

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

**CONSERVATION ET MODULATION DE SECTEURS DYNAMIQUES
FONCTIONNELS ENTRE RIBONUCLÉASES HOMOLOGUES**

Par
Laurie-Anne Charest

Mémoire présenté pour l'obtention du grade de
Maître ès sciences (M.Sc.) en microbiologie appliquée

Jury d'évaluation

Président du jury et
examineur interne

Pr David Chatenet, Ph.D.
INRS-Institut Armand-Frappier

Examineur externe

Pr Guillaume Lamoureux, Ph.D.
Department of Chemistry and Biochemistry
Concordia University, Montréal, QC

Directeur de recherche

Pr Nicolas Doucet, Ph.D.
INRS-Institut Armand-Frappier

REMERCIEMENTS

Au cours de mon passage à l'INRS-Institut Armand-Frappier, j'ai eu la chance de rencontrer plusieurs personnes qui ont largement contribué à la réalisation de ce mémoire et qui ont fait en sorte que cette expérience universitaire fut si enrichissante, autant sur le plan professionnel que personnel.

Je souhaite tout d'abord adresser mes plus sincères remerciements à mon directeur de recherche, Nicolas Doucet, qui s'est toujours montré à l'écoute et très disponible tout au long de la maîtrise. Sa passion et sa vision de la science ont été de grandes sources d'inspiration.

Je tiens à exprimer ma gratitude envers mes collègues du laboratoire Doucet : Benjamin, Hélène, Nhung, Philippe et spécialement Donald, qui m'a généreusement transmis ses connaissances concernant la purification des protéines et qui a répondu à mes innombrables questions. Mes remerciements s'adressent également à mes collègues de l'Institut, notamment ceux du laboratoire Guertin et Déziel. En plus de me guider au niveau scientifique, de belles amitiés sont nées à travers ces échanges. Un merci spécial à Audrey-Anne, Fabrice, Marie-Christine et Amélie pour le support moral, les fous rires, les anecdotes divertissantes et les fameux soupers.

Bien sûr, ce mémoire n'aurait pu voir le jour sans le soutien des organismes subventionnaires tels que la Fondation universitaire Armand-Frappier de l'INRS. Je remercie également les professeurs Éric Déziel et Claude Guertin pour leur contribution financière.

Enfin, je tiens à mentionner l'important soutien émotionnel et psychologique offert par tous mes amis et les membres de ma famille. Chacun à leur manière, ils ont contribué à mon développement personnel et à mon cheminement scientifique. J'exprime toute ma reconnaissance à mes parents et à mon frère pour leur appui et leur support inconditionnels à plusieurs niveaux. Merci à mon père, qui m'a inculqué la curiosité scientifique et dont la passion pour les sciences est si inspirante.

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 à 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}N -CPMG), d'analyse modelfree et de titrage RMN pour capturer et comparer le rôle de la flexibilité conformationnelle entre deux homologues structuraux de la famille des ribonucléases pancréatiques : la RNase A et la protéine cationique des éosinophiles humaine (ECP, ou RNase 3). En plus de conserver les mêmes résidus catalytiques et la même structure tridimensionnelle, les deux homologues montrent des secteurs de résidus dynamiques similaires à l'échelle de temps de la milliseconde, suggérant que la flexibilité conformationnelle peut être conservée parmi des protéines structurellement analogues qui affichent néanmoins une faible identité de séquence. Notre travail démontre également que la flexibilité conformationnelle de ECP peut être dynamiquement et fonctionnellement reproduite dans la structure de la RNase A, ce que nous démontrons par la création d'une chimère hybride entre les deux protéines. Ces résultats soutiennent l'hypothèse que la flexibilité conformationnelle est en partie nécessaire pour la fonction catalytique chez des enzymes homologues, soulignant en outre l'importance des secteurs dynamiques de résidus dans l'organisation structurale des protéines.

TABLE DES MATIÈRES

Remerciements.....	iii
Résumé.....	iv
Table des matières.....	v
Liste des figures.....	vi
Liste des abréviations.....	vii
CHAPITRE 1 : MISE EN CONTEXTE.....	8
1.1. La dynamique des protéines et l'allostérie	9
1.2. La régulation allostérique	10
1.3. Hypothèses et objectifs	12
1.4. Les ribonucléases pancréatiques : un système modèle	14
1.5. Analyse dynamique par dispersion de relaxation RMN (¹⁵ N-CPMG)	19
1.6. Titrage RMN.....	22
CHAPITRE 2 : ARTICLE SCIENTIFIQUE.....	25
2.1. Présentation de l'article scientifique.....	26
2.2. Contribution des auteurs	26
2.3. Résumé.....	27
2.4. Article scientifique.....	28
2.5. Conclusions et perspectives	69
RÉFÉRENCES.....	72

LISTE DES FIGURES

Figure 1. Échelle de temps des mouvements moléculaires observés chez les protéines	9
Figure 2. Modèle allostérique de contrôle dynamique chez les protéines	11
Figure 3. Mécanisme catalytique de la famille des RNases.....	15
Figure 4. Structure tridimensionnelle de RNases pancréatiques retrouvées chez les vertébrés ..	16
Figure 5. Comparaison structurale entre la RNase A et ECP	18
Figure 6. Principe de la dispersion de relaxation RMN (^{15}N -CPMG).....	22
Figure 7. Titrage RMN d'un ligand dans une protéine d'intérêt.....	24

LISTE DES ABRÉVIATIONS

CPMG	Carr-Purcell-Meiboom-Gill
ECP	Eosinophil Cationic Protein (RNase 3, protéine cationique des éosinophiles de l'humain)
HSQC	Heteronuclear Single-Quantum Coherence
ms	Milliseconde (10^{-3} s)
ns	Nanoseconde (10^{-9} s)
ps	Picoseconde (10^{-12} s)
R_1	Relaxation longitudinale
R_2	Relaxation transversale
RMN	Résonance magnétique nucléaire
RNase A	Ribonucléase A bovine
TROSY	Transverse Relaxation-Optimized Spectroscopy
μ s	Microseconde (10^{-6} s)
VIH	Virus de l'immunodéficience humaine

CHAPITRE 1 : MISE EN CONTEXTE

1.1. La dynamique des protéines et l'allostérie

Les jolies images qu'offrent les structures cristallines des protéines ne représentent en fait qu'une fraction de la réalité. En effet, cette vision 'statique' des protéines cache le fait que ces macromolécules bioactives sont en fait des molécules flexibles qui échantillonnent en permanence différents états conformationnels à plusieurs échelles de temps (Figure 1)¹. De plus, il est désormais reconnu que les mouvements des protéines ne sont pas exclusivement aléatoires, mais qu'ils sont souvent essentiels au contrôle du repliement, des interactions protéine-protéine (ou protéine-ligand), à la liaison de substrat(s), à l'événement catalytique, au relâchement des produits, etc. À cet effet, les mouvements des protéines qui se produisent à l'échelle de temps de la milliseconde (ms) sont particulièrement intéressants puisque leur vitesse coïncide avec la plupart des processus biologiques essentiels à la liaison des ligands et/ou à la vitesse catalytique des enzymes (Figure 1)¹⁻⁴. La modification de ces réseaux dynamiques conservés à l'échelle de la milliseconde a démontré un effet perturbateur sur l'événement catalytique de plusieurs enzymes⁵. Par exemple, nous savons que des mutations ou la liaison de ligands peuvent affecter la redistribution d'ensembles de populations dynamiques hétérogènes en solution, perturbant ainsi le ratio d'espèces conformationnelles actives vs inactives, ce qui affecte indubitablement la catalyse enzymatique⁶.

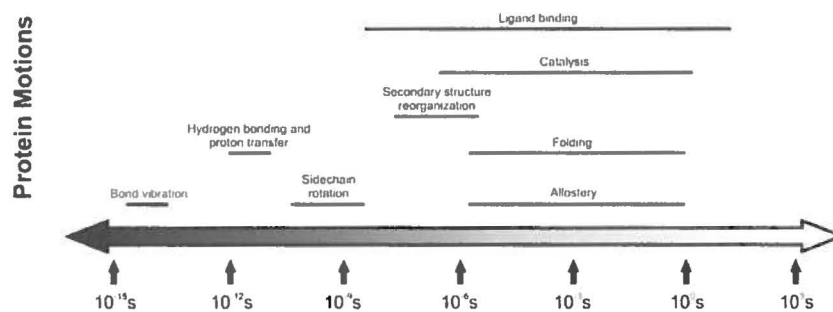


Figure 1. Échelle de temps des mouvements moléculaires observés chez les protéines. Une quantité importante de mouvements biologiquement pertinents au niveau cellulaire se produisent à l'échelle de temps μ s-ms. Figure tirée de réf.¹.

Les analyses désormais célèbres de Monod, Wyman et Changeux publiées dans *J. Mol. Biol.* en 1965 ont permis de définir le concept de l'allostérie, qui se résume à sa plus simple expression par l'effet ressenti au site X d'une protéine ou d'une enzyme lors de la liaison d'un ligand (nommé effecteur) à un site Y proximal ou distal⁷. Ces effets allostériques, qui impliquent la plupart du temps des changements conformationnels locaux ou globaux de la protéine, ont longtemps été considérés spécifiques à certains systèmes bien particuliers. Or, il a récemment été suggéré que toutes les protéines non-fibreuses puissent se comporter à titre de potentielles cibles allostériques dont la fonction s'appuie sur la dynamique de résidus spécifiques ou de réseaux conformationnels conservés liés à l'activité biologique de la protéine^{6,8,9}. Cela remet en question la définition classique de l'allostérie et suggère que toutes les surfaces protéiques pourraient éventuellement être soumises à un contrôle dynamique, une affirmation récemment supportée expérimentalement chez une β -lactamase permutée¹⁰.

1.2. La régulation allostérique

La régulation et le contrôle allostérique de l'activité des enzymes par de petites molécules effectrices représente une approche intéressante pour la conception de nouveaux médicaments¹¹. En effet, au lieu de concentrer les efforts pharmacologiques sur un seul site effecteur pour la conception de petites molécules (site actif, ou cavité de liaison), les inhibiteurs allostériques permettent d'offrir des sites de régulation supplémentaires pour aider à soulager l'émergence de variants protéiques résistants aux traitements médicamenteux. 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 2)^{11, 12}. Par conséquent, des efforts de recherche récents ont été consacrés au contrôle de mouvements visant à affecter la fonction enzymatique à l'aide d'une combinaison d'effecteurs allostériques et d'inhibiteurs compétitifs classiques, donnant lieu à un certain nombre de petites molécules actuellement sur le marché ou en étude clinique^{11, 13, 14}. Entre autres, les composés TIBO, qui sont des inhibiteurs non-nucléosidiques de la transcriptase inverse du VIH, inhibent de façon non-compétitive la polymérisation de l'ADN lors de la réplication du VIH en liant une poche distante du site actif principal de l'enzyme¹⁵.

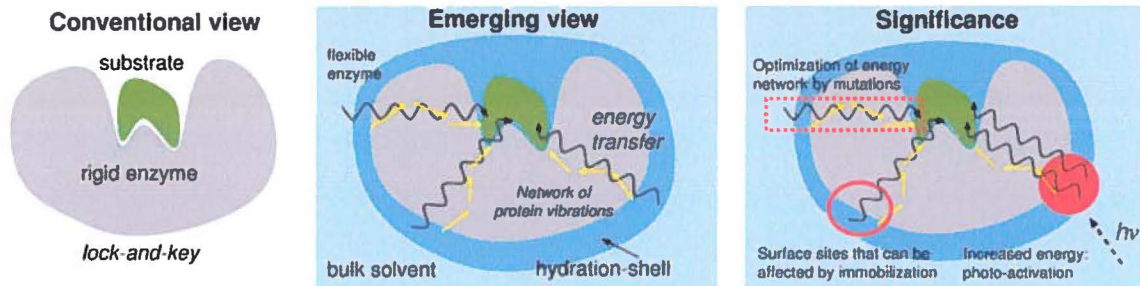


Figure 2. Modèle allostérique de contrôle dynamique chez les protéines. Contrairement au modèle conventionnel, dans lequel une seule conformation productive et rigide est retrouvée pour l'enzyme et le substrat (à gauche), le modèle allostérique illustre l'existence d'une gamme de mouvements pertinents pour l'activité catalytique d'une enzyme (au centre). Les mouvements concertés jouent un rôle dans l'activité par transfert énergétique aux atomes réactifs de l'enzyme et/ou des substrats¹⁶. Ce modèle offre de nouvelles possibilités pour la modulation allostérique de l'activité enzymatique, notamment *via* l'interruption ou l'amélioration du transfert d'énergie altérant la liaison ou la réactivité (à droite). Figure tirée de réf.¹⁷.

Or, la capacité d'identifier ces sites allostériques sur les protéines demeure un défi de taille, notamment en raison des limites physiques des techniques expérimentales actuellement employées pour étudier la dynamique moléculaire. Le design de nouveaux médicaments est traditionnellement axé sur la recherche de composés qui se lient au site actif principal d'une cible protéique donnée, fournissant une inhibition compétitive au ligand naturel. Récemment, il est devenu clair que cette méthodologie limite considérablement le spectre et l'étendue de la découverte de médicaments potentiellement prometteurs¹¹. Par conséquent, la recherche de sites alternatifs qui permettent la modulation de l'activité—notamment en offrant un contrôle distal par allostérie (Figure 2)—augmente la variété de groupements chimiques fonctionnels contre les cibles cliniques. De plus, une quantité importante de composés thérapeutiques originalement rejetés par manque d'affinité au site actif de cibles protéiques pourraient désormais s'avérer prometteurs à titre de médicaments ciblant de nouveaux sites allostériques distaux et fonctionnellement essentiels, offrant un nouveau moyen de contrôler l'activité d'une cible enzymatique en affectant sa flexibilité et/ou en perturbant des changements conformationnels biologiquement importants¹¹.

