

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

**ÉTUDE DE LA RELATION STRUCTURE-FONCTION-
FLEXIBILITÉ DE GLYCOSIDE HYDROLASES DES FAMILLES
11 et 12 CHEZ *STREPTOMYCES LIVIDANS***

Par
Louise D. ROUX

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

Charles Gauthier
INRS-Institut Armand-Frappier

Examineur externe

Steve Bourgault
Département de chimie
UQAM

Directeur de recherche

Nicolas Doucet
INRS-Institut Armand-Frappier

REMERCIEMENTS

En préambule de ce mémoire, je souhaite remercier toutes les personnes qui m'ont aidée, soutenue et accompagnée tout au long de cette aventure. Tout d'abord, mes remerciements vont au Pr. Nicolas Doucet, mon directeur de recherche qui m'a permis d'effectuer ma maîtrise dans son laboratoire, de participer à de nombreux congrès et d'avoir mon nom dans des publications scientifiques. Merci pour sa patience et son implication dans ce projet, tout en me laissant une grande autonomie et en prenant compte de mes opinions.

Ensuite, je tiens particulièrement à remercier l'assistante de recherche Myriam Létourneau, pour sa grande aide et son accompagnement dans mes travaux de recherche et de rédaction.

Je remercie bien évidemment toute l'équipe du laboratoire Doucet : Philippe, David, Marie-Aude, Jacinthe, Andrés, Yossef, Chitra, Lian et Hang ; qui ont su m'intégrer dans leur groupe, m'aider à réaliser mon projet et me soutenir moralement quand il le fallait. Chacun d'entre eux mérite un remerciement spécifique et sera toujours le bienvenu dans ma famille. J'espère que vous continuerez à célébrer les anniversaires à grand coups de gâteaux, trouverez un nouveau « Mafia-bar » et surtout que vous respecterez votre semaine d'esclavage !

Je souhaite aussi remercier Dr. Sameer Al-Abdul-Wahid et Dr. Tara Sprules pour l'acquisition des expériences de RMN de mon projet et leurs réponses toujours rapides et pertinentes à mes nombreuses demandes et questions. De même, je remercie Nathalie Drouin et Mélissa Morin pour leur aide précieuse en purification de protéines et en réparation de HPLC/FPLC.

Enfin, je remercie chaleureusement toutes les personnes rencontrées à l'INRS pour leur gentillesse et leurs conseils.

Merci à tous.

RÉSUMÉ

Les xylanases sont des enzymes qui catalysent le clivage hydrolytique de la liaison glycosidique de xylanes, des hémicelluloses que l'on retrouve principalement dans les parois cellulaires des plantes et des algues. Malgré les nombreux efforts dévoués à l'amélioration catalytique des xylanases à des fins d'ingénierie enzymatique, en particulier pour leur utilisation industrielle, la compréhension restreinte de leurs propriétés moléculaires limite les progrès biotechnologiques. Les premières études structurales de plusieurs xylanases de la famille 11 (GH11) ont proposé l'existence d'un mouvement « ouvert-fermé » *via* une boucle près du site actif (le « pouce »), qui semble jouer un rôle dynamique majeur dans la liaison au substrat, la libération du produit et/ou la catalyse. Dans le but de vérifier cette hypothèse et de tester l'importance fonctionnelle de la dynamique de cette enzyme à l'échelle de temps de son événement catalytique (millisecondes), nous avons entrepris une étude de changements conformationnels chez la xylanase B2 de *Streptomyces lividans* à l'aide d'expériences de dispersion de relaxation (¹⁵N-CPMG) en spectroscopie RMN. Pour tester la conservation évolutive d'événements dynamiques impliqués dans la catalyse chez cette famille d'enzymes, nous avons également entrepris des analyses semblables chez un homologue structural de XlnB2, à savoir la xylanase C (XlnC) de *S. lividans*, qui possède 77 % d'identité de séquence et une activité catalytique similaire à XlnB2. Nos données préliminaires semblent suggérer que XlnC ne montre pas d'échange conformationnel dans sa forme libre, alors que la présence du ligand xylobiose (X2) induit l'apparition d'un nombre restreint de résidus qui subissent des mouvements à cette échelle de temps. De manière analogue, la cellulase B2 (CelB2) de *S. lividans* a également été étudiée pour sa structure similaire à celle des GH11 et son absence de boucle en forme de « pouce ». Tout comme chez XlnC, les premiers résultats obtenus avec CelB2 semblent indiquer que cette enzyme ne présente pas d'échange conformationnel à l'échelle de temps de la milliseconde. Nos résultats laissent entrevoir que cette famille d'enzymes dépend de profils dynamiques distincts, qui pourraient à la fois être tributaires de la présence ou de l'absence de domaines de liaison aux sucres et/ou de différents profils de reconnaissance et d'affinité envers divers substrats. Nos données démontrent que les repliements protéiques ne peuvent être systématiquement considérés équivalents au niveau de leur potentiel d'adaptation en ingénierie enzymatique, malgré leur forte similarité de séquence, de structure et de fonction. Ces résultats ouvrent la voie à une meilleure compréhension moléculaire de l'utilisation, de la modification et de la modulation d'activités enzymatiques en ingénierie des protéines.

ABSTRACT

Xylanases are enzymes that catalyze the hydrolytic breakdown of glycosidic bonds in xylan, a group of hemicelluloses found in the cell walls of plants and algae. Despite many efforts devoted to the catalytic improvement of xylanases — particularly for their industrial use — limited understanding of their molecular properties currently limits their biotechnological function. Previous structural studies performed on several family 11 xylanases (GH11) proposed the existence of an “open-closed” movement occurring *via* a loop near the active site (*i.e.* the “thumb loop”), which appears to play a major dynamic role in substrate binding, product release and/or catalysis. To test this hypothesis and to investigate the functional importance of enzyme flexibility occurring on the time scale of catalysis in xylanases (*i.e.* milliseconds), we performed conformational exchange studies of xylanase B2 from *Streptomyces lividans* using ¹⁵N-CPMG relaxation dispersion NMR spectroscopy. To examine the potential evolutionary conservation of dynamic events involved in catalysis within this enzyme family, we also examined structural homologs of XlnB2, primarily focusing on xylanase C (XlnC) from *S. lividans*, which conserves 75 % sequence identity with XlnB2 and displays similar catalytic activity. Our preliminary data suggests that XlnC does not exhibit conformational exchange in its free form, whereas the presence of a xylobiose (X2) ligand induces conformational exchange on a limited number of residues. Similarly, cellulase B2 (CelB2) from *S. lividans* was also investigated for its analogous structure and absence of a “thumb loop”. Similar to XlnC, preliminary results on CelB2 indicate that this enzyme does not experience conformational exchange on the millisecond time scale. Our results suggest that this enzyme family relies on distinct dynamic profiles, which may depend on the presence or absence of sugar binding domains and/or different recognition and affinity patterns towards various substrates. Our data demonstrate that protein folds cannot systematically be considered equivalent in terms of potential for enzymatic engineering adaptation, despite their strong similarity in sequence, structure and function. These results pave the way for a better molecular understanding of the use, modification, and modulation of enzymatic activities in protein engineering.

TABLE DES MATIÈRES

REMERCIEMENTS	II
RÉSUMÉ	III
ABSTRACT	IV
TABLE DES MATIÈRES	V
LISTE DES TABLEAUX.....	VIII
LISTE DES FIGURES	IX
LISTE DES ABRÉVIATIONS	X
CHAPITRE 1 : INTRODUCTION	11
1.1 <i>Biomasse</i>	11
1.2 <i>Enzymes impliquées dans la dégradation des polysaccharides</i>	11
1.3 <i>Glycoside hydrolases</i>	12
1.3.1 Structure.....	12
1.3.2 Mécanisme catalytique.....	13
1.4 <i>Xylanases</i>	15
1.4.1 Xylanases GH11.....	15
1.4.2 Spécificité de substrat des GH11.....	16
1.4.3 Structure et homologie de séquence des GH11	17
1.4.4 Propriétés dynamiques des GH11	19
1.4.5 Applications des GH11.....	21
1.5 <i>Cellulases</i>	23

1.5.1	Cellulases GH12.....	23
1.5.2	Comparaison cellulases GH12/xylanases GH11	24
1.6	<i>Relation structure, fonction et flexibilité des protéines</i>	25
1.7	<i>Projet de recherche</i>	26
1.7.1	Problématique	27
1.7.2	Hypothèses et Objectifs	27
1.8	<i>Méthodologie</i>	28
1.8.1	Expression et purification des enzymes	28
1.8.2	Caractérisation des protéines par résonance magnétique nucléaire.....	30
1.8.3	Caractérisation RMN de l'échange conformationnel à l'échelle de temps de la catalyse	32
CHAPITRE 2 : MÉTHODOLOGIES RMN ET COMPUTATIONNELLES APPLIQUÉES À L'ÉTUDE DE LA DYNAMIQUE DES PROTÉINES.....		35
2.1	<i>Contribution des auteurs</i>	35
2.2	<i>Résumé</i>	36
CHAPITRE 3 : LA LIAISON D'UN LIGAND INDUIT L'APPARITION D'ÉCHANGE CONFORMATIONNEL CHEZ LA XYLANASE B2 DE STREPTOMYCES LIVIDANS.....		47
3.1	<i>Contribution des auteurs</i>	47
3.2	<i>Résumé</i>	48
CHAPITRE 4 : RÉSULTATS COMPLÉMENTAIRES		69
4.1	<i>Xylanase C de S. lividans</i>	69
4.1.1	Attribution des résonances RMN de la chaîne principale	69
4.1.2	Effet de la liaison d'un ligand	70

4.1.3	Changements conformationnels à l'échelle de la milliseconde	70
4.2	<i>Cellulase B2 de S. lividans</i>	72
4.2.1	Attribution des résonances RMN de la chaîne principale	72
4.2.2	Effet de la liaison d'un ligand	73
4.2.3	Changements conformationnels à l'échelle de la milliseconde	73
CONCLUSION ET PERSPECTIVES		75
BIBLIOGRAPHIE		80

LISTE DES TABLEAUX

Tableau 5.1 : Principales caractéristiques des enzymes sélectionnées pour de potentielles futures études.....	79
---	----

LISTE DES FIGURES

Figure 1.1 : Topologie du site actif des GH.....	13
Figure 1.2 : Nomenclature des sous-sites d'arrimage chez les GH.....	13
Figure 1.3 : Mécanismes d'hydrolyse de la liaison glycosidique.....	14
Figure 1.4 : Liaisons accessibles pour l'hydrolyse des hétéroxylanes.....	16
Figure 1.5 : Représentation schématisée et diagramme topologique d'une GH11 typique.....	17
Figure 1.6 : Résidus du sillon catalytique d'une GH11 typique.....	19
Figure 1.7 : Mouvements proposés des GH11 sur la base de structures cristallines.....	20
Figure 1.8 : Superposition des structures d'une xylanase GH11 et cellulase GH12.....	24
Figure 1.9 : Carte du plasmide pIJ702 avec sites de restriction.....	28
Figure 1.10 : Chromatographie échangeuse de cations.....	30
Figure 1.11 : Échelle de temps des mouvements dynamiques à l'échelle atomique chez les protéines.....	33
Figure 1.12 : Exemples de courbes de dispersion de relaxation ^{15}N -CPMG pour un résidu dynamique.....	34
Figure 4.1 : Spectre ^1H - ^{15}N HSQC attribué de XlnC de <i>S. lividans</i>	69
Figure 4.2 : Résidus de XlnC présentant une variation de déplacements chimiques ^1H - ^{15}N ($\Delta\delta$) lors de la liaison au xylobiose (X2).....	70
Figure 4.3 : Résidus présentant de l'échange conformationnel chez XlnB2 et XlnC en absence et en présence de xylobiose (X2).....	71
Figure 4.4 : Spectre ^1H - ^{15}N HSQC attribué de CelB2 de <i>S. lividans</i>	72
Figure 4.5 : Résidus de CelB2 présentant une variation de déplacements chimiques ^1H - ^{15}N ($\Delta\delta$) lors de la liaison au cellobiose (C2).....	73
Figure 4.6 : Résidus présentant de l'échange conformationnel chez CelB2, en absence et en présence de cellobiose (C2).....	74
Figure 5.1 : Arbre phylogénétique des GH11 bactériennes.....	77

LISTE DES ABRÉVIATIONS

^{13}C	Carbone-13
C2	Cellobiose
CAZy	Carbohydrate-Active enZYmes Database (http://www.cazy.org/)
CelB2	Domaine catalytique de la cellulase B de <i>S. lividans</i>
CPMG	Carr-Purcell-Meiboom-Gill
$\Delta\delta$	Variation de déplacements chimiques
GH	Glycoside hydrolase (ou glycosidase)
^1H	Hydrogène (proton)
HSQC	Heteronuclear Single Quantum Coherence
k_{cat}	Constante catalytique
K_{d}	Constante de dissociation
kDa	Kilodaltons
^{15}N	Azote-15
NOE	Nuclear Overhauser Effect
PDB	Protein Data Bank
pI	Point isoélectrique
R_1	Relaxation longitudinale
R_2	Relaxation transversale
RMN	Résonance magnétique nucléaire
<i>S. lividans</i>	<i>Streptomyces lividans</i>
Tx-Xyl	Xylanase GH11 de <i>Thermobacillus xylanilyticus</i>
X2	Xylobiose
XInB2	Domaine catalytique de la xylanase B de <i>S. lividans</i>
XInC	Domaine catalytique de la xylanase C de <i>S. lividans</i>

CHAPITRE 1 : INTRODUCTION

1.1 Biomasse

Le terme 'biomasse' désigne l'ensemble des matières organiques dérivées d'organismes vivants, majoritairement d'origine végétale (algues incluses), mais aussi d'origine animale ou fongique. Environ la moitié du carbone est incorporé dans la lignocellulose, qui compose la paroi des cellules végétales et le bois. Les plantes elles-mêmes ne réutilisent pas le carbone stocké dans les parois cellulaires, mais le laissent dans les feuilles mortes et la litière végétale. En raison des grandes quantités de biomasse végétale dans la biosphère, un large éventail d'organismes dégradant les plantes ont évolué pour utiliser la cellulose et l'hémicellulose comme source de carbone et d'énergie, recyclant ainsi le carbone et maintenant son cycle sur terre (1).

La biomasse est utilisée par l'être humain depuis qu'il maîtrise le feu. Elle demeure la première énergie renouvelable utilisée dans le monde, notamment pour le chauffage et la cuisson des plats de cuisine, mais essentiellement dans les pays peu industrialisés. En revanche, l'énergie tirée de la biomasse intéresse à nouveau les pays riches, qui sont confrontés au dérèglement climatique et à la perspective d'une crise des ressources en hydrocarbures fossiles. Il s'agit d'une filière en développement rapide, y compris sous des formes industrielles avec les biocarburants. Son introduction dans les systèmes énergétiques contribue à réduire (pour ce qui est du bilan global) les émissions de gaz à effet de serre, voire à restaurer certains puits de carbone. La biomasse est la plus grande et attrayante source d'énergie renouvelable, représentant 14 % de la demande totale d'énergie dans le monde (1).

1.2 Enzymes impliquées dans la dégradation des polysaccharides

La dégradation de la biomasse des plantes nécessite l'action concertée d'une multitude d'enzymes extracellulaires sécrétées par des organismes microbiens, champignons et bactéries, principalement des carbohydrases, dont la majorité sont des cellulases. En effet, la cellulose, polymère linéaire de β -(1,4)-D-glucose, est le polysaccharide le plus abondant sur terre, pouvant représenter 45 % de la matière sèche du bois. Les hémicelluloses viennent ensuite, pouvant constituer plus de 30 % de la matière sèche des parois végétales (2). Par rapport à la cellulose, l'hémicellulose ne contient pas que des unités de glucose, mais aussi du xylose, du mannose,

du galactose, du rhamnose ou de l'arabinose. Un intérêt grandissant est porté sur les enzymes qui dégradent ou favorisent la dégradation des hémicelluloses, car de nombreuses applications potentielles existent dans les industries agro-alimentaires, du papier ou pour la conversion en bio-carburants de ces composés à forte valeur ajoutée.

1.3 Glycoside hydrolases

Les glycoside hydrolases (ou glycosidases, GH) sont les enzymes actives sur les disaccharides, oligosaccharides et polysaccharides les mieux caractérisées. Elles catalysent l'hydrolyse enzymatique de la liaison glycosidique entre au moins deux groupements glucidiques ou entre un glucide et un groupement non-glucidique (3).

Les GH sont classées en plusieurs familles en fonction des similarités de séquence d'acides aminés, chaque famille reflétant ainsi une homologie de structure des protéines et de mécanisme catalytique, ainsi que les relations évolutives entre ces enzymes. Ce système de classification est disponible depuis 1998 à partir de la base de données en ligne CAZy (www.cazy.org) (4). À ce jour, la base de données CAZy fournit une liste de 145 familles de glycoside hydrolases.

1.3.1 Structure

La structure des protéines étant souvent mieux conservée que leur séquence, certaines familles peuvent être regroupées en clans, qui décrivent l'architecture de repliement des protéines. Par exemple, le clan GH-A est composé de 19 familles qui partagent le même repliement « $(\beta/\alpha)_8$ », alors que les clans GH-B et GH-C comprennent les enzymes avec un repliement en « β -jelly roll » (5).

Selon la spécificité de substrat et les mécanismes catalytiques, les glycosidases possèdent différentes conformations de leur site actif (Figure 1.1) (6). Dans la plupart des cas, un site actif organisé en sillon profond ou en tunnel est efficace pour des enzymes qui adoptent des mécanismes processifs lors de la dégradation des polysaccharides, leur permettant de libérer le produit tout en restant fermement liées à la chaîne de polysaccharides par plusieurs sous-sites. Dans les GH processives, les sillons profonds ou les tunnels sont généralement pavés de résidus aromatiques qui sont importants et assurent la liaison et le glissement du polymère à travers le

site actif. Les GH non processives, quant à elles, présentent un site actif sous forme de sillon plus ouvert ou de poche avec moins de résidus aromatiques (6).

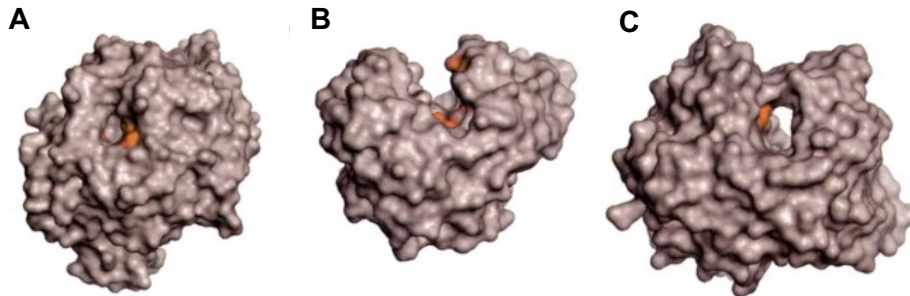


Figure 1.1 : Topologie du site actif des GH. A) poche, B) sillon, C) tunnel. Les résidus catalytiques sont en orange. D'après Davies and Henrissat, 1995 (6).

Davies *et al.* présentent les GH avec au moins cinq sous-sites, étiquetés -3, -2, -1 et +1, +2, de l'extrémité non-réductrice à l'extrémité réductrice du substrat, qui accueillent un long substrat à l'intérieur du site actif (Figure 1.2) (7). Selon cette nomenclature, le clivage du substrat se produit entre les sous-sites -1 et +1, quel que soit le nombre de sous-sites.

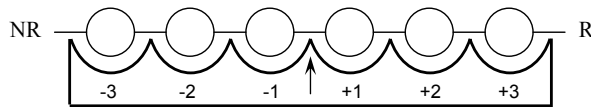


Figure 1.2 : Nomenclature des sous-sites d'arrimage chez les GH. De -3 à +3 de l'extrémité non-réductrice (NR) à l'extrémité réductrice (R). La flèche entre les sous-sites -1 et +1 indique la zone de clivage du substrat. D'après Davies *et al.*, 1997 (7).

1.3.2 Mécanisme catalytique

Les GH peuvent être classées comme endo- ou exo-enzymes, ce qui correspond à leur aptitude à cliver un substrat au milieu ou à la fin d'une chaîne (le plus souvent l'extrémité non-réductrice) (5). Les enzymes étudiées ici sont toutes des endoglucanases.

L'hydrolyse enzymatique de la liaison glycosidique a lieu par catalyse acide générale qui nécessite deux résidus : un donneur de proton acide/base et un nucléophile, habituellement des acides glutamiques ou aspartiques. Elle se produit par l'intermédiaire de deux mécanismes

majeurs donnant lieu soit à une rétention globale, soit à une inversion de la configuration anomérique, tel que décrit à l'origine par Koshland (Figure 1.3) (8).

Le mécanisme de rétention est un mécanisme en deux étapes. Chaque étape passe par un état de transition oxocarbénium. Le donneur de proton et le nucléophile sont situés à 5,5 Å l'un de l'autre. Dans la première étape appelée glycosylation, le résidu nucléophile attaque le carbone anomérique, tandis que l'acide cède son proton à l'oxygène glycosidique, clivant ainsi la liaison glycosidique. Dans la seconde étape ou dé-glycosylation, une molécule d'eau hydrolyse l'intermédiaire glycosylé, générant un produit avec la même stéréochimie que le substrat (6).

Le mécanisme d'inversion est un mécanisme en une étape, impliquant aussi un état de transition oxocarbénium. La réaction se produit avec l'assistance d'un donneur de proton et d'un nucléophile, qui sont généralement situés aux alentours de 6 à 11 Å l'un de l'autre. Dans ce mécanisme, la protonation de l'oxygène glycosidique et le départ de l'aglycone sont accompagnés par une attaque simultanée d'une molécule d'eau qui est activée par le nucléophile. Cette substitution nucléophile donne un produit unique avec une stéréochimie opposée au substrat (6).

Ce projet de maîtrise se concentre sur les xylanases de la famille GH11 et CelB2, une cellulase de la famille GH12. Ce sont des enzymes processives avec un mécanisme de rétention de l'anomérisation.

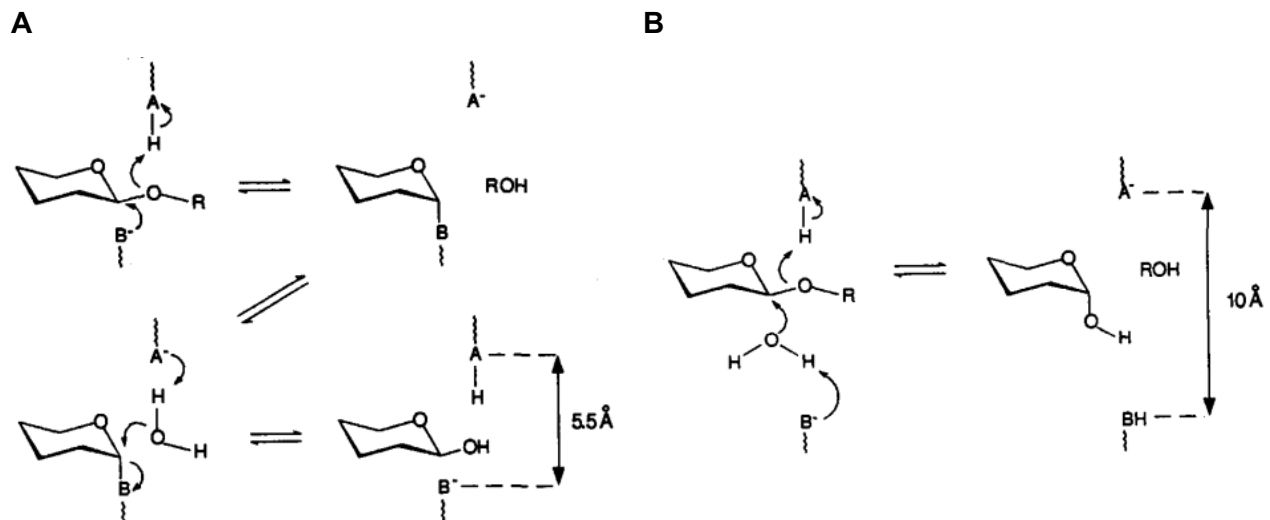


Figure 1.3 : Mécanismes d'hydrolyse de la liaison glycosidique. A) mécanisme avec rétention de l'anomérisation, B) mécanisme avec inversion de l'anomérisation. D'après Davies and Henrissat, 1995 (6).

1.4 Xylanases

Les xylanases (β -1,4-D-xylane xylanohydrolase, E.C. 3.2.1.8) sont les glycosidases qui hydrolysent la liaison glycosidique entre deux résidus D-xylopyranosyles liés en β -(1,4) dans l'unité xylane. Elles sont produites par de nombreux organismes, incluant des bactéries, algues, champignons, protozoaires, gastéropodes et arthropodes. La vaste diversité des xylanes a conduit à une grande variété de xylanases nécessitant une classification différente du simple système par spécificité de substrat. Depuis qu'une corrélation entre la masse moléculaire (MW) et le point isoélectrique (pI) a été découverte, une classification selon les propriétés physico-chimiques des xylanases a été proposée. Les xylanases sont ainsi divisées en deux groupes : la famille F, avec les membres possédant un haut poids moléculaire (>30 kDa) et un pI acide ; et la famille G, contenant les xylanases de faible poids moléculaire (<30 kDa) et un pI basique (9). Cependant, il existe des exceptions, particulièrement pour les xylanases nouvellement découvertes chez les champignons, qui ne peuvent être classifiées selon ce système.

L'utilisation du système CAZy, basé sur les similarités de séquence des domaines catalytiques, permet la hiérarchisation de toutes les xylanases. Il est donc devenu le système standard pour la classification de ces enzymes (4). Les xylanases sont retrouvées dans différentes familles de GH, mais les endo- β -1,4-xylanases avec un domaine catalytique unique sont limitées aux familles 5, 8, 10, 11 et 30. Ces enzymes diffèrent essentiellement par leur structure et leur spécificité de substrat. Les GH5, GH10 et GH30 affichent le même repliement « $(\beta/\alpha)_8$ », mais hydrolysent différents substrats, tandis que les xylanases GH8 présentent un repliement « $(\alpha/\alpha)_6$ » et ont une spécificité de substrat proche de celle des GH10. Enfin, les membres de la famille GH11, qui sont les plus petites xylanases avec une structure en « β -jelly-roll », sont hautement spécifiques et tolèrent mal les substitutions sur le squelette de xylane (10).

1.4.1 Xylanases GH11

Les xylanases GH11 sont souvent divisées en xylanases acides et xylanases alcalines en fonction de leur pH optimal (pH_{opt}) qui, malgré les similitudes structurales et de séquence, varie considérablement (pH de 2 à 11) (11). Plusieurs études ont démontré que le pH optimal des xylanases est en corrélation avec la nature du résidu à la position 46 (à moins d'indication contraire, tous les numéros de résidus proviennent de la numérotation de la xylanase B2 de

Streptomyces lividans, XlnB2). Pour les xylanases qui fonctionnent de manière optimale dans des conditions de pH < 5 (xylanases dites acides), ce résidu est un aspartate (Asp), alors que dans celles qui fonctionnent de manière optimale dans des conditions de pH ≥ 5, (xylanases dites alcalines), il s'agit d'une asparagine (Asn). En effet, en raison de sa proximité avec le résidu catalytique acide/base, le résidu Asp peut influencer l'état d'ionisation en créant une liaison hydrogène forte avec l'acide/base, ce qui rend plus facile la libération d'un proton (11,12). Les xylanases GH11 de bactéries ne présentent pas de pH optimal inférieur à 5,5.

En plus de la diversité des pH optimaux, les xylanases GH11 présentent des variations dans leur thermostabilité et leur thermoactivité. Les membres de la famille GH11 présentent des températures optimales (T_{opt}) allant de 35 °C à 85 °C, correspondant à des enzymes mésophiles, thermophiles et même hyper-thermophiles (13).

1.4.2 Spécificité de substrat des GH11

Les xylanases GH11 sont spécifiques aux hétéroxylanes. Elles sont incapables de cliver la cellulose ou des substrats dérivés de la cellulose. Elles clivent de préférence dans les régions non substituées, car elles ne peuvent pas attaquer une liaison xylosidique à partir de l'extrémité non-réductrice à côté d'un xylose ramifié (14,15). Par conséquent, les hétéroxylanes avec un haut degré de substitution ne sont pas complètement hydrolysés par les xylanases GH11, qui nécessitent trois résidus de xylose consécutifs non substitués pour l'hydrolyse (Figure 1.4). Cette observation traduit la relation entre la spécificité de substrat et le sillon étroit des GH11, par rapport à celle des GH10 par exemple, qui ont un sillon largement ouvert et plus grand (10,13,14).

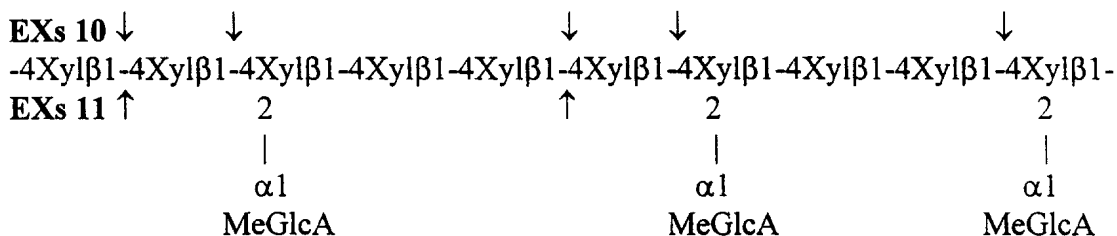


Figure 1.4 : Liaisons accessibles pour l'hydrolyse des hétéroxylanes. Les flèches indiquent les zones de clivage pour les xylanases GH10 (EXs10) ou 11 (EXs11). D'après Biely *et al.*, 1997 (14).

Les principaux produits non substitués libérés par les xylanases GH11 à partir des hétéroxylanes sont le xylobiose et le xylotriose. Les xylanases ne montrent aucune activité détectable sur ces produits et ont seulement une très faible activité sur le xylotétraose, tandis que le xylopentaose et xylohexaose sont clivés rapidement (15,16).

1.4.3 Structure et homologie de séquence des GH11

Les xylanases GH11 sont constituées d'un seul domaine catalytique avec une structure en « β -jelly-roll » qui ressemble à la forme d'une main droite partiellement fermée. Il se compose de deux feuillets β antiparallèles (A et B) tordus à presque 90° et d'une hélice α à trois tours (Figure 1.5). Les faces hydrophobes des deux feuillets β , placées l'une contre l'autre, constituent les « doigts », tandis que la partie tordue du feuillet B et l'hélice α forment la « paume », tandis que la partie tordue du feuillet B et l'hélice α forment la « paume » de la main. Le site actif avec les deux glutamates catalytiques est situé sur le côté concave de la paume. La longue boucle entre les brins B7 et B8 est appelée le « pouce », tandis que la boucle entre les brins B6 et B9, reliant les doigts avec la base du « pouce » et fermant partiellement le site actif, est appelée « cordon » (17).

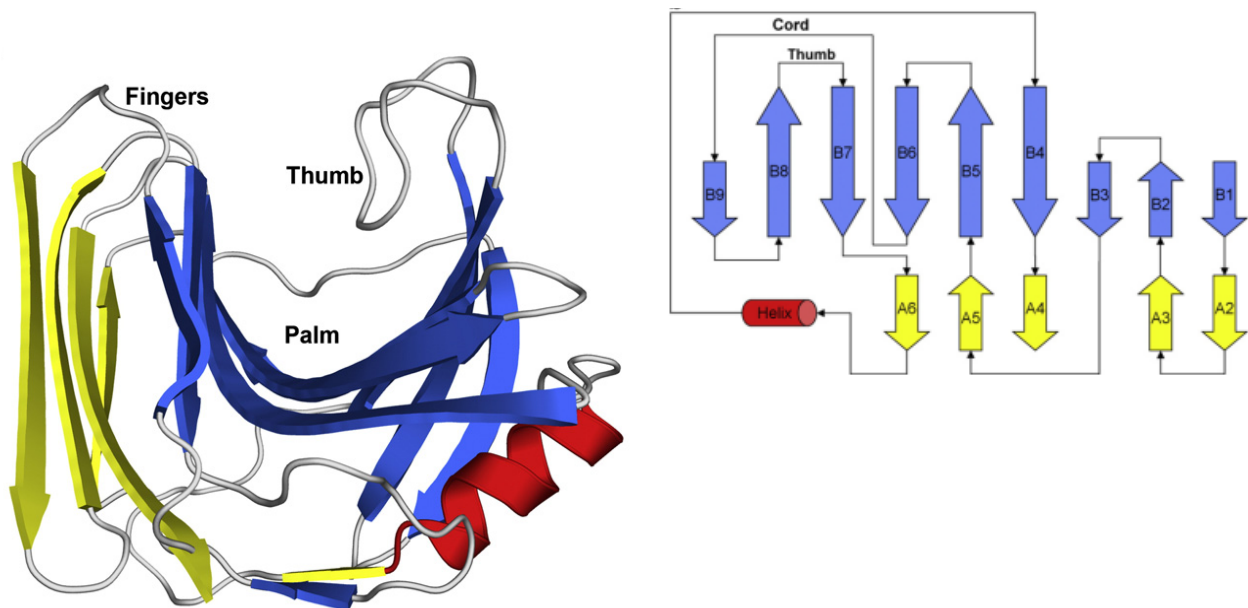


Figure 1.5 : Représentation schématisée (gauche) et diagramme topologique (droite) d'une GH11 typique : Tx-Xyl de *Thermobacillus xylanilyticus* (Uniprot ID : Q14RS0, structure résolue, mais non déposée dans RCSB PDB). D'après Paës *et al.*, 2012 (13).

Quelques variations par rapport à cette architecture standard existent. La plus commune est un feuillet β manquant (B1) à l'extrémité N-terminale. Les segments de plus haute similarité sont retrouvés dans des régions définies de la structure secondaire : les feuillets B5, B6, B8 et l'extrémité carboxyle de l'hélice alpha (11).

La région du « pouce », avec sa forme demi-pliée caractéristique, est très bien conservée parmi les xylanases GH11, avec une longueur de 11 résidus. Son sommet contient la séquence consensus PSIXG (conservée à plus de 90 %), où X peut correspondre à n'importe quel résidu acide aminé (11). Bien que l'ensemble des structures tridimensionnelles des xylanases GH11 soient hautement conservé, plusieurs insertions et délétions peuvent se produire au niveau du cordon. Plus particulièrement, sa longueur et sa conformation peuvent différer de façon significative. Le cordon se compose de cinq à dix résidus, dont une proline (Pro) conservée à 80 %, susceptible de jouer un rôle important (11,17).

En accord avec un mécanisme endo et une affinité pour les longs polymères de xylane non substitués, le sillon catalytique des xylanases GH11 est profond (~9 Å), étroit (4 Å) et long (~24 Å), composé pour la majorité d'au moins cinq sous-sites de -2 à +3 (14). Le site actif comprend principalement des résidus aromatiques et quelques résidus polaires appartenant aux brins B2, B3, B4 d'un côté et au brin B8 et au « pouce » de l'autre côté (Figure 1.6). Parmi les résidus aromatiques, beaucoup sont hautement conservés (à plus de 84 %, dont Trp20, Tyr74, Tyr78, Trp80 et Tyr89), tandis que d'autres semblent plus variables, notamment Tyr137 et Tyr179, qui peuvent être remplacés par un autre résidu aromatique, respectivement Phe et Trp. Les deux résidus catalytiques (Glu87 et Glu177) sont les mieux conservés (à 97 %). Le résidu en position 48, à côté de l'acide/base, est principalement occupé par une Val mais peut aussi correspondre à une Leu, un résidu de structure apparentée (13).

Dans l'ensemble, la structure des xylanases GH11 est hautement conservée, à la fois en séquence et dans l'espace tridimensionnel. Ceci est particulièrement observable au niveau du sillon catalytique, où plusieurs résidus hautement conservés dans les sous-sites -1 et +1 agissent de manière concertée pour affiner les propriétés des glutamates et stabiliser l'état de transition oxocarbénium.

