal Titration Calorimetry (Titration calorimétrique)
$k_B$	Constante de Boltzmann = $1,38 \times 10^{-23} \text{ m}^2 \cdot \text{kg s}^{-2} \cdot \text{K}^{-1}$
$k_{cat}$	Constante catalytique
$K_a, K_d$	Constantes thermodynamiques d'association et de dissociation d'un couple récepteur-ligand
$k_{ex}$	Rate of conformational exchange (Taux d'échange conformationnel)
$K_m$	Constante de Michaelis
$k_{on}, k_{off}$	Constantes cinétiques d'association et de dissociation d'un couple récepteur-ligand
LB Medium	Luria-Bertani broth (milieu Luria-Bertani)
min	Minute

mL	Millilitre
MW	Molecular Weight (Masse molaire)
NOE	Nuclear Overhauser Effect (Effet Overhauser Nucléaire)
NOESY	Nuclear Overhauser Effect Spectroscopy
pb	Base Pair (Paire de base)
ppm	Partie par million
PAGE	PolyAcrylamide Gel Electrophoresis (Électrophorèse sur gel de polyacrylamide)
PBS	Phosphate Buffered Saline (Tampon phosphate salin)
PCR	Polymerase Chain Reaction (Amplification en chaîne par polymérase)
PDB	Protein Database
$R_1, R_2$	Relaxations longitudinale et transverse
$R_{ex}$	Exchange contribution to relaxation (Contribution de l'échange pendant la relaxation)
RMN	Résonance magnétique nucléaire
$S^2$	Order parameter (Paramètre d'ordre)
SDS	Sodium Dodecyl Sulphate (Laurylsulfate de sodium)
UV	Ultra-Violet

## LISTE DES PUBLICATIONS

### Liste des publications incluses dans la thèse:

1. **Gagné, D.** & Doucet, N. (2013). "Structural and functional importance of local and global conformational fluctuations in the RNase A superfamily." *FEBS J* 280(22): 5596-5607.
2. **Gagné, D.**, French, R.L., Narayanan, C., Agarwal, P.K., Simonović, M. & Doucet, N. (2015). "Perturbing the conformational dynamics of an active-site loop impacts purine binding in RNase A." Soumis dans *J Am Chem Soc* le 13 mai 2015.
3. **Gagné, D.**, Charest L.A., Morin, S., Kovrigin, E.L. & Doucet, N. (2012). "Conservation of flexible residue clusters among structural and functional enzyme homologues." *J Biol Chem* 287(53): 44289-44300.
4. **Gagné, D.**, Narayanan, C. & Doucet, N. (2015). "Network of long-range concerted chemical shift displacements upon ligand binding to human angiogenin." *Protein Sci* 24(4): 525-533.

### Liste des publications non incluses dans la thèse:

1. **Gagné, D.** & Doucet, N. (2015). "Sequence-specific backbone  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  resonance assignments of human ribonuclease 4." *Biomol NMR Assign* 9(1): 181-185.
2. Grosset, A.A., Labrie, M., **Gagné, D.**, Vladoiu, M.C., Gadoury, L., Doucet, N. & St-Pierre, Y. (2014). "Cytosolic galectin-7 impairs p53 functions and induces chemoresistance in breast cancer cells." *BMC Cancer* 14: 801.
3. Gobeil, S. M., Clouthier, C.M., Park, J., **Gagné, D.**, Berghuis, A.M., Doucet, N. & Pelletier, J.N. (2014). "Maintenance of Native-like Protein Dynamics May Not Be Required for Engineering Functional Proteins." *Chem Biol* 21(10): 1330-1340.
4. Vladoiu, M., Labrie, M., Létourneau, M., Egesborg, P, Gagné, D., Billard, E., Grosset, A.A., Doucet, N., Chatenet, D. & St-Pierre, Y. (2015) "Design of a peptidic inhibitor that targets the dimer interface of a prototypic galectin." Soumis à *Oncotarget*.

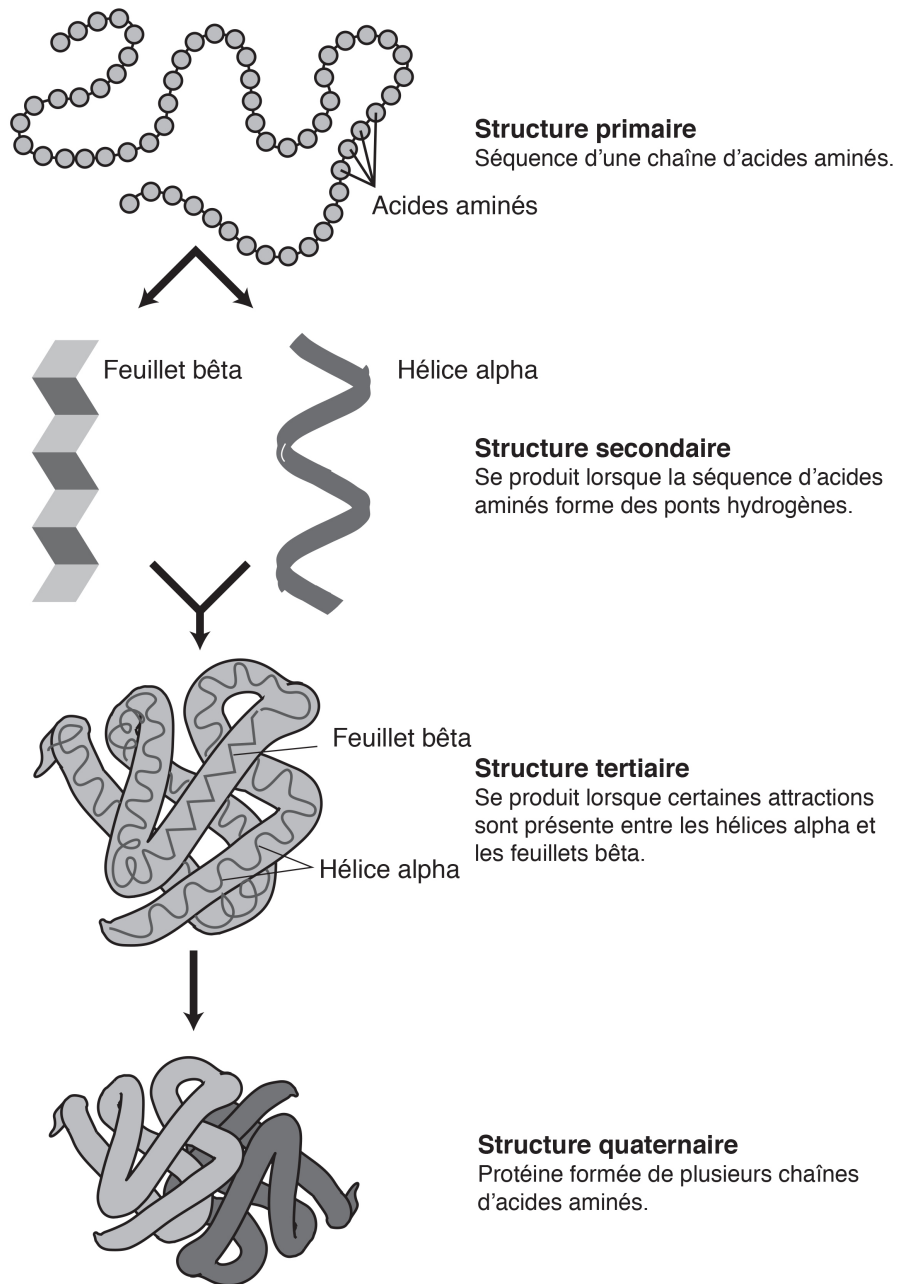
# 1 INTRODUCTION

## 1.1 Mise en contexte

Même si nous étudions les protéines depuis plus de 100 ans, nous ne comprenons toujours pas complètement comment elles fonctionnent. L'étude présentée dans cette thèse a pour objectif général de faire avancer notre compréhension du fonctionnement des enzymes. Au cours des dernières décennies, de nouveaux concepts ont été introduits dans le langage scientifique, et ils pourraient contribuer à comprendre le fonctionnement des enzymes. Ainsi, les concepts de dynamique des molécules, d'allostérie, et d'ingénierie des protéines y seront discutés plus en détails dans les pages qui suivent.

## 1.2 Les protéines

Le terme *protéine* a été introduit il y a près de 170 ans par Gerardus Johannes Mulder (1839), suite à ses études sur la composition de substances animales. Le terme provient du grec ancien *prôtos* qui signifie : *premier, essentiel*. Les protéines jouent un rôle central dans la cellule, qu'il soit structural, catalytique, ou autre. Elles sont composées d'une chaîne d'acides aminés liés entre eux par des liaisons peptidiques, formant la structure primaire de la protéine (Figure 1.1). Cette dernière se replie sur elle-même pour former des structures secondaires, tertiaires et quaternaires, qui varient selon la séquence en acides aminés, et qui leur donnent leurs fonctions. Les structures secondaires sont principalement composées des hélices alpha ( $\alpha$ ) et feuilletts bêta ( $\beta$ ), stabilisés par des liaisons hydrogènes entre les atomes. Les structures secondaires se lient ensemble via des liaisons hydrogènes, des ponts disulfures, des interactions électrostatiques, des interactions de type Van der Waals et des interactions d'empilement  $\pi$  pour former la structure tertiaire. Enfin, pour les protéines contenant plus d'une chaîne polypeptidique, l'interaction entre ces chaînes formera une structure quaternaire.



**Figure 1.1: Structure d'une protéine.** Structure primaire d'une protéine, composée d'une combinaison de 20 acides aminés formant une chaîne. La formation de ponts hydrogènes entre les acides aminés mèneront à la formation des structures secondaires, composées d'hélices alpha et de feuillets bêta. Les attractions résultant des forces ioniques, des ponts hydrogènes et des liens covalents conduiront à la formation de la structure tertiaire. Enfin, plusieurs chaînes d'acides aminés pourront s'assembler pour former la structure quaternaire. L'image a été modifiée du site <http://accessexcellence.org>.

Les protéines peuvent remplir différentes fonctions au sein de la cellule. Voici donc une liste de ces rôles :

- les protéines de structure, qui permettent à la cellule de maintenir son organisation dans l'espace, et qui sont les constituants du cytosquelette;

- les protéines de transport, qui assurent le transfert des différentes molécules dans et en dehors des cellules;
- les protéines régulatrices, qui modulent l'activité d'autres protéines ou qui contrôlent l'expression des gènes;
- les protéines de signalisation, qui captent les signaux extérieurs et assurent leur transmission dans la cellule ou l'organisme; il en existe plusieurs sortes;
- les protéines réceptrices, qui détectent les molécules messagères et les autres signaux pour que la cellule réponde en conséquence;
- les protéines sensorielles, qui détectent les signaux environnementaux (ex : lumière) et répondent en émettant des signaux dans la cellule;
- les récepteurs d'hormone, qui détectent les hormones et envoient des signaux à la cellule pour qu'elle agisse en conséquence (ex : l'insuline est une hormone qui, lorsqu'elle va être captée, va signaler à la cellule d'absorber et d'utiliser le sucre);
- les protéines motrices, qui permettent aux cellules ou organismes, ou même à certains éléments (cils), de se mouvoir ou de se déformer (ex : l'actine et la myosine permettent au muscle de se contracter);
- les protéines de défense, qui protègent la cellule contre les virus (ex. : les anticorps);
- les protéines de stockage, qui permettent la mise en réserve d'acides aminés pour pouvoir créer d'autres protéines (ex : l'ovalbumine, la principale protéine du blanc d'œuf, sert de stockage pour le développement des embryons de poulet);
- les enzymes, qui modifient la vitesse de presque toutes les réactions chimiques dans la cellule, sans être transformées dans la réaction.

La suite de la thèse portera essentiellement sur les enzymes.

### 1.3 La cinétique enzymatique

La catalyse est un processus qui permet d'augmenter la vitesse à laquelle une réaction parvient à l'équilibre en abaissant la barrière énergétique [1]. En conséquence, les enzymes sont d'excellents catalyseurs dus à leur spécificité de liaison au substrat. La cinétique enzymatique permet de décrire quantitativement les propriétés catalytiques des enzymes. Ces

mesures nous informent sur la spécificité et l'efficacité de la réaction enzymatique dans une condition donnée de pH, de température, et de concentration ionique.

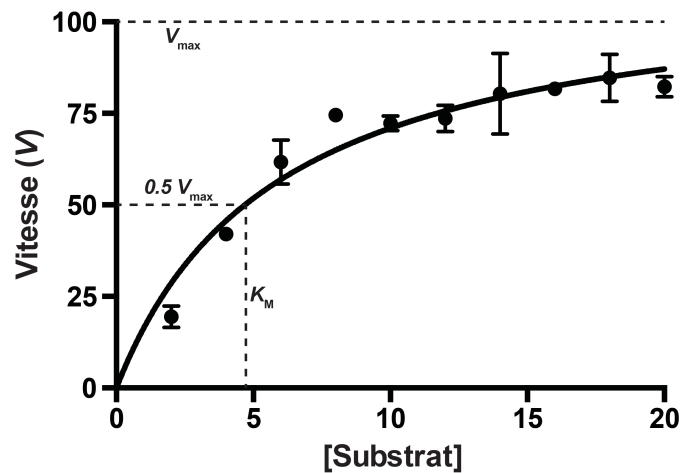
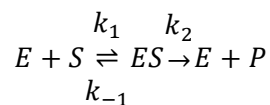


Figure 1.2: Effet de la concentration du substrat sur la vitesse de la réaction catalytique. La vitesse de la réaction augmente avec la concentration en substrat. À saturation, la vitesse approche  $V_{max}$ . Le  $K_M$  indique la concentration de substrat à  $V_{max}/2$ . L'image a été produite avec GraphPad Prism, version 6.

En 1913, Leonor Michaelis and Maud Menten postulèrent que l'enzyme (E) se fixe au substrat (S) pour former un complexe (ES). Le taux de formation du complexe est réversible et donné par les constantes  $k_1$  et  $k_{-1}$ . Par la suite, le complexe se brise à un taux  $k_2$  pour donner le produit de la réaction (P). Michaelis et Menten suppose que la formation du produit est l'étape limitante de la réaction:



La vitesse de la réaction augmente avec la concentration en substrat. À saturation, la vitesse de la réaction est maximale ( $V_{max}$ ). Étant donné que la courbe répond à une hyperbole pour la plupart des enzymes (Figure 1.2), la courbe peut s'exprimer par l'équation de Michaelis-Menten:

$$v_0 = \frac{V_{max}[S]}{K_M + [S]}$$

où  $v_0$  représente la vitesse initiale à une concentration donnée de substrat,  $V_{max}$  représente la vitesse maximale, et  $K_M$  représente la concentration en substrat pour laquelle une enzyme atteint la moitié de sa vitesse maximale. La constante catalytique  $k_{cat}$  peut être extrapolé comme suit:

$$V_{max} = k_{cat}[E_T]$$



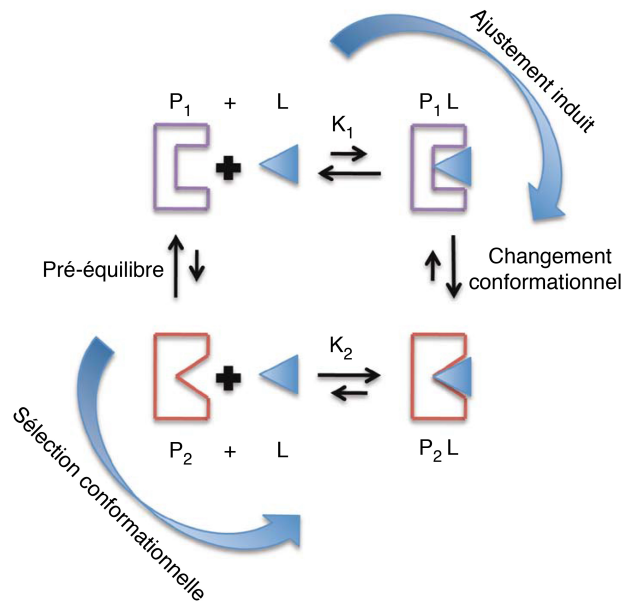
où  $[E_T]$  représente la quantité totale d'enzyme, soit  $[E_T] = [E] + [ES]$ . Ou encore, selon l'équation:

$$K_M = \frac{k_{off} + k_{cat}}{k_{on}}$$

où les constantes  $k_{on} = k_1$  et  $k_{off} = k_{-1}$ . Le rapport  $k_{cat}/K_M$  quantifie l'efficacité catalytique, puisqu'il correspond à une constante de vitesse pour de faibles concentrations en substrat, où  $[S] \ll K_M$  et  $[E] = [E_T]$ .

## 1.4 Les mouvements atomiques

En 1894, Hermann Emil Fischer proposa un modèle de "clef et serrure" afin d'expliquer la liaison entre une enzyme et son substrat [2]. Dans ce modèle, l'enzyme est rigide et c'est le substrat qui devait s'adapter parfaitement au site actif pour permettre la reconnaissance. Toutefois, dans les années qui ont suivi, il devenait évident que ce modèle ne pouvait en lui-même expliquer tous les aspects de la catalyse enzymatique. Ce n'est qu'en 1958 que Daniel E. Koshland Jr. suggéra un modèle d'ajustement induit, ou "induced fit", dans lequel la reconnaissance du substrat induit un changement conformationnel [3] (Figure 1.3). Il s'agissait du premier exemple où l'on proposait qu'une protéine pût être flexible. Par la suite, il devenait évident que le modèle proposé par Koshland était limité. En effet, l'ajustement induit n'est possible que s'il se produit *a priori* une interaction entre l'enzyme et son substrat. Plusieurs chercheurs ont proposé le mécanisme de sélection conformationnelle, ou "Conformational Selection", un modèle complémentaire à celui de l'ajustement induit [4-8]. Dans ce modèle, toutes les conformations de la protéine libre sont présentes en solution, et le substrat sélectionne la conformation la plus favorable énergétiquement (Figure 1.3).



**Figure 1.3: Cycle thermodynamique du mécanisme de reconnaissance moléculaire. Dans la sélection conformationnelle, la conformation de liaison ( $P_2$ ) est préexistante dans la solution avant la liaison du ligand ( $L$ ). Dans l'ajustement induit, la liaison du ligand induit un changement conformationnel (de  $P_1L$  à  $P_2L$ ) thermodynamiquement plus favorable. Les constantes cinétiques et thermodynamiques ( $K_1$  et  $K_2$ ) peuvent déterminer si la sélection conformationnelle ou ajustement induit est plus probable. L'image a été modifiée de la référence [9].**

De nos jours, le concept qu'une protéine peut être flexible est bien accepté. Ces mouvements atomiques peuvent être observés non seulement autour du site actif, mais également sur l'entièreté de la protéine. Il a été démontré que ces mouvements ne sont pas aléatoires, mais plutôt coordonnés avec une fonction donnée de l'enzyme. De nombreux exemples sont d'ailleurs présentés dans la littérature scientifique: la dihydrofolate réductase (DHFR) [10, 11], la ribonucléase A (RNase A) [12], l'adénylyl cyclase [13, 14] et la cyclophiline A (CypA) [15, 16]. Ces fluctuations peuvent se produire à des échelles de temps allant de la femtoseconde à plusieurs heures [17-19] (Figure 1.4). Les mouvements se produisant dans l'ordre de la microseconde à la milliseconde ( $\mu\text{s}$ - $\text{ms}$ ) sont particulièrement intéressants, puisqu'ils coïncident souvent avec la catalyse enzymatique, laquelle inclut la liaison du substrat, la réaction, et la libération du produit. Il a été démontré que la perturbation des mouvements à cette échelle de temps modifie considérablement la catalyse enzymatique [16].

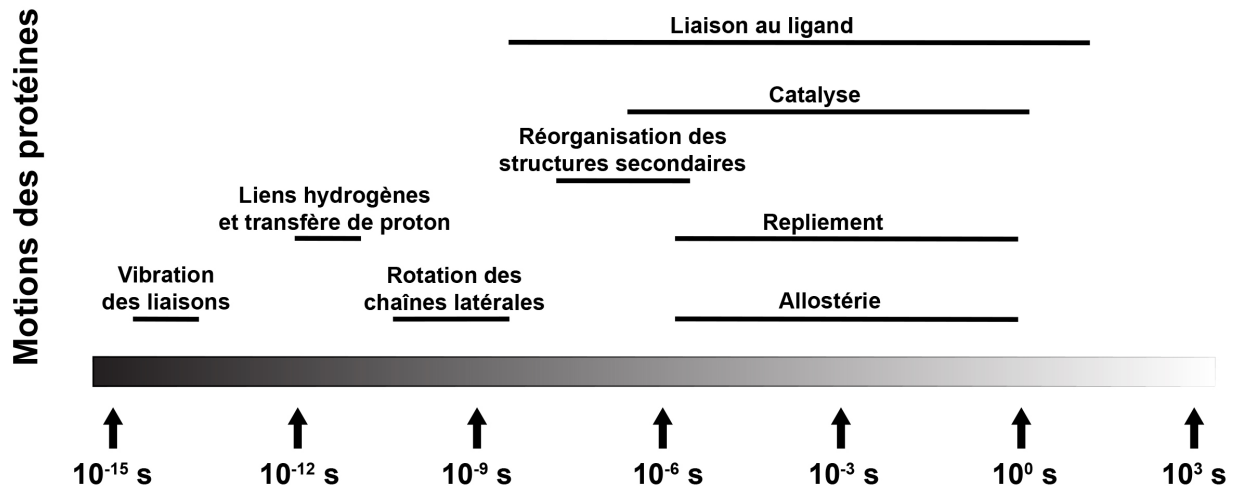


Figure 1.4: Échelles de temps d'évènements dynamiques. L'image a été modifiée de la référence [17].

### 1.4.1 L'allostérie

La théorie de l'allostérie découle du mécanisme de la sélection conformationnelle décrit plus tôt. Monod *et al.* [20], en 1965, ont été les premiers à décrire le concept d'allostérie. Le concept se résume par l'effet ressenti à un site A, suite à une perturbation à un site B [21]. La perturbation au site B peut être causée par la liaison d'un ligand (protéine, ARN, etc.), une mutation, ou encore par photoactivation (Figure 1.5). Ainsi, l'allostérie s'avère un mécanisme très important, qui permet d'adapter sa fonction selon des conditions environnementales changeantes.

La conception de nouveaux médicaments thérapeutiques est généralement basée sur le modèle statique de protéine et vise presque essentiellement le site actif. Bien qu'un nombre important de médicaments ait été identifié en bloquant le site actif, le taux de succès demeure faible, ce qui augmente considérablement le coût total relié à la découverte de nouvelles cibles. Ainsi, la régulation allostérique de macromolécules biologiques offre une nouvelle avenue [22]. En se liant à un site fonctionnel éloigné du site catalytique d'une enzyme ou du site actif principal d'une protéine, les effecteurs allostériques peuvent déclencher une cascade de changements conformationnels qui modifient la fonction et/ou l'activité enzymatique (Figure 1.5C). Plusieurs composés allostériques sont déjà sur le marché, ou sont en voie d'approbation par la FDA [22], tels l'Imatinib (Gleevec) [23], ou les inhibiteurs non nucléosidiques qui ciblent la transcriptase inverse du VIH [24] et la polymérase du virus C de l'herpès [25].

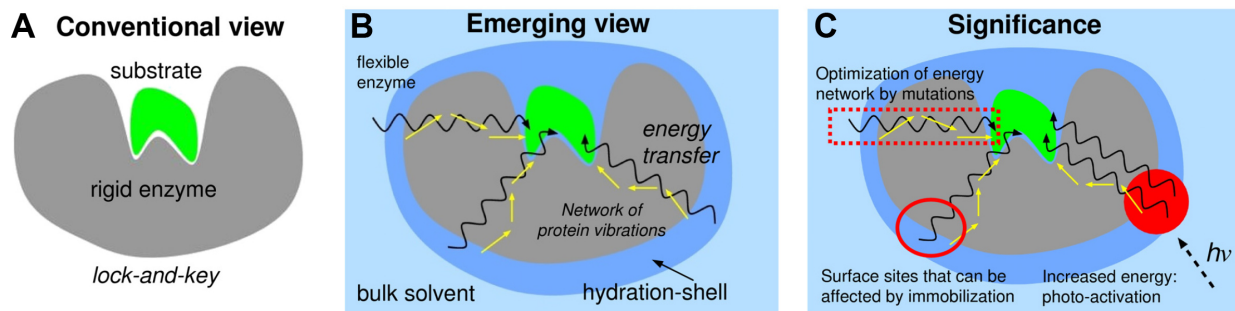


Figure 1.5: Interconnectivité entre la structure de la protéine, sa dynamique et sa fonction. (A) Représentation graphique de la théorie de "clef et serrure" proposée par Hermann Emil Fischer en 1894. (B) Plutôt que d'être rigide, les enzymes seraient flexibles. Les résidus seraient interconnectés via des réseaux dynamiques qui pourraient s'étendre de la surface de la protéine, en incluant le solvant, et ce, jusqu'au site actif. L'exemple peut s'appliquer à des mouvements à l'échelle de temps dit rapide ( $>\mu\text{s}$ ) ou lent ( $<\mu\text{s}$ ). (C) La fonction proposée de ces réseaux dynamiques serait de fournir un transfert thermodynamique entre les différentes régions de la protéine. Ce transfert d'énergie pourrait être induit par la photoactivation ou par la présence d'un ligand à un site allostérique situé à la surface de la protéine, ou être affecté par une mutation distante du site actif. La figure a été modifiée de la référence (Pratul Agarwal - Notes personnelles).

Bien que cette approche soit intéressante, il n'en demeure pas moins qu'il est très difficile d'identifier les sites allostériques. À ce titre, plusieurs méthodes sont particulièrement adaptées aux études d'allostérie et de mouvements atomiques en général. Ces méthodes seront discutées plus en profondeur dans les prochaines sections.

#### 1.4.2 Principales méthodes d'analyse de la dynamique moléculaire

Une variété de méthodes permet l'analyse des mouvements chez les protéines, et ce, à différentes échelles de temps. Reprenons le tableau de la figure 1.4, mais en y ajoutant plus de détails (Figure 1.6). Dans les paragraphes qui suivent, nous discuterons brièvement de quelques-unes de ces méthodes, et comment elles peuvent contribuer à l'analyse de la dynamique moléculaire.

La cristallographie conventionnelle utilise 3 dimensions, soit x, y et z, et présente une moyenne temporelle d'un ensemble de structures. Toutefois, en y ajoutant une 4e dimension (le temps), il devient possible de visualiser les changements structuraux en fonction du temps. En effet, un cristal contient, dans sa matrice, des canaux de solvant permettant la diffusion de petites molécules. Ces canaux peuvent être utilisés pour diffuser un substrat et activer une réaction enzymatique. Bien qu'intéressante, la méthode possède certaines limitations: 1. La protéine doit être biologiquement active à l'état cristallin; 2. Il doit être possible d'activer la réaction dans le cristal; 3. La réaction enzymatique ne doit pas perturber l'état cristallin et la diffraction des rayons X; 4. L'intermédiaire de réaction doit être détectable dans le cristal.

Depuis l'apparition de la première simulation de dynamique moléculaire en 1977 [26], la bio-informatique a pris une importance considérable. Un avantage majeur de l'utilisation des méthodes bio-informatiques, c'est qu'elles permettent d'aller là où d'autres méthodes ne le permettent pas techniquement, telle l'utilisation de substrats ou d'intermédiaires "instables" de réaction, et l'observation atomique en temps réel. Toutefois, les expériences étant basées sur des prédictions, les résultats doivent être interprétés avec précaution. Puisque les biologistes cherchent à comprendre les mouvements atomiques en relation avec la catalyse enzymatique, l'intérêt principal est de caractériser les mouvements à l'échelle de temps  $10^{-6}$  à  $10^0$  secondes (se référer à la figure 1.4 du présent document). Cependant, plus la fréquence des mouvements est faible, et plus le nombre de calculs requis est importants, Il s'agit d'une limitation majeure des méthodes bio-informatiques, une restriction causée par la puissance du "Central Processor Unit" (CPU). Ainsi, les bio-informaticiens estiment certains paramètres afin de réduire le nombre de calculs requis par le CPU, et ainsi s'approcher de l'échelle de temps d'intérêt. Malheureusement, cette approche réduit considérablement la précision de l'étude.

La résonance magnétique nucléaire (RMN) s'avère la plus versatile des méthodes expérimentales. Elle permet de quantifier les mouvements atomiques sur un large spectre d'échelles de temps. Le formalisme model-free, développé par Lipari et Szabo [27], permet l'étude des mouvements atomiques à l'échelle de temps de la picoseconde-nanoseconde (ps-ns). La méthode utilise les données expérimentales de  $R_1$ ,  $R_2$  et NOE afin de déterminer les amplitudes et les échelles de temps des mouvements intramoléculaires. Les mouvements sont décrits par le temps de corrélation global ( $\tau_m$ ), le paramètre d'ordre ( $S^2$ ), le temps de corrélation effectif ( $\tau_e$ ), et le paramètre d'échange conformationnel ( $R_{ex}$ ). Enfin, le Carr-Purcell-Meiboom-Gill ( $^{15}\text{N}$ -CPMG) permet la mesure de la flexibilité à l'échelle de temps de la catalyse enzymatique. Cette méthode permet d'extraire le taux d'échange conformationnel ( $k_{ex}$ ) de chacun des résidus présentant une courbe de dispersion de la relaxation. Puisque ces deux méthodes seront largement utilisées au cours de cette étude, une description plus détaillée est présentée dans les pages qui suivent.

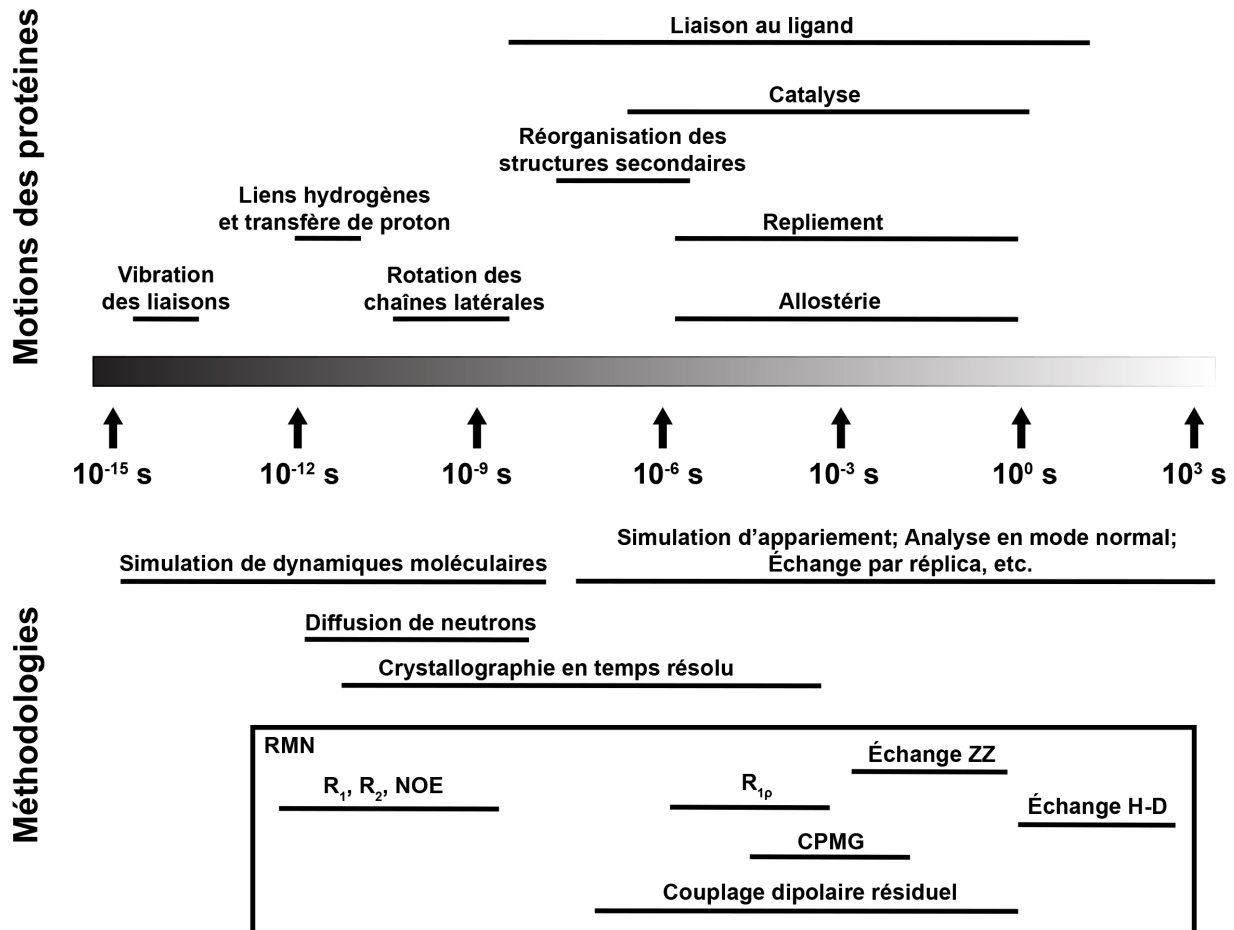


Figure 1.6: Techniques d'analyse d'évènements dynamiques en fonction de l'échelle de temps et du type de mouvements. L'image a été modifiée de la référence [17].

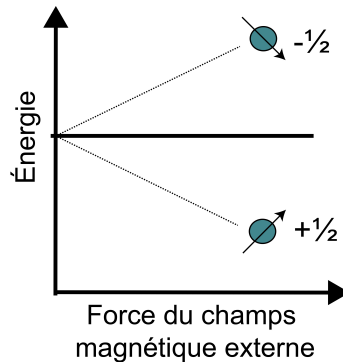
## 1.5 La dynamique des protéines

### 1.5.1 Principe de la RMN

La RMN est une technique qui utilise la propriété de spin non nul des noyaux des atomes. Lorsque placé dans un champ magnétique donné ( $B_0$ ), chaque atome absorbe l'énergie du rayonnement de radiofréquences auxquelles il est soumis, puis la relâche suite à une relaxation. Le signal obtenu permet de caractériser la structure et la dynamique d'une protéine d'intérêt à l'échelle atomique.

Le noyau d'un atome possède un moment magnétique, ou spin. La valeur de ce spin est une propriété intrinsèque au noyau, mais dépend du nombre de protons et de neutrons qui le composent. Un noyau possédant un nombre pair de nucléons aura un spin entier (0, 1, 2...)

alors qu'un noyau ayant un nombre impair aura un spin demi-entier ( $1/2, 3/2, 5/2\dots$ ). La RMN étudie généralement les atomes de spin  $1/2$ . Dans un champ magnétique, le nombre d'orientations possibles d'un noyau de spin  $I$  sera de  $2I+1$ , et varieront de  $-I$  à  $+I$  par incréments de 1. Le noyau de spin  $1/2$  peut être visualisé comme un petit aimant-barre, pouvant prendre 2 orientations possibles, soit  $+1/2$  et  $-1/2$  (Figure 1.7). Le signal qui est observé en RMN est une conséquence de la transition entre les 2 états suite à une stimulation par radiofréquence (RF).



**Figure 1.7: Orientation du moment magnétique d'un atome avec  $I = 1/2$  dans un champ magnétique  $B_0$ . La différence entre les 2 niveaux d'énergies est donnée par  $\Delta E$ .**

Le moment magnétique  $\mu$  est proportionnel au vecteur de rotation angulaire  $I$ , et du rapport gyromagnétique  $\gamma$ :

$$\mu = \gamma I$$

Les atomes possèdent des rapports gyromagnétiques différents, ce qui permet de les distinguer entre eux. La différence d'énergie entre les 2 niveaux sera donnée par l'équation:

$$\Delta E = h\nu = \hbar\gamma B_0$$

et sera proportionnelle au champ magnétique appliqué  $B_0$ . La fréquence de résonance ( $\nu$ ), ou précession de Larmor, sera calculée comme suit:

$$\nu = (\gamma B_0)/2\pi$$

Chaque atome aura une fréquence de résonance caractéristique, due à de petites perturbations causées par son environnement magnétique local. Ces perturbations sont influencées par le positionnement de l'atome dans la structure de la protéine et par rapport au solvant, via le positionnement des chaînes latérales, des résidus aromatiques, des angles de torsion des liaisons et des liaisons hydrogènes.

### 1.5.2 La relaxation des spins

La relaxation des spins permet la quantification des mouvements atomiques à l'échelle de temps de la picoseconde à la nanoseconde (ps-ns) et de la microseconde à la milliseconde ( $\mu$ s-ms). La sélection des atomes d'intérêts doit être au préalable déterminée. Pour les protéines, les atomes  $^1\text{H}$ ,  $^{13}\text{C}$ , et  $^{15}\text{N}$  sont les plus fréquemment utilisés. Puisque seul  $^1\text{H}$  est présent naturellement, les autres atomes doivent être incorporés lors de l'expression de la protéine dans un milieu minimal. L'incorporation du  $^{15}\text{N}$  est fréquemment utilisé, et permet l'analyse des groupements amides ( $^1\text{H}$ - $^{15}\text{N}$ ) de chaque résidu (excepté pour les prolines, qui ne possèdent pas de groupement amide) formant la protéine. Il est à noter que pour l'étude qui suit, seuls les groupements amides ont été analysés et par conséquent, la théorie traite de la relaxation des spins  $^1\text{H}$  et  $^{15}\text{N}$ .

Les deux principales interactions qui gouvernent le comportement de la relaxation des spins sont les interactions dipôle-dipôle avec les atomes voisins, et l'anisotropie du déplacement chimique ("Chemical Shift Anisotropy", CSA) [28]. La technique utilise généralement trois types de mesures, soit la relaxation longitudinale ( $R_1$ ), la relaxation transversale ( $R_2$ ) et l'approximation des états quasi stationnaires de l'effet Overhauser nucléaire ("Steady-State Heteronuclear Nuclear Overhauser Effect").

Le signal RMN qui est observé se nomme le FID ("Free Induction Decay"). Le signal obtenu, la magnétisation, est la somme de l'ensemble des moments magnétiques des atomes de la protéine en fonction du temps. En conséquence, le FID observé est la somme de tous les FID individuels pour chacun des atomes oscillant à une fréquence caractéristique. Le taux de décroissance de l'intensité du signal peut être mesuré par les expériences de  $R_1$  et  $R_2$  (ces expériences sont décrites dans les pages qui suivent). La transformation de Fourier permet ensuite de convertir le signal en fréquence (Figure 1.8). Trois paramètres peuvent être extraits, soit le déplacement chimique, l'intensité, et la largeur de bande. Ces paramètres seront décrits dans les pages qui suivent.



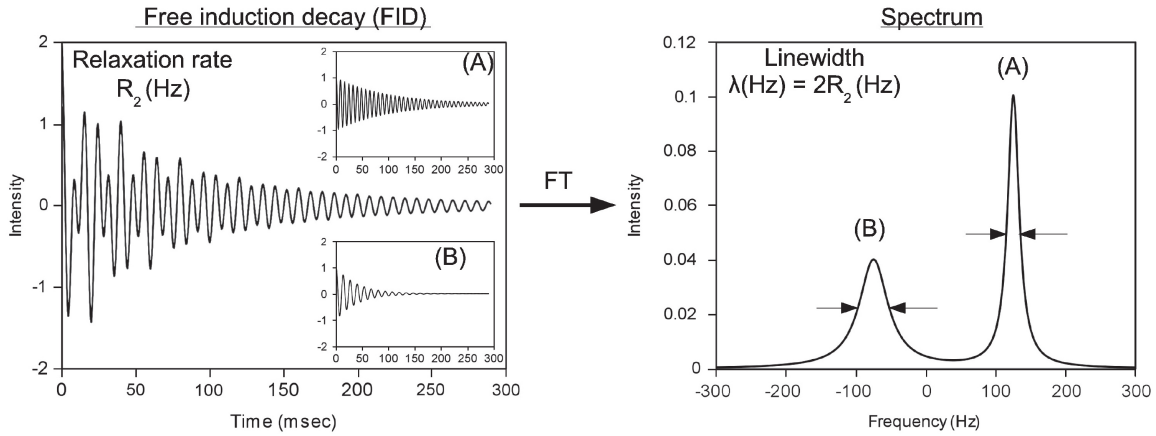


Figure 1.8: Description du signal RMN. Le signal RMN est observé sous la forme d'une fréquence en fonction du temps, aussi appelé FID ("Free Induction Decay"). Le FID (gauche) est la somme de tous les FIDs de chacun des atomes individuels présent dans la protéine et peut être quantifié par la mesure de  $R_2$ . La transformation de Fourier (FT) permet de convertir le signal en fréquence (droite). Ainsi, la largeur de bande et l'intensité du signal seront affectées par la fréquence et la décroissance du signal (A et B). L'image a été tiré de la référence [29].

### 1.5.3 Relaxation longitudinale $R_1$

La relaxation longitudinale  $R_1$  ( $R_1 = 1/T_1$ ) correspond à la constante de temps du retour à l'équilibre thermodynamique des spins qui ont été excités par l'impulsion radio-fréquence.  $R_1$  est décrit par l'équation suivantes [30]:

$$R_1 = d[J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)] + cJ(\omega_N)$$

où  $d$  représente la constante dipolaire:

$$d = \frac{1}{4} \left( \frac{\mu_0}{4\pi} \right)^2 \frac{(\hbar \gamma_N \gamma_H)^2}{r_{N-H}^6}$$

où  $c$  décrit l'anisotropie du déplacement chimique:

$$c = \frac{(\omega_N \Delta\sigma)^2}{3}$$

où  $J(\omega)$  représente la fonction de densité spectrale à différentes fréquences de Larmor,  $\mu_0$  symbolise la permittivité magnétique du vide,  $\hbar$ , la constante de Planck,  $\gamma_N$  et  $\gamma_H$ , les rapports gyromagnétiques pour  $^{15}\text{N}$  et  $^1\text{H}$ , respectivement,  $r_{N-H}$ , la vibration effective moyenne d'un lien N-H,  $\Delta\sigma$ , le CSA, et la fréquence de Larmor du  $^{15}\text{N}$  est décrite par l'équation:

$$\omega_N = 2\pi\nu_N = \gamma_N B_0$$

où  $\nu_N$  indique la fréquence de Larmor en Hz, et  $B_0$  décrit la force du champs magnétique.

L'approche de type "inversion-recovery" est utilisée pour l'enregistrement des données  $R_1$  du groupement amide (Figure 1.9). En bref, une première impulsion à  $180^\circ$  permet d'inverser la magnétisation. Après un délai, permettant aux spins de relaxer, une nouvelle impulsion à  $90^\circ$  permet de ramener la magnétisation dans le plan  $xy$ , et le signal  $R_1$  est enregistré. Puisque  $R_1$  est affecté par la fréquence de Larmor et le temps de corrélation effectif ( $\tau_c$ ), la variation du délai entre deux impulsions affectera l'amplitude de la magnétisation ( $A_t$ ) qui sera enregistrée selon l'équation:

$$A_t = A_0 e^{-R_1 t}$$

où  $A_0$  indique l'amplitude au temps 0, et  $t$ , le délai entre les deux impulsions.

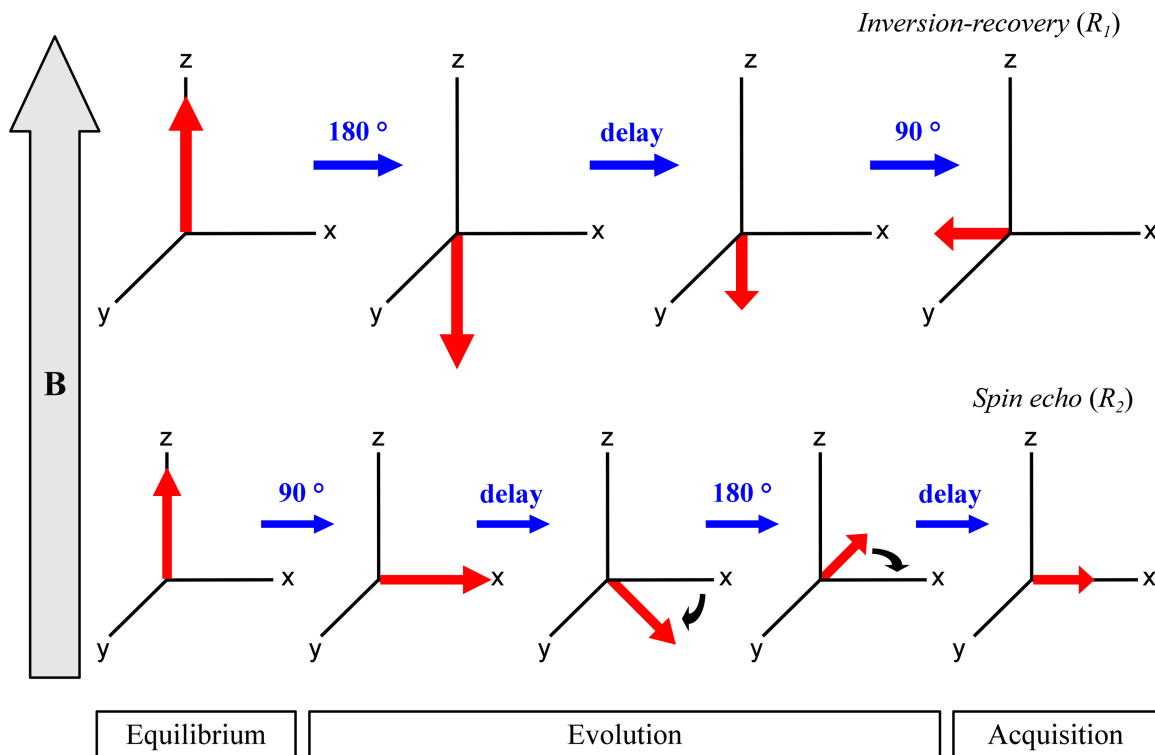


Figure 1.9: Approche expérimentale pour l'enregistrement du  $R_1$  (inversion recovery) et du  $R_2$  (spin echo) dans un champ magnétique  $B$ . L'état de magnétisation ( $M_0$ ) est représenté par la flèche rouge. La mesure du taux de relaxation est affecté par la variation des délais, qui affecte l'amplitude du signal. L'image a été tirée de la référence [31].

#### 1.5.4 Relaxation transversale $R_2$

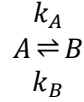
La relaxation transversale  $R_2$  ( $R_2 = 1/T_2$ ) correspond à la perte de cohérence des spins responsables de la magnétisation transversale (dans le plan  $xy$ ).  $R_2$  se définit comme suit:

$$R_2^0 = \frac{d}{2} [4J(0) + J(\omega_H + \omega_N) + 3J(\omega_N) + 6J(\omega_N) + 6J(\omega_N + \omega_N)] + \frac{c}{6} [4J(0) + 3J(\omega_N)]$$

et

$$R_2 = R_2^0 + R_{ex}$$

où  $R_2^0$  symbolise la relaxation transversale pure, et  $R_{ex}$ , la contribution de  $R_2$  pour les processus d'échange conformationnel se produisant à l'échelle de la  $\mu$ s-ms. Lors d'une échange entre deux états A et B:



où  $k_A$  et  $k_B$  représentent les taux d'échanges apparents entre les deux états,  $R_{ex}$  se définit comme suit:

$$R_{ex} \approx \frac{\rho_A(1 - \rho_A)k_{ex}}{1 + (k_{ex}/\Delta\omega)^2}$$

où  $\rho_A$  et  $\rho_B$  symbolisent les populations des deux états A et B, et  $\Delta\omega$  représente la différence de fréquence de résonance entre les deux états. Cependant, l'approximation de  $R_{ex}$  n'est valide que si  $\rho_A \gg \rho_B$  et que  $R_{2,A} \approx R_{2,B}$ .

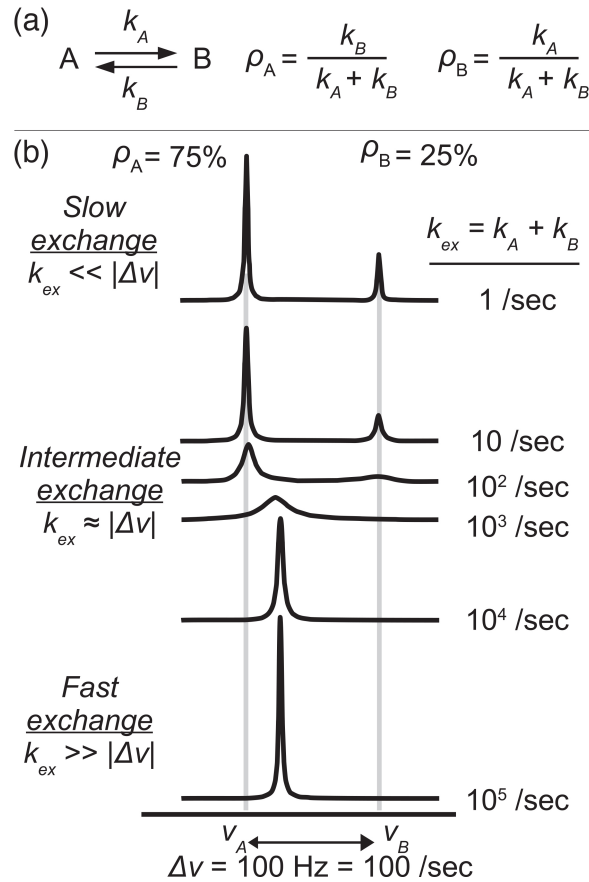
L'échange entre les deux états peut être lent, intermédiaire ou rapide, et variera en fonction de  $k_{ex}$  et  $\Delta\omega = 2\pi\Delta\nu$  (Figure 1.10). Le paramètre  $\alpha$  permet de calculer le régime de l'échange, selon l'équation [32]:

$$\alpha = \frac{2(k_{ex}/\Delta\omega)^2}{1 + (k_{ex}/\Delta\omega)^2}$$

où  $0 \leq \alpha < 1$  indique un échange lent,  $\alpha \approx 1$ , un échange intermédiaire, et  $1 < \alpha \leq 2$ , un échange rapide (Figure 1.10). Pour simplifier le calcul, le paramètre  $\alpha$  peut être approximé selon l'équation:

$$\alpha = \left( \frac{B_{0,2} + B_{0,1}}{B_{0,2} - B_{0,1}} \right) \left( \frac{R_{ex,2} - R_{ex,1}}{R_{ex,2} + R_{ex,1}} \right)$$

où  $B_{0,1}$  et  $B_{0,2}$  représentent la force des deux champs magnétiques 1 et 2, et  $R_{ex,1}$  et  $R_{ex,2}$  indiquent la contribution de l'échange conformationnel.



**Figure 1.10:** Échange conformationnel caractérisé par l'expérience de Carr-Purcell-Meiboom-Gill ( $^{15}\text{N}$ -CPMG). (A) Les variables  $k_A$  et  $k_B$  représentent les taux d'échanges conformationnels unidirectionnels des états A et B, dont les populations sont identifiées par  $\rho_A$  et  $\rho_B$ . (B) Le taux d'échange conformationnel ( $k_{ex}$ ) peut-être extrait à partir des variables  $k_A$  et  $k_B$ . La variation de fréquence ( $\Delta\nu$ ) entre les pics des deux états conformationnels permet d'identifier le type d'échange. L'image a été modifiée de la référence [29].

Les données  $R_2$  du groupement amide dans les protéines sont en général enregistrées en utilisant une approche de type "spin echo" intégrée à la séquence d'impulsions du  $^1\text{H}$ - $^{15}\text{N}$  HSQC (Figure 1.9). En bref, une première impulsion à  $90^\circ$ , suivie d'un délai, permet de transférer la magnétisation dans le plan  $xy$ . Une nouvelle impulsion à  $180^\circ$  est ensuite appliquée, suivi d'un délai de durée égal au précédent, ce qui permet d'inverser le signal et de créer un écho, soit le retour à la cohérence des spins de fréquence angulaire différente. Le signal  $R_2$  est enregistré après le second délai. Tout comme  $R_1$ , l'amplitude de la magnétisation dépendra de la fréquence de Larmor et de la durée des délais, selon l'équation:

$$A_t = A_0 e^{-R_2 t}$$

### 1.5.5 Effet Overhauser nucléaire NOE

L'effet Overhauser nucléaire (NOE) décrit une forme de relaxation causée par l'interaction entre deux spins à travers l'espace, et se définit comme suit:

$$NOE = 1 + \frac{\gamma_H}{\gamma_N} \frac{d}{R_1} [6J(\omega_H + \omega_N) - J(\omega_N - \omega_N)]$$

Le NOE est obtenu par l'acquisition de deux spectres, soit avec et sans pré-saturation des protons. La saturation des protons permet d'égaliser les populations de spins  $^1\text{H}$  transitionnels, et amène la magnétisation de  $^{15}\text{N}$  vers un état stable. Ainsi, le NOE correspond au rapport entre l'amplitude des spins en  $^{15}\text{N}$  avec ( $A_{sat}$ ) et sans ( $A_{ref}$ ) pré-saturation des protons:

$$NOE = \frac{A_{sat}}{A_{ref}}$$

### 1.5.6 Les expériences RMN

#### 1.5.6.1 La dispersion de relaxation Carr-Purcell-Meiboom-Gill (CPMG)

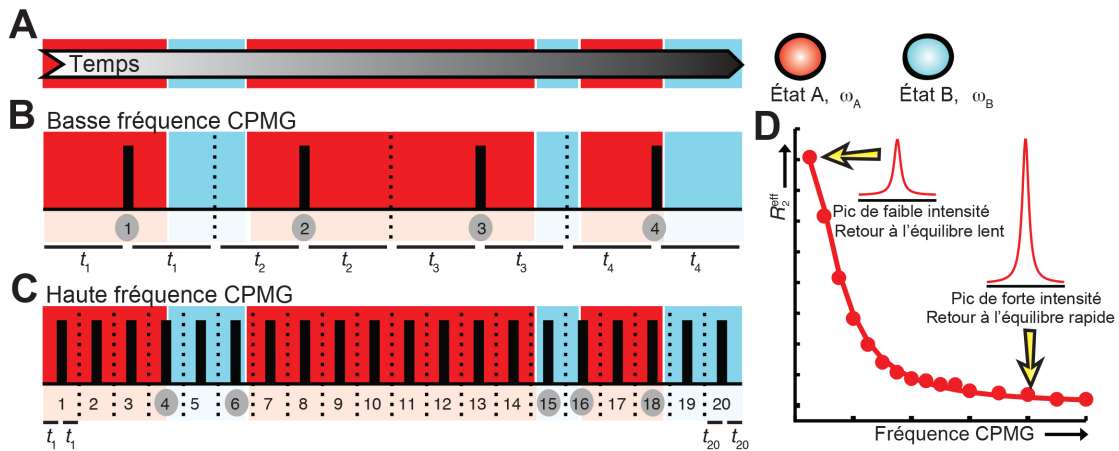
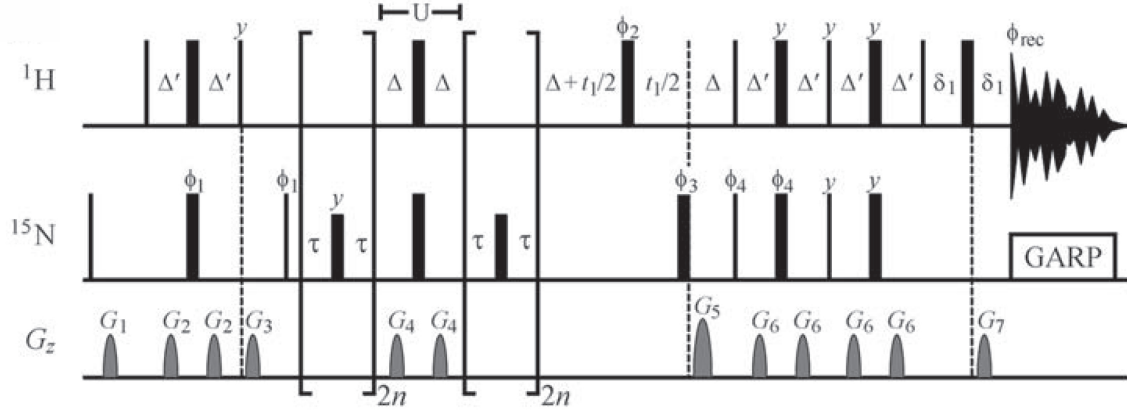


Figure 1.11: Représentation schématique d'une expérience de dispersion de relaxation  $^{15}\text{N}$ -CPMG. (A) Échange conformationnel entre deux états, A (rouge) et B (bleu), en fonction du temps. (B,C) Application d'une suite d'impulsions à haute et basse fréquences (avec des temps de relaxation  $\tau_x$ , où x représente la numérotation séquentielle du délai  $\tau$ ), permettant de détecter les mouvements de chacun des atomes. (D) Courbe de dispersion de relaxation, qui permet de détailler l'échange conformationnel d'un atome. L'image a été modifiée de la référence [33].

Les mouvements atomiques à l'échelle de la  $\mu\text{s}$ -ms correspondent à plusieurs évènements biochimiques, tels que la catalyse enzymatique ou la liaison aux substrats. L'expérience de CPMG permet de détecter les changements conformationnels se situant entre

50 et 3000 échanges  $A \leftrightarrow B$  par seconde sur la protéine. Les mouvements moléculaires à cette échelle de temps altèrent la fréquence de résonance et créent un élargissement des résonances, affectant du même coup le taux de relaxation transverse  $R_{2,eff}(1/\tau_{CP})$ . Ainsi, il est possible de mesurer la décroissance du signal  $R_{2,eff}(1/\tau_{CP})$  par CPMG en lançant une série d'impulsions à haute et basse fréquences (avec des temps de relaxation de durée  $\nu_{CPMG} = 1/\tau_{CP}$ ) (Figures 1.11 et 1.12).



**Figure 1.12: Expérience de dispersion de relaxation de Carr-Purcell-Meiboom-Gill (CPMG).** La séquence d'impulsions permet la quantification des mouvements atomiques à l'échelle de la  $\mu\text{s}$ -ms dans les protéines.  $\tau = \tau_{CP}$ ,  $\Delta = 1/(4J_{NH}) = 2.7 \text{ ms}$ ,  $\delta = 200 \text{ ms}$ , et le temps de relaxation total est  $t = (2n)(8\tau)$ . Le cycle de phase est  $\Phi_1 = x, -x$ ;  $\Phi_2 = 4(x), 4(-x)$ ;  $\Phi_3 = x, x, y, y, -x, -x, -y, -y$ ;  $\Phi_4 = y$ ; récepteur =  $x, -x, x, -x$ . La sélection de la cohérence est obtenue en inversant l'amplitude de  $G_5$  et  $\Phi_4$ ; la phase du récepteur et  $\Phi_1$  sont inversés pour chaque augmentation de  $t_1$ . L'image a été modifiée de la référence [28].

L'équation  $R_{2,eff}(1/\tau_{CP})$  peut s'exprimer comme suit [34]:

$$R_{2,eff}(1/\tau_{CP}) = \frac{1}{2} \left( R_{2A}^0 + R_{2B}^0 + k_{ex} - \frac{1}{\tau_{CP}} \cosh[D_+ \cosh(\eta_+) - D_- \cos(\eta_-)]^{-1} \right)$$

$$D_{\pm} = \frac{1}{2} \left[ \pm 1 + \frac{\Psi + 2\Delta\omega^2}{(\Psi^2 + \zeta^2)^{1/2}} \right]$$

$$\eta_{\pm} = \frac{\tau_{CP}}{\sqrt{2}} \left[ \pm \Psi + (\Psi^2 + \zeta^2)^{1/2} \right]^{1/2}$$

$$\Psi = (R_{2A}^0 - R_{2B}^0 - \rho_A k_{ex} + \rho_B k_{ex})^2 - \Delta\omega^2 + 4\rho_A \rho_B k_{ex}^2$$

$$\zeta = 2\Delta\omega (R_{2A}^0 - R_{2B}^0 - \rho_A k_{ex} + \rho_B k_{ex})$$

Où  $R_{A/B}$  et  $\rho_{A/B}$  font référence aux taux de relaxation transverse et aux populations (à l'équilibre) entre les sites A et B, séparés par un déplacement chimique de  $\Delta\omega$ . Lorsque l'échange est rapide, c'est-à-dire que  $k_{ex} \gg |\Delta\omega|$ , on est assuré que  $\rho_A \gg \rho_B$ . L'équation peut ensuite être simplifiée comme suit [34]:

$$R_2(1/\tau_{CP}) = R_2^0 + \Phi_{ex}/k_{ex} [1 - (2/k_{ex}\tau_{CP})\tanh(k_{ex}\tau_{CP}/2)]$$

Où  $\Phi_{ex} = (\omega_A - \omega_B)^2 \rho_A \rho_B$ . Ainsi, en traçant une courbe du signal  $R_{2,eff}(1/\tau_{CP})$  en fonction de  $1/\tau_{CP}$  (Figure 1.11D), il est possible d'en extraire mathématiquement la constante d'échange ( $k_{ex}$ ) entre conformères transitoires en solution, en plus de permettre l'obtention des ratios des populations de chacun des conformères ( $\rho_A, \rho_B$ ) et la différence de déplacement chimique entre l'état fondamental A et l'état excité B ( $\Delta\omega$ ). Ultimement, la technique permet également de résoudre la structure moléculaire de l'état excité (conformère B), même s'il s'avère *a priori* invisible dans le spectre  $^1\text{H}-^{15}\text{N}$  HSQC [33, 35, 36].

### 1.5.6.2 Formalisme model-free

Le formalisme model-free permet de détecter les mouvements atomiques des protéines à l'échelle de temps de la ps-ns. Une estimation des mouvements atomiques à l'échelle de temps de la  $\mu\text{s}$ -ms, provenant de la contribution de la relaxation transversale des spins, permet d'extraire  $R_{ex}$ . La technique a été introduite par Lipari et Szabo en 1982 [27, 37]. Son intérêt vient du fait qu'il soit possible de découpler la rotation globale de la molécule des mouvements locaux.

La fonction de corrélation  $C(\tau)$  est le produit de la fonction de corrélation brownienne sur la diffusion rotationnelle  $C_0(\tau)$  (rotation globale de la macromolécule) et de la fonction de corrélation de la dynamique interne  $C_I(\tau)$  (mouvements des liens N—H) [37]:

$$C(\tau) = C_0(\tau) \cdot C_I(\tau)$$

Cette approche permet de calculer la fonction de corrélation de la dynamique interne de façon indépendante de la rotation globale:

$$C_0(\tau) = \frac{1}{5} \sum_{i=-k}^k w_i \cdot e^{-\tau/\tau_i}$$

et

$$C_I(\tau) = S^2 + (1 - S_f^2)e^{\tau/\tau_f} + (S_f^2 - S^2)e^{\tau/\tau_s}$$

où  $w$  correspond à un facteur de pondération,  $S^2$  est le carré du paramètre d'ordre,  $\tau$  correspond à l'échelle de temps effectif (caractérisé par  $S^2$  dans le cas de mouvements internes), les indices  $s$  et  $f$  réfèrent aux mouvements lents (nanoseconde) et rapides (picoseconde), et l'indice  $i$  représente le nombre de termes exponentiels dans la fonction de corrélation et peut varier de -2 à +2. Dans un cas simple de rotation isotrope,  $k=0$  et  $\tau_i$  est

remplacé par le temps de corrélation global  $\tau_m$ , qui est semblable à la composante isotrope de diffusion  $\mathcal{D}_{iso}$ :

$$\frac{1}{\tau_m} = 6\mathcal{D}_{iso}$$

En absence de diffusion isotrope ( $k=2$ ), un tenseur de diffusion, qui adopte la forme d'un sphéroïde, est défini. Celui-ci peut adopter une forme oblate ou prolata (Figure 1.13).

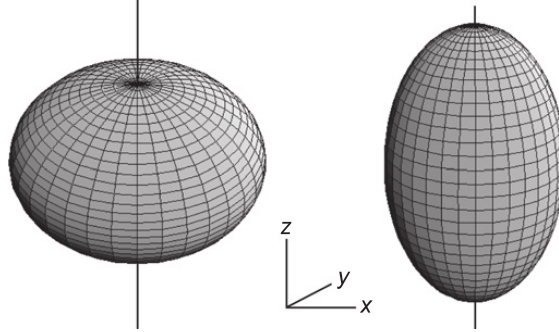


Figure 1.13: Tenseur de diffusion de symétrie axiale. La figure montre les tenseurs de diffusion pour des sphéroïdes oblat (gauche) et prolata (droite). Le sphéroïde oblat montre un tenseur de diffusion plus petit sur l'axe des z que dans le plan xy. À l'opposé, le sphéroïde prolata possède un tenseur de diffusion plus petit sur les axes xy que sur l'axe z. Les tenseurs de diffusion pour l'ellipsoïde (non montrée) sont tous différents, alors qu'ils sont identiques pour la sphère. L'image a été modifiée de la référence [31].

Des modèles "model-free" locaux peuvent être créés à partir des paramètres internes  $S^2$  et  $\tau$ :

$$m(0): \{ \}$$

$$m(1): \{ S^2 \}$$

$$m(2): \{ S^2, \tau_e \}$$

$$m(3): \{ S^2, R_{ex} \}$$

$$m(4): \{ S^2, \tau_e, R_{ex} \}$$

$$m(5): \{ S^2, S_f^2, \tau_s \}$$

$$m(6): \{ S^2, \tau_f, S_f^2, \tau_s \}$$

$$m(7): \{ S^2, S_f^2, \tau_s, R_{ex} \}$$

$$m(8): \{ S^2, \tau_f, S_f^2, \tau_s, R_{ex} \}$$

$$m(9): \{ R_{ex} \}$$



Le paramètre d'ordre se définit comme la combinaison des paramètres d'ordre à différentes échelles de temps:

$$S^2 = S_f^2 \cdot S_s^2$$

La valeur  $S^2$  varie entre 0 (mouvement isotrope non restreint) et 1 (absence de mouvement). Dans le formalisme model-free, le  $R_{ex}$  permet de compenser l'écart causé entre les valeurs expérimentales et d'extraire  $R_2$  lorsque des mouvements à l'échelle de temps de la  $\mu$ s-ms sont présents.

En assumant que la densité spectrale suit la forme Lorentzienne, la fonction de densité spectrale du model-free peut se définir comme suit:

$$J(\omega) = \frac{2}{5} \sum_{i=-k}^k w_i \cdot \tau_i \left( \frac{S^2}{1 + (\omega\tau_i)^2} + \frac{(1 - S_f^2)(\tau_f + \tau_i)\tau_f}{(\tau_f + \tau_i)^2 + (\omega\tau_f\tau_i)^2} + \frac{(S_f^2 - S^2)(\tau_s + \tau_i)\tau_s}{(\tau_s + \tau_i)^2 + (\omega\tau_s\tau_i)^2} \right)$$

Dans le cas d'une diffusion isotrope globale, où  $i=0$ ,  $w=1$  et  $\tau_i = \tau_m$ , la fonction peut être réduite à l'expression suivante:

$$J(\omega) = \frac{2}{5} \tau_m \left( \frac{S^2}{1 + (\omega\tau_m)^2} + \frac{(1 - S_f^2)(\tau_f + \tau_m)\tau_f}{(\tau_f + \tau_m)^2 + (\omega\tau_f\tau_m)^2} + \frac{(S_f^2 - S^2)(\tau_s + \tau_m)\tau_m}{(\tau_s + \tau_m)^2 + (\omega\tau_s\tau_m)^2} \right)$$

À partir de la densité spectrale, les valeurs  $R_1$ ,  $R_2$  et  $NOE$  sont extraites et comparées aux valeurs expérimentales pour la sélection du modèle.