Les effecteurs allostériques offrent un autre avantage important : les effets secondaires de nombreux médicaments sont souvent causés par la liaison de la molécule à des sites actifs similaires sur des protéines structurellement homologues, mais dont la fonction est différente. Or, les sites allostériques retrouvés sur les surfaces distales ont peu de chance d'être aussi bien conservés parmi les protéines fonctionnellement distinctes. Les effecteurs allostériques offrent ainsi la possibilité de réduire ou d'éliminer les effets secondaires attribuables à des liaisons non-spécifiques¹¹.

1.3. Hypothèses et objectifs

De toute évidence, le succès de modulateurs allostériques visant à réguler la fonction des enzymes est largement tributaire de notre capacité à identifier les sites allostériques et/ou les réseaux dynamiquement fonctionnels à la surface des protéines. Ceci nécessite une compréhension précise des mouvements moléculaires liant la structure et la fonction des protéines, principalement aux échelles de temps biologiquement pertinentes (μ s-ms). Les techniques expérimentales employées pour l'élucidation des mouvements de protéines à cette échelle de temps demeurent peu utilisées et/ou s'appuient sur les méthodologies récemment développées. Conséquemment, le profil dynamique de plusieurs familles de protéines et enzymes cibles d'intérêt clinique demeure généralement mal caractérisé ou tout simplement inconnu. Bien que des études théoriques aient illustré l'influence des changements conformationnels sur la fonction enzymatique, la validation expérimentale de ces hypothèses demeure à ce jour pratiquement inexplorée.

Malgré l'imposante littérature consacrée au sujet ces dernières années^{1-6, 8, 9, 11, 18-24}, il n'en demeure pas moins que l'importance de l'allostérie et la caractérisation des propriétés dynamiques et fonctionnellement essentielles chez plusieurs membres protéiques et enzymatiques distincts et/ou apparentés demeurent des concepts relativement obscurs. Dans la présente étude, nous analyserons des membres d'une famille d'enzymes structurellement et fonctionnellement homologues dans le but de mettre en lumière l'importance des mouvements dynamiques à l'échelle de temps de leur événement catalytique (ms). À cet effet, nous émettons les hypothèses suivantes :

Hypothèse no. 1 : Il existe chez des enzymes structurellement et fonctionnellement homologues des réseaux de résidus flexibles formés d'une hiérarchie de secteurs dynamiques contigus et dont les motions à l'échelle de temps de la milliseconde sont liées à l'activité catalytique de ces enzymes apparentées.

Hypothèse no. 2 : Ces réseaux dynamiques sont semblables chez des homologues structuraux et fonctionnels de la même famille d'enzyme. Leur existence ayant fait l'objet d'une pression sélective fonctionnelle, il en résulte une conservation évolutive de ces réseaux dynamiques. En d'autres termes, nous soulevons l'hypothèse que des mouvements non-aléatoires de résidus puissent être évolutivement sélectionnés comme partie intégrante de la relation structure-fonction chez de nombreuses enzymes homologues.

Hypothèse no. 3 : La perturbation de ces réseaux dynamiques conservés affecte l'activité catalytique optimale de ces biocatalyseurs, démontrant qu'ils ne sont pas uniquement fortuits et que leur présence est partie intégrante de la fonction enzymatique.

Dans le but de répondre à ces hypothèses, nous proposons d'entreprendre une série d'expériences dynamiques et fonctionnelles sur deux homologues de la superfamille des ribonucléases pancréatiques : la ribonucléase A bovine (RNase A) et la protéine cationique des éosinophiles de l'humain (ECP, ou RNase 3). Pour répondre aux trois hypothèses susmentionnées, nous tenterons d'atteindre les objectifs suivants :

Objectif no. 1 : Démontrer l'existence de résidus et de réseaux dynamiques conservés à l'échelle de temps de la milliseconde chez la RNase A et ECP en utilisant des techniques de dispersion de relaxation par résonance magnétique nucléaire (RMN), notamment la technique de Carr-Purcell-Meiboom-Gill (^{15}N -CPMG, *vide infra*).

Objectif no. 2 : Utiliser des techniques de titrage RMN et d'interaction protéine-ligand dans le but d'étudier l'effet de la liaison d'un ligand sur les événements dynamiques observés chez ces homologues enzymatiques.

Objectif no. 3 : Utiliser des techniques de mutagenèse dirigée dans le but de démontrer que les perturbations dynamiques d'homologues structuraux et fonctionnels affecte leur fonction biologique et catalytique.

Les sections subséquentes offriront des détails supplémentaires sur les différents systèmes enzymatiques à l'étude, en plus d'offrir un bref tour d'horizon de la méthodologie principale employée lors de ces analyses. Le chapitre 2 présentera ensuite les résultats obtenus sous la forme d'un article scientifique publié fin 2012 dans le périodique *J. Biol. Chem* (réf.²⁵).

1.4. Les ribonucléases pancréatiques : un système modèle

Les hypothèses et objectifs soulevés dans la précédente section requièrent l'étude de systèmes enzymatiques modèles qui répondent à plusieurs critères spécifiques et adaptés à la réalité expérimentale proposée. La superfamille des ribonucléases pancréatiques (RNases) constitue un excellent modèle pour étudier les processus évolutifs et dynamiques qui régulent la fonction catalytique entre différents homologues structuraux. Tout d'abord, ces RNases forment un vaste réseau d'enzymes fonctionnellement distinctes partageant un repliement caractéristique et des éléments catalytiques similaires, dont la stricte conservation des résidus impliqués dans l'activité enzymatique (Figure 3)²⁶. Le membre fondateur de cette famille, la ribonucléase A bovine (RNase A), est l'une des enzymes les plus étudiées du 20^e siècle²⁷. Conséquemment, même si sa dynamique fonctionnelle n'a été étudiée en détail que récemment par des techniques RMN²⁸⁻³³, une quantité importante d'informations structurales, fonctionnelles et mutationnelles sont néanmoins disponibles à des fins comparatives²⁶, ce qui constitue un avantage indéniable.

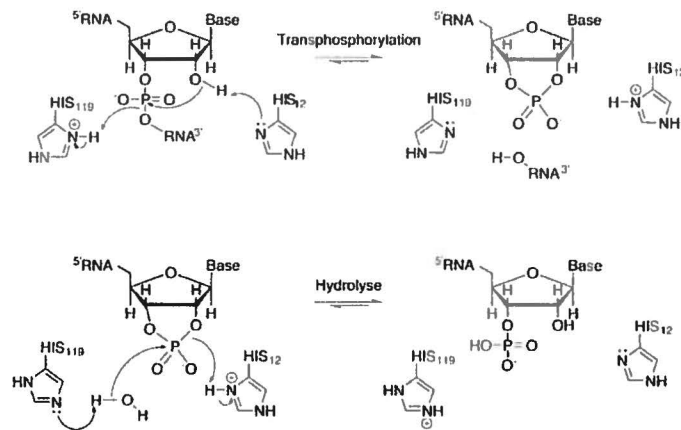


Figure 3. Mécanisme catalytique de la famille des RNases. Le mécanisme catalytique de la RNase A est illustré. Le résidu His12 agit à titre de base générale alors que His119 agit à titre d'acide général dans la première partie de la réaction (transphosphorylation). Les rôles sont ensuite inversés pour la seconde partie de la réaction (hydrolyse), permettant de régénérer l'enzyme libre et active.

En ce qui a trait aux homologues de la RNase A, des études génomiques, protéomiques et structurales ont permis l'identification de 8 membres chez l'humain (RNase 1-8)³⁴ et plus d'une centaine chez d'autres vertébrés (Figure 4)²⁶. Les membres de la superfamille des RNases catalysent tous la transphosphorylation et l'hydrolyse subséquente de brins d'ARN avec des efficacités catalytiques divergentes, en plus de bénéficier d'une large gamme d'activités biologiques encore peu caractérisées. À titre d'exemple, l'angiogénine (RNase 5) intervient dans la synthèse de nouveaux vaisseaux sanguins et son dérèglement cause une angiogenèse pathologique liée à la sclérose amyotrophique latérale, au diabète et à plus de 18 formes différentes de cancers humains, notamment le cancer du sein, du côlon, du foie, du pancréas et de la prostate^{35, 36}. De même, la RNase 1 humaine est impliquée dans la régulation de l'homéostasie vasculaire et neurologique, dans l'inactivation du VIH, en plus d'agir à titre d'agent anti-tumoral³⁷.

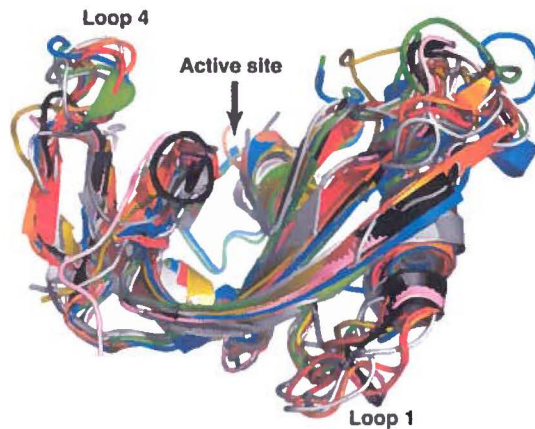


Figure 4. Structure tridimensionnelle de RNases pancréatiques retrouvées chez les vertébrés. Tous les membres de la superfamille des RNases pancréatiques conservent une structure tridimensionnelle similaire, avec une variabilité et une flexibilité principalement concentrées dans les boucles. Les membres représentatifs de la famille sont structurellement superposés sur la structure de la RNase A (rouge) comme suit : BS-RNase (jaune, PDB 3DJX); EDN (RNase 2) humaine (bleu, PDB 1GQV); ECP (RNase 3) humaine (vert, PDB 1QMT); RNase 4 humaine (brun, PDB 1RNF); Angiogénine humaine (noir, PDB 1ANG); RNase 1 de rat (orange, PDB 1RRA); Angiogénine bovine (rose, PDB 1AGI); Onconase de grenouille (gris, PDB 3PHN); et RNase ZF-3e de poisson (vert olive, PDB 2VQ9). Figure tirée de réf.²⁶.

Notre attention s'est portée sur la RNase 2 (neurotoxine dérivée des éosinophiles, ou EDN) et la RNase 3 (protéine cationique des éosinophiles, ou ECP), deux homologues humains qui possèdent 67 % d'identité de séquence (Figure 5). Ces protéines sont toutes deux situées dans les granules cytotoxiques des cellules éosinophiles et sont libérées conjointement suite à la dégranulation provoquée par des stimuli immunitaires. Ainsi, ECP est une enzyme impliquée dans la cytotoxicité, la coagulation, en plus de posséder une activité antibactérienne³⁸. Curieusement, malgré une structure très similaire (Figure 5A) et une conservation absolue des résidus impliqués dans la catalyse enzymatique, ECP et EDN possèdent des activités ribonucléolytiques qui sont de 2 à 4 ordres de grandeur plus faibles que celle de la RNase A³⁸. Nos précédents travaux de RMN sur la RNase A et ceux du laboratoire de J. Patrick Loria à l'Université Yale^{28-33, 39} (en plus d'indices obtenus de structures cristallines et d'études de modélisation moléculaire²⁶) nous laissent croire que certaines RNases homologues sont des enzymes potentiellement dynamiques pouvant faire l'objet de contrôle allostérique. En effet, les travaux de Doucet et Loria ont montré que les mouvements de résidus à l'échelle de temps de la milliseconde présents dans RNase A sont liés à son activité catalytique, et que la perturbation de

ces mouvements par mutagenèse affecte la fonction de l'enzyme^{30,31}. Il a notamment été observé que la perturbation des interactions moléculaires à un site distant reliant His48 à la boucle I (Figure 5) abolit complètement la dynamique et affecte la catalyse enzymatique de cette protéine. Ces études récentes ont dévoilé l'existence du premier site allostérique caractérisé chez un membre de la famille des RNases, fournissant des informations sur l'inhibition allostérique potentielle de cette famille d'enzymes. Fait intéressant, la boucle I impliquée dans ces mouvements allostériques chez la RNase A est beaucoup plus courte chez ECP (6 résidus au lieu de 12), suggérant que les mouvements de cette boucle d'ECP sont probablement plus restreints que ceux retrouvés chez la RNase A. Pour vérifier si ce potentiel est conservé chez cet homologue humain, l'incidence des mouvements conformationnels chez ECP sera caractérisée telle que décrit dans la section précédente.