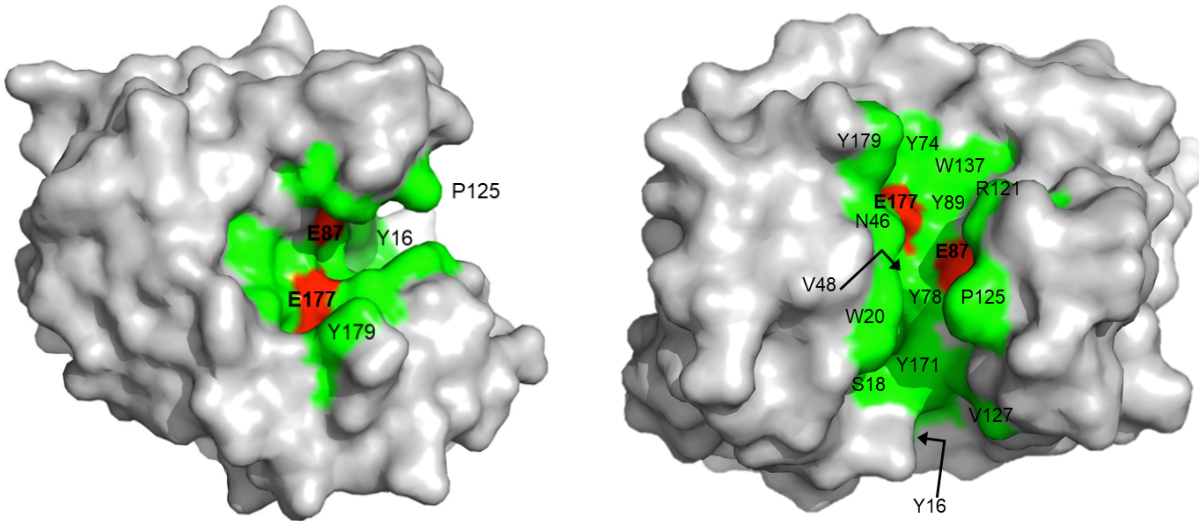
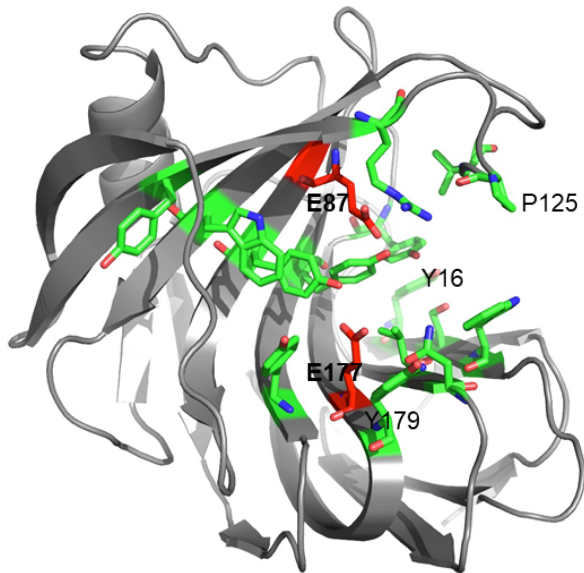


Figure 1.6 : Résidus du sillon catalytique d'une GH11 typique : XlnB2 de *Streptomyces lividans* (Uniprot ID : P26515, RCSB PDB : 5EJ3).

Les résidus du sillon catalytique sont en vert tandis que les deux résidus catalytiques sont en rouge.

En haut : représentation de surface. En bas : représentation « cartoon » avec les résidus du sillon catalytique représentés en bâtons.



1.4.4 Propriétés dynamiques des GH11

Plusieurs modèles de xylanases GH11, obtenus par le biais des structures cristallines (18-20), de simulations de dynamique moléculaire (19-21) et de quelques données provenant d'expériences de RMN (22,23), ont déjà démontré la présence d'une dynamique chez divers membres de cette famille d'enzymes.

Les xylanases GH11 peuvent exister sous trois conformations différentes qui suivent les coordonnées de la réaction catalytique. En premier lieu, l'enzyme est en conformation de liaison B (pour « Bound »), dans laquelle le ligand peut se fixer à l'enzyme, puis elle adopte la

conformation fermée C (pour « Closed »), durant laquelle le sillon catalytique est partiellement fermé pendant la catalyse. Finalement, la conformation relâchée L (pour « Loose ») est observée, permettant le relargage des produits (Figure 1.7) (19). De plus, une étude computationnelle a suggéré qu'une augmentation de la température pouvait induire des changements significatifs dans la conformation du « pouce ». Sous la température optimale, la boucle du « pouce » et la « paume » adopteraient une conformation fermée, tandis qu'à cette température les deux domaines se sépareraient, facilitant ainsi l'accès du substrat au site actif (20). Ce mouvement varierait aussi en fonction de la présence ou de l'absence du substrat : en absence de substrat, la xylanase présenterait des mouvements dépendants de la température, alors qu'en présence du substrat, le « pouce » limiterait l'accès au site actif, peu importe la température (21).

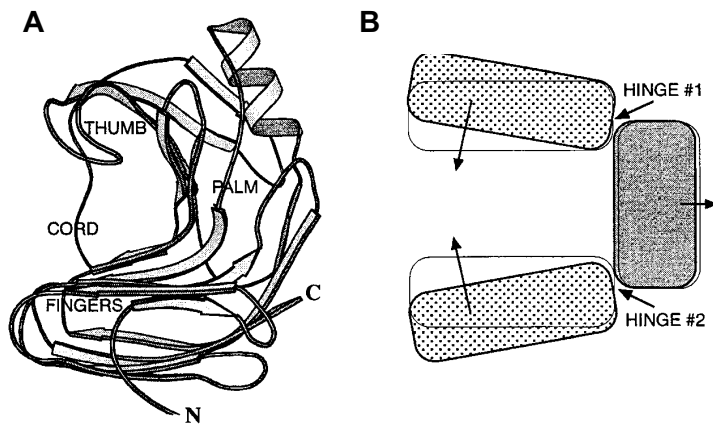


Figure 1.7: Mouvements proposés des GH11 sur la base de structures cristallines. A) Représentation « cartoon », B) Dessin schématique simplifié des différents éléments de structure. Les positions des deux charnières putatives sont indiquées. D'après Muilu *et al.*, 1998 (19).

Ces résultats tendent à suggérer que la boucle du « pouce » joue un rôle majeur dans la liaison du substrat et/ou la catalyse chez la xylanase GH11, tout en prenant compte du fait que ces résultats sont le fruit de calculs de dynamique moléculaire et non de démonstrations expérimentales. Afin d'étudier cette structure, des expériences de mutagenèse *in vitro* ont été effectuées sur la xylanase GH11 de *Thermobacillus xylanilyticus* (Tx-Xyl) (24). Un repositionnement du « pouce » grâce à la suppression des résidus Tyr111_{Tx-Xyl} et Thr121_{Tx-Xyl} (Val122 et Thr132 chez XlnB2) a produit des effets différents. La suppression de Thr121_{Tx-Xyl} a eu des conséquences dramatiques sur l'activité, alors que la suppression de Tyr111_{Tx-Xyl} a seulement affectée le k_{cat} . La mutagenèse a aussi été utilisée pour créer un variant sans « pouce », presque inactif, mais qui conserve la capacité à lier les substrats (xylotétraose et xylopentaose). En outre, à la différence de la Tx-Xyl sauvage, le mutant sans « pouce » a lié efficacement le cellotétraose, même si aucune activité d'hydrolyse de la cellulose n'a été

détectée. Dans l'ensemble, ces données montrent que le « pouce » est un facteur déterminant pour le choix du substrat et qu'il joue un rôle dans le processus catalytique. Une autre expérience de mutagenèse dirigée, ciblant la séquence conservée à l'extrémité du « pouce » Pro-Ser-Ile pour en faire un nouveau triplet Pro-Gly-Cys, a augmenté le k_{cat} de l'enzyme de 20 %. Un travail similaire sur la xylanase GH11 de *Penicillium griseofulvum* (PgXynB) a conduit à un mutant faisant preuve d'une efficacité catalytique augmentée vis-à-vis des arabinoxylanes et xylo-oligosaccharides (25). Ces mutations ont été choisies pour améliorer la flexibilité du « pouce », d'où la promotion de l'activité de l'enzyme. Une autre indication de la souplesse structurelle du « pouce » a été découverte chez la xylanase de *Bacillus subtilis*, où la présence d'une mutation spécifique a permis d'observer le pouce dans une conformation ouverte. La structure cristalline révèle une grande amplitude (plus de 15 Å) pour le mouvement d'ouverture-fermeture du « pouce » (26). Il a également été suggéré que les variations de conformation similaires peuvent exister dans d'autres xylanases GH11 (13).

1.4.5 Applications des GH11

Les hétéroxylanes et homoxylanes sont parmi les polysaccharides les plus abondants dans les plantes. Les xylanases microbiennes, qui peuvent cliver le squelette interne de xylane, ont gagné un intérêt constant au cours de ces dernières décennies dans de nombreuses applications, notamment en biotechnologie alimentaire, mais également dans l'industrie du papier et la production de biocarburants (13).

En biotechnologie alimentaire, les xylanases GH11 sont intéressantes dans les industries relatives aux céréales, comme la boulangerie et l'industrie brassicole. Des mélanges de xylanases de plusieurs familles peuvent être utilisés pour promouvoir la qualité du pain, notamment en augmentant la durée de conservation et le volume du pain, l'amélioration de la structure de la mie et la réduction du caractère collant de la pâte (27). Les GH11 ainsi que d'autres familles de xylanases sont également utilisées dans le processus de fabrication de la bière, en favorisant entre autres l'augmentation des performances de filtration du moût de bière pour diminuer la fraction visqueuse d'arabinoxylanes contenue dans le liquide (28). En plus des biotechnologies céréalières, les mélanges de xylanases sont utilisés pour clarifier les jus de fruits. Par exemple, un cocktail de xylanases de *S. sclerotiorum* aide à la clarification du jus en libérant des oligo- et monosaccharides et diminue jusqu'à 27 % le niveau de matières insolubles, donnant un meilleur rendement et facilitant la filtration du jus (29).

Une autre application économiquement importante des xylanases GH11 traite de l'alimentation animale. La plupart du temps, en raison de leur potentiel de synergie, des cocktails de xylanases sont utilisés comme suppléments alimentaires chez les volailles et ruminants pour améliorer les performances de croissance, l'assimilation des aliments, les paramètres immunitaires et la microflore intestinale (30,31).

Les enzymes robustes, notamment les GH11, sont d'une importance particulière dans l'industrie du papier pour réduire les quantités de chlore et le dioxyde de chlore utilisées dans le bio-blanchiment de la pâte à bois. Un prétraitement par un mélange de xylanases peut diminuer l'utilisation de produits chimiques de blanchiment par 10 à 20 % et se traduit généralement par une blancheur finale supérieure (32). De plus, la xylanase GH11 thermostable convient bien aux procédés à haute température rencontrés dans l'industrie des fibres et du papier. La faible masse moléculaire relative des GH11 peut aussi aider à pénétrer dans la paroi cellulaire interne, entraînant une meilleure suppression des polymères de lignine de la pâte à bois (33).

L'une des applications importantes des GH11 est leur utilisation dans la production de biocarburants. La conversion de la biomasse lignocellulosique issue des résidus de maïs, de la bagasse de la canne à sucre et des graminées produit un mélange de glucose, de xylose et d'arabinose qui peut encore être fermenté pour produire du bioéthanol comme biocarburant de deuxième génération (13,34). Dans cette production, les xylanases peuvent casser la structure de la matrice lignocellulosique et dépolymériser les polymères en sucres qui sont ensuite fermentés en éthanol, généralement par des levures. L'éthanol doit ensuite passer par des étapes de distillation, rectification et déshydratation pour répondre aux spécifications du carburant. Les résidus de maïs, la bagasse de la canne à sucre et les graminées sont des sources potentielles de biocarburants en raison de leur rendement élevé, de leur exigence minimale en éléments nutritifs et en eau, et le fait qu'ils peuvent être cultivés sans être en concurrence avec les cultures alimentaires actuelles.

Afin de produire un cocktail enzymatique optimisé de xylanases pour les applications mentionnées ci-dessus, il sera nécessaire de mieux comprendre la relation fonction/structure/flexibilité des GH11 au niveau moléculaire.

1.5 Cellulases

Les cellulases hydrolysent les liaisons β -(1,4) entre les résidus β -D-glucose de la cellulose et appartiennent à plusieurs familles de GH, mais sont traditionnellement divisées en deux groupes : les endoglucanases (EC 3.2.1.4), qui peuvent hydrolyser les liens internes des chaînes de cellulose, et les cellobiohydrolases (EC 3.2.1.91), qui agissent préférentiellement sur l'extrémité des chaînes. Leurs structures 3D sont très variées : des protéines tout en hélices α à celles tout en brins β , en passant par le tonneau $(\beta/\alpha)_8$, les cellulases couvrent huit familles de repliements (35). En commun avec d'autres enzymes qui agissent sur des glycosides, tels que les xylanases, les amylases et les chitinases, elles possèdent un long sillon de liaison du substrat où se lient au moins deux unités de glucose de chaque côté du centre catalytique. Le sillon est bordé par des résidus interagissant *via* des liaisons hydrogène. Quelques chaînes latérales de résidus aromatiques sont souvent trouvées en tant que plateformes hydrophobes qui forment des sous-sites de liaison au sucre. Les endoglucanases ont généralement un sillon de liaison ouvert, alors que les cellobiohydrolases ont des boucles étendues qui entourent la chaîne de cellulose de sorte qu'un tunnel de liaison au substrat est formé (35).

Les cellulases figurent parmi les enzymes les plus utilisées dans l'industrie. Elles interviennent notamment dans les industries du textile et du papier pour le lissage des fibres. De plus, elles auront un rôle essentiel en tant qu'enzyme de saccharification dans les futurs procédés de bioconversion de la biomasse lignocellulosique.

Seules les glucanases de la famille 12 des glycoside hydrolases seront présentées dans ce projet, car elles présentent une forte homologie structurale avec les GH11, partageant le même repliement en « β -jelly roll » et appartenant d'ailleurs au même clan C.

1.5.1 Cellulases GH12

Cette famille comporte majoritairement des endoglucanases, mais aussi quelques xyloglucane hydrolases et des licheninases.

La première structure de GH12 résolue est celle de CelB2 issue de *Streptomyces lividans*, à la fois sous forme libre (36) et avec un ligand lié de manière covalente (37). Le repliement global de l'enzyme montre une remarquable similitude avec celui de la famille GH11. Son architecture en « β -jelly-roll » est formée de deux feuillets antiparallèles A et B, composés respectivement de 6 et 9 brins β , arrangés en « sandwich ». Ces feuillets constituent un sillon assez long, profond et

large, pouvant accommoder environ cinq sous-sites comportant de nombreux résidus aromatiques. La catalyse s'effectue par hydrolyse acide, selon le mécanisme de rétention globale d'anomérisation, par l'intermédiaire de deux résidus : Glu120 (nucléophile) et Glu203 (acide/base).

1.5.2 Comparaison cellulases GH12/xylanases GH11

La superposition et la comparaison des résidus catalytiques et de liaison confirment que les familles 11 et 12 ont évolué à partir d'un ancêtre commun (Figure 1.8). Cependant, il existe quelques différences importantes entre les xylanases GH11 et les cellulases GH12. D'abord, des distinctions sont observables au niveau du repliement général, car même si l'architecture en « β -jelly-roll » est globalement parfaitement conservée, de nombreuses xylanases ne possèdent pas le premier feuillet B1. Puis, des distinctions sont également observables au niveau de la spécificité de substrat de ces enzymes. En effet, les cellulases dégradent la cellulose, un substrat homogène et stable grâce à de nombreuses liaisons hydrogène intra- et inter-moléculaires, en plus d'être plus ordonné que le xylane, l'unique substrat des GH11, qui existe plutôt sous forme d'hétéroxylane. Une activité notable sur le xylane a aussi été observée chez des cellulases GH12, mais pas pour CelB de *S. lividans* (37).

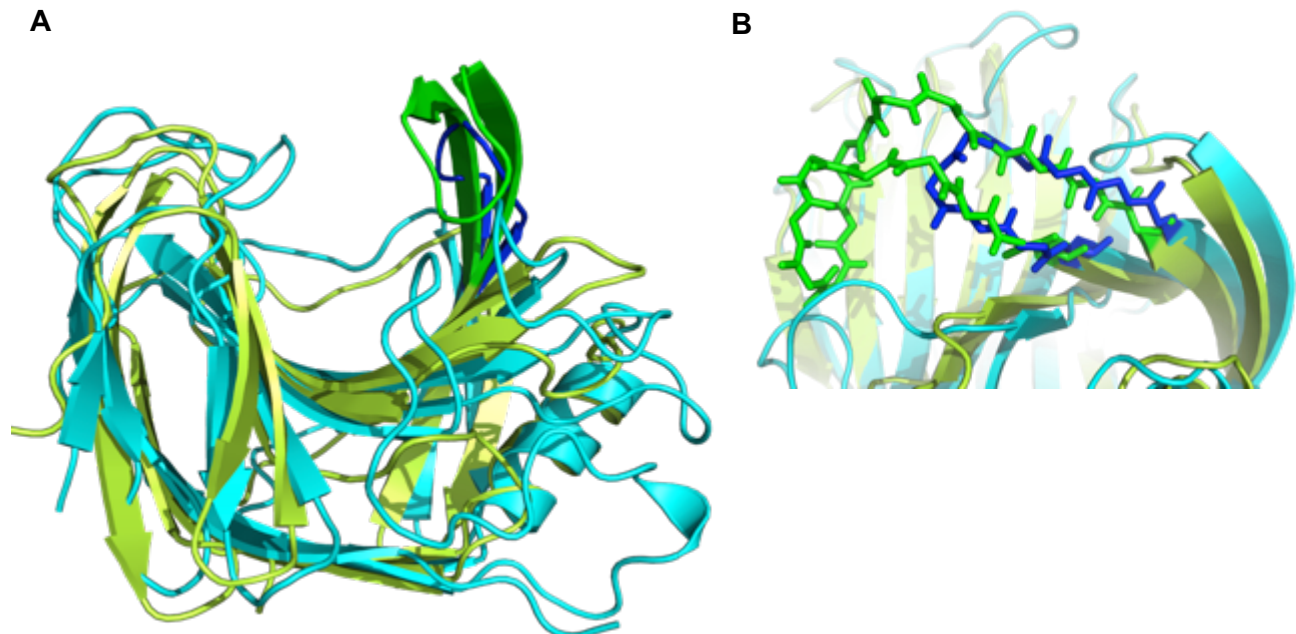


Figure 1.8 : Superposition des structures d'une xylanase GH11 (vert) et cellulase GH12 (bleu). La xylanase GH11 est la XlnB2 de *Streptomyces lividans* (Uniprot ID : P26515, RCSB PDB : 5EJ3), alors que la cellulase

GH12 est la CelB de *Streptomyces lividans* (Uniprot ID: Q54331, RCSB PDB : 1NLR). A) comparaison des repliements en « β -jelly-roll ». B) comparaison des chaînes principales des résidus du « pouce ».

Les éléments qui influencent vraisemblablement cette divergence de fonctionnalité se retrouvent d'une part au niveau des dyades catalytiques, qui sont topographiquement équivalentes, mais où les chaînes latérales des Glu acide/base n'ont pas la même orientation, et d'autre part au niveau de l'arrimage du substrat en lien avec la topographie et les propriétés électrostatiques du sillon. Les variations majeures entre ces enzymes se situent dans les régions en boucle qui flanquent le site actif, où la plupart des résidus responsables de la reconnaissance du substrat sont situés. Dans les xylanases, plusieurs des boucles semblent être réduites à une longueur minimale, alors qu'elles sont plus étendues dans les cellulases de la famille 12. Seulement une des boucles, B8-B7, est beaucoup plus longue dans la famille 11, créant ce qu'on appelle le « pouce » (17). La présence de ce « pouce » rend le sillon de liaison au substrat très profond et étroit au niveau des sous-sites (-1) et (-2). La présence à son extrémité de la triade Pro-Ser-Ile expose une surface plane hydrophobe, plus susceptible d'accueillir un substrat hydrophobe peu substitué plutôt que de la cellulose (24,37). Chez CelB, ce « pouce » est manquant et une large cavité débouche au niveau des sous-sites (-1) et (-2). De plus, le résidu aliphatique Val ou Leu conservé à la position 48, en plus de participer au maintien de l'orientation du résidu acide/base chez les Xyl-11, semble favoriser l'arrimage d'un substrat hydrophobe comme le xylose. Ce résidu empêcherait aussi la fixation de la cellulose par la mise en place de contraintes stériques en raison de sa chaîne latérale et aussi par l'absence de partenaires pour former une liaison hydrogène avec l'O₆ des substrats (37).

1.6 Relation structure, fonction et flexibilité des protéines

Historiquement, l'importance de la flexibilité des protéines dans l'événement catalytique chez les enzymes a d'abord été introduite dans la théorie de « l'ajustement induit » de Koshland (38). Récemment, le mécanisme de liaison de ligand couplé à des changements de conformation a été divisé en deux cas : « l'ajustement induit » et la « sélection conformationnelle » (39). Dans les deux cas, on suppose que toutes les conformations de la protéine pré-existent. Dans le mécanisme « d'ajustement induit », le ligand se lie à la forme libre prédominante de l'enzyme, qui va changer de conformation pour obtenir la conformation liée optimale. Dans le mécanisme de

« sélection conformationnelle », le ligand sélectionne une conformation optimale présente seulement en petites quantités.

Par ailleurs, il semble que pour certaines familles de protéines, la dynamique soit conservée parmi les homologues structuraux et fonctionnels similaires (40,41). Dans de nombreux autres cas, la dynamique des protéines semble évolutivement conservée, les changements évolutifs dans la séquence et la structure de la protéine étant souvent étroitement liés à la flexibilité (42). Toutefois, d'autres données indiquent que les protéines structurellement conservées ne partagent pas nécessairement les mêmes propriétés dynamiques, car ces dernières sont codées par la séquence primaire et non tertiaire (43).

Bien que les fluctuations atomiques sur différentes échelles de temps soient connues pour permettre un large éventail d'activités biologiques, telles que le repliement des protéines, la réorganisation structurelle, les interactions protéine-protéine et la liaison d'un ligand (44), leur rôle dans la catalyse enzymatique reste encore mal défini. Une indication suggérant fortement que la dynamique structurale joue un rôle important dans la catalyse enzymatique a été précédemment rapportée, ainsi que des mouvements fonctionnels liés au processus catalytique, comme par exemple chez la dihydrofolate réductase (45), la RNase (41) ou la triosphosphate isomérase (46).

Pour l'ingénierie de biocatalyseurs améliorés, il est nécessaire de découvrir comment les protéines exercent leurs fonctions et comment celles-ci sont déterminées par leur structure et leur flexibilité. Il est donc important d'étudier la dynamique des protéines afin d'obtenir une représentation complète de l'action de la protéine dans la liaison au substrat, la catalyse, et/ou la libération du produit.

1.7 Projet de recherche

Dans le but d'obtenir une meilleure compréhension de la relation structure-fonction-flexibilité et de sa conservation évolutive dans la grande famille des GH11, il convient de caractériser les changements conformationnels survenant lors de la catalyse chez différentes xylanases GH11 du règne bactérien, dans leur forme libre et en présence de leur substrat.

1.7.1 Problématique

La flexibilité atomique chez les enzymes se révèle d'une importance cruciale dans la conversion enzymatique de polysaccharides. Il a été démontré que les mouvements moléculaires de protéines à différentes échelles de temps peuvent être impliqués dans la régulation de l'activité enzymatique, en jouant notamment un rôle essentiel dans le contrôle de la relation entre la structure et la fonction chez plusieurs systèmes enzymatiques. Des études cristallographiques et des simulations de dynamique moléculaire sur plusieurs xylanases GH11 ont proposé l'existence d'un mouvement « ouvert-fermé », *via* une boucle près du site actif, le « pouce », qui peut jouer un rôle majeur dans la liaison au substrat, la libération du produit et/ou la catalyse. Bien que ces études fournissent des informations sur la flexibilité des protéines, elles n'offrent pas de détails sur les mouvements potentiels survenant lors de la réaction catalytique de l'enzyme, c'est à dire sur une échelle de temps de la micro à la milliseconde (μ s-ms).

Pour développer des biocatalyseurs hautement efficaces et obtenir une meilleure compréhension de l'importance et de la conservation des événements dynamiques impliqués dans la réaction catalytique des xylanases GH11, il est nécessaire de réaliser des études de dynamique de plusieurs protéines homologues à l'échelle de temps de la catalyse. La présente étude vise à mettre en place les outils moléculaires et à initier l'étude dynamique de cette famille d'enzymes, permettant à terme de comparer les similitudes et distinctions moléculaires dans la flexibilité évolutive des différents membres bactériens de la grande famille des xylanases GH11.

1.7.2 Hypothèses et Objectifs

En partant de la prémisse que la flexibilité atomique de certaines enzymes est impliquée dans la régulation de leur activité enzymatique, nous avons émis l'hypothèse que certaines GH11 présentent des changements conformationnels en présence de leur substrat et que ces mouvements jouent un rôle clé dans la réaction catalytique de l'enzyme.

Pour répondre à cette hypothèse, nous nous sommes donnés les objectifs suivants :

1. Cloner, exprimer et purifier les homologues structuraux de la famille des xylanases GH11 (XlnB2 et XlnC) et des cellulases GH12 (CelB2) de *Streptomyces lividans*.
2. Entreprendre des analyses dynamiques par résonance magnétique nucléaire (RMN) de XlnB2 de *S. lividans*.

3. Initier des analyses RMN préliminaires similaires sur les glycosidases XlnC et CelB2 nécessaires à la caractérisation comparative de leur dynamique évolutive avec XlnB2.

1.8 Méthodologie

1.8.1 Expression et purification des enzymes

Les gènes codant pour les enzymes étudiées ont été précédemment amplifiés et clonés dans des vecteurs d'expression développés dans les laboratoires de Rolf Morosoli, François Shareck et Claude Dupont (47). Les vecteurs possèdent le promoteur constitutif C109, un gène de résistance au thiostrepton (*tsrR*) et un gène de production de tyrosinase (*mel*) (Figure 1.9). Les gènes d'intérêt ont été introduits dans le plasmide pIJ702 au niveau des sites de restriction *KpnI* (qui reconnaît la séquence G_GTAC^C) et *BglII* (qui reconnaît A^GATC_T) (47).

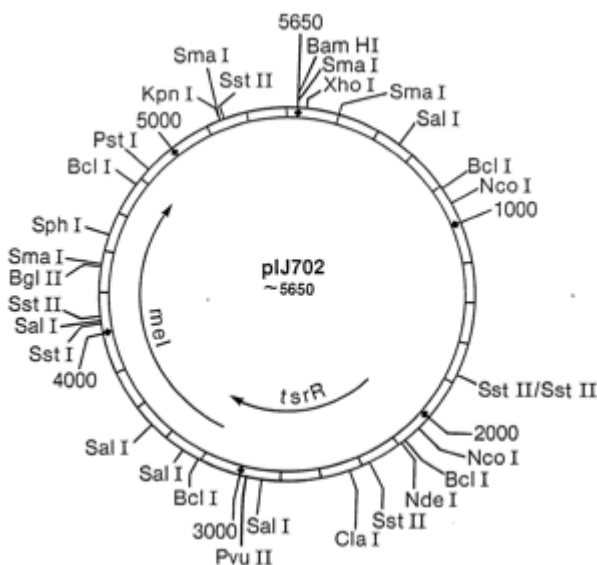


Figure 1.9 : Carte du plasmide pIJ702 avec sites de restriction. Le vecteur de clonage possède un gène de résistance au thiostrepton (*tsrR*) et un gène de production de tyrosinase (*mel+*) comme marqueurs. Issu de www.atcc.org.

Les constructions résultantes ont servi à transformer des souches *S. lividans* 10-164, un bon système d'expression au fort potentiel de sécrétion. Ces souches sont des mutants xylanase- et

cellulase-négatifs de la souche de type sauvage 1326, obtenues en interrompant le gène *msiK* requis pour la génération d'énergie des systèmes de transport de disaccharides.

L'intégration des plasmides dans les bactéries a été réalisée par incubation des protoplastes (bactéries libérées de leur paroi extracellulaire) en présence de l'ADN plasmidique recombinant et de polyéthylène glycol (PEG), qui semble faciliter l'entrée du plasmide à travers la membrane plasmique (48). La pénétration de l'ADN est stimulée par un choc thermique. La sélection des bactéries transformées, ayant donc intégré le gène de résistance au thiostrepton (*tsrR*), sera réalisée en ajoutant cet antibiotique dans le milieu de culture R5. Ce milieu solide riche est idéal pour la régénération des protoplastes. De plus, il est possible d'identifier les plasmides porteurs des gènes d'intérêt par rapport aux plasmides vides selon la couleur des colonies. En effet, en insérant les gènes entre les sites de restriction KpnI et BglII, le gène *mel* est tronqué et non fonctionnel, empêchant l'expression de la tyrosinase qui catalyse normalement la production de la mélanine, rendant les colonies blanches et non marrons. Les clones positifs ont également été confirmés par séquençage de l'ADN en utilisant les amorces 5'-CCGCTCGTCGCGTACTACATCGTCG-3' et 5'-CGACGATGTAGTACGCGACGAGCGG-3' pour XlnB2, 5'-GCCACTACCATCACC-3' et 5'-TCCGCTGACCGTGAT-3' pour XlnC et 5'-GACACCACATCTGC-3' et 5'-GGGGTCGGTGCCGCC-3' pour CelB2.

Les protéines ont été exprimées en cultivant les souches transformées à 34 °C en milieu minimum M14, supplémenté en xylose 1 % comme seule source de carbone.

Les protéines exprimées ont été purifiées par chromatographie d'échange d'ions. Les xylanases étudiées ayant un pI > 5, elles ont été purifiées par FPLC sur colonne échangeuse de cations (HiTrap CM FF, GE Healthcare) après concentration par filtration et dialyse dans un tampon citrique 20 mM à pH 4,5. La colonne est pré-équilibrée avec le même tampon et les enzymes sont éluées en utilisant un gradient de 0 à 1 M de NaCl. Dans ces conditions, les enzymes chargées positivement interagissent avec la phase stationnaire anionique de la colonne, tandis que les protéines chargées négativement ou non chargées ne sont pas retenues (Figure 1.10). Le gradient de NaCl permet d'éluier les protéines accrochées à la colonne en entrant en compétition avec les cations d'intérêt pour les sites anioniques de la phase stationnaire.

CelB2 ayant un pI de 4,2 (calculé par ExPASy ProtParam selon la séquence du domaine catalytique), cette enzyme a été purifiée par FPLC sur colonne échangeuse d'anions (HiTrap Q HP, GE Healthcare) dans un tampon Tris 20 mM à pH 7,2 avec une élution réalisée par un gradient de 0 à 1 M de NaCl. Le principe de cette chromatographie est le même que pour l'échange de cations, mais ici l'enzyme d'intérêt est chargée négativement et la phase

stationnaire positivement. Les fractions collectées contenant CelB2 ont ensuite été injectées sur une colonne de tamis moléculaire (Superdex 75 GL 10/300, GE Healthcare) et les protéines ont été éluées selon leur poids moléculaire par un tampon phosphate 100mM à pH 6.

Les fractions collectées contenant les protéines d'intérêt ont été analysées par SDS-PAGE pour en estimer la pureté. La concentration en protéine a été déterminée par spectroscopie UV en utilisant les coefficients d'extinction molaire (ϵ) prédits avec ExPASy ProtParam selon la séquence des enzymes.

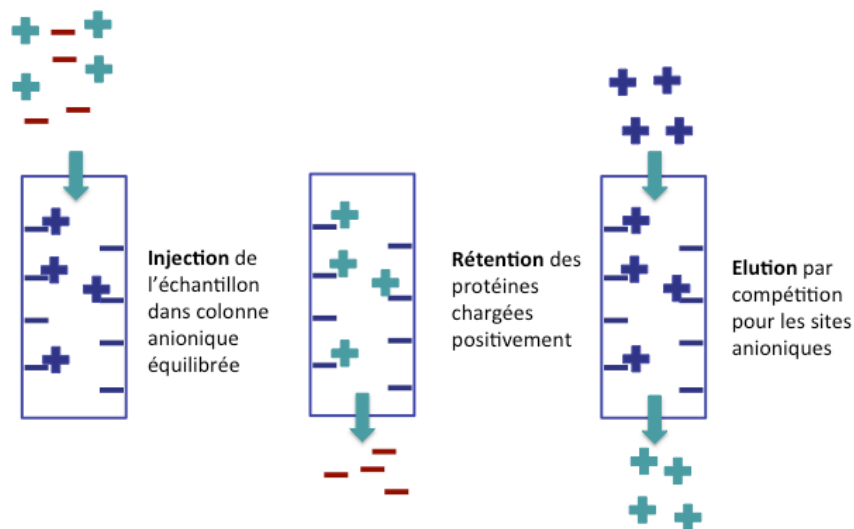


Figure 1.10 : Chromatographie échangeuse de cations : la colonne chargée négativement est équilibrée par les cations du tampon (en bleu foncé), les protéines de l'échantillon chargées positivement (en bleu clair) sont retenues par la colonne, tandis que les ions négatifs (en rouge) sont directement élués. Les flèches indiquent le sens d'écoulement.

1.8.2 Caractérisation des protéines par résonance magnétique nucléaire

La RMN désigne une propriété de certains noyaux atomiques lorsqu'ils sont placés dans un champ magnétique B_0 . Soumis à des impulsions électromagnétiques, ces noyaux peuvent absorber l'énergie du rayonnement puis la relâcher. L'énergie mise en jeu lors de ce phénomène de résonance correspond à une fréquence très précise, appelée déplacement chimique (δ), dépendant de la nature même de l'atome, du champ magnétique et de l'environnement atomique immédiat (49).

Comme la cristallographie, la RMN est une technique qui permet la détermination de la structure des protéines, mais elle a l'avantage de fournir des informations sur la dynamique interne des protéines à l'échelle atomique et sur une large gamme d'échelles de temps (49). Un autre avantage est qu'elle peut être effectuée en solution, ce qui simule l'environnement naturel des

protéines. Par conséquent, la RMN est une technique puissante pour déterminer la structure et la dynamique de petites protéines dans des conditions qui se rapprochent des conditions physiologiques. Les principales contraintes de cette technique sont la faible sensibilité (en terme de rapport signal sur bruit) et la limitation dans la taille des molécules observables avec la haute résolution spectroscopique nécessaire. En effet, même si des protéines de plusieurs centaines de kDa ont été analysées par RMN à l'aide notamment de techniques de marquage isotopique spécifique, il demeure plus simple d'étudier des protéines de moins de 25 kDa (50).

Les noyaux présentant la capacité de résonner et permettant l'étude des protéines par RMN sont principalement les ^1H , ^{15}N et ^{13}C . Seul le ^1H est présent en grande quantité dans la nature (99,98 %) dans une protéine. Les ^{15}N (0,04 %) et ^{13}C (1,07 %) doivent être ajoutés artificiellement. Ainsi, les enzymes sont enrichies en isotopes ^{15}N et/ou ^{13}C en cultivant les souches de *Streptomyces lividans* transformées dans du milieu minimum supplémenté avec respectivement du $(^{15}\text{NH}_4)_2\text{SO}_4$ comme seule source d'azote, et/ou du ^{13}C -glucose comme unique source de carbone.

Afin de pouvoir faire correspondre chaque résonance ^1H - ^{15}N retrouvée dans les spectres RMN 2D à son propre résidu dans la séquence de la protéine d'intérêt, les xylanases dont les résidus n'ont pas été attribués sont enrichies en isotopes ^{15}N et ^{13}C . Une série d'expériences 2D ^1H - ^{15}N (HSQC) et 3D ^1H - ^{13}C - ^{15}N (HNC_AC_B et $\text{C}_B\text{C}_A(\text{CO})\text{NH}$) ont été effectuées pour attribuer les pics de résonance visualisés sur les spectres à chaque acide aminé de l'enzyme. Les spectres HSQC, HNC_AC_B et $\text{C}_B\text{C}_A(\text{CO})\text{NH}$ permettent d'identifier les résidus selon le voisinage dans la séquence de la protéine des atomes d'hydrogène, azote et carbone (49).

Pour caractériser la cinétique et les constantes de liaison des ligands (K_d), des expériences de titrage RMN permettant de repérer les résidus du site actif et/ou les résidus impliqués dans la liaison du substrat au site actif ont été réalisées. Comme le déplacement chimique de chaque résidu de la protéine est dépendant de son environnement chimique, il est possible d'utiliser les spectres ^1H - ^{15}N HSQC pour déterminer avec précision le site de liaison d'un ligand sur une protéine enrichie en isotopes ^{15}N . Cette technique est appelée titrage RMN. Lorsqu'un ligand se lie à l'enzyme, il interagit avec un résidu donné et modifie l'environnement chimique de ce résidu, causant un repositionnement de son pic de résonance sur le spectre ^1H - ^{15}N HSQC. Ainsi, en ajoutant graduellement un ligand à un échantillon protéique et en enregistrant un spectre ^1H - ^{15}N HSQC à chaque ajout, il est possible de suivre la variation du déplacement chimique et de déterminer les résidus impliqués dans la liaison et/ou la discrimination du substrat. Les spectres ^1H - ^{15}N HSQC des xylanases et de la cellulase ont été acquis en utilisant respectivement comme ligand du xylobiose et du cellobiose, des produits de la réaction catalytique pouvant se lier au site

actif sans être hydrolysés. Les variations du déplacement chimique ($\Delta\delta$) ont été calculées en utilisant l'équation $\Delta\delta(\text{ppm}) = ((\Delta\delta_{\text{HN}}^2 + \Delta\delta_{\text{N}}^2/25)/2)^{1/2}$, tel que décrit dans le chapitre principal du mémoire (51). La constante de dissociation (K_d), correspondant à la concentration en ligand à 50 % de la $\Delta\delta$ maximale, a été déterminée pour chaque résidu possédant un déplacement chimique significatif ($\Delta\delta \geq 0,05$ ppm).