Plusieurs programmes ont été développés pour l'optimisation des paramètres "model-free". Ces programmes incluent *ModelFree* [38, 39], *FAST ModelFree* [40], *DASHA* [41], *Tensor2* [42, 43], *DYNAMICS* [44], *relax* [45, 46], et plusieurs autres. Dans le contexte de cette thèse, seul le programme *relax* a été utilisé.

### 1.5.6.3 Attribution des résonances d'une protéine

L'attribution des résonances en fonction des déplacements chimiques permet l'identification de chacun des pics dans le spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC. Chaque pic représente un vecteur amide de la chaîne principale de la protéine, et donc d'un acide aminé en particulier (Figure 1.14). Chaque acide aminé est alors identifié par une résonance correspondant au couplage entre le déplacement chimique de son proton en x et celui de son azote en y. Fait à noter, les prolines ne sont pas visibles sur ce spectre bidimensionnel puisqu'elles ne possèdent pas de groupement amide.

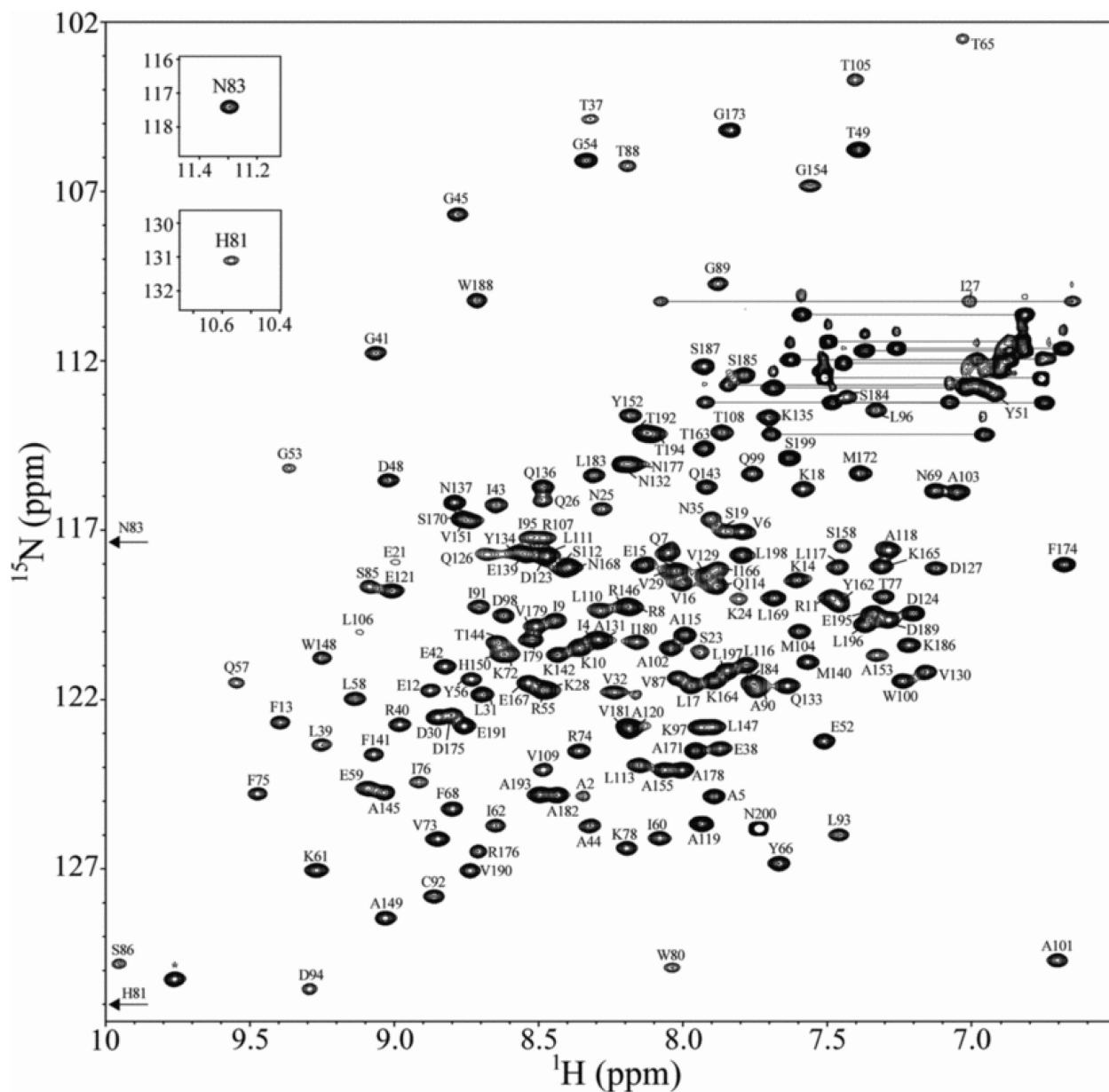
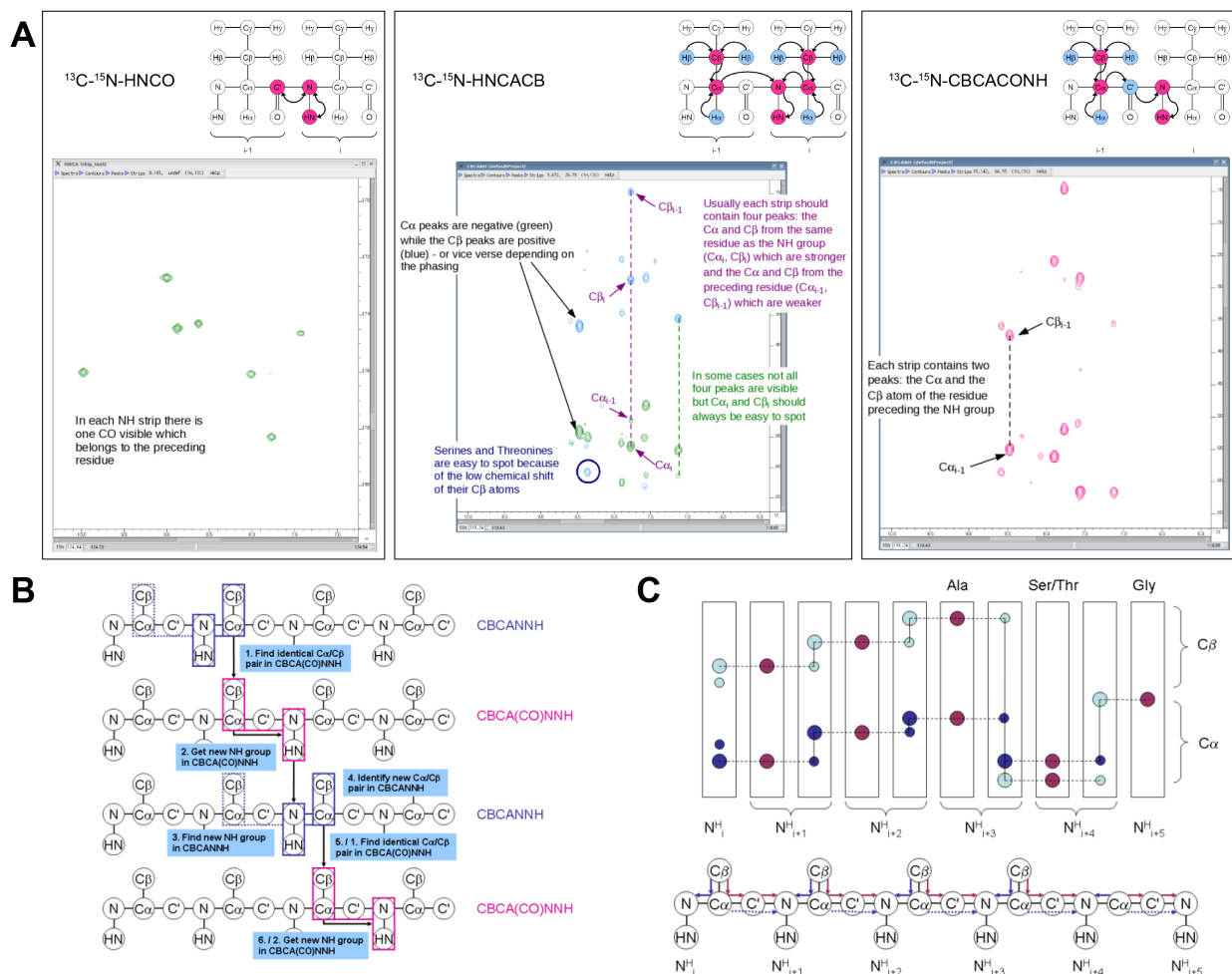


Figure 1.14: Attribution des résidus de la chaîne principale en  $^1\text{H}$  et  $^{15}\text{N}$  de la protéine HIP2. Vue du spectre  $^1\text{H}$ - $^{15}\text{N}$ -HSQC. Les résidus sont identifiés par le code à une lettre suivi du numéro dans la séquence. L'image a été tirée de la référence [47].

L'attribution séquentielle est la méthode la plus couramment utilisée pour l'attribution des résidus. Elle consiste en une série d'expériences qui permettront de corréler les atomes des acides aminés  $i$  (celui auquel appartient le NH) et  $i-1$  (le résidu précédant le NH). Les expériences les plus courantes sont HNCO, HNCACB et CBCA(CO)NH (Figure 1.15). Pour effectuer ces expériences, la protéine devra être marquée en  $^1\text{H}$ ,  $^{13}\text{C}$  et  $^{15}\text{N}$ . Le transfert de magnétisation se fait sur plusieurs atomes, représentés en rose (bleu) pour les atomes qui seront visibles (invisibles) dans le spectre. La corrélation des atomes visibles dans les trois

spectres permettra d'identifier les carbones  $C_{\alpha}$  et  $C_{\beta}$  associés au groupement NH de chacun des résidus dans la séquence. Le déplacement chimique de ces deux carbones permet d'identifier le type de résidu, et la connaissance de la structure primaire de la protéine permet l'identification du résidu en particulier. Même s'il est toujours possible d'exécuter cette tâche manuellement, il existe un bon nombre de logiciels qui permettent à l'utilisateur d'accélérer et de lui simplifier le travail.

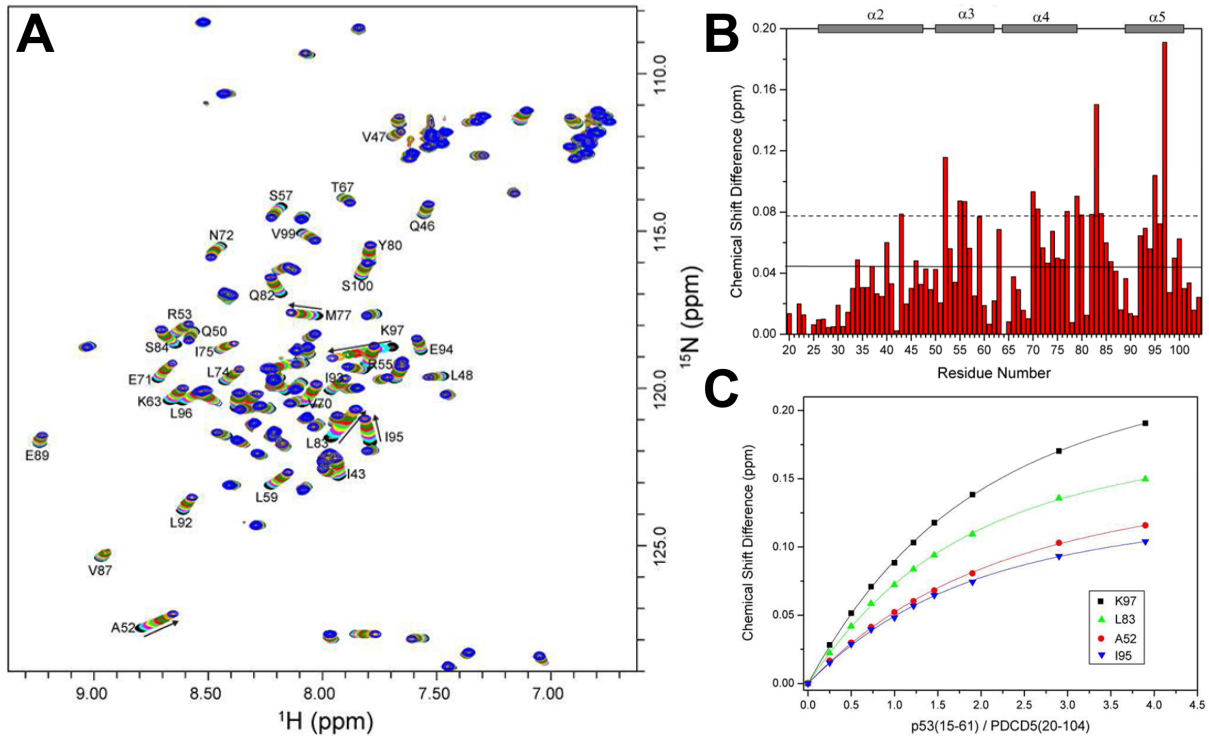


**Figure 1.15: Spectres RMN permettant l'attribution des résonances d'une protéine. (A) Description des spectres HNCO, HNCACB et CBCA(CO)NH. (B) Présentation de l'attribution séquentielle des résidus en six étapes, en utilisant les spectres HNCACB et CBCA(CO)NH. (C) Identification séquentielle des résidus dans la séquence. Les images ont été tirées du site <http://www.protein-nmr.org.uk/solution-nmr/spectrum-descriptions/>.**

#### 1.5.6.4 Titrage RMN

La fréquence de résonance d'un atome demeure inchangée pourvu que son environnement chimique ne soit pas modifié. Un changement dans l'environnement chimique peut être causé par une modification de la solution tampon (pH, sels, etc.), de la température

d'expérimentation, ou encore par la présence d'un ligand. Le titrage RMN tire profit de cet "inconvenient", en permettant d'aller chercher de l'information sur les résidus dynamiquement affectés par la liaison à des ligands. Ainsi, ces expériences de titrage permettent de caractériser les résidus qui sont en contact direct avec un ligand et/ou les résidus loin du site actif qui sont toutefois affecté par la liaison d'un ligand à la protéine. En plus de fournir des informations sur les résidus fonctionnels qui pourraient potentiellement être ciblés par mutagenèse, ces expériences de titrage peuvent donner des informations sur les potentiels sites allostériques distants, mais néanmoins fonctionnels. Le titrage RMN est effectué en suivant les déplacements chimiques observés dans un spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC lors de l'augmentation de la concentration d'un ligand d'intérêt par l'ajout progressif de celui-ci (Figures 1.16A-B).



**Figure 1.16:** Exemple d'une expérience de titrage RMN, avec la protéine PDCD5(20-104) marquée au  $^{15}\text{N}$  et le ligand non marqué p53(15-61). (A)  $^1\text{H}$ - $^{15}\text{N}$  HSQC montrant le déplacement chimique des résidus affectés par la présence du ligand. Le code de couleur (de bleu à violet) correspond au déplacement chimique suivant l'augmentation graduelle en concentration du ligand. (B) Histogramme représentant la variation du déplacement chimique par résidu. La ligne solide correspond à la moyenne des variations de déplacement chimique. La ligne pointillée correspond à la moyenne des variations de déplacement chimique additionnée à la déviation standard. (C) Graphique montrant la variation du déplacement chimique des résidus K97, L83, A52 et I95 en fonction du rapport des concentrations ligand/protéine (p53(15-61)/PDCD5(20-104)). L'image a été modifiée de la référence [48].

Parce que la variation du déplacement chimique  $\Delta\delta$  se produit en deux dimensions, soit  $^1\text{H}$  et  $^{15}\text{N}$ , une moyenne pondérée est utilisée pour le calcul:

$$\Delta\delta = [((^1\text{H}P_0 - ^1\text{H}P_{sat})^2 + (^{15}\text{N}P_0 - ^{15}\text{N}P_{sat})^2/25)/2]^{1/2}$$

où  $P_0$  et  $P_{sat}$  représente les positions dans le spectre (en ppm) du résidu libre et lié à saturation. Une constante d'affinité  $K_a$  peut être extraite en mesurant la différence observée ( $\Delta_{OBS}$ ) entre les intensités de chaque point et l'intensité du pic initial par rapport à l'intensité maximale ( $\Delta_{MAX}$ ; correspondant à la différence d'intensité entre le premier et le dernier pic). Un graphique correspondant à l'intensité de chaque point en fonction du rapport de la concentration du ligand (L) sur la concentration de la protéine (P) peut être préparé. Si on assume un rapport stoechiométrique de 1:1, le rapport  $\Delta_{OBS} / \Delta_{MAX}$  augmentera entre 0 et 1, et donnera la fraction liée de la protéine (FB; Fraction Bound). En calculant le rapport  $\Delta_{OBS} / \Delta_{MAX}$  en fonction de L/P, il est possible d'extraire  $K_a$  (Figure 1.16C):

$$\frac{\Delta_{OBS}}{\Delta_{MAX}} = FB = \frac{(1 + (L/P) + (K_a/P)) - \left( \left( 1 + (L/P) + (K_a/P) \right)^2 - 4(L/P) \right)^{1/2}}{2}$$

#### 1.5.6.5 Déplacement en collectivité

Les protéines existent dans un ensemble de conformations en solution. Lors de la liaison d'un ligand, il se produit un changement dans la distribution de la population. L'étude de l'équilibre des états conformationnels est importante pour la compréhension du mode de transmission de l'information à travers une protéine. Le changement conformationnel est accompagné par un déplacement en collectivité de certains résidus, formant un réseau d'interactions "longue distance" à l'intérieur de la protéine. Au cours des dernières années, plusieurs approches permettant de quantifier la direction et l'intensité du déplacement chimique ont vu le jour [49-52]. Par titrage RMN, il est possible d'observer et de quantifier individuellement ce déplacement chimique suite à la liaison d'un ligand en calculant le  $\cos(\theta)$  entre les déplacements des résonances (Figure 1.17):

$$\cos(\theta) = \frac{\vec{A} \cdot \vec{B}}{|\vec{A}| |\vec{B}|}$$

Ainsi, un  $\cos(\theta) \approx 1$  signifie que le déplacement des résidus est coordonné. De façon similaire, l'intensité du déplacement est déterminée comme suit:

$$\chi = \frac{\vec{A} \cdot \vec{B}}{|\vec{B}|^2}$$

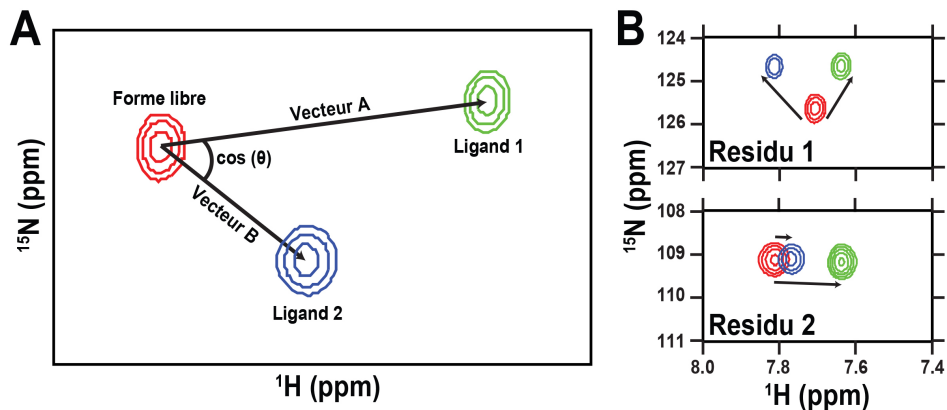


Figure 1.17: Analyse de projection du déplacement chimique (CHESPA). (A) Représentation graphique de l'approche CHESPA, décrite dans la référence [52]. La position  $^1\text{H}$ - $^{15}\text{N}$  d'un résidu sous forme libre est représentée par le pic rouge, alors que les formes liées de l'enzyme sont représentées par les pics vert (ligand 1) et bleu (ligand 2). Les flèches indiquent le déplacement chimique en  $^1\text{H}$ - $^{15}\text{N}$  (longueur et direction) pour chaque pic de son origine (en rouge) jusqu'en position saturée (bleu et vert). (B) Représentation d'une analyse de projection de deux résidus se déplaçant de façon indépendante (panneau du haut), ou de manière coordonnée (panneau du bas). L'image a été modifiée de la référence [52].

## 1.6 Hypothèses et objectifs

### 1.6.1 Hypothèses

L'importance des mouvements atomiques a déjà été montrée dans plusieurs systèmes protéiques [10-12, 53-62]. Toutefois, personne n'a encore démontré si ces mouvements, dans l'ordre de la  $\mu\text{s}$ - $\text{ms}$ , sont conservés pour des protéines de structure et de fonction similaires. En effet, si ces mouvements non-aléatoires de résidus clés sont essentielles à l'activité catalytique, nous émettons l'hypothèse qu'elles soient évolutivement sélectionnées comme faisant partie intégrante de la relation structure-fonction chez des homologues structuraux et fonctionnels. En conséquence, nous émettons les hypothèses suivantes:

- **Hypothèse no 1:** Les réseaux de résidus flexibles formés d'une hiérarchie de secteurs dynamiques contigus sont présents chez des enzymes structurellement et fonctionnellement homologues. Ces mouvements atomiques à l'échelle de temps de la  $\mu\text{s}$ - $\text{ms}$  sont encodées dans la séquence et sont liées à l'activité catalytique.
- **Hypothèse no 2:** Pour des homologues structuraux de fonctions similaires, ces réseaux dynamiques sont conservés à travers l'évolution. En effet, si on émet l'hypothèse que les mouvements sont essentielles à l'activité catalytique de l'enzyme, elles seront conservées au même titre que la séquence et la structure.

- **Hypothèse no 3:** La perturbation de ces réseaux dynamiques conservés à l'échelle de temps de la  $\mu$ s-ms affecte l'activité catalytique des enzymes.

### 1.6.2 Objectifs

L'objectif principal du présent projet vise la caractérisation de la dynamique fonctionnelle des RNases humaines par  $^{15}\text{N}$ -CPMG. Afin de démontrer les hypothèses mentionnées ci-haut, nous étudierons et comparerons les enzymes dans leur forme libre ou liée, et nous préparerons des mutants fonctionnels et/ou dynamiquement affectés.

- **Objectif no 1:** Étude de la conservation et de l'importance fonctionnelle de la dynamique chez les RNases humaines homologues. Pour ce faire, nous caractériserons les événements dynamiques des RNases à l'échelle de temps de la  $\mu$ s-ms par la technique de Carr-Purcell-Meiboom-Gill ( $^{15}\text{N}$ -CPMG). L'objectif vise à démontrer la présence de réseaux dynamiques conservés chez des homologues structuraux.
- **Objectif no 2:** Analyse fonctionnelle de certains résidus sur la reconnaissance des substrats et ligands. Par l'utilisation de la technique de titrage RMN, nous étudierons l'effet de la liaison d'un ligand sur les événements dynamiques.
- **Objectif no 3:** Perturbations des réseaux dynamiques. Nous utiliserons des techniques de mutagenèse dirigée dans le but de démontrer que les perturbations dynamiques d'homologues structuraux et fonctionnels affectent leur fonction biologique et catalytique.

Dans le but d'examiner la relation structure-séquence-dynamique, nous nous sommes arrêtés à l'étude des RNases pancréatiques humaines. Des études précédentes ont démontré que la RNase A dépend de mouvements atomiques concertés pour entreprendre les différentes étapes de sa réaction catalytique [63, 64]. En raison de leur très forte homologie structurale et de la conservation de leur activité catalytique, la famille des RNases humaines s'avère particulièrement intéressante pour tester les hypothèses ci-haut mentionnées. Les sections qui suivent offriront des détails sur les membres composant la famille RNases humaines.

## 1.7 Les ribonucléases pancréatiques : un système modèle

### 1.7.1 La RNase A

La RNase A (EC 3.1.27.5), une enzyme pancréatique de *Bos taurus*, a été cristallisée pour la première fois en 1939 par Kunitz [65, 66]. Elle a été la première enzyme à voir sa séquence identifiée [67, 68], et la troisième à voir sa structure tridimensionnelle résolue [69]. La RNase A est à l'origine de nombreuses études, que ce soit sur la stabilité, le repliement, la chimie des protéines ou l'enzymologie. Trois prix Nobel de chimie ont été octroyés en 1972 pour leur étude sur la RNase A, soit à Christian B. Anfinsen [70], Standord Moore [71], et William H. Stein [71], pour leurs travaux sur la relation entre la séquence et la structure. Un quatrième prix Nobel a été décerné en 1984 à Robert Bruce Merrifield [72], pour son développement de la méthode de synthèse chimique sur matrice solide.

#### 1.7.1.1 Structure de la RNase A

La RNase A est une petite protéine de 124 acides aminés, avec un poids moléculaire de 13.7 kDa [73]. Sa structure tridimensionnelle a été résolue et publiée à plusieurs reprises, dans sa forme libre ou liée. Prenant la forme d'un rein ("kidney-type structure"), la RNase A possède 4 longs brins  $\beta$  antiparallèles et 3 courtes hélices  $\alpha$  (Figure 1.18). Quatre ponts disulfures (Cys26-Cys84, Cys58-Cys110, Cys40-Cys95, Cys65-Cys72) permettent de stabiliser la structure. Avec une dimension approximative de 3.5 x 4.5 x 3.1 nm, la protéine possède une cavité logeant le site actif et une triade de résidus catalytiques: His12, His119 et Lys41 [74].

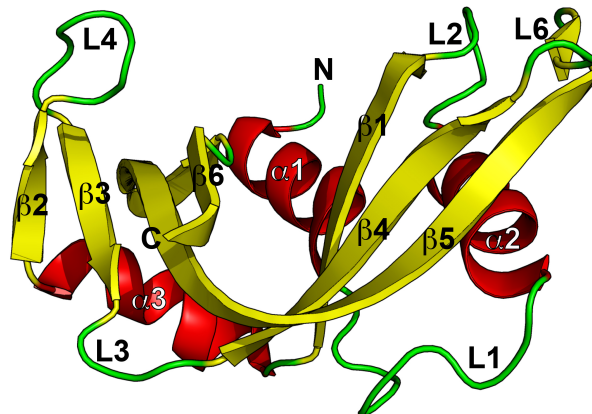
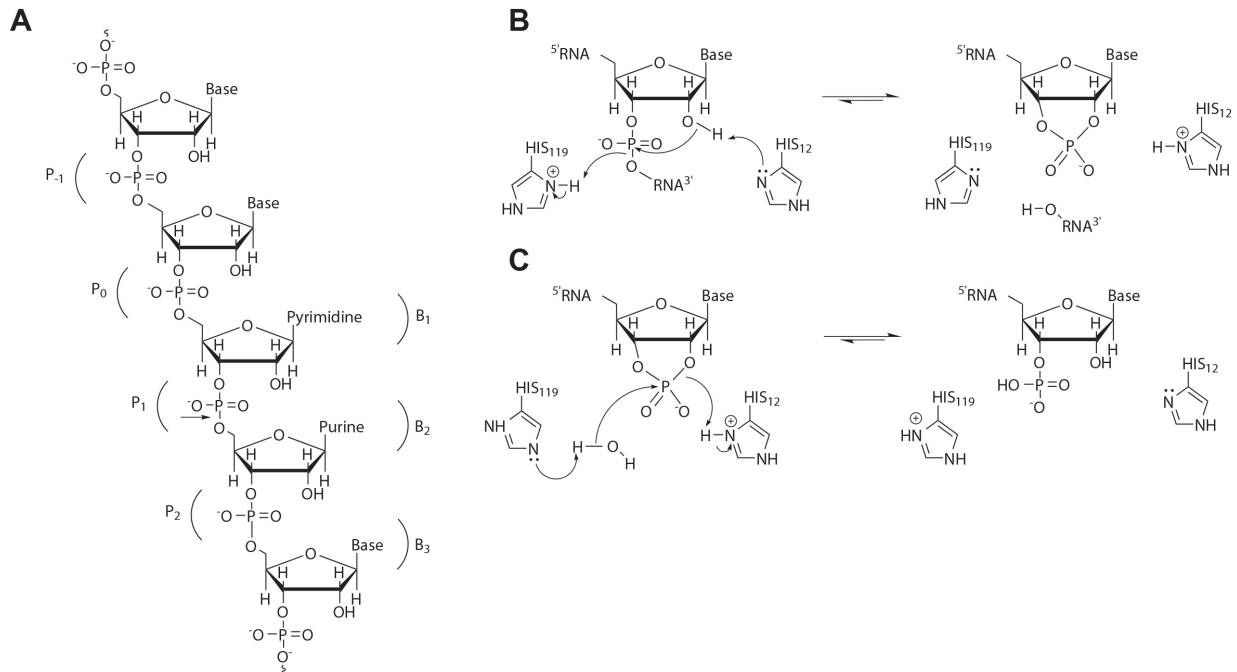


Figure 1.18: Représentation schématique de la structure tertiaire de la RNase A. Les structures secondaires sont identifiées selon la nomenclature de la référence [75], de même que les extrémités C- et N-terminales. Les différentes structures secondaires sont identifiées par couleur: Brins  $\beta$  (jaune); hélices  $\alpha$  (rouge); boucle (vert). La structure a été produite à l'aide du logiciel PyMOL en utilisant les données structurales de la Protein Data Bank (PDB 7RSA [75]).



### 1.7.1.2 Liaison du substrat et catalyse enzymatique

La RNase A possède une activité d'endoribonucléase qui permet de dégrader l'ARN en catalysant le clivage de la liaison P—O<sub>5</sub> de l'ARN simple brin situé après une pyrimidine. Dans le mécanisme catalytique (Figure 1.19B), His12 agit comme une base en acceptant un proton de l'oxygène en 2' pour faciliter l'attaque nucléophile de l'oxygène sur le phosphodiester du substrat. Pendant ce temps, His119 agit comme un acide en donnant un proton à l'oxygène en 5', favorisant ainsi la formation d'un intermédiaire 2',3'-cyclique. Cette réaction de transphosphorylation est assistée par la Lys41, qui permet de stabiliser l'état de transition [76, 77]. Une réaction d'hydrolyse (Figure 1.19C), où les His12 et His119 agissent comme acide et base respectivement (rôle inversé), permet de ramener l'enzyme à son état initial.



**Figure 1.19. Mécanisme catalytique de la famille des RNases. (A) Représentation d'un fragment d'ARN et identification des sous-sites permettant sa liaison à la RNase A. Le lien scissile est représenté par une flèche en P<sub>1</sub>. (B) Le résidu His12 agit à titre de base générale alors que His119 agit à titre d'acide général dans la première partie de la réaction (transphosphorylation). (C) Les rôles sont ensuite inversés pour la seconde partie de la réaction (hydrolyse), permettant de régénérer l'enzyme libre et active. L'image a été tirée de la référence [78].**

Plusieurs études ont démontré que la liaison de l'ARN s'étendait bien au-delà du lien scissile. Il a été démontré que la  $V_{max}$  d'une chaîne d'acide polyuracique (Poly(U)) est de 3 à 20 fois supérieur à celui obtenu avec une chaîne de 3-4 nucléotides [79]. En effet, il a été établi que durant les premières étapes de la réaction, l'enzyme favorise la liaison des fragments qui ont de préférence un minimum de 6-7 nucléotides à l'une des extrémités du lien scissile [74]. Cette liaison implique un certain nombre de résidus regroupés dans des sous-sites (Figure

1.19A et Tableau 1.1) [80-82]. Trois de ces sous-sites interagissent avec la base ( $B_1$ ,  $B_2$  et  $B_3$ ), alors que les autres sous-sites se lient plutôt au groupement phosphate ( $P_{-1}$ ,  $P_0$ ,  $P_1$  et  $P_2$ ) [74, 83]. Les sous-sites confèrent la spécificité du substrat pour l'enzyme. Ainsi, la RNase A préfère la liaison d'une pyrimidine en  $B_1$ , et d'une purine en  $B_2$ .

Tableau 1.1: Interactions entre les sous-sites de la RNase A et l'ARN [84, 85].

<i>Sous-sites</i>	<i>Résidus</i>
B <sub>1</sub>	Thr45, Asp83
B <sup>2</sup>	Asn71, Glu111
B <sub>3</sub>	?
P <sub>-1</sub>	Arg85
P <sub>0</sub>	Lys66
P <sub>1</sub>	Gln11, His12, Lys41, His119, Asp121
P <sub>2</sub>	Lys7, Arg10

La tableau 1.2 présente une liste des paramètres cinétiques provenant de la littérature pour la coupure de fragments d'ARN. Des études ont démontré que l'enzyme agit de façon non processive, c'est-à-dire qu'elle relâche le produit de la réaction pour lier un nouvel ARN, plutôt que de poursuivre la catalyse du même brin [86].

Plus de 250 séquences complètes chez les vertébrés possèdent un profil typique, caractéristique des membres de la superfamille des RNase A:

- Des séquences d'environ 130 acides aminés non interrompues par des introns;
- Une séquence signal pour la sécrétion en N-terminal;
- De six à huit cystéines formant des ponts disulfures;
- Trois résidus catalytiques, composés de deux histidines et d'une lysine;
- Un motif "CKXXNTF" auquel appartient la lysine catalytique [87].

Il est aussi à noter que les membres de la superfamille des RNase A ne partagent aucune homologie de séquences et de structure avec les familles des RNase H, RNase III et RNase P.

Tableau 1.2: Bref résumé des paramètres cinétiques pour la coupure d'ARN chez les membres de la famille des RNases humaines et chez la RNase A. poly(U): chaîne d'acide polyuracidique; poly(A): chaîne d'acide polyadénylique; poly(C): chaîne d'acide polycytidique; UpA: uridylyl (3',5')adénosine; CpA: cytidylyl (3',5')adénosine; ytRNA: Yeast Transfer RNA (ARN de transfert de levure).

<i>Substrat</i>	$k_{cat}$ ( $s^{-1}$ )	$K_M$ (mM)	$k_{cat}/K_M$ ( $10^3 M^{-1}s^{-1}$ )	<i>Ref.</i>
<b><u>RNase A</u></b>				
poly(U)	24 ± 15	0.06 ± 0.01	400 ± 300	[86]
poly(C)	21×10 <sup>4</sup> ± 3×10 <sup>4</sup>	0.033 ± 0.007	636 ± 69	[88]
poly(A)	0.023 ± 0.001	0.08 ± 0.009	0.28 ± 0.04	[86]
UpA	1400 ± 150	0.62 ± 0.09	2300 ± 400	[86]
<b><u>RNase 1</u></b>				
poly(C)	14×10 <sup>4</sup> ± 2×10 <sup>3</sup>	0.10 ± 0.013	1480 ± 110	[88]
<b><u>RNase 2</u></b>				
poly(U)	249 ± 12	53 ± 4	4.8 ± 0.5	[89]
poly(C)	170 ± 4	80 ± 7	2.125	[90]
ytRNA	31 ± 1	0.003 ± 0.0002	9340 ± 170	[89]
<b><u>RNase 3</u></b>				
poly(U)	36 ± 0.4	253 ± 10	0.14 ± 0.005	[91]
UpA	6.2 ± 0.4	5.4 ± 0.5	1.150	[92]
CpA	4.2 ± 0.3	2.4 ± 0.3	1.750	[92]
<b><u>RNase 4</u></b>				
UpA	Données non disponibles		250	[93]
CpA	Données non disponibles		0.63	[93]
<b><u>RNase 5</u></b>				
UpA	Données non disponibles		0.0012 ± 0.0002	[94]
CpA	Données non disponibles		0.0146 ± 0.0003	[94]
<b><u>RNase 6</u></b>				
ytRNA	0.13	0.005	26.0	[95]
<b><u>RNase 7</u></b>				
Aucune donnée comparative				
<b><u>RNase 8</u></b>				
Aucune donnée comparative				

## 1.7.2 La famille des RNases pancréatiques humaines

Chez l'humain, on dénombre huit RNases qui possèdent une activité ribonucléolytique [96, 97]. Tous les membres conservent les éléments principaux de la séquence en acides aminés, incluant les 4 ponts disulfures (sauf pour l'angiogénine, qui n'en possède que 3) et la triade de résidus catalytiques (Figure 1.20).

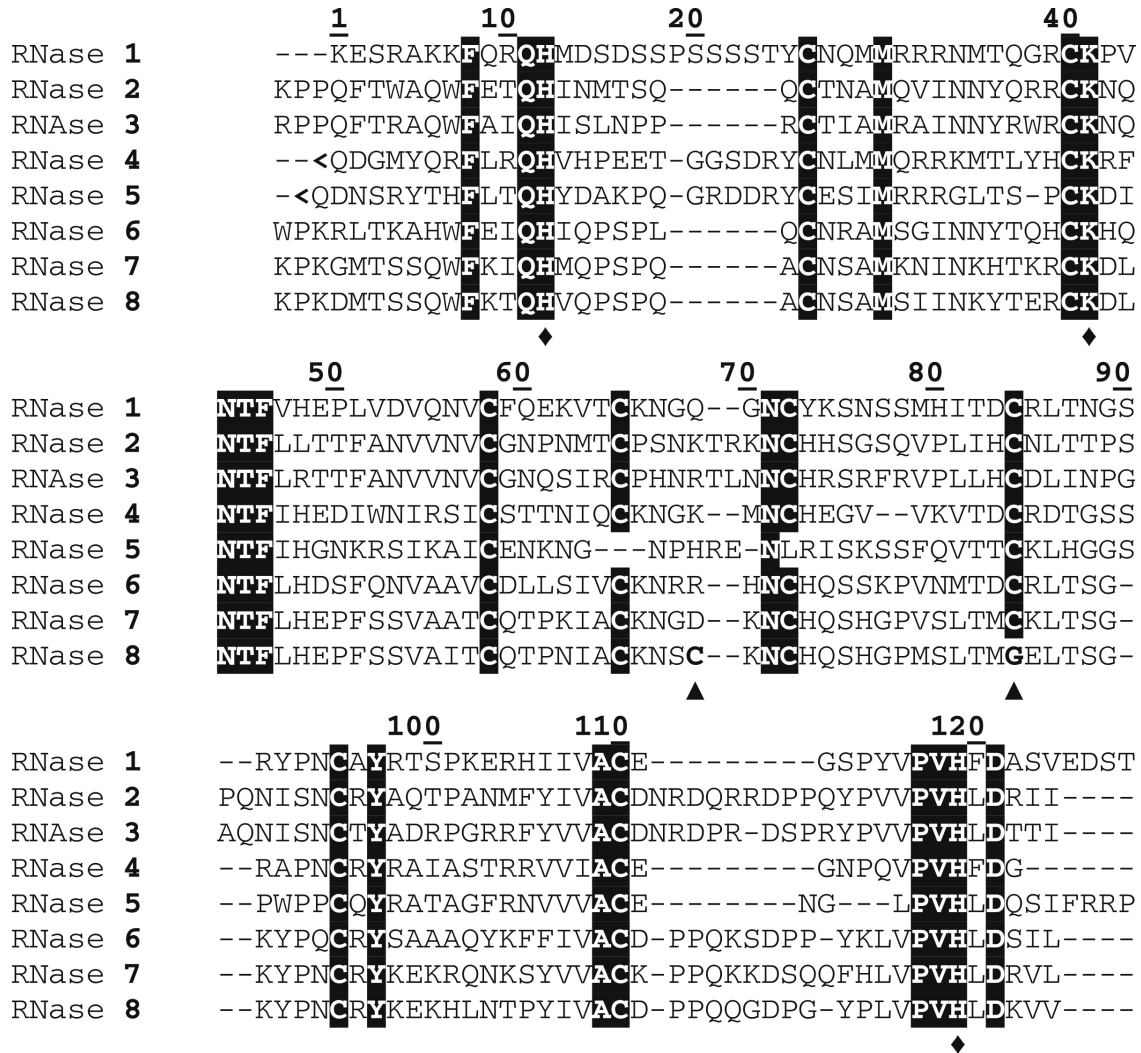
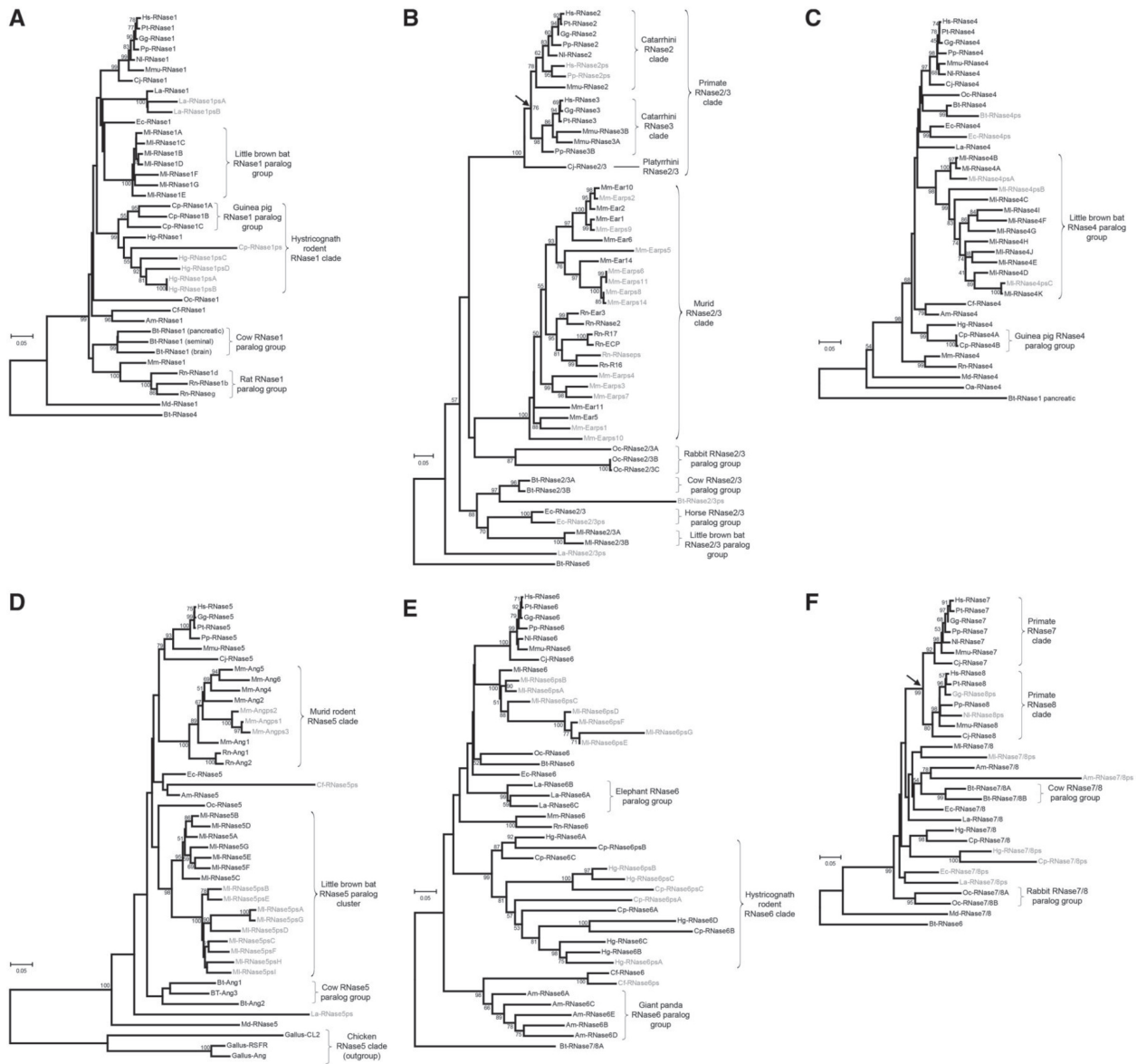


Figure 1.20. Alignement de séquences des huit RNases humaines. Les résidus conservés à 100% sont encadrés en noir. Le caractère (◆) indique la position dans la séquence des résidus catalytiques. Le changement de position de la cystéine en position 84 vers la position 69 chez la RNase 8 est indiqué par le caractère (▲). L'acide pyroglutamique des RNase 4 et RNase 5 est représenté par le caractère (<). L'alignement de séquence a été généré avec CLUSTALW2. L'image a été tirée de la référence [97].



**Figure 1.21.** Arbre phylogénétique montrant l'évolution des gènes et pseudogènes des RNases chez les vertébrés. Les RNases ont été subdivisées en 6 groupes représentant les RNases humaines (A) 1/A, (B) 2/3, (C) 4, (D) 5 (angiogénine), (E) 6 et (F) 7/8. Les pseudogènes sont indiqués en caractères gris. Les flèches indiquent le temps de répliation des RNases 2/3 et 7/8. Les valeurs numériques situées dans l'arbre représentent le niveau de confiance ("bootstrap") de l'estimation. L'image a été tirée de la référence [98].

Ces homologues structuraux se distinguent par le fait que leur activité catalytique varie selon un large spectre (Tableau 1.2), et que des fonctions biologiques ont été attribuées à plusieurs d'entre eux. Parce qu'ils sont aussi largement distribués dans plusieurs tissus, organes et fluides corporels, on retrouve plusieurs variantes de la même protéine avec différents profils de glycosylation.

D'un point de vue évolutif, l'analyse phylogénétique suggère que les RNases 2 (*Homo sapiens* RNase 2) et 3 (*Homo sapiens* RNase 3), de même que les RNases 7 (*Homo sapiens*

RNase 7) et 8 (*Homo sapiens* RNase 8) sont très rapprochées (Figure 21). Exception faite de ces deux groupes, les RNases humaines sont évolutivement plus proches de leur comparatif animal.

#### 1.7.2.1 *RNase 1 (Pancreatic Human Ribonuclease 1; HPR1)*

Chez l'humain, la RNase 1 est la plus proche parente de la RNase A, avec 70% d'homologie de séquence. Outre le pancréas, on retrouve l'enzyme dans de nombreux endroits, incluant l'urine [99], le cerveau [100], le sperme [101] et les reins [102]. À cause de cette diversité, différents niveaux de glycosylation ont été observés, dépendamment du tissu [103, 104]. Cependant, on note la présence de quelques caractéristiques structurales et fonctionnelles uniques à la RNase 1 [105, 106]: 1- forte activité ribonucléolytique pour l'ARN double brin; 2- Une extension C-terminale de 4 résidus EDST; 3- Une activité cytotoxique contre plusieurs lignées cellulaires immortelles de mammifères. L'enzyme a une préférence pour les chaînes de poly(C) plutôt que pour le poly(U), alors que son pH optimal se situe aux environs de 8.0 [107]. La RNase 1 est l'enzyme présentant la plus forte activité ribonucléolytique dans la famille des RNases humaines.

#### 1.7.2.2 *RNase 2 (Eosinophil Derived-Neurotoxin; EDN)*

La RNase 2 a été isolée de la rate [108], de cellules éosinophiles [109], du foie [110], du placenta [111], de l'urine [99] et des reins [102]. Elle est la plus proche parente d'ECP, partageant 66% d'homologie de séquence (Figures 1.20 & 1.21). Plusieurs résidus considérés comme importants à l'activité catalytique ne sont pas conservés chez la RNase 2: Lys7, Arg10, Lys66 et la Phe120. L'enzyme préfère les chaînes de poly(U) aux chaînes de poly(C), et aucune activité catalytique n'a été observée avec le poly(A) ou l'ARN double brin. Son pH optimal se situe entre 6.5 et 7.0 pour la dégradation de l'ARN de transfert de levure [112, 113]. En plus de sa fonction ribonucléolytique, la RNase 2 agirait en tant qu'agent neurotoxique [109] et helminthotoxique [114], et posséderait une action antivirale contre le Virus Respiratoire Syncytial (VRS) et le Virus de l'Immunodéficience Humaine (VIH) [115-118].

#### 1.7.2.3 *RNase3 (Eosinophil Cationic Protein; ECP)*

La RNase 3 conserve plusieurs similarités avec la RNase 2 décrites précédemment. Toutefois, à l'opposé de sa consœur, son activité catalytique pour l'ARN de transfert de levure est très faible [113, 119]. D'un point de vue structural, la RNase 3 contient 19 arginines, soit 14% de son total en acide aminé, ce qui lui donne un point isoélectrique théorique de 10.72. Alors qu'elle conserve une l'activité neurotoxique [109], antivirale, et helminthotoxique, des

fonctions antibactérienne et antiparasitaire ont également été observées, de même qu'une activité cytotoxique contre plusieurs lignées cellulaires de mammifères [120]. Les fonctions antibactérienne et cytotoxique seraient principalement attribuables à la présence des arginines et du Trp35 [121, 122]. Selon toute vraisemblance, son activité catalytique ne serait pas requise pour sa fonction cytotoxique, une particularité spécifique à la RNase 3 [123].

#### 1.7.2.4 *RNase 4*

Avec seulement 119 acides aminés, la RNase 4 est la plus petite des huit RNases humaines. Elle s'avère aussi être la plus conservée entre les espèces (humaine, porcine et bovine), avec 90% d'homologie de séquence [124], preuve qu'il y aurait eu une pression évolutive évidente pour le maintien d'une fonction biologique toujours non identifiée. La RNase 4 démontre une activité catalytique plus importante pour les chaînes de poly(U), une caractéristique observée uniquement chez les RNases 2 et 3 [125]. Son pH optimal se situe autour de 6.0 pour la dégradation du UpA [125]. La seule modification post-transcriptionnelle observée est l'ajout d'un acide pyroglutamique, un acide aminé non standard présent dans certaines protéines par cyclisation d'un résidu d'acide glutamique donnant un lactame, sur le premier résidu [126].

#### 1.7.2.5 *RNase 5 (Angiogénine)*

La RNase 5, mieux connue sous le nom d'angiogénine, a été isolée de cellules cancéreuses du colon [127] et du plasma [128]. La protéine induit la formation de nouveaux vaisseaux sanguins [127]. Bien qu'elle possède une activité ribonucléolytique très faible, celle-ci semble essentielle à sa fonction biologique [125]. Cette faible activité catalytique serait imputable à la présence de la chaîne latérale de la Gln117 qui nuit à la liaison de la pyrimidine, ainsi qu'à l'orientation peu favorable de la région en C-terminale [129-131]. Par ailleurs, une activité antimicrobienne permettrait de croire que la RNase 5 jouerait également un rôle de défense dans l'organisme [132, 133]. Tout comme chez la RNase 4, son premier acide aminé est modifié en acide pyroglutamique [134].

#### 1.7.2.6 *RNase 6 (RNase K6)*

Parce qu'elle partage 72% d'homologie de séquence avec la RNase K2 de bovin, la 6e RNase à être découverte chez l'humain porte aussi le nom de RNase K6. Chez l'humain, sa séquence se rapproche de celle de la RNase 2, avec 47% d'identité. Toutefois, son activité catalytique contre l'ARN de transfert de levure est 40 fois inférieure à celle de la RNase 2 [95]. N'ayant pas été entièrement caractérisée, peu d'informations sont encore disponibles sur la



RNase 6. La protéine n'a toujours pas été cristallisée, mais tout porte à croire que sa structure sera similaire à celles des membres de la famille des RNases. Nos études en ce sens le démontrent, quoique non-publiées.

#### 1.7.2.7 RNase 7

La RNase 7 partage 78% d'homologie de séquence avec la RNase 8. La protéine a été isolée de tissus épithéliaux, tels que la peau, l'intestin, et les voies respiratoire et urinaire [135]. Sa structure tridimensionnelle a été résolue par RMN [136]. Elle se caractérise par la présence de 18 lysines, ce qui lui confère un point isoélectrique théorique de 9.83. Tout comme la RNase 3, elle possède une activité antimicrobienne [120, 133, 135, 137, 138]. La RNase 7 agirait en perturbant localement la membrane, un mode d'action qui diffère de celui de la RNase 3 [139]. Son activité catalytique pour l'ARN de transfert de levure est d'environ 10 fois inférieure à celui de la RNase 1 [135].

#### 1.7.2.8 RNase 8

La RNase 8 a été identifiée dans le placenta. Bien qu'évolutivement proche de la RNase 7, son activité catalytique est beaucoup plus faible [135]. Sa structure tridimensionnelle n'a toujours pas été déterminée. Toutefois, malgré la présence des 8 cystéines caractéristiques des RNases, la cystéine en position 81 est mutée par une glycine, alors qu'une nouvelle cystéine émerge en position 66 (Figure 1.20). Une activité antimicrobienne contre plusieurs bactéries à Gram positif et négatif a été identifiée [140], ce qui laisse supposer qu'elle pourrait jouer un rôle dans la défense de l'organisme. Toutefois, aucune fonction antivirale ou antibactérienne n'a été attribuée à la RNase 8 [141].

## 1.8 Autres techniques

### 1.8.1 Production d'enzymes recombinantes

Les séquences d'ADN codant pour la RNase A et des RNases humaines ont été insérées dans des vecteurs d'expressions pET (Novagen, Wisconsin, US) ou pJexpress (DNA2.0, Californie, US), et transformées chez *Escherichia coli* BL21(DE3). Cette bactérie lysogène a été transformée avec un prophage (DE3) qui contient le gène codant pour l'ARN polymérase T7, ainsi que son promoteur lac. Les systèmes pET et pJexpress utilisent l'opéron *lac*, lequel est dit "inductible", car son expression est réprimée en temps normal, mais activée sous l'action d'un inducteur, soit l'allolactose (Figure 1.22). Ainsi, en absence d'allolactose, un

métabolite du lactose, le répresseur Lac se lie à l'opérateur *lac* et empêche du même coup la production de l'ARN polymérase T7 (et donc l'induction). En présence d'allolactose, ce dernier se liera au répresseur Lac, l'empêchant de se lier à l'opérateur et donc, l'ARN polymérase T7 sera produite et l'opéron *lac* sera activé. Puisque la séquence d'intérêt est située en aval de l'opéron, elle sera produite par la même occasion. L'allolactose étant un substrat de la  $\beta$ -galactosidase, il ne peut être utilisé puisqu'il sera rapidement dégradé dans la cellule. Donc, nous utiliserons un analogue, l'isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) pour induire l'expression de notre protéine. Contrairement à l'allolactose, l'IPTG ne peut être métabolisé. Donc, sa concentration demeurera constante pendant toute la durée de l'induction.

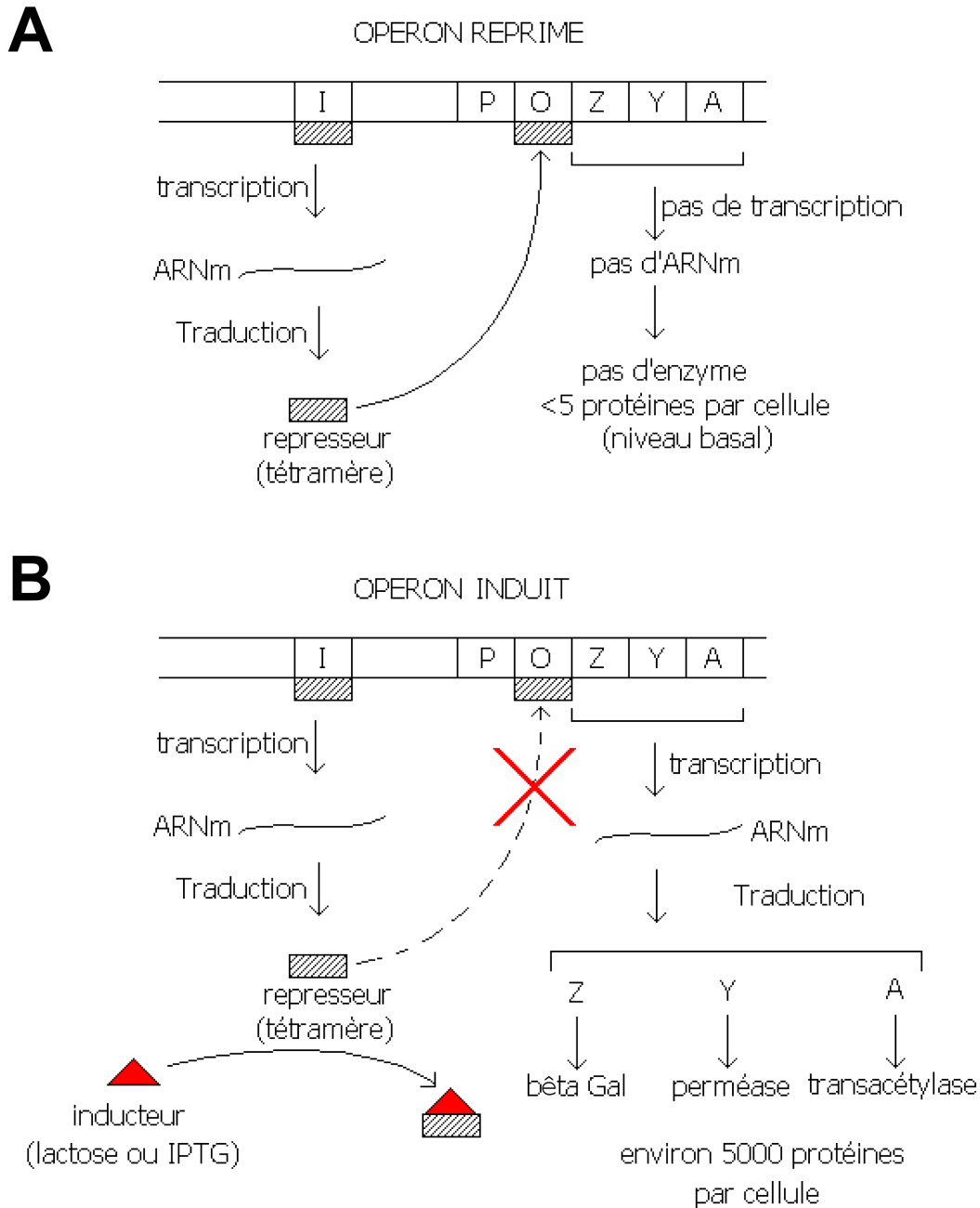
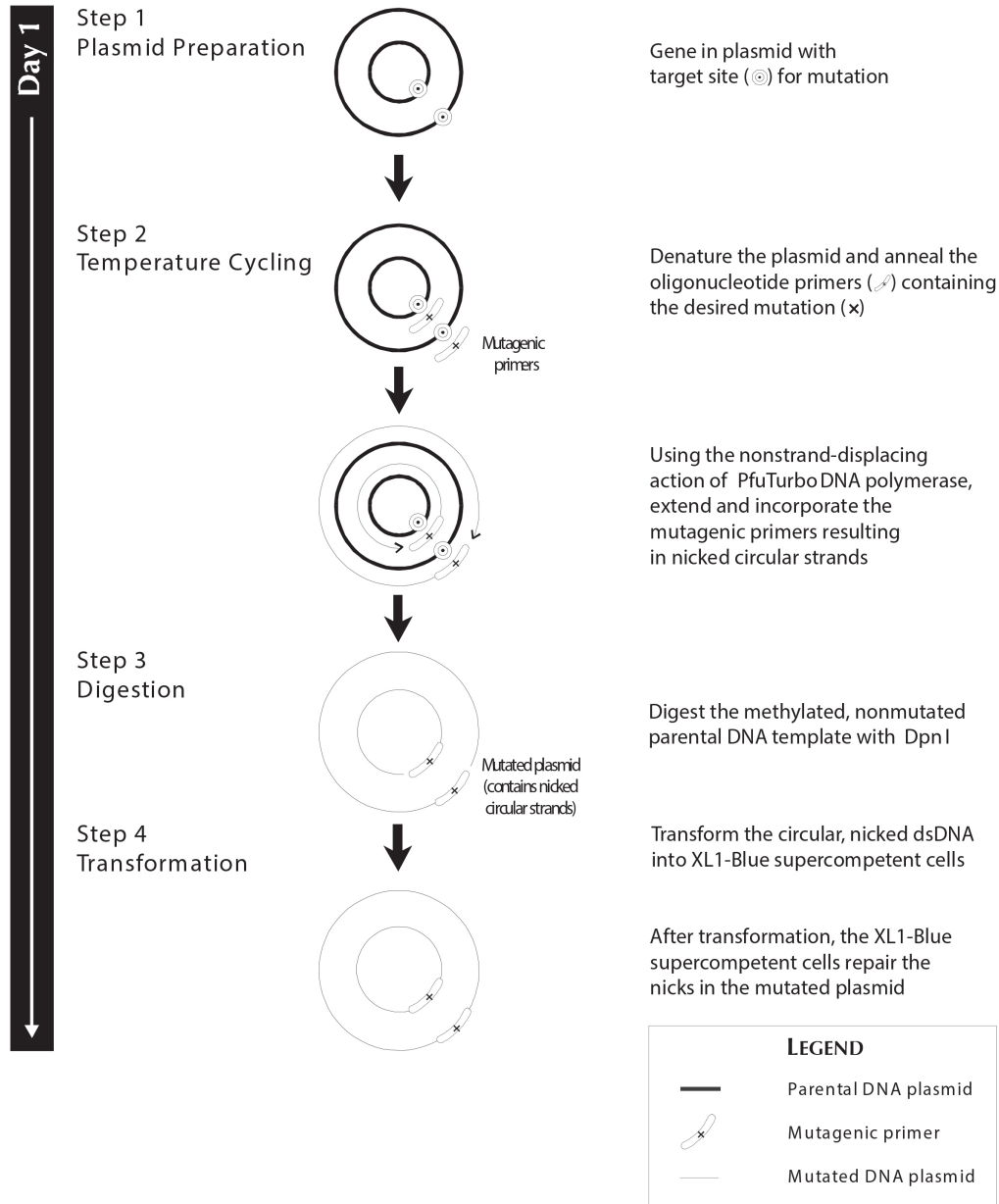


Figure 1.22. Régulation de l'opéron *lac*. (A) En l'absence d'IPTG, l'opéron est réprimé et la transcription ne se fait pas. (B) Lorsque l'IPTG est ajouté, il se fixe au répresseur, et la transcription des gènes codant pour les protéines bêta Gal, perméase et transacétylase est initiée. I: gène codant le répresseur; P: Promoteur; O: Opérateur; Z: gène codant la bêta-galactosidase; Y: gène codant pour une perméase; A: gène codant pour une transacétylase; IPTG: IsoPropylThioGalactoside, induit l'opéron lactose, mais n'est pas métabolisé. L'image a été tirée du site <http://www.cours-de-biochimie.fr/operons.php>.

### 1.8.2 La mutagenèse dirigée

La mutagenèse dirigée permet d'induire, de façon précise, une ou plusieurs mutations dans un génome. Plusieurs techniques de mutagenèse dirigée sont disponibles. Toutefois, ces méthodes sont souvent laborieuses, puisqu'elles nécessitent que l'ADN soit sous-cloné dans un vecteur M13 (ADN simple-brin) afin d'introduire la mutation, puis retransféré dans le vecteur d'expression. Dans le contexte de cette thèse, nous avons utilisé la technique du "QuikChange". Cette technique de mutagenèse dirigée permet d'éliminer les étapes de sous-clonage, simplifiant considérablement la procédure (Figure 1.23). Deux amorces complémentaires et contenant la mutation sont utilisées. Le vecteur d'ADN parental double-brin est amplifié avec la *pfu* ADN polymérase afin d'incorporer la mutation désirée. L'ADN est ensuite traité avec l'enzyme *DpnI*, laquelle digère l'ADN natif méthylé (l'ADN nouvellement amplifié n'étant pas méthylé). La nouvelle construction peut ensuite être transformée chez *Escherichia coli* BL21(DE3).



**Figure 1.23. Description des étapes de la méthode de mutagenèse dirigée QuikChange™. Deux amorces complémentaires contenant la mutation sont utilisées pour l'essai. L'ADN est amplifié avec de la PfuTurbo DNA polymérase (ou similaire) pendant 25-30 cycles. L'ADN parental est ensuite digéré avec de la DpnI. L'ADN est transformé chez des bactéries E. coli BL21(DE3) (ou équivalentes). Les colonies résistantes sont sélectionnées pour les expériences futures. L'image a été tirée du manuel d'instruction "QuikChange Site-Directed Mutagenesis Kit" produit par Stratagene (La Jolla, CA USA).**

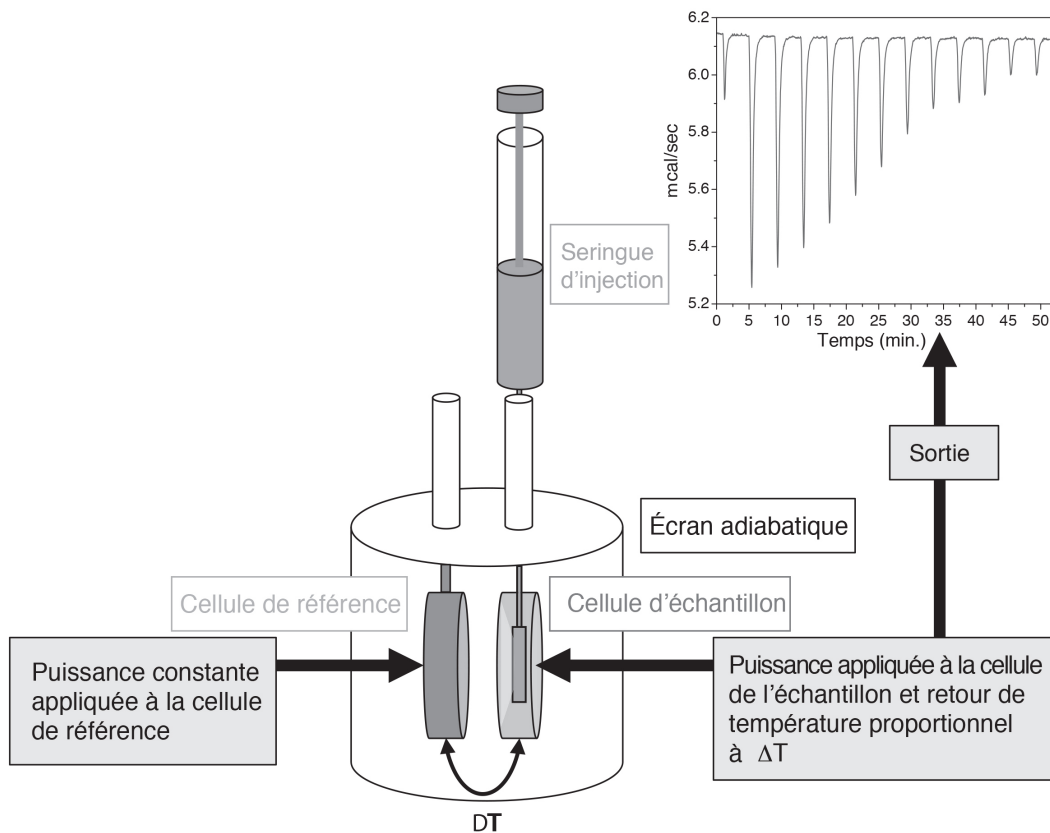
### 1.8.3 Titrage calorimétrique

L'ITC (tiré du terme anglais "Isothermal Titration Calorimetry") est une méthode biophysique utilisée pour caractériser la liaison entre un récepteur et un ligand. Le ligand peut être un substrat, un inhibiteur, une drogue, un cofacteur, une coenzyme, un groupe

prosthétique, un ion métal, un polypeptide, une protéine, un oligonucléotide, un acide nucléique, ou toute autre molécule qui pourrait interagir de façon non covalente avec le récepteur. La technique permet de mesurer directement les paramètres thermodynamiques de changements d'enthalpie ( $\Delta H$ ) et d'entropie ( $\Delta S$ ). À partir de ces valeurs, l'énergie libre de Gibbs ( $\Delta G$ ) et la constante de dissociation  $K_d$  peuvent être extraites en utilisant la relation:

$$\Delta G = RT \ln(K_d)^{-1} = \Delta H - T\Delta S$$

Où  $R$  représente la constante des gaz parfaits ( $8.314 \text{ JK}^{-1}\text{mol}^{-1}$ ),  $K_d$ , la constante de dissociation, et  $T$ , la température absolue (en Kelvin). Ainsi, une valeur élevée de  $K_d$  signifie que la liaison est faible. Cette liaison est directement dépendante de la température.



