

### Les sites de fixation des différents complexes

La protéine X fait partie des protéines qui lient la choline (choline binding proteins - ChBP). Le motif fixant la choline (choline binding motive : cbm) est composé d'une séquence particulière d'approximativement 20 acides aminés (dont au moins deux tryptophanes) (Garcia *et al.*, 1988). Ce motif est intégré de manière répétée dans la structure de la protéine, les sites cbm étant tous localisés dans la partie C-terminale de la protéine (à droite dans la figure 2.4.28). Les sites de fixation 1 à 5 correspondent à des sites cbm. Ils lient de préférence les complexes Gd-HPDO3A et Gd-DO3A et aussi le complexe Gd-DOTA-BOM. Le taux d'occupation des différents sites est différent pour les trois complexes (voir tableau 2.4.9).

Les sites 6 et 7 ne correspondent pas à des sites cbm. C'est sur ces deux sites que se lient les complexes Gd-DTPA et Gd-DOTMA. Chacun des deux complexes n'a qu'un seul site de fixation avec cette protéine. On remarque qu'il n'y a pas de site de fixation des complexes sur la partie N-terminale de la protéine, dont la surface est chargée très négativement.

Dans la discussion qui suit, les numéros de sites correspondent à la dénomination définie plus haut (figure 2.4.28) et non pas à la valeur du taux d'occupation.

### Mode de fixation des différents complexes

#### Gd-DTPA (charge : -2)

##### Site 6 (taux d'occupation : 0,99)

Le site est localisé vers le côté C-terminal de la protéine, correspondant à une partie de surface globalement chargée positivement (figure 2.4.29.a). On remarque la coïncidence d'une zone de surface de charge positive avec la fixation du seul complexe de charge -2. La fixation se fait dans une orientation où tous les groupements carboxyles sont tournés vers la surface. Le complexe se lie avec son côté ionique tourné vers la surface. Son orientation permet l'interaction de l'ion  $Gd^{3+}$  et de la plupart des groupements carboxyliques du ligand avec les résidus de la surface protéique. La chaîne du complexe est tournée vers le solvant. L'ion  $Gd^{3+}$  est coordonné par l'atome d'oxygène OD de l'Asn266 (Gd-O : 2,4 Å) qui agit comme neuvième ligand de l'ion. Les interactions directes entre les atomes d'oxygène du ligand et des résidus protéiques sont indiquées dans la figure 2.4.29.b. Deux atomes d'oxygène d'un même bras du ligand forment des liaisons H avec les atomes d'azote amide de la chaîne principale (O(loin Gd)-N(Asn269) : 2,8 Å, O(proche Gd)-N(Gly268) : 2,9 Å). Les atomes d'oxygène de deux autres bras forment une liaison H avec l'atome d'oxygène OH de la Tyr264 (O(ligand)-OH : 2,7 Å) et un pont salin avec l'atome d'azote NZ de la Lys273 (O(ligand)-NZ : 2,6 Å). De plus, un réseau de molécules d'eau ordonnées entre le ligand et la surface renforce la fixation (figure 2.4.30.a).

Cinq molécules d'eau sont directement liées aux atomes d'oxygène du ligand et en particulier aux atomes d'oxygène des deux bras du ligand qui n'interagissent pas directement avec des résidus protéiques. La densité électronique correspondant au ligand dans ce dérivé est si bien définie qu'il a été possible de construire le modèle du complexe à partir de la densité 1Fo-Fc, calculée sans le modèle du complexe, et ainsi de résoudre la structure du complexe jusque-là inconnue.