1.8.3 Caractérisation RMN de l'échange conformationnel à l'échelle de temps de la catalyse

Pour étudier la dynamique interne et son importance dans la catalyse enzymatique, XlnB2 enrichie en isotopes ^{15}N a été étudiée à l'aide d'expériences de dispersion de relaxation Carr-Purcell-Meiboom-Gill (CPMG). La technique de dispersion de relaxation CPMG permet d'examiner les échanges conformationnels sur une échelle de temps de la micro- à la milliseconde (μs - ms), ce qui correspond à l'échelle de temps de la liaison du ligand et de la catalyse chez la plupart des enzymes (52) (Figure 1.11). Cette échelle de temps est idéale puisqu'elle correspond ainsi à la vitesse de la constante catalytique k_{cat} calculée pour les enzymes de la famille GH11. En effet, les k_{cat} calculés de XlnB et XlnC de *S. lividans* sont de 1012 s^{-1} et 1108 s^{-1} respectivement (53). Si les mouvements des xylanases étudiées sont néanmoins plus lents ou rapides que ceux observables par CPMG, il demeure possible d'utiliser des techniques différentes pour leur caractérisation dynamique, comme notamment la méthode EXchange SpectroscopY (EXSY ou ZZ-exchange) ou R1 ρ , respectivement. De plus amples détails sur les différentes méthodologies employées pour étudier l'échelle de temps de la dynamique des protéines se retrouvent dans la revue de littérature présentée au chapitre 2, qui fut récemment publiée sous forme d'article de revue (54).

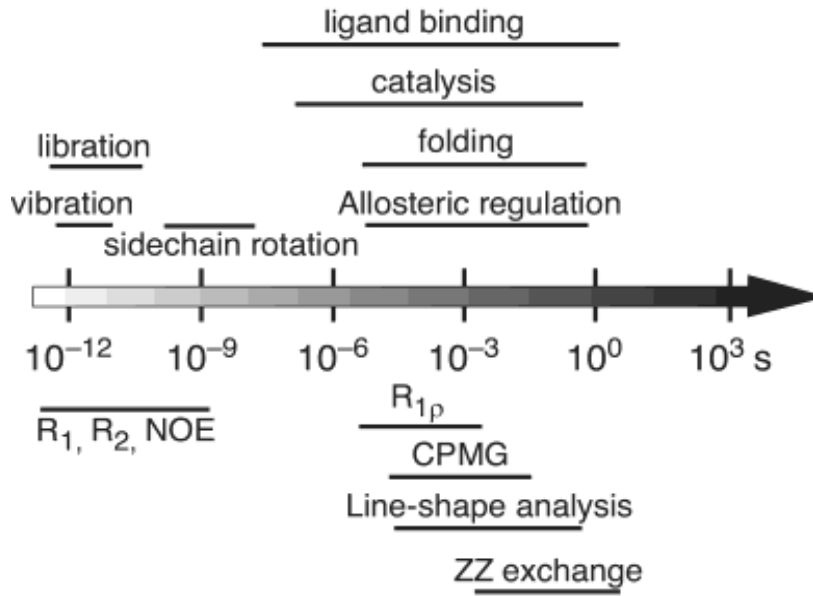


Figure 1.11 : Échelle de temps des mouvements dynamiques à l'échelle atomique chez les protéines. Les techniques RMN principalement employées pour les caractériser sont également indiquées. Figure tirée de Palmer, 2004 (55).

La technique de CPMG consiste dans un premier temps à déstabiliser les vecteurs amides (^1H - ^{15}N) d'un résidu par une série d'impulsions électromagnétiques. Le temps de retour à l'équilibre initial de la relaxation transversale (R_2) va alors dépendre du déplacement chimique moyen de chaque vecteur amide. Si le déplacement chimique d'un résidu varie en raison d'un changement de son environnement, sa résonance de relaxation ne sera pas la même qu'au repos, entraînant un changement du R_2 , ce qui signifie que le résidu subit de la dynamique à l'échelle de temps de la milliseconde. Cette série d'impulsions électromagnétiques est répétée en variant le temps de repos (τ_{cp}) entre chaque série d'impulsions. Les résultats obtenus permettent de générer des courbes de dispersion de relaxation présentant le taux de relaxation R_2 en fonction de l'inverse du τ_{cp} , discriminant ainsi les résidus dynamiques d'une protéine de ceux qui ne le sont pas à cette échelle de temps (Figure 1.12). La méthode est également quantitative et permet de calculer les populations des différents états stable (p_A) et excité (p_B), du déplacement chimique entre ces états ($\Delta\omega$), ainsi que leur vitesse d'échange (k_{ex}) (55).

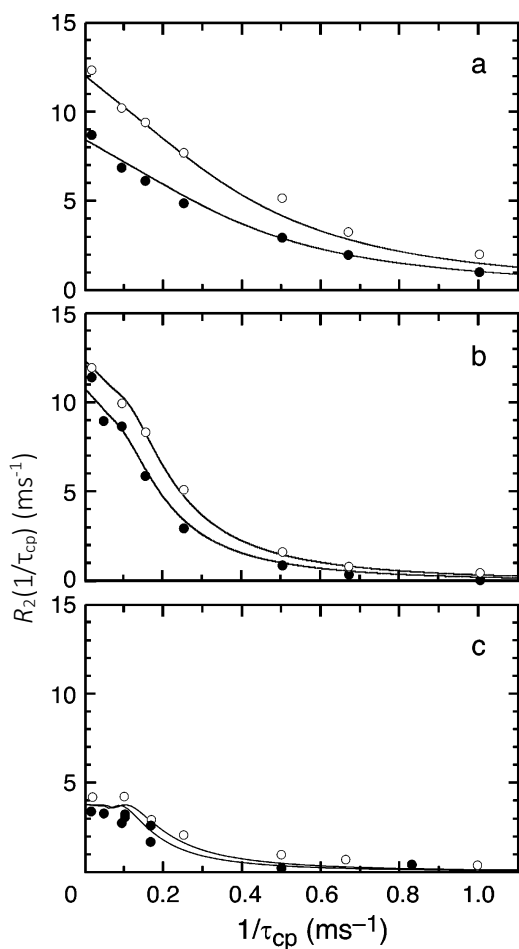


Figure 1.12 : Exemples de courbes de dispersion de relaxation ^{15}N -CPMG pour un résidu dynamique : Cys38 de l'inhibiteur de la trypsine pancréatique à (A) 313K, (B) 300K, et (C) 290K. Les valeurs de $R_2(1/\tau_{cp})$ à B_0 11,7 T (en noir) et B_0 14.1 T (en blanc) sont représentées à chaque température. L'échange est rapide à 313 K (A), intermédiaire à 300k (B) et lent à 290K (C). D'après Palmer, 2004 (55).

CHAPITRE 2 : MÉTHODOLOGIES RMN ET COMPUTATIONNELLES APPLIQUÉES À L'ÉTUDE DE LA DYNAMIQUE DES PROTÉINES

2.1 Contribution des auteurs

Les méthodologies RMN pertinentes au présent mémoire sont présentées dans ce chapitre sous la forme d'un article de revue publié dans *Archives of Biochemistry and Biophysics* en août 2017 sous la référence suivante :

Narayanan C, Bafna K, **Roux LD**, Agarwal PK & Doucet N (2017) Applications of NMR and computational methodologies to study protein dynamics. *Arch. Biochem. Biophys.*, 628, 71-80.

Chitra Narayanan, Khushboo Bafna et Pratul K. Agarwal ont écrit la section computationnelle de la revue touchant à l'étude de la dynamique des protéines par modélisation moléculaire (section 5). **Louise D. Roux a réalisé la revue de toutes les méthodologies pertinentes à l'étude des changements conformationnels des protéines par RMN (Tableau 1). À cet effet, cette revue a permis de sélectionner la technique de dispersion de relaxation CPMG comme méthode d'analyse de dynamique à privilégier pour l'étude de l'échelle de la milliseconde chez les glycoside hydrolases de *S. lividans*, ainsi que l'utilisation des techniques de relaxation R_1 , R_2 et NOE pour compléter l'analyse des mouvements moléculaires de XlnB2 à une échelle de temps plus rapide (picoseconde-nanoseconde), tel que présenté au chapitre 3 du présent mémoire (51).** Chitra Narayanan et Nicolas Doucet, **avec l'aide de Louise D. Roux**, ont écrit l'introduction, la conclusion, ainsi que les sections de la revue touchant à l'étude de la dynamique des protéines par RMN (sections 2 à 4). Nicolas Doucet a également corrigé et mis en page ce manuscrit, en plus d'en être l'auteur de correspondance.

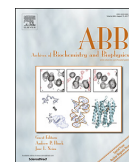
2.2 Résumé

De nombreuses observations illustrent désormais le rôle déterminant de la flexibilité des protéines à l'échelle atomique dans les événements biologiques tels que l'allostérie, la signalisation cellulaire et la catalyse enzymatique. Au fil des ans, la résonance magnétique nucléaire (RMN) a fourni des informations importantes à propos des mouvements structuraux des protéines à de multiples échelles de temps. Cet article de revue s'adressant à l'ensemble de la communauté permet d'illustrer à quel point cette technique, en plus d'être un outil indispensable pour étudier les mouvements à l'échelle atomique, continue de façonner de nombreux domaines de la science et de l'ingénierie des protéines. L'évolution constante des techniques de RMN et le développement de logiciels d'analyse de calculs complémentaires aux données expérimentales permettent désormais de fournir la direction spatiale et l'observation de représentations structurelles traditionnellement plus difficiles à obtenir en spectroscopie RMN. En plus de son rôle bien établi dans l'élucidation structurale, nous présentons des exemples récents qui illustrent la puissance combinée du marquage isotopique sélectif, des expériences de dispersion de relaxation, des analyses de déplacement chimique et des approches informatiques permettant la caractérisation de sous-états conformationnels chez des systèmes protéiques et enzymatiques.



Contents lists available at ScienceDirect

Archives of Biochemistry and Biophysics

journal homepage: www.elsevier.com/locate/yabbi

Review article

Applications of NMR and computational methodologies to study protein dynamics

Chitra Narayanan^a, Khushboo Bafna^b, Louise D. Roux^a, Pratul K. Agarwal^{c,d},
Nicolas Doucet^{a,e,f,*}^a INRS-Institut Armand-Frappier, Université du Québec, 531 Boul. des Prairies, Laval, QC H7V 1B7, Canada^b Genome Science and Technology, University of Tennessee, Knoxville, TN 37996, USA^c Department of Biochemistry, Cellular and Molecular Biology, University of Tennessee, Knoxville, TN 37996, USA^d Computational Biology Institute and Computer Science and Mathematics Division, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, TN 37830, USA^e PROTEO, The Quebec Network for Research on Protein Function, Structure, and Engineering, 1045 Avenue de la Médecine, Université Laval, Québec, QC G1V 0A6, Canada^f GRASP, The Groupe de Recherche Axé sur la Structure des Protéines, 3649 Promenade Sir William Osler, McGill University, Montréal, QC H3G 0B1, Canada

ARTICLE INFO

Article history:

Received 2 January 2017

Received in revised form

3 May 2017

Accepted 4 May 2017

Available online 5 May 2017

Keywords:

Relaxation dispersion
Quasi anharmonic analysis
Chemical shift analysis
Protein dynamics
Allostery
Conformational sub-states

ABSTRACT

Overwhelming evidence now illustrates the defining role of atomic-scale protein flexibility in biological events such as allostery, cell signaling, and enzyme catalysis. Over the years, spin relaxation nuclear magnetic resonance (NMR) has provided significant insights on the structural motions occurring on multiple time frames over the course of a protein life span. The present review article aims to illustrate to the broader community how this technique continues to shape many areas of protein science and engineering, in addition to being an indispensable tool for studying atomic-scale motions and functional characterization. Continuing developments in underlying NMR technology alongside software and hardware developments for complementary computational approaches now enable methodologies to routinely provide spatial directionality and structural representations traditionally harder to achieve solely using NMR spectroscopy. In addition to its well-established role in structural elucidation, we present recent examples that illustrate the combined power of selective isotope labeling, relaxation dispersion experiments, chemical shift analyses, and computational approaches for the characterization of conformational sub-states in proteins and enzymes.

© 2017 Elsevier Inc. All rights reserved.

Contents

1. Introduction	72
2. Selective isotope labeling: overcoming protein size limitation in relaxation	73
3. Investigating the microsecond to millisecond time frame to probe biologically relevant protein sub-states involved in allostery and enzyme function	73
4. Uncovering potential allosteric sites using the NMR chemical shift covariance (CHESCA) and projection (CHESPA) analyses	75
5. Enhancing the structural and time-evolution throughput of NMR spectroscopy using computational and theoretical approaches	76
5.1. Combining NMR observables and MD simulations to characterize protein motion	76
5.2. Conformational fluctuations and sub-states characterization using the Quasi Anharmonic Analysis (QAA)	77
6. Conclusion	79
Acknowledgements	79
References	79

* Corresponding author. INRS-Institut Armand-Frappier, Université du Québec, 531 Boul. des Prairies, Laval, QC H7V 1B7, Canada.
E-mail address: nicolas.doucet@iaf.inrs.ca (N. Doucet).

1. Introduction

Proteins are not rigid entities but flexible assemblies. Despite the broad range of growth and living conditions inside eukaryotic and prokaryotic cells, a number of studies have efficiently demonstrated that changes in molecular flexibility of proteins and enzymes on various time scales results in significant loss of biological function and/or phenotypic defects [1–4]. Each protein exists as an ensemble of conformations that undergo continuous exchange within several spatial and temporal scales, relying on these motions to achieve their biological function. To fully grasp, predict, and eventually engineer efficient protein behavior at the molecular level, it is therefore crucial to understand the effects of conformational exchange and atomic-scale protein flexibility on biological function. To this day, nuclear magnetic resonance (NMR) spectroscopy represents one of the most powerful techniques available to unravel and investigate such dynamic protein behaviors on several time frames (Table 1). Technological developments and continuous improvements in processing power have recently allowed computational approaches to significantly complement

NMR methodologies, providing spatial directionality and structural representations not easily or rapidly achievable by NMR. The present review article showcases several recent studies and methodologies characterizing protein and enzyme dynamics, primarily illustrating advances in NMR relaxation and computational approaches, coupled to other biophysical approaches. Recent examples demonstrating the combined power of selective isotope labeling, relaxation dispersion experiments, chemical shift analyses, and computational approaches are presented.

Excellent compendia have recently covered the theoretical aspects of NMR methodologies available for quantifying protein dynamics on a broad range of time frames, *i.e.* from picoseconds to hours [5–17]. To prevent undue repetition of the same theoretical concepts, the present report neither aims to offer a comprehensive overview of these NMR methodologies, nor to fully cover the wide array of protein systems exemplifying the importance of conformational flexibility in protein function. Rather, the goal is to illustrate to a general audience how the combination of newer developments with now established methods such as the NMR nuclear spin and relaxation dispersion experiments have become

Table 1
General outline of different NMR methodologies employed to quantify protein dynamics on a broad range of time frames.

Method	Description	Information provided	Time scale of protein motions investigated	Exchange rate ^a (k_{ex}) characterized	Selected references
H/D Exchange	Detection of hydrogen exchange with a fully deuterated solvent: disappearance of amide proton signals over time	Solvent accessibility, hydrogen bond strength, slow conformational exchange	Seconds- minutes		[91,92]
Real-Time NMR (RT NMR)	Direct detection of dynamic processes by quantifying the time-dependence of NMR signal intensity after a perturbation of the system	Protein folding, solvent hydrogen-exchange, slow conformational exchange	Seconds	$k_{ex} < 1 \text{ s}^{-1}$ $k_{ex} \ll \Delta\omega$	[93,94]
ZZ-exchange (Exchange Spectroscopy; EXSY)	Simultaneous determination of ¹⁵ N longitudinal relaxation and chemical exchange during a mixing delay	Slow conformational exchange such as domain movement, ligand binding and release, topological interconversion of secondary structure, <i>cis-trans</i> isomerization	Milliseconds-seconds	$k_{ex} \approx 0.2\text{--}100 \text{ s}^{-1}$ $k_{ex} \ll \Delta\omega$	[95,96]
Chemical- and Dark-state Exchange Saturation Transfer (CEST/DEST)	Saturation transfer by chemical exchange between invisible (minor, excited) state and visible (major, ground) state	Information on invisible states, populations of the states, exchange rates and slow-intermediate conformational exchange such as small domain movements	Milliseconds	$k_{ex} \approx 10\text{--}100 \text{ s}^{-1}$ $k_{ex} \leq \Delta\omega$	[6,28]
Carr-Purcell Meiboom-Gill Relaxation Dispersion (CPMG RD)	Attenuation of chemical exchange broadening by applying refocusing pulses: reduce transverse relaxation R_2 sensitive to motions	Kinetic, thermodynamic, and structural information for intermediate-fast exchange processes such as side chain reorientation, loop motions, secondary structure changes and hinged domain movements	Microseconds-milliseconds	$k_{ex} \approx 200\text{--}6000 \text{ s}^{-1}$ $k_{ex} \approx \Delta\omega$	[37,82,97]
$R_{1\rho}$ rotating frame relaxation dispersion (RF RD)	Attenuation of chemical exchange broadening by applying refocusing pulses: reduce relaxation $R_{1\rho}$ sensitive to motions	Fast-intermediate conformational exchange such as motions of loops, side chains and secondary structure elements	Microseconds-milliseconds	$k_{ex} \approx 10,000\text{--}50,000 \text{ s}^{-1}$ $k_{ex} \geq \Delta\omega$	[5]
Paramagnetic relaxation enhancement (PRE)	Increased rate of relaxation due to magnetic dipolar interaction between a nucleus and an unpaired electron of a spin label	Lowly populated states of macromolecules and their complexes, non-specific interactions, ligand binding sites, fast dynamic processes	Microseconds	$k_{ex} \approx 100,000 \text{ s}^{-1}$ $k_{ex} \gg \Delta\omega$	[15]
R_1 , R_2 and heteronuclear NOE relaxation (nuclear spin relaxation)	Analysis of longitudinal and transverse relaxation rates and the associated heteronuclear NOEs, sensitive to protein motions	Fast conformational exchange such as bond vibration and libration, side chain rotamer interconversion, random coil, loop motions, and backbone torsion angle rotation	Picoseconds - nanoseconds		[16,98]
Residual Dipolar Coupling (RDC)	Combined analysis of dipolar coupling constants measured in an array of experimental conditions	Structural states of a protein in solution, orientation of bond vectors, fast conformational exchange	Picoseconds-milliseconds		[17,99]
Chemical Shift Covariance (CHESCA) or Projection (CHESPA) Analysis	Analysis of residue networks that show correlated changes in chemical shifts due to perturbations in the system (mutations, ligand binding, etc.)	Allosteric networks, ligand binding sites			[43,44,49–51]

^a The exchange rate (k_{ex}) is proportional to the chemical shift difference (in Hertz) between two exchanging states ($\Delta\omega$). Partly adapted from Ref. [10].

indispensable tools for providing a uniquely thorough understanding of the historically elusive protein dynamics and allosteric processes occurring in many biological systems. Non-technical readers should nevertheless be informed that most recent advancements have seen rigorous targeting of distinct atomic probes in protein backbone and residue side chains (HN, H α , NH $_2$, C α , C', CH $_3$, etc.), in addition to effectively overcoming the classical size limitation of protein complexes amenable to NMR investigation [18,19]. Protein relaxation still requires dedicated access to high-field NMR spectrometers (typically between 500 MHz and 1 GHz ^1H frequencies) and for extended periods of time (hours/days). This is due to the relatively low signal sensitivity of the method, especially in light of long three-dimensional heteronuclear experiments. However, significant advances over the past 10–15 years have lifted many classical limitations of the NMR technique, effectively providing a more complete understanding of the time-dependent properties of exchange processes occurring during protein function and allosteric communication.

2. Selective isotope labeling: overcoming protein size limitation in relaxation

For a long time, protein NMR applications were limited by the size of the molecular complexes under investigation. This was primarily attributed to a number of experimental factors pertaining to physical limitations of the technique and isotope labeling methodologies [20]. One of the most important limitations in protein NMR is the fact that higher molecular weight protein monomers or multimeric complexes tumble (or rotate) more slowly in solution than their smaller counterparts. This affects relaxation parameters and results in spectral broadening (or dampening) of NMR signals. Effectively, if a high molecular weight protein is uniformly labeled with $^{13}\text{C}/^{15}\text{N}$ isotopes throughout its primary structure, the relevant spectral NMR resonances are likely to be overlapped with other resonances, and therefore masked (or completely invisible) to the experimentalist. The uniform isotope labeling methodology required to achieve NMR signal presents another limitation for NMR of large proteins. Signal resolution for any multidimensional NMR experiment on a uniformly labeled large protein or protein complex results in a significant percentage of resonance overlaps, preventing detailed atomic-scale analyses for many of the overlapping residues.

Recent years have seen significant improvements in selective isotope protein labeling methodologies, effectively eliminating some of the disadvantages of uniform labeling [20]. A convincing example of selective labeling coupled to recent protein dynamics developments was illustrated by the Kay group on high molecular weight systems, particularly the 20S core particle (CP) proteasome [21,22]. Focusing on 'metabolic tricks' limiting isotope labeling to the methyl groups of Ile, Leu, Met, and Val residues, the authors have efficiently characterized functionally relevant motions in this megadalton complex. The 20S core particle (CP) proteasome is a 670-kDa complex formed by a combination of 2 α and 2 β ring barrel structures that build a molecular machine involved in degrading cellular proteins that have become damaged over time. Due to massive size limitations, standard uniform isotope labeling prevents clear atomic-scale investigation. Using unlabeled (therefore invisible to NMR) beta protomers and highly deuterated, selectively labeled $^{13}\text{CH}_3$ -Methionine residues in the alpha protomers (5 of which are found in the N-terminal pore entrance), the authors elegantly characterized the functional gating mechanism of this macromolecular complex by combining mutagenesis with paramagnetic relaxation enhancement (PRE). The authors combined their NMR analyses with spin labels inside and outside the proteasome lumen, taking advantage

of the PRE effect to study the conformational changes in the molecular gating mechanism [21,22].

This example illustrates the power of combining selective protonation with mutagenesis, NMR relaxation, and PRE effects to investigate the dynamical behavior of high molecular weight protein complexes. Originally elusive, well-defined selective labeling protocols are now readily available to any skilled protein expression experimentalist [8,19], allowing one to significantly improve the investigation of conformational exchange for large protein complexes using NMR relaxation dispersion (see below).

3. Investigating the microsecond to millisecond time frame to probe biologically relevant protein sub-states involved in allostery and enzyme function

Arguably one of the most influential and now routinely employed methodologies of the past 15 years in solution-state NMR is the investigation of free and ligand-bound protein complexes using relaxation dispersion. Relaxation dispersion NMR allows the characterization of low populated, 'invisible' sub-states sampled by proteins in solution, either in their free states or as they perform their catalytic function in the presence of binding partners. For instance, the Carr-Purcell-Meiboom-Gill (CPMG) and $R_{1\rho}$ rotating frame experiments have been collectively employed to investigate conformational exchange rates (k_{ex}) occurring in proteins over time frames that roughly span ~ 100 to $\sim 50,000$ events per second (s^{-1}) [9,23] (Fig. 1). In addition to their relatively simple usage and implementation on multiple atomic probes (most of which are now available in standard and TROSY versions) [19,24], relaxation dispersion experiments are especially popular in light of the overlapping time scale of the characterized rates of motion with the slower microsecond-millisecond (μs - ms) time frame of relevant biological processes such as protein folding, enzyme catalysis, ligand binding, etc. Recent advances have emerged to improve the breadth of ^{15}N -labeled protein investigation by complementing the fast time scale (fast picosecond-nanosecond [ps-ns]) R_1 , R_2 , and heteronuclear NOE (hNOE) nuclear spin relaxation with CPMG and/or $R_{1\rho}$. This was done by selectively targeting ^{13}C , ^1H and/or ^2H dynamic spin probes and by improving spectral sensitivity [25–27].

More recent methodologies, such as the NMR chemical exchange saturation transfer (CEST) experiment [6,28,29], probe the same ms time frame and have provided improved sensitivity, in addition to allowing subtler conformational changes and better characterization of slower exchange processes (e.g. $k_{\text{ex}} = 50\text{--}150 \text{ s}^{-1}$). The CEST experiment can also simultaneously take advantage of the CEST-derived R_1 and R_2 data to extract S^2 order parameters reporting on the ps-ns time scale [30]. In spin-relaxation NMR, the order parameter S^2 is typically extracted from the model-free approach using experimental R_1 , R_2 , and hNOE NMR data [31–33]. This parameter is used to characterize the amplitude of the internal motions of a ^1H - ^{15}N bond vector on the ps-ns time scale. S^2 measures the degree of spatial restriction of each backbone ^1H - ^{15}N vector, describing an array of atomic-site motions ranging from fully unrestricted ($S^2 = 0$) to completely restricted ($S^2 = 1$) on this particular time frame. The CEST approach showed remarkably good agreement with standard R_1 , R_2 , and hNOE nuclear spin relaxation data, in addition to considerably improving NMR acquisition time (hours instead of days). The CPMG method can further be extended at the other end of the time spectrum, allowing the investigation of systems with exchange rates as high as 6000 s^{-1} [34]. The combined analysis of both amide ^{15}N and $^1\text{H}^{\text{N}}$ CPMG profiles and major state exchange-induced ^{15}N chemical shifts measured in $^1\text{H}^{\text{N}}$ - ^{15}N heteronuclear multiple-quantum coherence (HMQC) and heteronuclear single-quantum coherence (HSQC) data sets allows the accurate extraction of

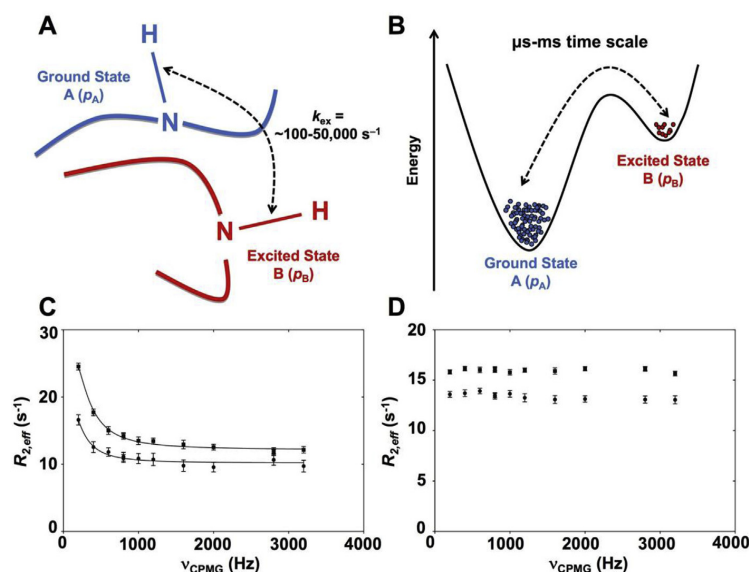


Fig. 1. Schematic depiction of exchange between two protein conformational sub-states on the μs -ms time scale. A) Schematic representation of a backbone residue HN vector experiencing conformational exchange between ground state A and excited state B on the μs -ms time scale. The popular Carr-Purcell-Meiboom-Gill (CPMG) and $R_{1\rho}$ rotating frame relaxation dispersion experiments have been collectively employed to investigate conformational exchange rates (k_{ex}) for residues experiencing conformational exchange in proteins over time frames that roughly span ~ 100 to $\sim 50,000$ events per second (s^{-1}), overlapping the time scale of relevant biological events. B) Energetic representation of the two-site exchange between ground state A (higher population, p_A), and excited state B (lower population p_B , often invisible on fast and intermediate NMR time scales) [23,53,100]. NMR relaxation dispersion experiments can provide rates of exchange (k_{ex}), populations (p_A, p_B), and chemical shifts between interconverting species ($\Delta\omega$). C) Representative experimental ^{15}N -CPMG NMR curves at 500 MHz (circles) and 800 MHz (squares) for a backbone HN vector experiencing conformational exchange on the μs -ms time scale in a protein. D) Flat experimental ^{15}N -CPMG NMR profiles at 500 MHz (circles) and 800 MHz (squares) for a backbone HN vector that does not experience conformational exchange on the μs -ms time scale. The somewhat limited atomistic details on directionality and length-scales provided by NMR for these movements can readily be complemented by MD simulations, which offer full atomistic details of the conformational populations.

exchange parameters and chemical shift difference between the interconversion states of systems exchanging with rates of $\sim 6000 \text{ s}^{-1}$.

In addition to the biological importance of the μs -ms time frame, and relative ease of technical implementation, recent examples of relaxation dispersion NMR have been employed with selective labeling and other biophysical methodologies to yield a complete portrait of protein allostery in various systems. A comprehensive enzymatic characterization also typically takes advantage of other biophysical techniques such as circular dichroism spectroscopy (CD) and/or isothermal titration calorimetry (ITC) [35,36]. A study on the allosteric mechanism of the heterodimeric 51-kDa glutamine amidotransferase imidazole glycerol phosphate synthase (IGPS) recently exemplified a subtle yet functionally relevant behavior observed through long-range allosteric motional transmission upon small-molecule effector binding to this enzyme involved in purine and histidine biosynthetic pathways in bacteria [37,38]. In this work, the authors combined ITC with $^1\text{H}, ^{13}\text{C}$ -TROSY-HMQC and methyl-TROSY multiple-quantum CPMG dispersion — all solely permitted because of selective isotope labeling of Ile, Leu, and Val methyl groups — to illustrate that the IGPS complexes with allosteric effectors distinctively respond to long-range ($>15 \text{ \AA}$) millisecond dynamics. Binding of allosteric effectors, monitored using ITC, revealed favorable enthalpic and entropic contributions to the free energy change, suggesting a shallow binding pocket for the effectors, consistent with docking and MD simulations. Interestingly, the authors demonstrate a direct linear correlation between the number of flexible residues on the μs -ms time scale and the catalytic efficiency (k_{cat}/K_m) of the glutaminase activity, further

confirming the dynamic behavior of amino acid residues that were previously shown to be critical for relaying allosteric information from independent biochemical studies [37,38].

These results highlight how protein-ligand complexes between IGPS and allosteric effectors with varying modulating efficiencies result in significantly distinct dynamic residue networks on the catalytic time frame of the enzyme. Further strengthening these observations, the authors observe that only the catalytically productive ternary complex between IGPS (the proper allosteric effector) and the PRFAR substrate yields a dynamic pathway with a shared, global exchange-rate constant (k_{ex}) for nearly all residues undergoing millisecond exchange in the enzyme. Although other IGPS allosteric activators also trigger millisecond motions, the exchange rates of individual dynamic residues are not concerted nor localized at selective residue positions essential for optimal allosteric communication. In addition to perturbing the unique profile of IGPS conformation, these suboptimal effectors also show concomitant gaps in their abilities to enhance glutamine hydrolysis [37,38].

This work further illustrates the power of combining biophysical, computational, and new NMR methodologies to improve our understanding of allostery, offering the means to observe 'invisible' conformational sub-states that would have remained experimentally elusive just a few years ago. Performing these experiments using classical non-TROSY $^1\text{H}, ^{15}\text{N}$ -HSQC titrations, coupled with standard backbone $^1\text{H}, ^{15}\text{N}$ -CPMG relaxation would have yielded overly crowded spectra without significant resolution to observe the subtle molecular and motional effects described in these studies. Also, reports have illustrated the lack of significant

structural changes caused by effector binding to IGPS (root-mean-square deviation of 0.41 Å between the apo and holo forms) [37]. As a result, previous crystal structures [39] and fluorescence quenching experiments examining solvent accessibility [40] were unable to provide definitive information on the intricate atomic-scale details underlying allosteric transmission in this enzyme system, further illustrating the power of the NMR methodology combination described here.

4. Uncovering potential allosteric sites using the NMR chemical shift covariance (CHESCA) and projection (CHESPA) analyses

Recent years have also seen the development of a variety of approaches that investigate NMR chemical shift perturbations to identify potential allosteric networks and structural dynamics in proteins [41–46]. The Melacini group introduced the chemical shift covariance analysis (CHESCA) approach [44] to identify dynamically driven intramolecular amino acid networks. CHESCA uses a combination of agglomerative clustering and singular value decomposition to identify amino acid networks that show correlated changes in chemical shifts due to perturbations arising from ligand binding or mutations.

The CHESCA approach involves the following steps: first, a matrix \mathbf{M} of combined ^1H and ^{15}N chemical shifts of a protein in the apo and multiple perturbed states is used to calculate inter-residue Pearson's correlation coefficients. The resulting symmetric matrix \mathbf{R} of pairwise correlation coefficients is analyzed using the agglomerative clustering (AC) method [47] to identify cluster(s) of coupled amino acid residues. AC assigns the first intracluster link for a residue pair with the highest correlation coefficient (r_{ij}). Subsequent intracluster links are assigned between the first pair and nearest neighbor based on the highest value of correlation coefficient to either of the residues in the original pair. These intracluster links, which can be visualized using dendrograms, identify cluster(s) of residues that show concerted response to perturbations. Clusters with more than three residues are grouped into 'networks'. At this point, however, it is unclear whether these networks play a role in allostery.

To assign a function to the identified network(s), a correlation coefficient matrix \mathbf{R}_i is computed using combined ^1H and ^{15}N chemical shifts for a subset of amino acid residues (\mathbf{M}_i) of the network(s) identified using AC, as described above. \mathbf{R}_i is analyzed again using AC to cluster functional states (such as active vs. inactive states) based on residues rather than grouping residues based on functional states. Selvaratnam et al. applied this approach to the cAMP binding domain of EPAC using chemical shifts of the apo and activator (analog) bound states. The authors identified two networks of amino acids, one separating active and inactive states of the protein, and another separating the bound and apo states. These observations indicate the role of these two networks in allostery. They further used singular value decomposition (SVD) [44,48] to validate the functional networks identified through hierarchical clustering of AC. By using combined chemical shift differences relative to an antagonist-bound form of EPAC as reference, the authors identified residue networks involved in ligand binding and allostery, respectively.

The Melacini group also introduced the chemical shift projection analysis (CHESPA) to quantify the size and direction of perturbations induced by ligand binding or point mutations (Fig. 2) [43]. In CHESPA, two residue-specific descriptors of this perturbation, i.e. the direction ($\cos\theta$) and magnitude (fractional shift X), are calculated based on the chemical shift changes observed in the perturbed system. The fractional shift X is calculated as the ratio of vector A along vector B and the magnitude of vector B:

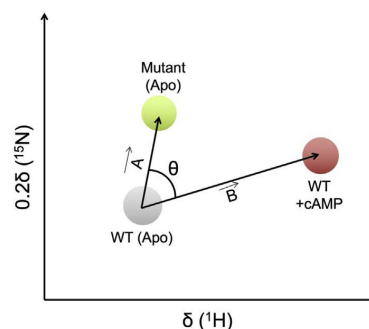


Fig. 2. Illustration of the CHESPA approach. ^1H - ^{15}N resonance assignments for apo, mutant and cAMP-bound forms of EPAC are represented as grey, green, and red circles, respectively. The compounded chemical shift upon mutation and ligand binding is calculated as the magnitude of vectors A and B, respectively (see text for details). θ represents the angle between vector A and B. Figure adapted from Ref. [43]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

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

The fractional shift X is a scalar quantity and is complemented by $\cos\theta$, which reports on the relative orientation of vectors A and B, θ being the angle between vectors A and B:

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

Of note is the fact that this approach is valid when the exchange in the free and ligand-bound conformations is fast on the NMR time scale. The presence of two or more states in the fast exchange regime could compromise the linear chemical shift pattern exploited by the CHESPA method, resulting in $\cos\theta$ values below unity. Thus, $\cos\theta$ values approaching unity suitably report on two-state equilibrium through the fractional shift X .

Much like CHESCA, the CHESPA approach was first applied on the EPAC model system by comparing the chemical shift changes arising from mutations and ligand binding to identify amino acid residues that (de)stabilize protein conformations associated with ligand binding [43]. In this study, the fractional shift X is positive if the mutation shifts the equilibrium towards the active state, and negative if the displacement is in the opposite direction. The value of X approaches 0 if variations caused by the mutation are negligible relative to those caused by ligand binding and/or if vectors A and B are orthogonal. Finally, $X \approx 1$ indicates that the magnitude of the chemical shift is identical for mutation and ligand binding events. The direction of movement is quantified using $\cos\theta$ (Fig. 2). Residues with $\cos\theta \approx 1$ show displacements in the same direction for both apo-mutant and ligand-bound states, whereas residues significantly affected by mutation display $\cos\theta < 1$. In this example, the L237W mutant of EPAC was shown to stabilize the inactive state, while cAMP binding activated wild type (WT) EPAC. The chemical shift difference between apo-WT and apo-mutant induced by amino acid substitution was calculated as the magnitude of vector A connecting apo-WT and apo-mutant peaks in the plane of ^1H and scaled ^{15}N spectra (Fig. 2). Similarly, the chemical shift difference induced upon ligand binding was computed as the magnitude of vector B connecting apo-WT and cAMP-bound WT peaks. The projection of vector A onto vector B is a measure of the shift of the conformational equilibrium between protein

conformations found in free and bound states. In a recent study, Selvaratnam et al. combined the CHESPA and CHESCA approaches to characterize the effects of mutations on the inactive vs. active conformational equilibria of apo EPAC [49]. The CHESPA analysis was used to remove the false positives in the allosteric clusters obtained through the hierarchical clustering of residue pairs in CHESCA.