**Figure 1.24. Représentation graphique d'un ITC.** La protéine et le tampon (sans la protéine) sont insérés dans la cellule d'échantillon et la cellule de référence, respectivement. Le ligand est placé dans la "seringue" et injecté à petites doses sur la protéine. Chaque injection produit un changement de température suite à la liaison de la protéine avec le ligand, et ce changement est enregistré à la sortie. L'image a été modifiée de la référence [142].

Son fonctionnement est assez simple, mais très précis (Figure 1.24). L'ITC utilise la méthode de mesure par compensation de puissance ("Power compensation"), ou souvent appelée isotherme. La protéine dans son tampon est insérée dans la cellule d'échantillon.

Lorsque le ligand est ajouté à la protéine, il s'ensuit un changement de température. La différence de température entre la solution dans la cellule de l'échantillon et celle dans la cellule de référence, qui contient le tampon vide (sans la protéine), est détectée par l'appareil. Celle-ci augmentera ou réduira la puissance afin de ramener la température au niveau initial. La puissance utilisée par l'appareil sera reconvertie en  $\mu\text{cal/sec}$ . Donc, lors d'une réaction exothermique, il y aura un dégagement de chaleur et donc, l'appareil compensera en réduisant la puissance, ce qui donnera un pic. Pour construire une courbe de titrage, le ligand sera injecté graduellement à la protéine, suivant le retour à l'équilibre. Au fur et à mesure que la concentration du ligand va augmenter, le(s) site(s) de liaison sur la protéine sera(ont) saturé(s) et le pic suivant sera plus faible. En traçant la courbe de l'aire sous le pic (sur l'axe des y) en fonction de la concentration en ligand (l'axe des x), les paramètres cinétiques pourront en être extraits.

#### **1.8.4 Analyse spectrale par dichroïsme circulaire**

Une macromolécule présente une propriété de dichroïsme circulaire lorsqu'elle a la capacité d'absorber différemment la lumière polarisée circulaire gauche ou circulaire droite. En raison de leur chiralité, cette propriété peut être facilement exploitée chez les protéines. En effet, les structures secondaires produisent des spectres caractéristiques (Figure 1.25) qui pourront être observés à l'aide d'un appareil de dichroïsme circulaire (CD, ou "Circular Dichroism"). Par exemple, l'analyse spectrale d'un mutant permettra d'identifier des changements conformationnels par rapport au type sauvage. De manière similaire, l'analyse structurale pourra se faire en faisant varier les conditions de solvant, de température, de pH, ou de salinité. Une analyse spectrale en fonction du temps permettra de calculer la température de dénaturation ( $T_m$ ) d'une protéine. En effet, la  $T_m$  correspond à la température où la variation d'absorbance est la moitié de la valeur maximale. Même si l'analyse spectrale donne moins d'information que la diffraction des rayons X ou que la RMN, le CD permet une analyse structurale rapide des protéines.

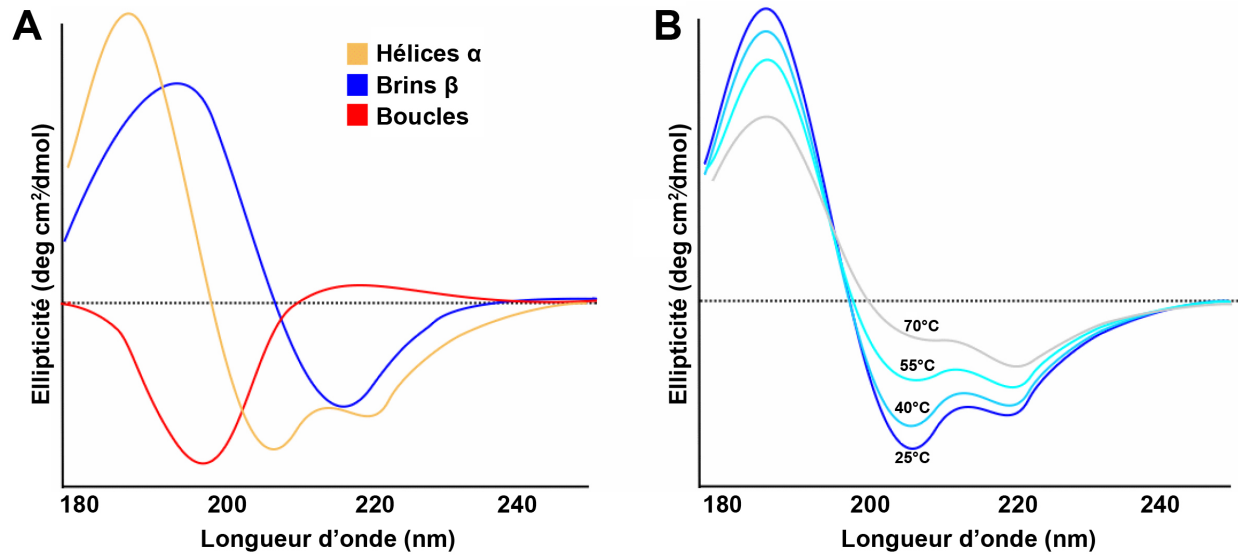


Figure 1.25. Analyse spectrale de dichroïsme circulaire. (A) Spectres caractéristiques des structures secondaires en hélices  $\alpha$ , brins  $\beta$  et boucles. (B) Changement structural observé suite à l'augmentation de la température, et conduisant à la dénaturation de la protéine. La température de dénaturation ( $T_m$ ) peut être extraite des données spectrales. L'image a été modifiée de la page web <http://www.proteinchemist.com/cd/cdspec.html>.



## **2 ARTICLE NO. 1: IMPORTANCE DES FLUCTUATIONS CONFORMATIONNELLES GLOBALES ET LOCALES SUR LA STRUCTURE ET LA FONCTION DES MEMBRES DE LA SUPERFAMILLE DE LA RNASE A**

Titre: "Structural and functional importance of local and global conformational fluctuations in the RNase A superfamily"

Auteurs: **Gagné, D.** & Doucet N.

### **2.1 Présentation de l'article scientifique**

Dans le présent article, nous avons criblé la littérature disponible à ce jour pour trouver des évidences démontrant l'importance de la dynamique fonctionnelle pour la catalyse enzymatique chez les RNases. Les données intègrent un bon nombre de techniques, telles que la cristallographie, la spectroscopie RMN et la simulation de dynamique moléculaire. Fait à noter, ces techniques se déroulent sur plusieurs échelles de temps. Les résultats démontrent une conservation évolutive de certains éléments structuraux, malgré une faible homologie de séquence.

Le manuscrit a été accepté le 13 mai 2013 dans la revue "FEBS Journal" (facteur d'impact (2013) de 3.986), et une première version de l'article a été publiée en ligne le 24 juin 2013. L'article a été publié dans le numéro 22 du volume 280 de l'édition de novembre 2013, aux pages 5596-5607, sous le numéro d'accèsion doi: 10.1111/febs.12371. En date du 8 février 2015, l'article a été cité 5 fois. La version formatée par l'éditeur sous format PDF et publiée en ligne est présentée dans le présent document.

### **2.2 Contribution des auteurs**

Le texte et les figures ont été écrits, préparés et édités par l'auteur principal (Donald Gagné), sous la supervision de Pr Nicolas Doucet. Pr Nicolas Doucet a soumis le manuscrit à l'éditeur.

## 2.3 Résumé

Il a été démontré chez plusieurs protéines et systèmes enzymatiques que les mouvements atomiques font partie intégrante de leur fonction biologique [10, 12-14, 55, 64, 143]. Ces mouvements moléculaires sont impliqués, entre autres, dans la liaison et la discrimination des substrats, le positionnement et/ou la libération du produit. Alors que de nombreuses études font un lien entre les mouvements atomiques et la fonction de la protéine, peu d'études se concentrent sur la conservation de la dynamique dans une même famille d'enzyme. En effet, l'hypothèse veut que si les mouvements sont importants pour la fonction de la protéine, elles soient conservées pour les membres d'une famille d'enzymes de même fonction et de structure similaire. Dans cet article, nous utilisons les membres de la famille des RNase A comme modèle. Cette famille d'enzymes, dont la fonction commune est de dégrader l'ARN à simple ou double brin, a su conserver sa structure et son activité catalytique tout au long de son évolution. Malgré ces similitudes, l'efficacité catalytique varie considérablement entre les membres. Les résultats présentés dans cet article proviennent de l'utilisation de différentes techniques (par exemple: spectroscopie RMN, simulations MD, cristallographie) et comparent le comportement dynamique des RNases sur plusieurs échelles de temps.

Bien que le nombre de sujets demeure limité, les résultats d'études montrent une conservation des mouvements de certains éléments structuraux chez plusieurs membres de la famille des RNases A. Cette conservation de la flexibilité atomique évoque le fait que ces mouvements ne sont pas aléatoires, mais plutôt coordonnés avec la fonction de la protéine.

## 2.4 Article scientifique

the **FEBS**  
Journal



REVIEW ARTICLE

### Structural and functional importance of local and global conformational fluctuations in the RNase A superfamily

Donald Gagné<sup>1</sup> and Nicolas Doucet<sup>1,2,3</sup>

<sup>1</sup> INRS-Institut Armand-Frappier, Université du Québec, Laval, Québec, Canada

<sup>2</sup> PROTEO, the Québec Network for Research on Protein Function, Structure, and Engineering

<sup>3</sup> GRASP, Groupe de Recherche Axé sur la Structure des Protéines, Montréal, Québec, Canada

#### Keywords

enzyme flexibility; evolutionary conservation; molecular dynamics; NMR relaxation; ribonucleases; RNase A; X-ray crystallography

#### Correspondence

N. Doucet, INRS-Institut Armand-Frappier, Université du Québec, 531 Boul. des Prairies, Laval, Québec, Canada H7V 1B7  
Fax: +1 450 686 5501  
Tel: +1 450 687 5010, ext. 4212  
E-mail: nicolas.doucet@iaf.inrs.ca  
Website: <http://www.profs.inrs.ca/ndoucet/>

(Received 25 February 2013, revised 10 May 2013, accepted 13 May 2013)

doi:10.1111/febs.12371

Understanding the relationship between protein structure and flexibility is of utmost importance for deciphering the tremendous rates of reactions catalyzed by enzyme biocatalysts. It has been postulated that protein homologs have evolved similar dynamic fluctuations to promote catalytic function, a property that would presumably be encoded in their structural fold. Using one of the best-characterized enzyme systems of the past century, we explore this hypothesis by comparing the numerous and diverse flexibility reports available for a number of structural and functional homologs of the pancreatic-like RNase A superfamily. Using examples from the literature and from our own work, we cover recent and historical evidence pertaining to the highly dynamic nature of this important structural fold, as well as the presumed importance of local and global concerted motions on the ribonucleolytic function. This minireview does not pretend to cover the overwhelming RNase A literature in a comprehensive manner; rather, efforts have been made to focus on the characterization of multiple timescale motions observed in the free and/or ligand-bound structural homologs as they proceed along the reaction coordinates. Although each characterized enzyme of this architectural fold shows unique motional features on a local scale, accumulating evidence from X-ray crystallography, NMR spectroscopy and molecular dynamics simulations suggests that global dynamic fluctuations, such as the functionally relevant hinge-bending motion observed in the prototypical RNase A, are shared between homologs of the pancreatic-like RNase superfamily. These observations support the hypothesis that analogous dynamic residue clusters are evolutionarily conserved among structural and functional homologs catalyzing similar enzymatic reactions.

#### Introduction

Tremendous progress has been made in understanding how proteins and enzymes work since the first 'lock and key' mechanism was proposed by Fischer in 1894 [1]. From their biosynthesis to their destruction by proteol-

ysis, enzyme biocatalysts are no longer perceived as static structures that merely act as rigid bodies accommodating stiff substrates over the course of a catalytic reaction. Recognition of the dynamic nature of

#### Abbreviations

BS-RNase, bovine seminal ribonuclease; ECP, eosinophil cationic protein (human RNase 3); EDN, eosinophil-derived neurotoxin (human RNase 2); hANG, human angiogenin (RNase 5); MD, molecular dynamics.

enzyme catalysis can easily be traced back to the ‘configurational adaptability’ hypothesis of Karush [2] and the induced-fit theory of Koshland [3], whereby equilibrium ensembles and changes in the 3D structure of the enzyme upon ligand binding were postulated to facilitate catalysis. These original concepts opened the door to the plasticity of protein structures, and have considerably helped in defining our current understanding of protein function, which now fully embraces the importance of biologically relevant atomic motions occurring on timescales that range from picoseconds to several minutes [4,5]. As a testament to this fact, Anfinsen was awarded the 1972 Nobel Prize in Chemistry for his work on the thermodynamics of folding in bovine pancreatic RNase A, demonstrating that the 3D structure – and thus proper folding – of a native protein is fully and solely determined by its amino acid sequence [6]. As a logical extension to Anfinsen’s thermodynamic hypothesis, one could just as easily envision that dynamic fluctuations essential for optimal catalysis would be determined by the structural fold of a given protein family (i.e. structure-determined motions), which could partly explain their high structural homology [7]. According to this hypothesis, structurally related enzyme homologs catalyzing similar reactions would show analogous protein flexibility on a wide range of timescales at critical active site positions and/or within close or distant structural motifs that are important for enzyme function.

Although atomic fluctuations on various timescales are known to enable a wide array of biologically relevant activities, such as protein folding, structural reorganization, protein–protein interactions, and ligand binding [4], their role in enzyme catalysis still remains poorly defined. Evidence that structural dynamics play an important role in optimal enzyme catalysis was previously reported, and functional motions have been linked to the catalytic process in many discrete enzyme systems, including (but not limited to) dihydrofolate reductase [8], triphosphate isomerase [9], and adenylate kinase [10]. Although comprehensive theoretical studies have recently correlated sequence evolution with structural dynamics [7], experimental evidence has yet to clearly show whether conserved 3D structures within functionally related protein homologs have been evolutionarily selected to preserve flexibility features that are essential to promote catalytic function. In the present article, we collect current and historical evidence gathered from X-ray crystallography, protein NMR spectroscopy, and computational molecular dynamics (MD) to investigate the catalytic importance of dynamic fluctuations shared between structural and functional homologs of the RNase A superfamily. The

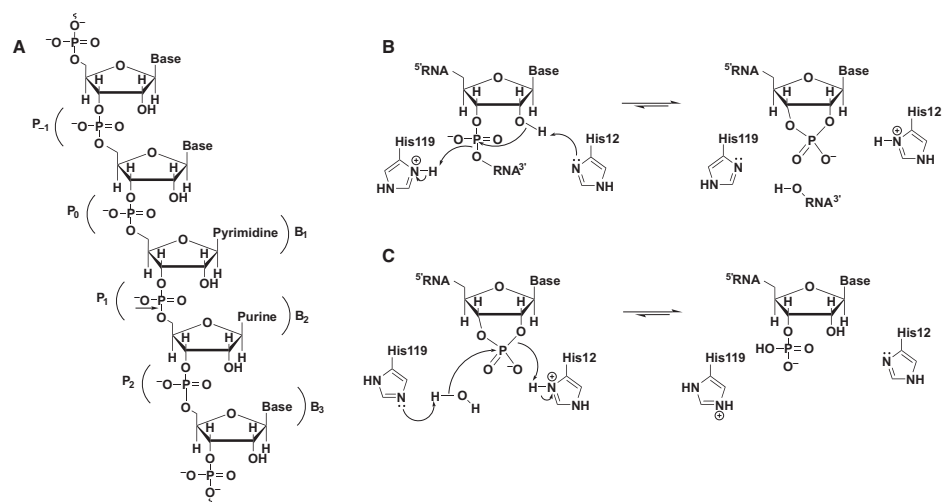
present minireview does not take into consideration the very important motions involved in folding mechanisms, nor does it consider the amply characterized swapped-dimer reorganization observed in many enzymes with this protein fold. Only motional events that were previously suggested to assist in ligand binding, stabilization, catalysis and/or product release in RNase A and structural homologs are considered. Experimental and theoretical reports show that pancreatic-like RNase homologs often show discretely divergent flexibility features on a local scale, which may be linked to their divergent biological activities. However, both local and global networks of residues experiencing similar clustering dynamics are conserved among several members of the superfamily, suggestive of a common role in catalytic function.

### **RNase A – conformational flexibility on subtle length scales**

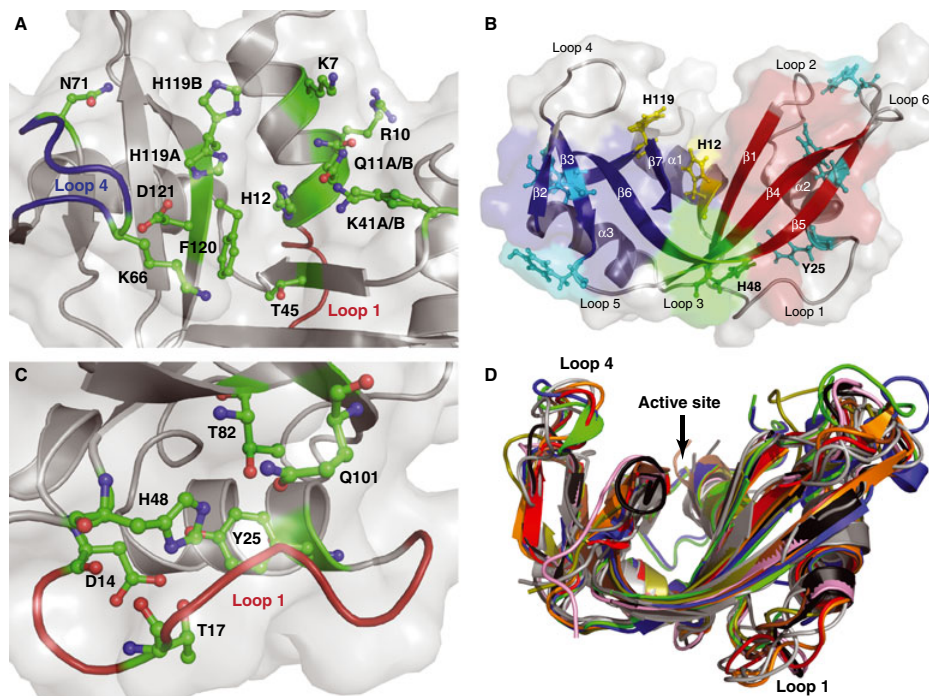
RNase was originally described between 1911 and 1920 by the biochemists Levene and Jones, who showed, in a series of seminal papers, that the ‘ferment’ of a boiled pig pancreas could digest yeast RNA [11–14]. Working from a variety of animal sources, Dubos at the Rockefeller Institute first purified the enzyme in 1937 [11]. It was subsequently named ‘ribonuclease’ by Dubos and Thompson in 1938 [15], and was first crystallized 3 years later by their colleague Kunitz [16]. Despite these early advances, RNase did not become the perfect enzyme system to investigate until Armour & Co. prepared over 1 kg of pure RNase A in the 1950s, and offered to supply any scientist of the biochemical community with a cheap and reliable source of this small (13.6 kDa), thermostable, and well-behaved enzyme [17,18]. Although early reports provided important information on its sequence [19] and catalytic mechanism – RNase A was the first enzyme for which a catalytic mechanism was ever elucidated [20] – it was not until the pioneering crystallographic work of Kartha *et al.* [21], Avey *et al.* [22] and Richards and Wyckoff [23] in the late 1960s that the true atomic details of its 3D structure became clearer. Through subsequent high-resolution X-ray crystallography studies of free and ligand-bound enzyme in the presence of substrate analogs, transition state analogs, products, and inhibitors, the first evidence of protein motions occurring over the course of the catalytic reaction became apparent [24]. Dozens of high-resolution crystals of RNase A have now been resolved in the presence of unique nucleotide ligands [24,25], providing a series of snapshots of the active enzyme as it moves along the reaction coordinates.

Although it has now been demonstrated that protein motions occurring on timescales ranging from picoseconds to milliseconds are required for optimal catalytic function in RNase A (see below), its 3D structure nevertheless remains relatively unperturbed as compared with the large conformational changes observed in other flexible systems [5]. Indeed, the structural changes observed in RNase A are characterized by modest deformations and small length-scale changes involving subtle reorganization of loops and secondary structure elements. As a broadly studied catalyst, RNase A originally served as a model system for investigating the importance of local and global residue motions in response to physical perturbations. For instance, Petsko *et al.* used it to demonstrate the existence of a dynamic transition at low temperature (220 K), below which the enzyme becomes too rigid to allow any ligand to fit in the ‘frozen-out’ active site, resulting in the loss of functionally important collective motions in the substrate cavity [26,27]. These experiments, in addition to the observation that the 2'-CMP inhibitor cannot be washed off from the active site at temperatures below the transition point, suggest that enzyme flexibility plays a crucial role in substrate recognition, catalysis, and/or product release.

Evidence of pH-induced motions in the vicinity of the active site environment and long-scale conformational effects have also been further characterized in detail. A number of crystallographic reports have now demonstrated the pH dependence of catalytic residue motions in RNase A [24,28–30]. RNase A catalyzes a well-characterized acid–base mechanism involving two histidines (His12 and His119) and a transition state stabilizing positive charge (Lys41). The transphosphorylation of a single-stranded RNA molecule by the enzyme yields a 2',3'-cyclic phosphomonoester intermediate, which can be expelled from the active site or hydrolyzed in a microscopic reverse reaction involving the same two histidines (Fig. 1) [20]. The catalytic base His119, which is also involved in transition state stabilization, is observed in two opposite conformations at acidic pH, shifting to the preferred conformation A at basic pH and/or in aqueous organic solvents [28] (Fig. 2A). Conformation B, which is favored in ionic salt solutions, has been shown to be incompatible with dinucleotide inhibitor binding [29]. MD simulations have suggested that His119 adopts both rotameric positions in solution, most likely experiencing conformational exchange over the course of a catalytic reaction [31,32]. In addition to its critical role in catalysis,



**Fig. 1.** RNA subsites and reaction mechanism catalyzed by RNase A. (A) Apparent molecular interactions between the RNA substrate and residues of the catalytic groove. Phosphate and base binding subsites are defined as P<sub>n</sub> and B<sub>n</sub>, with cytosine or uracil (adenine) favored at position B<sub>1</sub> (B<sub>2</sub>). The following residues are known to interact with each subsite: Lys66 (P<sub>0</sub>), Thr45 and Asp83 (B<sub>1</sub>), Gln11, His12, Lys41, His119, and Asp121 (P<sub>1</sub>), Asn71 and Glu111 (B<sub>2</sub>), and Lys7 and Arg10 (P<sub>2</sub>). The scissile bond is shown by an arrow. (B, C) Enzymatic reaction mechanism showing the transphosphorylation (B) and hydrolysis (C) reactions catalyzed by RNase A. Both panels are adapted from [89].



**Fig. 2.** Structural and motional features of RNase A. (A) Active site environment in RNase A. The main residues involved in ligand recognition, binding, stabilization and/or catalysis are displayed in ball-and-stick representation [Protein Data Bank (PDB) [1KF2](#)] [35]. The dynamic 'lid' loop 4 is shown in dark blue, and the distant flexible loop 1 is shown in red. Active site residues Gln11, Lys41 and His119 adopt more than one conformation (A/B), suggesting the importance of flexibility at specific sites. A standard coloring scheme is used for atomic representation: green for carbon, red for oxygen, and blue for nitrogen. (B) The V-shaped kidney structure of RNase A showing the two opposite domains defined by two antiparallel  $\beta$ -sheets, termed  $V_1$  (blue; residues 61–63, 71–75, 105–111, and 116–124) and  $V_2$  (red; residues 42–46, 82–87, and 96–101). The hinge region (green) is defined by residues 47–48, 80–81, and 102–104. Secondary structure elements are labeled, and so are the two catalytic histidines (His12 and His119) and the distal 'conformational switch' His48. The six tyrosines are shown in cyan (see text for details). (C) Zoomed view of the distal motional hinge environment in RNase A. A complex hydrogen-bonding network linking the flexible loop 1 to His48, Thr82, Gln101 and other residues of the  $V_2$   $\beta$ -strand was shown to act as an allosteric site whose motions control product release in RNase A [42–45]. (D) All known vertebrate members of the RNase A superfamily share a structurally similar 3D structure, with variability and flexibility being primarily found in loops. Representative members of the family are overlaid on the structure of RNase A (red) as follows: BS-RNase (yellow, PDB [3DJX](#)); human EDN (blue, PDB [1GOV](#)); human ECP (green, PDB [1QMT](#)); human RNase 4 (brown, PDB [1RNF](#)); hANG (black, PDB [1ANG](#)); rat RNase 1 (orange, PDB [1RRA](#)); bovine angiogenin (pink, PDB [1AGI](#)); frog Onconase (gray, PDB [3PHN](#)); and zebrafish RNase ZF-3e (olive, PDB [2VQ9](#)).

this observation suggests a potential role for His119 in ligand discrimination and/or stabilization. Whereas important functional residues, such as the catalytic acid His12, are notoriously rigid in the presence or absence of ligand [28,33,34], a number of active site residues involved in substrate recognition and/or transition state stabilization adopt multiple conformations in RNase A. High-resolution crystal comparisons

between the phosphate-free and sulfate-free enzyme and several ligand-bound complexes have indicated that as many as 21 residues could be modeled with two alternative side chain positions [33–35]. These include Gln11, Lys41, and Lys66, three active site residues involved in catalysis and proper ligand stabilization (Fig. 2A). Whereas Lys41 is involved in transition state stabilization, Gln11 is positioned to form a

hydrogen bond with the  $P_i$  substrate phosphate, and Lys66 contributes to the proper alignment of the catalytic His119 [36]. Loss of flexibility at these three critical active site positions could partly explain the severe effects observed on catalytic efficiency upon mutation [36].

Further evidence pertaining to the importance of global flexibility in the function of RNase A has been provided by reports on low-humidity crystals [37,38] and structural variations derived from distinct enzyme molecules in crystal asymmetric units [39]. RNase A shows a V-shaped kidney structure, with two opposite domains forming an elongated substrate-binding groove between two arms (Fig. 2B). These domains are defined by two antiparallel  $\beta$ -sheets termed  $V_1$  (residues 61–63, 71–75, 105–111, and 116–124) and  $V_2$  (residues 42–46, 82–87, and 96–101), which are linked at the base by a hinge region defined by residues 47–48, 80–81, and 102–104 (Fig. 2B) [39]. Recognition that long-scale motions in residues of the hinge region are coupled to the catalytic step of the reaction in RNase A dates back to the relaxation and stopped-flow temperature-jump experiments performed by Hammes *et al.* in the late 1960s [40,41]. Hammes recognized the existence of isomerization steps involving a restricted set of ionizable groups in the enzyme, one of which was later ascribed to His48 [40]. This imidazole side chain is located  $> 10 \text{ \AA}$  away from the catalytic site, and sits at the base of the hinge region (Fig. 2B,C). Ideally positioned to form a hydrogen-bonding network involving Thr82, Gln101, and the very flexible loop 1 (namely through Asp14 and Thr17) (Fig. 2C), His48 was recently shown to be an important conformational switch that controls distal motions that are essential for optimal product release [42–45]. The motional importance of this long-range effect at His48 was also indirectly confirmed by the change in the environment of tyrosines upon ligand binding, as well as by the detrimental effect on catalytic efficiency caused by tyrosine iodination [40]. This latter observation would remain a curiosity if not for the fact that no tyrosine is observed near the active site of the enzyme, in addition to the fact that all six tyrosines of the enzyme are located 10–20  $\text{\AA}$  from the catalytic His12 (Fig. 2B). These results strongly support the importance of allosteric motions linked to reactivity in RNase A. Using proton NMR spectroscopy, Markley later confirmed that Tyr25, which is in direct proximity to His48 (Fig. 2B,C), was coupled by an allosteric linkage to this important histidine of the hinge region [46]. It was also demonstrated that the protonation state of His48 is critical to promote movement of the nearby Gln101 side chain, which swings

out to uncover a distal pocket that can be filled by water molecules as the reaction proceeds [28]. We found that the identity and/or flexibility of Gln101, as a hinge residue distal from the active site, is also critical for the folding and/or stability of RNase A, as confirmed by the severely reduced stability of a Q101A mutant (Gagné and Doucet, unpublished).

Crystallographic evidence of the V-shaped enzyme showing conformational flexibility at the hinge region also suggested a potential correlation of this global architectural motion with ligand binding and release in RNase A. Crystal studies have shown rigid body motions of the  $V_1$  and  $V_2$  domains against each other, a bending motion that could account for the proper active site flexibility required for ligand recognition and/or release [37–39]. Although this bending motion of the two domains is extremely subtle on a global scale (ranging from  $1.6^\circ$  to  $2.2^\circ$  between the free and ligand-bound enzyme), it nevertheless correlates with the inability of RNase A to bind or release a ligand below the 220 K transition temperature [26,27,39].

This observation also confirms that the most important structural deformations in the enzyme involve hinge residues and the nearby flexible loop 1, which were recently characterized by Loria *et al.* in a series of studies combining enzyme kinetics, mutagenesis, NMR titration, kinetic solvent isotope effects, and NMR relaxation dispersion experiments [42–45,47,48]. The authors observed that RNase A exists in a dynamic pre-equilibrium composed of a total population that switches between two conformational forms (open and closed), which are affected upon ligand binding [47]. More importantly, these studies showed that the rate of conformational exchange ( $k_{ex}$ ) for these residues coincides with the rate-limiting product release in RNase A ( $k_{off}$ ), suggesting their interrelatedness. Mutations at or near the distant His48 were also observed to affect  $k_{cat}$ , in addition to uncoupling global long-scale motions occurring on the timescale of the catalytic reaction [42,43,45]. Investigation of a chimeric enzyme that perturbs the flexibility of the hinge/loop 1 environment was also performed. With the hybrid RNase A<sub>ECP</sub> variant, in which the short and rigid six-residue loop 1 from human RNase 3 [also known as eosinophil cationic protein (ECP)] replaces the 12-residue flexible loop 1 in RNase A, a 10-fold decrease in the product release rate constant  $k_{off}$  and a four-fold decrease in ligand affinity ( $K_d$ ) relative to the parent enzyme were observed [43]. Using mutagenesis and NMR relaxation dispersion experiments, the authors observed the existence of a hydrogen-bonding network transferring allosteric motions from the active site through the  $\beta_1$  and  $\beta_4$  strands, Thr82, His48, and

the distant loop 1 (Fig. 2B), confirming the functional importance of long-range conformational rearrangements in RNase A.

Recent years have seen the development of a number of NMR methods and a rapid improvement in MD simulations to sample protein flexibility, allowing the precise characterization of molecular motions over a large spectrum of timescales. Although the Lipari–Szabo formalism [49,50] remains the NMR method that is commonly used to study protein motions on the fast picosecond-to-nanosecond timescale, recent NMR techniques that rely on  $R_1$  and  $R_2$  relaxation, such as the off-resonance rotating frame experiment ( $R_{1\rho}$ ) [51] or the Carr–Purcell–Meiboom–Gill sequence used by Loria *et al.* on RNase A [42–45,47,52], now provide information on relevant millisecond motions occurring on the timescale of catalytic turnover ( $\sim 1000$ – $1500\text{ s}^{-1}$  in RNase A) [53,54]. Computational techniques that probe this particular millisecond catalytic time frame are also emerging as accurate methods for assessing local and global residue motions along the reaction coordinates [55]. Although most MD simulations performed on RNase A are still limited to the nanosecond timescale, they nevertheless provide necessary insights into the conformational changes associated with the different steps of the catalytic mechanism. Early MD simulations performed by Karpus *et al.* on free, substrate-bound and transition state-bound RNase A allowed the initial characterization of active site motions that cannot be seen with X-ray crystallography [32,56]. Coupled with the previously discussed crystallographic data, these simulations provided some of the first visual evidence that the catalytic His119 undergoes dihedral-angle transitions in the absence and presence of ligand, adopting the now well-characterized A and B conformations [28]. Other catalytic residues, such as His12 and Thr45, were shown to remain relatively immobile, and were similarly positioned in the free and ligand-bound protein. Interestingly, these studies also uncovered the existence of a long-range water network bridging charged residues that are important for transition state stabilization at the active site, involving Lys7, Lys41, Arg39, and Lys66, which were later demonstrated to be highly flexible. Recent MD studies have also confirmed that the hinge-bending motions involving the  $V_1$  and  $V_2$  domains constitute an intrinsic dynamic feature of the ligand-free enzyme [31], which NMR studies later characterized as the open and closed conformations of the free enzyme [47]. The highly conserved Thr45–His12 hydrogen bond was also suggested to contribute to the high rigidity of the active site region, whereas the functionally important adjacent loop 4

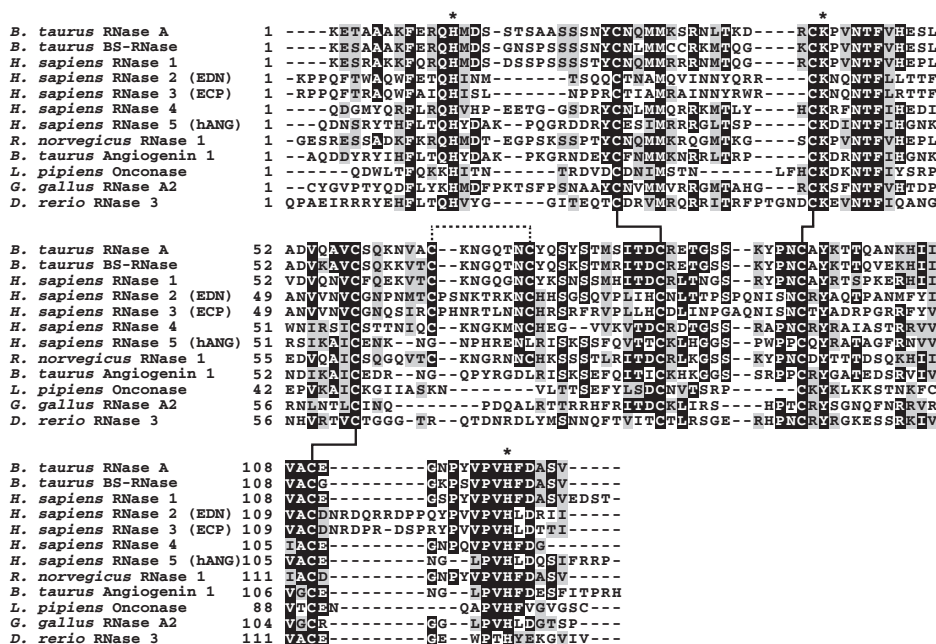
(residues 65–72) was shown to be one of the most flexible regions of the protein (Fig. 2A,B) [31]. Interestingly, Mazarella *et al.* have also observed that helix 3 (residues 50–70) (Fig. 2B), which joins the  $V_1$  and  $V_2$  domains, splits into two distinct segments that fluctuate independently on the nanosecond timescale, suggesting that this bending point could modulate breathing motions of the two domains in RNase A. Recent reports have suggested that important active site residues such as Lys41 could also play a role as general base/acid in lieu of His12 (see [57] and references therein).

### Conserved functional motions among structural homologs

The previous section is a testament to the number of theoretical and experimental reports that have independently highlighted the importance of local and global residue motions on multiple timescales in RNase A. A number of recent mutagenesis and dynamic reports have, in fact, convincingly demonstrated the existence of long-range allosteric motions in this enzyme, which cannot be perturbed without the loss of optimal catalysis [42,43,45]. As one of the best-characterized enzymes of the 20<sup>th</sup> century, RNase A is considered to be the prototypical member of a very large family that includes structural homologs found in all vertebrates, with known representatives in mammals, birds, amphibians, reptiles, and fish (Fig. 3) [58]. To date, all known structural homologs are endoribonucleases that have the ability to degrade single-stranded and/or double-stranded RNA, and maintain a number of important functional motifs, and, most importantly, share a very similar disulfide-bonded tertiary structure (Fig. 2D) [59]. In humans, the family includes eight canonical members, some of which have also evolved additional biological activities, often linked to innate host defense, neurotoxicity, angiogenesis, and immunosuppressive and/or antibacterial/antiviral activities [59,60]. For good reviews on the diversity of each human canonical member and the host defense capabilities of vertebrate RNases, see the excellent reviews of Sorrentino [60], Rosenberg [61], D'Alessio [58], and Zhang [62].

To date, well over 100 unique members of the vertebrate RNase superfamily have been identified, cloned, recombinantly expressed, and/or characterized [58]. However, only a restricted set have been subjected to dynamic investigation, considerably restricting flexibility comparisons between evolutionarily diverse enzymes with this structural fold. Although this is not an entirely comprehensive list, experimental and/or





**Fig. 3.** Sequence alignment between representative members of the vertebrate pancreatic-like RNase A superfamily. Alignment was performed with CLUSTALW2 [90,91], with representative structural homologs from cattle (*Bos taurus* RNase A, BS-RNase, angiogenin 1), human (*Homo sapiens* RNases 1, 2, 3, 4, and 5), rat (*Rattus norvegicus* RNase 1), frog [*Lithobates pipiens* ranpirinase (Onconase)], chicken (*Gallus gallus* RNase A2), and zebrafish (*Danio rerio* RNase 3). Residues showing > 50% identity (similarity) are highlighted in black (gray). The strictly conserved catalytic residues His12, Lys41 and His119 are starred (RNase A numbering). Filled connecting lines identify the three strictly conserved disulfide bridges between Cys26 and Cys84, Cys40 and Cys95, and Cys58 and Cys110. Dashed connecting lines identify the partially conserved disulfide bridge between Cys65 and Cys72 (RNase A numbering). Alignment was performed with BOXSHADE 3.21.

theoretical dynamic studies have been performed on free, ligand-bound, wild-type and/or mutant forms of human and rat RNase 1 [55,63], human eosinophil-derived neurotoxin (EDN, or RNase 2) [64], human ECP (or RNase 3) [65–68], human RNase 4 [64], bovine and human angiogenin (hANG, RNase 5) [69–74], the northern leopard frog ranpirinase (Onconase) [75–78], and monomeric derivatives of bovine seminal RNase (BS-RNase) [79,80]. The details of some of these dynamic studies will now be briefly described and compared with the local and global motional behavior of RNase A. For comparative purposes, the numbering of equivalent residues in bovine RNase A is provided in parentheses.

The range of ribonucleolytic activity, combined with the heterogeneity of their other biological functions, has led to renewed interest in human RNase A

homologs [60]. The existence of functional motions in human RNase 1, ECP (RNase 3) and angiogenin (RNase 5) has been well documented experimentally [63,65,74]. Despite showing low sequence identity (~30%), all structurally resolved human RNases share the same 3D fold stabilized by four disulfide bridges, except for angiogenin, which lacks the Cys65–Cys72 link found at the base of loop 4 in RNase A (Figs 2D and 3). Interestingly, to date, dynamic studies performed on human homologs have all confirmed local and global enzyme fluctuations that closely reflect those found in RNase A. NMR studies on hANG have characterized residue motions that mirror those observed in RNase A for several important residues of the active site [74]. For instance, His13 (His12) in hANG was found to be very rigid on the NMR time-scale, whereas Lys40 (Lys41) shows higher dispersion,

and His114 (His119) is in a dynamic equilibrium between two conformers differing in  $\chi^1$  and  $\chi^2$  angles, much as in RNase A [74]. Related observations have also been reported for active site residues in RNase 1 and EDN (RNase 2), whereby His119 in the former is flexible and undergoes chemical exchange [63], whereas His129 (His119) in the latter can adopt the B conformation that is also found in RNase A [81].

The most flexible secondary structure elements of the dynamically characterized human homologs were also shown to correlate with those of RNase A. For instance, considerable flexibility was observed in the short 66–68 segment of hANG (loop 4) (Fig. 2), which is a putative cell receptor binding site in this enzyme [74]. Residues in loop 4 were also shown to be involved in purine selection in the 5'-position of the phosphodiester cleavage site in RNase A [29,44]. NMR studies of loop 4 in RNase 1 have revealed that this segment is one of the most flexible environments of the enzyme in the absence of ligand, with considerable rigidity being imparted upon RNase inhibitor binding [63]. Our own recent NMR relaxation work focusing on the comparison between RNase A and ECP (RNase 3) also illustrates that the same loop 4 environment is the most flexible element in both enzymes, undergoing analogous conformational exchange on the catalytic timescale (milliseconds) [65]. We extended our analysis to demonstrate that each enzyme undergoes similar millisecond motions in the form of contiguous residue clusters that transcend the traditional definition of primary, secondary and tertiary protein architecture in these two homologs [65,82,83]. We also made similar observations in RNase 1, whereby conserved millisecond motions in loop 4 are consistent with motions observed in RNase A and ECP for this particular cluster involved in purine recognition and stabilization. This dynamic similarity for loop 4 residues in RNase 1 is also supported by the presence of lower-than-average order parameters ( $S^2$ ) – a measure of picosecond-to-nanosecond flexibility extracted from heteronuclear spin-relaxation model-free calculations [63] – and, in ECP, by nanosecond MD simulations confirming higher-than-average residue motion of this residue cluster [68]. Overall, accumulating evidence suggests that the conservation of flexibility in this particular loop structure overlooking the active site appears to be critical for enzymatic function in all pancreatic-like human RNase homologs characterized to date.

Focusing on the comparison between angiogenin and RNase A, Mazzarella *et al.* have also demonstrated the existence of similar global flexibility patterns in several members of the human RNase family

[64]. The two angiogenin domains defined by  $\beta$ -sheets  $V_1$  (residues 62–65, 69–72, 105–108, and 111–115) and  $V_2$  (residues 42–46, 76–84, and 93–101) were characterized by the presence of large breathing motions on the nanosecond timescale, similar to the hinge-bending motions observed in the reference enzyme (Fig. 2B) [64]. Proposing that this domain pincer motion may also play a functional role in structural homologs, the authors further analyzed complexed and ligand-free structures of EDN (RNase 2) and RNase 4. Small subdomain rearrangements were observed in the same  $\beta$ -sheet clusters (ranging from 1.9° to 2.2°), suggesting that the global dynamic behavior of homologs is preserved in the RNase architecture and is essential for optimal ligand binding, release, and/or enzyme activity [64]. Similarly, flexible clusters were also observed in other members of the vertebrate RNase superfamily. Namely, Lallemand *et al.* solved the NMR solution structure of human and bovine angiogenin, and demonstrated that both homologs are similar in terms of overall fold and flexibility [73,74]. The only noted functional exception was found for the catalytic His115 (His119) of the bovine enzyme, which, unlike its human counterpart, does not appear to adopt the B conformation.

Finally, of particular interest is the dynamic behavior of Onconase, an RNase homolog extracted from northern leopard frog oocytes [75–78]. The enzyme is currently under study for the treatment of various forms of cancer, and has the ability to evade the RNase inhibitor, which makes it a more effective drug than its RNase homologs [84]. Like other members of the family, Onconase has a similar fold but low sequence identity with RNase A (30%) [75]. The enzyme also has shorter loops in general, including the important loop 4 [78]. Nevertheless, motions for loop 4, Lys31 (Lys41) and His97 (His119), and conformational changes in the  $V_1$  and  $V_2$  domains, were also recently shown by crystallography to reflect those observed in structurally homologous RNases [78]. Similar conclusions were drawn for monomeric forms of BS-RNase, which appear to show analogous loop 4 flexibility and allosteric motions neighboring the distant loop 1 (residues 16–22) [79,80].

## Conclusions

It is now widely acknowledged that flexibility is crucial for proper biological function, and that perturbing molecular motions is a promising path for modulating protein function [85]. Although previous studies have shown the existence of similar collective protein motions among analogous members of protein

families [86–88], the link between residue dynamics and reactivity has been harder to establish for enzyme catalysis. Using a variety of reports focusing on the experimental and theoretical investigation of residue dynamics among members of the RNase A superfamily, we provide evidence that local and global collective motions are conserved among the structurally and functionally similar homologs of this enzyme clan. Analogous motions are observed for evolutionarily conserved and critical catalytic positions of the active site (namely Gln11, Lys41, Lys66, and His119), suggesting that dynamic plasticity is preserved for proper ribonucleolytic function among RNase A homologs. Also, conservation of global, collective flexibility manifested in the form of a functionally relevant hinge-bending motion between two protein domains has been well maintained and illustrated for several members of this enzyme fold. Strikingly, local and global dynamic fluctuations are preserved despite low sequence identity among several vertebrate RNases, supporting the hypothesis that protein architecture rather than primary sequence plays a determinant role in the intrinsic collective motions of a protein fold [86,88]. The functional effects of perturbing local and/or collective motions in RNase A illustrate that conservation of flexibility is not a trivial consequence of structural similarity [88]. Rather, these observations suggest that structural homologs of this enzyme fold may rely on global elasticity for proper enzyme catalysis.

### Acknowledgements

This work was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery grant (RGPIN 402623-2011) and a 'Fonds de Recherche Québec – Santé' (FRQS) Research Scholar Junior 1 Career Award (to N. Doucet). The authors also acknowledge support from the 'Fonds de recherche Québec – Nature et Technologie' (FRQNT) Strategic Cluster 'Regroupement Québécois de Recherche sur la Fonction, la Structure et l'Ingénierie des Protéines' (PROTEO) and the FRQS Strategic Cluster 'Groupe de Recherche Axé sur la Structure des Protéines' (GRASP). D. Gagné is the recipient of an NSERC Alexander Graham Bell Canada PhD scholarship.

### References

- Fischer E (1894) Einfluss der konfiguration auf die wirkung der enzvnie. *Ber Deutsch Chem Ges* **27**, 2985–2993.
- Karush F (1950) Heterogeneity of the binding sites of bovine serum albumin. *J Am Chem Soc* **72**, 2705–2713.
- Koshland DE (1958) Application of a theory of enzyme specificity to protein synthesis. *Proc Natl Acad Sci USA* **44**, 98–104.
- Doucet N & Pelletier JN (2009) Gaining insight into enzyme function through correlations with protein motions. In *Protein Engineering Handbook* (Lutz S & Bornscheuer UT, eds), pp. 187–211. Wiley-VCH, Weinheim.
- Henzler-Wildman K & Kern D (2007) Dynamic personalities of proteins. *Nature* **450**, 964–972.
- Anfinsen CB (1973) Principles that govern the folding of protein chains. *Science* **181**, 223–230.
- Liu Y & Bahar I (2012) Sequence evolution correlates with structural dynamics. *Mol Biol Evol* **29**, 2253–2263.
- Bhabha G, Lee J, Ekiert DC, Gam J, Wilson IA, Dyson HJ, Benkovic SJ & Wright PE (2011) A dynamic knockout reveals that conformational fluctuations influence the chemical step of enzyme catalysis. *Science* **332**, 234–238.
- Wang Y, Berlow RB & Loria JP (2009) Role of loop–loop interactions in coordinating motions and enzymatic function in triosephosphate isomerase. *Biochemistry* **48**, 4548–4556.
- Aden J, Verma A, Schug A & Wolf-Watz M (2012) Modulation of a pre-existing conformational equilibrium tunes adenylate kinase activity. *J Am Chem Soc* **134**, 16562–16570.
- Dubos RJ (1937) The decomposition of yeast nucleic acid by a heat resistant enzyme. *Science* **85**, 549–550.
- Jones EW (1920) The action of boiled pancreas extract on yeast nucleic acid. *Am J Physiol* **52**, 203–207.
- Levene PA & Medigreceanu F (1911) On nucleases. Second paper. *J Biol Chem* **9**, 389–402.
- Jones W (1912) On the formation of guanylic acid from yeast nucleic acid. *J Biol Chem* **12**, 31–35.
- Dubos RJ & Thompson RHS (1938) The decomposition of yeast nucleic acid by a heat-resistant enzyme. *J Biol Chem* **124**, 501–510.
- Kunitz M (1940) Crystalline ribonuclease. *J Gen Physiol* **24**, 15–32.
- Richards FM & Wyckoff HW (1971) Bovine pancreatic ribonuclease. In *The Enzymes* (Boyer PD, ed.), pp. 647–806. Academic Press, New York and London.
- Richards FM (1992) Linderstrom-Lang and the Carlsberg Laboratory: the view of a postdoctoral fellow in 1954. *Protein Sci* **1**, 1721–1730.
- Smyth DG, Stein WH & Moore S (1963) The sequence of amino acid residues in bovine pancreatic ribonuclease: revisions and confirmations. *J Biol Chem* **238**, 227–234.
- Cuchillo CM, Nogues MV & Raines RT (2011) Bovine pancreatic ribonuclease: fifty years of the first

- enzymatic reaction mechanism. *Biochemistry* **50**, 7835–7841.
- 21 Kartha G, Bello J & Harker D (1967) Tertiary structure of ribonuclease. *Nature* **213**, 862–865.
  - 22 Avey HP, Boles MO, Carlisle CH, Evans SA, Morris SJ, Palmer RA, Woolhouse BA & Shall S (1967) Structure of ribonuclease. *Nature* **213**, 557–562.
  - 23 Wyckoff HW, Hardman KD, Allewell NM, Inagami T, Johnson LN & Richards FM (1967) The structure of ribonuclease-S at 3.5 Å resolution. *J Biol Chem* **242**, 3984–3988.
  - 24 Gilliland GL (1997) Crystallographic studies of ribonuclease complexes. In *Ribonucleases: Structures and Functions* (D'Allesio G & Riordan JF, eds), pp. 305–341. Academic Press, New York.
  - 25 Boix E, Blanco JA, Nogues MV & Moussaoui M (2013) Nucleotide binding architecture for secreted cytotoxic endoribonucleases. *Biochimie* **95**, 1087–1097.
  - 26 Rasmussen BF, Stock AM, Ringe D & Petsko GA (1992) Crystalline ribonuclease A loses function below the dynamical transition at 220 K. *Nature* **357**, 423–424.
  - 27 Tilton RF Jr, Dewan JC & Petsko GA (1992) Effects of temperature on protein structure and dynamics: X-ray crystallographic studies of the protein ribonuclease-A at nine different temperatures from 98 to 320 K. *Biochemistry* **31**, 2469–2481.
  - 28 Berisio R, Lamzin VS, Sica F, Wilson KS, Zagari A & Mazzarella L (1999) Protein titration in the crystal state. *J Mol Biol* **292**, 845–854.
  - 29 Zegers I, Maes D, Dao-Thi MH, Poortmans F, Palmer R & Wyns L (1994) The structures of RNase A complexed with 3'-CMP and d(CpA): active site conformation and conserved water molecules. *Protein Sci* **3**, 2322–2339.
  - 30 de Mel VS, Doscher MS, Martin PD & Edwards BF (1994) The occupancy of two distinct conformations by active-site histidine-119 in crystals of ribonuclease is modulated by pH. *FEBS Lett* **349**, 155–160.
  - 31 Merlino A, Vitagliano L, Ceruso MA, Di Nola A & Mazzarella L (2002) Global and local motions in ribonuclease A: a molecular dynamics study. *Biopolymers* **65**, 274–283.
  - 32 Brünger AT, Brooks CL III & Karplus M (1985) Active site dynamics of ribonuclease. *Proc Natl Acad Sci USA* **82**, 8458–8462.
  - 33 Campbell RL & Petsko GA (1987) Ribonuclease structure and catalysis: crystal structure of sulfate-free native ribonuclease A at 1.5-Å resolution. *Biochemistry* **26**, 8579–8584.
  - 34 Wlodawer A, Svensson LA, Sjölin L & Gilliland GL (1988) Structure of phosphate-free ribonuclease A refined at 1.26 Å. *Biochemistry* **27**, 2705–2717.
  - 35 Berisio R, Sica F, Lamzin VS, Wilson KS, Zagari A & Mazzarella L (2002) Atomic resolution structures of ribonuclease A at six pH values. *Acta Crystallogr D Biol Crystallogr* **58**, 441–450.
  - 36 Chatani E & Hayashi R (2001) Functional and structural roles of constituent amino acid residues of bovine pancreatic ribonuclease A. *J Biosci Bioeng* **92**, 98–107.
  - 37 Kishan RV, Chandra NR, Sudarsanakumar C, Suguna K & Vijayan M (1995) Water-dependent domain motion and flexibility in ribonuclease A and the invariant features in its hydration shell. An X-ray study of two low-humidity crystal forms of the enzyme. *Acta Crystallogr D Biol Crystallogr* **51**, 703–710.
  - 38 Sadasivan C, Nagendra HG & Vijayan M (1998) Plasticity, hydration and accessibility in ribonuclease A. The structure of a new crystal form and its low-humidity variant. *Acta Crystallogr D Biol Crystallogr* **54**, 1343–1352.
  - 39 Vitagliano L, Merlino A, Zagari A & Mazzarella L (2002) Reversible substrate-induced domain motions in ribonuclease A. *Proteins* **46**, 97–104.
  - 40 Hammes GG (1968) Relaxation spectrometry of enzymatic reactions. *Acc Chem Res* **1**, 321–329.
  - 41 Hammes GG & Frederick GWJ (1969) Relaxation spectra of ribonuclease. VI. The interaction of ribonuclease with uridine 3'-monophosphate. *J Am Chem Soc* **91**, 7179–7186.
  - 42 Doucet N, Khirich G, Kovrigin EL & Loria JP (2011) Alteration of hydrogen bonding in the vicinity of histidine 48 disrupts millisecond motions in RNase A. *Biochemistry* **50**, 1723–1730.
  - 43 Doucet N, Watt ED & Loria JP (2009) The flexibility of a distant loop modulates active site motion and product release in ribonuclease A. *Biochemistry* **48**, 7160–7168.
  - 44 Watt ED, Rivalta I, Whittier SK, Batista VS & Loria JP (2011) Reengineering rate-limiting, millisecond enzyme motions by introduction of an unnatural amino acid. *Biophys J* **101**, 411–420.
  - 45 Watt ED, Shimada H, Kovrigin EL & Loria JP (2007) The mechanism of rate-limiting motions in enzyme function. *Proc Natl Acad Sci USA* **104**, 11981–11986.
  - 46 Markley JL (1975) Correlation proton magnetic resonance studies at 250 MHz of bovine pancreatic ribonuclease. II. pH and inhibitor-induced conformational transitions affecting histidine-48 and one tyrosine residue of ribonuclease A. *Biochemistry* **14**, 554–561.
  - 47 Beach H, Cole R, Gill ML & Loria JP (2005) Conservation of μs–ms enzyme motions in the apo- and substrate-mimicked state. *J Am Chem Soc* **127**, 9167–9176.
  - 48 Cole R & Loria JP (2002) Evidence for flexibility in the function of ribonuclease A. *Biochemistry* **41**, 6072–6081.
  - 49 Lipari G & Szabo A (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation

- in macromolecules. 1. Theory and range of validity. *J Am Chem Soc* **104**, 4546–4559.
- 50 Lipari G & Szabo A (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J Am Chem Soc* **104**, 4559–4570.
- 51 Akke M & Palmer AG (1996) Monitoring macromolecular motions on microsecond to millisecond time scales by R(1)rho-R(1) constant relaxation time NMR spectroscopy. *J Am Chem Soc* **118**, 911–912.
- 52 Loria JP, Rance M & Palmer AG III (1999) A relaxation-compensated Carr–Purcell–Meiboom–Gill sequence for characterizing chemical exchange by NMR spectroscopy. *J Am Chem Soc* **121**, 2331–2332.
- 53 Doucet N (2011) Can enzyme engineering benefit from the modulation of protein motions? Lessons learned from NMR relaxation dispersion experiments *Protein Pept Lett* **18**, 336–343.
- 54 Manley G & Loria JP (2012) NMR insights into protein allostery. *Arch Biochem Biophys* **519**, 223–231.
- 55 Ramanathan A & Agarwal PK (2011) Evolutionarily conserved linkage between enzyme fold, flexibility, and catalysis. *PLoS Biol* **9**, e1001193.
- 56 Haydock K, Lim C, Brunger AT & Karplus M (1990) Simulation analysis of structures on the reaction pathway of RNase A. *J Am Chem Soc* **112**, 3826–3831.
- 57 Formoso E, Matxain JM, Lopez X & York DM (2010) Molecular dynamics simulation of bovine pancreatic ribonuclease A-CpA and transition state-like complexes. *J Phys Chem B* **114**, 7371–7382.
- 58 Pizzo E & D'Alessio G (2007) The success of the RNase scaffold in the advance of biosciences and in evolution. *Gene* **406**, 8–12.
- 59 Dyer KD & Rosenberg HF (2006) The RNase A superfamily: generation of diversity and innate host defense. *Mol Divers* **10**, 585–597.
- 60 Sorrentino S (2010) The eight human 'canonical' ribonucleases: molecular diversity, catalytic properties, and special biological actions of the enzyme proteins. *FEBS Lett* **584**, 2194–2200.
- 61 Rosenberg HF (2008) RNase A ribonucleases and host defense: an evolving story. *J Leukoc Biol* **83**, 1079–1087.
- 62 Cho S, Beintema JJ & Zhang J (2005) The ribonuclease A superfamily of mammals and birds: identifying new members and tracing evolutionary histories. *Genomics* **85**, 208–220.
- 63 Kover KE, Bruix M, Santoro J, Batta G, Laurents DV & Rico M (2008) The solution structure and dynamics of human pancreatic ribonuclease determined by NMR spectroscopy provide insight into its remarkable biological activities and inhibition. *J Mol Biol* **379**, 953–965.
- 64 Merlino A, Vitagliano L, Ceruso MA & Mazzarella L (2003) Subtle functional collective motions in pancreatic-like ribonucleases: from ribonuclease A to angiogenin. *Proteins* **53**, 101–110.
- 65 Gagné D, Charest LA, Morin S, Kovrigin EL & Doucet N (2012) Conservation of flexible residue clusters among structural and functional enzyme homologues. *J Biol Chem* **287**, 44289–44300.
- 66 Sanjeev BS & Vishveshwara S (2002) Essential dynamics and sidechain hydrogen bond cluster studies on eosinophil cationic protein. *Eur Phys J D* **20**, 601–608.
- 67 Sanjeev BS & Vishveshwara S (2004) Conformational transitions in eosinophil cationic protein: a molecular dynamics study in aqueous environment. *J Biomol Struct Dyn* **22**, 171–182.
- 68 Sanjeev BS & Vishveshwara S (2005) Dynamics of the native and the ligand-bound structures of eosinophil cationic protein: network of hydrogen bonds at the catalytic site. *J Biomol Struct Dyn* **22**, 657–672.
- 69 Madhusudhan MS & Vishveshwara S (2001) Computer modeling of human angiogenin–dinucleotide substrate interaction. *Proteins* **42**, 125–135.
- 70 Madhusudhan MS, Sanjeev BS & Vishveshwara S (2001) Computer modeling and molecular dynamics simulations of ligand bound complexes of bovine angiogenin: dinucleotide topology at the active site of RNase a family proteins. *Proteins* **45**, 30–39.
- 71 Madhusudhan MS & Vishveshwara S (1999) Comparison of the dynamics of bovine and human angiogenin: a molecular dynamics study. *Biopolymers* **49**, 131–144.
- 72 Madhusudhan MS & Vishveshwara S (1998) Modeling of angiogenin–3'-NMP complex. *J Biomol Struct Dyn* **16**, 715–722.
- 73 Lequin O, Albaret C, Bontems F, Spik G & Lallemand JY (1996) Solution structure of bovine angiogenin by <sup>1</sup>H nuclear magnetic resonance spectroscopy. *Biochemistry* **35**, 8870–8880.
- 74 Lequin O, Thuring H, Robin M & Lallemand JY (1997) Three-dimensional solution structure of human angiogenin determined by <sup>1</sup>H,<sup>15</sup>N-NMR spectroscopy – characterization of histidine protonation states and pKa values. *Eur J Biochem* **250**, 712–726.
- 75 Merlino A, Mazzarella L, Carannante A, Di Fiore A, Di Donato A, Notomista E & Sica F (2005) The importance of dynamic effects on the enzyme activity: X-ray structure and molecular dynamics of onconase mutants. *J Biol Chem* **280**, 17953–17960.
- 76 Gorbatyuk VY, Tsai CK, Chang CF & Huang TH (2004) Effect of N-terminal and Met23 mutations on the structure and dynamics of onconase. *J Biol Chem* **279**, 5772–5780.
- 77 Callis M, Serrano S, Benito A, Laurents DV, Vilanova M, Bruix M & Ribo M (2013) Towards tricking a pathogen's protease into fighting infection: the 3D structure of a stable circularly permuted onconase variant cleaved by HIV-1 protease. *PLoS One* **8**, e54568.

- 78 Holloway DE, Singh UP, Shogen K & Acharya KR (2011) Crystal structure of onconase at 1.1 Å resolution – insights into substrate binding and collective motion. *FEBS J* **278**, 4136–4149.
- 79 Spadaccini R, Ercole C, Gentile MA, Sanfelice D, Boelens R, Wechselberger R, Batta G, Bernini A, Nicolai N & Picone D (2012) NMR studies on structure and dynamics of the monomeric derivative of BS-RNase: new insights for 3D domain swapping. *PLoS One* **7**, e29076.
- 80 Avitabile F, Alfano C, Spadaccini R, Crescenzi O, D'Ursi AM, D'Alessio G, Tancredi T & Picone D (2003) The swapping of terminal arms in ribonucleases: comparison of the solution structure of monomeric bovine seminal and pancreatic ribonucleases. *Biochemistry* **42**, 8704–8711.
- 81 Mosimann SC, Newton DL, Youle RJ & James MN (1996) X-ray crystallographic structure of recombinant eosinophil-derived neurotoxin at 1.83 Å resolution. *J Mol Biol* **260**, 540–552.
- 82 Halabi N, Rivoire O, Leibler S & Ranganathan R (2009) Protein sectors: evolutionary units of three-dimensional structure. *Cell* **138**, 774–786.
- 83 Reynolds KA, McLaughlin RN & Ranganathan R (2011) Hot spots for allosteric regulation on protein surfaces. *Cell* **147**, 1564–1575.
- 84 Leland PA & Raines RT (2001) Cancer chemotherapy – ribonucleases to the rescue. *Chem Biol* **8**, 405–413.
- 85 Lee GM & Craik CS (2009) Trapping moving targets with small molecules. *Science* **324**, 213–215.
- 86 Keskin O, Jernigan RL & Bahar I (2000) Proteins with similar architecture exhibit similar large-scale dynamic behavior. *Biophys J* **78**, 2093–2106.
- 87 Maguid S, Fernandez-Alberti S & Echave J (2008) Evolutionary conservation of protein vibrational dynamics. *Gene* **422**, 7–13.
- 88 Maguid S, Fernandez-Alberti S, Parisi G & Echave J (2006) Evolutionary conservation of protein backbone flexibility. *J Mol Evol* **63**, 448–457.
- 89 Raines RT (1998) Ribonuclease A. *Chem Rev* **98**, 1045–1065.
- 90 Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J & Lopez R (2010) A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res* **38**, W695–W699.
- 91 Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948.

## 2.5 Discussion

Dans cet article, nous présentons un résumé des mouvements atomiques observés chez plusieurs membres de la famille des RNases. Nous démontrons la conservation de mouvements locaux et globaux entre les protéines. Bien qu'il soit encore difficile d'établir un lien direct, les résultats semblent suggérer que cette conservation des mouvements atomiques entre les membres d'une même famille d'homologues structuraux partageant une fonction similaire pourrait être importante pour l'activité de la protéine. Malheureusement, il n'existe encore que peu de données expérimentales nous permettant de faire ce lien.

Des résultats précédents de cristallographie portant sur la RNase A sous formes libre et liée ont montré une légère oscillation de quelques degrés de la structure en V [144, 145]. Ce changement structural, quoique mineur, a été associé à la fonction de la protéine [145, 146]. Puisque cette perturbation des deux bras ( $V_1$  et  $V_2$ ) se produit à un taux d'échange conformationnel ( $k_{ex}$ ) similaire à celui de la libération des produits ( $k_{off}$ ), l'étape limitante de la réaction, il a été suggéré que les mouvements dans cet ordre de grandeur pouvaient être essentielles à la catalyse enzymatique [12]. En utilisant la mutagenèse dirigée, des études subséquentes ont confirmé l'implication de plusieurs résidus formant la charnière entre les secteurs  $V_1$  et  $V_2$  [64].

Bien que la famille des RNases comporte plus d'une centaine de membres, il n'y a que peu d'études caractérisant leur dynamique. Néanmoins, ces études démontrent une conservation des mouvements atomiques de certains éléments structuraux. Par exemple, une flexibilité de la boucle 4, qui est impliquée dans la sélection de la purine en 5' du lien phosphodiester, a été observée chez plusieurs RNases humaines, telles la RNase 1 et ECP, de même que chez la RNase A de bovin. De façon similaire, les oscillations entre les bras  $V_1$  et  $V_2$ , aperçues chez la RNase A, ont également été constatées chez plusieurs autres membres de la famille humaine des RNases. Plusieurs de ces résultats proviennent d'études bio-informatiques effectuées il y a plusieurs années. Toutefois, avec l'avancement des techniques de spectroscopie RMN, de bio-informatique, ou de cristallographie, l'accumulation de résultats devrait s'accélérer et permettre de mieux comprendre et interpréter les mouvements chez les protéines, en relation avec leur fonction.

## 2.6 Résultats supplémentaires

Depuis la parution de l'article, des études portant sur d'autres systèmes enzymatiques ont été publiées [147-149]. Ces études démontrent la conservation des mouvements au cours de l'évolution, et suggèrent l'importance de ces mouvements pour la fonction de la protéine. En appui à ces résultats, nous avons également poursuivi notre analyse de la flexibilité fonctionnelle chez les membres de la famille des RNases humaines (voir les chapitres 3-5).



### **3 ARTICLE NO. 2: ALTÉRATION DE L'ACTIVITÉ ENZYMATIQUE SUITE À LA PERTURBATION DE LA DYNAMIQUE CONFORMATIONNELLE D'UNE BOUCLE DU SITE ACTIF**

Titre: "Perturbation of the conformational dynamics of an active-site loop alters enzyme activity"

Auteurs: **Gagné, D.**, French, R.L., Narayanan, C., Agarwal, P.K., Simonović, M. & Doucet N.

#### **3.1 Présentation de l'article scientifique**

L'ingénierie des protéines est un champ d'application très prometteur, que ce soit en santé, en environnement, ou en industrie. Il consiste à modifier une protéine pour lui donner une fonction nouvelle ou améliorée. Bien qu'intéressant, il n'en demeure pas moins qu'il est très difficile de cibler les modifications à apporter à la protéine afin d'atteindre le but visé. En ce sens, les RNases, une famille d'homologues structuraux ayant pour fonction primaire la dégradation de l'ARN, possèdent des intérêts commerciaux intéressants. En effet, il a été démontré que plusieurs membres de la famille possédaient des fonctions biologiques spécifiques à chacune d'entre elles [97, 150, 151]. Par exemple, le potentiel cytotoxique de la RNase 1 humaine, ainsi que sa forte activité ribonucléolytique, pourrait être utilisé dans le traitement de nombreux cancers. Par contre, la protéine est reconnue dans le cytosol par l'inhibiteur des RNases et par les protéases, ce qui l'empêche d'atteindre et de dégrader l'ARN du cytosol et du noyau, et donc, de détruire la cellule (Figure 3.1). Depuis plusieurs années, le groupe de Raines [152-155] s'efforce de comprendre le mécanisme de liaison des RNases à son inhibiteur. Cet exemple démontre l'intérêt commercial que peuvent avoir certaines protéines, mais aussi, les problématiques à résoudre. Donc, avant de pouvoir envisager l'utilisation de ces protéines, nous nous devons de mieux comprendre leur fonctionnement. Avec cet objectif en tête, nous avons utilisé la RNase A de bovin, une enzyme relativement stable et qui se prête bien aux études par spectroscopie RMN, pour étudier le mécanisme de la catalyse enzymatique.