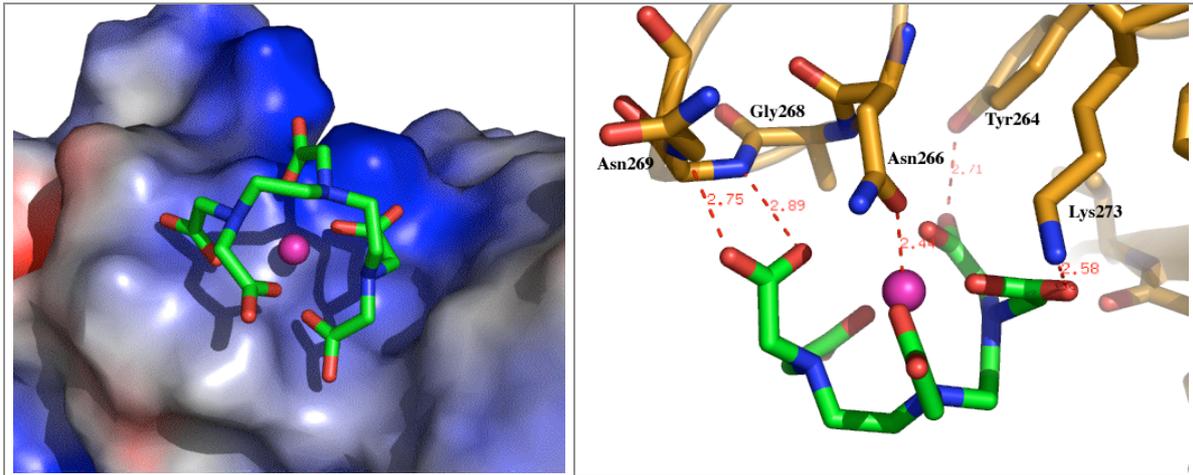


Figure 2.4.29 : Site de fixation 6 du dérivé protéine X + Gd-DTPA. a) Représentation en mode surface électrostatique. b) Interactions directes entre le complexe et la surface protéique.

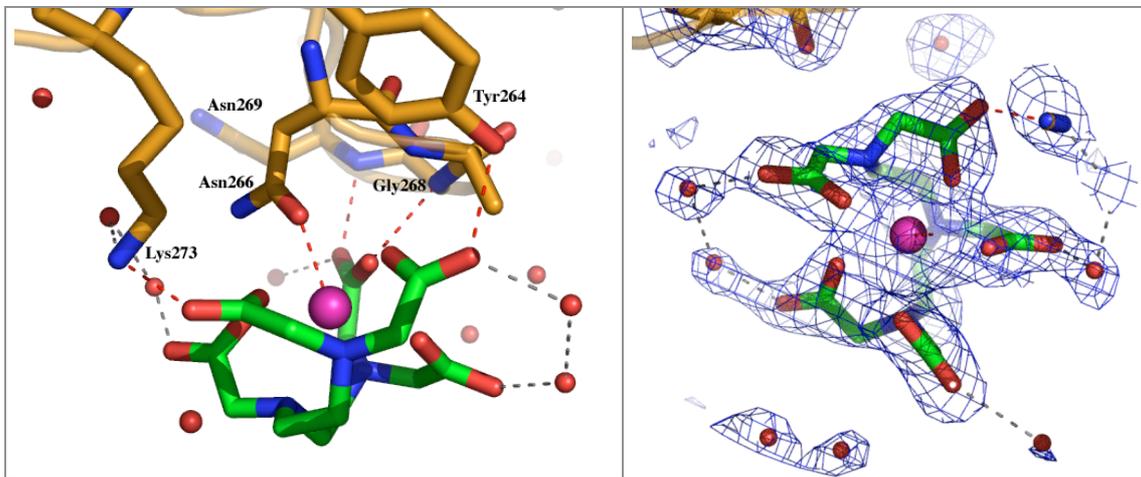


Figure 2.4.30 : Site de fixation 6 du dérivé protéine X + Gd-DTPA. a) Les molécules d'eau coordonnées par le ligand du complexe. b) La densité 2Fo-Fc du modèle affiné du complexe et des molécules d'eau est très bien définie.

### Gd-DOTMA (charge : -1)

#### Site 7 (taux d'occupation : 0,65)

Le complexe se lie par un côté formé par deux bras du ligand à une partie de surface légèrement concave et chargée positivement (figure 2.4.31.a). Sa fixation semble reposer sur le réseau de molécules d'eau ordonnées entre la surface protéique et le ligand, incluant quatre molécules d'eau coordonnées au ligand à des distances comprises entre 2 et 3 Å. Il n'y a pas de densité claire qui correspondrait à un atome agissant de neuvième ligand pour l'ion  $Gd^{3+}$ . Bien que la densité du modèle affiné ne soit pas bien définie pour la partie du ligand tournée vers le solvant, elle permet d'identifier sans ambiguïté l'orientation du complexe.