The CHESPA approach can also be used to investigate the effect of long-range residue network behavior occurring upon distinct ligands binding to an enzyme, as recently illustrated by the binding of 3'-UMP and 5'-AMP to human angiogenin (Fig. 3) [50]. In this study, the authors highlighted the existence of two distinct, long-range clusters of residue networks either displaying correlated or uncorrelated chemical shift trajectories upon binding of each ligand to the purine- or pyrimidine-specific subsites of the angiogenin catalytic cleft, suggesting a structural rearrangement or allosteric response linked to ligand binding in this enzyme. Finally, using alanine to glycine substitutions, Axe et al. probed structural and dynamical changes in the α -subunit of tryptophan synthase [51]. By combining the projection (CHESPA) and covariance (CHESCA) analyses for the free and ligand-bound forms of the wild-type and mutant enzymes, they identified point mutations that stabilized three conformational states, corresponding to the free form, glyceraldehyde-3-phosphate (product) bound and active (indole-

bound) states. The authors further illustrated that the dynamical properties of the allosteric networks of amino acid residues identified from the covariance analysis differed in each of these states.

5. Enhancing the structural and time-evolution throughput of NMR spectroscopy using computational and theoretical approaches

5.1. Combining NMR observables and MD simulations to characterize protein motion

Molecular dynamics (MD) simulations provide a time-evolution 'movie' of atomic-level motions based on Newton's laws of motions. Classical molecular mechanics describe the potential energy of a molecule, such as a protein, as a function of its atomic coordinates and dictate the interactions between atoms over the course of an MD simulation [52]. The various force-constants associated with inter-atomic interactions are collectively referred to as a force-field. MD trajectories, which correspond to a collection of conformational snapshots sampled during the simulation, are analyzed using a variety of approaches to identify conformational states that are potentially important for function [53]. While conformational motions in proteins are observed over a wide range of time scales — from the fast femtosecond-picosecond (fs-ps) to the slower

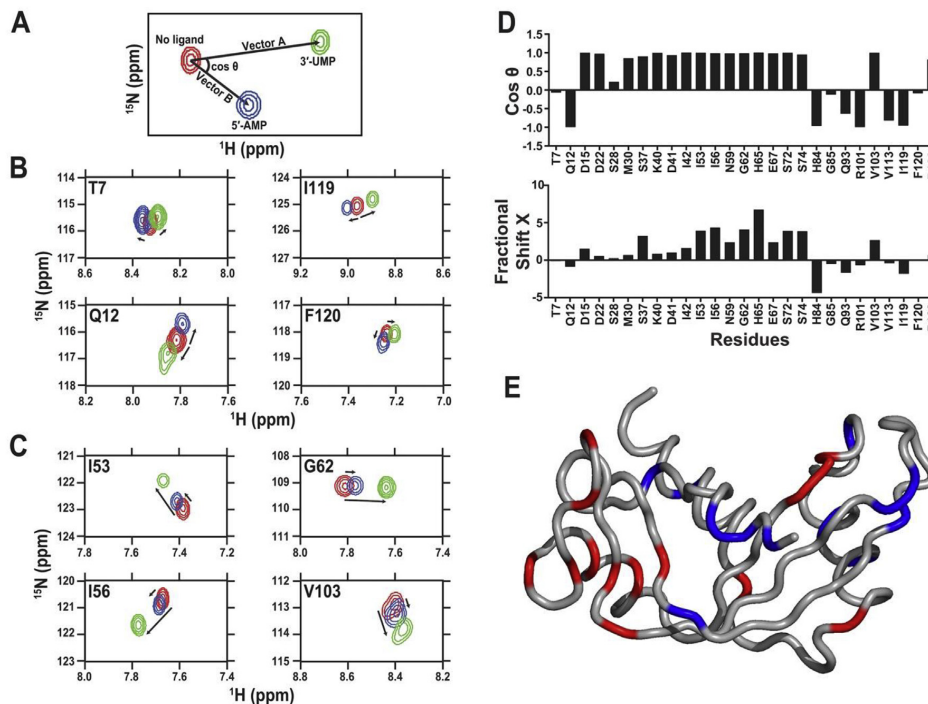


Fig. 3. Application of CHESPA to human angiogenin. Projection analysis describing independent and coordinated residue variations upon 5'-AMP and 3'-UMP binding to human angiogenin subsites. (A) Graphical representation of the CHESPA approach described in Ref. [43]. The ^1H - ^{15}N position of the peak is represented for the free (red) and bound forms (5'-AMP in blue and 3'-UMP in green) of the enzyme. Arrows indicate the movement of the ^1H - ^{15}N chemical shift (length and direction) for each peak from its origin (in red) to its saturated position (blue and green). Projection analysis of eight selected residues responding in (B) independent or (C) coordinated manner in human angiogenin. (D) Direction $\cos\theta$ and magnitude (fractional shift X) of the chemical shift perturbation of a subset of residues of angiogenin upon binding to saturation of the 5'-AMP and 3'-UMP. The $\cos\theta$ quantifies the angle between vectors A (3'-UMP) and B (5'-AMP) from its initial position in the free form of the enzyme. The fractional shift X represents the fractional composite of vectors A and B induced by the ligand-induced chemical shift change. Collectiveness was observed for residues with a $\cos\theta \sim 1$, as previously described [43]. (E) Residues showing coordinated displacements ($\cos\theta \sim 1$) are represented in red on the three-dimensional structure of human angiogenin, while residues showing uncoordinated displacements ($\cos\theta \neq 1$) are shown in blue. Reprinted with permission from Ref. [50], Copyright 2015 Wiley. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

millisecond-second (ms-s) — the range of time scales accessible by MD is currently limited due to the speed of central/graphical processing units and related computer hardware. Until recently, sampling from MD simulations was typically on the order of hundreds of nanoseconds. Enhanced sampling techniques, alongside advances in computational resources, now enable routine conformational sampling on the μ s time scale, with specialized supercomputing facilities reporting longer ms time scale simulations [54,55]. MD simulations have been used in combination with NMR to identify allosteric networks of conformational motions that play an important role in enzyme function [10,56,57].

A diverse array of other sequence and structure based computational approaches have also been used to identify and characterize protein flexibility on the fast time scale [58,59], in addition to characterizing allosteric networks that display motions on the slower μ s–ms time frame via dynamical coupling of distal residues in proteins [60–62]. For example, DynaMine is a sequence-based approach that provides a statistical and quantitative analysis of protein backbone flexibility on the fast time scale (ps–ns) for each amino acid of the protein [59]. Statistical coupling analysis is another approach that uses sequence datasets of protein families to identify co-evolving amino acid networks important for distinct biochemical functions [63,64]. Using this approach, Reynolds et al. identified a co-evolving amino acid network in the dihydrofolate reductase family undergoing millisecond conformational dynamics that was strongly correlated with enzyme catalysis [61].

Accurate prediction of NMR observables such as chemical shifts, residual dipolar couplings (RDCs), and order parameters (S^2) from protein coordinates has facilitated the characterization of protein dynamics through the comparison between NMR and molecular simulation ensembles [65]. Algorithms such as SHIFTX2, SHIFTS, CamShift, SPARTA+, and others have been used to predict chemical shifts from structural coordinates obtained from MD simulations [66–69]. Chemical shifts have also been used to compute order parameters (S^2) for MD ensemble coordinates. Comparison of the predicted S^2 with experimental observations facilitates the quantitative characterization of fast time scale dynamics of the protein backbone, providing a range of conformations associated with the observed order parameters [65,70,71]. Further, MD can complement structural observations that are missing from NMR relaxation [72,73].

In recent years, a number of studies have combined MD and NMR chemical shift analyses to probe conformational dynamics in proteins. For example, Robustelli et al. compared dynamically averaged chemical shifts predicted from MD simulation ensembles with experimental NMR shifts to characterize the effect of conformational dynamics on the observed chemical shift pattern in proteins. Using 100 ns and 1 μ s simulations of two homologs of ribonuclease H (RNase H), they showed that ensemble averaged chemical shifts calculated from MD are in better agreement with NMR observables than those predicted from static X-ray structures [42]. By probing the dynamical properties of regions displaying improved chemical shift agreement with experiment, the authors showed that these improvements result from: 1) a population weighted sampling of multiple conformational states, and 2) sampling within individual conformational basins. Further, analysis of averaged chemical shifts from MD simulations also identified erroneous conformations resulting from inadequacies of the force-field. These observations highlight the use of dynamically averaged chemical shifts to probe conformational dynamics in proteins.

NMR chemical shifts have also been used as replica-averaged structural restraints for MD simulations to characterize the dynamics of ribonuclease A (RNase A) [31–33]. Camilloni et al. showed that agreement between calculated and experimental chemical shifts was improved in the chemical shift restrained

ensembles relative to unrestrained MD ensembles [74]. Comparison of the free energy landscape of RNase A for the unrestrained and restrained MD ensembles showed that the restrained MD simulations sample the major and minor conformational states reported for RNase A, with the calculated population of the minor state consistent with previous experimental reports. These observations highlight the importance of using chemical shifts to bias the conformational landscape to reproduce dynamical properties of a protein in molecular dynamics simulations.

A number of NMR observables other than chemical shifts have also been used as restraints for MD simulations, either alone or by combining several observables [11,75]. Several studies have compared NMR residual dipolar couplings (RDCs) with predicted RDCs from MD ensembles to characterize backbone and side-chain dynamics over a range of time scales [12,13,76–79]. Markwick et al. combined plain and accelerated MD (AMD) simulations with experimental RDCs to characterize conformational dynamics of the GB3 protein over the faster ns and slower ms time scales [78]. In a subsequent study, the authors combined a novel AMD/SVD approach with extensive NMR observables to determine the time-resolved structural dynamics of ubiquitin over a broad time scale (ps–ms) [80]. SVD allowed the determination of optimal scalar J -coupling and RDC alignment tensor parameters. The population weighted average scalar couplings were shown to be in good agreement with experimental observations.

Computational MD simulations and experimentally extracted NMR relaxation parameters are now being routinely used as complementary approaches to characterize protein dynamics on multiple time scales. Among recent examples, this combined method was used to uncover amino acid residues associated with allosteric communication in Pin1 and IGPS (as reviewed above and in Ref. [14]). Cross-correlation analyses of protein motions from MD trajectories have highlighted changes in correlated motions upon binding of effector molecules to IGPS [38,81]. NMR and MD were also recently integrated to characterize the dynamical properties of the *Streptomyces lividans* xylanase B2 (XlnB2) in the free and ligand-bound states [82]. Ligand binding was shown to induce enhanced conformational dynamics of residues that interact with the ligand in the thumb loop and finger regions of the enzyme. Finally, by comparing order parameters from MD simulations with NMR observables, Fiset et al. showed that better agreement is observed with structured elements (α -helices and β -sheets) while sampling of loop regions by MD were insufficient to reproduce the experimental observations for these regions [83]. In a subsequent study, the authors characterized the dynamical properties of two β -lactamases upon substrate binding [84] (also reviewed in ref. [65]). By comparing order parameters obtained from MD and NMR for the free form, they showed excellent agreement for ordered regions while differences were observed for the less ordered loop regions. For the substrate-bound state, back-calculated order parameters from MD showed localized differences upon binding near catalytic residues.

5.2. Conformational fluctuations and sub-states characterization using the Quasi Anharmonic Analysis (QAA)

NMR techniques such as CPMG can provide the rate of exchange between distinct conformations constituting discrete conformational *sub-states*, enabled by concerted motions of residue networks that occur on the μ s–ms time scale. For instance, many residues in a loop can experience similar individual exchange rates (k_{ex}), suggesting a concerted group (or global) motion on this particular time frame. Rates of conformational exchange of surface loop regions have been shown to coincide with the catalytic turnover in several enzyme systems [85–87]. This has led to a significant interest in

linking events associated with these dynamic exchange rates with particular molecular mechanisms involved in protein function. Unfortunately, while possible, NMR is somewhat limited in its ability to provide rapid atomistic details on directionality and length-scales for these movements. Therefore, integration of computational and NMR techniques is becoming increasingly popular for the characterization of conformational sub-states and their populations. In combination with experimental NMR relaxation data, MD simulation techniques can be used to build a hierarchy of conformational states. With statistical sampling, MD can provide information on populations associated with each state and can also be extended to calculate the kinetic information about the interconversion between the states (Fig. 1B). The rates and magnitudes of interconversion are correlated with the CPMG-based relaxation parameters. Once validated, the computationally obtained conformational sub-states can provide detailed information about the orientation and other structural (and dynamical) properties of individual atoms, residues, and whole proteins.

A novel computational approach, the Quasi Anharmonic Analysis (QAA), was developed to identify sub-states in the conformational landscape of proteins [88,89]. QAA uses higher order statistics of positional variations associated with sampling of conformational states from MD simulations and provides insights into the role of conformational fluctuations on longer time scales. It permits the identification of conformational sub-states along a reaction pathway that exhibit structural and dynamical properties important for function. Although a number of alternate approaches exist, the advantage of QAA lies in its ability to correctly identify energetically homogeneous conformational sub-states that correlate well with experimental information. Using T4 lysozyme and ubiquitin as model systems, Ramanathan et al. applied QAA to identify and characterize the hierarchical organization of conformational sub-states in these proteins [88]. The authors also

characterized conformational sub-states associated with the *cis/trans* isomerization of proline peptide bonds catalyzed by the enzyme cyclophilin A (CypA) [88]. They identified a separate conformational sub-state corresponding to the transition state of the reaction pathway in CypA, where motions of clusters of residues were shown to adopt conformations that promote the transition state.

In combination with NMR relaxation dispersion experiments, QAA was recently used to identify the conformational sub-states associated with the substrate and product bound states of wild-type (WT) and a single point mutant form of RNase A [90]. MD simulation trajectories of the apo, reactant, and product-bound states were used for QAA. Analysis of the top QAA modes revealed reduced global conformational motions in the mutant, corresponding to the interconversion between the reactant to product states, relative to the WT enzyme. Further, enhanced conformational dynamics of the surface loop 4 was observed in the apo and ligand-bound states of the mutant form. These enhanced loop motions were shown to be consistent with significant changes observed in the ligand binding and catalytic properties of the enzyme. Further analysis revealed a lack of correlated motions of two surface loop regions (loops 1 and 4) in the mutant, suggesting a role of this correlated motion for optimal catalysis in the WT enzyme [90].

Further illustrating the power of this combined experimental-theoretical method, Gagné et al. used ^{15}N -CPMG, microsecond time-scale MD simulations, and QAA to investigate the conformational fluctuations associated with the binding of substrates in xylanase B2 from *Streptomyces lividans* (XlnB2) [82]. Using simulations of the apo enzyme and 6- and 9-unit xylan substrates bound to the enzyme as input for QAA, they characterized the dynamic modes and structural interactions of the enzyme that facilitate substrate binding (Fig. 4). Details of the role of structural

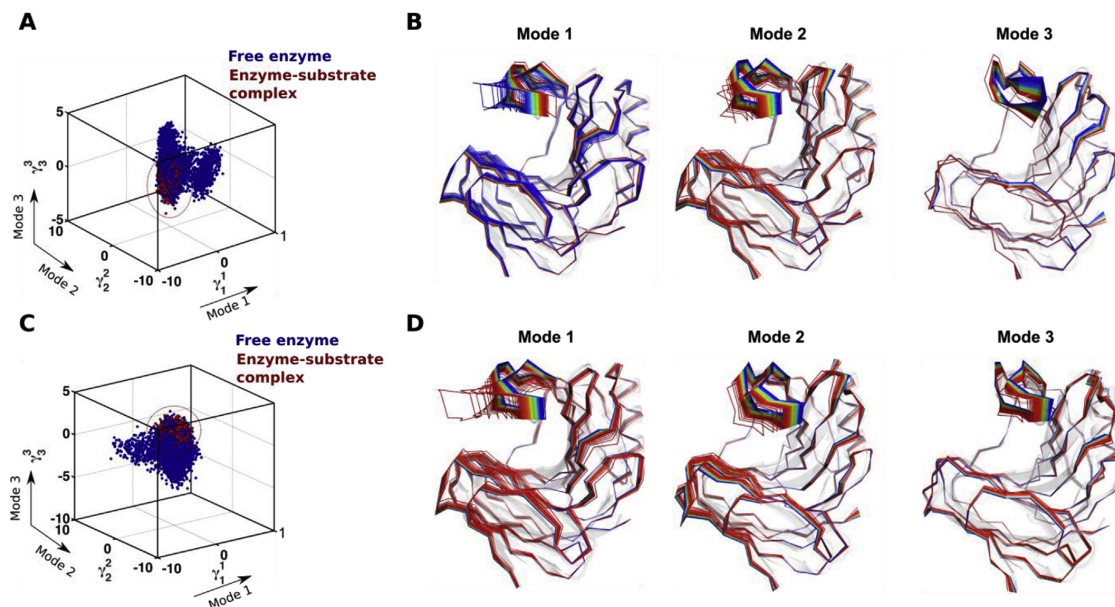


Fig. 4. Conformational sub-states of xylanase B2 (XlnB2) from *Streptomyces lividans* determined using computational simulations and QAA. Representative conformations along the top three modes for the interconversion between free and (A–B) X6-bound and (C–D) X9-bound XlnB2 binary complexes. A total of 200,000 conformational snapshots obtained from the MD simulations were used as input for QAA to identify the top QAA-independent component vectors for characterizing the primary dynamics associated with the substrate binding process in XlnB2. Reprinted with permission from Ref. [82]. Copyright 2016 American Chemical Society.

interactions of protein residues with ligand during the binding process and the range of motions (amplitude) that each residue undergoes were identified. It was discovered that the apo protein also samples the subset of conformations required for ligand binding. Quantitative characterization of these states would help in identifying the relationship between the rate of conformational exchange and the rate of ligand binding.

6. Conclusion

Technological advancements remain critical for improving NMR resolution and spectroscopic data acquisition when studying atomic-scale protein flexibility. The present review nevertheless illustrates that recent developments stem from the combination of clever isotope labeling schemes and carefully selected relaxation methods that either selectively target specific spin probes and/or improve spectral sensitivity, in addition to taking advantage of other biophysical techniques such as circular dichroism spectroscopy and/or isothermal titration calorimetry to characterize protein motions on multiple time frames. Finally, the combination of experimental NMR with computational methods aimed at predicting NMR chemical shifts, order parameters, and/or conformational sub-state populations (e.g. the quasi-anharmonic analysis, or QAA approach) now pave the way to an ever-increasing understanding of functionally relevant time-dependent motional events essential to biological function at the atomic level.

Acknowledgements

The authors thank David Bernard (INRS) for providing experimental NMR relaxation dispersion data. This work was supported in part by a multi-PI grant from the National Institute of General Medical Sciences of the National Institutes of Health under award number R01GM105978 to N.D. and P.K.A., and a Natural Sciences and Engineering Research Council of Canada Discovery Grant under award number RGPIN-2016-05557 (to N.D.). C.N. holds a Postdoctoral Fellowship from the Fondation Universitaire Armand-Frappier de l'INRS. N.D. holds a Fonds de Recherche Quebec – Sante (FRQS) Research Scholar Junior 2 Career Award.

References

- G. Bhabha, J. Lee, D.C. Ekiert, J. Gam, I.A. Wilson, H.J. Dyson, S.J. Benkovic, P.E. Wright, A dynamic knockout reveals that conformational fluctuations influence the chemical step of enzyme catalysis, *Science* 332 (2011) 234–238.
- S. Hammes-Schiffer, S.J. Benkovic, Relating protein motion to catalysis, *Annu. Rev. Biochem.* 75 (2006) 519–541.
- L.Y.P. Luk, J. Javier Ruiz-Pernía, W.M. Dawson, M. Roca, E.J. Loveridge, D.R. Glowacki, J.N. Harvey, A.J. Mulholland, I. Tuñón, V. Moliner, R.K. Allemann, Unraveling the role of protein dynamics in dihydrofolate reductase catalysis, *Proc. Natl. Acad. Sci. U. S. A.* 110 (2013) 16344–16349.
- LR. Warner, P.Z. Gatzeva-Topalova, P.A. Doerner, A. Pardi, M.C. Sousa, Flexibility in the periplasmic domain of BamA is important for function, *Structure* 25 (1) (2017) 94–106.
- A.G. Palmer, F. Massi, Characterization of the dynamics of biomacromolecules using rotating-frame spin relaxation NMR spectroscopy, *Chem. Rev.* 106 (2006) 1700–1719.
- N.J. Anthis, G.M. Clore, Visualizing transient dark states by NMR spectroscopy, *Q. Rev. Biophys.* 48 (2015) 35–116.
- M. Kovermann, P. Rogne, M. Wolf-Watz, Protein dynamics and function from solution state NMR spectroscopy, *Q. Rev. Biophys.* 49 (2016) e6.
- P. Lundström, P. Vallurupalli, D.F. Hansen, L.E. Kay, Isotope labeling methods for studies of excited protein states by relaxation dispersion NMR spectroscopy, *Nat. Protoc.* 4 (2009) 1641–1648.
- A. Sekhar, L.E. Kay, NMR paves the way for atomic level descriptions of sparsely populated, transiently formed biomolecular conformers, *Proc. Natl. Acad. Sci. U. S. A.* 110 (2013) 12867–12874.
- S. Boulton, G. Melacini, Advances in NMR methods to map allosteric sites: from models to translation, *Chem. Rev.* 116 (2016) 6267–6304.
- S. Grutsch, S. Brüschweiler, M. Tollinger, NMR methods to study dynamic allostery, *PLoS Comput. Biol.* 12 (2016) e1004620.
- L. Salmon, G. Bouvignies, P. Markwick, M. Blackledge, Nuclear magnetic resonance provides a quantitative description of protein conformational flexibility on physiologically important time scales, *Biochemistry* 50 (2011) 2735–2747.
- P. Guerry, L. Salmon, L. Mollica, J.-L. Ortega Roldan, P. Markwick, N.A.J. van Nuland, J.A. McCammon, M. Blackledge, Mapping the population of protein conformational energy sub-states from NMR dipolar couplings, *Angew. Chem. Int. Ed. Engl.* 52 (2013) 3181–3185.
- G.P. Lisi, J.P. Loria, Solution NMR spectroscopy for the study of enzyme allostery, *Chem. Rev.* 116 (2016) 6323–6369.
- G.M. Clore, J. Iwahara, Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes, *Chem. Rev.* 109 (2009) 4108–4139.
- V.A. Jarymowicz, M.J. Stone, Fast time scale dynamics of protein backbones: NMR relaxation methods, applications, and functional consequences, *Chem. Rev.* 106 (2006) 1624–1671.
- K. Chen, N. Tjandra, The use of residual dipolar coupling in studying proteins by NMR, *Top. Curr. Chem.* 326 (2012) 47–67.
- R. Sprangers, A. Velyvis, L.E. Kay, Solution NMR of supramolecular complexes: providing new insights into function, *Nat. Methods* 4 (2007) 697–703.
- Y. Xu, S. Matthews, TROSY NMR spectroscopy of large soluble proteins, *Top. Curr. Chem.* 335 (2013) 97–119.
- H. Zhang, H. van Ingen, Isotope-labeling strategies for solution NMR studies of macromolecular assemblies, *Curr. Opin. Struct. Biol.* 38 (2016) 75–82.
- L.E. Kay, New views of functionally dynamic proteins by solution NMR spectroscopy, *J. Mol. Biol.* 428 (2016) 323–331.
- R. Sprangers, L.E. Kay, Quantitative dynamics and binding studies of the 20S proteasome by NMR, *Nature* 445 (2007) 618–622.
- A.G. Palmer, Chemical exchange in biomacromolecules: past, present, and future, *J. Magn. Reson.* 241 (2014) 3–17.
- K. Pervushin, R. Riek, G. Wider, K. Wüthrich, Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution, *Proc. Natl. Acad. Sci. U. S. A.* 94 (1997) 12366–12371.
- T. Yuwen, P. Vallurupalli, L.E. Kay, Enhancing the sensitivity of CPMG relaxation dispersion to conformational exchange processes by multiple-quantum spectroscopy, *Angew. Chem. Int. Ed. Engl.* 55 (2016) 11490–11494.
- E. Rennella, A.K. Schuetz, L.E. Kay, Quantitative measurement of exchange dynamics in proteins via ¹³C relaxation dispersion of ¹³C/CHD2-labeled samples, *J. Biomol. NMR* 65 (2016) 59–64.
- R.B. Fenwick, D. Oyen, P.E. Wright, Multi-probe relaxation dispersion measurements increase sensitivity to protein dynamics, *Phys. Chem. Chem. Phys.* PCCP 18 (2016) 5789–5798.
- P. Vallurupalli, G. Bouvignies, L.E. Kay, Studying “invisible” excited protein states in slow exchange with a major state conformation, *J. Am. Chem. Soc.* 134 (2012) 8148–8161.
- G. Bouvignies, L.E. Kay, A 2D ¹³C-CEST experiment for studying slowly exchanging protein systems using methyl probes: an application to protein folding, *J. Biomol. NMR* 53 (2012) 303–310.
- Y. Gu, A.L. Hansen, Y. Peng, R. Brüschweiler, Rapid determination of fast protein dynamics from NMR chemical exchange saturation transfer data, *Angew. Chem. Int. Ed. Engl.* 55 (2016) 3117–3119.
- A.M. Mandel, M. Akke, A.G. Palmer, Backbone dynamics of Escherichia coli ribonuclease HI: correlations with structure and function in an active enzyme, *J. Mol. Biol.* 246 (1995) 144–163.
- G. Lipari, A. Szabo, Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity, *J. Am. Chem. Soc.* 104 (1982) 4546–4559.
- G. Lipari, A. Szabo, Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results, *J. Am. Chem. Soc.* 104 (1982) 4559–4570.
- P. Vallurupalli, G. Bouvignies, L.E. Kay, Increasing the exchange time-scale that can be probed by CPMG relaxation dispersion NMR, *J. Phys. Chem. B* 115 (2011) 14891–14900.
- L. Freiburger, K. Auclair, A. Mittermaier, Global ITC fitting methods in studies of protein allostery, *Methods* 76 (2015) 149–161.
- L.D. Hansen, M.K. Transtrum, C. Quinn, N. Demarse, Enzyme-catalyzed and binding reaction kinetics determined by titration calorimetry, *Biochim. Biophys. Acta* 1860 (2016) 957–966.
- G.P. Lisi, G.A. Manley, H. Hendrickson, I. Rivalta, V.S. Batista, J.P. Loria, Dissecting dynamic allosteric pathways using chemically related small-molecule activators, *Structure* 24 (2016) 1155–1166.
- I. Rivalta, G.P. Lisi, N.-S. Snoeberger, G. Manley, J.P. Loria, V.S. Batista, Allosteric communication disrupted by a small molecule binding to the imidazole glycerol phosphate synthase protein-protein interface, *Biochemistry* 55 (2016) 6484–6494.
- B.N. Chaudhuri, S.C. Lange, R.S. Myers, V.J. Davison, J.L. Smith, Toward understanding the mechanism of the complex cyclization reaction catalyzed by imidazole glycerolphosphate synthase: crystal structures of a ternary complex and the free enzyme, *Biochemistry* 42 (2003) 7003–7012.
- J.M. Lipchock, J.P. Loria, Nanometer propagation of millisecond motions in V-type allostery, *Structure* 18 (2010) 1596–1607.
- A. Cembran, J. Kim, J. Gao, G. Veglia, NMR mapping of protein conformational

- landscapes using coordinated behavior of chemical shifts upon ligand binding, *Phys. Chem. Chem. Phys.* PCCP 16 (2014) 6508–6518.
- [42] P. Robustelli, K.A. Stafford, A.G. Palmer, Interpreting protein structural dynamics from NMR chemical shifts, *J. Am. Chem. Soc.* 134 (2012) 6365–6374.
- [43] R. Selvaratnam, B. VanSchouwen, F. Fogolari, M.T. Mazhab-Jafari, R. Das, G. Melacini, The projection analysis of NMR chemical shifts reveals extended EPAC autoinhibition determinants, *Biophys. J.* 102 (2012) 630–639.
- [44] R. Selvaratnam, S. Chowdhury, B. VanSchouwen, G. Melacini, Mapping allostery through the covariance analysis of NMR chemical shifts, *Proc. Natl. Acad. Sci. U. S. A.* 108 (2011) 6133–6138.
- [45] S. Boulton, M. Akimoto, R. Selvaratnam, A. Bashiri, G. Melacini, A tool set to map allosteric networks through the NMR chemical shift covariance analysis, *Sci. Rep.* 4 (2014) 7306.
- [46] T. Konuma, Y.-H. Lee, Y. Goto, K. Sakurai, Principal component analysis of chemical shift perturbation data of a multiple-ligand-binding system for elucidation of respective binding mechanism, *Proteins* 81 (2013) 107–118.
- [47] M.B. Eisen, P.T. Spellman, P.O. Brown, D. Botstein, Cluster analysis and display of genome-wide expression patterns, *Proc. Natl. Acad. Sci. U. S. A.* 95 (1998) 14863–14868.
- [48] K. Sakurai, Y. Goto, Principal component analysis of the pH-dependent conformational transitions of bovine beta-lactoglobulin monitored by heteronuclear NMR, *Proc. Natl. Acad. Sci. U. S. A.* 104 (2007) 15346–15351.
- [49] R. Selvaratnam, M.T. Mazhab-Jafari, R. Das, G. Melacini, The auto-inhibitory role of the EPAC hinge helix as mapped by NMR, *PLoS One* 7 (2012) e48707.
- [50] D. Gagné, C. Narayanan, N. Doucet, Network of long-range concerted chemical shift displacements upon ligand binding to human angiogenin, *Protein Sci.* 24 (2015) 525–533.
- [51] J.M. Axe, D.D. Boehr, Long-range interactions in the α subunit of tryptophan synthase help to coordinate ligand binding, catalysis, and substrate channeling, *J. Mol. Biol.* 425 (2013) 1527–1545.
- [52] J.W. Ponder, D.A. Case, Force fields for protein simulations, *Adv. Protein Chem.* 66 (2003) 27–85.
- [53] P.K. Agarwal, N. Doucet, C. Chennubhotla, A. Ramanathan, C. Narayanan, Conformational sub-states and populations in enzyme catalysis, *Methods Enzymol.* 578 (2016) 273–297.
- [54] C. Narayanan, D. Bernard, N. Doucet, Role of conformational motions in enzyme function: selected methodologies and case studies, *Catalysts* 6 (2016) 81.
- [55] S. Hampton, P.K. Agarwal, S.R. Alam, P.S. Crozier, Towards microsecond biological molecular dynamics simulations on hybrid processors, in: *Proceedings of International Conference on High Performance Computing and Simulation, IEEE, 2010*, pp. 98–107.
- [56] J. Guo, H.-X. Zhou, Protein allostery and conformational dynamics, *Chem. Rev.* 116 (2016) 6503–6515.
- [57] G. Wei, W. Xi, R. Nussinov, B. Ma, Protein ensembles: how does nature harness thermodynamic fluctuations for life? The diverse functional roles of conformational ensembles in the cell, *Chem. Rev.* 116 (2016) 6516–6551.
- [58] H. Zhang, L. Kurgan, Sequence-based Gaussian network model for protein dynamics, *Bioinform. Oxf. Engl.* 30 (2014) 497–505.
- [59] E. Cilia, R. Pancsa, P. Tompa, T. Lenaerts, W.F. Vranken, From protein sequence to dynamics and disorder with DynaMine, *Nat. Commun.* 4 (2013) 2741.
- [60] V.A. Feher, J.D. Durrant, A.T. Van Wart, R.E. Amaro, Computational approaches to mapping allosteric pathways, *Curr. Opin. Struct. Biol.* 25 (2014) 98–103.
- [61] K.A. Reynolds, R.N. McLaughlin, R. Ranganathan, Hot spots for allosteric regulation on protein surfaces, *Cell* 147 (2011) 1564–1575.
- [62] A. Ramanathan, A.J. Savol, C.J. Langmead, P.K. Agarwal, C.S. Chennubhotla, Discovering conformational sub-states relevant to protein function, *PLoS One* 6 (2011) e15827.
- [63] N. Halabi, O. Rivoire, S. Leibler, R. Ranganathan, Protein sectors: evolutionary units of three-dimensional structure, *Cell* 138 (2009) 774–786.
- [64] S.W. Lockless, R. Ranganathan, Evolutionarily conserved pathways of energetic connectivity in protein families, *Science* 286 (1999) 295–299.
- [65] O. Fiset, P. Lagüe, S. Gagné, S. Morin, Synergistic applications of MD and NMR for the study of biological systems, *J. Biomed. Biotechnol.* 2012 (2012) 254208.
- [66] K.J. Kohlhoff, P. Robustelli, A. Cavalli, X. Salvatella, M. Vendruscolo, Fast and accurate predictions of protein NMR chemical shifts from interatomic distances, *J. Am. Chem. Soc.* 131 (2009) 13894–13895.
- [67] Y. Shen, A. Bax, SPARTA+: a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network, *J. Biomol. NMR* 48 (2010) 13–22.
- [68] S. Moon, D.A. Case, A new model for chemical shifts of amide hydrogens in proteins, *J. Biomol. NMR* 38 (2007) 139–150.
- [69] B. Han, Y. Liu, S.W. Ginzinger, D.S. Wishart, SHIFTX2: significantly improved protein chemical shift prediction, *J. Biomol. NMR* 50 (2011) 43–57.
- [70] Q. Liu, C. Shi, L. Yu, L. Zhang, Y. Xiong, C. Tian, General order parameter based correlation analysis of protein backbone motions between experimental NMR relaxation measurements and molecular dynamics simulations, *Biochem. Biophys. Res. Commun.* 457 (2015) 467–472.
- [71] M.V. Berjanskii, D.S. Wishart, The RCI server: rapid and accurate calculation of protein flexibility using chemical shifts, *Nucleic Acids Res.* 35 (2007) W531–W537.
- [72] N. Pastor, C. Amero, Information flow and protein dynamics: the interplay between nuclear magnetic resonance spectroscopy and molecular dynamics simulations, *Front. Plant Sci.* 6 (2015) 306.
- [73] M. Krepl, A. Cléry, M. Blatter, F.H.T. Allain, J. Sponer, Synergy between NMR measurements and MD simulations of protein/RNA complexes: application to the RRM, the most common RNA recognition motifs, *Nucleic Acids Res.* 44 (2016) 6452–6470.
- [74] C. Camilloni, A. Cavalli, M. Vendruscolo, Assessment of the use of NMR chemical shifts as replica-averaged structural restraints in molecular dynamics simulations to characterize the dynamics of proteins, *J. Phys. Chem. B* 117 (2013) 1838–1843.
- [75] S. Esteban-Martín, R. Bryn Fenwick, X. Salvatella, Synergistic use of NMR and MD simulations to study the structural heterogeneity of proteins, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 2 (2012) 466–478.
- [76] S.A. Showalter, E. Johnson, M. Rance, R. Brüschweiler, Toward quantitative interpretation of methyl side-chain dynamics from NMR by molecular dynamics simulations, *J. Am. Chem. Soc.* 129 (2007) 14146–14147.
- [77] G. Bouvignies, P.R.L. Markwick, M. Blackledge, Simultaneous definition of high resolution protein structure and backbone conformational dynamics using NMR residual dipolar couplings, *Chemphyschem* 8 (2007) 1901–1909.
- [78] P.R.L. Markwick, G. Bouvignies, M. Blackledge, Exploring multiple timescale motions in protein GB3 using accelerated molecular dynamics and NMR spectroscopy, *J. Am. Chem. Soc.* 129 (2007) 4724–4730.
- [79] C. Farès, N.-A. Lakomek, K.F.A. Walter, B.T.C. Frank, J. Meiler, S. Becker, C. Griesinger, Accessing ns-micros side chain dynamics in ubiquitin with methyl RDCs, *J. Biomol. NMR* 45 (2009) 23–44.
- [80] P.R.L. Markwick, G. Bouvignies, L. Salmon, J.A. McCammon, M. Nilges, M. Blackledge, Toward a unified representation of protein structural dynamics in solution, *J. Am. Chem. Soc.* 131 (2009) 16968–16975.
- [81] G. Manley, I. Rivalta, J.P. Loria, Solution NMR and computational methods for understanding protein allostery, *J. Phys. Chem. B* 117 (2013) 3063–3073.
- [82] D. Gagné, C. Narayanan, N. Nguyen-Thi, L.D. Roux, D.N. Bernard, J.S. Brunzelle, J.-F. Couture, P.K. Agarwal, N. Doucet, Ligand binding enhances millisecond conformational exchange in xylanase B2 from *Streptomyces lividans*, *Biochemistry* 55 (2016) 4184–4196.
- [83] O. Fiset, S. Morin, P.-Y. Savard, P. Lagüe, S.M. Gagné, TEM-1 backbone dynamics—insights from combined molecular dynamics and nuclear magnetic resonance, *Biophys. J.* 98 (2010) 637–645.
- [84] O. Fiset, S. Gagné, P. Lagüe, Molecular dynamics of class A β -lactamases—effects of substrate binding, *Biophys. J.* 103 (2012) 1790–1801.
- [85] R. Cole, J.P. Loria, Evidence for flexibility in the function of ribonuclease A, *Biochemistry* 41 (2002) 6072–6081.
- [86] K. Henzler-Wildman, D. Kern, Dynamic personalities of proteins, *Nature* 450 (2007) 964–972.
- [87] D.D. Boehr, D. McElheny, H.J. Dyson, P.E. Wright, The dynamic energy landscape of dihydrofolate reductase catalysis, *Science* 313 (2006) 1638–1642.
- [88] A. Ramanathan, P.K. Agarwal, Evolutionarily conserved linkage between enzyme fold, flexibility, and catalysis, *PLoS Biol.* 9 (2011) e1001193.
- [89] A. Ramanathan, A. Savol, V. Burger, C.S. Chennubhotla, P.K. Agarwal, Protein conformational populations and functionally relevant substates, *Acc. Chem. Res.* 47 (2014) 149–156.
- [90] D. Gagné, R.L. French, C. Narayanan, M. Simonović, P.K. Agarwal, N. Doucet, Perturbation of the conformational dynamics of an active-site loop alters enzyme activity, *Structure* 23 (2015) 2256–2266.
- [91] G.C. Li, A.K. Srivastava, J. Kim, S.S. Taylor, G. Veglia, Mapping the hydrogen bond networks in the catalytic subunit of protein kinase A using H/D fractionation factors, *Biochemistry* 54 (2015) 4042–4049.
- [92] J.J. Skinner, W.K. Lim, S. Bédard, B.E. Black, S.W. Englander, Protein dynamics viewed by hydrogen exchange, *Protein Sci.* 21 (2012) 996–1005.
- [93] A. Kumar, J. Balbach, Real-time protein NMR spectroscopy and investigation of assisted protein folding, *Biochim. Biophys. Acta* 1850 (2015) 1965–1972.
- [94] P. Rogne, T. Sparrman, I. Anugwom, J.-P. Mikkola, M. Wolf-Watz, Realtime (31)P NMR investigation on the catalytic behavior of the enzyme adenylate kinase in the matrix of a switchable ionic liquid, *ChemSusChem* 8 (2015) 3764–3768.
- [95] N.A. Farrow, O. Zhang, J.D. Forman-Kay, L.E. Kay, A heteronuclear correlation experiment for simultaneous determination of ¹⁵N longitudinal decay and chemical exchange rates of systems in slow equilibrium, *J. Biomol. NMR* 4 (1994) 727–734.
- [96] Y. Li, A.G. Palmer, TROSY-selected ZZ-exchange experiment for characterizing slow chemical exchange in large proteins, *J. Biomol. NMR* 45 (2009) 357–360.
- [97] G. Khirich, J.P. Loria, Complexity of protein energy landscapes studied by solution NMR relaxation dispersion experiments, *J. Phys. Chem. B* 119 (2015) 3743–3754.
- [98] P.J. Sapienza, A.L. Lee, Using NMR to study fast dynamics in proteins: methods and applications, *Curr. Opin. Pharmacol.* 10 (2010) 723–730.
- [99] V. Venditti, C.D. Schwieters, A. Grishaev, G.M. Clore, Dynamic equilibrium between closed and partially closed states of the bacterial Enzyme I unveiled by solution NMR and X-ray scattering, *Proc. Natl. Acad. Sci. U. S. A.* 112 (2015) 11565–11570.
- [100] J.G. Kempf, J.P. Loria, Measurement of intermediate exchange phenomena, *Methods Mol. Biol. Clifton N. J.* 278 (2004) 185–231.

