



**HAL**  
open science

# Cartographie structurale et fonctionnelle de la liaison entre la peptidyl-ARNt hydrolase et son substrat

Giorgi Laurent

► **To cite this version:**

Giorgi Laurent. Cartographie structurale et fonctionnelle de la liaison entre la peptidyl-ARNt hydrolase et son substrat. Microbiologie et Parasitologie. Ecole Polytechnique X, 2010. Français. NNT : . pastel-00572217

**HAL Id: pastel-00572217**

**<https://pastel.hal.science/pastel-00572217>**

Submitted on 1 Mar 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# THÈSE DE DOCTORAT DE L'ÉCOLE POLYTECHNIQUE

Présentée par

**Laurent GIORGI**

Pour obtenir le titre de docteur de l'école polytechnique

Discipline : Biologie

## **Cartographie structurale et fonctionnelle de la liaison entre la peptidyl-ARNt hydrolase et son substrat**

Soutenue le 27 novembre 2010 devant le jury composé de

Pr. Sylvain Blanquet	président du jury
Dr. François Bontems	examineur
Dr. Christophe Escudé	examineur
Dr. Marc Mirande	rapporteur
Dr. Pierre Plateau	directeur de thèse
Dr. Carine Tisé	rapporteur



« Selon une théorie, le jour où quelqu'un découvrira exactement à quoi sert l'univers et pourquoi il est là, ledit univers disparaîtra sur-le-champ pour être remplacé par quelque chose d'infiniment plus complexe et inexplicable. »

Douglas Noël Adams



La peptidyl-ARNt hydrolase est une enzyme qui hydrolyse les peptidyl-ARNt issus d'une terminaison prématurée de la traduction. Cette protéine est essentielle à la viabilité des bactéries, mais pas à celle des eucaryotes, ce qui fait d'elle une cible potentielle pour l'action d'anti-bactériens. Cela justifie également qu'on cherche à cartographier l'interaction de cette protéine avec son substrat, pour faciliter la conception d'inhibiteur.

Les tentatives d'obtention de cristaux de complexes enzyme:analogue de substrat étant restées vaines, nous avons choisi d'étudier de tels complexes en solution, par RMN. Grâce à un double marquage  $^{15}\text{N}/^{13}\text{C}$ , nous avons tout d'abord attribué les fréquences de résonance des atomes du squelette de la protéine et d'une grande partie des chaînes latérale. Nous avons ensuite étudié l'interaction entre la PTH d'*E. coli* et un analogue de son substrat synthétisé chimiquement, la diacétyl-Lys-(3'NH)-adénosine. Cette étude nous a permis de caractériser le rôle de nombreux résidus du site actif, notamment celui d'une phénylalanine (F66) interagissant via son cycle aromatique avec l'adénine 3'-terminale du substrat, celui d'une asparagine (N114) stabilisant une molécule d'eau responsable de l'hydrolyse du substrat et celui d'une autre asparagine (N10) permettant à la PTH de discriminer positivement les peptidyl-ARNt par rapport aux aminoacyl-ARNt.

Nous avons aussi étudié l'interaction entre la protéine et des mini-ARNt mimant la tige acceptrice et le bras T $\Psi$ C d'un ARNt. Ce travail a permis de cartographier la surface de la PTH où l'ARN s'ancre à la protéine. Il a confirmé l'importance de deux résidus basiques, la lysine K105 et l'arginine R133, pour la reconnaissance du phosphate en 5' de l'ARNt. Il a également révélé une interaction entre l'hélice C-terminale de la protéine et le bras T $\Psi$ C de l'ARNt, à 30 Å du site actif. La pertinence fonctionnelle de ce dernier contact a pu être établie par mutagenèse dirigée. L'ensemble de ces résultats permet de proposer un modèle complet de l'interaction entre la PTH et un peptidyl-ARNt.

Peptidyl-tRNA hydrolase (PTH) catalyzes the hydrolysis of peptidyl-tRNA molecules that have prematurely dissociated from the ribosome during protein translation. This protein is essential for bacterial viability whereas its eukaryotic homolog is dispensable. Therefore, PTH is an attractive target for the development of new antibacterials, and it seems important to map the interaction between this protein and its substrate in order to facilitate the design of inhibitors.

Because attempts to obtain crystals of PTH with substrate mimics remained unsuccessful, we chose to study such complexes in solution, by NMR spectroscopy. Spectra of  $^{15}\text{N}$ - $^{13}\text{C}$  uniformly labeled *Escherichia coli* PTH allowed us to assign the resonance frequencies of the protein backbone atoms and of many side-chain atoms. Then, we analyzed the interaction between PTH and a chemically synthesized small molecule, 3'-(*L*-[*N,N*-diacetyl-lysiny]amino-3'-deoxyadenosine, imitating the 3'-part of a peptidyl-tRNA. This study points out the role of many residues of the active center. In particular, the resulting enzyme-bound substrate model indicates (i) that a phenylalanine residue (Phe66) stacks with the adenine ring of the 3'-terminal adenosine, (ii) that an asparagine residue (Asn114) holds the water molecule involved in the hydrolysis of the substrate and (iii) that another asparagine residue (Asn10) confers to the PTH the capacity to distinguish peptidyl-tRNAs from aminoacyl-tRNAs.

We also characterized the interaction between PTH and mini-tRNAs derived from the acceptor end of tRNA<sup>His</sup>. These results confirmed the major role of Arg133 and Lys105 residues in the recognition of the 5'-phosphate of tRNA. They also showed an interaction between the C-terminal helix of the protein and the T $\Psi$ C arm of the tRNA, 30 Å away from the active site. The functional relevance of the latter contact could be established by site-directed mutagenesis. Altogether, our results allow us to propose a complete model of the interaction between PTH and a peptidyl-tRNA.

« Selon une autre théorie, ladite chose se serait en fait déjà produite. »

D.N.A. (1952-2001)





# **Remerciements**



## *Remerciements*

---

Je tiens à remercier tout d'abord Pierre Plateau pour m'avoir accueilli à ses côtés et pour m'avoir encadré et guidé tout au long de ma thèse en étant particulièrement présent et pédagogue.

Je tiens aussi à remercier Sylvain Blanquet, dont les conseils ont été déterminants tout au long de la réalisation de ce travail, et François Bontems, qui m'a fait découvrir la résonance magnétique nucléaire et qui a grandement contribué aux expériences de RMN et aux approches par modélisation présentées dans cette thèse.

Je remercie également Marc Mirande et Carine Tisné pour avoir accepté d'être rapporteurs de cette thèse et Christophe Escudé pour avoir bien voulu examiner ce travail.

Je souhaite ensuite remercier tous ceux qui ont apporté leur aide précieuse à mes travaux, Michel Fromant pour ses nombreux conseils, Caroline Launay pour tout ce qu'elle a fait, Jean-Louis Leroy, Sandrine Caputo et Aurélien Thureau de l'ICSN, ainsi que Guillaume van der Rest et Javier Perez qui m'ont fait découvrir leurs spécialités.

Je souhaite aussi remercier tous les membres du laboratoire que j'ai pu côtoyer et qui m'ont accueilli chaleureusement dans le monde de la recherche, Christine Schurdevin, Michel Panvert, Myriam Lazard, Yves Mechulam, Emmanuelle Schmitt, Thomas Simonson, Pascal et Pascale Delplanche, Annick Carpentier, Guillaume Heugas, Catherine Teillet, Catherine Morais, Françoise Humbert, Mélanie Monguillon, Marc Dauplais, Laura Fioni, François Cluzeau, Pierre-Damien Coureux, David Mignon, Gérald Peyroche et Alfonso Jaramillo.

Je tiens aussi à remercier particulièrement tous les thésards et post-doc du laboratoire, pour tous les bons moments passés ensemble, Laurent Guillon tout d'abord qui fut un vrai soutien, Agathe Tarze et Sandra Wydau, pour les donuts partagés, Laurent Mughherli, Marie Naveau, Laure Yatime, Romary Perrin, Lionel Cladière, Kader Ould Boubacar, Anne Lopes, Guillaume Launay, Josselin Noirel, Marcel Schmidt Am Busch, Alexey Aleksandrov, Najette Amara, Damien Thompson et Pablo Tortosa.

Je remercie aussi particulièrement ceux qui m'ont accueilli très chaleureusement dans leur équipe enseignante à Paris 7 et qui ont confirmé mon goût pour l'enseignement : Annick Méjean, Thérèse de Caldas, Fernando Rodriguez-Lima, Julien Dairou, Valérie Serre et Bertrand Friguet.

Je profite également de cette occasion pour remercier mes parents et tous les enseignants qui au cours de mes études m'ont donné le goût de la découverte scientifique, me poussant inexorablement dans la voie de la recherche pour en apprendre toujours plus.

## *Remerciements*

---

Je remercie enfin ma famille, mes amis et mes proches, parce qu'ils ont toujours été là dans les périodes les plus difficiles de ma thèse, pour me changer les idées, me soutenir, m'encourager, et m'apporter un peu de bonheur, ou beaucoup, en particulier Sara.

# **Sommaire**



---

	<i>Pages</i>
<b>Résumé</b>	<b>5</b>
<b>Abstract</b>	<b>6</b>
<b>Remerciements</b>	<b>9</b>
<b>Sommaire</b>	<b>13</b>
<b>Liste des figures et tableaux</b>	<b>25</b>
<b>Liste des abréviations</b>	<b>33</b>
<b>Introduction</b>	<b>37</b>
<b>1. Rôle des ARNt dans la synthèse protéique</b>	<b>41</b>
<b>1.1 Les ARNt</b>	<b>43</b>
<b>1.1.1 Structure des ARNt</b>	<b>43</b>
<b>1.1.2 Biosynthèse des ARNt</b>	<b>44</b>
<b>1.1.3 Biosynthèse des protéines</b>	<b>45</b>
<b>1.2 La correction d'erreurs dans la traduction</b>	<b>48</b>



---

	<i>Pages</i>
<b>2. Discrimination des acides aminés par les aminoacyl-ARNt synthétases</b>	<b>53</b>
<b>2.1 Spécificité des aminoacyl-ARNt synthétases</b>	<b>55</b>
2.1.1 Aminoacylation par les aminoacyl-ARNt synthétases	55
2.1.2 Spécificité des aminoacyl-ARNt synthétases	57
2.1.3 L'édition par les aminoacyl-ARNt synthétases	60
<b>2.2 L'édition par les aminoacyl-ARNt synthétases de classe I</b>	<b>62</b>
2.2.1 L'activité d'édition des synthétases de classe I	62
2.2.2 Structure des domaines CP1	64
<b>2.3 L'édition par les aminoacyl-ARNt synthétases de classe II</b>	<b>68</b>
2.3.1 L'édition pour les ARNt <sup>Pro</sup>	68
2.3.1.1 L'édition dans les ProRS	68
2.3.1.2 Edition en <i>trans</i>	69
2.3.1.3 ProX	70
2.3.1.4 YbaK	70
2.3.1.5 Différentes stratégies d'édition pour les ARNt <sup>Pro</sup>	71
2.3.2 L'édition pour les ARNt <sup>Thr</sup>	72
2.3.2.1 L'édition des ThrRS bactériennes et eucaryotes	72
2.3.2.2 L'édition par les ThrRS chez les archées	73
2.3.2.3 Une édition catalysée en <i>trans</i> chez quelques archées	74
2.3.3 L'édition des ARNt <sup>Ala</sup> mésaminoacylés	75
2.3.3.1 L'édition catalysée par les AlaRS	75

---

	<i>Pages</i>
2.3.3.2 Trans-édition des ARNt <sup>Ala</sup> mésaminoacylés	75
2.3.3.3 Structure des sites d'édition des AlaRS et de la protéine AlaX	76
2.3.4 L'édition pour les ARNt <sup>Phe</sup>	78
2.3.4.1 L'activité d'édition des PheRS	78
2.3.4.2 Le domaine d'édition des PheRS bactériennes	78
2.3.4.3 Le domaine d'édition des PheRS archéennes et eucaryotes	80
<b>3. Les D-aminoacyl-ARNt désacylases</b>	<b>81</b>
<b>3.1 Le problème des acides aminés D</b>	<b>83</b>
<b>3.2 D-aminoacyl-ARNt désacylases bactériennes (DTD1)</b>	<b>84</b>
<b>3.3 D-aminoacyl-ARNt désacylases archéennes (DTD2)</b>	<b>87</b>
<b>3.4 D-aminoacyl-ARNt désacylases chez les cyanobactéries (DTD3)</b>	<b>89</b>
<b>4. Le <i>drop-off</i></b>	<b>91</b>
<b>4.1 Edition ribosomale et <i>drop-off</i></b>	<b>93</b>

---

	<i>Pages</i>
<b>4.2 Terminaison précoce de la traduction impliquant les facteurs de terminaison</b>	<b>95</b>
<b>4.3 Contrôle du démarrage de la traduction</b>	<b>96</b>
<b>4.4 Trans-traduction et <i>drop-off</i></b>	<b>97</b>
<b><i>5. Les peptidyl-ARNt hydrolases</i></b>	<b>99</b>
<b>5.1 La peptidyl-ARNt hydrolase bactérienne</b>	<b>101</b>
5.1.1 Spécificité de la PTH bactérienne (PTH1)	101
5.1.2 Structure et mécanisme des PTH bactériennes	102
<b>5.2 La peptidyl-ARNt hydrolase archéenne</b>	<b>108</b>
5.2.1 Spécificité du substrat de PTH2	108
5.2.2 Structure et mécanisme des PTH archéennes	108
<b>5.3 Les peptidyl-ARNt hydrolases chez les eucaryotes</b>	<b>111</b>
5.3.1 Localisation intracellulaire des PTH1 et PTH2 eucaryotes	111
5.3.2 Autres fonctions des homologues de PTH1 et PTH2 chez les eucaryotes	111

---

	<i>Pages</i>
<b>5.4 Autres protéines agissant sur les peptidyl-ARNt</b>	<b>113</b>
<b>Objectifs de la thèse</b>	<b>115</b>
<b>Résultats</b>	<b>119</b>
<b>Chapitre 1 : Structure de la PTH d'<i>E. coli</i> en solution</b>	<b>121</b>
<b>1.1 Préparation de la protéine pour la RMN</b>	<b>123</b>
1.1.1 Production et purification de la PTH H20A	123
1.1.2 Choix des conditions pour la RMN	124
<b>1.2 Attribution des fréquences de résonance</b>	<b>128</b>
1.2.1 Analyse du spectre HSQC de la PTH H20A	128
1.2.2 Les systèmes de spins	130
1.2.3 Détermination de l'enchaînement des systèmes de spins	130
1.2.4 Identifications d'acides aminés spécifiques	131
1.2.5 Attribution des séquences grâce aux acides aminés spécifiques	132
1.2.6 Amélioration de l'attribution	133
1.2.7 Bilan de l'attribution	134

---

	<i>Pages</i>
<b>1.3 Etude structurale de la PTH en solution</b>	<b>136</b>
1.3.1 Structure secondaire de la PTH	136
1.3.2 Structure topologique de la PTH	137
<b>Chapitre 2 : Caractérisation de l'interaction entre la PTH et un analogue de son substrat, la diacétyl-Lys-(3'NH)-adénosine</b>	<b>143</b>
<b>2.1 Présentation du ligand étudié</b>	<b>145</b>
<b>2.2 Affinité de la diacétyl-Lys-(3'NH)-adénosine pour la PTH</b>	<b>146</b>
<b>2.3 Principe de la détermination des interactions protéine-ligand par RMN</b>	<b>147</b>
<b>2.4 Empreinte de la diacétyl-Lys-(3'NH)-adénosine sur la PTH</b>	<b>149</b>
<b>2.5 Positionnement de la diacétyl-Lys-(3'NH)-adénosine sur la PTH</b>	<b>151</b>
2.5.1 Construction des modèles	151
2.5.2 Interaction avec F66	153

---

	<i>Pages</i>
<b>2.5.3 Interaction avec N10, le résidu à la base de la discrimination entre aminoacyl-ARNt et peptidyl-ARNt</b>	<b>154</b>
<b>2.5.4 Interaction avec N68</b>	<b>156</b>
<b>2.5.5 Interaction avec N114</b>	<b>156</b>
<b>2.5.6 Interaction avec Tyr15</b>	<b>157</b>
<b>2.5.7 Autres interactions</b>	<b>157</b>
<b>2.5.8 Comparaison avec le modèle cristallographique</b>	<b>158</b>
<b>2.6 Conclusions</b>	<b>160</b>
<b>Chapitre 3 : Interaction entre la protéine et l'ARN</b>	<b>161</b>
<b>3.1 Interaction de la PTH avec l'ARNt<sup>His</sup></b>	<b>163</b>
<b>3.2 Etude de mini-substrats</b>	<b>165</b>
<b>3.2.1 Synthèse des mini-substrats</b>	<b>166</b>
<b>3.2.2 Etude catalytique des mini-substrats</b>	<b>167</b>
<b>3.2.3 Recherche d'un ligand pour l'étude RMN</b>	<b>167</b>
<b>3.3 Duplex<sup>His</sup></b>	<b>171</b>

---

	<i>Pages</i>
<b>3.3.1 Synthèse et aminoacylation du duplex<sup>His</sup></b>	<b>171</b>
<b>3.3.2 Etude catalytique de l'interaction du duplex<sup>His</sup> avec la PTH</b>	<b>172</b>
<b>3.4 Empreinte du duplex<sup>His</sup> sur la protéine</b>	<b>173</b>
3.4.1 Enregistrement des données	173
3.4.2 Analyse des spectres	174
3.4.3 Résultats obtenus à un ratio ARN/protéine de 0,4	175
<b>3.5 Analyse des régions affectées par la fixation du duplex<sup>His</sup></b>	<b>179</b>
3.5.1 Le site actif	179
3.5.2 La boucle-hélice au dessus du site actif	181
3.5.3 La région de la "pince" électrostatique formée par les résidus 105 et 133	181
3.5.4 L'hélice C-terminale	184
<b>3.6 Mutagenèse</b>	<b>185</b>
<b>3.7 Conclusion</b>	<b>186</b>

---

	<i>Pages</i>
<b>Chapitre 4 : Modélisation de l'interaction entre la PTH et son substrat</b>	<b>187</b>
<b>4.1 Modélisation de l'interaction entre la PTH et le duplex<sup>His</sup></b>	<b>189</b>
4.1.1 Paramètres de la modélisation	189
4.1.2 Modélisation sans orientation préalable	189
4.1.3 Modélisation avec orientation préalable du duplex	191
<b>4.2 Modélisation de l'interaction entre la PTH et un de ses substrats (duplex<sup>His</sup> aminoacylé et acétylé)</b>	<b>193</b>
4.2.1 Construction du modèle	193
4.2.2 Analyse des modèles obtenus	194
<b>4.3 Interaction de la PTH avec un ARNt complet</b>	<b>201</b>
<b>Conclusion</b>	<b>203</b>
<b>Annexe 1 : Séquence de la PTH H20A étiquetée</b>	<b>211</b>



---

	<i>Pages</i>
<b>Annexe 2 : Liste des expériences de RMN réalisées sur la PTH</b>	<b>215</b>
<b>Matériels et Méthodes</b>	<b>219</b>
<b>Bibliographie</b>	<b>231</b>

# **Liste des figures et tableaux**



	<i>Pages</i>
<b>Figure 1 :</b> Structure primaire des ARN	<b>43</b>
<b>Figure 2 :</b> Structure secondaire et numérotation des nucléotides	<b>44</b>
<b>Figure 3 :</b> Structure tridimensionnelle de l'ARNt <sup>Phe</sup> de levure	<b>44</b>
<b>Figure 4 :</b> Fidélité dans l'expression génétique	<b>49</b>
<b>Figure 5 :</b> Erreurs et corrections dans le cycle des ARNt	<b>51</b>
<b>Tableau 1 :</b> Répartition des aminoacyl-ARNt synthétases	<b>56</b>
<b>Figure 6 :</b> Structure des 20 acides aminés	<b>57</b>
<b>Tableau 2 :</b> Exemple de compétiteurs naturels des acides aminés	<b>58</b>
<b>Figure 7 :</b> Différentes voies d'édition par les synthétases	<b>61</b>
<b>Figure 8 :</b> Edition de l'homocystéine par la MetRS	<b>63</b>
<b>Figure 9 :</b> Structures tridimensionnelles de la LeuRS et de l'IleRS de <i>T. thermophilus</i>	<b>64</b>
<b>Figure 10 :</b> Alignement des domaines CP1	<b>65</b>
<b>Figure 11 :</b> Site d'édition de la LeuRS de <i>T. thermophilus</i>	<b>66</b>
<b>Figure 12 :</b> Edition pré-transfert de l'ornithine par la LysRS de classe II	<b>68</b>
<b>Figure 13 :</b> Mécanisme de l'édition du Ser-ARNt <sup>Thr</sup> par la ThrRS d' <i>E. coli</i>	<b>72</b>
<b>Figure 14 :</b> Site d'édition de la ThrRS de <i>P. abyssi</i> en présence d'un analogue du substrat de l'édition post-transfert	<b>74</b>
<b>Figure 15 :</b> Organisation des domaines de l'AlaRS, des protéines AlaX-L, AlaX-M et AlaX-S, et de la ThrRS	<b>76</b>
<b>Figure 16 :</b> Sérine dans le site d'édition de l'AlaX-S de <i>P. horikoshii</i>	<b>77</b>
<b>Figure 17 :</b> Mécanisme de l'édition post-transfert catalysée par la PheRS de <i>T. thermophilus</i>	<b>79</b>
<b>Figure 18 :</b> Structure de DTD1	<b>84</b>

	<i>Pages</i>
<b>Figure 19 :</b> Mécanisme catalytique de la D-Tyr-ARNt désacylase	<b>85</b>
<b>Figure 20 :</b> Représentation schématique des structures tridimensionnelles de la PTH1 d' <i>E. coli</i> et de la protéine DT2 d' <i>A. fulgidus</i>	<b>87</b>
<b>Figure 21 :</b> Interactions électrostatiques entre deux protéines voisines dans le cristal de la PTH d' <i>E. coli</i>	<b>103</b>
<b>Tableau 3 :</b> Paramètres catalytiques de mutants de la PTH d' <i>E. coli</i>	<b>104</b>
<b>Figure 22 :</b> Localisation des résidus d'asparagine et des résidus basiques à la surface de la PTH	<b>105</b>
<b>Figure 23 :</b> Structure de la PTH archéenne	<b>109</b>
<b>Figure 24 :</b> Analyse de la purification de la PTH par électrophorèse sur gel de polyacrylamide en conditions dénaturantes	<b>124</b>
<b>Figure 25 :</b> Spectre <sup>15</sup> N-HSQC de la PTH H20A	<b>126</b>
<b>Figure 26 :</b> Spectres <sup>15</sup> N-HSQC de la PTH sauvage non étiquetée	<b>127</b>
<b>Figure 27 :</b> Représentation schématique des atomes de deux acides aminés successifs dans une chaîne peptidique	<b>128</b>
<b>Figure 28 :</b> Acides aminés ayant un ou deux HN dans leur chaîne latérale	<b>129</b>
<b>Figure 29 :</b> Déplacements chimiques standard des atomes de carbones dans les protéines	<b>132</b>
<b>Tableau 4 :</b> Eléments de structure secondaire de la PTH d' <i>E. coli</i> prédits par la méthode CSI ou observés dans les cristaux de cette protéine	<b>136</b>
<b>Tableau 5 :</b> Observation des NOE prédits d'après la structure cristallographique de la PTH	<b>139</b>
<b>Figure 30 :</b> Comparaison de la structure topologique de la PTH dans le cristal et en solution	<b>141</b>
<b>Figure 31 :</b> Structure de la diacétyl-Lys-(3'NH)-adénosine	<b>146</b>

	<i>Pages</i>
<b>Figure 32 :</b> Spectres $^1\text{H}$ - $^{15}\text{N}$ -HSQC de la PTH H20A en présence ou en absence de diacétyl-(3'NH)-adénosine	149
<b>Figure 33 :</b> Variation du déplacement chimique des pics des $\text{H}^{\text{N}}$ -N de la protéine lors de l'ajout de diacétyl-Lys-(3'NH)-adénosine	150
<b>Figure 34 :</b> Empreinte de la diacétyl-Lys-(3'NH)-adénosine sur la PTH	150
<b>Figure 35 :</b> Modèle du complexe entre la diacétyl-Lys-(3'NH)-adénosine et la PTH	152
<b>Figure 36 :</b> Positions relatives de l'histidine 20 et de la diacétyl-Lys-(3'NH)-adénosine	153
<b>Tableau 6 :</b> Comparaison de la spécificité de la PTH sauvage et des mutants N10A et N10D	155
<b>Tableau 7 :</b> Distances des interactions entre la protéine et la diacétyl-Lys-(3'NH)-adénosine	155
<b>Figure 37 :</b> Interactions dans le site actif de la PTH	159
<b>Figure 38 :</b> Spectre HSQC du complexe entre la PTH 130 $\mu\text{M}$ et 13 $\mu\text{M}$ d'ARNt <sup>His</sup> d' <i>E. coli</i>	164
<b>Figure 39 :</b> Spectre HSQC du complexe entre la PTH 130 $\mu\text{M}$ et 260 $\mu\text{M}$ d'ARNt <sup>His</sup> d' <i>E. coli</i>	164
<b>Figure 40 :</b> Comparaison du complexe EF-Tu:GTP:Phe-ARNt <sup>Phe</sup> (A) et du modèle proposé pour la PTH (B)	165
<b>Figure 41 :</b> Représentation des différents ARN étudiés	166
<b>Tableau 8 :</b> Efficacités catalytiques de la PTH d' <i>E. coli</i> pour différents mini-substrats	167
<b>Figure 42 :</b> Spectres de RMN 1D de la tétraboucle <sup>His</sup>	168
<b>Figure 43 :</b> Appariement Hoogsteen G-C+	169
<b>Figure 44 :</b> Spectres $^1\text{H}$ - $^{15}\text{N}$ -HSQC de la PTH en présence de différentes concentrations de duplex <sup>His</sup>	173
<b>Figure 45 :</b> Modifications des pics de résonance de la PTH après ajout de 0,4 équivalent d'ARN	176

	<i>Pages</i>
<b>Tableau 9 :</b> Variations des pics de résonance des résidus de la PTH affectés par l'ajout de 0,4 équivalent de duplex <sup>His</sup>	177
<b>Figure 46 :</b> Représentation des résidus de la PTH affectés par la fixation du duplex <sup>His</sup>	178
<b>Figure 47 :</b> Représentation des régions de la PTH affectées par la fixation du duplex <sup>His</sup>	179
<b>Figure 48 :</b> Effet de l'addition de duplex <sup>His</sup> sur la région du spectre <sup>15</sup> N-HSQC correspondant aux couples H <sup>ε</sup> -N <sup>ε</sup> des arginines de la PTH	182
<b>Figure 49 :</b> Effet de la mutation R133A sur la région du spectre <sup>15</sup> N-HSQC correspondant aux couples H <sup>ε</sup> -N <sup>ε</sup> des arginines de la PTH	183
<b>Tableau 10 :</b> Paramètres catalytiques de différents mutants de la PTH d' <i>E. coli</i>	186
<b>Figure 50 :</b> Energies des modèles de complexes entre la PTH et le duplex obtenus sans orientation préalable	190
<b>Figure 51 :</b> Complexe entre la PTH et le duplex sans orientation préalable	190
<b>Figure 52 :</b> Energies des modèles de complexes entre la PTH et le duplex obtenus avec une orientation préalable du duplex	191
<b>Figure 53 :</b> Structure des modèles de complexes entre la PTH et le duplex obtenus avec une orientation préalable du duplex	191
<b>Tableau 11 :</b> Distances entre le duplex et la protéine dans les quatre meilleurs modèles obtenus	192
<b>Figure 54 :</b> Site actif de la PTH liée au diacétyl-Lys-duplex <sup>His</sup>	195
<b>Tableau 12 :</b> Distances entre la PTH et le diacétyl-Lys-duplex <sup>His</sup>	195
<b>Figure 55 :</b> Comparaison de la « pince » lysine-arginine dans la PTH et dans EF-Tu	197

---

	<i>Pages</i>
<b>Figure 56 :</b> Région de R131 dans le complexe PTH : diacétyl-Lys-duplex <sup>His</sup>	<b>198</b>
<b>Figure 57 :</b> Interactions hydrophobes entre les cytosines de l'extrémité 3'-terminale de l'ARN et la PTH	<b>199</b>
<b>Figure 58 :</b> Site actif de la PTH occupé par son substrat	<b>200</b>
<b>Figure 59 :</b> Modèle d'un complexe PTH : ARNt	<b>201</b>





# **Liste des abréviations**



aa-AMP : amimoacyl-adénylate  
aa-ARNt : amimoacyl-ARNt  
aaRS : aminoacyl-ARNt synthétase  
ADN : acide désoxyribonucléique  
AMP : adénosine monophosphate  
ARN : acide ribonucléique  
ARNm : acide ribonucléique messenger  
ARNt : acide ribonucléique de transfert  
ATP : adénosine triphosphate  
CAF : CRS2-associated factor  
CP1 : connective peptide 1  
CRS2 : chloroplast RNA splicing 2  
DO : densité optique  
DTD : D-aminoacyl-ARNt désacylase  
DTT : dithiothréitol  
EDTA : acide éthylène-diamine tétra-acétique  
EF-G : Elongation factor G  
EF-Tu : Elongation factor thermo unstable  
GMP : guanine monophosphate  
HPLC : High Pressure Liquid Chromatography  
HSQC : Heteronuclear Single Quantum Correlation  
IF1 : Initiation factor 1  
IF2 : Initiation factor 2  
IF3 : Initiation factor 3

## *Liste des abréviations*

---

IPTG : isopropyl- $\beta$ -D-thiogalactoside

Kd : constante de dissociation

Ki : constante d'inhibition

Km : constante de Michaelis

NTP : nucléoside triphosphate

NOE : Nuclear Overhauser Effect

PCR : polymerase chain reaction

PPi : pyrophosphate inorganique

ppm : partie par million (utilisé comme unité de mesure des déplacements chimiques)

PTH : peptidyl-ARNt hydrolase

RMN : résonance magnétique nucléaire

RRF : ribosome recycling factor

RF3 : release factor 3

SDS : dodécyl sulfate de sodium

SDS-PAGE : SDS polyacrylamide gel electrophoresis

# **Introduction**



En 1953, à Cambridge, James Watson et Francis Crick établissent la structure hélicoïdale de l'ADN (Watson and Crick, 1953). Les années suivantes ont vu naître des découvertes spectaculaires dans le fonctionnement moléculaire des organismes vivants. Véritable défi scientifique et épique de la fin du vingtième siècle, opposant le consortium public international mené entre autres par James Watson, et une société privée, Celera Genomics, dirigée par Craig Venter, le séquençage complet de l'ADN du génome humain est officiellement annoncé en juin 2000 par Bill Clinton, à la Maison Blanche. Mais les 20 000 gènes découverts (Consortium, 2004) ne sont que la source d'information nécessaire à la synthèse de macromolécules qui sont, elles, les véritables acteurs de toutes les fonctions biologiques. La compréhension du rôle et du mode d'action de toutes ces macromolécules et leurs réseaux, constitue un des défis du 21<sup>ème</sup> siècle.

Les macromolécules biologiques occupent une place et une fonction précise dans l'orchestre cellulaire. Comprendre le rôle de ces molécules, permet donc de comprendre comment s'organise la vie, comment elle peut être désorganisée par diverses maladies, et comment ces dysfonctionnements peuvent être évités ou corrigés. Les protéines, acteurs majeurs du monde vivant, sont de longs polymères constitués à partir de vingt acides aminés différents. L'information génétique, elle, est codée par quatre bases A, T, G, et C. Cette information peut être copiée pour permettre la duplication du matériel génétique avant la division cellulaire. Pour la synthèse protéique, l'ADN doit d'abord être transcrit en ARN messager, qui est, lui aussi, basé sur un langage à 4 lettres (A, U, G et C) et qui, comme son nom l'indique, porte le message nécessaire à la synthèse d'une protéine (à partir d'un gène).

La conversion du langage à quatre lettres de l'ARN au langage à 20 mots (les vingt acides aminés) des protéines est réalisée au niveau du ribosome. Cette étape, appelée traduction, nécessite des interprètes spécifiques de chaque acide aminé : ce sont les ARN de transfert ou ARNt. Leur existence avait été prédite dès 1955 par Francis Crick, avant qu'ils ne soient découverts par Hoagland et Zamecnik en 1958 (Hoagland *et al.*, 1958). Indispensables à la traduction du message génétique en protéines, ils sont nécessaires à la vie. Leur concentration totale a été estimée à 0,5 mM chez *Escherichia coli*, ce qui correspond à environ 200 000 molécules par cellule (Goodsell, 1991). Ces molécules clés et les différents mécanismes leur permettant de jouer efficacement leur rôle, en limitant les erreurs, vont être présentés dans cette introduction.





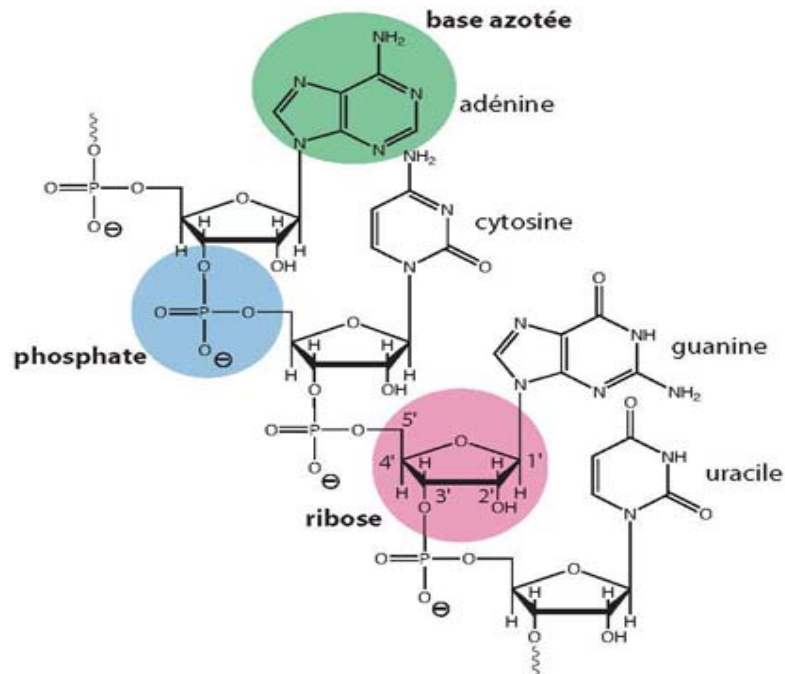
# **1. Rôle des ARNt dans la synthèse protéique**



## 1.1 Les ARNt

### 1.1.1 Structure des ARNt

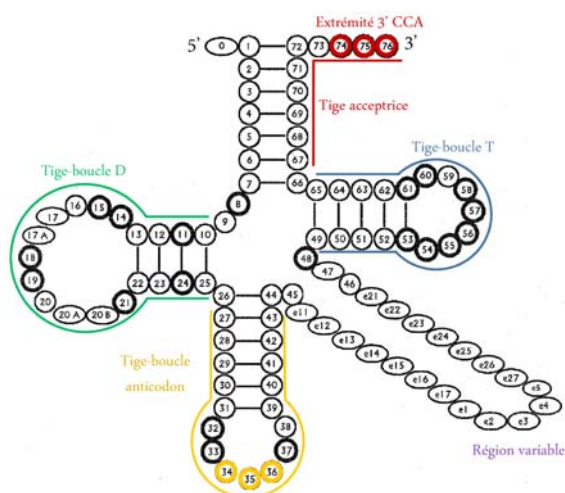
Les ARNt sont généralement composés de l'enchaînement de 74 à 98 nucléotides. Chaque nucléotide est composé d'un sucre, le ribose, d'un groupement phosphate et d'une base azotée qui peut être soit une purine, adénine (A) ou guanine (G), soit une pyrimidine, cytosine (C) ou uracile (U) (figure 1). Les nucléotides sont reliés entre eux par une liaison phosphodiester impliquant l'oxygène en 3' du ribose du nucléotide *i* et le phosphate en 5' du nucléotide *i*+1. Ils sont conventionnellement numérotés du numéro 1 à l'extrémité 5' au numéro 76 à l'extrémité 3'.



#### Figure 1 : Structure primaire des ARN

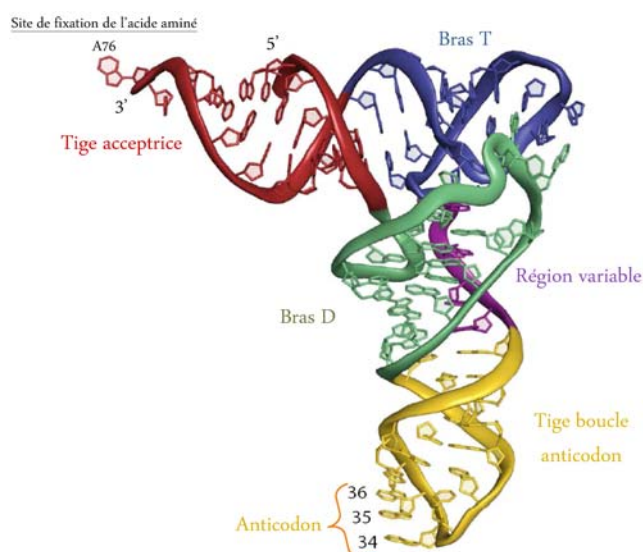
Les trois éléments constitutifs des nucléotides sont représentés sur fond de couleur : le ribose en rose, la base azotée en vert (reliée au ribose par une liaison N-glycosidique) et le groupement phosphate en bleu. Les carbones du cycle du ribose sont numérotés de 1' à 5'.

La plupart des ARNt ont une structure secondaire caractéristique « en feuille de trèfle » (figure 2) formée de quatre régions de bases appariées, séparées par des boucles (boucle T $\psi$ C, boucle variable, boucle de l'anticodon et boucle DHU). Ces molécules présentent une organisation tridimensionnelle en forme de L (figure 3), avec deux régions importantes à chaque extrémité, (i) la tige acceptrice se terminant par trois nucléotides (en rouge), C, C et A<sub>76</sub>, qui va accepter l'acide aminé et (ii) la boucle de l'anticodon (en jaune), spécifique de chaque type d'ARNt, qui s'appariera à l'ARN messenger sur le ribosome.



**Figure 2 :**  
**Structure secondaire et numérotation des nucléotides**

Les cercles représentent des nucléotides variables mais présents dans tous les ARNt, et les ovales des nucléotides qui ne sont pas toujours présents dans les ARNt. Les trois nucléotides C74, C75 et A76, à l'extrémité de la tige acceptrice, sont entourés de rouge. Les trois bases de l'anticodon sont entourées en jaune. D'après Sprinzl *et al.* (1998).



**Figure 3 :**  
**Structure tridimensionnelle de l'ARNt<sup>Phe</sup> de levure**

Les différentes régions de l'ARNt sont coloriées avec les mêmes couleurs que sur la figure 2. D'après Shi et Moore (2000) (code PDB 1EHZ).

### 1.1.2 Biosynthèse des ARNt

La biosynthèse des ARNt est un processus complexe qui nécessite l'intervention de plusieurs enzymes, variables d'un organisme à l'autre. Chez la levure *Saccharomyces cerevisiae*, Phizicky a estimé que plus de 60 protéines différentes interviennent dans la maturation des ARNt pour produire 3 à 6 millions d'ARNt par génération cellulaire (Phizicky, 2005). Après leur production par une ARN polymérase, les précurseurs d'ARNt vont subir plusieurs modifications biochimiques. Cette maturation se déroule en 5 étapes majeures :

1) digestion de leur extrémité 5' par une endonucléase, la ribonucléase P (RNase P), qui, de manière quasi-universelle, est un complexe ribonucléoprotéique où l'ARN porte l'activité catalytique (Evans *et al.*, 2006 ; Guerrier-Takada *et al.*, 1983 ; Torres-Larios *et al.*, 2006).

2) digestion endonucléolytique ou exonucléolytique de l'extrémité 3'. La voie endonucléolytique est catalysée par la RNase Z, présente dans la plupart, voire la totalité, des eucaryotes et des archées, et dans environ la moitié des bactéries dont le génome a été séquencé. La RNase Z clive le précurseur de l'ARNt au niveau de la base 73, générant un ARN prêt pour l'addition du CCA

par l'ARNt nucléotidyl-transférase (voir ci-dessous). La voie exonucléolytique commence par une coupure endonucléolytique par la RNase E, en aval de l'extrémité 3' de l'ARNt mature. Cette réaction est suivie par la digestion exonucléolytique des nucléotides restant en 3', grâce à l'action de plusieurs exonucléases aux fonctions redondantes (RNase PH, RNase T, RNase D et RNase II chez *E. coli*) (Hartmann *et al.*, 2009 ; Redko *et al.*, 2007).

3) addition de la séquence CCA à l'extrémité 3', dans les organismes où cette séquence n'est pas codée par les gènes des ARNt (les eucaryotes, la plupart des bactéries et certaines archées). Cette maturation est catalysée par l'ARNt nucléotidyl-transférase (Deutscher, 1982 ; Hecht *et al.*, 1958 ; Xiong and Steitz, 2004).

4) épissage des introns présents dans certains gènes d'ARNt eucaryotes, archéens et, parfois, bactériens. Chez les bactéries, les introns s'auto-épissent. Chez les eucaryotes et les archées, l'épissage fait intervenir une endonucléase qui excise les introns et une ligase qui joint les exons (Abelson *et al.*, 1998).

5) modification de nucléotides (une dizaine en moyenne par ARNt) (Rozenski *et al.*, 1999). Ces modifications peuvent être des réactions chimiques simples comme des méthylations, des thiolations ou des désaminations. Ce sont aussi parfois des modifications complexes impliquant la participation de plusieurs enzymes. A ce jour, 119 nucléotides modifiés distincts ont été identifiés dans les ARNt (Czerwoniec *et al.*, 2009).

Une fois mature, l'ARNt va pouvoir être reconnu par une aminoacyl-ARNt synthétase, qui va greffer covalamment un acide aminé à son extrémité 3'-OH. Après avoir été aminoacylé, l'ARNt pourra alors délivrer son acide aminé sur le ribosome qui réalise la synthèse protéique à partir de l'information portée par un ARN messenger.

### 1.1.3 Biosynthèse des protéines

Le ribosome est un complexe ribonucléoprotéique dont plusieurs structures à haute résolution ont été obtenues au début des années 2000 (Ban *et al.*, 2000 ; Schlutzenzen *et al.*, 2000 ; Wimberly *et al.*, 2000 ; Yusupov *et al.*, 2001). Il est formé de deux sous-unités :

1) la petite sous-unité (30S chez les bactéries et les archées, 40S chez les eucaryotes) qui contient le site de décodage où se fait l'interaction codon-anticodon entre un ARNt et un ARNm, au niveau de trois sites : le site A (pour Aminoacyl), le site P (pour Peptide) et le site E (pour Exit). Chaque site présente un groupe de trois nucléotides de l'ARN messenger, appelé codon, auquel peut venir se fixer un ARNt possédant l'anticodon correspondant. Au maximum, deux ARNt peuvent occuper simultanément ces trois sites.

2) la grande sous-unité (50S chez les bactéries et les archées, 60S chez les eucaryotes) qui contient le site actif de transfert peptidique nécessaire à l'élongation de la chaîne polypeptidique et où s'effectue l'hydrolyse de la liaison entre l'ARNt et la protéine néo-synthétisée, au moment de la terminaison de la traduction.

Lors du démarrage de la traduction, les deux sous-unités s'associent pour former le complexe de démarrage (70S chez les bactéries et les archées, 80S chez les eucaryotes) qui contient aussi un ARN messager et un ARNt initiateur méthionylé. Plusieurs facteurs de démarrage (IF1, IF2, IF3 chez les bactéries, 5 facteurs aIF chez les archées et 13 facteurs eIF chez les eucaryotes) interviennent lors de la formation de ce complexe, qui se déroule de manière différente chez les bactéries, chez les archées et chez les eucaryotes.

Chez tous les organismes, la traduction démarre par l'association d'un ARNt initiateur et d'un codon AUG (parfois, GUG ou UUG chez les bactéries et les archées) sur l'ARN messager, au niveau du site P. Cet ARNt initiateur porte à son extrémité 3' une méthionine, qui est formylée chez les bactéries. Les ARNt initiateurs comportent des caractéristiques (comme le mésappariement de la base 1-72 chez les procaryotes ou la présence d'une paire de bases A<sub>1</sub>U<sub>72</sub> chez les eucaryotes) qui leur permettent d'interagir avec les facteurs de démarrage et pas avec les facteurs d'élongation.

Pendant un cycle d'élongation, les aminoacyl-ARNt élongateurs sont reconnus et emmenés sur le ribosome par un facteur d'élongation (EF-Tu chez les bactéries, aEF1A chez les archées, eEF1A chez les eucaryotes). L' aminoacyl-ARNt se fixe alors au niveau du site A, avec appariement codon-anticodon, chassant l'ARNt éventuellement présent dans le site E, puis réagit avec le peptidyl-ARNt présent au niveau du site P, ce qui conduit au transfert du peptide sur l'acide aminé porté par l'ARNt dans le site A, allongeant ainsi la chaîne peptidique d'un acide aminé. Cette réaction de transfert peptidique est suivie par la translocation du nouveau peptidyl-ARNt du site A vers le site P et de l'ARNt désacylé du site P vers le site E, avec un décalage simultané de l'ARN messager sur le ribosome. Cette translocation est catalysée par un facteur d'élongation (EF-G chez les bactéries, EF-2 chez les archées et les eucaryotes) et s'accompagne de l'hydrolyse d'une molécule de GTP. Le site A est ainsi libéré et une nouvelle étape d'élongation peut commencer. L'ARNt expulsé au niveau du site E peut ensuite être à nouveau aminoacylé par une aminoacyl-ARNt synthétase et participer à un autre cycle d'élongation. Les diverses étapes de ce cycle d'élongation sont assez rapides pour qu'un seul ribosome d'une cellule eucaryote puisse ajouter 3 à 8 acides aminés par seconde à la chaîne polypeptidique en cours de synthèse et pour qu'un ribosome procaryote en ajoute jusqu'à 15 à 20 par seconde (Lodish and Jacobsen, 1972 ; Palmiter, 1974).

La traduction se termine normalement quand un codon de terminaison arrive au niveau du site A où il est reconnu par un facteur de terminaison de classe I (RF1 ou RF2 chez les bactéries, e/aRF1 chez les eucaryotes et les archées). Ces

protéines miment un ARNt aminoacylé (Nissen *et al.*, 2000). Elles provoquent l'hydrolyse du peptidyl-ARNt accroché au site P et, donc, la libération du polypeptide néo-synthétisé. Des facteurs de terminaison de classe II semblent aussi jouer un rôle dans la terminaison, sans reconnaître le codon stop (Freistroffer *et al.*, 2000) mais en interagissant avec les facteurs de classe I. Chez les bactéries, RF3 favorise ainsi la libération des facteurs de classe I (Freistroffer *et al.*, 1997). Il n'est cependant pas indispensable à la survie des bactéries. Chez les eucaryotes, le facteur eRF3 qui, lui, est essentiel, se lie à eRF1 pour permettre une hydrolyse rapide du peptidyl-ARNt (Frolova *et al.*, 1996 ; Zhouravleva *et al.*, 1995).

D'autres facteurs interviennent ensuite pour recycler le ribosome. Chez les bactéries, RRF (pour *Ribosome Recycling Factor* aussi connu sous le nom de RF4) et EF-G vont provoquer la dissociation des deux sous-unités du ribosome, puis, RF3 va agir pour libérer l'ARNt désacylé encore présent dans le site P (Kaji and Hirokawa, 2000). Bien que RRF soit universellement conservé chez les procaryotes et dans les organelles eucaryotes, aucun homologue de ce facteur n'est présent dans le cytoplasme des eucaryotes et des archées. Chez les eucaryotes, la dissociation des sous-unités ribosomales est catalysée par le facteur d'initiation eIF3, dont l'action est stimulée par les facteurs eIF3j (une sous-unité faiblement liée à eIF3), eIF1 et eIF1A (Pisarev *et al.*, 2007). Ensuite, eIF1 intervient pour éjecter l'ARNt désacylé, et eIF3j permet la dissociation de l'ARN messenger.

Au cours de ces diverses étapes, des erreurs peuvent se produire. Elles pourraient handicaper fortement la croissance cellulaire s'il n'existait pas également, comme nous allons le voir maintenant, des mécanismes permettant de les corriger.



## 1.2 La correction d'erreurs dans la traduction

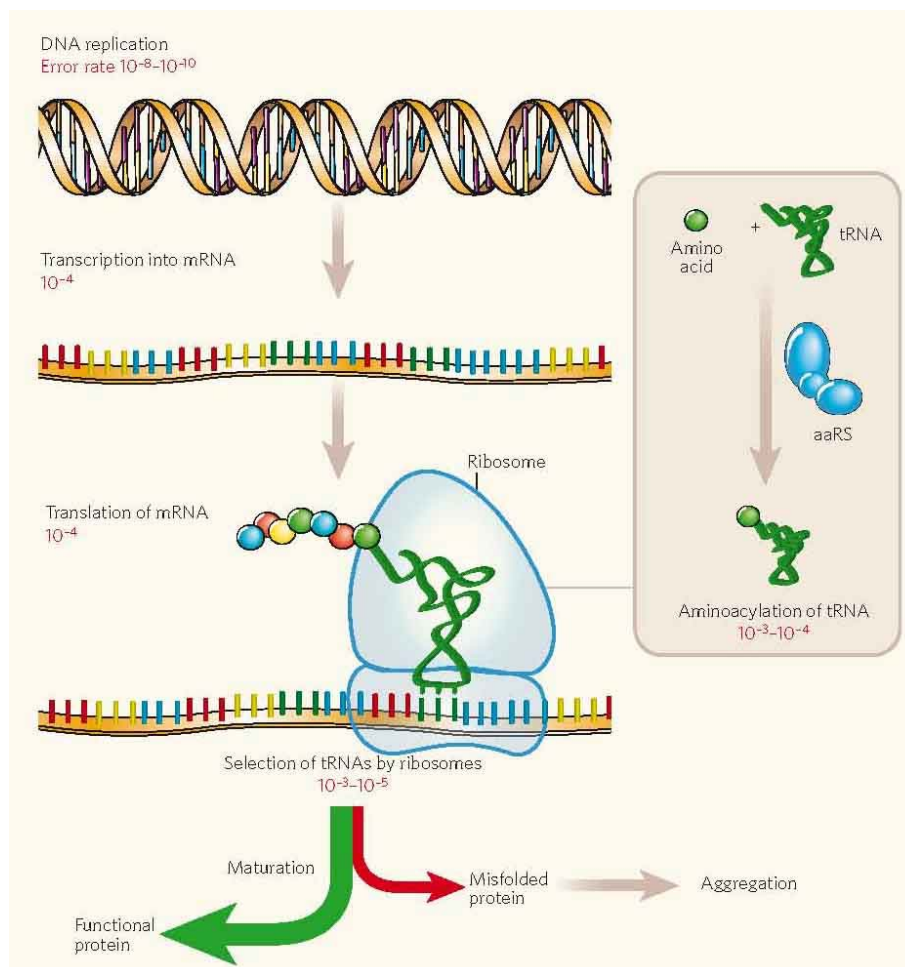
La vie sous la forme complexe que l'on connaît aurait été impossible sans une grande fidélité dans la réplication de l'ADN (environ une erreur sur 1 milliard de paires de bases répliquées (Roy and Ibbá, 2006)) et, donc, sans une grande précision dans la synthèse des protéines. En effet, l'information contenue dans l'ADN dicte la séquence des protéines et certaines de ces protéines maintiennent et dupliquent le matériel génétique. Notamment pour cette raison, la précision avec laquelle le message génétique est converti en protéines fonctionnelles est un aspect critique de la traduction. En effet, une erreur dans la synthèse de protéines peut induire des erreurs dans la réplication de l'ADN, par exemple si la protéine erronée est impliquée dans les mécanismes qui assurent la fidélité de la réplication de l'ADN. En conséquence, il va apparaître encore plus d'erreurs dans la synthèse protéique, d'où encore plus d'erreurs dans la synthèse de l'ADN. Un tel cycle catastrophique a été décrit par Orgel dès 1963, lorsqu'il étudiait la fidélité de l'ADN polymérase alpha (Orgel, 1963, 1973). Pour garantir à la synthèse protéique une fidélité suffisante, plusieurs mécanismes sont apparus au cours de l'évolution, permettant de corriger des erreurs à différentes étapes de la traduction.

Le taux moyen d'erreurs d'incorporation d'acide aminé dans les protéines est de un sur 1000 à un sur 10000 dans des conditions normales de croissance (Bouadloun *et al.*, 1983 ; Ibbá and Söll, 1999 ; Loftfield and Vanderjagt, 1972). Cette valeur cumule les erreurs dans la transcription de l'ADN en ARNm ( $\sim 10^{-4}$ ), dans la synthèse des aa-ARNt ( $\sim 10^{-4}$ ) et dans la traduction de l'ARNm par le ribosome ( $\sim 10^{-4}$ ) (Ibbá and Söll, 1999 ; Ogle and Ramakrishnan, 2005 ; Roy and Ibbá, 2006). Ce taux d'erreurs fait que la plupart des protéines produites sont pleinement fonctionnelles, d'autant plus que de nombreuses mutations ponctuelles n'altèrent pas la structure et la fonction de la protéine.

Comme les taux d'erreurs à chacune des étapes mentionnées ci-dessus (et résumées dans la figure 4) sont similaires et additifs, une augmentation du nombre d'erreurs à n'importe laquelle de ces étapes diminuerait la fidélité de toute la synthèse protéique. Une telle augmentation sera généralement néfaste pour la cellule, même si certains organismes peuvent augmenter volontairement ce taux d'erreurs pour résister à des conditions extrêmes. C'est le cas par exemple de cellules de mammifères où, lors d'un stress oxydant intense, le taux de méSAMinoacylation des ARNt par de la méthionine augmente d'un facteur 10 (Netzer *et al.*, 2009). Il en résulte une sur-incorporation de méthionine dans les protéines qui permettrait à ces cellules de mieux résister au stress oxydant grâce au pouvoir anti-oxydant de cet acide aminé puisque qui peut inactiver un agent oxydant en étant oxydé en sulfoxyde de méthionine puis réduit ensuite par une réductase (Luo and Levine, 2009).

Au niveau du ribosome, l'anticodon de chaque aa-ARNt est apparié avec le codon approprié de l'ARNm, positionnant l'acide aminé au niveau du centre

peptidyl-transférase du ribosome (Green and Noller, 1997). Acceptant tous les ARNt et transférant les 20 acides aminés (et même certains acides aminés non naturels (Giegé, 2003)), le ribosome contrôle l'appariement codon-anticodon, mais ne vérifie pas la bonne aminoacylation de l'ARNt. La synthèse d'un ARNt incorrectement aminoacylé peut donc conduire à une erreur d'incorporation d'acide aminé, ce qui fait que l'étape d'aminocyclation des ARNt doit être très précise.

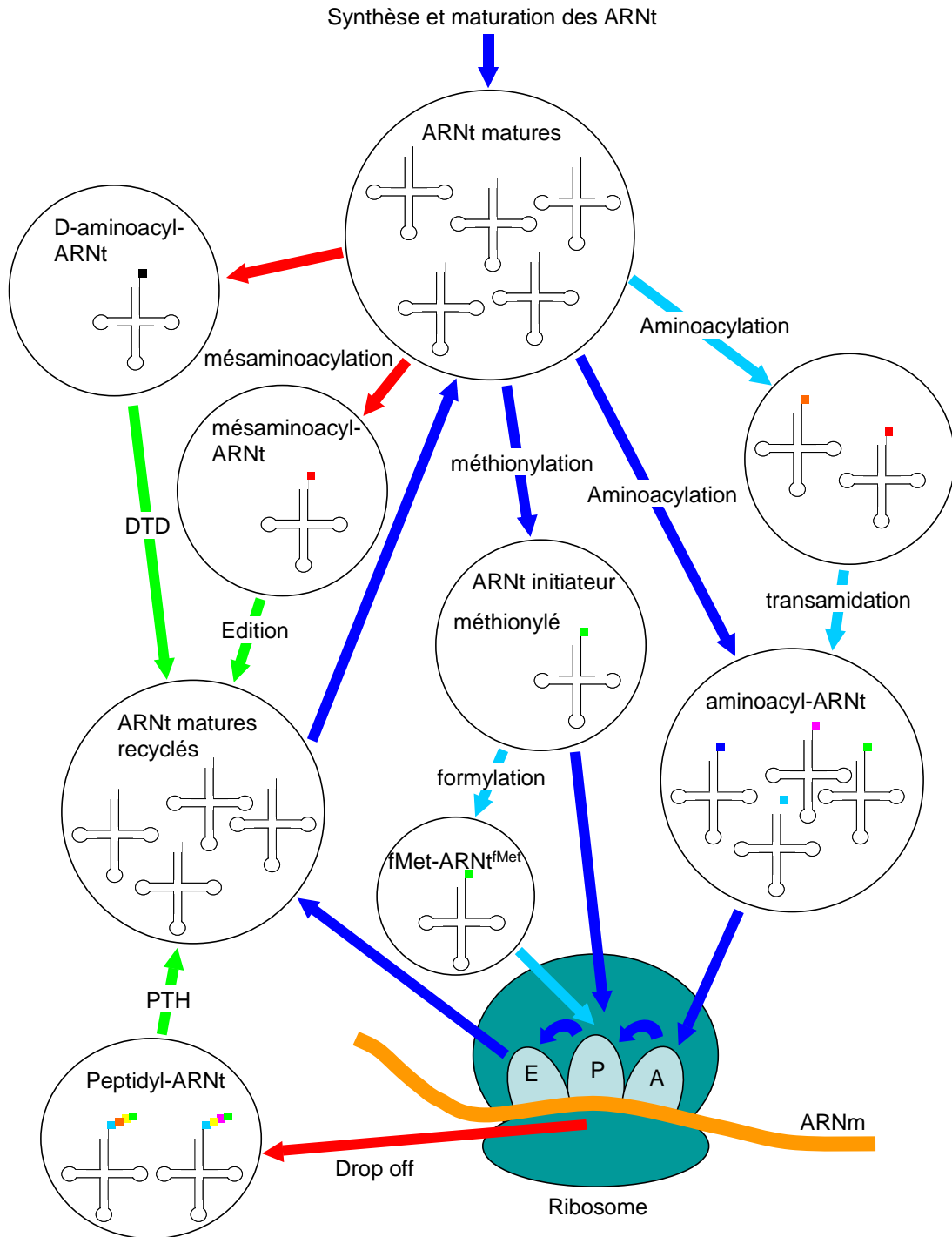


#### Figure 4 : Fidélité dans l'expression génétique

D'après Roy et Ibba (2006). Les différentes étapes de l'expression d'un gène et leur taux d'erreur sont représentés sur cette figure. Le transfert d'information de la séquence d'ADN d'un gène à la séquence protéique correspondante nécessite la transcription de cette séquence en ARN messager, puis, au niveau du ribosome, la traduction de cet ARNm en une séquence d'acides aminés, grâce à des ARN de transfert spécifiquement aminoacylés. Chacune de ces étapes est réalisée avec un taux d'erreurs d'environ  $10^{-4}$ . En cas d'erreur, les protéines peuvent adopter un repliement non fonctionnel. Elles seront alors généralement dégradées par la cellule ou peuvent conduire à un agrégat nocif.

En plus d'erreurs ponctuelles d'incorporation d'un acide aminé, il peut se produire sur le ribosome des erreurs de décalage du cadre de lecture qui conduisent à une protéine totalement différente de celle prévue par le code génétique. Comme nous le discuterons ci-après, il peut arriver également que la biosynthèse d'une protéine s'interrompe par la libération prématurée du peptidyl-ARNt lié au ribosome. Ce phénomène peut conduire à une séquestration des ARNt sous forme de peptidyl-ARNt non métabolisables et, finalement, à un blocage de toute la synthèse protéique.

Pour diverses raisons, il est donc important pour une cellule d'éviter que les ARNt ne se retrouvent estérifiés en 3' avec autre chose que l'acide aminé pour lequel ils sont prévus. Le cycle des ARNt, présenté sur le schéma ci-dessous, mentionne certaines erreurs qui peuvent survenir (flèches rouges sur la figure 5) et qui seront évoquées dans les paragraphes suivants, tout comme les moyens mis en place pour les éviter ou les corriger (flèches vertes).



**Figure 5 : Erreurs et corrections dans le cycle des ARNt**

Les deux sous-unités du ribosome en cours de traduction sont représentées en turquoise avec les trois sites de fixation (A, P et E) des ARNt. Un ARN messager en cours de traduction est représenté en orange. Les ARNt, sous différentes formes, sont regroupés dans des cercles reliés par des flèches. Les flèches en bleu foncé représentent les différentes étapes du cycle standard des ARNt. Les flèches en bleu clair représentent des voies alternatives pour la formation de certains aminoacyl-ARNt, présentes dans certains organismes (formylation du Met-ARNt<sup>fMet</sup> chez les bactéries ou transamidation (voir ci-dessous)). Les flèches en rouge correspondent à des erreurs et les flèches en vert à des mécanismes de corrections d'erreurs qui seront présentés dans la suite de l'introduction : édition des ARNt mésaminoacylés (paragraphe 2), désacylation des D-aminoacyl-ARNt (DTD, paragraphe 3) et hydrolyse des peptidyl-ARNt (PTH, paragraphe 5) issus du *drop off* (paragraphe 4).



## **2. Discrimination des acides aminés par les aminoacyl-ARNt synthétases**



## 2.1 Spécificité des aminoacyl-ARNt synthétases

### 2.1.1 Aminoacylation par les aminoacyl-ARNt synthétases

L'aminacylation des ARNt est catalysée par les aminoacyl-ARNt synthétases (aaRS). Généralement, un organisme donné comprend une aaRS pour chacun des 20 acides aminés canoniques.

La synthèse des aminoacyl-ARNt (aa-ARNt) par ces enzymes se déroule en deux étapes. L'enzyme lie d'abord une molécule d'ATP et une molécule d'acide aminé pour catalyser la formation d'un aminoacyl-adenylate (aa-AMP), qui reste lié à l'enzyme, avec libération d'un pyrophosphate (PPi) (Equation 1). Puis, l'acide aminé activé est transféré sur le groupement 2'- ou 3'-hydroxyle du ribose de l'adénosine 3'-terminale de l'ARNt, pour générer un aa-ARNt en libérant l'AMP (Equation 2) (Freist, 1989 ; Ibba and Söll, 2000).

Equation 1 : Acide aminé + ATP  $\rightarrow$  Aminoacyl-AMP + PPi

Equation 2 : Aminoacyl-AMP + ARNt  $\rightarrow$  Aminoacyl-ARNt + AMP

Les aaRS sont divisées en deux classes, se distinguant par leur repliement et leur domaine de liaison à l'ATP (Cusack *et al.*, 1991 ; Eriani *et al.*, 1990) (voir tableau 1).

Les synthétases de classe I sont presque toutes monomériques ( $\alpha$ ) ou dimériques ( $\alpha_2$ ). Leur site actif adopte une structure de type repliement de Rossmann, caractérisé par une série de brins  $\beta$  parallèles reliés entre eux par des hélices  $\alpha$ . Ces enzymes comportent des séquences peptidiques hautement conservées, HIGH (Ludmerer and Schimmel, 1987) et KMSKS (Hountondji *et al.*, 1986), qui interviennent dans la liaison de l'ATP et dans la stabilisation de l'état de transition, lors de l'activation de l'acide aminé. Les synthétases de classe II sont monomériques ( $\alpha$ ), dimériques ( $\alpha_2$ ) ou tétramériques ( $\alpha_4$  ou  $\alpha_2\beta_2$ ). Leur site actif est construit autour d'un feuillet  $\beta$  antiparallèle. Elles possèdent trois motifs peptidiques plus ou moins conservés, appelés motifs 1, 2 et 3 (Cusack *et al.*, 1990 ; Eriani *et al.*, 1995 ; Eriani *et al.*, 1990). Chaque classe est divisée en 3 sous-classes (a, b, c) en fonction des séquences primaires et de l'organisation structurale des aaRS qui les composent (Cusack, 1995 ; Landes *et al.*, 1995 ; Moras, 1992 ; Ribas de Pouplana and Schimmel, 2001).

Certains organismes ne possèdent pas le jeu complet des aaRS. C'est le cas, en particulier, d'un grand nombre de bactéries et de certaines archées, qui ne possèdent pas d'AsnRS et/ou de GlnRS et qui synthétisent l'Asn-ARNt<sup>Asn</sup> et le Gln-ARNt<sup>Gln</sup> en deux étapes (pour revue, Blanquet *et al.*, 2003). L'ARNt<sup>Asn</sup> (ou l'ARNt<sup>Gln</sup>) est d'abord aminoacylé par une AspRS (ou une GluRS) avec formation d'Asp-ARNt<sup>Asn</sup> (ou de Glu-ARNt<sup>Gln</sup>). Puis une amidotransférase catalyse une



réaction de transamidation à partir d'un donneur de groupement amide qui peut être la L-asparagine ou la L-glutamine. Cette réaction convertit l'acide aspartique (ou glutamique) lié à l'ARNt<sup>Asn</sup> (ou l'ARNt<sup>Gln</sup>) en asparagine (ou glutamine). On retrouve un tel mécanisme d'aminocyclation indirecte chez certaines archées pour la formation du Cys-ARNt<sup>Cys</sup>. Dans ce cas, il existe une phosphoséryl-ARNt synthétase (SepRS) qui produit du phosphoséryl-ARNt<sup>Cys</sup> avant que la phosphosérine liée à l'ARNt<sup>Cys</sup> ne soit convertie en cystéine par une enzyme appelée SepCysS (Sauerwald *et al.*, 2005). Une voie d'aminocyclation indirecte est aussi présente dans les organismes (des trois règnes du vivant) utilisant la sélénocystéine. L'ARNt<sup>Sec</sup> est d'abord aminoacylé par la SerRS avec une sérine. Puis cette sérine liée à l'ARNt<sup>Sec</sup> est transformée en sélénocystéine par une enzyme qui utilise le sélénophosphate comme donneur de sélénium. On peut noter que les aminoacyl-ARNt issus de la première étape de ces mécanismes sont peu reconnus par les facteurs d'élongation, ce qui empêche qu'ils soient conduits au ribosome avant d'être correctement aminoacylés.

Classe I		Classe II	
Ia	ArgRS $\alpha$ MetRS $\alpha, \alpha_2$ IleRS $\alpha$ ValRS $\alpha$ LeuRS $\alpha, \alpha\beta$ CysRS $\alpha, \alpha_2$	Iia	SerRS $\alpha_2$ ThrRS $\alpha_2$ ProRS $\alpha_2^*$ HisRS $\alpha_2$ GlyRS $\alpha_2$
Ib	GluRS $\alpha^*$ GlnRS $\alpha$ LysRS $\alpha^{**}$	Iib	AspRS $\alpha_2$ AsnRS $\alpha_2$ LysRS $\alpha_2^{**}$
Ic	TrpRS $\alpha_2$ TyrRS $\alpha_2$	Iic	PheRS $\alpha, \alpha_2, \alpha_2\beta_2$ GlyRS $\alpha_2\beta_2$ AlaRS $\alpha, \alpha_2, \alpha_4$ (SepRS <sup>***</sup> $\alpha_4$ ; PylRS <sup>****</sup> $\alpha_2$ )

**Tableau 1 : Répartition des aminoacyl-ARNt synthétases**

D'après Söll et Ibba (2003) et Nozawa *et al.* (2009). L'organisation oligomérique de chaque aaRS est indiquée. Certaines existent sous forme de monomères ( $\alpha$ ), d'homodimères ( $\alpha_2$ ), d'hétérodimères ( $\alpha\beta$ ) ou de tetramères ( $\alpha_2\beta_2$  ou  $\alpha_4$ ).

\* Chez les eucaryotes supérieurs, un polypeptide unique porte les activités GluRS et ProRS (Cerini *et al.*, 1991).

\*\* Il existe deux familles de LysRS. Les eucaryotes, la plupart des bactéries et quelques archées expriment une LysRS de classe II, tandis que la plupart des archées et quelques bactéries expriment une LysRS très différente, appartenant à la classe I. Certains organismes (par exemple, l'archée *Methanosarcina barkeri*) produisent à la fois une LysRS de classe I et une LysRS de classe II.

\*\*\* Certaines archées (23 sur les 69 entièrement séquencées) ne produisent pas de cystéinyl-ARNt synthétases. Elles possèdent une phosphoséryl-ARNt synthétase, (SepRS), qui estérifie l'ARNt<sup>Cys</sup> avec de la phosphosérine. Le phosphoséryl-ARNt<sup>Cys</sup> est ensuite converti en Cys-ARNt<sup>Cys</sup> par une enzyme appelée SepCysS (Sauerwald *et al.*, 2005).

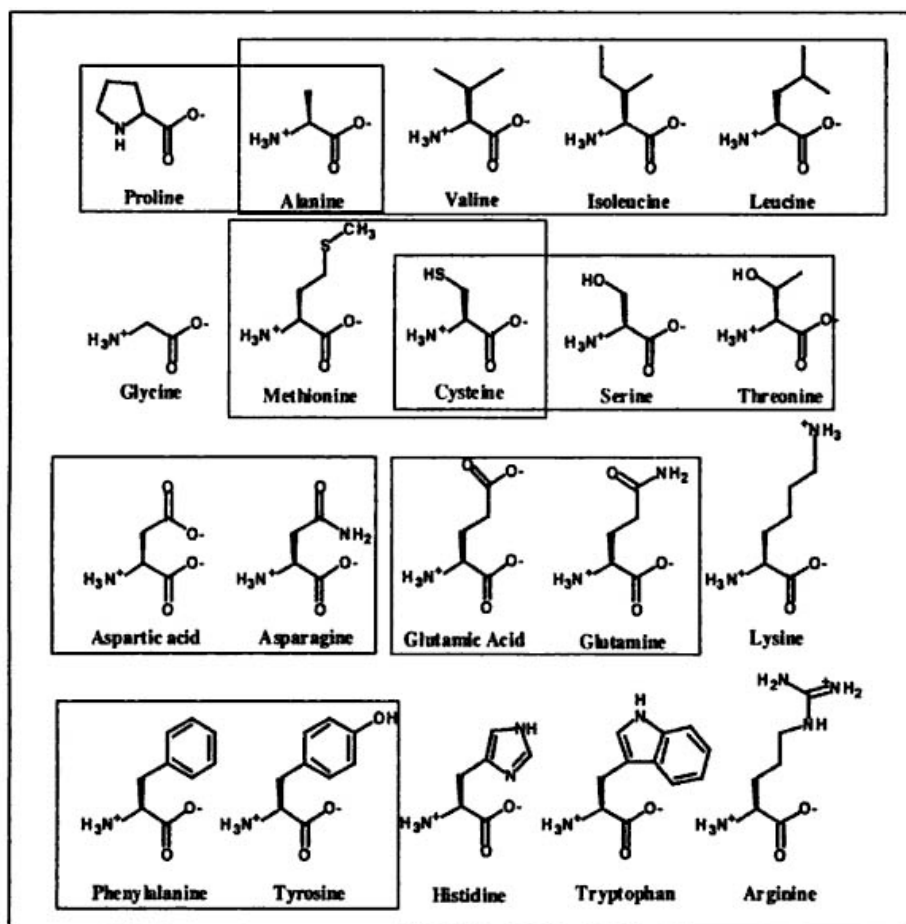
\*\*\*\* Quelques archées (4 sur les 69 entièrement séquencées) et quelques bactéries (par exemple *Desulfotobacterium hafniense*) insèrent de façon co-traductionnelle de la pyrrolysine dans leurs protéines. Elles possèdent pour cela, un ARNt dont l'anticodon est complémentaire du codon ambre UAG, et qui est aminoacylé par une aaRS particulière appelée pyrrolysyl-ARNt synthétase (PylRS) (Srinivasan *et al.*, 2002).