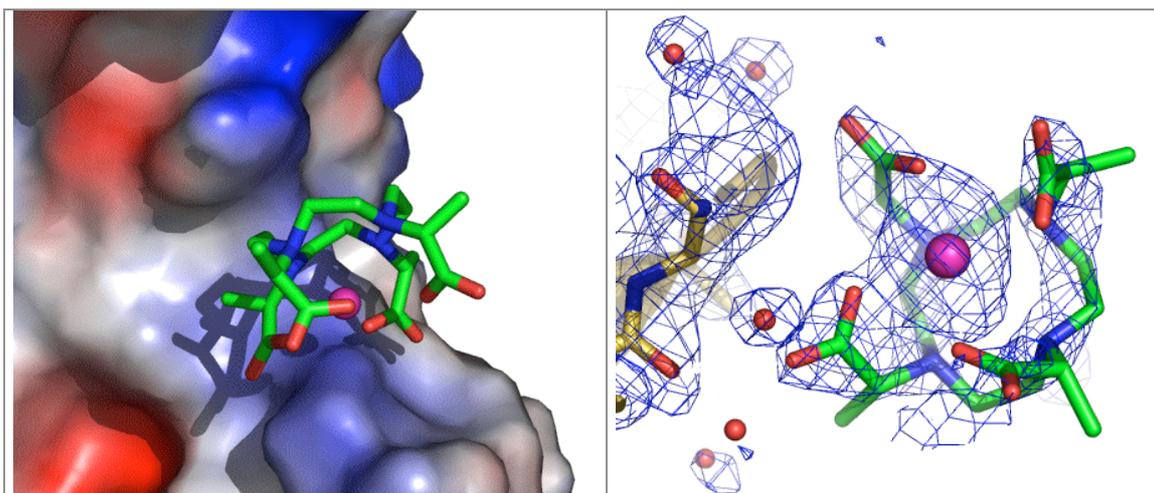


Figure 2.4.31 : Site de fixation 7 du dérivé protéine X + Gd-DOTMA. a) Représentation en mode surface électrostatique. b) La densité 2Fo-Fc du modèle affiné.

### **Gd-HPDO3A** (charge : 0)

Tous les sites de fixation du complexe correspondent à des cavités cbm. La charge de la surface des cavités est légèrement négative. Les cavités cbm sont formées par au moins deux tryptophanes dont les cycles aromatiques sont approximativement perpendiculaires. La fixation du complexe Gd-HPDO3A se fait par interaction hydrophobe entre le macrocycle du complexe et le cycle aromatique du tryptophane situé au fond de la cavité, les plans des deux cycles étant parallèles (distance entre plans parallèles  $\sim 3,7$  Å). Le cycle du deuxième tryptophane de la cavité est parallèle à un côté du complexe formé par deux bras du ligand. Pour les sites 1, 2 et 3, la densité électronique montre la présence d'une molécule d'eau liée à l'ion  $Gd^{3+}$ , agissant comme neuvième ligand de l'ion. Pour les sites 1 et 2, cette même molécule d'eau pont, par liaison H, l'atome d'azote ND2 d'une asparagine du bord de la cavité de fixation.

Par la suite, je décris d'éventuelles interactions supplémentaires propres aux différents sites.

### **Site 1** (taux d'occupation : 0,87) (figure 2.4.32)

La liaison du complexe implique une molécule de protéine symétrique, un atome d'oxygène du ligand formant un pont salin avec l'atome d'azote NZ de la Lys196, résidu qui appartient à la molécule symétrique (O-NZ : 2,7 Å). Un atome d'oxygène du ligand interagit avec l'atome d'azote de la même asparagine, qui est pontée par la molécule d'eau liée au  $Gd^{3+}$  (O(ligand)-OD : 2,7 Å). Un réseau de molécules d'eau entoure le ligand, dont cinq molécules directement liées au ligand. La densité 2Fo-Fc du modèle affiné est bien définie (figure 2.4.32.c).

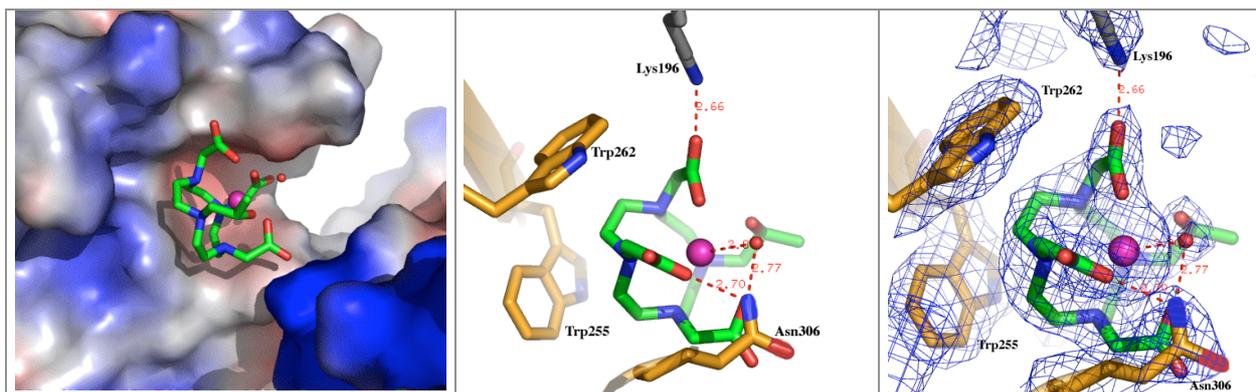


Figure 2.4.32 : Site de fixation 1 du dérivé protéine X + Gd-HPDO3A. a) Représentation en mode surface électrostatique de la cavité de fixation. En gris, la lysine 196 appartenant à une molécule de protéine symétrique. b) Interactions entre le complexe et la protéine et une molécule d'eau qui sert de neuvième ligand à l'ion  $Gd^{3+}$ . c) La densité 2Fo-Fc du modèle est bien définie.