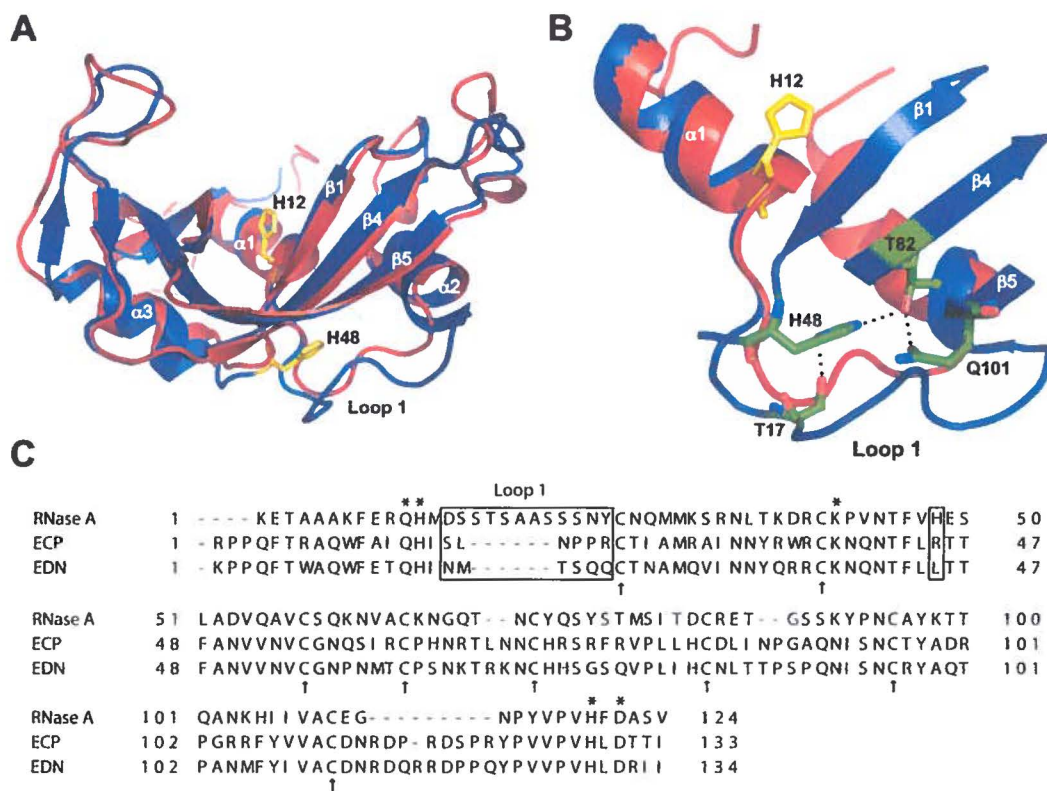


Figure 5. Comparaison structurale entre la RNase A et ECP. A) Superposition structurale de la RNase A (bleu, PDB 1FS3) et ECP (rouge, PDB 1DYT). His12 dénote la position du site actif par rapport à la position de His48 et de la boucle 1 chez la RNase A. B) Vue agrandie des interactions atomiques entre His48 et la boucle 1 chez la RNase A. À noter, la boucle 1 chez la RNase A fait 12 résidus (D₁₄SSTSAASSS₂₅) alors qu'elle ne fait que 6 résidus chez ECP (S₁₇LNPPR₂₂, numérotation ECP). C) Alignement de séquence de la RNase A, ECP et EDN. Les positions du site actif Gln11, His12, Lys41, His119 et Asp121 sont marquées d'un astérisque et les 8 cystéines formant les 4 ponts disulfures conservés de la structure sont identifiées par des flèches. La boucle 1 et la position His48 sont encadrées. Figure tirée de réf.³¹.

En plus du potentiel dynamique de cette famille d'enzymes, des contraintes expérimentales rendent l'étude de ces RNases homologues idéale pour évaluer les hypothèses précédemment soulevées. En effet, tous les homologues de RNases connus sont des monomères globulaires cytosoliques dont la taille se situe aux alentours de 15 kDa, faisant de ces enzymes des candidates de choix pour leur étude RMN. La trop grande taille des protéines et leur potentiel de multimérisation affectent à la fois la qualité des spectres ¹H-¹⁵N HSQC et la précision des

signaux obtenus, ce qui limite grandement le potentiel d'analyse de relaxation dynamique de la technique RMN. De ce fait, la trop grande taille d'une protéine multiplie la quantité de résonances N—H et augmente la probabilité de superposition des résonances amides dans le spectre ^1H - ^{15}N HSQC, rendant ainsi plusieurs résidus « invisibles » puisqu'ils se confondent les uns avec les autres dans le spectre. De plus, la taille trop élevée de la protéine et/ou sa tendance à multimériser affecte à la fois la vitesse de rotation en solution (*tumbling*) et les constantes de relaxation R_1 et R_2 , ce qui a pour effet d'élargir et de réduire la taille des pics dans un spectre ^1H - ^{15}N HSQC (*line broadening*). Les RNases étant solubles, petites et monomériques, ces problèmes sont ainsi évités lors de leur étude par RMN.

Sous forme recombinante, les RNases sont réputées pour être des protéines difficiles à exprimer dans un système d'expression bactérien puisqu'elles nécessitent la présence d'un environnement oxydant pour leur repliement. En effet, toutes les RNases pancréatiques requièrent la formation de 3 à 4 ponts disulfures pour être actives et bien repliées (Figure 5). Puisque le cytoplasme bactérien est un environnement réducteur, la formation de ces ponts disulfures demeure impossible *in vivo* et les RNases se retrouvent ainsi sous forme de corps d'inclusion insolubles lorsqu'exprimées chez *E. coli*. Or, des protocoles de repliement *in vitro* avec échange d'oxydo-réduction ont été développés pour permettre le repliement adéquat de ces enzymes^{38, 40}. Lorsque repliées et maintenues dans un tampon approprié, elles sont également réputées pour leur stabilité en solution et leur longévité à température ambiante, ce qui les rend d'autant plus appropriées pour des analyses de relaxation RMN. En effet, ces expériences peuvent souvent s'échelonner sur plusieurs jours, requérant des échantillons protéiques stables.

1.5. Analyse dynamique par dispersion de relaxation RMN (^{15}N -CPMG)

Les mouvements moléculaires à l'échelle de temps de la milliseconde chez les protéines se manifestent sous forme de changements conformationnels subtils ou majeurs qui peuvent être caractérisés en détail à l'aide d'expériences RMN de Carr-Purcell-Meiboom-Gill (^{15}N -CPMG). Cette technique, aussi appelée dispersion de relaxation, permet de fournir des informations dynamiques à l'échelle atomique pour chaque vecteur N—H isotopiquement enrichi sur le squelette peptidique d'une protéine. Ainsi, chacun des résidus de la protéine sera observable dans

un spectre ^1H - ^{15}N HSQC et les effets de la perturbation de la période de relaxation transversale (R_2) sera quantifiable par ^{15}N -CPMG, à l'exception des résidus proline, qui ne possèdent pas de vecteur N—H^{20, 22, 41}. Ayant été décrite et révisée à de multiples reprises depuis son développement et sa première utilisation dans un système protéique par Loria, Rance et Palmer il y a une quinzaine d'années⁴¹, les détails physiques et théoriques de cette méthodologie dépassent largement le cadre du présent mémoire. Le lecteur intéressé se fera volontiers diriger vers une multitude d'articles de revue et d'extensions de la technique appliquées à son amélioration et/ou à son utilisation dans un contexte RMN différent (introduction du TROSY, application à d'autres stratégies de marquages isotopiques, etc.)^{3, 5, 19-23, 41-45}. Une description vulgarisée du phénomène est néanmoins offerte dans le présent chapitre, afin de permettre au lecteur d'apprécier les avantages et inconvénients de la méthodologie principale présentée dans le chapitre 2.

En partant du postulat qu'une protéine ou une enzyme bien repliée et active en solution est largement peuplée par un seul conformère thermodynamiquement stable, certains conformères excités de plus haute énergie (peuplés à plus de 0.5%) peuvent néanmoins être observés pour une courte période de temps en solution, question de permettre à la protéine d'effectuer une action biologique quelconque (liaison d'un ligand, catalyse enzymatique, relâchement d'un produit, etc.). Le mouvement de va-et-vient répétitif du vecteur N—H causé par l'échange conformationnel entre l'état fondamental stable (conformation A) et l'état excité de plus haute énergie (conformation B) peut être caractérisé et quantifié avec précision par RMN pour chacun des vecteurs N—H accessibles dans le spectre ^1H - ^{15}N HSQC de la protéine isotopiquement enrichie. Ainsi, pour chacun des vecteurs N—H et pour un mouvement dont la vitesse d'échange se situe entre la microseconde et la milliseconde (μs - ms), l'échantillonnage de plus haute énergie (conformation B) se trouve en échange constant avec l'état fondamental (conformation A) à une vitesse donnée, dont l'échange peut être quantifié à l'aide de la constante d'échange k_{ex} ($k_{\text{ex}} = k_{\text{A}} + k_{\text{B}}$) (Figure 6A). Lorsqu'effectuée sur des échantillons protéiques isotopiquement enrichis (^2H , ^{13}C et/ou ^{15}N), la technique de CPMG est particulièrement adaptée à la caractérisation d'événements de flexibilité survenant à chaque résidu non-proline pour des changements conformationnels locaux se situant *grosso modo* entre 50 et 3000 échanges $k_{\text{A}} \leftrightarrow k_{\text{B}}$ par seconde, ce qui correspond à l'échelle de temps de la catalyse de la plupart des systèmes enzymatiques connus (Figure 2).

Ainsi, en permettant de caractériser et de quantifier les mouvements atomiques qui se produisent à cette échelle de temps biologiquement pertinente, la technique de ^{15}N -CPMG s'avère idéale pour identifier des groupes de résidus qui pourraient définir des sites allostériques inconnus et/ou potentiellement essentiels à l'événement catalytique d'une enzyme. Puisque cette méthode permet de caractériser la dynamique de la protéine résidu par résidu en fonction du vecteur N—H analysé, une courbe de dispersion de relaxation sera observable pour chaque vecteur N—H, et donc pour chaque résidu de la protéine qui subit un mouvement concerté à cette échelle de temps (Figure 6B, courbe noire). En revanche, un vecteur N—H immobile à cette échelle de temps n'expérimentera aucun effet R_2 en fonction des différentes impulsions de refocalisation (τ_{cp}) fixées dans une période de relaxation donnée (Figure 6B, ligne rouge). La dispersion de relaxation est le résultat de l'analyse mathématique quantitative de données de relaxation obtenues à deux champs magnétiques distincts (typiquement 500, 600, 700 et/ou 800 MHz) et permet d'extraire 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 (p_A , p_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 ^1H - ^{15}N HSQC^{20, 22, 45, 46}.

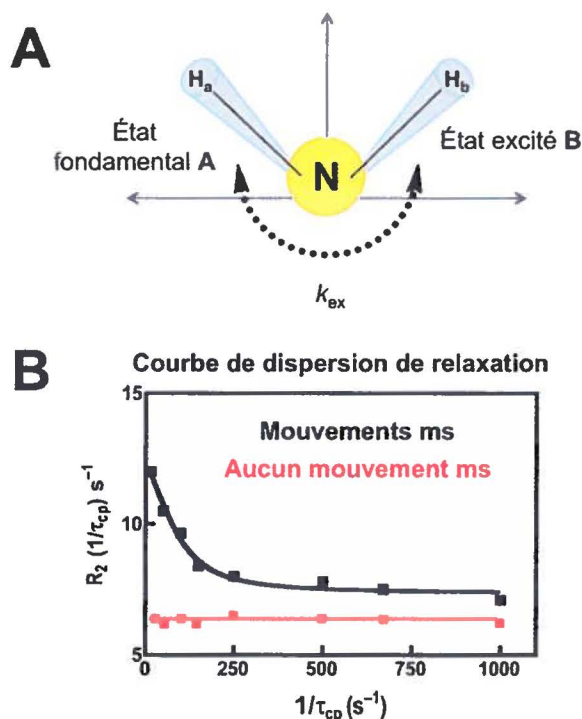


Figure 6. Principe de la dispersion de relaxation RMN (^{15}N -CPMG). A) Échange conformationnel entre un état fondamental (conformation A) et un état excité (conformation B) pour un vecteur N—H donné, dont la vitesse d'échange est caractérisée par la constante d'échange k_{ex} . B) Courbe de dispersion de relaxation typique démontrant l'effet de la dynamique d'un vecteur N—H sur la relaxation transversale (R_2) en fonction de l'inverse d'impulsions de refocalisation (τ_{cp}) fixées dans une période de relaxation donnée. La courbe noire représente un résidu expérimentant un échange conformationnel à l'échelle de temps du CPMG (50-3000 s^{-1}) alors que la courbe rouge dénote un résidu fixe à cette échelle de temps.

1.6. Titrage RMN

Certaines enzymes se comportent allostériquement selon le modèle Monod, Wyman et Changeux (MWC), lequel stipule que des structures en solution existent en échange conformationnel dynamique en l'absence de tout régulateur ou substrat (*conformational selection*), tandis que d'autres enzymes se comportent selon le modèle de Koshland, Nemethy et Filmer (KNF), énonçant plutôt que la liaison d'un ligand ou d'un substrat est nécessaire pour conditionner la conversion dynamique du conformère A en conformère B (*induced fit*)¹⁸. Ces deux modèles soulignent l'importance de caractériser les mouvements d'une protéine en l'absence et en présence de ligands, puisque l'échange conformationnel peut se voir activé ou

diminué en l'absence ou en présence de ligands. Dans le but de fournir des informations sur les résidus dynamiquement affectés par la liaison des ligands, nous allons effectuer des expériences de titrage RMN (Figure 7). 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 affectés 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 ^1H - ^{15}N HSQC lors de l'augmentation de la concentration d'un ligand d'intérêt par l'ajout successif de celui-ci (Figure 7).