## 2.1.2 Spécificité des aminoacyl-ARNt synthétases

Les aminoacyl-ARNt synthétases ont la difficile tâche de sélectionner leurs ARNt spécifiques et leur acide aminé parmi un ensemble de composés dont certains sont très similaires.

La sélection de l'ARNt est facilitée par la grande taille de ces molécules qui fait que différentes parties de ces molécules peuvent être utilisées comme "éléments d'identité", c'est-à-dire comme déterminants structuraux assurant la reconnaissance par l'aaRS spécifique et/ou empêchant la reconnaissance par les aaRS non spécifiques. Les éléments d'identité peuvent ainsi être situés dans différentes régions de l'ARNt comme le bras accepteur, le bras anticodon ou la boucle variable (Giegé *et al.*, 1998).

Les acides aminés sont beaucoup plus petits et ne peuvent être discriminés que par la nature de leur chaîne latérale.



**Figure 6 : Structure des 20 acides aminés**

D'après Giegé et Frugier (2003). Les 20 acides aminés protéinogènes sont regroupés dans des cases selon leurs ressemblances structurales. Ces ressemblances posent un challenge pour leur reconnaissance moléculaire par les aminoacyl-ARNt synthétases. Les acides aminés, plus petits que celui qui correspond à une aaRS particulière, posent aussi un problème aux aaRS puisque ceux-ci peuvent se fixer dans le site actif de l'enzyme sans qu'il y ait de conflit stérique.

Or, comme on peut le voir sur la figure 6, certains acides aminés ont des structures très proches, rendant difficile leur reconnaissance à un niveau moléculaire. Par exemple, la thréonyl-ARNt synthétase doit distinguer la thréonine de la valine, qui lui est isostérique, et de la sérine, dont la seule différence par rapport à la thréonine est l'absence d'un groupe méthyle.

Certaines propriétés chimiques spécifiques permettent parfois de différencier des acides aminés proches (Fersht and Dingwall, 1979a). Ainsi, la CysRS comporte un ion zinc dans son site actif qui lie spécifiquement la charge négative du groupement thiolate de la cystéine et ne lie pas la sérine dont la structure est pourtant très proche de celle de la cystéine (Newberry *et al.*, 2002). Cependant, dans la plupart des cas, il est thermodynamiquement impossible d'obtenir un site actif permettant de distinguer efficacement des acides aminés proches (voir tableau 2).

aaRS	Compétiteurs naturels de l'acide aminé spécifique
Alanyl-ARNt synthétase	Glycine, sérine, cystéine, acide $\alpha$ -aminobutyrique
Arginyl-ARNt synthétase	Lysine
Asparagyl-ARNt synthétase	Acide aspartique, glutamine
Aspartyl-ARNt synthétase	Asparagine, acide glutamique
Cystéinyl-ARNt synthétase	Sérine, alanine, acide $\alpha$ -aminobutyrique
Glutamyl-ARNt synthétase	Glutamine, acide aspartique
Glutaminyl-ARNt synthétase	Acide glutamique, asparagine
Glycyl-ARNt synthétase	Aucun
Histidyl-ARNt synthétase	Aucun
Isoleucyl-ARNt synthétase	Valine, leucine, alanine, thréonine, O-méthylthréonine
Leucyl-ARNt synthétase	Valine, alanine, norvaline, méthionine, isoleucine
Lysyl-ARNt synthétase	Méthionine, ornithine
Méthionyl-ARNt synthétase	Homocystéine
Phénylalanyl-ARNt synthétase	Tyrosine
Prolyl-ARNt synthétase	Alanine, cystéine
Séryl-ARNt synthétase	Cystéine, acide $\alpha$ -aminobutyrique
Thréonyl-ARNt synthétase	Valine, sérine, cystéine, acide $\alpha$ -aminobutyrique
Tryptophanyl-ARNt synthétase	Aucun
Tyrosyl-ARNt synthétase	Phénylalanine
Valyl-ARNt synthétase	Thréonine, sérine, alanine, acide $\alpha$ -aminobutyrique

**Tableau 2 : Exemple de compétiteurs naturels des acides aminés**

D'après Hendrickson et Schimmel (2003). Pour chaque aminoacyl-ARNt synthétase (colonne de gauche), les compétiteurs naturels de l'acide aminé spécifique de cette synthétase sont reportés dans la colonne de droite.

Prenons l'exemple de l'isoleucine et la valine. La chaîne isopropyle de la valine diffère de la chaîne isobutyle de l'isoleucine par l'absence d'un groupement méthyle. Comme les gênes stériques peuvent mettre en jeu des énergies considérables, il est facile d'imaginer des géométries empêchant efficacement que l'isoleucine, plus grosse que la valine, vienne se fixer dans un site actif prévu pour la valine. Par contre, la valine, qui est plus petite que l'isoleucine, peut se loger dans tout site actif conçu pour reconnaître l'isoleucine. A la fin des années 1950, Linus Pauling avait ainsi estimé que la différence d'énergie libre entre le

complexe formé avec l'isoleucine et le complexe formé avec la valine ne pouvait excéder 1 à 2 kcal/mol (Pauling, 1957), ce qui correspond à un rapport entre les constantes d'association de l'isoleucine et de la valine inférieur à 100. Ce rapport est très inférieur au taux de 1/3000 mesuré *in vivo* au niveau de la traduction des codons de l'isoleucine (Loftfield, 1963 ; Loftfield and Vanderjagt, 1972). Il doit donc exister des mécanismes, en aval de la simple reconnaissance de l'acide aminé, pour expliquer le haut taux de spécificité assuré par les synthétases dans la réaction d'aminacylation de leurs ARNt.

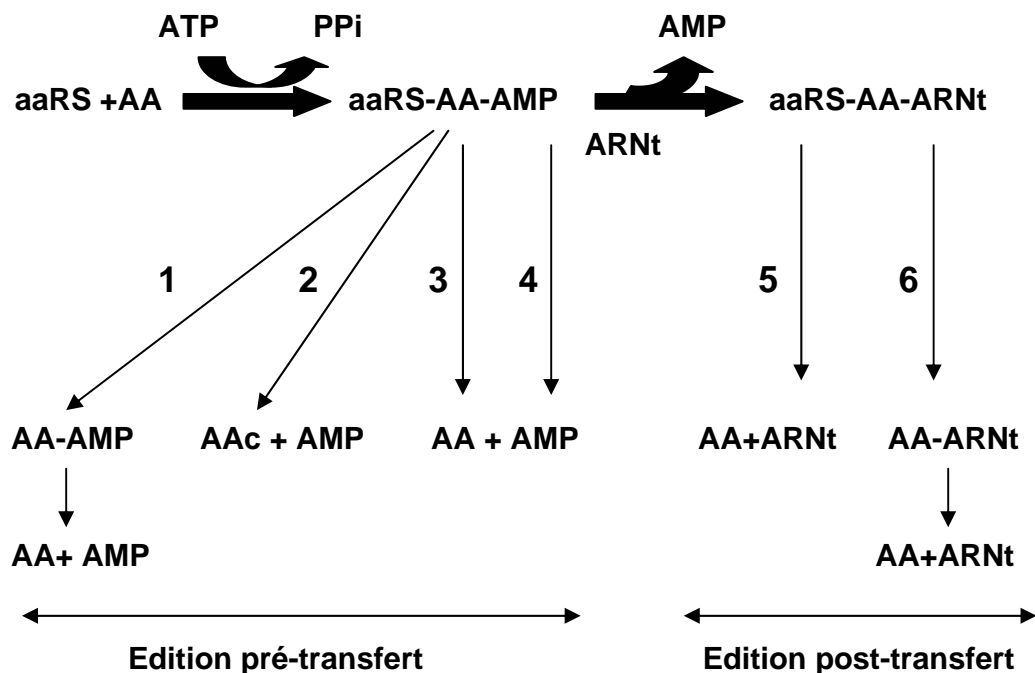
### 2.1.3 L'édition par les aminoacyl-ARNt synthétases

Conformément aux prédictions de Pauling, il a été démontré expérimentalement que l'IleRS pouvait fixer la valine et catalyser la formation de Val-AMP (Baldwin and Berg, 1966). Dans les années 1970, il a également été proposé qu'après l'activation de l'acide aminé par une synthétase, un mécanisme de correction pouvait intervenir, appelé relecture (*proofreading* en anglais) ou édition (Hopfield, 1974 ; Ninio, 1975). Ce mécanisme est basé sur l'existence d'une étape irréversible dans l'amination de l'ARNt : l'hydrolyse de l'ATP lors de l'activation de l'acide aminé suivie, *in vivo*, par l'hydrolyse du pyrophosphate par la pyrophosphatase. Cette étape irréversible permet une deuxième sélection, laquelle s'effectuerait par hydrolyse enzymatique soit de l'aminocyl-adénylate, soit de l'aminocyl-ARNt, l'acide aminé incorrect étant préférentiellement éliminé. Ce mécanisme explique pourquoi la formation de Val-AMP par l'IleRS ne conduit que rarement à la formation de Val-ARNt<sup>lle</sup>.

L'édition est un mécanisme qui consomme de l'énergie puisque, si on est en présence d'un acide aminé incorrect, l'hydrolyse de la molécule d'ATP ayant servi à la formation de l'aminocyl-adénylate ne sera généralement pas suivie par la synthèse d'une molécule d'aminocyl-ARNt. C'est d'ailleurs cette hydrolyse d'ATP en présence d'un acide aminé incorrect qui a permis de repérer expérimentalement l'existence de ces réactions d'édition.

La résolution des structures de nombreuses aaRS, parfois en présence d'analogues de substrats, a permis de préciser les bases moléculaires qui permettent aux aaRS de catalyser ces réactions d'édition. Comme on va le voir ci-dessous, elles ont souvent lieu au niveau d'un site catalytique porté par un domaine de la protéine distinct du domaine où s'effectuent l'activation de l'acide aminé et le transfert de celui-ci sur l'ARNt. Dans certains cas, il existe aussi des protéines isolées qui possèdent une activité d'édition. Il est d'ailleurs vraisemblable que les domaines d'édition des aaRS aient été acquis par celles-ci au cours de l'évolution, leur permettant de corriger elles-mêmes leurs erreurs juste après qu'elles se soient produites, plutôt que de faire appel à une enzyme de correction agissant en *trans*, en prenant le risque que des ARNt mésaminocylés soient pris en charge par le facteur d'élongation, EF-Tu ou a/eEF1a, et amenés au ribosome avant d'avoir été hydrolysés par cette enzyme de correction.

Il existe plusieurs mécanismes d'édition. D'une façon générale, lorsque c'est l'aminocyl-adénylate incorrect qui est éliminé, on parle d'édition "pré-transfert" tandis qu'on appelle édition "post-transfert" l'hydrolyse d'un aminocyl-ARNt incorrect. Dans l'édition pré-transfert, comme on peut le voir sur la figure 7, les aminocyl-adénylates incorrects peuvent être expulsés avant l'arrivée de l'ARNt (voie 1 sur la figure 7), ou hydrolysés en absence (voies 2 et 3) ou en présence de l'ARNt (voie 4). L'édition post-transfert des aa-ARNt peut avoir lieu en *cis* (voie 5), sur la synthétase comme dans le cas de l'IleRS, ou en *trans* (voie 6). Elle est alors catalysée par une protéine distincte de la synthétase.



**Figure 7 : Différentes voies d'édition par les synthétases**

Dans la première voie, illustrée par le cas de la ProRS (Hati *et al.*, 2006 ; Splan *et al.*, 2008), l'adénylate (AA-AMP) est relâché pour être ensuite hydrolysé en solution, sans action de la synthétase. La deuxième voie, qui ne nécessite pas la liaison d'ARNt, est illustrée par le cas de la cyclisation de l'homocystéinyl-adénylate par la MetRS avec formation d'homocystéinyl-thiolactone (acide aminé cyclique, AAc) qui est ensuite expulsée du site actif (Fersht and Dingwall, 1979b ; Jakubowski, 1991 ; Jakubowski and Fersht, 1981). Enfin, l'hydrolyse sélective du mauvais adénylate peut être catalysée par l'enzyme elle-même en présence (voie 3) ou absence d'ARNt (voie 4) comme dans le cas de la ValRS ou de la GlnRS (Fersht and Dingwall, 1979c ; Gruic-Sovulj *et al.*, 2005). Les voies 5 et 6, qui correspondent à une édition post-transfert en *cis* ou en *trans*, respectivement, seront étudiées en détail ci-après.

Actuellement, des mécanismes d'édition ont été décrits dans le cas de 9 des 20 (ou 21, voire 22, si on compte la SepRS et la PylRS) aminoacyl-ARNt synthétases (Hendrickson and Schimmel, 2003). La PheRS, la ThrRS et la ValRS possèdent une activité d'édition post-transfert alors que la LysRS de classe II possède une activité d'édition pré-transfert. D'autres synthétases possèdent à la fois une activité d'édition pré- et post-transfert (IleRS, LeuRS, MetRS, AlaRS, ProRS), mais généralement, une des deux prédomine. Ces deux mécanismes peuvent être indépendants ou non. Ainsi des mutations peuvent inactiver de manière sélective un seul des deux mécanismes, notamment dans l'IleRS (Hendrickson *et al.*, 2000) ou, dans le cas par exemple de la LeuRS, supprimer à la fois l'édition pré- et post-transfert (Lincecum *et al.*, 2003), ou même inactiver un des deux mécanismes et stimuler l'autre (Williams and Martinis, 2006).

Des mécanismes variés d'édition sont donc employés pour améliorer la fidélité de la synthèse des aminoacyl-ARNt. Nous allons maintenant les détailler pour les différentes aminoacyl-ARNt synthétases.

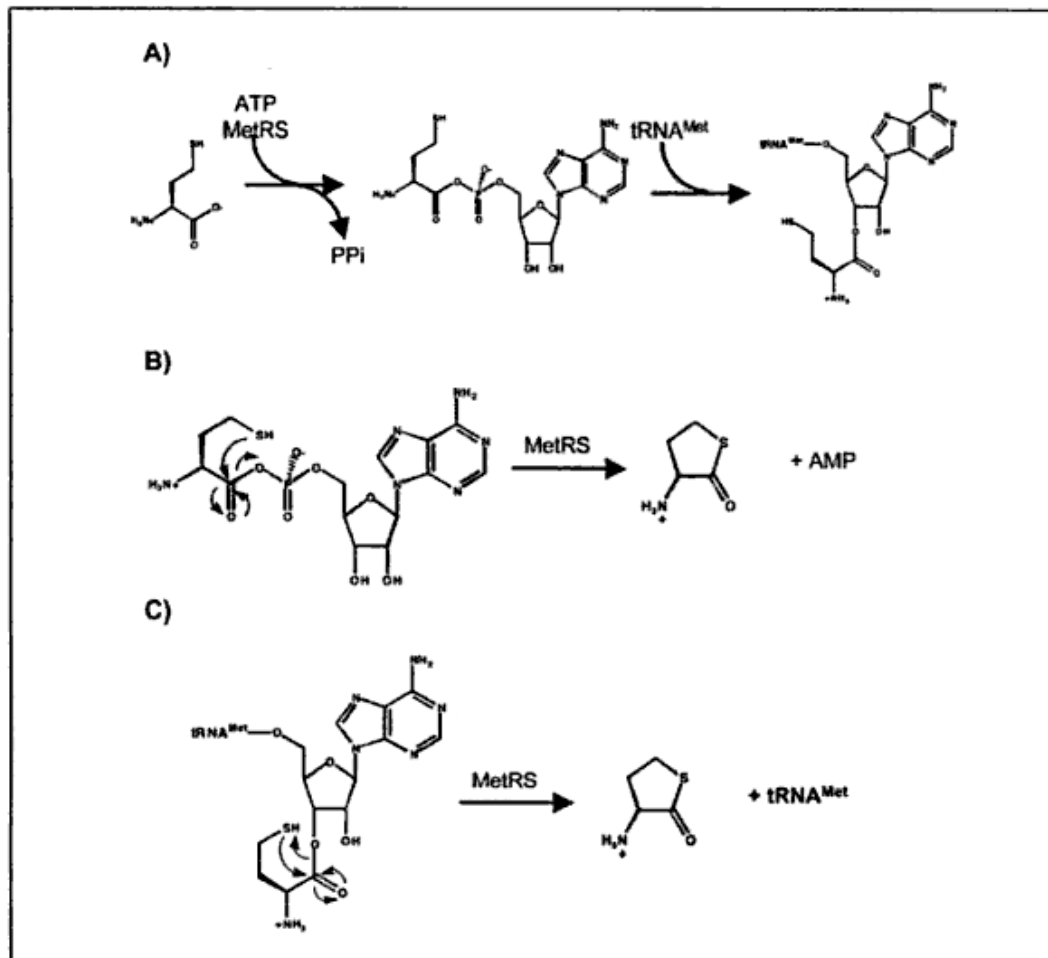
## 2.2 L'édition par les aminoacyl-ARNt synthétases de classe I

### 2.2.1 L'activité d'édition des synthétases de classe I

L'activité d'édition pré-transfert de la GlnRS a lieu dans le site actif de la synthétase (Gruic-Sovulj *et al.*, 2005). Les autres enzymes de classe I possédant une activité d'édition, LeuRS, IleRS, ValRS et MetRS, appartiennent toutes à la sous-classe Ia. Les LeuRS, IleRS et ValRS possèdent un domaine appelé CP1 (*connecting peptide 1*), qui est impliqué dans leur activité d'édition. En effet, des mutations dans cette région peuvent supprimer l'activité d'édition de ces synthétases (Schmidt and Schimmel, 1994, 1995). De plus, les domaines CP1 de plusieurs LeuRS, ValRS et IleRS ont été produits sous forme isolée, indépendamment du reste de l'enzyme. Ils sont alors toujours capables d'hydrolyser les ARNt spécifiques mésaminoacylés. Par contre, ils n'hydrolysent pas leurs ARNt correctement aminoacylés. Ainsi, le domaine CP1 de la LeuRS d'*Aquifex aeolicus* hydrolyse l'Ile-ARNt<sup>Leu</sup> mais pas le Leu-ARNt<sup>Leu</sup> (Zhao *et al.*, 2005), le domaine CP1 de la ValRS de *Bacillus stearothermophilus* hydrolyse le Thr-ARNt<sup>Val</sup>, mais pas le Val-ARNt<sup>Val</sup> et le domaine CP1 de l'IleRS d'*E. coli* hydrolyse le Val-ARNt<sup>Ile</sup> mais pas l'Ile-ARNt<sup>Ile</sup> (Lin and Schimmel, 1996). Seul le domaine CP1 de la LeuRS d'*E. coli* n'est pas capable d'hydrolyser l'Ile-ARNt<sup>Leu</sup> (Chen *et al.*, 2000), mais il devient actif si on lui adjoint les deux brins  $\beta$  (voir figure 9) qui le relient au domaine catalytique de l'enzyme (Betha *et al.*, 2007).

Par ailleurs, l'inactivation par mutagenèse du site d'édition (Schmidt and Schimmel, 1995), ainsi que des expériences de délétion de ce domaine, montrent qu'il n'est pas indispensable à l'activation de l'acide aminé ni à son transfert sur l'ARNt. Cela souligne l'indépendance des deux activités de ces synthétases et suggère que ces enzymes ont pu n'acquérir cette fonction d'édition que tardivement dans l'évolution.

La CysRS et la MetRS possèdent, elles aussi, un domaine CP1 mais celui-ci est plus petit que celui des LeuRS, ValRS et IleRS, et il ne fait pas d'édition. La CysRS ne possède aucune activité d'édition connue et, pour la MetRS, l'édition se fait au niveau du site actif (Fersht and Dingwall, 1979b ; Jakubowski, 1991 ; Jakubowski and Fersht, 1981). Elle concerne essentiellement la discrimination de l'homocystéine, un acide aminé faisant partie des voies de biosynthèse et de dégradation de la méthionine. L'édition conduit à la formation d'homocystéinylthiolactone qui est expulsée du site actif (voir figure 8).



**Figure 8 : Edition de l'homocystéine par la MetRS**

D'après Hendrickson et Schimmel (2003). Les différentes réactions représentent l'activation de l'homocystéine par la MetRS et son transfert sur l'ARNt<sup>Met</sup> (A), la cyclisation pré-transfert de l'homocystéinyl-adénylate (B) ou la cyclisation post-transfert de l'homocystéinyl-ARNt<sup>Met</sup> (C). L'édition de l'homocystéine par la MetRS conduit à la formation d'homocystéinyl-thiolactone.

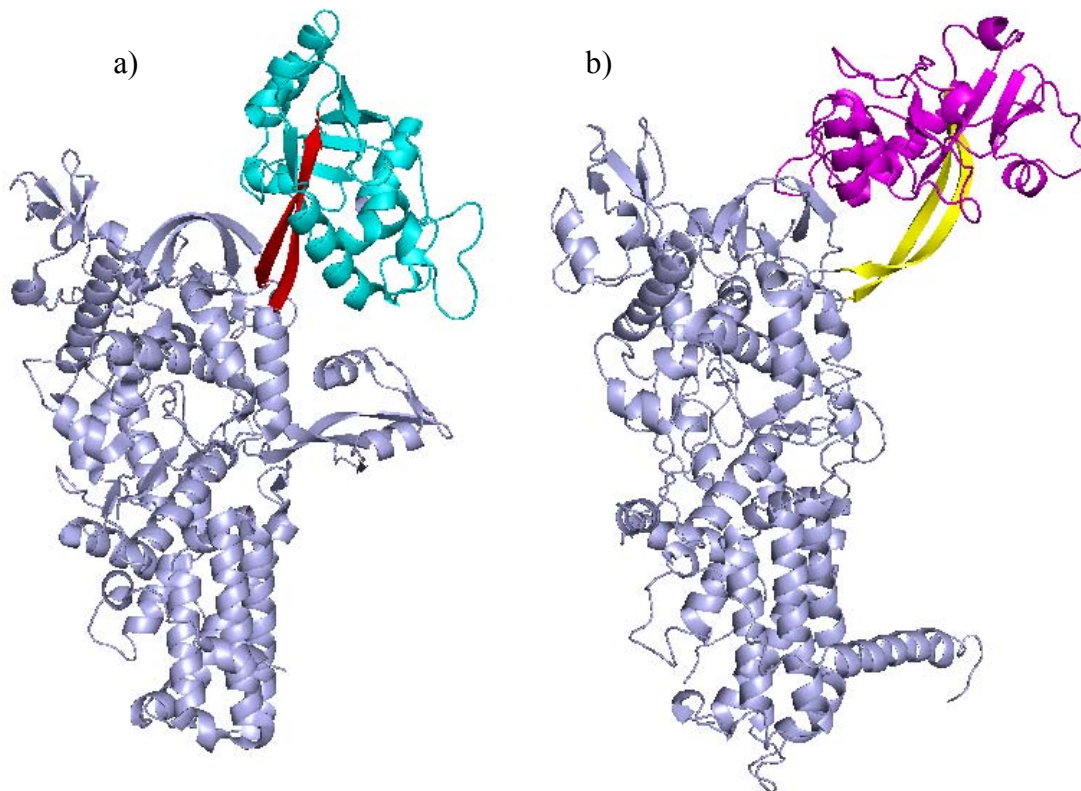
L'activité d'édition semble essentiellement post-transfert pour la ValRS et la LeuRS (Fersht and Kaethner, 1976 ; Jakubowski and Fersht, 1981). Pour l'IleRS, les premières études biochimiques ont mis en évidence un mécanisme d'édition pré-transfert (Fersht, 1977), mais des études plus récentes montrent qu'il existe aussi un mécanisme d'édition post-transfert (Dulic *et al.*, 2010 ; Hendrickson *et al.*, 2002 ; Lin and Schimmel, 1996).

Lors de l'édition post-transfert, la translocation entre le site d'amination et le site d'édition semble due à un mouvement de l'extrémité CCA de l'ARNt qui reste lié au corps de l'enzyme. Un ARNt ne peut donc pas être aminoacylé si un aminoacyl-ARNt est en train d'être édité. Ceci est à la base du pouvoir antifongique de certains composés comme l'AN2690 (5-fluoro-1,3-dihydro-1-hydroxy-2,1-benzoxaborole), qui inhibe la LeuRS en formant des liaisons covalentes entre le bore et les groupements 2'- et 3'-hydroxyle de l'adénosine 3'-terminale de l'ARNt<sup>Leu</sup>. Cette réaction bloque l'ARNt dans le site d'édition, rendant l'enzyme incapable de fixer un autre ARNt pour l'amination (Rock *et al.*, 2007).



### 2.2.2 Structure des domaines CP1

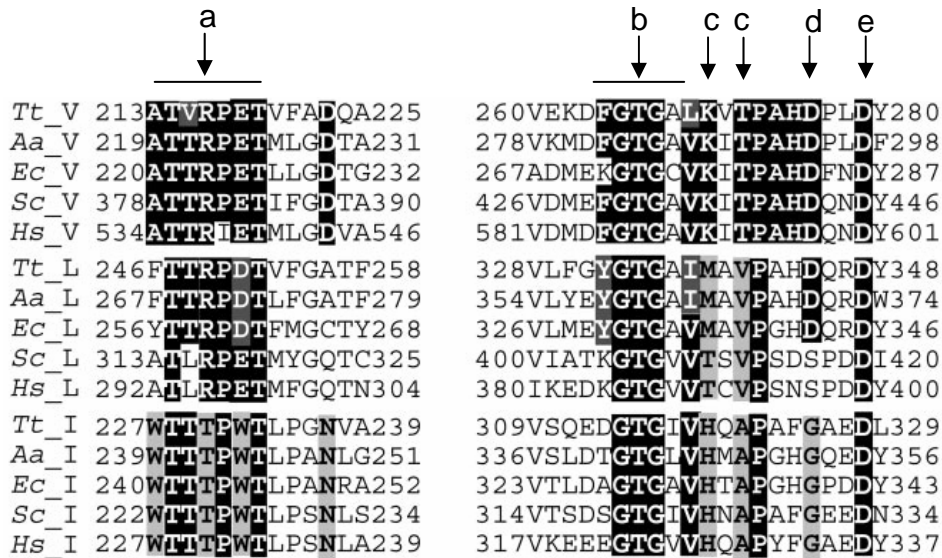
Les domaines CP1 des LeuRS, IleRS et ValRS ont des séquences et des structures homologues (Fukunaga and Yokoyama, 2005a). Le domaine CP1 est inséré dans le corps principal de l'enzyme et connecté à l'enzyme par 2 brins  $\beta$ , le site d'édition étant situé à près de 30 Å du site d'aminacylation. Les domaines CP1 des LeuRS bactériennes et mitochondriales ont un point d'insertion légèrement différent de celui des IleRS, ValRS et LeuRS archéennes (Cusack *et al.*, 2000). Il en résulte une différence entre l'orientation de ce domaine dans la LeuRS et dans l'IleRS, par exemple de *T. thermophilus* (voir figure 9), ce qui fait que, pour accéder aux sites d'édition, les extrémités 3' des ARNt doivent être dans des positions différant par une rotation de près de 180°.



**Figure 9 : Structures tridimensionnelles de la LeuRS et de l'IleRS de *T. thermophilus***

Les structures cristallographiques de la LeuRS (a) (code PDB 1OBC) et de l'IleRS (b) (code PDB 1ILE) sont représentées ci-dessus à l'aide du logiciel Pymol. Le corps principal des enzymes est représenté en gris et les domaines d'édition en cyan (a) et magenta (b), les deux brins  $\beta$  liant les domaines CP1 au reste de l'enzyme sont représentés en rouge (a) et jaune (b).

Les sites catalytiques des domaines d'édition des LeuRS, IleRS et ValRS possèdent aussi de nombreuses similarités, comme on peut le voir sur l'alignement ci-dessous.



**Figure 10 : Alignement des domaines CP1**

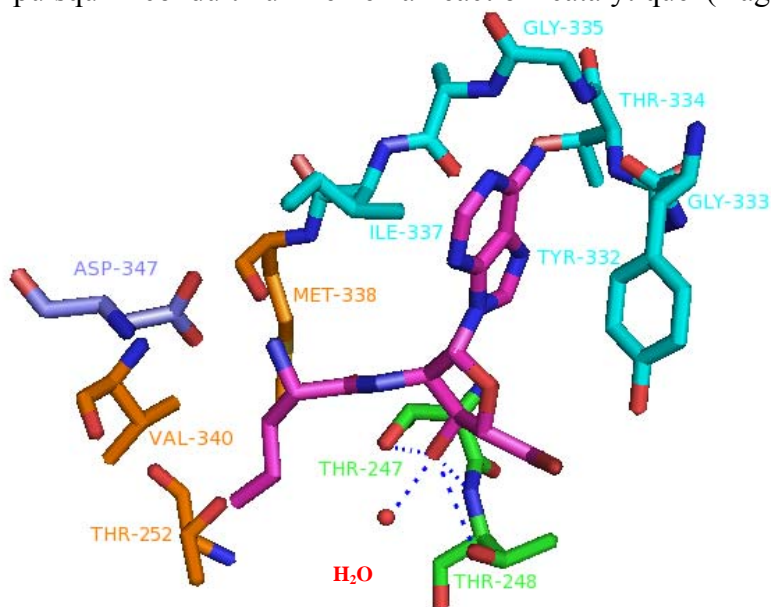
D'après Fukunaga et Yokoyama (2005b). Deux régions de la séquence des ValRS, LeuRS et IleRS, sont alignées pour différents organismes : *T. thermophilus*, *Aquifex aeolicus*, *E. coli*, *S. cerevisiae* et *Homo sapiens*. Les résidus indiqués possèdent des particularités détaillées dans le texte :

- Région riche en thréonines formant un réseau de liaisons hydrogène avec le substrat.
- Motif GTG et deux résidus en interaction avec l'adénine du substrat.
- Résidus impliqués dans la spécificité du site d'édition.
- Acide aspartique conservé dans les ValRS et certaines LeuRS. Il interagit avec le groupement hydroxyle du substrat.
- Acide aspartique catalytique universellement conservé.

Dans la région centrale, ces domaines possèdent un acide aspartique (figure 10.e) universellement conservé, qui forme un pont salin avec le groupement amino du substrat. La mutation de cet acide aspartique supprime l'activité d'édition de l'IleRS, de la ValRS et de la LeuRS (Dock-Bregeon *et al.*, 2004 ; Fukunaga *et al.*, 2004 ; Fukunaga and Yokoyama, 2005a ; Lincecum *et al.*, 2003).

Une région riche en thréonines (figure 10.a) est conservée dans ces trois enzymes. Des mutations simultanées de plusieurs de ces thréonines suppriment l'activité d'édition de l'IleRS et de la LeuRS de *T. thermophilus* (Fukunaga *et al.*, 2004 ; Zhai and Martinis, 2005). Dans la structure cristallographique du complexe entre la LeuRS de *T. thermophilus* et la 2'-(L-norvalyl)amino-2'-désoxyadénosine (un analogue du substrat de l'édition post-transfert), deux thréonines conservées de cette région (T247 et T248) participent à un réseau de liaisons hydrogène stabilisant l'état de transition de la réaction d'édition (Lincecum *et al.*, 2003 ; Zhai and Martinis, 2005).

Ce réseau de liaisons hydrogène implique également le groupement 3'-OH du ribose du ligand et le groupement carbonyle de l'acide aminé (Lincecum *et al.*, 2003). Ce groupement 3'-OH est aussi stabilisé dans la ValRS de *T. thermophilus* par une liaison hydrogène avec la chaîne latérale de l'Asp276, conservé dans les domaines CP1 des ValRS et de certaines LeuRS (Fukunaga and Yokoyama, 2005b)(figure 10.d). La stabilisation du groupement 3'-hydroxyle de l'ARNt par des liaisons hydrogène pourrait lui permettre de jouer un rôle dans la catalyse de la réaction, en lui donnant la possibilité d'activer une molécule d'eau au voisinage du site de coupure (voir figure 11). Ceci est d'autant plus vraisemblable que des expériences de mutagenèse dirigée systématique dans le site d'édition de la LeuRS n'ont pas permis d'identifier un résidu catalytique pouvant activer cette molécule d'eau (Nordin and Schimmel, 2002). Des calculs énergétiques ont d'ailleurs confirmé récemment cette hypothèse, qui confère donc à l'ARNt un rôle de ribozyme puisqu'il conduit lui-même la réaction catalytique (Hagiwara *et al.*, 2010).



**Figure 11 : Site d'édition de la LeuRS de *T. thermophilus***

La structure cristallographique du site d'édition de la LeuRS de *T. thermophilus* en complexe avec un analogue du substrat d'édition post-transfert, la 2'-(L-norvalyl)amino-2'-désoxyadénosine, est représentée ci-dessus à l'aide du logiciel Pymol et à partir du fichier PDB : 1OBC. L'analogue du substrat est représenté en magenta. Son groupement 3'-OH participe à un réseau de liaisons hydrogène (en pointillé) avec les thréonines 247 et 248, (représentées en vert), et permet la stabilisation d'une molécule d'eau dans le site d'édition. L'acide aspartique universellement conservé au sein des domaines CP1 (Asp347) est représenté en bleu. La tyrosine (Tyr332) et l'isoleucine (Ile337) prenant l'adénine du substrat en sandwich sont représentées en cyan ainsi que le motif GTG aux extrémités duquel elles se trouvent. La thréonine (Thr252), la valine (Val340) et la méthionine (Met338) jouant un rôle dans la reconnaissance spécifique du substrat sont représentées en orange.

Par ailleurs, une boucle comportant un motif GTG (figure 10.b), très conservé dans les domaines CP1, intervient pour positionner, dans le site d'édition, l'adénosine de l'extrémité 3' de l'ARNt mésaminoacylé (Fukunaga and Yokoyama, 2005b ; Lincecum *et al.*, 2003). Ainsi, dans la LeuRS de *T. thermophilus*, deux résidus voisins de cette boucle, la Tyr332 et l'Ile337, prennent

en sandwich l'adénine terminale de l'ARNt. Dans la ValRS de *T. thermophilus*, les résidus équivalents, la Phe264 et la Leu269, ont le même rôle. Les interactions sont possibles grâce aux deux glycines du motifs GTG, qui adoptent des conformations que ne pourraient prendre les autres acides aminés (Lincecum *et al.*, 2003). Dans le cas de l'IleRS de *T. thermophilus*, l'adénine interagit via des liaisons de van der Waals avec la Val318, qui est le résidu équivalent à l'Ile337 de la LeuRS et à la Leu269 de la ValRS. Dans l'IleRS, il n'y a pas de résidu aromatique comparable à la Tyr332 et à la Phe264 des LeuRS et ValRS. Par contre, il existe des liaisons hydrogène entre les atomes N1 et N6 de la base et différents groupements du squelette peptidique (Fukunaga and Yokoyama, 2006).

D'autres résidus, conservés différemment entre les synthétases, contribuent à la spécificité de celles-ci envers les divers acides aminés. Ainsi, la Lys270 et la Thr272 (figure 10.c) de la ValRS de *T. thermophilus*, strictement conservées dans les ValRS mais pas dans les LeuRS ou les IleRS, interagissent avec le groupement hydroxyle de la chaîne latérale de la thréonine éditée, mais pas avec la valine qui ne possède pas de groupement hydroxyle (Fukunaga and Yokoyama, 2005b). La mutation de cette lysine chez *E. coli* (Lys277) rend possible l'hydrolyse du Val-ARNt<sup>Val</sup> (Hountondji *et al.*, 2002), ce qui indique que cette lysine sert aussi à repousser la valine et sa chaîne latérale hydrophobe.

Dans les IleRS, on trouve systématiquement une histidine à la position occupée par la Lys270 dans la ValRS de *T. thermophilus*. Sur la structure de l'IleRS de *T. thermophilus* complexée avec la L-valine (Fukunaga *et al.*, 2004), on constate que cette histidine empêcherait, par gêne stérique, la fixation de l'isoleucine dans le site d'édition (Fukunaga *et al.*, 2004). En accord avec cette conclusion, la mutation de cette histidine chez *E. coli* rend l'IleRS capable d'hydrolyser l'Ile-ARNt<sup>Ile</sup> (Hendrickson *et al.*, 2002). Ce résultat a aussi été observé chez *T. thermophilus* pour cette histidine (His379), mais aussi pour une thréonine (Thr233) appartenant à la région riche en thréonine de l'IleRS, qui participe donc aussi à la spécificité de l'IleRS (Fukunaga and Yokoyama, 2006).

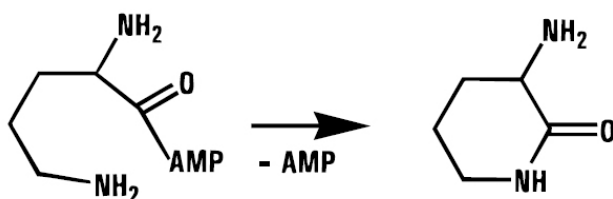
Dans la LeuRS de *T. thermophilus*, la Thr252 (conservée dans le motif riche en thréonine de tous les domaines CP1), la Met338 (spécifique de certaines LeuRS) et la Val340 (spécifique de toutes les LeuRS) forment une cavité trop petite pour accepter la leucine, mais dans laquelle les acides aminés édités (Ile et Met essentiellement) sont stabilisés par un réseau de liaisons hydrogène (Liu *et al.*, 2006 ; Zhai *et al.*, 2007).

Ainsi, certains résidus interagissent directement avec les substrats des sites d'édition tandis que d'autres empêchent l'acide aminé correct de se fixer dans ce site, souvent pour des raisons stériques. Ceci est à rapprocher du double criblage qui a été évoqué pour décrire la façon dont ces synthétases sélectionnent l'acide aminé qui leur correspond. Un premier criblage est effectué au niveau du site d'activation, qui élimine les acides aminés plus gros que l'acide aminé correct. Un deuxième criblage est effectué au niveau du site d'édition qui, au contraire, n'accepte que les acides aminés plus petits que l'acide aminé correct.

## 2.3 L'édition par les aminoacyl-ARNt synthétases de classe II

Si les enzymes de classe I possèdent un site d'édition assez similaire, les enzymes de classe II qui catalysent une réaction d'édition utilisent pour cela des domaines structurellement très différents.

Pour les enzymes de classe II, des études ont montré que la ProRS, la ThrRS, l'AlaRS et la PheRS possèdent des activités d'édition que nous allons détailler ci-dessous. La LysRS de classe II semble aussi avoir une activité d'édition, mais uniquement pré-transfert, via un mécanisme de cyclisation similaire à celui de la MetRS. Il a vraisemblablement lieu dans le site où se fait l'activation de l'acide aminé (Ataide and Ibba, 2004 ; Jakubowski, 1997, 1999). Il permet à la LysRS, qui est capable d'activer l'ornithine, d'éviter que celle-ci ne soit transférée sur l'ARNt<sup>Lys</sup>. Il semble qu'après activation de l'ornithine, le groupement NH<sub>3</sub><sup>+</sup> de la chaîne latérale de cet acide aminé soit capable d'attaquer le groupement carbonyle de l'ornithyl-adénylate, ce qui conduit à la formation d'un composé cyclique.



**Figure 12 : Edition pré-transfert de l'ornithine par la LysRS de classe II**

D'après Jakubowski (1999). L'édition pré-transfert de l'ornithyl-adénylate par la LysRS de classe II conduit, par cyclisation, à la formation d'ornithyl-lactame, en libérant l'AMP.

### 2.3.1 L'édition pour les ARNt<sup>Pro</sup>

#### 2.3.1.1 L'édition dans les ProRS

De nombreuses ProRS utilisent un mécanisme d'édition pré-transfert pour exclure l'alanine (Hati *et al.*, 2006). Cela a été démontré dans le cas de ProRS bactériennes (*E. coli*), archéennes (*M. jannaschii*) ou eucaryotes (*S. cerevisiae*). La réaction se fait selon la voie 1 présentée à la figure 7, c'est-à-dire par un départ de l'enzyme de l'alanyl-adénylate, qui est ensuite hydrolysé spontanément en solution.

De plus, un grand nombre de ProRS bactériennes possèdent une activité d'édition post-transfert qui contribue à écarter l'alanine (Ahel *et al.*, 2002 ; Beuning and Musier-Forsyth, 2000). Un domaine (INS) d'environ 180 acides aminés, identifié par mutagenèse dirigée, inséré entre les motifs 2 et 3 de l'enzyme est, en effet, capable d'hydrolyser l'Ala-ARNt<sup>Pro</sup> (Wong *et al.*, 2002 ; Wong *et al.*, 2003).

Les ProRS de certains eucaryotes inférieurs possèdent un domaine analogue à INS mais situé à l'extrémité N-terminale de la protéine. Ces domaines présentent des séquences assez variables d'un organisme à l'autre. Dans certains cas, comme chez *Plasmodium falciparum*, ce domaine confère à la ProRS, une activité d'hydrolyse de l'Ala-ARNt<sup>Pro</sup>. Dans d'autres cas, comme chez *S. cerevisiae*, la ProRS ne semble pas posséder cette activité malgré la présence de ce domaine d'insertion (SternJohn *et al.*, 2007). Les ProRS des archées et du cytoplasme des eucaryotes supérieurs ne possèdent ni ce domaine d'insertion ni cette activité d'édition. Ce domaine INS est aussi assez fréquemment absent dans les ProRS bactériennes puisque certaines bactéries (*Bacteroidetes*, groupe *Thermus-Deinococcus*, *Mollicutes*, certains spirochètes, certaines protéobactéries  $\delta$ , ...) expriment une ProRS de type archéen, dépourvue de domaine INS, et que quelques bactéries (protéobactéries  $\alpha$ , protéobactérie  $\epsilon$  *Sulfurimonas denitrificans*) expriment une ProRS de type bactérien, mais sans domaine INS. Enfin, on peut noter que certaines bactéries (par exemple *Chloroflexus aurantiacus* et *Clostridium botulinum*) expriment à la fois une ProRS de type bactérien avec un domaine INS et une ProRS de type archéen sans domaine INS.

Dans la ProRS d'*E. coli*, des expériences de mutagenèse dirigée ont permis d'identifier plusieurs résidus jouant un rôle important dans l'activité d'édition post-transfert de l'Ala-ARNt<sup>Pro</sup>. Ainsi, la mutation de la lysine 279 en alanine engendre une perte de 95% de l'activité d'hydrolyse de l'Ala-ARNt<sup>Pro</sup> par cette enzyme (Wong *et al.*, 2002). Les mutations T257A, D350A et H369A ont aussi un effet important (environ 80% de perte d'activité). Par ailleurs, l'histidine 369 semble en partie responsable de la spécificité du site d'édition, en empêchant probablement la proline de s'y fixer. En effet, les mutants H369A et H369C possèdent une forte activité d'hydrolyse du Pro-ARNt<sup>Pro</sup>.

### 2.3.1.2 Edition en *trans*

Dans les trois domaines du vivant, les ProRS activent la cystéine et la transfèrent sur l'ARNt<sup>Pro</sup> avec une efficacité catalytique *in vitro* proche de celle obtenue avec la proline. Cependant, les ProRS ne semblent pas posséder d'activité permettant l'édition spécifique du Cys-ARNt<sup>Pro</sup> (Ahel *et al.*, 2002 ; An and Musier-Forsyth, 2004 ; Beuning and Musier-Forsyth, 2001).

L'absence d'activité d'édition des ProRS, envers la cystéine pourrait donc poser un problème de fidélité, s'il n'existait pas d'autres mécanismes évitant la mésincorporation de cystéine. D'autre part, comme on vient de le voir, seulement certaines ProRS possèdent une activité d'édition post-transfert vis-à-vis de l'alanine. Or de nombreux organismes expriment des protéines homologues au domaine INS des ProRS bactériennes. La question s'est donc posée de savoir si ces protéines indépendantes n'avaient pas une activité d'édition leur permettant d'hydrolyser Cys-ARNt<sup>Pro</sup> et/ou les Ala-ARNt<sup>Pro</sup> (Zhang *et al.*, 2000).

### 2.3.1.3 ProX

Ainsi la bactérie *Clostridium sticklandii* exprime une ProRS de type archéen, sans domaine INS, et une protéine indépendante appelée ProX (ou PrdX), homologue au domaine INS. Il a été montré *in vitro* que cette protéine ProX était capable d'hydrolyser l'Ala-ARNt<sup>Pro</sup> mais pas le Cys-ARNt<sup>Pro</sup> (Ahel *et al.*, 2003). On peut donc imaginer que la présence de la protéine ProX permet de compenser l'absence d'activité post-transfert de la ProRS de *C. sticklandii* vis-à-vis de l'Ala-ARNt<sup>Pro</sup>, ce qui ferait de ProX un exemple de protéine autonome capable d'édition en *trans*.

La spécificité de la protéine ProX semble due, comme celle du domaine INS des ProRS bactériennes, à la présence d'une histidine conservée qui restreint l'accès du site actif aux petits acides aminés comme l'alanine. Cette histidine n'est pas alignée avec celle des domaines INS, mais une analyse structurale comparative entre le domaine INS de la ProRS d'*Enterococcus faecalis* et la protéine ProX de *Caulobacter crescentus* (Crepin *et al.*, 2006) montre que les deux histidines occupent des positions identiques dans les deux sites actifs, au voisinage de la lysine catalytique.

### 2.3.1.4 YbaK

Dans différentes bactéries et archées, il existe des protéines homologues au domaine INS, mais qui ne possèdent pas l'histidine conservée évoquée ci-dessus. La première protéine de cette famille à avoir été caractérisée est la protéine YbaK d'*Haemophilus influenzae* (Wong *et al.*, 2003). A l'origine, cette protéine a été décrite comme une hydrolase capable de désacyler l'Ala-ARNt<sup>Pro</sup> *in vitro*, mais pas le Pro-ARNt<sup>Pro</sup>. Plus tard, il a été montré que la protéine YbaK d'*H. influenzae* hydrolysait le Cys-ARNt<sup>Pro</sup> environ 100 fois plus efficacement que l'Ala-ARNt<sup>Pro</sup> (An and Musier-Forsyth, 2004). Mais les concentrations de protéine mises en jeu pour détecter de l'hydrolyse du Cys-ARNt<sup>Pro</sup> *in vitro* restaient élevées (~1 µM) et la spécificité de la protéine semblait faible puisque tous les Cys-ARNt testés, y compris le Cys-ARNt<sup>Cys</sup> (an2005 ruan2005), sont hydrolysés par YbaK *in vitro* (An and Musier-Forsyth, 2005 ; Ruan and Söll, 2005). Il n'était donc pas certain que YbaK intervienne bien *in vivo* pour hydrolyser le Cys-ARNt<sup>Pro</sup>.

Le rôle d'YbaK *in vivo* a pu être confirmé chez *E. coli*, grâce à l'utilisation d'une souche portant une mutation faux-sens Cys -> Pro dans le gène de la thymidylate synthase. Cette mutation affecte une cystéine essentielle de la protéine et rend donc la souche auxotrophe pour la thymine. Dans ce contexte génétique, l'inactivation du gène *ybaK* permet une croissance en absence de thymine, ce qui suggère qu'il y a alors insertion de cystéine en réponse à un codon proline et donc, en fin de compte, que YbaK est bien une Cys-ARNt<sup>Pro</sup> désacylase *in vivo* (Ruan and Söll, 2005).

Une autre expérience explique peut-être pourquoi la protéine semble peu spécifique *in vitro* alors que son rôle *in vivo* pourrait être d'hydrolyser spécifiquement les Cys-ARNt<sup>Pro</sup>. En effet, il a été montré que la protéine YbaK formait un complexe d'assez haute affinité avec la ProRS ( $K_d$  de 550 nM en absence d'ARNt<sup>Pro</sup> et de 45 nM en présence d'ARNt<sup>Pro</sup>) (An and Musier-Forsyth, 2005). Il se pourrait donc qu'*in vivo*, YbaK agisse sous une forme complexée à la ProRS, ce qui lui permettrait, d'une part, d'hydrolyser préférentiellement le Cys-ARNt<sup>Pro</sup> et, d'autre part, d'intervenir avant que cet ARN mésaminoacylé n'ait été pris en charge par EF-Tu.

Structuralement très proche du domaine INS des ProRS bactériennes, YbaK est une protéine d'environ 180 résidus, contenant un feuillet bêta de 7 brins entouré par six petites hélices alpha (Zhang *et al.*, 2000). On retrouve une lysine (Lys46) fortement conservée, à une position équivalente à celle de la Lys279 de la ProRS d'*E. coli* (Wong *et al.*, 2002), dans une poche comprenant un trou oxyanion similaire à celui des protéases à sérine. Cette lysine est indispensable à l'activité d'hydrolyse du Cys-ARNt<sup>Pro</sup> et de l'Ala-ARNt<sup>Pro</sup> (An and Musier-Forsyth, 2004).

### 2.3.1.5 Différentes stratégies d'édition pour les ARNt<sup>Pro</sup>

De manière surprenante, il n'existe pas de corrélation entre l'absence de domaine INS dans la ProRS d'un organisme donné et la présence de protéines du type ProX-YbaK. Même dans les bactéries phylogénétiquement proches, on rencontre des situations très diverses. Ainsi, parmi les actinobactéries (bactéries à Gram positif, à forte teneur en GC), on trouve des espèces qui expriment une ProRS avec un domaine INS et une (ou plusieurs) protéine(s) du type ProX-YbaK (par exemple, *Mycobacterium tuberculosis*), des espèces qui expriment une ProRS avec domaine INS et aucune protéine de type ProX-YbaK (par exemple *Mycobacterium avium*), des espèces qui expriment une ProRS sans domaine INS et une (ou plusieurs) protéine(s) du type ProX-YbaK (par exemple, *Rubrobacter xylanophilus*), et des espèces qui expriment une ProRS sans domaine INS et aucune protéine de type ProX-YbaK (par exemple *Frankia* sp. EAN1pec). Il reste donc des organismes où aucune protéine susceptible d'hydrolyser le Cys-ARNt<sup>Pro</sup> et/ou l'Ala-ARNt<sup>Pro</sup> n'est connue.

Il est possible que les concentrations intracellulaires de cystéine, d'alanine et de proline fassent que la mésaminoacylation de l'ARNt<sup>Pro</sup> reste faible et qu'un mécanisme d'édition soit inutile dans ces organismes. Ou alors, il existe d'autres protéines, encore inconnues, qui empêchent la formation de Cys-ARNt<sup>Pro</sup> et/ou d'Ala-ARNt<sup>Pro</sup>, ou qui catalysent l'hydrolyse de tels ARNt mésaminoacylés.



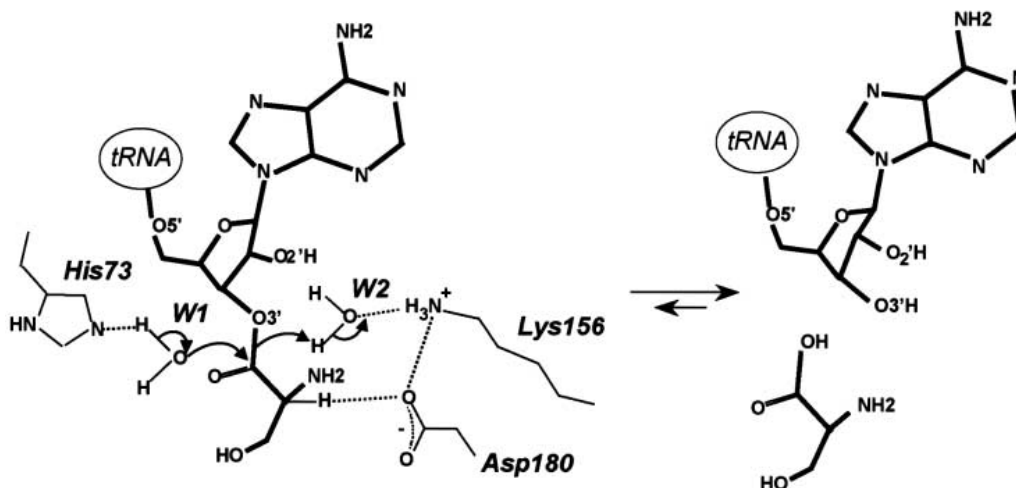
## 2.3.2 L'édition pour les ARNt<sup>Thr</sup>

### 2.3.2.1 L'édition des ThrRS bactériennes et eucaryotes

Les ThrRS doivent être capables de distinguer la thréonine de la valine et de la sérine qui en sont très proches. Une étude structurale réalisée avec la ThrRS d'*E. coli*, montre qu'un ion Zn interagit, dans le site actif de la protéine, avec le groupement hydroxyle de la thréonine. Il en résulte que la valine, qui ne possède pas de groupement hydroxyle, ne peut pas se fixer dans ce site actif. Par contre, la sérine peut interagir avec l'ion zinc, et cet acide aminé peut effectivement être activé puis transféré sur l'ARNt<sup>Thr</sup> (Sankaranarayanan *et al.*, 2000 ; Sankaranarayanan *et al.*, 1999).

Chez les bactéries et les eucaryotes, les ThrRS sont capables d'hydrolyser le Ser-ARNt<sup>Thr</sup> grâce à un domaine N-terminal N2, dont on verra ci-dessous qu'il présente des homologies de séquence avec le domaine d'édition des AlaRS (Dock-Bregeon *et al.*, 2000 ; Dock-Bregeon *et al.*, 2004).

Le site d'édition N2 des ThrRS présente un motif HxxxH fortement conservé. Une étude par mutagenèse dirigée chez *E. coli* (Dock-Bregeon *et al.*, 2004) a montré que les deux histidines de ce motif, ainsi que l'acide aspartique en position 180, jouaient un rôle important dans le mécanisme d'édition puisque le double mutant H73A-H77A est incapable de désacyler le Ser-ARNt<sup>Thr</sup>, tout comme le mutant D180A.



**Figure 13 : Mécanisme de l'édition du Ser-ARNt<sup>Thr</sup> par la ThrRS d'*E. coli***

D'après Dock-Bregeon *et al.* (2004). Ce mécanisme de l'hydrolyse du Ser-ARNt<sup>Thr</sup> dans le site d'édition de la ThrRS a été déduit des structures cristallographiques de la ThrRS d'*E. coli* en présence de sérine et d'analogues du substrat de la réaction d'édition (5'-O-(N-L-thréonyl)-sulfamoyl adénosine ou séryl-3'-adénosine (SerA76)).

La structure de la ThrRS d'*E. coli* en présence d'analogues du substrat de la réaction d'édition (5'-O-(N-L-thréonyl)-sulfamoyl adénosine (SerAMS) ou séryl-3'-adénosine (SerA76)) a permis de proposer un mécanisme qui implique deux molécules d'eau (W1 et W2 sur la figure 13) (Dock-Bregeon *et al.*, 2004). Dans ce modèle, l'histidine 73 stabilise la molécule d'eau W1 dans une position qui lui permet de réaliser une attaque nucléophile sur le carbone du groupement carbonyle de la sérine. La deuxième molécule d'eau W2 agirait comme donneur de proton, cédant un proton à l'oxygène O3' de A76 pour cliver la liaison ester entre l'ARNt et l'acide aminé.

Par ailleurs, les trois résidus His77, Tyr104 (non représentés sur la figure 13) et Asp180, semblent participer à la sélection de la sérine vis-à-vis de la thréonine, en empêchant stériquement la thréonine chargée sur l'ARNt de se fixer dans le site d'édition (Dock-Bregeon *et al.*, 2004).

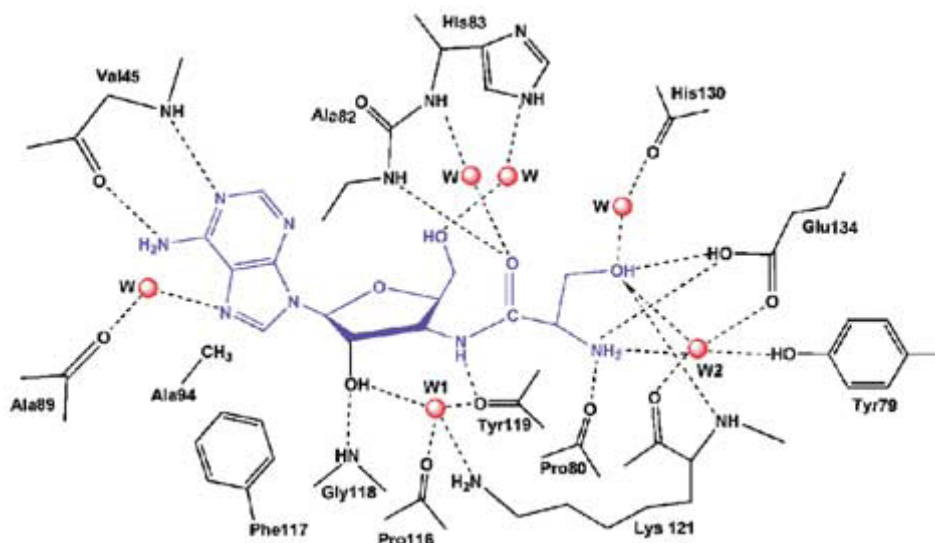
### 2.3.2.2 L'édition par les ThrRS chez les archées

Si quelques archées (par exemple, *Thermoplasma acidophilum* ou *Ferroplasma acidarmanus*) utilisent un domaine N2 du type bactérien/eucaryote, la plupart des ThrRS archéennes possèdent à la place un domaine N-terminal différent de N2 mais qui joue le même rôle (Beebe *et al.*, 2004 ; Korencic *et al.*, 2004).

Ce domaine d'édition, et notamment celui de *Pyrococcus abyssi* (Pab-NTD), possède une forte homologie structurale avec la D-Tyr-ARNt<sup>Tyr</sup> désacylase d'*E. coli* (DTD1), dont nous parlerons dans la suite de cette introduction, même si les deux polypeptides ne possèdent que 14% d'identité de séquence (Dwivedi *et al.*, 2005).

Le site actif du domaine d'édition de Pab-NTD a été étudié par cristallographie avec un analogue du substrat de l'édition post-transfert, la 3'-(L-séryl)-amino-3'-deoxyadénosine (Ser3AA). Le mécanisme n'est pas clairement établi mais un grand nombre d'interactions entre la protéine et cet analogue, faisant intervenir plusieurs molécules d'eau, a été identifié (Hussain *et al.*, 2006).

Les résidus Val45, Ala94 et Phe117, hautement conservés, forment une cavité dans laquelle se place l'adénine 3'-terminale de l'ARNt. Une molécule d'eau W1 (voir figure 14) semble idéalement placée dans le site actif pour l'hydrolyse de la liaison ester entre l'ARNt et la sérine. La seule chaîne latérale basique dans ce site actif est celle de la lysine 121, mais l'enzyme garde son activité d'édition si cette lysine est mutée en méthionine, dont la chaîne latérale n'est pas basique. Aussi, la seule base restante dans le site actif au voisinage de W1, et donc susceptible de l'activer pour son attaque nucléophile, est le groupement 2'-OH du ribose, suggérant un mécanisme catalytique assisté par le substrat (Hussain *et al.*, 2006) et un état de transition stabilisé par le trou oxyanion formé par l'alanine 82 et l'histidine 83.



**Figure 14 : Site d'édition de la ThrRS de *P. abyssi* en présence d'un analogue du substrat de l'édition post-transfert**

D'après Hussain *et al.* (2006). Ce schéma représentant les interactions entre la ThrRS de *Pyrococcus abyssi* (Pab-NTD) et un analogue de son substrat d'édition post-transfert, la 3'-(L-séryl) amino-3'-deoxyadénosine (Ser3AA), a été réalisé à partir de la structure cristallographique du complexe (code PDB 2HL0).

La thréonine ne semble pas pouvoir se fixer au site d'édition, à cause d'un conflit stérique entre son groupement méthyle, le groupement carbonyle de la tyrosine 119 et la chaîne latérale de la tyrosine 120, fortement conservée dans les ThrRS archéennes.

### 2.3.2.3 Une édition catalysée en *trans* chez quelques archées

Dans les ThrRS de quelques espèces d'archées (par exemple *Sulfolobus solfataricus*, *Sulfolobus tokadaii*, *Aeropyrum pernix*), ce domaine d'édition manque, mais il y a production d'une protéine indépendante ressemblant beaucoup à ce domaine (Beebe *et al.*, 2004 ; Korencic *et al.*, 2004). Ainsi, *Sulfolobus solfataricus* possède une ThrRS qui comprend uniquement un domaine d'aminocyclation de type bactérien et un domaine de liaison de l'anticodon (ThrRS-cat). Cette protéine ne comporte donc pas de domaine d'édition (Shimizu *et al.*, 2009) et, effectivement, elle n'hydrolyse pas le Ser-ARNt<sup>Thr</sup>. Par contre, il existe dans cette archée une autre protéine (ThrRS-ed) formée du domaine d'édition des archées fusionné à un domaine de reconnaissance de l'anticodon. Cette protéine hydrolyse le Ser-ARNt<sup>Thr</sup>, mais ni le Thr-ARNt<sup>Thr</sup> ni le Ser-ARNt<sup>Ser</sup> (Korencic *et al.*, 2004). L'importance de la ThrRS-ed a été montrée *in vivo* puisque la croissance d'une souche où son gène est inactivé est ralentie dans un milieu riche en sérine, ce qui suggère que la production de Ser-ARNt<sup>Thr</sup> est alors suffisamment élevée pour affecter la croissance (Korencic *et al.*, 2004).

### 2.3.3 L'édition des ARNt<sup>Ala</sup> mésaminoacylés

#### 2.3.3.1 L'édition catalysée par les AlaRS

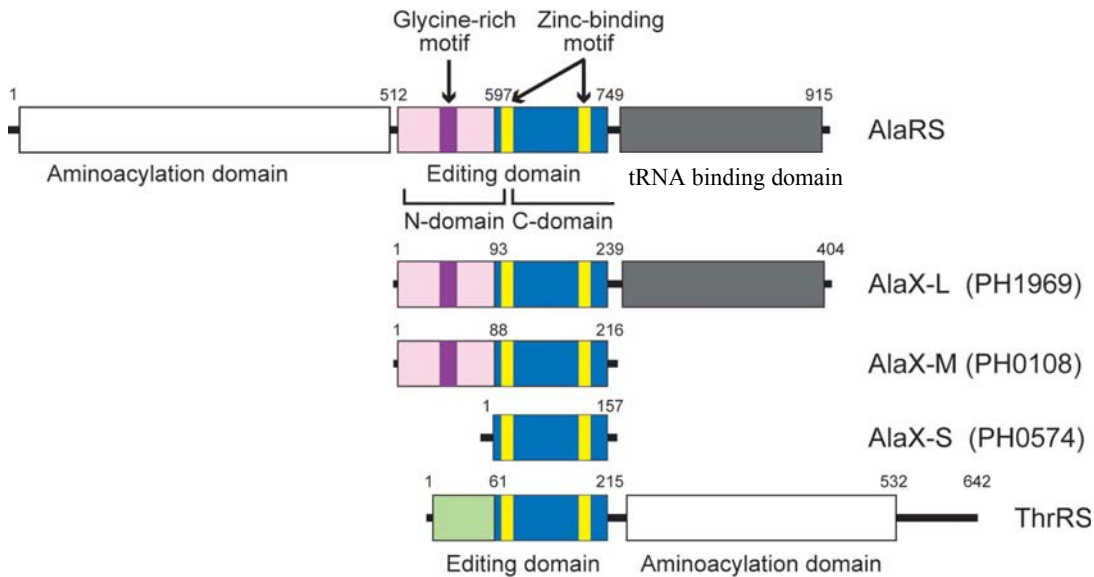
A quelques rares exceptions près, comme par exemple l'AlaRS de *Nanoarchaeum equitans* (Sokabe *et al.*, 2005), les AlaRS possèdent une activité d'édition post-transfert contre la sérine et la glycine (Tsui and Fersht, 1981). Cette activité est portée par un domaine d'édition C-terminal homologue au domaine N2-terminal des ThrRS bactériennes et eucaryotes (Beebe *et al.*, 2004 ; Beebe *et al.*, 2003).

Les AlaRS sont capables d'activer la glycine et la sérine avec des efficacités seulement 250 et 500 fois plus faibles, respectivement, que celle qui est obtenue pour l'alanine (Tsui and Fersht, 1981). L'activation de la sérine, plus volumineuse que l'alanine, est surprenante. Une étude structurale sur un fragment actif de l'AlaRS d'*Aquifex aeolicus* privé de son site d'édition (Swairjo and Schimmel, 2005) montre que la sérine peut outrepasser l'exclusion stérique dans le site actif via un changement conformationnel de l'enzyme, qui permet au groupement amide de l'Asn194 de former un pont hydrogène avec le groupement hydroxyle de la sérine alors qu'il interagit avec le groupement carboxyle de l'alanine ou de la glycine. Cette asparagine est strictement conservée chez les bactéries et chez les eucaryotes, mais elle n'est présente que dans quelques AlaRS archéennes. Dans les autres AlaRS d'archées, elle est remplacée par une thréonine qui, elle aussi, semble en mesure d'établir une liaison hydrogène avec le groupement hydroxyle de la sérine (Sokabe *et al.*, 2009).

#### 2.3.3.2 Trans-édition des ARNt<sup>Ala</sup> mésaminoacylés

Par ailleurs, de nombreux organismes expriment des protéines appelées AlaX, qui possèdent une grande homologie avec le domaine d'édition de l'AlaRS (Ahel *et al.*, 2003). Il en existe trois types, AlaX-L, AlaX-M et AlaX-S. AlaX-L possède un domaine homologue au domaine d'édition des AlaRS, ainsi qu'un domaine homologue à celui servant à l'ancrage de l'ARNt (Guo *et al.*, 2009) tandis que AlaX-M, AlaX-S ne comportent pas ce domaine et qu'AlaX-S ne comporte pas non plus l'extrémité N-terminale du domaine d'édition qui comprend entre autres un motif riche en glycine (figure 15).

Les protéines AlaX-L et AlaX-M sont largement distribuées dans les trois domaines du vivant alors qu'AlaX-S n'est produite que par quelques archées. Certains organismes (cyanobactéries, mycoplasmes, *E. coli*...) n'expriment cependant aucune protéine AlaX. Les archées des genres *Thermococcus* et *Pyrococcus* sont les seuls organismes connus possédant ces trois protéines AlaX (Fukunaga and Yokoyama, 2007).



**Figure 15 : Organisation des domaines de l'AlaRS, des protéines AlaX-L, AlaX-M et AlaX-S, et de la ThrRS**

D'après Fukunaga et Yokoyama (2007). La numérotation correspond à celle des protéines de *Pyrococcus horikoshii* pour l'AlaRS, AlaX-L, AlaX-M et AlaX-S, et à celle de la ThrRS et d'*E. coli* pour la ThrRS.

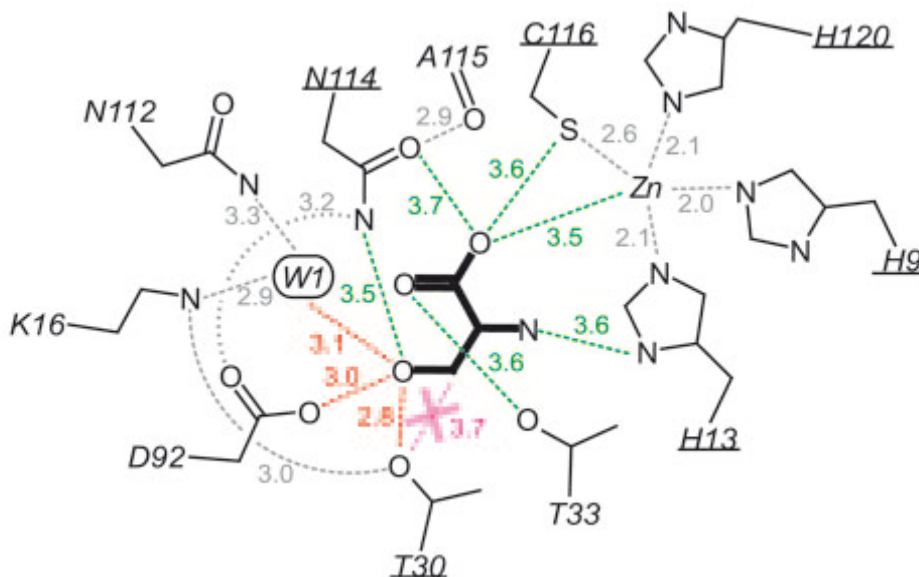
Les protéines AlaX-M de *Methanosarcina barkeri* et de *S. solfataricus* désacylent effectivement le Ser-ARNt<sup>Ala</sup> et le Gly-ARNt<sup>Ala</sup> mais pas l'Ala-ARNt<sup>Ala</sup> ni le Ser-ARNt<sup>Ser</sup>, ni le Ser-ARNt<sup>Thr</sup>, malgré l'homologie avec le domaine d'édition des ThrRS (Ahel *et al.*, 2003). Ainsi, même sans le domaine des AlaRS qui permet la reconnaissance des ARNt, les protéines AlaX sont capables de désacyler spécifiquement les ARNt<sup>Ala</sup> mésacylés. Chez *Pyrococcus horikoshii*, les protéines AlaX-M et AlaX-S sont capables d'hydrolyser le Ser-ARNt<sup>Ala</sup> mais seule la protéine AlaX-M hydrolyse le Gly-ARNt<sup>Ala</sup> (Fukunaga and Yokoyama, 2007 ; Sokabe *et al.*, 2005). La protéine AlaX-L de *S. cerevisiae* ne possède pas d'activité d'édition et sa fonction reste actuellement inconnue (Ahel *et al.*, 2003).

### 2.3.3.3 Structure des sites d'édition des AlaRS et de la protéine AlaX

Etant donné leurs fortes homologies, les protéines AlaX et les sites d'édition des AlaRS emploient certainement un mécanisme d'édition semblable (Sokabe *et al.*, 2005). Ces protéines possèdent un motif HXXXH et un motif CXXXH, tous les deux très conservés (Beebe *et al.*, 2003). Ces motifs stabilisent un ion Zn<sup>2+</sup> dans le site actif, mais il n'est pas certain que ce dernier soit nécessaire à l'activité d'édition (Sokabe *et al.*, 2005 ; Sokabe *et al.*, 2009). Ces motifs sont d'ailleurs conservés dans le domaine d'édition des ThrRS où ils semblent ne pas lier d'ion Zn<sup>2+</sup> et où ils seraient, par contre, en interaction directe avec le substrat (Dock-Bregeon *et al.*, 2004).

La structure du complexe entre la protéine AlaX-S de *Pyrococcus horikoshii* et la sérine montre qu'une thréonine, la Thr30, joue un rôle clé dans la fixation de la sérine et dans l'exclusion de l'alanine (Sokabe *et al.*, 2005). Avec

l'Asn114 et l'Asp92 (voir figure 16), cette thréonine participe avec une molécule d'eau à un réseau de liaisons hydrogène stabilisant le groupement hydroxyle de la chaîne latérale de la sérine dans le site d'édition. Ainsi, le site actif de la protéine AlaX-S rejette l'alanine à cause d'un conflit stérique entre la chaîne latérale de l'alanine et la Thr30, et accepte la sérine, car alors, la gêne stérique due à la chaîne latérale est compensée énergétiquement par la formation de plusieurs liaisons hydrogène avec le groupement hydroxyle de la chaîne latérale.