### Site 2 (taux d'occupation : 0,66) (figure 2.4.33.a)

La densité électronique, non montrée ici, est relativement bien définie et permet d'identifier l'orientation du complexe. De manière analogue au site 1, l'asparagine impliquée dans la coordination de la molécule d'eau liée au gadolinium interagit également avec un atome d'oxygène du ligand. Du réseau de molécules d'eau entourant le complexe cinq molécules sont directement liées au ligand.

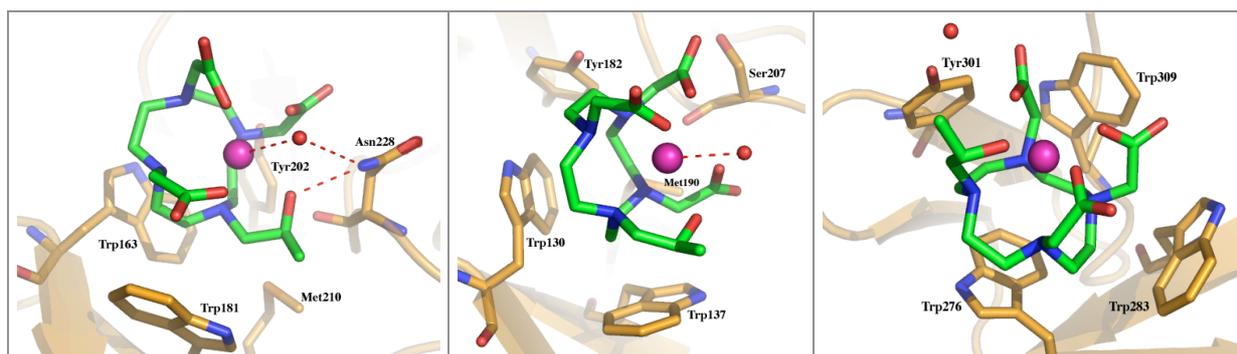


Figure 2.4.33 : Sites de fixation du dérivé protéine X + Gd-HPDO3A : a) site 2, b) site 3, c) site 4.

### Site 3 (taux d'occupation : 0,60) (figure 2.4.33.b)

Un réseau de molécules d'eau entourant le complexe inclut 3 molécules directement liées au ligand.

### Site 4 (taux d'occupation : 0,72) (figure 2.4.33.c)

Le site 4 possède trois tryptophanes perpendiculaires. Le cycle d'un des tryptophanes est parallèle au macrocycle du ligand et les cycles des deux autres sont parallèles à deux côtés du complexe respectivement formés par deux bras du ligand. Un autre côté de la cavité est formé par une tyrosine, perpendiculaire à deux des tryptophanes. L'ion  $Gd^{3+}$  pointe vers le solvant. Une molécule d'eau est directement liée au ligand.

### Gd-DO3A (charge : 0)

Tous les sites de fixation du complexe sont des cavités cbm. Dans tous les sites de fixation, le complexe Gd-DO3A est lié par interaction hydrophobe, son macrocycle étant parallèle au cycle aromatique d'un des tryptophanes formant la cavité, à une distance légèrement supérieure à 3 Å. Pour deux des complexes affinés (site 5 et site 2), le cycle du deuxième tryptophane de la cavité est approximativement parallèle au côté du complexe formé par deux

bras de ligand. Pour ces deux mêmes sites, l'ion  $Gd^{3+}$  est coordonné par l'atome d'oxygène OD de l'asparagine du bord de la cavité, qui fournit ainsi un huitième ligand à l'ion. Cette coordination engendre un mouvement du résidu vers l'ion (voir Figure 2.4.35.b). La densité électronique ne permet pas d'identifier un neuvième ligand éventuel de l'ion. Pour aucun des sites, la fixation n'implique la formation de liaisons hydrogène ou ioniques entre le ligand et des résidus protéiques. Des caractéristiques supplémentaires des différents sites de fixation, sont données ci-dessous

#### Site 5 (taux d'occupation : 0,83)

Le site le plus occupé par le complexe Gd-DO3A correspond à un site faiblement occupé (13%) du complexe Gd-HPDO3A. L'orientation du complexe Gd-HPDO3A n'a pas été affinée, mais la densité 1Fo-Fc indique que le complexe se lie dans une orientation perpendiculaire au complexe Gd-DO3A, alors que pour les autres sites leur fixation se fait dans la même orientation. Le taux de fixation plus élevé du complexe Gd-DO3A pourrait être dû à la taille plus petite de cette cavité. Quatre molécules d'eau sont liées au ligand, dont une qui ponte un atome d'oxygène amide de la chaîne principale.