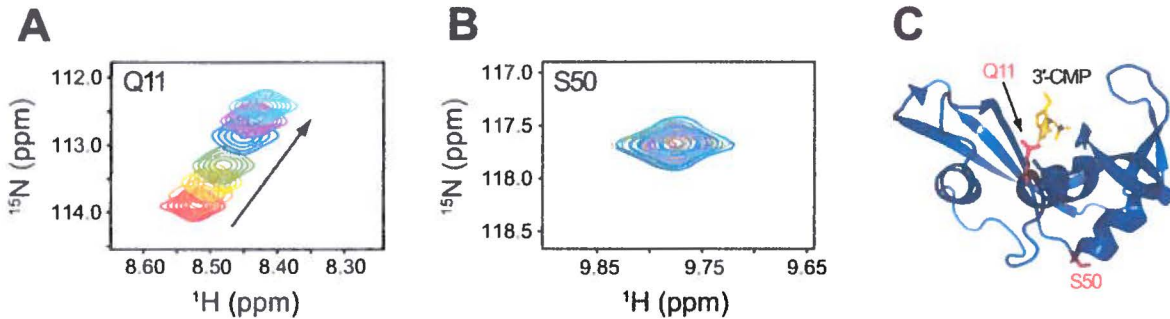


Figure 7. Titration RMN d'un ligand dans une protéine d'intérêt. Le titrage RMN permet l'observation directe des résidus en contact avec un ligand et/ou des résidus dynamiques affectés par la liaison d'un ligand dans une protéine. L'exemple illustre le titrage du 3'-CMP dans la RNase A^{30,31}. L'interaction entre la protéine et le ligand est suivie par les changements de déplacements chimiques subis par le squelette peptidique de la protéine pour chacun des résidus observés dans un spectre ¹H-¹⁵N HSQC et pour divers ratios molaires ligand:protéine allant de 0 (rouge) à 12 (cyan) équivalents. A et B) sections du spectre ¹H-¹⁵N HSQC montrant les changements de déplacement chimique que subissent les résidus (A) Gln11 et (B) Ser50 lors de la liaison du 3'-CMP. Puisque Gln11 est un résidu du site actif en contact direct avec le substrat de l'enzyme, un important déplacement chimique du pic de ce résidu est observé lors de l'augmentation de la concentration en 3'-CMP. En revanche, puisque Ser50 est situé loin du site actif et est non affecté par la liaison du ligand, aucun déplacement du pic de résonance ¹H-¹⁵N n'est observé. C) cartographie structurale de Gln11 et Ser50 sur la structure tridimensionnelle de la RNase A. Gln11 et Ser50 sont représentés en bâtons rouges et 3'-CMP en bâtons jaunes.

CHAPITRE 2 : ARTICLE SCIENTIFIQUE

2.1. Présentation de l'article scientifique

Les résultats du présent chapitre illustrent l'importance de secteurs de résidus dynamiques chez des protéines structurellement et fonctionnellement homologues, en plus de démontrer qu'il nous est possible de contrôler le mouvement de résidus dynamiques semblables et constitués de réseaux contigus à l'aide de la mutagenèse. Les résultats sont présentés sous la forme d'un article scientifique publié dans l'édition du 28 décembre 2012 de la revue *The Journal of Biological Chemistry* (volume 287, numéro 53, pages 44289-44300, DOI 10.1074/jbc.M112.394866), un périodique avec un facteur d'impact de 4,651 (Thomson Reuters, juin 2013). La version publiée de l'éditeur sous format PDF est disponible en ligne sur le site web du périodique *J. Biol. Chem*²⁵. Or, dans le but de préserver la mise en page de ce mémoire, la version finale de l'article sous sa forme manuscrite est ici présentée.

2.2. Contribution des auteurs

Les travaux exposés dans cet article scientifique ont été majoritairement réalisés par Donald Gagné, étudiant au doctorat dans le laboratoire de Nicolas Doucet. En tant qu'étudiante à la maîtrise sous la supervision directe de Donald Gagné, j'ai été impliquée dans toutes les étapes scientifiques et intellectuelles de cette étude, en plus d'avoir préparé des figures et d'avoir aidé à la rédaction de certaines sections du manuscrit. J'ai participé aux étapes d'expression et de purification des RNases, aux expériences de titrage RMN et à l'analyse des expériences de relaxation ¹⁵N-CPMG, acquérant par le fait même toute l'expertise méthodologique décrite dans le présent manuscrit. Dr Sébastien Morin (Swiss Institute of Bioinformatics and Biozentrum, Basel, Suisse) a été impliqué dans la mise au point de l'analyse modelfree (dynamique ps-ns) et Pr Evgenii L. Kovrigin (Marquette University, Milwaukee, Wisconsin, États-Unis) a contribué à l'analyse RMN de « line shape ». L'interprétation des résultats et la rédaction du manuscrit ont été effectués par Nicolas Doucet et Donald Gagné, avec mon aide.

2.3. 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 à 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}N -CPMG), d'analyse modelfree et de titrage RMN pour capturer et comparer le rôle de la flexibilité conformationnelle entre deux homologues structuraux de la famille des ribonucléases pancréatiques : la RNase A et la protéine cationique des éosinophiles humaine (ECP, ou RNase 3). En plus de conserver les mêmes résidus catalytiques et la même structure tridimensionnelle, les deux homologues montrent des secteurs de résidus dynamiques similaires à l'échelle de temps de la milliseconde, suggérant que la flexibilité conformationnelle peut être conservée parmi des protéines structurellement analogues qui affichent néanmoins une faible identité de séquence. Notre travail démontre également que la flexibilité conformationnelle de ECP peut être dynamiquement et fonctionnellement reproduite dans la structure de la RNase A, ce que nous démontrons par la création d'une chimère hybride entre les deux protéines. Ces résultats soutiennent l'hypothèse que la flexibilité conformationnelle est en partie nécessaire pour la fonction catalytique chez des enzymes homologues, soulignant en outre l'importance des secteurs dynamiques de résidus dans l'organisation structurale des protéines.

2.4. Article scientifique

Conservation of flexible residue clusters among structural and functional enzyme homologues

Donald Gagné¹, Laurie-Anne Charest¹, Sébastien Morin², Evgenii L.
Kovrigin³ & Nicolas Doucet^{1*}

¹From INRS-Institut Armand-Frappier, Université du Québec, 531 Boulevard des Prairies, Laval,
QC H7V 1B7, CANADA

²From the Swiss Institute of Bioinformatics and Biozentrum, University of Basel,
Klingelbergstrasse 50/70, 4056 Basel, SWITZERLAND.

³From the Chemistry Department, Marquette University, P.O. Box 1881, Milwaukee, Wisconsin
53201, USA

***J. Biol. Chem.* 2012. 287(53): 44289-44300.**

*To whom correspondence should be addressed: INRS-Institut Armand-Frappier, Université du Québec, 531 Boulevard des Prairies, Laval, QC H7V 1B7, CANADA, Tel: (450) 687-5010, ext. 4212; Fax: (450) 686-5501; E-mail: nicolas.doucet@iaf.inrs.ca

Keywords: ribonuclease A, eosinophil cationic protein, NMR relaxation, CPMG, conformational flexibility, protein dynamics, mutagenesis, protein sectors.

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.

SUMMARY

Conformational flexibility between structural ensembles is an essential component of enzyme function. While the broad dynamical landscape of proteins is known to promote a number of functional events on multiple timescales, 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 timescale 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 (ECP, 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 ECP 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 importance 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). While 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 timescale 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 superfamilies (11) and among members of the same enzyme fold (12). While 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), is yet unknown whether residue motions on multiple timescales 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 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 preserving the same catalytic triad and displaying 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. Though 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, 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). While 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 timescale of k_{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 ps-ns timescale have also been shown 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 (H48-T82), which are located on the adjacent β 1 and β 4 strands (23). These molecular motions are transmitted over a length-scale of more than 20 Å from loop 1 to active site residues T45 and D83 (also located in β 1 and β 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). While the dynamically important residue pair H48-T82 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 picosecond-to-nanosecond (ps-ns) and the slower microsecond-to-millisecond timescales (μ s-ms). 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 analyses 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, while motional clusters are structurally conserved between the two homologues, their respective global rates of conformational exchange (k_{ex}) vary significantly. As predicted by the absence of the functionally dynamic H48-T82 residue pair in ECP, this enzyme lacks the aforementioned loop1- β 1- β 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_{ECP} chimera (22).

Our current motional investigation demonstrates that both ECP and a chimeric hybrid between RNase A and ECP (RNase A_{ECP} (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 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). ¹⁵N-labelled 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 M⁻¹cm⁻¹ (9,800 M⁻¹cm⁻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 ¹⁵N-labelled ECP in 15 mM sodium acetate, 10% ²H₂O, pH 5.0; or 0.7 mM ¹⁵N-labelled RNase A in 5 mM MES-NaOH, 7 mM NaCl, 0.01% NaN₃, 10% ²H₂O, pH 6.4.

NMR experiments were carried out on Varian (Agilent) 500 MHz 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 (BMRB entry 15757) and further confirmed with a ^1H - ^{15}N TOCSY HSQC experiment.

Relaxation dispersion experiments (μs - ms motions)—Backbone amide ^{15}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 τ_{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), 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 timescale motions (ps - ns)— ^{15}N - R_1 , ^{15}N - R_2 and steady state heteronuclear NOE experiments were performed in an interleaved fashion at 500 MHz, as previously described (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 timescale were analyzed by fitting the NMR spin relaxation rates to the model-free formalism (reviewed in (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 1QMT) and hydrogen atoms were added in DS Visualizer 3.5 (Accelrys). Values for the ^{15}N chemical shift anisotropy (CSA) and N-H bond length were set at -172 ppm and 1.02 Å, respectively. Model selection for each residue and diffusion tensors were tested, optimized and selected using Akaike information criterion (AIC) 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. ^1H - ^{15}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. ^1H - ^{15}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 addition of 0.01M 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 (IDAP) (40). The 1D datasets for fitting were obtained as slices through the fast-exchange peaks in ^1H - ^{15}N HSQC spectra using Sparky (33) with a custom Python extension (41). The 1D 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, the dissociation rate constant as well as the frequency of the bound complex using the IDAP code implemented in MATLAB (MathWorks). The standard errors were estimated from fitting of multiple line shape datasets generated with 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 IDAP 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 uracil-3'-monophosphate (3'-UMP) and adenosine-5'-monophosphate (5'-AMP) ligands (Fig. 1b,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}N -labeled proteins and binding was followed by measuring chemical shift variations observed in ^1H - ^{15}N Heteronuclear Single-Quantum Coherence (HSQC) spectra until saturation was reached. Affinities of 3'-UMP and 5'-AMP binding to ECP were estimated using line shapes observed in the ^1H - ^{15}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, likely due to weak affinity

of the nucleotides to a second non-specific site. To estimate affinity of the binding interaction with the cognate site in the 3'-UMP datasets, we selected residues that only responded to the main binding event in ECP, *i.e.* V125 and V127 (Fig. S1). The 5'-AMP binding has much weaker contribution from the second binding process, so residues demonstrating the most significant shifts were analyzed (H15, A110, and H128). 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 (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 B₁ and B₂ 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) ¹H-¹⁵N weighted average chemical shift variations upon 3'-UMP binding to ECP, relative to 17 residues in RNase A (Fig. 2a,b). While the chemically affected residues are mostly located in the vicinity of the B₁ subsite and are directly involved in binding and/or catalysis in RNase A (*e.g.* H12, K41, T45, H119) (Fig. 2b), their functional equivalents in ECP (H15, K39, T42 and H128) do not show significant chemical shift perturbations upon 3'-UMP binding (Fig. 2a). If 3'-UMP does bind to the B₁ 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 T45 and D83 are required for optimal product release in RNase A (24). However, no D83 equivalent can be found in ECP, and T42 (the T45 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. 2b).

Based on their motional behavior (*vide infra*) and secondary structure elements (44), the two structural homologues were divided into four distinct regional residue clusters encompassing loop 4 (cluster 1), the β 2 sheet (cluster 2), the β 1 sheet and 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 (T45, F46, T82, T100, Q101, all part of cluster 3) also show important conformational exchange on the millisecond timescale 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 Q58, R75 and R77, located respectively 20.2, 18.0 and 14.7 Å away from the ligand—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 (β 1-sheet and adjacent loop 1) are almost completely devoid of significant chemical shift perturbations upon 3'-UMP binding to the B₁ subsite, suggesting that concerted long-range interactions linking loop 1 to the active site are not functionally important in ECP.

Purine binding to the B₂ subsite also yields important chemical shift differences between these two structural homologues. While 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. 2c-d). The magnitude in the ¹H-¹⁵N weighted average chemical shift variations is also stronger in RNase A than ECP. Only the immediate vicinity of the B₂ subsite is chemically affected in presence of 5'-AMP in ECP, including the important catalytic residues H64 and H128. While H128 is required for the hydrolytic step of the reaction, H64 (located in loop 4) directly interacts with P₀ during turnover and is thought to act as the functional equivalent to K66 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). While residues of loop 4 show considerable chemical shift variations upon 5'-AMP binding to RNase A, only H64 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. 2d). 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 (*vide infra*).

Fast timescale analysis (ps-ns)—While 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 timescale 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 ^{15}N - R_1 , ^{15}N - R_2 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 out of the 121 non-proline residues of ECP could be reliably quantitated. Residues F43, R45, N53, Q58, N70, R73, F76, C83 and L129 were removed from the analysis because of low signal intensity, spectral overlapping and/or because they were absent from the ^1H - ^{15}N HSQC. The average values of the relaxation rates are $R_1 = 1.13 \pm 0.13$, $R_2 = 14.04 \pm 1.78$, and $\text{NOE} = 0.478 \pm 0.71$ (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^2) are used to characterize the amplitude of the internal motions of the ^1H - ^{15}N vector on the ps-ns timescale, and thus provide a measure of atomic scale residue flexibility on this particular time frame (46-48). As a result, a completely unrestricted ^1H - ^{15}N bond vector would show a S^2 value of 0 while a fully rigid one would display an order parameter of 1. The generalized S^2 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 $S^2 = 0.740 \pm 0.052$ indicates that ECP is a bit more flexible than the very rigid RNase A on the ps-ns timescale, which displays an average $S^2 = 0.910 \pm 0.051$ (35). Interestingly, the catalytic residue K38 in ECP, which is strictly conserved among pancreatic RNases, displays one of the highest order parameters in the protein ($S^2 = 0.848 \pm 0.026$), along with its very rigid neighbor N39 ($S^2 = 0.907 \pm 0.055$). K38 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), which may require rigidity in ECP. On the contrary, the two most flexible residues, V78 (S^2

= 0.571 ± 0.062) and D84 ($S^2 = 0.452 \pm 0.039$), are at the beginning and the end of the $\beta 2$ 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, a R_{ex} exchange parameter is calculated to achieve an adequate fit to the model-free analysis. This R_{ex} parameter accounts for contributions to R_2 that potentially describe motions occurring on the slower μ s-ms timescale (Fig. 4). Much like with RNase A, a large number of residues throughout the ECP sequence require the addition of this exchange contribution (26 out of the 112 residues analyzed), suggesting the existence of motions occurring on the slower μ s-ms timescale in this enzyme. This observation is in line with our slow timescale analysis of ECP performed by ^{15}N -CPMG (*vide infra*), most notably for residues of loop 4 (cluster 1, Fig. 3), which display some of the most important R_{ex} values in the protein (Fig. 4). Interestingly, the highest R_{ex} observed is for residue H64 ($R_{ex} = 4.22 \pm 1.53 \text{ s}^{-1}$), the structural equivalent to K66 in RNase A, which is also one of the most flexible residues of RNase A on the millisecond timescale (Fig. 5c). Overall, while 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 timescale, in addition to similar flexible residue sectors on the millisecond timescale (*vide infra*).