**Figure 16 : Sérine dans le site d'édition de l'AlaX-S de *P. horikoshii***

D'après Sokabe *et al.* (2005). Ce schéma a été obtenu à partir de la structure cristallographique du complexe entre une sérine et l'AlaX-S de *P. horikoshii*. Un réseau de liaisons hydrogène stabilisant la sérine (produit de la réaction d'édition) dans le site actif est représenté en pointillé. La gêne stérique (en rouge) entre les chaînes latérales de la sérine et de la Thr30 est compensée par des liaisons hydrogène (en orange) entre le groupement hydroxyle de la sérine et la protéine.

Cette thréonine est conservée dans la plupart des protéines AlaX-S et AlaX-L et l'on peut noter que la protéine AlaX-L de *S. cerevisiae*, qui ne possède pas d'activité d'édition, ne possède pas non plus cette thréonine. Dans le domaine d'édition des AlaRS, ainsi que dans certaines protéines AlaX-M, le rôle joué par cette thréonine est tenu par une glutamine (Gln 584 dans l'AlaRS d'*E. coli*) (Sokabe *et al.*, 2005). Cette glutamine peut écarter l'alanine du site d'édition comme le fait la thréonine, en créant un conflit stérique avec le groupement méthyle de l'alanine. De plus, elle pourrait contribuer à stabiliser la glycine, via une liaison hydrogène faible avec un hydrogène  $\alpha$  de la glycine (Sokabe *et al.*, 2005). Une telle discrimination entre la glycine et l'alanine a d'ailleurs déjà été décrite dans le cas du site d'activation de la GlyRS de *T. thermophilus*, où le Glu359 n'occupe pas seulement la place d'une chaîne latérale plus grande que celle de la glycine dans le site actif, mais interagit aussi avec un hydrogène  $\alpha$  de la glycine (Arnez *et al.*, 1999).

A propos de la reconnaissance du Gly-ARNt<sup>Ala</sup> par l'AlaRS et la protéine AlaX-M, mais pas par la protéine AlaX-S, il faut aussi noter que l'AlaRS et la protéine AlaX-M possèdent une région riche en glycines qui n'existe pas dans la protéine AlaX-S (voir figure 15). Cette région, qui forme une boucle juste au-dessus du site actif (Fukunaga and Yokoyama, 2007), pourrait être importante pour la reconnaissance du Gly-ARNt<sup>Ala</sup>.

Dans le cas de l'AlaRS, l'ARNt semble devoir se dissocier du site d'aminacylation avant de pouvoir se fixer dans le site d'édition (Sokabe *et al.*, 2009). L'édition de l'AlaRS pourrait ainsi être moins efficace que l'édition par d'autres aminoacyl-ARNt synthétases, ce qui pourrait expliquer la conservation, au moins dans certains organismes, de protéines faisant de l'édition en *trans* (Sokabe *et al.*, 2009).

### 2.3.4 L'édition pour les ARNt<sup>Phe</sup>

#### 2.3.4.1 L'activité d'édition des PheRS

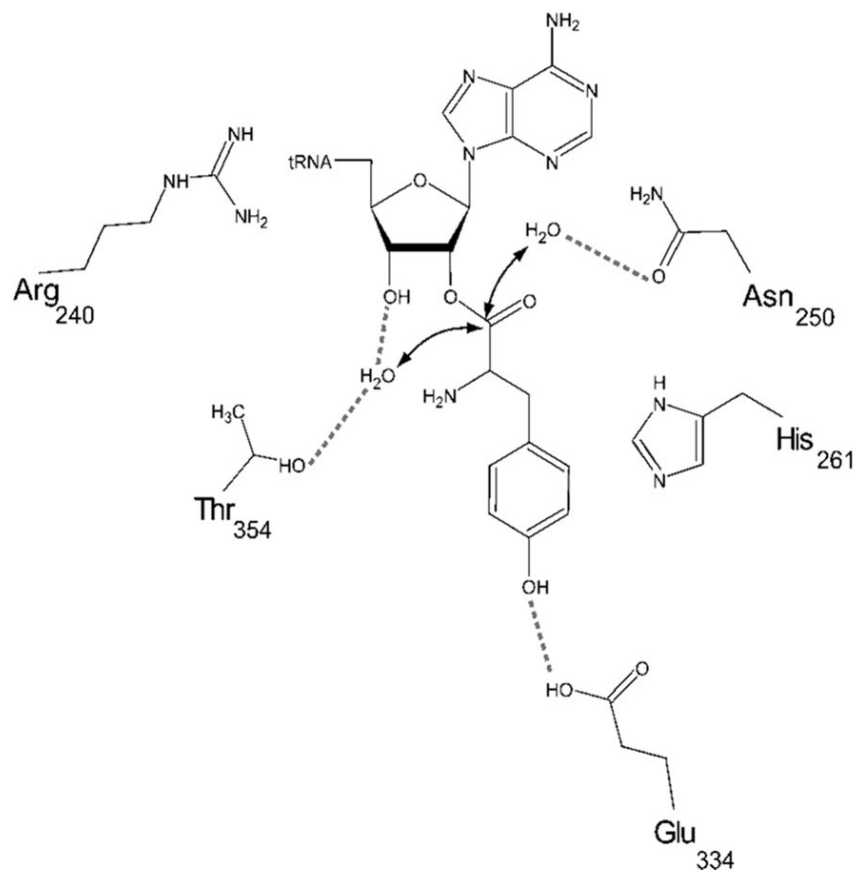
La PheRS d'*E. coli* contrôle sa fidélité via des mécanismes d'édition pré- et post-transfert qui lui permettent d'éviter la formation de Tyr-ARNt<sup>Phe</sup>. En particulier, elle possède une activité qui lui permet d'hydrolyser le Tyr-ARNt<sup>Phe</sup> (Roy *et al.*, 2004). Cette activité est très spécifique de la tyrosine, puisque d'autres aminoacyl-ARNt, où l'acide comprend un cycle aromatique substitué en position *para* (p-chlorophénylalanine, p-bromophénylalanine), ne sont pas hydrolysés efficacement par la PheRS (Ibba *et al.*, 1994 ; Sharma *et al.*, 2000).

Le domaine responsable de l'activité d'édition post-transfert de la PheRS d'*E. coli* a été identifié par des expériences de mutagenèse dirigée portant sur les résidus fortement conservés dans les différentes PheRS bactériennes (Roy *et al.*, 2004). Il est situé à 35-40 Å du site d'activation et fait partie de la sous-unité  $\beta$  de la protéine. Ce dernier résultat est en accord avec le fait que la PheRS mitochondriale de *S. cerevisiae* ne possède pas de sous-unité  $\beta$  et ne possède pas non plus de fonction d'édition (Reynolds *et al.*, 2010 ; Roy *et al.*, 2005).

#### 2.3.4.2 Le domaine d'édition des PheRS bactériennes

Le domaine d'édition des PheRS ne ressemble à aucun autre domaine d'édition connu. Chez les bactéries, il comporte une cavité hydrophobe dans laquelle vient se loger la chaîne latérale de la tyrosine. Au fond de cette cavité, un acide glutamique conservé (Glu 334 dans la PheRS de *T. thermophilus*) permet l'ancrage de la tyrosine dans le site d'édition via une liaison hydrogène entre sa chaîne latérale et le groupement p-hydroxyle de la tyrosine (voir figure 17), mais ne permet pas la fixation de la phénylalanine (Kotik-Kogan *et al.*, 2005 ; Ling *et al.*, 2007 ; Roy *et al.*, 2004).

Par ailleurs, dans le site d'édition, une histidine (His265 dans la PheRS d'*E. coli*) est proche de la liaison ester hydrolysée. Cependant, cette histidine ne joue pas un rôle majeur dans la catalyse puisque sa substitution par une alanine ne diminue que d'un facteur 3 l'activité d'édition de la protéine (Ling *et al.*, 2007). Elle ne jouerait donc pas du tout le même rôle que l'His73 de la ThrRS d'*E. coli* et interviendrait plutôt dans le positionnement du substrat. En fait, aucun résidu du site d'édition ne semble avoir un rôle déterminant dans la catalyse (Ling *et al.*, 2007), ce qui conduit à penser que l'attaque nucléophile d'une molécule d'eau est assistée par un groupement hydroxyle du ribose de l'adénosine 3'-terminale (ici, le groupement 3'-OH), comme dans le cas de la ThrRS archéenne. En accord avec cette hypothèse, le remplacement de ce groupement 3'-OH par un atome d'hydrogène diminue environ 300 fois l'efficacité catalytique de la réaction d'édition (Ling *et al.*, 2007).



**Figure 17 : Mécanisme de l'édition post-transfert catalysée par la PheRS de *T. thermophilus***

D'après Ling *et al.* (2007). Le Tyr-ARNt<sup>Phe</sup> est hydrolysé dans le site d'édition de la PheRS de *T. thermophilus*, grâce à une molécule d'eau stabilisée, en particulier, par le groupement 3'-OH du substrat. L'His261 est l'équivalent de l'His265 de la PheRS d'*E. coli*.



### 2.3.4.3 Le domaine d'édition des PheRS archéennes et eucaryotes

Si les structures des domaines d'édition des PheRS archéennes et eucaryotes ressemblent à celle des domaines des PheRS bactériennes (voir figure 20), les résidus interagissant avec la tyrosine sont généralement différents de ceux que l'on trouve dans les PheRS bactériennes (Sasaki *et al.*, 2006).

Le principe de la reconnaissance de la chaîne latérale de la tyrosine est néanmoins le même, avec une cavité hydrophobe au fond de laquelle se trouve un résidu d'acide aspartique (Asp234 dans la PheRS de *P. horikoshii*), qui interagit avec le groupement p-hydroxyle de la tyrosine et qui joue donc un rôle équivalent à celui du Glu334 dans la PheRS de *T. thermophilus* (Sasaki *et al.*, 2006).

Contrairement à ce qu'on a vu ci-dessus dans le cas des PheRS bactériennes, il existe dans les PheRS archéennes/eucaryotes des résidus, au voisinage de la liaison ester hydrolysée, dont la mutation décroît considérablement l'activité d'édition de l'enzyme (par exemple, la mutation NA217 dans la PheRS de *P. horikoshii* diminue au moins 100 fois la vitesse d'hydrolyse du Tyr-ARNt<sup>Phe</sup> (Sasaki *et al.*, 2006)). Comme, par ailleurs, aucune expérience n'a été réalisée sur cette enzyme avec du Tyr-3'-dA76-ARNt<sup>Phe</sup> (ARNt sans groupement hydroxyle à la position 3' du ribose terminal), on ne sait pas actuellement si, dans les PheRS archéennes et eucaryotes, le mécanisme catalytique est assisté par le groupement 3'-OH terminal de l'ARNt.

### **3. Les D-aminoacyl-ARNt désacylases**



### 3.1 Le problème des acides aminés D

Nous venons d'étudier plusieurs mécanismes visant à faire en sorte qu'une aminoacyl-ARNt synthétase reconnaisse son ARNt et l'aminoacyle avec l'acide aminé correspondant, tout en écartant les autres acides aminés, notamment grâce à la différence de taille et de composition chimique de leur chaîne latérale. Cependant, comme les acides aminés sont des molécules chirales (à l'exception de la glycine), il existe deux formes possibles, L ou D, qui possèdent les mêmes chaînes latérales.

Dans le monde vivant, les protéines, produites par les ribosomes, sont uniquement composées d'acides aminés L. Pourtant, les acides aminés D n'ont pas disparu des systèmes biologiques. Ils sont présents dans de nombreux organismes, et ils y jouent un rôle actif. Les acides aminés D sont ainsi des constituants de la paroi cellulaire de nombreuses bactéries. Certains antibiotiques, produits par des bactéries ou des champignons, contiennent aussi des acides aminés D (comme le D-Asp, D-Trp, D-Val et D-Cys (Friedman, 1999)). Chez les mammifères, des acides aminés D sont présents sous forme libre et interviennent vraisemblablement dans la différenciation cellulaire (Takarada *et al.*, 2009) et la neurotransmission (Wolosker *et al.*, 1999). Il doit donc exister des mécanismes qui permettent d'éviter que les acides aminés D soient incorporés dans les protéines.

Une première sélection est effectuée par les aminoacyl-ARNt synthétases, qui sont spécifiques des acides aminés L. Mais cette spécificité n'est pas absolue et, par exemple, chez *E. coli*, la D-tyrosine, le D-tryptophane, le D-aspartate, la D-lysine et la D-histidine peuvent être transférés *in vitro* sur l'ARNt correspondant, par la synthétase correspondante (Calendar and Berg, 1966, 1967 ; Sheoran and First, 2008 ; Sheoran *et al.*, 2008 ; Soutourina *et al.*, 2000a ; Soutourina *et al.*, 2000b ; Takayama *et al.*, 2005). Les vitesses mesurées *in vitro* avec les acides aminés D sont de 15 à 2000 fois plus faibles que celles qui sont obtenues avec les acides aminés L. De telles vitesses seraient cependant suffisantes pour provoquer une accumulation de D-aminoacyl-ARNt dans les cellules, s'il n'existait pas des protéines, les D-aminoacyl-ARNt désacylases (DTD), qui hydrolysent spécifiquement les ARNt aminoacylés avec un acide aminé D.

### 3.2 D-aminoacyl-ARNt désacylases bactériennes (DTD1)

Une activité capable de catalyser l'hydrolyse des D-Tyr-ARNt<sup>Tyr</sup>, mais pas celle des L-Tyr-ARNt<sup>Tyr</sup>, a été observée dès 1967 chez *E. coli*, dans la levure, dans les réticulocytes de lapin et dans le foie de rat (Calendar and Berg, 1967). Plus récemment, (1999) le gène codant la D-aminoacyl-ARNt-désacylase d'*E. coli* (*dtd*) a été identifié (Soutourina *et al.*, 1999). Cette enzyme (DTD1) possède un spectre d'action très étendu puisqu'elle est capable d'hydrolyser les ARNt aminoacylés avec des acides aminés D très différents, tels que le D-aspartate et le D-tryptophane (Soutourina *et al.*, 2000b).

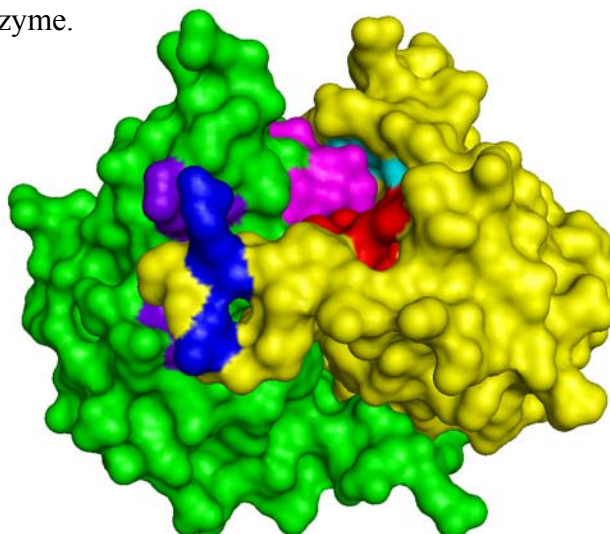
Chez *E. coli*, il a été montré que les souches n'exprimant pas la D-aminoacyl-ARNt désacylase ont une croissance ralentie lorsque le milieu de culture contient des acides aminés D (D-Tyr, D-Trp, D-Asp, D-Ser ou D-Gln) (Soutourina *et al.*, 2000b ; Soutourina *et al.*, 1999). Il a également été montré que, lorsque ces souches sont cultivées en présence de 2,4 mM de D-tyrosine, près de 50% des ARNt<sup>Tyr</sup> cellulaires se trouvent convertis en D-Tyr-ARNt<sup>Tyr</sup> (Soutourina *et al.*, 2004). Ces résultats indiquent que la D-aminoacyl-ARNt désacylase protège le colibacille contre les effets toxiques des acides aminés D en hydrolysant les D-aminoacyl-ARNt produits par les aminoacyl-ARNt synthétases.

Les structures des protéines DTD1 d'*E. coli* (code PDB 1jke (Ferri-Fioni *et al.*, 2001)), d'*Haemophilus influenzae* (code PDB 1j7g (Lim *et al.*, 2003)), d'*Aquifex aeolicus* (code PDB 2dbo), de *Leishmania major* (code PDB 1tc5) et de l'homme (code PDB 2okv) ont été résolues par cristallographie et diffraction aux rayons X. Toutes ces protéines sont homodimériques.

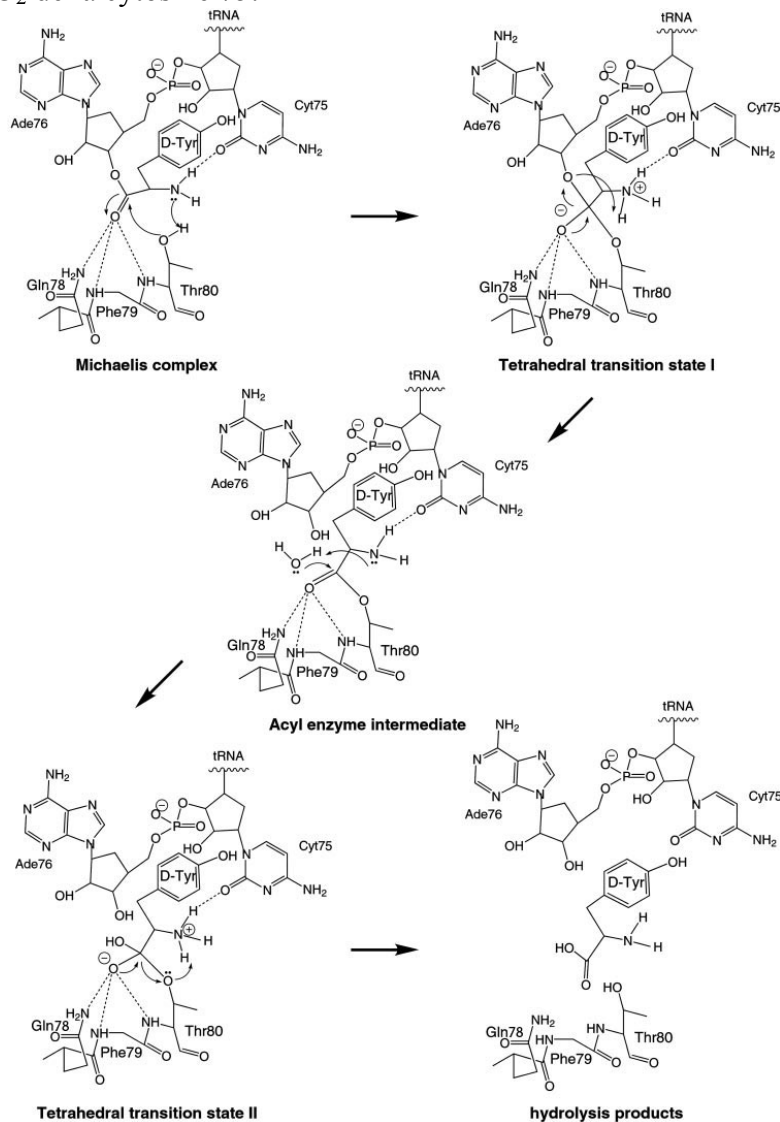
A l'interface des deux sous-unités de la protéine DTD1 d'*E. coli*, on trouve deux régions hautement conservées, la région 77-81 (SQFTL) (en rouge sur la figure 18) d'un protomère et la région 135-140 (NXGP(V/F)T) (en magenta) du second protomère. Ces deux régions forment une cavité qui constitue vraisemblablement le site actif de l'enzyme.

#### Figure 18 : Structure de DTD1

Représentation à l'aide du logiciel Pymol de la surface du dimère DTD1 d'*E. coli* (1JKE.pdb). En jaune le protomère A, en vert le B. En rouge, la région 77-81 (SQFTL) du protomère A. En magenta, la région 135-140 (NXGP(V/F)T) du protomère B. En cyan, la Met129 du protomère A. En bleu, l'Arg87 et l'Arg 90 du protomère A. En violet, l'Arg48 et l'Arg53 du protomère B.



La structure du complexe entre DTD1 et son substrat n'a pas encore été obtenue. Néanmoins, sur la base de la structure tridimensionnelle de la protéine DTD1 d'*Haemophilus influenzae*, un modèle de mécanisme catalytique a été proposé (Lim *et al.*, 2003) (voir figure 19). Le groupement hydroxyle de la Thr80 de l'enzyme, interviendrait comme groupement nucléophile pour attaquer le groupement carbonyle de la D-tyrosine. Un trou oxyanion, formé par les groupements NH de la Thr80 et de la Phe79 ainsi que celui de la chaîne latérale de la Gln78, stabiliserait l'état de transition, tétraédrique, chargé négativement. Cette réaction serait assistée par le substrat lui-même, via le groupement amino de la D-tyrosine qui servirait de base acceptant le proton du groupement hydroxyle de la Thr80 qui deviendrait alors nucléophile. Le groupement amino de la D-tyrosine alors chargé positivement serait, lui, stabilisé par une liaison hydrogène avec l'atome O<sub>2</sub> de la cytosine 75.



**Figure 19 : Mécanisme catalytique de la D-Tyr-ARNt désacylase**

D'après Lim *et al.* (2003). Ce mécanisme proposé par Lim *et al.* a été déduit d'un modèle d'ancrage du D-Tyr-ARNt dans le site actif de l'enzyme.

Dans ce modèle, qui n'a pas encore été validé par des expériences de mutagenèse, le positionnement proposé pour la D-tyrosine est en accord avec la stéréospécificité de DTD1. En effet, la chaîne latérale de la D-tyrosine est exposée vers le solvant, ce qui permet à la protéine d'accepter une grande variété de chaînes latérales. Au contraire, un tel positionnement pour un acide aminé L entraînerait des conflits stériques avec la protéine (Lim *et al.*, 2003). La méthionine 129, très conservée, jouerait un rôle dans cette discrimination puisque sa mutation en lysine dans la protéine DTD1 d'*E. coli*, permet à celle-ci de fixer des acides aminés L dans son site actif (Hussain *et al.*, 2006).

Par ailleurs, il a été démontré que la D-Tyr-adénosine n'était pas clivée par DTD1 (Calendar and Berg, 1967), ce qui suppose l'existence d'interactions entre l'ARNt et la protéine au-delà de l'adénosine terminale. Des résidus situés au voisinage du site actif, R48 et R53 d'un protomère et R87 et R90 de l'autre protomère chez *E. coli* (en violet et en bleu sur la figure 18), tous conservés en résidus chargés positivement dans les protéines DTD1 des différents organismes, forment une chaîne cationique qui pourrait permettre à la protéine d'établir des contacts avec l'extrémité polyanionique de l'ARNt lors de la fixation du D-aminoacyl-ARNt.

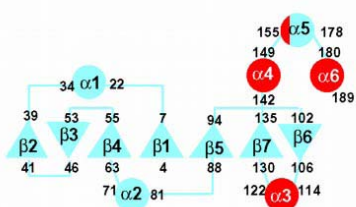
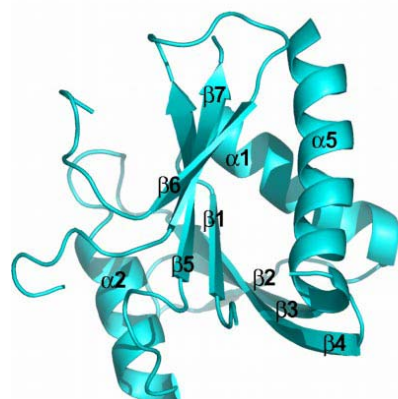
La structure de DTD1 est très homologue à celle du domaine N-terminal d'édition des ThrRS archéennes même si l'identité de séquence entre ces polypeptides n'est que de 14% (Dwivedi *et al.*, 2005). Il est donc possible que le module d'édition de la ThrRS archéenne (Pab-NTD) ait été ajouté à la ThrRS au cours de l'évolution à partir d'un ancêtre commun avec cette désacylase. Dans DTD1, on trouve une cavité formée par les résidus Val33, Ala102 et Phe 125, qui est structuralement très proche de celle qui est formée par les résidus Val45, Ala94 et Phe117 de Pab-NTD. D'après la structure d'un complexe entre Pab-NTD et un analogue du substrat, cette cavité accueille l'adénosine terminale de l'ARNt (Hussain *et al.*, 2006). Celle de DTD1, qui est au voisinage du site actif présumé, pourrait donc y jouer le même rôle.

### 3.3 D-aminoacyl-ARNt désacylases archéennes (DTD2)

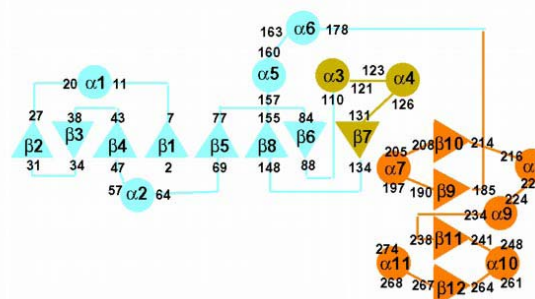
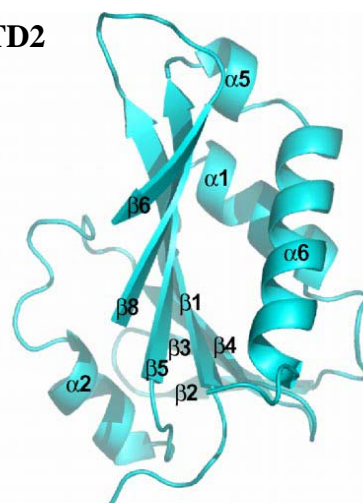
Il existe des homologues de *dtd* dans environ les deux tiers des bactéries et chez la plupart des eucaryotes dont le génome a été séquencé. Par contre, aucun homologue de ces gènes n'est présent chez les archées (à une exception près : *Methanospheera stadtmannae*). Chez ces organismes, il existe une enzyme (DTD2) dont la séquence est complètement différente de celle de DTD1, mais dont les caractéristiques fonctionnelles sont similaires. Ainsi, DTD2 complète une souche d'*E. coli* déficiente en *dtd*. Des homologues de la protéine DTD2 existent chez la plupart des archées et chez les plantes.

Les structures des protéines DTD2 d'*Archaeoglobus fulgidus* (code PDB 1yqe) et de *Pyrococcus horikoshii* (code PDB 2gfg) ont été résolues par cristallographie lors d'approches de génomique structurale, avant que la fonction de ces protéines ne soit connue. Ces deux structures, très proches entre elles, sont très différentes de la structure de DTD1. Elles possèdent par contre des ressemblances avec celle de la peptidyl-ARNt hydrolase bactérienne que nous étudierons plus tard (voir figure 20).

PTH1



DTD2



**Figure 20 : Représentation schématique des structures tridimensionnelles de la PTH1 d'*E. coli* et de la protéine DT2 d'*A. fulgidus***

Pour mieux faire apparaître la ressemblance entre les structures tridimensionnelles des protéines PTH1 et DTD2, le squelette polypeptidique des résidus (10-20, 114-122, 137-159 et 180-193) de



PTH1, qui n'ont pas d'équivalents dans DTD2, n'est pas représenté. De même, seul le squelette peptidique correspondant aux résidus 1-88 et 148-182 de DTD2 est représenté. Une représentation schématique des structures secondaires de chaque enzyme est tracée sous leur structure 3D, avec les brins bêta sous formes de triangles et les hélices sous forme de cercles. Les éléments de structure secondaire de PTH1 qui n'ont pas d'équivalent sur DTD2 sont colorés en rouge et ceux de DTD2 qui n'ont pas d'équivalent sur PTH1 sont en jaune et orange. Ces deux protéines peuvent être superposées avec un RMSD (*root mean square deviation*) de 1,9 Å pour les atomes C<sup>α</sup> de 97 résidus. D'après Ferri-Fioni *et al.* (2006).

DTD2 n'est cependant pas capable d'hydrolyser les peptidyl-ARNt ce qui n'est pas étonnant puisque les résidus catalytiques de la PTH (voir paragraphe 5 ci-dessous) ne sont pas conservés. Contrairement à la PTH et à la D-aminoacyl-ARNt-désacylase DTD1, elle contient deux atomes de zinc. La région où se fixent ces atomes de zinc semble jouer un rôle majeur dans l'activité de l'enzyme. Dans cette région, un certain nombre de résidus conservés (His80, His92, Glu137, His140, His141, Glu154, His197 et Tyr198) ont en effet été mutés, ce qui a eu pour effet de diminuer l'activité catalytique de l'enzyme et/ou sa capacité à fixer du zinc (Ferri-Fioni *et al.*, 2006). En particulier, la mutation de l'histidine 140 en alanine inactive complètement l'enzyme. Cette région ne correspond pas au site actif de la PTH, suggérant que le repliement commun à la PTH et à DTD2 a été utilisé pour produire des protéines ayant des fonctions et des propriétés catalytiques distinctes.

### 3.4 D-aminoacyl-ARNt désacylases chez les cyanobactéries (DTD3)

Dans la sous-classe des cyanobactéries, la plupart des espèces ne possèdent ni homologue de DTD1 ni homologue de DTD2. Une troisième D-aminoacyl-ARNt-désacylase y a été identifiée récemment (DTD3) (Wydau *et al.*, 2009).

La protéine DTD3 appartient à une grande famille d'hydrolases dont l'activité requiert un ou plusieurs ions métalliques (Holm and Sander, 1997). Les structures de plusieurs protéines présentant de fortes homologies de séquence avec DTD3 sont connues (codes PDB 1xwy, 1zzm, 1zix, 1j6o, 2gzx). Elles adoptent toutes le même repliement, ce qui suggère que DTD3 l'adopte également. Il est très différent de celui des deux autres DTD connues.

Comme DTD2, DTD3 peut fixer deux ions métalliques mais la forme à deux ions zinc n'est que faiblement active. Pour que l'enzyme soit pleinement active, il faut que le centre bimétallique soit occupé par un atome de zinc et soit un atome de cobalt, soit un atome de nickel, soit un atome de manganèse, ou alors qu'il soit occupé par deux atomes de cobalt.

DTD3 complémente une souche d'*E. coli*  $\Delta dtd$  et possède des homologues dans de très nombreux organismes, ce qui suggère que la plupart des cellules contiennent une D-aminoacyl-ARNt désacylase. Dans plusieurs organismes, il existe même à la fois un homologue de DTD3 et une autre DTD (voir (Wydau *et al.*, 2009) pour revue). La défense des organismes contre les acides aminés D semble donc universelle, même si les diverses enzymes responsables de cette activité sont très différentes.



## **4. Le *drop-off***



## 4.1 Edition ribosomale et *drop-off*

Nous venons de voir qu'un grand nombre de mécanismes sont mis en place dans la cellule afin d'éviter que les ARNt ne soient mésaminoacylés, ce qui peut conduire à l'incorporation d'un mauvais acide aminé dans la synthèse protéique, ou à la séquestration d'ARNt non fonctionnels. Un autre mécanisme, désigné en anglais sous le nom de *drop-off*, va aussi entraîner la séquestration d'ARNt, cette fois-ci sous forme de peptidyl-ARNt. Le *drop-off* est un mécanisme de terminaison prématurée de la traduction par lequel un peptide en cours d'élongation, accroché à un ARNt, se détache du ribosome. La peptidyl-ARNt hydrolase (PTH) que nous allons évoquer dans la dernière partie de cette introduction, assure le recyclage des ARNt ainsi séquestrés en hydrolysant la liaison ester entre l'ARNt et le peptide.

L'existence du phénomène de *drop-off* a été découverte dans les années 70, notamment grâce aux travaux de Menninger (Menninger, 1976). Pour l'instant, il n'a été mis en évidence que chez les bactéries. Dès 1972, Atherly et Menninger ont isolé une souche d'*E. coli* thermosensible qui, à température non permissive, accumule rapidement des peptidyl-ARNt (Atherly and Menninger, 1972). Cette accumulation est létale pour la cellule car elle entraîne une séquestration des ARNt (Menninger, 1979) et notamment de l'ARNt<sup>Lys</sup> (Heurgué-Hamard *et al.*, 1996). Cette souche thermosensible porte une mutation dans le gène de la PTH (*pth*) ce qui a permis de conclure, d'une part, que la libération de peptidyl-ARNt par le ribosome était un phénomène fréquent et, d'autre part, que la PTH permettait d'éviter l'accumulation de ces peptidyl-ARNt dans la cellule.

Le taux de dissociation des peptidyl-ARNt des ribosomes est assez élevé puisque Menninger estime que la probabilité qu'un peptidyl-ARNt se détache lors d'un cycle d'élongation est en moyenne de 1/2600. Ainsi, la probabilité qu'un peptidyl-ARNt se détache avant la fin de la synthèse d'une protéine de 300 acides aminés serait de 11%. Dans des conditions où la synthèse protéique est ralentie, par exemple par le passage à une température non permissive d'une souche *pth*<sup>ts</sup>, la probabilité de dissociation d'un peptidyl-ARNt lors d'un cycle d'élongation augmente, pouvant atteindre la valeur de 1/90 (Menninger, 1976).

Les travaux de Menninger montrent que la taille des peptides qui se décrochent du ribosome, toujours liés à un ARNt, peut être grande (Menninger, 1976) (au moins de 800 à 10 000 daltons). Il a alors été suggéré que le *drop-off* pourrait être un mécanisme de correction d'erreur, qui interviendrait quand le peptidyl-ARNt ne correspond pas au codon de l'ARN messenger auquel il est apparié, par exemple à la suite d'un mauvais décalage du ribosome lors du passage d'un codon au suivant.

Le *drop-off* serait ainsi un mécanisme de correction d'erreur que l'on pourrait qualifier de "post-transfert" par opposition aux mécanismes de correction qui ont lieu dans le site A, avant la formation de la nouvelle liaison peptidique. Il

pourrait donc être complémentaire d'un autre mécanisme d'édition post-transfert, récemment identifié par Zaher et Green (Zaher and Green, 2009a, b). En étudiant la traduction *in vitro* de petits ARNm modèles, ces chercheurs ont en effet montré qu'un mauvais appariement codon-anticodon au niveau du site P, ou à la fois au niveau du site P et au niveau du site E, pouvait provoquer la liaison d'un facteur de terminaison RF1 ou RF2 au niveau du site A, et, finalement, la terminaison de la traduction et la libération du peptide.

## 4.2 Terminaison précoce de la traduction impliquant les facteurs de terminaison

Si la preuve d'une édition ribosomale conduisant au *drop-off* n'a pas encore été apportée, des études récentes montrent que le *drop-off* n'est pas seulement spontané mais qu'il peut être catalysé par des facteurs de terminaison de la traduction et/ou par des facteurs d'initiation.

Ainsi, la recherche de révertants d'une mutation *pth<sup>ts</sup>* chez *E. coli*, a abouti à la sélection de souches où l'expression de RRF est diminuée (Heurgué-Hamard *et al.*, 1998). Au contraire, une surproduction de RRF augmente l'accumulation de peptidyl-ARNt dans un mutant *pth<sup>ts</sup>* cultivé à une température un peu inférieure à la température non permissive. Le recyclage de peptidyl-ARNt par RRF requiert la participation de EF-G (Rao and Varshney, 2001) et est stimulé par IF3 (Singh *et al.*, 2005). Par ailleurs, le *drop-off* est facilité par RF3 puisque l'inactivation du gène *prfC* (codant RF3) supprime partiellement la thermosensibilité d'un mutant *pth<sup>ts</sup>* chez *E. coli* et que la combinaison d'une mutation dans *frr* (codant RRF) et de l'inactivation de *prfC* a un effet additif sur la suppression de la mutation *pth<sup>ts</sup>* (Gong *et al.*, 2007 ; Singh *et al.*, 2008).



### 4.3 Contrôle du démarrage de la traduction

Le *drop-off* semble également pouvoir être catalysé par d'autres facteurs de la traduction, notamment au début de la biosynthèse d'une protéine.

Les facteurs de démarrage de la traduction, IF1 et IF2, interviennent normalement pour garantir le bon positionnement de l'ARNt initiateur dans le site P, lors de la formation du complexe 70S de démarrage de la traduction. L'hydrolyse de la molécule de GTP liée à IF2, (Benne *et al.*, 1973) entraîne ensuite le départ de ces facteurs, et le complexe 70S peut alors entrer dans une phase d'élongation. Le recrutement continu d'aminocyl-ARNt au niveau du site A empêche ensuite ces facteurs d'initiation de se lier à nouveau au ribosome (Das and Varshney, 2006). Cependant, dans le cas de ribosomes bloqués, soit au niveau d'un codon stop par l'absence de facteur de terminaison, soit au niveau d'un codon sens par l'absence de l'aminocyl-ARNt correspondant au codon du site A, IF1 et IF2 pourraient se lier à nouveau au ribosome, ce qui pourrait déstabiliser la liaison entre le ribosome et le peptidyl-ARNt dans le site P et conduire à son *drop-off*.

Effectivement, dans un système de traduction *in vitro*, l'augmentation simultanée de la concentration en IF1 et IF2 accroît considérablement la libération de peptidyl-ARNt dans le cas de ribosomes bloqués (Karimi *et al.*, 1998). De plus, la surexpression simultanée de ces deux facteurs dans une souche d'*E. coli* exprimant une PTH thermosensible, inhibe la multiplication cellulaire à une température de 38°C, alors que la bactérie peut se développer à cette température lorsque ces deux facteurs ne sont pas surexprimés (Karimi *et al.*, 1998). Il semble donc qu'*in vivo* aussi, le *drop-off* soit catalysé par IF1 et IF2.

Cette action simultanée d'IF1 et IF2 se produit préférentiellement pendant les premiers cycles d'élongation. Elle pourrait donc avoir un rôle dans la régulation de la traduction de certains ARN messagers contenant des codons rares au début de leur cadre de lecture (Chen and Inouye, 1990 ; Chen and Inouye, 1994). Un tel mécanisme pourrait être utile en conditions de carence, quand il peut être avantageux pour la cellule d'arrêter la synthèse d'un certain nombre de protéines, et d'augmenter ainsi ses chances de terminer la synthèse d'autres protéines.

## 4.4 Trans-traduction et *drop-off*

Le *drop-off* pourrait également participer au déblocage des ribosomes bloqués sur un ARN messager sans codon de terminaison. De tels ARNm sont le fruit d'une transcription incomplète ou d'une dégradation de l'ARN messager. Si un ARNm sans codon de terminaison est traduit, le ribosome reste séquestré sur l'ARN messager au niveau du dernier codon et ne peut donc plus être utilisé pour la traduction d'une autre protéine. Ceci est d'autant plus néfaste que plusieurs ribosomes peuvent alors se trouver bloqués à la suite du premier, sur le même ARN messager.

Afin de recycler ces ribosomes bloqués, il existe un mécanisme (la trans-traduction) découvert récemment, faisant intervenir l'ARNtm. Cet ARN est une molécule hybride, ARNt/ARNm, possédant une région mimant l'ARNt<sup>Ala</sup>, pouvant être aminoacylée par l'AlaRS, et une autre région codant un court polypeptide (Komine *et al.*, 1994 ; Tu *et al.*, 1995). Une fois alanylé, l'ARNtm peut être recruté au niveau du site A du ribosome bloqué. Le peptide en cours d'élongation est alors transféré sur l'alanine de l'ARNtm, comme dans un cycle d'élongation normal. La partie ARNm de l'ARNtm est ensuite traduite à la suite de ce peptide et une terminaison classique peut avoir lieu au niveau du codon stop de l'ARNtm.

Il est possible que le *drop-off* contribue également à la libération des ribosomes bloqués sur des ARNm tronqués. En effet, la surexpression de l'ARNtm permet la survie de souche d'*E. coli pth<sup>ts</sup>* à une température non permissive, ce qui suggère que la libération de peptidyl-ARNt est diminuée lorsque l'ARNtm est surproduit. Réciproquement, l'inactivation du gène codant l'ARNtm (*ssrA*) aggrave le phénotype thermosensible de cette souche, ce qui indique qu'en absence d'ARNtm, la quantité de peptidyl-ARNt libérée par *drop-off* est accrue (Singh and Varshney, 2004).

On peut donc imaginer que le phénomène de *drop-off* est concurrentiel de la trans-traduction et que la présence d'ARN messagers tronqués est à l'origine de la libération de peptidyl-ARNt dans la cellule.



## **5. Les peptidyl-ARNt hydrolases**



Même si les divers mécanismes conduisant au *drop-off* ne sont pas encore clairement établis à ce jour, il paraît certain que c'est un phénomène fréquent au moins chez les bactéries. Le fait que de nombreux facteurs puissent le catalyser suggère qu'il n'est pas seulement un phénomène accidentel mais qu'il peut être utile, vraisemblablement pour éviter ou corriger des erreurs. Il pose cependant le problème de la production de peptidyl-ARNt dans la cellule, et nous allons maintenant nous intéresser à la PTH qui est l'enzyme responsable du recyclage des ARNt ainsi séquestrés.

## 5.1 La peptidyl-ARNt hydrolase bactérienne

Les peptidyl-ARNt produits lors de la traduction, par le *drop-off*, ne semblent avoir aucune fonction dans la cellule. Ils pourraient être dégradés entièrement en nucléotides et en acides aminés mais cela serait coûteux en énergie et il est sûrement plus avantageux de recycler les ARNt afin qu'ils puissent servir à de nouveaux cycles d'élongation. Il existe pour cela une enzyme, la peptidyl-ARNt hydrolase (PTH), capable d'hydrolyser la liaison ester entre l'ARNt et le peptide, pour régénérer un ARNt fonctionnel. Une telle activité a été observée il y a une quarantaine d'années chez *E. coli* et chez la levure (Cuzin *et al.*, 1967 ; Kossel and RajBhandary, 1968). Depuis, la protéine d'*E. coli* a fait l'objet de nombreuses études biochimiques et structurales et son gène a été identifié en 1991 (Garcia-Villegas *et al.*, 1991). Des homologues de cette PTH (PTH1) sont présents dans toutes les bactéries mais également dans de nombreux eucaryotes.

### 5.1.1 Spécificité de la PTH bactérienne (PTH1)

La protéine PTH1 d'*E. coli* reconnaît les aminoacyl-ARNt dont le groupement N-terminal de l'acide aminé est modifié par l'addition d'un ou plusieurs autres acides aminés ou simplement, par celle d'un groupement acétyle ou formyle. Son activité spécifique augmente avec la taille de la chaîne peptidique accrochée à l'ARNt, jusqu'à une longueur de 3 ou 4 liaisons peptidiques (Shiloach *et al.*, 1975a). Ceci suggère que la protéine PTH1 interagit avec les 3 ou 4 premiers acides aminés des peptidyl-ARNt.

Les aminoacyl-ARNt dont le groupement N-terminal de l'acide aminé n'est pas modifié peuvent aussi être hydrolysés par la PTH1 mais avec une efficacité catalytique très faible ( $0,006 \mu\text{M}^{-1} \text{s}^{-1}$  pour le Tyr-ARNt<sup>Tyr</sup> contre  $0,47 \mu\text{M}^{-1} \text{s}^{-1}$  pour le N-acétyl-Tyr-ARNt<sup>Tyr</sup> (Soutourina *et al.*, 1999). Il est donc peu probable qu'une telle hydrolyse ait lieu de manière significative *in vivo* où les aminoacyl-ARNt sont, après leur synthèse, rapidement captés par EF-Tu.

Comme on l'a vu ci-dessus, PTH1 possède une architecture tridimensionnelle qui est proche de la D-aminoacyl-ARNt désacylase DTD2.

Pourtant, PTH1 n'est pas une D-aminoacyl-ARNt désacylase. Elle hydrolyse les aminoacyl-ARNt aminoacylés par des acides aminés D avec une efficacité catalytique 400 fois plus faible que celle de DTD1. D'autre part, PTH1 agit beaucoup plus efficacement lorsque l'acide aminé estérifié à l'ARNt fait partie de la série L (Kossel and RajBhandary, 1968). Le rapport des efficacités catalytiques pour un N-acétyl-L-aminoacyl-ARNt et pour un N-acétyl-D-aminoacyl-ARNt est ainsi égal à 150 (Soutourina *et al.*, 1999).

La protéine PTH1 est capable d'hydrolyser efficacement tous les L-aminoacyl-ARNt N-modifiés à l'exception du fMet-ARNt<sup>fMet</sup> (initiateur chez les bactéries) (Cuzin *et al.*, 1967 ; Kössel and RajBhandary, 1968). La présence de deux bases non appariées (C<sub>1</sub>A<sub>72</sub>) au sommet de la tige acceptrice de cet ARNt<sup>fMet</sup> est responsable de cette spécificité (Schulman and Pelka, 1975). En effet, si on remplace les bases C<sub>1</sub>A<sub>72</sub> de l'ARNt<sup>fMet</sup> par une paire C<sub>1</sub>G<sub>72</sub>, G<sub>1</sub>C<sub>72</sub>, A<sub>1</sub>U<sub>72</sub> ou U<sub>1</sub>A<sub>72</sub>, le fMet-ARNt<sup>fMet</sup> devient un bon substrat de la PTH. Réciproquement, l'introduction d'une paire C<sub>1</sub>A<sub>72</sub> dans l'acétyl-méthionyl-ARNt<sup>mMet</sup> ou dans l'acétyl-valyl-ARNt<sup>Val</sup> ralentit nettement l'hydrolyse de ces ARN par la PTH (Dutka *et al.*, 1993). La prise en charge rapide de l'ARNt<sup>fMet</sup> par IF2 (Guillon *et al.*, 1996 ; Thanedar *et al.*, 2000) participe aussi vraisemblablement au fait que le fMet-ARNt<sup>fMet</sup> n'est pas hydrolysé par la PTH *in vivo*.

Le phosphate à l'extrémité 5' des ARNt semble être un élément important de reconnaissance des ARNt par la PTH (Schulman and Pelka, 1975). Un ARNt porteur d'un groupement 5'-OH est ainsi hydrolysé 17 fois moins efficacement par la PTH qu'un ARNt possédant un groupement 5'-phosphate (Fromant *et al.*, 1999). Par contre, dans le cas du fMet-ARNt<sup>fMet</sup>, cet effet du phosphate 5'-terminal est aboli. Ce phosphate servirait donc de relais entre l'enzyme et l'état, apparié ou non, de la paire de bases 1-72 (Dutka *et al.*, 1993).

L'ARNt<sup>His</sup>, qui a la spécificité de posséder une base supplémentaire appariée en 5' est néanmoins un bon substrat de la PTH (Fromant *et al.*, 2000). Dans ce cas, le rôle joué par le phosphate 5'-terminal dans les autres ARNt est joué par le phosphate situé entre les nucléotides +1 et -1.

La PTH est donc vraisemblablement capable d'hydrolyser tous les peptidyl-ARNt issus du *drop-off*. Parmi les deux produits de la réaction, seul l'ARNt est un inhibiteur efficace du substrat (Shiloach *et al.*, 1975a ; Shiloach *et al.*, 1975b). Ceci suggère que l'ARNt établit les interactions primaires avec la PTH alors que la région peptide interagit avec la PTH pour positionner précisément le substrat dans le site actif de l'enzyme (Das and Varshney, 2006).

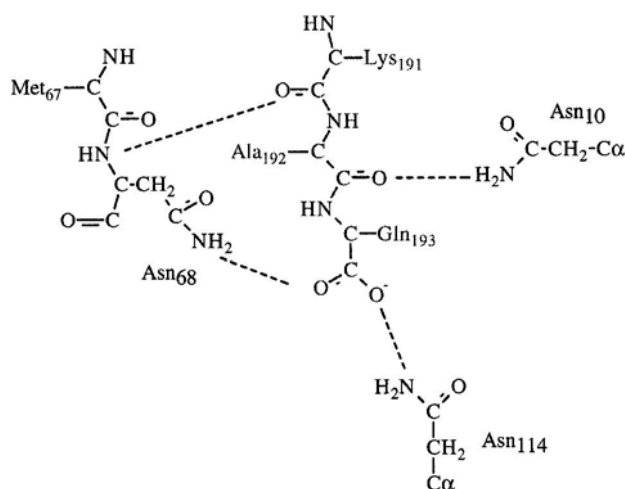
### 5.1.2 Structure et mécanisme des PTH bactériennes

La première PTH1 dont la structure a été résolue par radiocristallographie est celle d'*E. coli*, à une résolution de 1,2 Å (Schmitt *et al.*, 1997). Par la suite, la structure de la PTH de *Mycobacterium tuberculosis* a aussi été résolue par radiocristallographie et par RMN. La PTH1 forme un domaine globulaire  $\alpha/\beta$

construit autour d'un feuillet  $\beta$  entouré d'hélices  $\alpha$ . Des résidus strictement conservés dans les PTH1 délimitent un canal occupé dans le cristal de la PTH d'*E. coli* par l'extrémité C-terminale d'une molécule voisine. Des groupements carbonyles des trois derniers résidus de cette autre molécule établissent des interactions électrostatiques (liaison hydrogène) avec trois asparagines (N10, N68 et N114) (voir figure 21). La mutation de ces asparagines en alanines, diminue fortement la vitesse d'hydrolyse du diacétyl-lysyl-ARNt<sup>Lys</sup> par la PTH (d'environ deux ordres de grandeur). La mutation de N10 en acide glutamique a un effet encore plus important sur l'efficacité catalytique de l'enzyme (plus de trois ordres de grandeur, voir tableau 3), ce qui suggère que l'interaction de cette asparagine avec le substrat est due au groupement NH<sub>2</sub> de sa chaîne latérale qui permet la formation de liaisons hydrogène (voir figure 21). L'ensemble de ces résultats suggère que l'interaction de ces deux molécules de PTH dans le cristal pourrait mimer la formation d'un complexe entre la PTH et un des produits de la réaction d'hydrolyse, le peptide.

**Figure 21 :**  
**Interactions électrostatiques entre deux protéines voisines dans le cristal de la PTH d'*E. coli***

Les chaînes latérales des asparagines Asn10, Asn68 et Asn114 et la chaîne principale des trois résidus (191-193) de l'extrémité C-terminale d'une molécule voisine sont représentées en deux dimensions. Des pointillés représentent les liaisons hydrogène entre ces deux molécules (Fromant *et al.*, 1999).



Les chaînes latérales de N68 et N114 forment des liaisons hydrogène avec l'acide aminé C-terminal de la molécule voisine (Q193), tandis que celle de N10 interagit avec l'avant-dernier acide aminé (A192). L'importance de N10 dans l'hydrolyse du diacétyl-lysyl-ARNt<sup>Lys</sup>, où un seul acide aminé est lié à l'ARNt, suggère que Q193 se positionne au même endroit que la lysine de ce substrat modèle, permettant au groupement acétyle porté par le groupement  $\alpha$ -amino de ce substrat d'interagir avec N10. Ces asparagines joueraient donc un rôle important dans le mécanisme catalytique, en positionnant précisément la partie peptidique du substrat dans le site actif. La faible influence de leurs mutations en alanine sur le  $K_m$  confirme par ailleurs l'idée que l'affinité du substrat pour la PTH est essentiellement due aux interactions entre la protéine et l'ARNt.



Enzyme	k <sub>cat</sub> (s <sup>-1</sup> )	K <sub>m</sub> (μM)	k <sub>cat</sub> /K <sub>m</sub> relatif (c)	$\frac{k_{cat}/K_m \text{ (ARNt 5'-P)}}{k_{cat}/K_m \text{ (ARNt 5'-OH)}}$ (d)
sauvage (a)	3,6	6	100	17
N10A (a)	0,028	6,6	0,71	nd
Y15A (a)	0,93	9,6	16	17
T18A (a)	0,67	9,8	11	14
R19A (a)	0,13	28	0,77	16
H20A (a)	nm	nm	<0,8	nd
N21A (a)	0,16	9,5	2,8	14
R33A (a)	3,7	5,5	112	21
R35A (a)	4,3	7,9	91	16
F66A (a)	0,9	6,9	22	nd
M67A (a)	0,17	4,1	6,9	nd
N68A (a)	0,075	12,5	1	13
R82A (a)	4,5	7	107	18
D93A (a)	0,05	10	0,83	nd
K103A (a)	> 0,7	> 40	2,7	14
K105A (a)	> 0,25	> 40	1	2,9
H113A (a)	1,2	8,8	23	nd
N114A (a)	0,055	4,9	1,9	18
K117A (a)	1,6	21	13	16
R133A (a)	0,021	14	0,25	4,7
K142A (a)	> 1,8	> 30	10	13
K144A (a)	3,8	14	45	15
R170A (a)	3,5	7,5	78	18
R186A (a)	2,8	9,3	50	15
R133A-K105A (a)	0,008	22	0,061	1,7
sauvage (b)	9,3	5,5	100	nd
N10D (b)	0,005	7,9	0,037	nd
H20N (b)	nm	nm	< 0,0001	nd
H20Q (b)	nm	nm	< 0,0001	nd
M67E (b)	0,045	3,6	0,74	nd
D93N (b)	0,37	3,8	5,8	nd
K103R (b)	6,3	4,1	91	nd
K103Q (b)	5	23	13	nd
K103S (b)	2,6	37	4,2	nd
K105R (b)	2,4	8,7	16	nd
K105Q (b)	1,9	36	3,1	Nd
K105S (b)	1,5	30	3	Nd
H113Q (b)	9,1	24	22	Nd

**Tableau 3 : Paramètres catalytiques de mutants de la PTH d'*E. coli***

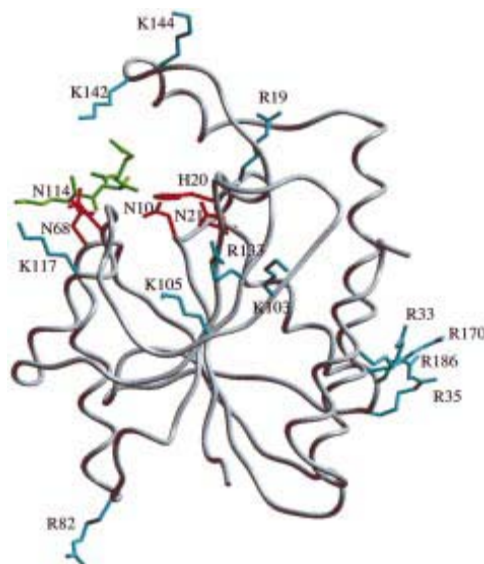
Pour chaque mutant, les valeurs  $k_{cat}$  et  $K_m$  proviennent de (a) (Fromant *et al.*, 1999) et (Schmitt *et al.*, 1997) et (b) (Goodall *et al.*, 2004) et concernent l'hydrolyse (a) de diacétyl-Lys-ARNt<sup>Lys</sup> et (b) du substrat Oregon Green-méthionyl-ARNt<sup>mMet</sup>. (c) Une valeur de 100 a été donnée au  $k_{cat}/K_m$  relatif obtenu avec la PTH sauvage et le substrat utilisé dans chacune des deux expériences. (d) Pour chaque mutant, la valeur indiquée est le rapport entre la valeur de  $k_{cat}/K_m$  obtenue pour l'hydrolyse de diacétyl-ARNt<sup>Lys</sup> 5'-phosphate et celle obtenue pour du diacétyl-ARNt<sup>Lys</sup> 5'-OH. nm : non mesurable. nd : non déterminé.

Comme on vient de le voir, dans les cristaux de la PTH d'*E. coli*, le résidu Q193 de la molécule de PTH voisine pourrait occuper la position de l'acide aminé estérifié à l'ARNt dans un peptidyl-ARNt. La position du groupement carboxyle de Q193 indique donc la position de liaison ester hydrolysée par la PTH. À moins de 5 Å du groupement carboxyle de Q193, se trouve une histidine, H20 chez *E. coli*, strictement conservée. La mutation de cette histidine a pour effet de diminuer fortement l'activité catalytique de la PTH, à tel point qu'elle devient indétectable (voir tableau 3). Dans le cas d'une mutation en alanine, cela correspond à une baisse d'activité de plus de 3 ordres de grandeurs. Dans le cas d'une mutation en asparagine, ou en glutamine, il a été conclu que la diminution de l'activité était supérieure à 6 ordres de grandeurs (voir tableau 3), alors que la protéine est toujours capable de fixer le substrat (Goodall *et al.*, 2004). Cette histidine rendue basique par son interaction avec D93, lui aussi strictement conservé, pourrait activer la molécule d'eau responsable de l'hydrolyse du substrat (Goodall *et al.*, 2004).

Dans la région du site actif putatif de la PTH, d'autres résidus (Y15, T18, R19, N21, F66, M67, H113, K117) semblent aussi jouer un rôle dans l'activité de la PTH puisque leur mutation en alanine affecte l'efficacité catalytique de cette enzyme (voir tableau 3). Ainsi, les mutations R19A et K117A entraînent une perte d'affinité assez importante pour le diacétyl-Lys-ARNt<sup>Lys</sup>, qui indique que les résidus positifs Arg19 et Lys117 jouent un rôle fort dans la stabilisation du complexe enzyme-substrat. Si R19 est localisé dans le site actif, au voisinage de H20, K117 est situé à la limite de ce site actif, du côté opposé à la région de fixation du peptide, suggérant une interaction avec une des bases de l'extrémité CCA de l'ARNt.

### Figure 22 : Localisation des résidus d'asparagine et des résidus basiques à la surface de la PTH

Le squelette de la PTH d'*E. coli* est représenté en gris. Les 3 résidus de l'extrémité C-terminale d'une molécule voisine sont représentés en vert. Les chaînes latérales des résidus N10, H20, N21, N68 et N114, appartenant au site actif putatif, sont représentées en rouge. Les chaînes latérales en bleu indique la position des résidus basiques situés à la surface de l'enzyme et étudiés par mutagenèse dirigée (tableau 3). Cette figure a été dessinée à l'aide du logiciel Setor (Evans, 1993) à partir du fichier PDB (2PTH) de la PTH d'*E. coli* (Fromant *et al.*, 1999).



Au-dessus du site actif, une région comprenant une boucle et une hélice (voir figure 22) adopte une conformation ouverte dans le cristal de la PTH de *Mycobacterium tuberculosis* (Selvaraj *et al.*, 2007), différente de la conformation fermée de cette région dans les cristaux de la PTH d'*E. coli*. Le site catalytique de

la PTH étant inoccupé dans le cristal de la PTH de *M. tuberculosis*, il se pourrait que les deux conformations reflètent une différence entre l'état lié et libre de l'enzyme. Une analyse par RMN de la PTH de *M. tuberculosis* a d'ailleurs montré une grande flexibilité de cette partie de la protéine en solution (Pulavarti *et al.*, 2008). Dans cette région, la lysine 142 chez *E. coli*, quoique peu conservée, et, dans une moindre mesure, la lysine 144, semblent jouer un rôle important dans la reconnaissance du substrat puisque leur mutation en alanine entraîne une augmentation du  $K_m$  pour le diacétyl-Lys-ARNt<sup>Lys</sup> d'un facteur 2,5 pour la mutation K144A et supérieur à 5 pour la mutation K142A. La lysine 142 est située à l'opposé de la région de fixation présumée du peptide, par rapport au site actif, à environ 10-15 Å de celui-ci. Ce résidu basique est donc un bon candidat pour interagir avec un groupement phosphate de l'extrémité de la tige acceptrice de l'ARNt.

Comme on l'a vu ci-dessus, le phosphate 5'-terminal de l'ARNt est un élément de reconnaissance majeur du substrat par PTH1. Ce phosphate semble stabilisé dans une pince formée par les chaînes latérales de la lysine 105 et l'arginine 133, chez *E. coli*. En effet, d'une part, ces deux résidus sont situés à une distance entre 15 et 20 Å du site actif présumé, ce qui correspond à une distance possible entre l'extrémité 3'-OH d'un ARNt et son groupement 5'-phosphate. D'autre part, les mutations en alanine des résidus 105 et 133 réduisent fortement l'activité catalytique de l'enzyme, d'un facteur 100 dans le cas de la mutation K105A et d'un facteur 400 dans le cas de la mutation R133A, ce qui souligne leur rôle majeur dans l'activité catalytique de la PTH. Enfin, la baisse d'efficacité catalytique due à l'absence du phosphate 5'-terminal est plus faible avec les mutants K105A et R133A qu'avec la PTH sauvage (voir tableau 3(d)). Dans le cas du double mutant K105A-R133A, cet effet est encore plus faible. Ainsi, l'activité catalytique de ce double mutant sur un substrat déphosphorylé n'est qu'1,7 fois plus faible que l'activité sur un mutant possédant un phosphate en 5', alors qu'elle est 17 fois plus faible pour la PTH sauvage. Cela montre donc que la contribution positive amenée par la présence du phosphate 5'-terminal est faible en l'absence des chaînes latérales de K105 et R133. Les chaînes latérales de ces deux résidus interagiraient donc directement avec le phosphate 5'-terminal de l'ARNt pour stabiliser le complexe enzyme:substrat. Cette interaction serait semblable à celle qui existe dans le cas d'EF-Tu où, là aussi, le phosphate 5'-terminal de l'ARNt interagit avec les chaînes latérales d'une lysine et d'une arginine (Nissen *et al.*, 1995). Cette similarité dans les modes de reconnaissance est sans doute à rapprocher du fait que EF-Tu, comme la PTH, doit distinguer l'ARNt initiateur, ne possédant pas ce phosphate, des autres ARNt.

Dans la région des résidus 105 et 133, la mutation de la lysine 103 en alanine, chez *E. coli*, entraîne une augmentation du  $K_m$  supérieure à un facteur 5 alors que sa mutation en arginine n'a que peu d'effet (voir tableau 3). Cependant, cette lysine ne semble pas en interaction avec le phosphate 5'-terminal, puisque la baisse d'activité due à la mutation K103A est la même pour un ARNt déphosphorylé en 5' et pour un ARNt portant un groupe phosphate en 5'. Vu sa

position, la lysine 103 pourrait interagir avec un autre phosphate du bras accepteur de l'ARNt.

Chez *E. coli*, d'autres arginines (R33, R35, R82, R170, R186) situées à la surface de la protéine ont été mutées en alanine. Aucune ne semble jouer un rôle important dans la reconnaissance du substrat.

Les acides aminés C-terminaux de la PTH forment, dans les cristaux de l'enzyme d'*E. coli*, une hélice  $\alpha$ . Par analogie avec la structure du complexe entre EF-Tu et le Phe-ARNt<sup>Phe</sup> (Nissen *et al.*, 1995), cette hélice pourrait être en contact avec la région de la tige T $\Psi$ C de l'ARNt (Fromant *et al.*, 1999). On peut cependant noter que cette hélice n'a pas été observée dans l'étude par cristallographie de la PTH de *M. tuberculosis*, suggérant que cette région était désordonnée ou qu'elle pouvait adopter plusieurs conformations dans le cristal (Selvaraj *et al.*, 2007). De plus, l'extrémité C-terminale est assez variable d'une PTH à l'autre, même si l'on y retrouve toujours un certain nombre de résidus basiques susceptibles d'interagir avec l'ARNt. Son rôle dans la fixation de l'ARNt doit donc encore être démontré.

## 5.2 La peptidyl-ARNt hydrolase archéenne

Des recherches dans les banques de séquences montrent que, chez les archées, il n'existe aucun gène présentant des homologies significatives avec les PTH bactériennes. Les archées ne sont cependant pas dépourvues d'activité PTH, mais elles expriment une PTH (PTH2) dont la séquence est très différente de celles des PTH bactériennes. La protéine PTH2 est néanmoins fonctionnellement interchangeable avec la PTH bactérienne PTH1, puisque l'expression de PTH2 chez *E. coli* permet de compléter l'absence de PTH endogène (Fromant *et al.*, 2003 ; Rosas-Sandoval *et al.*, 2002).

### 5.2.1 Spécificité du substrat de PTH2

Comme la PTH bactérienne, la PTH archéenne hydrolyse les aminoacyl-ARNt dont le groupement N-terminal de l'acide aminé est bloqué (Rosas-Sandoval *et al.*, 2002). Si l'acide aminé n'est pas bloqué, la vitesse d'hydrolyse est 200 fois plus faible (Fromant, M., Plateau, P., Blanquet S., résultats non publiés). Les Met-ARNt<sup>Met</sup> initiateurs des archées ne sont pas formylés et ne possèdent pas le mésappariement C1-A72 décrit ci-dessus (Bult *et al.*, 1996). La PTH archéenne n'a donc pas à exclure le fMet-ARNt<sup>fMet</sup>. Ainsi, contrairement à la PTH bactérienne, la PTH archéenne est capable d'hydrolyser le fMet-ARNt<sup>fMet</sup>. Les mécanismes de reconnaissance des ARNt par les PTH bactérienne et archéenne sont donc différents.

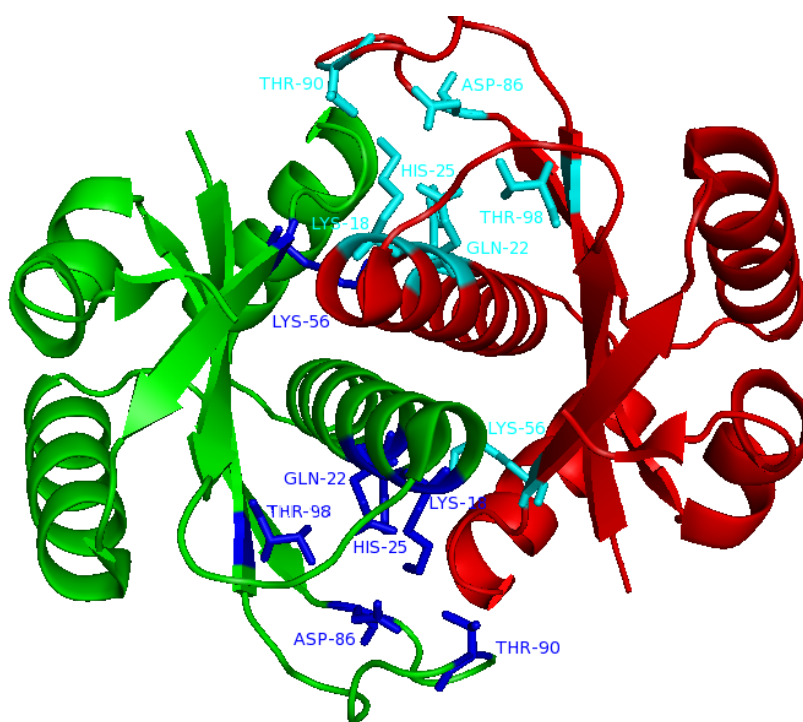
Par ailleurs, l'ARNt non aminoacylé est un inhibiteur bien plus efficace pour l'activité de la PTH2 que pour celle de la PTH1 (il existe un facteur supérieur à 100 entre les 2  $K_i$  (Fromant *et al.*, 2003)), ce qui suggère que la PTH archéenne possède des interactions plus fortes avec les ARNt ou la partie ARNt des peptidyl-ARNt.

### 5.2.2 Structure et mécanisme des PTH archéennes

Bien que les PTH2 n'aient été découvertes qu'en 2002 (Fromant *et al.*, 2003 ; Rosas-Sandoval *et al.*, 2002), plusieurs structures de PTH de type 2 ont été résolues par cristallographie ou par RMN. On dispose ainsi de celles des PTH2 de *Sulfolobus solfataricus* (code PDB 1XTY) (Fromant *et al.*, 2005), de *Thermoplasma acidophilum* (code PDB 1RLK), d'*Archaeoglobus fulgidis* (code PDB 1RZW) (Powers *et al.*, 2005), de *Pyrococcus horikoshii* (code PDB 2D3K) (Shimizu *et al.*, 2008), de *Methanocaldococcus jannaschii* (code PDB 2ZV3), ainsi que de la structure d'une protéine humaine (Bit1) homologue aux PTH archéennes (code PDB 1Q7S) (De Pereda *et al.*, 2004). Toutes ces structures sont très proches les unes des autres. Ces protéines forment un dimère symétrique, chaque protomère étant constitué d'un feuillet  $\beta$  entouré de deux groupes de deux

hélices alpha (voir figure 23). L'interface des deux sous-unités est formée par des contacts de van der Waals entre des résidus hydrophobes de l'hélice  $\alpha 1$  de chaque protomère, très conservés dans les PTH2, notamment A20, A24, A27, V28, V31, I135 et K16, via la partie aliphatique de sa chaîne latérale (numérotation *S. solfataricus*).

Comme pour la PTH1, aucune structure du complexe entre la PTH2 et son substrat n'a été obtenue à ce jour. Les mécanismes de reconnaissance et d'hydrolyse du substrat ne sont donc pas encore totalement élucidés. Cependant, un certain nombre de résidus très conservés (K18, Q22, H25, K56, D86, T90 et T98 chez *S. solfataricus*) (Fromant *et al.*, 2005) sont regroupés dans une région comprenant l'extrémité N-terminale de l'hélice  $\alpha 1$  et la boucle entre le feuillet  $\beta 3$  et  $\beta 4$  d'un protomère, ainsi que les brins  $\beta 2$  et  $\beta 5$  de l'autre protomère. Lorsque ces résidus sont mutés en alanine, l'activité de la PTH2 est réduite, ce qui montre que cette région correspond vraisemblablement au site actif.



**Figure 23 : Structure de la PTH archéenne**

Représentation à l'aide du logiciel Pymol, de la structure cristallographique de la PTH de *S. solfataricus* (code PDB 1XTY). Le protomère A est coloré en vert et le protomère B en rouge. Les chaînes latérales de résidus conservés (K18, Q22, H25, K56, D86, T90 et T98) sont représentées en bleu foncé pour le protomère A et en bleu clair pour le protomère B.

Les résidus K18, D86 et T90 semblent jouer un rôle prépondérant dans le mécanisme catalytique puisque la vitesse des PTH2 où ces résidus sont mutés en alanine est réduite de deux (pour T90) à trois ordres de grandeur (pour K18 et D86), alors que le  $K_m$  pour le diacétyl-Lys-ARNt<sup>Lys</sup> est peu affecté (Fromant *et al.*, 2005). K18 et D86 pourraient jouer un rôle via leurs fonctions acido-basiques alors que le groupement hydroxyle de T90 pourrait agir en tant que nucléophile.

La proximité des résidus K18 et D86, dont les chaînes latérales interagissent via un pont salin, et le fait que le double mutant KA18-DA86 n'est pas plus affecté que les simples mutants, suggèrent que D86 pourrait diminuer la valeur du  $pK_a$  de K18, ce qui renforcerait sa réactivité pour l'attaque du groupement carbonyle de la liaison ester du substrat. Cette activation pourrait rappeler l'activation de H20 par D93 dans la PTH d'*E. coli*, mais la ressemblance de ces sites actifs s'arrête là, puisqu'ils sont structurellement très différents et qu'on ne retrouve pas ici les asparagines qui interagissent chez *E. coli* avec la partie peptidique du substrat.

Dans les cristaux de la PTH2 de *S. solfataricus* et de *T. acidophilum*, un ion sulfate semble se loger au voisinage du site actif, stabilisé par les chaînes latérales de Q22, de D86 (via une molécule d'eau), de K56 (via deux molécules d'eau) et par le groupement amide de A87. Cet ion sulfate pourrait mimer un phosphate de l'extrémité CCA du substrat qui serait ainsi stabilisé dans le site actif.

La forte concentration de résidus chargés positivement sur une face de la PTH2, au voisinage de ce site actif, pourrait jouer un rôle dans la forte fixation de l'ARNt sur la PTH2, mais cela reste à confirmer. Par contre, à la différence de la PTH1 où le phosphate de l'extrémité 5' de l'ARNt est spécifiquement reconnu par une pince formée de deux résidus basiques très conservés (R133 et K105 chez *E. coli*), aucun résidu basique pouvant jouer ce rôle n'est conservé à la surface de la PTH2. On peut d'ailleurs noter que le  $K_m$  de la PTH2 pour un diacétyl-Lys-ARNt<sup>Lys</sup> déphosphorylé est légèrement inférieur à celui d'un diacétyl-Lys-ARNt<sup>Lys</sup> possédant un phosphate en 5' (Fromant *et al.*, 2003), ce qui indique que ce phosphate n'a pas le rôle important qu'il joue dans l'interaction avec PTH1.