CHAPITRE 3 : LA LIAISON D'UN LIGAND INDUIT L'APPARITION D'ÉCHANGE CONFORMATIONNEL CHEZ LA XYLANASE B2 DE *STREPTOMYCES LIVIDANS*

3.1 Contribution des auteurs

Le cœur des résultats de ce mémoire sont présentés dans ce chapitre sous la forme d'un article scientifique publié dans le périodique *Biochemistry* en juillet 2016 sous la référence suivante :

Gagné D, Narayanan C, Nguyen-Thi N, **Roux LD**, Bernard DN, Brunzelle JS, Couture JF, Agarwal PK & Doucet N (2016) Ligand binding enhances millisecond conformational exchange in xylanase B2 from *Streptomyces lividans*. *Biochemistry*, 55(30), 4184–4196.

Donald Gagné a réalisé l'acquisition et l'analyse des données de cristallographie de XlnB2 avec l'aide de Jean-François Couture et Joseph S. Brunzelle. Il a également réalisé l'attribution initiale des résonances de XlnB2 avec l'aide de Nhung Nguyen-Thi. Il a entrepris les analyses RMN CHESPA, CPMG et effectué le titrage RMN avec les ligands. Chitra Narayanan a réalisé les études de modélisation moléculaire de XlnB2 en présence de différents substrats avec l'aide de Pratul K. Agarwal. Nhung Nguyen-Thi a démarré le projet et a exprimé et purifié XlnB2 enrichie en ^{15}N et/ou ^{13}C . **Louise D. Roux a réalisé une partie des analyses RMN de dispersion de relaxation CPMG, en plus d'effectuer toutes les analyses de relaxation R_1 , R_2 et NOE de XlnB2 avec l'aide de David N. Bernard. Elle a également repris l'expression et la purification de l'enzyme pour refaire des échantillons nécessaires à la suite du projet et à la révision de l'article.** Nicolas Doucet, Donald Gagné, Chitra Narayanan, Jean-François Couture et Pratul K. Agarwal ont rédigé la majeure partie de l'article, avec la contribution de tous les auteurs. En plus d'être l'auteur de correspondance, Nicolas Doucet a dirigé l'ensemble du projet et a été responsable de son financement.

3.2 Résumé

Les xylanases catalysent l'hydrolyse du xylane, une source abondante de carbone et d'énergie dont les ramifications commerciales sont importantes. Malgré d'énormes efforts dédiés à l'amélioration catalytique des xylanases, leur succès reste limité en raison de notre compréhension restreinte de leurs propriétés moléculaires. Des études antérieures ont suggéré un rôle potentiel de la dynamique des résidus à l'échelle atomique dans la modulation de l'activité catalytique des xylanases GH11. Cependant, ces études dynamiques ont été effectuées à des échelles de temps plus rapides que l'événement catalytique. Ici, nous avons utilisé des expériences de titrage RMN et de dispersion de relaxation ^{15}N -CPMG en combinaison avec la cristallographie des rayons-X et des simulations de modélisation moléculaire dans le but de sonder les changements conformationnels se produisant sur une échelle de temps catalytiquement pertinente (milliseconde) chez la xylanase B2 de *Streptomyces lividans* 66 (XlnB2) et chez son mutant non fonctionnel E87A. Nos résultats montrent des propriétés dynamiques distinctes pour les états de l'enzyme apo et liée à son ligand. La forme apo de XlnB2 présente de l'échange conformationnel pour les résidus dans les doigts et les régions du sillon catalytique. En revanche, le variant E87A, qui est catalytiquement altéré, affiche une dynamique uniquement dans les doigts, ce qui démontre l'effet à longue échelle de la mutation sur la flexibilité globale de l'enzyme. La liaison au ligand induit un échange conformationnel accru pour les résidus interagissant avec le ligand, situés dans les doigts et les régions de la boucle du pouce. Ces résultats soulignent le rôle potentiel des mouvements dans les doigts et les régions de la boucle du pouce pour la reconnaissance, le positionnement, la processivité et/ou la stabilisation des ligands chez XlnB2. À notre connaissance, ce travail représente la première caractérisation expérimentale de la dynamique à l'échelle de temps de la milliseconde chez un membre de la famille des xylanases GH11. Ces résultats offrent un éclairage supplémentaire sur le rôle potentiel de l'échange conformationnel chez les enzymes de la famille GH11, fournissant des informations dynamiques essentielles à l'amélioration des enzymes par ingénierie des protéines.

Ligand Binding Enhances Millisecond Conformational Exchange in Xylanase B2 from *Streptomyces lividans*

Donald Gagné,^{†,●} Chitra Narayanan,[†] Nhung Nguyen-Thi,^{†,▽} Louise D. Roux,[†] David N. Bernard,[†] Joseph S. Brunzelle,[‡] Jean-François Couture,^{§,||,⊥} Pratul K. Agarwal,^{#,Ⓜ} and Nicolas Doucet^{*,†,||,⊥}

[†]INRS-Institut Armand-Frappier, Université du Québec, 531 Boul. des Prairies, Laval, Québec H7V 1B7, Canada

[‡]Department of Molecular Pharmacology and Biological Chemistry, Feinberg School of Medicine, Northwestern University, 320 East Superior Street, Chicago, Illinois 60611, United States

[§]Ottawa Institute of Systems Biology, Department of Biochemistry, Microbiology and Immunology, University of Ottawa, 451 Smyth Road, Ottawa, Ontario K1H 8M5, Canada

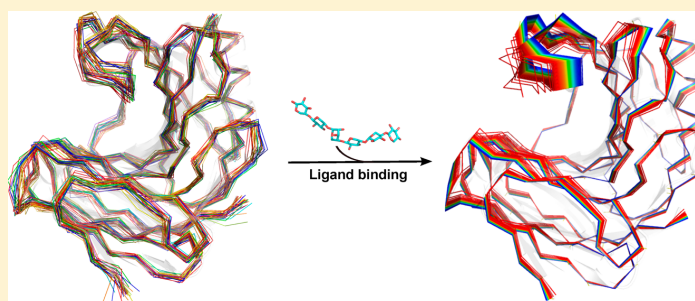
^{||}PROTEO, Québec Network for Research on Protein Function, Engineering, and Applications, Université Laval, 1045 Avenue de la Médecine, Québec, Québec G1V 0A6, Canada

[⊥]GRASP, Groupe de Recherche Axé sur la Structure des Protéines, McGill University, 3649 Promenade Sir William Osler, Montréal, Québec H3G 0B1, Canada

[#]Computational Biology Institute and Computer Science and Mathematics Division, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, Tennessee 37830, United States

[Ⓜ]Department of Biochemistry, Cellular and Molecular Biology, University of Tennessee, Knoxville, Tennessee 37996, United States

S Supporting Information



ABSTRACT: Xylanases catalyze the hydrolysis of xylan, an abundant carbon and energy source with important commercial ramifications. Despite tremendous efforts devoted to the catalytic improvement of xylanases, success remains limited because of our relatively poor understanding of their molecular properties. Previous reports suggested the potential role of atomic-scale residue dynamics in modulating the catalytic activity of GH11 xylanases; however, dynamics in these studies was probed on time scales orders of magnitude faster than the catalytic time frame. Here, we used nuclear magnetic resonance titration and relaxation dispersion experiments (¹⁵N-CPMG) in combination with X-ray crystallography and computational simulations to probe conformational motions occurring on the catalytically relevant millisecond time frame in xylanase B2 (XlnB2) and its catalytically impaired mutant E87A from *Streptomyces lividans* 66. Our results show distinct dynamical properties for the apo and ligand-bound states of the enzymes. The apo form of XlnB2 experiences conformational exchange for residues in the fingers and palm regions of the catalytic cleft, while the catalytically impaired E87A variant displays millisecond dynamics only in the fingers, demonstrating the long-range effect of the mutation on flexibility. Ligand binding induces enhanced conformational exchange of residues interacting with the ligand in the fingers and thumb loop regions, emphasizing the potential role of residue motions in the fingers and thumb loop regions for recognition, positioning, processivity, and/or stabilization of ligands in XlnB2. To the best of our knowledge, this work represents the first experimental characterization of millisecond dynamics in a GH11 xylanase family member. These results offer new insights into the potential role of conformational exchange in GH11 enzymes, providing essential dynamic information to help improve protein engineering and design applications.

Cellulosic biomass is considered an inexhaustible energy source for a number of renewable energy applications relying on enzymatic bioconversion, including fuel production,

Received: February 12, 2016

Revised: July 5, 2016

Published: July 7, 2016

food preparation, pulp and paper transformation, and medical treatment. Cellulosic polysaccharides account for more than 50% of all plant biomass, representing one of the most abundant organic carbon sources on the planet.^{1,2} Xylanases are involved in the degradation of heteroxylan (hemicellulose), one of the major components of plant biomass. On the basis of amino acid sequence similarity and hydrophobic cluster analysis, the Carbohydrate-Active Enzyme Database (CAZy, <http://cazy.org/>) classifies xylanases into six glycoside hydrolase (GH) families: GH5, GH8, GH10, GH11, GH30, and GH43. Among them, xylanases from the GH11 family have primarily been used in biobleaching applications.³ This family has been thoroughly characterized and is considered a true xylanase family because of its high substrate specificity toward xylan.^{1,4,5} As a result, several GH11 members have been structurally and functionally characterized as β -glycosidase models.^{6–15}

The typical structural feature of the GH11 family is the overall conservation of the β -jellyroll domain, which shows a closed right-handed architecture and an active site cleft covered by a “thumb loop”.¹ Like many bacteria and fungi that concomitantly secrete a number of hydrolytic enzymes to degrade plant cell walls, *Streptomyces lividans* 66 secretes two isoforms of the GH11 xylanase B (XlnB), namely, XlnB1 and XlnB2. XlnB1 is composed of two discrete structural and functional units, an N-terminal catalytic domain and a C-terminal xylan-binding domain, while XlnB2 contains only the β -jellyroll catalytic domain. Both forms have the same specific activity against xylan and retain endoacting activity profiles toward this substrate, generating large amounts of oligoxylosides with degrees of polymerization ranging from one to eight xylose units.^{16,17} Despite the high degree of structural homology within the family, GH11 xylanases display significant biophysical variability, including an optimal catalytic pH ranging from 1 to 9, and distinctive thermostability.¹ Interestingly, few structural differences account for these vast functional disparities, emphasizing the importance of understanding the molecular and mechanistic determinants responsible for the diverse properties defining optimal catalytic efficiency among GH11 xylanases.⁸

Dynamic properties of xylanases were previously observed in crystal structures and from molecular dynamic (MD) simulations performed on a number of GH11 enzymes, suggesting the existence of a thumb loop motion that may play a major role in substrate binding, product release, and/or catalysis.^{18–22} Connelly et al. showed that GH11 xylanase from *Bacillus circulans* (BCX) is rigid on the picosecond-to-nanosecond time scale in the absence and presence of ligand,²³ while Havukainen et al. illustrated that the width of the active site cavity of the *Trichoderma reesei* xylanase II (XYNII) is reduced upon binding to epoxy-alkyl- β -D-xyloside analogues.²⁴ Using novel *in silico* modeling approaches combined with *in vitro* studies, Paës et al. showed the important role of thumb loop residues in product release and (de)glycosylation.²⁵ While these studies provided important information about protein flexibility, they nevertheless report on motions occurring on time scales much faster than that of the catalytic reaction. Indeed, crystal structure investigation provides access to limited snapshots of possible xylanase conformations, while classical MD simulations explore motions on a much faster time scale (nanoseconds), which is far from the millisecond time frame of catalytic turnover in these enzymes.

In the study presented here, we characterized the conformational dynamics of XlnB2 and a catalytically impaired variant (XlnB2-E87A) from *S. lividans* in the apo and ligand-bound states over a broad range of time scales using ¹⁵N-CPMG NMR relaxation dispersion measurements and molecular dynamics (MD) simulations to probe conformational fluctuations on the catalytic time frame. Mutation of the general acid/base residue E87 (which performs hydrolysis by adding or removing a proton) to an alanine blocks the first step of the hydrolytic reaction,²⁶ resulting in a catalytically impaired enzyme variant. This mutation prevents distortion of the substrate conformation in the active site cleft,²⁶ providing information about the structural complex between xylanase and longer xyloside ligands. The XlnB2-E87A variant was, therefore, also used to investigate the conformational exchange experienced by a Michaelis–Menten complex between the enzyme and a typical oligomeric substrate. Our results show that XlnB2 and its catalytically impaired E87A variant experience distinct conformational exchange behavior in the apo and ligand-bound states. Enhanced conformational dynamics in the thumb loop region was observed upon binding of the xylobiose (X2) product and xylopentaose (X5) substrate. Analysis of the conformational substates of the top dynamic modes associated with the interconversion between the free and substrate-bound states further revealed a “twist type” motion involving the thumb loop and fingers of XlnB2. These observations provide important insights into the potential role of functionally relevant millisecond time-scale dynamics in XlnB2.

■ EXPERIMENTAL PROCEDURES

Cloning, Mutagenesis, and Protein Purification. The gene encoding XlnB2 was cloned as described previously.²⁷ The XlnB2-E87A variant was generated by site-directed mutagenesis using standard procedures. The sequence encoding the XlnB2-E87A variant (without the signal sequence) was amplified from vector pLAF935, which contained the *xlnB2* gene as a template, using the following DNA primers: 5′-CCGGTACCGACACG-GTCGTCACGA-3′ (forward primer, with the *KpnI* site underlined) and 5′-CCGAGCTCTCACCCGCCGACGTTG-ATGCT-3′ (reverse primer, with the *SacI* site underlined). The point mutation E87A was introduced by overlap extension polymerase chain reaction (PCR),²⁸ using the following DNA primers: 5′-CCGCTCGTCGCGTACTACATCGTCG-3′ (forward) and 5′-CGACGATGTAGTACGGACGACGCGG-3′ (reverse). The 593 bp *KpnI*–*SacI*-digested PCR product containing the E87A mutation was cloned into the same restriction sites in plasmid pC109 (NEB, Ipswich, MA), which carries a thiostrepton resistance marker. The resulting construct was subsequently transformed into *S. lividans* 66 10-164 (American Type Culture Collection, Manassas, VA), a xylanase- and cellulose-negative protoplast, and plated on R5 medium without an antibiotic for 16 h at 34 °C. The agar plates were then flooded with a thiostrepton solution and further incubated for 3–4 days at 34 °C. The resulting plasmid was isolated, and positive clones were confirmed by DNA sequencing.

XlnB2 and XlnB2-E87A were enriched by growing the transformed *S. lividans* strains in M14 medium supplemented with (¹⁵NH₄)₂SO₄ (Sigma-Aldrich, St. Louis, MO) and 1% (w/v) xylose as the sole nitrogen and carbon sources, respectively. For multidimensional NMR experiments, xylose was replaced with [¹³C]glucose (w/v) (Sigma-Aldrich). It is known that glucose represses the secretion of extracellular enzymes in *S. lividans*.¹⁶ However, expression yields of the targeted enzymes

with [^{13}C]glucose were shown to be sufficient for NMR investigation. After incubation for 72 h, proteins were obtained with final yields of 63 and 36 mg/L for ^{13}C - and ^{15}N -labeled XlnB2 and XlnB2-E87A, respectively. The proteins were concentrated by ultrafiltration and dialyzed against 20 mM citric buffer (pH 4.5), prior to being loaded on a CM FF column (GE Healthcare, Baie d'Urfe, QC) pre-equilibrated with 20 mM citrate buffer (pH 4.5). The enzymes were eluted with a linear gradient of 20 mM citrate buffer (pH 4.5) and 1 M NaCl. Fractions containing XlnB2 or XlnB2-E87A were collected, dialyzed in water, and lyophilized. Mass spectrometry was also employed to confirm sequence identity and integrity. The protein concentration was determined by UV-vis spectroscopy (between 250 and 350 nm) using a predicted extinction coefficient value (ϵ_{280}) of $63830 \text{ M}^{-1} \text{ cm}^{-1}$ (ExpASY ProtParam). Protein purity was assessed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis using ImageJ (National Institute of Health, Bethesda, MD).

Crystallization and Structure Determination of XlnB2.

Crystals of XlnB2 were obtained by vapor diffusion by mixing equal volumes of the purified XlnB2 and a mother liquor composed of 0.2 M zinc acetate dehydrate, 0.1 M sodium acetate (pH 4.5), and 10% (w/v) PEG 3000 at room temperature. Rod-shaped crystals were harvested and soaked in the mother liquor supplemented with 20% glycerol. A single diffraction data set was collected by the Life-Science-Collaborative Access Team at Argonne National Laboratory (Lemont, IL). The structure was determined by molecular replacement using Xyl1 from *Streptomyces* sp. S38 as a search model [Protein Data Bank (PDB) entry 1HIX].²⁹ The structure of XlnB2 was completed using iterative rounds of model building and refinement using Coot³⁰ and Phaser,³¹ respectively. The quality of the structure was assessed using MolProbity.³²

Resonance Assignments. NMR samples were prepared by solubilizing lyophilized, uniformly ^{15}N -labeled or ^{15}N - and ^{13}C -labeled proteins in 20 mM MES (pH 6.5), 20 mM NaCl, 0.01% NaN_3 , and 10% $^2\text{H}_2\text{O}$. All NMR experiments were recorded at 298 K on samples containing 0.6–0.8 mM protein in 5 mm Shigemi NMR tubes. NMR experiments were performed on 600 and 800 MHz Inova NMR spectrometers (Agilent, Santa Clara, CA) equipped with triple-resonance cold probes and pulsed-field gradients. For backbone assignments of XlnB2, a series of standard Agilent BioPack ^1H , ^{15}N two-dimensional (HSQC) and ^1H , ^{13}C , ^{15}N three-dimensional experiments were performed.³³ All spectra were processed with NMRPipe³⁴ and analyzed using NMRViewJ version 8 and Sparky, in addition to a series of in-house scripts. The backbone assignments of XlnB2-E87A were achieved by overlaying the ^1H - ^{15}N HSQC spectrum with the corresponding assigned spectrum of wild type (WT) XlnB2. Peaks showing chemical shift variations were confirmed using three-dimensional (3D) assignment experiments.

NMR Titration Experiments. All NMR titration experiments were performed in 20 mM MES (pH 6.5), 20 mM NaCl, and 0.01% NaN_3 . Concentrations of 657 μM (661 μM) of ^{15}N -labeled XlnB2 (XlnB2-E87A) were used for the titration of unlabeled xylobiose (xylopentaose). ^1H - ^{15}N HSQC experiments were performed at 800 MHz with 256 and 8192 points in the t_1 and t_2 dimensions, respectively. ^1H - ^{15}N HSQC spectra of XlnB2 were recorded by gradually increasing the ligand:enzyme molar ratios to 0.3, 1.2, 3, 6, 12, 20, 30, 40, 60, 80, and 100. Similarly, ^1H - ^{15}N HSQC spectra of XlnB2-E87A

were recorded with increasing ligand:enzyme molar ratios of 0.3, 1.2, 3, 6, 12, 20, 30, 60, 80, 90, and 100. Chemical shift variations ($\Delta\delta$) were calculated using the equation described by Grzesiek et al.³⁵ $\Delta\delta$ (ppm) = $[(\Delta\delta_{\text{HN}}^2 + \Delta\delta_{\text{N}}^2/25)/2]^{1/2}$. The dissociation constant (K_d) was determined for each residue showing a $\Delta\delta$ of ≥ 0.05 ppm and averaged, as previously described.³⁶ Residues Trp20, Thr120, Glu128, Thr130, and Tyr171 were used for determining the K_d , which corresponds to the intercept on the x -axis (ligand concentration) at 50% the total variation of $\Delta\delta$.

^{15}N Spin Relaxation. ^{15}N T_1 , T_2 , and $^{15}\text{N}\{^1\text{H}\}$ steady-state NOE spectra for free WT XlnB2 were acquired with water-selective pulse sequences described by Farrow et al.³⁷ ^1H - ^{15}N HSQC spectra recorded to measure the T_1 and T_2 series were collected with 216 complex points in the ^{15}N dimension and 1666 points in the proton dimension and zero-filled to 6664. Sixteen transients were recorded for each measurement. Relaxation delays of 0.01, 0.08, 0.16, 0.32, 0.64, 0.96, 1.28, and 1.92 s and 0.01, 0.03, 0.05, 0.07, 0.09, 0.11, and 0.15 s were used to measure T_1 and T_2 , respectively. A $^{15}\text{N}\{^1\text{H}\}$ NOE spectrum with and without a 5.0 s proton presaturation period following a 5.0 s relaxation delay was used compared to a single 10 s relaxation delay in the reference spectrum. NOE spectra were recorded with 192 complex points in the ^{15}N dimension but were otherwise identical to the T_1 and T_2 spectra. The sample was dissolved in potassium phosphate buffer (pH 6.5), and all spectra were recorded at 800 MHz and 25 $^\circ\text{C}$.

NMR Relaxation Dispersion Experiments (millisecond motions). Backbone amide ^{15}N -CPMG relaxation dispersion experiments were performed at 25 $^\circ\text{C}$ for XlnB2 (657 μM) and XlnB2-E87A (661 μM) in the apo and saturated ligand-bound forms (100 molar equivalent ligand concentration) with B_0 fields of 14.1 T (600 MHz) and 18.77 T (800 MHz) using published sequences.³⁸ The constant time delay was set to 20 ms, and interleaved two-dimensional spectra were recorded 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. Global residue fits and model analyses were performed by fitting 600 and 800 MHz CPMG dispersion data to the single-quantum CPMG equation^{33,39,40} using GraphPad Prism 5. Global residue fits were performed using the following residues from the fingers region: Gln8, Phe19, Trp20, Val48, Gly71, Asn72, Thr176, and Gly178 (XlnB2) and Gly71, Asn72, Thr176, Gly178, and Tyr179 (XlnB2-E87A). Residue selection for global k_{ex} calculation was based on (1) the quality of dual-field individual fits (600 and 800 MHz) for residues showing similar k_{ex} values (within error) and (2) similar residue position correspondence between the two proteins (XlnB2 and XlnB2-E87A).

Chemical Shift Projection Analysis (CHESPA). A CHESPA was used to characterize the effect of binding of two ligands to XlnB2-E87A using the protocol described by Axe et al.⁴¹ For each selected residue, the fractional shift (X) = $\frac{\bar{A}\cdot\bar{B}}{|\bar{B}|^2}$ and $\cos(\theta) = \frac{\bar{A}\cdot\bar{B}}{|\bar{A}|\cdot|\bar{B}|}$ were calculated as described by Selvaratnam et al.⁴² using the ^1H and ^{15}N coordinates of the peak in its initial state (apoenzyme) and at saturation upon binding two ligands, X2 and X5.

Computational Modeling of the Apo and Substrate-Bound States of XlnB2. Molecular dynamics (MD) simulations were performed to model XlnB2 in the free and substrate-bound (binary complex) forms, both with explicit

water solvent. Xylopentaose (X5), xylohexaose (X6), and xyloheptaose (X7) substrates were used for the ligand-bound simulations. Model preparation and simulations were performed using the AMBER version 14⁴³ suite of programs for biomolecular simulations. AMBER's 99SB_ildn force field was used for all simulations. MD simulations were performed using NVIDIA graphical processing units (GPUs) and AMBER's pmemd.cuda simulation engine (with the default SPFP precision).

For the free form, the coordinates of the XlnB2 enzyme were taken from the crystal structure determined in this study (PDB entry 5EJ3). The system was neutralized by addition of counterions (two Cl⁻ ions), and the resulting system was then immersed in a rectangular box of SPC/E water, with a 10 Å minimum distance between the protein and the edge of the periodic box. A similar protocol was used to prepare the system for the binary complex with a model substrate. The crystal structure of substrate-bound xylanase II from *T. reesei* (PDB entry 4HK8) was used to prepare the substrate-bound conformations of XlnB2. The longer chain substrate (X9) was created by extending the coordinates of the X6 substrate through duplication of three xylose units (+3, +2, +1) and translating them to attach at the end of the chain (-3 xylose unit) using PyMOL visualization software. Analysis of MD trajectories showed that X6 and X9 were most stable in the binding pocket during the long MD runs and were therefore selected for further structural and dynamic characterization.

The prepared systems were equilibrated using a protocol described previously.⁴⁴ The equilibrated systems were then used to run 1 μs of production MD under constant energy conditions (NVE ensemble). Coordinates stored every 10 ps for a total of 100000 conformational snapshots were used for analysis, collected for each of the two systems (free form and binary complex).

Quasi-Anharmonic Analysis (QAA). The 200000 conformational snapshots obtained from the (apo and substrate-bound complex) MD simulations were used for QAA. The top QAA-independent component vectors were analyzed in detail to characterize the primary dynamics associated with the substrate binding process in XlnB2. Note that QAA with 100000 (and more) conformational snapshots was used to provide reproducible and accurate conformational fluctuations, and only the protein Cα atom positions were used for the analysis. To obtain an accurate description of the conformational dynamics associated with the substrate binding process, good sampling is required. We therefore used the X6 and X9 substrates with six and nine xylose units, respectively, for the simulations and subsequent characterization of conformational dynamics in XlnB2.

Enzyme–Substrate Interactions. The energy for the enzyme–substrate interactions was calculated as a sum of electrostatic and van der Waals energies between atom pairs. We developed this protocol previously to investigate other enzyme systems.⁴⁵

$$E_{\text{enz-sub}} = \sum (E_{\text{el}} + E_{\text{vdw}})$$

where E_{el} is the electrostatic contribution, E_{vdw} is the van der Waals term, and the summation runs over all atom pairs for the enzyme–substrate complex. The E_{el} and E_{vdw} terms were computed as follows

$$E_{\text{el}} = \frac{q_i q_j}{\epsilon(r) r_{ij}} \text{ and } E_{\text{vdw}} = \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6}$$

where q_i values are partial charges and A_{ij} and B_{ij} are Lennard-Jones parameters. These parameters were obtained from the AMBER force field. A distance-dependent dielectric function was used:

$$\epsilon(r_{ij}) = A + \frac{B}{1 + k \exp(-\lambda B r_{ij})}$$

where $B = \epsilon_{\infty} - A$, $\epsilon_{\infty} = 78.4$ for water, $A = -8.5525$, $\lambda = 0.003627$, and $k = 7.7839$.

All xylanase and substrate atom pairs were included in the calculations, and resulting interaction energies were summed up per residue pair. The energies were calculated for 1000 snapshots, every 1 ns, sampled during the full 1 μs simulation and were averaged over these 1000 snapshots.

RESULTS

Crystal Structure of XlnB2. To generate a working model for XlnB2, we determined the crystal structure of WT XlnB2 (PDB entry 5EJ3) by molecular replacement using Xyl1 as a search model (Table 1). Two molecules were placed in the asymmetric unit and refined to 1.31 Å. The final model was composed of two chains, each comprising residues 1–191 of XlnB2. The structure folds as a β-jellyroll domain displaying two β-sheets of five and nine β-strands (termed sheets A and B, respectively). The single α-helix is positioned under β-sheet B. Comparison of the XlnB2 crystal structure with two canonical

Table 1. Data Collection and Refinement Statistics for the Crystal Structure of WT XlnB2 from *S. lividans* 66^a

Data Collection ^b	
space group	P2 ₁
cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	41.0, 66.9, 65.4
β (deg)	94.1
resolution (Å)	46.7–1.3 (1.44–1.3)
<i>R</i> _{sym}	6.7 (55.2)
<i>I</i> / <i>σI</i>	12.2 (2.6)
completeness (%)	97.5 (96.4)
redundancy	4.2 (4.2)
Refinement	
resolution (Å)	46.7–1.3
no. of reflections	81872
<i>R</i> _{work} / <i>R</i> _{free}	12.9–16.6
no. of atoms	
XlnB2	3055
water	410
<i>B</i> factor (Å ²)	
XlnB2	15.5
water	31.8
rmsd	
bond lengths (Å)	0.014
bond angles (deg)	1.42
Molprobrity score	0.98
Ramachandran plot (%)	
outliers	0
favored	98

^aThe data set has been collected on a single crystal. ^bData for the highest-resolution shell are given in parentheses.

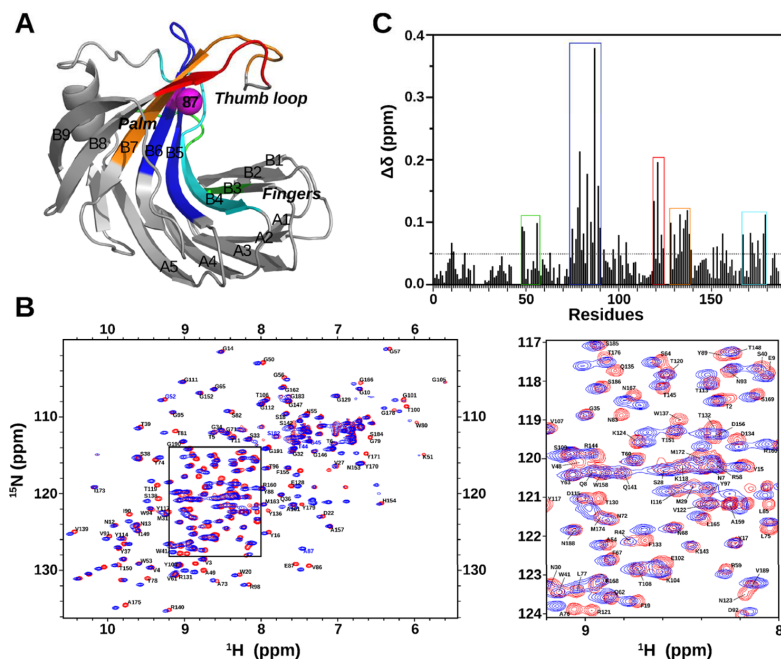


Figure 1. Chemical shift variations experienced by XlnB2 upon introduction of the E87A mutation. (A) Cartoon representation of the crystal structure of XlnB2 (PDB entry 5EJ3). The E87A point mutation is displayed as a magenta sphere. Residue clusters presenting the most significant chemical shift changes are color-coded on the structure of XlnB2. (B) ^1H - ^{15}N HSQC overlay of WT XlnB2 (red) and XlnB2-E87A (blue). A close-up view of the central spectral region is shown on the right. Amino acids are identified by the single-letter code followed by sequential residue numbers. (C) Bar graph representing chemical shift variations ($\Delta\delta$) between XlnB2 and its E87A variant. The dotted line indicates the altered threshold of chemical shift variations considered significant ($\Delta\delta > 0.05$ ppm). $\Delta\delta$ values were calculated using the equation $\Delta\delta$ (ppm) = $[(\Delta\delta_{\text{HN}}^2 + \Delta\delta_{\text{N}}^2/25)/2]^{1/2}$ as described by Gresiek et al.³⁵ (see Experimental Procedures).

GHI1 xylanases from *T. reesei*⁸ and *B. circulans*⁴⁶ shows a high degree of structural homology, with root-mean-square deviation (rmsd) values of 0.8 and 0.4 Å for all atoms, respectively. These results demonstrate that XlnB2 shares the common structural properties of other xylanases, such as a closed right-handed architecture and a substrate-binding site covered by a β -hairpin (termed the thumb loop) formed by β -strands B8 and B7 (Figure 1A). A set of six β -strands, also termed fingers, form the other side of the substrate-binding cleft. The cord loop, which joins β -strands B6 and B9, is 12 residues long in XlnB2.

NMR Assignments and Spectral Comparison of XlnB2 and the E87A Variant. The ^1H - ^{15}N HSQC (heteronuclear single quantum coherence) spectra of XlnB2 and XlnB2-E87A show well-dispersed resonances, indicating well-folded proteins (Figure 1B). We used a combination of two-dimensional (^1H - ^{15}N HSQC) and three-dimensional [3D HNCOC, HNCACB, CBCA(CO)NH, and HN(CA)CO] NMR experiments to sequentially assign the ^1H , ^{13}C , and ^{15}N backbone resonances of XlnB2 and XlnB2-E87A. Peak assignments were confirmed with a ^{15}N -NOESY experiment. A total of 172 (171) backbone ^1H - ^{15}N peaks were assigned of the 186 peaks expected for XlnB2 (XlnB2-E87A), excluding the five prolines and the N-terminal residue, corresponding to 92.5% (91.9%) of all non-proline residues. Resonance assignments were obtained for 84.4% (84.4%) of non-proline C α atoms and 69.9% (69.9%) of non-proline and non-glycine C β atoms. Most of the unassigned resonances were located on finger loops, suggesting that these residues experience signal broadening and potential

dynamics in solution. In both enzymes, the NH cross peaks from arginine residues and the indole NH cross peaks from the tryptophan side chains were not assigned. Chemical shift assignments have been deposited in the BioMagResBank (<http://www.bmrb.wisc.edu>) as entry 26761.

A comparison of the chemical shift changes and predicted order parameters of the WT and E87A variant shows that the point mutation does not significantly perturb the three-dimensional structure of the enzyme (Figure S1). The chemical environment of residues affected by the mutation was characterized by quantifying the chemical shift variations ($\Delta\delta$) observed between XlnB2 and the E87A variant. An overlay of the ^1H - ^{15}N HSQC spectra of XlnB2 and the E87A variant showed several residues with $\Delta\delta$ values of >0.05 ppm between the two proteins (Figure 1B,C). As expected, Glu87 ($\Delta\delta = 0.37$ ppm) and the surrounding residues in β -strands B5 and B6 showed the largest chemical shift variations (blue box, Figure 1A,C). Sizable chemical shift changes were also observed for residues of β -strands B3, B4, and B7 (and associated loops), as well as in the thumb (red and orange boxes, Figure 1A,C).