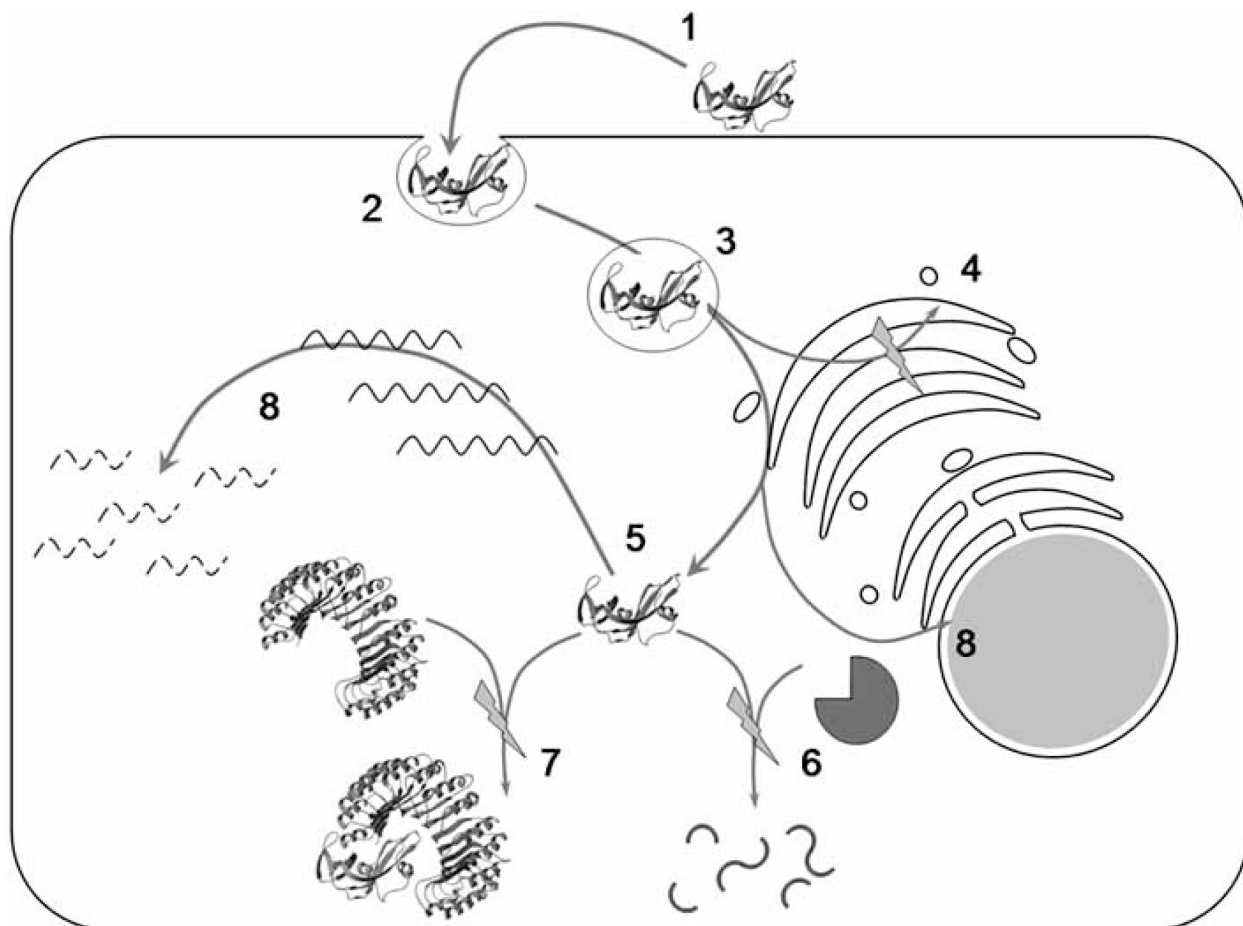


Figure 3.1. Cytotoxicité des RNases. La protéine se fixe à la membrane cellulaire (1) et entre par endocytose (2). À l'intérieur de l'endosome (3), la protéine sera transportée vers les organites, comme les lysosomes (4), où relâchée dans le cytosol (5). La protéine doit être en mesure de résister aux attaques par les protéases (6) et de s'évader de l'inhibiteur des RNases (7) pour espérer dégrader l'ARN dans le cytosol et dans le noyau (8). L'image a été modifiée de la référence [151].

Le manuscrit a été soumis le 13 mai 2015 dans la revue "Journal of the American Chemical Society" (facteur d'impact (2013) de 11.444). La version soumise à l'éditeur sous format PDF est présentée dans le présent document.

### 3.2 Contribution des auteurs

Le travail expérimental a été réalisé par l'auteur principal (Donald Gagné). L'analyse et l'interprétation des résultats de  $^{15}\text{N}$ -CPMG et du titrage RMN ont été faites par l'auteur principal (Donald Gagné). La formation et l'analyse des cristaux ont été effectuées par Rachel L. French, étudiante au doctorat dans le laboratoire de Pr Miljan Simonović. Enfin, Chitra Narayanan et le Pr Pratul Agarwal ont mené les expériences d'analyses de simulations de dynamique moléculaire. Le texte et les figures ont été écrits, préparés et édités par l'auteur principal

(Donald Gagné), avec une contribution soutenue de Chitra Narayanan, sous la supervision de Pr Nicolas Doucet, sauf exception des résultats et figures touchant les simulations de dynamique moléculaire, préparé par le Pr Pratul Agarwal. Pr Nicolas Doucet a soumis le manuscrit à l'éditeur.

### 3.3 Résumé

L'ingénierie des protéines est un champ d'application en pleine progression. Toutefois, avant de pouvoir facilement améliorer ou modifier la fonction d'une enzyme, il est essentiel de bien comprendre son fonctionnement. Au cours des dernières années, les mouvements atomiques ont été associés à la fonction des protéines, au même titre que la séquence et la structure [17, 78, 156]. Pour cette étude, nous avons utilisé la RNase A, une petite enzyme qui se prête bien aux études RMN. Bien que cette enzyme ait fait l'objet de nombreuses recherches, notre compréhension de sa fonction est loin d'être complète. Afin de mieux comprendre l'implication de la flexibilité atomique dans la fonction des protéines, nous avons étudié le rôle de l'alanine en position 109.

L'Ala109 est située sur le brin  $\beta 5$ , et est conservée à 88% chez les membres de la famille des RNases. Bien que ce résidu ait été associé à l'initiation de la chaîne de repliement [157, 158], il a été suggéré qu'il pourrait aussi avoir une influence sur la catalyse enzymatique. Par cristallographie, il a été démontré que Ala109 pouvait former une interaction de type Van der Waals avec l'histidine 119, ainsi qu'avec l'inhibiteur 3'-N-piperidine-4-carboxyl-3'-deoxy-ara-uridine [159]. Par ailleurs, des mouvements dans l'échelle de temps de la  $\mu\text{s}$ -ms avaient déjà été observées par spectroscopie RMN chez le variant Ala109 [160]. Dans cet article, nous explorons le rôle potentiel de Ala109 chez la RNase A, en mutant le résidu pour une glycine, un changement mineur consistant à enlever un groupement méthyle. Nous montrons que la mutation crée un repositionnement du ligand 5'-AMP en  $B_2$ , un changement qui semble causé par une augmentation du taux d'échange conformationnel et une extension du mouvement dans la boucle 4.

## 3.4 Article scientifique

### ***Perturbation of the conformational dynamics of an active-site loop alters enzyme activity***

Donald Gagné<sup>1</sup>, Rachel L. French<sup>2</sup>, Chitra Narayanan<sup>1</sup>, Pratul K. Agarwal<sup>3,4</sup>, Miljan Simonović<sup>2</sup> & Nicolas Doucet<sup>1,5,6\*</sup>

<sup>1</sup>INRS-Institut Armand-Frappier, Université du Québec, 531 Boul. des Prairies, Laval, Québec H7V 1B7, Canada.

<sup>2</sup>Department of Biochemistry and Molecular Genetics, University of Illinois at Chicago, 900 S. Ashland, Chicago, IL 60607, USA.

<sup>3</sup>Computational Biology Institute and Computer Science and Mathematics Division, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, TN 37830, USA.

<sup>4</sup>Department of Biochemistry, Cellular and Molecular Biology, University of Tennessee, Knoxville, TN 37996, USA.

<sup>5</sup>PROTEO, the Quebec Network for Research on Protein Function, Structure, and Engineering, 1045 Avenue de la Médecine, Université Laval, Québec, QC, G1V 0A6, Canada.

<sup>6</sup>GRASP, the Groupe de Recherche Axé sur la Structure des Protéines, 3649 Promenade Sir William Osler, McGill University, Montréal, QC, H3G 0B1, Canada.

---

**ABSTRACT:** The role of internal dynamics in enzyme function is a highly debated topic. In particular, how small changes in structure of regions far away from the active site alter protein dynamics (over a wide range of timescales) and overall enzyme mechanisms is of wide interest. Using ribonuclease A (RNase A) as a model system, we demonstrate that elimination of a single side-chain methyl group located >10 Å away from the site of reaction can significantly alter the conformational integrity and ligand binding properties. This A109G mutation does not perturb the thermodynamic stability of RNase A or cause significant differences in the structure, both in the apo and ligand-bound states. However, significant enhancement in conformational dynamics was observed for the ligand-bound forms of the A109G variant, as probed over nano- to milli-second timescales using a combination of NMR relaxation dispersion experiments and computational simulations. In addition, a major repositioning of the 5'-AMP was observed in the ligand-bound form in the mutant, displaying two binding orientations on the enzyme. The A109G replacement induced long-range flexibility perturbations in the active-site loop 4 (residues 65-72), resulting in alteration of purine binding and reduction in catalytic efficiency. These results illustrate the large effects caused by seemingly small changes in structure on the long-range conformational dynamics and ligand binding specificities within proteins. This study supports the emerging paradigm of the importance of preserving wild-type dynamics in enzyme systems that rely on flexibility for function, and therefore has wider significance for protein engineering and enzyme design.

---

#### **1. INTRODUCTION**

Knowledge of how proteins and enzymes function has broad implications for both fundamental and applied research. Increasing evidence suggests that proteins exist as inter-converting conformers and sample distinct conformations enabled by structural fluctuations that occur over a wide range of timescales<sup>1,2</sup>. Conformational sub-states sampled by proteins have been correlated with promoting function such as binding and catalysis in numerous enzyme systems<sup>3-8</sup>. Advances in biophysical and computational techniques now provide a significant means to probe and gain

atomistic insights into the dynamics of proteins over a wide range of timescales.

Ribonuclease A (RNase A, EC 3.1.27.5), which catalyzes the hydrolysis of single stranded RNA, has been studied extensively for over five decades and has served as a model for countless biochemical and biophysical analyses. RNase A is a small monomeric protein composed of 124 amino acids with a basic pI of 8.6. It is characterized by a V-shaped kidney structure formed by a core of  $\beta$ -sheets surrounded by loops and  $\alpha$ -helices, with the active site lying in the cleft at the bottom of the inverted  $\beta$ -sheet<sup>9</sup>. This cleft contains subsites for interactions with nucleotide bases (B<sub>1</sub>, B<sub>2</sub>,

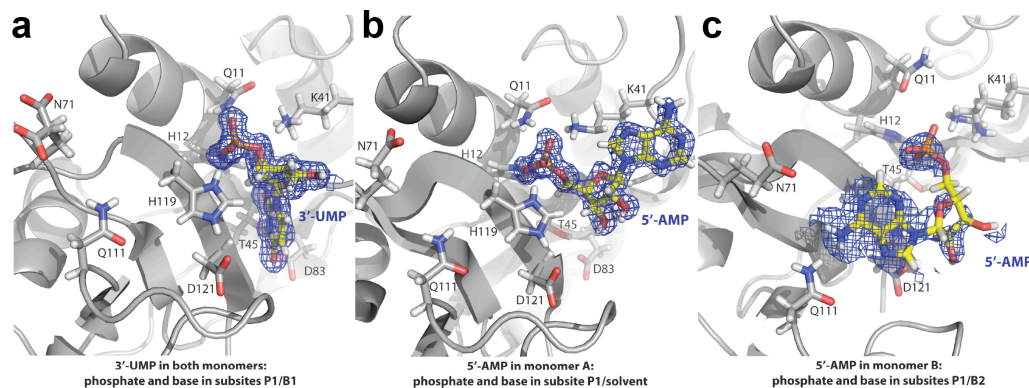
**Table 1.** Crystallographic and structure refinement statistics for the apo, 3'-UMP-, and 5'-AMP-bound forms of RNase A-A109G.

Crystal	Apo RNase A-A109G	Complex with 3'-UMP	Complex with 5'-AMP
Data Collection			
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2	P2 <sub>1</sub> 2 <sub>1</sub> 2	P2 <sub>1</sub> 2 <sub>1</sub> 2
Unit cell parameters (Å)	a = 76.2 b = 51.4 c = 58.2	a = 76.5 b = 51.8 c = 57.9	a = 76.8 b = 51.7 c = 58.2
Resolution (Å)	1.80	1.45	1.50
Total reflections	321,513	374,299	603,029
Unique reflections	21,146	39,580	35,341
Rsym (%)	17.7 (22.7)	13.3 (74.8)	14.2 (43.4)
Rpim (%)	13.4 (21.5)	6.8 (4.5)	6.8 (22.0)
CC1/2 (%)	(88.1)	(50.3)	(87.3)
I/σI	6.3 (2.4)	11.7 (1.4)	10.6 (4.1)
Completeness (%)	97.7 (91.0)	93.1 (56.9)	94.6 (87.0)
Redundancy	2.2 (1.5)	4.6 (3.0)	4.8 (4.5)
Structure Refinement			
Rwork/Rfree (%)	0.145/0.203	0.127/0.170	0.151/0.186
Protomers per ASU	2	2	2
Amino acids per ASU	249	249	249
Ligands per ASU	0	2	2
Solvent molecules	382	589	425
RMSDs			
Bond lengths (Å)	0.011	0.010	0.010
Bond angles (°)	1.288	1.445	1.297
Average B factor (Å <sup>2</sup> )	18.6	19.1	17.3
Ramachandran			
Favored/allowed (%)	98.0/2.0	97.7/2.3	97.0/3.0

B<sub>3</sub>) and phosphoryl groups (P<sub>0</sub>, P<sub>1</sub>, P<sub>2</sub>) of single-stranded RNA substrates<sup>10</sup>. B<sub>1</sub> and B<sub>2</sub> subsites preferentially bind to pyrimidine and purine bases, respectively, with a marked preference for adenine in the B<sub>2</sub> subsite<sup>11,12</sup>. Residues His12, Lys41 and His119 in the P<sub>1</sub> subsite form the catalytic triad and are crucial for catalytic activity<sup>13</sup>. In addition to these catalytic residues, a number of residues in the surrounding environment are essential to the optimal ribonucleolytic activity of the enzyme, where changes in the chemistry are known to alter substrate specificity, binding and optimal catalysis.

In recent years, NMR studies have investigated the conformational dynamics of RNase A, suggesting a link between flexibility and enzyme catalysis<sup>4,14-17</sup>. The rate constant for conformational exchange ( $k_{ex}$ ) in RNase A was shown to correlate with the rate of enzyme turnover ( $k_{cat}$ ), implicating long-range secondary structure motions with substrate turnover<sup>10,18</sup>. Long-range residue mutations have also been shown to drastically alter the catalytic activity, demonstrating the

importance of correlated residue sector dynamics in this enzyme. Alanine at position 109 is located on the β5-strand within the V<sub>1</sub> domain and is part of the hydrophobic core of the enzyme. This residue is 88% conserved within members of the RNase A superfamily and 100% conserved in human<sup>19</sup>. While it has been associated with the chain-folding initiation site<sup>20</sup>, other studies suggest a more direct involvement of this residue in the catalytic reaction, with Ala109 potentially making van der Waals interactions with the catalytic residue, His119. Using a relaxation-compensated Carr-Purcell-Meiboom-Gill (<sup>15</sup>N-CPMG) NMR experiment, Cole and Loria showed that Ala109 experiences millisecond (ms) motions in the apo form of RNase A<sup>4</sup>. This is further supported by NMR titration studies, which demonstrated that the milli-second dynamics of Ala109 is affected by the presence of ligands<sup>21,22</sup>. Mutation of A109 to a glycine has been shown to slightly lower the conformational stability of RNase A<sup>19</sup>, suggesting that flexibility at this position may be important for enzyme function.



**Figure 1.** Schematic representation of the mutant A109G in complex with 3'-UMP and 5'-AMP. Positioning of the ligands in the density map for (a) 3'-UMP (PDB 4WYZ) and (b,c) 5'-AMP (PDB 4WYP) on the 3D structure of the mutant A109G. The positioning of the 5'-AMP in c represents a low occupancy. Key residues (Gln11, His12, Lys41, Thr45, Asn71, Gln83, Gln111, His119 and Asp121) are shown in gray stick representation. Nitrogen, oxygen and phosphate atoms are colored blue, red and orange, respectively.

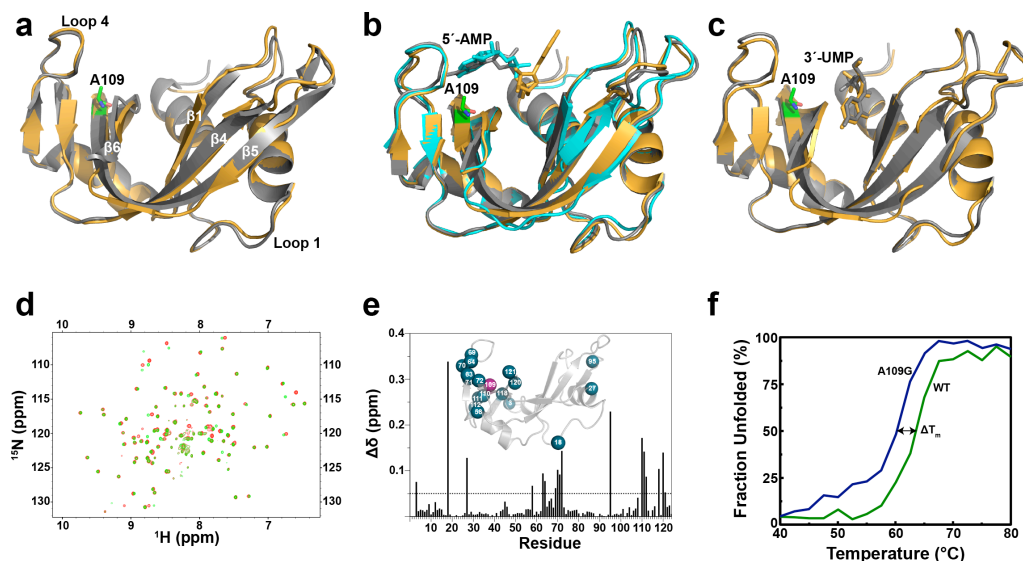
In this study, we mutated alanine for a glycine at position 109 to characterize the effect of the mutation on the global structure, dynamics, and catalytic activity in RNase A. We determined the crystal structure of the mutated apo form and in complex with two single nucleotide ligands, 5'-AMP and 3'-UMP, which bind specifically to the purine and pyrimidine binding sites in WT. These ligands are end-products of the catalytic reaction and are known to bind to the catalytic site of ribonucleases<sup>11</sup>. Conformational dynamics of the WT and A109G variant over the nano- to milli-second timescales were characterized using MD simulations and NMR <sup>15</sup>N-CPMG experiments. We show that there is no significant difference in the three-dimensional structure of the mutant relative to the WT enzyme in the apo and ligand-bound states. Nonetheless, the mutation induces a repositioning of the 5'-AMP purine, resulting in altered binding affinity caused by changes in the dynamic behavior of the enzyme. Despite minimal effects on the structure, the mutation considerably affects the rate of conformational exchange for residues in the V<sub>1</sub> domain, altering loop 4 dynamics in the nano- to milli-second timescales. These results highlight the importance of preserving WT conformational dynamics in enzyme systems that rely on flexibility for function. While single mutational replacements often have neutral effects on protein structure, they can significantly alter function by perturbing long-scale conformational dynamics and by disrupting important motional sectors on multiple timescales.

## 2. RESULTS

### X-ray crystallography and structural comparison between RNase A and A109G

We determined crystal structures of the A109G mutant in its apo form (PDBID 4WYN), as well as in complex with two single nucleotides adenosine-5'-monophosphate (5'-AMP) (PDBID 4WYP) and uracil-3'-monophosphate (3'-UMP) (PDBID 4WYZ) that bind to the purine and pyrimidine subsites, respectively, to characterize the effect of the mutation on the three-dimensional structure of RNase A (Figure 1). The crystals of the A109G mutant of RNase A belong to the orthorhombic crystal system and contain two enzyme molecules in the asymmetric unit (Table 1). Our results show that 3'-UMP binds to the canonical B<sub>1</sub> subsite in the catalytic groove of the mutant enzyme in both monomers (Figures 1A & 2C). While most of the hydrogen bonding interactions between 3'-UMP and the enzyme were conserved, a closer examination showed that an additional interaction with His119 was acquired in the mutant enzyme, presumably due to a slight repositioning of the phosphate group (Figure 2C).

In striking contrast, we found a variation in the position of 5'-AMP in the catalytic pocket of the A109G mutant, with the ligand adopting two distinct orientations (Figures 1B-C). While it is abundantly documented that the adenine base of single-stranded RNA substrates bind to the B<sub>2</sub> subsite in RNase A<sup>10</sup>, our crystal structure shows that the base of 5'-AMP faces the solvent and not the B<sub>2</sub> subsite when bound to monomer A (Figures 1B & 2B). In this particular complex, the adenine and ribose rings of 5'-AMP were



**Figure 2.** Structural comparison between WT RNase A and mutant A109G. Schematic overlays of (a) apo forms of WT RNase A (grey, PDB 7RSA) and A109G (orange, PDB 4WYN), (b) 5'-AMP-bound WT RNase A (grey, PDB 1Z6S), soaked A109G (orange, PDB 4WYP), (c) 3'-UMP-bound WT RNase A (grey, PDB 1O0N) and A109G (orange, PDB 4WYZ). The A109G site is indicated using green sticks representation. 5'-AMP and 3'-UMP are also displayed using stick representation. (d) Overlay of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra for the WT RNase A (red) and A109G mutant (green). (e) Mapping of chemical shift differences ( $\Delta\delta$ ) resulting from the A109G mutation on the primary structure of RNase A.  $^1\text{H}$  and  $^{15}\text{N}$  composite chemical shift differences ( $\Delta\delta$ ) were calculated according to the following equation<sup>23</sup>:  $\Delta\delta$  (ppm) =  $[(\Delta\delta_{\text{HN}}^2 + \Delta\delta_{\text{N}}^2/25)/2]^{1/2}$ . The inset shows residues with  $\Delta\delta > 0.05$  ppm (blue spheres) highlighted on the three-dimensional structure of A109G (PDB 4WYN). (f) Temperature unfolding for WT (green) and A109G mutant (blue) RNase A determined by circular dichroism.

rotated  $\sim 180^\circ$  relative to the canonical conformation observed in the complex with the wild-type enzyme. Instead of being in the  $B_2$  subsite, the adenine ring sits atop the side chain of Lys41. Consequently, hydrogen bonds between the purine base and residues Asn67, Gln69 and Asn71, and the stacking interaction between the purine base and the imidazole ring of His119 were absent in the crystal of the A109G mutant. Intriguingly, 5'-AMP was bound in a distinct orientation in the second monomer of the asymmetric unit, which we shall refer to as monomer B (Figures 1C & 2B). During later stages of the structure refinement, relatively strong positive peaks in the mFo-DFc difference electron density map were observed near subsites  $B_2$  and  $P_1$  in monomer B. We interpreted the density as adenine ring and phosphate, and thus modeled and refined 5'-AMP. Because the occupancy of this particular ligand refined to 0.76, it is reasonable to suggest that this alternate orientation is not a preferred conformation of 5'-AMP by the A109G mutant of RNase A.

#### Chemical shift variations between RNase A and A109G

An overlay of the NMR  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of both the WT and A109G RNases showed variations in the

chemical shifts for a number of residues (Figure 2D). Chemical shift variations were calculated as differences between shifts of the mutant and WT RNases, with variations  $>0.05$  ppm being considered significant. Residues in close proximity to the mutation site displayed the largest variation (Figure 2E). These include Asn71 and Asp121, residues important for interaction with the ligands in the  $B_2$  and  $P_1$  subsites, respectively. Further, residues in loop 4 (residues 65-72), which are highly dynamic<sup>22</sup> and have been shown to be important for the binding of the substrate<sup>18</sup>, showed significant variations relative to WT.

#### Thermal unfolding of RNase A and A109G

Characterization of the WT and A109G apo forms was also performed using circular dichroism (CD) at 25°C. Both proteins exhibited identical spectral characteristic of  $\alpha$ -helical and  $\beta$ -sheet structures (not shown)<sup>24,25</sup>. Thermal unfolding was performed by increasing the temperature up to 80°C, with changes in the secondary and tertiary structures monitored using CD at 210 nm. A melting temperature ( $T_m$ ) of  $63.9 \pm 0.4^\circ\text{C}$  was observed for the WT protein, consistent with values reported previously<sup>24,25</sup>. A comparable melting temperature of  $61.4 \pm 1.1^\circ\text{C}$  was observed for the A109G mutant, suggesting that the mutation does not

**Table 2.** Kinetic parameters for WT RNase A and variant A109G against yeast tRNA at 25°C.

	$K_m$ ( $\mu\text{M}$ )	$K_m$ relative to WT	$k_{\text{cat}}$ ( $\text{s}^{-1}$ )	$k_{\text{cat}}$ relative to WT	$k_{\text{cat}}/K_m$ ( $\text{M}^{-1} \text{s}^{-1}$ ) $\times 10^6$	$k_{\text{cat}}/K_m$ relative to WT
WT	$64 \pm 1$	1	$3759 \pm 193$	1	$58 \pm 2$	1
A109G	$45 \pm 5$	0.7	$930 \pm 237$	0.15	$21 \pm 6$	0.36

induce significant effects on the stability of the protein (Figure 2F).

#### Binding kinetics and catalytic efficiency of RNase A and A109G

We used isothermal titration calorimetry (ITC) to characterize the kinetics of binding for WT and A109G in presence of both 5'-AMP and 3'-UMP ligands. The thermodynamic parameters and the dissociation constant  $K_d$  were similar in the presence of 3'-UMP (WT- $K_d$ (3'-UMP) =  $67 \pm 6 \mu\text{M}$ ; A109G- $K_d$ (3'-UMP) =  $50 \pm 8 \mu\text{M}$ ), a reflection of the minor changes observed in the 3D structure. In contrast, binding of 5'-AMP is associated with a reduced binding affinity (WT- $K_d$ (5'-AMP) =  $105 \pm 1 \mu\text{M}$ ; A109G- $K_d$ (5'-AMP) =  $440 \pm 26 \mu\text{M}$ ) as well as an increase in  $\Delta H$ , corresponding to the higher energy required for binding to the ligand. We also performed steady-state kinetics with yeast transfer RNA (ytRNA) to complement the ITC data (Table 2). The catalytic efficiency ( $k_{\text{cat}}/K_m$ ) of the mutant decreased 3-fold, a result caused by a 4-fold reduction in the catalytic rate ( $k_{\text{cat}}$ ). While the binding of a longer RNA strand did not seem to be impaired by the mutation, this result confirms an indirect involvement of alanine 109 in catalysis.

#### Conformational dynamics on the microsecond-to-millisecond timescale

Many residues in RNase A experience a conformational exchange that can be described by a two-state global motion inter-conversion process with an exchange rate constant ( $k_{\text{ex}}$ ) of  $\sim 1700 \text{ s}^{-1}$ . This  $k_{\text{ex}}$  is similar to the catalytic constant ( $k_{\text{cat}}$ ) and the rate-limiting product release ( $k_{\text{off}}$ ) in RNase A, suggesting a potential correlation between conformational exchange and catalytic function in this enzyme<sup>14-16</sup>. In an effort to provide insight into the microsecond-to-millisecond ( $\mu\text{s}$ -ms) dynamic changes induced by the A109G replacement, we performed relaxation-compensated Carr-Purcell-Meiboom-Gill (<sup>15</sup>N-CPMG) experiments on the apo form and in complex with the two ligands (5'-AMP and 3'-UMP) (Figure 3). In the free forms, a total of 17 residues showed conformational exchange in both the WT and A109G variant. Residues experiencing motion on this timescale show clear relaxation dispersion profiles that are localized in clusters 1 and 3 in both proteins, as defined previously<sup>22</sup>. The global exchange rate ( $k_{\text{ex}}$ ) of  $1621 \pm 152 \text{ s}^{-1}$  calculated for the free WT RNase A is consistent with results previously reported<sup>14-16</sup>. The A109G variant displayed enhanced global conformational exchange, a shift attributed to increased ms dynamics in the loop 4 environment (cluster 1) (Figure 3, Table 3). Indeed, while a few residues of loop 4 in A109G either become

rigid or experience motions outside the <sup>15</sup>N-CPMG time frame, global  $k_{\text{ex}}$  for the dynamic residues of this important active-site loop that are involved in purine recognition (Lys66, Gln69, Asn71) is significantly increased in A109G ( $k_{\text{ex}} = 3649 \pm 473 \text{ s}^{-1}$ ) relative to WT ( $k_{\text{ex}} = 1548 \pm 82 \text{ s}^{-1}$ ) (Table 3).

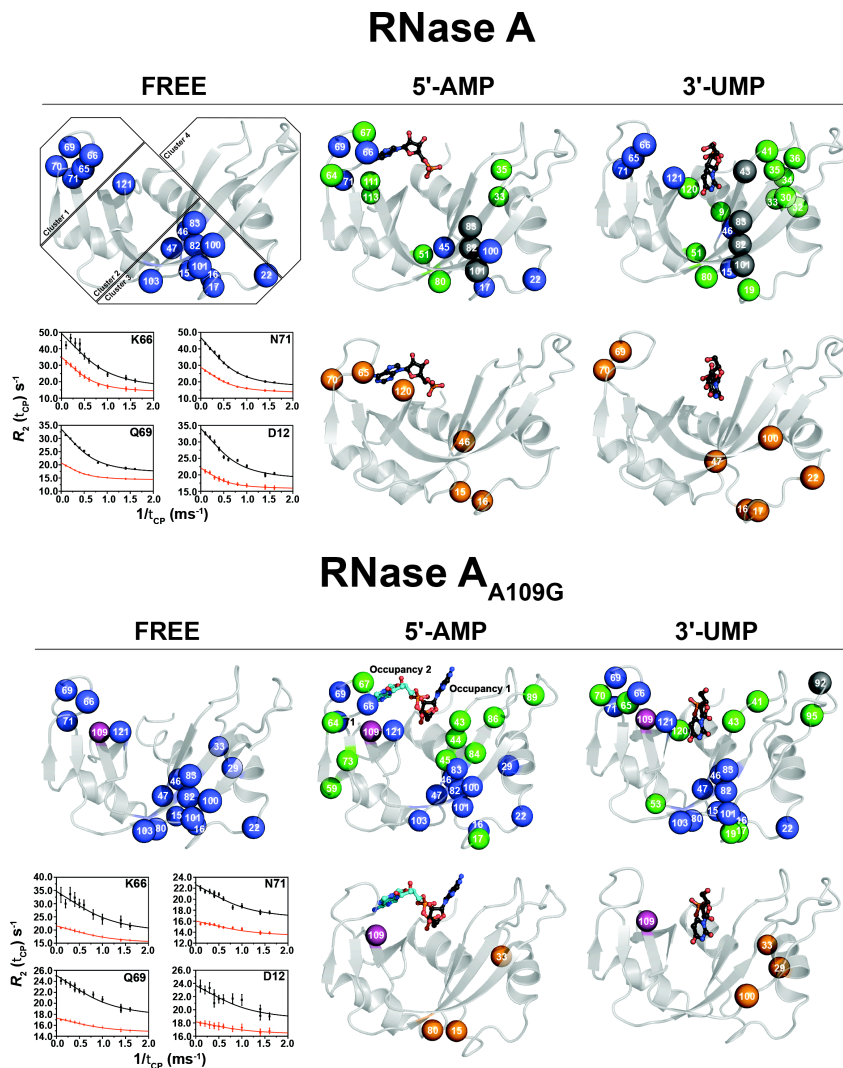
A total of 16 and 22 residues exhibited conformational dynamics in the 5'-AMP-bound WT and mutant RNases, respectively (Figure 3). A gain in ms dynamics is predominantly observed in clusters 1, 3 and 4 in both proteins. This is particularly true for the mutant, with residues Val43, Asn44, Thr45, Cys84, Glu86 and Ser89 showing relaxation dispersion profiles upon ligand binding (Figure 3, green spheres). Repositioning of the 5'-AMP ligand modified the network of dynamic residues in cluster 4, especially on  $\beta_1$  (residues 43-47), which now faces the 5'-AMP. The global  $k_{\text{ex}}$  for the A109G protein variant shows a four-fold increase relative to WT. While there is a gain in dynamics for a number of residues in cluster 3 of the mutant, we note that the average  $k_{\text{ex}}$  is comparable for both proteins. The difference in  $k_{\text{ex}}$  is quite striking for cluster 1, which shows a two-fold increase in the ms dynamics for the mutant (Table 3). In accordance with the repositioning of 5'-AMP in the A109G variant, these results support the fact that enhanced ms dynamics in cluster 1 (loop 4) hinder the proper positioning of 5'-AMP in the B<sub>2</sub> subsite, resulting in the repositioning of the ligand to the B<sub>1</sub> subsite.

A major reorganization of the residues displaying significant flexibility on the ms time frame is observed in the 3'-UMP-bound RNase A, a result not seen in the A109G mutant (Figure 3). In the WT, only 6 residues preserved ms motions when compared to the apo form (blue spheres), while a gain of dynamics upon ligand binding is observed for 12 additional residues (green spheres). These additional residues are located primarily on  $\alpha_2$  (residues 30-36 on cluster 3), a region devoid of ms dynamics in the apo form. The mutant protein shows a gain in conformational exchange for 9 residues (green spheres), in addition to 14 residues that retain the motions of the apo form (blue spheres), most of them localized in cluster 3. As observed with the free and 5'-AMP-bound forms, cluster 1 in A109G shows a two-fold increase in  $k_{\text{ex}}$ , while the average dynamics in cluster 3 are comparable to WT (Table 3).

#### Conformational dynamics on the nanosecond-to-microsecond timescale

Molecular dynamics (MD) simulations were performed for WT and A109G in complex with 5'-AMP and 3'-UMP. Ligands with an intact phosphodiester bond between the single nucleotides were prepared to





**Figure 3.** Effects of 5'-AMP and 3'-UMP binding on the millisecond dynamics of WT RNase A and mutant A109G. Millisecond timescale (ms) conformational exchange experienced by apo, 5'-AMP- and 3'-UMP-bound forms of WT RNase A and mutant A109G. Motions were probed by  $^{15}\text{N}$ -CPMG relaxation dispersion experiments at 500 MHz and 800 MHz (298 K). Residues were considered for further analysis only if the difference in measured  $R_2$  ( $1/T_{cp}$ ) values at fast ( $T_{cp} = 0.625$  ms) and slow ( $T_{cp} = 10$  ms) pulsing rates was greater than  $2 \text{ s}^{-1}$ , similar to previous reports<sup>4,15,22</sup>. Blue spheres: residues showing  $^{15}\text{N}$ -CPMG dispersion profiles with  $\Delta R_2$  ( $1/T_{cp}$ )  $> 2 \text{ s}^{-1}$ ; orange spheres: residues showing no conformational exchange relative to the apo form (*i.e.*, dampened ms dynamics upon ligand binding) or outside of the  $^{15}\text{N}$ -CPMG time frame; green spheres: residues gaining conformational exchange [ $\Delta R_2$  ( $1/T_{cp}$ )  $> 2$ ] upon ligand binding; gray spheres: assigned residues in the apo form that cannot be assigned in the ligand-bound form due to line broadening; purple spheres: position 109. Residues are highlighted on the 3D structure of WT RNase A (apo form: PDB 7RSA; 3'-UMP-bound: PDB 1OON; 5'-AMP-bound: PDB 1Z6S) and mutant A109G (apo form: PDB 4WYN; 3'-UMP-bound: PDB 4WYZ; 5'-AMP-bound: PDB 4WYP). Insets show representative dispersion curves acquired at 500 MHz (red) and 800 MHz (black) for a subset of loop 4 (cluster 1) in apo forms of WT and A109G. Residue clusters are identified according to<sup>22</sup>.

**Table 3.** Conformational exchange rates ( $k_{ex}$ ) for WT RNase A and mutant A109G at 25°C.

	$k_{ex}(s^{-1})$		
	Cluster 1 (loop 4) <sup>1</sup>	Cluster 3	Global
<u>Apo</u>			
WT	1548 ± 82	1438 ± 125	1621 ± 152
A109G	3649 ± 473	1115 ± 119	3145 ± 1651
<u>5'-AMP</u>			
WT	687 ± 41	1398 ± 108	750 ± 111
A109G	1580 ± 143	1669 ± 279	3328 ± 543
<u>3'-UMP</u>			
WT	1136 ± 46	1639 ± 128	1375 ± 50
A109G	2234 ± 523	1189 ± 154	1463 ± 216

RNase A dynamic residue clusters are defined as in ref. 22.

simulate the reactant state while single nucleotides without the phosphodiester bond were used to mimic the product state. Root mean square fluctuations (RMSF) were calculated for each residue, over a 1- $\mu$ s simulation trajectory for WT and variant A109G. To quantitatively assess the difference in global conformational dynamics at nanosecond to microsecond (ns- $\mu$ s) time-scale for the WT and A109G, we characterized the top ten eigenmodes, using quasi-harmonic analysis<sup>26,27</sup>. Primary flexibility of the WT and A109G proteins in the reactant and product states are shown in Figure 4. On the  $\mu$ s time-scale, the mutant showed reduction of dynamics in loop 4, while significantly enhanced dynamics was observed for  $\alpha$ 2 and part of loop 1 relative to the WT for the reactant state. In contrast, the product state showed dynamics similar to that of the WT for most of the protein. There was a marginal reduction in the conformational dynamics of loop 1 (residues 16-25) in the mutant enzyme (Figure 4), consistent with previous studies which showed that the dynamics of loop 1 are correlated with efficient product release in RNase A<sup>15</sup>. We note that MD simulation results were obtained using dinucleotide substrate and corresponding cleaved products states while the NMR relaxation dispersion experiments were performed using mononucleotide ligands. Overall, the computational studies indicate that the A109G mutant shows significant increase in flexibility of the distal regions of the enzyme in the ground state for the bound reactant complex on the  $\mu$ s time-scale. Further studies are required to investigate the role of increased flexibility in the ground reactant in accessing functionally relevant conformational sub-states that have catalytically competent structural and dynamical features for the substrate cleavage<sup>28</sup>.

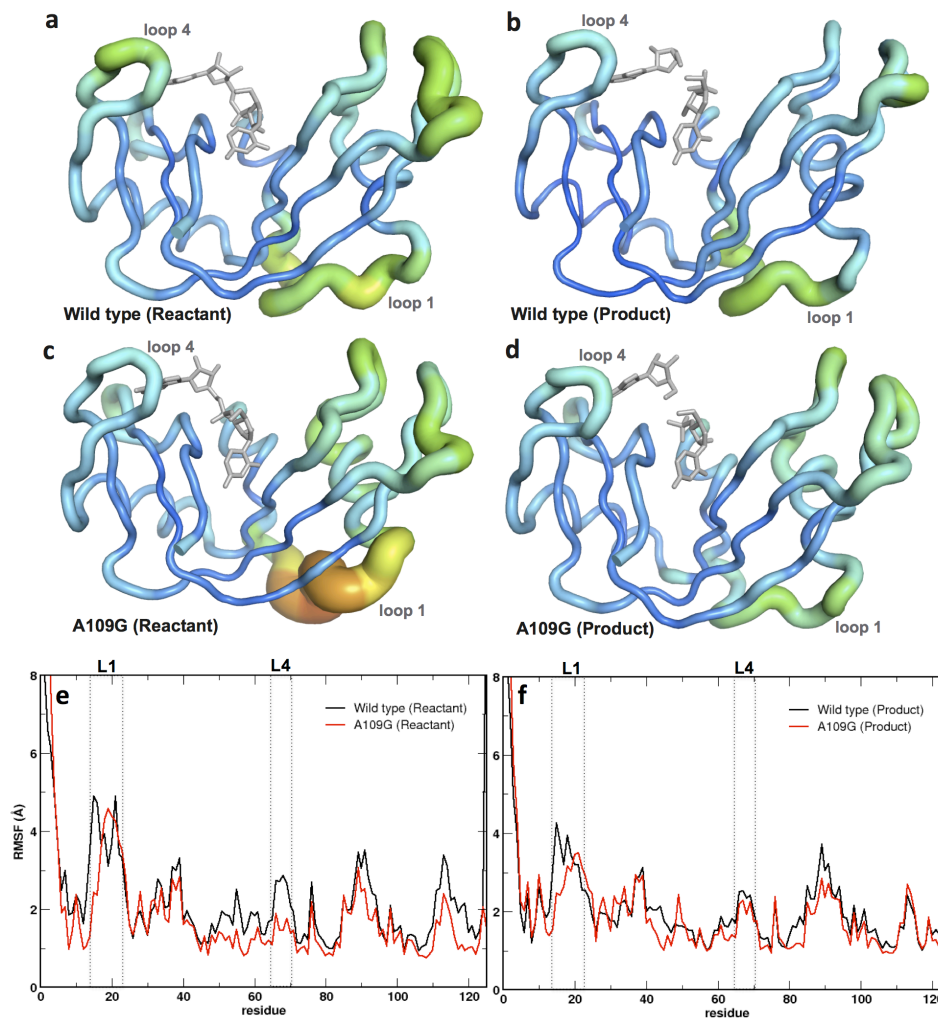
### 3. DISCUSSION

Mutations located away from the active site have been shown to drastically alter the catalytic activity of

enzymes. In this study, we characterized the effect of a single residue replacement of the alanine at position 109 to glycine on the structure, dynamics and catalytic activities of RNase A. We obtained crystal structures of the A109G mutant in the apo form and in presence of two end-products of the reaction, 5'-AMP and 3'-UMP, to characterize the effect of the mutation on ligand binding. Our results show that the mutation does not significantly perturb the three-dimensional structure of the protein, with only minor differences in loop 1, which has been shown to exert long-range effects on RNase A catalysis<sup>14,22</sup>.

While the structural differences between the WT and mutant enzymes were minimal, we observed a major repositioning of 5'-AMP ligand from the B<sub>2</sub> subsite located near loop 4, to a location where the adenine base faces the solvent rather than the enzyme. As a result of this repositioning, the purine base now interacts with Lys41 instead of its usual partners (Asn67, Gln69 and Asn71). Interactions between the phosphate group of the nucleotide and residues Gln11, His12, His119 and Phe120 are preserved in the mutant (Figures 1A & 1B). It is important to note that a smaller fraction of 5'-AMP interacts with the mutant enzyme in a manner reminiscent of the canonical mode of binding. However, even in this orientation, the adenine base is not optimally positioned within the B<sub>2</sub> subsite, but rather adopts a position partially between the B<sub>1</sub> and B<sub>2</sub> subsites. Apart from the His119, all other interactions are lost in this orientation of the ligand (Figures 1A & 1C). These results emphasize the large effect of a single distal residue substitution on the ligand binding affinity of the protein.

We characterized the effect of the A109G mutation on the conformational dynamics of the protein by performing a series of biophysical experiments and computational analyses. We used relaxation dispersion<sup>15</sup>N-CPMG experiments to extract the rate exchange ( $k_{ex}$ ) for each residue experiencing conformational dynamics on the millisecond timescale. Previous studies demonstrated that <sup>15</sup>N-CPMG provides relevant



**Figure 4.** Root mean square fluctuations (RMSFs) for WT RNase A and A109G mutant for a 1- $\mu$ s molecular dynamics simulation. Enzyme backbone flexibility was determined from the top ten eigenmodes of the RMSFs. (a-d) RMSF values are shown using tube representations for the WT (a, b) and A109G mutant (c, d) RNase A for the reactant (left) and product (right) states. Thickness of the tube corresponds to the flexibility of the residues, with thicker tubes (red end of the spectrum) representing more flexible regions and thinner tubes (blue end of the spectrum) denoting less flexible regions in the protein. (e-f) Comparison of the RMSFs as a function of sequence for the reactant (e) and product (f) states for the WT (black) and mutant (red) proteins. Loop 1 and loop 4 (labeled L1 and L4, respectively) regions are represented using dotted lines. Panels A-D depict RMSF based on  $C_{\alpha}$  while panels E-F are based on all atom calculations for each residue.

information in relation to the function of RNase A<sup>3,4,14,22</sup>. Significant changes in the dynamic profiles are observed for the 5'-AMP-bound state of the A109G mutant, a reflection of ligand repositioning (Figure 3). We note that the 3'-UMP-bound form displays very

different dynamics for the mutant protein, a change that is not reflected in the three-dimensional structure of the protein. Interestingly, chemical shift differences for the main catalytic residue His12 upon binding to both ligands are considerably affected in the mutant. His12

is known to be rigid on the fast timescale<sup>29</sup>, which might be a requirement for proper catalysis. These results suggest the existence of different conformational pre-equilibria between the ligand-bound forms in the WT and mutant RNase A.

We performed global fit  $k_{ex}$  calculations for residues located in pre-defined clusters corresponding to previously described dynamic sectors in RNase A<sup>22</sup>.  $k_{ex}$  values provide valuable information on the conformational dynamics of individual residue clusters, which can be correlated with the catalytic activity of the protein. We observed a significant increase of the  $k_{ex}$  in cluster 1 (loop 4) for the mutant in both the apo and ligand-bound forms. These results are supported by lower  $R_{ex}$  for all residues in this cluster displaying significant conformational flexibility (Figure 3). Interestingly, similarly enhanced conformational dynamics of loop 3 (cluster 1) were also observed on the nano- to micro-second timescales probed using MD simulations (Figure 4). These results suggest that the enhanced dynamics observed for loop 4 in the mutant may hinder the interaction of the ligand with its hydrogen-bonding partners (Asn67, Gln69 and Asn71) in the B<sub>2</sub> subsite. It was previously shown that Thr45 forms hydrogen-bonding interactions with the pyrimidine base while sterically excluding the purine base from this subsite<sup>30</sup>. As a result, the 5'-AMP adopts a position energetically unfavorable between the B<sub>1</sub> and B<sub>2</sub> subsites and exposed to solvent. We found no significant effect of the enhanced dynamics of loop 4 on the location of 3'-UMP within the structure of the A109G variant, demonstrating that pyrimidine binding is specific to B<sub>1</sub> and independent from purine binding in B<sub>2</sub>.

#### 4. CONCLUSION

In conclusion, our results show that mutation of alanine at position 109 to a glycine residue does not perturb the 3D backbone structure of RNase A in the apo form and in complex with two single nucleotides, 5'-AMP and 3'-UMP. Nevertheless, a major repositioning of 5'-AMP to the solvent between the B<sub>1</sub> pyrimidine-specific and B<sub>2</sub> purine-specific subsites is observed, supporting the role of multiple timescale changes in the dynamics of the protein induced by this single alanine to glycine replacement. While there is no significant change in structure, the mutation induced large changes in the dynamics of the mutant relative to the WT RNase A. Enhanced dynamics of distal protein regions are observed on the nano- and milli-second timescales for the mutant system determined from MD simulations and NMR relaxation experiments. In particular, conformational dynamics of loop 4 is also significantly enhanced in the apo and ligand-bound forms of the A109G mutant, offering an explanation for the modest yet significant changes observed in binding and catalysis. These results showcase the large effects caused by seemingly modest residue substitutions on

the long-range conformational dynamics and residue binding specificities within proteins. This study further underlines the importance of preserving wild type dynamics in enzyme systems that rely on flexibility for function, which is of critical significance to protein engineering and enzyme design. This study highlights the effect of single-point mutation in RNase A on the shift in the conformational dynamics, and its correlation with the catalytic efficiency. As preliminary evidence already indicates, these results also provide new insights suggesting that modulating conformational dynamics may be a mechanism used by nature for regulating enzyme activity, thus providing new opportunities for developing new enzyme engineering strategies<sup>31</sup>.

#### 5. METHODS

**DNA Constructs.** Oligonucleotide synthesis was performed by Bio Basic (Markham, ON). The WT, 372-bp bovine RNase A gene inserted into pET22b(+) and codon-optimized for *Escherichia coli* expression was a generous gift from J. Patrick Loria (Yale University, New Haven, CT). The A109G mutation was generated by the QuikChange site-directed mutagenesis protocol (Agilent Technologies, Mississauga, ON) using the two following complementary primers: A109G-Forward, 5'-GCGAATAAACATATTATTGTTGGCTGCGAAGGCAA TCCGTATGTG-3'; A109G-Reverse, 5'-CACCTACGGATTGCCTTCGCAGCCAACAATAATAT GTTTATTTCGC-3'. The amplified PCR product was *DpnI*-treated and transformed into *E. coli* BL21(DE3) for all protein expression procedures. Colonies were selected and the presence of the mutation was confirmed by DNA sequencing (G enome Qu ebec, Montr al, QC).

**Protein Expression and Purification.** <sup>15</sup>N-labeled samples were prepared by growing *E. coli* BL21(DE3) in M9 minimal medium and enzymes were purified as previously described<sup>15</sup>. Protein concentration was determined using an extinction coefficient of 9,800 M<sup>-1</sup>cm<sup>-1</sup><sup>32</sup>.

**Crystallization and structure determination.** Crystals of apo RNase A-A109G were obtained by the sitting drop vapor-diffusion method. Equal volumes of the protein solution (~30 mg/mL) and the well buffer were mixed, and the drops were incubated at room temperature (21 C). Crystals were obtained using the NeXtal PACT Suite (Qiagen). The high-resolution crystals of RNase A-A109G were obtained from either 0.1M MIB Buffer, pH 8.0 and 25% (w/v) PEG 1500; 0.1M MIB Buffer, pH 6.0 and 25% (w/v) PEG 1500; or 0.1M HEPES, pH 7.0, 0.2M NaCl and 20% (w/v) PEG 6000. The binary complex crystals were obtained by soaking apo enzyme crystals with 50mM of either 3'-UMP (Chem-Impex International) or 5'-AMP (Sigma-Aldrich). All crystals were cryoprotected either with 15% (w/v) PEG 400 or 20% ethylene glycol prior to X-ray exposure. Data were collected at liquid nitrogen

temperature at the Southeast Regional Collaborative Access Team 22-ID (SER-CAT) beamline at the Advanced Photon Source, Argonne National Laboratory. The diffraction data were processed in HKL2000<sup>33</sup>. The crystal structure of apo RNase A-A109G was determined by molecular replacement in Phaser<sup>34</sup> using the wild-type bovine RNase A structure (PDB ID code 1KF5) as a search model. The refined apo RNase A-A109G was used to phase other RNase A-A109G crystals. The structure refinement was performed in Phenix<sup>35</sup>, and the model building was done in Coot<sup>36,37</sup>. All figures were produced in PyMOL (The PyMOL Molecular Graphics System, Version 1.2, Schrödinger, LLC).

**NMR Experiments.** All NMR experiments were carried out on Agilent 500 MHz and 800 MHz NMR spectrometers equipped with triple-resonance cold probes and pulsed-field gradients. NMR spectra were recorded at 298 K on samples containing 0.75 mM <sup>15</sup>N-labeled RNase A (or the A109G variant) in 5 mM MES-NaOH buffer (pH 6.4) with 7 mM NaCl, 0.01% Na<sub>3</sub>, and 10% <sup>2</sup>H<sub>2</sub>O. Backbone resonances were assigned by standard multidimensional experiments<sup>38</sup>. Relaxation-compensated <sup>15</sup>N-CPMG experiments were acquired in an interleaved fashion with  $\tau_{cp}$  <sup>15</sup>N-CPMG repetition delays of 0.625, 0.714 ( $\times 2$ ), 1.0, 1.25, 1.67, 2.0, 2.50 ( $\times 2$ ), 3.33, 5.0, and 10 ms, using a total relaxation period of 40 ms. All NMR spectra were processed using NMRPipe<sup>39</sup>, in-house scripts and analyzed with Sparky<sup>40</sup>. Global residue fits and model analyses were performed by fitting 500 MHz and 800 MHz relaxation dispersion data to the full single-quantum <sup>15</sup>N-CPMG equation<sup>41</sup> using GraphPad Prism 5. NMR titration experiments were performed as previously described<sup>22</sup>, using the single-nucleotide ligands 3'-UMP and 5'-AMP reconstituted in the NMR buffer.

**Molecular Dynamics Simulations.** RNase A-ligand complexes were modeled from the PDB structure 1U1B, using only the coordinates of chain A. The ligands were modeled based on a related PDB structure 1RCN. Residue 109 was computationally mutated from alanine to glycine by removing the methyl group to generate the mutant structure. Pre-processing steps and simulations were performed using the AMBER simulation package<sup>42</sup>. Simulations were performed with AMBER ff98 force-field and SPC/E water model. For system preparation, the protein-ligand complex was solvated and the system was neutralized through the addition of four Cl<sup>-</sup> counterions. After the pre-processing steps, the system was equilibrated using the protocol described previously<sup>43</sup>. All production runs were performed at 300K under NVE (constant volume and energy) conditions. The WT and mutant structures in complex with ligands were simulated for 1  $\mu$ s for each system (after 50 ns of additional equilibration). AMBER's GPU-enabled *pmemd* simulation engine was used for equilibration and production runs. Backbone (C) and all-atom

flexibility of simulation trajectories was determined from the RMSF, computed by aggregating the magnitude of displacement eigenmodes computed using the quasi-harmonic analysis (QHA) in the *ptraj* analysis module in AMBER. As described previously<sup>44</sup>, only the top 10 QHA modes were used in the analysis to focus on the principal dynamics or long time-scale fluctuations in the proteins.

**Isothermal Titration Calorimetry (ITC).** WT RNase A and mutant A109G enzymes were prepared in 20 mM MES, pH 6.4 at a concentration of 0.2 mM. Freshly prepared 3'-UMP and 5'-AMP were reconstituted in the same buffer at a concentration of 3 mM (15-fold higher than the protein concentration). All experiments were performed on a MicroCal MCS microcalorimeter (GE Healthcare, Piscataway, NJ) at 25°C. A minimum of three repeats was performed for each ligand/enzyme combination. A total of 16 injections (2.5  $\mu$ L each) were used for the experiment, with a delay of 130 seconds between each injection. Values for  $n$ ,  $\Delta H$ ,  $\Delta S$  and  $K_d$  were calculated by fitting  $\Delta Q$  for each injection to a sigmoidal curve in Origin 7.0 (OriginLab, Northampton, MA) using a one-site model, as previously described<sup>45</sup>.

**Circular dichroism spectroscopy (CD).** CD spectra were recorded using a Jasco J-815 spectropolarimeter equipped with a thermostated cell holder. Proteins were reconstituted at a concentration of 50  $\mu$ M in phosphate-buffered saline solution (PBS, pH 7.4) in the mid-UV (170-290 nm) in a quartz cuvette with 0.1-cm optical path length. CD spectra were recorded at a scan speed of 100 nm/min, a 2-nm bandwidth, and a response time of 2 sec at a temperature of 25°C. The sample compartment was purged with nitrogen, and spectra were averaged over 5 scans. The contribution of the solvent to the spectra was subtracted using the Spectral Management software (Jasco).

**Thermal Unfolding.** CD measurements were made with a Jasco J-815 spectropolarimeter (Jasco Analytical Instruments) using a quartz cell with 0.1-cm optical path lengths. Proteins were reconstituted at a concentration of 50  $\mu$ M in Phosphate-Buffered Saline (pH 7.4). The ellipticity at 210 nm was measured at 2.5°C intervals, with 5 seconds equilibration time, up to 80°C. The denaturation temperature,  $T_m$ , corresponding to the temperature by which half of the tertiary structure at a specific wavelength (210 nm) is denatured, was determined with Spectral Analysis (Jasco Analytical Instruments).

**Steady-state kinetics on yeast tRNA.** The kinetics of yeast tRNA (Sigma-Aldrich) cleavage was performed by isothermal titration calorimetry (ITC) using the continuous technique and according to the method described by Todd & Gomez<sup>46</sup>. Each assay was run by injecting 1 mM ytRNA into 1 pM (WT) or 100 pM (A109G) enzyme. Both enzymes and substrate were solubilized in 20 mM MES buffer at pH 6.4. The assays were run at 25°C for 5 minutes. The steady-state kinetic parameters were calculated based on the

method described in ref. <sup>47</sup>, assuming an average molecular weight of 28,100 for the tRNA <sup>48</sup>.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [nicolas.doucet@iaf.inrs.ca](mailto:nicolas.doucet@iaf.inrs.ca), Phone: (450) 687-5010, ext. 4212.

### Funding Sources

This work was supported by the National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health (NIH) under award number R01GM105978 (to N.D. and P.K.A.) and a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant under award number RGPIN 402623-2011 (to N. D.). D.G. holds a NSERC Alexander Graham Bell Canada Graduate Scholarship. N.D. holds a Fonds de Recherche Québec – Santé (FRQS) Research Scholar Junior 1 Career Award.

## ACKNOWLEDGEMENTS

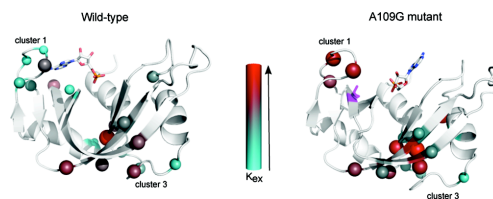
The authors thank Tara Sprules and Sameer Al-Abdul-Wahid of the Quebec/Eastern Canada High Field NMR Facility (McGill University) for their excellent technical assistance, in addition to Laurie-Anne Charest for her help with NMR titration experiments.

## REFERENCES

- (1) Frauenfelder, H.; Sligar, S.; Wolynes, P. *Science* **1991**, *254*, 1598.
- (2) Frauenfelder, H.; Leeson, D. T. *Nature structural biology* **1998**, *5*, 757.
- (3) Beach, H.; Cole, R.; Gill, M. L.; Loria, J. P. *J Am Chem Soc* **2005**, *127*, 9167.
- (4) Cole, R.; Loria, J. P. *Biochemistry* **2002**, *41*, 6072.
- (5) Agarwal, P. K. *J Am Chem Soc* **2005**, *127*, 15248.
- (6) Eisenmesser, E. Z.; Millet, O.; Labeikovsky, W.; Korzhnev, D. M.; Wolf-Watz, M.; Bosco, D. A.; Skalicky, J. J.; Kay, L. E.; Kern, D. *Nature* **2005**, *438*, 117.
- (7) Wang, L.; Tharp, S.; Selzer, T.; Benkovic, S. J.; Kohen, A. *Biochemistry* **2006**, *45*, 1383.
- (8) Baldwin, A. J.; Walsh, P.; Hansen, D. F.; Hilton, G. R.; Benesch, J. L.; Sharpe, S.; Kay, L. E. *J Am Chem Soc* **2012**, *134*, 15343.
- (9) Wlodawer, A.; Svensson, L. A.; Sjolín, L.; Gilliland, G. L. *Biochemistry* **1988**, *27*, 2705.
- (10) Gagne, D.; Doucet, N. *FEBS J* **2013**.
- (11) Raines, R. T. *Chem Rev* **1998**, *98*, 1045.
- (12) Cuchillo, C. M.; Moussaoui, M.; Barman, T.; Travers, F.; Nogue, M. V. *Protein Sci* **2002**, *11*, 117.
- (13) Pares, X.; Nogue, M. V.; de Llorens, R.; Cuchillo, C. M. *Essays Biochem* **1991**, *26*, 89.
- (14) Doucet, N.; Khirich, G.; Kovrigin, E. L.; Loria, J. P. *Biochemistry* **2011**, *50*, 1723.
- (15) Doucet, N.; Watt, E. D.; Loria, J. P. *Biochemistry* **2009**, *48*, 7160.
- (16) Watt, E. D.; Shimada, H.; Kovrigin, E. L.; Loria, J. P. *Proc Natl Acad Sci U S A* **2007**, *104*, 11981.
- (17) Khirich, G.; Loria, J. P. *J Phys Chem B* **2015**, *119*, 3743.
- (18) Loria, J. P.; Berlow, R. B.; Watt, E. D. *Acc Chem Res* **2008**, *41*, 214.

- (19) Smith, B. D.; Raines, R. T. *J Mol Biol* **2006**, *362*, 459.
- (20) Torrent, J.; Rubens, P.; Ribo, M.; Heremans, K.; Vilanova, M. *Protein Sci* **2001**, *10*, 725.
- (21) Toiron, C.; Gonzalez, C.; Bruix, M.; Rico, M. *Protein Sci* **1996**, *5*, 1633.
- (22) Gagne, D.; Charest, L. A.; Morin, S.; Kovrigin, E. L.; Doucet, N. *J Biol Chem* **2012**, *287*, 44289.
- (23) Grzesiek, S.; Stahl, S. J.; Wingfield, P. T.; Bax, A. *Biochemistry* **1996**, *35*, 10256.
- (24) Seshadri, S.; Oberg, K. A.; Fink, A. L. *Biochemistry* **1994**, *33*, 1351.
- (25) Robertson, A. D.; Baldwin, R. L. *Biochemistry* **1991**, *30*, 9907.
- (26) Levy, R. M.; Karplus, M.; Kushick, J.; Perahia, D. *Macromolecules* **1984**, *17*, 1370.
- (27) Agarwal, P. K.; Geist, A.; Gorin, A. *Biochemistry* **2004**, *43*, 10605.
- (28) Ramanathan, A.; Savol, A.; Burger, V.; Chennubhotla, C. S.; Agarwal, P. K. *Acc Chem Res* **2014**, *47*, 149.
- (29) Merlino, A.; Vitagliano, L.; Ceruso, M. A.; Di Nola, A.; Mazzarella, L. *Biopolymers* **2002**, *65*, 274.
- (30) delCardayre, S. B.; Raines, R. T. *Biochemistry* **1994**, *33*, 6031.
- (31) Agarwal, P. K.; Schultz, C.; Kalivretanos, A.; Ghosh, B.; Broedel, S. E. *J Phys Chem Lett* **2012**, *3*, 1142.
- (32) Sela, M.; Anfinsen, C. B. *Biochim Biophys Acta* **1957**, *24*, 229.
- (33) Otwinowski, Z.; Minor, W. *Method Enzymol* **1997**, *276*, 307.
- (34) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J Appl Crystallogr* **2007**, *40*, 658.
- (35) Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. *Acta Crystallogr D* **2010**, *66*, 213.
- (36) Emsley, P.; Cowtan, K. *Acta Crystallogr D* **2004**, *60*, 2126.
- (37) Krissinel, E.; Henrick, K. *Acta Crystallogr D* **2004**, *60*, 2256.
- (38) Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G.; Rance, M.; Skelton, N. J. *Protein NMR Spectroscopy: Principles and Practice*; Elsevier Academic, San Diego, 2007.
- (39) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J Biomol NMR* **1995**, *6*, 277.
- (40) Goddard, T. D.; Kneller, D. G., University of California.
- (41) Manley, G.; Loria, J. P. *Arch Biochem Biophys* **2012**, *15*, 223-31.
- (42) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J Comput Chem* **2005**, *26*, 1668.
- (43) Agarwal, P. K. *Proteins* **2004**, *56*, 449.
- (44) Ramanathan, A.; Agarwal, P. K. *PLoS Biology* **2011**, *9*, e1001193.
- (45) Freyer, M. W.; Lewis, E. A. *Methods Cell Biol* **2008**, *84*, 79.
- (46) Todd, M. J.; Gomez, J. *Anal Biochem* **2001**, *296*, 179.
- (47) Spencer, S. D.; Raffa, R. B. *Pharm Res* **2004**, *21*, 1642.
- (48) Rosenberg, H. F.; Domachowske, J. B. *Methods Mol Biol* **2001**, *160*, 355.

## Table of Contents (TOC) Graphic



### 3.5 Discussion

Dans cet article, nous avons tenté d'identifier le rôle de l'alanine en position 109 chez la RNase A, un résidu conservé à 100% chez l'humain. Ala109 se situe sur le brin  $\beta_5$ , dans le secteur  $V_1$ , une région généralement associée à l'initiation de la chaîne de repliement [161]. Nous démontrons que sa mutation pour une glycine cause un repositionnement du ligand 5'-AMP. Dans l'unité cristalline, le groupement phosphate du ligand a principalement été observé entre les sous-sites  $B_1$  et  $B_2$ , alors que la base faisait face au solvant. Bien que moins fréquemment, nous avons aussi aperçu le 5'-AMP dans le sous-site  $B_2$ . Toutefois, son positionnement était non favorable à la liaison. Alternativement, le 3'-UMP est demeuré à sa position naturelle en  $B_1$ . Ce résultat suppose donc que la protéine possède un réseau d'interactions, où la perturbation d'un point affecte des secteurs distants du site muté. Ces canaux d'interactions, auxquels fait référence la figure 1.4 du présent document, ont déjà été démontrés dans plusieurs systèmes enzymatiques [64, 162-165]. Chez la RNase A, il a été montré que l'histidine en position 48, un résidu situé sur le brin  $\beta_1$  dans le secteur  $V_2$ , et situé à une distance de 18 Å du site actif, pouvait moduler la dynamique et affecter négativement l'activité de l'enzyme. Un résultat similaire a été observé avec la création d'une chimère RNase  $A_{ECP}$ , où la boucle 1 de la RNase A, située à plus de 20 Å du site actif, a été remplacée par celle d'ECP. Par spectroscopie RMN, la chimère RNase  $A_{ECP}$  affiche un portrait considérablement modifié de sa dynamique, avec un taux de changement conformationnel et une activité fonctionnelle réduite. [64, 166].

Nous avons utilisé la spectroscopie RMN afin de caractériser ce réseau d'interactions menant au changement de position du ligand 5'-AMP. Par  $^{15}\text{N}$ -CPMG, le profil dynamique du mutant dans sa forme libre est demeuré sensiblement le même que celui du type sauvage, montrant une concentration des résidus dynamiques dans deux régions distinctes. Alors que le taux d'échange conformationnel ( $k_{ex}$ ) de la région 3 (correspondant aux brins  $\beta_1/\beta_4/\beta_5$  dans le secteur  $V_2$ ) est demeuré sensiblement le même, le  $k_{ex}$  de la boucle 4 (secteur  $V_1$ ) est passé de  $1548 \pm 82 \text{ s}^{-1}$  à  $3649 \pm 473 \text{ s}^{-1}$  chez le mutant. Ce changement dans le taux d'échange conformationnel est supporté par des résultats de simulation de dynamique moléculaire, qui démontrent une valeur moyenne quadratique pondérée, ou RMSF ("Root Mean Square Fluctuation"), plus importante dans la boucle 4. Ces résultats supportent le fait que la mutation du résidu Ala109 pour une glycine a causé une déstabilisation de la boucle 4, la rendant du même coup plus flexible. Tout porte à croire que cette plus grande flexibilité empêche le ligand 5'-AMP de se lier correctement à sa position naturelle en  $B_2$ . Ce résultat s'accompagne d'une



réduction dans la stabilité du mutant, passant de 64 à 61°C, compatibles avec les observations mentionnées précédemment. Par ailleurs, les résultats structuraux démontrent que la liaison du ligand 3'-UMP s'effectue normalement au sous-site B<sub>1</sub>, malgré une très légère réduction de l'affinité.

En enlevant un groupement méthyle, nous démontrons que Ala109 joue un rôle indirect dans la catalyse. Par <sup>15</sup>N-CPMG, il nous a été possible d'identifier un changement du taux d'échange conformationnel dans la boucle 4, malgré un profil dynamique global sensiblement inchangé. Cette étude montre qu'il existe une implication directe ou indirecte des mouvements atomiques dans la fonction de la protéine. L'utilisation combinée de la spectroscopie RMN, de la cristallographie et de la bio-informatique nous a permis de démontrer l'importance de Ala109 dans le dynamisme de la RNase A.