## 5.3 Les peptidyl-ARNt hydrolases chez les eucaryotes

### 5.3.1 Localisation intracellulaire des PTH1 et PTH2 eucaryotes

Chez les eucaryotes, il existe fréquemment plusieurs gènes homologues à ceux des PTH bactériennes ou archéennes. Ainsi, la levure *Saccharomyces cerevisiae* possède un homologue de la PTH bactérienne (PTH1) et un homologue de la PTH archéenne (PTH2), tous deux capables de compléter une souche d'*E. coli* déficiente en PTH. Il est bien sûr envisageable que l'une de ces PTH soit mitochondriale et l'autre cytoplasmique, mais il n'est pas sûr que ce soit le cas. Ainsi, l'inactivation du gène codant la PTH1 ralentit fortement la croissance de la levure dans des conditions où une activité mitochondriale est nécessaire (Fromant *et al.*, 2003). De plus, la caractérisation systématique du protéome mitochondrial de *S. cerevisiae* indique que PTH1 en fait partie (Sickmann *et al.*, 2003 ; Steinmetz *et al.*, 2002). On peut aussi remarquer que dans la levure *Schizosaccharomyces pombe* qui possède, elle aussi, un homologue de PTH1 et un homologue de PTH2, l'homologue de PTH1 possède une extension N-terminale suggérant un motif d'adressage mitochondrial. Quant à la PTH2, elle possède chez les eucaryotes une extension N-terminale, absente chez les archées, qui pourrait être une séquence d'adressage (Rosas-Sandoval *et al.*, 2002). De plus, elle a été retrouvée dans le protéome mitochondrial de *S. cerevisiae* (Sickmann *et al.*, 2003). Il se pourrait donc que la PTH1 et la PTH2 soient toutes les deux mitochondriales.

### 5.3.2 Autres fonctions des homologues de PTH1 et PTH2 chez les eucaryotes

La redondance de protéines possédant une activité PTH chez les eucaryotes s'explique peut-être par le fait que certaines de ces enzymes interviennent dans des processus très différents de l'hydrolyse des peptidyl-ARNt.

Ainsi, chez l'homme, la protéine Bit1 (*Bcl2-inhibitor of transcription*), est homologue à la PTH2 archéenne. *In vitro*, cette protéine est capable d'hydrolyser le diacétyl-lysyl-ARNt<sup>Lys</sup> avec une bonne efficacité catalytique ( $k_{cat}/K_m = 5,9 \mu\text{M}^{-1} \text{s}^{-1}$ ) (De Pereda *et al.*, 2004). La protéine Bit1 possède un domaine d'adressage mitochondrial à son extrémité N-terminale. Effectivement, des expériences de microscopie confocale sur des cellules exprimant une protéine de fusion entre Bit1 et la GFP (*green fluorescent protein*) indiquent que Bit1 est mitochondriale. Mais cette protéine est localisée dans l'espace intermembranaire (Jan *et al.*, 2004), c'est-à-dire dans un compartiment où on ne s'attend pas à avoir de production de peptidyl-ARNt.

En fait, la protéine Bit1 semble jouer un rôle dans l'apoptose quand elle est libérée dans le cytoplasme. Elle forme alors un complexe pro-apoptotique avec le



régulateur transcriptionnel AES (*Amino-terminal Enhancer of Split*), qui est un membre de la famille des facteurs de transcription Groucho. Par un mécanisme non encore complètement élucidé (Jan *et al.*, 2004 ; Kairouz-Wahbe *et al.*, 2008), ce complexe déclenche alors la mort cellulaire par une voie indépendante de celle des caspases. Actuellement aucun indice ne permet d'imaginer un lien entre l'activité PTH de Bit1 et son rôle dans l'apoptose (Kairouz-Wahbe *et al.*, 2008), ce qui fait que Bit1 est considérée comme une protéine multi-fonctionnelle (*moonlighting protein* en anglais, du verbe *to moonlight* : avoir un second métier, travailler au noir). On ignore cependant quand l'activité peptidyl-ARNt hydrolase de Bit1 pourrait être utile à la cellule.

Chez les plantes, il existe fréquemment plusieurs homologues de *pth1* et de *pth2* dans le même organisme (4 homologues de *pth1* et 3 homologues de *pth2* chez *Arabidopsis thaliana*). Certains de ces homologues pourraient posséder d'autres fonctions que l'hydrolyse de peptidyl-ARNt. C'est le cas par exemple de la protéine CRS2 du maïs (*Chloroplast RNA Splicing 2*), homologue aux PTH bactériennes, qui intervient dans l'épissage de certains introns chloroplastiques. CRS2 possède 50% d'identité avec la PTH d'*E. coli*. Néanmoins, elle ne semble pas avoir d'activité peptidyl-ARNt hydrolase, puisqu'elle ne peut pas compléter une souche d'*E. coli* déficiente en PTH (Jenkins and Barkan, 2001), sans que l'on sache exactement pourquoi. Les structures des deux enzymes sont en effet très proches, notamment au niveau du site actif, où la plupart des résidus importants pour l'activité de la PTH d'*E. coli* (N10, Y15, T18, R19, H20, N21, M67, N68, D93, H113, N114) sont conservés dans CRS2 et positionnés de manière identique. Une différence existe au niveau de la pince K105-R133 qui, chez *E. coli*, interagit avec le phosphate 5'-terminal des ARNt, et qui est absente dans CRS2. A la place de la lysine 105, on trouve une glutamine dans CRS2 et, à la place de l'arginine 133, on trouve une sérine. Cependant, il ne semble pas que ces substitutions expliquent la différence de spécificité des deux enzymes. En effet, ni CRS2 sauvage, ni le double mutant Q105K-S133R, ne complémente une souche d'*E. coli pth<sup>ts</sup>* alors qu'une PTH d'*E. coli* portant les deux mutations K105Q et R133S complémente cette souche. Cette différence ne semble donc pas responsable de la perte de fonction de CRS2 envers le recyclage des peptidyl-ARNt.

Dans les chloroplastes du maïs, la protéine CRS2 s'associe à la protéine CAF1 (*CRS2-associated factor 1*) ou à la protéine CAF2, et cette association est indispensable pour que l'activité d'épissage de CRS2 puisse s'exercer (Ostheimer *et al.*, 2003). Il est donc possible que CRS2 possède une activité PTH uniquement lorsque les complexes CRS2-CAF1 ou CRS2-CAF2 se forment. Pour tester cette hypothèse, une expérience a été réalisée afin de déterminer si la co-expression des protéines CRS2 et CAF2 chez *E. coli* permettait de compléter une mutation *pth<sup>ts</sup>*. Aucune complémentation n'a été observée, mais ce résultat n'est pas vraiment concluant dans la mesure où, dans cette expérience, la production de la protéine CAF2 chez *E. coli* n'a pas pu être mise en évidence (Ostheimer *et al.*, 2003).

## 5.4 Autres protéines agissant sur les peptidyl-ARNt

Dans les réticulocytes de lapin, Gross *et al.* (1992) ont décrit une activité différente des PTH connues, capable de cliver les peptidyl-ARNt entre les deux nucléotides C75 et A76 de leur extrémité 3'-terminale. Cette enzyme serait donc une phosphodiesterase et un mécanisme de recyclage des ARNt impliquant cette enzyme nécessiterait ensuite l'action d'une ARNt nucléotidyl-transférase pour régénérer une extrémité 3' complète et, donc, un ARNt fonctionnel. Le gène codant la protéine responsable de cette activité phosphodiesterase n'a cependant pas été identifié à ce jour. Le rôle de cette enzyme *in vivo* n'est donc pas clairement établi.

La délétion des gènes *pth1* et *pth2*, séparément ou simultanément, n'empêche pas la levure de se développer. Il pourrait exister une autre enzyme capable de recycler les peptidyl-ARNt chez la levure, notamment dans son cytosol. Des extraits bruts d'une souche de *S. cerevisiae* où les deux gènes de la PTH1 et de la PTH2 ont été inactivés possèdent en effet toujours une activité capable d'hydrolyser un diacétyl-lysyl-ARNt<sup>Lys</sup> (Giorgi, L., Plateau, P., Blanquet, S., résultats non publiés). Nos tentatives de purification des protéines responsables de cette activité, réalisées pendant mon Master, ont montré que deux enzymes pourraient contribuer à cette activité : la RNase T2 et la carboxypeptidase Y. Mais comme la RNase T2 est une enzyme sécrétée et que la carboxypeptidase Y est localisée dans la vacuole, il est peu probable que ces enzymes interviennent pour hydrolyser des peptidyl-ARNt dans le cytoplasme.

La protéinase K, une protéase produite par le champignon *Engyodontium album* (précédemment appelé *Tritirachium album*) est, elle aussi, capable de cliver les aminoacyl-ARNt N-bloqués, en libérant un ARNt fonctionnel (Vidales *et al.*, 1979). Par contre, elle n'agit que faiblement sur les aminoacyl-ARNt non bloqués. Cependant, comme dans le cas de la RNase T2, il s'agit d'une enzyme sécrétée qui n'a donc sans doute pas de rôle dans l'hydrolyse des peptidyl-ARNt cytoplasmiques.

Récemment, Takaku *et al.* ont montré que, contrairement à la plupart des RNases Z, la RNase Z d'*E. coli* et la RNase Z<sup>L</sup> humaine pouvaient cliver l'extrémité CCA d'ARNt matures ou d'ARNt aminoacylés (Takaku and Nashimoto, 2008). *In vivo*, lorsque la RNase Z d'*E. coli* est fortement surproduite, la croissance des cellules est ralentie et il y a alors une accumulation, à l'intérieur des cellules, d'ARNt raccourcis à leur extrémité 3' (Takaku and Nashimoto, 2008). Si cette enzyme peut hydrolyser les aminoacyl-ARNt, il est vraisemblable qu'elle puisse agir également sur les peptidyl-ARNt, mais aucune expérience n'a encore été réalisée pour tester cette hypothèse.



# **Objectifs de la Thèse**



Deux types structurellement très différents de PTH ont été caractérisés : la PTH1, présente dans les bactéries, et la PTH2 retrouvée chez les archées. Ces deux types d'enzyme sont fonctionnellement interchangeable et pourtant, il n'y a apparemment pas eu de transferts de gènes horizontaux entre les bactéries et les archées. Toutes les bactéries expriment systématiquement une protéine du type PTH1, alors que toutes les archées expriment une protéine de type PTH2. Cette situation, qui contraste avec les nombreux transferts de gènes qui ont eu lieu dans le cas des autres enzymes de corrections d'erreurs évoqués dans les paragraphes précédents ou dans le cas des aminoacyl-ARNt synthétases, suggère qu'il existe des différences significatives non encore identifiées entre les machineries de traduction des bactéries et des archées, qui font que les PTH bactériennes sont mieux adaptées à la machinerie bactérienne et que les PTH archéennes conviennent mieux à la machinerie archéenne.

Les eucaryotes expriment généralement à la fois des PTH de type bactérien et des PTH de type archéen. Certaines de ces enzymes, voire toutes, sont localisées dans des organites intracellulaires (mitochondries, chloroplastes). Chez la levure, les homologues de PTH1 et de PTH2 ne sont pas indispensables à la croissance puisqu'une souche  $\Delta pth1 \Delta pth2$  est viable. Il est donc possible qu'une troisième forme de PTH existe dans le cytoplasme des eucaryotes. Cependant, il n'existe pas de raison justifiant *a priori* l'existence d'une telle protéine. En effet, le ribosome possède une activité d'hydrolyse des peptidyl-ARNt, utilisée lors de la terminaison de la traduction. Il est donc envisageable que, chez les eucaryotes, ce soit cette activité qui soit responsable du recyclage des peptidyl-ARNt et, donc, que la machinerie de traduction eucaryote puisse se passer d'une PTH équivalente à celle des bactéries ou des archées.

Dans tous les cas, il apparaît que la protéine PTH1 est absolument indispensable à la survie des bactéries, alors que cette protéine n'est pas essentielle à la croissance des eucaryotes. Ceci fait de PTH1 une cible antibactérienne prometteuse pour le développement d'inhibiteurs spécifiques. Pour concevoir un tel inhibiteur, il serait utile d'avoir une idée précise du mécanisme de fixation du substrat et le mécanisme de sa catalyse par PTH1. Il est nécessaire pour cela de connaître la structure haute résolution du complexe entre PTH1 et son substrat. Or, jusqu'à présent, aucun laboratoire n'a réussi à obtenir des cristaux d'un complexe entre PTH1 et son substrat. C'est pourquoi, nous avons choisi d'étudier un tel complexe en solution, par RMN.

La RMN permet de connaître les fréquences de résonance des atomes de la protéine possédant un spin nucléaire  $\frac{1}{2}$  (donc de ses atomes d'hydrogène, et d'azote et/ou de carbone si elle est isotopiquement marquée au  $^{15}\text{N}$  et/ou au  $^{13}\text{C}$ ). Via des transferts d'aimantation, la RMN permet aussi de déterminer quels atomes sont proches les uns des autres, ce qui en fait un bon outil d'étude structurale. La fréquence de résonance de chaque atome est spécifique et dépend de sa nature et de son environnement chimique. Comme celui-ci peut être modifié lors de l'interaction de la protéine avec un ligand, la RMN permet aussi d'étudier la formation de complexes.

Ce travail de thèse a donc eu pour but, dans un premier temps, d'identifier les fréquences de résonance des différents atomes de la PTH d'*E. coli* et d'analyser sa structure en solution. Ensuite, nous avons étudié l'interaction de la PTH avec différents substrats toujours par RMN ainsi que grâce à des méthodes biochimiques, ce qui nous permet finalement de proposer un modèle du complexe entre la PTH et son substrat et de déterminer le rôle joué par un grand nombre de résidus dans l'activité catalytique de cette enzyme. Ces résultats pourront être utilisés pour une recherche guidée d'inhibiteurs de la PTH bactérienne.

# Résultats





# **Chapitre 1**

## **Structure de la PTH d'*E. coli* en solution**



## 1.1 Préparation de la protéine pour la RMN

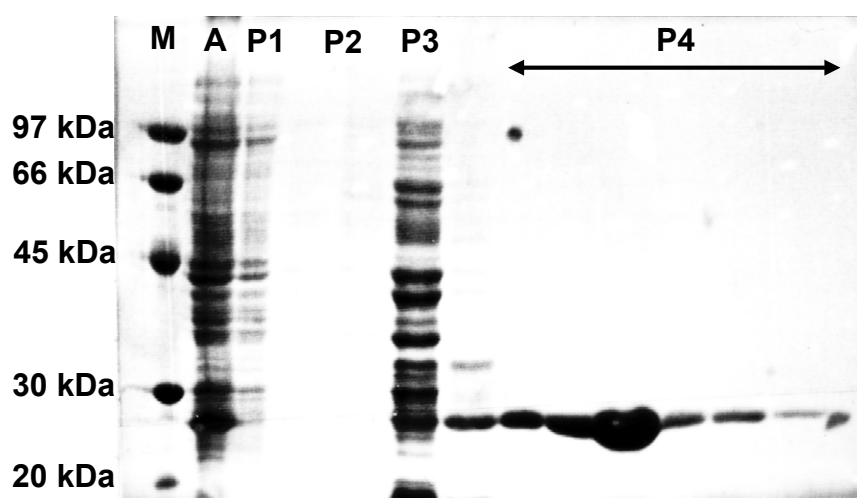
Avant de réaliser des expériences de RMN sur le complexe entre la peptidyl-ARNt hydrolase et des analogues de son substrat, nous avons étudié la protéine seule, en solution. Nous avons choisi de travailler sur une protéine mutante, dans laquelle une histidine du site actif (His20) est remplacée par une alanine. Cette mutation diminue de plusieurs ordres de grandeur le  $k_{cat}$  de la réaction d'hydrolyse catalysée par la PTH (Goodall *et al.*, 2004 ; Schmitt *et al.*, 1997) mais n'affecte pas la fixation des peptidyl-ARNt par la protéine (Goodall *et al.*, 2004). De plus, Bal *et al.* qui ont étudié par RMN l'impact d'une telle mutation sur la PTH de *M. tuberculosis*, ont conclu que ce mutant était structuralement semblable à la PTH sauvage (Bal *et al.*, 2007). L'intérêt du variant H20A vient du fait qu'il rend possible des expériences avec des composés qui seraient hydrolysés par l'enzyme sauvage pendant les expériences de RMN qui peuvent durer plusieurs heures.

### 1.1.1 Production et purification de la PTH H20A

Nous avons donc entrepris la production du mutant H20A de la PTH d'*E. coli*. Le gène codant ce variant avait déjà été construit au laboratoire (Schmitt *et al.*, 1997). Pour simplifier la purification de la protéine, nous avons ajouté à ce gène un fragment d'ADN codant un peptide contenant 6 histidines consécutives, permettant à la protéine d'être purifiée par affinité sur une colonne portant des ions cobalt.

Des études précédentes ayant suggéré que l'hélice C-terminale de la PTH pourrait jouer un rôle dans la fixation de son substrat, nous avons choisi de placer ce peptide de poly-histidine à l'extrémité N-terminale de la protéine. Entre les six histidines et la protéine, nous avons inséré les acides aminés correspondant à un site de reconnaissance de la thrombine. Une séquence de 20 acides aminés a donc été ajoutée à l'extrémité N-terminale : MGSSHHHHHSSGLVPRGSH.

Nous avons constaté que la surproduction de PTH H20A ainsi étiquetée n'était pas toxique pour la cellule. A partir d'un litre de culture (dans un milieu minimum) et après purification, nous obtenons environ 2 mg de PTH H20A, avec une pureté supérieure à 95%, d'après une analyse par électrophorèse sur gel de polyacrylamide en conditions dénaturantes (voir figure 24).



**Figure 24 : Analyse de la purification de la PTH par électrophorèse sur gel de polyacrylamide en conditions dénaturantes**

Gel d'électrophorèse coloré au bleu de Coomassie, réalisé avec 6  $\mu\text{l}$  de différents échantillons prélevés au cours de la purification de la PTH. M : marqueur de poids moléculaire. A : protéines appliquées sur la colonne Talon. P1, P2, P3, P4 : échantillons des fractions d'éluion de la chromatographie correspondent aux protéines se décrochant à une concentration en imidazole de 0 mM, 5 mM, 25 mM, 50 mM, respectivement. Pour P4, 6  $\mu\text{l}$  de différentes fractions, obtenues lors de l'éluion à 50 mM d'imidazole, ont été déposées séparément sur le gel.

Pour déterminer si le peptide de poly-histidine perturbait la fixation du substrat, nous avons aussi construit une souche produisant une protéine sauvage fusionnée à ce même peptide. Nous avons ensuite comparé l'efficacité catalytique ( $k_{\text{cat}}/K_m$ ) de la PTH sauvage étiquetée ou non. Dans les mêmes conditions de dosage (28°C, 20 mM Tris-HCl (pH 7,5), 10 mM  $\text{MgCl}_2$ , 0,1 mM EDTA et 0,1 mM dithiothréitol), nous avons constaté que les paramètres  $k_{\text{cat}}$  et  $K_m$  de l'enzyme étiquetée étaient diminués d'un facteur 2 à 3 seulement par rapport à ceux de l'enzyme sauvage non étiquetée. Les efficacités catalytiques ( $k_{\text{cat}}/K_m$ ) de ces deux enzymes sont très proches :  $0,59 \text{ s}^{-1} \mu\text{M}^{-1}$  pour la PTH étiquetée et de  $0,56 \text{ s}^{-1} \mu\text{M}^{-1}$  pour la PTH non étiquetée. Ces résultats nous ont montré que la PTH étiquetée était bien active et qu'elle pouvait donc être utilisée pour notre étude par RMN.

### 1.1.2 Choix des conditions pour la RMN

Pour cartographier par RMN les sites d'interaction entre une protéine et un ligand de celle-ci, il est indispensable de disposer de l'attribution des fréquences de résonance des noyaux observables : ceux des protons  $^1\text{H}$ , ceux des atomes d'azote  $^{15}\text{N}$  et, éventuellement, ceux des atomes de carbone  $^{13}\text{C}$ .

Pour des petites protéines (de quelques dizaines de résidus) l'attribution des déplacements chimiques des protons, pouvait être faite à partir d'expériences de RMN homonucléaires selon la méthode proposée par Wüthrich dans son

ouvrage « NMR of Proteins and Nucleic Acids » (Wüthrich, 1986). Mais quand la protéine est plus grande, la superposition des signaux des différents protons rend cette analyse impossible. De plus, la vitesse de relaxation du signal augmente considérablement avec la taille de la protéine, ce qui provoque une importante diminution de l'intensité des signaux et leur élargissement.

Il est alors nécessaire d'utiliser des expériences de RMN hétéronucléaires qui corrélient les fréquences de résonance d'atomes de nature différente liés covalamment. Cela permet un éclatement des signaux dans plusieurs dimensions, ce qui limite le problème de la superposition des pics. Ces expériences ont été rendues possibles grâce à l'utilisation de sondes multi-noyaux permettant d'exciter simultanément les atomes de  $^1\text{H}$ , de  $^{13}\text{C}$  et de  $^{15}\text{N}$  et grâce au marquage isotopique des atomes de carbone et d'azote de protéines recombinantes produites par des organismes cultivés dans des milieux dont les sources de carbone et/ou d'azote contiennent uniquement du  $^{13}\text{C}$  ou du  $^{15}\text{N}$ . C'est donc ce que nous avons fait.

Ainsi, l'expérience RMN de référence pour l'étude structurale des protéines, la  $^{15}\text{N}$ -HSQC (pour *Heteronuclear Single Quantum Coherence*), corrélie les protons et les atomes de  $^{15}\text{N}$  de chaque groupement amide de la protéine. Son résultat peut être visualisé sous forme de spectres à deux dimensions où le déplacement chimique des protons est porté en abscisse et celui des atomes d'azote en ordonnée (voir figure 25). Des pics de résonance apparaissent pour chaque couple  $\text{H}^{\text{N}}-\text{N}$  de la protéine pour lesquels le proton et l'azote sont liés covalamment. Elle permet une sélection et un éclatement dans une deuxième dimension des signaux des différents  $\text{H}^{\text{N}}$ , ce qui permet généralement de les distinguer.

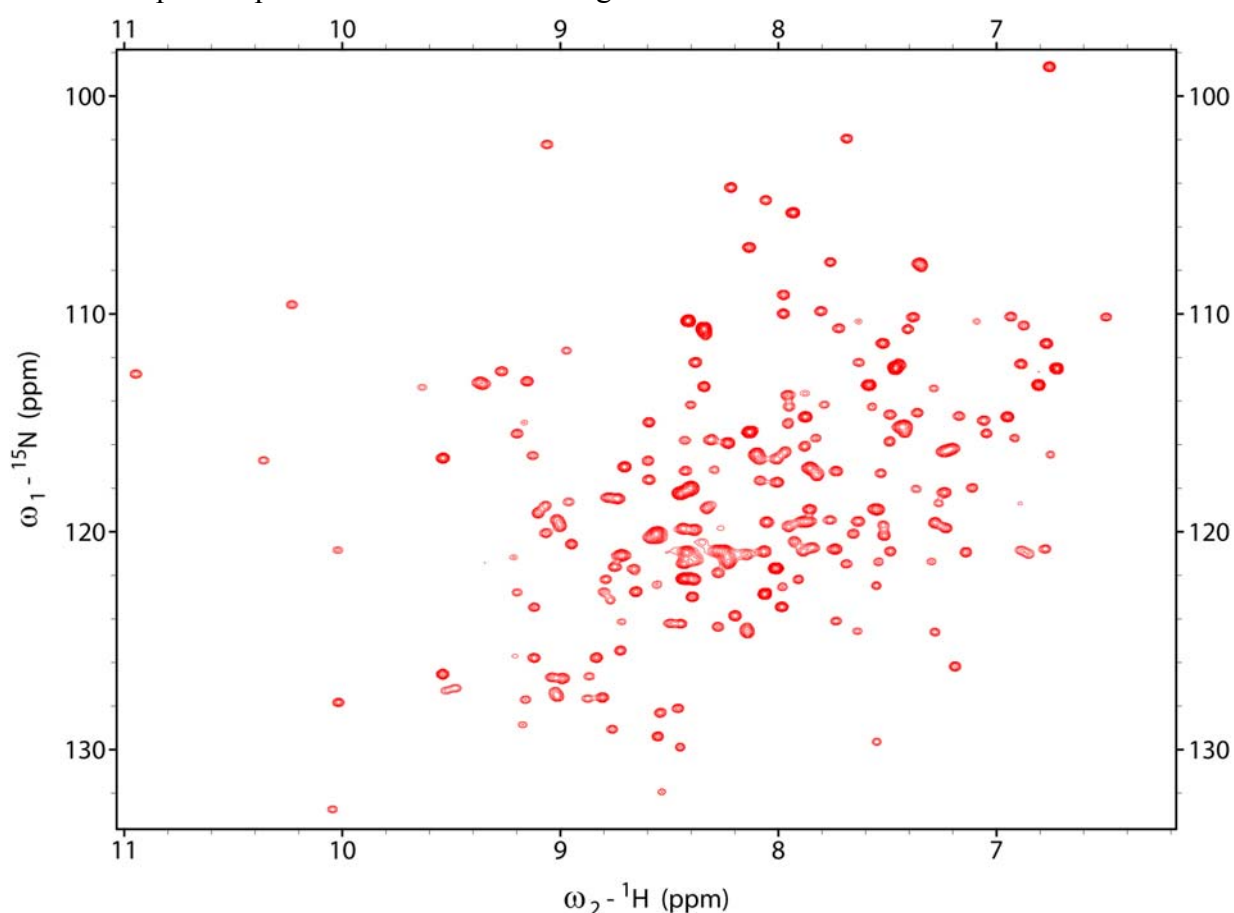
Les spectres de RMN dépendent des caractéristiques du tampon dans lequel la protéine est diluée (notamment sa force ionique et son pH) et de la température à laquelle est réalisée l'expérience. Il était donc important de choisir dès le début un tampon adapté à ce que nous voudrions faire plus tard, pour ne pas avoir à refaire l'attribution.

A pH acide, la vitesse d'échange des protons de la protéine avec ceux de l'eau est diminuée, ce qui améliore la qualité des spectres. Nous avons donc choisi un tampon contenant 50 mM d'acétate de sodium à pH 6,0 et nous avons vérifié que la PTH était active dans ce tampon. On perd seulement un facteur 3 dans l'efficacité catalytique par rapport à un dosage effectué en présence de 50 mM de Tris-HCl à pH 7,5. Nous avons aussi choisi de travailler à 17°C pour ne pas dénaturer la protéine au cours des expériences, les pics étant assez fins à cette température.

Souhaitant plus tard étudier un complexe entre la PTH et des analogues de son substrat, nous avons aussi cherché à trouver des conditions dans lesquelles une telle étude est possible. A forte concentration de protéine (nécessaire pour la RMN) et d'ARNt, nous avons observé l'apparition d'un précipité. En testant

différents tampons dans de petits volumes observés à la loupe binoculaire, nous nous sommes rendu compte que l'ajout de force ionique diminuait ce phénomène. Nous avons donc ajouté 200 mM de NaCl dans notre tampon. Nous avons vérifié que, conservée dans de telles conditions, (200 mM NaCl, 50 mM d'acétate de sodium pH 6,0 et à 17°C), la PTH sauvage conservait son activité catalytique pendant plusieurs jours sans être dégradée.

Malgré la taille de la protéine (193 résidus + 20 pour le tag), l'ensemble de ces conditions nous a permis d'obtenir des spectres  $^{15}\text{N}$ -HSQC de très bonne qualité. Comme on peut le voir sur la figure 25, la plupart des pics présents sur ces spectres peuvent en effet être distingués.



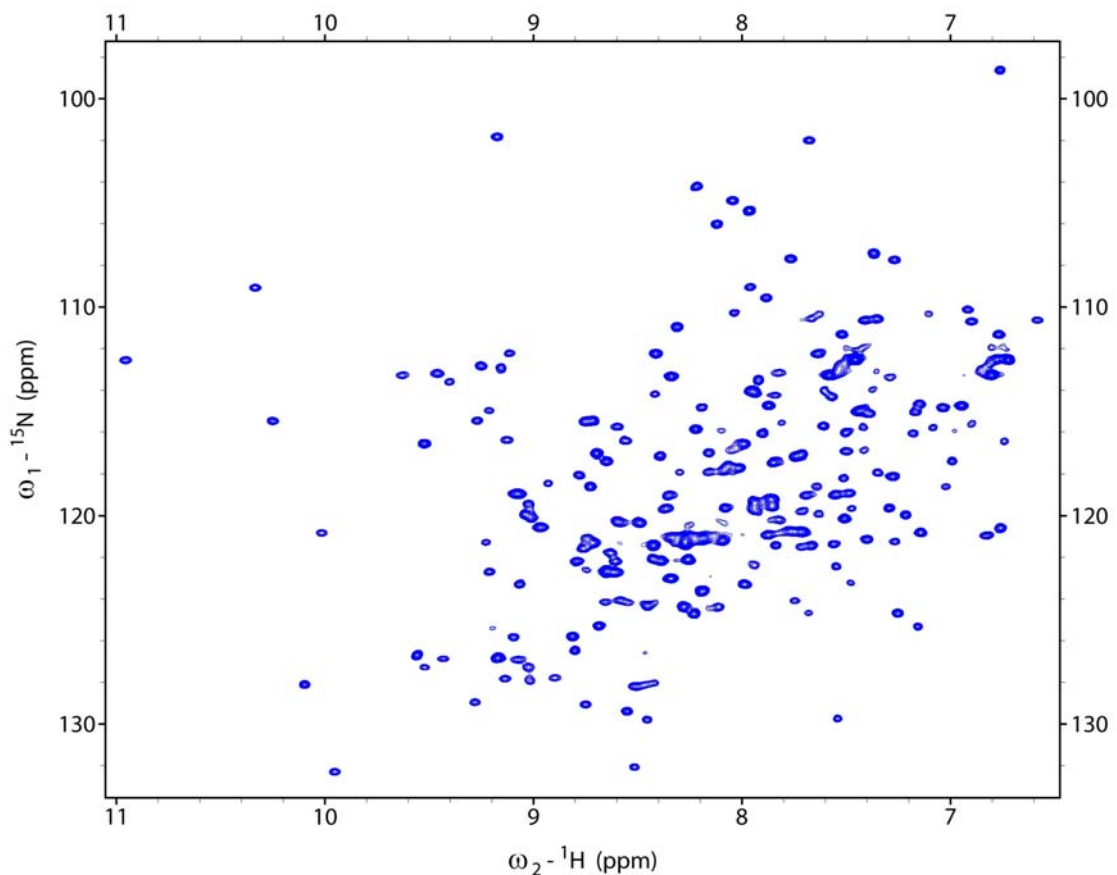
**Figure 25 : Spectre  $^{15}\text{N}$ -HSQC de la PTH H20A**

Après transformée de Fourier des signaux obtenus, on obtient un pic de résonance pour chaque couple H-N lié covalamment dans la protéine (150  $\mu\text{M}$  de PTH H20A dans 50 mM d'acétate de sodium pH 6, 200 mM NaCl). Les coordonnées de ces pics correspondent au déplacement chimique du proton H (en abscisse) et au déplacement chimique de l'atome d'azote N qui lui est lié covalamment (en ordonnée), pour les couples  $\text{H}^{\text{N}}-\text{N}$  du squelette de la protéine, pour les couples  $\text{H}^{\epsilon}-\text{N}^{\epsilon}$  des chaînes latérales des glutamines et des tryptophanes et, pour les couples  $\text{H}^{\delta}-\text{N}^{\delta}$  des asparagines.

Nous avons aussi comparé le spectre ci-dessus (qui est celui de la PTH H20A étiquetée avec un peptide de poly-histidine) avec celui de la PTH sauvage, non étiquetée (figure 26). Nous avons observé que plusieurs pics avaient disparu

entre ces deux spectres. Nous nous sommes rendu compte par la suite (après l'attribution) qu'il s'agissait des pics des couples  $H^N$ -N des résidus de l'extension N-terminale de la PTH H20A ainsi que du résidu A20, ce qui est logique puisque ces résidus n'existent pas dans la PTH sauvage. D'autres pics se déplacent de quelques dixièmes de ppm, et nous avons aussi vérifié par la suite qu'il s'agissait de résidus proches de H20, comme (N10, T18, R19, N21, G23, A24, W25, N68, D93 ou N114), ou proche de l'extrémité N-terminale de la PTH (T1, I2, K3, L4), ce qui est aussi logique puisque la mutation H20A ou l'ajout de l'extension N-terminale modifie l'environnement des résidus proches de ces positions.

Cependant, les positions de la plupart des autres pics varient très peu entre les deux spectres. Cela confirme les résultats obtenus par Bal *et al.* sur la PTH de *M. tuberculosis* (Bal *et al.*, 2007), qui montrent que la PTH H20A est structuralement très proche de la PTH sauvage. Si la PTH H20A était déstructurée ou repliée différemment de la PTH sauvage, l'environnement des  $H^N$  et des N du squelette peptidique aurait été différent dans les deux protéines et leurs fréquences de résonance auraient donc été éloignées sur les deux spectres. Cela indique également que la présence de l'extension N-terminale ne perturbe pas le repliement de la PTH. La PTH H20A semble donc un bon modèle pour l'étude structurale de la PTH en solution puis pour l'étude de son interaction avec des analogues de son substrat.



**Figure 26 : Spectres  $^{15}N$ -HSQC de la PTH sauvage non étiquetée**

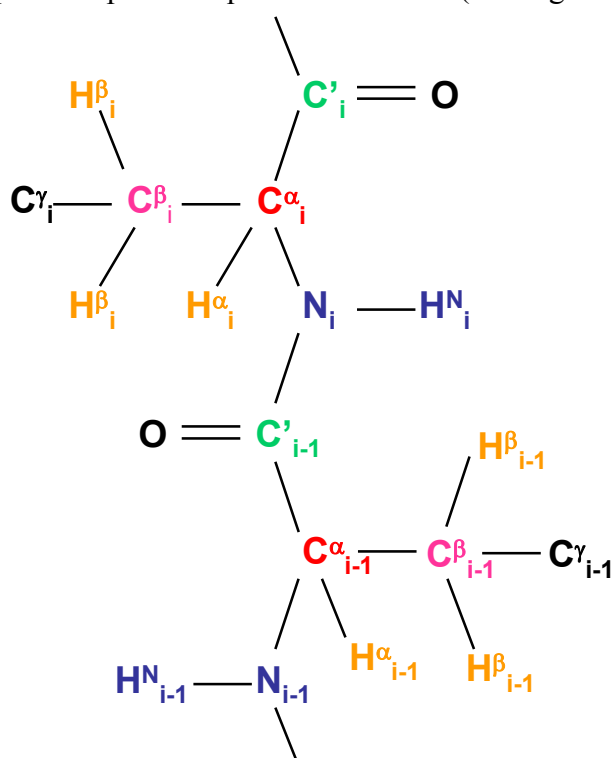
Expérience réalisée à 90  $\mu M$  de PTH dans 50 mM d'acétate de sodium pH 6, 200 mM NaCl.



## 1.2 Attribution des fréquences de résonance

### 1.2.1 Analyse du spectre HSQC de la PTH H20A

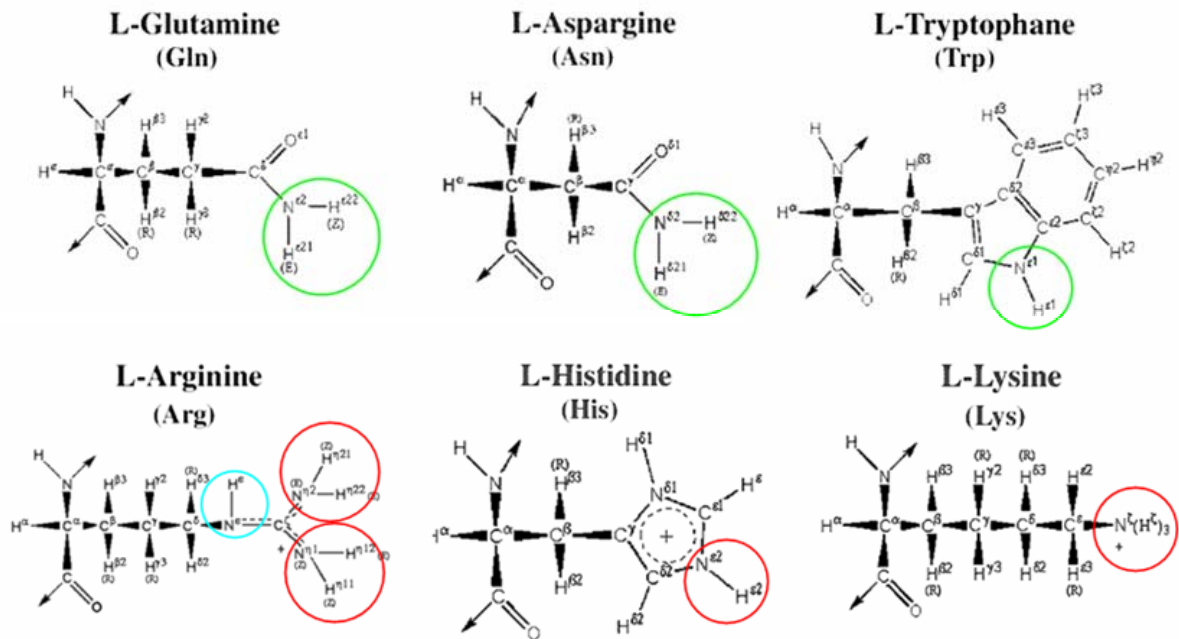
Comme nous venons de le voir, le spectre  $^{15}\text{N}$ -HSQC d'une protéine corrèle, en particulier, les fréquences de résonance des  $\text{H}^{\text{N}}_i$  et  $\text{N}_i$  du squelette de la protéine pour chaque acide aminé  $i$ . (voir figure 27).



**Figure 27 : Représentation schématique des atomes de deux acides aminés successifs dans une chaîne peptidique**

Les couples  $\text{H}^{\text{N}}_i\text{-N}_i$  et  $\text{H}^{\text{N}}_{i-1}\text{-N}_{i-1}$ , dont on observe les déplacements chimiques sur une expérience  $^{15}\text{N}$ -HSQC, sont représentés en bleu. Pour chaque système de spins, les autres atomes sont colorés selon leur nature.

Ce spectre fait aussi apparaître des pics de résonance pour les couples H-N des chaînes latérales de certains acides aminés comme les asparagines, les glutamines et les tryptophanes, qui résonnent généralement à des fréquences voisines de celles des couples  $\text{H}^{\text{N}}\text{-N}$  du squelette, ainsi que les couples  $\text{H}^{\epsilon}\text{-N}^{\epsilon}$  des arginines si l'on augmente la fenêtre d'observation (ce que l'on a fait par la suite). Par contre, avec les réglages que nous avons utilisés, nous n'observons pas les pics de résonance correspondant aux couples H-N des chaînes latérales des lysines et des histidines ni aux couples  $\text{H}^{\eta}\text{-N}^{\eta}$  des arginines, car leurs protons s'échangent généralement trop rapidement avec l'eau, ce qui empêche la détection de leurs pics de résonance.



**Figure 28 : Acides aminés ayant un ou deux H<sup>N</sup> dans leur chaîne latérale**

En vert sont entourés les couples H-N des chaînes latérales de la glutamine, de l'asparagine et du tryptophane, dont les pics de résonance sont visibles sur les spectres HSQC que nous avons enregistrés. En bleu est entouré le couple H<sup>ε</sup>-N<sup>ε</sup> de l'arginine dont le pic de résonance peut être facilement observé en augmentant la fenêtre d'observation. En rouge sont entourés les couples H-N des chaînes latérales de l'histidine, de la lysine et de l'arginine, dont les pics de résonance ne sont pas visibles sur nos spectres. D'après <http://www.bmrb.wisc.edu/referenc/commonaa.php>

La protéine étudiée comporte 213 résidus, dont 11 prolines, 10 asparagines, 2 glutamines et 2 tryptophanes (voir séquence en annexe 1). Les prolines ne possédant pas de H<sup>N</sup>, on ne verra pas de pics pour ces 11 résidus. Pour les asparagines et les glutamines les deux protons des groupements NH<sub>2</sub> des chaînes latérales peuvent résonner à des fréquences différentes. Pour chacun de ces groupements, on peut donc s'attendre à observer deux pics de résonance distincts avec la même fréquence azote mais une fréquence proton différente. Dans la région des fréquences de résonance des couples H<sup>N</sup>-N, on peut donc observer théoriquement  $213 - 11 + 2 * 10 + 2 * 2 + 2 = 228$  pics.

Malgré la grande abondance de pics dans la région centrale du spectre (voir figure 25), nous avons réussi à repérer plus de 200 taches. La différence avec le nombre de pics théorique vient du fait que certains pics sont superposés et difficilement discernables en 2 dimensions.

Ainsi, malgré la relativement grande taille de la protéine, la plupart des pics semblent identifiables et possèdent une intensité similaire. Les déplacements chimiques des protons s'étalent sur plus de 4 ppm, ce qui montre que la protéine en solution est fortement structurée et qu'une étude par RMN de cette protéine est possible.

Grâce à une expérience HSQC filtrée  $\text{NH}_2$  (qui ne montre pas les pics des  $\text{NH}_2$  des asparagines et des glutamines), nous avons pu repérer un certain nombre de pics correspondant aux couples H-N des chaînes latérales des Asn et des Gln qui, dans cette expérience, disparaissaient deux par deux avec la même fréquence azote. Nous avons ainsi pu établir une liste de couples H-N appartenant exclusivement au squelette de la protéine ou aux chaînes latérales des deux tryptophanes de la protéine.

### 1.2.2 Les systèmes de spins

Pour déterminer à quel acide aminé correspond chaque couple  $\text{H}^{\text{N}}\text{-N}$  de la liste obtenue, nous avons cherché à déterminer, pour chaque acide aminé  $i$  de la PTH H20A, les fréquences de résonance des atomes du système de spins  $i$  regroupant les atomes  $\text{H}^{\text{N}}_i$ ,  $\text{N}_i$ ,  $\text{C}^\alpha_i$ ,  $\text{C}^\beta_i$ ,  $\text{C}^\alpha_{i-1}$  et  $\text{C}^\beta_{i-1}$  (voir figure 27).

Ceci est possible grâce à l'acquisition de spectres 3D, HNCA, HN(CO)CA, HNCACB, CBCA(CO)NH (voir annexe 2) qui corrélient, grâce à des transferts d'aimantation par couplage scalaire, les fréquences de résonance d'un  $\text{H}^{\text{N}}_i$ , du  $\text{N}_i$  correspondant et d'un ou plusieurs atomes du même résidu ( $\text{C}^\alpha_i$ ,  $\text{C}^\beta_i$ ) ou du résidu précédent ( $\text{C}^\alpha_{i-1}$ ,  $\text{C}^\beta_{i-1}$ ).

Ces expériences nous ont permis d'identifier les deux couples H-N des chaînes latérales des tryptophanes ainsi que plus de 200 systèmes de spins différents. Cependant, une vingtaine de ces systèmes de spins n'étaient pas complètement définis à cause d'une ambiguïté survenant lorsque les pics  $\text{H}^{\text{N}}\text{-N}$  de deux résidus sont superposés (même fréquence  $\text{H}^{\text{N}}$  et même fréquence N). Dans ce cas, on trouve alors deux valeurs possibles pour les déplacements chimiques de chacun des atomes  $\text{C}^\alpha_i$ ,  $\text{C}^\beta_i$ ,  $\text{C}^\alpha_{i-1}$  et  $\text{C}^\beta_{i-1}$ . Il est impossible de savoir auquel des deux résidus correspondent ces déplacements chimiques. Il y a donc 16 systèmes de spins possibles pour ces deux résidus, correspondant aux 16 combinaisons possibles pour les quadruplets ( $\text{C}^\alpha_i$ ;  $\text{C}^\beta_i$ ;  $\text{C}^\alpha_{i-1}$ ;  $\text{C}^\beta_{i-1}$ ). Cette ambiguïté sera levée par la suite.

### 1.2.3 Détermination de l'enchaînement des systèmes de spins

Nous avons ensuite cherché des couples  $\text{C}^\alpha_i\text{-C}^\beta_i$  ayant des déplacements chimiques identiques à ceux de couples  $\text{C}^\alpha_{j-1}\text{-C}^\beta_{j-1}$  (à 0,1 ppm près). En effet, cela nous permet alors de déduire que l'acide aminé du système de spins  $i$  précède l'acide aminé du système de spins  $j$ . Dans certains cas, ceci a permis de diminuer l'ambiguïté présentée ci-dessus puisque certains couples  $\text{C}^\alpha_i\text{-C}^\beta_i$  ne correspondaient à aucun couple  $\text{C}^\alpha_{j-1}\text{-C}^\beta_{j-1}$ . Les combinaisons correspondantes pouvaient donc être éliminées des possibilités. Au mieux, cela limitait l'ambiguïté à 4 possibilités, correspondant aux 4 combinaisons possibles entre les deux couples  $\text{C}^\alpha_i\text{-C}^\beta_i$  et les deux couples  $\text{C}^\alpha_{i-1}\text{-C}^\beta_{i-1}$  restants.

On a trouvé ainsi de nombreux enchaînements de systèmes de spins, mais il est cependant impossible, par cette méthode, d'obtenir l'enchaînement complet de tous les systèmes de spins de la protéine.

On ne peut en effet pas toujours trouver le résidu précédent et, parfois, plusieurs choix sont possibles, soit parce que plusieurs systèmes de spins ont des couples  $C^\alpha_i-C^\beta_i$  ou  $C^\alpha_{i-1}-C^\beta_{i-1}$  avec des déplacements chimiques identiques, soit à cause de l'ambiguïté présentée ci-dessus. Le premier phénomène est aggravé dans le cas des glycines qui ne possèdent pas de  $C^\beta$ , où la double corrélation est donc impossible, et pour lesquelles il est fréquent de trouver des déplacements chimiques de  $C^\alpha$  identiques ou très proches puisque ce sont des résidus de nature identique. L'existence de prolines pour lesquelles les systèmes de spins n'ont pu être observés (ne possédant pas de  $H^N$ ) est aussi un frein à l'élaboration des enchaînements.

Nous avons alors enregistré d'autres spectres 3D (HBHA(CO)NH, HBHANH), qui corrélient les fréquences de résonance d'un  $H^N_i$ , du  $N_i$  correspondant, et, suivant l'expérience, des  $H^\alpha_i$ ,  $H^\alpha_{i-1}$ ,  $H^\beta_i$  et  $H^\beta_{i-1}$  (voir annexe 2). Ceci nous a permis d'utiliser les déplacements chimiques des  $H^\alpha_i$ ,  $H^\alpha_{i-1}$  et  $H^\beta_i$ ,  $H^\beta_{i-1}$  pour déterminer l'enchaînement des systèmes de spins. Mais, à cause de la qualité relativement faible des spectres, certains déplacements chimiques (notamment un grand nombre de  $H^\alpha_{i-1}$ ,  $H^\beta_i$  et  $H^\beta_{i-1}$ ) n'ont pu être déterminés avec certitude et cette méthode n'a pas permis d'aboutir à une identification complète de l'enchaînement des systèmes de spins. Elle ne lève pas non plus l'ambiguïté pré-citée.

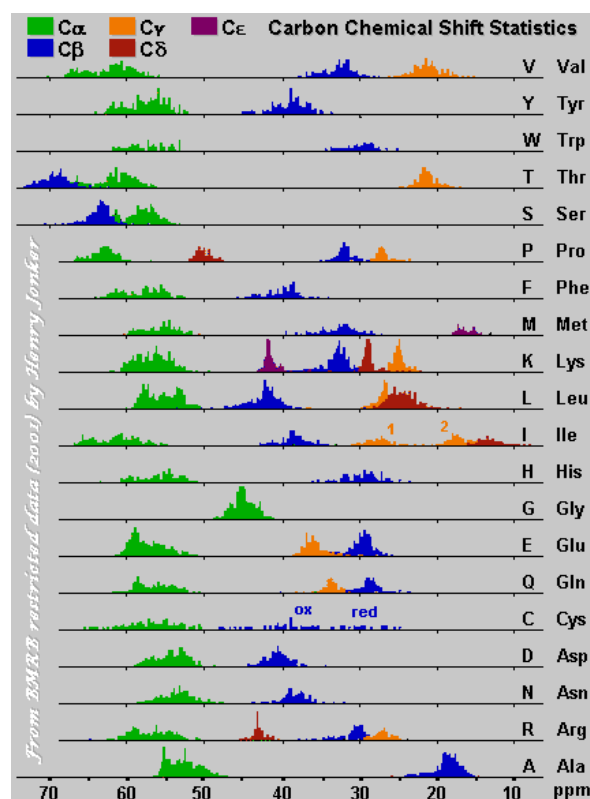
### 1.2.4 Identifications d'acides aminés spécifiques

Parmi les 20 acides aminés, certains présentent des déplacements chimiques caractéristiques (voir figure 29). Nous les avons donc identifiés, pour pouvoir ensuite faire correspondre certains enchaînements de systèmes de spins à des séquences d'acides aminés dans la protéine. Ainsi, les sérines (S) et les thréonines (T) ont des déplacements chimiques très élevés pour leur  $C^\beta$ , (ce sont les seuls au-dessus de 60 ppm). Nous avons identifié 20 systèmes de spin ayant un  $C^\beta$  avec déplacement chimique élevé, correspondant donc aux dix sérines et aux dix thréonines (avec une préférence pour T si la valeur est très élevée), ainsi que les 20 systèmes de spins des acides aminés suivant ces sérines et ces thréonines. A l'opposé, les alanines (A) ayant un déplacement chimique très faible pour leur  $C^\beta$  (ce sont les seuls en dessous de 25 ppm), nous avons donc pu identifier les 22 alanines, ainsi que les 22 systèmes de spins des acides aminés suivant ces alanines.

D'autre part, nous avons pu identifier un certain nombre de glycines (G) en utilisant le fait que ces résidus n'ont pas de  $C^\beta$  et que leur  $C^\alpha$  possède un faible déplacement chimique (c'est le seul en dessous de 45 ppm).

**Figure 29 : Déplacements chimiques standard des atomes de carbone dans les protéines**

Sur cette figure, on peut voir une représentation statistique de la valeur des déplacements chimiques observés dans les protéines pour les différents carbones des chaînes latérales de chacun des 20 acides aminés. On peut aussi trouver, dans la BMRB (*Biological Magnetic Resonance Data Bank*), des statistiques sur les fréquences de résonance des atomes d'hydrogène et d'azote. D'après ([http://www.bmrb.wisc.edu/ref\\_info/statful.htm](http://www.bmrb.wisc.edu/ref_info/statful.htm)).



Pour l'ensemble des autres systèmes de spins, nous avons déterminé à chaque fois la nature des résidus compatibles avec les déplacements chimiques observés. Par exemple, un résidu ayant pour déplacement chimique de 60 ppm pour le C<sup>α</sup> et de 40 ppm pour le C<sup>β</sup>, peut être une isoleucine (I), une cystéine (C), une phénylalanine (P), une tyrosine (Y) ou encore, moins probablement, une leucine (L) ou un acide aspartique (D), mais certainement pas un des autres résidus (voir figure 29).

**1.2.5 Attribution des séquences grâce aux acides aminés spécifiques**

Comme on l'a vu ci-dessus, on peut relier le résidu i au résidu précédent i-1 pour identifier des enchaînements de systèmes de spins, et on peut aussi déterminer la nature d'un certain nombre d'acides aminés dans chaque enchaînement obtenu. Les enchaînements sont plus ou moins longs pour les raisons citées précédemment. A ce stade, la plupart des résidus faisaient partie d'enchaînements de 2 à 15 acides aminés et la nature d'une cinquantaine de résidus (sur 193) était identifiée.

Grâce à ces identifications, nous avons ensuite pu placer plusieurs enchaînements de systèmes de spins sur la séquence de la protéine, en commençant par placer ceux dont la position était certaine (par exemple, des séquences contenant deux résidus identifiés dont la distance correspondait à une seule position possible dans la protéine). Quand plusieurs solutions étaient possibles, la compatibilité de la séquence de la protéine avec les déplacements

chimiques des résidus non identifiés, a parfois été déterminante. Cette procédure a aussi permis de lever l'ambiguïté concernant les résidus ayant des déplacements chimiques de leurs  $H^N$  et N proches, en éliminant les enchaînements ne correspondant à aucune séquence dans la protéine.

Nous avons ainsi réussi à attribuer environ 60% des acides aminés de la protéine avec certitude et à faire des hypothèses plausibles pour 20% des acides aminés. Beaucoup de possibilités existaient encore pour les 20% restants, soit parce que les enchaînements déterminés étaient trop courts, soit parce qu'ils ne possédaient pas ou pas assez d'acides aminés identifiés.

### 1.2.6 Amélioration de l'attribution

Nous avons réalisé trois types d'expériences pour améliorer nos connaissances de la nature des acides aminés correspondant à certains systèmes de spins et pour nous assurer de leur succession.

Nous avons tout d'abord enregistré les spectres d'une expérience TOCSY HSQC, corrélant les  $H^N_i$ ,  $N_i$  et les  $H_i$  (tous les atomes d'hydrogène de la chaîne latérale du résidu  $i$ ) pour améliorer notre attribution notamment des  $H^\alpha$  et confirmer ou trouver des nouveaux enchaînements de systèmes de spins.

Ensuite, une expérience HCCH TOCSY a été enregistrée. Elle permet, pour les atomes n'appartenant pas à un cycle aromatique, de corrélérer entre eux les déplacements chimiques d'un carbone de la chaîne latérale d'un acide aminé, d'un proton qui lui est lié et des autres protons de cette chaîne latérale, grâce à des transferts d'aimantation par couplage scalaire. Il est ainsi théoriquement possible d'obtenir le nombre d'atomes de carbone dans une chaîne latérale donnée et les valeurs des déplacements chimiques des atomes de carbone et d'hydrogène de cette chaîne latérale. Ces valeurs sont souvent caractéristiques du type d'acide aminé. Nous avons ainsi pu trouver ou confirmer la grande majorité des déplacements chimiques des  $C^\alpha$ ,  $C^\beta$ ,  $H^\alpha$  et  $H^\beta$ . La qualité des spectres que nous avons obtenus (beaucoup de superpositions de pics, signal peu intense), ne nous a pas permis d'attribuer les déplacements chimiques de l'ensemble des atomes des chaînes latérales pour tous les résidus. Cependant, à partir de ceux que nous avons caractérisés avec certitude, nous avons pu déterminer la nature de la plupart des acides aminés que nous n'avions pas pu identifier jusqu'alors. Les corrélations  $C^\alpha$ - $C^\beta$  étant ici indépendantes des fréquences de résonance des  $H^N$  et N, l'analyse de cette expérience nous a aussi permis de lever l'ambiguïté restant sur les couples  $C^\alpha$ - $C^\beta$  possédant un pic  $H^N$ -N identique.

Enfin, nous avons réalisé une expérience  $^{15}N$ -NOESY-HSQC, qui permet de corrélérer entre eux les déplacements chimiques d'un  $H^N$ , d'un N et des protons situés à moins de 5 Å de ce  $H^N$ , grâce à un transfert d'aimantation par effet Overhauser, NOE (*Nuclear Overhauser Effect*). Les  $H^N$  de résidus successifs étant à une distance pouvant autoriser un NOE, nous avons pu observer des pics

correspondant aux déplacements chimiques du couple  $N_i-H_i^N$  et du  $H_{i-1}^N$  et/ou  $H_{i+1}^N$ . L'intensité de ces pics dépend des éléments structuraux auxquels appartiennent les résidus concernés mais la plupart d'entre eux ont pu être repérés. Cette étude nous a permis de confirmer la plupart des enchaînements de systèmes de spins déjà identifiés.

Cette expérience nous a aussi permis d'identifier spécifiquement les déplacements chimiques des couples  $H^\delta-N^\delta$  des asparagines et des couples  $H^\epsilon-N^\epsilon$  des chaînes latérales des glutamines, repérés précédemment grâce aux expériences HSQC. En effet, leur  $H_i^N$  du squelette est souvent voisin (à moins de 5 Å) d'au moins un des  $H^\delta$  (ou  $H^\epsilon$ ) de leurs chaînes latérales. Nous avons donc pu observer des pics corrélant les couples  $H_i^N-N_i$  de la chaîne principale et un ou deux  $H^\delta$  (ou  $H^\epsilon$ ) de la chaîne latérale, ainsi que des pics corrélant les couples  $H^\delta-N^\delta$  (ou  $H^\epsilon-N^\epsilon$ ) de la chaîne latérale et à la fois le  $H_i^N$  de la chaîne principale et le deuxième  $H^\delta$  (ou  $H^\epsilon$ ) (systématiquement très proche du premier  $H^\delta$  (ou  $H^\epsilon$ )). Pour les 10 asparagines et les deux glutamines, nous avons ainsi pu attribuer les déplacements chimiques de tous les couples  $H^\delta-N^\delta$  et  $H^\epsilon-N^\epsilon$ . Cette étude a été confirmée par la suite (après l'étude topologique et structurale de la PTH), par l'observation d'autres NOE interrésiduels, correspondant à des atomes d'hydrogènes voisins des  $H^\delta$  ou des  $H^\epsilon$  des chaînes latérales des asparagines ou des glutamines, sur la structure de la PTH.

Par ailleurs, les déplacements chimiques des carbones  $C'$  ont été attribués grâce à une expérience (HNCO) corrélant les  $H_i^N$ ,  $N_i$  et  $C'_{i-1}$ . Pour les acides aminés possédant des couples  $H^N-N$  ayant des déplacements chimiques très proches, cette attribution a été facilitée par la bonne qualité des spectres. Cependant les carbones  $C'$  du résidu C-terminal et des résidus précédant une proline n'ont pas pu être identifiés par cette méthode puisque les résonances de ces atomes ne sont pas corrélées à celles d'un couple  $H_i^N-N_i$ .

### 1.2.7 Bilan de l'attribution

En fin de compte, grâce à ces nouvelles connaissances, nous avons réussi à identifier les systèmes de spins de tous les résidus à l'exception des prolines et des 15 résidus de l'extension N-terminale situés en amont de la proline en position 16. Nous avons déterminé les déplacements chimiques de la plupart des atomes du squelette et ceux de certains des atomes des chaînes latérales. Ce travail a été compliqué par la taille de la protéine qui élargit les signaux, notamment des chaînes latérales et qui augmente le nombre de pics à analyser et le nombre de pics se chevauchant.

Pour les 193 résidus de la PTH, en ne comptant pas l'extension N-terminale, nous avons donc attribué :

-les 183 déplacements chimiques de tous les N de la chaîne principale. (193 résidus moins 10 prolines).

-les déplacements chimiques des 183 H<sup>N</sup> de la chaîne principale.

-181 déplacements chimiques de C', n'ayant pu observer le pic du C' de G111 ni ceux des acides aminés précédant les 10 prolines, ni celui du C' du résidu C-terminal.

-190 déplacements chimiques de C<sup>α</sup>. Il ne manque que ceux de 3 prolines. Les résonances des C<sup>α</sup><sub>i</sub> des prolines sont en effet plus difficilement identifiables puisqu'elles ne peuvent être attribuées que grâce à la connaissance des déplacements chimiques des C<sup>α</sup><sub>i-1</sub> des résidus qui les suivent et une corrélation avec les C<sup>β</sup><sub>i</sub> sur l'HCCH-TOCSY.

-169 déplacements chimiques de C<sup>β</sup> sur 172. Grâce aux déplacements chimiques des C<sup>β</sup><sub>i-1</sub> et grâce à l'HCCH-TOCSY, nous avons identifié les résonances des C<sup>β</sup> pour 7 prolines sur les dix, ainsi que pour l'ensemble des autres résidus, à l'exception des 21 glycines qui ne possèdent pas de C<sup>β</sup>.

-197 déplacements chimiques de H<sup>α</sup> sur 214. Nous n'avons pas pu attribuer ceux de L8, P63, P98, P99, P127, H138, P139, P153, P154, L180. Pour les 21 glycines, qui possèdent deux H<sup>α</sup> avec des déplacements chimiques pouvant être différents, nous n'avons pu attribuer les résonances de G137, G151 et G179, ni celle du deuxième H<sup>α</sup> de G147.

-208 déplacements chimiques de H<sup>β</sup>. Les signaux des H<sup>β</sup> sont souvent plus faibles que ceux des H<sup>α</sup>. De plus, la plupart des acides aminés possèdent plusieurs H<sup>β</sup> qui peuvent résonner à une fréquence identique (notamment les 3 H<sup>β</sup> des alanines). On peut donc difficilement évaluer le nombre de déplacements chimiques différents observables en théorie. Cependant, il existe seulement 18 résidus pour lesquels aucun H<sup>β</sup> n'a pu être déterminé avec certitude (dont les 10 prolines).

Pour ce qui est des autres atomes des chaînes latérales, nous avons attribué 100% des pics de résonance des groupements NH<sub>2</sub> des chaînes latérales des asparagines et glutamines. L'attribution des résonances de tous les atomes des chaînes latérales n'a pas été terminée car elle n'était pas nécessaire à la suite de notre travail. Elle nécessiterait l'enregistrement d'autres spectres, notamment pour les résidus aromatiques, et d'augmenter la qualité de nos spectres, ce qui est envisageable. Nous avons travaillé à 130 μM de PTH et nos spectres ont été enregistrés sur un spectromètre RMN de 600 MHz. Il serait possible d'augmenter encore la concentration protéique et de travailler sur un spectromètre plus puissant.



## 1.3 Etude structurale de la PTH en solution

### 1.3.1 Structure secondaire de la PTH

Pour connaître la structure secondaire d'une protéine en solution, on peut utiliser les valeurs des déplacements chimiques de certains atomes ( $H^\alpha$ ,  $C^\alpha$ ,  $C^\beta$  et  $C^O$ ). En effet, le type de structure auquel appartient un résidu (boucle, hélice alpha ou brin bêta) influence les fréquences de résonance de ces atomes.

Nous avons utilisé la méthode CSI (*Chemical Shift Index*) (Wishart and Sykes, 1994 ; Wishart *et al.*, 1995), qui compare les valeurs des déplacements chimiques des atomes  $H^\alpha$ ,  $C^\alpha$ ,  $C^\beta$  et  $C'$  à celles des « *random coil* » et attribue, pour chaque atome, un indice (0,1,-1) selon que le déplacement chimique mesuré est compatible, au-dessus ou au-dessous du déplacement chimique « *random* ». Ensuite, le programme applique un certain nombre de filtres en regardant les indices des différents atomes d'un résidu et de ces résidus voisins, pour en déduire le type de structure secondaire auquel appartient chaque résidu. Les résultats obtenus sont présentés dans le tableau 4 ci-dessous.

	Hélice $\alpha$	Hélice $\alpha$		Brin $\beta$	Brin $\beta$
	RMN CSI	Cristal		RMN CSI	Cristal
$\alpha 1$	22-33	22-34	$\beta 1$	3-6	4-7
$\alpha 2$	70-80	71-81	$\beta 2$		39-41
$\alpha 3$	108-122	114-122	$\beta 3$	46-53	46-53
$\alpha 4$	145-149	142-149	$\beta 4$	57-62	55-63
$\alpha 5$	155-175	155-178	$\beta 5$	87-92	88-94
$\alpha 6$	180-187	180-189	$\beta 6$	101-106	102-106
			$\beta 7$	129-133	130-135
				150-152	

**Tableau 4 : Eléments de structure secondaire de la PTH d'*E. coli* prédits par la méthode CSI ou observés dans les cristaux de cette protéine**

Les numéros des résidus indiquent la limite des éléments de structure secondaire obtenus par la méthode CSI ou observés dans la structure cristallographique de la protéine. Les numéros des hélices et des brins sont ceux qui ont été définis par Schmitt *et al.* pour décrire la structure cristallographique (Schmitt *et al.*, 1997). En rouge sont indiqués les éléments de structure secondaire prédits ou observés par l'une des deux méthodes mais pas par l'autre.

Cette méthode nous a permis d'identifier six hélices alpha, correspondant aux six hélices alpha de la structure cristallographique de la PTH. On trouve quelques différences au niveau des limites de ces hélices, notamment pour l'hélice  $\alpha 3$  qui semble plus longue dans la structure en solution. D'après les données de RMN, les résidus 108 à 114 appartiendraient à cette hélice alors que la structure

cristallographique présente cette séquence de résidus comme une boucle reliant l'hélice à un brin. Les autres différences ne concernent qu'un petit nombre de résidus. On peut noter que l'hélice C-terminale semble structurée en solution. La structure de cette région pourrait donc ressembler à celle des cristaux de la PTH d'*E. coli* et non à celle des cristaux de la PTH de *M. tuberculosis*, dans laquelle la partie C-terminale de la protéine est désordonnée (Selvaraj *et al.*, 2007).

Nous avons aussi identifié sept régions formant des brins étendus. Six d'entre elles correspondent à six des sept brins du feuillet bêta de la structure cristallographique de la PTH. Là aussi, les limites de ces éléments structuraux varient légèrement entre la structure cristallographique et la structure en solution. Un brin étendu supplémentaire (150-152) correspond à un brin étendu observé également dans la structure cristallographique mais qui ne fait pas partie du feuillet bêta. Enfin, un petit brin bêta formé par les résidus 39 à 41 dans la structure cristallographique, n'a pas été identifié comme tel lors de l'analyse effectuée par la méthode CSI.

Les petites différences dans les limites des éléments structuraux et dans l'identification de petits éléments peuvent être dues au fait que la méthode CSI est une technique statistique qui prédit avec une grande probabilité l'existence de structures secondaires, mais n'apporte pas de certitude. Ces différences ne sont donc pas forcément significatives.

La structure secondaire de la PTH en solution semble donc proche de celle dans les cristaux même si, à ce stade, on ne peut exclure qu'il y ait certaines différences localisées. Cette ressemblance apporte aussi une confirmation à l'attribution des fréquences de résonance que nous avons réalisée.

### 1.3.2 Structure topologique de la PTH

Pour compléter notre identification des éléments de structure secondaire de la PTH en solution, nous avons recherché la présence de NOE caractéristiques de ces structures : NOE intenses entre les protons  $H^N_i$  et  $H^N_{i-1}$  dans les hélices, NOE entre les protons  $H^N_i$  et  $H^{\alpha}_{i-3}$  dans les hélices, NOE intenses entre les protons  $H^N_i$  et les  $H^{\alpha}_{i-1}$  dans les brins bêta, et NOE observables entre les protons  $H^N_i$  et  $H^N_j$  ou entre les protons  $H^N_i$  et  $H^{\alpha}_j$ , dans les feuillets bêta, où  $j$  désigne un résidu faisant parti d'un brin voisin du brin auquel appartient le résidu  $i$ . Les NOE peuvent être mis en évidence grâce à une expérience  $^{15}\text{N}$ -NOESY-HSQC qui corrèle pour chaque  $H^N$ , les déplacements chimiques de ce  $H^N$ , du N correspondant, et de tout atome d'hydrogène présent à moins de 4 à 5 Å de ce  $H^N$ . Nous avons déjà utilisé cette expérience pour attribuer les déplacements chimiques des couples  $H^{\delta}$ - $N^{\delta}$  et  $H^{\epsilon}$ - $N^{\epsilon}$  des chaînes latérales des asparagines et glutamines et pour confirmer les enchaînements de certains systèmes de spin.

Dans une protéine bien structurée comme la PTH, chaque  $H^N$  est généralement proche d'un nombre assez important (pouvant aller jusqu'à une vingtaine) d'atomes d'hydrogène (appartenant à la chaîne latérale de son résidu ou aux résidus voisins). Ceci entraîne un recouvrement de nombreux pics, d'autant plus important que la protéine possède plus de 200 résidus. L'identification complète de tous les pics observés est donc difficile et nous avons préféré comparer les NOE observés à ceux prédits par la structure cristallographique de la PTH. Pour cette étude, nous nous sommes restreints aux protons  $H^N$ ,  $H^\alpha$  et  $H^\beta$ , car ils étaient attribués plus complètement que les autres protons.

Au fichier PDB de la structure cristallographique de la PTH d'*E. coli* (code PDB : 2PTH), nous avons ajouté les coordonnées des atomes d'hydrogène grâce au programme REFMAC (Winn *et al.*, 2003). Nous avons ensuite écrit un programme en Java pour déterminer la liste des NOE théoriquement observables si la structure de la PTH en solution est identique à la structure dans les cristaux.

Ce programme prend comme données d'entrée les coordonnées géométriques des  $H^N$ ,  $H^\alpha$  et  $H^\beta$  de la protéine dans les cristaux, ainsi que les valeurs des déplacements chimiques de ces atomes et des atomes d'azote. Il calcule ensuite les distances entre les  $H^N$  et les divers atomes d'hydrogène de la protéine. Il donne en sortie un fichier contenant une liste de pics potentiels, avec, pour chacun d'eux, les déplacements chimiques d'un  $H^N$ , du N correspondant et d'un atome d'hydrogène situé à une distance inférieure à une distance paramétrable (4 ou 5 Å par exemple) de ce  $H^N$ . Cette liste est directement lisible par Sparky, ce qui permet de récupérer l'intensité du signal à la position attendue pour chacun de ces pics sur le spectre NOESY de la protéine.

A partir des 183  $H^N$ , 197  $H^\alpha$  et 208  $H^\beta$  attribués, notre programme a prédit 2485 pics possibles entre un  $H^N$  et soit un  $H^N$ , soit un  $H^\alpha$ , soit un  $H^\beta$  situé à moins de 5 Å de celui-ci (1399 pour 4 Å et 872 pour 3 Å). Pour analyser ces résultats de manière systématique, nous avons considéré qu'un pic était significatif lorsque son intensité était au moins trois fois supérieure au bruit moyen.

Comme on peut le voir sur le tableau 5 ci-dessous, un fort pourcentage des pics attendus est observé (79%, 93% et 96% si on considère une distance de 5 Å, de 4 Å ou de 3 Å, respectivement). Ces pourcentages sont encore plus élevés pour les atomes d'hydrogène d'un même résidu, ce qui valide notre attribution, ainsi que pour les atomes d'hydrogène de résidus situés à une distance comprise entre 1 et 4 liaisons peptidiques dans la séquence de la PTH (dans ce cas, les NOE observés sont caractéristiques des structures secondaires). Le pourcentage est un peu plus faible pour les NOE longue distance (entre résidus espacés de plus de 4 liaisons peptidiques) entre des atomes d'hydrogènes situés à moins de 5 Å (61%) mais il augmente nettement quand on restreint l'analyse aux atomes situés à moins de 4 Å (83%) ou à moins de 3 Å (93%).

5 Å	Nombre de NOE prédits	Nombre de NOE observés	Pourcentage de NOE observés / prédits
$\lambda=0$	579	555	96
$0<\lambda<5$	1507	1157	77
$\lambda>4$	399	246	61
Total	2485	1958	79

4 Å	Nombre de NOE prédits	Nombre de NOE observés	Pourcentage de NOE observés / prédits
$\lambda=0$	578	554	96
$0<\lambda<5$	705	646	92
$\lambda>4$	116	96	83
Total	1399	1296	93

3 Å	Nombre de NOE prédits	Nombre de NOE observés	Pourcentage de NOE observés / prédits
$\lambda=0$	500	480	96
$0<\lambda<5$	358	347	97
$\lambda>4$	14	13	93
Total	872	841	96

**Tableau 5 : Observation des NOE prédits d'après la structure cristallographique de la PTH**

Pour chacune des distances considérées (5, 4 ou 3 Å), ce tableau indique le nombre de NOE prédits à partir de la structure cristallographique de la PTH, le nombre de NOE observés parmi les NOE prédits, ainsi que le pourcentage correspondant. Les résultats sont présentés en fonction du paramètre  $\lambda$ , égal au nombre de liaisons peptidiques séparant les résidus auxquels appartiennent les deux protons susceptibles d'interagir par NOE. Un pic a été considéré comme significatif lorsque son intensité était au moins trois fois supérieure au bruit moyen, ce qui correspond environ au dixième de l'intensité moyenne de tous les pics prédits pour 5 Å.