Comparison of the Effect of Ligand Binding on XlnB2 and the E87A Variant. Changes in the chemical environment upon ligand binding can provide valuable information about potential binding site(s), as well as the kinetics and ligand affinity for the enzyme. We performed NMR chemical shift titration measurements to characterize and compare the binding properties of XlnB2 and its E87A variant with two xyloside ligands. XlnB2 is an endoxylanase that hydrolyzes

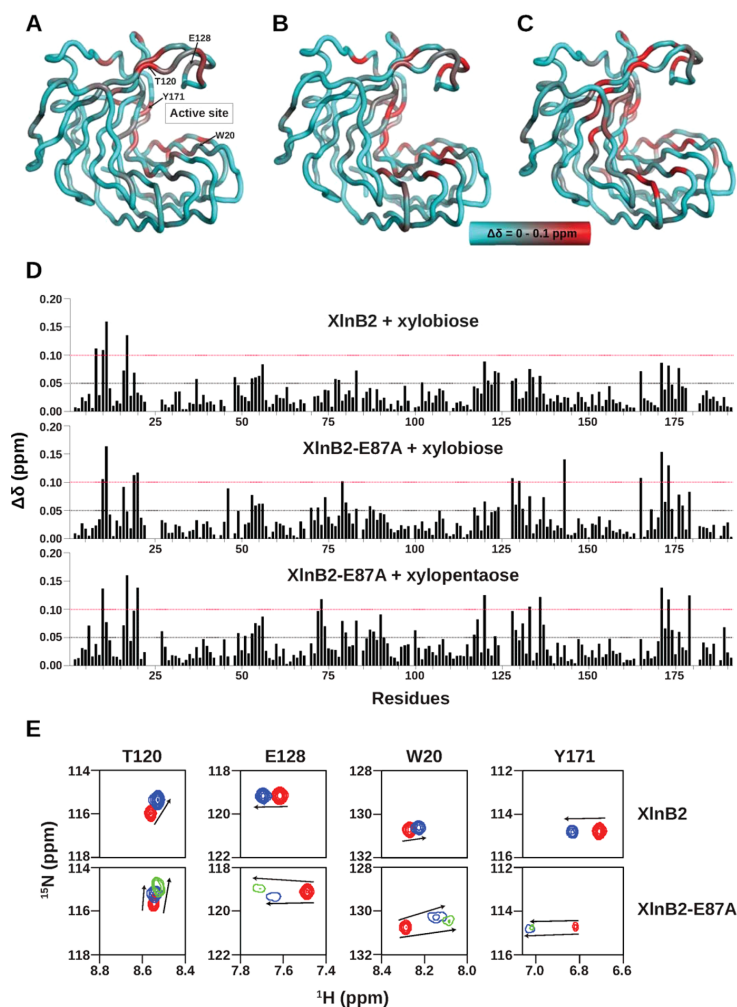


Figure 2. ^1H – ^{15}N composite chemical shift variations experienced by XlnB2 and XlnB2-E87A upon oligoxyloside binding. Structural mapping of chemical shift variations ($\Delta\delta$) observed for (A) XlnB2 induced upon xylobiose (X2) binding, (B) XlnB2-E87A induced upon xylobiose binding, and (C) XlnB2-E87A induced upon xylopentaose (X5) binding. The $\Delta\delta$ was calculated between free and ligand-saturated enzymes as described by the equation $\Delta\delta$ (ppm) = $[(\Delta\delta_{\text{HN}}^2 + \Delta\delta_{\text{N}}^2/25)/2]^{1/2}$ and reported on the three-dimensional structure of XlnB2 using a cyan-to-red color scale with a 0.0–0.1 ppm range.³⁵ The locations of the active site and key residues are reported on the 3D structure in panel A. (D) Bar graph representation of the chemical shift variations observed for XlnB2 and XlnB2-E87A upon binding to xylobiose (X2) and xylopentaose (X5). $\Delta\delta > 0.05$ ppm and $\Delta\delta > 0.1$ ppm chemical shift changes are represented by black and red dotted lines, respectively. (E) ^1H – ^{15}N chemical shift direction and magnitude for residues Trp20, Thr120, Glu128, and Tyr171 in XlnB2 and XlnB2-E87A observed upon ligand binding (see the text for details). The ^1H – ^{15}N peak position of the free enzyme and that upon saturation with X2 (X5) is shown in red and blue (green), respectively. All images were prepared with PyMOL, and graphs were produced with GraphPad Prism.

xylan to oligoxylosides with varying lengths, ranging from one to eight xylose units.¹⁶ Xylotriase (X3) is the smallest substrate of this endocleavage reaction, releasing xylose (X1) and xylobiose (X2) as uncleavable products. Here, we titrated XlnB2 and the E87A variant with incremental concentrations of X2 and X5 (E87A only), up to saturation, to determine ligand binding effects on the chemical environment of the enzyme. All residues displaying $\Delta\delta$ values of >0.05 ppm are highlighted on the 3D structure of XlnB2 and the E87A variant (Figure 2A–C). A total of four residues (Gln8, Gly10, Thr11, and Tyr17) and 11 residues (Gly10, Thr11, Phe19, Trp20, Gly79, Glu128,

Thr130, Lys143, Leu165, Tyr171, and Ile173) displayed $\Delta\delta$ values of >0.1 ppm, defined as 2σ , upon binding to X2 in XlnB2 and the E87A variant, respectively (Figure 2D). Both proteins exhibited weak binding affinities (K_i) of 7.07 and 8.71 mM, respectively, toward X2, suggesting that the mutation did not significantly perturb binding to product X2. Although the number of residues with a $\Delta\delta$ value of >0.1 ppm is larger for XlnB2-E87A, most residues affected by the presence of the ligand were located in the vicinity of the active site.

Previous studies suggested that the active site of GH11 family members could accommodate linear xylan substrates

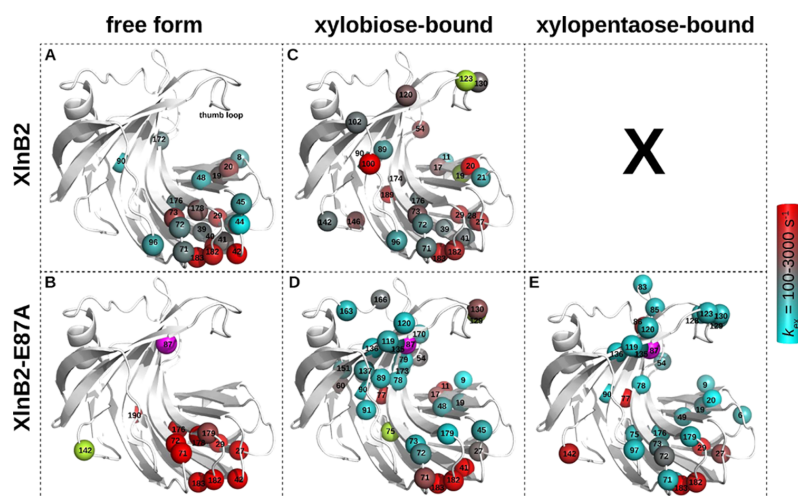


Figure 3. Residues experiencing conformational exchange in the free and ligand-bound forms of XlnB2 and XlnB2-E87A. Conformational exchange was probed using ^{15}N -CPMG NMR relaxation dispersion for XlnB2 and its catalytically impaired variant E87A in the (A and B) free, (C and D) xylobiose-saturated (X2), and (E) xylopentaose-saturated (X5) forms. Residues with a $\Delta R_2(1/\tau_{cp})$ of $>1.8\text{ s}^{-1}$ are reported as spheres on the 3D structure of the enzyme. The color-graded scale represents the rate of conformational exchange (k_{ex}) experienced by the backbone ^1H - ^{15}N vector, calculated on a per residue basis using the single-quantum ^{15}N -CPMG equation.⁶³ Residues with a $\Delta R_2(1/\tau_{cp})$ of $>1.8\text{ s}^{-1}$ for which a k_{ex} could not be extracted are colored light green. The mutation site (E87A) is represented as a magenta sphere. All images were prepared with PyMOL. Representative relaxation dispersion curves for each system are shown in Figure S2.

ranging from five to six xylose units. On the basis of these observations, we titrated XlnB2-E87A with a xylopentaose (X5) ligand (five xylose units), effectively mimicking a substrate-bound Michaelis–Menten reaction complex. A total of 10 residues (Gly10, Tyr17, Trp20, Ala73, Thr120, Phe133, Tyr136, Tyr171, Ile173, and Tyr179) showed $\Delta\delta$ values of >0.1 ppm in the E87A variant upon X5 binding (Figure 2D). Residues displaying the largest $\Delta\delta$ were located in the active site region (Figure 2C), in addition to residues in the thumb loop hinge (Thr120 and Tyr136) and fingers (Tyr17 and Tyr179). Despite titration with a longer ligand, binding of X5 to the E87A variant did not induce significantly different chemical shift variations relative to X2, both in terms of the number and magnitude shift of affected residues, and showed a similar low binding affinity for X5 ($K_d = 7.93\text{ mM}$). Our chemical shift analyses do not suggest the presence of a secondary binding site (SBS), as was observed in xylanases from other organisms to enhance xylan hydrolysis and substrate binding.^{47,48} These results suggest that the enzyme may primarily bind and stabilize a limited number of xylose units in the active site, regardless of xylose substrate unit length.

Recent studies showed that the direction and magnitude of the ^1H and ^{15}N chemical shift perturbations could provide valuable information about the nature and behavior of ligand binding events between protein variants.^{41,42,49} Using this CHESPA approach, we investigated whether the E87A mutation in XlnB2 affected thumb loop behavior with respect to ligand binding. Although the E87A mutation did not extensively perturb the general chemical shift profile induced by X2 binding relative to XlnB2, Glu128 and Thr130 showed some of the most significant thumb loop $\Delta\delta$ variations upon ligand binding in both proteins (Figure 1). Interestingly, these two variable residues of the thumb loop also showed significant ^1H and ^{15}N chemical shift variations upon introduction of the

E87A mutation (Figure 1A,C). We determined that the ^1H and ^{15}N chemical shift magnitude and direction for Glu128 were identical between XlnB2 and its mutational variant upon binding to X2 (Figure 2E). Furthermore, the direction and intensity of the movement were also similar upon binding to both X2 and X5, suggesting that the length of the ligand did not affect the magnetic environment or direction of the thumb loop movement. When ligand binds to XlnB2 and XlnB2-E87A, we also observed chemical shift variations for Thr120 (Figure 2D), a residue close to the hinge of the thumb loop. Much like Glu128, Thr120 showed coordinated chemical shift changes upon binding to X2 and X5 (Figure 2E), further suggesting that the movement and orientation of the thumb loop are the same in XlnB2 and its E87A variant. Two additional residues, Trp20 and Tyr171, located on opposite sides of the active site, were also highly affected upon ligand binding. These conserved residues play an important role in substrate binding in GH11 family members.^{1,46,50} Our results demonstrate a similar behavior, showing comparable K_d values upon ligand binding, and coordinated chemical shift responses (Figure 2E). These results suggest that the catalytically inactive XlnB2-E87A variant preserves a similar WT thumb loop movement and orientation and that these conformational properties may be important for substrate positioning, discrimination, or processivity. However, because the direction and magnitude of chemical shifts are both preserved in WT XlnB2 and the catalytically inactive E87A variant upon ligand binding, these results suggest that thumb loop movement is not involved in the actual chemical step of XlnB2 catalysis.

Conformational Exchange on the Millisecond Time Scale. It was previously demonstrated that protein flexibility and coordinated residue motions play an important role in catalysis for a variety of enzyme systems, such as dihydrofolate reductase,⁵¹ RNA-dependent RNA polymerase,⁵² and ribonu-

tease.⁵³ A number of studies have also investigated the fast time-scale dynamics in GH11 xylanases.^{1,18,22,24,25} For comparison with homologous members of the family, we also acquired ¹⁵N T_1 , T_2 , and ¹⁵N{¹H} NOE data to investigate fast time-scale dynamics in apo WT XlnB2 (Figure S2). Much like previous observations for the *B. circulans* xylanase,²³ we found that the measured T_1 , T_2 , and NOE parameters for the backbone amides do not vary significantly. A number of ¹⁵N{¹H} heteronuclear NOEs still show significantly smaller values relative to the protein average, including one thumb loop residue (G129) and three residues of the cord (T96, Y97, and G101), indicating that these positions undergo larger-amplitude fluctuations on the picosecond-to-nanosecond time scale. While a full Lipari–Szabo investigation or reduced spectral density mapping analysis in the presence and absence of ligand(s) could potentially yield information about fast time-scale dynamics in this enzyme, our XlnB2 relaxation parameters are relatively uniform across the sequence, suggesting an organized tertiary structure on this fast time scale. However, conformational fluctuations in XlnB2 occurring on the millisecond time scale could potentially be coupled to substrate recognition, stabilization, processivity (substrate sliding), and/or a direct catalytic event in GH11 xylanases. To characterize these dynamic events in solution, and to further understand the functional importance of conformational fluctuations occurring on the catalytic time scale in GH11 xylanases, we performed ¹⁵N-Carr–Purcell–Meiboom–Gill (¹⁵N-CPMG) relaxation dispersion measurements of the apo and ligand-bound forms of XlnB2 and the catalytically impaired XlnB2-E87A variant. This methodology is particularly suited to identifying residues experiencing conformational exchange (motions occurring on the millisecond time scale) in enzymes, overlapping with the rate of catalysis.⁵⁴ While Connelly et al. reported on potential microsecond-to-millisecond motions extracted from the rate of chemical exchange (R_{ex}) determined using the Lipari–Szabo model-free formalism on the *B. circulans* xylanase,²³ to the best of our knowledge, the work presented here represents the first dedicated experimental validation of the conformational fluctuations experienced by a GH11 xylanase on the catalytically relevant millisecond time frame.

All residues experiencing significant conformational exchange [$R_2(1/\tau_{CP}) > 1.8 \text{ s}^{-1}$ (Figure S3)] are shown as spheres on the structure of XlnB2 and its XlnB2-E87A variant (Figure 3). In the apo forms (Figure 3A,B), a total of 22 (12) residues showed significant millisecond motions in XlnB2 (XlnB2-E87A). In WT XlnB2 (Figure 3A), residues experiencing conformational exchange are predominantly clustered in the fingers, and few others in the palm region of the enzyme. These results demonstrate that other regions of the protein, including the thumb loop, are rigid in the apo form on this particular time scale. While the global profile remains similar in the free XlnB2-E87A variant (with visible loss of millisecond dynamics in the B1–B2–B3 sheets), the global rate of conformational exchange (k_{ex}) in the fingers and palm region is 2-fold greater for the mutant (for WT XlnB2, $k_{ex} = 1220 \pm 392 \text{ s}^{-1}$; for XlnB2-E87A, $k_{ex} = 2794 \pm 399 \text{ s}^{-1}$).

The product-bound enzyme complex between XlnB2 and X2 showed a total of 29 and 32 residues in XlnB2 and XlnB2-E87A, respectively, demonstrating an increased level of conformational exchange relative to the apo state (Figure 3C,D). Residues in the thumb loop and the region facing the active site cleft acquire conformational fluctuations upon binding to X2. We note that more residues in the palm region

facing the active site cleft experience millisecond dynamics in the XlnB2-E87A mutant relative to WT XlnB2. The substrate-bound complex with X5 in the XlnB2-E87A variant showed 33 residues displaying millisecond conformational exchange (Figure 3E), with enhanced conformational motions in the thumb loop and distal regions in the vicinity of this loop. Residues in the finger loop regions, such as Ser182 and Gly183, showed significantly greater conformational exchange rates ($k_{ex} > 3000 \text{ s}^{-1}$) for free and ligand-bound states, consistent with line broadening observations for some residues of this region (Figure 1B). A comparison of the chemical shift variations ($\Delta\delta_N$) from titration experiments with the $\Delta\omega$ obtained from ¹⁵N-CPMG showed no significant correlation, suggesting that the chemical shift variations observed upon ligand binding do not exclusively report on the existence of interconverting conformations between a ground and excited state in apo XlnB2 in solution (Figure S4).

Interestingly, while the thumb loop and the palm region facing the active site cleft are practically devoid of millisecond dynamics in the free enzyme, ligand binding (substrate and product) induces considerable conformational exchange in these regions for both XlnB2 and the E87A variant. Previous studies, based on crystal structures and MD simulations probing the nanosecond time scale, showed a dynamic shift of the thumb loop upon ligand binding.^{24,55} In contrast, our results reveal that the thumb loop region is rigid on the millisecond time scale in the free form of both XlnB2 and the E87A variant, while enhanced conformational exchange (including the thumb loop) was observed upon ligand binding in both enzymes.

Computational Simulations of the Free and Substrate-Bound Forms of XlnB2. To obtain insights into the potential relevance of longer time-scale residue dynamics in the function of XlnB2, we performed molecular dynamics (MD) simulations of XlnB2 in the free form (apoenzyme), and as a binary complex with substrates of different lengths (see Experimental Procedures). Conformations of the enzyme and enzyme–substrate complexes collected during a 1 μs MD simulation were analyzed using quasi-anharmonic analysis (QAA) to identify the conformational substates associated with the interconversion between the apo and binary complex forms. This type of analysis has been used previously for other enzyme systems to provide unique insights into the role of conformational substates and longer time-scale fluctuations.^{44,56} MD simulation results showed that the X5 substrate displayed highly dynamic behavior, where the substrate moved around the active site and adopted different orientations. This behavior made it difficult to focus on the functionally relevant motions. Therefore, we characterized the conformational behavior of X6 and X9 bound to XlnB2.

The all-atom root mean square deviations (rmsd), calculated using the first structure as reference (Figure S5), was stable throughout the simulation, indicating the stability of the proteins and convergence of the simulations. Average enzyme–substrate interaction energies (van der Waals and electrostatic) for X6- and X9-bound XlnB2 are shown in panels A and B of Figure 4, respectively, where the corresponding binary complex structures are shown in the bottom panel. Regions displaying negative average interaction energies (blue end of the spectrum) correspond to those showing the strongest favorable interactions between the enzyme and substrate. Table S1 shows the top residues with the most favorable enzyme–substrate interaction energy for X6- and X9-bound enzyme (averaged

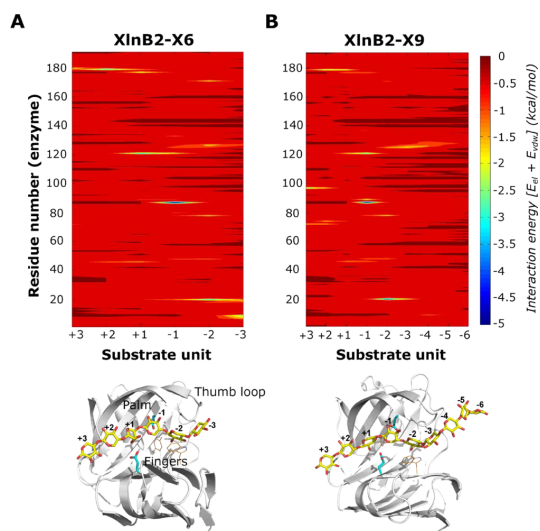


Figure 4. Protein–ligand interactions of the XlnB2–substrate binary complexes. Protein–substrate interaction energy maps for (A) XlnB2–X6 and (B) XlnB2–X9 binary complexes displaying the average interaction energies between XlnB2 and the substrate. Interaction energies were calculated as the sum of van der Waals and electrostatic energies averaged over all simulation snapshots (see [Experimental Procedures](#)). Favorable interactions correspond to the blue end of the spectrum, while unfavorable interaction energies correspond to the red end of the spectrum. The bottom panel shows the structure of the binary complex between XlnB2 and X6/X9 substrate (yellow sticks) used for the MD simulations. Residues within 4 Å of the ligand that display millisecond conformational exchange in the apo and ligand-bound states, determined from NMR relaxation dispersion experiments, are displayed as orange lines. Catalytic residues are shown as cyan sticks. The individual sugar units of the substrate are numbered according to the standard nomenclature for the reducing and nonreducing units.

over 1000 conformations). Not surprisingly, catalytic residues E87 and E177 showed strong interactions with the substrate. Other residues with favorable interaction energies encompass the fingers, thumb loop, and palm regions of XlnB2, highlighting the role of residues in these regions in ligand binding and catalysis, consistent with observations from our NMR titration experiments. Interestingly, the most favorable interactions with the enzyme are observed for substrate units occupying the –2 to +2 docking sites in both the X6- and X9-bound XlnB2 systems (Figure 4 and Table S1), consistent with previous studies indicating the importance of these docking subsites in this enzyme family.¹

We used QAA to characterize the conformational substates associated with the interconversion between the apo and substrate-bound states. QAA is a computational methodology that allows accurate identification of conformational fluctuations associated with protein function, particularly at longer time scales.⁵⁷ In particular, the use of higher-order statistics has allowed the identification of conformational dynamics (and functionally related substates) associated with enzyme catalysis and biomolecular recognition for a number of protein systems. More details regarding QAA are available in ref 57. Note that the preference for use of QAA over second-order methods (such as normal-mode analysis, principal component analysis,

or quasi-harmonic analysis) is due to its important advantage of providing quantitatively accurate and reproducible conformational motions at longer time scales. As has been previously discussed, methods such as normal-mode analysis depend on the starting structure and are known to provide only qualitative results that can be difficult to match with experimental data.⁵⁸

Our QAA results show large-scale conformational fluctuations associated with the interconversion between the apo and binary complex of XlnB2 (Figure 5A,C). These fluctuations correspond to the transition between conformational substates corresponding to populations in lower- and higher-energy regions of the conformational landscape. The apo form of XlnB2 was observed to sample conformations that are also visible in the substrate-bound state (represented by the red dashed region in Figure 5A,C), indicating that conformations associated with substrate recognition and binding are presampled in the apo form on the microsecond time scale. The top three QAA-independent component vectors (or modes) were analyzed in detail to characterize the primary dynamics associated with substrate binding in the X6 and X9 XlnB2 systems (panels B and D of Figure 5, respectively). These modes correspond to the intrinsic motions or conformational fluctuations associated with protein function⁵⁶ and correspond to the interconversion between the apo and substrate-bound states in this case. For the interconversion between free and X6-bound states (Figure 5B), the top three modes displayed large amplitudes of motions in the thumb loop, with the β -strands and loops in the fingers showing significant motions in modes 1 and 2. Specifically, modes 1 and 2 showed significant dynamics in the thumb loop, β -strands B1–4 and A2–4 in the fingers region of XlnB2, and B5–8 of the palm region, while mode 3 showed dynamics predominantly in the thumb loop (the β -strand numbering follows that in Figure 1A). Similar large amplitude motions of the thumb loop and fingers regions were observed for the interconversion between free and X9-bound states (Figure 5D). Interestingly, modes 1 and 2 in both X6 and X9 systems showed a twisting motion (as seen from the shift from the red to blue spectrum of the conformations that track the first to last conformations, respectively, in Figure 5B,D). As has been demonstrated using QAA for a number of other proteins,^{56,57} these motions are an intrinsic property or are the built-in motions of the protein. Therefore, these motions provide insights into how the movement of the thumb and palm regions may favor substrate binding. Overall, our computational results illustrate the potential role of conformational dynamics of the β -strands (and associated loops) of the fingers and thumb loop regions in the transition between the free and substrate-bound complex forms.

DISCUSSION

The use of enzyme biocatalysts in industrial chemical processes provides a sustainable and environmentally friendly alternative to a number of chemical transformations. However, the narrow range of desired characteristics such as substrate specificity and thermal stability, among others, currently limits their extensive commercial usage. The field of protein engineering has been devoted to developing strategies for designing structural and biochemical features in enzymes to impart them with new and/or improved function. While the theoretical principles of protein engineering remain simple, significant challenges remain with respect to targeting functionally relevant residues and characterizing the role of flexibility in enzyme activity. With

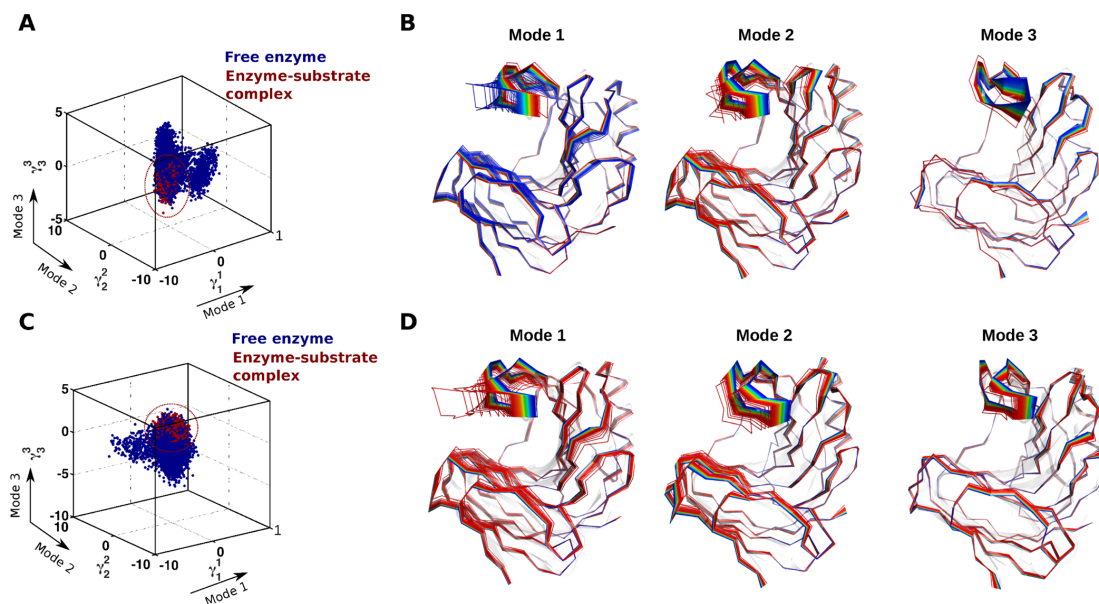


Figure 5. Conformational substates of XlnB2 determined using computational simulations and QAA. Representative conformations along the top three modes 1–3 for the interconversion between free and (A–B) X6-bound and (C–D) X9-bound XlnB2 binary complexes. A total of 200000 conformational snapshots obtained from the MD simulations were used for QAA to identify the top QAA-independent component vectors for characterizing the primary dynamics associated with the substrate binding process in XlnB2 (see [Experimental Procedures](#)).

advances in biocomputational approaches and NMR spectroscopy methodologies, the role of conformational fluctuations in substrate binding and product release, either directly or indirectly involved in the catalytic step, has been partially characterized.^{59,60} In the work presented here, we aimed to characterize conformational flexibility on the catalytically relevant millisecond time scale for XlnB2 and its catalytically impaired E87A variant.

The role of conformational flexibility in modulating enzyme catalysis has been well established for a variety of proteins.^{51,61} Previous studies of GH11 family enzymes have shown the existence of fast nanosecond time-scale motions in some xylanase members.¹ MD studies based on 1 ns simulations of xylanase in *T. reesei* suggested that flexibility of the thumb loop may be important for allowing access of substrate to the active site,⁵⁵ while simulations of xylanase from *Bacillus subtilis* showed that the extent of thumb loop movement is reduced in the presence of X2, which restricts access to the active site.¹⁸ However, these studies explored the dynamical properties of xylanases on the nanosecond time scale, which is several orders of magnitude faster than the catalytically relevant millisecond time scale of substrate turnover in these enzymes. On the basis of observations of fitting to the model-free formalism, Connelly et al.²³ suggested that several residues in the apo form of BCX may display longer time-scale microsecond-to-millisecond mobility.

In this study, we provide the first experimental validation of conformational exchange processes occurring on the catalytically relevant millisecond time scale in xylanases, using NMR relaxation dispersion experiments to probe the potential influence of dynamics on ligand binding. Our results show that ligand binding enhances millisecond dynamics in the thumb loop and fingers of XlnB2. The apo form of the enzyme

displayed a rigid thumb loop on the millisecond time scale, as previously suggested on the basis of fits from model-free analyses performed on the *B. circulans* xylanase.²³ Ligand binding enhanced millisecond dynamics in the thumb loop region, in contrast to previous studies in which binding of X2 was shown to reduce mobility of the thumb loop on the fast nanosecond time scale, restricting access to the active site.¹⁸ Interestingly, the number of residues displaying dynamics increased upon ligand binding, in contrast to previous model-free observations.²³ The enhanced conformational exchange in the thumb loop region upon ligand binding indicates that this motion may play an important role in ligand processing, consistent with observations from mutations in the thumb loop region leading to decreased activity of the xylanase from *Thermobacillus xylanilyticus*.^{19,22} The thumb loop has been shown to play important roles in substrate selectivity,¹⁹ product release, and (de)glycosylation.²⁵ Residues of the thumb loop, including Glu128 and Thr120 near the hinge region of the loop, show significant chemical shift variations upon ligand binding in XlnB2. Further, enhanced conformational exchange in residues Asn123, Gly128, and Thr130 of the thumb loop upon ligand binding may allow the enzyme to adopt conformations that (destabilize) stabilize the binary complex with the (product) substrate. Changes in the conformational motions and chemical shift variations upon ligand binding are also observed for residues, such as Trp20 and Tyr171, in the fingers regions. Modifications of these residues through point mutations will provide important insights into their role in substrate binding and conformational fluctuations affecting function in XlnB2.

Our results reveal the presence of similar millisecond motions in the fingers and palm regions, triggered by binding of ligands of different sizes ([Figure 3](#)), suggesting that dynamics

does not play a significant role in substrate selectivity. The length of the ligand does not seem to play a significant role in enhanced conformational exchange observed in the fingers and palm regions, as illustrated by similar rates of conformational exchange observed in the presence of X2 and X5 ligands (Figure 3). Similar conformational exchange rates for the catalytically impaired XlnB2-E87A variant suggest that the mutation did not affect ligand binding in XlnB2, as evidenced from similar K_d values and comparable thumb loop CHESPA behavior (see above). Analysis of the interconversion between the free and ligand-bound states of XlnB2, using quasi-harmonic analysis of conformations obtained from MD simulations, shows substates that display enhanced dynamics in the fingers and thumb loop regions, consistent with our results from NMR relaxation experiments (Figure 3). Some of the residues displaying dynamics in the apo form and enhanced dynamics upon ligand binding are located within 4 Å of the ligand in the active site (orange sticks in the 3D structures shown in Figure 4), suggesting a direct effect of these conformational fluctuations upon ligand binding. QAA further revealed a twisting motion of the thumb loop and fingers regions in the top dynamic modes 1 and 2. We speculate that this twisting movement presents an alternate robust motion compared to the open–close motion for a processive enzyme like XlnB2. We note that the open–close motion has been observed for the thumb loop in the presence of substrate in the *B. circulans* xylanase.¹

In summary, we characterized the broad dynamic profile experienced by the *S. lividans* 66 XlnB2 and its catalytically impaired E87A variant on the slow (millisecond) time scale using MD simulations and NMR relaxation dispersion experiments, in addition to determining its crystal structure. Our results show a distinct millisecond dynamic behavior for these enzymes in the apo and ligand-bound states. A comparison of these two states revealed that the ligand-bound protein experiences enhanced conformational exchange in the fingers and thumb loop regions, suggesting a potential role of these fluctuations for optimal ligand recognition, positioning, discrimination, processivity, and/or catalysis. This observation suggests a ligand-induced conformational shift in line with a potential induced fit mechanism of molecular recognition in GH11 xylanases, which was previously suggested for GH8 xylanases.⁶² This provides an avenue to further investigate the relative importance of conformational selection and induced fit in the catalytic mechanism of GH11 xylanases.

■ ASSOCIATED CONTENT

● Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.6b00130.

Comparisons between the structural properties of WT XlnB2 and XlnB2-E87A (Figure S1), ^{15}N T_1 , T_2 , and $^{15}\text{N}\{^1\text{H}\}$ NOEs for free XlnB2 (Figure S2), representative relaxation dispersion curves (Figure S3), a comparison between $\Delta\omega$ and $\Delta\delta$ (Figure S4), the convergence of MD simulations (Figure S5), and average interaction energies between XlnB2 and the xylohexaose (X6) and xylononaose (X9) substrates (Table S1) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

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

Present Addresses

[∇]N.N.-T.: Department of Biotechnology, Institute of New Technology, Institute of Military Science and Technology, 17 Hoang Sam, Hanoi, Vietnam.

●D.G.: Structural Biology Initiative, CUNY Advanced Science Research Center, 85 St. Nicholas Terrace, New York, NY 10031.

Funding

This work was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant via Grant RGPIN-2016-05557 (to N.D.) and the National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health (NIH) via Grant R01GM105978 (to N.D. and P.K.A.). N.D. holds a Fonds de Recherche Québec-Santé (FRQS) Research Scholar Junior 2 Career Award, and D.G. held an NSERC Alexander Graham Bell Canada Graduate Scholarship.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Tara Sprules and Sameer Al-Abdul-Wahid of the Quebec/Eastern Canada High Field NMR Facility (McGill University, Montreal, QC) for their excellent technical assistance, in addition to Claude Dupont (INRS-Institut Armand-Frappier, Université du Québec) for providing the original XlnB2 construct and for transferring his vast *Streptomyces* knowledge to our lab.