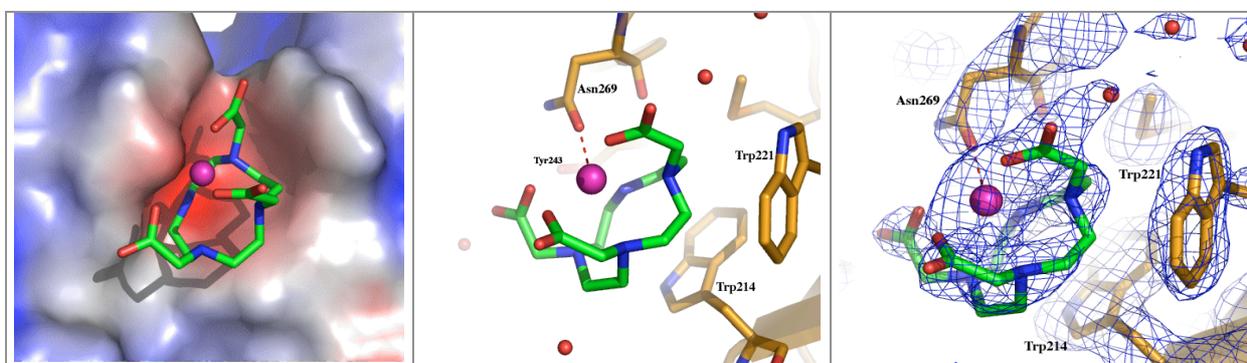


Figure 2.4.34 : Site de fixation 5 du dérivé protéine X + Gd-DO3A. a) Représentation en mode surface électrostatique. b) Coordination de l'ion  $Gd^{3+}$  par l'Asp269. c) Densité 2Fo-Fc du modèle affiné.

#### Site 2 (taux d'occupation : 0,66) (figure 2.4.35.a, b et c)

La densité n'indique aucune interaction par liaison H du ligand ni avec des résidus protéiques ni avec des molécules d'eau.

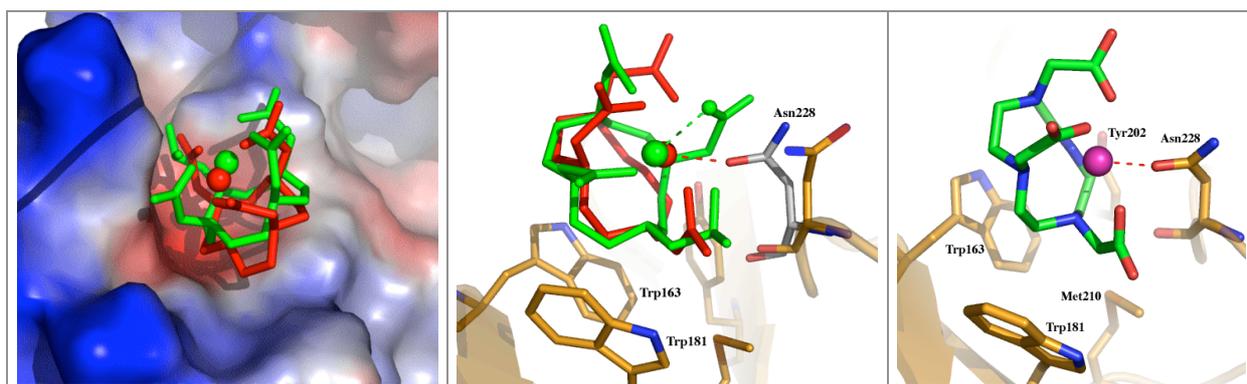


Figure 2.4.35 : Site de fixation 2 du dérivé protéine X + Gd-DO3A. a) et b) En rouge : le modèle de Gd-DO3A, en vert le modèle de Gd-HPDO3A. a) Représentation en mode surface de la cavité de fixation (la surface est calculée à partir de la conformation des résidus dans le dérivé du complexe Gd-HPDO3A). b) En gris : la conformation du résidu Asn228 dans le dérivé de Gd-DO3A, en or : sa conformation dans le dérivé de Gd-HPDO3A. Pour la fixation du complexe Gd-DO3A, le mouvement du résidu Asn228 vers l'ion  $Gd^{3+}$  permet la coordination directe de l'ion  $Gd^{3+}$ . c) Fixation du complexe Gd-DO3A.