Il est important de noter que l'absence d'un pic correspondant à un NOE prédit ne signifie pas forcément que les atomes correspondants sont plus éloignés en solution que dans le cristal. Un certain nombre de NOE ne sont en effet pas observables à cause de la compétition au niveau du transfert d'aimantation entre les atomes d'hydrogène cibles des NOE. Ils échappent ainsi à notre statistique. De plus, une analyse plus détaillée nous a permis de voir que les pics prédits mais non observés ne sont pas regroupés dans une même région et ne sont donc pas représentatifs de différences structurales majeures. Le fort pourcentage de NOE prédits et effectivement observés indique donc que la structure de la PTH en solution est très proche de celle de la PTH dans le cristal.

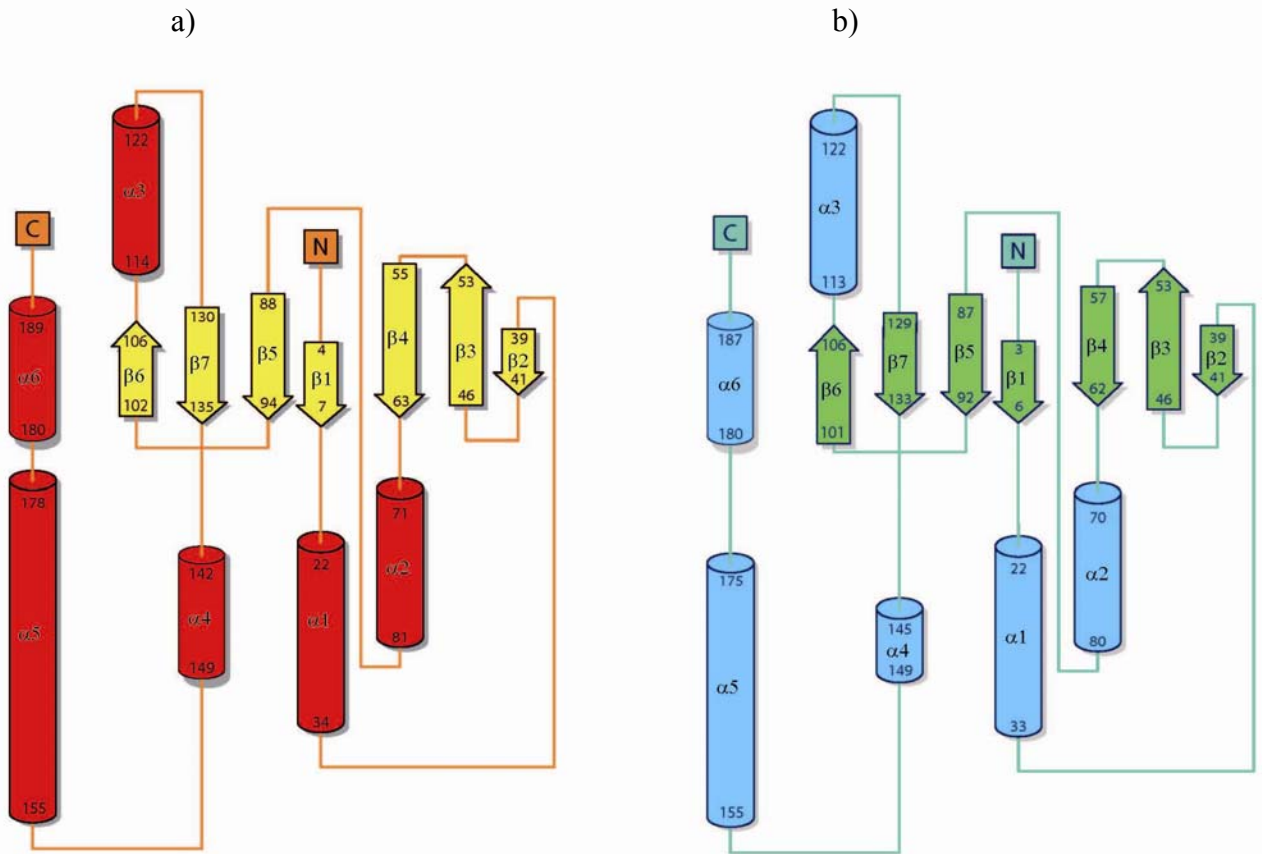
Nous avons parallèlement vérifié la présence de NOE caractéristiques des hélices  $\alpha$ , ce qui nous a permis de compléter les résultats fournis par la méthode CSI. Pour l'hélice  $\alpha_3$ , on retrouve des NOE de type  $H^N_i-H^N_{i+1}$  pour l'ensemble des résidus 108 à 122 mais des NOE de type  $H^N_i-H^{\alpha}_{i-3}$  seulement pour les valeurs de  $i$  allant de 116 à 122 et pas pour celles entre 111 et 115, ce qui semble indiquer que la structure adoptée par les résidus 108 à 113 est plutôt une boucle, comme dans le cristal.

Les autres prédictions CSI ont été confirmées par cette approche. Notamment, pour les résidus 142 à 144 de l'hélice  $\alpha_4$ , on ne retrouve pas les NOE typiques des hélices, ce qui confirme que cette hélice est probablement plus courte en solution que dans le cristal. On peut d'ailleurs noter qu'elle est aussi plus courte dans la structure cristallographique de la PTH de *M. tuberculosis* et dans la structure par RMN de cette enzyme. La conformation légèrement différente de cette région dans les cristaux de la PTH d'*E. coli*, pourrait venir du fait que dans ces cristaux, une partie du site actif de la protéine est occupé par l'extrémité C-terminale d'une molécule voisine. Comme on l'a déjà mentionné ci-dessus, l'extrémité C-terminale de la PTH adopte une structure en hélice  $\alpha$  dans les cristaux de la PTH d'*E. coli* mais pas dans ceux de la PTH de *M. tuberculosis*. Les NOE observés indiquent sans ambiguïté l'existence de l'hélice C-terminale en solution dans le cas de la PTH d'*E. coli*.

Parmi les 246 NOE prédits à longue distance et observés ( $\lambda > 4$ ), 120 d'entre eux sont des NOE caractéristiques de la formation du feuillet bêta de la protéine. On retrouve en effet 14 NOE entre  $\beta_1$  et  $\beta_4$ , 17 entre  $\beta_1$  et  $\beta_5$ , 33 entre  $\beta_3$  et  $\beta_4$ , 14 entre  $\beta_5$  et  $\beta_7$  et 20 entre  $\beta_6$  et  $\beta_7$ . On observe également 21 NOE entre les résidus 39-41 et le brin  $\beta_3$ , ce qui suggère que ces résidus forment un brin bêta appartenant au feuillet, comme dans la structure cristallographique de la protéine (brin  $\beta_2$ ), mais contrairement à la prédiction CSI. Le feuillet  $\beta$  semble donc se replier de manière similaire en solution et dans le cristal.

Certains des 126 autres NOE observés entre résidus éloignés dans la séquence de la protéine correspondent à des résidus situés dans des boucles aux extrémités de deux brins bêta voisins. D'autres, reflètent le rapprochement de certains éléments de structure secondaire. Ils ne concernent cependant pas l'ensemble de la protéine et leur nombre n'aurait pas été suffisant pour déterminer, *ab initio*, la structure de la protéine en solution. Une telle étude aurait nécessité, par exemple, l'observation de NOE entre les protons des chaînes latérales des acides aminés.

L'ensemble de ces résultats permet cependant de conclure que la structure topologique de la PTH en solution est très proche de celle dans le cristal (voir figure 30). Les légères différences peuvent être dues à notre méthode, mais peuvent aussi simplement refléter la plasticité de la molécule en solution. On peut en déduire que la structure tridimensionnelle de la PTH en solution est proche de celle de la PTH dans le cristal. C'est pourquoi, dans la suite de notre étude, nous nous servirons de la structure cristallographique de la PTH comme modèle de base pour étudier l'interaction de cette protéine avec des analogues de son substrat.



**Figure 30 : Comparaison de la structure topologique de la PTH dans le cristal et en solution**

a) Structure topologique de la PTH d'*E. coli* dans le cristal (code pdb : 2PTH) : les cylindres en rouges représentent les hélices alpha, les flèches en jaune représentent les brins bêta. b) Structure topologique de la PTH d'*E. coli* en solution. Les cylindres en bleu représentent les hélices alpha, les flèches en vert représentent les brins bêta.



## **Chapitre 2**

**Caractérisation de l'interaction  
entre la PTH et un analogue de son  
substrat, la diacétyl-Lys-(3'NH)-  
adénosine**





Nous avons tout d'abord étudié l'interaction entre la PTH et un analogue du substrat de la PTH correspondant à la région d'un peptidyl-ARNt se logeant dans le site actif de l'enzyme.

## 2.1 Présentation du ligand étudié

Comme nous l'avons vu précédemment, l'interaction entre la PTH et son substrat (notamment le diacétyl-Lys-ARNt<sup>Lys</sup>) a déjà été étudiée par mutagenèse dirigée. Certains résidus jouant un rôle soit dans l'activité catalytique, soit dans la fixation du substrat ont pu être identifiés. Nous avons voulu poursuivre cette étude afin d'établir un modèle structural de l'interaction entre la PTH et son substrat. Comme nous l'avons déjà précisé, cette étude n'a pu être réalisée par cristallographie puisque aucun cristal du complexe entre la protéine et son substrat ou un analogue de son substrat n'a pu être obtenu.

Nous avons donc étudié, par RMN, la liaison à la PTH de la diacétyl-Lys-(3'NH)-adénosine (3'-(L-[N,N-diacetyl-lysiny]amino-3'-désoxyadénosine). Ce composé, synthétisé par Gavin O'Mahony et Morten Grøtli (*Department of Chemistry, University of Gothenburg, Göteborg, Suède*), mime la région d'un peptidyl-ARNt où le peptide est accroché à l'ARNt. Il comprend une adénosine, correspondant à la dernière base d'un ARNt, reliée par une liaison amide à la forme acétylée d'un acide aminé. Comme acide aminé, nous avons choisi la lysine puisque la plupart des études biochimiques sur la PTH, pour déterminer l'importance de ses résidus catalytiques, ont été menées en utilisant le diacétyl-Lys-ARNt<sup>Lys</sup>. La présence d'une liaison amide, à la place de la liaison ester qui existe dans un peptidyl-ARNt, permet d'éviter l'hydrolyse spontanée du composé pendant les expériences de RMN qui peuvent durer plusieurs heures.

Par contre, la liaison amide ne permet vraisemblablement pas d'empêcher l'hydrolyse de ce composé par la PTH puisque Shiloach *et al.* ont montré, dans le cas du (Gly)<sub>3</sub>-Phe-ARNt<sup>Phe</sup>, que le remplacement de la liaison ester par un liaison amide ne diminuait que d'un facteur 30 la valeur du  $k_{cat}/K_m$  de la PTH pour ce substrat (Shiloach *et al.*, 1975). C'est pour ralentir l'hydrolyse enzymatique de notre substrat, que nous avons utilisé le variant H20A de cette enzyme, comme décrit au début du chapitre 1. Nous avons d'ailleurs vérifié qu'aucune hydrolyse de la diacétyl-Lys-(3'NH)-adénosine n'était détectable après 15 jours d'incubation à une concentration de 900  $\mu\text{M}$ , en présence de 150  $\mu\text{M}$  de PTH H20A, d'après un spectre <sup>1</sup>H de ce composé avant et après incubation dans les conditions de la RMN.

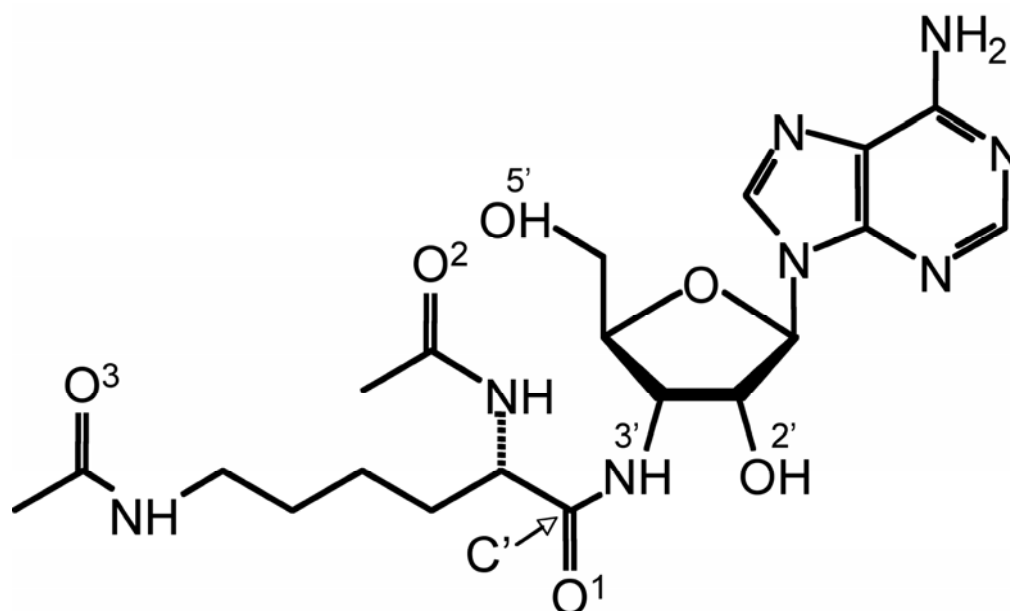


Figure 31 : Structure de la diacétyl-Lys-(3'NH)-adénosine

## 2.2 Affinité de la diacétyl-Lys-(3'NH)-adénosine pour la PTH

Afin de savoir si la diacétyl-Lys-(3'NH)-adénosine interagissait avec la PTH, nous avons déterminé si elle inhibait l'hydrolyse du diacétyl-Lys-ARNt<sup>Lys</sup> catalysée par cette enzyme. Cette expérience a été réalisée dans les conditions de la RMN (pH 6,0, en présence de 50 mM d'acétate de sodium et de 200 mM de NaCl). Nous avons obtenu un  $K_i$  de 12 mM, c'est-à-dire de 3 ordres de grandeurs plus élevé que celui obtenu pour l'ARNt<sup>His</sup> qui est de 15  $\mu$ M dans les mêmes conditions. Cette valeur reflète la constante de dissociation du complexe formé entre la diacétyl-Lys-(3'NH)-adénosine et la PTH. L'affinité relativement faible de ce ligand est cependant compatible avec une étude par RMN de son interaction avec la PTH.

## 2.3 Principe de la détermination des interactions protéine-ligand par RMN

La cartographie des zones d'interaction par RMN repose sur l'analyse et la comparaison de spectres  $^{15}\text{N}$ -HSQC en présence et en absence de ligand. Chaque pic observé sur ces spectres correspond à un couple  $\text{H}^{\text{N}}-\text{N}$  du squelette peptidique de la protéine et est donc caractéristique d'un résidu de la protéine. Ces spectres permettent aussi d'observer l'évolution des pics des couples  $\text{H}^{\text{N}}-\text{N}$  des chaînes latérales de certains résidus comme les asparagines et les glutamines.

Les pics des résidus impliqués dans l'interaction peuvent subir une variation de leur déplacement chimique à cause d'un changement de leur environnement chimique. L'identification de ces résidus permet donc d'obtenir une empreinte du ligand sur la protéine. Plus précisément, on peut définir une distance ( $\Delta_i$ ) qui reflète la variation des déplacements chimiques du pic du résidu  $i$  entre la forme libre et la forme liée. Cette distance prend en compte les déplacements chimiques des atomes  $\text{H}^{\text{N}}$  et  $\text{N}$  en fonction de leur rapport gyromagnétique moyen :

$$\Delta_i = [(\Delta\delta_{\text{N}}/10)^2 + (\Delta\delta_{\text{H}})^2]^{1/2}$$

$\Delta\delta_{\text{N}}$  : différence des déplacements chimiques  $^{15}\text{N}$  de l'atome  $\text{N}$  du résidu  $i$  entre l'état libre et l'état lié (en ppm).

$\Delta\delta_{\text{H}}$  : différence des déplacements chimiques  $^1\text{H}$  du proton  $\text{H}^{\text{N}}$  du résidu  $i$  entre l'état libre et l'état lié (en ppm).

On considère ensuite que si la distance  $\Delta_i$  est supérieure à une certaine valeur, correspondant par exemple au double de l'écart-type de la distribution obtenue pour les  $\Delta_i$ , la variation de déplacement chimique est assez significative pour que le résidu soit impliqué dans l'interaction.

Les pics des résidus impliqués dans l'interaction peuvent aussi subir une diminution d'intensité mesurée par l'indice de perturbation d'intensité ( $\text{IPI}_i$ ) qui reflète la variation d'intensité du pic du résidu  $i$  entre la forme libre et la forme liée.

$$\text{IPI}_i = (\text{I}_i \text{ libre} - \text{I}_i \text{ liée}) / \text{I}_i \text{ libre}$$

$\text{I}_i \text{ libre}$  : intensité du pic du résidu  $i$  dans la forme libre de la protéine

$\text{I}_i \text{ liée}$  : intensité du pic du résidu  $i$  dans la forme liée de la protéine

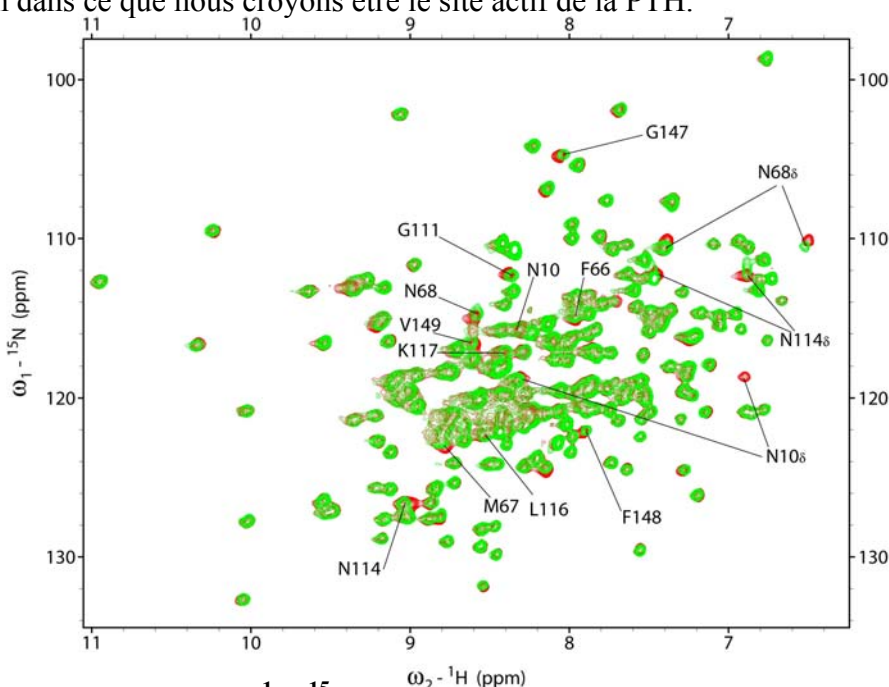
Si le régime d'échange entre les deux molécules est assez rapide ( $2*\pi * \nu_{\text{échange}} \gg |\omega_{\text{libre}} - \omega_{\text{liée}}|$ , où  $\nu_{\text{échange}}$  est la fréquence de l'échange et  $\omega_{\text{libre}}$  et  $\omega_{\text{liée}}$  sont, respectivement, les pulsations d'un atome dans l'état libre et dans l'état lié de la protéine), nous n'observerons pas deux pics (l'un correspondant à la résonance de la forme libre et l'autre à celle de la forme liée) mais un seul pic à

une fréquence intermédiaire. Si le régime d'échange est lent, les deux formes libre et liée coexisteront et on verra les pics correspondant à ces deux formes avec des intensités proportionnelles à leurs concentrations respectives ( $2 * \pi * \nu_{\text{échange}} \gg |\omega_{\text{libre}} - \omega_{\text{liée}}|$ ). Ainsi, aucun écart dans les déplacements chimiques ne sera observé pour le pic initial et l'IPI se révélera le meilleur indice pour identifier les résidus affectés par la présence du ligand. Pour certaines valeurs de  $\nu_{\text{échange}}$ , il est possible que, pour une partie des atomes de la protéine, on soit en échange rapide et que, pour l'autre, on soit en échange lent. Dans le cas d'un échange intermédiaire, certains pics pourront donc à la fois se déplacer et varier en intensité alors que d'autres ne feront que l'un ou l'autre.

## 2.4 Empreinte de la diacétyl-Lys-(3'NH)-adénosine sur la PTH

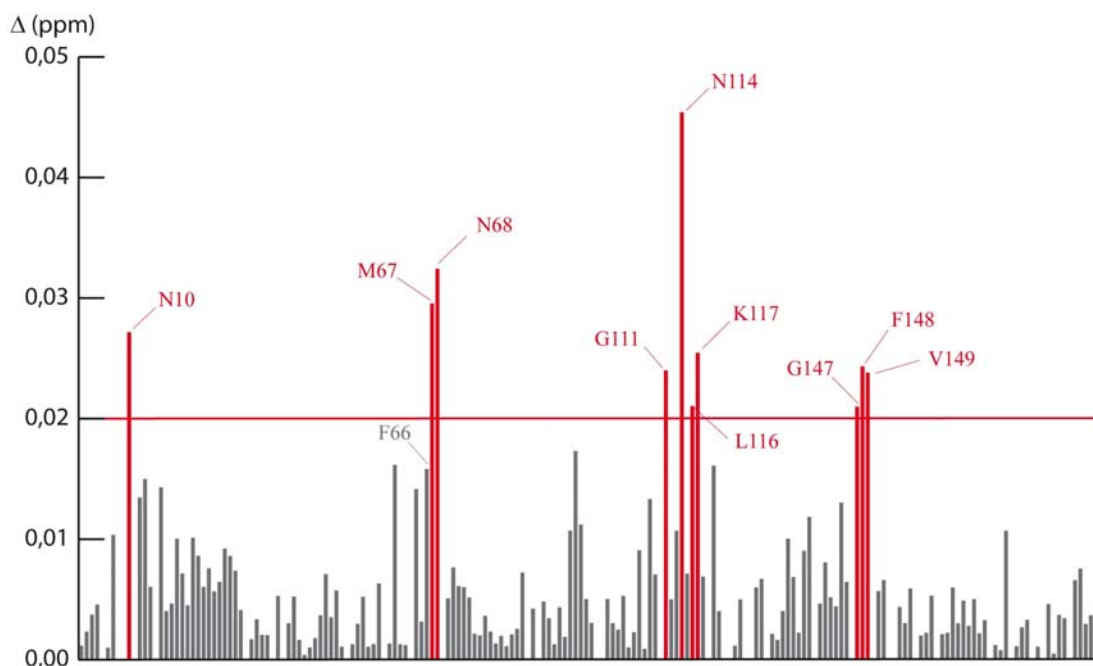
Nous avons donc enregistré des spectres  $^1\text{H}$ - $^{15}\text{N}$ -HSQC de la protéine à 150  $\mu\text{M}$  en présence ou en absence de ligand. Comme on peut le voir sur la figure 32, la plupart des pics restent parfaitement superposables quand on compare les spectres enregistrés en absence ou en présence de 900  $\mu\text{M}$  de diacétyl-Lys-(3'NH)-adénosine. D'autres pics se déplacent, indiquant un échange rapide entre les formes libre et liée de la protéine. Si l'on considère que le  $K_d$  du ligand pour la protéine est de 12 mM (identique au  $K_i$ ), le taux de saturation de la protéine (à 150  $\mu\text{M}$ ) par le ligand (à 900  $\mu\text{M}$ ) est alors d'environ 7%.

Dix résidus ont un déplacement chimique affecté de manière significative par la présence de diacétyl-Lys-(3'NH)-adénosine (en rouge sur la figure 33), même si leur  $\Delta\delta$  ne dépasse pas 0,05 ppm à cause de la faible saturation de la protéine par le ligand. Ils appartiennent tous à la région réputée être celle du site actif (N10, M67, N68, G111, N114, L116, K117) ou de la boucle-hélice qui surplombe ce site actif (G147, F148, V149) (voir figure 34). De plus, les résonances des couples  $\text{H}^\delta$ - $\text{N}^\delta$  des chaînes latérales des asparagines N10, N68 et N114 sont très fortement affectées par l'addition de diacétyl-Lys-(3'NH)-adénosine. Ces résultats indiquent que la diacétyl-Lys-(3'NH)-adénosine se fixe bien dans ce que nous croyons être le site actif de la PTH.



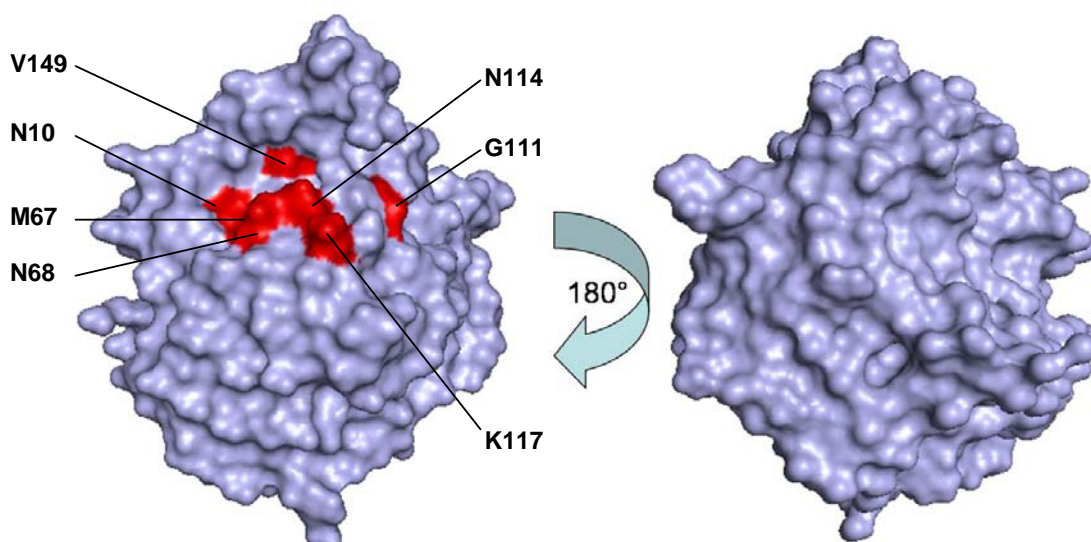
**Figure 32 : Spectres  $^1\text{H}$ - $^{15}\text{N}$ -HSQC de la PTH H20A en présence ou en absence de diacétyl-Lys-(3'NH)-adénosine**

Spectre  $^1\text{H}$ - $^{15}\text{N}$ -HSQC obtenu en présence de 150  $\mu\text{M}$  de PTH et 900  $\mu\text{M}$  de diacétyl-Lys-(3'NH)-adénosine) (vert), superposé à celui obtenu en présence de 150  $\mu\text{M}$  de PTH seule (rouge). Les expériences ont été réalisées à 17°C, dans un tampon contenant 50 mM d'acétate de sodium (pH 6,0) et 200 mM de chlorure de sodium. Pour les résidus les plus affectés, des flèches indiquent les positions des pics de leurs couples  $\text{H}^N$ - $\text{N}$  et des couples  $\text{H}^\delta$ - $\text{N}^\delta$  pour les asparagines.



**Figure 33 : Variation du déplacement chimique des pics des  $H^N-N$  de la protéine lors de l'ajout de diacétyl-Lys-(3'NH)-adénosine**

Pour chaque résidu  $i$  (1 à 193) de la protéine, la distance  $\Delta_i$  entre les spectres  $^{15}N$ -HSQC de la PTH en présence ou en absence de 900  $\mu M$  de diacétyl-Lys-(3'NH)-adénosine est indiquée sur l'histogramme. Les résidus possédant un  $\Delta_i$  supérieur à 0,02 ppm sont représentés en rouge et annotés, ainsi que F66. Les  $\Delta_i$  des deux  $H^\delta-N^\delta$  des chaînes latérales de N10, N68 et N114 ne sont pas représentés sur cette figure. Ils valent, respectivement, 0,029 et 0,04 ppm ; 0,042 et 0,037 ppm ; 0,046 et 0,043 ppm. Les  $\Delta_i$  des  $H^\delta-N^\delta$  et des  $H^\epsilon-N^\epsilon$  des chaînes latérales des autres asparagines et glutamines sont inférieures 0,01 ppm à l'exception de celui d'un  $H^\delta-N^\delta$  de N21 qui vaut 0,031 ppm. Une valeur de 0 ppm a été donnée pour les prolines et pour les pics superposés pour lesquels il a été impossible de mettre en évidence un déplacement éventuel.



**Figure 34 : Empreinte de la diacétyl-Lys-(3'NH)-adénosine sur la PTH**

Représentation 3D de la surface de la PTH d'*E. coli* (code PDB 2PTH), où les résidus sensibles à la fixation de la diacétyl-Lys-(3'NH)-adénosine ( $\Delta_i > 0,02$  ppm) sont colorés en rouge. Tous ces résidus sont concentrés dans une même région. Leu116, Gly157 et Phe148, légèrement enfouis, ne sont pas visibles sur cette représentation. Cette figure a été réalisée à l'aide du logiciel Pymol.

## 2.5 Positionnement de la diacétyl-Lys-(3'NH)-adénosine sur la PTH

### 2.5.1 Construction des modèles

Pour obtenir un modèle du complexe entre la protéine et la diacétyl-Lys-(3'NH)-adénosine, le programme HADDOCK (*High Ambiguity Driven biomolecular DOCKing*) a été mis en œuvre (Dominguez *et al.*, 2003). Pour définir la structure d'un complexe protéine–ligand, ce logiciel utilise des données énergétiques intra- et intermoléculaires, mais également des données expérimentales comme, par exemple, celles qui sont fournies par des expériences de RMN. Un de ses gros avantages est qu'il permet de définir des contraintes floues spécifiant qu'une liste choisie de résidus de la protéine est en contact avec le ligand, mais ne spécifiant pas précisément ces contacts.

Comme on l'a vu ci-dessus, les expériences de RMN indiquent que les pics  $H^N-N$  de 10 résidus de la PTH se déplacent d'une distance  $\Delta$  supérieure à 0,02 ppm en présence de diacétyl-Lys-(3'NH)-adénosine. Cette variation de déplacement chimique reflète soit un contact direct entre le ligand et le résidu de la protéine auquel appartiennent ces atomes N et  $H^N$ , soit une perturbation structurale de ce résidu induite par un contact entre le ligand et un autre résidu. Comme il était difficile de savoir *a priori* laquelle de ces deux hypothèses était la bonne, nous avons tout d'abord inclus, dans la liste des résidus devant être au contact du ligand, l'ensemble des résidus fortement affectés par la présence de diacétyl-Lys-(3'NH)-adénosine : N10, M67, N68, G111, N114, L116, K117, G147, F148 et V149.

Pendant une première minimisation énergétique en corps-rigide, le programme a généré 1000 modèles et les 200 modèles avec l'énergie d'interaction intermoléculaire la plus basse ont été utilisés par HADDOCK pour l'affinement semi-flexible.

Les 200 modèles affinés comportent un grand nombre de disparités, ce qui reflète vraisemblablement l'impossibilité pour la diacétyl-Lys-(3'NH)-adénosine de se fixer sur la protéine en respectant toutes les contraintes imposées. En particulier, les résidus G111 et K117, excentrés par rapport aux autres (en jaune sur la figure 35), ne sont en contact avec le ligand que dans 10 et 27 modèles, respectivement. Cette sous-représentation des contacts entre le ligand et ces deux résidus suggère que ces résidus ne sont pas en contact direct avec le ligand et que les variations de leurs déplacements chimiques résultent d'un réarrangement structural. Cette conclusion est compatible avec une ouverture du site actif, lors de l'occupation de celui-ci, telle qu'elle a été suggérée par Selvaraj *et al.* (Selvaraj *et al.*, 2007).

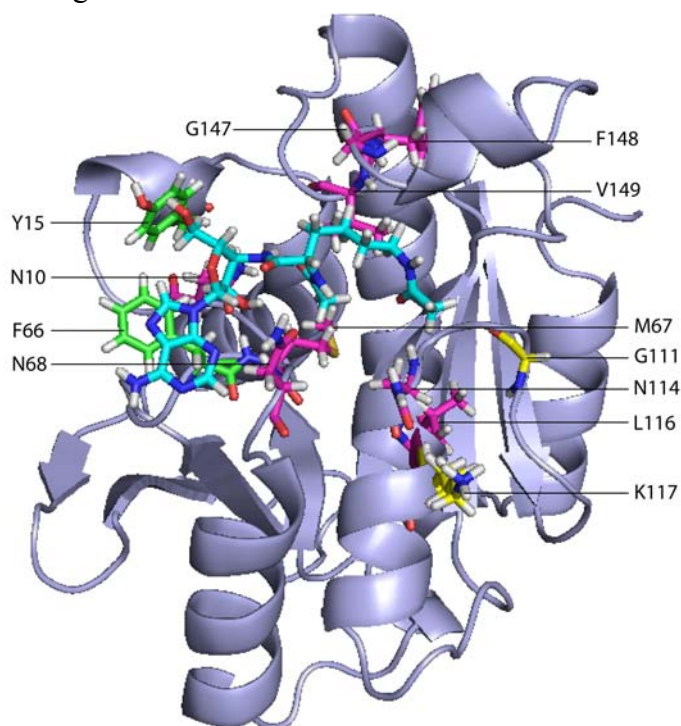


Pour améliorer la cohérence des modèles obtenus, nous avons donc réalisé un deuxième calcul en supprimant les contraintes sur G111 et K117. Les 200 modèles obtenus après affinement possèdent encore des différences mais la variabilité est nettement plus faible qu'après la première série de calculs. On retrouve d'ailleurs une même orientation de la diacétyl-Lys-(3'NH)-adénosine sur la PTH dans environ 50% des modèles, dont celui possédant le meilleur score HADDOCK. Ce score HADDOCK prend en compte, pour chaque modèle, à la fois l'énergie totale d'interaction du complexe et le respect des contraintes floues. La proportion des modèles ayant cette orientation est plus grande encore lorsque l'on considère ceux qui ont les meilleurs scores puisqu'elle atteint 80% pour les 50 meilleurs modèles. Nous avons donc retenu les 20 modèles possédant les meilleurs scores pour les étudier en détail.

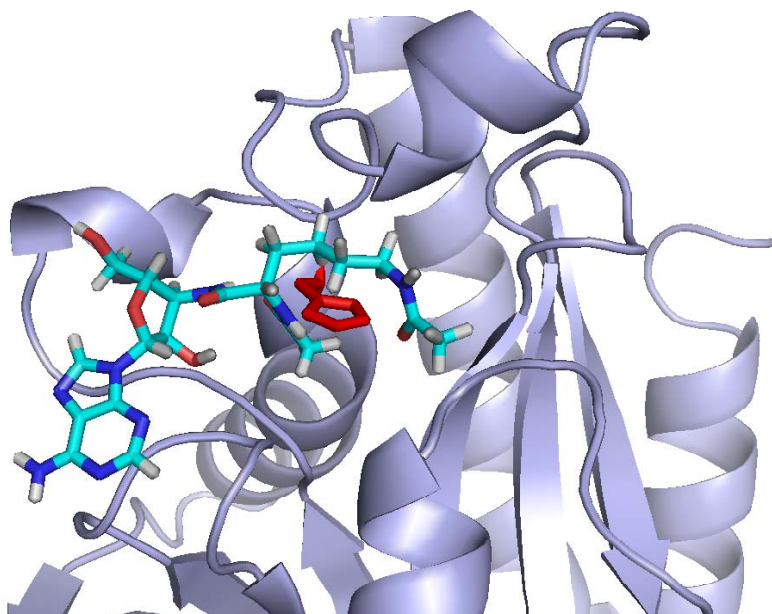
Pour 19 de ces 20 modèles, dont celui représenté sur la figure 35 qui possède le meilleur score HADDOCK, la diacétyl-Lys-(3'NH)-adénosine se positionne dans le site actif putatif de la PTH avec son adénine au voisinage de F66 et l'extrémité de la chaîne latérale de sa lysine acétylée au voisinage de N114. On peut noter que cette orientation correspond aussi au modèle possédant le meilleur score dans la première modélisation réalisée, incluant les contraintes sur G111 et K117. Dans ces 19 modèles, la position du carbone C' (voir figure 31) du groupement carbonyle de la liaison amide entre l'adénosine et la lysine diacétylée est constante. La distance moyenne de ce carbone à celui dans le modèle possédant le meilleur score est de 1,2 Å (avec un écart-type de 0,75 Å) (voir tableau 7, dernière colonne). Un seul parmi les 20 modèles est orienté dans le sens opposé, avec l'adénine très éloignée de F66. Comme on ne retrouve une telle orientation que dans une dizaine d'autres modèles sur les 200 calculés, nous ne nous y sommes pas intéressés davantage.

**Figure 35 : Modèle du complexe entre la diacétyl-Lys-(3'NH)-adénosine et la PTH**

Pour le modèle possédant le meilleur score HADDOCK, la protéine est représentée en gris à l'aide du logiciel Pymol (en mode *cartoon*). Les chaînes latérales des résidus de la liste de contraintes floues utilisée dans le deuxième calcul HADDOCK sont représentées en magenta, celles de G111 et K117 sont représentées en jaune et celles de F66 et Y15 en vert. La diacétyl-Lys-(3'NH)-adénosine est représentée en cyan.



Dans ces 19 modèles, nous avons ajouté la chaîne latérale de l'histidine 20 à la position qu'elle occupe dans la structure cristallographique de la PTH d'*E. coli*. A l'exception de 2 modèles, la position du ligand est clairement compatible avec la présence d'un résidu d'histidine, même s'il n'avait pas été inclus dans notre calcul réalisé avec le variant H20A. A titre d'exemple, la figure 36 montre la position de l'histidine 20 dans le cas du modèle possédant le meilleur score HADDOCK.



**Figure 36 : Positions relatives de l'histidine 20 et de la diacétyl-Lys-(3'NH)-adénosine**

La chaîne latérale de l'histidine 20 (en rouge) a été rajoutée sur le modèle du complexe entre la diacétyl-Lys-(3'NH)-adénosine (en cyan) et la PTH (en gris) possédant le meilleur score HADDOCK. L'orientation de la structure de la PTH est la même qu'à la figure 35.

### 2.5.2 Interaction avec F66

Les 19 modèles considérés ci-dessus indiquent une interaction hydrophobe entre le cycle aromatique de F66 de la protéine et l'adénine du ligand, qui se retrouvent à chaque fois à une distance inférieure à 4 Å l'un de l'autre. Dans l'expérience de RMN, F66 fait partie des résidus affectés par la présence de ligand (voir figure 33), même s'il n'a pas été pris en compte dans la liste des contraintes floues puisque son delta est de 0,016 ppm, légèrement inférieur à la limite de 0,02 ppm que nous avons retenue. On peut noter qu'en revanche, M67 et N68 dans le voisinage direct de F66 font, eux, partie de cette liste. Un rôle de F66 dans la réaction d'hydrolyse catalysée par la PTH est confirmé par le fait qu'une mutation de ce résidu en alanine diminue l'efficacité catalytique d'un facteur 5 (voir tableau 3 dans l'Introduction). D'autre part, un alignement disponible sur le site Pfam (<http://pfam.sanger.ac.uk>), réalisé à partir des 1607 séquences de PTH, montre que la position 66 (par rapport à la numérotation de la séquence de la PTH d'*E. coli*) est occupée par une phénylalanine dans 47% des cas et par une tyrosine dans

49% des cas. Ceci suggère que la présence à cet endroit d'une chaîne latérale aromatique est importante pour le fonctionnement de la PTH.

L'idée d'une proximité entre le cycle aromatique de F66 et le ligand est soutenue par le résultat de l'expérience de RMN  $^{15}\text{N}$ -NOESY-HSQC. En effet, la corrélation NOESY entre un proton résonnant à 7,3 ppm et le  $\text{H}^{\text{N}}$  de F66 disparaît lors de l'addition de diacétyl-Lys-(3'NH)-adénosine. Or, si on considère les protons attribués qui, d'après la structure cristallographique de la PTH d'*E. coli*, sont situés au voisinage du  $\text{H}^{\text{N}}$  de F66, on constate qu'aucun d'entre eux ne possède cette valeur de déplacement chimique et que, parmi les protons non attribués, seuls ceux du cycle de F66 pourraient avoir cette valeur. Ceci suggère que le pic à 7,3 ppm correspond à un proton  $\text{H}^{\delta}$  de F66. Un pic à 7,3 ppm est aussi visible sur une expérience 3D-NOESY dans  $\text{D}_2\text{O}$  à partir du  $\text{H}^{\alpha}$  de F66 (à 3,4 Å du  $\text{H}^{\delta}$  dans le cristal) et du  $\text{H}^{\beta}$  de F66 (à 2,5 Å du  $\text{H}^{\delta}$  dans le cristal), ce qui confirme cette attribution. La disparition de ce pic lors de l'addition du ligand indique que l'environnement de F66 est fortement modifié par la présence du ligand.

### **2.5.3 Interaction avec N10, le résidu à la base de la discrimination entre aminoacyl-ARNt et peptidyl-ARNt**

Dans 6 modèles, dont les deux possédant le meilleur score HADDOCK, un atome d'hydrogène  $\text{H}^{\delta}$  du groupement  $\text{NH}_2$  de la chaîne latérale de N10 se trouve à une distance inférieure à 4 Å d'un atome d'oxygène  $\text{O}^2$  du groupement acétate porté par le groupement  $\alpha\text{-NH}_2$  de la lysine du ligand (voir tableau 7 et figure 37(A)). Cette distance est compatible avec la formation d'une liaison hydrogène. Une telle interaction correspond à celle qui a été mise en évidence dans les cristaux de la PTH d'*E. coli*, (voir figure 37 ci-dessous) où l'extrémité C-terminale (K191-A192-Q193) d'une molécule de protéine voisine dans le cristal occupe le site actif, mimant ainsi le produit de la réaction d'hydrolyse de la PTH. Dans ce cas, le groupement  $\text{NH}_2$  de N10 est proche de l'oxygène du groupement carbonyle du deuxième résidu en partant de l'extrémité C-terminale, c'est-à-dire de l'oxygène qui correspond à  $\text{O}^2$  sur la diacétyl-Lys-(3'NH)-adénosine.

Ce résultat indique que N10 interagit avec la liaison amide que doit posséder un aminoacyl-ARNt dont le groupement N-terminal de l'acide aminé est modifié, pour être un substrat de la PTH. Pour confirmer ce résultat, nous avons comparé la vitesse initiale à laquelle le diacétyl-Lys-ARNt<sup>Lys</sup> et le Lys-ARNt<sup>Lys</sup> étaient hydrolysés par la PTH sauvage, ou par ses mutants N10A ou N10D (voir tableau 6). L'enzyme sauvage hydrolyse le diacétyl-Lys-ARNt<sup>Lys</sup> 360 fois plus efficacement que le Lys-ARNt<sup>Lys</sup> alors que, pour les mutants N10A et N10D, le rapport de la vitesse d'hydrolyse du diacétyl-Lys-ARNt<sup>Lys</sup> sur celle du Lys-ARNt<sup>Lys</sup> n'est plus que de 3,6 et 0,8, respectivement. Ce résultat confirme que ce résidu, N10, qui est strictement conservé dans les séquences des PTH, confère à l'enzyme la capacité de discriminer positivement les aminoacyl-ARNt dont le groupement N-terminal de l'acide aminé est modifié.

PTH	Diacétyl-Lys-ARNt <sup>Lys</sup>	Lys-ARNt <sup>Lys</sup>
sauvage	720	2,0
N10A	2,6	0,73
N10D	0,92	1,1

**Tableau 6 : Comparaison de la spécificité de la PTH sauvage et des mutants N10A et N10D**

Les vitesses initiales d'hydrolyse, mesurées à 28°C en présence de 20 mM Tris-HCl (pH 7.5), 10 mM MgCl<sub>2</sub>, 0,1 mM EDTA, 0,1 mM dithiothreitol et 1,5 μM de substrat, sont indiqués pour chaque mutant et chaque substrat en 10<sup>-3</sup> s<sup>-1</sup> avec une marge d'erreur de ±15%.

	score haddock	Distance N68 (H <sup>δ</sup> ) O1	Distance N10 (H <sup>δ</sup> ) O2	Distance N10 (H <sup>δ</sup> ) O2'	Distance N114 (H <sup>N</sup> ) O3	Distance Y15 (O <sup>H</sup> ) O5	Distance N10 (H <sup>δ</sup> ) O1	Distance N68 (H <sup>N</sup> ) O5'	Distance C' C'(1)*
1	-44,3	<b>2,68</b>	<b>2,72</b>	<b>3,01</b>	<b>2,54</b>	<b>3,32</b>	5,28	8,83	0
2	-42,42	<b>1,71</b>	<b>2,97</b>	<b>2,56</b>	<b>2,26</b>	<b>2,7</b>	5,19	7,87	0,6
3	-42,4	<b>1,74</b>	5,83	<b>2,61</b>	4,24	<b>2,82</b>	5,83	4,12	1,56
4	-41,8	6,91	5,64	4,71	<b>2,77</b>	7,64	<b>1,58</b>	6,25	1,44
5	-41,1	8,68	7,29	5,23	<b>3,51</b>	12,38	4,88	6,59	3,34
6	-41	8,32	8,66	4,68	<b>2,81</b>	7,61	4,78	<b>2,18</b>	0,84
7	-40,6	8,95	6,98	5,46	<b>2,94</b>	8,6	<b>2,9</b>	<b>2,77</b>	0,6
8	-40,29	7,28	4,73	5	<b>1,6</b>	<b>3,27</b>	6,82	6,61	1,19
9	-40,13	<b>1,64</b>	7,2	<b>3,59</b>	5,65	<b>3,29</b>	6,11	8,77	1,3
10	-39	9,16	7,14	5,49	<b>3,09</b>	8	<b>3,06</b>	<b>2,43</b>	0,81
11	-38,58	<b>1,81</b>	6,99	<b>3,24</b>	<b>1,74</b>	<b>4,75</b>	5,34	9,83	0,67
12	-38,5	4,16	6,87	<b>2,43</b>	<b>2,83</b>	7,95	3,99	<b>3,19</b>	1,22
13	-38,3	7,22	8,17	6,09	<b>1,93</b>	7,48	4,79	<b>3,44</b>	2,88
14	-38,14	<b>1,75</b>	<b>3,51</b>	<b>2,56</b>	<b>1,8</b>	<b>4,44</b>	5,19	8,86	1,16
16	-37,8	<b>1,65</b>	4,31	<b>2,81</b>	<b>1,81</b>	6,4	6,19	9,41	1,34
17	-37,7	<b>1,69</b>	<b>3,95</b>	<b>2,49</b>	<b>1,66</b>	5,94	5,86	9,06	1,49
18	-36,9	<b>1,64</b>	<b>2,61</b>	<b>3,01</b>	4,32	7,03	6,01	10,2	1,2
19	-36,7	7,76	8,87	5,97	<b>2,94</b>	8,77	5,53	5,99	0,73
20	-36,4	9,59	<b>3,28</b>	5,32	<b>2,94</b>	7,89	<b>3,28</b>	<b>2,24</b>	1,01

**Tableau 7 : Distances des interactions entre la protéine et la diacétyl-Lys-(3'NH)-adénosine**

Dans ce tableau, on considère les 20 modèles possédant le meilleur score HADDOCK, à l'exception du 15<sup>ème</sup>, dont l'orientation du ligand est très différente des autres. Les modèles sont classés suivant leur score HADDOCK, donné à titre indicatif dans une unité arbitraire (colonne 2). Pour chaque modèle, les distances entre certains atomes de la protéine et certains atomes du ligand (voir première ligne), sont données en Å. Les distances permettant une liaison hydrogène entre les deux molécules sont écrites en gras dans les cases grisées (<4 Å dans le cas d'une interaction entre un atome d'oxygène et un proton, <5 Å dans le cas d'une interaction entre deux atomes d'oxygène ; cette distance de 5 Å prend en compte la présence d'un atome d'hydrogène entre les deux atomes d'oxygène (O—H.....O)). Pour les deux protons des chaînes latérales des asparagines H<sup>δ</sup>, la distance donnée est celle entre le H<sup>δ</sup> le plus proche du ligand et l'atome concerné du ligand.

\* La dernière colonne montre la distance entre le C' du ligand considéré et celui du modèle possédant le meilleur score, après superposition des deux complexes.

Dans 5 des 6 modèles évoqués ci-dessus, dont les deux meilleurs scores, le groupement  $\text{NH}_2$  de la chaîne latérale de N10 est à une distance permettant une liaison hydrogène à la fois avec  $\text{O}^2$ , appartenant à la région peptidique du ligand et avec l'atome  $\text{O}^2$  du ribose du ligand. Cela suggère qu'en plus de son implication dans la reconnaissance spécifique des peptidyl-ARNt, N10 pourrait aussi participer à la stabilisation du substrat dans le site actif, via  $\text{O}^2$ .

#### **2.5.4 Interaction avec N68**

Dans 9 modèles sur les 19 dont les 3 possédant le meilleur score, le groupement  $\text{NH}_2$  de la chaîne latérale de N68 est situé à une distance inférieure à 2,8 Å de l'atome d'oxygène  $\text{O}^1$  du groupement amide du ligand. Cette interaction rappelle, là aussi, celle qui existe dans le cristal, entre le groupement  $\text{NH}_2$  de N68 et un atome d'oxygène du groupement carboxylique C-terminal de la protéine voisine qui mime le produit de la réaction. La conservation de N68 dans presque toutes les séquences de Pth disponibles est en faveur de l'importance de ce contact. Rappelons également que des expériences de mutagenèse dirigée ont montré qu'un variant N68A avait une efficacité catalytique 100 fois inférieure à celle de l'enzyme sauvage.

Dans le cristal, le N de la chaîne principale de N68 est en contact avec le groupement carbonyle de la lysine 191, qui est le troisième résidu en partant de l'extrémité C-terminale de la protéine voisine présente dans le site actif de la PTH. Dans nos modèles, cette interaction est impossible car l'empilement de l'adénine du ligand sur F66 positionne le ribose du ligand à la place qu'occupent la lysine 191 et l'alanine 192 de la protéine voisine, dans le cristal.

#### **2.5.5 Interaction avec N114**

Dans 16 cas sur 19, le  $\text{H}^{\text{N}}$  de N114 est situé à une distance comprise entre 1,6 et 3,5 Å de  $\text{O}^3$ , l'oxygène du groupement acétyle en position  $\epsilon$  de la chaîne latérale de la lysine du ligand (voir tableau 7 et figure 37(A)), ce qui est compatible avec une interaction de type liaison hydrogène. Dans les 3 autres modèles, la distance est plus grande, tout en restant toujours inférieure à 5,7 Å. Il n'y a cependant pas d'équivalent de ce groupement acétyle dans le cas d'un peptidyl-ARNt naturel et cette interaction n'existe donc pas dans le cas d'un substrat naturel.

Dans le cristal, on observe aussi une interaction entre le groupement  $\text{NH}_2$  de N114 et un atome d'oxygène du groupement carboxyle C-terminal de la protéine mimant le produit de la réaction (voir figure 37(B)). Un des deux atomes d'oxygène de ce groupement correspond à celui qui est présent dans le substrat ( $\text{O}^1$ ), l'autre ( $\text{O}^4$ ) est fourni par la molécule d'eau qui attaque la liaison ester dans le substrat. Si on admet que, dans le cristal, N114 se lie à  $\text{O}^4$ , il est logique de ne pas retrouver cette interaction dans le cas d'un analogue de substrat ne possédant

pas l'équivalent de ce deuxième atome d'oxygène. Le résidu N114 étant au coeur du site actif putatif, l'environnement chimique de son groupement NH<sub>2</sub> est certainement sensible à la fixation de la diacétyl-Lys-(3'NH)-adénosine, ce qui peut expliquer une variation de son déplacement chimique. Dans certains modèles, N114 semble d'ailleurs pouvoir interagir avec K117 ou N68 via des liaisons hydrogène entre leurs chaînes latérales.

### **2.5.6 Interaction avec Tyr15**

Dans 7 des 19 modèles du complexe, dont les 2 possédant le meilleur score, on observe une distance assez faible (de 2,7 à 4,75 Å), entre l'atome d'oxygène du groupement hydroxyle de Y15 et celui du groupement hydroxyle porté par le carbone 5' du ribose du ligand. Avec un authentique diacétyl-Lys-ARNt<sup>Lys</sup>, il pourrait donc se former une liaison hydrogène entre la Tyr15 et le groupement phosphate en 5' de l'adénosine 3'-terminale de l'ARNt, ce qui pourrait expliquer la perte d'un facteur 6 dans la valeur du  $k_{cat}/K_m$  quand on mute la Tyr15 en alanine (voir tableau 3). Dans le cristal, le groupement hydroxyle d'Y15 semble former une liaison hydrogène avec l'oxygène de la chaîne principale de l'histidine 188, le résidu n-5 du peptide présent dans le site actif. Il semble cependant peu probable que cette interaction intervienne lors de l'hydrolyse par la PTH d'un peptidyl-ARNt. En effet le  $k_{cat}/K_m$  de la PTH envers un diacétyl-Lys-ARNt<sup>Lys</sup> diminue d'un facteur 6 quand cette tyrosine est mutée en alanine alors que ce substrat mime un dipeptidyl-ARNt ne possédant pas de résidu n-5 avec lequel Y15 pourrait interagir. Ce résultat confirme que cette tyrosine possède un autre rôle que celui de reconnaître le résidu n-5 de la région peptidique du substrat. De plus, la forte conservation de cette tyrosine (93%) souligne l'importance de ce résidu.

### **2.5.7 Autres interactions**

Enfin, les positions des résidus G147, F148 et V149, dont les résonances sont affectées par la fixation de la diacétyl-Lys-(3'NH)-adénosine, semblent compatibles avec une interaction hydrophobe avec la chaîne latérale de la lysine du ligand.

Comme on peut le voir sur la figure 35, les résidus apolaires G147, F148 et V149, qui appartiennent à la région de la boucle-hélice surplombant le site actif, sont au voisinage de la partie aliphatique de la chaîne latérale de la lysine du ligand, ce qui suggère une interaction hydrophobe favorable. Cependant, la chaîne latérale du premier acide aminé estérifié à l'ARN sur un peptidyl-ARNt peut être, en principe, aussi bien apolaire que polaire et nous pensons donc que cette interaction est anecdotique. En présence d'un substrat contenant un ARNt, il est possible que ces 3 résidus interagissent avec les pyrimidines à l'extrémité 3' du polynucléotide.

Le groupement acétate porté par le groupement  $\alpha$ -NH<sub>2</sub> de la lysine (qui mime l'extrémité d'un acide aminé supplémentaire) adopte deux positions distinctes dans les différents modèles. Il est orienté vers N10 dans 7 modèles dont les deux meilleurs scores et à l'opposé dans les 12 autres. Dans les deux cas, la position de ce groupement est compatible avec l'ajout d'une chaîne peptidique à l'extrémité N-terminale de la lysine, sans qu'il y ait de conflit stérique avec la structure de la PTH.

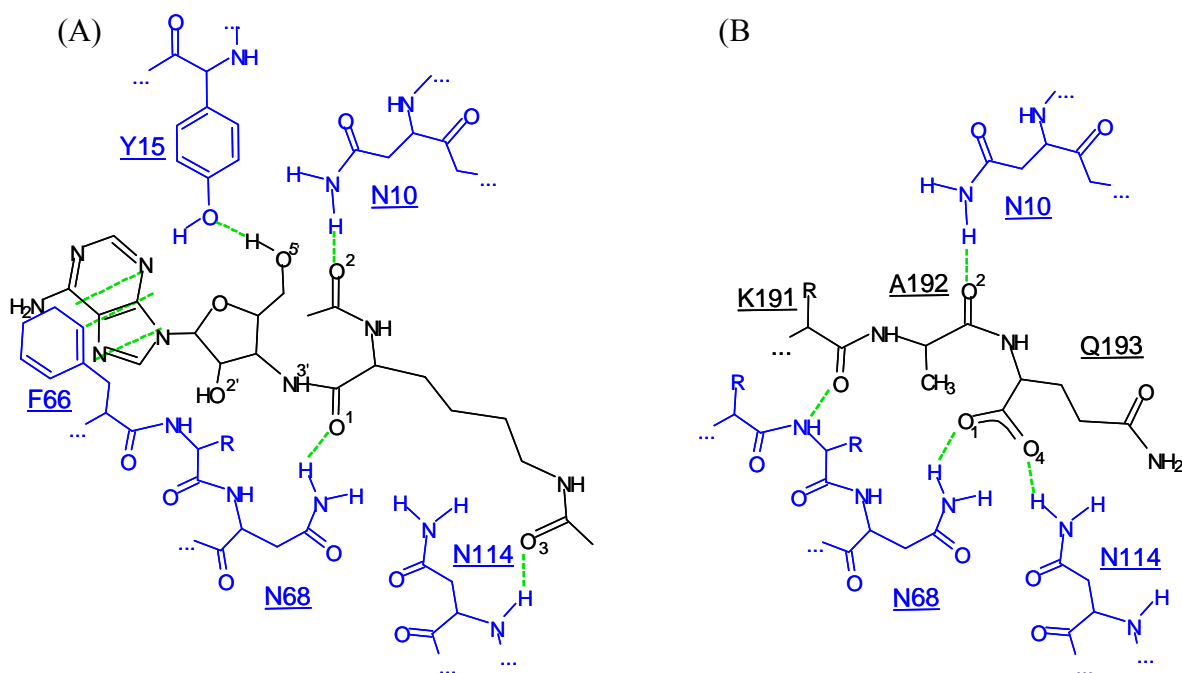
### **2.5.8 Comparaison avec le modèle cristallographique**

A cause de la valeur élevée du  $K_d$  de la diacétyl-Lys-(3'NH)-adénosine pour la protéine on pourrait s'attendre à ce qu'il n'y ait que très peu de contacts entre le ligand et la protéine. Cependant, le modèle possédant le meilleur score HADDOCK montre une demi-douzaine d'interactions entre la protéine et le ligand. Ces interactions sont retrouvées, ensemble ou en partie, dans beaucoup des 19 modèles retenus après notre analyse avec HADDOCK. Il est possible que le ligand puisse utiliser différentes combinaisons d'interactions avec la protéine, sans que toutes ces interactions soient établies en même temps. A l'équilibre, différents conformères enzyme:ligand, de potentiels thermodynamiques voisins s'échangeraient rapidement. Ainsi, les contacts détectés par RMN intégreraient les contacts spécifiques de chaque conformation. Une autre possibilité est que la demi-douzaine d'interactions détectées coexistent dans un même complexe. L'énergie de liaison ainsi récupérée, payant pour des réarrangements géométriques et électronique mutuels, contribuerait peu à la stabilité du complexe protéine : ligand.

La géométrie du complexe telle que nous la déduisons est tout à fait compatible avec la structure cristalline de la PTH liée à son produit putatif. Rappelons-le, le site actif de l'enzyme dans le cristal est occupé par l'extrémité C-terminale d'une molécule de PTH voisine. Dans le schéma ci-dessous (figure 37), nous récapitulons les interactions possibles déduites de l'approche RMN et des études cristallographiques.

Si on passe en 3 dimensions, la position de la lysine de la diacétyl-Lys-(3'NH)-adénosine est assez proche de celle de l'acide aminé C-terminal (Q193) de la molécule simulant le produit de la réaction dans le cristal, puisque leurs C <sup>$\alpha$</sup>  se retrouvent à une distance de 1,13 Å, leurs C <sup>$\beta$</sup>  à 1,06 Å, leurs C <sup>$\gamma$</sup>  à 0,89 Å et leurs C <sup>$\delta$</sup>  à 0,89 Å. Les positions du C' de ces deux résidus sont, quant à elles, un peu plus écartées (3,37 Å). La liaison C <sup>$\alpha$</sup> -C' est en effet orientée différemment (près de 120° de différence) dans les deux modèles (RMN et cristallin). Cette différence peut être due au fait que l'un des complexes est construit avec un analogue du substrat et l'autre avec l'un des deux produits de la réaction d'hydrolyse. Dans le cas de l'analogue du substrat, les interactions entre l'adénosine, d'une part, et la Phe66 et la Tyr15, d'autre part, orientent le C' vers la région de ces deux derniers résidus. Dans le cas du cristal, le C' n'est plus orienté vers cette région puisque l'adénosine n'est pas là, mais est orienté vers N114, grâce à une liaison hydrogène

entre le groupement  $\text{NH}_2$  de ce résidu et l'atome d'oxygène  $\text{O}^4$  sur le groupement carboxyle C-terminal du produit.



**Figure 37 : Interactions dans le site actif de la PTH**

(A) entre la diacétyl-Lys-(3'NH)-adénosine et la PTH HA20. (B) entre la PTH et les trois résidus de l'extrémité C-terminale d'une protéine voisine dans le cristal de la PTH d'*E. coli*. Les résidus de la protéine sont en bleu, la diacétyl-Lys-(3'NH)-adénosine est en noir, les interactions entre la protéine et les ligands sont en vert. Pour la clarté de la figure, la liaison hydrogène possible entre le groupement  $\text{NH}_2$  de N10 et l'atome  $\text{O}^2$  du ribose, n'a pas été représentée.



## **2.6 Conclusions**

Le modèle que nous proposons est en accord avec les résultats de la RMN ainsi qu'avec les prédictions précédemment faites sur la nature du site actif de la PTH. Il suggère que la diacétyl-Lys-(3'NH)-adénosine se lie à la PTH à une position équivalente à celle qu'occupent l'adénosine 3'-terminale et le résidu d'acide aminé dans une molécule de substrat (peptidyl-ARNt). Il indique que la Phe66 joue un rôle important dans le positionnement du substrat de la PTH, en interagissant avec l'adénine 3'-terminale de l'ARNt. La Tyr15 pourrait, elle, participer au positionnement du ribose 3'-terminal de l'ARNt. L'Asn 114 interviendrait pour stabiliser la molécule d'eau activée par l'His20 et responsable de l'attaque nucléophile de la liaison ester dans le substrat. Le modèle est enfin en accord avec l'idée que l'interaction entre deux molécules de PTH dans le cristal reflète bien la liaison de la PTH avec un des produits de sa réaction catalytique. En effet, on retrouve les mêmes liaisons hydrogène impliquant N10 et N68. N10 est, par ailleurs, en grande partie responsable de la discrimination entre un peptidyl-ARNt qui est substrat de la PTH et les aminoacyl-ARNt présents dans la cellule et qui ne sont pas substrats de la PTH.

Cette étude par RMN confirme qu'une petite molécule comme la diacétyl-(3'NH)-adénosine peut établir des interactions spécifiques dans le site actif de la PTH, sans avoir besoin pour cela d'être reliée au corps de l'ARNt. Ce résultat est encourageant si on pense à la possibilité de concevoir des inhibiteurs de la PTH ayant des propriétés anti-bactériennes.

## **Chapitre 3**

# **Interaction entre la protéine et l'ARN**



Après avoir caractérisé l'interaction entre la PTH et un analogue de son substrat venant se loger dans le site catalytique, nous avons cherché à cartographier la région de cette enzyme qui lie la partie nucléique d'un peptidyl-ARNt. Pour cette étude, nous avons choisi d'étudier des ARN dérivés de l'ARNt<sup>His</sup> car l'histidyl-ARNt synthétase est une des rares aminoacyl-ARNt synthétases capables d'aminocycler efficacement des petites hélices d'ARN (Francklyn and Schimmel, 1990 ; Musier-Forsyth and Schimmel, 1993). Cette capacité est en partie due à la présence d'une base supplémentaire G<sub>-1</sub>-C<sub>73</sub>, (voir figure 41) à l'extrémité de la tige acceptrice de l'ARNt<sup>His</sup>, qui est spécifique aux ARNt<sup>His</sup> et qui joue un rôle important pour la reconnaissance par l'HisRS (Francklyn and Schimmel, 1990). Cela donne la possibilité de déterminer comment ces petites hélices, une fois aminoacylées et acétylées, sont reconnues par la PTH. Du point de vue de la RMN, travailler avec des ARN plus petits qu'un ARNt entier peut être intéressant car cela permet de diminuer le poids moléculaire du complexe étudié et, donc, de diminuer la largeur des raies de résonance de ce complexe. De plus, comme on le verra ci-dessous, il se trouve que des mini-ARNt dérivés de l'ARNt<sup>His</sup> d'*E. coli* avaient été synthétisés au laboratoire par Michel Fromant et que, après aminoacylation et acétylation, l'action de la PTH d'*E. coli* sur ces mini-substrats avait déjà été évaluée. Enfin, on sait que l'acétyl-His-ARNt<sup>His</sup> est un bon substrat de la PTH, ce qui montre que la paire de base additionnelle que comporte cet ARN est compatible avec une reconnaissance efficace par la PTH (Fromant *et al.*, 2000).

### 3.1 Interaction de la PTH avec l'ARNt<sup>His</sup>

Une première série d'expériences a été réalisée avec un ARNt<sup>His</sup> entier. Nous avons donc produit et purifié cet ARNt puis nous avons enregistré des spectres HSQC de la PTH à 130  $\mu$ M seule puis en présence de différentes concentrations d'ARNt<sup>His</sup>. Cet ARNt a été ajouté soit sous forme lyophilisée, soit à partir d'une solution concentrée. Lors de l'addition de l'ARNt, notamment à faible concentration, nous avons observé un précipité. Même si les conditions de la RMN (tampon et pH) ont été modifiées pour tenter de réduire ce phénomène, il n'a pu être complètement supprimé. Nous avons quand même pu observer deux spectres HSQC du complexe en mélangeant dans l'un 130  $\mu$ M de PTH et 13  $\mu$ M d'ARNt<sup>His</sup>, et dans l'autre 130  $\mu$ M de PTH et 260  $\mu$ M d'ARNt<sup>His</sup>. Cependant, comme on peut le voir sur les figures 38 et 39, l'évolution des pics (notamment ceux appartenant à des massifs) n'a pu être analysée comme dans le cas de la diacétyl-Lys-(3'NH)-adénosine à cause de la diminution générale de l'intensité des pics due à l'augmentation de la taille de l'objet étudié (23 kDa pour la protéine seule et 48 kDa pour le complexe entre la protéine et l'ARNt) et aussi, certainement, à la précipitation partielle de la protéine.

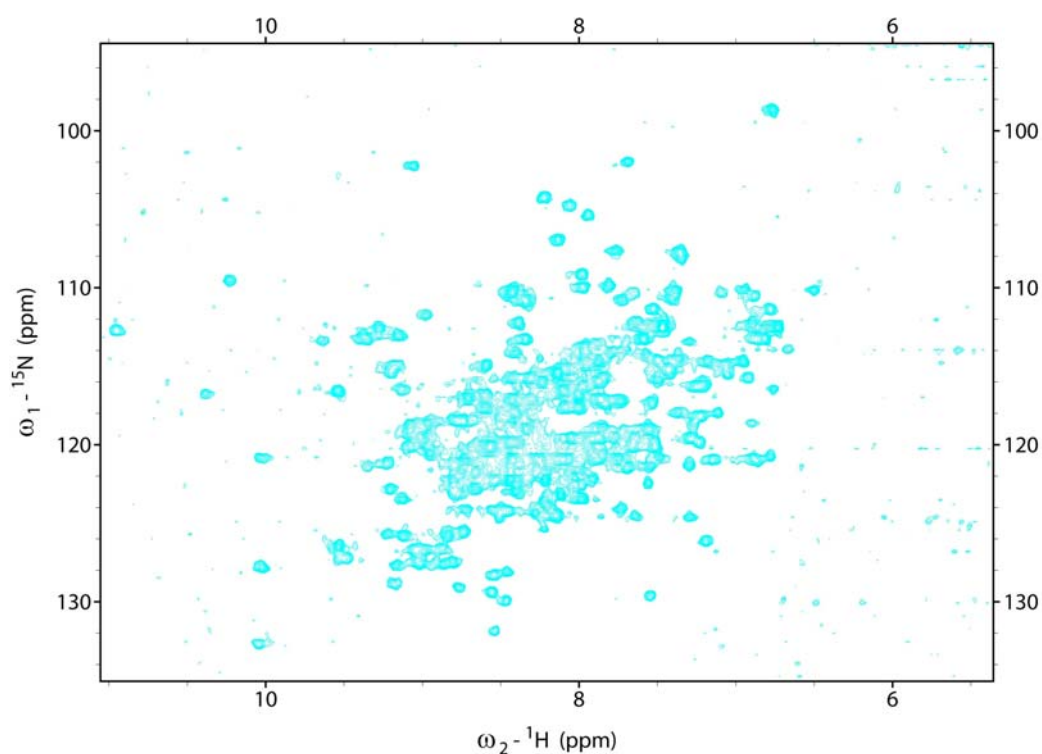


Figure 38 : Spectre HSQC d'un échantillon contenant 130  $\mu\text{M}$  de PTH et 13  $\mu\text{M}$  d'ARNt<sup>His</sup> d'*E. coli*

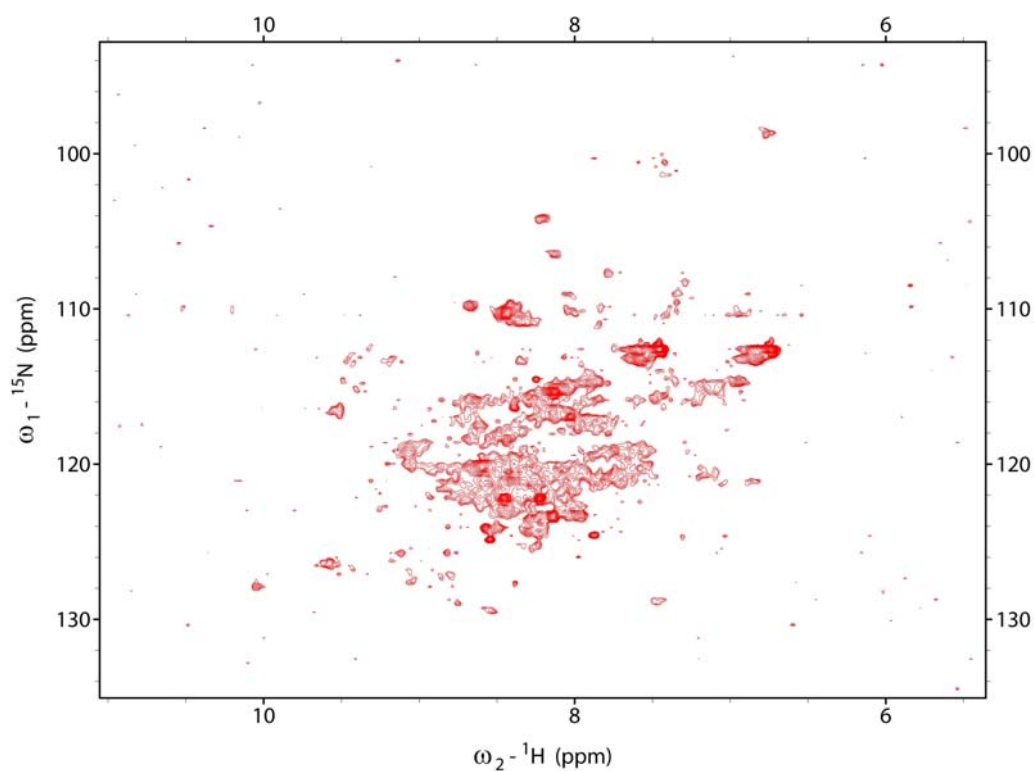
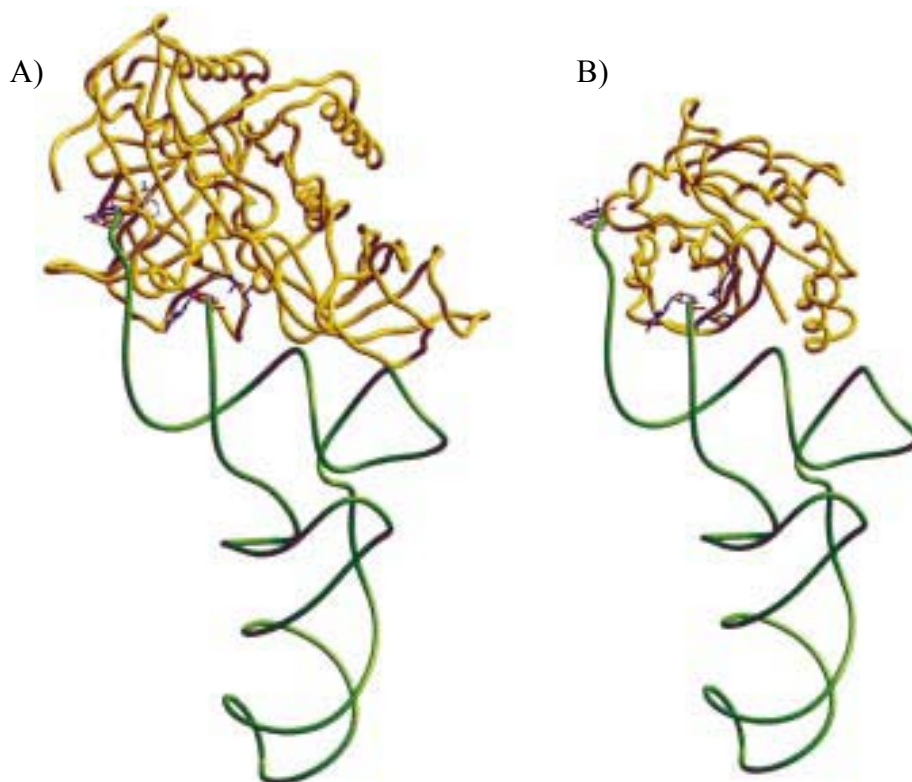


Figure 39 : Spectre HSQC d'un échantillon contenant 130  $\mu\text{M}$  de PTH et 260  $\mu\text{M}$  d'ARNt<sup>His</sup> d'*E. coli*

### 3.2 Etude de mini-substrats

Par analogie avec le modèle de fixation de l'ARNt sur EF-Tu (voir figure 40), un modèle d'interaction entre la PTH et un ARNt avait été proposé (Fromant *et al.*, 1999). Dans ce modèle, la PTH interagit uniquement avec la tige acceptrice et le bras TΨC de l'ARNt. Cette hypothèse nous a incités à travailler avec des ARN plus petits qu'un ARNt entier.