■ REFERENCES

- (1) Paës, G., Berrin, J. G., and Beaugrand, J. (2012) GH11 xylanases: Structure/function/properties relationships and applications. *Biotechnol. Adv.* 30, 564–592.
- (2) Uffen, R. L. (1997) Xylan degradation: a glimpse at microbial diversity. *J. Ind. Microbiol. Biotechnol.* 19, 1–6.
- (3) Xia, T., and Wang, Q. (2009) Directed evolution of *Streptomyces lividans* xylanase B toward enhanced thermal and alkaline pH stability. *World J. Microbiol. Biotechnol.* 25, 93–100.
- (4) Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, R. C., Jr., and Warren, R. A. (1991) Domains in microbial beta-1, 4-glycanases: sequence conservation, function, and enzyme families. *Microbiol. Rev.* 55, 303–315.
- (5) Henrissat, B., and Davies, G. (1997) Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol.* 7, 637–644.
- (6) Campbell, R. L., R. D. R., Wakarchuk, W. W., To, R., Sung, W., and Yaguchi, M. (1993) A comparison of the structures of the 20 kDa xylanases from *Trichoderma harzianum* and *Bacillus circulans*. In *Proceedings of the second TRICEL symposium on Trichoderma reesei cellulases and other hydrolases* (Suominen, P. R. T., Ed.) pp 63–72.
- (7) Törrönen, A., Harkki, A., and Rouvinen, J. (1994) Three-dimensional structure of endo-1,4-beta-xylanase II from *Trichoderma reesei*: two conformational states in the active site. *EMBO J.* 13, 2493–2501.
- (8) Törrönen, A., and Rouvinen, J. (1995) Structural comparison of two major endo-1,4-xylanases from *Trichoderma reesei*. *Biochemistry* 34, 847–856.
- (9) McIntosh, L. P., Hand, G., Johnson, P. E., Joshi, M. D., Korner, M., Plesniak, L. A., Ziser, L., Wakarchuk, W. W., and Withers, S. G. (1996) The pKa of the general acid/base carboxyl group of a

- glycosidase cycles during catalysis: a ^{13}C -NMR study of bacillus circulans xylanase. *Biochemistry* 35, 9958–9966.
- (10) Connelly, G. P., and McIntosh, L. P. (1998) Characterization of a buried neutral histidine in *Bacillus circulans* xylanase: internal dynamics and interaction with a bound water molecule. *Biochemistry* 37, 1810–1818.
- (11) Plesniak, L. A., Connelly, G. P., Wakarchuk, W. W., and McIntosh, L. P. (1996) Characterization of a buried neutral histidine residue in *Bacillus circulans* xylanase: NMR assignments, pH titration, and hydrogen exchange. *Protein Sci.* 5, 2319–2328.
- (12) Wakarchuk, W. W., Campbell, R. L., Sung, W. L., Davoodi, J., and Yaguchi, M. (1994) Mutational and crystallographic analysis of the active site residues of the *Bacillus circulans* xylanase. *Protein Sci.* 3, 467–475.
- (13) Wakarchuk, W. W., Sung, W. L., Campbell, R. L., Cunningham, A., Watson, D. C., and Yaguchi, M. (1994) Thermostabilization of the *Bacillus subtilis* xylanase by the introduction of disulfide bonds. *Protein Eng., Des. Sel.* 7, 1379–1386.
- (14) Lawson, S. L., Wakarchuk, W. W., and Withers, S. G. (1997) Positioning the acid/base catalyst in a glycosidase: studies with *Bacillus circulans* xylanase. *Biochemistry* 36, 2257–2265.
- (15) Törrönen, A., and Rouvinen, J. (1997) Structural and functional properties of low molecular weight endo-1,4-beta-xylanases. *J. Biotechnol.* 57, 137–149.
- (16) Kluepfel, D., Vats-Mehta, S., Aumont, F., Shareck, F., and Morosoli, R. (1990) Purification and characterization of a new xylanase (xylanase B) produced by *Streptomyces lividans* 66. *Biochem. J.* 267, 45–50.
- (17) Hurtubise, Y., Shareck, F., Kluepfel, D., and Morosoli, R. (1995) A cellulase/xylanase-negative mutant of *Streptomyces lividans* 1326 defective in cellobiose and xylobiose uptake is mutated in a gene encoding a protein homologous to ATP-binding proteins. *Mol. Microbiol.* 17, 367–377.
- (18) Vieira, D. S., Degreve, L., and Ward, R. J. (2009) Characterization of temperature dependent and substrate-binding cleft movements in *Bacillus circulans* family 11 xylanase: a molecular dynamics investigation. *Biochim. Biophys. Acta, Gen. Subj.* 1790, 1301–1306.
- (19) Paës, G., Tran, V., Takahashi, M., Boukari, I., and O'Donohue, M. J. (2007) New insights into the role of the thumb-like loop in GH-11 xylanases. *Protein Eng., Des. Sel.* 20, 15–23.
- (20) Cervera Tison, M., Andre-Leroux, G., Lafond, M., Georis, J., Juge, N., and Berrin, J. G. (2009) Molecular determinants of substrate and inhibitor specificities of the *Penicillium griseofulvum* family 11 xylanases. *Biochim. Biophys. Acta, Proteins Proteomics* 1794, 438–445.
- (21) Murakami, M. T., Arni, R. K., Vieira, D. S., Degreve, L., Ruller, R., and Ward, R. J. (2005) Correlation of temperature induced conformation change with optimum catalytic activity in the recombinant G/11 xylanase A from *Bacillus subtilis* strain 168 (1A1). *FEBS Lett.* 579, 6505–6510.
- (22) Pollet, A., Vandermarliere, E., Lammertyn, J., Strelkov, S. V., Delcour, J. A., and Courtin, C. M. (2009) Crystallographic and activity-based evidence for thumb flexibility and its relevance in glycoside hydrolase family 11 xylanases. *Proteins: Struct., Funct., Genet.* 77, 395–403.
- (23) Connelly, G. P., Withers, S. G., and McIntosh, L. P. (2000) Analysis of the dynamic properties of *Bacillus circulans* xylanase upon formation of a covalent glycosyl-enzyme intermediate. *Protein Sci.* 9, 512–524.
- (24) Havukainen, R., Torronen, A., Laitinen, T., and Rouvinen, J. (1996) Covalent binding of three epoxyalkyl xylosides to the active site of endo-1,4-xylanase II from *Trichoderma reesei*. *Biochemistry* 35, 9617–9624.
- (25) Paës, G., Cortes, J., Simeon, T., O'Donohue, M. J., and Tran, V. (2012) Thumb-loops up for catalysis: a structure/function investigation of a functional loop movement in a GH11 xylanase. *Comput. Struct. Biotechnol. J.* 1, e201207001.
- (26) Vandermarliere, E., Bourgeois, T. M., Rombouts, S., van Campenhout, S., Volckaert, G., Strelkov, S. V., Delcour, J. A., Rabijns, A., and Courtin, C. M. (2008) Crystallographic analysis shows substrate binding at the -3 to $+1$ active-site subsites and at the surface of glycoside hydrolase family 11 endo-1,4- β -xylanases. *Biochem. J.* 410, 71–79.
- (27) Gauthier, C., Li, H., and Morosoli, R. (2005) Increase in xylanase production by *Streptomyces lividans* through simultaneous use of the Sec- and Tat-dependent protein export systems. *Appl. Environ. Microbiol.* 71, 3085–3092.
- (28) Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* 77, 51–59.
- (29) Wouters, J., Georis, J., Engher, D., Vandenhoute, J., Dusart, J., Frère, J. M., Depiereux, E., and Charlier, P. (2001) Crystallographic analysis of family 11 endo-beta-1,4-xylanase Xyl11 from *Streptomyces sp.* S38. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 57, 1813–1819.
- (30) Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 60, 2126–2132.
- (31) McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* 40, 658–674.
- (32) Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 66, 12–21.
- (33) Cavanagh, J., Fairbrother, W., Palmer, A. G., III, Skelton, N., and Rance, M. (2006) *Protein NMR Spectroscopy: Principles and Practice*, 2nd ed., Elsevier Academic, San Diego.
- (34) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* 6, 277–293.
- (35) Grzesiek, S., Stahl, S. J., Wingfield, P. T., and Bax, A. (1996) The CD4 determinant for downregulation by HIV-1 Nef directly binds to Nef. Mapping of the Nef binding surface by NMR. *Biochemistry* 35, 10256–10261.
- (36) Mittermaier, A., and Meneses, E. (2013) Analyzing protein-ligand interactions by dynamic NMR spectroscopy. *Methods Mol. Biol.* 1008, 243–266.
- (37) Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry* 33, 5984–6003.
- (38) 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.
- (39) Carver, J. P., and Richards, R. E. (1972) A general two-site solution for the chemical exchange produced dependence of T_2 upon the Carr–Purcell pulse separation. *J. Magn. Reson.* 6, 89–105.
- (40) Palmer, A. G., 3rd. (2004) NMR characterization of the dynamics of biomacromolecules. *Chem. Rev.* 104, 3623–3640.
- (41) Axe, J. M., and Boehr, D. D. (2013) Long-range interactions in the alpha subunit of tryptophan synthase help to coordinate ligand binding, catalysis, and substrate channeling. *J. Mol. Biol.* 425, 1527–1545.
- (42) Selvaratnam, R., VanSchouwen, B., Fogolari, F., Mazhab-Jafari, M. T., Das, R., and Melacini, G. (2012) The projection analysis of NMR chemical shifts reveals extended EPAC autoinhibition determinants. *Biophys. J.* 102, 630–639.
- (43) Case, D. A., Darden, T. A., Cheatham, T. E., III, Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B., Hayik, S., Roitberg, A., Seabra, G., Swails, J., Götz, A. W., Kolossváry, I., Wong, K. F., Paesani, F., Vanicek, J., Wolf, R. M., Liu, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M.-J., Cui, G., Roe, D. R., Mathews, D. H., Seetin, M. G., Salomon-Ferrer, R., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., and Kollman, P. A. (2012) AMBER 12, University of California, San Francisco.

- (44) Ramanathan, A., and Agarwal, P. K. (2011) Evolutionarily Conserved Linkage between Enzyme Fold, Flexibility, and Catalysis. *PLoS Biol.* 9, e1001193.
- (45) Agarwal, P. K. (2004) Computational Studies of the mechanism of cis/trans isomerization in HIV-1 catalyzed by cyclophilin A. *Proteins: Struct., Funct., Genet.* 56, 449–463.
- (46) Wakarchuk, W. W., Campbell, R. L., Sung, W. L., Davoodi, J., and Yaguchi, M. (1994) Mutational and crystallographic analyses of the active site residues of the *Bacillus circulans* xylanase. *Protein Sci.* 3, 467–475.
- (47) Ludwiczek, M. L., Heller, M., Kantner, T., and McIntosh, L. P. (2007) A secondary xylan-binding site enhances the catalytic activity of a single-domain family 11 glycoside hydrolase. *J. Mol. Biol.* 373, 337–354.
- (48) Cuyvers, S., Dornez, E., Delcour, J. A., and Courtin, C. M. (2012) Occurrence and functional significance of secondary carbohydrate binding sites in glycoside hydrolases. *Crit. Rev. Biotechnol.* 32, 93–107.
- (49) Gagné, D., Narayanan, C., and Doucet, N. (2015) Network of long-range concerted chemical shift displacements upon ligand binding to human angiogenin. *Protein Sci.* 24, 525–533.
- (50) Ludwiczek, M. L., D'Angelo, I., Yalloway, G. N., Brockerman, J. A., Okon, M., Nielsen, J. E., Strynadka, N. C., Withers, S. G., and McIntosh, L. P. (2013) Strategies for modulating the pH-dependent activity of a family 11 glycoside hydrolase. *Biochemistry* 52, 3138–3156.
- (51) Hanoian, P., Liu, C. T., Hammes-Schiffer, S., and Benkovic, S. (2015) Perspectives on electrostatics and conformational motions in enzyme catalysis. *Acc. Chem. Res.* 48, 482–489.
- (52) Boehr, D. D., Liu, X., and Yang, X. (2014) Targeting structural dynamics of the RNA-dependent RNA polymerase for anti-viral strategies. *Curr. Opin. Virol.* 9, 194–200.
- (53) 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.
- (54) Kempf, J. G., and Loria, J. P. (2004) Measurement of intermediate exchange phenomena. *Methods Mol. Biol.* 278, 185–231.
- (55) Muilu, J., Torronen, A., Perakyla, M., and Rouvinen, J. (1998) Functional conformational changes of endo-1,4-xylanase II from *Trichoderma reesei*: a molecular dynamics study. *Proteins: Struct., Funct., Genet.* 31, 434–444.
- (56) Ramanathan, A., Savol, A. J., Agarwal, P. K., and Chennubhotla, C. S. (2012) Event detection and sub-state discovery from biomolecular simulations using higher-order statistics: application to enzyme adenylate kinase. *Proteins: Struct., Funct., Genet.* 80, 2536–2551.
- (57) Ramanathan, A., Savol, A. J., Langmead, C. J., Agarwal, P. K., and Chennubhotla, C. S. (2011) Discovering conformational sub-states relevant to protein function. *PLoS One* 6, e15827.
- (58) Ramanathan, A., and Agarwal, P. K. (2009) Computational identification of slow conformational fluctuations in proteins. *J. Phys. Chem. B* 113, 16669–16680.
- (59) Palmer, A. G., 3rd. (2015) Enzyme dynamics from NMR spectroscopy. *Acc. Chem. Res.* 48, 457–465.
- (60) Kohen, A. (2015) Role of dynamics in enzyme catalysis: substantial versus semantic controversies. *Acc. Chem. Res.* 48, 466–473.
- (61) Hammes, G. G., Benkovic, S. J., and Hammes-Schiffer, S. (2011) Flexibility, diversity, and cooperativity: pillars of enzyme catalysis. *Biochemistry* 50, 10422–10430.
- (62) De Vos, D., Collins, T., Nerinckx, W., Savvides, S. N., Claeysens, M., Gerday, C., Feller, G., and Van Beeumen, J. (2006) Oligosaccharide binding in family 8 glycosidases: crystal structures of active-site mutants of the beta-1,4-xylanase pXyl from *Pseudoalteromonas haloplanktis* TAH3a in complex with substrate and product. *Biochemistry* 45, 4797–4807.
- (63) Manley, G., and Loria, J. P. (2012) NMR insights into protein allostery. *Arch. Biochem. Biophys.* 519, 223–231.

Ligand Binding Enhances Millisecond Conformational Exchange in Xylanase B2 from *Streptomyces lividans*

Donald Gagné^{1,#}, Chitra Narayanan¹, Nhung Nguyen-Thi^{1,&}, Louise D. Roux¹, David N. Bernard¹, Joseph S. Brunzelle², Jean-François Couture^{3,4,5}, Pratul K. Agarwal^{6,7} & Nicolas Doucet^{1,4,5*}

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

²Department of Molecular Pharmacology and Biological Chemistry, Feinberg School of Medicine, Northwestern University, 320 East Superior Street, Chicago, IL, 60611, USA

³Ottawa Institute of Systems Biology, Department of Biochemistry, Microbiology and Immunology, University of Ottawa, 451 Smyth Road, Ottawa, Ontario, K1H 8M5, Canada.

⁴PROTEO, the Québec Network for Research on Protein Function, Engineering, and Applications, 1045 Avenue de la Médecine, Université Laval, Québec, QC, G1V 0A6, Canada.

⁵GRASP, the Groupe de Recherche Axé sur la Structure des Protéines, 3649 Promenade Sir William Osler, McGill University, Montréal, QC, H3G 0B1, Canada.

⁶Computational Biology Institute and Computer Science and Mathematics Division, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, TN, 37830, USA.

⁷Department of Biochemistry, Cellular and Molecular Biology, University of Tennessee, Knoxville, TN, 37996, USA.

*Corresponding author: Email: nicolas.doucet@iaf.inrs.ca; Phone: (450) 687-5010, ext. 4212.

&Present address: Department of Biotechnology, Institute of New Technology, Institute of Military Science and Technology, 17 Hoang Sam, Hanoi, Vietnam.

#Present address: Structural Biology Initiative, CUNY Advanced Science Research Center, 85 St. Nicholas Terrace, New York, NY, 10031, USA.

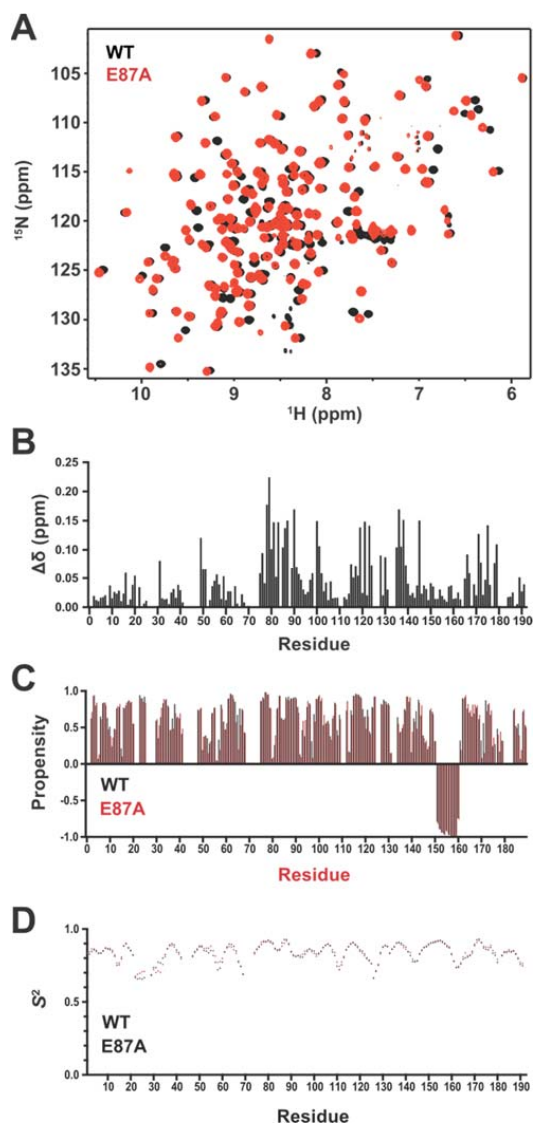


Figure S1. Comparison of structural changes between XlnB2 and XlnB2-E87A. A) Overlay of ^1H - ^{15}N HSQC spectra of XlnB2 (black) and XlnB2-E87A (red). B) Composite ^1H - ^{15}N chemical shift differences between the E87A mutant and WT XlnB2 (see Experimental Procedures for details). C) Secondary structure propensity for the XlnB2 (black) and XlnB2-E87A (red) predicted using TALOS. D) Order parameters (S^2) for XlnB2 (black) and XlnB2-E87A (red) predicted using TALOS.

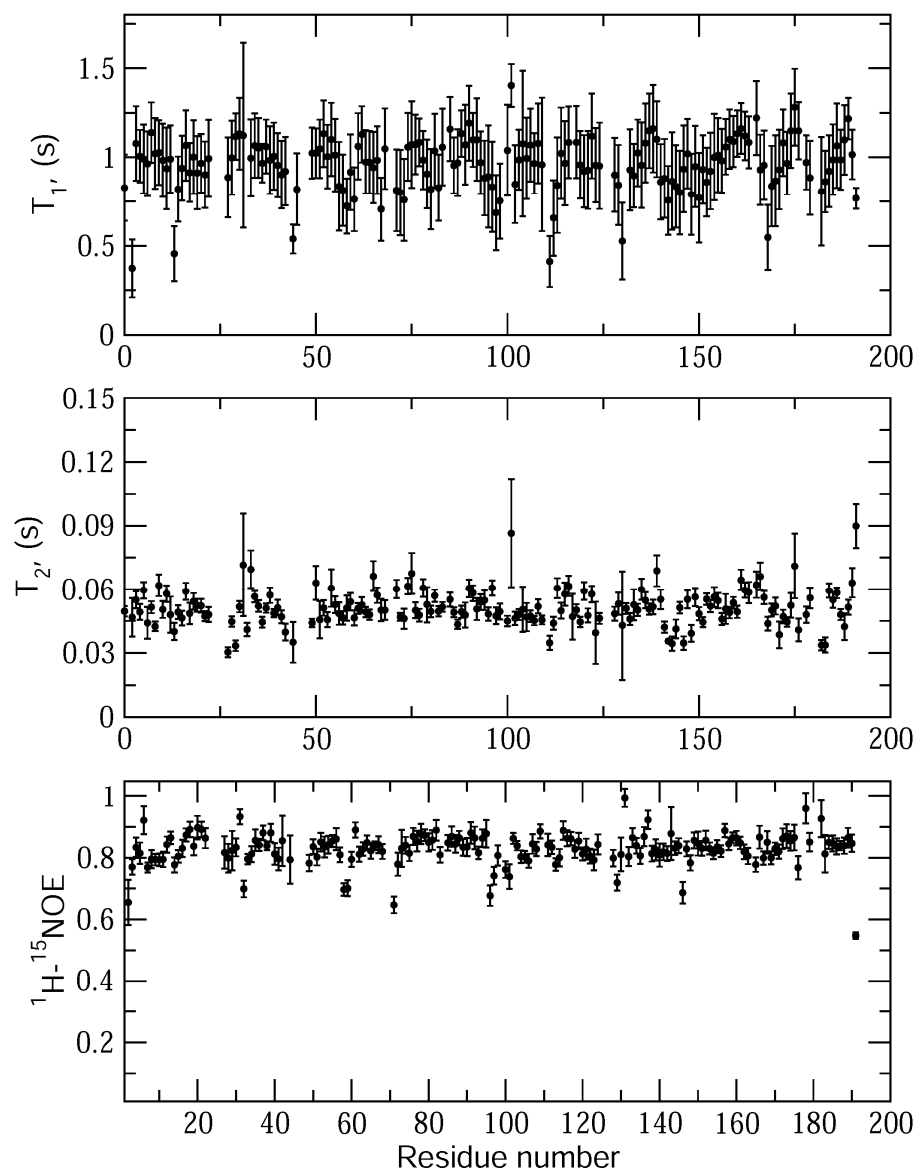


Figure S2. ^{15}N T_1 , T_2 and NOE values for the main chain amides of the apo form of XlnB2. The mean values for all residues are shown as filled circles, with the standard deviation represented using error bars. Two residues (32 and 131 were removed from the T_1 plot due to aberrant values).

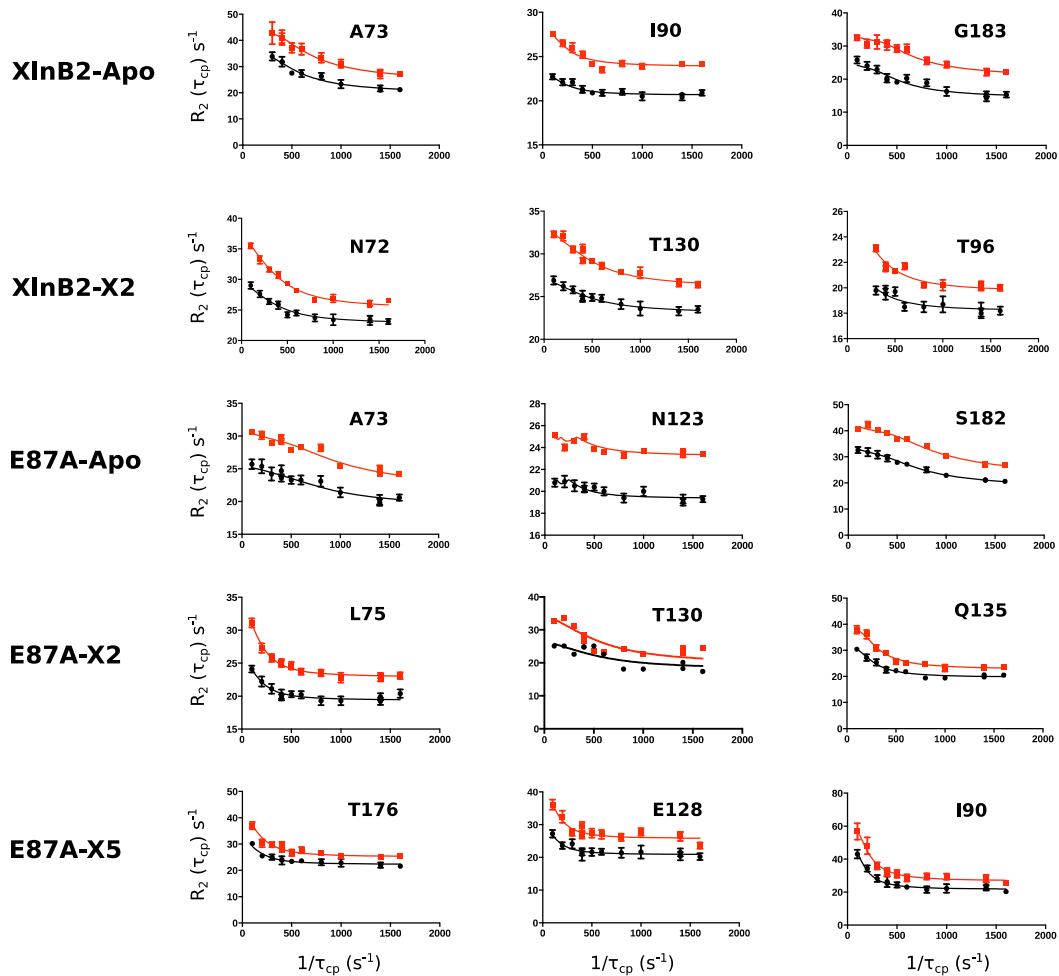


Figure S3. Representative relaxation dispersion curves at 600 MHz (black) and 800 MHz (red) (298K) are shown for residues experiencing ^1H - ^{15}N conformational exchange on the millisecond timescale in the *apo* and ligand-bound states of XlnB2 and XlnB2-E87A. See Experimental Procedures.

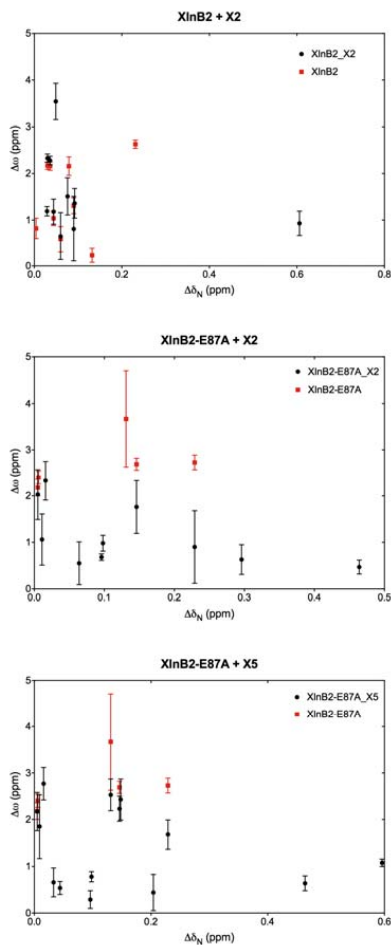


Figure S4. Chemical shift fluctuations (in the nitrogen dimension only, $\Delta\delta_N$) observed from titration experiments relative to $\Delta\omega$ values extracted from single-residue dual fits extracted from ^{15}N -CPMG relaxation dispersion experiments. This comparison provides information on whether the conformational exchange observed by ^{15}N -CPMG results from the exchange between the apo and bound-like conformations in absence of ligand. $\Delta\delta_N$ from XlnB2 titration with X2 were compared to $\Delta\omega$ values from ^{15}N -CPMG experiments performed on both the apo and X2-bound XlnB2 (red squares and black circles in A, respectively). Residues in the binding pocket whose chemical shifts might be influenced by the spatial proximity to the ligand were excluded from this comparison, as only the effect of conformational exchange on $\Delta\delta_N$ was investigated. Similar comparisons were made for the XlnB2 E87A titrations with X2 and X5 (B and C, respectively). Results show no significant correlation between these parameters, suggesting that the chemical shift variations observed upon ligand binding do not exclusively report on the existence of interconverting conformations between a ground and excited state in free XlnB2 in solution.

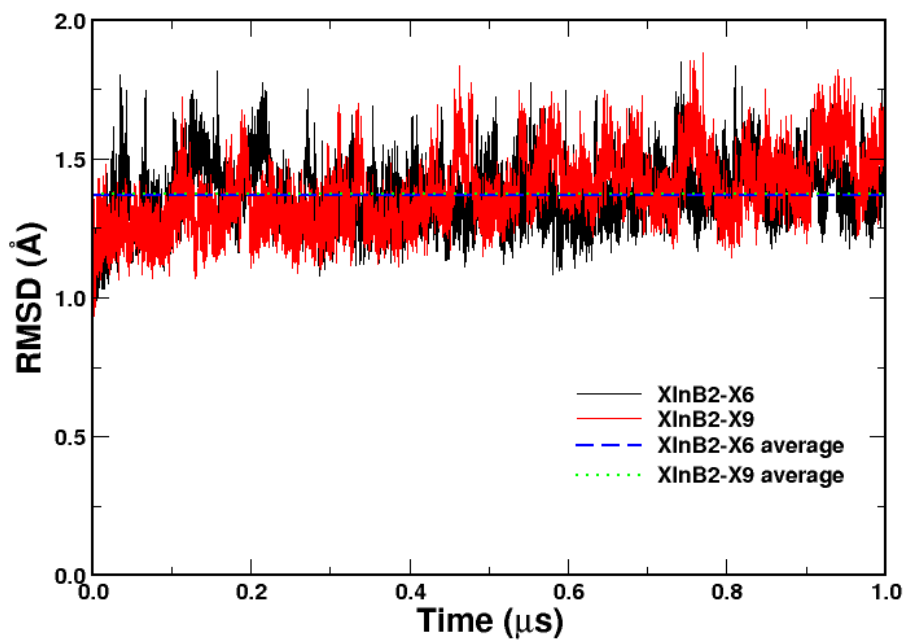


Figure S5. All-atom root mean square deviation (RMSD) as a function of simulation time for XlnB2 in the apo (black) and xylononaose-bound (X9) binary complex (red). RMSDs were calculated using the first structure as the reference structure. Running averages are shown as blue dashed and green dotted lines for XlnB2-X6 and XlnB2-X9, respectively.

Table S1. Average enzyme-substrate interaction energies for the interaction between XlnB2 and xylohexaose (X6) and xyloonaose (X9) substrates. Interaction energies were calculated as the sum of van der Waals and electrostatic energies averaged over all simulation snapshots (see Experimental Procedures).

XlnB2-X6			XlnB2-X9		
Enzyme residue	Substrate unit	Energy (kcal/mol)	Enzyme residue	Substrate unit	Energy (kcal/mol)
87	-1	-5.13	87	-1	-5.24
121	+1	-3.30	20	-2	-4.38
179	+2	-3.08	121	-1	-2.86
20	-2	-3.0	97	+3	-2.54
9	-3	-2.4	179	+2	-2.45
177	+1	-1.95	89	-1	-1.83
8	-3	-1.71	179	+1	-1.78
121	-1	-1.67	78	-2	-1.62
171	-2	-1.5	125	-3	-1.59
78	-2	-1.57	20	-3	-1.56

CHAPITRE 4 : RÉSULTATS COMPLÉMENTAIRES

4.1 Xylanase C de *S. lividans*

4.1.1 Attribution des résonances RMN de la chaîne principale

Le spectre ^1H - ^{15}N HSQC de XlnC présente des résonances bien dispersées en dimensions proton et azote, offrant une indication que la protéine est correctement repliée (Figure 4.1). En utilisant une combinaison d'expériences RMN 2D ^1H - ^{15}N (HSQC) et 3D ^1H - ^{13}C - ^{15}N (HNCACB et $\text{C}_{\text{B}}\text{C}_{\text{A}}(\text{CO})\text{NH}$), il a été possible d'attribuer les résonances des ^1H , ^{13}C et ^{15}N de la chaîne principale de l'enzyme. Au total, 187 résonances ^1H - ^{15}N de XlnC ont été attribuées, soit 100 % des résidus (en excluant les quatre prolines et le résidu N-terminal).

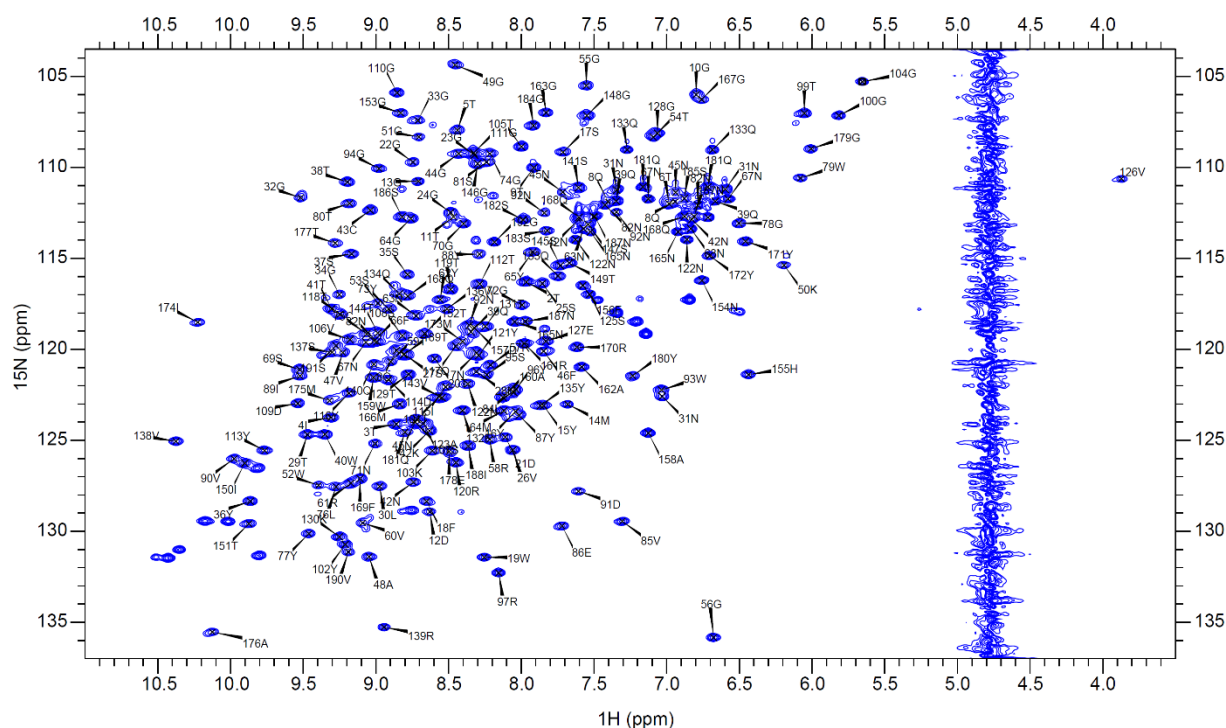


Figure 4.1 : Spectre ^1H - ^{15}N HSQC attribué de XlnC de *S. lividans*. Les résidus attribués sont identifiés en utilisant la numérotation séquentielle des acides aminés de l'enzyme.

4.1.2 Effet de la liaison d'un ligand

Des expériences de titrage RMN ont permis de repérer les résidus impliqués dans la liaison d'un ligand au site actif. Des spectres ^1H - ^{15}N HSQC de XInC ont été obtenus en utilisant des concentrations croissantes en xylobiose (X2) jusqu'à saturation (un produit non-clivable de cette endoxylanase), tel que présenté au chapitre 3. Au total, 11 résidus (Thr9, Ser17, Phe18, Phe46, Ala48, Gly49, Ser125, Glu127, Gly128, Ile174 et Thr177) présentent une variation de déplacements chimiques ^1H - ^{15}N ($\Delta\delta$) $\geq 0,05$ ppm entre l'enzyme libre et l'enzyme saturée par le ligand. Tel qu'attendu, les résidus affectés par la liaison du ligand sont retrouvés dans une petite zone autour du site actif, incluant des résidus de la boucle du « pouce » (Figure 4.2).

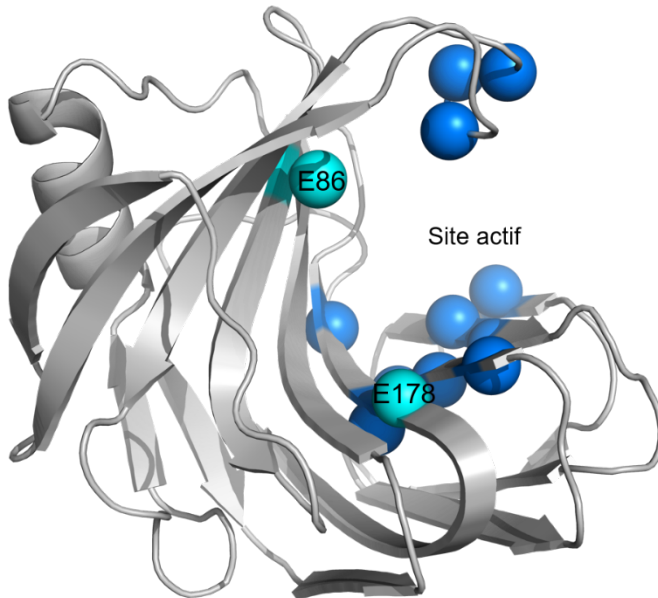


Figure 4.2 : Résidus de XInC présentant une variation de déplacements chimiques ^1H - ^{15}N ($\Delta\delta$) lors de la liaison au xylobiose (X2).

Les sphères bleues représentent les résidus présentant un déplacement chimique ($\Delta\delta$) $\geq 0,05$ ppm. Les $\Delta\delta$ ont été calculés entre l'enzyme libre et l'enzyme saturée par le ligand, selon l'équation $\Delta\delta(\text{ppm}) = ((\Delta\delta_{\text{HN}}^2 + \Delta\delta_{\text{N}}^2/25)/2)^{1/2}$ (tel que décrit au chapitre 3). Les résidus catalytiques sont représentés par des sphères cyan.

La structure de XInC de *S. lividans* est un modèle d'homologie construit à partir de son homologue XInB2.

4.1.3 Changements conformationnels à l'échelle de la milliseconde

XInB2 et XInC de *S. lividans* partagent une forte identité de séquence (77 %), en plus de propriétés physico-chimiques et thermodynamiques similaires, ce qui laisse entrevoir la possibilité qu'elles puissent également partager des propriétés dynamiques. Tel que démontré par des expériences de dispersion de relaxation ^{15}N -CPMG présentées au chapitre 3, la xylanase B2 démontre une flexibilité dans les régions de la « paume » et des « doigts » en absence de liaison au ligand, en plus de changements conformationnels apparaissant dans la région du

« pouce » lors de la liaison au xylobiose (Figure 4.3, haut). Au contraire, les expériences de dispersion de relaxation ^{15}N -CPMG réalisées sur la xylanase C montrent qu'en absence de ligand, cette protéine ne présente pas d'échange conformationnel à l'échelle de temps de la milliseconde. Cependant, en présence de xylobiose (X2), des résidus de la boucle du « pouce » acquièrent de la flexibilité (Figure 4.3, bas). De plus, ces résidus dynamiques présentent des taux d'échange conformationnel (k_{ex}) plus lents que ceux de XlnB2.

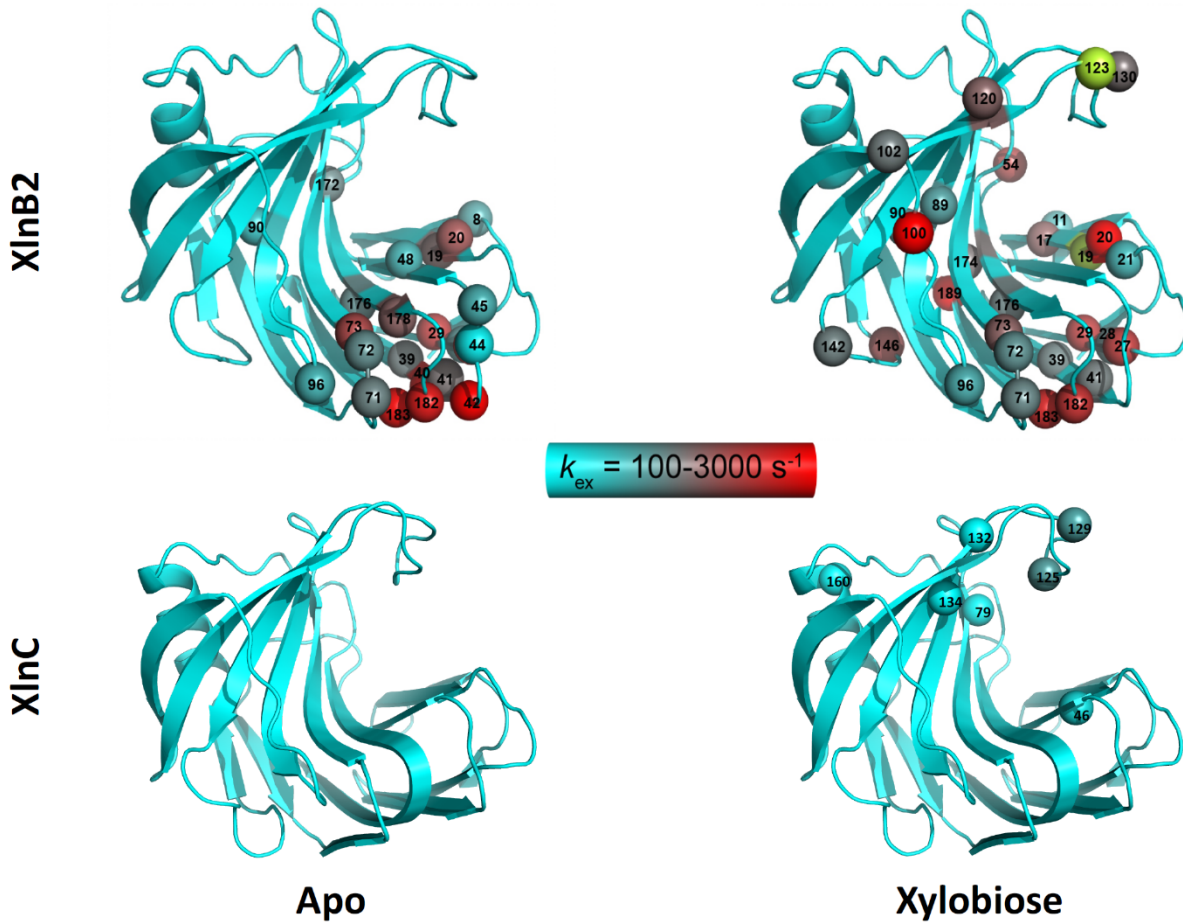


Figure 4.3 : Résidus présentant de l'échange conformationnel chez XlnB2 et XlnC en absence et en présence de xylobiose (X2). Les résidus démontrant une valeur $\Delta R_2 (1/\tau_{\text{cp}}) > 1,8 \text{ s}^{-1}$ sont présentés sous la forme de sphères sur la structure de XlnB2 et XlnC. L'échelle colorée représente le taux d'échange conformationnel (k_{ex}) subi par le vecteur 1H-15N de la chaîne principale, calculé en utilisant l'équation ^{15}N -CPMG⁵³ avec des données acquises à 600 et 800 MHz. Les résidus présentant un $\Delta R_2 (1/\tau_{\text{cp}}) > 1,8 \text{ s}^{-1}$ pour lesquels aucune valeur fiable de k_{ex} n'a pu être extraite sont représentés par des sphères vertes (voir chapitre 3 pour les détails expérimentaux).