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 (^{15}N -CPMG) (10). ^{15}N -CPMG experiments allow the characterization of an equilibrium exchange process experienced by the ^1H - ^{15}N bond vector by recording the transverse relaxation rate constant, R_2 , as a function of τ_{cp} , an inter-pulse delay in the CPMG pulse train (49). ^{15}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 (9,50-52)). 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 timescale (Fig. 5). On the contrary, ^{15}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 structural information of the excited state (differences in chemical shifts, $\Delta\omega$)

as well as population dynamics (exchange rates, k_{ex} , and equilibrium populations, p_{APB}) (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 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. 5a), a region that is almost completely devoid of ms dynamics in RNase A (Fig. 5c). Simultaneous fitting of the all-timescale, 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, while 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 $1438 \pm 125 \text{ s}^{-1}$ (Fig. 5c). 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. While the global

exchange rate of loop 4 ($k_{\text{ex}} = 1548 \pm 82 \text{ s}^{-1}$) parallels the product release and catalytic rates in RNase A ($k_{\text{off}} = k_{\text{cat}}$), the global exchange rate for loop 4 is more than three times slower in ECP ($k_{\text{ex}} = 504 \pm 44 \text{ s}^{-1}$) 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. 2b), 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}N chemical shift variations ($\Delta\delta$ ppm) with ^{15}N chemical shift differences between the major and the minor excited state obtained from ^{15}N -CPMG ($\Delta\omega$) remained inconclusive (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}N chemical shift differences of the NMR titrations ($\Delta\delta$ ppm) and the $\Delta\omega$ ppm obtained from ^{15}N -CPMG. The absence of such linear correlation implies additional perturbations caused by the arrival of the ligands and/or that the conformational exchange we observe by ^{15}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 (*vide infra*).

The RNase A_{ECP} 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 β strands 1 and 4, both of them containing the important active site residues T45 and D83 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_{ECP} 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 vs. 12 in the latter), ECP also lacks the important H48-T82

residue pair that we showed 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. 5a), highlighting a different dynamic landscape in this protein.

Originally not apparent from the original 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_{ECP}, Fig. 5a,b). Despite conserving 90% sequence identity with RNase A, the RNase A_{ECP} chimera almost perfectly mirrors the ms dynamics of ECP. Indeed, the conformational exchange of cluster 3 residues is completely dampened in the RNase A_{ECP} hybrid. Conversely, millisecond dynamics are acquired for contiguous residues in cluster 2 (Fig. 5b), a sector that is almost completely devoid of ms motions in RNase A (Fig. 5c). 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 one 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 ¹⁵N-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 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 clusters (1 and 2) such that k_{ex} and populations were shared while the

frequency difference was individual for each residue. The results were filtered on the basis of the coefficient of determination ($R^2 > 0.9$) to ensure that only good fits are used in the analysis. Due to spectral overlap induced by ligand binding and variable data quality, only residues C62 and T67 (cluster 1) were available in the free and ligand-bound forms for comparison. For C62, we determined that the frequency difference between major and minor conformers was in the 95% confidence interval of 320–450 s^{-1} for apo ECP and 300–530 s^{-1} for the ECP-3'-UMP complex. For T67, the corresponding ranges for the same complexes were 180–240 s^{-1} and 190–290 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 out 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, while 3'-UMP and 5'-AMP considerably rigidify the ECP structure upon binding (Fig. 6a,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. 6c,d). A total of 11 (8) and 14 (9) residues rigidify upon 3'-UMP and 5'-AMP binding to ECP (RNase A), respectively (Figs. S4 and S5). However, the increase in conformational exchange upon ligand binding is considerably more pronounced 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 1H - ^{15}N HSQC backbone resonances due to line broadening in the hinge region of cluster 3 (Fig. 6c,d, grey spheres). Overall, these observations are consistent with the previously suggested ligand-dependent induction of hinge motions between the two β -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 (Figs. S4 and S5). For instance, the strictly conserved catalytic lysine (K41 in RNase A, K38 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, while the flexible K41 in RNase A shows significant chemical shift variations in the presence of 3'-UMP, K38 is not at all affected by ligand binding and does not show any conformational exchange in ECP (Fig. S4). While K41 gains conformational exchange in the RNase A-nucleotide 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 ^1H - ^{15}N bond vector of K38 on the ps-ns timescale. Also, it is interesting to point out that F46, one of the most critical aromatic residues for RNase A stability and folding (55-57), is both flexible in the apo and 3'-UMP-bound forms of RNase A (Fig. S4). Although very highly conserved in the RNase family and positioned in the exact same orientation, no such flexibility is observed for F43 in ECP. These observations suggest that very similar functional homologues may still use local dynamics for very different purposes in defining function and stability.

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 (NMA) to infer catalytic timescale motional similarities between protein folds (58). Generally, these studies also involve the direct comparison of simulated data to crystallographically- 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 timescale.

Theoretical studies have delivered a breadth of relevant information on slow timescale 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 timescale. The present study also validates the existence of dynamic clusters—similar to the evolutionary sectors of Ranganathan and coworkers (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 presence of a dynamic subdomain (cluster 3) that transfers millisecond motions essential to product release in RNase A is absent from ECP. Despite conserving the same dynamic cluster architecture as that of RNase A, experimental evidence illustrates that ECP is a more conformationally restrained homologue on the ms timescale.

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 timescales that can be probed by MD simulations (ps-ns), a time frame that is still very far from enzyme turnover (ms). To overcome such barrier, Ramanathan and Agarwal recently used quasi-harmonic analysis (QHA) to model large-scale conformational fluctuations occurring on slower timescales in this enzyme family, providing the most complete theoretical comparison of slow dynamics between ribonuclease homologues so far (12). Focusing on the motional comparison of three pancreatic-like ribonuclease homologues from human, bovine and rat, 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 coworkers 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 (60). 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 3D structure of the S1A serine protease family underlies the biological function of its members, a property they also observed in the PDZ, PAS, SH2 and SH3 domain families (60). These groups of coevolving residues—termed sectors (60,61)—show striking physical connectivity in the 3D 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 coworkers also observe a direct correlation between these coevolving sectors and the networks of residues undergoing conformational fluctuations associated with enzyme catalysis in dihydrofolate reductase (DHFR) (61). Similar to the experimentally characterized clusters in the present work, the authors show that dynamic motions associated with catalysis in DHFR 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 crosstalk on the catalytic timescale (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, further lending 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).

REFERENCES

1. Branden, C., and Tooze, J. (1999) *Introduction to Protein Structure*, 2nd ed., Garland Science, New York
2. Eiben, C. B., Siegel, J. B., Bale, J. B., Cooper, S., Khatib, F., Shen, B. W., Players, F., Stoddard, B. L., Popovic, Z., and Baker, D. (2012) *Nat. Biotechnol.* **30**, 190-192
3. 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) *Science* **329**, 309-313
4. Baker, D. (2010) *Protein Sci.* **19**, 1817-1819
5. Lassila, J. K. (2010) *Curr. Opin. Chem. Biol.* **14**, 676-682
6. Ma, B., and Nussinov, R. (2010) *Curr. Opin. Chem. Biol.* **14**, 652-659
7. Henzler-Wildman, K., and Kern, D. (2007) *Nature* **450**, 964-972
8. Baldwin, A. J., and Kay, L. E. (2009) *Nat. Chem. Biol.* **5**, 808-814
9. Doucet, N. (2011) *Protein Pept. Lett.* **18**, 336-343
10. Loria, J. P., Berlow, R. B., and Watt, E. D. (2008) *Acc. Chem. Res.* **41**, 214-221
11. Carnevale, V., Raugei, S., Micheletti, C., and Carloni, P. (2006) *J. Am. Chem. Soc.* **128**, 9766-9772
12. Ramanathan, A., and Agarwal, P. K. (2011) *PLoS Biol.* **9**, e1001193
13. Cuchillo, C. M., Nogues, M. V., and Raines, R. T. (2011) *Biochemistry* **50**, 7835-7841
14. Marshall, G. R., Feng, J. A., and Kuster, D. J. (2008) *Biopolymers* **90**, 259-277
15. Raines, R. T. (1998) *Chem. Rev.* **98**, 1045-1066
16. Dyer, K. D., and Rosenberg, H. F. (2006) *Mol. Divers.* **10**, 585-597

17. Sorrentino, S. (2010) *FEBS Lett.* **584**, 2194-2200
18. Boix, E. (2001) *Methods Enzymol.* **341**, 287-305
19. 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) *J. Biol. Chem.* **282**, 4626-4633
20. Boix, E., and Nogues, M. V. (2007) *Mol. Biosyst.* **3**, 317-335
21. Jarymowycz, V. A., and Stone, M. J. (2006) *Chem. Rev.* **106**, 1624-1671
22. Doucet, N., Watt, E. D., and Loria, J. P. (2009) *Biochemistry* **48**, 7160-7168
23. Doucet, N., Khirich, G., Kovrigin, E. L., and Loria, J. P. (2011) *Biochemistry* **50**, 1723-1730
24. Watt, E. D., Rivalta, I., Whittier, S. K., Batista, V. S., and Loria, J. P. (2011) *Biophys. J.* **101**, 411-420
25. Watt, E. D., Shimada, H., Kovrigin, E. L., and Loria, J. P. (2007) *Proc. Natl. Acad. Sci. USA* **104**, 11981-11986
26. Butterwick, J. A., Loria, J. P., Astrof, N. S., Kroenke, C. D., Cole, R., Rance, M., and Palmer, A. G., 3rd. (2004) *J. Mol. Biol.* **339**, 855-871
27. Henzler-Wildman, K. A., Lei, M., Thai, V., Kerns, S. J., Karplus, M., and Kern, D. (2007) *Nature* **450**, 913-916
28. Lee, G. M., and Craik, C. S. (2009) *Science* **324**, 213-215
29. Davey, J. A., and Chica, R. A. (2012) *Protein. Sci.* **21**, 1241-1252
30. Sela, M., and Anfinsen, C. B. (1957) *Biochim. Biophys. Acta* **24**, 229-235
31. Loria, J. P., Rance, M., and Palmer, A. G., 3rd. (1999) *J. Am. Chem. Soc.* **121**, 2331-2332
32. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR* **6**, 277-293
33. Goddard, T. D., and Kneller, D. G., University of California, San Francisco

34. Carver, J. P., and Richards, R. E. (1972) *J. Magn. Reson.* **6**, 89-105
35. Cole, R., and Loria, J. P. (2002) *Biochemistry* **41**, 6072-6081
36. Doucet, N., Savard, P. Y., Pelletier, J. N., and Gagné, S. M. (2007) *J. Biol. Chem.* **282**, 21448-21459
37. Morin, S. (2011) *Prog. Nucl. Magn. Reson. Spectrosc.* **59**, 245-262
38. d'Auvergne, E. J., and Gooley, P. R. (2008) *J. Biomol. NMR* **40**, 121-133
39. d'Auvergne, E. J., and Gooley, P. R. (2008) *J. Biomol. NMR* **40**, 107-119
40. Kovrigin, E. L. (2012) *J. Biomol. NMR* **53**, 257-270
41. Kovrigin, E. L., and Loria, J. P. (2006) *Biochemistry* **45**, 2636-2647
42. Kaplan, J. I., and Fraenkel, G. (1980) *NMR of Chemically Exchanging Systems*, Academic Press
43. Rao, B. D. N. (1989) *Methods Enzymol.*, 279-311
44. Mallorqui-Fernandez, G., Pous, J., Peracaula, R., Aymami, J., Maeda, T., Tada, H., Yamada, H., Seno, M., de Llorens, R., Gomis-Ruth, F. X., and Coll, M. (2000) *J. Mol. Biol.* **300**, 1297-1307
45. Boix, E., Leonidas, D. D., Nikolovski, Z., Nogues, M. V., Cuchillo, C. M., and Acharya, K. R. (1999) *Biochemistry* **38**, 16794-16801
46. Lipari, G., and Szabo, A. (1982) *J. Am. Chem. Soc.* **104**, 4559-4570
47. Lipari, G., and Szabo, A. (1982) *J. Am. Chem. Soc.* **104**, 4546-4559
48. Mandel, A. M., Akke, M., and Palmer, A. G., 3rd. (1995) *J. Mol. Biol.* **246**, 144-163
49. Manley, G., and Loria, J. P. (2012) *Arch. Biochem. Biophys.* **519**, 233-231
50. Kay, L. E. (2005) *J. Magn. Reson.* **173**, 193-207
51. Palmer, A. G., 3rd. (2001) *Annu. Rev. Biophys. Biomol. Struct.* **30**, 129-155
52. Palmer, A. G., 3rd. (2004) *Chem. Rev.* **104**, 3623-3640