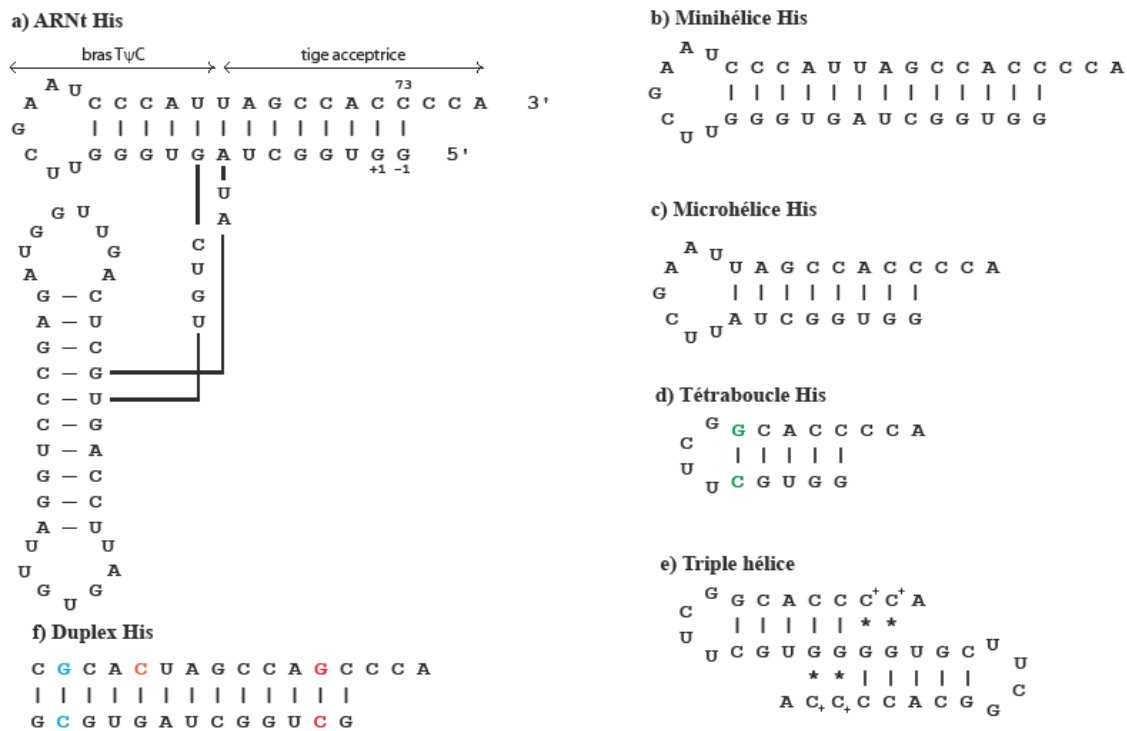
**Figure 40 : Comparaison du complexe EF-Tu:GTP:Phe-ARNt<sup>Phe</sup> (A) et du modèle proposé pour la PTH (B)**

D'après Fromant *et al.* (1999). L'ARNt est représenté en vert et les deux bases 3' et 5' terminales sont représentées en rouge. Les protéines sont représentées en jaune. Le complexe EF-Tu:GTP:Phe-ARNt<sup>Phe</sup> (A) a été obtenu par cristallographie, et le modèle entre la PTH et le Phe-ARNt<sup>Phe</sup> (B) a été construit en supposant que les extrémités 3' et 5' de l'ARNt se fixent de façon analogue sur la PTH et sur EF-Tu.

Comme on l'a mentionné ci-dessus, l'action de la PTH sur de petits ARN dérivés de l'ARNt<sup>His</sup> avait déjà été évaluée. Les résultats obtenus dans cette étude, brièvement présentés ci-dessous, confirment que des mini-ARNt peuvent interagir efficacement avec la PTH.

### 3.2.1 Synthèse des mini-substrats

Trois molécules d'ARN, présentées sur la figure 41b, c et d ) ont été synthétisées in vitro par transcription T7 (Milligan and Uhlenbeck, 1989). La minihélice<sup>His</sup> est basée sur la tige acceptrice et la tige-boucle TΨC de l'ARNt<sup>His</sup> d'*E. coli*. Elle comprend donc 13 paires de bases, l'extrémité CCA et une boucle de 7 nucléotides. La microhélice<sup>His</sup> correspond à la tige acceptrice et à la boucle TΨC de l'ARNt<sup>His</sup>. Elle comprend 8 paires de bases, l'extrémité CCA et la même boucle de 7 nucléotides que la minihélice<sup>His</sup>. La troisième molécule, la tétraboucle<sup>His</sup>, comprend l'extrémité CCA, les 4 dernières paires de bases de la tige acceptrice de l'ARNt<sup>His</sup>, et une paire de bases supplémentaire (G<sub>-1</sub>-C<sub>73</sub>, en vert sur la figure 41d) permettant de stabiliser une boucle de 4 nucléotides. Les 4 nucléotides de cette boucle ont été choisis spécialement pour que cette structure en tige-boucle soit thermodynamiquement stable malgré sa petite taille (Martinis and Schimmel, 1992 ; Shi *et al.*, 1992).



**Figure 41 : Représentation des différents ARN étudiés**

Les différents ARN étudiés dans ce chapitre sont représentés ici en 2 dimensions afin de visualiser leur composition nucléotidique. Les bases appariées sont reliées par des traits. a) L'ARNt<sup>His</sup> représenté est celui d'*E. coli*. b) La minihélice<sup>His</sup> est composée des nucléotides de la tige acceptrice et du bras TΨC de l'ARNt<sup>His</sup> complet. c) La microhélice<sup>His</sup> est composée des nucléotides de la tige acceptrice et de la boucle du bras TΨC de l'ARNt<sup>His</sup>. d) La tétraboucle<sup>His</sup> est composée de l'extrémité de la tige acceptrice de l'ARNt<sup>His</sup> complétée par une paire de bases supplémentaire (en vert) et d'une boucle de 4 nucléotides. e) la conformation en triple hélice (voir 3.2.3) de la tétraboucle est représentée avec des astérisques (\*) représentant des appariements de Hoogsten. f) Le duplex<sup>His</sup> (voir 3.3) est composé des nucléotides de la tige acceptrice et de la tige du bras TΨC de l'ARNt<sup>His</sup> à l'exception des paires de bases G-C en bleu et en rouge remplaçant deux paires de bases C-G, et d'une cytosine à la position 65 (en orange) remplaçant une uracile.

Possédant tous les trois la paire de bases supplémentaire de la tige acceptrice, ces ARN sont aminoacylables par l'HisRS d'*E. coli*. Cependant, l'efficacité catalytique ( $k_{cat}/K_m$ ) est réduite de 2 à 3 ordres de grandeur dans le cas de la mini- ou de la micro-hélice et de 5 ordres de grandeur dans le cas de la tétraboucle, par rapport à celle que l'on obtient avec l'ARNt<sup>His</sup> entier (Francklyn and Schimmel, 1990 ; Shi *et al.*, 1992).

### 3.2.2 Etude catalytique des mini-substrats

Ces ARN ont été aminoacylés et acétylés. Puis, leur hydrolyse par la PTH a été étudiée (voir tableau 8).

Substrat	nb de bases	$k_{cat}/K_m$ ( $\mu\text{M}^{-1}\cdot\text{s}^{-1}$ )	$k_{cat}/K_m$ relatif
acétyl-[ <sup>14</sup> C]His-ARNt <sup>His</sup>	76	0,1	100
acétyl-[ <sup>14</sup> C]His-minihélice <sup>His</sup>	36	0,05	50
acétyl-[ <sup>14</sup> C]His-microhélice <sup>His</sup>	26	0,0058	5.8
acétyl-[ <sup>14</sup> C]His-tétraboucle <sup>His</sup>	17	0,0054	5.4
acétyl-[ <sup>14</sup> C]His-duplex <sup>His</sup>	29	0,015	15

**Tableau 8 : Efficacités catalytiques de la PTH d'*E. coli* pour différents mini-substrats**

Pour chaque substrat, les efficacités catalytiques ont été mesurées à 28°C en utilisant un mélange réactionnel contenant 20 mM de Tris-HCl (pH 7,5), 10 mM de MgCl<sub>2</sub>, 0,15  $\mu\text{M}$  de substrat et 2 à 60 nM de PTH. Les  $k_{cat}/K_m$  relatifs ont été calculés en donnant une valeur de 100 à la valeur obtenue pour l'ARNt<sup>His</sup>.

L'efficacité catalytique obtenue avec la minihélice<sup>His</sup> est deux fois plus faible que celle que l'on obtient avec l'ARNt<sup>His</sup>. Cette faible différence suggère que la formation du complexe entre la PTH et l'ARNt<sup>His</sup> n'implique pas d'autres régions que la tige acceptrice et le bras T $\Psi$ C. La perte du bras T $\Psi$ C, dans le cas de la minihélice et de la tétraboucle, entraîne une diminution de l'efficacité catalytique de la PTH d'un facteur 10 par rapport à la minihélice. Ce résultat montre que l'interaction avec la PTH peut avoir lieu en l'absence du bras T $\Psi$ C. Cependant, la présence de ce bras améliore la catalyse et on peut donc penser qu'il interagit avec la PTH.

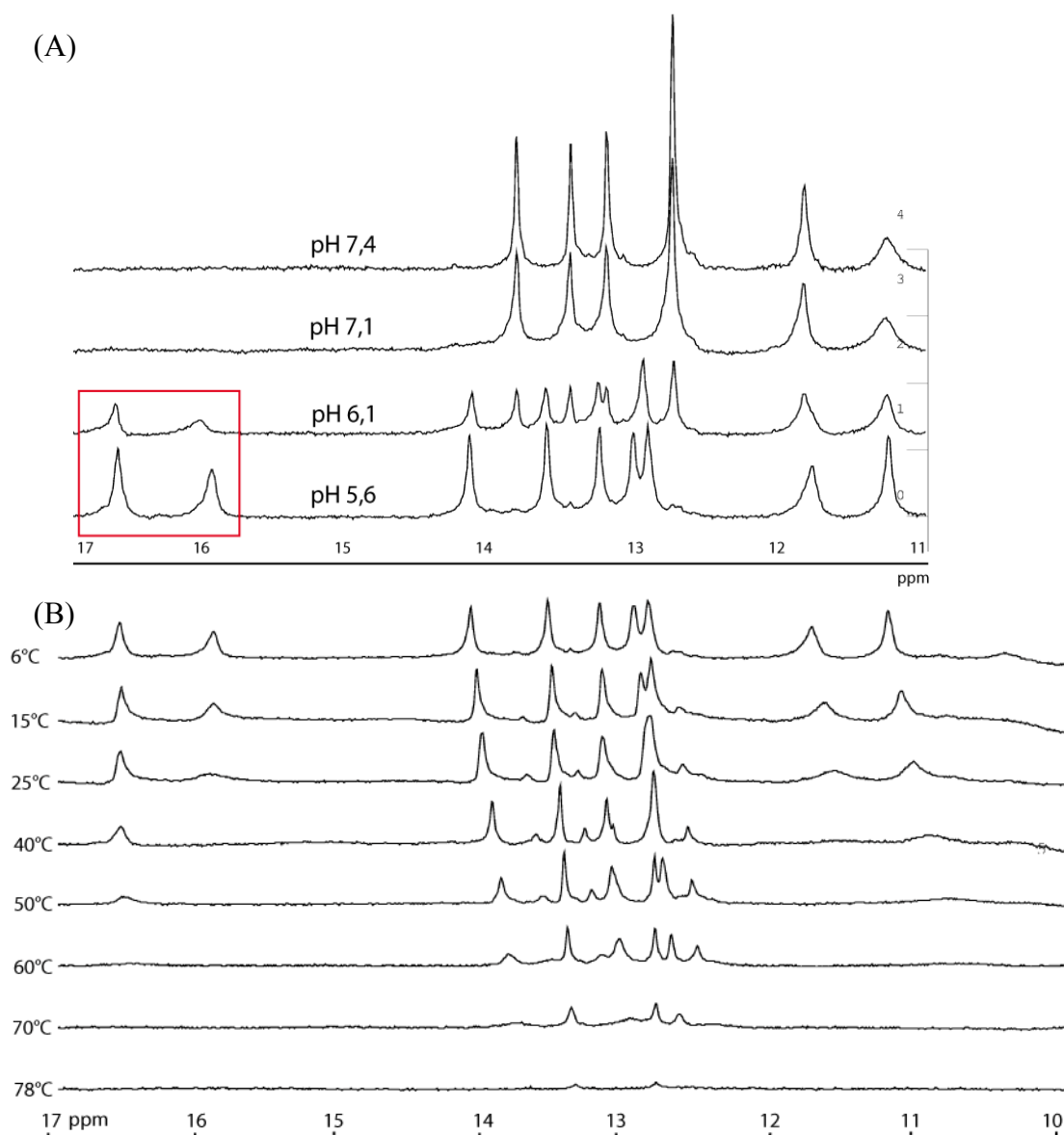
### 3.2.3 Recherche d'un ligand pour l'étude RMN

D'après ces résultats, la tétraboucle apparaissait comme un bon candidat pour étudier par RMN la formation du complexe entre un ARN et la PTH. Elle est reconnue par la PTH, et sa petite taille devait permettre d'augmenter la qualité des spectres de RMN du complexe PTH-ARN.

Avec l'aide de Jean-Louis Leroy, la tétraboucle a été synthétisée chimiquement à l'ICSN (Gif-sur-Yvette), afin d'en obtenir une quantité suffisante

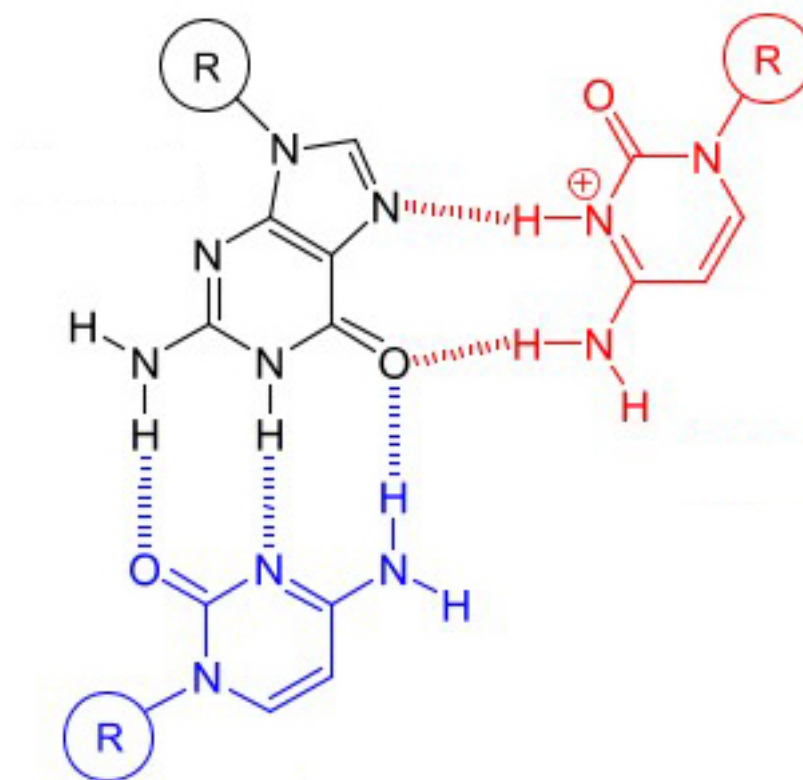


pour une étude par RMN. Des spectres de RMN  $^1\text{H}$  de la tétraboucle (500  $\mu\text{M}$ ) ont ensuite été enregistrés à différents pH (5,6 à 7,4) et à différentes températures (entre 6 et 70°C). Cette étude a révélé qu'à pH acide, la tétraboucle adoptait, au moins partiellement, une conformation sous forme de triple hélice. Comme le montre la figure 42, nous avons en effet observé des déplacements chimiques de protons imino caractéristiques de cytosines protonées, suggérant la formation d'appariements Hoogsteen ( $\text{G-C}^+$ ) (voir figure 43). L'intensité de ces pics diminue quand le pH ou la température augmente.



**Figure 42 : Spectres de RMN 1D de la tétraboucle<sup>His</sup>**

(A) Spectres RMN 1D de la tétraboucle à 0,5 mM et 6°C, à différents pH. Les pics correspondant aux protons de l'appariement Hoogsteen, encadrés en rouge, sont visibles uniquement à pH acide.  
 (B) Spectres RMN 1D de la tétraboucle à 0,5 mM et pH 5,6, à différentes températures. L'intensité des pics correspondant aux protons de l'appariement Hoogsteen diminue quand la température augmente.



**Figure 43 : Appariement Hoogsteen G-C+**

Ce schéma montre qu'une guanine (en noir) peut établir simultanément une interaction de type Watson-Crick avec une cytosine (en bleu) et une interaction de type Hoogsteen avec une autre cytosine (en rouge). Ceci permet la formation d'une hélice à 3 brins. Les traits en pointillés montrent les liaisons hydrogène de l'appariement Watson-Crick (en bleu) et de l'appariement Hoogsteen (en rouge).

La disparition de ces signaux quand le pH augmente, est un résultat cohérent avec le fait que ce type d'appariement nécessite la protonation de cytosines favorisée à pH acide. La figure 41e montre un exemple de triple hélice que la tétraboucle pourrait former à pH acide. Il est à noter que la formation de triple hélice ne se fait pas à pH 7,5. Les résultats sur l'hydrolyse par la PTH de mini-ARNt aminoacylés et acétylés, réalisés à ce pH (7,5) n'ont donc pas été perturbés par la formation d'une telle structure.

Etant données les séquences similaires des extrémités acceptrices de la tétraboucle<sup>His</sup>, des mini- et microhélice<sup>His</sup> et de l'ARNt<sup>His</sup> entier, nous nous sommes demandé si l'ARNt<sup>His</sup> ne pouvait pas, lui aussi, se dimériser à pH 6,0 grâce à des appariements en triple hélice. Nous avons donc enregistré des spectres RMN d'ARNt<sup>His</sup> à différents pH (entre 5,5 et 7,5), mais nous n'avons jamais observé de raies de résonance pouvant correspondre à des protons imino de cytosines protonées.

La différence entre la tétraboucle<sup>His</sup> et l'ARNt<sup>His</sup> vient peut-être du fait que, synthétisée chimiquement, elle n'était pas phosphorylée en 5', ce qui pourrait

favoriser la formation de cette triple hélice. Nous n'avons cependant pas poursuivi la caractérisation de cette différence et avons préféré utiliser, pour les études d'interaction avec la PTH, un ARN dont la séquence permettait d'éviter la formation de triples hélices.

### 3.3 Duplex<sup>His</sup>

#### 3.3.1 Synthèse et aminoacylation du duplex<sup>His</sup>

Nous avons donc choisi un duplex formé par deux oligonucléotides en partie complémentaires dont les séquences sont basées sur celle de la tige acceptrice et de la tige TΨC de l'ARNt<sup>His</sup>, mais avec quelques modifications (voir figure 41f).

Pour améliorer la stabilité du duplex, la base U<sub>65</sub> (numérotation de l'ARNt<sup>His</sup> complet) a été changée en C pour renforcer l'appariement avec G8 (en orange sur la figure 41f). Pour éviter une possible hybridation de l'extrémité 3' d'un oligonucléotide avec l'extrémité 3' de l'autre oligonucléotide, nous avons aussi changé la paire de bases G<sub>11</sub>-C<sub>62</sub> en C<sub>11</sub>-G<sub>62</sub> (en bleu sur la figure 41f). Enfin, pour éviter la formation d'une triple hélice, nous avons changé la paire de bases G<sub>1</sub>-C<sub>72</sub> en C<sub>1</sub>-G<sub>72</sub>. Cette paire de base G<sub>1</sub>-C<sub>72</sub> est présente dans la plupart des ARNt d'*E. coli* (Sprinzl *et al.*, 1989), mais des études sur des variants de l'ARNt<sup>Met</sup> ont montré qu'avec une paire de bases C<sub>1</sub>-G<sub>72</sub>, l'ARNt<sup>Met</sup> méthionylé et acétylé était aussi bien reconnu par la PTH que l'ARNt possédant une paire G<sub>1</sub>-C<sub>72</sub>, et environ 20 fois mieux que l'ARNt<sup>Met</sup> sauvage (Dutka *et al.*, 1993). Une telle paire de bases est donc compatible avec une bonne reconnaissance par la PTH.

Le duplex ne comporte pas de phosphate à son extrémité 5' mais, possédant la base supplémentaire caractéristique des ARNt<sup>His</sup>, il dispose bien d'un phosphate à la position +1, indispensable à la reconnaissance par la PTH. Les deux brins du duplex ont été synthétisés chimiquement avec l'aide de Sandrine Caputo, à l'ICSN de Gif-sur-Yvette, puis assemblés. Comme attendu, le spectre RMN <sup>1</sup>H du duplex, ne présente pas de raie spécifique d'une cytosine protonée et confirme l'appariement des deux brins d'ARN.

Pour vérifier que le duplex était bien reconnu par la PTH nous avons ensuite préparé de l'acétyl-His-duplex<sup>His</sup> afin de savoir si ce substrat pouvait être hydrolysé par cette enzyme. L'acétylation du duplex a été réalisée enzymatiquement en utilisant l'HisRS d'*E. coli*. Pour cela, le duplex a d'abord été phosphorylé en 5' puisque ce phosphate est un élément de reconnaissance du substrat par cette enzyme (Fromant *et al.*, 2000). Il a ensuite été aminoacylé dans un mélange réactionnel contenant 20 mM de Tris-HCl (pH 7,5) et des concentrations variables d'ATP, d'histidine radioactive, de duplex et d'HisRS. Nous avons fait varier ces paramètres ainsi que le temps de l'incubation afin d'acétyler le plus de duplex possible. Nous avons en effet déterminé que l'efficacité catalytique de cette réaction était 800 fois moins bonne que celle obtenue avec l'ARNt<sup>His</sup> et l'hydrolyse spontanée du produit de cette réaction limitait le plateau de la réaction d'acétylation. Nous avons ensuite acétylé l'His-duplex<sup>His</sup> obtenu et l'avons purifié. Nous avons déterminé que la quantité de duplex<sup>His</sup> finalement aminoacylée, acétylée et purifiée était supérieure à 10% de la quantité de duplex mise en jeu dans cette expérience.

### 3.3.2 Etude catalytique de l'interaction du duplex<sup>His</sup> avec la PTH

L'acétyl-His-duplex<sup>His</sup> a ensuite été exposé à la PTH dans des conditions standard pour déterminer l'efficacité catalytique de son hydrolyse. Il est à noter que cette mesure n'a pas été perturbée par une éventuelle inhibition de la part du duplex non aminoacylé présent dans la préparation d'acétyl-His-duplex<sup>His</sup>. En effet, la concentration en substrat (0,15  $\mu\text{M}$ ) et donc en duplex non aminoacylé (moins de 1,5  $\mu\text{M}$ ) est largement inférieure au  $K_i$  mesuré pour le duplex dans les mêmes conditions (20  $\mu\text{M}$ ), ce qui garantit que l'inhibition est totalement négligeable.

L'efficacité catalytique obtenue pour l'hydrolyse de l'acétyl-His-duplex<sup>His</sup> par la PTH ( $0,015 \mu\text{M}^{-1} \cdot \text{s}^{-1}$ ) est plus faible que pour la minihélice mais plus élevée que pour la microhélice (voir tableau 8). Ce résultat est cohérent avec le fait que le duplex mime la tige acceptrice mais aussi la tige T $\Psi$ C alors que la microhélice ne mime que la tige acceptrice. La différence entre le duplex<sup>His</sup> et la minihélice<sup>His</sup> suggère aussi un éventuel rôle joué par la boucle T $\Psi$ C qui pourrait stabiliser l'extrémité double brin de l'hélice, mais cette différence pourrait aussi provenir des modifications apportées à la séquence du duplex<sup>His</sup>.

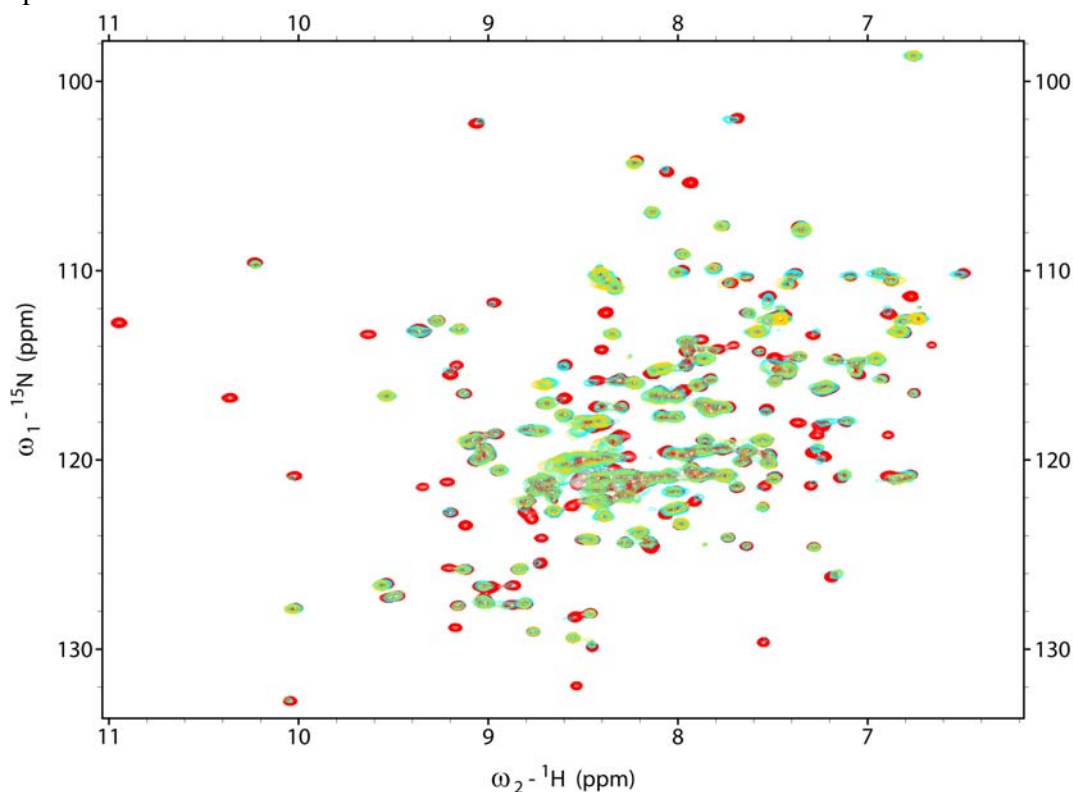
Dans le but d'étudier son interaction avec la PTH par RMN, nous avons aussi voulu savoir si le duplex<sup>His</sup> non aminoacylé interagissait avec la PTH dans les conditions de la RMN. Pour cela, nous avons déterminé s'il inhibait la réaction catalysée par la PTH. Ces expériences ont été réalisées dans les conditions de la RMN, c'est-à-dire à pH 6,0 en présence de 50 mM d'acétate de sodium et de 200 mM de NaCl en utilisant des concentrations de duplex allant jusqu'à 130  $\mu\text{M}$ . Nous avons montré que la présence de duplex<sup>His</sup> inhibait l'hydrolyse du diacétyl-Lys-ARNt<sup>Lys</sup> par la PTH et que cette inhibition était compétitive. La constante d'inhibition, qui reflète la valeur de la constante de dissociation  $K_d$  du complexe PTH-duplex<sup>His</sup>, est de 45  $\mu\text{M}$ . Elle est assez proche de la constante d'inhibition de l'ARNt<sup>His</sup> non aminoacylé, qui est, dans les mêmes conditions, de 15  $\mu\text{M}$ . Le duplex<sup>His</sup> apparaît donc comme un bon candidat pour étudier l'interaction de la PTH avec un ARN, d'autant plus qu'il nous sera possible de travailler à des concentrations en duplex<sup>His</sup> de l'ordre de cette valeur de  $K_d$ .

### 3.4 Empreinte du duplex<sup>His</sup> sur la protéine

Comme dans le cas de la diacétyl-Lys-(3'NH)-adénosine, nous avons utilisé la RMN pour identifier les résidus de la PTH impliqués dans la reconnaissance du duplex<sup>His</sup> décrit précédemment.

#### 3.4.1 Enregistrement des données

Nous avons enregistré plusieurs spectres HSQC de la PTH (130  $\mu\text{M}$ ), en présence de concentrations croissantes de duplex<sup>His</sup> (de 0 à 130  $\mu\text{M}$ ), obtenues par ajouts successifs de duplex concentré. Nous avons remarqué que, si ce mélange était effectué à 4°C, il avait tendance à se troubler, notamment à faible concentration d'ARN, vraisemblablement à cause d'une précipitation partielle de la PTH, comme dans le cas de l'ajout d'ARNt<sup>His</sup>. Nous avons donc toujours réalisé ces mélanges à température ambiante, avant de refroidir les échantillons jusqu'à 17°C.



**Figure 44 : Spectres  $^1\text{H}$ - $^{15}\text{N}$ -HSQC de la PTH en présence de différentes concentrations de duplex<sup>His</sup>**

Spectres  $^1\text{H}$ - $^{15}\text{N}$ -HSQC obtenus en présence de 130  $\mu\text{M}$  de PTH et de 0  $\mu\text{M}$  (rouge), 52  $\mu\text{M}$  (bleu) ou 104  $\mu\text{M}$  (jaune) de duplex<sup>His</sup> (17°C, acétate de sodium 50 mM, pH 6,0, NaCl 200 mM).

Si certains pics semblent disparaître ou se déplacer lors de l'addition de l'ARN, l'invariance de la position de la majorité des pics sur les différents spectres HSQC (figure 44) montre que l'ensemble des spectres a été enregistré dans des conditions reproductibles.

Cependant, même si ce n'est pas visible sur la figure 44 à cause d'une normalisation du signal, l'intensité de la majorité des pics diminue quand la quantité de duplex ajouté augmente. Sur le spectre correspondant à un rapport stœchiométrique entre le duplex et la PTH, l'intensité de la plupart des pics est 2 à 3 fois plus faible que sur le spectre de référence enregistré en l'absence de PTH. Une telle diminution systématique de l'intensité des pics est vraisemblablement causée par l'augmentation de la taille du système étudié (22 kDa pour la protéine libre, 32 kDa pour le complexe duplex et la protéine). On ne peut cependant pas écarter la possibilité qu'une partie des protéines se soient agrégées au cours de l'addition d'ARN, ce qui pourrait avoir diminué la concentration des protéines contribuant au signal de RMN.

Pour vérifier que l'ARN ne se dégradait pas pendant les expériences de RMN, un contrôle a été effectué en incubant plusieurs jours de l'ARNt<sup>His</sup> seul ou en présence de PTH H20A, dans les conditions de la RMN, à 17°C et à 4°C. La capacité de l'ARNt<sup>His</sup> à être aminoacylé par l'HisRS a ensuite été mesurée. Elle était identique pour l'ARNt<sup>His</sup> non incubé, pour l'ARNt<sup>His</sup> incubé seul ou pour l'ARNt<sup>His</sup> incubé en présence de la PTH H20A, quelle que soit la température, ce qui garantit que notre préparation ne comportait pas de nucléase dégradant les ARN.

### 3.4.2 Analyse des spectres

Malgré la complexité des spectres dans certaines régions, nous avons pu observer le déplacement et la variation d'intensité de la plupart des pics. Certains pics se déplacent de plus en plus au cours de l'addition d'ARN, d'autres ne se déplacent pas mais voient leur intensité diminuer jusqu'à leur disparition. D'autres, les plus nombreux, semblent invariants (par rapport à l'intensité moyenne de chaque spectre). Il semble donc bien y avoir une interaction spécifique entre le duplex et la PTH. Comme certains pics se déplacent et que d'autres s'élargissent sans se déplacer, nous avons considéré que nous étions dans un type d'échange rapide à intermédiaire et qu'il serait donc judicieux d'analyser à la fois le  $\Delta$  et l'IPI (voir le chapitre 2 pour la définition de ces termes).

L'analyse simultanée des spectres à différents ratios a permis de suivre l'évolution de certains pics, ce qui aurait été plus difficile si nous avions comparé chaque spectre uniquement avec le spectre de référence. Par exemple, un pic semblant avoir disparu au ratio 1:1 peut ne plus être visible parce qu'il s'est déplacé vers un autre et a fusionné avec lui. Par contre, son déplacement peut être repéré sur les autres spectres.

Sans détailler l'ensemble de ce travail, il faut cependant noter que cette analyse peut être en partie biaisée quand deux pics sont très proches l'un de l'autre. Par exemple, si le signal d'un des deux pics diminue, il est difficile de déterminer précisément ce qui se passe au niveau du deuxième pic à cause de la contribution du premier pic à l'intensité du deuxième. Réciproquement, l'évolution de l'intensité du premier pic est difficile à évaluer précisément à cause de la contribution du deuxième pic.

De même, il est possible qu'un massif évolue peu parce que les différents pics qui le composent se déplacent de manière symétrique. Certains déplacements ont donc pu être manqués. Cependant, si les déplacements avaient été importants, l'analyse des spectres à des concentrations élevées de ligand aurait permis de les détecter. Ce sont donc uniquement de faibles déplacements que nous avons éventuellement manqués.

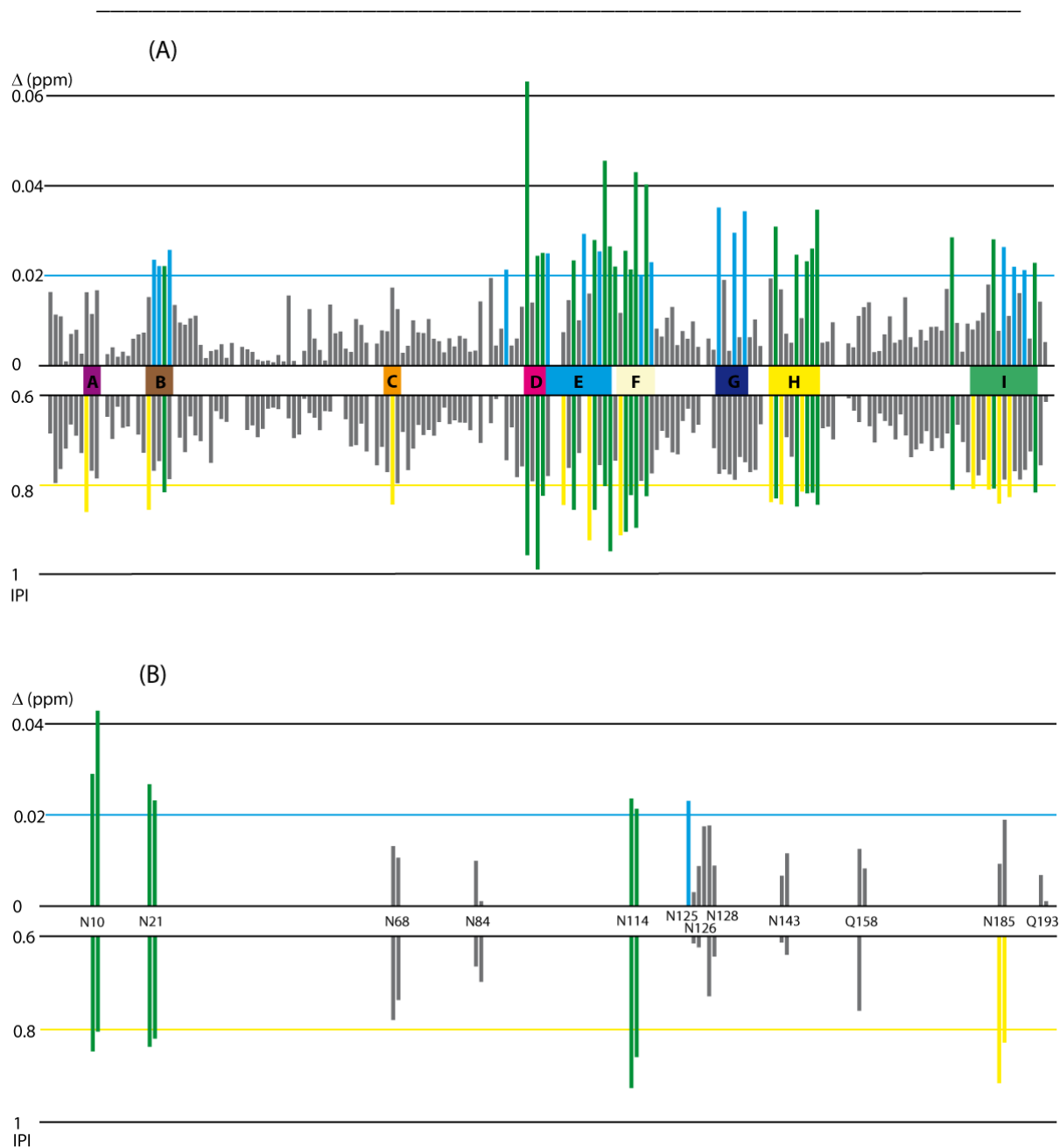
De plus, heureusement, les cas où les pics sont superposés sont assez peu nombreux, si bien que les biais éventuels ne concernent qu'un petit nombre de pics et, généralement, seulement des pics peu affectés.

Nous avons donc pu déterminer le  $\Delta$  et l'IPI de la majorité des pics pour les différents ratios ARN/protéine. Si les déplacements des pics sont plus importants pour de fortes concentrations d'ARN (notamment au ratio ARN/protéine de 0,8 ou de 1), la forte diminution d'intensité de certains pics, accompagnée de certaines superpositions, rendaient plus difficile leur suivi. Pour minimiser les erreurs d'interprétation, nous avons donc choisi d'analyser en détail le spectre correspondant à un ratio ARN/protéine de 0,4, pour lequel aucun pic n'a totalement disparu, et qui correspond aussi à une concentration de duplex voisine de la constante de dissociation du complexe PTH-duplex (45  $\mu$ M, voir 3.3.2). A ces concentrations (52  $\mu$ M pour le duplex et 130  $\mu$ M pour la PTH), si l'on considère que la constante de dissociation ( $K_d$ ) du duplex pour la PTH est de 45  $\mu$ M (identique au  $K_i$ ), le taux de saturation de la protéine par le duplex est d'environ 27%.

### 3.4.3 Résultats obtenus à un ratio ARN/protéine de 0,4

Les  $\Delta$  et les IPI entre le spectre de référence et le spectre correspondant à un ratio ARN/protéine de 0,4 sont présentés ci-dessous sous forme d'histogramme pour les pics correspondant aux couples H–N du squelette peptidique de la protéine (figure 45(A)) et des chaînes latérales des asparagines et glutamines (figure 45(B)).





**Figure 45 : Modifications des pics de résonance de la PTH après ajout de 0,4 équivalent d'ARN**

En fonction de la position de chaque résidu dans la séquence primaire de la protéine, les valeurs des  $\Delta$  et des IPI des pics de résonance des couples  $H^N-N$  (A) et des deux  $H^\delta-N^\delta$  ou  $H^\epsilon-N^\epsilon$  des chaînes latérales des asparagines et glutamines (B), en présence d'un ratio duplex<sup>His</sup>/PTH de 0,4 ont été mesurées sur les spectres  $^1H-^{15}N$ -HSQC présentés à la figure 44 et reportées ici sous la forme d'un histogramme. Les pics dont le  $\Delta$  est supérieur à 0,02 ppm sont indiqués en bleu, ceux dont l'IPI est supérieur à 0,8 sont indiqués en jaune. Ceux pour lesquels à la fois le  $\Delta$  est supérieur à 0,02 ppm et l'IPI est supérieur à 0,8 sont indiqués en vert. Sur l'histogramme du haut (A), les régions les plus affectées, discutées dans la partie 3.5 sont indiquées selon le code couleur de la figure 47 : A, résidus 8 à 10, en violet ; B, résidus 20 à 24, en marron ; C, résidus 66 à 68, en orange ; D, résidus 93 à 96, en rose ; E, résidus 97 à 109, en cyan ; F, résidus 111 à 117, en jaune clair ; G, résidus 130 à 135, en bleu ; H, résidus 140 à 149, en jaune ; I, résidus 179 à 191, en vert. Sur l'histogramme du bas (B), les numéros des asparagines et des glutamines considérées sont indiqués directement.

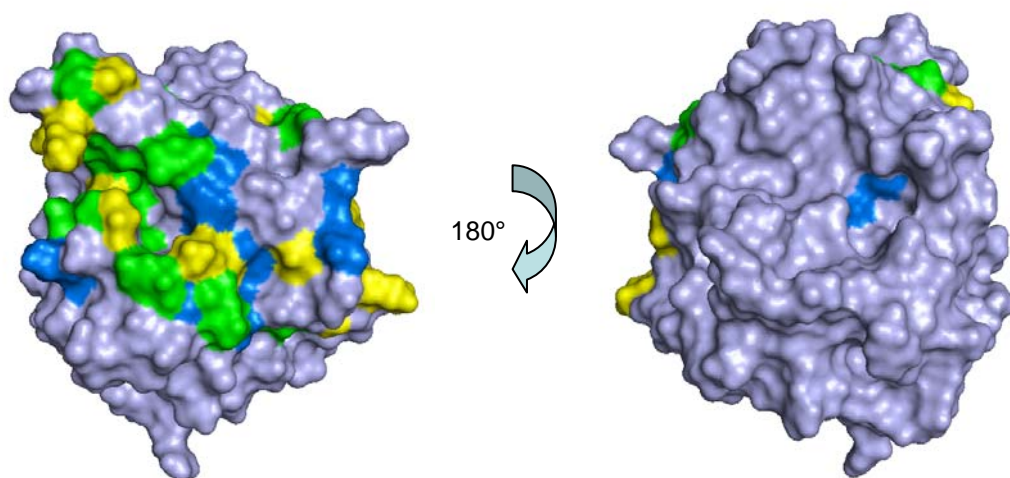
On peut voir que 20 pics de  $H^N-N$  de la chaîne peptidique et 6 pics de  $H^\delta-N^\delta$  appartenant aux chaînes latérales de 3 asparagines (en vert) ont un  $\Delta$  supérieur à 0,02 ppm et un IPI supérieur à 0,8, qui sont les deux valeurs significatives choisies. 15 pics de  $H^N-N$  peptidiques et un pic de  $H^\delta-N^\delta$  d'asparagine (en bleu) ont un  $\Delta$  supérieur à 0,02 ppm et une faible variation d'intensité. Enfin, 13 pics de  $H^N-N$  et 2 pics  $H^\delta-N^\delta$  de la même asparagine (en jaune) ont une faible variation de déplacement chimique mais un IPI supérieur à 0,8. Les valeurs exactes des  $\Delta$  et des IPI des résidus les plus affectés sont données dans le tableau ci-contre (tableau 9) avec le même code couleur.

**Tableau 9 : Variations des pics de résonance des résidus de la PTH affectés par l'ajout de 0,4 équivalent de duplex<sup>His</sup>**

Pour les résidus les plus affectés par la présence du duplex<sup>His</sup> (à un ratio duplex<sup>His</sup>/PTH de 0,4), les valeurs des  $\Delta$  et des IPI des pics de résonance du couple  $H^N-N$  (chaîne principale) et des deux couples  $H^\delta-N^\delta$  (chaîne latérale) pour les asparagines sont présentés ici, dans l'ordre décroissant, du plus affecté au moins affecté, avec le même code couleur que celui utilisé pour la figure 45. Les variations du pic correspondant au couple  $H^e-N^e$  de R133 (voir 3.5.3) sont aussi présentées avec le même code couleur. Les variations des pics des autres arginines, non attribués, sont plus faibles (en dessous de 0,01 ppm pour le  $\Delta$  et en dessous de 0,75 pour l'IPI).

résidu	$\Delta$ (ppm)	résidu	IPI
$H^N-N$			
D93	0,0631	L95	0,9890
G108	0,0455	D93	0,9572
N114	0,0430	G109	0,9482
L116	0,0403	K105	0,9233
H130	0,0352	G111	0,9120
V149	0,0347	G112	0,9041
G135	0,0343	N114	0,8950
D141	0,0309	L8	0,8595
R133	0,0295	L106	0,8549
F104	0,0293	A102	0,8546
W175	0,0285	A20	0,8545
A183	0,0281	V145	0,8472
L106	0,0280	G100	0,8438
G109	0,0265	V149	0,8434
N185	0,0264	M67	0,8427
F148	0,0260	K142	0,8425
A24	0,0258	T184	0,8408
G112	0,0256	G140	0,8375
G107	0,0254	D141	0,8291
D96	0,0251	R186	0,8261
L97	0,0250	L116	0,8240
V145	0,0247	H110	0,8238
L95	0,0244	D96	0,8232
A102	0,0234	H113	0,8215
N21	0,0233	G147	0,8178
G147	0,0232	F148	0,8163
K117	0,0230	K191	0,8155
K191	0,0229	G23	0,8153
G23	0,0222	V146	0,8137
A22	0,0221	W175	0,8098
H110	0,0221	K182	0,8091
L187	0,0220	G179	0,8072
H113	0,0214	A183	0,8068
L89	0,0214	G108	0,8014
A189	0,0213	N21	0,7968
G115	0,0200	N68	0,7949
D86	0,0195	I2	0,7946
G140	0,0194	E94	0,7902
R131	0,0191	G115	0,7893
K182	0,0181	R133	0,7869
M67	0,0174	H188	0,7866
M174	0,0171	N185	0,7864
K142	0,0170	A24	0,7853
N10	0,0168	N10	0,7842
L8	0,0164	F66	0,7698
$H^\delta-N^\delta$			
N10	0,0430	N114	0,9526
N10	0,0291	N185	0,9165
N21	0,0268	N114	0,8600
N114	0,0236	N10	0,8476
N21	0,0232	N21	0,8374
N125	0,0231	N185	0,8285
N114	0,0214	N21	0,8201
N185	0,0189	N10	0,8048
$H^e-N^e$			
R133	0,0274	R133	0,8024

Nous avons reporté cette liste de résidus affectés sur la structure cristallographique de la PTH, en coloriant leur surface avec le même code couleur (figure 46). Comme on peut le voir, les résidus affectés forment une zone continue sur une face de la protéine. L'interaction entre la PTH et le duplex semble donc être spécifique. De plus, aucune région de cette zone ne possède de résidus affectés uniquement via leur  $\Delta$  ou leur IPI. Il ne semble donc pas y avoir une zone où tous les atomes sont en échange rapide et une autre où tous les atomes sont en échange lent, ce qui aurait pu révéler une différence de fixation suivant les zones d'interaction.



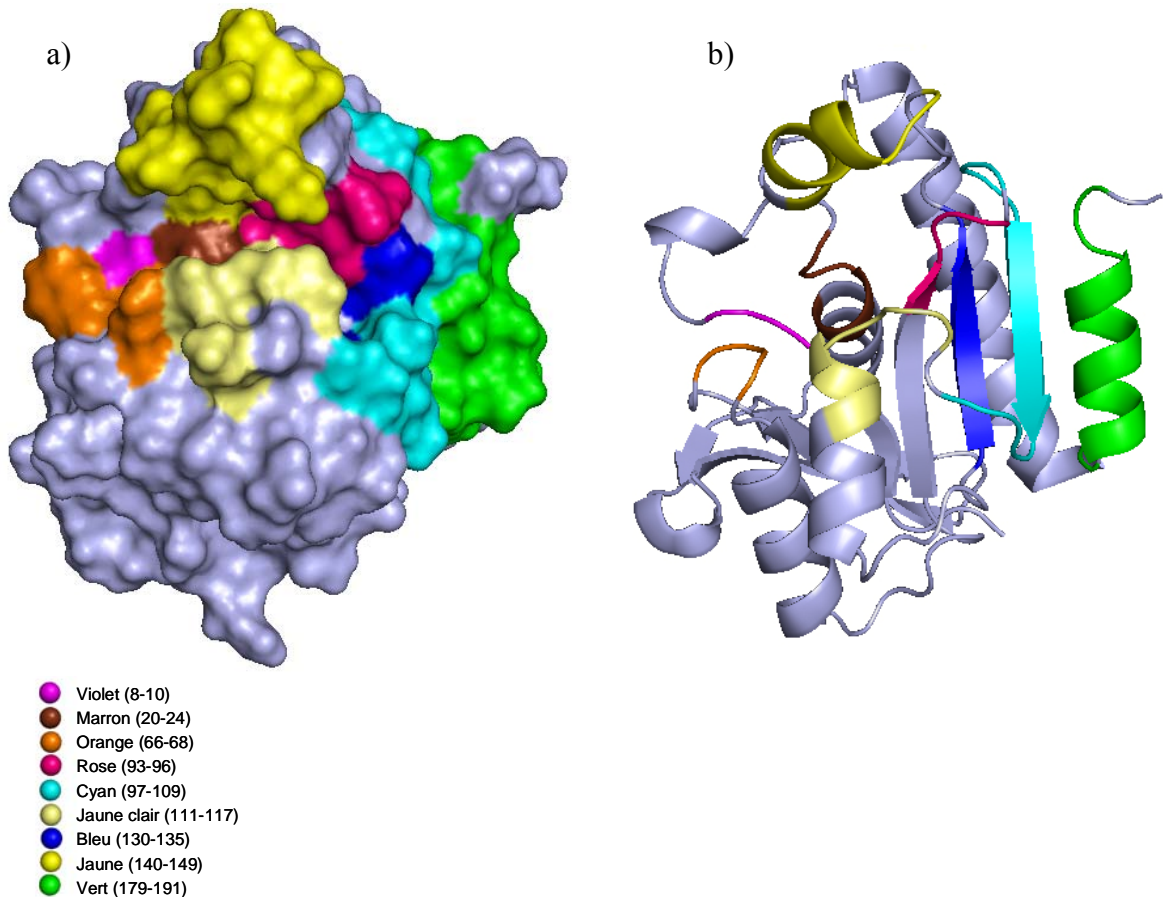
**Figure 46 : Représentation des résidus de la PTH affectés par la fixation du duplex<sup>His</sup>**

La surface de la PTH d'*E. coli* est représentée à partir du fichier PDB 2PTH, grâce au logiciel Pymol, selon deux orientations différant de 180°. Les résidus dont les pics H<sup>N</sup>-N sont affectés par la fixation du duplex<sup>His</sup> ont été coloriés selon le même code couleur que celui utilisé pour les histogrammes de la figure 45.

Certains résidus appartenant à la zone de contact ne semblent pas affectés alors que leurs résidus voisins le sont. Il n'est pas obligatoire qu'un changement d'environnement chimique se traduise par un changement de déplacement chimique assez important pour apparaître dans notre sélection (la fréquence de résonance d'un atome dans la forme libre peut être proche de celle dans la forme liée même si l'atome est dans la zone de contact). De plus, comme nous l'avons expliqué ci-dessus, nous avons pu manquer certaines variations d'intensité ou de déplacement chimique dans le cas où plusieurs pics de résonance sont superposés. Rappelons aussi que les prolines n'ont pas pu être étudiées, puisqu'elles ne possèdent pas de proton amide. Il nous a donc paru logique d'analyser les résultats par régions d'interactions dans lesquelles la plupart des résidus sont affectés, mais pas forcément tous.

### 3.5 Analyse des régions affectées par la fixation du duplex<sup>His</sup>

Nous allons maintenant décrire les régions de la PTH qui sont affectées par la fixation du duplex. Elles sont représentées avec différentes couleurs sur la figure 47 ci-dessous.



**Figure 47 : Représentation des régions de la PTH affectées par la fixation du duplex<sup>His</sup>**

La surface (a) et la structure secondaire (b) de la PTH d'*E. coli* sont représentées à partir du fichier PDB 2PTH, grâce au logiciel Pymol. Les régions affectées par la fixation du duplex<sup>His</sup> ont été coloriées selon la légende ci-dessus.

#### 3.5.1 Le site actif

(Résidus 8 à 10, 20 à 24, 66 à 68, 93 à 96 et 111 à 117)

Entre les spectres de la protéine avec ou sans le duplex, de fortes variations de déplacements chimiques ou d'intensité ont été observées pour les pics de N10 (surtout ceux de sa chaîne latérale), de A20 (qui remplace l'histidine présente dans la PTH sauvage), de D93, de N114, et, dans une moindre mesure, de N68. Ces

résidus font partie du site actif. La plupart des résidus dans leur voisinage sont aussi affectés par la fixation du duplex.

Parmi eux, en marron sur la figure 47, les résidus Ala22, Gly23 et Ala24 sont situés à une extrémité de l'hélice  $\alpha 1$  qui est reliée au résidu 20 par l'intermédiaire de l'Asn21, elle aussi affectée, notamment au niveau de sa chaîne latérale. Tous ces résidus n'étaient que peu affectés lors de la fixation de la diacétyl-Lys-(3'NH)-adénosine étudiée au chapitre 2. Or ces résidus sont enfouis à l'intérieur de la protéine et sont donc vraisemblablement éloignés de la surface d'interaction avec l'ARN. On peut même les voir sur la vue de la protéine montrant la face opposée du site actif (figure 46). Cela suggère un mouvement d'une partie de cette hélice en réponse à la fixation de l'ARN. Un tel mouvement pourrait permettre de positionner l'histidine 20 à l'endroit précis où elle peut efficacement jouer le rôle essentiel qu'est le sien dans le mécanisme catalytique, puisque, rappelons-le, ce résidu semble activer la molécule d'eau responsable de l'hydrolyse du substrat (Goodall *et al.*, 2004) et sa mutation en alanine His  $\rightarrow$  Ala réduit de plus de 3 ordres de grandeurs l'efficacité catalytique de la protéine (voir tableau 3).

Les résidus 8 à 10 (en violet sur la figure 47) ainsi que, 66 à 68 (en orange), font partie de ceux qui étaient déjà affectés par la fixation la diacétyl-Lys-(3'NH)-adénosine. Par homologie avec le modèle de l'interaction entre ce ligand et la protéine, on peut penser que F66 interagit avec l'ARN via une interaction hydrophobe entre sa chaîne latérale aromatique et l'adénine 3'-terminale de l'ARN. Comme nous l'avons souligné dans cette étude, les autres résidus de ces régions, notamment N68 et N10 jouent aussi un rôle important dans l'activité catalytique de la PTH et il est donc logique de voir leur environnement affecté par la fixation du duplex même s'ils ne sont pas forcément en contact direct avec celui-ci.

Les résidus D96 et G111 sont considérés par Selvaraj *et al.* comme étant à la limite entre le site de fixation de l'ARN et le site de fixation du peptide, ces deux sites étant reliés par un canal qui s'ouvrirait par un mouvement concerté de ces résidus lors de l'occupation du site actif (Selvaraj *et al.*, 2007). Cette conclusion est basée sur la comparaison entre (i) la structure cristalline de la PTH d'*E. coli* possédant l'extrémité C-terminale d'une molécule voisine dans le site de fixation du peptide et (ii) la structure cristalline de la PTH de *M. tuberculosis*, où ce site est inoccupé. La distance interatomique minimale entre D96 et G111 est de 6,7 Å sur la structure de la PTH d'*E. coli* et de 3,1 Å (permettant un contact) entre D98 et G113, les résidus homologues dans la PTH de *M. tuberculosis*. Il semblerait donc que la protéine puisse « s'ouvrir » dans cette région pour accueillir son ligand. D'après nos résultats sur la fixation de la diacétyl-Lys-(3'NH)-adénosine, ils seraient, en fait, à la limite entre le site actif de la protéine, englobant le site de fixation du peptide et de l'extrémité 3' de l'ARN, et la région de la protéine qui fixe la partie double brin de l'ARN. Il est donc logique que les résidus voisins de D96 (93-96 en rose) et de G111 (112-117 en jaune clair) soient sensibles à la fixation de l'ARN. Une telle ouverture du site actif lors de son

occupation pourrait d'ailleurs être responsable du fait que K117 et G111 sont affectés par la fixation de la diacétyl-Lys-(3'NH)-adénosine sans être en contact direct avec ce ligand.

### 3.5.2 La boucle-hélice au dessus du site actif

(Résidus 140 à 149)

Au-dessus du site actif, en jaune sur la figure 47, une autre région a été identifiée lors de l'interaction avec la diacétyl-Lys-(3'NH)-adénosine, comme jouant un rôle important dans la reconnaissance du substrat. Il s'agit des résidus 140 à 149 dont la lysine 142 fait partie. Cette lysine, dont l'importance a été révélée par mutagenèse dirigée pourrait se lier à l'ARNt via un pont salin entre sa chaîne latérale et un des phosphates de l'ARNt. La région 140-149 est formée d'une hélice accolée à une boucle. Comme nous l'avons déjà mentionné (paragraphe 5 de l'introduction), chez *M. tuberculosis*, cette région semble très flexible (Pulavarti *et al.*, 2008). De plus, sa conformation est différente (plus « ouverte ») dans la structure cristallographique de la PTH de *M. tuberculosis* et dans la structure de la PTH d'*E. coli* où le site actif est en partie occupé. Notre étude de la structure secondaire sur la PTH d'*E. coli* semble aussi montrer des différences dans cette région, entre les structures en solutions et dans les cristaux, notamment dans la taille de l'hélice. Elle pourrait donc adopter une configuration différente entre la forme libre et liée de la PTH, parallèlement à l'interaction entre K142 et l'ARNt. Pour ces raisons, il est logique d'observer que la plupart des résidus de cette région (notamment K142) sont affectés par la fixation du duplex.

### 3.5.3 La région de la "pince" électrostatique formée par les résidus 105 et 133

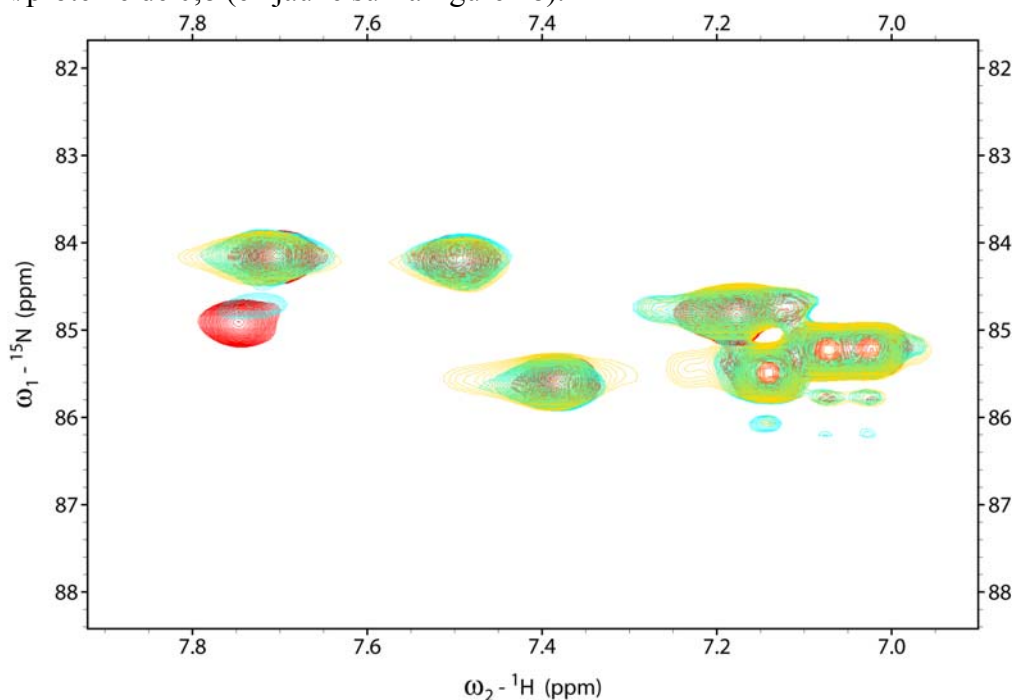
(Résidus 97 à 109 et 130 à 135)

A côté du site actif et de l'autre côté de la « limite » formée par D96 et G111, une région (en bleu et cyan) est fortement affectée par la fixation du duplex, alors qu'elle n'était pas affectée par la fixation de la diacétyl-Lys-(3'NH)-adénosine. Elle est composée de deux brins  $\beta$  ( $\beta_6$  et  $\beta_7$ ) et des boucles autour d'eux. Elle comporte un certain nombre de résidus chargés positivement (K103, K105, R131 et R133), susceptibles d'interagir avec l'ARNt chargé négativement.

Les résidus 97 à 110 (en cyan), comprenant ceux du brin  $\beta_6$ , sont tous affectés et présentent en grande majorité soit un IPI supérieur à 0,8, soit un  $\Delta$  supérieur à 0,02 ppm (à part les prolines 98 et 99 qui n'ont pas de H<sup>N</sup> et dont l'implication n'a pas pu être analysée). Les résidus H130, R133 et G135 de  $\beta_7$ , possèdent un  $\Delta$  largement supérieur à 0,02 ppm (et R131 possède un  $\Delta$  voisin de 0,02 ppm) et tous les résidus de 130 à 135 (en bleu) possèdent un IPI proche de 0,8.

Cette région entoure les résidus K105 et R133 qui, d'après des études de mutagenèse dirigée, formeraient une pince électrostatique interagissant avec le phosphate 5' terminal (+1) des ARNt. Cette interaction proviendrait de la formation de ponts salins entre les charges positives portées par les chaînes latérales de ces résidus et la charge négative de ce groupement phosphate. Le fait que la grande majorité des résidus des éléments structuraux autour de K105 et R133 soient affectés par la fixation du duplex est en accord avec l'idée que ces deux résidus jouent un rôle important dans la reconnaissance de l'ARN.

L'observation des déplacements chimiques des couples  $H^{\zeta}-N^{\zeta}$  des lysines et  $H^{\eta}-N^{\eta}$  des arginines (voir figure 28) est difficile à cause de la grande vitesse d'échange des protons  $H^{\zeta}$  avec l'eau. En revanche, en choisissant une fenêtre de travail adaptée, on peut aisément observer les couples  $H^{\epsilon}-N^{\epsilon}$  des arginines. Pour confirmer l'implication de la chaîne latérale du résidu R133 dans la fixation de l'ARN, nous avons donc observé la région du spectre HSQC correspondant aux couples  $H^{\epsilon}-N^{\epsilon}$  des chaînes latérales des arginines, en présence ou en absence de duplex<sup>His</sup>. Comme on peut le voir sur la figure 48, un pic est fortement affecté par l'addition d'ARN. A un ratio ARN/protéine de 0,4 (en bleu sur la figure 48), son IPI est de 0,8 et son  $\Delta$  est de 0,027. Ce pic devient indétectable à un ratio ARN/protéine de 0,8 (en jaune sur la figure 48).

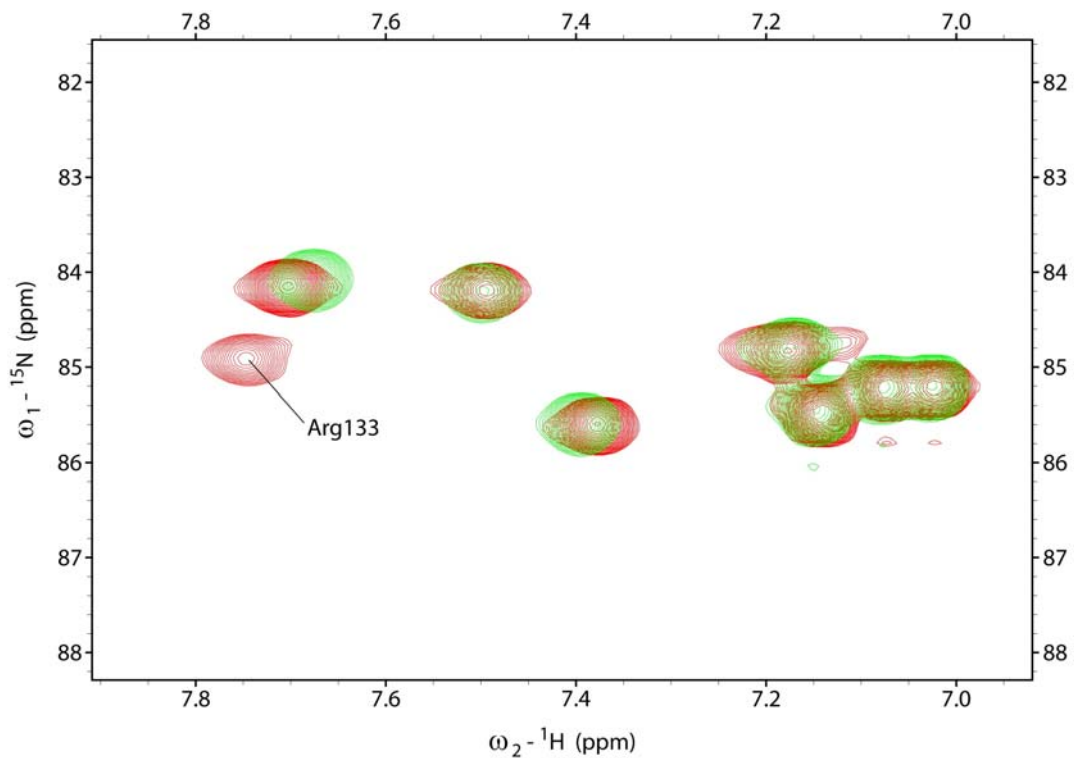


**Figure 48 : Effet de l'addition de duplex<sup>His</sup> sur la région du spectre <sup>15</sup>N-HSQC correspondant aux couples  $H^{\epsilon}-N^{\epsilon}$  des arginines de la PTH**

Les spectres ont été enregistrés en présence de 130  $\mu$ M de PTH et en absence de duplex<sup>His</sup> (en rouge) ou à un ratio duplex<sup>His</sup>/PTH de 0,4 (en bleu) ou de 0,8 (en jaune). Deux pics ayant des fréquences de résonance <sup>1</sup>H plus élevées et un pic ayant une fréquence de résonance <sup>1</sup>H plus faible, n'apparaissent pas sur cette figure. Ces pics sont peu modifiés par l'addition de duplex<sup>His</sup>.

Sur l'expérience NOESY, nous n'avons pas pu observer de NOE nous permettant d'attribuer les pics des chaînes latérales des arginines de la même

manière que ce que nous avons fait pour les asparagines et les glutamines. Pour vérifier que le pic fortement affecté par l'addition de duplex<sup>His</sup> correspondait bien au couple H<sup>ε</sup>-N<sup>ε</sup> de R133, nous avons comparé le spectre HSQC du double mutant H20A-R133A et le spectre de référence qui est celui du simple mutant PTH-H20A. Comme on peut le voir sur la figure 49, le pic qui se déplace lors de la fixation du duplex, est absent sur le spectre du double mutant, ce qui confirme, premièrement que ce pic est bien celui de R133 et, deuxièmement, que la chaîne latérale de R133 est impliquée dans la reconnaissance de l'ARN.



**Figure 49 : Effet de la mutation R133A sur la région du spectre <sup>15</sup>N-HSQC correspondant aux couples H<sup>ε</sup>-N<sup>ε</sup> des arginines de la PTH**

Les pics correspondant au mutant PTH HA20 sont en rouges. Les pics correspondant au double mutant PTH H20A-R133A sont en vert. Le pic qui disparaît, et qui correspond donc à l'Arg 133, est indiqué.

Notre étude confirme donc que la région de R133 et de K105 est engagée dans la reconnaissance de l'ARNt. La diacétyl-Lys-(3'NH)-adénosine mimant uniquement l'extrémité 3'-terminale d'un peptidyl-ARNt, il est logique que la région de la pince K105-R133 ne soit pas affectée par la fixation de ce ligand. Nous avons d'ailleurs vérifié que le pic correspondant au couple H<sup>ε</sup>-N<sup>ε</sup> de R133 ne l'était pas non plus.



### 3.5.4 L'hélice C-terminale

(résidus 175 et 179 à 191)

Enfin, une région située à côté de la précédente, et donc à plus de 30 Å du site actif, est fortement affectée par la fixation du duplex. Cette région comprend l'hélice  $\alpha_6$ , C-terminale. En particulier, les résidus W175, G179, K182, A183, T184, N185, R186 et A189 possèdent un IPI élevé et parfois aussi un  $\Delta$  supérieur à 0,02 ppm. Cette région n'était pas affectée par la fixation de la diacétyl-Lys-(3'NH)-adénosine, ce qui est en accord avec son éloignement du site actif.

D'après des études sur la PTH de *M. tuberculosis*, cette hélice est très mobile (Pulavarti *et al.*, 2008), ce qui lui permettrait de se déplacer au cours de la fixation de l'ARNt. Nos résultats suggèrent que cette région est effectivement en contact avec l'ARN.