## 3.6 Résultats supplémentaires

### 3.6.1 Comparaison des structures native et mutée

Dans l'article présenté dans ce chapitre, les structures libres et liées de la RNase A et de la RNase A<sub>A109G</sub> ont été visuellement comparées. Aucune différence significative n'a été observée, bien que des variations sont visibles dans plusieurs boucles. Pour faire suite à cette analyse, nous avons voulu faire une comparaison quantitative des structures avec l'outil bio-informatique FATCAT [167]. L'application permet d'identifier les similarités et les différences entre deux structures. Avec des RMSD de 0.6 Å pour la forme libre, l'analyse FATCAT a révélé que les deux structures étaient relativement similaires. Tel qu'observé visuellement, des variations dans les boucles 1, 2 et 6 sont visibles, à en juger par l'apparition des couleurs rouge et bleu sur la matrice représentant les distances atomiques (Figure 3.2A). Cette différence est cependant beaucoup plus prononcée dans la boucle 1, une fois la protéine liée aux ligands (Figures 3.2B-C). Malgré les variations observées dans quelques boucles, les structures demeurent relativement similaires.

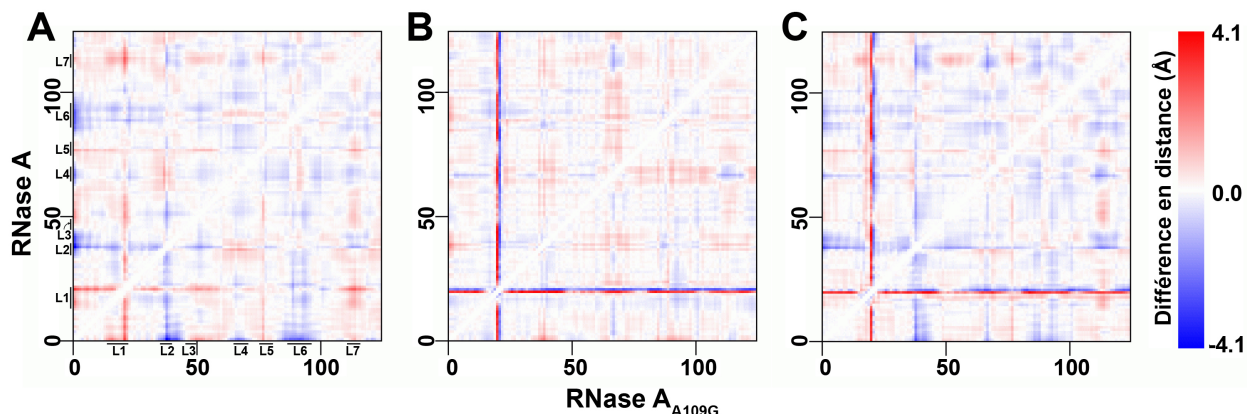
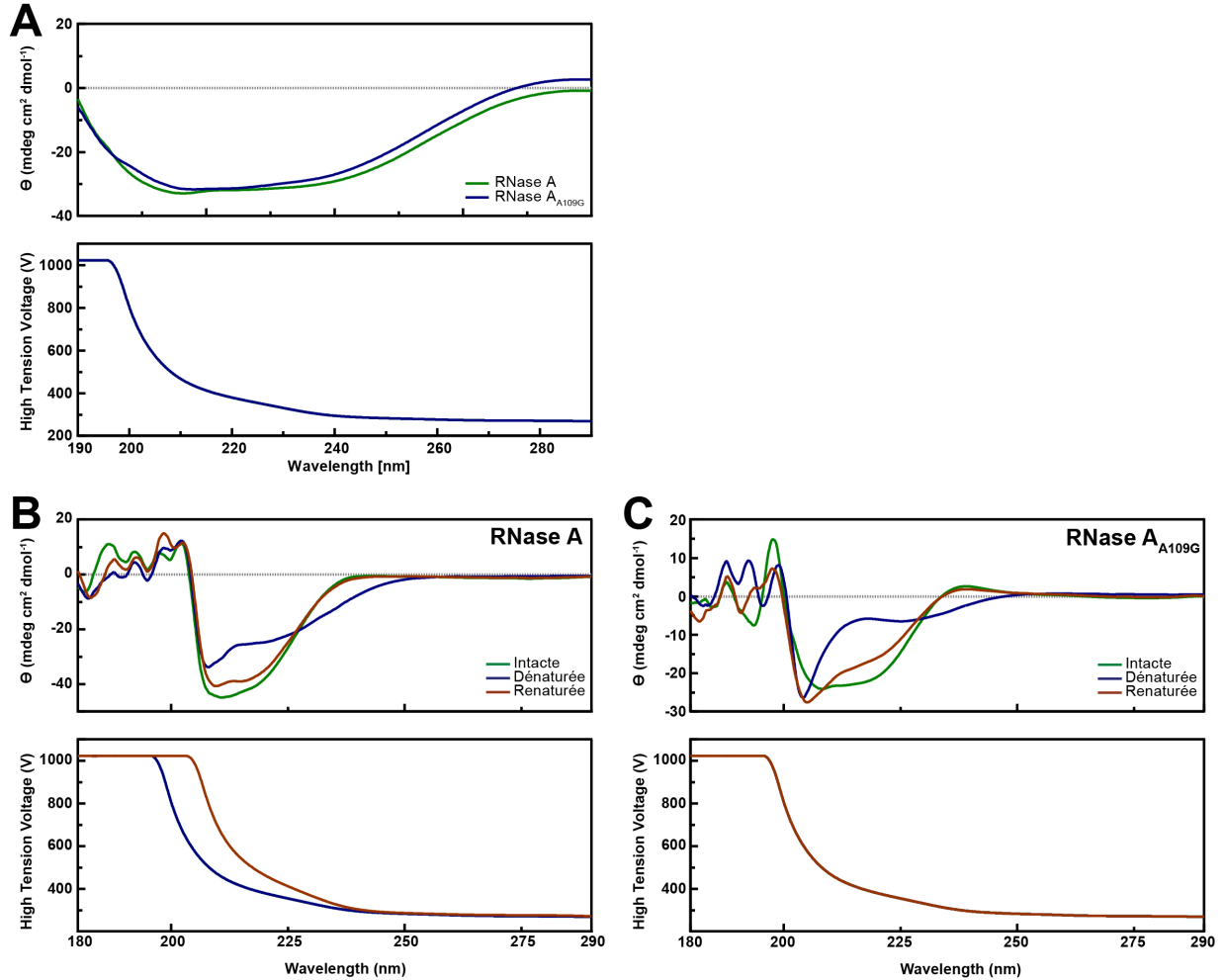


Figure 3.2. Comparaison de la structure tridimensionnelle de la RNase A et du mutant RNase A<sub>A109G</sub>. L'analyse comparative a été faite avec l'outil bio-informatique FATCAT [167]. La différence entre les distances des deux structures est présentée sous forme de matrice entre les RNase A et RNase A<sub>A109G</sub> dans la forme (A) libre, dans la forme liée avec le ligand (B) 5'-AMP (cocristal) ou (C) 3'-UMP. L1-7 montre le positionnement des boucles sur la matrice.

### 3.6.2 Analyse structurale du mutant RNase A<sub>A109G</sub>

Nous avons montré, par comparaison des structures cristallines de la RNase A et du mutant RNase A<sub>A109G</sub>, qu'elle n'était que peu affectée par la mutation. Ce résultat est également supporté par la comparaison des spectres UV lointains (Figure 3.3A). Malgré un alignement presque parfait des spectres UV lointains des deux protéines, la température de dénaturation est passée de 64 à 61°C, une légère baisse en stabilité chez le mutant. La dénaturation à 80°C (Figure 3.3B, bleu) des RNases A et A<sub>A109G</sub> montre une déstabilisation des feuillets bêta (215 nm) et des hélices alpha (208 et 222 nm). On remarquera toutefois que le profil des spectres entre les deux protéines une fois dénaturées diffère légèrement, un résultat qui supporte les données précédentes voulant que le résidu Ala109 joue un rôle dans la stabilité de la protéine. Une superposition des spectres de la protéine native (vert) avec la protéine renaturée (rouge) montre que le repliement de la RNase A est presque complet, un résultat qui avait été démontré précédemment [70]. En contrepartie, le mutant affiche une différence, principalement visible entre 205 et 225 nm. Ce résultat va dans le même sens que ce qui a été présenté dans l'article, c'est-à-dire que Ala109 déstabilise la boucle 4, laquelle affecte le repositionnement du ligand au sous-site B<sub>2</sub>.



**Figure 3.3. Stabilité des RNases A et A<sub>A109G</sub>.** (A) Spectre UV lointain des RNases A (vert) et A<sub>A109G</sub> (bleu). Les protéines ont été solubilisées dans un tampon 10 mM phosphate à pH 7.4 à une concentration de 50  $\mu$ M. Les spectres ont été acquis 25°C en utilisant une vitesse de balayage de 50 nm/minute. Courbes de dénaturation à 220 nm des spectres UV lointains des (B) RNases A et (C) RNase A<sub>A109G</sub>. Les spectres ont été acquis à 25°C (vert), après dénaturation à 80°C (bleu), et après renaturation (rouge).



## 4 ARTICLE NO. 3: CONSERVATION DE RÉGIONS FLEXIBLES ENTRE HOMOLOGUES STRUCTURAUX ET FONCTIONNELS

Titre: "Conservation of flexible residue clusters among structural and functional enzyme homologues"

Auteurs: **Gagné, D.**, Charest, L.A., Morin, S., Kovrigin, E.L. & Doucet N.

### 4.1 Présentation de l'article scientifique

L'importance de la flexibilité conformationnelle pour la catalyse enzymatique a été démontrée dans plusieurs systèmes enzymatiques [12, 63, 160, 168, 169]. On peut donc penser que si ces mouvements atomiques sont essentielles à l'activité de la protéine, elles seront conservées entre les membres d'une même famille d'homologues structuraux de fonction similaire. Malheureusement, encore peu d'études portent sur le sujet, comme en fait foi l'article présenté au chapitre 2 du présent document. Nous avons donc entrepris une étude qui compare deux membres de la famille des RNases, soit la RNase A et ECP. Malgré leur similitude structurale, les deux protéines ne partagent que 30% d'homologie de séquence. Afin de caractériser les mouvements atomiques, nous avons utilisé une combinaison de titrage RMN et de  $^{15}\text{N}$ -CPMG. Les résultats obtenus supportent l'hypothèse de conservation des mouvements atomiques dans l'échelle de temps de la  $\mu\text{s}$ -ms, au même titre que la séquence et la structure.

Le manuscrit a été accepté le 1<sup>er</sup> novembre 2012 dans la revue "Journal of Biological Chemistry" (facteur d'impact (2013) de 4.600), et une première version de l'article a été publiée en ligne le 7 novembre 2012. L'article a été publié dans le numéro 53 du volume 287 de l'édition du 28 décembre 2012, aux pages 44289-44300, sous le numéro d'accèsion doi: 10.1074/jbc.M112.394866. À ce jour, l'article a été cité 9 fois. La version formatée par l'éditeur sous format PDF et publiée en ligne est présentée dans le présent document.

## 4.2 Contribution des auteurs

Le travail expérimental a été réalisé par l'auteur principal (Donald Gagné), avec l'aide technique de Laurie-Anne Charest. L'analyse et l'interprétation des résultats de  $^{15}\text{N}$ -CPMG et de titrage RMN ont été faites par l'auteur principal (Donald Gagné). Sébastien Morin a effectué l'analyse biocomputationnelle du "Model-Free" et participé à son interprétation. L'analyse de type "Line Shape" et l'interprétation de ces résultats ont été réalisées par Pr Evgenii Kovrigin. Le texte et les figures ont été écrits, préparés et édités par l'auteur principal (Donald Gagné), sous la supervision de Pr Nicolas Doucet. Pr Nicolas Doucet a soumis le manuscrit à l'éditeur.

## 4.3 Résumé

L'ingénierie des protéines, c'est-à-dire le pouvoir de modifier une enzyme dans l'optique d'améliorer ou de lui donner de nouvelles fonctions, s'avère d'un grand intérêt avec des applications commerciales, pharmaceutiques et industrielles. Toutefois, elle nécessite une bonne compréhension du fonctionnement de l'enzyme. Même si les scientifiques étudient le comportement des enzymes depuis plusieurs décennies, nous réalisons encore aujourd'hui qu'il nous manque des morceaux au casse-tête. Nous savons maintenant que les protéines sont des entités flexibles, un modèle originalement présenté par Kosland en 1958 [3]. En utilisant plusieurs systèmes enzymatiques, il a été démontré que ces mouvements atomiques pouvaient être essentielles à la fonction et/ou au repliement de la protéine [11, 14, 63, 160, 162, 163, 169]. Des études théoriques ont postulé l'existence de régions de résidus dynamiques [143, 170]. Ces secteurs seraient conservés entre des homologues structuraux exerçant la même fonction.

Dans cet article, nous avons utilisé deux homologues structuraux de la famille des RNases, soit la RNase A (*Bos taurus*) et ECP (*Homo sapiens*). Les deux enzymes partagent la même fonction de digérer l'ARN, malgré leur faible identité de séquence. La RNase A est l'une des protéines les plus étudiées des 50 dernières années. D'ailleurs, elle a donné son nom à la famille, qui est composée de plus de 250 séquences complètes. ECP fait quant à elle partie d'une famille de huit enzymes canoniques chez l'humain. En plus de dégrader l'ARN, d'autres fonctions lui ont été attribuées: antibactérienne, helminthotoxique et cytotoxique [171]. Les mouvements atomiques à l'échelle de temps de la  $\mu\text{s}$ -ms ont précédemment été décrites chez la RNase A [12, 160]. Ces études ont permis d'identifier un groupe de résidus dynamiques situés dans la région L1/ $\beta$ 1/ $\beta$ 4/ $\beta$ 5 du secteur  $V_2$ . La perturbation de ces résidus était corrélée

avec une diminution de l'activité ribonucléolytique de la RNase A [64]. Par ailleurs, Cole *et al.* [160] avait démontré que le taux d'échange conformationnel ( $k_{ex}$ ) chez la RNase A était corrélé avec le  $k_{cat}$ , un résultat qui pouvait suggérer une implication directe des mouvements moléculaires durant la catalyse enzymatique.

En nous basant sur ces observations, nous voulions vérifier si les secteurs de résidus dynamiques identifiés chez la RNase A étaient également conservés chez les homologues structuraux de fonction similaire. Pour ce faire, nous avons utilisé une combinaison de titrage RMN, de  $^{15}\text{N}$ -CPMG, et de méthodes biophysiques afin de comparer les mouvements atomiques présentes chez ECP et la RNase A. Les résultats démontrent que les mouvements globaux dans l'ordre de la  $\mu\text{s}$ -ms sont conservés chez ECP. Bien que la dynamique globale ait été préservée, nous avons observé un réarrangement des réseaux locaux de résidus dynamiques. Par ailleurs, ECP montre une vitesse d'échange globale de ses régions fonctionnelles inférieures à celle de la RNase A, résultat qui concorde avec son activité catalytique réduite. Ce résultat est également supporté par la construction d'une chimère de la RNase A dans laquelle sa boucle 1 (dynamique) est remplacée par celle d'ECP (non-dynamique). Cette chimère, RNase A<sub>ECP</sub>, présente des caractéristiques fonctionnelles et dynamiques semblables à celles d'ECP, confirmant la possibilité d'utiliser l'ingénierie enzymatique dans le but de contrôler la dynamique et la fonction chez les enzymes.

Supplemental Material can be found at:  
<http://www.jbc.org/content/suppl/2012/11/07/M112.394866.DC1.html>

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 287, NO. 53, PP. 44289–44300, DECEMBER 28, 2012  
 © 2012 BY THE AMERICAN SOCIETY FOR BIOCHEMISTRY AND MOLECULAR BIOLOGY, INC. PUBLISHED IN THE U.S.A.

## Conservation of Flexible Residue Clusters among Structural and Functional Enzyme Homologues<sup>\*[5]</sup>

Received for publication, July 3, 2012, and in revised form, November 1, 2012. Published, JBC Papers in Press, November 7, 2012, DOI 10.1074/jbc.M112.394866

Donald Gagné<sup>†1</sup>, Laurie-Anne Charest<sup>†1</sup>, Sébastien Morin<sup>§</sup>, Evgenii L. Kovrigin<sup>¶</sup>, and Nicolas Doucet<sup>‡2</sup>

From the <sup>†</sup>Institut National de la Recherche Scientifique-Institut Armand-Frappier, Université du Québec, Laval, Québec H7V 1B7, Canada, the <sup>§</sup>Swiss Institute of Bioinformatics and Biozentrum, University of Basel, Klingelbergstrasse 50/70, 4056 Basel, Switzerland, and the <sup>¶</sup>Chemistry Department, Marquette University, Milwaukee, Wisconsin 53201

**Background:** It remains unclear whether structural homologues rely on similar concerted motions to promote enzyme function.

**Results:** Ribonuclease homologues display similar, contiguous clustering motions that can be modulated by mutagenesis.

**Conclusion:** Conformational flexibility can be conserved between distant structural homologues.

**Significance:** Controlling dynamics to modulate function has broad implications in protein engineering and allosteric drug design.

Conformational flexibility between structural ensembles is an essential component of enzyme function. Although the broad dynamical landscape of proteins is known to promote a number of functional events on multiple time scales, it is yet unknown whether structural and functional enzyme homologues rely on the same concerted residue motions to perform their catalytic function. It is hypothesized that networks of contiguous and flexible residue motions occurring on the biologically relevant millisecond time scale evolved to promote and/or preserve optimal enzyme catalysis. In this study, we use a combination of NMR relaxation dispersion, model-free analysis, and ligand titration experiments to successfully capture and compare the role of conformational flexibility between two structural homologues of the pancreatic ribonuclease family: RNase A and eosinophil cationic protein (or RNase 3). In addition to conserving the same catalytic residues and structural fold, both homologues show similar yet functionally distinct clusters of millisecond dynamics, suggesting that conformational flexibility can be conserved among analogous protein folds displaying low sequence identity. Our work shows that the reduced conformational flexibility of eosinophil cationic protein can be dynamically and functionally reproduced in the RNase A scaffold upon creation of a chimeric hybrid between the two proteins. These results support the hypothesis that conformational flexibility is partly required for catalytic function in homologous enzyme folds, further highlighting the impor-

ance of dynamic residue sectors in the structural organization of proteins.

The role of sequence and structure in defining enzyme function is a broadly accepted biological dogma (1). However, structure-function analyses remain very limited in their ability to predict and design new enzyme catalysts. Indeed, information about sequence and structure remains largely insufficient in providing a complete description of the intricate complexities that govern protein folding and enzyme function. Despite recent advancements in predicting folding mechanisms and catalytic function (2, 3), current limitations partly stem from our lack of understanding of the molecular mechanisms that define protein behavior in space and time (*i.e.* as flexible and dynamic macromolecules (4, 5)). Although classical computational design methodologies have typically considered proteins as homogeneously static structures (5), overwhelming theoretical and experimental evidence now advocates a more complex view of their reaction cycle. Among a few other methodologies, NMR relaxation experiments recently uncovered the existence of functionally relevant, low populated conformational states playing essential roles in the thermodynamics of ligand binding, substrate discrimination, active-site reorganization, and product release over the complete time course of an enzyme turnover (6–10). From an enzyme engineering perspective, designing efficient biocatalysts requires the modulation of flexibility events at the atomic scale to exert some control over enzyme function, further validating the need for a better characterization of the motional states found in natural enzymes, thus leading to a superior understanding of their precise role in catalysis (4, 5). In many enzyme systems, conformational exchange has been shown to correlate with the time scale of enzyme turnover, suggesting that flexible networks of concerted residue motions are an integral part of the structure-function relationship in numerous protein architectures (7).

Theoretical studies have postulated the existence of functionally relevant clusters of dynamic residues, which appear to be conserved across evolutionary distant clans of enzyme

\* This work was supported in part by Natural Sciences and Engineering Research Council of Canada Discovery Grant RGPIN 402623-2011 (to N. D.) and a Fonds de Recherche du Québec Santé Research Scholar-Junior 1 career award (to N. D.).

[5] This article contains supplemental Figs. S1–S5.

<sup>†</sup> Recipient of a "Fondation Universitaire Armand-Frappier de l'INRS" Ph.D. scholarship.

<sup>‡</sup> Supported by the Fonds de Recherche du Québec, Nature et Technologies, Strategic Cluster "Regroupement Québécois de Recherche sur la Fonction, la Structure et l'Ingénierie des Protéines" (PROTEO) and the Fonds de Recherche du Québec Santé (FRQS) Strategic Cluster "Groupe de Recherche Axé sur la Structure des Protéines" (GRASP). To whom correspondence should be addressed: INRS-Institut Armand-Frappier, Université du Québec, 531 Blvd. des Prairies, Laval, Québec H7V 1B7, Canada. Tel.: 450-687-5010 (ext. 4212); Fax: 450-686-5501; E-mail: nicolas.doucet@iaf.inrs.ca.



## Flexible Residue Clusters in Enzyme Homologues

superfamilies (11) and among members of the same enzyme fold (12). Although these studies put forth compelling evidence suggesting that enzyme function is at least partly controlled by conformational exchange, the experimental validation of such hypotheses remains largely unexplored. Although the role of flexibility in enzyme function has been amply demonstrated on a case-by-case basis (7, 9), it is yet unknown whether residue motions on multiple time scales are evolutionary traits conserved among structurally and functionally similar homologues.

Having been studied for more than 50 years as a *de facto* protein model for biochemical and biophysical analyses (13–15), ribonucleases represent an excellent model system to investigate the evolutionary and dynamic processes that regulate catalytic function among enzyme homologues. RNase A is the founding member of the mammalian and vertebrate superfamily that bears its name, which comprises an extensive network of functionally distinct enzymes sharing invariant structural and catalytic elements (16). These enzymes all catalyze the transphosphorylation and subsequent hydrolysis of single-stranded RNA molecules (Fig. 1), but they also carry broad and yet relatively uncharacterized biological activities. Initial sequencing of the human genome identified eight canonical members of this family (RNases 1–8), all of which preserve the same catalytic triad and display the analogous kidney-shaped tertiary fold of RNase A (17). These structurally similar enzymes catalyze such diversified activities as neurotoxicity, angiogenesis, immunosuppressivity, or anti-pathogenicity while still preserving varying degrees of ribonucleolytic activity. Although they do not all degrade RNA with the same catalytic efficiency, human ribonucleases require the strict conservation of important catalytic residues to perform their non-catalytic function (16). Interestingly, one of the few exceptions to this rule is eosinophil cationic protein (ECP,<sup>3</sup> or RNase 3), a human canonical member for which a functional active site is not always required for biological function (18), much like in RNase 7 (19). ECP has been the subject of many studies over the past few years, mainly because of its potential use as an antibiotic and as a cytotoxic agent (20). Although its biological function remains unclear, ECP has been shown to display antibacterial, neurotoxic, helminthotoxic, antiviral, and cytotoxic activities, some of them independent of its ribonucleolytic activity (18).

Short and long range concerted residue motions occurring on the time scale of  $k_{\text{cat}}$  ( $\sim 10^{-3}$  s) have been shown to correlate with substrate-to-product conversion and rearrangement in a number of divergent enzyme systems and thus are postulated to be essential for optimal enzyme function (6, 7). Motions on the faster picosecond-to-nanosecond (ps-ns) time scale have also been shown to be important in enzyme function, namely by destabilizing the ground state and affecting the rates of barrier crossing associated with chemical exchange (21). We and others have shown that RNase A requires concerted millisecond (ms) dynamics for efficient catalysis through motions involving

a distant loop (loop 1, residues 14–26) located more than 20 Å away from the reactive center (22). During turnover, the flexible loop 1 propagates ms motions to the active site through a highly conserved pair of hydrogen-bonded residues (His<sup>48</sup>-Thr<sup>82</sup>), which are located on the adjacent  $\beta 1$  and  $\beta 4$  strands (23). These molecular motions are transmitted over a length scale of more than 20 Å from loop 1 to active site residues Thr<sup>45</sup> and Asp<sup>83</sup> (also located in  $\beta 1$  and  $\beta 4$ ), providing important pyrimidine interactions with RNA ligands during catalysis (23, 24). As was demonstrated by mutagenesis (22, 23, 25) and unnatural amino acid modifications (24), this concerted dynamic network is essential for optimal catalysis in RNase A, acting by modulating the rate-limiting step of the reaction (product release). Although the dynamically important residue pair His<sup>48</sup>-Thr<sup>82</sup> is 86% conserved in the mammalian and vertebrate ribonuclease superfamily (23), ECP is one of only two human ribonucleases lacking this important interaction, coinciding with a steep decrease in ribonucleolytic activity (18).

Based on these observations, it is tempting to verify whether evolutionarily conserved motional networks between homologous ribonucleases could partially account for their divergent catalytic activities. In the present study, we aimed at validating whether highly homologous structural and functional ribonucleases catalyzing the same transphosphorylation reaction also display conserved dynamic behaviors both on the fast ps-ns and the slower microsecond-to-millisecond ( $\mu\text{s}$ -ms) time scales. We challenge this hypothesis by investigating the conformational exchange and the conservation of dynamic clusters between ECP and RNase A, two structural homologues of the ribonuclease superfamily. We used NMR relaxation dispersion, model-free analysis, and titration experiments to show that both enzymes display similar millisecond dynamic residue clusters near their active sites, with relatively rigid and homogeneous backbone fluctuations on the ps-ns time frame. Interestingly, although motional clusters are structurally conserved between the two homologues, their respective global rates of conformational exchange ( $k_{\text{ex}}$ ) vary significantly. As predicted by the absence of the functionally dynamic His<sup>48</sup>-Thr<sup>82</sup> residue pair in ECP, this enzyme lacks the aforementioned loop 1- $\beta 1$ - $\beta 4$  network of coupled residue motions previously shown to be involved in the modulation of product release in RNase A. Additionally, this functionally essential network can be eliminated in RNase A by swapping loop 1 for that of ECP in the RNase A<sub>ECP</sub> chimera (22).

Our current motional investigation demonstrates that both ECP and a chimeric hybrid between RNase A and ECP (RNase A<sub>ECP</sub> (22)) show very similar dynamic behaviors, further highlighting the possibility of using mutagenesis to control residue flexibility and function through the dynamic conversion of one enzyme (RNase A) into another (ECP). Finally, our conformational investigation of ECP confirms previous RNase A observations suggesting that this fold can be dynamically divided into several independent regional flexible subdomains (or dynamic clusters).

The current study represents one of the very few experimental validations of similar conformational exchange in structurally related enzymes (26, 27) and the first to compare similar mesophile homologues, confirming theoretical observations

<sup>3</sup> The abbreviations used are: ECP, eosinophil cationic protein; ms, millisecond; ps-ns, picosecond-to-nanosecond;  $\mu\text{s}$ -ms, microsecond-to-millisecond; HSQC, heteronuclear single-quantum coherence; CPMG, Carr-Purcell-Meiboom-Gill.

suggesting that specific motions can be evolutionarily conserved within and among structurally similar protein folds. The present work also demonstrates the importance of controlling millisecond dynamics to modulate protein function, a central concept with broad implications in protein engineering and allosteric drug design (4, 5, 28, 29).

## EXPERIMENTAL PROCEDURES

**Cloning, Expression, and Purification**—*Escherichia coli* codon-optimized sequences of ECP and RNase A (GenScript) were subcloned into NdeI/HindIII-digested expression vector pET22b(+) (EMD Biosciences) and transformed into *E. coli* BL21(DE3).  $^{15}\text{N}$ -Labeled samples were prepared by growing *E. coli* BL21(DE3) in M9 minimal medium, and enzymes were purified as described by Boix (18) (ECP) and Doucet *et al.* (22) (RNase A). Protein concentration was determined using an extinction coefficient of  $17,460 \text{ M}^{-1} \text{ cm}^{-1}$  ( $9,800 \text{ M}^{-1} \text{ cm}^{-1}$ ) for ECP (18) (RNase A (30)).

**Solution NMR Experiments**—All NMR experiments were recorded at 298 K on samples containing 0.2–0.7 mM  $^{15}\text{N}$ -labeled ECP in 15 mM sodium acetate, 10%  $^2\text{H}_2\text{O}$ , pH 5.0 or containing 0.7 mM  $^{15}\text{N}$ -labeled RNase A in 5 mM MES-NaOH, 7 mM NaCl, 0.01% NaN<sub>3</sub>, 10%  $^2\text{H}_2\text{O}$ , pH 6.4. NMR experiments were carried out on Varian (Agilent) 500- and 800-MHz NMR spectrometers equipped with triple-resonance cold probes and pulsed field gradients. Backbone resonance assignments for apo-ECP were taken from the Biological Magnetic Resonance Data Bank (entry 15757) and further confirmed with a  $^1\text{H}$ - $^{15}\text{N}$  total correlation spectroscopy heteronuclear single-quantum coherence (TOCSY-HSQC) experiment.

**Relaxation Dispersion Experiments ( $\mu\text{s}$ -ms Motions)**—Backbone amide  $^{15}\text{N}$  CPMG relaxation dispersion experiments were acquired on apo and saturated enzyme complexes using published sequences (31) and methods (22). Interleaved two-dimensional spectra were collected in a constant time manner with  $\tau_{\text{CP}}$  CPMG repetition delays of 0.625, 0.714 ( $\times 2$ ), 1.0, 1.25, 1.67, 2.0, 2.50 ( $\times 2$ ), 3.33, 5.0, and 10 ms, using a total relaxation period of 40 ms. All NMR spectra were processed using NMRPipe (32) and in-house CPMG scripts and analyzed with Sparky (33). Global residue fits and model analyses were performed by fitting 500- and 800-MHz CPMG dispersion data to the full single-quantum CPMG equation (34) using GraphPad Prism 5.

**Fast Time Scale Motions (ps-ns)**— $^{15}\text{N}$ - $R_1$ ,  $^{15}\text{N}$ - $R_2$  and steady state heteronuclear NOE experiments were performed in an interleaved fashion at 500 MHz, as described previously (35, 36). The  $R_1$  experiments were performed with relaxation delays of 10, 50, 100 ( $\times 2$ ), 200, 350, 700, 1100, and 1400 ms. The  $R_2$  experiments were performed with relaxation delays of 10, 30, 50 ( $\times 2$ ), 70, 90, 110, and 130 ms. The motional parameters on the faster ps-ns time scale were analyzed by fitting the NMR spin relaxation rates to the model-free formalism (reviewed in Ref. 37), using the methodology for the dual optimization of the model-free parameters and the global diffusion tensor proposed by d'Auvergne (38) and implemented in the program *relax* 2.1.0 (38, 39). The crystallographic structure of free ECP was used as starting coordinates (PDB entry 1QMT), and hydrogen atoms were added in DS Visualizer 3.5 (Accelrys).

Values for the  $^{15}\text{N}$  chemical shift anisotropy and N–H bond length were set at  $-172$  ppm and  $1.02 \text{ \AA}$ , respectively. Model selection for each residue and diffusion tensors was tested, optimized, and selected using the Akaike information criterion as described in the *relax* manual. Errors were obtained from 500 Monte Carlo simulations.

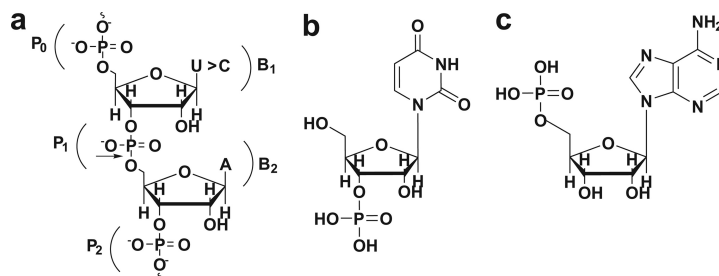
**NMR Titration Experiments**—Freshly prepared 3'-UMP and 5'-AMP ligands were dissolved in ECP or RNase A NMR buffers.  $^1\text{H}$ - $^{15}\text{N}$  sensitivity-enhanced HSQC experiments were acquired at 800 MHz using spectral widths (points) of 1600 Hz (256) and 7000 Hz (8192) in the  $t_1$  and  $t_2$  dimensions, respectively.  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra were collected for titration points of ligand-enzyme ratios of 0, 0.174, 0.393, 0.691, 1.31, 2.71, 6, and 12. The pH was maintained constant by the addition of 0.01 M HCl, when necessary.

**NMR Line Shape Analysis**—Analysis of NMR line shapes originating from titrations of ECP with 3'-UMP and 5'-AMP ligands was performed using the NMR line shape analysis module of the Integrative Data Analysis Platform (40). The one-dimensional data sets for fitting were obtained as slices through the fast exchange peaks in  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra using Sparky (33) with a custom Python extension (41). The one-dimensional NMR line shape was simulated using Bloch-McConnell equations for a spin exchanging between two magnetic environments as described earlier (42, 43). The line shapes were optimized to fit the experimental data by varying the thermodynamic equilibrium constant and the dissociation rate constant as well as the frequency of the bound complex using the Integrative Data Analysis Platform code implemented in MATLAB (MathWorks). The S.E. values were estimated from fitting of multiple line shape data sets generated with the addition of the Gaussian noise to the spectral intensities and solution concentrations. The uncertainty of the best-fit parameters was determined by the boundaries enclosing 95% of all best-fit results collected in the runs with Gaussian perturbations. The Integrative Data Analysis Platform software is available upon request from Evgenii L. Kovrigin.

## RESULTS

**Ligand Binding Interactions in ECP and RNase A**—The structural and functional binding differences between ECP and RNase A were assessed by NMR titration experiments using the 3'-UMP and 5'-AMP ligands (Fig. 1, *b* and *c*). 3'-UMP and 5'-AMP are single nucleotide homologues of the end products of the UpA dinucleotide RNA substrate, providing good structural assessments of the individual ribonuclease subsites involved in substrate binding to the active site cavity in both enzymes. Ligands were individually titrated in  $^{15}\text{N}$ -labeled proteins, and binding was followed by measuring chemical shift variations observed in  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra until saturation was reached. Affinities of 3'-UMP and 5'-AMP binding to ECP were estimated using line shapes observed in the  $^1\text{H}$ - $^{15}\text{N}$  HSQC titration experiments (41–43). Observation of ligand binding to the cognate sites in ECP was complicated by the apparent secondary binding events observed in both titrations, probably due to weak affinity of the nucleotides to a second nonspecific site. To estimate affinity of the binding interaction with the cognate site in the 3'-UMP data sets, we selected residues that

## Flexible Residue Clusters in Enzyme Homologues



**FIGURE 1. Substrates and products of a ribonuclease reaction.** *a*, schematic representation of a single-stranded RNA molecule binding to the active site of ECP. Phosphate and base binding subsites for RNA substrates are defined as  $P_n$  and  $B_n$ , of which the  $P_{0-2}$  and  $B_{1-2}$  subsites are shown (18). The phosphodiester bond scission occurs at the  $P_1$  subsite (arrow) and is universally catalyzed by two strictly conserved histidine residues: His<sup>15</sup> and His<sup>128</sup> (ECP numbering). In ECP, uracil is preferred over cytosine in the  $B_1$  subsite (U > C), a preference that is reversed in RNase A (C > U). Adenosine is almost universally required at the  $B_2$  subsite. *b*, schematic representation of the uracil-3'-monophosphate ligand (3'-UMP). *c*, schematic representation of the adenosine-5'-monophosphate ligand (5'-AMP).

only responded to the main binding event in ECP (*i.e.* Val<sup>125</sup> and Val<sup>127</sup>) (supplemental Fig. S1). The 5'-AMP binding has much weaker contribution from the second binding process, so residues demonstrating the most significant shifts were analyzed (His<sup>15</sup>, Ala<sup>110</sup>, and His<sup>128</sup>). Despite significant scatter of the global dissociation rate constants obtained in the individual fitting, the equilibrium dissociation constant ( $K_d$ ) was accurately estimated in all cases (supplemental Fig. S1). Therefore, global fitting to obtain  $K_d$  within each group was performed on ECP, resulting in dissociation constants of  $340 \pm 30 \mu\text{M}$  for 5'-AMP and  $460 \pm 100 \mu\text{M}$  for 3'-UMP (Table 1).

In both enzymes, the 3'-UMP and 5'-AMP nucleotides preferentially bind to the  $B_1$  and  $B_2$  subsites of the active site cavity, respectively. This provided a unique opportunity to characterize distinct atomic interactions involved in ligand discrimination in both structural homologues (Fig. 2). A total of 12 residues showed significant ( $>0.1$  ppm)  $^1\text{H}$ - $^{15}\text{N}$  weighted average chemical shift variations upon 3'-UMP binding to ECP, relative to 17 residues in RNase A (Fig. 2, *a* and *b*). Although the chemically affected residues are mostly located in the vicinity of the  $B_1$  subsite and are directly involved in binding and/or catalysis in RNase A (*e.g.* His<sup>12</sup>, Lys<sup>41</sup>, Thr<sup>45</sup>, and His<sup>119</sup>) (Fig. 2*b*), their functional equivalents in ECP (His<sup>15</sup>, Lys<sup>39</sup>, Thr<sup>42</sup>, and His<sup>128</sup>) do not show significant chemical shift perturbations upon 3'-UMP binding (Fig. 2*a*). If 3'-UMP does bind to the  $B_1$  subsite of ECP as expected, this would illustrate that the high sequence conservation of active site residues in the two homologues does not translate into similar pyrimidine recognition, partly accounting for the significantly reduced binding affinity of 3'-UMP in ECP (Table 1). Pyrimidine contacts with Thr<sup>45</sup> and Asp<sup>83</sup> are required for optimal product release in RNase A (24). However, no Asp<sup>83</sup> equivalent can be found in ECP, and Thr<sup>42</sup> (the Thr<sup>45</sup> equivalent) is completely unaffected by 3'-UMP binding. Additionally, residues displaying important chemical shifts upon 3'-UMP binding are scattered throughout the ECP structure, contrasting with the relatively bundled group of clustering residues in RNase A (Fig. 2*b*).

Based on their motional behavior (see below) and secondary structure elements (44), the two structural homologues were divided into four distinct regional residue clusters encompassing loop 4 (cluster 1), the  $\beta_2$  sheet (cluster 2), the  $\beta_1$  sheet and

**TABLE 1**  
Binding affinities of 3'-UMP and 5'-AMP to RNase A and ECP

	$K_d$ 3'-UMP	$K_d$ 5'-AMP
RNase A	$9.7 \pm 0.9^a$ $\mu\text{M}$	$124.0 \pm 0.9^b$ $\mu\text{M}$
ECP	$460 \pm 100^c$	$340 \pm 30^c$

<sup>a</sup> Taken from Ref. 65.

<sup>b</sup> Taken from Ref. 66.

<sup>c</sup> Estimated from NMR line shape analysis (see "Experimental Procedures").

adjacent loop 1 (cluster 3), and loops 2 and 6 (cluster 4) (Fig. 3). Interestingly, the contiguous residues displaying important chemical shift variations upon pyrimidine binding to RNase A (Thr<sup>45</sup>, Phe<sup>46</sup>, Thr<sup>82</sup>, Thr<sup>100</sup>, Gln<sup>101</sup>, all part of cluster 3) also show important conformational exchange on the millisecond time scale and are known to be involved in the propagation of motions between loop 1 and the active site during turnover (22, 23). Coupled to the chemical shift differences induced by 3'-UMP binding, this observation strongly suggests that long scale millisecond motions connecting the active site to loop 1 (cluster 3) are involved in the binding/release of pyrimidine ligands in RNase A. Although long scale chemical shift variations are also observed in ECP upon pyrimidine binding (*e.g.* residues Gln<sup>58</sup>, Arg<sup>75</sup>, and Arg<sup>77</sup>, located 20.2, 18.0, and 14.7 Å away from the ligand, respectively), residues of cluster 3 in ECP are not similarly responsive, and no contiguous residue network is affected upon ligand binding to this enzyme (Fig. 2). Additionally, residues of the ECP cluster 3 ( $\beta_1$ -sheet and adjacent loop 1) are almost completely devoid of significant chemical shift perturbations upon 3'-UMP binding to the  $B_1$  subsite, suggesting that concerted long range interactions linking loop 1 to the active site are not functionally important in ECP.

Purine binding to the  $B_2$  subsite also yields important chemical shift differences between these two structural homologues. Whereas the chemical environment of only 4 residues is significantly perturbed upon 5'-AMP binding to ECP, 14 residues are affected when this ligand binds to RNase A (Fig. 2, *c* and *d*). The magnitude in the  $^1\text{H}$ - $^{15}\text{N}$  weighted average chemical shift variations is also stronger in RNase A than ECP. Only the immediate vicinity of the  $B_2$  subsite is chemically affected in presence of 5'-AMP in ECP, including the important catalytic residues

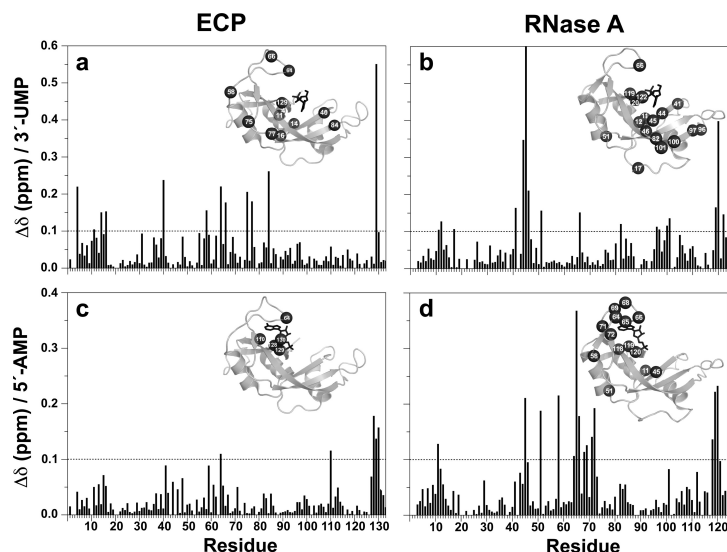


FIGURE 2.  $^1\text{H}$ - $^{15}\text{N}$  chemical shift variations induced by 3'-UMP and 5'-AMP binding to ECP and RNase A. Weighted average chemical shift differences ( $\Delta\delta$ ) for 3'-UMP (a and b) and 5'-AMP (c and d) are mapped on the primary sequence of ECP (a and c) and RNase A (b and d). The  $^1\text{H}$ - $^{15}\text{N}$  weighted average composite chemical shift differences ( $\Delta\delta$ ) were calculated between free and ligand-saturated enzymes according to the equation (67),  $\Delta\delta$  (ppm) =  $(\Delta\delta_{\text{HN}}^2 + \Delta\delta_{\text{C}\alpha}^2/25/2)^{1/2}$ . The position of 3'-UMP and 5'-AMP (black sticks) is displayed on the three-dimensional structure of RNase A (PDB entry 1O0N) and is predicted on ECP (PDB entry 1H1H) based on a structural overlay with RNase A. Black spheres represent residues with  $\Delta\delta > 0.1$  ppm.

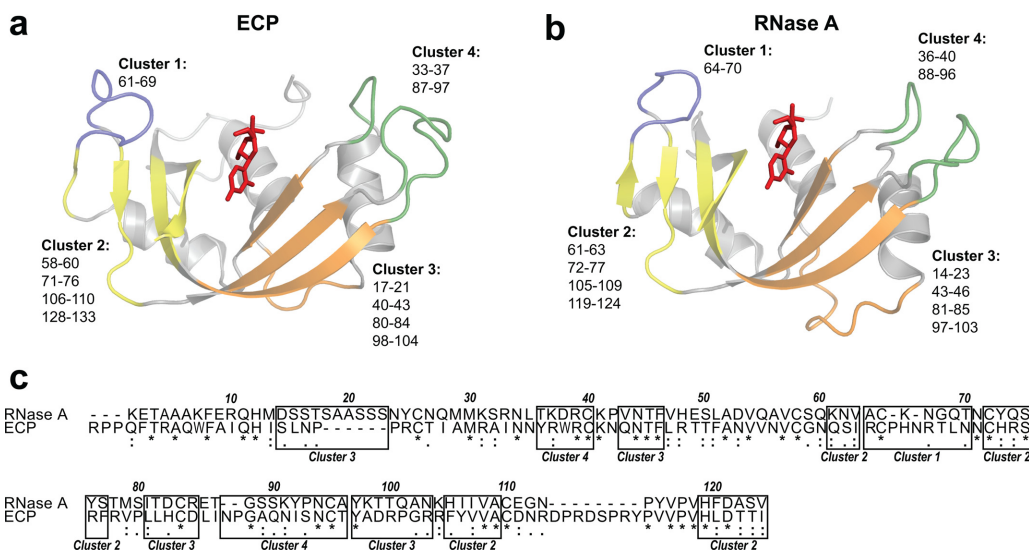


FIGURE 3. Dynamic residue clusters in ECP and RNase A. The two ribonuclease homologues are divided into four distinct residue clusters covering the following secondary structure elements (44): loop 4 (cluster 1, purple),  $\beta$ 2-sheet (cluster 2, yellow),  $\beta$ 1-sheet with adjacent loop 1 (cluster 3, orange), and loops 2 and 6 (cluster 4, green). The four color-coded clusters are highlighted on the three-dimensional structure of ECP (PDB entry 1H1H) (a) and RNase A (PDB entry 1O0N) (b), with encompassed residues listed. The 3'-UMP ligand is shown in red. c, Espresso (68) sequence alignment of ECP and RNase A showing the position of conserved residues and secondary structure elements forming the dynamic clusters. Sequence numbering is that of RNase A.

## Flexible Residue Clusters in Enzyme Homologues

His<sup>64</sup> and His<sup>128</sup>. Whereas His<sup>128</sup> is required for the hydrolytic step of the reaction, His<sup>64</sup> (located in loop 4) directly interacts with P<sub>0</sub> during turnover and is thought to act as the functional equivalent to Lys<sup>66</sup> in RNase A (45). The most striking differences observed upon 5'-AMP binding to ECP and RNase A occur for residues of clusters 1 and 2 (Fig. 3), more specifically for residues in loop 4 (cluster 1). Although residues of loop 4 show considerable chemical shift variations upon 5'-AMP binding to RNase A, only His<sup>64</sup> is affected when this purine binds to ECP. This loop faces the active site and acts as an arm that holds the purine ligand in position (Fig. 2*d*). Interestingly, loop 4 is the only dynamic cluster experiencing conformational exchange in both the free forms of ECP and RNase A, albeit with distinct conformational exchange rates (see below).

**Fast Time Scale Analysis (ps-ns)**—Although most proteins interact and catalyze reactions on a time frame much slower than nanoseconds, the kinetics and thermodynamics underlying such larger conformational motions may rely on faster time scale dynamics (21). In an effort to decipher the importance of such fast motions in ECP and to compare them with the previously characterized ps-ns motions in RNase A (35), we fitted the <sup>15</sup>N-*R*<sub>1</sub>, <sup>15</sup>N-*R*<sub>2</sub> and heteronuclear NOE spin relaxation data to the new approach for the dual optimization of the model-free parameters and the global diffusion tensor proposed by d'Auvergne and Gooley (38, 39). The spin relaxation behavior of 112 of the 121 non-proline residues of ECP could be reliably quantitated. Residues Phe<sup>43</sup>, Arg<sup>45</sup>, Asn<sup>53</sup>, Gln<sup>58</sup>, Asn<sup>70</sup>, Arg<sup>73</sup>, Phe<sup>76</sup>, Cys<sup>83</sup>, and Leu<sup>129</sup> were removed from the analysis because of low signal intensity and/or spectral overlapping and/or because they were absent from the <sup>1</sup>H-<sup>15</sup>N HSQC. The average values of the relaxation rates are *R*<sub>1</sub> = 1.13 ± 0.13 s<sup>-1</sup> and *R*<sub>2</sub> = 14.04 ± 1.78 s<sup>-1</sup>, and NOE = 0.478 ± 0.71 (supplemental Fig. S2). Much like in RNase A, the ECP heteronuclear NOE values at the N and C terminus of the protein are similar to those in the rest of the protein. This was expected because of the structural similarity in the protein architecture, which relies on the N-terminal helix and C-terminal sheet packing for binding and activity (35).

The model-free calculated order parameters (*S*<sup>2</sup>) are used to characterize the amplitude of the internal motions of the <sup>1</sup>H-<sup>15</sup>N vector on the ps-ns time scale and thus provide a measure of atomic scale residue flexibility on this particular time frame (46–48). As a result, a completely unrestricted <sup>1</sup>H-<sup>15</sup>N bond vector would show an *S*<sup>2</sup> value of 0, whereas a fully rigid one would display an order parameter of 1. The generalized *S*<sup>2</sup> parameters determined from model-free fitting are plotted as a function of residue number in Fig. 4. The calculated order parameters are relatively high and uniform across the sequence, much like with RNase A (35). However, the average value of *S*<sup>2</sup> = 0.740 ± 0.052 indicates that ECP is a bit more flexible than the very rigid RNase A on the ps-ns time scale, which displays an average *S*<sup>2</sup> = 0.910 ± 0.051 (35). Interestingly, the catalytic residue Lys<sup>38</sup> in ECP, which is strictly conserved among pancreatic RNases, displays one of the highest order parameters in the protein (*S*<sup>2</sup> = 0.848 ± 0.026), along with its very rigid neighbor Asn<sup>39</sup> (*S*<sup>2</sup> = 0.907 ± 0.055). Lys<sup>38</sup> is thought to be involved in the stabilization of the excess negative charge on the phosphoryl oxygens in the transition state during RNA cleavage

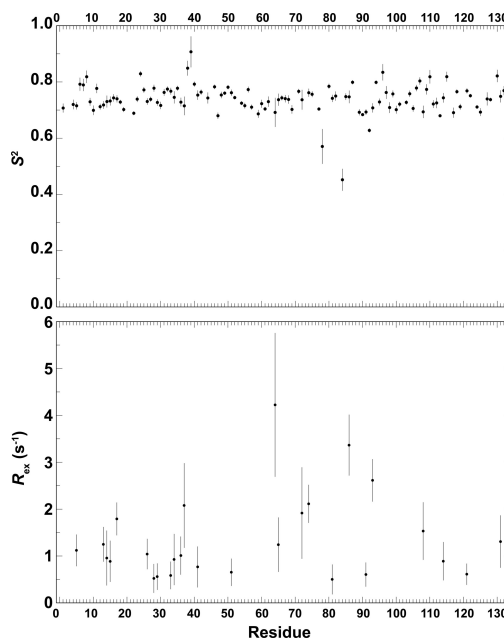
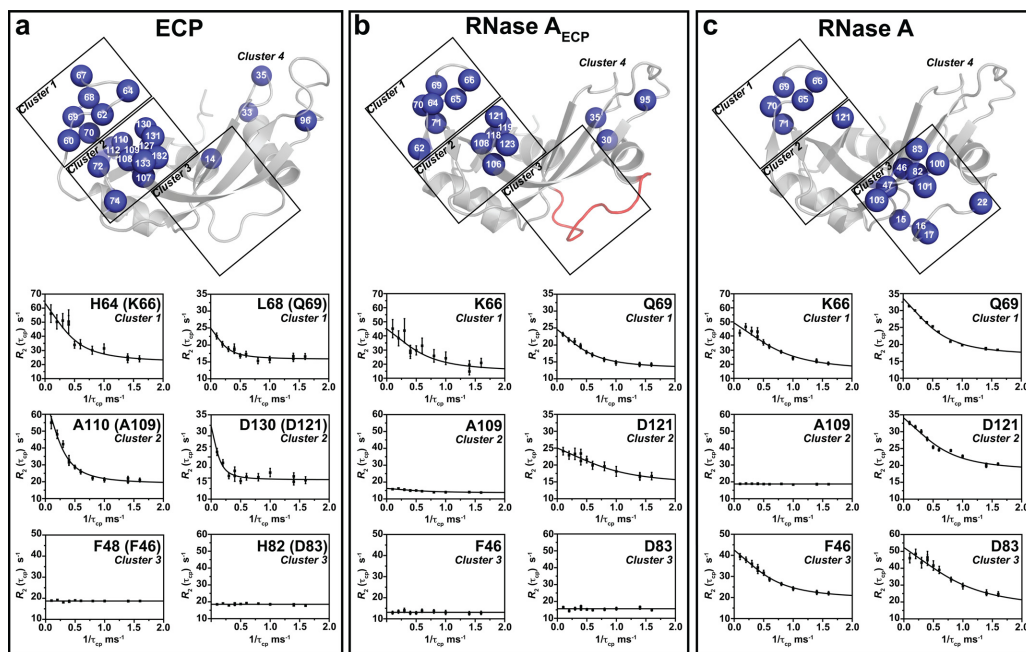


FIGURE 4. Model-free calculated *S*<sup>2</sup> and *R*<sub>ex</sub> parameters plotted on the sequence of ECP. The generalized order parameters (*S*<sup>2</sup>) provide a measure of atomic scale flexibility of the <sup>1</sup>H-<sup>15</sup>N bond vector on the ps-ns time scale. The predicted *R*<sub>ex</sub> parameter accounts for contributions to *R*<sub>2</sub> that potentially describe motions occurring on the slower μs-ms time scale.

(15), which may require rigidity in ECP. On the contrary, the two most flexible residues, Val<sup>78</sup> (*S*<sup>2</sup> = 0.571 ± 0.062) and Asp<sup>84</sup> (*S*<sup>2</sup> = 0.452 ± 0.039), are at the beginning and the end of the β2 strand, respectively. Neither was shown to play any particular role in enzyme function.

For a number of residues that do not fit to simple calculation models, an *R*<sub>ex</sub> exchange parameter is calculated to achieve an adequate fit to the model-free analysis. This *R*<sub>ex</sub> parameter accounts for contributions to *R*<sub>2</sub> that potentially describe motions occurring on the slower μs-ms time scale (Fig. 4). Much like with RNase A, a large number of residues throughout the ECP sequence require the addition of this exchange contribution (26 of the 112 residues analyzed), suggesting the existence of motions occurring on the slower μs-ms time scale in this enzyme. This observation is in line with our slow time scale analysis of ECP performed by <sup>15</sup>N CPMG (see below), most notably for residues of loop 4 (cluster 1; Fig. 3), which display some of the most important *R*<sub>ex</sub> values in the protein (Fig. 4). Interestingly, the highest *R*<sub>ex</sub> observed is for residue His<sup>64</sup> (*R*<sub>ex</sub> = 4.22 ± 1.53 s<sup>-1</sup>), the structural equivalent to Lys<sup>66</sup> in RNase A, which is also one of the most flexible residues of RNase A on the millisecond time scale (Fig. 5*c*). Overall, although ECP is a bit more flexible than RNase A on the ps-ns time frame, both enzymes show relatively rigid and homogeneous backbone fluctuations on this particular time scale, in addi-





**FIGURE 5. Conformational exchange is confined to four independent yet contiguous dynamic clusters in structural and functional ribonuclease homologues.** Conformational exchange was investigated in the apo forms of ECP, RNase A, and the chimeric hybrid RNase A<sub>ECP</sub>, in which the RNase A loop 1 (<sup>14</sup>DSSTSAASSSN<sup>25</sup>, in red) was replaced by that of ECP (<sup>17</sup>SLNPPR<sup>22</sup>, ECP numbering) (22). Catalytic time scale (ms) residue motions were probed using <sup>15</sup>N CPMG relaxation dispersion experiments at 500 and 800 MHz (298 K) for the apo forms of ECP (a), RNase A<sub>ECP</sub> (b), and RNase A (c). Residues were considered for further analysis only if the difference in measured  $R_2$  ( $\Delta R_2$  ( $1/\tau_{cp}$ )) values at fast ( $\tau_{cp} = 0.625$  ms) and slow ( $\tau_{cp} = 10$  ms) pulsing rates was greater than  $2\text{ s}^{-1}$ , similar to previous reports (22, 35). Blue spheres represent residues displaying <sup>15</sup>N CPMG dispersion with  $\Delta R_2$  ( $1/\tau_{cp}$ )  $> 2\text{ s}^{-1}$ . Residues are highlighted on the three-dimensional structure of ECP (PDB entry 1H1H) (a) and RNase A (PDB entry 1OON) (b and c). Representative relaxation dispersion curves (800 MHz, 298 K) are shown for structurally equivalent positions of ECP (RNase A) found in cluster 1 (His<sup>64</sup> (Lys<sup>66</sup>) and Leu<sup>68</sup> (Gln<sup>69</sup>)), cluster 2 (Ala<sup>110</sup> (Ala<sup>109</sup>) and Asp<sup>130</sup> (Asp<sup>121</sup>)), and cluster 3 (Phe<sup>48</sup> (Phe<sup>46</sup>) and His<sup>82</sup> (Asp<sup>83</sup>)). Flexible residues experiencing <sup>1</sup>H-<sup>15</sup>N bond vector motions on the millisecond time scale display a relaxation dispersion curve, whereas residues with no motion on this particular time frame display flat relaxation profiles. Swapping loop 1 of RNase A with that of ECP converts the dynamic clusters of the former into those of the latter. RNase A<sub>ECP</sub> data were taken from Ref. 22.

tion to similar flexible residue sectors on the millisecond time scale (see below).

**Conserved Networks of Flexible Millisecond Residue Clusters—**We investigated the apo and ligand-bound millisecond motions experienced by both structural homologues using solution NMR relaxation dispersion experiments (<sup>15</sup>N CPMG) (10). <sup>15</sup>N CPMG experiments allow the characterization of an equilibrium exchange process experienced by the <sup>1</sup>H-<sup>15</sup>N bond vector by recording the transverse relaxation rate constant,  $R_2$ , as a function of  $\tau_{cp}$ , an interpulse delay in the CPMG pulse train (49). <sup>15</sup>N CPMG experiments thus enable the detection of millisecond dynamics in proteins caused by internal conformational exchange, ligand binding, and/or the chemical reaction (for reviews, see Refs. 9 and 50–52). When  $R_2$  values are plotted as a function of  $1/\tau_{cp}$ , the characteristic dispersion of  $R_2$  values reveals residues experiencing conformational exchange on the millisecond time scale (Fig. 5). On the contrary, <sup>15</sup>N nuclear spins that do not experience motions on this particular time frame do not show any dispersion (flat line profiles). In addition to this qualitative assessment, the fitted relaxation dispersion curves can further be quantitatively analyzed to extract struc-

tural information of the excited state (differences in chemical shifts,  $\Delta\omega$ ) as well as population dynamics (exchange rates ( $k_{ex}$ ) and equilibrium populations ( $p_A p_B$ )) (see Ref. 49 and references therein).

Our results demonstrate that ECP and RNase A reveal four clusters of dynamic residues closely mapping the structural clusters shown in Fig. 3. We also observed that conformational exchange in these clusters is significantly different between the two proteins (Fig. 5). In both homologues, similar contiguous residues experiencing conformational exchange interact with their structural neighbors to form bundles of cross-talking networks that transfer millisecond motions over structurally distinct clusters and subdomains of the protein. Although ECP and RNase A display different catalytic activities and binding properties, these two structural homologues nevertheless constrain millisecond dynamics to these four dynamic clusters instead of displaying randomly scattered ms dynamics throughout the protein structure. These localized dynamic clusters suggest that conformational exchange may have evolved through the formation of confined and contiguous networks of coupled motions to support common function and/or that ms dynamics

## Flexible Residue Clusters in Enzyme Homologues

may play distinct roles in the biological function of these structural and functional homologues. This clustering of ms dynamics into contiguous yet independent motional sectors may thus serve as a point of comparison to analyze functional differences between enzyme homologues.

A direct comparison of the dynamic clusters between RNase A and ECP indicates that some residue clusters experiencing conformational exchange in one protein homologue are almost entirely silenced in the other. This can be observed for the core residues of clusters 2 and 3. For instance, 12 residues of cluster 2 show considerable conformational exchange in ECP (Fig. 5*a*), a region that is almost completely devoid of ms dynamics in RNase A (Fig. 5*c*). Simultaneous fitting of the all-time scale, two-state Carver-Richards equation (34) to the spin relaxation data obtained at 500 and 800 MHz for all ECP residues of cluster 2 yielded a global exchange rate constant ( $k_{ex}$ ) of  $728 \pm 104 \text{ s}^{-1}$ . Inversely, although cluster 3 motions are entirely absent from ECP, 11 nearly contiguous residues linking the active site to loop 1 in RNase A show a global  $k_{ex}$  of  $1,438 \pm 125 \text{ s}^{-1}$  (Fig. 5*c*). Cluster 1 is the only protein sector showing important conformational exchange in both structural homologues. Relaxation dispersion is observed for 5 and 6 residues of loop 4 in RNase A and ECP, respectively. Whereas the global exchange rate of loop 4 ( $k_{ex} = 1548 \pm 82 \text{ s}^{-1}$ ) parallels the product release and catalytic rates in RNase A ( $k_{off} = k_{cat}$ ), the global exchange rate for loop 4 is more than 3 times slower in ECP ( $k_{ex} = 504 \pm 44 \text{ s}^{-1}$ ) and appears to be unrelated to  $k_{cat}$  (18). Interestingly, cluster 3 in RNase A encompasses the same long scale residues experiencing chemical shift variations upon pyrimidine binding (Fig. 2*b*), in line with previous hypotheses linking motions of this cluster to product release in RNase A (22–25). However, efforts to correlate free and 3'-UMP- or 5'-AMP-bound  $^{15}\text{N}$  chemical shift variations ( $\Delta\delta$  ppm) with  $^{15}\text{N}$  chemical shift differences between the major and the minor excited state obtained from  $^{15}\text{N}$  CPMG ( $\Delta\omega$ ) remained inconclusive (supplemental Fig. S3). In an ideal case, *i.e.* when the excited state of the enzyme corresponds to the product-bound form obtained by a shift in the reaction equilibrium (in this case saturation with 3'-UMP or 5'-AMP product analogues), one should expect a perfect linear correlation between the  $^{15}\text{N}$  chemical shift differences of the NMR titrations ( $\Delta\delta$  ppm) and the  $\Delta\omega$  ppm obtained from  $^{15}\text{N}$  CPMG. The absence of such linear correlation implies additional perturbations caused by the arrival of the ligands and/or implies that the conformational exchange we observe by  $^{15}\text{N}$  CPMG is unrelated to transitions between free and product-bound states in both ECP and RNase A. The observation of exchange at certain sites may be reflecting or providing a measure of ground state destabilization that is essential for catalysis but unrelated to the functional importance of the excited state in product release. The integrity of such motional networks nevertheless remains critical for optimal enzyme catalysis in both homologues, as was previously demonstrated elsewhere (22–25, 41) and in the present study (see below).

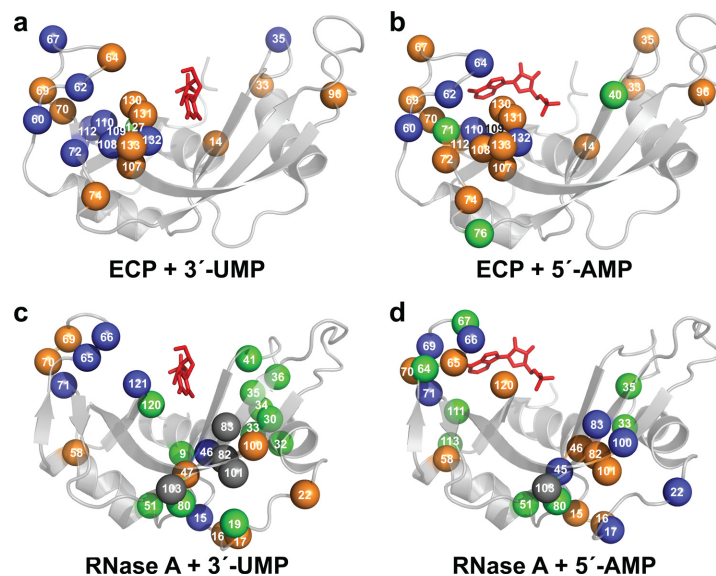
**The RNase A<sub>ECP</sub> Chimera; Translating RNase A Dynamics into Those of ECP**—Some of us have previously shown that cluster 3 residues in RNase A propagate motions from loop 1 (residues 14–24) to  $\beta$  strands 1 and 4, both of them containing the

important active site residues Thr<sup>45</sup> and Asp<sup>83</sup> involved in RNA pyrimidine stabilization at the active site (22–25). These motions correlate to those of the distant loop 1 and affect the rate-limiting step of RNase A catalysis by limiting the rate of product release. This was demonstrated by creating a chimeric hybrid between RNase A and ECP in which loop 1 in the former is replaced by loop 1 of the latter (generating the RNase A<sub>ECP</sub> chimera), affecting both the millisecond dynamics and activity of the enzyme (22). Not only is loop 1 conformationally restrained and much shorter in ECP than in RNase A (6 residues in the former *versus* 12 in the latter); ECP also lacks the important His<sup>48</sup>-Thr<sup>82</sup> residue pair that we showed to be essential for the propagation of ms motions in this dynamic cluster (23). As mentioned above, cluster 3 residues are completely devoid of conformational exchange in ECP (Fig. 5*a*), highlighting a different dynamic landscape in this protein.

Originally not apparent from the first chimera study (22), one of the most impressive observations of the global dynamic comparison between ECP and RNase A is the motional similarity between ECP and the swapped mutant of RNase A (RNase A<sub>ECP</sub>; Fig. 5, *a* and *b*). Despite conserving 90% sequence identity with RNase A, the RNase A<sub>ECP</sub> chimera almost perfectly mirrors the ms dynamics of ECP. Indeed, the conformational exchange of cluster 3 residues is completely dampened in the RNase A<sub>ECP</sub> hybrid. Conversely, millisecond dynamics are acquired for contiguous residues in cluster 2 (Fig. 5*b*), a sector that is almost completely devoid of ms motions in RNase A (Fig. 5*c*). Even a few residues of cluster 4 now experience conformational exchange in this chimeric hybrid, a protein cluster that is otherwise very rigid in RNase A.

Consequently, replacing the distant loop 1 in RNase A not only reduces ligand binding by 1 order of magnitude (22); it also affects the entire conformational flexibility of the enzyme by translating the dynamic landscape of WT RNase A into that observed in WT ECP. These results demonstrate that mutagenesis can be used to selectively modulate conformational exchange in proteins, a prerequisite to exert efficient control over motional residue networks affecting function in protein engineering and drug design applications (4, 5, 29).

**Relationship between Ligand Binding and Internal Dynamics**—To clarify the effect of ligand binding on the collective dynamics of each structural homologue, we also performed  $^{15}\text{N}$  CPMG experiments in the presence of saturating concentrations of 3'-UMP and 5'-AMP (Fig. 6). For both enzymes, ligand-bound complexes show that conformational exchange is still segregated to the same four clusters as those observed for the unbound forms. Using relaxation dispersion measurements, it may be possible to show how internal conformational equilibrium is shifted by the addition of a ligand. This “population shift” (6) has been demonstrated in RNase A by comparing frequency differences between major and minor dynamic states between the free and ligand-bound enzyme (41). This analysis requires very high quality relaxation data because the frequency difference of exchanging conformers is most difficult to obtain in the fast exchange regime (53). To attempt this analysis in ECP, we performed global fittings of residues in each dynamic individual cluster (1 and 2) such that  $k_{ex}$  and populations were shared while the frequency difference was individual for each



**FIGURE 6. Conformational exchange induced by 3'-UMP and 5'-AMP binding to ECP and RNase A.** Catalytic time scale (ms) residue dynamics were probed using  $^{15}\text{N}$  CPMG relaxation dispersion experiments for 3'-UMP-bound (a and c) and 5'-AMP-bound (b and d) forms of ECP (a and b) and RNase A (c and d). *Blue spheres*, residues showing  $^{15}\text{N}$  CPMG dispersion profiles with a  $\Delta R_2(1/\tau_{cp}) > 2 \text{ s}^{-1}$ ; *orange spheres*, residues showing no conformational exchange relative to the apo form (i.e. dampened ms dynamics upon ligand binding); *green spheres*, residues gaining conformational exchange ( $\Delta R_2(1/\tau_{cp}) > 2 \text{ s}^{-1}$ ) upon ligand binding; *gray spheres*, assigned residues in the apo form that cannot be assigned in the ligand-bound form due to line broadening. Residues are highlighted on the three-dimensional structure of ECP (PDB entry 1H1H) and RNase A (PDB entry 1O0N). The 3'-UMP and 5'-AMP ligands are shown in red.

residue. The results were filtered on the basis of the coefficient of determination ( $R^2 > 0.9$ ) to ensure that only good fits were used in the analysis. Due to spectral overlap induced by ligand binding and variable data quality, only residues Cys<sup>62</sup> and Thr<sup>67</sup> (cluster 1) were available in the free and ligand-bound forms for comparison. For Cys<sup>62</sup>, we determined that the frequency difference between major and minor conformers was in the 95% confidence interval of 320–450  $\text{s}^{-1}$  for apo-ECP and 300–530  $\text{s}^{-1}$  for the ECP-3'-UMP complex. For Thr<sup>67</sup>, the corresponding ranges for the same complexes were 180–240 and 190–290  $\text{s}^{-1}$ , respectively. Similarity of the frequency differences between conformations in the free and bound forms of ECP may indicate that the same type of population shift is induced by the ligand in ECP as the one observed in the earlier RNase A report (41).