53. Kovrigin, E. L., Kempf, J. G., Grey, M. J., and Loria, J. P. (2006) *J. Magn. Reson.* **180**, 93-104
54. Vitagliano, L., Merlino, A., Zagari, A., and Mazzarella, L. (2002) *Proteins* **46**, 97-104
55. Chatani, E., Nonomura, K., Hayashi, R., Balny, C., and Lange, R. (2002) *Biochemistry* **41**, 4567-4574
56. Kadonosono, T., Chatani, E., Hayashi, R., Moriyama, H., and Ueki, T. (2003) *Biochemistry* **42**, 10651-10658
57. Vila, R., Benito, A., Ribo, M., and Vilanova, M. (2009) *Biopolymers* **91**, 1038-1047
58. Bahar, I., and Rader, A. J. (2005) *Curr. Opin. Struct. Biol.* **15**, 586-592
59. Marcos, E., Crehuet, R., and Bahar, I. (2010) *PLoS Comput. Biol.* **6**, e1000738
60. Halabi, N., Rivoire, O., Leibler, S., and Ranganathan, R. (2009) *Cell* **138**, 774-786
61. Reynolds, K. A., McLaughlin, R. N., and Ranganathan, R. (2011) *Cell* **147**, 1564-1575
62. Merlino, A., Vitagliano, L., Ceruso, M. A., Di Nola, A., and Mazzarella, L. (2002) *Biopolymers* **65**, 274-283
63. Merlino, A., Vitagliano, L., Ceruso, M. A., and Mazzarella, L. (2003) *Proteins* **53**, 101-110
64. Gunasekaran, K., Ma, B., and Nussinov, R. (2004) *Proteins* **57**, 433-443
65. Fisher, B. M., Schultz, L. W., and Raines, R. T. (1998) *Biochemistry* **37**, 17386-17401
66. Doucet, N., Jayasundera, T. B., Simonovic, M., and Loria, J. P. (2010) *Proteins* **78**, 2459-2468

67. Grzesiek, S., Stahl, S. J., Wingfield, P. T., and Bax, A. (1996) *Biochemistry* **35**, 10256-10261
68. Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B., Keduas, V., and Notredame, C. (2006) *Nucleic Acids Res.* **34**, W604-608

ACKNOWLEDGMENTS

The authors would like to 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_{ECP}. The authors also thank Tara Sprules and Sameer Al-Abdul-Wahid from the Québec/Eastern Canada High Field NMR Facility for their excellent NMR technical assistance, as well as Éric Déziel and Claude Guertin (INRS) for their kind generosity.

FOOTNOTES

This work was supported by NSERC Discovery grant RGPIN 402623-2011 (to N.D.) and a FRQS Research Scholar Junior 1 Career Award (to N.D). N.D. also acknowledges support from the 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.G. and L.A.C. are recipients of a “Fondation Universitaire Armand-Frappier de l'INRS” Ph.D. scholarship.

Table 1. Binding affinities of 3'-UMP and 5'-AMP to RNase A and ECP.

	K_d 3'-UMP (μM)	K_d 5'-AMP (μM)
RNase A	9.7 ± 0.9^a	124.0 ± 0.9^b
ECP	460 ± 100^c	340 ± 30^c

^aTaken from ref. (65); ^bTaken from ref. (66); ^cEstimated from NMR line shape analysis (see Experimental Procedures).

FIGURES

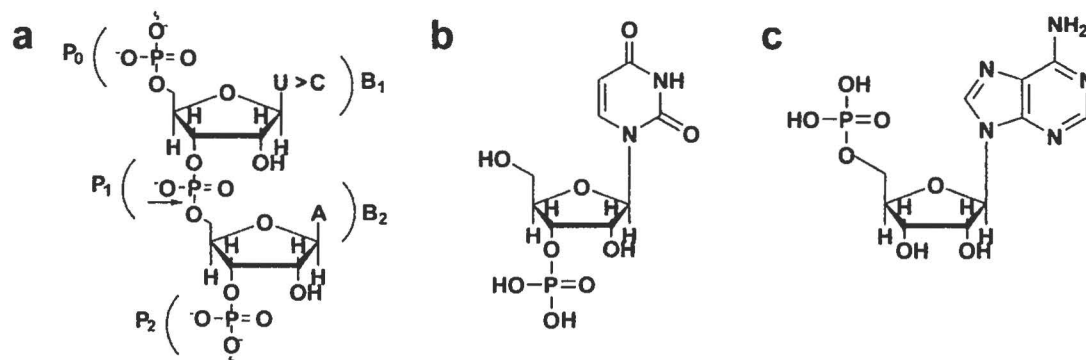


FIGURE 1. Substrates and products of a ribonuclease reaction. (a) Schematic representation of a single-stranded RNA molecule binding to the active site of eosinophil cationic protein (ECP). Phosphate and base binding subsites for RNA substrates are defined as P_n and B_n , of which the P_{0-2} and B_{1-2} 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: His15 and His128 (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).

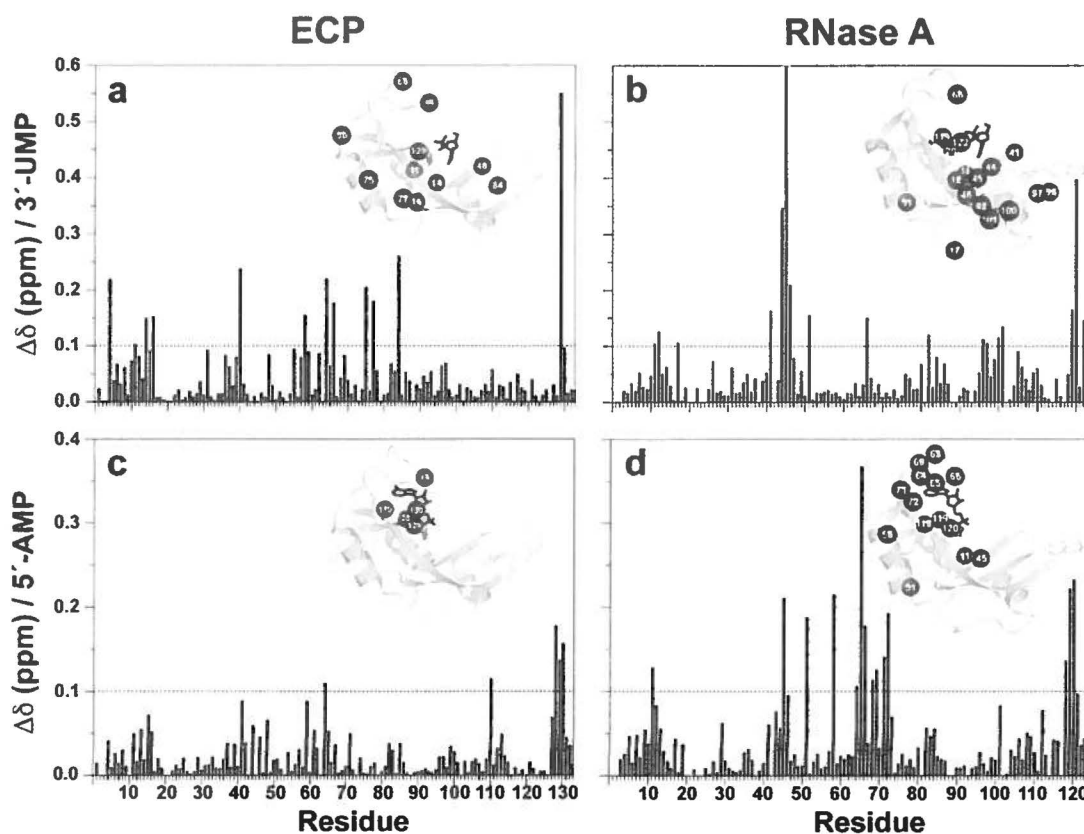


FIGURE 2. ^1H - ^{15}N chemical shift variations induced by 3'-UMP and 5'-AMP binding to ECP and RNase A. Weighted average chemical shift differences ($\Delta\delta$) for (a, b) 3'-UMP and (c, d) 5'-AMP are mapped on the primary sequence of (a, c) ECP and (b, d) RNase A. The ^1H - ^{15}N weighted average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

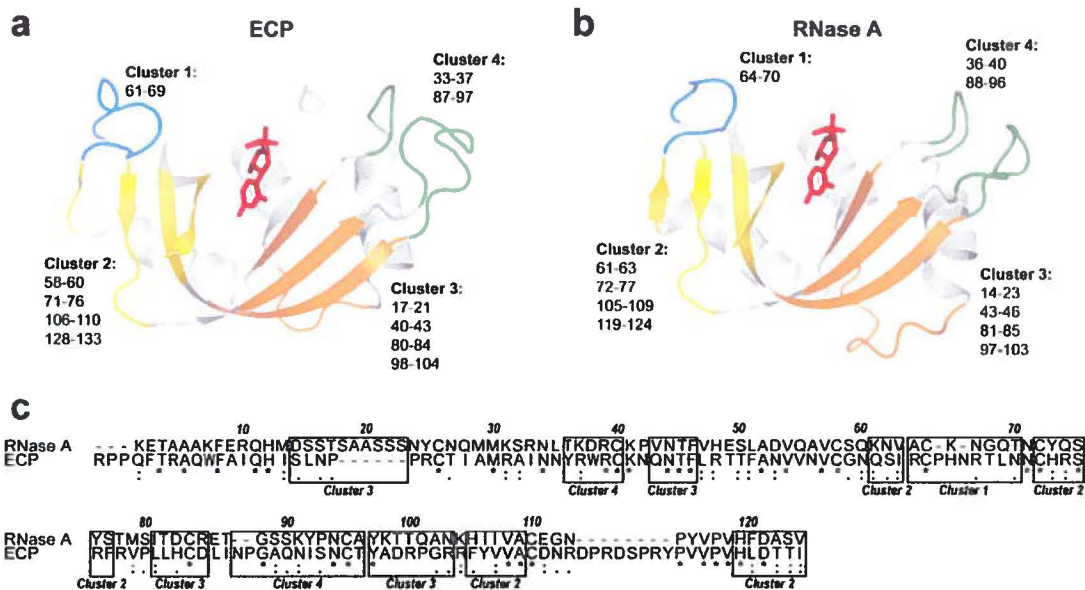


FIGURE 3. Dynamic residue clusters in ECP and RNase A. The two ribonuclease homologues are divided into four distinct residue clusters covering the following secondary structure elements (44): loop 4 (cluster 1, purple), β 2-sheet (cluster 2, yellow), β 1-sheet with adjacent loop 1 (cluster 3, orange), and loops 2 and 6 (cluster 4, green). The four color-coded clusters are highlighted on the 3D structure of (a) ECP (PDB 1H1H) and (b) RNase A (PDB 1O0N), with encompassing residues listed. The 3'-UMP ligand is shown in red. (c) Expresso (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.

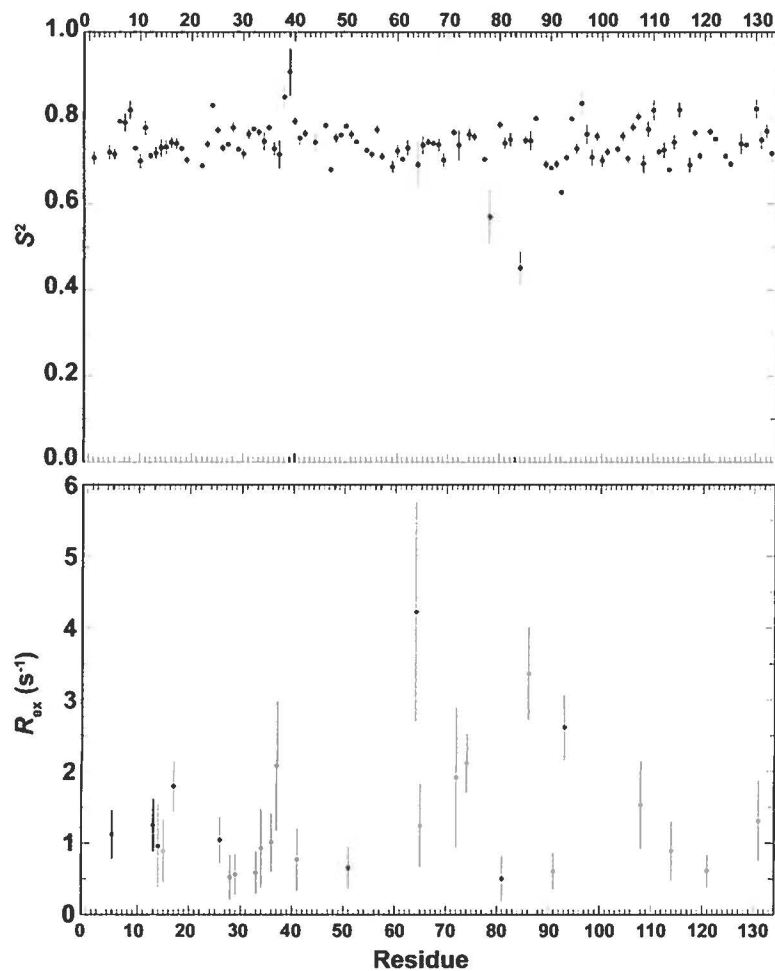


FIGURE 4. Model-free calculated S^2 and R_{ex} parameters plotted on the sequence of ECP. The generalized order parameters (S^2) provide a measure of atomic scale flexibility of the ^1H - ^{15}N bond vector on the ps-ns timescale. The predicted R_{ex} parameter accounts for contributions to R_2 that potentially describe motions occurring on the slower μs -ms timescale.