### 3.6 Mutagenèse

Comme nous l'avons vu dans l'introduction (paragraphe 5), plusieurs expériences de mutagenèse dirigée ont déjà permis d'identifier un certain nombre de résidus impliqués dans la fixation du substrat et/ou dans la catalyse de la réaction d'hydrolyse du substrat par la PTH. Notre étude par RMN suggère que d'autres régions pourraient jouer un rôle important dans le fonctionnement de cette enzyme : l'hélice C-terminale et la région autour de L95, qui est le résidu dont l'IPI est le plus affecté par la fixation de l'ARN. Nous avons donc entrepris des expériences de mutagenèse dirigée afin de préciser le rôle de ces deux régions. Nous avons aussi modifié un résidu basique très conservé, R131, situé lui aussi dans une région affectée par la fixation de l'ARN. Ainsi, chacun des résidus H92, L95, D96, R131, K182 et N185 a été muté en alanine. Chaque protéine mutante a été purifiée par affinité grâce à son étiquette de poly-histidine et ses constantes de Michaelis ont été déterminées en utilisant le diacétyl-Lys-ARNt<sup>Lys</sup> comme substrat.

Comme on peut le voir sur le tableau 10, plusieurs mutations affectent seulement légèrement le  $k_{cat}$  et/ou le  $K_m$  de la réaction d'hydrolyse catalysée par la PTH. Ainsi, les mutations H92A et N185A diminuent d'un facteur 3 l'efficacité catalytique de l'enzyme. Avec la mutation D96A, les valeurs du  $k_{cat}$  et du  $K_m$  sont toutes les deux environ deux fois plus faibles que celles de la PTH sauvage, ce qui fait que l'efficacité catalytique de l'enzyme mutante est identique à celle de l'enzyme sauvage.

Trois mutations (L95A, R131A, K182A) ont un effet plus prononcé sur l'activité de la PTH (voir tableau 10). Dans le cas de la mutation L95A, la valeur du  $k_{cat}$  est diminuée d'un facteur 25 par rapport à l'enzyme sauvage. La valeur du  $K_m$  du substrat n'est pas modifiée. Ce résultat suggère que ce résidu est impliqué dans la structuration du site actif ou dans le positionnement du substrat dans le site actif. Le caractère hydrophobe de L95 semble d'ailleurs important puisque, si ce résidu n'est conservé que dans 52 % des cas (voir tableau 10), on retrouve, à la position 95, un résidu hydrophobe aliphatique (Leu, Ile, Val ou Met) dans 92% des séquences.

La mutation R131A entraîne une diminution de la valeur du  $k_{cat}/K_m$  de deux ordres de grandeurs. Cela confirme donc l'importance de ce résidu dans l'activité catalytique de la PTH, d'autant plus qu'il est conservé dans 91% des séquences. Cependant, parmi les pics correspondant aux couples H<sup>ε</sup>-N<sup>ε</sup> des chaînes latérales des arginines, il n'y a que celui de R133 qui est fortement affecté par la fixation du duplex<sup>His</sup>. Il n'est donc pas certain que R131 interagisse directement avec le substrat. Son rôle serait donc plutôt structural.

Dans l'hélice C-terminale, la mutation K182A entraîne une diminution de l'activité catalytique d'un facteur 55. En particulier, le  $K_m$  de ce mutant pour le substrat est 8 fois plus faible que celui de la protéine sauvage. Ce résultat

confirme que l'hélice C-terminale de la PTH est impliquée dans la reconnaissance de son substrat et montre que K182, appartenant à cette hélice, est fortement impliquée dans cette reconnaissance. Malgré l'importance de ce résidu pour l'activité catalytique de la PTH d'*E. coli*, celui-ci n'est que faiblement conservé dans les autres PTH (20%). Par contre, il y a presque toujours au moins un résidu basique aux positions 181, 182 ou 186, susceptible d'interagir avec l'ARN chargé négativement. L'interaction entre l'hélice C-terminale et l'ARNt pourrait donc être conservée dans la plupart des PTH, même s'il existe quelques organismes, comme *Saccharomyces cerevisiae*, où cette hélice n'existe pas.

	$k_{cat}$ ( $s^{-1}$ )	$K_m$ ( $\mu M$ )	$k_{cat}/K_m$ relatif	IPI	$\Delta$	conservation
WT	1,06	1,8	100			
HA92	0,65	3,7	30	0,757	0,013	55%
LA95	0,043	1,8	4,0	0,989	0,024	52%
DA96	0,41	0,7	99	0,823	0,025	72%
RA131	0,069	9,4	1,2	0,764	0,019	91%
KA182	0,16	15	1,8	0,809	0,018	20%
NA185	0,60	3,0	34	0,786	0,026	47%

**Tableau 10 : Paramètres catalytiques de différents mutants de la PTH d'*E. coli***

Pour chaque enzyme, les valeurs de  $k_{cat}$  et de  $K_m$ , pour le diacétyl-lysyl-ARNt<sup>Lys</sup> ont été mesurées comme cela est décrit dans la partie Matériel et Méthodes. Pour l'enzyme sauvage, une valeur de 100 a été donnée arbitrairement au  $k_{cat}/K_m$  relatif. Toutes ces expériences ont été réalisées avec des protéines étiquetées avec un peptide de poly-histidine. Pour chaque résidu muté en alanine, le tableau rappelle aussi les valeurs des IPI et des  $\Delta$  déterminés lors de l'expérience de RMN réalisée avec un ratio duplex<sup>His</sup>/PTH de 0,4 (voir 3.4). Le pourcentage de conservation des résidus mutés est celui donné par l'alignement disponible sur le site Pfam (<http://pfam.sanger.ac.uk>), réalisé à partir de 116 séquences de PTH représentatives de l'ensemble des séquences de PTH déterminées à ce jour.

### 3.7 Conclusion

Grâce à la RMN, nous avons donc identifié une zone étendue de la PTH où l'environnement des groupements amides change lorsqu'on ajoute de l'ARN. Le fait que cette région soit d'un seul tenant confirme la spécificité de l'interaction entre l'ARN et la PTH dans les conditions de la RMN. Cette région contient tous les résidus qui, d'après des résultats antérieurs de mutagenèse dirigée, étaient soupçonnés d'interagir avec l'ARN. Par ailleurs, l'importance de la chaîne latérale de R133 dans la reconnaissance de l'ARN a été confirmée. Enfin, les expériences de RMN ont indiqué que l'hélice C-terminale de la PTH jouait un rôle dans l'ancrage de l'ARNt, ce qui a ensuite été confirmé par mutagenèse dirigée. Nous allons par la suite utiliser les résidus ainsi identifiés pour établir un modèle de l'interaction de la PTH avec le duplex<sup>His</sup> et avec un diacétyl-Lys-duplex<sup>His</sup>.

## **Chapitre 4**

# **Modélisation de l'interaction entre la PTH et son substrat**



## 4.1 Modélisation de l'interaction entre la PTH et le duplex<sup>His</sup>

### 4.1.1 Paramètres de la modélisation

Nous avons cherché à construire un modèle du complexe entre la protéine et le duplex. Pour cela, nous nous sommes servis d'un ensemble de scripts écrit pour CNS. Un ensemble de scripts similaires, élaboré par François Bontems, avait déjà été utilisé pour étudier la fixation de la thymosine  $\beta 4$  sur de l'actine (Domanski *et al.*, 2004). Tout comme HADDOCK, (utilisé ci-dessus pour l'étude de l'interaction entre la PTH et la diacétyl-Lys-(3'NH)-adénosine), il permet de faire évoluer un système vers un minimum d'énergie en prenant en compte des contraintes d'interactions, mais il était plus simple et plus rapide à mettre en place qu'HADDOCK pour cette étude de l'interaction entre la PTH et le duplex et, surtout, pour la modélisation que nous allons faire également de l'interaction entre la PTH et un duplex aminoacylé et acétylé.

Nous sommes partis de la structure cristallographique de la PTH sauvage (contenant l'histidine 20) à laquelle nous avons ajouté les atomes d'hydrogène manquants, et de la structure du duplex construite grâce au logiciel XDNA.

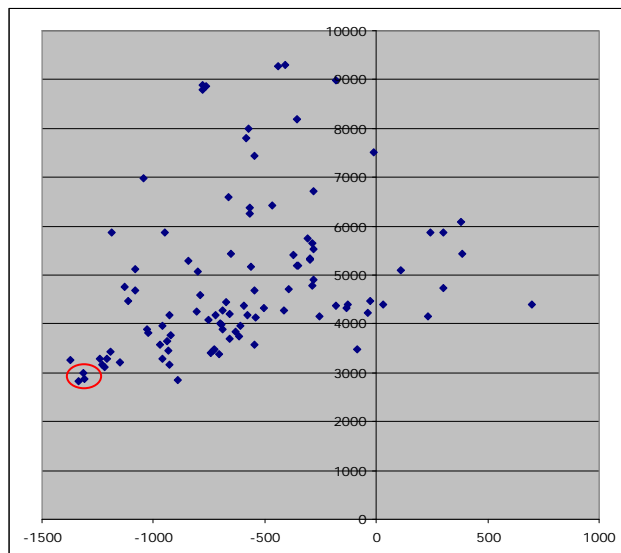
Une liste de "contraintes floues" a été définie à partir d'une liste de résidus de la protéine affectés par la liaison au duplex (chaque contrainte spécifiant qu'un résidu de cette liste doit être au contact du duplex). Les résidus choisis sont ceux qui, sur l'expérience HSQC réalisée avec un rapport ARN/PTH H2O de 0,4, ont un IPI supérieur à 0,8 ou un delta supérieur à 0,02 ppm et qui sont à la surface de la protéine, c'est-à-dire les résidus suivants : 20, 67, 68, 96, 105, 110, 114, 117, 133, 141, 142, 145, 149, 182, 184, 186, 189 et 191. Cette liste a été utilisée pour construire un ensemble de modèles du complexe protéine/duplex en donnant une pénalité de  $10 \text{ kcal.mol}^{-1}.\text{\AA}^{-2}$  à chaque fois que la distance entre un résidu de la liste et l'ARN était supérieure à 4 Å.

### 4.1.2 Modélisation sans orientation préalable

Dans un premier temps, le duplex était éloigné de la protéine en face de la région affectée, à une distance variable comprise entre 10 et 30 Å et avec une orientation variable. 100 positions de départ aléatoires ont permis d'obtenir 100 modèles de complexe.

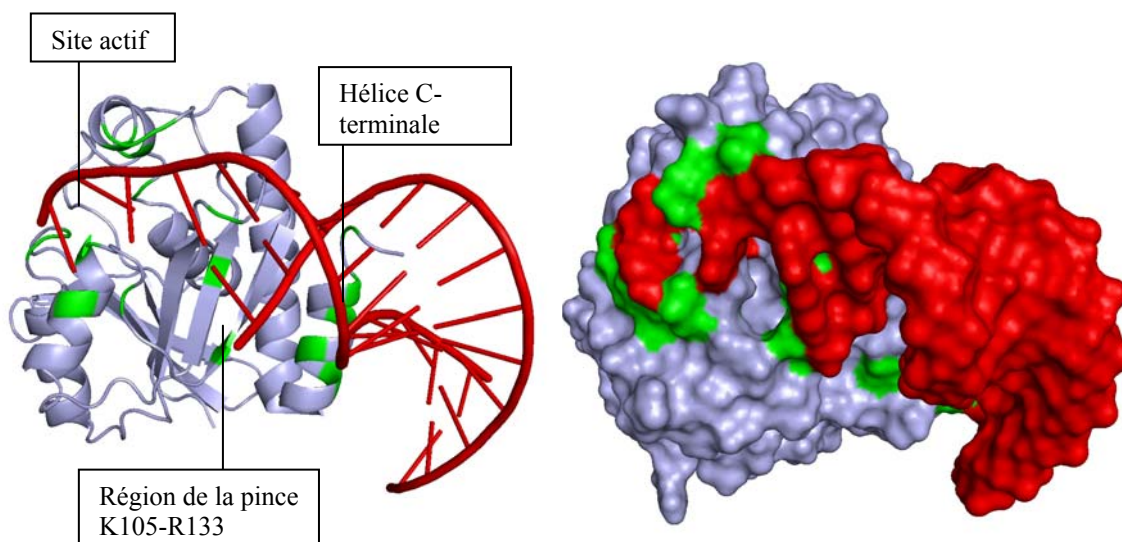
L'énergie totale de chacun de ces modèles a ensuite été calculée, ainsi qu'un terme d'énergie de contraintes qui reflète le non-respect des contraintes floues imposées. Trois solutions avec le meilleur compromis entre l'énergie totale et l'énergie liée au contraintes ont été retenues (voir figure 50). Dans ces trois modèles, l'ARN est bien en contact avec la zone d'interaction prédite par

l'expérience de RMN, et l'ARN s'oriente avec son extrémité 3' du côté du site actif de la PTH dans deux de ces modèles dont celui représenté sur la figure 51. Dans le troisième, l'orientation de l'ARN est inverse, avec l'extrémité CCA éloignée du site actif.



**Figure 50 : Energies des modèles de complexes entre la PTH et le duplex obtenus sans orientation préalable**

Pour chacun des 100 modèles obtenus, un losange bleu indique l'énergie totale (en abscisse) et l'énergie liée aux contraintes floues (en ordonnée). Les 3 solutions retenues sont entourées en rouge.

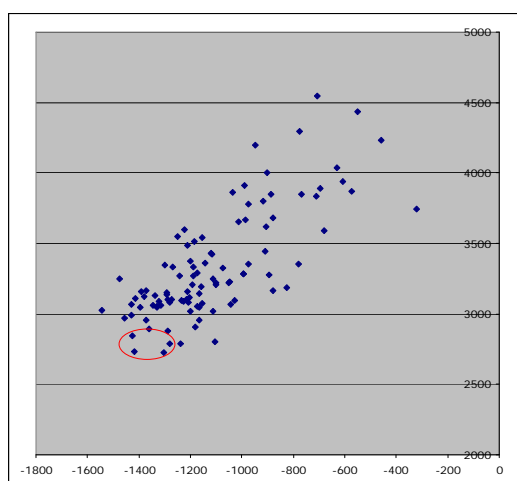


**Figure 51 : Complexe entre la PTH et le duplex sans orientation préalable**

Représentation d'un des meilleurs modèles du complexe entre la PTH et le duplex. Les résidus de la liste de contraintes floues utilisées pour l'obtention de ce modèle sont colorés en vert, le duplex avec son extrémité 3' à gauche est en rouge et la protéine en gris. Les deux représentations montrent la même vue du complexe en mode « cartoon » et surface, respectivement.

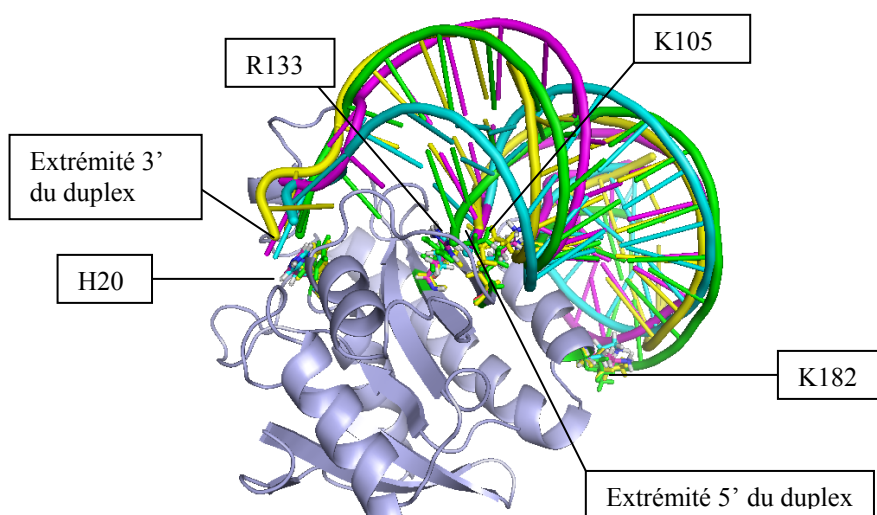
### 4.1.3 Modélisation avec orientation préalable du duplex

Pour pouvoir faire converger nos modèles rapidement vers les complexes d'énergie minimale, nous avons relancé une série de 100 calculs en positionnant à nouveau le duplex à une distance variable de la protéine ( $20 \pm 10 \text{ \AA}$ ), mais en partant toujours d'une orientation identique à celle du meilleur modèle obtenu ci-dessus. Les 4 solutions présentant le meilleur compromis entre l'énergie totale et l'énergie liée aux contraintes floues (entourées sur la figure 52) sont présentées sur la figure 53.



**Figure 52 : Energies des modèles de complexes entre la PTH et le duplex obtenus avec une orientation préalable du duplex**

Pour chacun des 100 modèles obtenus, un losange bleu indique l'énergie totale (en abscisse) et l'énergie liée aux contraintes floues (en ordonnée). Les 4 solutions retenues sont entourées en rouge.



**Figure 53 : Structure des modèles de complexes entre la PTH et le duplex obtenus avec une orientation préalable du duplex**

La structure de la PTH est représentée en gris. Les positions du duplex dans les quatre meilleurs modèles (entourés sur la figure 52) sont représentées avec quatre couleurs également utilisées pour représenter les chaînes latérales de H20, K105, R133 et K182.



	H20 - O <sup>3'</sup>	K105 - O <sup>5'</sup>	R133 - O <sup>5'</sup>	K182 - OP
1	2,55	2,38	3,33	2,22
2	8,78	3,99	7,6	2,21
3	8,09	4,99	7,58	2,22
4	8,7	7,35	2,228	2,89

**Tableau 11 : Distances entre le duplex et la protéine dans les quatre meilleurs modèles obtenus**

Toutes les distances sont données en Å. La 2<sup>ème</sup> colonne indique la distance entre le O<sub>3'</sub> de l'adénosine 3'-terminale du duplex et le N<sup>ε</sup> de l'histidine 20 de la PTH. La 3<sup>ème</sup> colonne indique la distance entre un H<sup>δ</sup> de K105 et un atome d'oxygène du groupement phosphate +1 du duplex. La 4<sup>ème</sup> colonne indique la distance entre un H<sup>η</sup> du groupement guanidium de R133 et un atome d'oxygène du groupement phosphate +1 du duplex. La 5<sup>ème</sup> colonne indique la distance entre un H<sup>δ</sup> de K182 et un atome d'oxygène d'un groupement phosphate du duplex. Pour le modèle 1, le groupement phosphate le plus proche de K182 est celui porté par la guanosine 53 du duplex (selon la numérotation de l'ARNt<sup>His</sup>) qui correspond à la dernière base en amont (du côté 5') de la boucle TψC d'un ARNt. Pour les modèles 2, 3 et 4, le groupement phosphate le plus proche de K182 est porté par la cytidine 52 du duplex.

Dans l'ensemble de ces modèles, i) l'extrémité 3'-acceptrice de l'ARN se positionne dans la région du site actif de l'enzyme, au voisinage de l'histidine 20 qui est le résidu catalytique le plus important (voir tableau 11), ii) le phosphate +1 de l'ARN se trouve proche des résidus Lys105 et Arg133, ce qui est en accord avec l'existence d'une pince formée par les chaînes latérales de ces résidus, autour de ce phosphate, iii) l'hélice C-terminale de la protéine se positionne dans le grand sillon du bras TψC, où le groupement amino de K182 est situé à moins de 3 Å d'un groupement phosphate de l'ARN. Dans le modèle 1, ce phosphate est celui de la guanosine 53 (selon la numérotation de l'ARNt<sup>His</sup>). Dans les modèles 2, 3 et 4, il s'agit du phosphate de la cytidine 52. Les modèles obtenus sont compatibles avec l'ensemble des résultats fournis par la RMN, ce qui confirme la pertinence de l'interaction entre la PTH et le duplex décrite par ces modèles.

Du côté du site actif, on aurait pu s'attendre à retrouver un empilement de l'adénine 3'-terminale de l'ARNt sur F66 et une interaction entre le 5'-phosphate de l'adénosine 76 et Y15, comme nous avons conclu dans le cas de la diacétyl-Lys-(3'NH)-adénosine, mais cette possibilité ne semble pas avoir été exploitée par notre programme. Vu le positionnement du duplex sur nos modèles, il semble cependant possible d'obtenir de telles interactions en modifiant seulement la conformation de l'extrémité CCA du duplex, sans toucher au reste de la molécule.

## 4.2 Modélisation de l'interaction entre la PTH et un de ses substrats (duplex<sup>His</sup> aminoacylé et acétylé)

### 4.2.1 Construction du modèle

Les résultats acquis avec la diacétyl-Lys-(3'NH)-adénosine et le duplex<sup>His</sup> nous ont permis de tirer des conclusions sur :

(i) la position du site actif au voisinage de H20, l'interaction de N114 interagissant avec une molécule d'eau responsable de la catalyse et l'implication des chaînes latérales de N10, Y15, F66 et N68, dans la stabilisation du substrat, grâce à des interactions avec le premier acide aminé ou avec l'adénosine 3'-terminale du substrat ;

(ii) le rôle de la chaîne latérale de l'arginine 133 et de la lysine 105 dans l'interaction avec le phosphate +1 du substrat ;

(iii) l'implication de l'hélice C-terminale de la PTH dans la fixation de la partie ARN du substrat, notamment via la lysine 182.

Pour vérifier que ces résultats étaient bien compatibles entre eux et avec la formation d'un complexe qui permet l'hydrolyse du substrat, nous avons cherché à construire un modèle incluant la protéine, un de ses substrats (le diacétyl-Lys-duplex<sup>His</sup> et une molécule d'eau dans le site actif. A partir des résultats précédents, nous avons défini un certain nombre de contraintes de distances, décrites ci-dessous.

#### Site actif :

Au niveau du site actif, une première série de contraintes a été suggérée par l'étude de l'interaction entre la PTH et la diacétyl-Lys-(3'NH)-adénosine (voir chapitre 2) :

- distance de 3 Å entre l'atome d'oxygène O<sup>5'</sup> du groupement 5'-phosphate de l'adénine 76 du substrat et l'oxygène de la chaîne latérale de Y15.

- distance de 4 Å entre l'atome C<sup>8</sup> de l'adénine A76 et les atomes C<sup>5</sup> ou C<sup>6</sup> du cycle aromatique de F66, ainsi qu'entre les atomes C<sup>4</sup> et N<sup>3</sup> de l'adénine A76 et C<sup>7</sup> de F66, pour favoriser une interaction entre l'adénine 3'-terminale du substrat et la phénylalanine 66 de la protéine.

- distance de 2 Å entre un des atomes H<sup>δ</sup> de la chaîne latérale de l'asparagine N10, et soit l'atome O<sup>2'</sup> (voir figure 58) du ribose 3'-terminal du substrat, soit l'atome O<sup>3</sup> (voir figure 58) du groupement acétate porté par le groupement α-NH<sub>2</sub> de la lysine du substrat (les différents modèles obtenus avec la diacétyl-Lys-(3'NH)-adénosine suggéraient ces deux possibilités).

- distance de 2 Å entre un des atomes H<sup>δ</sup> de la chaîne latérale de l'asparagine N68 et l'atome O<sup>1</sup> du groupement oxyle de la fonction ester du ligand.

#### Molécule d'eau catalytique :

Nous avons ajouté trois contraintes afin de vérifier si H2O peut bien activer la molécule d'eau responsable de l'attaque nucléophile sur le groupement carbonyle du substrat entraînant son hydrolyse. La stabilisation de cette molécule d'eau par N114 est suggérée par l'étude de la fixation de la diacétyl-Lys-(3'NH)-adénosine sur la PTH (voir chapitre 2).

- distance de 2 Å entre l'oxygène de la molécule d'eau et au moins un des atomes H<sup>δ</sup> de la chaîne latérale de l'asparagine N114,
- distance de 2 Å entre le N<sup>ε</sup> de H2O et un atome d'hydrogène de cette molécule d'eau,
- distance de 2 Å entre l'oxygène de cette molécule d'eau et l'atome de carbone du groupement carbonyle de la lysine.

Nous avons de plus rajouté une contrainte permettant d'orienter la chaîne latérale de H2O de manière que l'hydrogène H<sup>δ1</sup> de H2O soit proche (à 2 Å) d'un oxygène de la chaîne latérale du glutamate 93 ce qui permettrait la déprotonation de H2O par D93 suggérée par Goodall *et al.* (Goodall *et al.*, 2004).

Interactions

#### Interaction avec le phosphate +1 :

- distance de 2 Å entre un des atomes H<sup>n</sup> de R133 et un des atomes d'oxygène du groupement 5'phosphate de la cytidine 1 du duplex.
- distance de 2 Å entre un des atomes H<sup>ζ</sup> de K105 et un des atomes d'oxygène du groupement 5'phosphate de la cytidine 1 du duplex.
- distance supérieure à 4 Å entre les atomes H<sup>n</sup> de R133 les atomes H<sup>ζ</sup> de K105

#### Interaction avec l'hélice C-terminale :

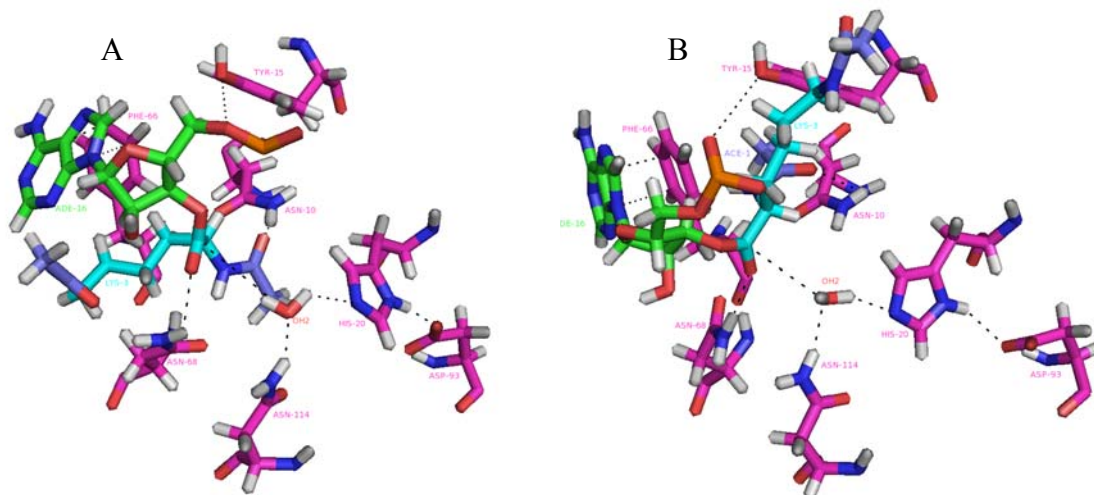
- distance de 4 Å entre un des atomes H<sup>ζ</sup> de K182 et un atome d'oxygène d'un des quatre derniers groupements phosphate de la région de l'ARN mimant la tige TΨC d'un ARNt.

Les modèles ont été construits selon le même protocole que pour la modélisation du complexe entre la PTH et le duplex<sup>His</sup> en commençant avec une préorientation du substrat et en attribuant une pénalité de 10 kcal.mol<sup>-1</sup>.Å<sup>-2</sup> à chaque contrainte pour laquelle la distance imposée était dépassée. Ils ont ensuite subi une relaxation sans contraintes qui leur a permis d'atteindre un état d'équilibre énergétique.

### **4.2.2 Analyse des modèles obtenus**

Les cinq modèles de plus basse énergie obtenus montrent que toutes ces interactions peuvent être établies simultanément, même si les distances sont légèrement différentes de celles qui étaient définies par les contraintes (voir tableau 12). Les modèles sont très proches les uns des autres. La principale différence concerne la position de la chaîne latérale de la lysine qui s'étend dans la direction de N68 sauf dans le modèle 4, où elle s'étend vers Y15 (voir figure

54). Dans les deux cas cette position est différente de celle occupée par la lysine dans le cas du complexe entre la PTH et la diacétyl-Lys-(3'NH)-adénosine où elle était orientée vers N114.



**Figure 54 : Site actif de la PTH liée au diacétyl-Lys-duplex<sup>His</sup>**

Certains résidus du site actif de la PTH sont représentés en magenta, l'adénosine 3'-terminale du diacétyl-Lys-duplex<sup>His</sup> est représentée en vert, sa lysine en cyan et les deux groupements acétyle en violet. Dans le modèle 4 (B) la lysine est orientée vers Y15 alors que dans les 4 autres modèles, tous très proches du modèle 1 (A), elle est orientée vers l'extérieur de la protéine, dans la direction de N68. Des pointillés représentent les interactions décrites ci-après.

		1	2	3	4	5	Moyenne	Ecart type
1	N68 - O <sup>1</sup>	3,29	3,3	2,37	1,89	1,79	2,528	0,733
2	N10 - O <sup>2</sup>	1,77	1,9	1,53	2,73	2,02	1,99	0,451
3	N10 - O <sup>2'</sup>	7,3	7,22	8,08	7,19	7,03	7,364	0,412
4	N114 - O <sup>w</sup>	2,38	2,25	2,09	2,11	2,35	2,236	0,133
5	H20 - H <sup>w</sup>	2,28	2,49	2,98	2,06	2,16	2,394	0,364
6	C' - O <sup>w</sup>	3,88	3,26	4,26	3,61	4,22	3,846	0,421
7	Y15 - O <sup>5'</sup>	3,11	3,01	3,56	5,07	3,27	3,604	0,845
8	F66 - A76	3,07	3,04	3,42	3,26	3,48	3,254	0,198
9	K182 - O <sub>P</sub>	2,89	3,47	3,28	3,82	3,234	3,3388	0,340
10	K105 - O <sub>P+1</sub>	2,5	2,73	2,57	2,43	2,56	2,558	0,111
11	R133 - O <sub>P+1</sub>	2,7	3	2,06	2,64	2,94	2,668	0,372
12	K105 - O <sub>P</sub>	2,23	2,32	2,27	2,29	2,35	2,283	0,061
13	C' - C' <sub>Cter</sub>	4,85	5,21	4,74	4,63	4,44	4,774	0,286
14	C' <sub>-1</sub> - C' <sub>-1Cter</sub>	2,04	2,37	0,76	4,04	0,97	2,036	1,312
15	H20 - D93	2,39	2,51	1,76	2,49	1,67	2,164	0,413

**Tableau 12 : Distances entre la PTH et le diacétyl-Lys-duplex<sup>His</sup>**

Pour les 5 meilleurs modèles obtenus lors de la modélisation du complexe entre la PTH et le diacétyl-Lys-duplex<sup>His</sup>, les distances minimales entre divers atomes de ces deux composés sont données en Å ainsi que la distance moyenne et l'écart-type (voir figure 58 pour la numérotation des atomes).

Ligne 1 : entre un H<sup>δ</sup> de N68 et O<sub>1</sub> du diacétyl-Lys-duplex<sup>His</sup>

Ligne 2 : entre un H<sup>δ</sup> de N10 et O<sub>2</sub> du diacétyl-Lys-duplex<sup>His</sup>

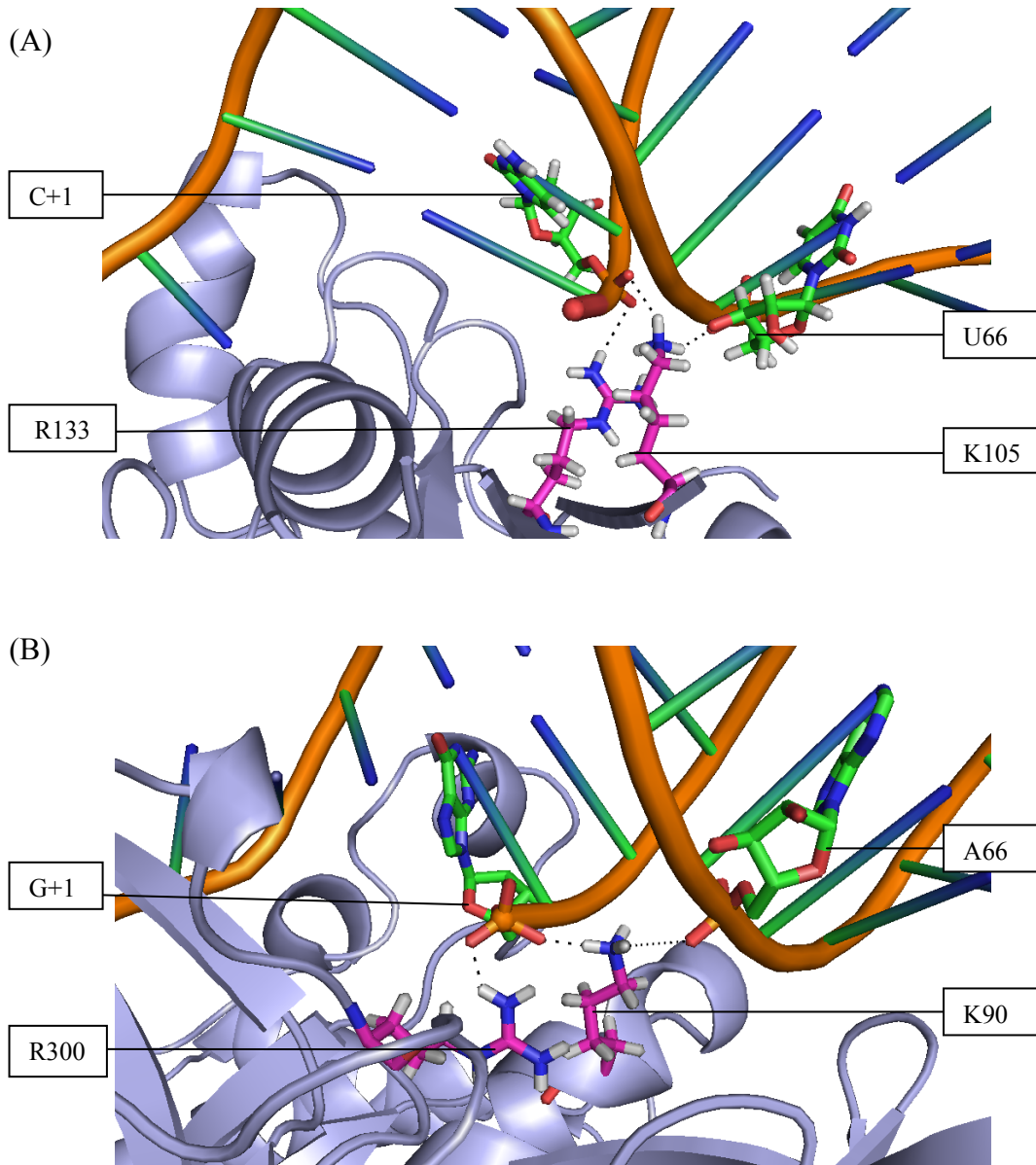
- Ligne 3 : entre un H<sup>δ</sup> de N10 et O<sub>2</sub>' de l'adénosine 76 du diacétyl-Lys-duplex<sup>His</sup>
- Ligne 4 : entre un H<sup>δ</sup> de N114 et l'atome d'oxygène de la molécule d'eau présente dans le site actif.
- Ligne 5 : entre le N<sup>ε</sup> de H20 et un atome d'hydrogène de la molécule d'eau présente dans le site actif.
- Ligne 6 : entre le C' du diacétyl-Lys-duplex<sup>His</sup> et l'atome d'oxygène de la molécule d'eau.
- Ligne 7 : entre l'oxygène O<sup>n</sup> de la Tyr15 et un atome d'oxygène du groupement phosphate de l'adénosine 3'-terminale du diacétyl-Lys-duplex<sup>His</sup>.
- Ligne 8 : entre un carbone du cycle aromatique de F66 et un carbone de l'adénine 3'-terminale du diacétyl-Lys-duplex<sup>His</sup>.
- Ligne 9 : entre un H<sup>ζ</sup> de K182 et un atome d'oxygène d'un groupement phosphate du diacétyl-Lys-duplex<sup>His</sup>, celui de G53 (selon la numérotation de l'ARNt<sup>His</sup>) pour les modèles 1,2 et 5 et celui de C52 pour les modèles 3 et 4.
- Ligne 10 : entre un H<sup>ζ</sup> de K105 et un atome d'oxygène du groupement phosphate +1 du diacétyl-Lys-duplex<sup>His</sup>.
- Ligne 11 : entre un H<sup>n</sup> de R133 et un atome d'oxygène du groupement phosphate +1 du diacétyl-Lys-duplex<sup>His</sup>.
- Ligne 12 : entre un H<sup>ζ</sup> de K105 et un atome d'oxygène du groupement phosphate de l'uracile 66 (selon la numérotation de l'ARNt<sup>His</sup>) du diacétyl-Lys-duplex<sup>His</sup>.
- Ligne 13 : entre l'atome C' du diacétyl-Lys-duplex<sup>His</sup> dans nos modèles et, après superposition des deux PTH, l'atome C' du résidu C-terminal (Q193) de l'extrémité de la protéine voisine présente dans le site actif de la PTH dans la structure cristallographique.
- Ligne 14 : entre l'atome C'<sub>1</sub> du diacétyl-Lys-duplex<sup>His</sup> (appartenant au groupement acétyle mimant une liaison peptidique supplémentaire) dans nos modèles et, après superposition des deux PTH, l'atome C'<sub>1</sub> du résidu A192 de l'extrémité de la protéine voisine présente dans le site actif de la PTH dans la structure cristallographique.
- Ligne 15 : entre le N<sup>δ</sup> de H20 et un O<sup>δ</sup> de l'aspartate 93 de la PTH.

En particulier :

- L'hélice C-terminale est en mesure d'interagir avec le bras TψC de l'ARNt, notamment via K182 dont les H<sup>ζ</sup> se trouvent à environ 3 Å d'un oxygène d'un groupement phosphate de l'ARN (voir ligne 10 du tableau ci-dessus). Ce phosphate est celui de G53 dans 3 des 5 modèles et celui de C52 dans les deux autres cas. Ce résultat explique l'importance de ce résidu et de cette hélice que nous avons mise en évidence par mutagenèse dirigée.

- le phosphate 5'-terminal (ou +1 dans le cas de l'ARNt<sup>His</sup>) peut être positionné dans une pince formée par les chaînes latérales des résidus K105 et R133, avec une distance inférieure à 3 Å, dans tous nos modèles, entre un oxygène de ce groupement phosphate et à la fois un H<sup>ζ</sup> de K105 et un H<sup>n</sup> de R133. K105 semble aussi pouvoir établir une liaison hydrogène via un autre H<sup>ζ</sup> avec un oxygène d'un groupement phosphate de l'autre brin (celui de U66), puisqu'on les retrouve aussi à une distance inférieure à 3 Å dans nos 5 modèles. Ceci suggère que K105 est intercalé entre les deux brins de l'ARN. Le mode de fixation et de reconnaissance du phosphate 5' d'un ARNt, qui permet à la PTH de reconnaître spécifiquement les ARNt élongateurs, a déjà été comparé à celui de EF-Tu (Fromant *et al.*, 1999) qui reconnaît aussi ce groupement phosphate grâce à une arginine et une lysine. Sur la structure cristallographique du complexe entre EF-Tu et un Phe-ARNt<sup>Phe</sup>, on voit d'ailleurs (sur la figure 55) que la lysine de

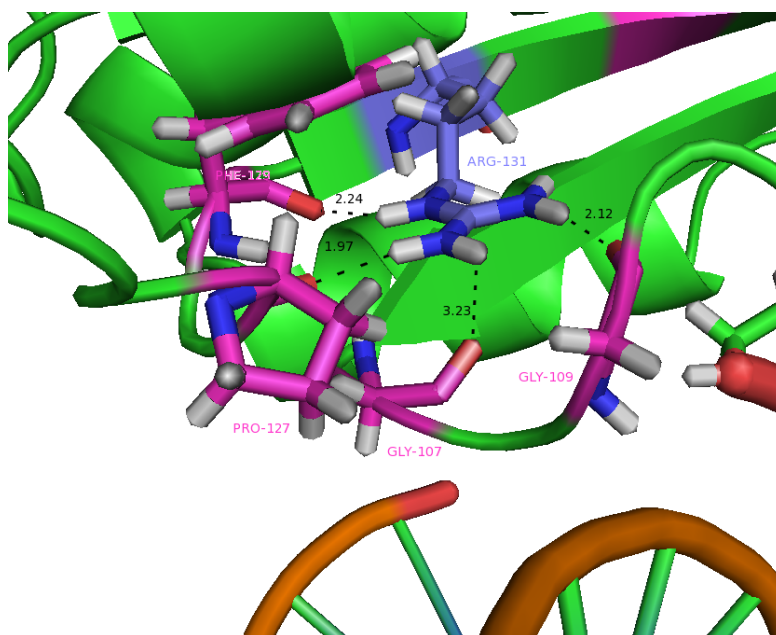
cette pince est aussi intercalée entre les deux brins de l'ARN au niveau du même nucléotide (A66 pour l'ARNt<sup>Phe</sup>).



**Figure 55 : Comparaison de la « pince » lysine-arginine dans la PTH et dans EF-Tu**

La figure (A) représente la région de la pince K105-R133 dans le modèle 1 du complexe entre la PTH et le diacétyl-Lys-duplex<sup>Lys</sup>. La figure (B) représente la région de la pince K90-R300 dans la structure cristallographique du complexe EF-Tu : Phe-ARNt<sup>Phe</sup> (code PDB : 1TTT). Dans chacune des deux figures la protéine est représentée en gris et l'ARN en orange. Les chaînes latérales des résidus K105 et R133 de la PTH et K90 et R300 de EF-Tu sont représentées en magenta. Les nucléotides portant les groupements phosphate interagissant avec ces résidus sont représentés en vert.

Au voisinage de cette région, on trouve l'arginine R131 qui a un rôle important puisque sa mutation en alanine diminue de deux ordres de grandeur l'efficacité catalytique de la PTH. Cependant, la chaîne latérale de cette arginine est située assez loin de l'ARN, le groupement phosphate le plus proche étant le phosphate -1 présent uniquement sur l'ARNt<sup>His</sup>. Par contre, les atomes d'hydrogène de son groupement guanidium sont à moins de 3 Å du C' de plusieurs résidus : G107, G109, P127, F129 (voir figure 56). Ce résidu semble donc jouer un rôle structural dans la région de la pince K105-R133.

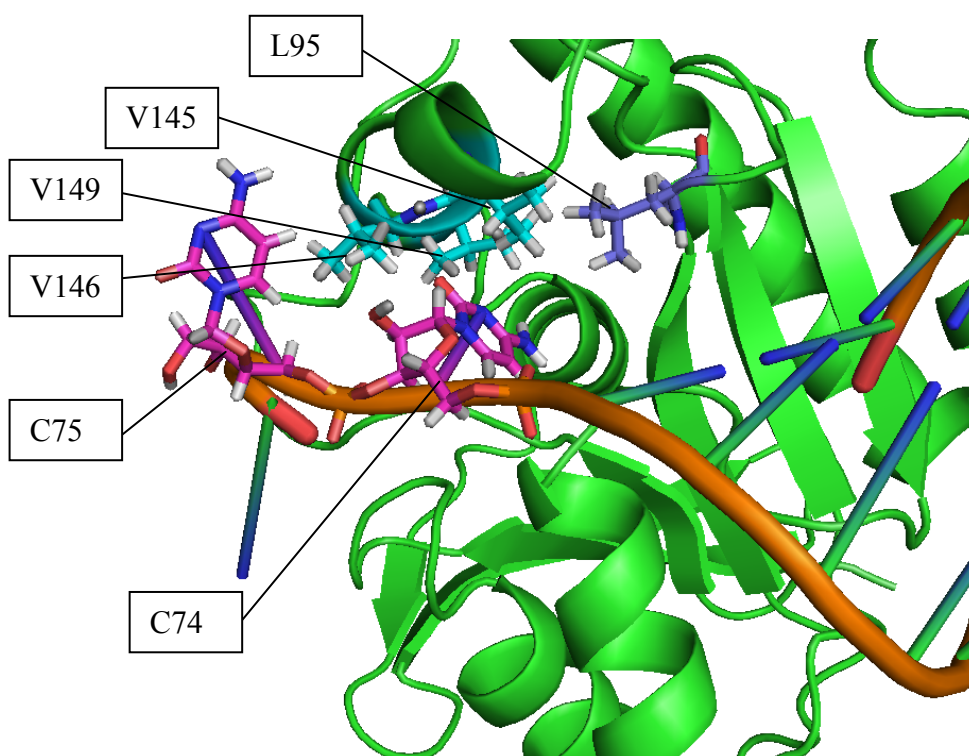


**Figure 56 : Région de R131 dans le complexe PTH : diacétyl-Lys-duplex<sup>His</sup>**

Un zoom de la structure 3D (mode « cartoon ») du modèle 1 du complexe entre la PTH et le diacétyl-Lys-duplex<sup>His</sup> est représenté en vert. La chaîne latérale de R131 est représentée en bleu et celles de G107, G109, P127 et F129 en magenta. L'extrémité 5'-terminale du diacétyl-Lys-duplex<sup>His</sup> apparaît en rouge à droite de l'image, assez éloigné de R131. Les distances entre le groupement guanidium de R131 et les C' des différents résidus sont indiquées en Å (en noir).

- Dans le site actif on retrouve un positionnement du substrat en accord avec la position de la diacétyl-Lys-(3'NH)-adénosine décrite plus haut. On peut noter que nous avons laissé la liberté à la chaîne latérale de N10 d'interagir avec O<sub>2</sub>' ou avec O<sub>2</sub> et que nos modèles ont tous convergé vers une interaction avec O<sub>2</sub> (voir ligne 2 et 3 du tableau 12), interaction que l'on trouvait aussi dans le cristal. Dans le cristal, la position du C' du résidu A192 dans le site actif de la PTH est d'ailleurs très proche (environ 2 Å en moyenne, voir ligne 14 du tableau 12) de celle du C'<sub>-1</sub> du groupement acétyle du diacétyl-Lys-duplex<sup>His</sup> qui est son homologue. Le C' du résidu Q193 est un peu plus éloigné (environ 4 Å) de la position du C' de la lysine du diacétyl-Lys-duplex<sup>His</sup>. On peut expliquer cela par le fait que le peptide mime un produit de la réaction alors que le diacétyl-Lys-duplex<sup>His</sup> est un substrat qui, dans nos modèles, est proche de la position lui permettant d'être hydrolysé par la PTH.

- Au voisinage du site actif, on note que l'extrémité 3'-terminale du duplex vient se loger dans un canal notamment formé par les valines 145,146 et 149 appartenant à la région de la boucle hélice surplombant le site actif, ce qui rend possible une interaction hydrophobe entre ces résidus et les cytosines de l'ARN. Dans chacun de nos modèles, au moins une de ces valines est en effet à moins de 3 Å d'une des deux cytosines de l'extrémité 3' de l'ARN (voir figure 57). L95, dont nous avons révélé l'importance par mutagenèse dirigée semble aussi participer à cette interaction hydrophobe avec l'ARN.



**Figure 57 : Interactions hydrophobes entre les cytosines de l'extrémité 3'-terminale de l'ARN et la PTH**

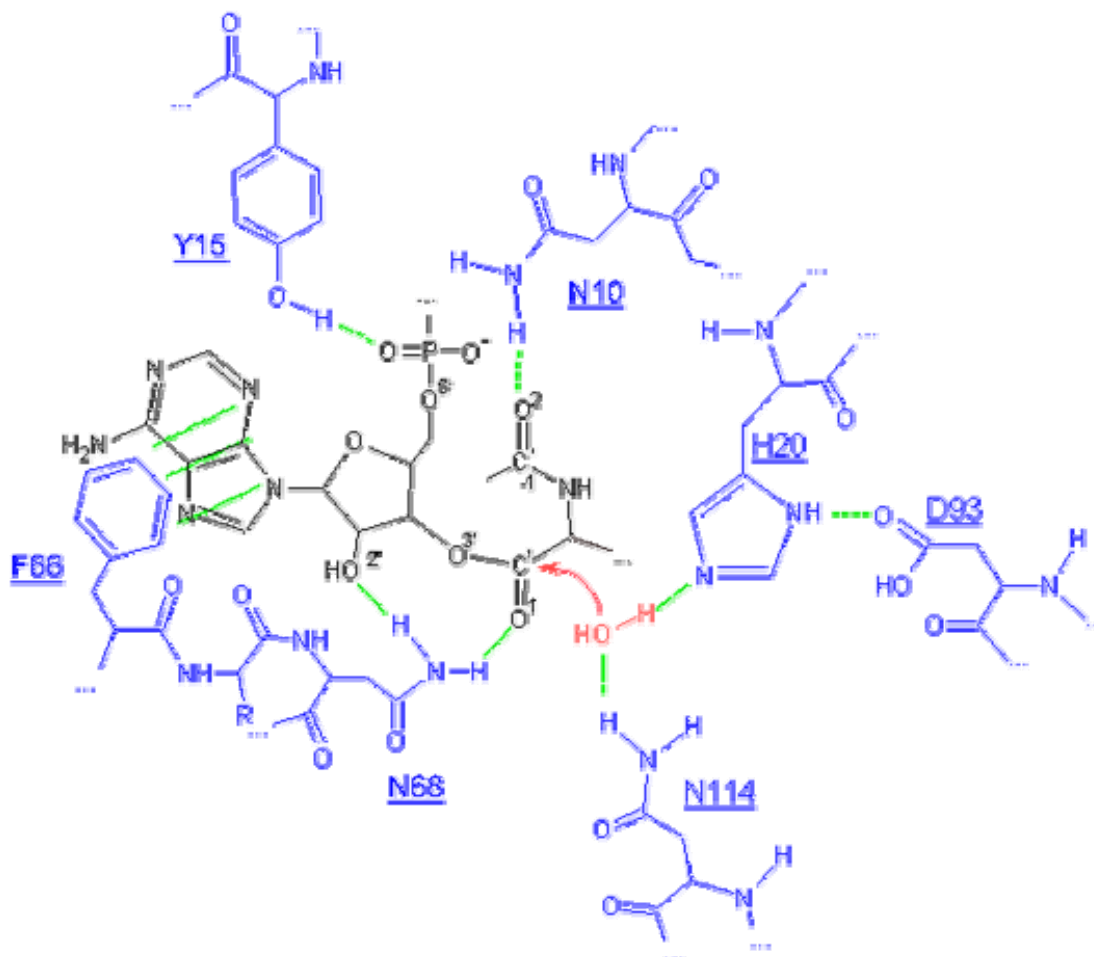
Sur le modèle 4 de l'interaction entre la PTH et le diacétyl-Lys-duplex<sup>His</sup>, représenté en mode « *cartoon* », les chaînes latérales des valines V145, V146 et V149, appartenant à la région de la boucle-hélice surplombant le site actif, sont représentées en cyan, L95 est représenté en bleu. Les cytidines C74 et C75 de l'extrémité 3'-terminale de l'ARN sont représentées en magenta.

- Dans l'ensemble de nos modèles, la molécule d'eau se positionne dans le site actif, stabilisée par une interaction avec un H<sup>δ</sup> de N114 (à moins de 2,4 Å de l'oxygène de la molécule d'eau), à moins de 3 Å d'un atome d'azote du cycle imidazole de His20 qui peut ainsi l'activer, et à environ 4 Å du carbone de la liaison ester entre le peptide et l'ARNt. Elle est donc idéalement placée pour l'attaque nucléophile de l'hydrolyse du peptidyl-ARNt (voir figure 58). Un tel positionnement soutient donc l'hypothèse présentée au chapitre 2 concernant le rôle de N114 dans le mécanisme catalytique de la PTH. On peut remarquer que l'interaction entre le H<sup>N</sup> de N114 et l'oxygène O<sup>3</sup> du groupement acétyle de la chaîne latérale de la lysine du substrat, que l'on observait avec la diacétyl-Lys-



(3'NH)-adénosine, a disparue puisque la position de la lysine est différente dans les résultats de ces deux modélisations, comme décrit ci-dessus. La déprotonation de l'histidine 20 par l'aspartate 93 est également compatible avec notre modèle puisque l'atome O<sup>δ</sup> du groupement carboxylique de cet aspartate se retrouve à une distance d'environ 2 Å du proton H<sup>δ</sup> de His20.

- Nos modèles valident l'hypothèse, elle aussi issue de l'étude de l'interaction entre la protéine et la diacétyl-Lys-(3'NH)-adénosine, concernant le rôle joué par Phe66. Un empilement du cycle hydrophobe de F66 avec l'adénosine à une distance comprise entre 3 et 4 Å (voir figure 54) est donc bien possible avec le substrat. De la même manière, nos modèles sont en accord avec une interaction entre Y15 et le groupement phosphate en 5' de l'adénosine 3'-terminale (ligne 7 du tableau 12). La position de l'adénosine 3'-terminale adoptée dans le cas de la diacétyl-Lys-(3'NH)-adénosine est donc bien compatible avec les autres points d'ancrage de l'ARN sur la protéine.

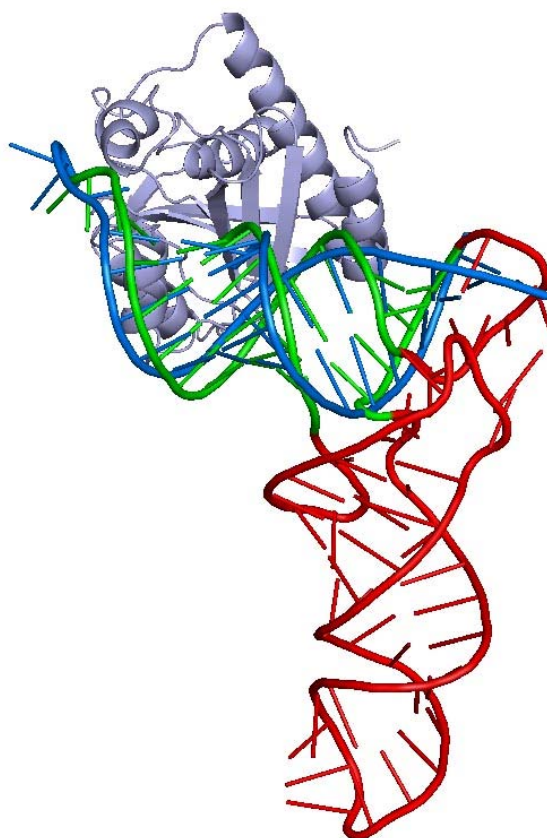


**Figure 58 : Site actif de la PTH occupé par son substrat**

Les résidus du site actif de la PTH impliqués dans la stabilisation du diacétyl-Lys-duplex<sup>His</sup> (représenté en noir) et de la molécule d'eau (en rouge) responsable de l'hydrolyse du substrat sont représentés en bleu. Des pointillés verts indiquent les différentes interactions décrites dans le texte ci-dessus. Une flèche rouge indique l'attaque nucléophile réalisée par la molécule d'eau responsable de l'hydrolyse du substrat.

### 4.3 Interaction de la PTH avec un ARNt complet

A partir des modèles obtenus, nous avons ensuite tenté de voir comment pouvait se faire l'amarrage d'un ARNt complet à la PTH. Pour cela, nous avons superposé la tige acceptrice et le bras T $\psi$ C d'un ARNt complet sur le diacétyl-lys-duplex<sup>His</sup>. Nous avons utilisé pour cela l'ARNt<sup>Phe</sup> dans la configuration qu'il adopte dans le complexe EF-Tu : Phe-ARNt<sup>Phe</sup> (code PDB : 1TTT). Comme on peut le voir sur la figure 59, cela ne provoque aucune gêne stérique. Aucune autre région de la PTH n'est au contact de l'ARNt.



**Figure 59 : Modèle d'un complexe PTH : ARNt**

En mode « *cartoon* », la PTH est représentée en gris, le diacétyl-lys-duplex<sup>His</sup> correspondant au modèle 1 décrit ci-dessus est représenté en bleu, et la région de l'ARNt<sup>Phe</sup> (sa tige acceptrice et son bras T $\psi$ C) superposée au diacétyl-lys-duplex<sup>His</sup> est représentée en vert, le reste de l'ARNt en rouge.



# **Conclusion**



La PTH bactérienne est indispensable chez *E. coli* alors qu'elle ne l'est pas chez les eucaryotes. Cette particularité en fait une cible antibactérienne très intéressante dans la perspective de l'élaboration d'une nouvelle classe d'antibiotiques qui viserait à inhiber cette enzyme. Plusieurs brevets, concernant l'utilisation des PTH de *Pseudomonas aeruginosa* (Edwards *et al.*, 2002) et *Enterococcus faecalis* (Edward *et al.*, 2003) comme cibles antibactériennes, ont d'ailleurs été déposés. Il est donc important de trouver un inhibiteur efficace de ces enzymes. La PTH permet, en effet, de recycler des ARNt séquestrés dans la cellule sous forme de peptidyl-ARNt, à la suite d'erreurs dans la traduction des protéines, afin qu'il puissent à nouveau jouer leur rôle indispensable dans la synthèse des protéines. Sans elle, la pénurie en ARNt libres entraîne le blocage de la machinerie de la traduction et la mort cellulaire. Bien que les structures cristallographiques des PTH d'*Escherichia coli* et de *Mycobacterium tuberculosis* aient été résolues, ainsi que celle d'un homologue de cette enzyme (CRS2) chez le maïs, aucune structure d'un complexe entre une de ces enzymes et un ligand n'a pu être déterminée.

La cristallographie n'ayant pas permis d'obtenir de telles structures, nous avons entrepris d'étudier l'interaction entre la PTH et son substrat grâce à la RMN. Nous avons tout d'abord utilisé cette technique pour étudier la PTH d'*E. coli* en solution. Nous avons trouvé que la structure de cette enzyme en solution était assez proche de sa structure cristallographique et noté quelques variations dans la structure secondaire de la protéine en solution, reflétant sa plasticité. Notamment, une région située au-dessus du site actif, composée d'une boucle et d'une hélice qui pourrait être dans une position ouverte en solution, serait, dans le cristal, refermée sur le site actif occupé par l'extrémité C-terminale d'une molécule voisine. Une telle idée avait déjà été mise en avant lorsque la structure cristallographique de la PTH d'*E. coli* dont le site actif est occupé, avait été comparée à celle de *M. tuberculosis*, dont le site actif est libre (Selvaraj *et al.*, 2007). Ce changement structural pourrait servir à verrouiller le substrat dans le site actif, une fois celui-ci reconnu par l'enzyme.

L'attribution des fréquences de résonance des couples H<sup>N</sup>-N de l'ensemble des résidus de la protéine nous a ensuite permis de cartographier l'interaction de cette protéine avec deux analogues de substrat.

Nous avons d'abord étudié l'interaction de la PTH avec la diacétyl-Lys-(3'NH)-adénosine. Ce composé mime la région d'un peptidyl-ARNt où le peptide est accroché à l'ARNt. A partir d'une liste de résidus dont les fréquences de résonance étaient affectées par la présence de ce ligand, le logiciel HADDOCK nous a permis de proposer un modèle de l'interaction entre la PTH et ce ligand qui suggère que la diacétyl-Lys-(3'NH)-adénosine se lie à la PTH à une position équivalente à celle qu'occupent l'adénosine 3'-terminale et le résidu d'acide aminé C-terminal dans une molécule de substrat (peptidyl-ARNt). Ce modèle indique que la Phe66 joue un rôle important dans le positionnement du substrat de la PTH, en interagissant avec l'adénine 3'-terminale de l'ARNt, leurs cycles aromatiques se recouvrant partiellement. On retrouve une telle interaction entre la

chaîne latérale d'une phénylalanine et une adénine dans plusieurs complexes, dont la structure a été résolue à ce jour, et, notamment, dans des complexes où une aminoacyl-ARNt synthétase est liée à de l'ATP (par exemple pour la GlyRS, code PDB : 1B76) à un aminoacyl-adénylate (par exemple pour l'HisRS, code PDB : 1ADY) ou à un ARNt (pour lequel l'interaction avec la phénylalanine se fait via son adénine 3'-terminale par exemple dans le cas de l'AspRS, code PDB : 1ASZ). Dans les cas les plus favorables, un tel recouvrement peut correspondre à un gain énergétique de plusieurs kcal.mol<sup>-1</sup> (Copeland *et al.*, 2008), ce qui confirme l'importance de cette interaction pour la PTH d'*E. coli* même si ce recouvrement n'est que partiel. Notre étude sur les sites d'éditions des synthétases, présentée en introduction, nous a aussi permis de voir que ce type d'interaction, entre un acide aminé aromatique et l'adénine 3'-terminale d'un aminoacyl-ARNt, peut, tout comme dans le cas de la PTH, être un élément important pour la catalyse. C'est notamment le cas de la phénylalanine 117 du domaine d'édition de la ThrRS de *Pyrococcus abyssi*, hautement conservée, (Hussain *et al.*, 2006). Dans le domaine d'édition CP1 des LeuRS et des IleRS on retrouve aussi une tyrosine ou une phénylalanine jouant un tel rôle (Fukunaga and Yokoyama, 2006).

La Tyr15 pourrait, elle, participer au positionnement du ribose 3'-terminal de l'ARNt en interagissant avec son groupement 5'-phosphate. Dans notre modèle, on retrouve les mêmes liaisons hydrogène impliquant N10 et N68 que celles qui existent entre deux molécules de PTH dans le cristal. Ceci confirme que ces deux résidus jouent donc un rôle crucial dans le positionnement de la région peptidique du substrat de la PTH. L'interaction de N10 avec le groupement carbonyle de l'avant-dernier résidu de l'extrémité C-terminale du peptide est, par ailleurs, assez importante pour que la PTH puisse faire la différence entre un peptidyl-ARNt et un aminoacyl-ARNt. Nous avons vérifié ce rôle de N10 au niveau biochimique. Ainsi, un mutant N10A ou N10D de la PTH hydrolyse un aminoacyl-ARNt aussi efficacement qu'un peptidyl-ARNt.

Au cours de cette étude nous avons aussi remarqué que deux résidus (G111 et K117) semblaient affectés par la fixation de la diacétyl-(3'NH)-adénosine alors qu'ils ne semblent pas être au contact direct de celle-ci. Ces deux résidus font partie d'un canal, qui relie le site actif et le site de reconnaissance du phosphate 5' du substrat (au niveau de la pince formée par les résidus K105 et R133 de la PTH). Ce canal est dans une position ouverte quand le site actif de la PTH d'*E. coli* est occupé par un peptide (dans le cristal) alors qu'il est dans une position fermée dans le cristal de la PTH de *M. tuberculosis* où le site actif est vide (Selvaraj *et al.*, 2007). On peut donc imaginer que la fixation de la diacétyl-(3'NH)-adénosine dans le site actif entraîne une ouverture de ce canal.

Comme on l'a vu dans l'introduction, la PTH doit pouvoir hydrolyser tous les aminoacyl-ARNt N-bloqués, à l'exception du formylméthionyl-ARNt<sup>fMet</sup> qui sert au démarrage de la traduction. Pour épargner le fMet-ARNt<sup>fMet</sup>, la PTH se sert d'un caractère exceptionnel de cet ARNt à l'extrémité de sa tige acceptrice : l'absence d'un appariement Watson-Crick entre les bases 1 et 72. D'autre part, il avait été suggéré que deux résidus basiques de la PTH, K105 et R133, formaient

une pince électrostatique interagissant avec le phosphate 5'-terminal de ses peptidyl-ARNt substrats. Si ce phosphate est au sommet de la tige acceptrice de l'ARNt<sup>fMet</sup>, caractérisée par l'absence d'appariement entre les bases 1 et 72, ce phosphate ne peut pas atteindre son site récepteur. Le phosphate en 5' des ARNt informerait donc la PTH de la stabilité de la paire de bases terminant le bras accepteur (Fromant *et al.*, 1999). Pour confirmer le rôle joué par cette pince, et pour cartographier toute la région d'interaction entre la protéine et son substrat, nous avons étudié le complexe formé par la protéine et un duplex d'ARN. Nous avons trouvé qu'une région étendue de résidus, regroupés sur une face de la protéine, étaient fortement affectée par la présence de ce duplex. Ceci confirme la spécificité de l'interaction entre l'ARN et la PTH dans les conditions de la RMN. Cette région contient tous les résidus qui, d'après des résultats antérieurs de mutagenèse dirigée, étaient soupçonnés d'interagir avec l'ARN, notamment dans la région du site actif où pourrait venir se loger l'extrémité 3'-terminale de l'ARN ainsi que dans la région de K105 et R133.

Par ailleurs, l'importance de la chaîne latérale de R133 dans la reconnaissance de l'ARN a été confirmée. En effet, sur la région du spectre HSQC de la PTH correspondant aux H<sup>ε</sup>-N<sup>ε</sup> des arginines, un pic se déplace fortement lors de la titration de la PTH par le duplex, s'élargit et finit par disparaître, ce qui confirme que la chaîne latérale d'une arginine interagit avec le duplex. L'attribution des H<sup>ε</sup>-N<sup>ε</sup> des arginines n'étant pas évidente, nous avons rajouté une mutation R133A à notre PTH et avons enregistré un spectre <sup>15</sup>N-HSQC de ce mutant, sur lequel nous avons vu, dans la région du spectre correspondant aux H<sup>ε</sup>-N<sup>ε</sup> des arginines, les mêmes pics qu'en absence de la mutation R133A à part justement celui qui se déplaçait en présence de duplex. Ainsi, nous avons pu conclure qu'il s'agissait du pic de résonance du couple H<sup>ε</sup>-N<sup>ε</sup> de la chaîne latérale de R133, ce qui confirme que son environnement chimique est fortement modifié lors de l'interaction entre la PTH et le duplex, en accord avec une interaction entre R133 et le phosphate +1 du duplex.

Enfin, les expériences de RMN ont indiqué que l'hélice C-terminale de la PTH jouait un rôle dans l'ancrage de l'ARNt, ce qui a ensuite été confirmé par mutagenèse dirigée, notamment dans le cas de K182.

Les résidus ainsi identifiés nous ont permis de proposer un modèle de l'interaction de la PTH avec ce duplex qui valide l'implication de ces résidus dans la reconnaissance du duplex. On retrouve bien une interaction entre l'ARN et l'hélice C-terminale de la PTH, ainsi que l'implication de la pince K105-R133 dans la reconnaissance du phosphate 5'-terminal (+1) de l'ARN. Certains résidus affectés par la présence de duplex, comme cette pince, les résidus basiques de l'hélice C-terminale ou K142, appartenant à la boucle-hélice surplombant le site actif, semblent être directement au contact de l'ARN alors que d'autres, comme R131 (ou ceux de l'hélice α1) font sans doute l'objet d'un réarrangement structural.



En regroupant les divers résultats acquis avec la diacétyl-(3'NH)-adénosine et avec le duplex<sup>His</sup>, nous avons pu aboutir à un modèle de l'interaction entre la PTH et son substrat, un diacétyl-Lys-duplex<sup>His</sup> et nous pouvons donc proposer un modèle de la reconnaissance et de l'action de la PTH sur son substrat. Le diacétyl-Lys-duplex<sup>His</sup> ou, plus généralement, n'importe quel peptidyl-ARNt serait, d'une part, reconnu au niveau de sa région ARN par le biais d'interactions avec l'hélice C-terminale de la PTH et, spécifiquement, au niveau de la pince K105-R133 qui se lie à son phosphate 5'-terminal. Dans un deuxième temps, ou peut-être de manière concertée, un réarrangement structural permettrait à la boucle-hélice surplombant le site actif de se refermer sur le substrat, celui-ci étant alors positionné précisément grâce à plusieurs interactions dans le site actif. En particulier, l'adénine 3'-terminale du substrat s'empile sur le cycle aromatique de F66 et la tyrosine 15 participe au positionnement de l'adénosine 3'-terminale du substrat en interagissant via son groupement hydroxyle avec le groupement 5'-phosphate de cette adénosine, et en positionnant le ribose de telle sorte que le groupement amino de N68 forme une liaison hydrogène avec l'atome d'oxygène du groupement oxyle de l'extrémité C-terminale du peptide estérifié à l'ARN. En outre, la partie peptide du substrat est maintenue en place dans le site actif grâce à une liaison hydrogène entre le groupement carbonyle de sa première liaison peptidique et le groupement amino de N10. Ainsi positionnée dans le site actif, la liaison ester entre le ribose et le premier acide aminé du peptide, peut être hydrolysée par une molécule d'eau stabilisée dans le site actif, à proximité du site de coupure, par N114 qui établit avec elle une liaison hydrogène. Cette stabilisation de la molécule d'eau lui permet d'être activée par l'histidine 20, précédemment déprotonée par l'acide aspartique 93, un peu comme dans le cas de la fin du mécanisme des protéases à sérine. Le carbone C-terminal de la région peptidique du substrat subit alors une attaque nucléophile de la part de cette molécule d'eau qui va libérer un ARNt pleinement fonctionnel pouvant rejoindre la machinerie de la traduction.

Les modélisations que nous avons réalisées ont utilisé comme point de départ la structure cristallographique de la PTH d'*E. coli*, qui, si l'on suppose que la PTH peut adopter plusieurs conformations, est sans doute très proche de la forme complexée à son substrat, du fait de l'occupation de son site actif par un de ses produits. Nous pourrions cependant étudier de manière plus précise la structure de la PTH en solution en attribuant davantage les fréquences de résonance des atomes des chaînes latérales de l'ensemble des résidus de la PTH et en analysant les NOE entre ces atomes.

Nous pourrions ensuite faire de même en présence de nos différents analogues de substrats, pour étudier les changements conformationnels que l'ensemble des résultats connus aujourd'hui suggèrent, notamment ceux qui entraîneraient la fermeture du site actif. L'attribution des fréquences de résonance de certains atomes de nos différents substrats, pourrait aussi nous permettre d'observer par RMN des NOE entre la protéine et le substrat qui nous permettrait aussi d'augmenter encore davantage la précision de nos modèles.

Dans notre introduction, nous avons remarqué que d'autres protéines connues agissent, elles aussi, sur des ARNt reliés par une liaison peptidique à un ou plusieurs acides aminés et leur étude nous a permis de chercher des similarités entre les modes de reconnaissance et d'action de ces enzymes et ceux de la PTH. Plusieurs types de réactions sont employés par ces enzymes pour hydrolyser la liaison ester entre un ARN et un acide aminé. Certaines de ces réactions sont catalysées par les ARNt eux-mêmes puisque leur groupement 3'-hydroxyle, libre à leur extrémité 3'-terminale, est parfois capable d'activer la molécule d'eau responsable de la catalyse. La molécule d'eau responsable de l'hydrolyse est par ailleurs souvent stabilisée dans le site actif par un réseau de liaisons hydrogène impliquant les chaînes latérales de plusieurs résidus comme des asparagines, des histidines, des lysines et des thréonines, et parfois des ions métalliques comme le zinc. Cette molécule d'eau est ensuite responsable de l'attaque nucléophile sur la liaison ester entre l'ARNt et l'acide aminé ou le peptide, ou dans certains cas (comme pour DTD1) sur un acyl-enzyme qui est un intermédiaire de la réaction. En fin de réaction, l'atome d'oxygène de cette molécule d'eau se retrouve à chaque fois à l'extrémité C-terminale du peptide ainsi libéré, comme dans le modèle que l'on propose pour la PTH. Cependant le mécanisme employé par la PTH, mis à part la stabilisation de la molécule d'eau dans le site actif par une asparagine, ne semble pas être utilisé par une autre de ces enzymes. Le rôle joué par l'histidine (H20) qui active la molécule d'eau responsable de l'hydrolyse, après avoir été rendu basique par l'intervention d'un acide aspartique (D93) n'est cependant pas sans rappeler la fin du mécanisme des protéases à sérine. Dans ce mécanisme, la sérine catalytique intervient tout d'abord pour former une liaison ester avec l'extrémité C-terminale du peptide coupé, et former un acyl-enzyme. La suite du mécanisme des protéases à sérine ressemble beaucoup à ce que nous proposons pour la PTH, puisqu'une molécule d'eau activée par une histidine elle aussi rendue basique par un acide aspartique, vient hydrolyser la liaison ester de l'acyl-enzyme, de la même manière qu'est hydrolysée la liaison ester entre l'ARN et le peptide dans le cas de la PTH.