Our NMR results nevertheless point to significant motional distinctions in the behavior of ECP and RNase A upon ligand binding. The most striking effect is the gaining and dampening of conformational exchange in the aforementioned motional clusters regardless of the RNA ligand investigated. Indeed, whereas 3'-UMP and 5'-AMP considerably rigidify the ECP structure upon binding (Fig. 6, a and b), the opposite is true for RNase A, for which an increase in the total amount of residues displaying relaxation dispersion can be observed (Fig. 6, c and d). A total of 11 (8) and 14 (9) residues rigidify upon 3'-UMP and 5'-AMP binding to ECP (RNase A), respectively (supplemental Figs. S4 and S5). However, the increase in conformational exchange upon ligand binding is considerably more pro-

nounced in RNase A than ECP. This is particularly true for the 3'-UMP complex (Fig. 6c), in which ligand binding induces conformational exchange for all residues of  $\alpha 2$  facing the hydrophobic core of the enzyme (residues 25–34, cluster 4). Globally, only 1 and 3 residues gain ms dynamics when 3'-UMP and 5'-AMP bind to ECP, respectively. This number jumps to 12 and 8 residues when the same two ligands bind to RNase A. The increase in RNase A ms dynamics upon ligand binding is also supported by the disappearance of several  $^1\text{H}$ - $^{15}\text{N}$  HSQC backbone resonances due to line broadening in the hinge region of cluster 3 (Fig. 6, c and d, *gray spheres*). Overall, these observations are consistent with the previously suggested ligand-dependent induction of hinge motions between the two  $\beta$ -sheets (54), a distinctive feature of RNase A that is undetectable in ECP.

The comparison between residues experiencing nitrogen chemical shift variations and the effect they exert on conformational exchange upon ligand binding also yields interesting distinctions between ECP and RNase A (supplemental Figs. S4 and S5). For instance, the strictly conserved catalytic lysine (Lys<sup>41</sup> in RNase A, Lys<sup>38</sup> in ECP) is thought to be involved in the stabilization of the excess negative charge on the phosphoryl oxygens in the transition state during RNA cleavage (15). Surprisingly, although the flexible Lys<sup>41</sup> in RNase A shows significant chemical shift variations in the presence of 3'-UMP, Lys<sup>38</sup> is not at all affected by ligand binding and does not show any conformational exchange in ECP (supplemental Fig. S4). Although Lys<sup>41</sup> gains conformational exchange in the RNase A-nucleotide



## Flexible Residue Clusters in Enzyme Homologues

complex, this result is not observed in ECP. This different binding and dynamic behavior for such an important catalytic residue may partly explain the important differences observed in the catalytic rates of these two RNases (18), which is also supported by the restricted  $^1\text{H}$ - $^{15}\text{N}$  bond vector of Lys<sup>38</sup> on the ps-ns time scale. Also, it is interesting to point out that Phe<sup>46</sup>, one of the most critical aromatic residues for RNase A stability and folding (55–57), is flexible in both the apo and 3'-UMP-bound forms of RNase A (supplemental Fig. S4). Although very highly conserved in the RNase family and positioned in the exact same orientation, no such flexibility is observed for Phe<sup>43</sup> in ECP. These observations suggest that very similar functional homologues may still use local dynamics for very different purposes in defining function and stability.

### DISCUSSION

Most of the work investigating the importance of conserved functional dynamics among enzyme family homologues was historically achieved through computational analyses. These studies typically use coarse-grained models, such as normal mode analysis, to infer catalytic time scale motional similarities between protein folds (58). Generally, these studies also involve the direct comparison of simulated data with crystallographically resolved and/or NMR-resolved enzymes in apo and ligand-bound forms, inferring dynamic information through, but not limited to, residual dipolar coupling measurements and/or Debye-Waller factors (*B*-factors) (59). The latter provides a rough approximation of residue flexibility in protein crystals, albeit with no clear definition of dynamic time scale.

Theoretical studies have delivered a breadth of relevant information on slow time scale conservation of functionally relevant motions in many enzyme families and protein folds, including, but not limited to, the amino acid kinase family (59), aspartate proteases (11), oxidoreductases, and peptidyl-prolyl isomerases (12). Nonetheless, experimental validation of such theoretical observations remains elusive, and very few experimental studies have yet provided a clear portrait of catalytically relevant dynamic conservation between structural and functional protein homologues. The present work partly provides such validation by showing that two members of the pancreatic-like ribonuclease family sharing similar catalytic mechanisms and protein folds also retain similar ps-ns dynamics and conformational exchange on the ms time scale. The present study also validates the existence of dynamic clusters, similar to the evolutionary sectors of Ranganathan and co-workers (60, 61), which transcend the classical definition of primary, secondary, or tertiary protein architectures. We observe that millisecond dynamic clusters are sequestered into distinct protein subdomains and, as previously suggested, appear to be structurally encoded and "fine tuned" by the protein fold (Fig. 5) (60). The dynamic subdomain (cluster 3) that transfers millisecond motions essential to product release in RNase A is absent from ECP. Despite having the same dynamic cluster architecture as that of RNase A, experimental evidence illustrates that ECP is a more conformationally restrained homologue on the ms time scale.

Previous theoretical studies have suggested the existence of collective motions in the pancreatic-like ribonuclease family

using molecular dynamics simulations (62, 63). Until very recently, such analyses remained confined to the very fast time scales that can be probed by MD simulations (ps-ns), a time frame that is still very far from enzyme turnover (ms). To overcome such a barrier, Ramanathan and Agarwal (12) recently used quasiharmonic analysis to model large scale conformational fluctuations occurring on slower time scales in this enzyme family, providing the most complete theoretical comparison of slow dynamics between ribonuclease homologues so far. Focusing on the motional comparison of three pancreatic-like ribonuclease homologues from humans, bovines, and rats, the authors show remarkable interspecies similarities in the slow motions of surface loops and distal motifs in this enzyme family, observations that transcend sequence identity. The present investigation offers the first experimental validation of the existence of such theoretical networks between two members of the pancreatic ribonuclease family.

The conservation of clustering sectors of coevolving residues among family members and functional homologues has been suggested as a potentially important structural determinant defining biological function in proteins. Ranganathan and co-workers (60) have used statistical coupling analyses of sequence alignments to argue that the classical hierarchy of primary, secondary, tertiary, and quaternary structures used to define protein architecture remains largely insufficient to explain the three-dimensional cooperativity observed between coevolving residues among protein homologues. Much like the confined assemblies of dynamic residue clusters highlighted in the present study, the investigators argue that non-random correlations in the physical connectivity between coevolving groups of residues in the three-dimensional structure of the S1A serine protease family underlie the biological function of its members, a property they also observed in the PDZ, PAS, Src homology 2, and Src homology 3 domain families (60). These groups of coevolving residues, termed sectors (60, 61), show striking physical connectivity in the three-dimensional structure of protein families while completely transcending classical subdivisions of primary, secondary, or tertiary subdomain architectures. These sectors, which show stark three-dimensional similarity with the dynamic clusters observed in the present work, are shown to be structurally independent and confined to functional regions among the serine protease family, encompassing residues involved in binding or catalysis (60). Most interestingly, Ranganathan and co-workers (61) also observe a direct correlation between these coevolving sectors and the networks of residues undergoing conformational fluctuations associated with enzyme catalysis in dihydrofolate reductase. Similar to the experimentally characterized clusters in the present work, the authors show that dynamic motions associated with catalysis in dihydrofolate reductase extend well beyond the active site environment, dynamically connecting surface sites on both ends of the enzyme through long range networks of dynamic residues that cross-talk on the catalytic time scale (61).

The present study, coupled to our previous demonstration that RNase A relies on a 20-Å communication link between the active site and loop 1 for optimal catalysis (22–25), suggests that the Ranganathan sectors of coevolving residues may represent a

conserved mechanism to initiate allosteric regulation on protein surfaces (61, 64). We experimentally uncovered a similar architecture between two homologues of the ribonuclease fold, lending further support to the potential evolutionary conservation of motional networks defining function in protein families. Systematic analysis of correlations between dynamics and function for important protein families is likely to provide novel insights into the structure-function-flexibility relationship that defines protein architectures. This knowledge is required for successful *de novo* enzyme design (4, 5, 29) as well as for the development of new allosteric drugs (28).

*Acknowledgments*—We thank Gennady Khirich and J. Patrick Loria (Yale University) for helpful discussions and for providing part of the NMR assignments of RNase A and RNase A<sub>ECP</sub>. We also thank Tara Sprules and Sameer Al-Abdul-Wahid from the Québec/Eastern Canada High Field NMR Facility for excellent NMR technical assistance as well as Éric Déziel and Claude Guertin (Institut National de la Recherche Scientifique) for kind generosity.

## REFERENCES

- Branden, C., and Tooze, J. (1999) *Introduction to Protein Structure*, 2nd Ed., Garland Science, New York
- Eiben, C. B., Stegel, J. B., Bale, J. B., Cooper, S., Khatib, F., Shen, B. W., Players, F., Stoddard, B. L., Popovic, Z., and Baker, D. (2012) Increased Diels-Alderase activity through backbone remodeling guided by Foldit players. *Nat. Biotechnol.* **30**, 190–192
- Siegel, J. B., Zanghellini, A., Lovick, H. M., Kiss, G., Lambert, A. R., St Clair, J. L., Gallaher, J. L., Hilvert, D., Gelb, M. H., Stoddard, B. L., Houk, K. N., Michael, F. E., and Baker, D. (2010) Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction. *Science* **329**, 309–313
- Baker, D. (2010) An exciting but challenging road ahead for computational enzyme design. *Protein Sci.* **19**, 1817–1819
- Lassila, J. K. (2010) Conformational diversity and computational enzyme design. *Curr. Opin. Chem. Biol.* **14**, 676–682
- Ma, B., and Nussinov, R. (2010) Enzyme dynamics point to stepwise conformational selection in catalysis. *Curr. Opin. Chem. Biol.* **14**, 652–659
- Henzler-Wildman, K., and Kern, D. (2007) Dynamic personalities of proteins. *Nature* **450**, 964–972
- Baldwin, A. J., and Kay, L. E. (2009) NMR spectroscopy brings invisible protein states into focus. *Nat. Chem. Biol.* **5**, 808–814
- Doucet, N. (2011) Can enzyme engineering benefit from the modulation of protein motions? Lessons learned from NMR relaxation dispersion experiments. *Protein Pept. Lett.* **18**, 336–343
- Loria, J. P., Berlow, R. B., and Watt, E. D. (2008) Characterization of enzyme motions by solution NMR relaxation dispersion. *Acc. Chem. Res.* **41**, 214–221
- Carnevale, V., Raugel, S., Micheletti, C., and Carloni, P. (2006) Convergent dynamics in the protease enzymatic superfamily. *J. Am. Chem. Soc.* **128**, 9766–9772
- Ramanathan, A., and Agarwal, P. K. (2011) Evolutionarily conserved linkage between enzyme fold, flexibility, and catalysis. *PLoS Biol.* **9**, e1001193
- Cuchillo, C. M., Nogués, M. V., and Raines, R. T. (2011) Bovine pancreatic ribonuclease. Fifty years of the first enzymatic reaction mechanism. *Biochemistry* **50**, 7835–7841
- Marshall, G. R., Feng, J. A., and Kuster, D. J. (2008) Back to the future. Ribonuclease A. *Biopolymers* **90**, 259–277
- Raines, R. T. (1998) Ribonuclease A. *Chem. Rev.* **98**, 1045–1066
- Dyer, K. D., and Rosenberg, H. F. (2006) The RNase A superfamily. Generation of diversity and innate host defense. *Mol. Divers.* **10**, 585–597
- Sorrentino, S. (2010) The eight human “canonical” ribonucleases. Molecular diversity, catalytic properties, and special biological actions of the enzyme proteins. *FEBS Lett.* **584**, 2194–2200
- Boix, E. (2001) Eosinophil cationic protein. *Methods Enzymol.* **341**, 287–305
- Huang, Y. C., Lin, Y. M., Chang, T. W., Wu, S. J., Lee, Y. S., Chang, M. D., Chen, C., Wu, S. H., and Liao, Y. D. (2007) The flexible and clustered lysine residues of human ribonuclease 7 are critical for membrane permeability and antimicrobial activity. *J. Biol. Chem.* **282**, 4626–4633
- Boix, E., and Nogués, M. V. (2007) Mammalian antimicrobial proteins and peptides. Overview on the RNase A superfamily members involved in innate host defence. *Mol. Biosyst.* **3**, 317–335
- Jarymowicz, V. A., and Stone, M. J. (2006) Fast time scale dynamics of protein backbones. NMR relaxation methods, applications, and functional consequences. *Chem. Rev.* **106**, 1624–1671
- Doucet, N., Watt, E. D., and Loria, J. P. (2009) The flexibility of a distant loop modulates active site motion and product release in ribonuclease A. *Biochemistry* **48**, 7160–7168
- Doucet, N., Khirich, G., Kovrigin, E. L., and Loria, J. P. (2011) Alteration of hydrogen bonding in the vicinity of histidine 48 disrupts millisecond motions in RNase A. *Biochemistry* **50**, 1723–1730
- Watt, E. D., Rivalta, I., Whittier, S. K., Batista, V. S., and Loria, J. P. (2011) Reengineering rate-limiting, millisecond enzyme motions by introduction of an unnatural amino acid. *Biophys. J.* **101**, 411–420
- Watt, E. D., Shimada, H., Kovrigin, E. L., and Loria, J. P. (2007) The mechanism of rate-limiting motions in enzyme function. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 11981–11986
- Butterwick, J. A., Loria, J. P., Astrof, N. S., Kroenke, C. D., Cole, R., Rance, M., and Palmer, A. G., 3rd. (2004) Multiple time scale backbone dynamics of homologous thermophilic and mesophilic ribonuclease HI enzymes. *J. Mol. Biol.* **339**, 855–871
- Henzler-Wildman, K. A., Lei, M., Thai, V., Kerns, S. J., Karplus, M., and Kern, D. (2007) A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* **450**, 913–916
- Lee, G. M., and Craik, C. S. (2009) Trapping moving targets with small molecules. *Science* **324**, 213–215
- Davey, J. A., and Chica, R. A. (2012) Multistate approaches in computational protein design. *Protein Sci.* **21**, 1241–1252
- Sela, M., and Anfinsen, C. B. (1957) Some spectrophotometric and polarimetric experiments with ribonuclease. *Biochim. Biophys. Acta* **24**, 229–235
- Loria, J. P., Rance, M., and Palmer, A. G., 3rd (1999) A relaxation-compensated Carr-Purcell-Meiboom-Gill sequence for characterizing chemical exchange by NMR spectroscopy. *J. Am. Chem. Soc.* **121**, 2331–2332
- Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293
- Goddard, T. D., and Kneller, D. G. (2008) *Sparky*, version 3, University of California, San Francisco, CA
- Carver, J. P., and Richards, R. E. (1972) A general two-site solution for the chemical exchange produced dependence of  $T_2$  upon the Carr-Purcell pulse separation. *J. Magn. Reson.* **6**, 89–105
- Cole, R., and Loria, J. P. (2002) Evidence for flexibility in the function of ribonuclease A. *Biochemistry* **41**, 6072–6081
- Doucet, N., Savard, P. Y., Pelletier, J. N., and Gagné, S. M. (2007) NMR investigation of Tyr<sup>105</sup> mutants in TEM-1  $\beta$ -lactamase. Dynamics are correlated with function. *J. Biol. Chem.* **282**, 21448–21459
- Morin, S. (2011) A practical guide to protein dynamics from <sup>15</sup>N spin relaxation in solution. *Prog. Nucl. Magn. Reson. Spectrosc.* **59**, 245–262
- d’Auvergne, E. J., and Gooley, P. R. (2008) Optimisation of NMR dynamic models II. A new methodology for the dual optimisation of the model-free parameters and the Brownian rotational diffusion tensor. *J. Biomol. NMR* **40**, 121–133
- d’Auvergne, E. J., and Gooley, P. R. (2008) Optimisation of NMR dynamic models I. Minimisation algorithms and their performance within the model-free and Brownian rotational diffusion spaces. *J. Biomol. NMR* **40**, 107–119
- Kovrigin, E. L. (2012) NMR line shapes and multi-state binding equilibria. *J. Biomol. NMR* **53**, 257–270
- Kovrigin, E. L., and Loria, J. P. (2006) Enzyme dynamics along the reaction coordinate. Critical role of a conserved residue. *Biochemistry* **45**,

## Flexible Residue Clusters in Enzyme Homologues

- 2636–2647
42. Kaplan, J. L., and Fraenkel, G. (1980) *NMR of Chemically Exchanging Systems*, Academic Press, Inc., New York
43. Rao, B. D. (1989) Nuclear magnetic resonance line-shape analysis and determination of exchange rates. *Methods Enzymol.* **176**, 279–311
44. Mallorquí-Fernández, G., Pous, J., Peracaula, R., Aymami, J., Maeda, T., Tada, H., Yamada, H., Seno, M., de Llorens, R., Gomis-Rüth, F. X., and Coll, M. (2000) Three-dimensional crystal structure of human eosinophil cationic protein (RNase 3) at 1.75 Å resolution. *J. Mol. Biol.* **300**, 1297–1307
45. Boix, E., Leonidas, D. D., Nikolovski, Z., Nogués, M. V., Cuchillo, C. M., and Acharya, K. R. (1999) Crystal structure of eosinophil cationic protein at 2.4 Å resolution. *Biochemistry* **38**, 16794–16801
46. Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules II. Analysis of experimental results. *J. Am. Chem. Soc.* **104**, 4559–4570
47. Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules I. Theory and range of validity. *J. Am. Chem. Soc.* **104**, 4546–4559
48. Mandel, A. M., Akke, M., and Palmer, A. G., 3rd. (1995) Backbone dynamics of *Escherichia coli* ribonuclease HI. Correlations with structure and function in an active enzyme. *J. Mol. Biol.* **246**, 144–163
49. Manley, G., and Loria, J. P. (2012) NMR insights into protein allostery. *Arch. Biochem. Biophys.* **519**, 223–231
50. Kay, L. E. (2005) NMR studies of protein structure and dynamics. *J. Magn. Reson.* **173**, 193–207
51. Palmer, A. G., 3rd (2001) NMR probes of molecular dynamics. Overview and comparison with other techniques. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 129–155
52. Palmer, A. G., 3rd (2004) NMR characterization of the dynamics of biomacromolecules. *Chem. Rev.* **104**, 3623–3640
53. Kovrigin, E. L., Kempf, J. G., Grey, M. J., and Loria, J. P. (2006) Faithful estimation of dynamics parameters from CPMG relaxation dispersion measurements. *J. Magn. Reson.* **180**, 93–104
54. Vitagliano, L., Merlino, A., Zagari, A., and Mazzarella, L. (2002) Reversible substrate-induced domain motions in ribonuclease A. *Proteins* **46**, 97–104
55. Chatani, E., Nonomura, K., Hayashi, R., Balny, C., and Lange, R. (2002) Comparison of heat- and pressure-induced unfolding of ribonuclease A. The critical role of Phe<sup>46</sup> which appears to belong to a new hydrophobic chain-folding initiation site. *Biochemistry* **41**, 4567–4574
56. Kadosono, T., Chatani, E., Hayashi, R., Moriyama, H., and Ueki, T. (2003) Minimization of cavity size ensures protein stability and folding. Structures of Phe<sup>46</sup>-replaced bovine pancreatic RNase A. *Biochemistry* **42**, 10651–10658
57. Vila, R., Benito, A., Ribó, M., and Vilanova, M. (2009) Mapping the stability clusters in bovine pancreatic ribonuclease A. *Biopolymers* **91**, 1038–1047
58. Bahar, I., and Rader, A. J. (2005) Coarse-grained normal mode analysis in structural biology. *Curr. Opin. Struct. Biol.* **15**, 586–592
59. Marcos, E., Crehuet, R., and Bahar, I. (2010) On the conservation of the slow conformational dynamics within the amino acid kinase family. NAGK the paradigm. *PLoS Comput. Biol.* **6**, e1000738
60. Halabi, N., Rivoire, O., Leibler, S., and Ranganathan, R. (2009) Protein sectors. Evolutionary units of three-dimensional structure. *Cell* **138**, 774–786
61. Reynolds, K. A., McLaughlin, R. N., and Ranganathan, R. (2011) Hot spots for allosteric regulation on protein surfaces. *Cell* **147**, 1564–1575
62. Merlino, A., Vitagliano, L., Ceruso, M. A., Di Nola, A., and Mazzarella, L. (2002) Global and local motions in ribonuclease A. A molecular dynamics study. *Biopolymers* **65**, 274–283
63. Merlino, A., Vitagliano, L., Ceruso, M. A., and Mazzarella, L. (2003) Subtle functional collective motions in pancreatic-like ribonucleases. From ribonuclease A to angiogenin. *Proteins* **53**, 101–110
64. Gunasekaran, K., Ma, B., and Nussinov, R. (2004) Is allostery an intrinsic property of all dynamic proteins? *Proteins* **57**, 433–443
65. Fisher, B. M., Schultz, L. W., and Raines, R. T. (1998) Coulombic effects of remote subsites on the active site of ribonuclease A. *Biochemistry* **37**, 17386–17401
66. Doucet, N., Jayasundera, T. B., Simonović, M., and Loria, J. P. (2010) The crystal structure of ribonuclease A in complex with thymidine-3'-monophosphate provides further insight into ligand binding. *Proteins* **78**, 2459–2468
67. Grzesiek, S., Stahl, S. J., Wingfield, P. T., and Bax, A. (1996) The CD4 determinant for downregulation by HIV-1 Nef directly binds to Nef. Mapping of the Nef binding surface by NMR. *Biochemistry* **35**, 10256–10261
68. Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B., Keduas, V., and Notredame, C. (2006) Expresso. Automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic Acids Res.* **34**, W604–W608

## **Supporting Information**

### **Conservation of flexible residue clusters among structural and functional enzyme homologues**

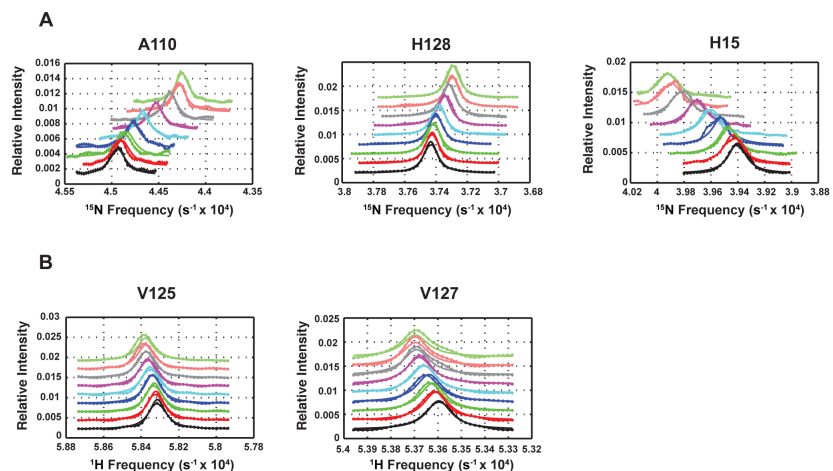
**Donald Gagné<sup>1</sup>, Laurie-Anne Charest<sup>1</sup>, Sébastien Morin<sup>2</sup>  
Evgenii L. Kovrigin<sup>3</sup> & Nicolas Doucet<sup>1\*</sup>**

<sup>1</sup>INRS-Institut Armand-Frappier, Université du Québec, 531 Boulevard des Prairies, Laval, QC H7V 1B7, CANADA

<sup>2</sup>Swiss Institute of Bioinformatics and Biozentrum, University of Basel, Klingelbergstrasse 50/70, 4056 Basel, SWITZERLAND

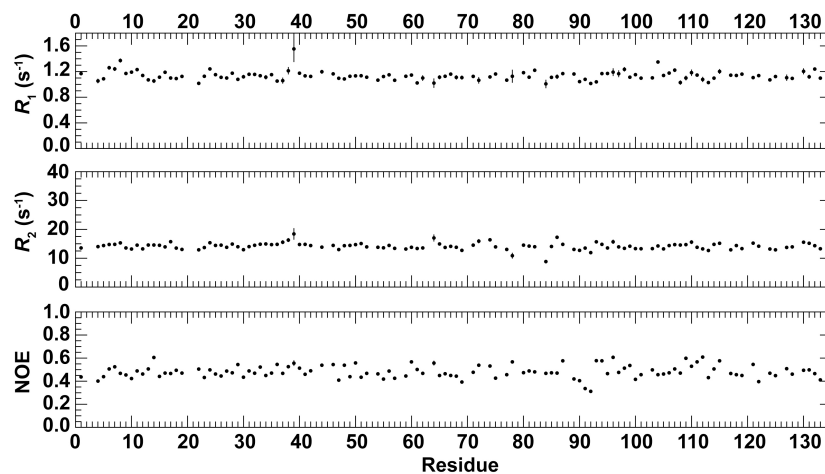
<sup>3</sup>Chemistry Department, Marquette University, P.O. Box 1881, Milwaukee, Wisconsin 53201, USA

\*Corresponding author: [nicolas.doucet@iaf.inrs.ca](mailto:nicolas.doucet@iaf.inrs.ca)

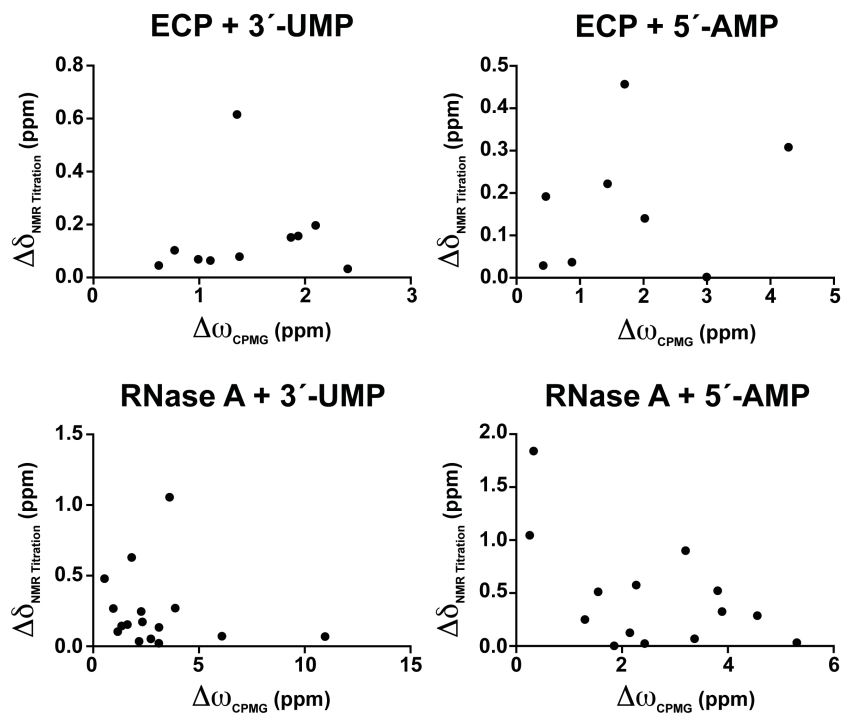


**Figure S1. NMR line shape analysis of ECP titrations with 3'-UMP and 5'-AMP.**

The best-fit results for the ECP titrations with (a) 5'-AMP and (b) 3'-UMP are presented. One-dimensional slices extracted from experimental  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra are shown with circles (connected with straight lines to guide the eye). The best-fit of the two-state binding mechanism is shown as a solid line. The bottom trace in all graphs corresponds to the ligand-free ECP spectrum. The spectra from subsequent titration points are displaced vertically for easier viewing corresponding to the ligand-to-protein molar ratios of 0,  $0.18 \pm 0.02$ ,  $0.39 \pm 0.03$ ,  $0.69 \pm 0.05$ ,  $1.31 \pm 0.09$ ,  $2.7 \pm 0.2$ ,  $6.0 \pm 0.4$ ,  $12.0 \pm 0.8$ , and  $18 \pm 1$ .



**Figure S2.  $R_1$ ,  $R_2$  and steady state heteronuclear NOE relaxation parameters plotted on the sequence of ECP.**  $^{15}\text{N}$ - $R_1$ ,  $^{15}\text{N}$ - $R_2$  and steady state heteronuclear NOE experiments were performed in an interleaved fashion. The  $R_1$  experiments were performed with relaxation delays of 10, 50, 100 ( $\times 2$ ), 200, 350, 700, 1100, and 1400 ms. The  $R_2$  experiments were performed with relaxation delays of 10, 30, 50 ( $\times 2$ ), 70, 90, 110, and 130 ms. All NMR experiments were recorded at 500 MHz on a sample containing 0.2 mM  $^{15}\text{N}$ -labelled ECP in 15 mM sodium acetate, 10%  $^2\text{H}_2\text{O}$ , pH 5.0 at 298K.



**Figure S3. Comparison between the  $^{15}\text{N}$   $\Delta\delta_{\text{NMR titration}}$ , determined by the difference in the chemical shift between the apo and ligand bound forms of ECP (RNase A), and  $\Delta\omega_{\text{CPMG}}$ , determined by the difference in the chemical shifts of the ground and excited states of ECP (RNase A) extracted by  $^{15}\text{N}$ -CPMG. The absence of a linear correlation implies additional perturbations caused by the arrival of the ligands and/or that the conformational exchange we observe by  $^{15}\text{N}$ -CPMG is unrelated to transitions between free and product-bound states in both ECP and RNase A.**

A

ECP + 3'-UMP			
Residues showing $^{15}\text{N}-\Delta\delta > 0.5$ ppm upon 3'-UMP binding	Residues showing $^{15}\text{N}$ -CPMG dispersion with $\Delta R_2 (1/\tau_{cp}) > 2 \text{ s}^{-1}$ in 3'-UMP-bound ECP	Residue gaining conformational exchange upon 3'-UMP binding to ECP [ $\Delta R_2 (1/\tau_{cp}) > 2 \text{ s}^{-1}$ ]	Residues showing dampened ms dynamics upon 3'-UMP binding to ECP
Q4	W35	V127	Q14
F11	I60		Y33
Q14	C62		H64
N31	T67		N69
F48	H72		N70
C55	V108		S74
C62	V109		C96
H64	A110		Y107
N69	D112		D130
R75	T132		T131
C96			I133
L129			
D130			

B

RNase A + 3'-UMP			
Residues showing $^{15}\text{N}-\Delta\delta > 0.5$ ppm upon 3'-UMP binding	Residues showing $^{15}\text{N}$ -CPMG dispersion with $\Delta R_2 (1/\tau_{cp}) > 2 \text{ s}^{-1}$ in 3'-UMP-bound RNase A	Residue gaining conformational exchange upon 3'-UMP binding to RNase A [ $\Delta R_2 (1/\tau_{cp}) > 2 \text{ s}^{-1}$ ]	Residues showing dampened ms dynamics upon 3'-UMP binding to RNase A
Q11	S15	E9	S16
T17	F46	A19	T17
K41	C65	M30	S22
N44	K66	S32	V47
T45	N71	R33	C58
F46	D121	N34	Q69
V47		L35	T70
L51		T36	T100
T82		K41	
A96		L51	
Q101		S80	
H119		F120	
A122			



**Figure S4. Comparison between nitrogen chemical shift mapping and conformational exchange induced by 3'-UMP binding to ECP and RNase A.** (A)  $^{15}\text{N}$  chemical shift variations ( $\Delta\delta$ ) compared to the conformational exchange behavior induced by 3'-UMP binding to ECP. (B)  $^{15}\text{N}$  chemical shift variations ( $\Delta\delta$ ) compared to the conformational exchange behavior induced by 3'-UMP binding to RNase A. Orange boxes: residues with significant chemical shift variations that also experience dampened conformational exchange upon 3'-UMP binding. Blue boxes: residues with significant chemical shift variations that conserve a similar conformational exchange between the apo and 3'-UMP-bound enzyme. Green boxes: residues with significant chemical shift variations that gain conformational exchange upon 3'-UMP binding.

A

ECP + 5'-AMP			
Residues showing $^{15}\text{N}$ - $\Delta\delta > 0.5$ ppm upon 5'-AMP binding	Residues showing $^{15}\text{N}$ -CPMG dispersion with $\Delta R_2 (1/\tau_{cp}) > 2 \text{ s}^{-1}$ in 5'-AMP-bound ECP	Residue gaining conformational exchange upon 5'-AMP binding to ECP [ $\Delta R_2 (1/\tau_{cp}) > 2 \text{ s}^{-1}$ ]	Residues showing dampened ms dynamics upon 5'-AMP binding to ECP
W10	I60	Q40	Q14
S59	C62	C71	Y33
H128	H64	F76	W35
L129	V109		T67
D130	A110		N69
	T132		N70
			H72
			S74
			C96
			Y107
			V108
			D112
			D130
			T131

B

RNase A + 5'-AMP			
Residues showing $^{15}\text{N}$ - $\Delta\delta > 0.5$ ppm upon 5'-AMP binding	Residues showing $^{15}\text{N}$ -CPMG dispersion with $\Delta R_2 (1/\tau_{cp}) > 2 \text{ s}^{-1}$ in 5'-AMP-bound RNase A	Residue gaining conformational exchange upon 5'-AMP binding to RNase A [ $\Delta R_2 (1/\tau_{cp}) > 2 \text{ s}^{-1}$ ]	Residues showing dampened ms dynamics upon 5'-AMP binding to RNase A
Q11	T17	R33	S15
V43	S22	L35	S16
T45	V47	L51	F46
F46	K66	A64	C58
L51	Q69	N67	C65
C58	N71	S80	T70
A64	D83	E111	T82
C65	T100	N113	Q101
K66			D121
G68			
Q69			
N71			
Q101			
H119			
F120			

**Figure S5. Comparison between nitrogen chemical shift mapping and conformational exchange induced by 5'-AMP binding to ECP and RNase A.** (A)  $^{15}\text{N}$  chemical shift variations ( $\Delta\delta$ ) compared to the conformational exchange behavior induced by 5'-AMP binding to ECP. (B)  $^{15}\text{N}$  chemical shift variations ( $\Delta\delta$ ) compared to the conformational exchange behavior induced by 5'-AMP binding to RNase A. Orange boxes: residues with significant chemical shift variations that also experience dampened conformational exchange upon 5'-AMP binding. Blue boxes: residues with significant chemical shift variations that conserve a similar conformational exchange between the apo and 5'-AMP-bound enzyme. Green boxes: residues with significant chemical shift variations that gain conformational exchange upon 5'-AMP binding.

## 4.5 Discussion

Encore peu d'études portent sur la conservation de la flexibilité atomique chez les membres d'une même famille d'homologues structuraux de fonction similaire. La plupart des études disponibles proviennent de simulations de dynamique moléculaire à des échelles de temps plus rapides que celles correspondant à la catalyse enzymatique, ou encore de la cristallographie. Dans cet article, nous avons utilisé le  $^{15}\text{N}$ -CPMG pour valider le modèle de conservation de la dynamique chez des homologues structuraux. Les résultats démontrent que les mouvements atomiques globales dans l'ordre de la  $\mu\text{s}$ - $\text{ms}$  sont préservées entre la RNase A et ECP, qui ne partagent pourtant que 30% d'homologie de séquence. De plus, l'étude nous a aussi permis de valider la présence de régions dynamiques [172, 173]. De telles régions de résidus dynamiques avaient été observées par Ranganathan *et al.* chez d'autres systèmes enzymatiques en utilisant une méthode statistique d'évolution corrélée entre les acides aminés [172]. Selon les auteurs, ces régions, encodées dans la forme tridimensionnelle de la protéine, auraient une fonction distincte et seraient conservées entre les membres d'une même famille de protéines de fonctions identiques, mais de séquences divergentes. Dans cette étude, nous avons observé la présence de mouvements atomiques dans la boucle 4 et ce, chez les deux protéines. Des études de mutagenèse dirigée ont démontré que la boucle réduisait l'affinité du ligand pour la protéine, sans toutefois affecter l'activité enzymatique [91]. De façon intéressante, alors que la RNase A présente un réseau dynamique de résidus dans la région 3 correspondant aux structures secondaires L1/ $\beta$ 1/ $\beta$ 4/ $\beta$ 5, ECP demeure rigide dans la région correspondante. À l'inverse, la région 2 chez ECP se veut flexible, une région complètement rigide chez la RNase A. Il est postulé que ces régions, qui sont encodées dans la structure, sont importantes pour affiner la fonction de la protéine.

La préparation d'une chimère (RNase A<sub>ECP</sub>), où la boucle 1 (située dans la région 3) de la RNase A a été remplacée par celle d'ECP, nous a permis de démontrer qu'il était possible de modifier le profil dynamique de la protéine. En procédant à une analyse dynamique par  $^{15}\text{N}$ -CPMG, nous avons observé une réorganisation du profil dynamique, avec un déplacement de la dynamique de la région 3 (L1/ $\beta$ 1/ $\beta$ 4/ $\beta$ 5) à la région 2 ( $\beta$ 2/ $\beta$ 3/ $\beta$ 6/ $\beta$ 7/ $\alpha$ 3). Nous démontrons qu'il est possible de moduler la dynamique et d'influencer la fonction de la protéine par le changement d'une boucle distante du site actif. La compréhension des mouvements atomiques couplées à la fonction nous sera nécessaire pour le design *de novo* de protéines.

## 4.6 Résultats supplémentaires

Les données présentées dans l'article démontrent une conservation de la dynamique globale à l'échelle de temps de la  $\mu$ s-ms entre ECP et la RNase A, une observation qui va dans le même sens que les résultats d'études précédentes qui ont été résumés dans l'article présenté au chapitre 2 du présent document [78]. Nous avons montré que, malgré la présence de mouvements globaux, nous observons des réseaux de résidus dynamiques distincts à chaque protéine, similaire à ce qui avait été observé par Ranganathan avec d'autres systèmes enzymatiques [172]. La modulation de ces secteurs dynamiques affectait la fonction de la protéine, ce qui pouvait suggérer que les mouvements jouent un rôle de premier plan dans la fonction de la protéine. Bien que ces résultats soient intéressants, il n'en demeure pas moins que le nombre limité d'homologues structuraux étudiés jusqu'à maintenant ne permet pas d'obtenir un portrait complet. Dans les sections suivantes, nous présentons les résultats d'autres RNases humaines en utilisant la même approche scientifique que pour les articles précédents.

### 4.6.1 Caractérisation biophysique de la RNase 4 humaine

#### 4.6.1.1 Introduction

La RNase 4 fait partie des huit enzymes canoniques chez l'humain (Figure 4.1B). Sa structure tridimensionnelle a été obtenue en 1999 [174]. Elle est l'enzyme la plus conservée entre les espèces, ce qui laisse penser qu'elle joue un rôle primordial chez les vertébrés [126]. Comme pour la RNase A et ECP, elle conserve les résidus catalytiques et fonctionnels, les quatre ponts disulfures, de même que la structure tridimensionnelle caractéristique des RNases (Figure 4.1) [174]. Avec 119 acides aminés, elle est la plus petite des RNases humaines. Comme pour l'angiogénine, la RNase 4 possède un résidu pyroglutaminye en N-terminal [126]. C'est aussi le seul membre de la famille des RNases à n'avoir aucun site de glycosylation. La séquence contient une délétion de deux résidus entre les positions 74 et 75 et une extrémité C-terminale tronquée. La RNase 4 possède une activité catalytique plus élevée pour le UpA que pour le CpA [93]. Il a été proposé que l'arginine en position 101 jouerait un rôle similaire à celui de la sérine 123 chez la RNase A, en favorisant la liaison de l'uracile [175, 176]. Toutefois, d'autres études ont plutôt suggéré que le groupe libre du carboxyle en C-terminal pourrait réduire la liaison de la cytosine [126].

**A**

```

RNase4  ---QDGMVQRFLRQHVHFEET-GGSDRYCNLMQRRKMTLYHCKRFNTFHEHEDIWNIIRSI CSTTNIQCK--NGKMNCHEG--VVKVVDQRDIT
RNase1  ---KESRAKKFQRQHMDS DSSPSSSSTYCNOMRRRRNMTOGRCKP VNTFVHEPLVDVQNVCFQEKVTCK--NGQGNKYKSNSSMHIDCRLIT
RNase2  KPPQFTWAQWFETOHINMTS-----OQCTNAMQVINNYQRCKNONTFLRTTFANVVNVCGNPNMTCP SNKTRKNCHHSGSQVPLIHCDLIT
RNase3  RPPQFTWAQWFETOHISLNF-----PRCTIARAINNYRWCKNONTFLRTTFANVVNVCGNQSIHRC PHNRTLNNCHRSRFRVPLHCDLIT
RNase5  -Q-DNSRYTHFLTOHYDAKPO-GRDDRYCESIMRRRG-LTSPCKDINTFHEGKRSIKAI CENKNGNP--HR--ENLRISKSSFQVITCKLIT
RNase6  WPKRLTKAHWFEIOHIQSPF-----LQCNRAMSGTINNYTQCKKHONTFLHDSFQNVAAVCDLLSITVCK--NRRRNCHOSKSPVNMTCRLIT
RNase7  KPKGM TSSQWFKIOHMOPSP-----QACNSAMKNINKHTRCKDLNTFLHEPFS SVAATCQTFKIACK--NGDKNCHOSHGA VSLTMCKLIT
RNase8  KEKDM TSSQWFKIOHVQSPF-----QACNSAMSIINKYTERCKDLNTFLHEPFS SVAITCQTFNIACK--NSCKNCHOSHGPM SLMGELIT
RNaseA  ---KETAARAFERQHMDSSTS AASSSNYCNQMMSRRLTKDRCKPVNTFVHESLADVQAVCSQKNVACK--NGQTNQVQSYSTM SIVDCREIT

RNase4  90  100  110
RNase4  GS--S-R--APNCRYRAIASTRRVVIVACEG-----NPOVPVHFDG-----
RNase1  NG--S-R--YPCA VRTSPKERHILVACEG-----SPYVPVHFDASVEDST
RNase2  TP--S-PQNISNCRYAQTPANMFYIVACDNRDQRRDPPQYVPVHLDRII
RNase3  NPGAQ--N--ISNCRYADRPGRFVYVACDN-RDPRDSPRYVPVHLDTTI
RNase5  GG--S-P--WPFCCQYRATAGFRNVVACEN-----GLPVHLDQSI FRRP
RNase6  SG--KYP--QCRYSA AAYKFFIVACDP-PQKSDPP-YKLVVHLD SIL
RNase7  SG--KHP--NCRYKEKRONKSYVIVACKP-PQKSDQFHLVPHLD RVL
RNase8  SG--KYP--NCRYKEKHLNTFYIVACDP-PQQGDPG-YPLVPHLDKVV
RNaseA  GS--S-K--YPCA YKTTQANKHILVACEG-----NPIYVPVHFDASV

```

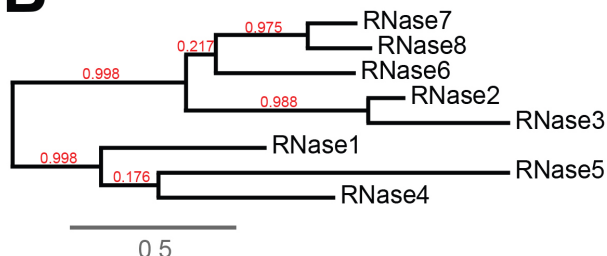
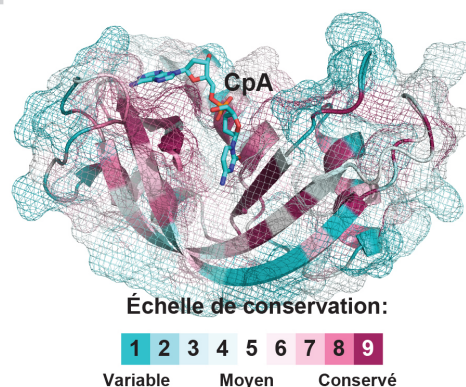
**B****C**

Figure 4.1. Conservation évolutive de la RNase 4 humaine. (A) Alignement de séquence de la RNase 4 par Expresso avec les RNases canoniques 1-8 et la RNase A (*Bos taurus*). Les résultats ont été transférés dans Boxshade pour une impression de meilleure qualité. Les résidus identiques sont représentés en noir. Les résidus non conservés, mais similaires sont affichés en gris. La numérotation des résidus est basée sur la séquence de la RNase 4 humaine. (B) Analyse phylogénétique des membres de la famille des RNases humaines. L'arbre a été construit avec l'application en ligne PhyML [177], et en utilisant les séquences en acides aminés des RNases. Le nombre (0,5) indique les substitutions nucléotidiques par site. (C) Représentation de la conservation évolutive des résidus de la RNase 4 humaine. L'échelle de conservation des résidus a été calculée avec l'application en ligne ConSurf [178], et ajoutée sur la structure tridimensionnelle de la RNase 4 humaine (PDB 1RNF). La position prédite du ligand UpA (uridylyl-3',5'-adenosine; cyan) a été déterminée par superposition de la structure PDB 1RPG sur la structure PDB 1RNF.

En utilisant l'expérience de  $^{15}\text{N}$ -CPMG, nous avons montré précédemment que deux membres de la famille des RNases, soit la RNase A bovine et ECP humaine, présentaient des mouvements atomiques à l'échelle de temps de la  $\mu\text{s}$ -ms [166]. Les secteurs dynamiques pouvaient être modulés par mutagenèse dirigée, ce qui affectait l'activité de la protéine, et *de facto*, laissait supposer que la flexibilité est importante pour l'activité de l'enzyme. Afin de valider notre hypothèse selon laquelle les mouvements atomiques à l'échelle de temps de la  $\mu\text{s}$ -ms seraient conservés chez les membres d'une même famille d'homologues structuraux de fonctions similaires, nous avons reconduit l'expérience en utilisant la RNase 4 humaine. Les résultats montrent que la RNase 4 conserve une flexibilité atomique globale, malgré des

mouvements locaux distinctes. La présence de ligands rigidifie l'enzyme de façon considérable, un résultat qui diverge de ce qui a été observé chez les autres RNases.

#### 4.6.1.2 Résultats

Même si sa structure a été identifiée depuis plus d'une décennie, peu d'études portent sur la caractérisation biophysique de la RNase 4 humaine. En conséquence, la première étape était d'identifier les résidus sur le spectre  $^1\text{H}$ - $^{15}\text{N}$ -HSQC. Pour ce faire, nous avons conduit une série d'expériences RMN, telle que décrite dans la section 1.5.6 du présent document. Les résultats sont présentés dans l'article (Figure 4.2): Gagné D. & Doucet N. (2015) "Sequence-specific backbone  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  resonance assignments of human ribonuclease 4" *Biomol NMR Assign*, 9(1):181-185 [179].

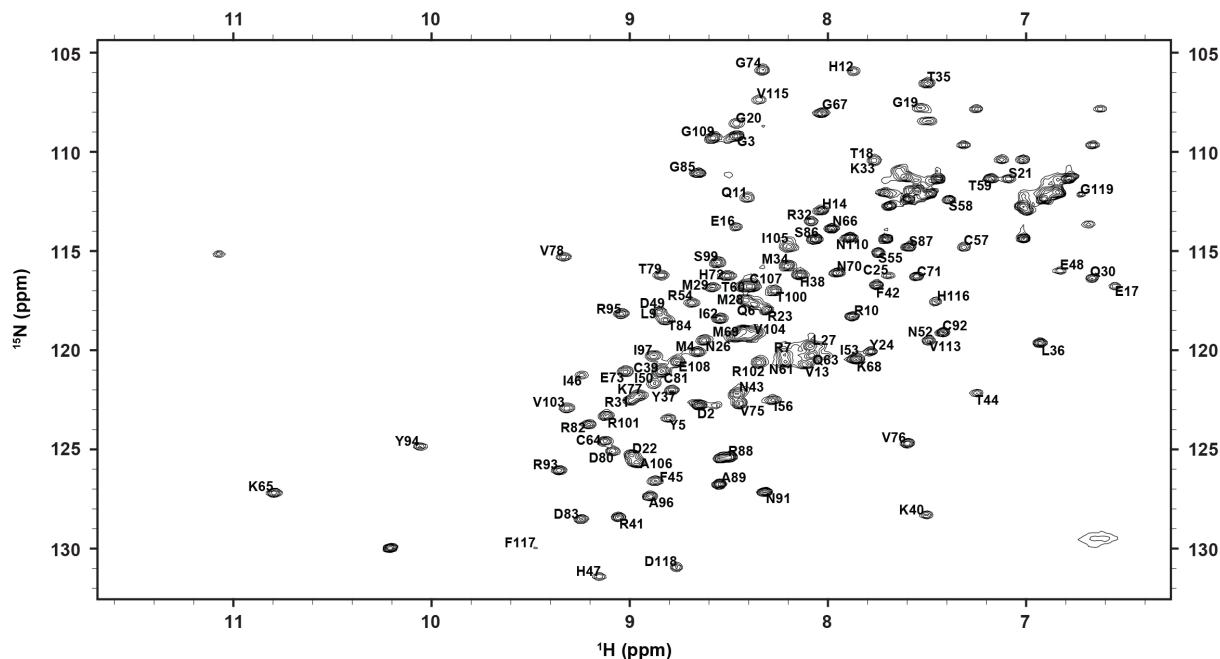
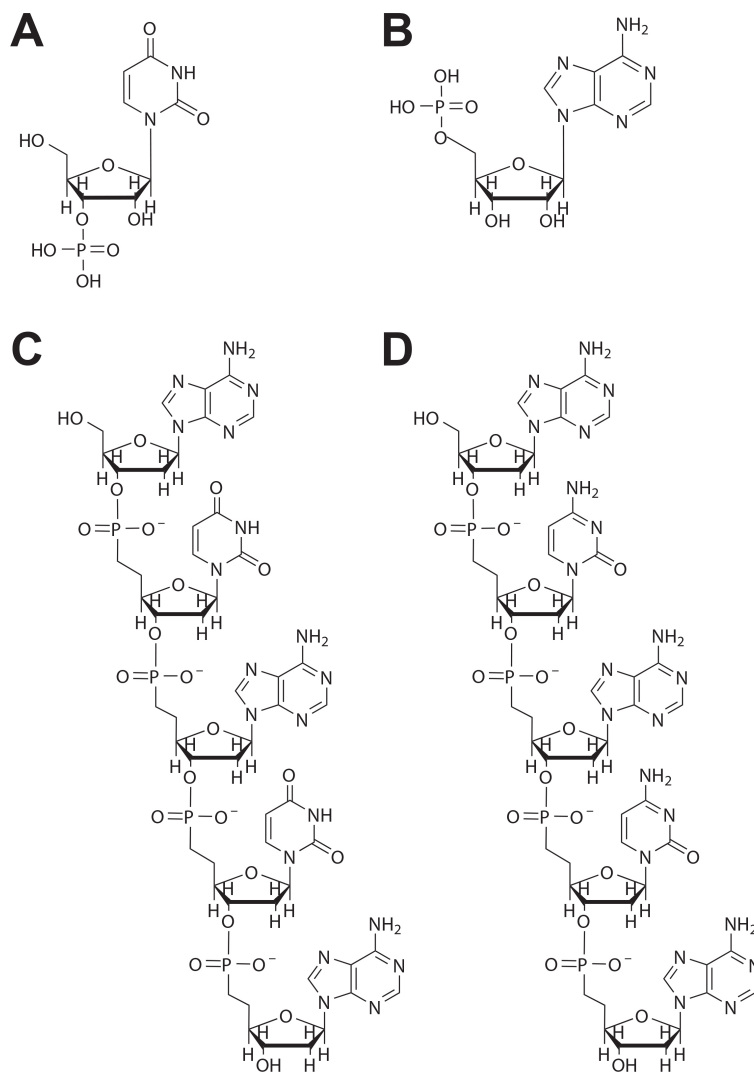


Figure 4.2. Spectre  $^1\text{H}$ - $^{15}\text{N}$ -HSQC à deux dimensions de la RNase 4 humaine. Tous les acides aminés ont été identifiés sur le pic correspondant par la lettre simple et le numéro du résidu dans la séquence. Le spectre a été produit avec Sparky 3.115.

Dans le but de caractériser les mouvements atomiques de la RNase 4 humaine, les expériences qui suivent ont été conduites sur l'enzyme libre et liée. Pour ce faire, nous avons utilisé les deux mêmes ligands que pour l'article décrit au chapitre 4, soit le 3'-UMP et le 5'-AMP. Ces nucléotides simples représentent les plus petits produits de la réaction, et se fixent respectivement aux sous-sites  $B_1$  et  $B_2$ , situés de part et d'autre du lien scissile (Figure 1.19A). Nous avons aussi ajouté deux oligonucléotides (ADN simple brin), qui s'apparentent au substrat (Figure 4.3).



**Figure 4.3. Représentation schématique d'analogues substrats et produits de la réaction d'une ribonucléase utilisés pour les expériences de spectroscopie RMN. Les plus petits produits de la réaction sont représentés par les ARN simples brins (A) 3'-UMP et (B) 5'-AMP. Chez la RNase A, les 3'-UMP et 5'-AMP se lient aux sous-sites B<sub>1</sub> et B<sub>2</sub> respectivement [73]. À l'opposé, les ADN simples brins (C) dATATA-3' et (D) dACACA-3' représentent des analogues du substrat. Les structures chimiques ont été produites avec ChemDraw.**

La liaison d'un ligand va induire un changement conformationnel. Afin de caractériser cette transition, nous avons conduit des expériences de titrage RMN chez la RNase 4 avec quatre ligands (Figure 4.4). Situés de part et d'autre de la cavité catalytique, les résidus situés sur les brins  $\beta$ 1 et  $\beta$ 4 montrent les plus fortes variations de déplacements chimiques ( $\Delta\delta$ ) (Figure 4.4). L'importance du rôle joué par ces deux brins  $\beta$  est appuyée par un taux de conservation élevée dans ces régions (Figure 4.1C). Toutefois, ce résultat diffère de ce qui avait été observé chez ECP, où les résidus situés sur le brin  $\beta$ 1 n'étaient pas ou peu affectés par la présence du ligand. On note aussi la présence de plusieurs résidus situés sur l'hélice  $\alpha$ 1, une région relativement bien conservée chez la RNase 4 humaine (Figure 4.1C). Tel qu'observé



avec ECP et la RNase A, la lysine située à l'extrémité de la boucle 4 montre un  $\Delta\delta$  significatif en présence de ligands. Il a été proposé que cette lysine interagissait avec le groupement phosphate du ligand en  $P_0$  [158]. Toutefois, chez les membres de la famille des RNases humaines, Lys66 est remplacée par des équivalents chez EDN (Ser64) et ECP (His64), alors que la RNase 5 ne possède aucun substitut. Malgré que l'histidine possède des caractéristiques chimiques considérablement différentes d'une lysine, un  $\Delta\delta$  significatif est également observé chez ECP (se référer à la figure 1.2 du présent article). Par ailleurs, des études ont démontré que la mutation de la Lys66 réduisait par 3 fois l'activité ribonucléolytique de la RNase A [180]. Ces résultats appuient fortement l'hypothèse que le résidu présent à cette position joue un rôle de premier plan dans la liaison du ligand.

Les résultats présentés dans l'article de la section 4.4 suggèrent une conservation de la flexibilité atomique globale à l'échelle de temps de la  $\mu$ s-ms chez les homologues structuraux partageant une fonction similaire. Toutefois, nous avons observé la présence de réseaux de résidus dynamiques spécifiques à chaque protéine. En utilisant la même approche, nous voulions vérifier si la RNase 4 humaine répondait de manière similaire. Pour ce faire, nous avons conduit une série d'expériences  $^{15}\text{N}$ -CPMG sur l'enzyme libre et liée (Figure 4.5). Dans sa forme libre, un total de 53 résidus (soit 46% du nombre total de résidus) ont montré des mouvements avec un  $\Delta R_2(1/\tau_{CP}) > 1.5 \text{ s}^{-1}$ . Bien que les mouvements atomiques soient répartis globalement, on notera la présence de deux secteurs avec des  $k_{ex}$  de  $1595 \pm 164 \text{ s}^{-1}$  et  $1528 \pm 218 \text{ s}^{-1}$ , et qui correspondent à la boucle 4 et à l'hélice  $\alpha 2$  (et autour) de la RNase 4, respectivement. Ce résultat peut suggérer que ces secteurs dynamiques agissent en corrélation pour une fonction donnée de la protéine. Avec seulement 12, 14 et 2 résidus montrant un  $\Delta R_2(1/\tau_{CP}) > 1.5 \text{ s}^{-1}$  en présence de 3'-UMP, 5'-AMP et dATATA respectivement, ces données montrent une diminution de la flexibilité de la structure de la protéine (Figure 4.5). De plus, on notera l'absence de secteurs de résidus dynamiques, tel qu'observé chez la forme libre. De la même manière, le  $k_{ex}$  de ces résidus dynamiques varient considérablement, ce qui peut laisser supposer que cette flexibilité atomiques: 1- n'est pas requise pour la fonction catalytique de l'enzyme; 2- permet à l'enzyme de s'adapter au substrat (à la séquence), la rendant ainsi plus catalytiquement plus efficace.

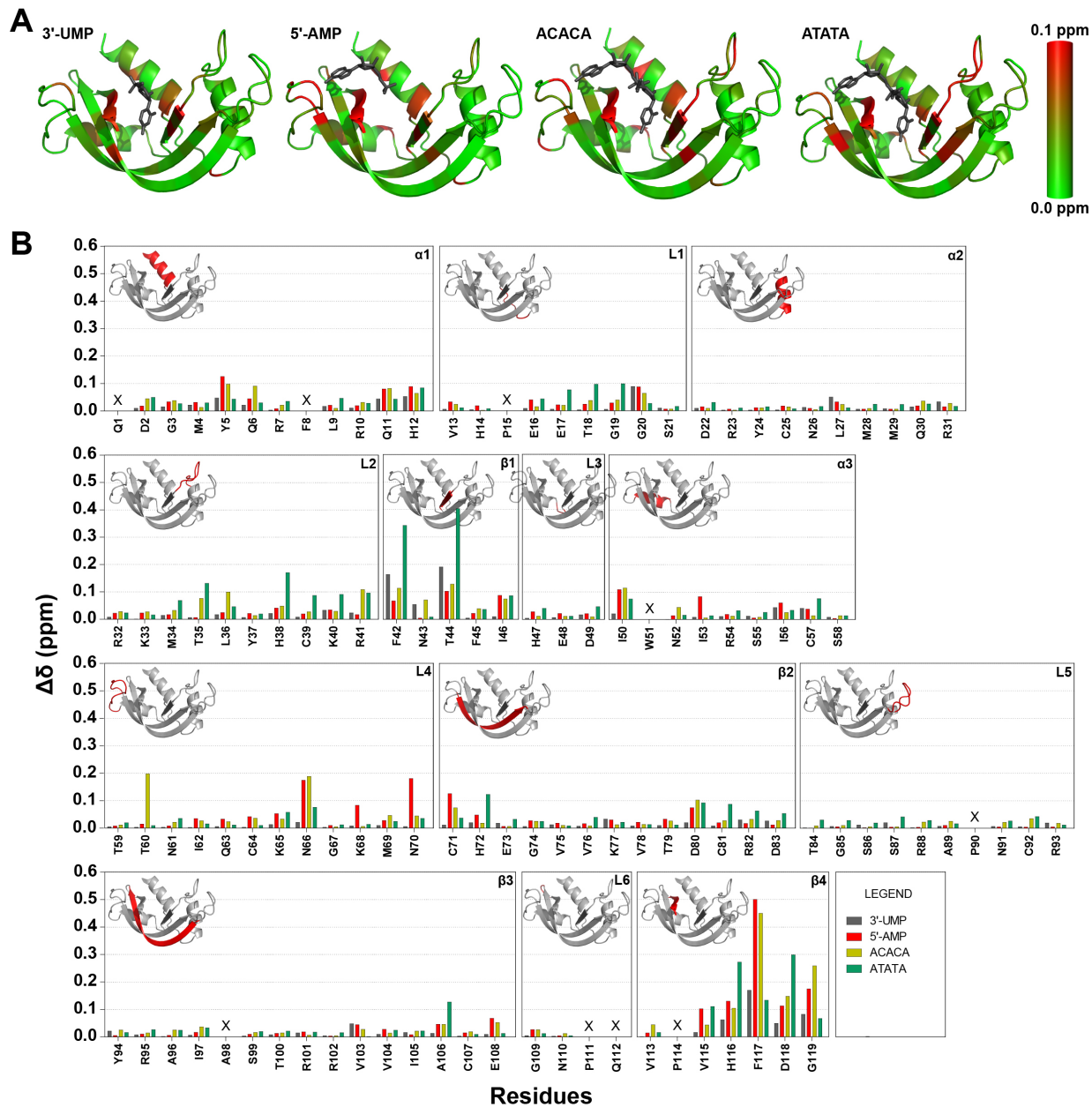


Figure 4.4. Variation des déplacements chimiques ( $\Delta\delta$ ) en  $^1\text{H}$  et  $^{15}\text{N}$  de la RNase 4 humaine en présence de ligands. (A) Le  $\Delta\delta$  entre 0.0 et 0.1 ppm est représenté sous forme d'un gradient de couleur sur la structure tridimensionnelle de la RNase 4 (PDB 1RNF). La position prédite des ligands (gris foncé) a été déterminée par superposition des structures: 3'-UMP (PDB 1O0N), 5'-AMP (PDB 1Z6S), dACACA (PDBs 1Z6S & 1RNN) et dATATA (PDBs 1Z6S & 1O0N). (B) Graphique montrant les  $\Delta\delta$  à 1:12 (ratio molaire) protéine:ligand par structure secondaire, tel que décrit par Terzyan et al. [174], pour chacun des résidus de la RNase 4 en présence de 3'-UMP (gris), de 5'-AMP (rouge), de dACACA (jaune) ou de dATATA (vert). Le  $\Delta\delta$  de chaque résidu a été calculé entre la protéine apo et sa forme liée, en utilisant l'équation  $\Delta\delta(\text{ppm}) = [(\Delta\delta_H^2 + \Delta\delta_N^2/25)/2]^{1/2}$  [181]. L'emplacement de chacune des structures secondaires est indiqué en rouge sur la structure tridimensionnelle de la RNase 4. Les images des structures tridimensionnelles ont été préparées avec PyMOL.

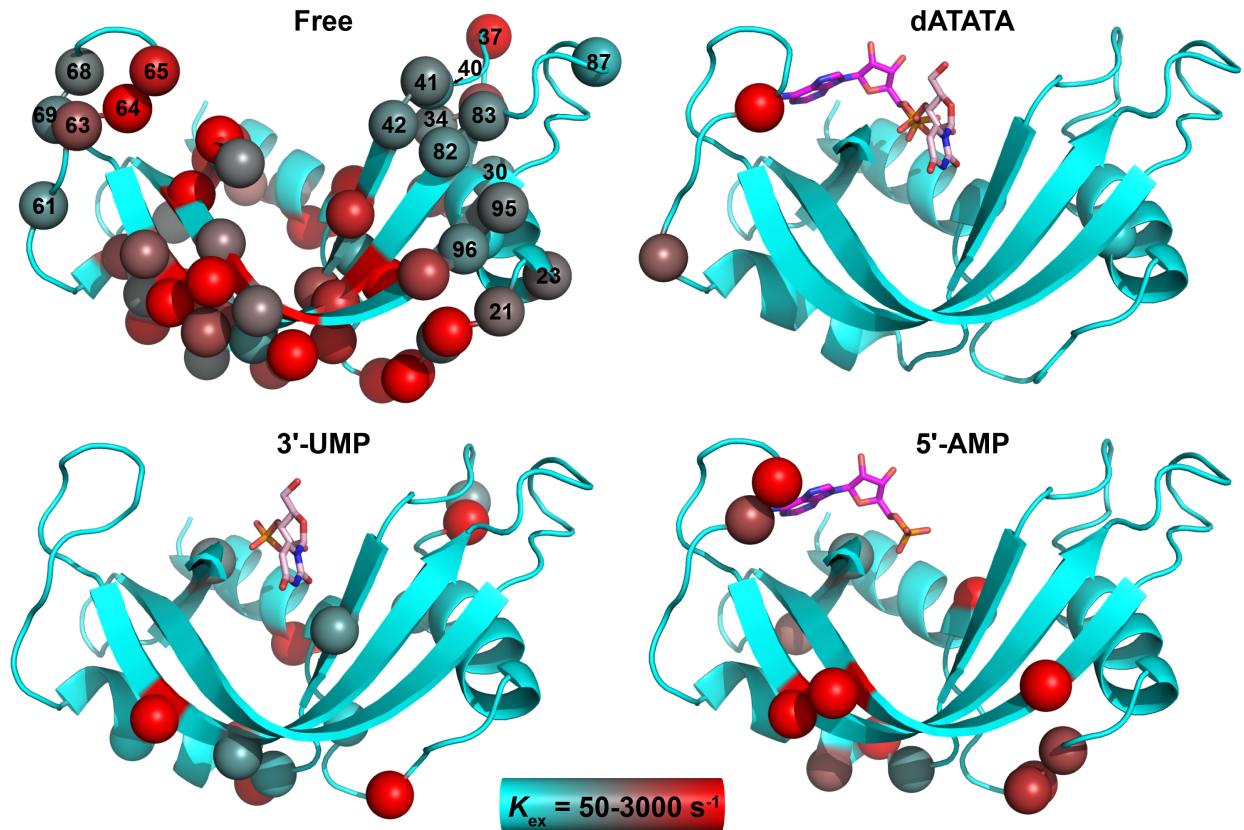


Figure 4.5: Échange conformationnel à l'échelle de temps de la  $\mu\text{s}$ -ms chez la RNase 4 humaine. Le  $C\alpha$  des résidus dynamiques avec un  $\Delta R_2(1/\tau_{CP}) = 1.5 \text{ s}^{-1}$  sont représentés par des sphères sur la structure tridimensionnelle de la RNase 4 (PDB 1RNF) dans sa forme libre ou liée. Les sphères sont colorées selon une échelle de couleur correspondant au taux d'échange conformationnel ( $k_{ex}$ ). Les chiffres indiquent le numéro du résidu dans la séquence en acides aminés. La position prédite des ligands 3'-UMP (PDB 1O0N) et 5'-AMP (PDB 1Z6S) a été déterminée par la superposition des structures sur celle de la RNase 4 (PDB 1RNF). Les structures tridimensionnelles ont été préparées avec PyMOL.