**Site 4** (taux d'occupation : 0,63)

Cette cavité est formée par trois tryptophanes et une tyrosine dont les cycles sont perpendiculaires. Les deux tryptophanes qui ne sont pas parallèles au macrocycle du ligand se trouvent chacun face à un bras du ligand. L'ion  $Gd^{3+}$  pointe vers le solvant. La densité de ce site n'est pas très bien définie et ne permet pas de détecter d'éventuels huitième et neuvième ligand de l'ion  $Gd^{3+}$ .

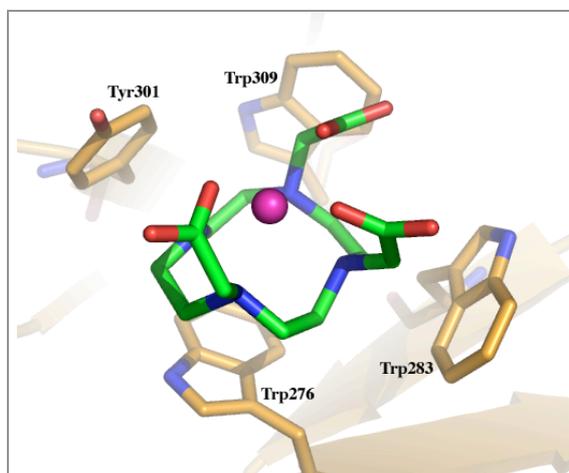


Figure 2.4.36 : Site de fixation 4 du dérivé protéine X + Gd-DO3A.

Les trois sites de fixation mineurs du complexe (taux d'occupation ~30%) se trouvent également dans des cavités cbm.

**Gd-DOTA-BOM** (charge -1)

**Site 4** (taux d'occupation : 0,95)

Le complexe se lie avec un taux d'occupation très élevé dans la cavité cbm du site 4. Cette cavité, relativement large, est formée par trois tryptophanes et une tyrosine perpendiculaires. La densité électronique expérimentale et 1Fo-Fc autour de l'ion  $Gd^{3+}$  est très forte, mais elle ne semble pas correspondre à une orientation ordonnée du complexe (figure 2.4.37.a) et ne permet pas de modéliser le complexe fixé.

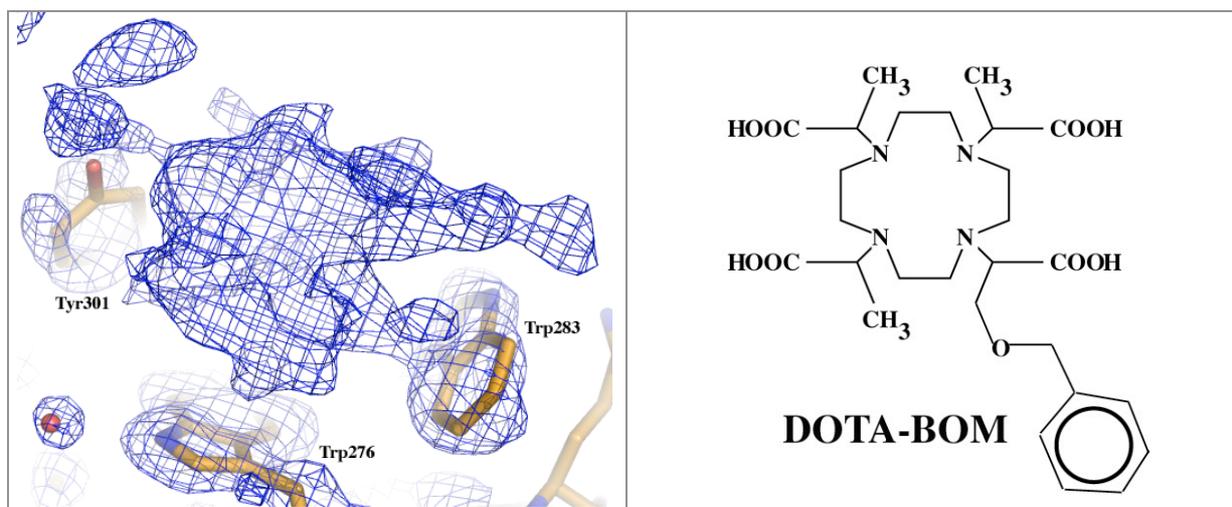


Figure 2.4.37 : a) Densité électronique expérimentale au site 4 du dérivé protéine X + Gd-DOTA-BOM.  
b) Formule du ligand DOTA-BOM.

### **Gd-DTPA-BMA (charge 0) :**

Le complexe se lie avec des taux d'occupation de l'ordre de 30% dans les cavités cbm du site 2 et du site 4.