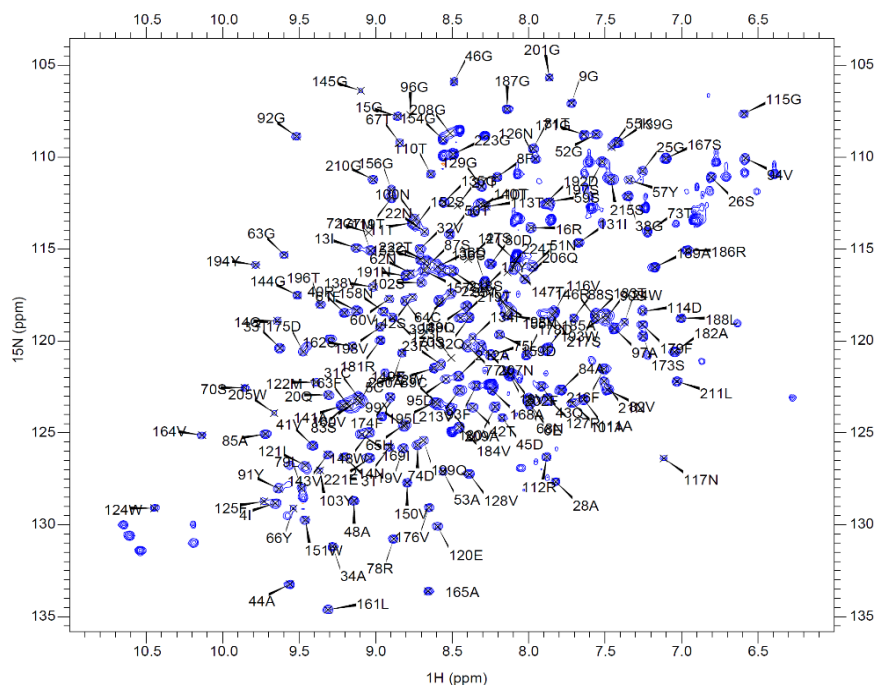
4.2 Cellulase B2 de *S. lividans*

La cellulase B2 (CelB2) correspond au domaine catalytique de de la cellulase B (CelB) de *S. lividans*. Cette enzyme présente peu d'identité de séquence avec XlnB2 (16 %), mais une structure très similaire en « β -jelly roll ». La différence majeure entre les structures des xylanases GH11 et cette cellulase GH12 est la longueur réduite de la boucle du « pouce » chez la cellulase. Les expériences de dispersion de relaxation ^{15}N -CPMG réalisées sur XlnB2 et XlnC laissent présager l'importance du « pouce » et de sa dynamique à l'échelle de temps de la milliseconde dans l'événement catalytique. Il semble donc intéressant d'étudier la dynamique structurale de CelB2 à cette échelle de temps.

4.2.1 Attribution des résonances RMN de la chaîne principale

Le spectre ^1H - ^{15}N HSQC de CelB2 présente des résonances bien dispersées en dimensions proton et azote, offrant une indication que la protéine est correctement repliée (Figure 4.4). En utilisant une combinaison d'expériences RMN 2D ^1H - ^{15}N (HSQC) et 3D ^1H - ^{13}C - ^{15}N (HNC_AC_B et $\text{C}_B\text{C}_A(\text{CO})\text{NH}$), il a été possible d'attribuer les résonances des ^1H , ^{13}C et ^{15}N de la chaîne principale de l'enzyme. Au total 200 résonances ^1H - ^{15}N de CelB2 ont été attribuées, soit 92,6 % des 216 résidus (en excluant les dix-sept prolines et le résidu N-terminal).

Figure 4.4 : Spectre ^1H - ^{15}N HSQC attribué de CelB2 de *S. lividans*. Les résidus attribués sont identifiés en utilisant la numérotation séquentielle des acides aminés de l'enzyme.



4.2.2 Effet de la liaison d'un ligand

Comme pour XlnB2 et XlnC, des expériences de titrage RMN ont permis de repérer les résidus impliqués dans la liaison du substrat au site actif. Des spectres ^1H - ^{15}N HSQC de CelB2 ont été obtenus en utilisant des concentrations croissantes en cellobiose (C2) jusqu'à saturation (un produit non-clivable de cette endocellulase), selon les méthodes expérimentales décrites au chapitre 3. Au total, 15 résidus (Asn22, Arg23, Gly24, Asp45, Gly46, Thr50, Asn51, Ala53, Lys55, Ser56, Tyr66, Ile121, Arg127, Glu149 et Val150) présentent des déplacements chimiques ($\Delta\delta$) $\geq 0,05\text{ppm}$ entre l'enzyme libre et l'enzyme saturée par le ligand. Ces résidus sont retrouvés répartis dans les « doigts » de l'enzyme, principalement autour du site actif (Figure 4.5).

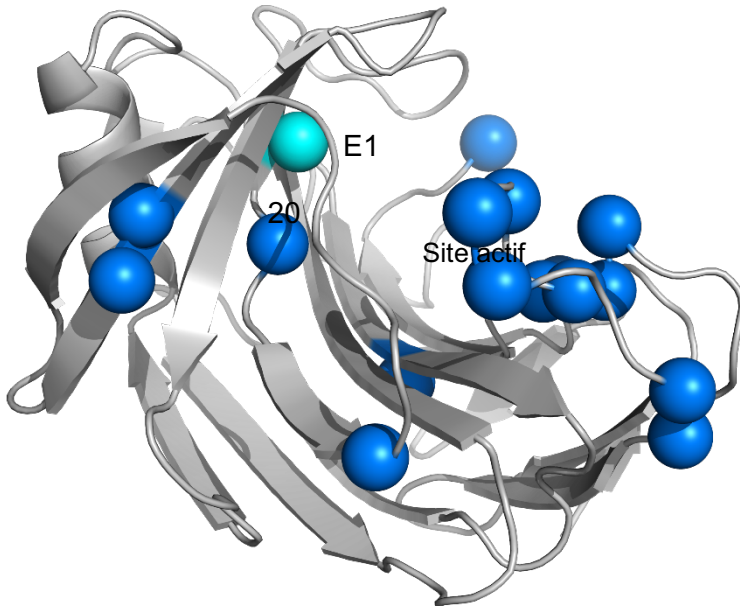


Figure 4.5: Résidus de CelB2 présentant une variation de déplacements chimiques ^1H - ^{15}N ($\Delta\delta$) lors de la liaison au cellobiose (C2).

Les sphères bleues représentent les résidus présentant un déplacement chimique ($\Delta\delta$) $\geq 0,05\text{ ppm}$. Les $\Delta\delta$ ont été calculés entre l'enzyme libre et l'enzyme saturée par le ligand, selon l'équation $\Delta\delta(\text{ppm}) = ((\Delta\delta_{\text{HN}}^2 + \Delta\delta_{\text{N}}^2/25)/2)^{1/2}$ (tel que décrit au chapitre 3). Les résidus catalytiques sont représentés par des sphères cyan.

Uniprot ID: Q54331, RCSB PDB : 1NLR

4.2.3 Changements conformationnels à l'échelle de la milliseconde

Les expériences de dispersion de relaxation ^{15}N -CPMG réalisées sur la cellulase B2 montrent que cette protéine présente peu d'échanges conformationnels à l'échelle de temps de la milliseconde dans sa forme libre, alors que de nombreux résidus des « doigts » et de la « paume » apparaissent dynamiques à cette échelle de temps en présence de cellobiose (C2) (Figure 4.6). Ces régions flexibles en présence de ligand correspondent aux régions impliquées dans la liaison et la reconnaissance du substrat, mises en évidence par les expériences de titrage RMN précédentes. Ces résultats suggèrent que les mouvements à l'échelle de temps de la

milliseconde sont impliqués dans la reconnaissance, la stabilisation et/ou le relâchement du produit chez cette enzyme.

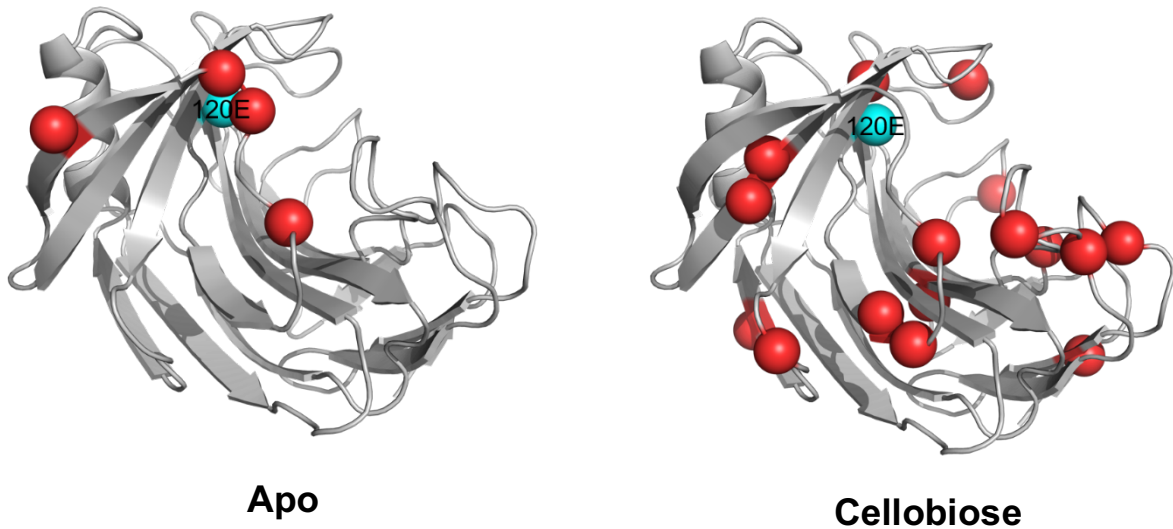


Figure 4.6 : Résidus présentant de l'échange conformationnel chez CelB2, en absence et en présence de cellobiose (C2). Les résidus dynamiques sont présentés sous la forme de sphères rouges et le résidu catalytique sous la forme d'une sphère cyan.

CONCLUSION ET PERSPECTIVES

Dans le présent mémoire, nous avons utilisé un certain nombre de cas de figure pour illustrer que la flexibilité atomique se révèle d'une importance cruciale dans l'activité catalytique de certaines enzymes. En effet, dans des cas particuliers, il a été démontré que les mouvements moléculaires à différentes échelles de temps pouvaient être impliqués dans la régulation de l'activité enzymatique, en jouant notamment un rôle essentiel dans le contrôle de la relation entre la structure et la fonction chez plusieurs systèmes enzymatiques (voir chapitre 2). De surcroît, des études cristallographiques et des simulations de dynamique moléculaire sur plusieurs xylanases GH11 ont d'ailleurs proposé l'existence d'un mouvement « ouvert-fermé », *via* une boucle près du site actif (le « pouce »), dont la dynamique à l'échelle de temps de la catalyse joue potentiellement un rôle majeur dans la liaison au substrat, la libération du produit et/ou la catalyse. Ces études précédentes n'ont toutefois pas offert de détails précis sur la localisation moléculaire de ces mouvements potentiels survenant lors de la réaction catalytique chez ces enzymes, c'est-à-dire sur une échelle de temps de la micro à la milliseconde (μs - ms).

Dans le but de tester ces hypothèses et pour offrir une meilleure compréhension de la potentielle conservation des événements dynamiques impliqués dans la réaction catalytique chez les xylanases GH11, nous avons réalisé un certain nombre d'études de relaxation RMN chez plusieurs homologues de cette famille, notamment la XlnB2, la XlnC et la CelB2 de *S. lividans*. Bien que ces études demeurent préliminaires, nous avons été en mesure de déceler l'existence de certains changements conformationnels conservés chez la boucle du « pouce » chez certains de ces homologues, en plus de permettre l'observation beaucoup plus vaste de résidus subissant des mouvements conservés en absence et en présence de ligands dans les régions des doigts et de la paume du repliement de type « β -jelly roll » qui caractérise ces glycoside hydrolases. Nos observations permettent également d'observer un certain nombre de similitudes et de distinctions moléculaires observées dans la flexibilité évolutive d'un nombre limité de membres bactériens de la grande famille des xylanases et cellulases GH11/12. Nos résultats laissent toutefois entrevoir que cette famille d'enzymes dépend de profils dynamiques distincts, qui pourraient à la fois être tributaires de la présence ou de l'absence de domaines de liaison aux sucres et/ou de différents profils de reconnaissance et d'affinité envers divers substrats. Nos données démontrent ainsi que les repliements protéiques ne peuvent être systématiquement considérés équivalents au niveau de leur potentiel d'adaptation en ingénierie enzymatique, malgré leur forte similarité de séquence, de structure et de fonction. En effet, même si certaines enzymes font partie de la même famille

et qu'elles catalysent des réactions similaires, elles démontrent malgré tout des profils dynamiques distincts en fonction de leur séquence et de la reconnaissance du/des ligand(s).

L'analyse individuelle de chacun de ces systèmes, en plus d'études mutationnelles supplémentaires plus poussées, permettraient d'offrir de plus amples informations quant au(x) réseau(x) dynamique(s) impliqué(s) dans la transmission allostérique de ces mouvements moléculaires à l'échelle atomique. De surcroît, nous considérons qu'une extension de ces études chez un plus grand nombre d'homologues structuraux permettrait probablement d'extraire un profil dynamique plus global, à la manière d'une « cartographie dynamique généralisée » du repliement de type « β -jelly roll » qui caractérise les xylanases et cellulases de la famille GH11/12. Dans le but d'obtenir un profil évolutif plus complet, il serait d'ailleurs intéressant d'approfondir ces résultats en étudiant les xylanases Xyl11 de *B. agaradhaerens*, BCX de *B. circulans*, Xyn11B de *C. japonicus* et XynB de *C. thermocellum*, qui représentent toutes différents niveaux de relations évolutives avec XlnB2 et XlnC de *S. lividans* (Figure 5.1). Dans cette vaste famille, les avantages de ces membres particuliers sont multiples. Tout d'abord, les résonances des résidus des domaines catalytiques de Xyl11 de *B. agaradhaerens* et BCX de *B. circulans* ont déjà été attribués par RMN et sont disponibles (22,56). De plus, toutes les xylanases choisies ne possèdent qu'un seul domaine catalytique de taille inférieure à 25-30 kDa, une taille idéale pour l'analyse RMN. Le Tableau 5.1 recense les principales caractéristiques des enzymes choisies, leurs identifiants, leurs masses moléculaires (MW) totales telles qu'offertes par UniProt, la position des différents domaines qui les composent, leurs points isoélectriques (pI) et leur homologie avec XlnB2 de *S. lividans*. Finalement, dans le but de contrôler et de moduler l'activité catalytique à des fins d'ingénierie enzymatique, il pourrait également être intéressant de construire des xylanase 11 ancestrales sur la base de l'identité de séquence de toutes les xylanases bactériennes connues. Ces études supplémentaires permettraient d'offrir une meilleure compréhension de l'importance de la dynamique moléculaire à l'échelle atomique chez ces biocatalyseurs, en plus d'offrir des outils uniques de contrôle de l'échange conformationnel à des fins de perturbation allostérique d'enzymes spécifiques utilisées dans nombre d'applications industrielles et biomédicales.

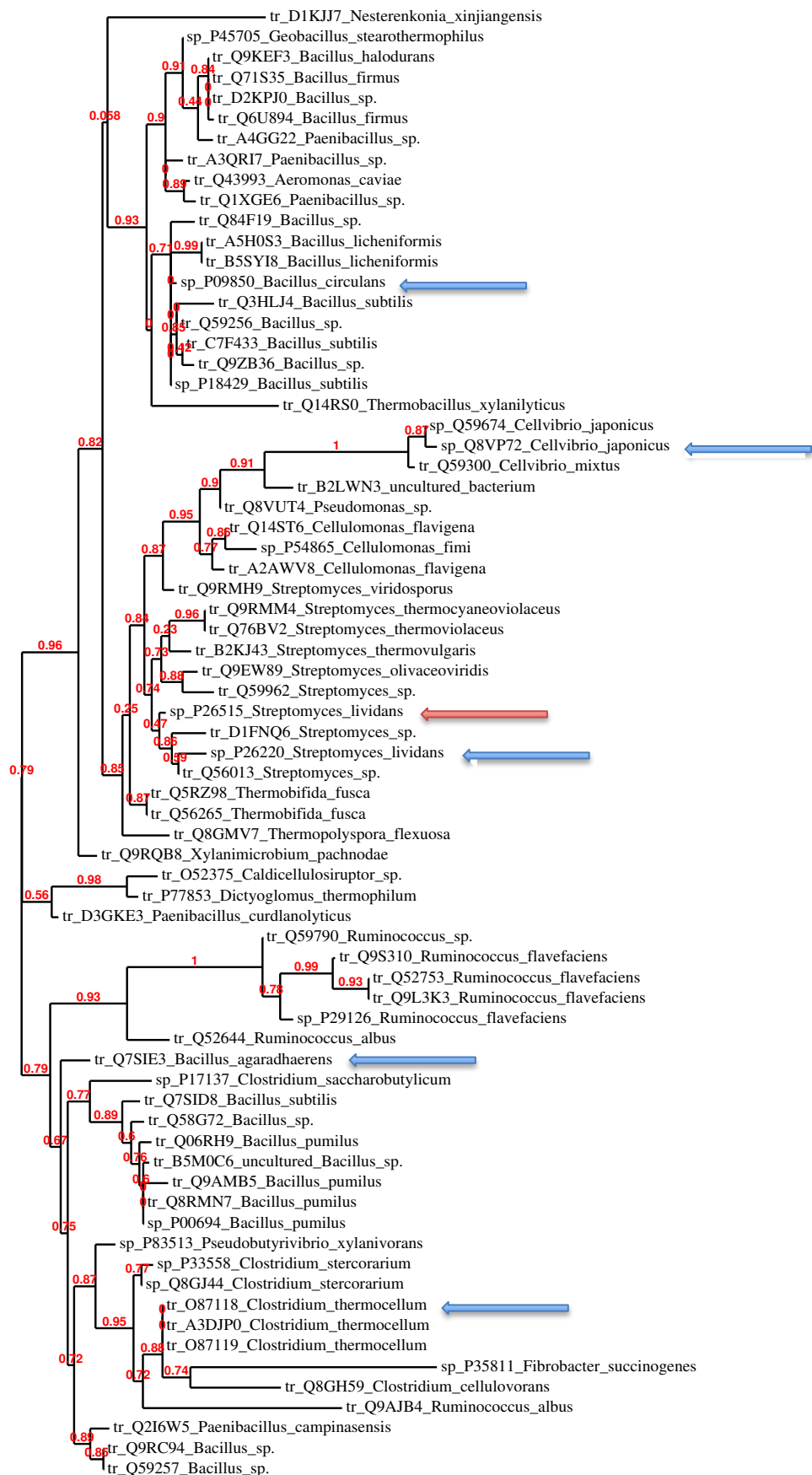


Figure 5.1 : Arbre phylogénétique des GH11 bactériennes. Le code précédant l'espèce correspond à l'identifiant UniProt. La flèche rouge indique XlnB2 et les flèches bleues indiquent les xylanases 11 sélectionnées. D'après la référence (13).

Tableau 5.1 : Principales caractéristiques des enzymes sélectionnées pour de potentielles futures études.

Organisme	Nom ^a	Uniprot ID	Code PDB/BMRB	MW (kDa)	Positions des domaines	pI ^b	Pourcentage identité/similarité ^c	Références
<i>Bacillus agaradhaerens</i> AC13	Xyl11 Rec	Q7SIE3	1QH7/5352	23,152	2-194* : Catalytique	7,98	51 %/66 %	Sabini <i>et al.</i> , 1999 (57) Betz <i>et al.</i> , 2002 (56)
<i>Bacillus circulans</i>	BCX Rec	P09850	1XNB/4704	23,359	29-213 : Catalytique	9,05	63 %/73 %	Wakarchuk <i>et al.</i> , 1994 (58) Connely <i>et al.</i> , 2000 (22)
<i>Streptomyces lividans</i> 66	XlnC Nat	P26220	–	25,673	50-240 : Catalytique	6,75	75 %/82 %	Kluepfel <i>et al.</i> , 1992 (59)
<i>Streptomyces lividans</i> 66	XlnB Nat	P26515	–	35,575	42-230 : Catalytique 231-249 : Linker 250-335 : CBM2	8,7	–	Kluepfel <i>et al.</i> , 1990 (53)
<i>Cellvibrio japonicus</i>	Xyn11B Rec	Q8VP72	–	37,912	30-225* : Catalytique 226-357 : Linker + CBM10	8,78	49 %/68 %	Emami <i>et al.</i> , 2002 (60)
<i>Clostridium thermocellum</i> F1/YS	XynB Rec	O87118	–	49,884	29-222* : Catalytique 223-457 : Linker + CBM6 + Dockerin	6,93	52 %/67 %	Hayashi <i>et al.</i> , 1999 (61)

^aNat ou Rec indiquent que l'enzyme listée correspond à la xylanase 11 native ou recombinante, respectivement.

^bPourcentages d'identité ou de similarité par rapport à XlnB2 de *S.lividans*, obtenus par alignement des séquences BLASTp.

*Position des domaines catalytiques basée sur l'alignement BLASTp contre XlnB2 de *S.lividans*.

**pI calculé par ExPASy ProtParam selon la séquence du domaine catalytique.

BIBLIOGRAPHIE

1. Gallezot, P. (2012) Conversion of biomass to selected chemical products. *Chem. Soc. Rev.* **41**, 1538-1558
2. Varner, J. E., and Lin, L. S. (1989) Plant cell wall architecture. *Cell* **56**, 231-239
3. Davies, G. J., Planas, A., and Rovira, C. (2012) Conformational analyses of the reaction coordinate of glycosidases. *Acc. Chem. Res.* **45**, 308-316
4. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, D490-495
5. Henrissat, B., and Davies, G. (1997) Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol.* **7**, 637-644
6. Davies, G., and Henrissat, B. (1995) Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853-859
7. Davies, G. J., Wilson, K. S., and Henrissat, B. (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem. J.* **321**, 557-559
8. Koshland, D. E. (1953) Stereochemistry and the mechanism of enzymatic reactions. *Biol. Rev.* **28**, 416-436
9. Wong, K. K., Tan, L. U., and Saddler, J. N. (1988) Multiplicity of beta-1,4-xylanase in microorganisms: functions and applications. *Microbiol. Rev.* **52**, 305-317
10. Pollet, A., Delcour, J. A., and Courtin, C. M. (2010) Structural determinants of the substrate specificities of xylanases from different glycoside hydrolase families. *Crit. Rev. Biotechnol.* **30**, 176-191
11. Sapag, A., Wouters, J., Lambert, C., de Ioannes, P., Eyzaguirre, J., and Depiereux, E. (2002) The endoxylanases from family 11: computer analysis of protein sequences reveals important structural and phylogenetic relationships. *J. Biotechnol.* **95**, 109-131

12. Joshi, M. D., Sidhu, G., Pot, I., Brayer, G. D., Withers, S. G., and McIntosh, L. P. (2000) Hydrogen bonding and catalysis: a novel explanation for how a single amino acid substitution can change the pH optimum of a glycosidase. *J. Mol. Biol.* **299**, 255-279
13. Paes, G., Berrin, J. G., and Beaugrand, J. (2012) GH11 xylanases: Structure/function/properties relationships and applications. *Biotechnol. Adv.* **30**, 564-592
14. Biely, P., Vrsanska, M., Tenkanen, M., and Kluepfel, D. (1997) Endo-beta-1,4-xylanase families: differences in catalytic properties. *J. Biotechnol.* **57**, 151-166
15. Katapodis, P., Vrsanska, M., Kekos, D., Nerinckx, W., Biely, P., Claeysens, M., Macris, B. J., and Christakopoulos, P. (2003) Biochemical and catalytic properties of an endoxylanase purified from the culture filtrate of *Sporotrichum thermophile*. *Carbohydr. Res.* **338**, 1881-1890
16. Vardakou, M., Dumon, C., Murray, J. W., Christakopoulos, P., Weiner, D. P., Juge, N., Lewis, R. J., Gilbert, H. J., and Flint, J. E. (2008) Understanding the structural basis for substrate and inhibitor recognition in eukaryotic GH11 xylanases. *J. Mol. Biol.* **375**, 1293-1305
17. Torronen, A., and Rouvinen, J. (1997) Structural and functional properties of low molecular weight endo-1,4-beta-xylanases. *J. Biotechnol.* **57**, 137-149
18. Havukainen, R., Torronen, A., Laitinen, T., and Rouvinen, J. (1996) Covalent binding of three epoxyalkyl xylosides to the active site of endo-1,4-xylanase II from *Trichoderma reesei*. *Biochemistry* **35**, 9617-9624
19. Muilu, J., Torronen, A., Perakyla, M., and Rouvinen, J. (1998) Functional conformational changes of endo-1,4-xylanase II from *Trichoderma reesei*: a molecular dynamics study. *Proteins* **31**, 434-444
20. Murakami, M. T., Arni, R. K., Vieira, D. S., Degreve, L., Ruller, R., and Ward, R. J. (2005) Correlation of temperature induced conformation change with optimum catalytic activity in the recombinant G/11 xylanase A from *Bacillus subtilis* strain 168 (1A1). *FEBS Lett.* **579**, 6505-6510
21. Vieira, D. S., Degreve, L., and Ward, R. J. (2009) Characterization of temperature dependent and substrate-binding cleft movements in *Bacillus circulans* family 11 xylanase: a molecular dynamics investigation. *Biochim. Biophys. Acta* **1790**, 1301-1306

22. Connelly, G. P., Withers, S. G., and McIntosh, L. P. (2000) Analysis of the dynamic properties of *Bacillus circulans* xylanase upon formation of a covalent glycosyl-enzyme intermediate. *Protein Sci.* **9**, 512-524
23. Ludwiczek, M. L., Heller, M., Kantner, T., and McIntosh, L. P. (2007) A secondary xylan-binding site enhances the catalytic activity of a single-domain family 11 glycoside hydrolase. *J. Mol. Biol.* **373**, 337-354
24. Paes, G., Tran, V., Takahashi, M., Boukari, I., and O'Donohue, M. J. (2007) New insights into the role of the thumb-like loop in GH-11 xylanases. *Protein Eng. Des. Sel.* **20**, 15-23
25. Cervera Tison, M., Andre-Leroux, G., Lafond, M., Georis, J., Juge, N., and Berrin, J. G. (2009) Molecular determinants of substrate and inhibitor specificities of the *Penicillium griseofulvum* family 11 xylanases. *Biochim. Biophys. Acta* **1794**, 438-445
26. Pollet, A., Vandermarliere, E., Lammertyn, J., Strelkov, S. V., Delcour, J. A., and Courtin, C. M. (2009) Crystallographic and activity-based evidence for thumb flexibility and its relevance in glycoside hydrolase family 11 xylanases. *Proteins* **77**, 395-403
27. Driss, D., Bhiri, F., Siela, M., Bessess, S., Chaabouni, S., and Ghorbel, G. (2013) Improvement of breadmaking quality by xylanase GH11 from *Penicillium occitanis* Pol6. *J. Texture Stud.* **44**, 75–84
28. Zhao, L., Meng, K., Bai, Y., Shi, P., Huang, H., Luo, H., Wang, Y., Yang, P., Song, W., and Yao, B. (2013) Two family 11 xylanases from *Achaetomium* sp. Xz-8 with high catalytic efficiency and application potentials in the brewing industry. *J. Agric. Food Chem.* **61**, 6880-6889
29. Olfa, E., Mondher, M., Issam, S., Ferid, L., and Nejib, M. M. (2007) Induction, properties and application of xylanase activity from *Sclerotinia sclerotiorum* S2 fungus. *J. Food Biochem.* **31**, 96–107
30. Zhang, L., Xu, J., Lei, L., Jiang, Y., Gao, F., and Zhou, G. H. (2014) Effects of xylanase supplementation on growth performance, nutrient digestibility and non-starch polysaccharide degradation in different sections of the gastrointestinal tract of broilers fed wheat-based diets. *Asian-Australas J. Anim. Sci.* **27**, 855-861

31. Yang, H. J., and Xie, C. Y. (2010) Assessment of fibrolytic activities of 18 commercial enzyme products and their abilities to degrade the cell wall fraction of corn stalks in *in vitro* enzymatic and ruminal batch cultures. *Anim. Feed Sci. Technol.* **159**, 110–121
32. Kulkarni, N., Shendye, A., and Rao, M. (1999) Molecular and biotechnological aspects of xylanases. *FEMS Microbiol. Rev.* **23**, 411-456
33. Beaugrand, J., Paës, G., Reis, D., Takahashi, M., Debeire, P., O'Donohue, M., and Chabbert, B. (2005) Probing the cell wall heterogeneity of micro-dissected wheat caryopsis using both active and inactive forms of a GH11 xylanase. *Planta* **222**, 246-257
34. Dodd, D., and Cann, I. K. (2009) Enzymatic deconstruction of xylan for biofuel production. *Glob. Change Biol. Bioenergy* **1**, 2-17
35. Sandgren, M., Stahlberg, J., and Mitchinson, C. (2005) Structural and biochemical studies of GH family 12 cellulases: improved thermal stability, and ligand complexes. *Prog. Biophys. Mol. Biol.* **89**, 246-291
36. Sulzenbacher, G., Shareck, F., Morosoli, R., Dupont, C., and Davies, G. J. (1997) The *Streptomyces lividans* family 12 endoglucanase: construction of the catalytic cre, expression, and X-ray structure at 1.75 Å resolution. *Biochemistry* **36**, 16032-16039
37. Sulzenbacher, G., Mackenzie, L. F., Wilson, K. S., Withers, S. G., Dupont, C., and Davies, G. J. (1999) The crystal structure of a 2-fluorocellobiosyl complex of the *Streptomyces lividans* endoglucanase CelB2 at 1.2 Å resolution. *Biochemistry* **38**, 4826-4833
38. Koshland, D. E. (1958) Application of a theory of enzyme specificity to protein synthesis. *Proc. Natl. Acad. Sci. USA* **44**, 98-104
39. Hammes, G. G., Chang, Y. C., and Oas, T. G. (2009) Conformational selection or induced fit: a flux description of reaction mechanism. *Proc. Natl. Acad. Sci. USA* **106**, 13737-13741
40. Bhabha, G., Ekiert, D. C., Jennewein, M., Zmasek, C. M., Tuttle, L. M., Kroon, G., Dyson, H. J., Godzik, A., Wilson, I. A., and Wright, P. E. (2013) Divergent evolution of protein conformational dynamics in dihydrofolate reductase. *Nat. Struct. Mol. Biol.* **20**, 1243-1249
41. Narayanan, C., Bernard, D. N., Bafna, K., Gagné, D., Chennubhotla, C. S., Doucet, N., and Agarwal, P. K. (2018) Conservation of dynamics associated with biological function in an enzyme superfamily. *Structure* **26**, 426-436

42. Osuna, S., Jimenez-Oses, G., Noey, E. L., and Houk, K. N. (2015) Molecular dynamics explorations of active site structure in designed and evolved enzymes. *Acc. Chem. Res.* **48**, 1080-1089
43. Fang, J., Nevin, P., Kairys, V., Venclovas, C., Engen, J. R., and Beuning, P. J. (2014) Conformational analysis of processivity clamps in solution demonstrates that tertiary structure does not correlate with protein dynamics. *Structure* **22**, 572-581
44. Gutteridge, A., and Thornton, J. (2004) Conformational change in substrate binding, catalysis and product release: an open and shut case? *FEBS Lett.* **567**, 67-73
45. Bhabha, G., Lee, J., Ekiert, D. C., Gam, J., Wilson, I. A., Dyson, H. J., Benkovic, S. J., and Wright, P. E. (2011) A dynamic knockout reveals that conformational fluctuations influence the chemical step of enzyme catalysis. *Science* **332**, 234-238
46. Wang, Y., Berlow, R. B., and Loria, J. P. (2009) Role of loop-loop interactions in coordinating motions and enzymatic function in triosephosphate isomerase. *Biochemistry* **48**, 4548-4556
47. Gauthier, C., Li, H., and Morosoli, R. (2005) Increase in xylanase production by *Streptomyces lividans* through simultaneous use of the Sec- and Tat-dependent protein export systems. *Appl. Environ. Microbiol.* **71**, 3085-3092
48. Liu, Y. C., and Vidali, L. (2011) Efficient polyethylene glycol (PEG) mediated transformation of the moss *Physcomitrella patens*. *J. Vis. Exp.*, 2560
49. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., 3rd, Rance, M., and Skelton, N. J. (2007) *Protein NMR Spectroscopy (Second Edition) - Principles and Practice*, Academic Press
50. Nucci, N. V., Valentine, K. G., and Wand, A. J. (2014) High-resolution NMR spectroscopy of encapsulated proteins dissolved in low-viscosity fluids. *J. Magn. Reson.* **241**, 137-147
51. Gagné, D., Narayanan, C., Nguyen-Thi, N., Roux, L. D., Bernard, D. N., Brunzelle, J. S., Couture, J. F., Agarwal, P. K., and Doucet, N. (2016) Ligand binding enhances millisecond conformational exchange in xylanase B2 from *Streptomyces lividans*. *Biochemistry* **55**, 4184-4196
52. Palmer, A. G., 3rd. (2015) Enzyme dynamics from NMR spectroscopy. *Acc. Chem. Res.* **48**, 457-465

53. Kluepfel, D., Vats-Mehta, S., Aumont, F., Shareck, F., and Morosoli, R. (1990) Purification and characterization of a new xylanase (xylanase B) produced by *Streptomyces lividans* 66. *Biochem. J.* **267**, 45-50
54. Narayanan, C., Bafna, K., Roux, L. D., Agarwal, P. K., and Doucet, N. (2017) Applications of NMR and computational methodologies to study protein dynamics. *Arch. Biochem. Biophys.* **628**, 71-80
55. Palmer, A. G., 3rd. (2004) NMR characterization of the dynamics of biomacromolecules. *Chem. Rev.* **104**, 3623-3640
56. Betz, M., Löhr, F., Wienk, H., and Rüterjans, H. (2002) ¹H, ¹³C and ¹⁵N chemical shift assignment of *Bacillus agaradhaerens* family 11 xylanase. *J. Biomol. NMR* **23**, 333-334
57. Sabini, E., Sulzenbacher, G., Dauter, M., Dauter, Z., Jorgensen, P. L., Schulein, M., Dupont, C., Davies, G. J., and Wilson, K. S. (1999) Catalysis and specificity in enzymatic glycoside hydrolysis: a 2,5B conformation for the glycosyl-enzyme intermediate revealed by the structure of the *Bacillus agaradhaerens* family 11 xylanase. *Chem. Biol.* **6**, 483-492
58. Wakarchuk, W. W., Campbell, R. L., Sung, W. L., Davoodi, J., and Yaguchi, M. (1994) Mutational and crystallographic analyses of the active site residues of the *Bacillus circulans* xylanase. *Protein Sci.* **3**, 467-475
59. Kluepfel, D., Daigneault, N., Morosoli, R., and Shareck, F. (1992) Purification and characterization of a new xylanase (xylanase C) produced by *Streptomyces lividans* 66. *Appl. Environ. Microbiol.* **36**, 626-631
60. Emami, K., Nagy, T., Fontes, C. M., Ferreira, L. M., and Gilbert, H. J. (2002) Evidence for temporal regulation of the two *Pseudomonas cellulosa* xylanases belonging to glycoside hydrolase family 11. *J. Bacteriol.* **184**, 4124-4133
61. Hayashi, H., Takehara, M., Hattori, T., Kimura, T., Karita, S., Sakka, K., and Ohmiya, K. (1999) Nucleotide sequences of two contiguous and highly homologous xylanase genes xynA and xynB and characterization of XynA from *Clostridium thermocellum*. *Appl. Microbiol. Biotechnol.* **51**, 348-357