#### 4.6.1.3 Discussion

La RNase A (*Bos taurus*) a servi de modèle pour un bon nombre d'études portant sur différents champs d'applications [73]. Toutefois, les autres membres de la famille ont été jusqu'à maintenant peu caractérisés. Nous avons précédemment montré qu'ECP, un membre de la famille des RNases humaines, présentait des mouvements atomiques globaux à l'échelle de temps de la  $\mu\text{s}$ -ms, un résultat similaire à ce qui avait été observé chez la RNase A [160, 166]. Malgré cette similitude, le taux d'échange conformationnel variait considérablement, de manière cohérente avec l'efficacité catalytique. La présence de réseaux de résidus dynamiques spécifiques à chaque protéine laisse supposer que chaque enzyme possède un caractère unique. Dans le but d'étendre notre étude à d'autres membres de la famille, nous avons caractérisé la RNase 4 humaine.

Les résultats montrent que la RNase 4 est dynamique à l'échelle de temps de la  $\mu$ s-ms. Contrairement à ce qui avait été observé auparavant chez les RNase A et ECP, la flexibilité s'étend sur toute la structure. Malgré la présence accrue de résidus dynamiques, la stabilité de la protéine demeure similaire aux autres RNases [182]. La faible quantité de données cinétiques nous empêche malheureusement de déterminer si les mouvements atomiques pourraient jouer un rôle dans la catalyse enzymatique. Des  $k_{cat}/K_M$  de  $0.25 \times 10^{-6} \text{ s}^{-1}\text{M}^{-1}$  et  $1.7 \times 10^{-6} \text{ s}^{-1}\text{M}^{-1}$  ont été publiées pour les RNase 4 et RNase A, avec uridyle(3'→5')adenosine (UpA) comme substrat [183, 184]. En comparaison, des  $k_{ex}$  de  $1595 \pm 164 \text{ s}^{-1}$  et  $1803 \pm 71 \text{ s}^{-1}$  ont été mesurés dans la boucle 4 chez les RNase 4 et RNase A, respectivement. Ces résultats vont dans le même sens que ce qui avait été observé auparavant, c'est-à-dire qu'il semble y avoir une corrélation entre le taux d'échange conformationnel et l'efficacité catalytique. En effet, plus le taux d'échange conformationnel ( $k_{ex}$ ) est élevé, et plus l'enzyme est catalytiquement efficace ( $k_{cat}/K_M$ ).

Dans le but de mieux comprendre l'effet de la liaison de ligands sur la dynamique, nous avons répété l'expérience de  $^{15}\text{N}$ -CPMG sur l'enzyme dans sa forme liée. Les données montrent une rigidification de la structure à l'échelle de temps de la  $\mu$ s-ms, un résultat qui avait aussi été observé chez ECP [166]. L'absence de mouvements en présence de l'oligonucléotide dATATA corrèle avec sa faible affinité ( $K_D > 2 \text{ mM}$ ). Il a été suggéré que l'absence de mouvements dans la région correspondant à l'intersection entre les deux secteurs  $V_1$  et  $V_2$  pourrait être responsable de cette diminution d'affinité [145]. Malgré une réduction de la flexibilité, les résidus dynamiques sont distribués à travers toute la surface de la protéine, et présentent des taux d'échange conformationnel variable. Cette variabilité en fonction du ligand pourrait toutefois s'avérer un élément positif, en ce sens que cela pourrait signifier que l'enzyme est adaptable. Bien que les RNases semblent favoriser certaines séquences, il n'en demeure pas moins qu'elles sont capables de s'ajuster à toutes les séquences en ARN. Toutefois, l'efficacité catalytique, et fort probablement le  $k_{ex}$ , varieront en conséquence.

En résumé, les résultats suggèrent une conservation des mouvements atomiques à l'échelle de temps de la  $\mu$ s-ms entre les homologues structuraux partageant une fonction similaire, et ce, malgré une faible identité de séquence. Toutefois, des différences significatives sont observées entre les homologues structuraux et fonctionnels, entre autre dans le taux d'échange conformationnel et dans la distribution locale des résidus dynamiques. Ces écarts pourraient être associées aux variations dans leur affinité aux ligands et/ou efficacité catalytique. Afin d'apporter des clarifications, des études cinétiques devront être entreprises.

## 4.6.2 Comparaison biophysique d'EDN et d'ECP

### 4.6.2.1 Introduction

Dans cette étude, nous avons comparé les mouvements moléculaires de deux membres de la famille des RNases humaines, soit ECP et EDN (Figure 4.6). Avec une similarité de séquence en acides aminés de 67%, ces deux protéines ont récemment évolué et partagent plusieurs fonctions communes [97]. Malgré ces similarités, l'activité ribonucléolytique d'ECP est de 50-100 fois plus faible que celle de son homologue [119]. De plus, les deux protéines affichent une préférence marquée pour le poly(U) plutôt que le poly(C), elles sont inactives envers le poly(A) et l'ARN double brin, et elles montrent un pH catalytique optimal de 6.5-7.0 pour l'ARN de transfert de levure [112, 113]. On notera un élément distinctif, soit la présence d'une sérine (histidine) en position 64 chez EDN (ECP), plutôt qu'une lysine (Lys66 chez la RNaseA) (Figure 4.6C). Il a été proposé que la Lys66 interagirait avec le groupement phosphate en  $P_o$ . Précédemment, nos études ont d'ailleurs démontré que le résidu situé à cette position était significativement affecté par la présence d'un ligand, lui conférant par le fait même un rôle prédominant dans la liaison du substrat (se référer aux sections 4.4 et 4.6.1.2 du présent document).

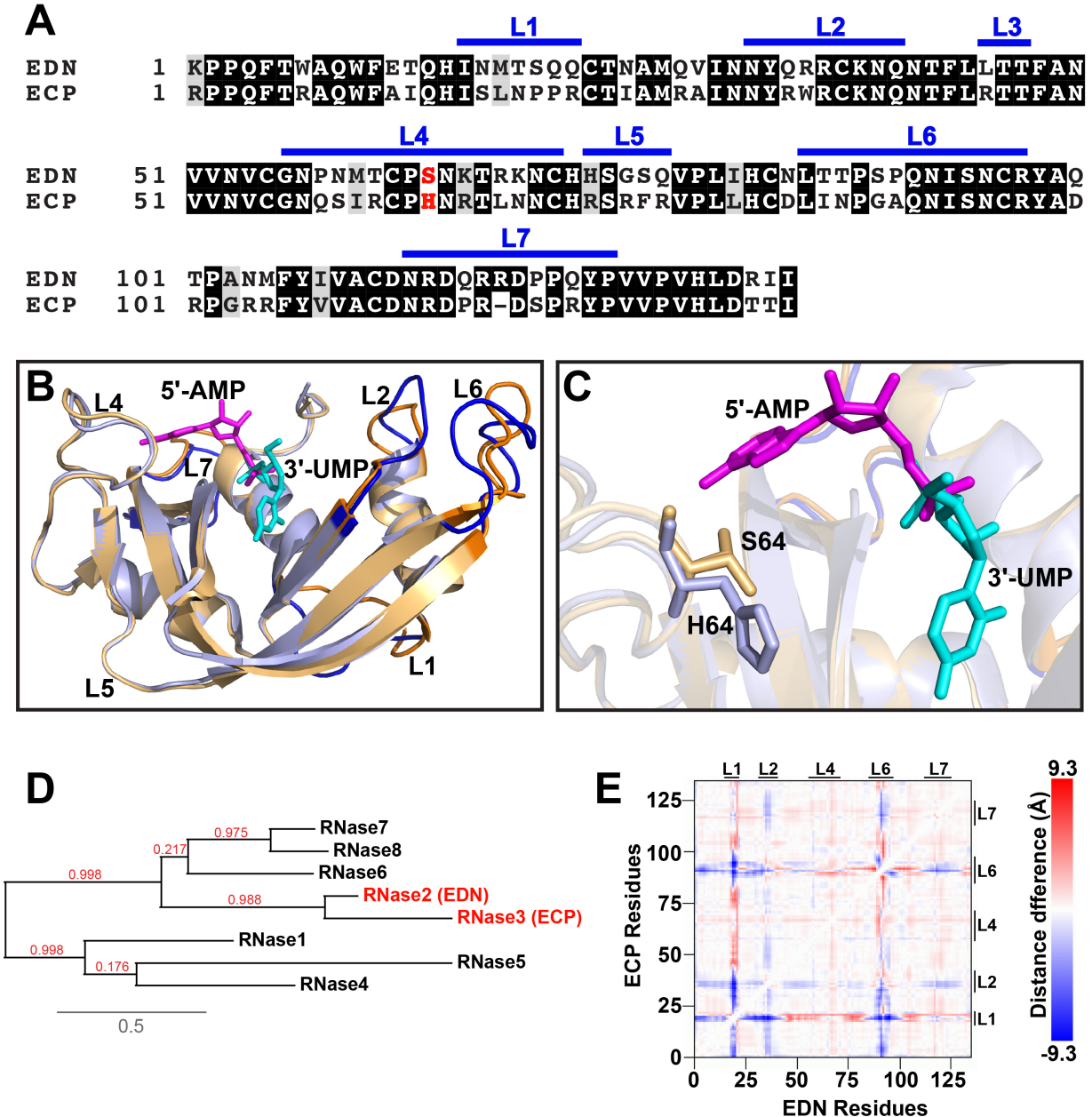


Figure 4.6: Homologie structurale entre EDN (Eosinophil-Derived Neurotoxin; RNase 2) et ECP (Eosinophil Cationic Protein; RNase 3). (A) Alignement de séquences d'EDN et d'ECP. L'alignement a été réalisé par ClustalW2 [185] et graphiquement édité avec BoxShade ([http://embnet.vital-it.ch/software/BOX\\_form.html](http://embnet.vital-it.ch/software/BOX_form.html)). Les zones noires (grises) représentent les résidus identiques (similaires). (B) Superposition des structures tridimensionnelles d'EDN (orange) et d'ECP (bleu). Les variations structurales les plus importantes sont visibles dans les boucles 1, 2, 6 et 7 (couleurs foncées). La position prédite des ligands 5'-AMP (cyan) et le 3'-UMP (magenta) a été ajoutée manuellement sur la structure par la superposition des PDB 1Z6S (5'-AMP) et 1O0N (3'-UMP). (C) Vue agrandie de la boucle 4 et de la position prédite des ligands dans le site actif. La Ser64 (His64) est représentée sous la forme de bâtonnet pour EDN (ECP). (D) Arbre phylogénétique des huit RNases canoniques chez l'humain. La reconstruction phylogénétique a été faite avec l'application en ligne PhyML, en utilisant la méthode des distances génétiques. Les chiffres représentent le niveau de confiance ("Bootstrap") des regroupements. (E) Matrice de distance entre les structures tridimensionnelles d'EDN (PDB 1GQV) et d'ECP (PDB 1QMT), construites avec l'application en ligne FATCAT. La matrice de distance est comprise entre 9.3 et -9.3 Å.

#### 4.6.2.2 Résultats

Bien qu'EDN ait fait l'objet de nombreuses études fonctionnelles, aucun groupe n'a encore étudié les mouvements atomiques de cet intrigant membre de la famille des RNases humaines. En conséquence, la première étape de cette étude était d'attribuer les résidus sur le spectre  $^1\text{H}$ - $^{15}\text{N}$ -HSQC de la protéine (Figure 4.7). Comme auparavant, nous avons conduit une série d'expériences RMN permettant d'attribuer les résidus sur le spectre (se référer à la section 1.5.6 du présent document). Un total de 99.2% des résidus (excluant les prolines) ont été attribués.

Afin de caractériser l'état de transition lors de la liaison de ligands, nous avons procédé à des expériences de titrage RMN. Pour ce faire, nous avons utilisé les mêmes nucléotides simples et oligonucléotides que précédemment (Figure 4.3). À saturation, les résultats chez EDN montrent une réponse importante dans la boucle 4, de même que pour l'hélice  $\alpha_1$  et le brin  $\beta_5$  (en C-terminal), alors que la variation du déplacement chimique ( $\Delta\delta$ ) est spécifique à chaque ligand pour les brins  $\beta_1$  et  $\beta_3$  (Figure 4.8). Malgré plusieurs similitudes, les  $\Delta\delta$  sont beaucoup moins prononcés chez ECP (Figure 4.9). Par ailleurs, on notera que le résidu en position 64 est affecté par la liaison des ligands aux sous-sites  $B_1$  et  $B_2$ , un résultat qui avait également été observé chez d'autres RNases. Il est aussi intéressant de constater que la direction du déplacement chimique (Figures 4.8B et 4.9B) varie selon le ligand, mais reste similaire entre les résidus Ser64 (EDN) et His64 (ECP) pour un même ligand.

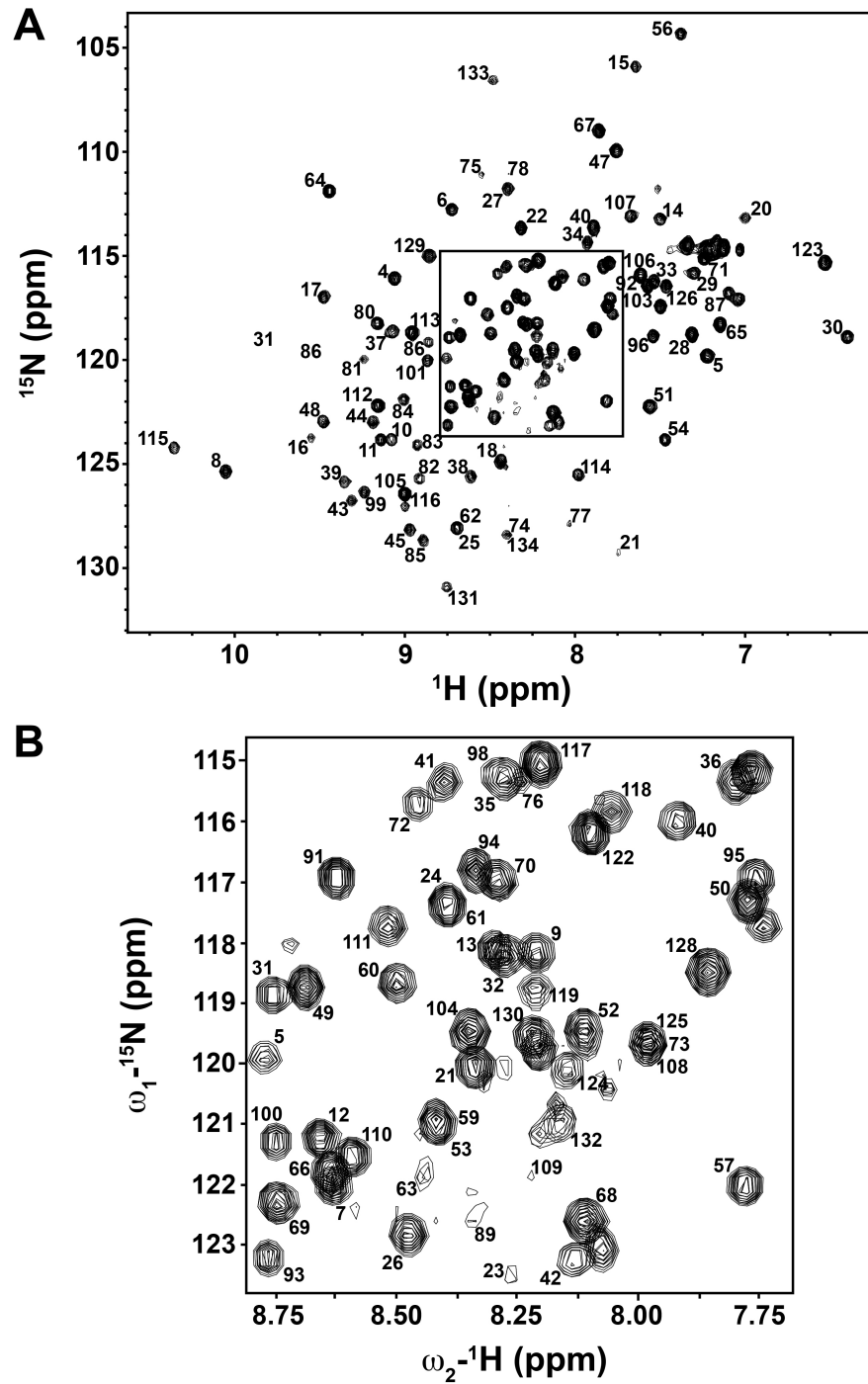


Figure 4.7. Spectre  $^1\text{H}$ - $^{15}\text{N}$ -HSQC à deux dimensions d'EDN. Tous les acides aminés ont été identifiés sur le pic correspondant par la lettre simple et le numéro du résidu dans la séquence. (A) Vue complète du spectre. Pour une meilleure visibilité, la région du centre a été agrandie (B). Le spectre a été produit avec NMRView 8.



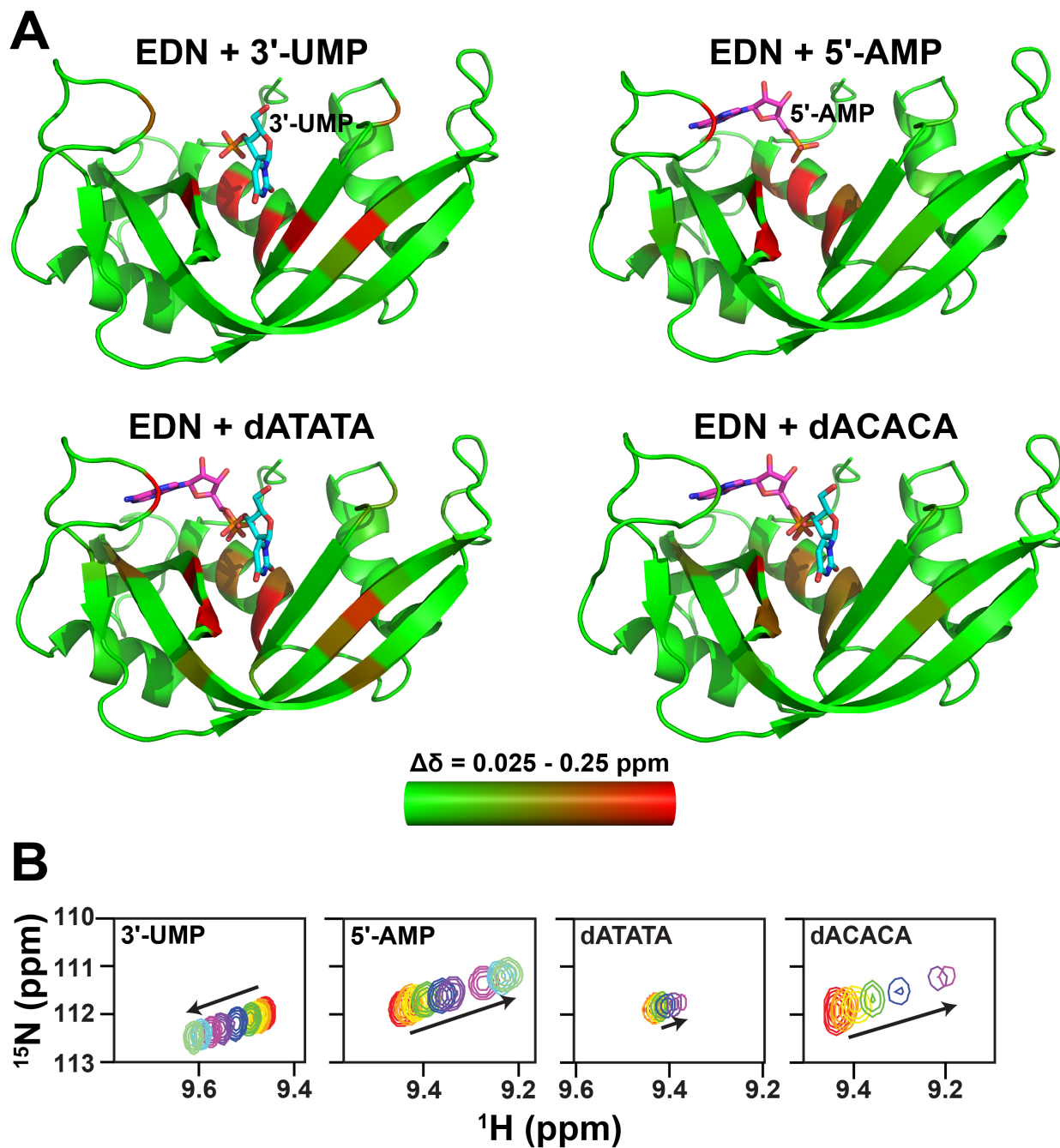


Figure 4.8: Titration RMN d'EDN en présence de ligands. (A) Variations des déplacements chimiques ( $\Delta\delta$ ) en  $^1\text{H}$ - $^{15}\text{N}$  induites par la liaison de ligands. Le  $\Delta\delta$  de chaque résidu a été calculé entre la protéine apo et sa forme liée, en utilisant l'équation  $\Delta\delta(\text{ppm}) = [(\Delta\delta_{\text{H}}^2 + \Delta\delta_{\text{N}}^2/25)/2]^{1/2}$  [181]. L'intensité du  $\Delta\delta$  est représentée par un gradient de couleur (du vert au rouge) pour chacun des résidus sur la structure tridimensionnelle d'EDN (PDB 1GQV). La position prédite des ligands 5'-AMP (magenta) et le 3'-UMP (cyan) a été ajoutée manuellement sur la structure par la superposition des PDB 1Z6S (5'-AMP) et 1O0N (3'-UMP). (B) Spectres en  $^1\text{H}$ - $^{15}\text{N}$  du  $\Delta\delta$  pour le résidu Ser64, situé sur la boucle 4. Les points représentent le ratio molaire protéine:ligand: 0 (rouge), 0,174 (orange), 0,393 (jaune), 0,691 (vert), 1,31 (bleu), 2,71 (violet), 6 (magenta), 12 (cyan) et 18 (vert clair). Les flèches indiquent la direction du déplacement chimique avec l'augmentation de la concentration des ligands.

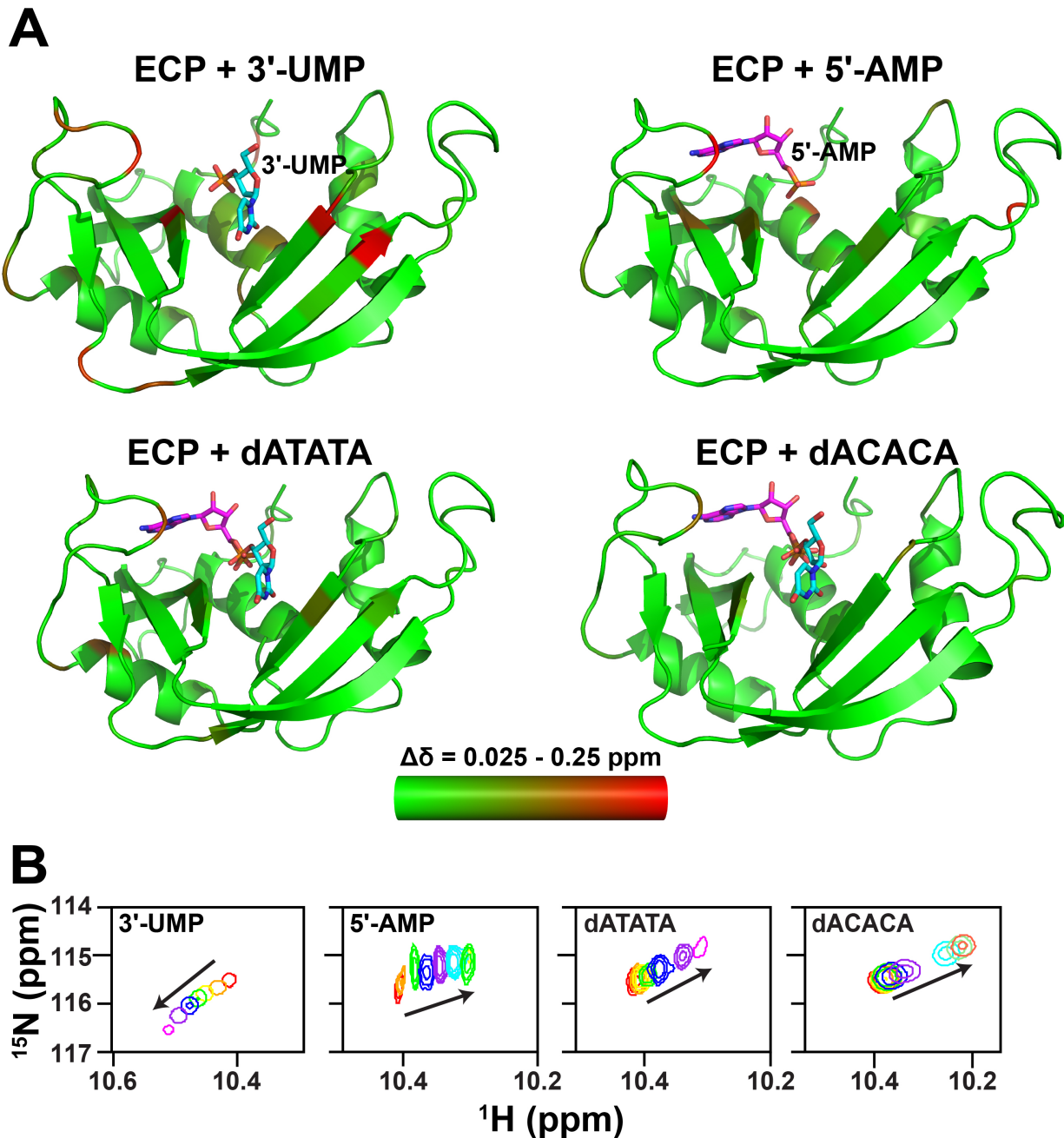


Figure 4.9: Titrage RMN d'ECP en présence de ligands. (A) Variations des déplacements chimiques ( $\Delta\delta$ ) en  $^1\text{H}$ - $^{15}\text{N}$  induites par la liaison de ligands. Le  $\Delta\delta$  de chaque résidu a été calculé entre la protéine apo et sa forme liée, en utilisant l'équation  $\Delta\delta(\text{ppm}) = [(\Delta\delta_{\text{H}}^2 + \Delta\delta_{\text{N}}^2/25)/2]^{1/2}$  [181]. L'intensité du  $\Delta\delta$  est représentée par un gradient de couleur (du vert au rouge) pour chacun des résidus sur la structure tridimensionnelle d'ECP (PDB 1QMT). La position prédite des ligands 5'-AMP (magenta) et le 3'-UMP (cyan) a été ajoutée manuellement sur la structure par la superposition des PDB 1Z6S (5'-AMP) et 1O0N (3'-UMP). (B) Spectres en  $^1\text{H}$ - $^{15}\text{N}$  du  $\Delta\delta$  pour le résidu His64, situé sur la boucle 4. Les points représentent le ratio molaire protéine:ligand: 0 (rouge), 0,174 (orange), 0,393 (jaune), 0,691 (vert), 1,31 (bleu), 2,71 (violet), 6 (magenta), 12 (cyan) et 18 (vert clair). Les flèches indiquent la direction du déplacement chimique avec l'augmentation de la concentration des ligands.

Jusqu'à présent, nous avons montré que la dynamique globale des membres de la famille des RNases humaine était conservée à l'échelle de temps de la  $\mu$ s-ms. Toutefois, la présence de réseaux de résidus dynamiques spécifiques à chaque protéine démontre certaines distinctions qui pourraient être liées à leur efficacité catalytique (qui varie de plusieurs ordres de grandeur entre les différentes RNases), ou encore, à cause de la présence d'activité biologique spécifique à certaines RNases. Il est important de noter que les enzymes analysées jusqu'à présent sont évolutivement très éloignées (Figure 1.21) et font partie de groupes différents, selon la subdivision des membres établies par Goo et Cho [98]. Dans cette étude, nous avons voulu savoir si des membres évolutivement rapprochés pourraient conserver non seulement la dynamique globale, tel qu'observé chez ECP et la RNase A [166], mais aussi les secteurs dynamiques. Par  $^{15}\text{N}$ -CPMG, EDN et ECP montrent une dynamique globale et locale relativement similaire, avec la présence de réseaux dynamiques visibles dans le secteur  $V_1$  (Figure 4.10). Un total de 10 (18) résidus montrent un  $R_2(1/\tau_{CP}) \geq 1.5 \text{ s}^{-1}$ . Le taux d'échange conformationnel ( $k_{ex}$ ) global calculé pour EDN (ECP) est de  $1087 \pm 174 \text{ s}^{-1}$  ( $534 \pm 80 \text{ s}^{-1}$ ). Ce résultat va dans le même sens que les études précédentes, qui démontraient qu'un  $k_{ex}$  plus élevé corrélait avec une efficacité catalytique supérieure. On notera toutefois l'absence de mouvements pour les résidus situés à l'extrémité de la boucle 4 chez EDN, malgré l'implication de la Ser64 lors de la liaison des ligands.

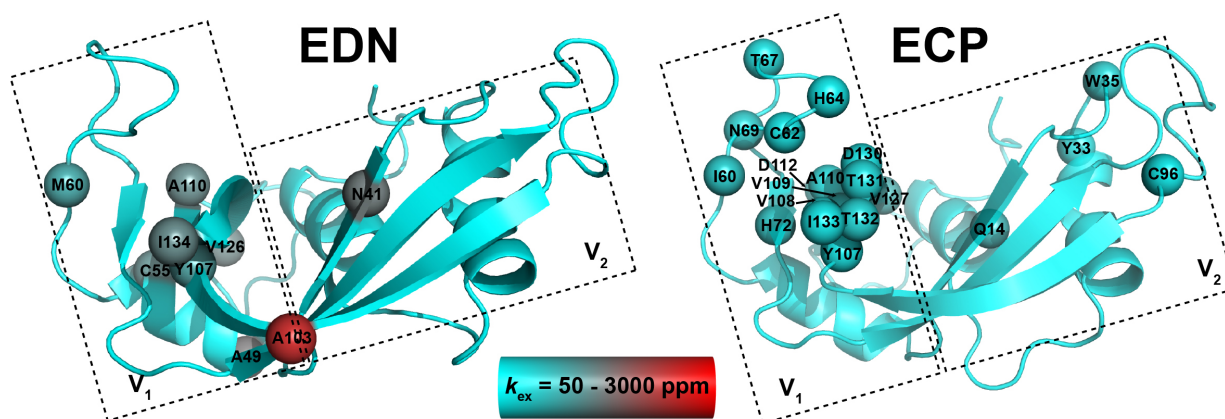


Figure 4.10: Comparaison de l'échange conformationnel entre EDN et ECP dans leur forme libre. Les résidus avec une courbe de dispersion de relaxation dont le  $\Delta R_2(1/\tau_{CP}) \geq 1.5 \text{ s}^{-1}$  sont représentés par des sphères sur la structure tridimensionnelle (EDN: PDB 1GQV; ECP: PDB 1QMT). Les sphères sont colorées selon une échelle de couleur correspondant au taux d'échange conformationnel ( $k_{ex}$ ). Les protéines sont subdivisées en deux secteurs, soit  $V_1$  et  $V_2$ , tel que décrit dans l'article [78]. Les structures tridimensionnelles ont été produites avec PyMOL.

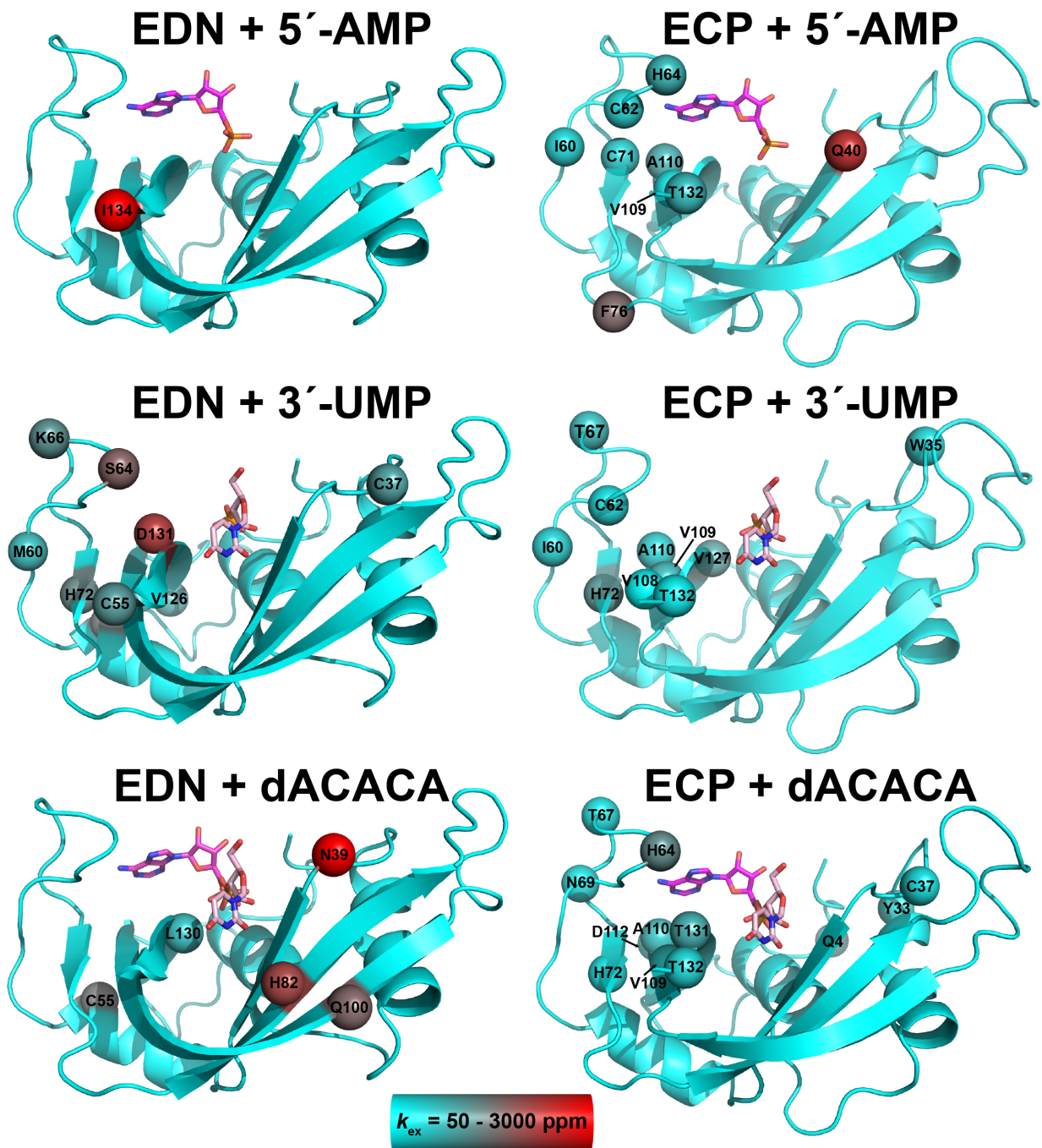


Figure 4.11: Comparaison de l'échange conformationnel entre EDN et ECP en présence des produits de la réaction 5'-AMP et 3'-UMP, et de l'oligonucléotide dACACA. Les résidus avec une courbe de dispersion de relaxation dont le  $\Delta R_2(1/\tau_{CP}) \geq 1.5 s^{-1}$  sont représentés par des sphères sur la structure tridimensionnelle (EDN: PDB 1GQV; ECP: PDB 1QMT). La position prédite des ligands 5'-AMP (magenta) et le 3'-UMP (cyan) a été ajoutée manuellement sur la structure par la superposition des PDB 1Z6S (5'-AMP, magenta) et 1O0N (3'-UMP, rose). Les sphères sont colorées selon une échelle de couleur correspondant au taux d'échange conformationnel ( $k_{ex}$ ). Les structures tridimensionnelles ont été produites avec PyMOL.

Alternativement, un total de 1 (9), 8 (10) et 5 (12) résidus montrent un  $R_2(1/\tau_{CP}) \geq 1.5 s^{-1}$  chez EDN (ECP) en présence des ligands 5'-AMP, 3'-UMP et dACACA, respectivement

(Figure 4.11). De façon générale, on remarquera que le profil des secteurs de résidus dynamiques chez EDN varient considérablement entre les ligands. En contrepartie, les secteurs dynamiques demeurent sensiblement identiques chez ECP, avec des réseaux de résidus dynamiques dans le secteur  $V_1$ , incluant la boucle 4 et les résidus des brins  $\beta$ . Il en est de même pour le taux d'échange conformationnel ( $k_{ex}$ ) des résidus dynamiques chez EDN, qui varie énormément d'un ligand à l'autre, comparativement à ce qui est observé chez ECP. Il est à noter que la liaison de l'oligonucléotide dATATA conduit à la précipitation des deux protéines.

#### 4.6.2.3 Discussion

Les récentes données expérimentales indiquent que plusieurs mouvements moléculaires concertés qui se produisent sur l'échelle de temps de la réaction catalytique jouent un rôle important dans la promotion de la catalyse dans de nombreux systèmes enzymatiques [14, 163, 168, 186]. Toutefois, il est encore difficile de lier ces mouvements à l'évènement catalytique. *A priori*, il est permis de penser que si ces mouvements atomiques sont essentielles à l'activité de la protéine, elles seront conservées chez des homologues structuraux ayant la même fonction, malgré une faible similarité dans la séquence en acides aminés. Jusqu'à présent, nous avons montré une conservation de la flexibilité à l'échelle de temps de la  $\mu$ -ms chez la RNase A, ECP et la RNase 4. Malgré cette conservation de la flexibilité globale, nous avons observé des réseaux de résidus dynamiques distincts à chaque membre. De plus, les protéines possèdent des comportements particuliers en présence de ligands, démontrant le caractère unique associé à chacune des enzymes.

Pour cette étude, nous voulions caractériser les mouvements atomiques de deux membres qui ont récemment évolutivement divergé. Nous nous attendions à ce que les deux membres présentent une dynamique locale relativement similaire. Comme prévu, nous démontrons qu'ECP et EDN, en plus de conserver les mouvements globaux, affichent un profil dynamique similaire dans leur forme libre. En effet, même si les deux protéines montrent une efficacité catalytique très différente, les résidus présentant des mouvements atomiques significatifs sont confinés dans deux régions du secteur  $V_1$  (se référer à l'article présenté au Chapitre 2 du présent document). La présence de telles régions avait d'ailleurs fait l'objet d'études par le groupe de Ranganathan dans d'autres systèmes enzymatiques [172]. En effet, en utilisant une analyse statistique de corrélation de l'évolution entre les acides aminés pour les membres d'une même famille d'enzymes de fonction similaire, il a observé une évolution corrélée de certains groupes de résidus. Ces "secteurs" pourraient être associés à des fonctions distinctes dans la protéine.



En présence de l'oligonucléotide dACACA, qui mime la liaison du substrat, on observe un changement dans la répartition des résidus. Malgré la réduction du nombre de résidus dynamiques, les résultats suggèrent une réorganisation conformationnelle favorisant la liaison du ligand. Toutefois, l'absence de résidus dynamiques dans la région située entre les deux secteurs pourrait défavoriser la liaison du substrat, un résultat qui avait été également observé chez la RNase 4. Afin d'étudier le relâchement du produit, nous avons lié les enzymes avec deux oligonucléotides simples, soit 5'-AMP et 3'-UMP. Bien qu'un réarrangement des résidus soit visible chez EDN en présence du 3'-UMP, le changement le plus important se produit suite à la liaison du 5'-AMP (Figure 4.11). Alors qu'ECP conserve les réseaux locaux de résidus dynamiques, EDN devient complètement rigide. Cette réorganisation des réseaux de résidus dynamiques en présence de différents ligands peut suggérer qu'EDN est plus "adaptable" qu'ECP, ce qui la rend catalytiquement plus efficace.

En bref, nous avons démontré une conservation des mouvements atomiques à l'échelle de temps de la  $\mu$ s-ms chez deux membres provenant de gènes qui ont récemment divergé, soit EDN et ECP. Contrairement à ce qui avait été observé chez la RNase A et la RNase 4, la conservation se situe non seulement à l'échelle globale, mais aussi, par la présence de réseaux de résidus dynamiques dans le secteur  $V_1$ .

## 5 ARTICLE NO. 4: RÉSEAUX DE CORRÉLATIONS À LONGUE DISTANCE DE DÉPLACEMENTS CHIMIQUES CONCERTÉS SUITE À LA LIAISON DE LIGANDS À L'ANGIOGÉNINE HUMAINE

Titre: "Network of long-range concerted chemical shift displacements upon ligand binding to human angiogenin"

Auteurs: **Gagné, D.**, Narayanan, C. & Doucet N.

### 5.1 Présentation de l'article scientifique

Les modèles de sélection conformationnelle et d'ajustement induit dépendent de la flexibilité de la protéine et des réarrangements structuraux afin de permettre la liaison du substrat. Par spectroscopie RMN, il est possible d'étudier ces changements conformationnels. Dans cet article, nous avons utilisé les techniques de titrage RMN et de  $^{15}\text{N}$ -CPMG et les avons appliquées à l'angiogénine. Cette protéine de 123 acides aminés, aussi appelée RNase 5, fait partie de la grande famille des RNases humaines. Bien qu'elle ne partage que 32% d'homologie de séquence avec la RNase A, l'angiogénine affiche une activité catalytique ( $k_{cat}/K_M$ )  $10^5$  fois inférieure à celle de son homologue structural. En plus de digérer l'ARN, l'angiogénine a acquis une fonction dans la régénération des vaisseaux sanguins. Alors que le  $^{15}\text{N}$ -CPMG a montré que les formes libre et liée de la protéine étaient rigides à l'échelle de temps de la  $\mu\text{s}$ -ms, nous avons observé la présence de déplacements chimiques coordonnés pour certains groupes de résidus. En utilisant la méthode d'analyse CHESPA, nous avons été en mesure d'identifier deux groupes de résidus: le premier groupe correspondant aux résidus 40-42, et le second groupe correspondant au domaine  $V_1$  de l'angiogénine.

Le manuscrit a été accepté le 25 novembre 2014 dans la revue "Protein Science" (facteur d'impact (2013) de 2.861), et une première version de l'article a été publiée en ligne le 26 décembre 2014. L'article a été publié dans le numéro 4 du volume 24 de l'édition d'avril 2015, aux pages 525-533, sous le numéro d'accèsion doi: 10.1002/pro.2613. À ce jour, l'article

a été cité 1 fois. La version formatée par l'éditeur sous format PDF et publiée en ligne est présentée dans le présent document.

## 5.2 Contribution des auteurs

Le travail expérimental a été réalisé par l'auteur principal (Donald Gagné). L'analyse et l'interprétation des résultats de  $^{15}\text{N}$ -CPMG et titrage RMN ont été faites par l'auteur principal (Donald Gagné). Chitra Narayanan a fait l'analyse CHESPA et participé à l'interprétation des résultats associés à l'analyse. Le texte et les figures ont été écrits, préparés et édités par l'auteur principal (Donald Gagné), sous la supervision de Pr Nicolas Doucet. Pr Nicolas Doucet a soumis le manuscrit à l'éditeur.

## 5.3 Résumé

Kosland a été le premier à soumettre l'hypothèse qu'une protéine pouvait subir un changement conformationnel d'ajustement induit pour lier le substrat [3]. Plus récemment, il a été proposé que plusieurs conformères de la protéine dans sa forme libre seraient présents en solution. Selon un mécanisme de sélection conformationnelle, le substrat sélectionnerait alors la conformation la plus favorable énergétiquement. Avec le développement de techniques en spectroscopie RMN et en bio-informatique, de nouvelles avancées sur notre compréhension du mécanisme de reconnaissance du substrat ont pu être faites. En effet, ces deux techniques permettent d'identifier les conformères faiblement peuplés et de haute énergie, en absence ou en présence de substrats.

Au cours d'études précédentes (voir les Chapitres 2-4 de la présente thèse), nous avons démontré la présence de conformères en solution chez plusieurs membres de la famille des RNases. La présence de mouvements dans les formes libre et liée pouvait suggérer que les protéines utilisent les mécanismes de sélection conformationnelle et d'ajustement induit. Afin de mieux comprendre les mécanismes qui régulent la catalyse enzymatique chez les membres de la famille des RNases, nous avons poursuivi notre analyse avec l'angiogénine. L'angiogénine, ou RNase 5, fait partie des huit RNases humaines ayant la fonction commune de digérer l'ARN. Cette enzyme de 123 acides aminés, principalement connue pour sa fonction angiogénique, conserve les principaux constituants des membres de la famille, soit une triade de résidus directement impliqués dans la catalyse enzymatique, et une structure en forme de V. Contrairement à ses homologues, elle ne possède que 3 ponts disulfures, plutôt que les 4 ponts



habituels. Avec 32% d'homologie de séquence avec la RNase A, son activité catalytique s'avère l'une des plus faibles dans la famille des RNases humaines, avec une efficacité ( $k_{cat}/K_M$ ) inférieure à celle de la RNase A d'un facteur  $10^5$  [187].

Par spectroscopie RMN, nous avons effectué des expériences de  $^{15}\text{N}$ -CPMG et de titrage, combinées à une analyse de projection de type CHESPA, pour étudier le mécanisme de reconnaissance des nucléotides chez l'angiogénine. Pour ce faire, nous avons utilisé deux produits finaux de la réaction, soit 3'-UMP et 5'-AMP. Ces deux ligands se fixent à des sous-sites spécifiques, identifiés  $B_1$  (3'-UMP) et  $B_2$  (5'-AMP). Nous démontrons que la protéine, bien que rigide à l'échelle de temps de la  $\mu\text{s}$ -ms, subit un changement conformationnel suite à la liaison des ligands. Ce réarrangement structural induit une réponse collective chez plusieurs résidus. Cette réponse coordonnée de secteurs de résidus confirme la présence de réseaux d'interaction liés à la fonction de la protéine.



# Network of long-range concerted chemical shift displacements upon ligand binding to human angiogenin

Donald Gagné,<sup>1</sup> Chitra Narayanan,<sup>1</sup> and Nicolas Doucet<sup>1,2,3\*</sup>

<sup>1</sup>INRS-Institut Armand-Frappier, Université du Québec, 531 Boulevard des Prairies, Laval, Québec H7V 1B7, Canada

<sup>2</sup>PROTEO, the Québec Network for Research on Protein Function, Structure, and Engineering, 1045 Avenue de la Médecine, Université Laval, Québec, Québec G1V 0A6, Canada

<sup>3</sup>GRASP, the Groupe de Recherche Axé sur la Structure des Protéines, 3649 Promenade Sir William Osler, McGill University, Montréal, Québec H3G 0B1, Canada

Received 26 September 2014; Accepted 25 November 2014

DOI: 10.1002/pro.2613

Published online 00 Month 2014 proteinscience.org

**Abstract:** Molecular recognition models of both induced fit and conformational selection rely on coupled networks of flexible residues and/or structural rearrangements to promote protein function. While the atomic details of these motional events still remain elusive, members of the pancreatic ribonuclease superfamily were previously shown to depend on subtle conformational heterogeneity for optimal catalytic function. Human angiogenin, a structural homologue of bovine pancreatic RNase A, induces blood vessel formation and relies on a weak yet functionally mandatory ribonucleolytic activity to promote neovascularization. Here, we use the NMR chemical shift projection analysis (CHESPA) to clarify the mechanism of ligand binding in human angiogenin, further providing information on long-range intramolecular residue networks potentially involved in the function of this enzyme. We identify two main clusters of residue networks displaying correlated linear chemical shift trajectories upon binding of substrate fragments to the purine- and pyrimidine-specific subsites of the catalytic cleft. A large correlated residue network clusters in the region corresponding to the V<sub>1</sub> domain, a site generally associated with the angiogenic response and structural stability of the enzyme. Another correlated network (residues 40–42) negatively affects the catalytic activity but also increases the angiogenic activity. <sup>15</sup>N-CPMG relaxation dispersion experiments could not reveal the existence of millisecond timescale conformational exchange in this enzyme, a lack of flexibility supported by the very low-binding affinities and catalytic activity of angiogenin. Altogether, the current report potentially highlights the existence of long-range dynamic reorganization of the structure upon distinct subsite binding events in human angiogenin.

**Keywords:** allostery; Carr-Purcell-Meiboom-Gill; chemical shift projection analysis; NMR; protein dynamics; relaxation dispersion; ribonuclease

*Abbreviations:* 5'-AMP, adenosine 5'-monophosphate; CHESPA, NMR chemical shift projection analysis; <sup>15</sup>N-CPMG, <sup>15</sup>N-Carr-Purcell-Meiboom-Gill relaxation dispersion experiments; 3'-UMP, uracil 3'-monophosphate; ECP, eosinophil cationic protein; DHFR, dihydrofolate reductase.; RNase A, bovine pancreatic ribonuclease A; HSQC, heteronuclear single quantum coherence spectroscopy.

\*Correspondence to: Nicolas Doucet, INRS-Institut Armand-Frappier, Université du Québec, 531 Boulevard des Prairies, Laval, Québec H7V 1B7, Canada. E-mail: nicolas.doucet@iaf.inrs.ca

## Introduction

The hypothesis that proteins undergo conformational changes to perform their function was first described in the “induced fit” model of molecular recognition proposed by Koshland in 1958.<sup>1</sup> In this model, slight reorganization of the protein backbone and side chains are required to promote binding, suggesting that proteins adopt their functional state only after interacting with their binding partner(s).<sup>2</sup> Recent advances in biocomputational and NMR methodologies are providing new insights into conformational diversity, challenging the

assumption that molecular recognition automatically proceeds through an induced fit mechanism. Computational and NMR methods provide information on weakly populated, higher energy conformers sampled by the free protein in the absence of ligand, revealing an ensemble of conformations that pre-exist in solution. The “conformational selection” model of molecular recognition postulates that binding selectivity will only occur with the most favorable ensemble substate(s), shifting the population toward the ligand-bound conformation.<sup>2,3</sup> While structural differences between the free and bound forms of a protein do not provide information on whether it follows an induced fit or conformational selection pathway, both models have been validated for numerous protein systems, including (but not limited to) bovine pancreatic ribonuclease (RNase A),<sup>4</sup> dihydrofolate reductase (DHFR),<sup>5,6</sup> and adenylate cyclase.<sup>7,8</sup>

Recently, we showed that RNase A and eosinophil cationic protein (ECP, or human RNase 3)—two structural homologues sharing a similar ribonucleolytic function—experience analogous conformational fluctuations in their free and bound states.<sup>9</sup> Our results demonstrated that both enzymes pre-exist as multiple, yet distinct conformational substates in the presence and absence of ligand. The occurrence of such conformational ensembles may suggest that both RNase A and ECP rely on conformational selection and/or induced fit mechanisms to promote their ribonucleolytic and/or other biological function. To improve our understanding of the molecular recognition mechanisms governing the function of pancreatic-like ribonucleases, we further extend our analysis to human angiogenin (or RNase 5), a 123-residue structural homologue first isolated in 1985<sup>10</sup> sharing 32% and 30% sequence identity with RNase A and ECP, respectively. The three aforementioned RNase structural homologues catalyze the same ribonucleolytic reaction, but the activity of human angiogenin is significantly weaker than that of its counterparts, with a catalytic efficiency ( $k_{cat}/K_m$ ) 10<sup>5</sup>-fold lower than that of RNase A.<sup>11</sup> Interestingly, angiogenin is the only pancreatic-like RNase to promote neovascularization, a biological activity that requires a functional RNA-degrading activity catalyzed by the enzyme.<sup>12,13</sup>

The overarching goal of our approach is to investigate whether enzymes sharing similar folds and function also rely on analogous dynamic behavior or sample parallel conformational changes to execute their function, effectively supporting the hypothesis that flexibility is structurally and evolutionarily conserved to preserve and/or promote function within protein families.<sup>14</sup> In this study, we use the NMR chemical shift projection analysis approach (CHESPA)<sup>15</sup> to characterize the global effects of molecular recognition in human angiogenin upon binding of fragment-based single-nucleotide analogues of a simple dinucleotide (UpA) RNA hydrolysis reaction:

3'-UMP and 5'-AMP. Both analogues selectively bind to two different subsites of the active site (the purine B<sub>1</sub> and pyrimidine B<sub>2</sub> subsites, respectively<sup>14</sup>), effectively reporting on two distinct binding events. Since the CHESPA approach is ideally suited for sensing changes between the apo versus bound equilibrium in the fast-exchange regime,<sup>15</sup> analyzing the binding mode of two distinct subsite fragments may allow us to gain some insight into the functional relevance of long-range residue networks affected by and/or implicated in ligand-binding events. This particular approach allows us to deconvolute the individual effects of B<sub>1</sub> and B<sub>2</sub> subsite binding, information normally hidden from classical substrate analogue-binding analyses. In other words, uncovering which chemical shift resonances are correlated upon binding of each individual “substrate fragment” (the product analogues), may provide information on clustering residues that are functionally important to stabilize biologically relevant RNA ligands and/or to uncover structural rearrangements and dynamic/allosteric sectors linked to protein function.<sup>16,17</sup> Our results show that some residue networks display high levels of chemical shift correlations in human angiogenin, suggesting the existence of long-range concerted residue rearrangements occurring during RNA ligand binding.

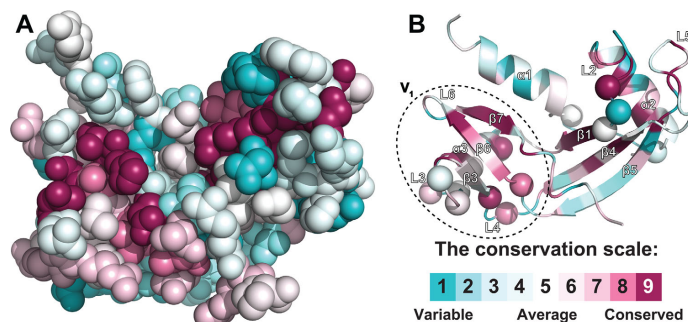
## Results

### Evolutionary conservation analysis

Sequence conservation profiles provide valuable information on the structural and functional importance of conserved positions in proteins. We performed a ConSurf<sup>18</sup> analysis to calculate the evolutionary conservation of angiogenin among vertebrate species (Fig. 1). We observed two groups of highly conserved residues, corresponding to the  $\alpha 3/\beta 3/\beta 6/\beta 7/L3$  and  $\alpha 3/\beta 1/\beta 4/L2/L5$  regions, which form the V-shape of the active-site cleft. Many of these residues are directly involved in either binding of the substrate or the catalytic reaction. A comparable conservation profile was detected for RNase 4, a structural homologue with similar ribonucleolytic function that shares 41% amino acid sequence identity with angiogenin.<sup>19</sup> However, there are marked differences between the two homologues, especially in the  $\alpha 1$  and  $\alpha 2$  helices, which contain some of the essential components of the binding cleft and residues of the folding initiation site, respectively.<sup>20</sup> While sections of  $\alpha 1$  are very highly conserved in RNase 4,<sup>19</sup> this helix is significantly variable among vertebrate angiogenin members, suggesting functional differences in substrate binding and/or distinct catalytic relevance between these structural RNase homologues.

### Comparison of <sup>1</sup>H-<sup>15</sup>N chemical shifts in the presence of ligands

Binding of a ligand alters the chemical environment of the target protein. This change in the chemical



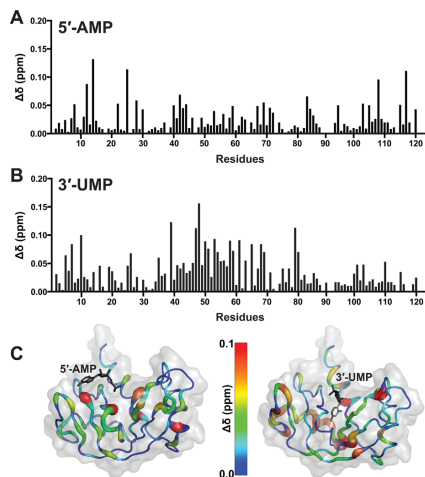
**Figure 1.** Evolutionary conservation of amino acid residues in human angiogenin. (A) Spheres and (B) cartoon representation of the sequence conservation profile. Color-coded scores were determined using the ConSurf server.<sup>18</sup> The conservation score is defined as the evolutionary rate of the site relative to the average evolutionary rate across all sites.<sup>56</sup> Site-specific rates are unitless. Three-dimensional structures were prepared with PyMOL using the Protein Data Bank (PDB) identification number 1ANG. For clarity, the protein is shown in the same orientation in panels A and B. Secondary structure identification was based on the published crystal structure of reference.<sup>57</sup> In panel B, residues showing CHESPA-coordinated chemical shift displacements ( $\cos(\theta) \approx 1$ ) are shown as spheres (refer to text and Fig. 4).

environment induces changes in the  $^1\text{H}$ - $^{15}\text{N}$ -resonance frequencies of the affected residues, detected as a shift in the position of the peaks in a  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum. Analysis of chemical shift changes induced by ligand binding can provide insights into the conformational changes associated with ligand binding. Here, we performed titrations of angiogenin with incremental concentrations, up to saturation, of 5'-AMP [Fig. 2(A)] and 3'-UMP [Fig. 2(B)] to analyze the chemical shift changes associated with the binding of the two ligands. Figure 2(C) shows the chemical shift differences ( $\Delta\delta$ ) between the 5'-AMP- and 3'-UMP-bound forms of angiogenin. A rainbow color scale and thicker cartoon (tube) structure are used to identify  $\Delta\delta$  values with the largest variation ( $> 0.075$  ppm) appearing in the orange-to-red end of the spectrum [Fig. 2(C)]. A total of five (Gln12, Tyr14, Tyr25, Ala106, and Gln117) and 14 (Thr7, Leu10, Cys39, His47, Gly48, Lys50, Arg51, Ile53, Asn61, Glu58, His65, Asn68, Thr79, and Thr97) residues showed  $\Delta\delta$  (ppm)  $> 0.075$  ppm for the 5'-AMP- and 3'-UMP-bound states, respectively. A majority of these residues are located near the binding site of each of the two ligands, consistent with crystal structures of the two ligand-bound states in RNase homologues [Fig. 2(C)]. We extracted the NMR-calculated dissociation constant ( $K_d$ ) for each of the above residues and subsequently averaged it.<sup>21,22</sup> A lower  $K_d$  value indicates a greater affinity of the residues for their ligand. We obtained  $K_d$  values of 1.75 and 4.28 mM for the 5'-AMP- and 3'-UMP-bound states, respectively. These values are significantly higher than those observed for RNase A and ECP, whose 5'-AMP (3'-UMP)  $K_d$  values are 0.1 (0.01) mM and 0.34 (0.46) mM, respectively.<sup>9</sup> Despite lower

ligand affinity, NMR titrations show that nucleotide binding to angiogenin is specific and that purine versus pyrimidine subsite preference is preserved (Fig. 2). The binding affinities are slightly better for 5'-AMP than 3'-UMP, similar to preferences observed in RNase A and ECP. These weaker binding affinities (larger  $K_d$ ) relative to other structural homologues are consistent with the very low ribonucleolytic activity of angiogenin.<sup>11</sup> As previously proposed, this lower catalytic activity may result from steric clashes with Gln117, which interferes with ligand binding in angiogenin, thus considerably affecting the catalytic efficiency of the enzyme.<sup>23</sup>

#### Coordinated behavior of chemical shifts upon ligand binding

Displacement in the magnitude and direction of NMR chemical shifts during ligand titration and/or upon mutation can provide important insights into structural and/or dynamic changes correlated with ligand binding. A variety of approaches such as principal component analysis (PCA), chemical shift projection analysis (CHESPA), chemical shift covariance analysis (CHESCA), and coordinated chemical shifts behavior (CONCISE)<sup>15,16,24-27</sup> have been developed in recent years to analyze chemical shift changes due to mutations and/or ligand binding. Here, we used the CHESPA approach<sup>15,16</sup> to uncover coordinated chemical shift behaviors upon 5'-AMP and 3'-UMP binding to human angiogenin, providing means to clarify the long-range and global effects of molecular recognition in this enzyme [Fig. 3(A)]. A total of 27 residues showing  $\Delta\delta > 0.025$  ppm in either of the two ligand-bound states were selected for further analysis. For each selected residue, the



**Figure 2.**  $^1\text{H}$ - $^{15}\text{N}$  chemical shift difference upon binding of 5'-AMP and 3'-UMP to human angiogenin. Bar graph representing the weighted average chemical shift difference ( $\Delta\delta$ , in ppm) of the residues of angiogenin following the binding to saturation of (A) 5'-AMP and (B) 3'-UMP.  $\Delta\delta$  was calculated as a difference between the free and bound forms of angiogenin according to the equation  $\Delta\delta(\text{ppm}) = ((\Delta\delta_{\text{NH}}^2 + \Delta\delta_{\text{N}}^2)/25)^{1/2}$ .<sup>58</sup> (C)  $\Delta\delta$  for each residue is displayed using a rainbow color scale with a 0.0–0.1 ppm range, reported on the three-dimensional structure of angiogenin (PDB 1ANG). Predicted positions of each ligand (in dark gray) on the structure of angiogenin was performed by overlapping the RNase A 5'-AMP (PDB 1Z6S) and 3'-UMP (PDB 1O0N) complexes onto the corresponding three-dimensional structure of free angiogenin.

corresponding  $^1\text{H}$ - $^{15}\text{N}$  peak for the enzyme in its free form was mapped together with the 5'-AMP- and 3'-UMP-bound peaks upon ligand saturation [Figs. 3(B,C)].

Our results show that the chemical shift displacement of the ligand-bound states relative to the free form is along the same direction (correlated) for some residues (e.g., Ile53, Ile56, Gly62, and Val103), while other residues show uncorrelated displacements of the peaks corresponding to the two ligand-bound states (e.g., Thr7, Gln12, Ile119, and Phe120). To quantify the direction of the movement of the chemical shift peaks, we calculated  $\cos(\theta)$ , which represents the angle between vectors **A** (3'-UMP) and **B** (5'-AMP) [Fig. 3(A)]. Residues with a  $\cos(\theta) \approx 1$  show displacements along the same direction for the two ligand-bound states. A total of 14 residues showed correlated displacements (red color in Fig. 4), while the displacements for the remaining 13 residues were uncorrelated (blue color in Fig. 4). Interestingly, our results show that the majority of

correlated residues are localized to the  $\alpha 3/\beta 3/\beta 6/\beta 7/\text{L}3$  region, which corresponds to the  $V_1$  domain of pancreatic-like RNases.<sup>14</sup>

To quantitate the chemical shift displacement for the two ligand-bound states, we calculated the fractional shift  $X$ , which corresponds to the projection of the shift induced by 5'-AMP on the shift induced by 3'-UMP [Fig. 3(D)]. For the calculations, a scaling factor of 0.2 was applied to the  $^{15}\text{N}$ -chemical shift, as described.<sup>15</sup>  $X \approx 0$  indicates that the magnitude of the chemical shifts is identical for the binding for both ligands.  $X > 0$  corresponds to a larger displacement of the chemical shift for 3'-UMP relative to that of 5'-AMP with both peaks moving along the same direction, while  $X < 0$  is indicative of a displacement in opposite directions. Interestingly, all 14 residues showing coordinated behavior also have values of  $X > 0$ . This cluster of residues including Ile53, Ile56, Gly62, Ser72, Ser74, and Val103 have  $\cos(\theta) \approx 1$  and show a large value of  $X$ . This strong collective response to the binding of the two ligands indicates a network of residues activated by ligand binding that could potentially modulate a coordinated local structural rearrangement or a dynamic/allosteric response linked to function in human angiogenin.

#### Absence of conformational exchange on the millisecond timescale

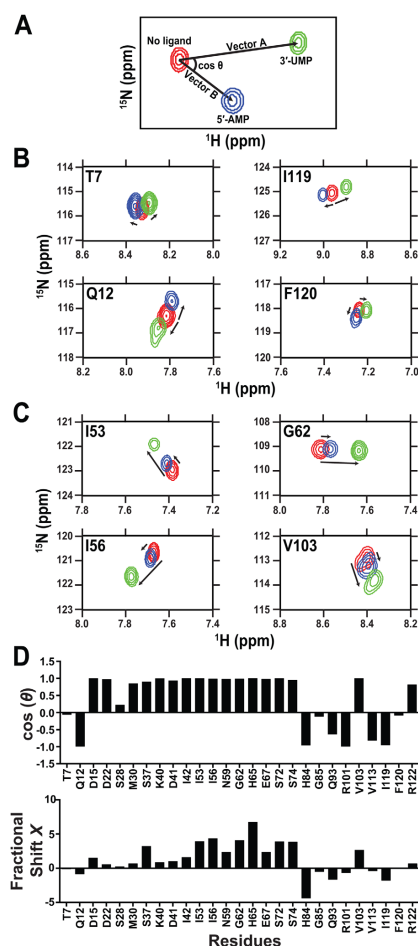
The above results suggest that angiogenin experiences conformational rearrangements upon ligand binding. Chemical shift analyses may report on subtle structural changes occurring in the enzyme upon ligand binding but do not necessarily provide details on the existence of collective motions experienced by residue networks undergoing conformational exchange in solution. To probe the conformational dynamics of angiogenin, we performed relaxation-compensated  $^{15}\text{N}$ -CPMG experiments on the free and 5'-AMP-bound states of the enzyme.  $^{15}\text{N}$ -CPMG experiments have previously been used to explore the conformational flexibility of RNases, effectively uncovering millisecond timescale motions clustered into dynamic residue networks essential to the ribonucleolytic function.<sup>4,9,28–32</sup> Interestingly, our results indicate that angiogenin is rigid in both free and bound states on the millisecond timescale. The absence of conformational exchange and the very low catalytic activity of angiogenin correlates with previous observations suggesting that flexibility on the millisecond timescale may be essential to proceed through some of the steps of the reaction coordinates in this enzyme family.<sup>4,9,14,28,30–33</sup> Angiogenin also displays a truncated functional and dynamic loop found in other highly active structural homologues (loop 4, residues 64–71 in RNase A).<sup>34</sup> If flexibility on the millisecond time frame is a requirement for optimal catalytic

activity in the fast-acting members of the pancreatic-like RNase fold (e.g., RNase A), the absence of catalytically relevant conformational exchange sampled by  $^{15}\text{N}$ -CPMG and the truncation of a normally dynamic loop in angiogenin could be expected, as the ribonucleolytic reaction catalyzed by this enzyme proceeds on a much slower timescale (seconds).<sup>11</sup> We note that faster nanosecond timescale motions in the loop regions of angiogenin have been reported previously.<sup>35,36</sup> These results suggest that the enzyme experiences side chain rearrangements on faster timescales, while the backbone remains rigid on the slower millisecond timescale detected by  $^{15}\text{N}$ -CPMG.