## **II.4.11. Résultats sur la fixation des complexes**

L'affinement des structures dérivées montre que ces dernières sont assez isomorphes avec la structure native. Généralement, la fixation des complexes engendre peu de changement de la structure protéique, des paramètres de maille et de la forme cristalline. Pour certains dérivés la liaison du complexe engendre des mouvements de chaîne latérale de résidus en interaction directe avec le complexe. La fixation des complexes sur les protéines est non-covalente et de faible affinité. Généralement, elle résulte de multiples interactions faibles, impliquant le plus souvent un réseau de molécules d'eau coordonnées entre le complexe et la surface protéique et elle dépend de la forme de la surface protéique.

La fixation des complexes peut combiner différents types d'interaction dont les suivants :

**1)** L'interaction directe entre l'ion  $Gd^{3+}$  et des atomes chargés négativement des résidus de la surface protéique. Pour des complexes où l'ion  $Gd^{3+}$  est entouré par 8 atomes de coordination du ligand, l'ion interagit avec un seul atome de la surface protéique, par exemple avec l'atome d'oxygène de la chaîne latérale des résidus Asn, Gln, Asp ou Glu. Pour le Gd-DO3A fournissant 7 atomes de coordination, l'ion pourra interagir avec deux atomes chargés négativement (les deux atomes d'oxygène carboxyliques de l'Asp ou du Glu).

**1b)** L'ion  $Gd^{3+}$  coordonne une molécule d'eau. Cette même molécule d'eau pontre l'ion  $Gd^{3+}$  à un ou plusieurs résidus protéiques.

**2)** Des ponts salins ou des liaisons H directes entre les atomes d'oxygène des bras des ligands (les atomes d'oxygène et d'azote pour Gd-DTPA-BMA) et des atomes d'oxygène ou d'azote de la surface protéique.

**3)** L'interaction hydrophobe entre le macrocycle (ou "pseudo-macrocycle") et des résidus aromatiques de la surface protéique.

**4)** L'établissement d'un réseau de liaisons H à travers la coordination de molécules d'eau par les atomes d'oxygène (et d'azote) des bras des ligands. Ces molécules d'eau établissent éventuellement des ponts vers la surface protéique.

**5)** L'interaction entre deux complexes menant à la fixation coopérative de deux molécules de complexe voisines (lysozyme).

Le tableau 2.4.9 résume quelle sorte d'interaction décrite ci-dessus intervient lors de la fixation des complexes pour les différents dérivés.

Gd-HPDO3A (0)		Gd-DO3A (0)				
lysozyme	protéine X	lysozyme	protéine X	YGGV	yeaZ	GI
3 4 5	3 1b 2 4	1 3 4 5	1 3 4	1	1 2	1 2 4
4 5	3 1b 2 4	1 3 4 5	1 3			
	3 4		3			
	3 4					

Gd-DOTMA (-1)				Gd-DTPA-BMA (0)	
protéine X	urate	Yggv	yeaZ	urate	thaumatine
4	3 2 4	2 4	1 4	1 2 4	3
				1 2 4	
				1b 2 4	

Gd-DTPA (-2)			Gd-DOTA (-1)		Gd-DOTA-BOM (-1)	
protéine X	GI	thaumatine	lysozyme	thaumatine	lysozyme	thaumatine
1 2 4	1b 2 4	3 2 4	3 1 2 5	3 2	3 2 4	3
			3 5			2 4

Tableau 2.4.9 : Interactions intervenant pour la fixation des complexes avec les différentes protéines (entre parenthèses la charge totale du complexe). Plusieurs lignes correspondent à différents sites de fixation.  
Abréviations des noms de protéine : GI : glucose isomérase, urate : urate oxydase.

Les complexes se lient souvent dans des cavités plus ou moins prononcées, formées par la surface de la protéine (thaumatine) ou par l'empilement cristallin (lysozyme), ce qui indique l'importance de facteurs stériques pour la fixation. La représentation des surfaces électrostatiques permet de constater que les complexes chargés négativement ont plus tendance à se lier sur des parties de surface chargées positivement que les complexes neutres. La représentation de la surface montre que le complexe se lie de préférence selon une orientation qui accorde sa forme à la forme de la surface, typiquement avec les bras carboxyliques ou méthylamides qui se nichent dans des creux de la surface. On remarque que deux complexes de structure très proche peuvent se lier selon des modes de fixation tout à fait différents (exemples : Gd-DOTA-BOM - Gd-DOTMA dans les protéines lysozyme, urate oxydase et protéine X, Gd-HPDO3A - Gd-DOTA dans le lysozyme et la thaumatine). La possibilité de liaison par interaction hydrophobe entre un complexe particulier et le cycle aromatique de tryptophanes ou de phénylalanines semble dépendre, outre de la nature du résidu aromatique, de l'environnement du résidu aromatique ou peut-être des conditions physico-chimiques. Ainsi dans le lysozyme les complexes Gd-HPDO3A et Gd-DO3A interagissent avec le(s) tryptophane(s), mais non pas le Gd-DOTMA, alors que c'est l'inverse pour l'urate oxydase. Les complexes qui interagissent avec la phénylalanine dans les dérivés de thaumatine sont d'autres complexes, dont les deux complexes avec un ligand linéaire qui, eux, ne semblent pas interagir avec les tryptophanes.