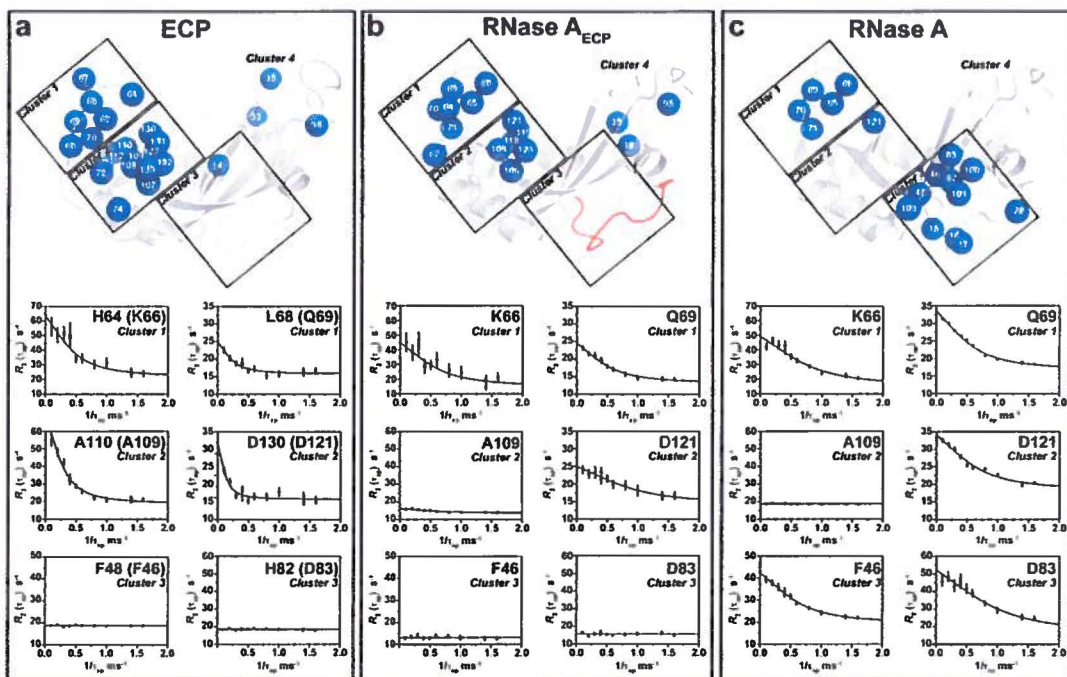


FIGURE 5. Conformational exchange is confined to four independent yet contiguous dynamic clusters in structural and functional ribonuclease homologues. Conformational exchange was investigated in the apo forms of ECP, RNase A and the chimeric hybrid RNase A_{ECP}, in which the RNase A loop 1 (D₁₄SSTSAASSSN_{Y25}, in red) was replaced by that of ECP (S₁₇LNPPR₂₂, ECP numbering) (22). Catalytic timescale (ms) residue motions were probed using ¹⁵N-CPMG relaxation dispersion experiments at 500 and 800 MHz (298K) for the apo forms of (a) ECP, (b) RNase A_{ECP}, and (c) RNase A. Residues were considered for further analysis only if the difference in measured R_2 (ΔR_2 ($1/\tau_{cp}$)) values at fast ($\tau_{cp}=0.625$ ms) and slow ($\tau_{cp}=10$ ms) pulsing rates was greater than 2 s^{-1} , similar to previous reports (22,35). Blue spheres represent residues displaying ¹⁵N-CPMG dispersion with ΔR_2 ($1/\tau_{cp}$) $> 2\text{ s}^{-1}$. Residues are highlighted on the 3D structure of (a) ECP (PDB 1H1H) and (b, c) RNase A (PDB 1O0N). Representative relaxation dispersion curves (800 MHz, 298K) are shown for structurally equivalent positions of ECP (RNase A) found in cluster 1: H64 (K66) and L68 (Q69); cluster 2: A110 (A109) and D130 (D121); and cluster 3: F48 (F46) and H82 (D83). Flexible residues experiencing ¹H-¹⁵N bond

vector motions on the millisecond timescale display a relaxation dispersion curve while residues with no motion on this particular time frame display flat relaxation profiles. Swapping loop 1 of RNase A by that of ECP converts the dynamic clusters of the former into those of the latter. RNase A_{ECP} data was taken from ref. (22).

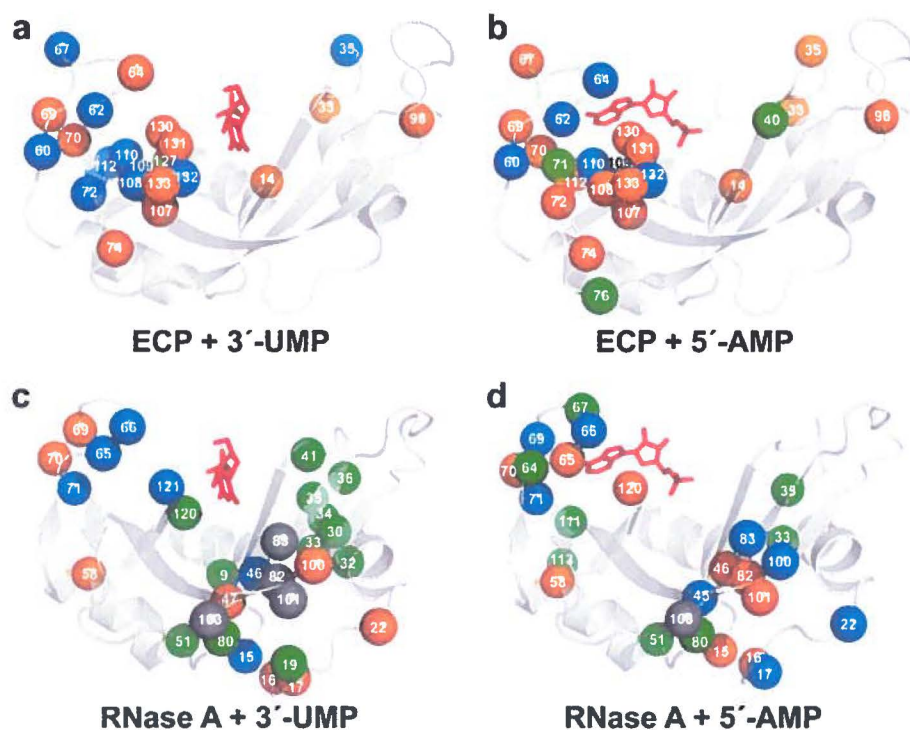


FIGURE 6. Conformational exchange induced by 3'-UMP and 5'-AMP binding to ECP and RNase A. Catalytic timescale (ms) residue dynamics were probed using ^{15}N -CPMG relaxation dispersion experiments for (a, c) 3'-UMP-bound and (b, d) 5'-AMP-bound forms of (a, b) ECP and (c, d) RNase A. Blue spheres: residues showing ^{15}N -CPMG dispersion profiles with a ΔR_2 ($1/\tau_{\text{cp}} > 2 \text{ s}^{-1}$); 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 [ΔR_2 ($1/\tau_{\text{cp}} > 2 \text{ s}^{-1}$)] upon ligand binding; grey 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 3D structure of ECP (PDB 1H1H) and RNase A (PDB 1O0N). The 3'-UMP and 5'-AMP ligands are shown in red.

SUPPORTING INFORMATION

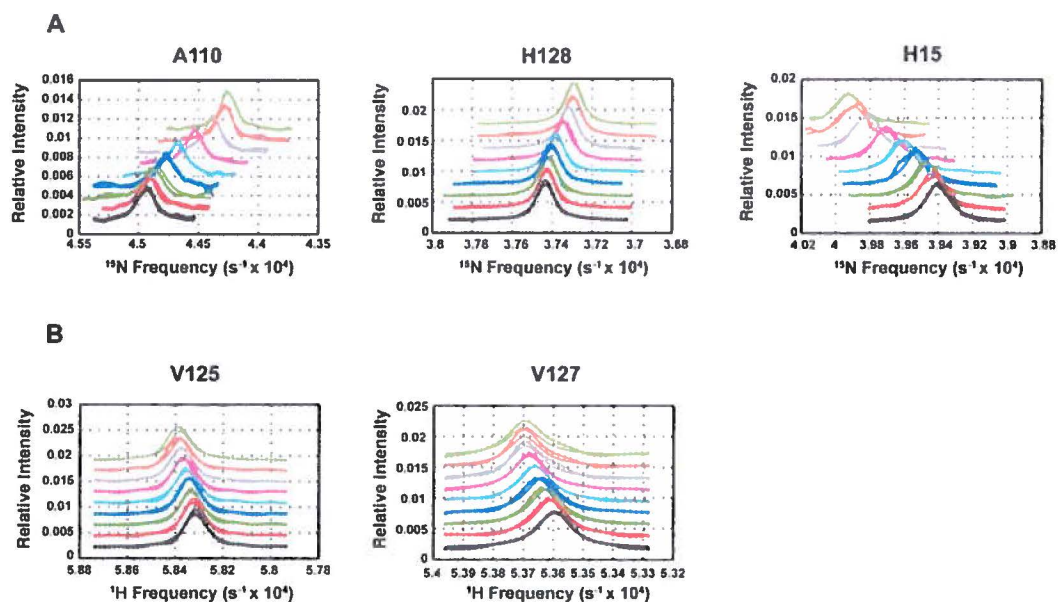


Figure S1. NMR line shape analysis of ECP titrations with 3'-UMP and 5'-AMP. The best-fit results for the ECP titrations with (a) 5'-AMP and (b) 3'-UMP are presented. One-dimensional slices extracted from experimental ^1H - ^{15}N HSQC spectra are shown with circles (connected with straight lines to 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 ± 0.02 , 0.39 ± 0.03 , 0.69 ± 0.05 , 1.31 ± 0.09 , 2.7 ± 0.2 , 6.0 ± 0.4 , 12.0 ± 0.8 , and 18 ± 1 .

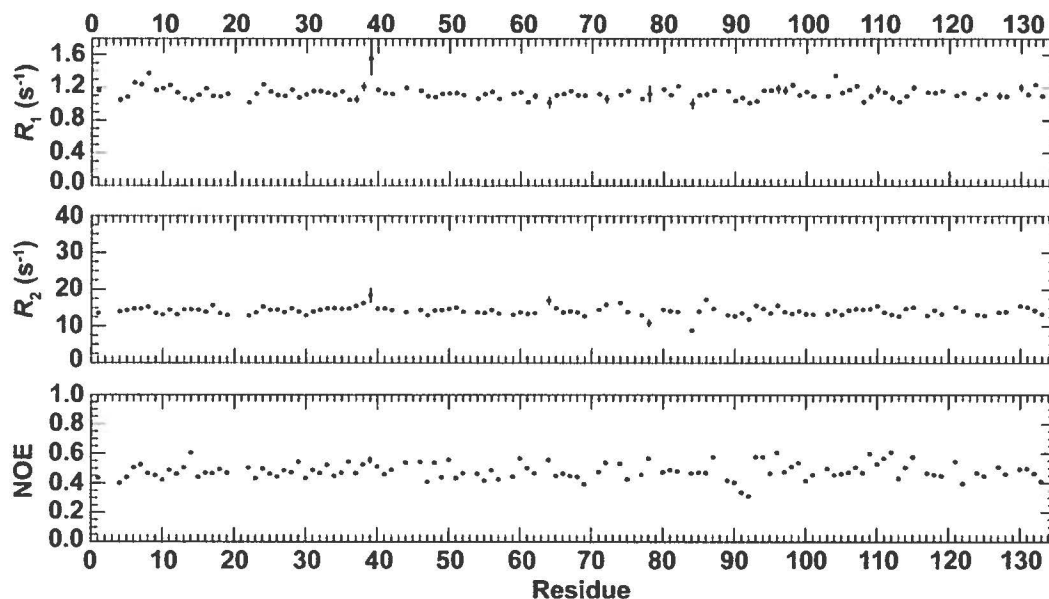


Figure S2. R_1 , R_2 and steady state heteronuclear NOE relaxation parameters plotted on the sequence of ECP. ^{15}N - R_1 , ^{15}N - R_2 and steady state heteronuclear NOE experiments were performed at in an interleaved fashion. The R_1 experiments were performed with relaxation delays of 10, 50, 100 ($\times 2$), 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}N -labelled ECP in 15 mM sodium acetate, 10% $^2\text{H}_2\text{O}$, pH 5.0 at 298K.