## Discussion

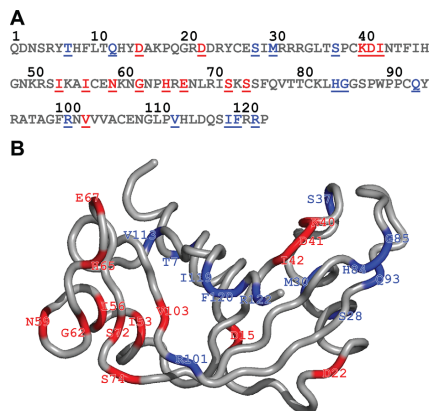
Conformational dynamics associated with ligand binding to an enzyme can involve large-scale structural fluctuations occurring over timescales that range from milliseconds to seconds (e.g., domain motions), or through small-scale changes (e.g., side chain) over shorter time frames, typically picoseconds to nanoseconds.<sup>37</sup> The hypothesis that millisecond motions are essential for the activity of enzymes is supported by an increasing number of studies that have reported this observation, a time frame that correlates with  $k_{cat}$  in most enzymes.<sup>4-7,38-47</sup> RNase A has been shown to experience millisecond motions that correlate with the  $k_{cat}$  of the ribonucleolytic reaction, that is, motions of residue clusters essential for product release.<sup>4,9,14,28,30-33</sup> The evolutionary conservation of these functionally relevant motional networks in the RNase superfamily remains elusive, but previous studies on ECP<sup>9</sup> and preliminary observations with other RNases suggests that the rate of conformational exchange on the catalytic timescale decreases together with catalytic activity. While this observation may be coincidental, preliminary data suggest that the more flexible RNase homologues—and those showing pre-organized dynamic residue clusters on the millisecond timescale—also appear to be the most catalytically active.<sup>9</sup>

In this study, we used  $^{15}\text{N}$ -CPMG relaxation dispersion experiments to probe the conformational dynamics of angiogenin on the millisecond timescale. Our results suggest that human angiogenin is largely rigid on this timescale, correlating with the very weak catalytic activity reported for this



**Figure 3.** NMR chemical shift projection analysis (CHESPA) approach describing independent and coordinated residue variations upon 5'-AMP and 3'-UMP binding to human angiogenin subsites. (A) Graphical representation of the CHESPA approach described in reference 15. The  $^1\text{H}$ - $^{15}\text{N}$  position of the peak is represented for the free (red) and bound forms (5'-AMP in blue and 3'-UMP in green) of the enzyme. The arrows indicate the movement of the  $^1\text{H}$ - $^{15}\text{N}$  chemical shift (length and direction) for each peak from its origin (in red) to its saturated position (blue and red). Projection analysis of eight selected residues responding in (B) independent or (C) coordinated manner. (D) Direction  $\cos(\theta)$  and magnitude (fractional shift  $X$ ) of the chemical shift perturbation of a subset of residues of angiogenin upon binding to saturation of the 5'-AMP and 3'-UMP. The  $\cos(\theta) = \frac{\vec{A} \cdot \vec{B}}{|\vec{A}| |\vec{B}|}$  quantifies the angle between vectors **A** (3'-UMP) and **B** (5'-AMP) from its initial position in the free form of the enzyme. The fractional shift  $X = \frac{\vec{A} \cdot \vec{B}}{|\vec{B}|^2}$  represents the fractional composite of vectors **A** and **B** induced by the ligand-induced chemical shift change. Bar graphs were prepared with GraphPad Prism. Collectiveness was observed for residues with a  $\cos(\theta) \approx 1$ , as previously described.<sup>15,16</sup>





**Figure 4.** Projection analysis of human angiogenin in the free and ligand-bound states. Only residues with a  $\Delta\delta > 0.025$  ppm were selected for the analysis. Residues showing coordinated displacements ( $\cos(\theta) \approx 1$ ) are represented in red on (A) the sequence (B) the three-dimensional structure of angiogenin. Alternatively, residues with uncoordinated displacements [ $\cos(\theta) \neq 1$ ] are shown in blue. The structure was prepared with PyMOL using PDB 1ANG.

enzyme.<sup>11,13,48</sup> Contrary to homologous RNases, angiogenin does not appear to rely on the same conformational exchange mechanism to promote its catalytic and/or biological neovascularization function. The enzyme could retain functionally relevant motions on longer timescales, that is, beyond the detection limit of the CPMG experiments. In addition to backbone rearrangements, binding of ligands to angiogenin might also trigger side chain reorganization, which occurs on the faster nanosecond timescale.<sup>36</sup> NMR flexibility work<sup>35</sup> and MD simulations<sup>14,36</sup> reporting on this faster nanosecond timescale have highlighted the importance of breathing motions between the  $V_1$  and  $V_2$  domains in angiogenin, similar to the hinge-bending motions observed in RNase A.

Using the CHESPA approach<sup>15,16</sup> to probe the effect of 3'-UMP and 5'-AMP binding relative to the free form of angiogenin, we showed that a number of residues exhibit chemical shift changes upon binding of two ligands at different subsites of the catalytic cleft. While some residues showed correlated displacements for the two distinct ligand-bound states—suggesting coordinated network behavior—chemical shift displacements for other residues were uncorrelated, suggesting uncoordinated behavior (Fig. 3). A mapping of these residues on the three-dimensional structure of angiogenin shows a spatial localization of residues displaying correlated displacements (Fig. 4). We note that residues located in or near the catalytic cleft, including the essential

catalytic residue His12, do not show correlated displacements. While the active-site residues display uncorrelated chemical shift changes upon binding the two ligands, clustering residues in the  $\alpha 3/\beta 3/\beta 6/\beta 7/L3$  regions show a correlated response to binding (Fig. 4). Interestingly, these residues cluster in the  $V_1$  domain of pancreatic RNases,<sup>14</sup> which shows high evolutionary conservation among vertebrate angiogenins [Figs. (1 and 4)]. A number of correlated residues belong to the hydrophobic core of angiogenin (Ile53, Ile56, Ser72, and Val103) and are essential for the conformational stability of the enzyme.<sup>20</sup> These results suggest that residues involved in maintaining the structural stability show correlated responses upon ligand binding, while the catalytic residues are perturbed differently upon binding to distinct ligands. Clustered residue networks could thus be functionally retained to preserve allosteric or structural reorganization movements essential for the binding of several nucleic acid ligands recognized by this protein fold,<sup>14</sup> while uncorrelated residues closer to the active site could offer finer specificities adapted to the function of each RNase A family member. We could also speculate that these correlated residue networks in angiogenin could be involved in neovascularization, a mechanism primarily involving residues of the  $V_1$  domain.<sup>49</sup> This observation is supported by previous NMR structural flexibility analyses on human and bovine angiogenins, where the human enzyme was noted to be more flexible in some regions of the  $V_1$  domain corresponding to a residue network presumably involved in binding to an endothelial cell receptor.<sup>35</sup>

A chimeric angiogenin protein was previously synthesized, in which residues 58–70 were swapped with the corresponding residues of RNase A. This hybrid showed reduced angiogenic activity, together with improved ribonucleolytic activity.<sup>49</sup> Also, segment 38–41 (located in the  $V_2$  domain) was shown to decrease the  $pK_a$  of Lys40, which contributed to increasing the catalytic activity of the enzyme.<sup>50</sup> It is intriguing to note that our CHESPA analysis shows that residues of these two clusters display a collective response upon binding of two distinct ligands to functionally different subsites of the catalytic cleft. These results could suggest the existence of at least two conformational substates (free and bound) that respond synergistically toward a specific protein function.

In summary, we have used the CHESPA approach to demonstrate that networks of clustering residues respond in a correlated fashion upon distinct subsite binding events in human angiogenin. Many residues displaying this collective behavior cluster in the  $V_1$  domain and are known to be involved in the hydrophobic core of the protein, suggesting the existence of local unfolding events and/or dynamic reorganization of the structure upon

distinct subsite binding events. Contrary to fast-acting structural homologues experiencing collective residue network reorganization on the millisecond timescale upon ligand binding, conformational exchange on this particular time frame was not observed in human angiogenin. The lack of millisecond dynamics is supported by the extremely low binding affinities and catalytic activity of angiogenin, which requires several seconds to cleave a single RNA molecule.<sup>11</sup> The present study provides further information on the intramolecular long-range residue networks that may control the function of this enzyme. Further analyses will offer additional details on this functional mechanism in angiogenin, in addition to providing insights into the interplay between its catalytic activity and the angiogenic response.

## Materials and Methods

### Evolutionary conservation of amino acid positions

The evolutionary conservation of amino acid positions based on the phylogenetic relations between homologous sequences was performed using the ConSurf Server (<http://consurf.tau.ac.il/>).<sup>18</sup> Human angiogenin was submitted to a Basic Local Alignment Search Tool (BLAST) and 150 homologous sequences with a respective minimum and maximum conservation score of 35% and 95% were selected. The sequences were aligned with MAFFT.<sup>51</sup> The conservation score was calculated by a Bayesian method, and a JTT was used as a substitution model. Evolutionarily conserved positions and regions were visualized with PyMOL.

### Cloning, expression, and purification

The sequence of angiogenin was acquired from UniProt under accession number P03950. Residues 25–147 (123 amino acids) were codon-optimized for *Escherichia coli* and subcloned into NdeI/HindIII-digested expression vector pJexpress411 (DNA2.0, Menlo Park, CA). The construct was transformed into *E. coli* BL21(DE3) and selected using kanamycin. Human angiogenin was expressed in M9 minimal medium for 6 hours at 37°C and purified as previously described.<sup>9,19</sup> Protein concentration was estimated using a theoretical extinction coefficient of 11,835 M<sup>-1</sup>cm<sup>-1</sup>, as calculated by ProtParam (ExPASy).

### <sup>15</sup>N-Carr-Purcell-Meiboom-Gill (CPMG) NMR relaxation dispersion experiments

All NMR experiments were performed at 298 K on samples containing 0.32 mM <sup>15</sup>N-labelled angiogenin (free or ligand-saturated forms) in 15 mM sodium acetate at pH 5.0 supplemented with 10% D<sub>2</sub>O. NMR experiments were carried out on Varian (Agilent) 500 MHz (11.7 T) and 800 MHz (18.8 T) NMR

spectrometers equipped with triple-resonance cold probes and pulsed field gradients. Backbone resonances were assigned using standard multidimensional experiments.<sup>52</sup> Interleaved two-dimensional spectra were collected in a constant time manner with  $\tau_{cp}$  CPMG repetition delays of 0.625, 0.714 ( $\times 2$ ), 1.0, 1.25, 1.67, 2.0, 2.50 ( $\times 2$ ), 3.33, 5.0, and 10 ms, using a total relaxation period of 40 milliseconds. All NMR spectra were processed using NMRPipe,<sup>53</sup> in-house CPMG scripts and analyzed with Sparky.<sup>54</sup>

### NMR titration experiments

All NMR titration experiments were conducted in 15 mM sodium acetate at pH 5.0. The pH was carefully monitored throughout the experiments and readjusted with Tris base or acetic acid, when necessary. Ligands 3'-UMP (Chemical Impex Intl Inc., Wood Dale, IL) and 5'-AMP (BioBasic Inc., Markham, ON, Canada) were reconstituted in the same buffer as the protein. <sup>1</sup>H-<sup>15</sup>N sensitivity-enhanced HSQC experiments were acquired at 800 MHz (18.8 T) using spectral widths (points) of 2025 and 8000 in the  $t_1$  and  $t_2$  dimensions, respectively. Binding kinetics was investigated by the titration of each ligand with enzyme/ligand molar ratios of 0, 0.174, 0.393, 0.691, 1.31, 2.71, 6, 12, 18, 24, and 30. The dissociation constant ( $K_d$ ) was extracted for each residue with a  $\Delta\delta > 0.075$  ppm and averaged.<sup>21,22</sup> For each residue, a curve corresponding to the chemical shift variation as a function of ligand concentration was plotted. The  $K_d$  corresponds to the intercept on the x-axis (ligand concentration) at 50% variation of the chemical shift. The  $K_d$  for all residues was averaged to provide a global  $K_d$  value with error, as previously described.<sup>55</sup>

### NMR chemical shift projection analysis (CHESPA)

The CHESPA analysis was carried out according to the protocol described by Axe *et al.*<sup>16</sup> A total of 27 residues showing a chemical shift variation  $\Delta\delta > 0.025$  ppm with both ligands was selected for further analysis. For each selected residue, the cosine  $(\theta) = \frac{\vec{A} \cdot \vec{B}}{|\vec{A}| |\vec{B}|}$  and the fractional shift  $X = \frac{\vec{A} \cdot \vec{B}}{|\vec{B}|^2}$  were calculated according to Selvaratnam *et al.*<sup>15</sup> using the <sup>1</sup>H and <sup>15</sup>N coordinates of the peak in its initial position (free enzyme) or at saturation upon binding with 3'-UMP (Vector **A**) and 5'-AMP (Vector **B**). Bar graphs were prepared with GraphPad Prism.

### Acknowledgments

The authors thank Tara Sprules of the Quebec/Eastern Canada High Field NMR Facility (McGill University) for her excellent technical assistance, and



Pratul K. Agarwal (Oak Ridge National Laboratory) for helpful comments. This study was supported by the National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health (NIH) under award number R01GM105978. D.G. holds a Natural Sciences and Engineering Research Council of Canada (NSERC) Alexander Graham Bell Canada Graduate Scholarship. N.D. holds a Fonds de Recherche Québec – Santé (FRQS) Research Scholar Junior 1 Career Award.

## References

- Koshland DE (1958) Application of a theory of enzyme specificity to protein synthesis. *Proc Natl Acad Sci USA* 44:98–104.
- Boehr DD, Nussinov R, Wright PE (2009) The role of dynamic conformational ensembles in biomolecular recognition. *Nat Chem Biol* 5:789–796.
- Ramanathan A, Savol A, Burger V, Chennubhotla CS, Agarwal PK (2014) Protein conformational populations and functionally relevant substates. *Acc Chem Res* 47:149–156.
- Beach H, Cole R, Gill ML, Loria JP (2005) Conservation of mus-ms enzyme motions in the apo- and substrate-mimicked state. *J Am Chem Soc* 127:9167–9176.
- McElheny D, Schnell JR, Lansing JC, Dyson HJ, Wright PE (2005) Defining the role of active-site loop fluctuations in dihydrofolate reductase catalysis. *Proc Natl Acad Sci USA* 102:5032–5037.
- Boehr DD, McElheny D, Dyson HJ, Wright PE (2006) The dynamic energy landscape of dihydrofolate reductase catalysis. *Science* 313:1638–1642.
- Wolf-Watz M, Thai V, Henzler-Wildman K, Hadjipavlou G, Eisenmesser EZ, Kern D (2004) Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nat Struct Mol Biol* 11:945–949.
- Henzler-Wildman KA, Thai V, Lei M, Ott M, Wolf-Watz M, Fenn T, Pozharski E, Wilson MA, Petsko GA, Karplus M, Hubner CG, Kern D (2007) Intrinsic motions along an enzymatic reaction trajectory. *Nature* 450:838–844.
- Gagné D, Charest LA, Morin S, Kovrigin EL, Doucet N (2012) Conservation of flexible residue clusters among structural and functional enzyme homologues. *J Biol Chem* 287:44289–44300.
- Strydom DJ, Fett JW, Lobb RR, Alderman EM, Bethune JL, Riordan JF, Vallee BL (1985) Amino acid sequence of human tumor derived angiogenin. *Biochemistry* 24:5486–5494.
- Park C, Kelemen BR, Klink TA, Sweeney RY, Behlke MA, Eubanks SR, Raines RT (2001) Fast, facile, hypersensitive assays for ribonucleolytic activity. *Methods Enzymol* 341:81–94.
- Shapiro R, Vallee BL (1989) Site-directed mutagenesis of histidine-13 and histidine-114 of human angiogenin. Alanine derivatives inhibit angiogenin-induced angiogenesis. *Biochemistry* 28:7401–7408.
- Curran TP, Shapiro R, Riordan JF (1993) Alteration of the enzymatic specificity of human angiogenin by site-directed mutagenesis. *Biochemistry* 32:2307–2313.
- Gagné D, Doucet N (2013) Structural and functional importance of local and global conformational fluctuations in the RNase A superfamily. *FEBS J* 280:5596–5607.
- Selvaratnam R, VanSchouwen B, Fogolari F, Mazhab-Jafari MT, Das R, Melacini G (2012) The projection analysis of NMR chemical shifts reveals extended EPAC autoinhibition determinants. *Biophys J* 102:630–639.
- Axe JM, Boehr DD (2013) Long-range interactions in the alpha subunit of tryptophan synthase help to coordinate ligand binding, catalysis, and substrate channeling. *J Mol Biol* 425:1527–1545.
- Axe JM, Yezdimer EM, O'Rourke KF, Kerstetter NE, You W, Chang CE, Boehr DD (2014) Amino acid networks in a (beta/alpha)<sub>8</sub> barrel enzyme change during catalytic turnover. *J Am Chem Soc* 136:6818–6821.
- Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res* 38:W529–W533.
- Gagné D, Doucet N (2014) (in press) Sequence-specific backbone <sup>1</sup>H, <sup>13</sup>C, and <sup>15</sup>N resonance assignments of human ribonuclease 4. *Biomol NMR Assign*.
- Smith BD, Raines RT (2006) Genetic selection for critical residues in ribonucleases. *J Mol Biol* 362:459–478.
- Fielding L (2003) NMR methods for the determination of protein-ligand dissociation constants. *Curr Top Med Chem* 3:39–53.
- Williamson MP (2013) Using chemical shift perturbation to characterise ligand binding. *Prog Nucl Mag Res Sp* 73:1–16.
- Russo N, Shapiro R, Acharya KR, Riordan JF, Vallee BL (1994) Role of glutamine-117 in the ribonucleolytic activity of human angiogenin. *Proc Natl Acad Sci USA* 91:2920–2924.
- Sakurai K, Goto Y (2007) Principal component analysis of the pH-dependent conformational transitions of bovine beta-lactoglobulin monitored by heteronuclear NMR. *Proc Natl Acad Sci USA* 104:15346–15351.
- Selvaratnam R, Chowdhury S, VanSchouwen B, Melacini G (2011) Mapping allostery through the covariance analysis of NMR chemical shifts. *Proc Natl Acad Sci USA* 108:6133–6138.
- Konuma T, Lee YH, Goto Y, Sakurai K (2013) Principal component analysis of chemical shift perturbation data of a multiple-ligand-binding system for elucidation of respective binding mechanism. *Proteins* 81:107–118.
- Cembran A, Kim J, Gao J, Veglia G (2014) NMR mapping of protein conformational landscapes using coordinated behavior of chemical shifts upon ligand binding. *Phys Chem Chem Phys* 16:6508–6518.
- Cole R, Loria JP (2002) Evidence for flexibility in the function of ribonuclease A. *Biochemistry* 41:6072–6081.
- Watt ED, Shimada H, Kovrigin EL, Loria JP (2007) The mechanism of rate-limiting motions in enzyme function. *Proc Natl Acad Sci USA* 104:11981–11986.
- Doucet N, Watt ED, Loria JP (2009) The flexibility of a distant loop modulates active site motion and product release in ribonuclease A. *Biochemistry* 48:7160–7168.
- Doucet N, Khirich G, Kovrigin EL, Loria JP (2011) Alteration of hydrogen bonding in the vicinity of histidine 48 disrupts millisecond motions in RNase A. *Biochemistry* 50:1723–1730.
- Watt ED, Rivalta I, Whittier SK, Batista VS, Loria JP (2011) Reengineering rate-limiting, millisecond enzyme motions by introduction of an unnatural amino acid. *Biophys J* 101:411–420.
- Watt ED, Shimada H, Kovrigin EL, Loria JP (2007) The mechanism of rate-limiting motions in enzyme function. *Proc Natl Acad Sci USA* 104:11981–11986.
- Ramanathan A, Agarwal PK (2011) Evolutionarily conserved linkage between enzyme fold, flexibility, and catalysis. *PLoS Biol* 9:e1001193.

35. Lequin O, Thuring H, Robin M, Lallemand JY (1997) Three-dimensional solution structure of human angiogenin determined by  $^1\text{H}$ ,  $^{15}\text{N}$ -NMR spectroscopy—characterization of histidine protonation states and pKa values. *Eur J Biochem* 250:712–726.
36. Merlino A, Vitagliano L, Ceruso MA, Mazzarella L (2003) Subtle functional collective motions in pancreatic-like ribonucleases: from ribonuclease A to angiogenin. *Proteins* 53:101–110.
37. Doucet N, Pelletier JN. Gaining insight into enzyme function through correlations with protein motions. In: Lutz S, Bornscheuer UT, Eds. (2009) *Protein engineering handbook*. Wiley-VCH, Weinheim, pp 187–211.
38. Mulder FA, Mittermaier A, Hon B, Dahlquist FW, Kay LE (2001) Studying excited states of proteins by NMR spectroscopy. *Nat Struct Biol* 8:932–935.
39. Rozovsky S, Jogl G, Tong L, McDermott AE (2001) Solution-state NMR investigations of triosephosphate isomerase active site loop motion: ligand release in relation to active site loop dynamics. *J Mol Biol* 310:271–280.
40. Korzhnev DM, Orekhov VY, Dahlquist FW, Kay LE (2003) Off-resonance R1rho relaxation outside of the fast exchange limit: an experimental study of a cavity mutant of T4 lysozyme. *J Biomol NMR* 26:39–48.
41. Eisenmesser EZ, Millet O, Labeikovsky W, Korzhnev DM, Wolf-Watz M, Bosco DA, Skalicky JJ, Kay LE, Kern D (2005) Intrinsic dynamics of an enzyme underlies catalysis. *Nature* 438:117–121.
42. Massi F, Wang C, Palmer AG, III (2006) Solution NMR and computer simulation studies of active site loop motion in triosephosphate isomerase. *Biochemistry* 45:10787–10794.
43. Vallurupalli P, Kay LE (2006) Complementarity of ensemble and single-molecule measures of protein motion: a relaxation dispersion NMR study of an enzyme complex. *Proc Natl Acad Sci USA* 103:11910–11915.
44. Berlow RB, Igumenova TI, Loria JP (2007) Value of a hydrogen bond in triosephosphate isomerase loop motion. *Biochemistry* 46:6001–6010.
45. Doucet N, Savard PY, Pelletier JN, Gagné SM (2007) NMR investigation of Tyr105 mutants in TEM-1 beta-lactamase: dynamics are correlated with function. *J Biol Chem* 282:21448–21459.
46. Kempf JG, Jung JY, Ragain C, Sampson NS, Loria JP (2007) Dynamic requirements for a functional protein hinge. *J Mol Biol* 368:131–149.
47. Labeikovsky W, Eisenmesser EZ, Bosco DA, Kern D (2007) Structure and dynamics of pin1 during catalysis by NMR. *J Mol Biol* 367:1370–1381.
48. Russo N, Acharya KR, Vallee BL, Shapiro R (1996) A combined kinetic and modeling study of the catalytic center subsites of human angiogenin. *Proc Natl Acad Sci USA* 93:804–808.
49. Wade Harper J, Vallee BL (1989) A covalent angiogenin/ribonuclease hybrid with a fourth disulfide bond generated by regional mutagenesis. *Biochemistry* 28:1875–1884.
50. Holloway DE, Chavali GB, Hares MC, Baker MD, Subbarao GV, Shapiro R, Acharya KR (2004) Crystallographic studies on structural features that determine the enzymatic specificity and potency of human angiogenin: Thr44, Thr80, and residues 38–41. *Biochemistry* 43:1230–1241.
51. Katoh K, Standley DM (2013) MAFPT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–780.
52. Cavanagh J, Fairbrother WJ, Palmer AG, Rance M, Skelton NJ (2007) *Protein NMR spectroscopy: principles and practice*, Elsevier Academic, San Diego.
53. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–293.
54. Goddard TD, Kneller DG Sparky 3. University of California, San Francisco.
55. Mittermaier A, Meneses E (2013) Analyzing protein-ligand interactions by dynamic NMR spectroscopy. *Methods Mol Biol* 1008:243–266.
56. Mayrose I, Mitchell A, Pupko T (2005) Site-specific evolutionary rate inference: taking phylogenetic uncertainty into account. *J Mol Evol* 60:345–353.
57. Acharya KR, Shapiro R, Allen SC, Riordan JF, Vallee BL (1994) Crystal structure of human angiogenin reveals the structural basis for its functional divergence from ribonuclease. *Proc Natl Acad Sci USA* 91:2915–2919.
58. Grzesiek S, Stahl SJ, Wingfield PT, Bax A (1996) The CD4 determinant for downregulation by HIV-1 Nef directly binds to Nef. Mapping of the Nef binding surface by NMR. *Biochemistry* 35:10256–10261.

## 5.5 Discussion

Les études chez la RNase A ont démontré la présence de mouvements atomiques à l'échelle de temps de la  $\mu$ s-ms [160, 166]. Le taux d'échange conformationnel mesuré par  $^{15}\text{N}$ -CPMG correspondait au  $k_{cat}$  de la réaction catalytique, un résultat qui laissait supposer que les mouvements pouvaient être nécessaires à la fonction de la protéine [160]. Cette hypothèse était supportée par un bon nombre d'études portant sur d'autres systèmes enzymatiques [10, 11, 14, 53, 58-60, 62, 188]. Ces résultats laissent supposer que si les mouvements sont importants à la fonction de la protéine, elles seront conservées chez des protéines de structure et de fonction similaires.

Par cristallographie et simulation de dynamique moléculaire, Merlino *et al.* avait montré des changements conformationnels dans l'ordre de la nanoseconde chez l'angiogénine [189]. Les réarrangements structuraux à cette échelle de temps sont en général associés avec la rotation des chaînes latérales, et peuvent être impliqués dans la liaison du substrat. En nous basant sur cette étude, nous avons voulu voir si l'angiogénine subissait également des mouvements à l'échelle de temps de la  $\mu$ s-ms. Pour ce faire, nous avons effectué des expériences de  $^{15}\text{N}$ -CPMG en utilisant la même séquence d'impulsions que dans nos études précédentes (voir les chapitres 3 à 6). Nous avons observé que l'angiogénine est complètement rigide à l'échelle de temps de la  $\mu$ s-ms. Ce résultat n'est peut-être pas si surprenant, et corrèle avec la faible activité catalytique affichée par l'enzyme. En effet, nous avons précédemment proposé que les RNases offrant la plus forte efficacité catalytique montraient aussi un  $k_{ex}$  élevé. Ce résultat peut suggérer la présence de mouvements dynamiques à plus faible échelle de temps, non mesurables par l'expérience de  $^{15}\text{N}$ -CPMG.

Nous avons utilisé le CHESPA afin de quantifier ces mouvements (direction et amplitude) reliés aux déplacements chimiques suite à la liaison de deux produits de la réaction, 3'-UMP et 5'-AMP. Ainsi, nous montrons la présence de deux réseaux distincts de résidus où les déplacements chimiques sont coordonnés. Ce résultat est d'autant plus surprenant du fait que ces déplacements coordonnés se produisent suite à la liaison de deux ligands se liant à deux sous-sites distincts. Le premier réseau, situé dans le secteur  $V_1$ , est composé de 14 résidus et inclut les structures secondaires  $\alpha 3/\beta 3/\beta 6/\beta 7/L3$ . Cette région a précédemment été associée à la fonction de néovascularisation chez l'angiogénine [190]. En effet, une réduction de l'activité angiogénique, combinée à une augmentation de l'activité catalytique, a été observée suite à la mutation des résidus 58-70, suggérant la présence d'interactions à longue

distance. Plusieurs résidus faisant partie du noyau hydrophobe (Ile53, Ile56, Ser72 et Val103) sont aussi présents dans cette région, qui pourrait également jouer un rôle dans la stabilité de la protéine [161]. En plus du secteur  $V_1$ , le segment correspondant aux résidus 40-42 (secteur  $V_2$ ) a démontré une réaction coordonnée suite à la liaison des ligands. Ce secteur inclut la Lys40, un résidu catalytique essentiel et qui a pour fonction de stabiliser l'état de transition [158]. La mutation du segment 38-41 cause une réduction du  $pK_a$  de la Lys40, contribuant à augmenter l'activité catalytique [191]. Ces résultats pourraient suggérer la présence d'au moins deux états conformationnels (libre et lié) qui répondent en synergie pour une fonction biologique donnée.

## 6 CONCLUSIONS ET PERSPECTIVES

Les enzymes sont des macromolécules qui jouent le rôle de catalyseurs biologiques. Elles abaissent l'énergie d'activation et accélèrent les réactions chimiques qui autrement seraient trop longues. Étant "propres" pour l'environnement, elles s'avèrent très intéressantes pour de nombreuses applications industrielles. Toutefois, leurs applications potentielles nécessiteront qu'elles soient au préalable modifiées afin qu'elles répondent exactement aux besoins donnés. Ces modifications sont encore complexes et souvent aléatoires, principalement dû à notre manque de connaissance du mécanisme de fonctionnement de ces enzymes. En effet, bien qu'elles soient étudiées depuis plus d'un siècle, nous réalisons qu'il nous manque encore des informations pour bien comprendre leur fonctionnement. Daniel E. Kosland a été le premier, en 1958, à proposer que les enzymes pouvaient subir des changements conformationnels pour accomplir sa fonction [3]. Depuis plusieurs années, un bon nombre de techniques permettant l'étude de la flexibilité atomique chez les protéines ont vu le jour. Même si nous n'en sommes encore qu'au début, le champ progresse très rapidement. Cependant, il demeure encore difficile de relier l'activité d'une enzyme avec les mouvements moléculaires, bien qu'il y ait de plus en plus d'évidences en ce sens.

En partant du concept que la dynamique soit requise pour la catalyse enzymatique, il est facile de conclure que la perturbation de ces mouvements atomiques affecterait l'activité de la protéine. De manière similaire, on pourrait émettre l'hypothèse que ces mouvements seront conservées entre des homologues structuraux partageant la même fonction. Pour cette étude, nous avons voulu tester ces hypothèses en utilisant les membres de la famille des RNases humaines. Nous démontrons que la dynamique globale à l'échelle de temps de la  $\mu$ s-ms est conservée entre plusieurs homologues structuraux. Les résultats suggèrent qu'un taux d'échange conformationnel ( $k_{ex}$ ) plus rapide coïncide avec une activité catalytique plus élevée, bien qu'un lien direct n'ait pu être établi. Nous avons observé que les résidus dynamiques, plutôt que d'être étendus à l'ensemble de la structure, étaient confinés dans des secteurs, lesquels étaient spécifiques à chaque enzyme. Ce concept de "secteurs dynamiques" supporte les résultats du groupe de Ranganathan [172], qui avait prédit une évolution corrélée de secteurs de résidus contigus chez la protéase à sérine S1A. Il avait proposé que ces secteurs puissent répondre à une fonction spécifique de l'enzyme. Par la création d'une chimère (RNase A<sub>ECP</sub>), nous avons modifié le profil de ces secteurs dynamiques et affecté la fonction de la protéine. De manière similaire, nous avons montré qu'une modification mineure, soit le retrait d'un

groupement méthyle, pouvait perturber la flexibilité de la protéine et créer un repositionnement du ligand. Ces deux mutants appuient l'hypothèse voulant que la dynamique soit importante pour la fonction de la protéine. Bien que ces mouvements atomiques soient conservés entre les homologues structuraux de fonctions similaires, des distinctions particulières sont visibles chez les protéines qui ont évolué il y a très longtemps. C'est le cas de le RNase A et d'ECP, qui présentent des secteurs dynamiques distincts, malgré la présence de mouvements globaux à l'échelle de temps de la  $\mu$ s-ms chez les deux protéines. La caractérisation d'EDN et d'ECP humaine permet d'appuyer ce résultat, puisque les deux enzymes, qui sont le produit d'une duplication récente du gène, donnent des profils relativement semblables à l'état libre.

Nos études comparatives de dynamique moléculaire chez les membres d'une même famille d'homologues structuraux accomplissant des fonctions similaires sont novatrices. En combinant plusieurs approches, telles que la spectroscopie RMN, la biologie moléculaire et la bio-informatique, nous avons été en mesure de répondre aux objectifs cités à la section 1.6.2 du présent document. Ces études constituent en quelque sorte la base pour de nouvelles recherches qui nous permettront d'améliorer notre compréhension du rôle biologique des RNases, mais aussi de celui des enzymes en général. Cette compréhension des mécanismes catalytiques est essentielle pour le développement de nouvelles enzymes qui pourront avoir un impact considérable sur notre quotidien.

## 7 RÉFÉRENCES

1. Lehninger, A.L., *Principles of biochemistry*. Worth Publishers, Inc., 1982. **New York, NY USA**: p. 211-220.
2. Fischer, E., *Einfluss der konfiguration auf die wirkung der enzvnie*. Ber. Deutsch. Chem. Ges., 1894. **27**: p. 2985-2993.
3. Koshland, D.E., *Application of a Theory of Enzyme Specificity to Protein Synthesis*. Proc Natl Acad Sci U S A, 1958. **44**(2): p. 98-104.
4. Foote, J. and C. Milstein, *Conformational isomerism and the diversity of antibodies*. Proc Natl Acad Sci U S A, 1994. **91**(22): p. 10370-4.
5. Leder, L., et al., *Spectroscopic, calorimetric, and kinetic demonstration of conformational adaptation in peptide-antibody recognition*. Biochemistry, 1995. **34**(50): p. 16509-18.
6. Tsai, C.J., B. Ma, and R. Nussinov, *Folding and binding cascades: shifts in energy landscapes*. Proc Natl Acad Sci U S A, 1999. **96**(18): p. 9970-2.
7. Van Regenmortel, M.H., *Transcending the structuralist paradigm in immunology-affinity and biological activity rather than purely structural considerations should guide the design of synthetic peptide epitopes*. Biomed Pept Proteins Nucleic Acids, 1995. **1**(3): p. 109-16.
8. Tsai, C.J., et al., *Folding funnels, binding funnels, and protein function*. Protein Sci, 1999. **8**(6): p. 1181-90.
9. Boehr, D.D., R. Nussinov, and P.E. Wright, *The role of dynamic conformational ensembles in biomolecular recognition*. Nat Chem Biol, 2009. **5**(11): p. 789-96.
10. Boehr, D.D., et al., *The dynamic energy landscape of dihydrofolate reductase catalysis*. Science, 2006. **313**(5793): p. 1638-42.
11. McElheny, D., et al., *Defining the role of active-site loop fluctuations in dihydrofolate reductase catalysis*. Proc Natl Acad Sci U S A, 2005. **102**(14): p. 5032-7.
12. Beach, H., et al., *Conservation of mus-ms enzyme motions in the apo- and substrate-mimicked state*. J Am Chem Soc, 2005. **127**(25): p. 9167-76.
13. Henzler-Wildman, K.A., et al., *Intrinsic motions along an enzymatic reaction trajectory*. Nature, 2007. **450**(7171): p. 838-44.
14. Wolf-Watz, M., et al., *Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair*. Nat Struct Mol Biol, 2004. **11**(10): p. 945-9.
15. Eisenmesser, E.Z., et al., *Enzyme dynamics during catalysis*. Science, 2002. **295**(5559): p. 1520-3.
16. Henzler-Wildman, K. and D. Kern, *Dynamic personalities of proteins*. Nature, 2007. **450**(7172): p. 964-72.
17. Doucet, N. and J.N. Pelletier, *Gaining insight into enzyme function through correlations with protein motions*, in *Protein engineering handbook*, S. Lutz and U.T. Bornscheuer, Editors. 2009, Wiley-VCH: Weinheim. p. 187-211.
18. Agarwal, P.K., *Enzymes: An integrated view of structure, dynamics and function*. Microb Cell Fact, 2006. **5**: p. 2.

19. Boehr, D.D., H.J. Dyson, and P.E. Wright, *An NMR perspective on enzyme dynamics*. Chem Rev, 2006. **106**(8): p. 3055-79.
20. Monod, J., J. Wyman, and J.P. Changeux, *On the Nature of Allosteric Transitions: A Plausible Model*. J Mol Biol, 1965. **12**: p. 88-118.
21. Feixas, F., et al., *Exploring the role of receptor flexibility in structure-based drug discovery*. Biophys Chem, 2014. **186**: p. 31-45.
22. Lee, G.M. and C.S. Craik, *Trapping moving targets with small molecules*. Science, 2009. **324**(5924): p. 213-5.
23. Martinez-Cajas, J.L. and M.A. Wainberg, *Antiretroviral therapy : optimal sequencing of therapy to avoid resistance*. Drugs, 2008. **68**(1): p. 43-72.
24. Martins, S., M.J. Ramos, and P.A. Fernandes, *The current status of the NNRTI family of antiretrovirals used in the HAART regime against HIV infection*. Curr Med Chem, 2008. **15**(11): p. 1083-95.
25. De Francesco, R. and A. Carfi, *Advances in the development of new therapeutic agents targeting the NS3-4A serine protease or the NS5B RNA-dependent RNA polymerase of the hepatitis C virus*. Adv Drug Deliv Rev, 2007. **59**(12): p. 1242-62.
26. McCammon, J.A., B.R. Gelin, and M. Karplus, *Dynamics of folded proteins*. Nature, 1977. **267**(5612): p. 585-90.
27. Lipari, G. and A. Szabo, *Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results*. J Am Chem Soc, 1982. **104**: p. 4559-4570.
28. Kempf, J.G. and J.P. Loria, *Protein dynamics from solution NMR: theory and applications*. Cell Biochem Biophys, 2003. **37**(3): p. 187-211.
29. Kleckner, I.R. and M.P. Foster, *An introduction to NMR-based approaches for measuring protein dynamics*. Biochim Biophys Acta, 2010.
30. Abragam, A., *The principles of nuclear magnetism*. The International series of monographs on physics. 1961, Oxford,: Clarendon Press. 599 p.
31. Morin, S., *A practical guide to protein dynamics from 15N spin relaxation in solution*. Prog Nucl Magn Reson Spectrosc, 2011. **59**(3): p. 245-62.
32. Millet, O., et al., *The static magnetic field dependence of chemical exchange linebroadening defines the NMR chemical shift time scale*. J. Am. Chem. Soc., 2000. **122**: p. 2867-2877.
33. Baldwin, A.J. and L.E. Kay, *NMR spectroscopy brings invisible protein states into focus*. Nat Chem Biol, 2009. **5**(11): p. 808-14.
34. Loria, J.P., R.B. Berlow, and E.D. Watt, *Characterization of enzyme motions by solution NMR relaxation dispersion*. Acc Chem Res, 2008. **41**(2): p. 214-21.
35. Hansen, D.F., P. Vallurupalli, and L.E. Kay, *Using relaxation dispersion NMR spectroscopy to determine structures of excited, invisible protein states*. J Biomol NMR, 2008. **41**: p. 113-120.
36. Skrynnikov, N.R., F.W. Dahlquist, and L.E. Kay, *Reconstructing NMR spectra of "invisible" excited protein states using HSQC and HMQC experiments*. J Am Chem Soc, 2002. **124**(41): p. 12352-60.



37. Lipari, G. and A. Szabo, *Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity*. J Am Chem Soc, 1982. **104**: p. 4546-4559.
38. Mandel, A.M., M. Akke, and A.G. Palmer, 3rd, *Backbone dynamics of Escherichia coli ribonuclease HI: correlations with structure and function in an active enzyme*. J Mol Biol, 1995. **246**(1): p. 144-63.
39. Palmer, A.G.r., M. Rance, and P.E. Wright, *Intramolecular motions of a zinc finger DNA-binding domain Xfin characterized by proton-detected natural abundance <sup>13</sup>C heteronuclear NMR spectroscopy*. J Am Chem Soc, 1991. **113**: p. 4371-4380.
40. Cole, R. and J.P. Loria, *FAST-Modelfree: a program for rapid automated analysis of solution NMR spin-relaxation data*. J Biomol NMR, 2003. **26**(3): p. 203-13.
41. Orekhov, V.Y., et al., *Processing of heteronuclear NMR relaxation data with the new software DASHA*. Appl Magn Reson, 1995. **9**: p. 581-588.
42. Cordier, F., et al., *Solution structure, rotational diffusion anisotropy and local backbone dynamics of Rhodobacter capsulatus cytochrome c2*. J Mol Biol, 1998. **281**(2): p. 341-61.
43. Dosset, P., et al., *Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data*. J Biomol NMR, 2000. **16**(1): p. 23-8.
44. Fushman, D., S. Cahill, and D. Cowburn, *The main-chain dynamics of the dynamin pleckstrin homology (PH) domain in solution: analysis of <sup>15</sup>N relaxation with monomer/dimer equilibration*. J Mol Biol, 1997. **266**(1): p. 173-94.
45. d'Auvergne, E.J. and P.R. Gooley, *Optimisation of NMR dynamic models II. A new methodology for the dual optimisation of the model-free parameters and the Brownian rotational diffusion tensor*. J Biomol NMR, 2008. **40**(2): p. 121-33.
46. d'Auvergne, E.J. and P.R. Gooley, *Optimisation of NMR dynamic models I. Minimisation algorithms and their performance within the model-free and Brownian rotational diffusion spaces*. J Biomol NMR, 2008. **40**(2): p. 107-19.
47. Cook, B.W. and G.S. Shaw, *Architecture of the catalytic HPN motif is conserved in all E2 conjugating enzymes*. Biochem J, 2012. **445**(2): p. 167-74.
48. Yao, H., et al., *NMR studies of the interaction between human programmed cell death 5 and human p53*. Biochemistry, 2012. **51**(13): p. 2684-93.
49. Selvaratnam, R., et al., *Mapping allostery through the covariance analysis of NMR chemical shifts*. Proc Natl Acad Sci U S A, 2011. **108**(15): p. 6133-8.
50. Selvaratnam, R., et al., *The projection analysis of NMR chemical shifts reveals extended EPAC autoinhibition determinants*. Biophys J, 2012. **102**(3): p. 630-9.
51. Cembran, A., et al., *NMR mapping of protein conformational landscapes using coordinated behavior of chemical shifts upon ligand binding*. Phys Chem Chem Phys, 2014. **16**(14): p. 6508-18.
52. Axe, J.M. and D.D. Boehr, *Long-range interactions in the alpha subunit of tryptophan synthase help to coordinate ligand binding, catalysis, and substrate channeling*. J Mol Biol, 2013. **425**(9): p. 1527-45.
53. Doucet, N., et al., *NMR investigation of Tyr105 mutants in TEM-1 beta-lactamase: dynamics are correlated with function*. J Biol Chem, 2007. **282**(29): p. 21448-59.

54. Eisenmesser, E.Z., et al., *Intrinsic dynamics of an enzyme underlies catalysis*. Nature, 2005. **438**(7064): p. 117-21.
55. Labeikovsky, W., et al., *Structure and dynamics of pin1 during catalysis by NMR*. J Mol Biol, 2007. **367**(5): p. 1370-81.
56. Berlow, R.B., T.I. Igumenova, and J.P. Loria, *Value of a hydrogen bond in triosephosphate isomerase loop motion*. Biochemistry, 2007. **46**(20): p. 6001-10.
57. Kempf, J.G., et al., *Dynamic requirements for a functional protein hinge*. J Mol Biol, 2007. **368**(1): p. 131-49.
58. Massi, F., C. Wang, and A.G. Palmer, 3rd, *Solution NMR and computer simulation studies of active site loop motion in triosephosphate isomerase*. Biochemistry, 2006. **45**(36): p. 10787-94.
59. Rozovsky, S., et al., *Solution-state NMR investigations of triosephosphate isomerase active site loop motion: ligand release in relation to active site loop dynamics*. J Mol Biol, 2001. **310**(1): p. 271-80.
60. Korzhnev, D.M., et al., *Off-resonance R1rho relaxation outside of the fast exchange limit: an experimental study of a cavity mutant of T4 lysozyme*. J Biomol NMR, 2003. **26**(1): p. 39-48.
61. Mulder, F.A., et al., *Measurement of slow (micro-s-ms) time scale dynamics in protein side chains by (15)N relaxation dispersion NMR spectroscopy: application to Asn and Gln residues in a cavity mutant of T4 lysozyme*. J Am Chem Soc, 2001. **123**(5): p. 967-75.
62. Vallurupalli, P. and L.E. Kay, *Complementarity of ensemble and single-molecule measures of protein motion: a relaxation dispersion NMR study of an enzyme complex*. Proc Natl Acad Sci U S A, 2006. **103**(32): p. 11910-5.
63. Doucet, N., E.D. Watt, and J.P. Loria, *The flexibility of a distant loop modulates active site motion and product release in ribonuclease A*. Biochemistry, 2009. **48**(30): p. 7160-8.
64. Doucet, N., et al., *Alteration of hydrogen bonding in the vicinity of histidine 48 disrupts millisecond motions in RNase A*. Biochemistry, 2011. **50**(10): p. 1723-30.
65. Kunitz, M., *Crystalline Ribonuclease*. J Gen Physiol, 1940. **24**(1): p. 15-32.
66. Kunitz, M., *Isolation from Beef Pancreas of a Crystalline Protein Possessing Ribonuclease Activity*. Science, 1939. **90**(2327): p. 112-3.
67. Hirs, C.H., S. Moore, and W.H. Stein, *The sequence of the amino acid residues in performic acid-oxidized ribonuclease*. J Biol Chem, 1960. **235**: p. 633-47.
68. Smyth, D.G., W.H. Stein, and S. Moore, *The sequence of amino acid residues in bovine pancreatic ribonuclease: revisions and confirmations*. J Biol Chem, 1963. **238**: p. 227-34.
69. Kartha, G., J. Bello, and D. Harker, *Tertiary structure of ribonuclease*. Nature, 1967. **213**(5079): p. 862-5.
70. Anfinsen, C.B., *Principles that govern the folding of protein chains*. Science, 1973. **181**(4096): p. 223-30.
71. Moore, S. and W.H. Stein, *Chemical structures of pancreatic ribonuclease and deoxyribonuclease*. Science, 1973. **180**(4085): p. 458-64.

72. Merrifield, B., *Solid phase synthesis*. Science, 1986. **232**(4748): p. 341-7.
73. Raines, R.T., *Ribonuclease A*. Chem Rev, 1998. **98**(3): p. 1045-1066.
74. Pares, X., et al., *Structure and function of ribonuclease A binding subsites*. Essays Biochem, 1991. **26**: p. 89-103.
75. Wlodawer, A., et al., *Structure of phosphate-free ribonuclease A refined at 1.26 Å*. Biochemistry, 1988. **27**(8): p. 2705-17.
76. Usher, D.A., D.I. Richardson, Jr., and F. Eckstein, *Absolute stereochemistry of the second step of ribonuclease action*. Nature, 1970. **228**(5272): p. 663-5.
77. Thompson, J.E. and R.T. Raines, *Value of general Acid-base catalysis to ribonuclease a*. J Am Chem Soc, 1994. **116**(12): p. 5467-8.
78. Gagne, D. and N. Doucet, *Structural and functional importance of local and global conformational fluctuations in the RNase A superfamily*. FEBS J, 2013. **280**(22): p. 5596-607.
79. Irie, M., et al., *Kinetic studies on the cleavage of oligouridylic acids and poly U by bovine pancreatic ribonuclease A*. J Biochem, 1984. **96**(1): p. 89-96.
80. Record, M.T., Jr., J.H. Ha, and M.A. Fisher, *Analysis of equilibrium and kinetic measurements to determine thermodynamic origins of stability and specificity and mechanism of formation of site-specific complexes between proteins and helical DNA*. Methods Enzymol, 1991. **208**: p. 291-343.
81. Jensen, D.E. and P.H. von Hippel, *DNA "melting" proteins. I. Effects of bovine pancreatic ribonuclease binding on the conformation and stability of DNA*. J Biol Chem, 1976. **251**(22): p. 7198-214.
82. Record, M.T., Jr., M.L. Lohman, and P. De Haseth, *Ion effects on ligand-nucleic acid interactions*. J Mol Biol, 1976. **107**(2): p. 145-58.
83. Nogues, M.V., M. Vilanova, and C.M. Cuchillo, *Bovine pancreatic ribonuclease A as a model of an enzyme with multiple substrate binding sites*. Biochim Biophys Acta, 1995. **1253**(1): p. 16-24.
84. Cuchillo, C.M., et al., *The exo- or endonucleolytic preference of bovine pancreatic ribonuclease A depends on its subsites structure and on the substrate size*. Protein Sci, 2002. **11**(1): p. 117-28.
85. Fisher, B.M., J.E. Grilley, and R.T. Raines, *A new remote subsite in ribonuclease A*. J Biol Chem, 1998. **273**(51): p. 34134-8.
86. delCardayre, S.B. and R.T. Raines, *Structural determinants of enzymatic processivity*. Biochemistry, 1994. **33**(20): p. 6031-7.
87. Beintema, J.J. and R.G. Kleineidam, *The ribonuclease A superfamily: general discussion*. Cell Mol Life Sci, 1998. **54**(8): p. 825-32.
88. Dey, P., et al., *Role of unique basic residues of human pancreatic ribonuclease in its catalysis and structural stability*. Biochem Biophys Res Commun, 2007. **360**(4): p. 809-14.
89. Sikriwal, D., D. Seth, and J.K. Batra, *Role of catalytic and non-catalytic subsite residues in ribonuclease activity of human eosinophil-derived neurotoxin*. Biol Chem, 2009. **390**(3): p. 225-34.

90. Sikriwal, D., et al., *Human eosinophil-derived neurotoxin: involvement of a putative non-catalytic phosphate-binding subsite in its catalysis*. Mol Cell Biochem, 2007. **303**(1-2): p. 175-81.
91. Singh, A. and J.K. Batra, *Role of unique basic residues in cytotoxic, antibacterial and antiparasitic activities of human eosinophil cationic protein*. Biol Chem, 2011. **392**(4): p. 337-46.
92. Boix, E., et al., *Kinetic and product distribution analysis of human eosinophil cationic protein indicates a subsite arrangement that favors exonuclease-type activity*. J Biol Chem, 1999. **274**(22): p. 15605-14.
93. Hofsteenge, J., A. Vicentini, and O. Zelenko, *Ribonuclease 4, an evolutionarily highly conserved member of the superfamily*. Cell Mol Life Sci, 1998. **54**(8): p. 804-10.
94. Shapiro, R., *Structural features that determine the enzymatic potency and specificity of human angiogenin: threonine-80 and residues 58-70 and 116-123*. Biochemistry, 1998. **37**(19): p. 6847-56.
95. Rosenberg, H.F. and K.D. Dyer, *Molecular cloning and characterization of a novel human ribonuclease (RNase k6): increasing diversity in the enlarging ribonuclease gene family*. Nucleic Acids Res, 1996. **24**(18): p. 3507-13.
96. Cho, S., J.J. Beintema, and J. Zhang, *The ribonuclease A superfamily of mammals and birds: identifying new members and tracing evolutionary histories*. Genomics, 2005. **85**(2): p. 208-20.
97. Sorrentino, S., *The eight human "canonical" ribonucleases: molecular diversity, catalytic properties, and special biological actions of the enzyme proteins*. FEBS Lett, 2010. **584**(11): p. 2194-200.
98. Goo, S.M. and S. Cho, *The expansion and functional diversification of the mammalian ribonuclease a superfamily epitomizes the efficiency of multigene families at generating biological novelty*. Genome Biol Evol, 2013. **5**(11): p. 2124-40.
99. Iwama, M., et al., *Purification and properties of human urine ribonucleases*. J Biochem, 1981. **89**(4): p. 1005-16.
100. Yasuda, T., et al., *Two distinct secretory ribonucleases from human cerebrum: purification, characterization and relationships to other ribonucleases*. Biochem J, 1993. **296** ( Pt 3): p. 617-25.
101. De Prisco, R., et al., *A ribonuclease from human seminal plasma active on double-stranded RNA*. Biochim Biophys Acta, 1984. **788**(3): p. 356-63.
102. Mizuta, K., et al., *Purification and characterization of three ribonucleases from human kidney: comparison with urine ribonucleases*. Arch Biochem Biophys, 1990. **281**(1): p. 144-51.
103. Beintema, J.J., et al., *The amino acid sequence of human pancreatic ribonuclease*. Anal Biochem, 1984. **136**(1): p. 48-64.
104. Beintema, J.J., et al., *Differences in glycosylation pattern of human secretory ribonucleases*. Biochem J, 1988. **255**(2): p. 501-5.
105. Gaur, D., S. Swaminathan, and J.K. Batra, *Interaction of human pancreatic ribonuclease with human ribonuclease inhibitor. Generation of inhibitor-resistant cytotoxic variants*. J Biol Chem, 2001. **276**(27): p. 24978-84.

106. D'Alessio, G., *New and cryptic biological messages from RNases*. Trends Cell Biol, 1993. **3**(4): p. 106-9.
107. Sorrentino, S. and M. Libonati, *Structure-function relationships in human ribonucleases: main distinctive features of the major RNase types*. FEBS Lett, 1997. **404**(1): p. 1-5.
108. Yasuda, T., et al., *Purification and characterization of a ribonuclease from human spleen. Immunological and enzymological comparison with nonsecretory ribonuclease from human urine*. Eur J Biochem, 1990. **191**(2): p. 523-9.
109. Gleich, G.J., et al., *Biochemical and functional similarities between human eosinophil-derived neurotoxin and eosinophil cationic protein: homology with ribonuclease*. Proc Natl Acad Sci U S A, 1986. **83**(10): p. 3146-50.
110. Sorrentino, S., G.K. Tucker, and D.G. Glitz, *Purification and characterization of a ribonuclease from human liver*. J Biol Chem, 1988. **263**(31): p. 16125-31.
111. Shapiro, R. and B.L. Vallee, *Interaction of human placental ribonuclease with placental ribonuclease inhibitor*. Biochemistry, 1991. **30**(8): p. 2246-55.
112. Sorrentino, S. and M. Libonati, *Human pancreatic-type and nonpancreatic-type ribonucleases: a direct side-by-side comparison of their catalytic properties*. Arch Biochem Biophys, 1994. **312**(2): p. 340-8.
113. Sorrentino, S. and D.G. Glitz, *Ribonuclease activity and substrate preference of human eosinophil cationic protein (ECP)*. FEBS Lett, 1991. **288**(1-2): p. 23-6.
114. Barker, R.L., et al., *Eosinophil cationic protein cDNA. Comparison with other toxic cationic proteins and ribonucleases*. J Immunol, 1989. **143**(3): p. 952-5.
115. Domachowske, J.B., et al., *Recombinant human eosinophil-derived neurotoxin/RNase 2 functions as an effective antiviral agent against respiratory syncytial virus*. J Infect Dis, 1998. **177**(6): p. 1458-64.
116. Domachowske, J.B., et al., *Evolution of antiviral activity in the ribonuclease A gene superfamily: evidence for a specific interaction between eosinophil-derived neurotoxin (EDN/RNase 2) and respiratory syncytial virus*. Nucleic Acids Res, 1998. **26**(23): p. 5327-32.
117. Rosenberg, H.F. and J.B. Domachowske, *Eosinophils, eosinophil ribonucleases, and their role in host defense against respiratory virus pathogens*. J Leukoc Biol, 2001. **70**(5): p. 691-8.
118. Rugeles, M.T., et al., *Ribonuclease is partly responsible for the HIV-1 inhibitory effect activated by HLA alloantigen recognition*. AIDS, 2003. **17**(4): p. 481-6.
119. Slifman, N.R., et al., *Ribonuclease activity associated with human eosinophil-derived neurotoxin and eosinophil cationic protein*. J Immunol, 1986. **137**(9): p. 2913-7.
120. Pulido, D., et al., *Towards the rational design of antimicrobial proteins: single point mutations can switch on bactericidal and agglutinating activities on the RNase A superfamily lineage*. FEBS J, 2013. **280**(22): p. 5841-52.
121. Carreras, E., et al., *Both aromatic and cationic residues contribute to the membrane-lytic and bactericidal activity of eosinophil cationic protein*. Biochemistry, 2003. **42**(22): p. 6636-6644.

122. Carreras, E., et al., *Surface-exposed amino acids of eosinophil cationic protein play a critical role in the inhibition of mammalian cell proliferation*. Mol Cell Biochem, 2005. **272**(1-2): p. 1-7.
123. Rosenberg, H.F., *Recombinant human eosinophil cationic protein. Ribonuclease activity is not essential for cytotoxicity*. J Biol Chem, 1995. **270**(14): p. 7876-81.
124. Sorrentino, S., *Human extracellular ribonucleases: multiplicity, molecular diversity and catalytic properties of the major RNase types*. Cell Mol Life Sci, 1998. **54**(8): p. 785-94.
125. Shapiro, R., et al., *Isolation and characterization of a human colon carcinoma-secreted enzyme with pancreatic ribonuclease-like activity*. Biochemistry, 1986. **25**(23): p. 7255-64.
126. Zhou, H.M. and D.J. Strydom, *The amino acid sequence of human ribonuclease 4, a highly conserved ribonuclease that cleaves specifically on the 3' side of uridine*. Eur J Biochem, 1993. **217**(1): p. 401-10.
127. Fett, J.W., et al., *Isolation and characterization of angiogenin, an angiogenic protein from human carcinoma cells*. Biochemistry, 1985. **24**(20): p. 5480-6.
128. Shapiro, R., et al., *Isolation of angiogenin from normal human plasma*. Biochemistry, 1987. **26**(16): p. 5141-6.
129. Acharya, K.R., et al., *Crystal structure of human angiogenin reveals the structural basis for its functional divergence from ribonuclease*. Proc Natl Acad Sci U S A, 1994. **91**(8): p. 2915-9.
130. Russo, N., et al., *Role of glutamine-117 in the ribonucleolytic activity of human angiogenin*. Proc Natl Acad Sci U S A, 1994. **91**(8): p. 2920-4.
131. Leonidas, D.D., et al., *Refined crystal structures of native human angiogenin and two active site variants: implications for the unique functional properties of an enzyme involved in neovascularisation during tumour growth*. J Mol Biol, 1999. **285**(3): p. 1209-33.
132. Hooper, L.V., et al., *Angiogenins: a new class of microbicidal proteins involved in innate immunity*. Nat Immunol, 2003. **4**(3): p. 269-73.
133. Simanski, M., et al., *Antimicrobial RNases in cutaneous defense*. J Innate Immun, 2012. **4**(3): p. 241-7.
134. Chavali, G.B., et al., *The crystal structure of human angiogenin in complex with an antitumor neutralizing antibody*. Structure, 2003. **11**(7): p. 875-85.
135. Zhang, J., K.D. Dyer, and H.F. Rosenberg, *Human RNase 7: a new cationic ribonuclease of the RNase A superfamily*. Nucleic Acids Res, 2003. **31**(2): p. 602-7.
136. Huang, Y.C., et al., *The flexible and clustered lysine residues of human ribonuclease 7 are critical for membrane permeability and antimicrobial activity*. J Biol Chem, 2007. **282**(7): p. 4626-33.
137. Harder, J. and J.M. Schroder, *RNase 7, a novel innate immune defense antimicrobial protein of healthy human skin*. J Biol Chem, 2002. **277**(48): p. 46779-84.
138. Schuster, C., et al., *Prenatal human skin expresses the antimicrobial peptide RNase 7*. Arch Dermatol Res, 2013. **305**(6): p. 545-9.

139. Torrent, M., et al., *Comparison of human RNase 3 and RNase 7 bactericidal action at the Gram-negative and Gram-positive bacterial cell wall*. FEBS J, 2010. **277**(7): p. 1713-25.
140. Rudolph, B., et al., *Identification of RNase 8 as a novel human antimicrobial protein*. Antimicrob Agents Chemother, 2006. **50**(9): p. 3194-6.
141. Zhang, J., K.D. Dyer, and H.F. Rosenberg, *RNase 8, a novel RNase A superfamily ribonuclease expressed uniquely in placenta*. Nucleic Acids Res, 2002. **30**(5): p. 1169-75.
142. Freyer, M.W. and A.E. Lewis, *Isothermal Titration Calorimetry: Experimental Design, Data Analysis, and Probing Macromolecule/Ligand Binding and Kinetic Interactions*. Methods in cell biology, 2008. **86**: p. 79-113.
143. Ramanathan, A. and P.K. Agarwal, *Evolutionarily conserved linkage between enzyme fold, flexibility, and catalysis*. PLoS Biol, 2011. **9**(11): p. e1001193.
144. Kishan, R.V., et al., *Water-dependent domain motion and flexibility in ribonuclease A and the invariant features in its hydration shell. An X-ray study of two low-humidity crystal forms of the enzyme*. Acta Crystallogr D Biol Crystallogr, 1995. **51**(Pt 5): p. 703-10.
145. Vitagliano, L., et al., *Reversible substrate-induced domain motions in ribonuclease A*. Proteins, 2002. **46**(1): p. 97-104.
146. Rasmussen, B.F., et al., *Crystalline ribonuclease A loses function below the dynamical transition at 220 K*. Nature, 1992. **357**(6377): p. 423-4.
147. Luebbing, E.K., et al., *Conservation of functionally important global motions in an enzyme superfamily across varying quaternary structures*. J Mol Biol, 2012. **423**(5): p. 831-46.
148. Zorzini, V., et al., *Structural and biophysical characterization of Staphylococcus aureus SaMazF shows conservation of functional dynamics*. Nucleic Acids Res, 2014. **42**(10): p. 6709-25.
149. Warren, N., et al., *Comparison of the intrinsic dynamics of aminoacyl-tRNA synthetases*. Protein J, 2014. **33**(2): p. 184-98.
150. Boix, E. and M.V. Nogues, *Mammalian antimicrobial proteins and peptides: overview on the RNase A superfamily members involved in innate host defence*. Mol Biosyst, 2007. **3**(5): p. 317-35.
151. Arnold, U., *Aspects of the cytotoxic action of ribonucleases*. Curr Pharm Biotechnol, 2008. **9**(3): p. 161-8.
152. Leland, P.A., et al., *Ribonuclease A variants with potent cytotoxic activity*. Proc Natl Acad Sci U S A, 1998. **95**(18): p. 10407-12.
153. Leland, P.A. and R.T. Raines, *Cancer chemotherapy--ribonucleases to the rescue*. Chem Biol, 2001. **8**(5): p. 405-13.
154. Lee, J.E., et al., *Structural basis for catalysis by onconase*. J Mol Biol, 2008. **375**(1): p. 165-77.
155. Ellis, G.A., et al., *Ribonuclease-Activated Cancer Prodrug*. ACS Med Chem Lett, 2012. **3**(4): p. 268-272.

156. Gagne, D. and N. Doucet, *Structural and functional importance of local and global conformational fluctuations in the RNase A superfamily*. FEBS J, 2013.
157. Torrent, J., et al., *Pressure versus temperature unfolding of ribonuclease A: an FTIR spectroscopic characterization of 10 variants at the carboxy-terminal site*. Protein Sci, 2001. **10**(4): p. 725-34.
158. Chatani, E. and R. Hayashi, *Functional and structural roles of constituent amino acid residues of bovine pancreatic ribonuclease A*. J Biosci Bioeng, 2001. **92**(2): p. 98-107.
159. Leonidas, D.D., et al., *The binding of 3'-N-piperidine-4-carboxyl-3'-deoxy-ara-uridine to ribonuclease A in the crystal*. Bioorg Med Chem, 2006. **14**(17): p. 6055-64.
160. Cole, R. and J.P. Loria, *Evidence for flexibility in the function of ribonuclease A*. Biochemistry, 2002. **41**(19): p. 6072-81.
161. Smith, B.D. and R.T. Raines, *Genetic selection for critical residues in ribonucleases*. J Mol Biol, 2006. **362**(3): p. 459-78.
162. Watt, E.D., et al., *The mechanism of rate-limiting motions in enzyme function*. Proc Natl Acad Sci U S A, 2007. **104**(29): p. 11981-6.
163. Venkitakrishnan, R.P., et al., *Conformational changes in the active site loops of dihydrofolate reductase during the catalytic cycle*. Biochemistry, 2004. **43**(51): p. 16046-55.
164. Benkovic, S.J. and S. Hammes-Schiffer, *Biochemistry. Enzyme motions inside and out*. Science, 2006. **312**(5771): p. 208-9.
165. Wong, K.F., et al., *Impact of distal mutations on the network of coupled motions correlated to hydride transfer in dihydrofolate reductase*. Proc Natl Acad Sci U S A, 2005. **102**(19): p. 6807-12.
166. Gagne, D., et al., *Conservation of flexible residue clusters among structural and functional enzyme homologues*. J Biol Chem, 2012. **287**(53): p. 44289-300.
167. Ye, Y. and A. Godzik, *FATCAT: a web server for flexible structure comparison and structure similarity searching*. Nucleic Acids Res, 2004. **32**(Web Server issue): p. W582-5.
168. Doucet, N. and J.N. Pelletier, *Simulated annealing exploration of an active-site tyrosine in TEM-1 beta-lactamase suggests the existence of alternate conformations*. Proteins, 2007. **69**(2): p. 340-8.
169. Bhabha, G., et al., *A dynamic knockout reveals that conformational fluctuations influence the chemical step of enzyme catalysis*. Science, 2011. **332**(6026): p. 234-8.
170. Carnevale, V., et al., *Convergent dynamics in the protease enzymatic superfamily*. J Am Chem Soc, 2006. **128**(30): p. 9766-72.
171. Boix, E., et al., *The antipathogen activities of eosinophil cationic protein*. Curr Pharm Biotechnol, 2008. **9**(3): p. 141-52.
172. Halabi, N., et al., *Protein sectors: evolutionary units of three-dimensional structure*. Cell, 2009. **138**(4): p. 774-86.
173. Merlino, A., et al., *Global and local motions in ribonuclease A: a molecular dynamics study*. Biopolymers, 2002. **65**(4): p. 274-83.



174. Terzyan, S.S., et al., *The three-dimensional structure of human RNase 4, unliganded and complexed with d(Up), reveals the basis for its uridine selectivity.* J Mol Biol, 1999. **285**(1): p. 205-14.
175. Doucet, N., et al., *The crystal structure of ribonuclease A in complex with thymidine-3'-monophosphate provides further insight into ligand binding.* Proteins, 2010. **78**(11): p. 2459-68.
176. Borkakoti, N., *The active site of ribonuclease A from the crystallographic studies of ribonuclease-A-inhibitor complexes.* Eur J Biochem, 1983. **132**(1): p. 89-94.
177. Guindon, S., et al., *New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0.* Syst Biol, 2010. **59**(3): p. 307-21.
178. Landau, M., et al., *ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures.* Nucleic Acids Res, 2005. **33**(Web Server issue): p. W299-302.
179. Gagne, D. and N. Doucet, *Sequence-specific backbone (1)H, (13)C, and (15)N resonance assignments of human ribonuclease 4.* Biomol NMR Assign, 2015. **9**(1): p. 181-5.
180. Fisher, B.M., J.H. Ha, and R.T. Raines, *Coulombic forces in protein-RNA interactions: binding and cleavage by ribonuclease A and variants at Lys7, Arg10, and Lys66.* Biochemistry, 1998. **37**(35): p. 12121-32.
181. Grzesiek, S., et al., *The CD4 determinant for downregulation by HIV-1 Nef directly binds to Nef. Mapping of the Nef binding surface by NMR.* Biochemistry, 1996. **35**(32): p. 10256-61.
182. Pearce, F.G., M.D.W. Griffin, and J.A. Gerrard, *Does domain swapping improve the stability of RNase A?* Biochemical and Biophysical Research Communications, 2009. **382**(1): p. 114-118.
183. Fisher, B.M., L.W. Schultz, and R.T. Raines, *Coulombic effects of remote subsites on the active site of ribonuclease A.* Biochemistry, 1998. **37**(50): p. 17386-401.
184. Hofsteenge, J., et al., *A single amino acid substitution changes ribonuclease 4 from a uridine-specific to a cytidine-specific enzyme.* Biochemistry, 1998. **37**(26): p. 9250-7.
185. Larkin, M.A., et al., *Clustal W and Clustal X version 2.0.* Bioinformatics, 2007. **23**(21): p. 2947-8.
186. Boehr, D.D., *Promiscuity in protein-RNA interactions: conformational ensembles facilitate molecular recognition in the spliceosome: conformational diversity in U2AF(6)(5) facilitates binding to diverse RNA sequences.* Bioessays, 2012. **34**(3): p. 174-80.
187. Park, C., et al., *Fast, facile, hypersensitive assays for ribonucleolytic activity.* Methods Enzymol, 2001. **341**: p. 81-94.
188. Mulder, F.A., et al., *Studying excited states of proteins by NMR spectroscopy.* Nat Struct Biol, 2001. **8**(11): p. 932-5.
189. Merlino, A., et al., *Subtle functional collective motions in pancreatic-like ribonucleases: from ribonuclease A to angiogenin.* Proteins, 2003. **53**(1): p. 101-10.

190. Harper, J.W. and B.L. Vallee, *A covalent angiogenin/ribonuclease hybrid with a fourth disulfide bond generated by regional mutagenesis*. *Biochemistry*, 1989. **28**(4): p. 1875-84.
191. Holloway, D.E., et al., *Crystallographic studies on structural features that determine the enzymatic specificity and potency of human angiogenin: Thr44, Thr80, and residues 38-41*. *Biochemistry*, 2004. **43**(5): p. 1230-41.