#### II.4.11.a. Quelques caractéristiques récurrentes de la fixation des différents complexes

##### Gd-DO3A

Le mode de fixation principal du complexe correspond à la coordination de l'ion  $Gd^{3+}$  par les deux atomes d'oxygène carboxyliques d'un aspartate ou glutamate de la surface protéique. Un mode de fixation alternatif est l'interaction hydrophobe du macrocycle avec le cycle aromatique de tryptophanes.

## **Gd-HPDO3A et Gd-DOTA-BOM**

Les deux complexes ont tendance à établir des liaisons hydrophobes avec les tryptophanes.

Les complexes Gd-DTPA-BMA, Gd-DTPA, Gd-DOTA, Gd-DOTMA et Gd-DOTA-BOM interagissent de différentes manières avec les protéines.

Le mode d'interaction peut, en partie, expliquer l'influence de la présence du complexe sur la solubilité de la protéine. L'interaction hydrophobe pourrait solubiliser la protéine, car, dans ces conditions, les complexes couvrent des parties hydrophobes de la protéine. Inversement, la fixation d'un complexe à des résidus ioniques de la surface par son côté ionique et tournant son côté apolaire vers le solvant pourrait diminuer la solubilité de la protéine.

## **Conclusions**

L'utilisation du complexe Gd-DO3A mène souvent à des dérivés avec des sites de fixation multiples et des taux de fixation moyens, mais suffisants pour un bon phasage. L'utilisation du complexe Gd-HPDO3A est indiquée pour des protéines avec un taux élevé de tryptophanes dans la séquence, pourvu que ces derniers soient supposés situés en surface.

La forte fixation des complexes Gd-HPDO3A et Gd-DO3A sur le motif structural cbm en fait des candidats de premier choix pour préparer des dérivés avec des protéines qui lient la choline (ChBP). Il serait envisageable d'introduire des sites cbm dans d'autres protéines par mutagenèse dirigée. Pourvu qu'une telle introduction n'agisse trop fortement sur le repliement, la solubilité et la cristallisation de la protéine les sites cbm pourraient alors agir comme sites de fixation des complexes.

## **II.5. Conclusions**

L'étude cristallographique de la fixation des complexes dans les cristaux de protéines a été menée avec un grand nombre de dérivés. Pour toutes les protéines testées, il a été possible d'obtenir des cristaux dérivés avec un fort pouvoir de phasage avec au moins un des complexes testés.

Pendant ma thèse, l'utilisation des complexes a par ailleurs permis la résolution de la structure de quatre nouvelles protéines.

Les résultats de l'étude confirment le potentiel des complexes pour se fixer fortement dans les cristaux et permettre ainsi l'obtention de dérivés à fort pouvoir de phasage. Ils confirment la facilité d'utilisation des complexes et leur applicabilité générale dans de nombreuses conditions testées.

Pour une combinaison de protéine et de complexe quelconque, le taux de succès pour obtenir un dérivé performant est de l'ordre de 50%.

Pour préparer un dérivé anomal avec une nouvelle protéine, en choisissant le complexe à utiliser de manière ciblée, tenant compte des caractéristiques de la protéine, et éventuellement en testant deux ou trois complexes différents, il est très probable d'obtenir un bon dérivé.

Les résultats des affinements du mode de fixation des différents complexes permettent de mieux orienter le choix du complexe à utiliser en fonction des propriétés de la protéine. Néanmoins, sauf pour des cas exceptionnels, comme celui des protéines ChBP, il n'est pas possible de prévoir avec certitude la fixation d'un complexe sur une protéine. Cela est dû au fait que la fixation des complexes résulte généralement de nombreuses interactions faibles, impliquant, entre autres, le réseau de molécules d'eau ordonnées à la surface de la protéine en combinaison avec des conditions stériques favorables à la liaison du complexe.

Il serait donc d'un grand intérêt de trouver un méthode relativement simple et facile à mettre en œuvre comme alternative à la cristallographie pour détecter la fixation ou non d'un complexe avec une protéine.