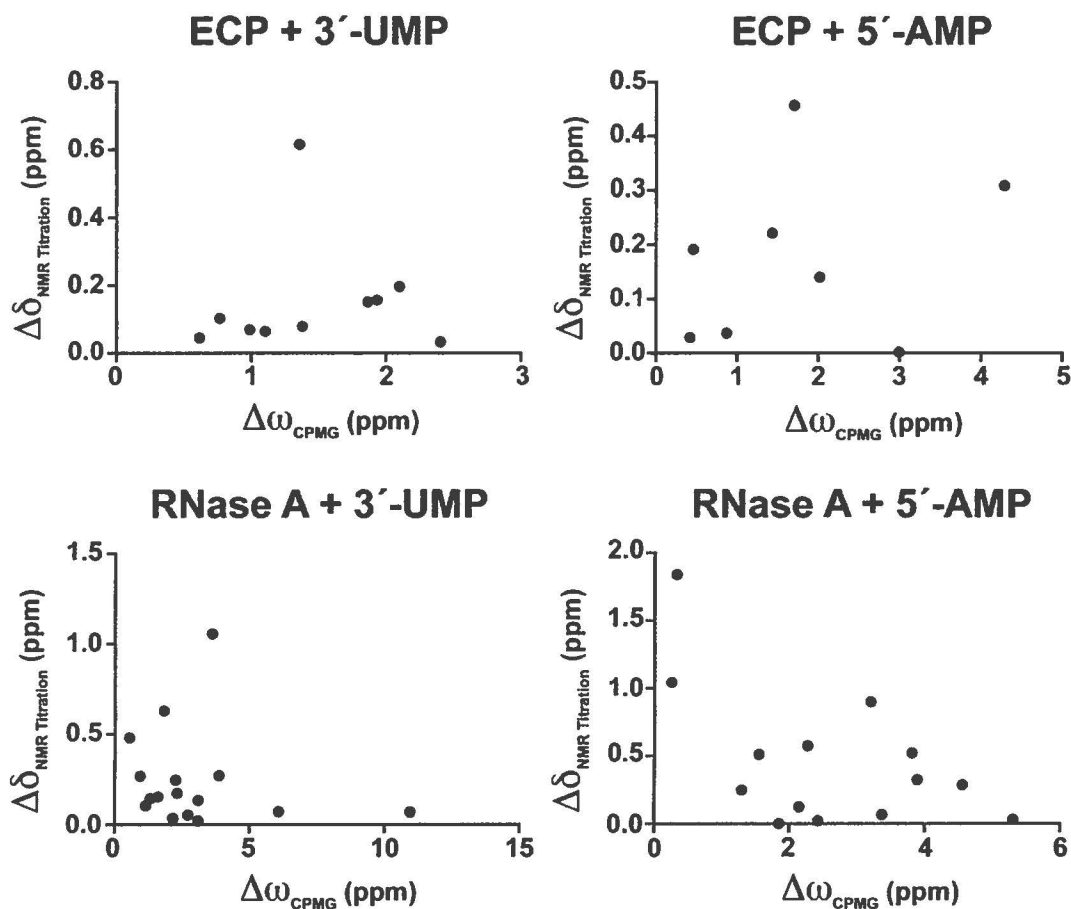


Figure S3. Comparison between the ^{15}N $\Delta\delta_{\text{NMR titration}}$, determined by the difference in the chemical shift between the apo and ligand bound forms of ECP (RNase A), and $\Delta\omega_{\text{CPMG}}$, determined by the difference in the chemical shifts of the ground and excited states of ECP (RNase A) extracted by ^{15}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}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}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}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}N chemical shift variations ($\Delta\delta$) compared to the conformational exchange behavior induced by 3'-UMP binding to ECP. (B) ^{15}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}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}N - $\Delta\delta > 0.5$ ppm upon 5'-AMP binding	Residues showing ^{15}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}N chemical shift variations ($\Delta\delta$) compared to the conformational exchange behavior induced by 5'-AMP binding to ECP.

(B) ^{15}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.

2.5. Conclusions et perspectives

L'étude ici présentée a permis de répondre aux hypothèses soulevées en introduction, à savoir qu'il existe chez des enzymes structurellement et fonctionnellement homologues des réseaux de résidus flexibles formés d'une hiérarchie de secteurs dynamiques contigus qui semblent encodés dans la structure et qui sont indépendants de la séquence primaire²⁶. Ces réseaux dynamiques transcendent les définitions classiques des structures secondaires et tertiaires, en plus d'être semblables entre les deux homologues structuraux et fonctionnels de la famille des RNases ici étudiés, suggérant l'existence de secteurs dynamiques évolutifs potentiellement conservés chez cette famille d'enzymes. Nous démontrons également que la perturbation de ces réseaux dynamiques conservés affecte la reconnaissance des ligands et qu'ils sont significativement perturbés lors de la liaison de ces ligands, suggérant que les réseaux dynamiques sont partie intégrante de la fonction biologique de ces enzymes. Nos études de mutagenèse nous ont également permis de construire un hybride entre la RNase A et ECP (remplacement de la boucle 1 dynamique de la RNase A par celle d'ECP), permettant de reproduire les réseaux dynamiques de ECP dans le contexte d'une chimère composée à 95% de la séquence de la RNase A. Ces résultats démontrent qu'il est possible de contrôler, de perturber et de moduler les réseaux dynamiques d'une protéine à l'aide de techniques de mutagenèse, ouvrant la porte au design et à l'ingénierie dynamique des protéines.

La présente étude s'avère évidemment partielle en ce qui a trait à la démonstration de la conservation évolutive de ces mouvements coordonnés chez toutes les RNases de vertébrés. Les étapes subséquentes doivent inévitablement être concentrées sur la comparaison dynamique de plusieurs autres homologues de cette famille enzymatique, non seulement en ce qui a trait aux autres membres humains (RNases 1-8), mais également pour plusieurs homologues fonctionnels retrouvés chez le rat, la souris, la grenouille, le poulet, le poisson, etc. À cet effet, seule une comparaison dynamique systématique de protéines effectuant la même action biologique (par exemple plusieurs versions vertébrées de ECP) permettront de faire ressortir la réelle conservation évolutive d'événements dynamiques liés à l'action biologique de ces enzymes. En effet, bien qu'elles conservent des structures tridimensionnelles et des activités catalytiques similaires, les différentes RNases demeurent malgré tout relativement distantes évolutivement³⁴. Des études sont actuellement en cours dans le laboratoire Doucet visant à la fois la comparaison dynamique de RNases humaines, mais également de plusieurs homologues biologiquement similaires retrouvés chez les vertébrés.

Des études d'inhibition spécifique de sites allostériques distants (notamment l'environnement de la boucle 1 chez la RNase A) permettront de vérifier l'importance de ces sites distaux dans la perturbation de l'activité catalytique de ces enzymes, assurant une modulation de la flexibilité qui n'est pas strictement dépendante de la mutagenèse (une technique évidemment inutilisable en milieu clinique). À cet effet, en raison de leur potentiel non-toxique et de leur grande application en milieu clinique, il pourrait être envisageable de considérer le design de peptides perturbant spécifiquement cet environnement allostérique. Des études semblables sont actuellement en cours sur d'autres systèmes enzymatiques dans le laboratoire Doucet, en collaboration avec l'équipe du professeur David Chatenet (INRS-Institut Armand-Frappier).

Finalement, à l'instar des travaux de Rama Ranganathan^{47, 48}, l'analyse systématique de corrélations entre la dynamique et la fonction chez des familles non-homologues de protéines est susceptible de fournir de nouveaux indices quant à la relation entre la structure, la fonction et la flexibilité chez l'ensemble des architectures protéiques, de manière à permettre une meilleure compréhension des secteurs dynamiques conservés dans l'ensemble du règne protéique. Cette

connaissance s'avérera nécessaire pour le succès de la conception *de novo* d'enzymes d'intérêts industriel et/ou pharmaceutique⁴⁹⁻⁵¹, en plus d'être essentielle au développement de nouveaux médicaments allostériques¹¹.

RÉFÉRENCES

1. Doucet, N., and Pelletier, J. N. (2009) Gaining insight into enzyme function through correlations with protein motions, In *Protein Engineering Handbook* (Lutz, S., and Bornscheuer, U. T., Eds.) 1st ed., pp 187-211, Wiley-VCH, Weinheim.
2. Agarwal, P. K. (2006) Enzymes: an integrated view of structure, dynamics and function, *Microb. Cell. Fact.* 5, 2.
3. Boehr, D. D., Dyson, H. J., and Wright, P. E. (2006) An NMR perspective on enzyme dynamics, *Chem. Rev.* 106, 3055-3079.
4. Liu, C. T., Stewart, R. J., and Benkovic, S. J. (2013) Too fast for catalysis, *Biocatal. Biotransform.* 31, 269-271.
5. Henzler-Wildman, K., and Kern, D. (2007) Dynamic personalities of proteins, *Nature* 450, 964-972.
6. Goodey, N. M., and Benkovic, S. J. (2008) Allosteric regulation and catalysis emerge via a common route, *Nat. Chem. Biol.* 4, 474-482.
7. Monod, J., Wyman, J., and Changeux, J. P. (1965) On the nature of allosteric transitions: a plausible model, *J. Mol. Biol.* 12, 88-118.
8. Mitternacht, S., and Berezovsky, I. N. (2011) A geometry-based generic predictor for catalytic and allosteric sites, *Protein Eng. Des. Sel.* 24, 405-409.
9. Gunasekaran, K., Ma, B., and Nussinov, R. (2004) Is allostery an intrinsic property of all dynamic proteins?, *Proteins* 57, 433-443.
10. Mathieu, V., Fastrez, J., and Soumillon, P. (2010) Engineering allosteric regulation into the hinge region of a circularly permuted TEM-1 β -lactamase, *Protein Eng. Des. Sel.* 23, 699-709.
11. Lee, G. M., and Craik, C. S. (2009) Trapping moving targets with small molecules, *Science* 324, 213-215.

12. Huang, R., Martinez-Ferrando, I., and Cole, P. A. (2010) Enhanced interrogation: emerging strategies for cell signaling inhibition, *Nat. Struct. Mol. Biol.* *17*, 646-649.
13. de Béthune, M. P. (2010) Non-nucleoside reverse transcriptase inhibitors (NNRTIs), their discovery, development, and use in the treatment of HIV-1 infection: a review of the last 20 years (1989-2009), *Antiviral Res.* *85*, 75-90.
14. Keov, P., Sexton, P. M., and Christopoulos, A. (2011) Allosteric modulation of G protein-coupled receptors: a pharmacological perspective, *Neuropharmacology* *60*, 24-35.
15. Martins, S., Ramos, M. J., and Fernandes, P. A. (2008) The current status of the NNRTI family of antiretrovirals used in the HAART regime against HIV infection, *Curr. Med. Chem.* *15*, 1083-1095.
16. Agarwal, P. K. (2005) Role of protein dynamics in reaction rate enhancement by enzymes, *J. Am. Chem. Soc.* *127*, 15248-15256.
17. Agarwal, P. K. (2014) A biophysical perspective on enzyme catalysis, *ACS Catal.*, In Press.
18. Cui, Q., and Karplus, M. (2008) Allostery and cooperativity revisited, *Protein Sci.* *17*, 1295-1307.
19. 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.
20. Doucet, N., and Loria, J. P. (2011) Relaxation dispersion studies of enzyme motions, In *Biomedical Applications of Nuclear Magnetic Resonance Spectroscopy* (Dingley, A. J., and Pascal, S. M., Eds.), pp 185-202.
21. 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.
22. Mittermaier, A. K., and Kay, L. E. (2009) Observing biological dynamics at atomic resolution using NMR, *Trends Biochem. Sci.* *34*, 601-611.

23. Peng, J. W. (2009) Communication breakdown: protein dynamics and drug design, *Structure* 17, 319-320.
24. Smock, R. G., and Gierasch, L. M. (2009) Sending signals dynamically, *Science* 324, 198-203.
25. Gagné, D., Charest, L. A., Morin, S., Kovrigin, E. L., and Doucet, N. (2012) Conservation of flexible residue clusters among structural and functional enzyme homologues, *J. Biol. Chem* 287, 44289-44300.
26. Gagné, D., and Doucet, N. (2013) Structural and functional importance of local and global conformational fluctuations in the RNase A superfamily, *FEBS J.* 280, 5596-5607.
27. Raines, R. T. (1998) Ribonuclease A, *Chem Rev* 98, 1045-1066.
28. Beach, H., Cole, R., Gill, M. L., and Loria, J. P. (2005) Conservation of mus-ms enzyme motions in the apo- and substrate-mimicked state, *J. Am. Chem. Soc.* 127, 9167-9176.
29. Cole, R., and Loria, J. P. (2002) Evidence for flexibility in the function of ribonuclease A, *Biochemistry* 41, 6072-6081.
30. 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.
31. 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.
32. 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.
33. 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. USA* 104, 11981-11986.
34. 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.

35. Gao, X., and Xu, Z. (2008) Mechanisms of action of angiogenin, *Acta Biochim. Biophys. Sin. (Shanghai)* 40, 619-624.
36. Yoshioka, N., Wang, L., Kishimoto, K., Tsuji, T., and Hu, G. F. (2006) A therapeutic target for prostate cancer based on angiogenin-stimulated angiogenesis and cancer cell proliferation, *Proc. Natl. Acad. Sci. USA* 103, 14519-14524.
37. Kover, K. E., Bruix, M., Santoro, J., Batta, G., Laurents, D. V., and 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.
38. Boix, E. (2001) Eosinophil cationic protein, *Methods Enzymol.* 341, 287-305.
39. Doucet, N., Jayasundera, T. B., Simonovic, 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.
40. 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.
41. 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.
42. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., Rance, M., and Skelton, N. J. (2007) *Protein NMR Spectroscopy: Principles and Practice*, Elsevier Academic, San Diego.
43. Mittermaier, A., and Kay, L. E. (2006) New tools provide new insights in NMR studies of protein dynamics, *Science* 312, 224-228.
44. Manley, G., and Loria, J. P. (2012) NMR insights into protein allostery, *Arch. Biochem. Biophys* 519, 223-231.
45. Palmer, A. G., 3rd, Kroenke, C. D., and Loria, J. P. (2001) Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules, *Methods Enzymol.* 339, 204-238.

46. Korzhnev, D. M., Religa, T. L., Banachewicz, W., Fersht, A. R., and Kay, L. E. (2010) A transient and low-populated protein-folding intermediate at atomic resolution, *Science* 329.
47. Halabi, N., Rivoire, O., Leibler, S., and Ranganathan, R. (2009) Protein sectors: evolutionary units of three-dimensional structure, *Cell* 138, 774-786.
48. Reynolds, K. A., McLaughlin, R. N., and Ranganathan, R. (2011) Hot spots for allosteric regulation on protein surfaces, *Cell* 147, 1564-1575.
49. Baker, D. (2010) An exciting but challenging road ahead for computational enzyme design, *Protein Sci.* 19, 1817-1819.
50. Davey, J. A., and Chica, R. A. (2012) Multistate approaches in computational protein design, *Protein Sci.* 21, 1241-1252.
51. Lassila, J. K. (2010) Conformational diversity and computational enzyme design, *Curr Opin. Chem. Biol.* 14, 676-682.