Nos résultats actuels sur le mécanisme de la PTH mais surtout sur le positionnement précis du substrat dans le site actif ouvrent de nouvelles pistes pour la recherche d'un inhibiteur efficace de la PTH bactérienne. Cette étude par RMN confirme qu'une petite molécule comme la diacétyl-Lys-(3'NH)-adénosine peut établir des interactions spécifiques dans le site actif de la PTH, sans avoir besoin pour cela d'être reliée au corps de l'ARNt. Ce résultat est encourageant si on pense à la possibilité de concevoir des inhibiteurs de la PTH ayant des propriétés anti-bactériennes. Cependant l'affinité de la diacétyl-Lys-(3'NH)-adénosine pour la PTH reste très faible et une telle découverte nécessiterait de le modifier pour qu'il puisse établir des interactions plus fortes avec l'enzyme. Nous avons remarqué que la chaîne latérale de la lysine acétylée de cette molécule pouvait interagir avec l'asparagine N114 de la PTH, ce qui a certainement pour conséquence d'améliorer l'affinité de cette molécule pour la PTH. Or la PTH reconnaît tous les peptidyl-ARNt, et, donc n'interagit pas spécifiquement avec la chaîne latérale du premier acide aminé lié à l'ARNt, on ne retrouve d'ailleurs pas cette interaction avec la chaîne latérale de la lysine, dans le cas du diacétyl-Lys-

## Conclusion

---

duplex<sup>His</sup>. Ceci montre que, pour améliorer l'affinité d'un nouvel inhibiteur de la PTH, il ne faut pas forcément chercher à reproduire uniquement les interactions entre la protéine et son substrat naturel, mais qu'il est certainement possible de garder les interactions que l'on trouve dans le cas de la diacétyl-Lys-(3'NH)-adénosine en rajoutant à cette molécule des groupements fonctionnels susceptibles d'interagir avec d'autres résidus du site actif de la PTH. Une telle recherche pourrait être faite grâce à des méthodes de criblage *in silico* et *in vitro*.

## **Annexe 1**

### **Séquence de la PTH H20A étiquetée**



M<sub>20</sub>GSSHHHHH S<sub>10</sub>SGLVPRGSH

M<sub>0</sub>TIKLIVGLA N<sub>10</sub>PGA EYAATR

A<sub>20</sub>NAGAWFVDL L<sub>30</sub>AERLRAPLR

E<sub>40</sub>EAKFFGYTS R<sub>50</sub>VTGGEDVR

L<sub>60</sub>LVPTTFMNL S<sub>70</sub>GKAVAAMAS

F<sub>80</sub>FRINPDEIL V<sub>90</sub>AHDELDLPP

G<sub>100</sub>VAKFKLGGG H<sub>110</sub>GGHNGLKDI

I<sub>120</sub>SKLGNNPNF H<sub>130</sub>RLRIGIGHP

G<sub>140</sub>DKNKVVGFV L<sub>150</sub>GKPPVSEQK

L<sub>160</sub>IDEAIDEAA R<sub>170</sub>CTEMWFTDG

L<sub>180</sub>TKATNRLHA F<sub>190</sub>KAQ



## **Annexe 2**

### **Liste des expériences de RMN réalisées sur la PTH**





**HSQC** : (*Heteronuclear Single Quantum Correlation*) ( $N_i$  ;  $H^N_i$ ) : Cette expérience corrèle les fréquences de résonance des couples H-N. Elle permet d'observer les fréquences de résonance des couples  $H^N$ -N de chaque résidu ainsi que des couples  $H^E$ - $N^E$  des asparagines,  $H^D$ - $N^D$  des glutamines,  $H^{\epsilon 1}$ - $N^{\epsilon 1}$  des tryptophanes et  $H^E$ - $N^E$  des arginines (si la fenêtre d'observation est assez large). Les fréquences de résonance des couples  $H^{\epsilon 2}$ - $N^{\epsilon 2}$  des histidines,  $H^{\zeta}$ - $N^{\zeta}$  des lysines et  $H^{\eta}$ - $N^{\eta}$  des arginines ne sont pas détectables dans cette expérience.

**HSQC filtrée NH<sub>2</sub>** : ( $N_i$  ;  $H^N_i$ ) : Cette expérience corrèle uniquement les fréquences de résonance des atomes H et N des groupements NH, mais pas ceux des groupements NH<sub>2</sub>. Les pics correspondant aux groupements NH<sub>2</sub> des chaînes latérales des asparagines et des glutamines, n'apparaissent donc pas sur cette expérience.

**HN(CO)CA** : ( $C^{\alpha}_{i-1}$  ;  $N_i$  ;  $H^N_i$ ) : Cette expérience corrèle les fréquences de résonance des couples  $H^N$ -N à celle du  $C^{\alpha}$  du résidu précédant par transfert d'aimantation via un C' (non édité en fréquence). Elle permet donc de compléter les systèmes de spins en y ajoutant les fréquences des  $C^{\alpha}_{i-1}$ .

**HNCA** : ( $C^{\alpha}_{i-1}$  et  $C^{\alpha}_i$  ;  $N_i$  ;  $H^N_i$ ) : Cette expérience corrèle les fréquences de résonance des couples  $H^N$ -N à celles du  $C^{\alpha}$  du résidu correspondant et du résidu précédent. Elle permet donc de compléter les systèmes de spins en y ajoutant les fréquences des  $C^{\alpha}$ .

**HNCACB** : ( $C^{\alpha}_i$  et  $C^{\alpha}_{i-1}$  ;  $N_i$  ;  $H^N_i$ ) et en négatif : ( $C^{\beta}_i$  et  $C^{\beta}_{i-1}$  ;  $N_i$  ;  $H^N_i$ ). Cette expérience corrèle les fréquences de résonance des couples  $H^N$ -N à celles du  $C^{\alpha}$  et du  $C^{\beta}$  du résidu correspondant et à celles du  $C^{\alpha}$  et du  $C^{\beta}$  du résidu précédent ; elle permet donc de compléter les systèmes de spins par les fréquences des  $C^{\beta}_i$  et  $C^{\beta}_{i-1}$  identifiées parallèlement sur l'expérience CBCA(CO)NH..

**CBCA(CO)NH** : ( $C^{\alpha}_{i-1}$  et  $C^{\beta}_{i-1}$  ;  $N_i$  ;  $H^N_i$ ). Cette expérience corrèle les fréquences de résonance des couples  $H^N$ -N à celles du  $C^{\alpha}$  et du  $C^{\beta}$  du résidu précédent par transfert d'aimantation via un atome C'. Pour les carbones du résidu précédent, elle est plus sensible que la HNCACB.

**HNCO** : ( $C'$  ;  $N_i$  ;  $H^N_i$ ). Cette expérience corrèle les fréquences de résonance des couples  $H^N$ -N à celle du C' du résidu précédent. Elle permet donc de compléter les systèmes de spins par les fréquences des  $C'_{i-1}$ .

**HBHA(CO)NH** : ( $H^{\alpha}_{i-1}$  et  $H^{\beta}_{i-1}$  ;  $N_i$  ;  $H^N_i$ ) : Cette expérience corrèle les fréquences de résonance des couples  $H^N$ -N à celle des atomes  $H^{\alpha}$  et  $H^{\beta}$  du résidu précédent. Elle permet de déterminer, pour chaque système de spin,, les fréquences de résonance des protons  $H^{\alpha}$  et  $H^{\beta}$  du résidu précédent.

**HBHANH** : ( $H^{\alpha}_i$  et  $H^{\beta}_i$  ;  $N_i$  ;  $H^N_i$ ) : Cette expérience corrèle les fréquences de résonance des couples  $H^N$ -N à celles des atomes  $H^{\alpha}$  et  $H^{\beta}$  du résidu correspondant.

Elle permet en théorie de compléter les systèmes de spins par les fréquences des  $H^{\alpha}_i$  et  $H^{\beta}_i$ , mais elle était peu sensible.

**TOCSY HSQC** : ( $H_i$  ;  $N_i$  ;  $H^{N_i}$ ) : Cette expérience corrèle les fréquences de résonance des couples  $H^N$ -N à celles de tous les atomes  $H_i$  du résidu correspondant (en particulier les  $H^{\alpha}_i$  et les  $H^{\beta}_i$ ).

**HCCH TOCSY** : ( $H^j_i$  ;  $C^j_i$  ;  $H^k_i$ ) :  $C^j_i$  est un carbone de la chaîne latérale du résidu  $i$ ,  $H^j_i$ , un des hydrogènes de ce carbone, et  $H^k_i$  un hydrogène de la chaîne latérale de ce résidu. Cette expérience polluée, par le signal de l'eau dont la fréquence de résonance est proche de celle des  $H^{\alpha}$ , a aussi été réalisée dans  $D_2O$ . Cela est possible car, contrairement aux expériences précédentes, elle ne fait pas apparaître les fréquences de résonance des  $H^N$  des résidus qui s'échangent avec le deutérium dans  $D_2O$ . Elle corrèle des couples ( $H^j$  ;  $C^j$ ) (par exemple ( $H^{\alpha}$  ;  $C^{\alpha}$ ) ou ( $H^{\beta}$  ;  $C^{\beta}$ ), ( $H^{\gamma}$  ;  $C^{\gamma}$ )) à l'ensemble des hydrogènes de la chaîne latérale du résidu correspondant. Elle permet en théorie de retrouver toutes les fréquences  $H^j$  et  $C^j$  des chaînes latérales. En pratique, elle nous a permis de retrouver la plupart des déplacements chimiques des  $C^{\alpha}$ ,  $C^{\beta}$ ,  $H^{\alpha}$  et  $H^{\beta}$ , et pour certains résidus étudiés spécifiquement, ceux des autres carbones et hydrogènes de leur chaîne latérale.

**$^{15}N$ -NOESY-HSQC** : ( $H^j$  ;  $N_i$  ;  $H_i$ ), avec  $H^j$  et  $H_i$  à une distance généralement inférieure à 5 Å) : Cette expérience corrèle les fréquences de résonance des couples  $H_i$ - $N_i$ , à celle des atomes H voisins appartenant ou non au résidu  $i$ . Contrairement aux expériences précédentes, le transfert d'aimantation entre  $H^j$  et  $H_i$  ne passe pas par des liaisons covalentes mais grâce à un effet Overhauser (NOE, *Nuclear Overhauser Effect*), généralement possible entre deux atomes à moins de 5 Å l'un de l'autre.

**$^{13}C$ -NOESY-HSQC** : ( $H^j_i$  ;  $C^j_i$  ;  $H^k$ ), avec  $H^j_i$  et  $H^k$  à une distance généralement inférieure à 5 Å : Cette expérience, réalisée dans  $D_2O$ , corrèle les fréquences de résonance des couples  $H^j_i$ - $C^j_i$  (par exemple  $H^{\alpha}_i$ - $C^{\alpha}_i$ ), à celle des atomes H voisins appartenant ou non au résidu  $i$ , grâce à un effet Overhauser.

# **Matériels et Méthodes**



### Production et purification de la PTH H20A étiquetée

Le gène de la PTH liée à une étiquette de poly-histidine, portant la mutation H20A, a été produit par PCR en utilisant le plasmide pUCPth exprimant le mutant H20A (Schmitt *et al.*, 1997b) comme matrice, et les oligonucléotides 5'-GGT GGT CAT ATG ACG ATT AAA TTG ATT GTC GG-3' et 5'-CCG CTC GAG TTA TTG CGC TTT AAA GGC GT-3' comme amorces. Des modifications (soulignées) ont été introduites dans les séquences des amorces pour créer des sites de restrictions NdeI et XhoI. Le fragment ainsi amplifié a ensuite été purifié avec le *PCR Purification Kit* (50) de Qiagen, puis il a été digéré par NdeI et XhoI. Le produit de la digestion a été inséré dans le plasmide pET15blpa (Guillon *et al.*, 2005), entre les sites de restrictions NdeI et XhoI, en aval d'une séquence codant une étiquette de poly-histidine et un site de clivage par la thrombine, pour donner le plasmide pET15blpa::PTHH20A. La séquence du gène ainsi cloné a été vérifiée par séquençage ADN.

La souche K37ΔrecAλDE3 d'*E. coli* (Ferri-Fioni *et al.*, 2006) a été transformée avec le plasmide pET15blpa::PTHH20A. Les cellules ainsi transformées ont été cultivées à 37° dans un litre de milieu minimum M9 (7,5 g/l Na<sub>2</sub>HPO<sub>4</sub>, 3 g/l KH<sub>2</sub>PO<sub>4</sub>, 0,5 g/l NaCl, 1 mM MgSO<sub>4</sub>, 0,1 mM CaCl<sub>2</sub>, 5 mg/l thiamine), complété par du <sup>15</sup>NH<sub>4</sub>Cl (1,0 g/l) et du glucose (2 g/l) pour marquer à l'isotope <sup>15</sup>N l'ensemble des protéines ou par du <sup>15</sup>NH<sub>4</sub>Cl (1,0 g/l) et du <sup>13</sup>C-glucose (2 g/l) pour le double marquage isotopique <sup>13</sup>C-<sup>15</sup>N. Quand la densité optique de la culture à 650 nm a atteint 0,6, nous avons ajouté de l'isopropyl-1-thio-β-D-galactopyranoside (IPTG) à une concentration finale d'1 mM pour induire l'expression de la PTH. La culture des cellules a ensuite été poursuivie pendant 15 heures à 18°C. Les cellules ont été récupérées par centrifugation (20 min, 3000 g) et resuspendues dans 20 ml de tampon A (10 mM Hepes, pH 7,5, 500 mM NaCl). Puis elles ont été cassées par sonication (10 min, 0°C) et les débris cellulaires ont été éliminés par centrifugation (20 min, 8000 g). Les acides nucléiques ont été précipités par ajout de sulfate de streptomycine (3% poids/vol) dans le surnageant qui a ensuite été centrifugé (20 min, 8000 g). Le surnageant a ensuite été précipité par du sulfate d'ammonium (70%) et centrifugé (20 min, 8000 g), pour éliminer les petites molécules et les impuretés ne précipitant pas. Le culot a été dissous dans 10 ml de tampon A et dialysé contre le même tampon. Cet extrait brut de protéines a ensuite été appliqué sur une colonne Talon (Clontech) de 3 ml équilibrée dans le tampon A, sur laquelle s'accroche la protéine grâce à son étiquette de poly-histidine. La colonne a été lavée avec du tampon A contenant des concentrations croissantes d'imidazole (0, 5 mM et 25 mM) et éluée avec du tampon A contenant 50 mM d'imidazole, concentration à laquelle la PTH se décroche. Les fractions contenant la PTH ont été identifiées grâce à une analyse par électrophorèse sur gel de polyacrylamide en présence de SDS. Elles ont été réunies puis concentrées par précipitation au sulfate d'ammonium (70%) et par centrifugation (20 min, 8000 g). Le culot a ensuite été repris dans 0,5 ml de tampon B (50 mM acétate de sodium, pH 6,0, 200 mM NaCl) et dialysé intensivement contre ce même tampon. Cette méthode de purification permet de

produire environ 2 mg de protéine dont l'homogénéité, estimée par une analyse sur gel de polyacrylamide en présence de SDS, est d'environ 95 %.

Pour vérifier l'absence de RNase dans notre préparation, de l'ARN<sup>t</sup><sup>His</sup> (50 µM) a été incubé pendant 15 jours en présence de PTH (75 µM) issue de cette préparation, à 4°C et à 17°C. Ensuite, il a été précipité à l'éthanol, centrifugé pour être aminoacylé par l'histidyl-ARNt synthétase dans des conditions décrites précédemment (Fromant *et al.*, 2000). Aucune dégradation de l'ARN n'a été constatée.

### **Production et purification de PTH sauvage et des mutants étiquetés**

Pour obtenir les mutants de la PTH étiquetée, le gène *pth* codant la PTH sauvage a été introduit dans les plasmides pET15blpa de la même manière que dans le cas du mutant HA20, en utilisant un plasmide pUC<sub>pth</sub> comme matrice dans l'expérience de PCR. Le plasmide obtenu a été appelé pET15blpa::*PTH*. Pour la mutagenèse dirigée, des plasmides pET15blpa contenant les mutations désirées (H92A, L95A, D96A, R131A, K182A ou N185A) ont été synthétisés par polymérisation d'ADN en utilisant pET15blpa::*PTH* comme matrice et deux amorces complémentaires portant la mutation, à l'aide de l'ADN polymérase Pfu-Turbo (Stratagene). Après plusieurs cycles thermiques, l'ADN initial, méthylé, a été coupé par DpnI avant la transformation de la souche XL1-Blue (Stratagene). Les séquences des gènes *pth* mutés ont systématiquement été vérifiées. Pour chaque mutant, la souche K37ΔrecAλDE3 d'*E. coli* a été transformée avec le plasmide pET15blpa::*PTH* portant la mutation puis cultivée dans un litre de milieu riche 2xTY. L'induction de l'expression des variants PTH et la purification des protéines ont été réalisées comme cela a été décrit plus haut pour la production de PTH H20A.

Le gène codant le double mutant PTH H20A R133A a été obtenu de la même manière, en introduisant la mutation R133A dans le plasmide pET15blpa::*PTH*H20A. La PTH H20A R133A, comportant l'étiquette de poly-histidine et marquée à l'isotope <sup>15</sup>N a été produite et purifiée comme la PTH H20A utilisée dans les expériences de RMN.

Les mutants N10A et N10D de la PTH d'*E. coli* ont été exprimés dans la souche K37Δp<sub>th</sub>Tr comme cela a été décrit pour d'autres mutants (Schmitt *et al.*, 1997b).

La PTH sauvage non étiquetée a été exprimée dans des cellules JM101TR (Hirel *et al.*, 1988) portant le plasmide pUC<sub>pth</sub> et purifiée comme décrit précédemment (Schmitt *et al.*, 1997a).

## Préparation des ARN

La minihélice<sup>His</sup>, la microhélice<sup>His</sup> et la tétraboucle<sup>His</sup> ont été préparées par Michel Fromant, en utilisant la procédure de transcription *in vitro* décrite par Milligan et Uhlenbeck (Milligan and Uhlenbeck, 1989). Toutes les transcriptions ont été réalisées pendant 180 min dans 5 ml de mélange réactionnel contenant 40 mM de Tris-HCl (pH 8,0), 200 nM des oligonucléotides matrice et amorce, 0,2 mM de chaque NTP (N = A, C, G, or U), 1 mM de GMP, 10 mM de MgCl<sub>2</sub>, 1 mM spermidine, 4,8 mM dithiothréitol, 0,1% de Triton X100 (v/v) et 15 µg/ml d'ARN polymérase du phage T7 (Uhlenbeck *et al.*, 1989). Les oligonucléotides utilisés comme matrice étaient : 5'-TGG GGT GGC TAA TGG GAT TCG AAC CCA CTA GCC ACC TAT AGT GAG TCG TAT TA-3' (minihélice<sup>His</sup>), 5'-TGG GGT GGC TAA TTC GAA TAG CCA CCT ATA GTG AGT CGT ATT A-3' (microhélice<sup>His</sup>) et 5'-TGG GGT GCC GAA GCA CCT ATA GTG AGT CGT ATT A-3' (tétraboucle<sup>His</sup>). La séquence de l'oligonucléotide amorce était 5'-TAA TAC GAC TCA CTA TAG-3'.

Après une extraction au phénol, les transcrits obtenus ont été purifiés par une chromatographie sur une colonne de tamis moléculaire (GF05, 1,1x27 cm) puis par une électrophorèse préparative sur un gel de polyacrylamide (18% poids/vol, 40x20x0,3 cm, 600 V, 24 h). Les transcrits ont été extraits du gel, précipités à l'éthanol et resuspendus dans une solution contenant 5 mM d'acétate de sodium (pH 4,8), 0,1 mM d'EDTA et 10 mM de MgCl<sub>2</sub>.

Les concentrations des minihélice<sup>His</sup>, microhélice<sup>His</sup> et tétraboucle<sup>His</sup> ont été estimées en utilisant des coefficients d'absorption molaire à 260 nm de 0,325, 0,245 et 0,156 µM<sup>-1</sup>.cm<sup>-1</sup> respectivement. Ces coefficients ont été obtenus avec l'aide du programme OligoCalc calculator (Kibbe, 2007), en prenant en compte les régions double brin indiquées à la figure 41.

L'ARNt<sup>His</sup> d'*E. coli* a été obtenu à partir d'une souche d'*E. coli* surproduisant cet ARNt, en utilisant la procédure décrite précédemment (Fromant *et al.*, 2000).

L'ARNt<sup>His</sup> et ses analogues (minihélice<sup>His</sup>, microhélice<sup>His</sup>, tétraboucle<sup>His</sup>) ont été aminoacylés pendant 30 min à 28°C, dans un milieu réactionnel contenant 20 mM Tris-HCl (pH 7.5), 7 mM MgCl<sub>2</sub>, 0,1 mM EDTA, 2 mM ATP, 10-50 µM L-[<sup>14</sup>C]histidine (327 Ci/mol), 1-8 µM d'histidyl-ARNt synthétase purifiée (Fromant *et al.*, 2000) et 5-50 µM d'ARN. Après acétylation par de l'anhydride acétique, les échantillons d'ARN aminoacylés et N-substitués ont été incubés en présence de sulfate de cuivre (10 mM), pour désacyler les ARNt aminoacylés mais non acétylés (Dutka *et al.*, 1993). Puis ils ont été purifiés par des chromatographies sur des colonnes Chelex 100 (Biorad) et GF05 (IBF) (Fromant *et al.*, 2000). Le diacétyl-[<sup>14</sup>C]lysyl-ARNt<sup>Lys</sup> (50 Ci/mol) a été préparé comme décrit précédemment (Schmitt *et al.*, 1997b).



Pour l'utiliser en grande quantité dans des expériences de RMN, la tétraboucle<sup>His</sup> (5'-GGU GCU UCG GCA CCC CA-3') a été synthétisée chimiquement avec l'aide de Jean-Louis Leroy (ICSN, Gif-sur-Yvette) sur un synthétiseur Amersham Biosciences LKB Gene Assembler, comme décrit précédemment (Snoussi *et al.*, 2001). Le rendement final était supérieur à 95%.

Pour préparer le duplex<sup>His</sup>, les oligoribonucléotides 5'-CGC ACU AGC CAG CCC A-3' et 5'-GCU GGC UAG UGC G-3' ont été synthétisés chimiquement comme la tétraboucle<sup>His</sup>, avec l'aide de Sandrine Caputo (ICSN, Gif-sur-Yvette). Les deux oligonucléotides ont été dialysés contre de l'eau. Leurs concentrations ont été calculées en utilisant des coefficients d'absorbance molaire à 260 nm de 0,185  $\mu\text{M}^{-1}\cdot\text{cm}^{-1}$  et 0,155  $\mu\text{M}^{-1}\cdot\text{cm}^{-1}$ , respectivement (Kibbe, 2007). Les deux oligonucléotides complémentaires ont été mélangés à une concentration de 100  $\mu\text{M}$  chacun, lyophilisés et conservés à -20°C. L'échantillon était ensuite dissous à une concentration finale de 1 mM, dans le tampon B (de RMN), incubé une heure à 60°C puis refroidi lentement pendant 12 h pour permettre l'hybridation des deux brins.

Pour aminoacyler ce duplex par l'histidyl-ARNt synthétase, nous l'avons d'abord phosphorylé à ses extrémités 5' et, pour cela, nous l'avons incubé pendant 30 min à 37°C en présence de 50 mM de Tris-HCl (pH 8.0), de 10 mM de  $\text{MgCl}_2$ , 5 mM de dithiothréitol, de 2 mM d'ATP et de 0,5 U/ $\mu\text{l}$  de polynucléotide kinase du phage T4 (Roche Applied Science). L'aminocyclation et l'acétylation de ce duplex ont été réalisées comme décrit ci-dessus pour l'ARNt<sup>His</sup>.

### Synthèse de la diacétyl-Lys-(3'NH)-adénosine

La diacétyl-Lys-(3'NH)-adénosine (3'-(L-[N,N-diacetyl-lysiny]amino-3'-deoxyadenosine) a été synthétisée par Gavin O'Mahony et Morten Grøtli (*Department of Chemistry, University of Gothenburg, Göteborg, Suède*), de la manière suivante (M. Grøtli, communication personnelle) :

La 3'-amino-3'-désoxyadénosine a été synthétisée comme décrit précédemment (Botta *et al.*, 1998) puis purifiée par une chromatographie sur colonne de gel de silice, éluée avec un mélange méthanol:acétate d'éthyle (12:88, v/v). Ce composé a été obtenu sous forme de poudre incolore avec un rendement global de 5,4 % (pour 10 étapes). La 3'-amino-3'-désoxyadénosine a ensuite été traitée avec du chlorure de *tert*-butyldiméthylsilyl (TBDMSiCl) (2,2 éq) dans de la pyridine (8 ml). La solution a été évaporée puis purifiée par chromatographie sur une colonne de gel de silice éluée avec un mélange acétate d'éthyle:hexane (7:3, v/v) pour obtenir la 3'-amino-2',5'-bis-O-(*tert*-butyldiméthyl-silyl)-3'-désoxyadénosine (avec 89% de rendement). La diacétyl-Lys-(3'NH)-adénosine a ensuite été obtenue par un protocole en deux étapes (O'Mahony *et al.*, 2007). Une liaison amide a d'abord été créée entre le composé obtenu précédemment et de la N,N-diacétyl-L-lysine commerciale (Bachem, Suisse) en utilisant, comme agents de couplage, le N-hydroxybenzotriazole (HOBt) et le 1-éthyl-3-(3-diméthylaminopropyl) carbodiimide (EDCI). Puis le composé ainsi obtenu a été

purifié par chromatographie sur une colonne de gel de silice (éluée avec un mélange CH<sub>2</sub>Cl<sub>2</sub>/méthanol 9:1, v/v) et déprotégé par du fluorure d'aluminium (10éq) dans le méthanol. Le solvant a été éliminé sous pression réduite et le résidu a été purifié sur une colonne préparative C18 HPLC (25mm), éluée avec un gradient d'eau (0%-65%) dans de l'acétonitrile contenant 0,1% d'acide trifluoroacétique. Le rendement total a été de 2,5% (pour 13 étapes). La structure du composé a été confirmée par des spectres RMN (<sup>1</sup>H et <sup>13</sup>C) et par spectrométrie de masse à haute résolution (rapport m/z [M+H<sup>+</sup>] calculé pour C<sub>19</sub>H<sub>28</sub>N<sub>8</sub>O<sub>6</sub> : 450,210 ; observé : 450,212). La pureté de la préparation a été confirmée par analyse élémentaire (pourcentages calculés pour C<sub>19</sub>H<sub>28</sub>N<sub>8</sub>O<sub>6</sub> : C 50,6 ; H 6,3 ; N 24,9. Observés : 50,8 ; H 6,6 ; N 24,9).

### Essais enzymatiques

Les mesures d'activité des diverses PTH ont été réalisées en incubant, pendant 10 min à 28°C, 0,1 à 0,3 μM des substrats radioactifs avec des quantités catalytiques de PTH (2-60 nM), dans 100 μl d'une solution (tampon C) contenant 20 mM de Tris-HCl (pH 7,5), 10 mM de MgCl<sub>2</sub>, 0,1 mM d'EDTA et 0,1 mM de dithiothréitol. Pour déterminer les valeurs des constantes k<sub>cat</sub> et K<sub>m</sub> des PTH sauvages et mutantes pour le diacétyl-Lys-ARNt<sup>Lys</sup>, les vitesses initiales d'hydrolyse ont été mesurées en présence de différentes concentrations de ce substrat (0,25 à 30 μM ; 50 Ci/mol). Pour comparer les vitesses initiales d'hydrolyse du diacétyl-[<sup>14</sup>C]lysyl-ARNt<sup>Lys</sup> ou du [<sup>14</sup>C]lysyl-ARNt<sup>Lys</sup> (50 Ci/mol), ces substrats ont été incubés à une concentration de 1,5 μM, en présence de quantités catalytiques (2-400 nM) de PTH sauvage, de PTH N10A ou de PTH N10D.

Pour la détermination des K<sub>i</sub>, des quantités variables de l'inhibiteur étudié (0,25 à 130 μM d'ARNt<sup>His</sup> ou de duplex<sup>His</sup>, 1 à 12 mM de diacétyl-Lys-(3'NH)-adénosine), ont été ajoutées au milieu réactionnel contenant, pour l'ARNt<sup>His</sup> et le duplex<sup>His</sup>, 0,15 à 12 μM de diacétyl-[<sup>14</sup>C]lysyl-ARNt<sup>Lys</sup> (50 Ci/mol) et 2 à 40 nM de PTH, et pour la diacétyl-Lys-(3'NH)-adénosine, 0,15 μM de diacétyl-[<sup>14</sup>C]lysyl-ARNt<sup>Lys</sup> (50 Ci/mol) et 2 nM de PTH. L'inhibiteur, le substrat et l'enzyme étaient préalablement resuspendus ou dilués dans le tampon C décrit ci-dessus ou dans le tampon B (tampon de la RMN). L'incubation a été réalisée dans le même tampon, pendant 10 min à 28°C ou 17°C (température de la RMN), respectivement.

Dans tous les cas, la réaction enzymatique a été arrêtée par addition de 340 μl d'éthanol froid, de 14 μl d'acétate de sodium (3 M, pH 4.8) et de 20 μl d'ARN entraîneur (ARN de levure à 4 mg/ml). Les échantillons ont ensuite été centrifugés. La radioactivité soluble (portée par l'acide aminé libéré), se retrouvant dans le surnageant, a été mesurée par comptage à scintillation, comme décrit précédemment (Dutka *et al.*, 1993).

Les valeurs de  $K_i$ ,  $K_m$  et  $k_{cat}$  ont été calculées par une méthode de régression non linéaire utilisant l'algorithme de Levenberg-Marquardt pour ajuster au mieux les données expérimentales à l'équation théorique de Michaelis (Dardel, 1994).

### Expériences de RMN

Les structures secondaires de la tétraboucle<sup>His</sup> et du duplex<sup>His</sup> ont été étudiées par RMN du proton à une dimension. Des spectres de la tétraboucle ont été enregistrés à une concentration de 0,5 mM à différents pH (entre 5,6 et 7,4) et à différentes températures (entre 6°C et 70°C), dans un tampon contenant 20 mM de phosphate de potassium et 100 mM de NaCl. Ces expériences de RMN ont été réalisées sous la conduite de Jean-Louis Leroy (ICSN, Gif-sur-Yvette) sur un spectromètre Varian Inova 500 MHz, en utilisant la séquence « *jump and return* » pour la suppression du signal de l'eau (Plateau and Guéron, 1982). Un spectre 1D du duplex a été enregistré à 130  $\mu$ M, dans les conditions de l'étude de son interaction avec la PTH, à 17°C dans le tampon B.

Pour la PTH, tous les spectres ont été enregistrés sur le spectromètre Bruker Advance II 600 MHz de l'antenne de l'ICSN à l'Ecole polytechnique, équipé d'une cryo-sonde TXI. Toutes les expériences ont été analysées à l'aide du logiciel Sparky (T. D. Goddard et D. G. Kneller, Sparky 3, University of California, San Francisco, USA). Des spectres <sup>15</sup>N-HSQC ont été enregistrés régulièrement tout au long des expériences de RMN (environ 40 jours au total). Aucune évolution du spectre n'a pu être notée, ce qui a permis de contrôler que la protéine ne s'était pas dégradée au cours de ces expériences. La liste des expériences de RMN réalisées, ainsi qu'un bref descriptif de chacune d'elle, sont présentés dans l'annexe 2.

L'attribution des fréquences de résonance des atomes du squelette de la PTH H20A a été réalisée à partir du résultat des expériences triple résonance HNCO, HNCA, HN(CO)CA, HNCACB, CBCA(CO)NH, HBHANH et HBHA(CO)NH. Les échantillons contenaient 150  $\mu$ M de protéine uniformément marquée au <sup>13</sup>C et au <sup>15</sup>N, dans le tampon B (50 mM acétate de sodium, pH 6,0, 200 mM NaCl, 90% H<sub>2</sub>O, 10% D<sub>2</sub>O). L'attribution des résonances de certains atomes des chaînes latérales des résidus de la protéine a été complétée grâce à une expérience HCCH-TOCSY enregistrée dans les mêmes conditions sur un échantillon dissous dans un tampon préparé en utilisant 100% de D<sub>2</sub>O. Une expérience <sup>15</sup>N-NOESY-HSQC avec un temps de mélange de 60 ms a aussi été enregistrée sur un échantillon de PTH marquée à l'azote <sup>15</sup>N, pour confirmer l'attribution de certains atomes d'hydrogène, établir la structure topologique de la PTH et attribuer les protons H<sup>δ</sup> des Asn et H<sup>ε</sup> des Gln. Enfin, pour analyser les corrélations entre les atomes d'hydrogène aromatiques et aliphatiques, une expérience <sup>13</sup>C-NOESY-HSQC tridimensionnelle a été enregistrée dans D<sub>2</sub>O sur un échantillon de PTH H20A marquée au <sup>13</sup>C, en optimisant l'observation des signaux aromatiques et aliphatiques.

L'interaction entre la PTH et la diacétyl-Lys-(3'NH)-adénosine a été caractérisée en enregistrant des spectres  $^{15}\text{N}$ -HSQC de la PTH (150  $\mu\text{M}$ ) en absence ou en présence de 150  $\mu\text{M}$  et 900  $\mu\text{M}$  de diacétyl-Lys-(3'NH)-adénosine. En pratique, la diacétyl-Lys-(3'NH)-adénosine a été aliquotée puis lyophilisée. Après avoir enregistré le spectre  $^{15}\text{N}$ -HSQC de référence (sans ligand), l'échantillon de PTH a été extrait du tube de RMN, utilisé pour dissoudre la quantité nécessaire de diacétyl-Lys-(3'NH)-adénosine lyophilisée avant d'être remis dans le tube de RMN. Une expérience  $^{15}\text{N}$ -NOESY-HSQC a aussi été enregistrée sur l'échantillon contenant 900  $\mu\text{M}$  de diacétyl-Lys-(3'NH)-adénosine.

L'interaction entre le PTH et l'ARNt<sup>His</sup> d'*E. coli* a été étudiée en enregistrant plusieurs spectres HSQC du même type que ci-dessus, dans le tampon B, avec 130  $\mu\text{M}$  de PTH ainsi que 0, 13 ou 260  $\mu\text{M}$  d'ARNt<sup>His</sup>. L'ARNt<sup>His</sup> a été ajouté dans l'échantillon à partir d'une solution concentrée, ou préalablement lyophilisée.

Pour l'étude de l'interaction entre la PTH et le duplex<sup>His</sup>, des aliquots de duplex<sup>His</sup> concentré ont été ajoutés progressivement avant chaque enregistrement de spectre HSQC. En pratique, chaque point de la titration a été obtenu en sortant l'échantillon précédent du tube de RMN (400  $\mu\text{L}$ ) et en le mélangeant à un petit volume d'ARN (5 à 12  $\mu\text{L}$  à 1 mM) avant de le replacer dans le tube. Sept expériences HSQC ont été enregistrées, avec des rapports ARN/protéine de 0, 0.1, 0.2, 0.4, 0.6, 0.8 et 1.

### Modélisation des complexes entre la PTH et les analogues de son substrat

Les coordonnées initiales des atomes de la PTH ont été extraites du fichier PDB 2PTH (structure cristallographique de la PTH d'*E. coli*). Les atomes d'hydrogènes ont été ajoutés et la structure ainsi obtenue a été rectifiée par minimisation d'énergie dans CNS en utilisant le champ de force *allhdg5.3*.

Les coordonnées de la diacétyl-Lys-(3'NH)-adénosine ont été construites en utilisant Pymol (W. L. DeLano, DeLano Scientific, San Carlo, USA) puis optimisées par minimisation d'énergie dans CNS. Les fichiers de topologie et de paramètres nécessaires ont été construits en se servant de termes extraits des champs de forces *dna-rna-allatom* (pour la partie adénosine) et *allhdg5.3* (pour la diacetyl-lysine).

Les coordonnées du duplex<sup>His</sup> ont été construites en utilisant le logiciel XDNA. En pratique, nous avons construit une hélice d'ADN de type A et, en utilisant Pymol, nous avons manuellement ajouté les groupements 2'-hydroxyles et enlevé les groupements 5-méthyles des thymine, manuellement. La géométrie de la molécule d'ARN obtenue a été rectifiée par une minimisation énergétique utilisant un script de CNS. Le diacétyl-Lys-duplex<sup>His</sup> a été construit en ajoutant la

diacétyl-lysine au duplex<sup>His</sup>, comme pour la construction de la diacétyl-Lys-(3'NH)-adénosine.

Les modèles du complexe entre la PTH H20A et la diacétyl-Lys-(3'NH)-adénosine ont été réalisés par Francois Bontems (ICSN) en utilisant une version autonome du logiciel HADDOCK (Dominguez *et al.*, 2003 ; van Dijk *et al.*, 2006) sur un cluster Macintosh de l'ICSN. Une version de la procédure d'arrimage (*docking*), optimisée pour les petites molécules, a été utilisée (Aurelien Thureau, communication personnelle). La première étape d'approche des deux molécules en corps rigide (it0) n'a pas été modifiée. Dans la seconde étape d'optimisation du complexe (it1), la première dynamique (en corps rigide) a été supprimée. En revanche le nombre de pas d'intégration des deux phases suivantes de recuit simulé a été augmenté (de 1000 à 5000 pas). La température initiale de ces deux phases a été abaissée (de 1000 à 500 K pour la première, de 500 à 300 K pour la seconde). Par ailleurs, la diacétyl-Lys-(3'NH)-adénosine a été laissée complètement flexible pendant ces deux étapes. 1000 complexes ont été construits, les 200 meilleurs ont été conservés à la fin de la procédure et ont été affinés en présence d'une couche de molécules d'eau.

Un ensemble de modèles du complexe entre la PTH et le duplex<sup>His</sup> a aussi été construit, par François Bontems, en utilisant d'un ensemble de scripts écrit pour CNS et en considérant que les résidus de la PTH H20A affectés au cours de la titration RMN par le duplex<sup>His</sup> et situés à la surface de la protéine sont en contact avec ce ligand. Pour chaque résidu affecté, une contrainte floue a été définie, stipulant qu'au moins un des atomes lourds (atomes n'étant pas des atomes d'hydrogène) de ce résidu devait être à une distance inférieure à 4 Å d'un atome lourd de l'ARN, en utilisant le potentiel *NOE soft-square* avec l'option *R-6 mean* et une constante de force de 10 kcal. mol<sup>-1</sup>. Å<sup>-2</sup>. Chaque complexe PTH-duplex a été calculé en deux étapes. Premièrement, le duplex d'ARN était disposé de manière aléatoire ou semi-aléatoire à une distance de la PTH l'empêchant d'être en contact avec la protéine (entre 10 et 30 Å). Ensuite, le complexe a été soumis à 3 étapes de dynamique cartésienne (à 300, 100 et 0K) en présence des contraintes intermoléculaires, en utilisant le champ de force standard *charmm 19* mais sans interactions électrostatiques. Pour éviter de trop larges déformations de la protéine et du duplex, les atomes lourds du squelette de la protéine ont été soumis à des contraintes harmoniques (100 kcal. mol<sup>-1</sup>. Å<sup>-2</sup> pour les résidus non affectés, et 10 kcal. mol<sup>-1</sup>. Å<sup>-2</sup> pour les résidus affectés) et la géométrie de l'ARN a été maintenue par un ensemble de contraintes de distances (pour les liaisons hydrogènes entre les bases appariées), planaires (pour les plans des bases) et d'angles dièdres (pour la géométrie du squelette sucre-phosphate), dérivées du script *dna-rna-restraint.def*. Deux séries de modèles ont été construites. Dans la première (disposition aléatoire), l'ARN était orienté aléatoirement et situé dans une demi-sphère en face de la surface d'interaction de la PTH, avant l'étape dynamique. Dans la seconde (disposition semi-aléatoire) l'ARN était situé comme dans la première série mais avec une orientation pré-définie, avec le bras CCA dirigé vers le site actif putatif de la PTH. A la fin, les modèles ont été évalués en utilisant deux critères, l'énergie évaluant le respect des contraintes floues et

l'énergie totale du complexe (calculée avec les termes de covalence, de van der Waals et d'électrostatique).

Les modèles de l'interaction entre la PTH et le diacétyl-Lys-duplex<sup>His</sup> ont été calculés en utilisant une série de contraintes avec une pénalité de 10 kcal.mol<sup>-1</sup>.Å<sup>-2</sup> à chaque fois que la distance imposée était dépassée. Nous avons introduit une contrainte de distance (de 3 Å) entre l'atome d'oxygène O<sup>5'</sup> du groupement 5'-phosphate de l'adénine 76 du substrat, et l'oxygène de la chaîne latérale de Y15. Nous avons aussi ajouté trois contraintes (de 4 Å entre les atomes C<sup>8</sup>, C<sup>4</sup> et N<sup>3</sup> de l'adénine A76 et les atomes C<sup>ζ</sup> ou C<sup>ε</sup>, et C<sup>γ</sup> du cycle aromatique de F66). Une autre contrainte stipulait qu'au moins un des atomes H<sup>δ</sup> de la chaîne latérale de l'asparagine N10 est proche (2 Å) soit du O<sup>2'</sup> (voir figure 58) du ribose 3'-terminal du substrat, soit de O<sup>3</sup> (voir figure 58), l'atome d'oxygène du groupement acétate porté par le groupement α-NH<sub>2</sub> de la lysine du substrat. Une autre contrainte (de 2 Å) visait à rapprocher au moins un des atomes H<sup>δ</sup> de la chaîne latérale de l'asparagine N68 et O<sup>1</sup>, l'atome d'oxygène du groupement oxyde de la fonction ester du ligand. Nous avons aussi imposé trois contraintes de distances (de 2 Å) entre l'atome d'oxygène de la molécule d'eau et au moins un des atomes H<sup>δ</sup> de la chaîne latérale de l'asparagine N114, entre le N<sup>ε</sup> de H20 et un atome d'hydrogène de cette molécule d'eau, et entre l'atome d'oxygène de cette molécule d'eau et l'atome de carbone du groupement carbonyle de la lysine. Nous avons de plus rajouté une contrainte de manière à ce que l'hydrogène H<sup>δ1</sup> de H20 soit proche (à 2 Å) d'un atome d'oxygène de la chaîne latérale du glutamate 93. Nous avons par ailleurs spécifié qu'un des atomes H<sup>n</sup> de R133 et un des atomes H<sup>ς</sup> de K105 (de l'extrémité de leurs chaînes latérales) devait former une pince autour du groupement 5' phosphate de la cytosine 1 du duplex, en étant situés à 2 Å d'un oxygène de ce groupement et en étant, par ailleurs, éloignés l'un de l'autre d'au moins 4 Å. Nous avons aussi forcé un des atomes H<sup>ς</sup> de K182 à être au voisinage d'un atome d'oxygène d'un des quatre derniers groupements phosphate de la région de l'ARN mimant la tige TΨC d'un ARNt. Les modèles ont été construits selon le même protocole que pour la modélisation du complexe entre la PTH et le duplex<sup>His</sup> en commençant avec une préorientation du substrat. Ils ont ensuite subi une relaxation sans contraintes qui a permis de minimiser leur énergie.

Le modèle de l'interaction entre la PTH et l'ARNt<sup>Phe</sup> a été construit en optimisant, grâce au logiciel Pymol, la superposition des phosphates de l'ARNt<sup>Phe</sup> (dont la structure est issue du fichier PDB (1TTT) où l'ARNt<sup>Phe</sup> est en complexe avec EF-Tu) et ceux du diacétyl-Lys-duplex<sup>His</sup> en complexe avec la PTH, dans le meilleur modèle obtenu lors de la modélisation précédente.



# **Bibliographie**





- Abelson, J., Trotta, C.R. and Li, H. (1998). tRNA splicing. *J. Biol. Chem.* **273**, 12685-12688.
- Ahel, I., Korencic, D., Ibba, M. and Söll, D. (2003). Trans-editing of mischarged tRNAs. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 15422-15427.
- Ahel, I., Stathopoulos, C., Ambrogelly, A., Sauerwald, A., Toogood, H., Hartsch, T. and Söll, D. (2002). Cysteine activation is an inherent in vitro property of prolyl-tRNA synthetases. *J. Biol. Chem.* **277**, 34743-34748.
- An, S. and Musier-Forsyth, K. (2004). Trans-editing of Cys-tRNA<sup>Pro</sup> by *Haemophilus influenzae* YbaK protein. *J. Biol. Chem.* **279**, 42359-42362.
- An, S. and Musier-Forsyth, K. (2005). Cys-tRNA<sup>Pro</sup> editing by *Haemophilus influenzae* YbaK via a novel synthetase:YbaK:tRNA ternary complex. *J. Biol. Chem.* **280**, 34465-34472.
- Arnez, J.G., Dock-Bregeon, A.C. and Moras, D. (1999). Glycyl-tRNA synthetase uses a negatively charged pit for specific recognition and activation of glycine. *J. Mol. Biol.* **286**, 1449-1459.
- Ataide, S.F. and Ibba, M. (2004). Discrimination of cognate and noncognate substrates at the active site of class II lysyl-tRNA synthetase. *Biochemistry* **43**, 11836-11841.
- Atherly, A.G. and Menninger, J.R. (1972). Mutant *E. coli* strain with temperature sensitive peptidyl-transfer RNA hydrolase. *Nat. New Biol.* **240**, 245-246.
- Bal, N.C., Agrawal, H., Meher, A.K. and Arora, A. (2007). Characterization of peptidyl-tRNA hydrolase encoded by open reading frame Rv1014c of *Mycobacterium tuberculosis* H37Rv. *Biol. Chem.* **388**, 467-479.
- Baldwin, A.N. and Berg, P. (1966). Transfer ribonucleic acid-induced hydrolysis of valyladenylate bound to isoleucyl ribonucleic acid synthetase. *J. Biol. Chem.* **241**, 839-845.
- Ban, N., Nissen, P., Hansen, J., Moore, P.B. and Steitz, T.A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289**, 905-920.
- Beebe, K., Merriman, E., Ribas De Pouplana, L. and Schimmel, P. (2004). A domain for editing by an archaebacterial tRNA synthetase. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 5958-5963.
- Beebe, K., Ribas De Pouplana, L. and Schimmel, P. (2003). Elucidation of tRNA-dependent editing by a class II tRNA synthetase and significance for cell viability. *EMBO J.* **22**, 668-675.

- Benne, R., Naaktgeboren, N., Gubbens, J. and Voorma, H.O. (1973). Recycling of initiation factors IF-1, IF-2 and IF-3. *Eur. J. Biochem.* **32**, 372-380.
- Betha, A.K., Williams, A.M. and Martinis, S.A. (2007). Isolated CP1 domain of *Escherichia coli* leucyl-tRNA synthetase is dependent on flanking hinge motifs for amino acid editing activity. *Biochemistry* **46**, 6258-6267.
- Beuning, P.J. and Musier-Forsyth, K. (2000). Hydrolytic editing by a class II aminoacyl-tRNA synthetase. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8916-8920.
- Beuning, P.J. and Musier-Forsyth, K. (2001). Species-specific differences in amino acid editing by class II prolyl-tRNA synthetase. *J. Biol. Chem.* **276**, 30779-30785.
- Blanquet S., Mechulam, Y., Schmitt, E. And Vial, L. (2003). Aminoacyl-Transfer RNA Maturation. In : *Translational Mechanism*, eds. Lapointe J. & Brakier-Gringas L., Landes Sciences, 65-79
- Botta, O., Moyroud, E., Lobato, C. and Strazewski, P. (1998). Synthesis of 3'-azido- and 3'-amino-3-deoxyadenosine in both enantiomeric forms. *Tetrahedron* **54**, 13529-13546.
- Bouadloun, F., Donner, D. and Kurland, C.G. (1983). Codon-specific missense errors in vivo. *EMBO J.* **2**, 1351-1356.
- Bult, C.J., White, O., Olsen, G.J., Zhou, L., Fleischmann, R.D., Sutton, G.G., Blake, J.A., FitzGerald, L.M., Clayton, R.A., Gocayne, J.D. *et al.* (1996). Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* **273**, 1058-1073.
- Calendar, R. and Berg, P. (1966). The catalytic properties of tyrosyl ribonucleic acid synthetases from *Escherichia coli* and *Bacillus subtilis*. *Biochemistry* **5**, 1690-1695.
- Calendar, R. and Berg, P. (1967). D-Tyrosyl RNA: formation, hydrolysis and utilization for protein synthesis. *J. Mol. Biol.* **26**, 39-54.
- Cerini, C., Kerjan, P., Astier, M., Gratecos, D., Mirande, M. and Semeriva, M. (1991). A component of the multisynthetase complex is a multifunctional aminoacyl-tRNA synthetase. *EMBO J.* **10**, 4267-4277.
- Chen, G.F. and Inouye, M. (1990). Suppression of the negative effect of minor arginine codons on gene expression; preferential usage of minor codons within the first 25 codons of the *Escherichia coli* genes. *Nucleic Acids Res.* **18**, 1465-1473.

- Chen, G.T. and Inouye, M. (1994). Role of the AGA/AGG codons, the rarest codons in global gene expression in *Escherichia coli*. *Genes Dev.* **8**, 2641-2652.
- Chen, J.F., Guo, N.N., Li, T., Wang, E.D. and Wang, Y.L. (2000). CP1 domain in *Escherichia coli* leucyl-tRNA synthetase is crucial for its editing function. *Biochemistry* **39**, 6726-6731.
- Consortium, I.H.G.S. (International Human Genome Sequencing), (2004). Finishing the euchromatic sequence of the human genome. *Nature* **431**, 931-945.
- Copeland, K.L., Anderson, J.A., Farley, A.R., Cox, J.R., and Tschumper, G.S. (2008). Probing phenylalanine/adenine pi-stacking interactions in protein complexes with explicitly correlated and CCSD(T) computations. *J. Phys. Chem. B* **112**, 14291-14295.
- Crepin, T., Yaremchuk, A., Tukalo, M. and Cusack, S. (2006). Structures of two bacterial prolyl-tRNA synthetases with and without a cis-editing domain. *Structure* **14**, 1511-1525.
- Cusack, S. (1995). Eleven down and nine to go. *Nat. Struct. Biol.* **2**, 824-831.
- Cusack, S., Berthet-Colominas, C., Hartlein, M., Nassar, N. and Leberman, R. (1990). A second class of synthetase structure revealed by X-ray analysis of *Escherichia coli* seryl-tRNA synthetase at 2.5 Å. *Nature* **347**, 249-255.
- Cusack, S., Hartlein, M. and Leberman, R. (1991). Sequence, structural and evolutionary relationships between class 2 aminoacyl-tRNA synthetases. *Nucleic Acids Res.* **19**, 3489-3498.
- Cusack, S., Yaremchuk, A. and Tukalo, M. (2000). The 2 Å crystal structure of leucyl-tRNA synthetase and its complex with a leucyl-adenylate analogue. *EMBO J.* **19**, 2351-2361.
- Cuzin, F., Kretchmer, N., Greenberg, R.E., Hurwitz, R. and Chapeville, F. (1967). Enzymatic hydrolysis of N-substituted aminoacyl-tRNA. *Proc. Natl. Acad. Sci. U.S.A.* **58**, 2079-2086.
- Czerwoniec, A., Dunin-Horkawicz, S., Purta, E., Kaminska, K.H., Kasprzak, J.M., Bujnicki, J.M., Grosjean, H. and Rother, K. (2009). MODOMICS: a database of RNA modification pathways. 2008 update. *Nucleic Acids Res.* **37**, D118-121.
- Dardel, F. (1994). MC-Fit: using Monte-Carlo methods to get accurate confidence limits on enzyme parameters. *Comput. Appl. Biosci.* **10**, 273-275.
- Das, G. and Varshney, U. (2006). Peptidyl-tRNA hydrolase and its critical role in protein biosynthesis. *Microbiology* **152**, 2191-2195.

De Pereda, J.M., Waas, W.F., Jan, Y., Ruoslahti, E., Schimmel, P. and Pascual, J. (2004). Crystal structure of a human peptidyl-tRNA hydrolase reveals a new fold and suggests basis for a bifunctional activity. *J. Biol. Chem.* **279**, 8111-8115.

Deutscher, M.P. (1982). tRNA nucleotidyltransferase. *The Enzymes* **15**, 183–215.

Dock-Bregeon, A., Sankaranarayanan, R., Romby, P., Caillet, J., Springer, M., Rees, B., Francklyn, C.S., Ehresmann, C. and Moras, D. (2000). Transfer RNA-mediated editing in threonyl-tRNA synthetase. The class II solution to the double discrimination problem. *Cell* **103**, 877-884.

Dock-Bregeon, A.C., Rees, B., Torres-Larios, A., Bey, G., Caillet, J. and Moras, D. (2004). Achieving error-free translation; the mechanism of proofreading of threonyl-tRNA synthetase at atomic resolution. *Mol. Cell* **16**, 375-386.

Domanski, M., Hertzog, M., Coutant, J., Gutsche-Perelroizen, I., Bontems, F., Carlier, M.F., Guittet, E. and van Heijenoort, C. (2004). Coupling of folding and binding of thymosin  $\beta$ 4 upon interaction with monomeric actin monitored by nuclear magnetic resonance. *J. Biol. Chem.* **279**, 23637-23645.

Dominguez, C., Boelens, R. and Bonvin, A.M. (2003). HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **125**, 1731-1737.

Dulic, M., Cvetic, N., Perona, J.J., and Gruic-Sovulj, I. (2010). Partitioning of tRNA-dependent editing between pre- and post-transfer pathways in class I aminoacyl-tRNA synthetases. *J. Biol. Chem.* **285**, 23799-23809.

Dutka, S., Meinnel, T., Lazennec, C., Mechulam, Y. and Blanquet, S. (1993). Role of the 1-72 base pair in tRNAs for the activity of *Escherichia coli* peptidyl-tRNA hydrolase. *Nucleic Acids Res.* **21**, 4025-4030.

Dwivedi, S., Kruparani, S.P. and Sankaranarayanan, R. (2005). A D-amino acid editing module coupled to the translational apparatus in archaea. *Nat. Struct. Mol. Biol.* **12**, 556-557.

Edwards M., Dharamsi A, Domagala M., Pinder B., Alam M.Z., Vedadi M., Wrezel O., Houston S., Kimber M., Vallee F., Awrey D. and Beattie B. (2002). Novel purified polypeptides from *Pseudomonas aeruginosa*. *International application published under the patent cooperation treaty (PCT)*. WO 03/0555904.

Edwards M., Dharamsi A, Vedadi M., Domagala M., Arrowsmith C., Mansoury K., Houston S., Richards D. and Beattie B. (2003). Novel purified polypeptides from *Enterococcus faecalis*. *International application published under the patent cooperation treaty (PCT)*. WO 2004/0114491.

- Eriani, G., Cavarelli, J., Martin, F., Ador, L., Rees, B., Thierry, J.C., Gangloff, J. and Moras, D. (1995). The class II aminoacyl-tRNA synthetases and their active site: evolutionary conservation of an ATP binding site. *J. Mol. Evol.* **40**, 499-508.
- Eriani, G., Delarue, M., Poch, O., Gangloff, J. and Moras, D. (1990). Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature* **347**, 203-206.
- Evans, D., Marquez, S.M. and Pace, N.R. (2006). RNase P: interface of the RNA and protein worlds. *Trends Biochem. Sci.* **31**, 333-341.
- Evans, S.V. (1993). SETOR: hardware-lighted three-dimensional solid model representations of macromolecules. *J. Mol. Graph.* **11**, 134-138, 127-138.
- Ferri-Fioni, M.L., Fromant, M., Bouin, A.P., Aubard, C., Lazennec, C., Plateau, P. and Blanquet, S. (2006). Identification in archaea of a novel D-Tyr-tRNA<sup>Tyr</sup> deacylase. *J. Biol. Chem.* **281**, 27575-27585.
- Ferri-Fioni, M.L., Schmitt, E., Soutourina, J., Plateau, P., Mechulam, Y. and Blanquet, S. (2001). Structure of crystalline D-Tyr-tRNA<sup>Tyr</sup> deacylase. A representative of a new class of tRNA-dependent hydrolases. *J. Biol. Chem.* **276**, 47285-47290.
- Fersht, A.R. (1977). Editing mechanisms in protein synthesis. Rejection of valine by the isoleucyl-tRNA synthetase. *Biochemistry* **16**, 1025-1030.
- Fersht, A.R. and Dingwall, C. (1979a). Cysteinyl-tRNA synthetase from *Escherichia coli* does not need an editing mechanism to reject serine and alanine. High binding energy of small groups in specific molecular interactions. *Biochemistry* **18**, 1245-1249.
- Fersht, A.R. and Dingwall, C. (1979b). An editing mechanism for the methionyl-tRNA synthetase in the selection of amino acids in protein synthesis. *Biochemistry* **18**, 1250-1256.
- Fersht, A.R. and Dingwall, C. (1979c). Establishing the misacylation/deacylation of the tRNA pathway for the editing mechanism of prokaryotic and eukaryotic valyl-tRNA synthetases. *Biochemistry* **18**, 1238-1245.
- Fersht, A.R. and Kaethner, M.M. (1976). Enzyme hyperspecificity. Rejection of threonine by the valyl-tRNA synthetase by misacylation and hydrolytic editing. *Biochemistry* **15**, 3342-3346.
- Francklyn, C. and Schimmel, P. (1990). Enzymatic aminoacylation of an eight-base-pair microhelix with histidine. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 8655-8659.

- Freist, W. (1989). Mechanisms of aminoacyl-tRNA synthetases: a critical consideration of recent results. *Biochemistry* **28**, 6787-6795.
- Freistroffer, D.V., Kwiatkowski, M., Buckingham, R.H. and Ehrenberg, M. (2000). The accuracy of codon recognition by polypeptide release factors. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2046-2051.
- Freistroffer, D.V., Pavlov, M.Y., MacDougall, J., Buckingham, R.H. and Ehrenberg, M. (1997). Release factor RF3 in *E. coli* accelerates the dissociation of release factors RF1 and RF2 from the ribosome in a GTP-dependent manner. *EMBO J.* **16**, 4126-4133.
- Friedman, M. (1999). Chemistry, nutrition, and microbiology of D-amino acids. *J. Agric. Food Chem.* **47**, 3457-3479.
- Frolova, L., Le Goff, X., Zhouravleva, G., Davydova, E., Philippe, M. and Kisselev, L. (1996). Eukaryotic polypeptide chain release factor eRF3 is an eRF1- and ribosome-dependent guanosine triphosphatase. *RNA* **2**, 334-341.
- Fromant, M., Ferri-Fioni, M.L., Plateau, P. and Blanquet, S. (2003). Peptidyl-tRNA hydrolase from *Sulfolobus solfataricus*. *Nucleic Acids Res.* **31**, 3227-3235.
- Fromant, M., Plateau, P. and Blanquet, S. (2000). Function of the extra 5'-phosphate carried by histidine tRNA. *Biochemistry* **39**, 4062-4067.
- Fromant, M., Plateau, P., Schmitt, E., Mechulam, Y. and Blanquet, S. (1999). Receptor site for the 5'-phosphate of elongator tRNAs governs substrate selection by peptidyl-tRNA hydrolase. *Biochemistry* **38**, 4982-4987.
- Fromant, M., Schmitt, E., Mechulam, Y., Lazennec, C., Plateau, P. and Blanquet, S. (2005). Crystal structure at 1.8 Å resolution and identification of active site residues of *Sulfolobus solfataricus* peptidyl-tRNA hydrolase. *Biochemistry* **44**, 4294-4301.
- Fukunaga, R., Fukai, S., Ishitani, R., Nureki, O. and Yokoyama, S. (2004). Crystal structures of the CP1 domain from *Thermus thermophilus* isoleucyl-tRNA synthetase and its complex with L-valine. *J. Biol. Chem.* **279**, 8396-8402.
- Fukunaga, R. and Yokoyama, S. (2005a). Crystal structure of leucyl-tRNA synthetase from the archaeon *Pyrococcus horikoshii* reveals a novel editing domain orientation. *J. Mol. Biol.* **346**, 57-71.
- Fukunaga, R. and Yokoyama, S. (2005b). Structural basis for non-cognate amino acid discrimination by the valyl-tRNA synthetase editing domain. *J. Biol. Chem.* **280**, 29937-29945.

- Fukunaga, R. and Yokoyama, S. (2006). Structural basis for substrate recognition by the editing domain of isoleucyl-tRNA synthetase. *J. Mol. Biol.* **359**, 901-912.
- Fukunaga, R. and Yokoyama, S. (2007). Structure of the AlaX-M trans-editing enzyme from *Pyrococcus horikoshii*. *Acta Crystallogr. D Biol. Crystallogr.* **63**, 390-400.
- Garcia-Villegas, M.R., De La Vega, F.M., Galindo, J.M., Segura, M., Buckingham, R.H. and Guarneros, G. (1991). Peptidyl-tRNA hydrolase is involved in lambda inhibition of host protein synthesis. *EMBO J.* **10**, 3549-3555.
- Giegé, R. (2003). Genetic code expansion. *Nat. Struct. Biol.* **10**, 414-416.
- Giegé R. and Frugier, M. (2003). Transfer RNA structure and identity. In : *Translational Mechanism*, eds. Lapointe J. & Brakier-Gringas, L., Landes Sciences, 1-24
- Giegé, R., Sissler, M. and Florentz, C. (1998). Universal rules and idiosyncratic features in tRNA identity. *Nucleic Acids Res.* **26**, 5017-5035.
- Gong, M., Cruz-Vera, L.R. and Yanofsky, C. (2007). Ribosome recycling factor and release factor 3 action promotes TnaC-peptidyl-tRNA Dropoff and relieves ribosome stalling during tryptophan induction of *tna* operon expression in *Escherichia coli*. *J. Bacteriol.* **189**, 3147-3155.
- Goodall, J.J., Chen, G.J. and Page, M.G. (2004). Essential role of histidine 20 in the catalytic mechanism of *Escherichia coli* peptidyl-tRNA hydrolase. *Biochemistry* **43**, 4583-4591.
- Goodsell, D.S. (1991). Inside a living cell. *Trends Biochem. Sci.* **16**, 203-206.
- Green, R. and Noller, H.F. (1997). Ribosomes and translation. *Annu. Rev. Biochem.* **66**, 679-716.
- Gross, M., Crow, P., and White, J. (1992a). The site of hydrolysis by rabbit reticulocyte peptidyl-tRNA hydrolase is the 3'-AMP terminus of susceptible tRNA substrates. *J. Biol. Chem.* **267**, 2080-2086.
- Gross, M., Starn, T.K., Rundquist, C., Crow, P., White, J., Olin, A., and Wagner, T. (1992b). Purification and initial characterization of peptidyl-tRNA hydrolase from rabbit reticulocytes. *J. Biol. Chem.* **267**, 2073-2079.
- Gruic-Sovulj, I., Uter, N., Bullock, T. and Perona, J.J. (2005). tRNA-dependent aminoacyl-adenylate hydrolysis by a nonediting class I aminoacyl-tRNA synthetase. *J. Biol. Chem.* **280**, 23978-23986.



- Guillon, J.M., Heiss, S., Soutourina, J., Mechulam, Y., Laalami, S., Grunberg-Manago, M. and Blanquet, S. (1996). Interplay of methionine tRNAs with translation elongation factor Tu and translation initiation factor 2 in *Escherichia coli*. *J. Biol. Chem.* **271**, 22321-22325.
- Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35**, 849-857.
- Guillon, L., Schmitt, E., Blanquet, S. and Mechulam, Y. (2005). Initiator tRNA binding by e/aIF5B, the eukaryotic/archaeal homologue of bacterial initiation factor IF2. *Biochemistry* **44**, 15594-15601.
- Guo, M., Chong, Y.E., Beebe, K., Shapiro, R., Yang, X.L. and Schimmel, P. (2009). The C-Ala domain brings together editing and aminoacylation functions on one tRNA. *Science* **325**, 744-747.
- Hagiwara, Y., Field, M.J., Nureki, O. and Tatenno, M. (2010). Editing mechanism of aminoacyl-tRNA synthetases operates by a hybrid ribozyme/protein catalyst. *J. Am. Chem. Soc.* **132**, 2751-2758.
- Hartmann, R.K., Gossringer, M., Spath, B., Fischer, S. and Marchfelder, A. (2009). The making of tRNAs and more - RNase P and tRNase Z. *Prog. Mol. Biol. Transl. Sci.* **85**, 319-368.
- Hati, S., Ziervogel, B., Sternjohn, J., Wong, F.C., Nagan, M.C., Rosen, A.E., Siliciano, P.G., Chihade, J.W. and Musier-Forsyth, K. (2006). Pre-transfer editing by class II prolyl-tRNA synthetase: role of aminoacylation active site in "selective release" of noncognate amino acids. *J. Biol. Chem.* **281**, 27862-27872.
- Hecht, L.I., Zamecnik, P.C., Stephenson, M.L., and Scott, J.F. (1958). Nucleoside tri-phosphates as precursors of ribonucleic acid end groups in a mammalian system. *J. Biol. Chem.* **233**, 954-963.
- Hendrickson, T.L., Nomanbhoy, T.K., de Crecy-Lagard, V., Fukai, S., Nureki, O., Yokoyama, S. and Schimmel, P. (2002). Mutational separation of two pathways for editing by a class I tRNA synthetase. *Mol. Cell* **9**, 353-362.
- Hendrickson, T.L., Nomanbhoy, T.K. and Schimmel, P. (2000). Errors from selective disruption of the editing center in a tRNA synthetase. *Biochemistry* **39**, 8180-8186.
- Hendrickson, T. L. and Schimmel, P. (2003). In : *Translational Mechanism*, eds. Lapointe J. & Brakier-Gringas, L., Landes Sciences, p34-64
- Heurgué-Hamard, V., Karimi, R., Mora, L., MacDougall, J., Leboeuf, C., Grentzmann, G., Ehrenberg, M. and Buckingham, R.H. (1998). Ribosome release

factor RF4 and termination factor RF3 are involved in dissociation of peptidyl-tRNA from the ribosome. *EMBO J.* **17**, 808-816.

Heurgué-Hamard, V., Mora, L., Guarneros, G. and Buckingham, R.H. (1996). The growth defect in *Escherichia coli* deficient in peptidyl-tRNA hydrolase is due to starvation for Lys-tRNA<sup>Lys</sup>. *EMBO J.* **15**, 2826-2833.

Hirel, P.H., Leveque, F., Mellot, P., Dardel, F., Panvert, M., Mechulam, Y. and Fayat, G. (1988). Genetic engineering of methionyl-tRNA synthetase: in vitro regeneration of an active synthetase by proteolytic cleavage of a methionyl-tRNA synthetase-beta-galactosidase chimeric protein. *Biochimie* **70**, 773-782.

Hoagland, M.B., Stephenson, M.L., Scott, J.F., Hecht, L.I. and Zamecnik, P.C. (1958). A soluble ribonucleic acid intermediate in protein synthesis. *J. Biol. Chem.* **231**, 241-257.

Holm, L. and Sander, C. (1997). An evolutionary treasure: unification of a broad set of amidohydrolases related to urease. *Proteins* **28**, 72-82.

Hopfield, J.J. (1974). Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc. Natl. Acad. Sci. U.S.A.* **71**, 4135-4139.

Hountondji, C., Dessen, P. and Blanquet, S. (1986). Sequence similarities among the family of aminoacyl-tRNA synthetases. *Biochimie* **68**, 1071-1078.

Hountondji, C., Lazennec, C., Beauvallet, C., Dessen, P., Pernollet, J.C., Plateau, P. and Blanquet, S. (2002). Crucial role of conserved lysine 277 in the fidelity of tRNA aminoacylation by *Escherichia coli* valyl-tRNA synthetase. *Biochemistry* **41**, 14856-14865.

Hussain, T., Kruparani, S.P., Pal, B., Dock-Bregeon, A.C., Dwivedi, S., Shekar, M.R., Sureshbabu, K. and Sankaranarayanan, R. (2006). Post-transfer editing mechanism of a D-aminoacyl-tRNA deacylase-like domain in threonyl-tRNA synthetase from archaea. *EMBO J.* **25**, 4152-4162.

Ibba, M., Kast, P. and Hennecke, H. (1994). Substrate specificity is determined by amino acid binding pocket size in *Escherichia coli* phenylalanyl-tRNA synthetase. *Biochemistry* **33**, 7107-7112.

Ibba, M. and Söll, D. (1999). Quality control mechanisms during translation. *Science* **286**, 1893-1897.

Ibba, M. and Söll, D. (2000). Aminoacyl-tRNA synthesis. *Annu. Rev. Biochem.* **69**, 617-650.

Jakubowski, H. (1991). Proofreading in vivo: editing of homocysteine by methionyl-tRNA synthetase in the yeast *Saccharomyces cerevisiae*. *EMBO J.* **10**, 593-598.

Jakubowski, H. (1997). Aminoacyl thioester chemistry of class II aminoacyl-tRNA synthetases. *Biochemistry* **36**, 11077-11085.

Jakubowski, H. (1999). Misacylation of tRNA<sup>Lys</sup> with noncognate amino acids by lysyl-tRNA synthetase. *Biochemistry* **38**, 8088-8093.

Jakubowski, H. and Fersht, A.R. (1981). Alternative pathways for editing non-cognate amino acids by aminoacyl-tRNA synthetases. *Nucleic Acids Res.* **9**, 3105-3117.

Jan, Y., Matter, M., Pai, J.T., Chen, Y.L., Pilch, J., Komatsu, M., Ong, E., Fukuda, M. and Ruoslahti, E. (2004). A mitochondrial protein, Bit1, mediates apoptosis regulated by integrins and Groucho/TLE corepressors. *Cell* **116**, 751-762.

Jenkins, B.D. and Barkan, A. (2001). Recruitment of a peptidyl-tRNA hydrolase as a facilitator of group II intron splicing in chloroplasts. *EMBO J.* **20**, 872-879.

Kairouz-Wahbe, R., Biliran, H., Luo, X., Khor, I., Wankell, M., Besch-Williford, C., Pascual, J., Oshima, R. and Ruoslahti, E. (2008). Anoikis effector Bit1 negatively regulates Erk activity. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 1528-1532.

Kaji A. and Hirokawa G. (2000) Disassembly of post termination complex by RRF (ribosome recycling factor), a possible new target for antimicrobial agents. In *The Ribosome: Structure, Function, Antibiotics and Cellular Interactions*. eds. Garret,R.A., Douthwaite,S.R., Liljas,A., Matheson,A.T., Moore,P.B. and Noller,H.F., American Society of Microbiology Press, Washington, DC, 527-539.

Karimi, R., Pavlov, M.Y., Heurgué-Hamard, V., Buckingham, R.H. and Ehrenberg, M. (1998). Initiation factors IF1 and IF2 synergistically remove peptidyl-tRNAs with short polypeptides from the P-site of translating *Escherichia coli* ribosomes. *J. Mol. Biol.* **281**, 241-252.

Kibbe, W.A. (2007). OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res.* **35**, W43-46.

Komine, M., Freedberg, I.M. and Blumenberg, M. (1994). Interleukin-1 alpha is released during transfection of keratinocytes. *J. Invest. Dermatol.* **103**, 580-582.

Korencic, D., Ahel, I., Schelert, J., Sacher, M., Ruan, B., Stathopoulos, C., Blum, P., Ibba, M. and Söll, D. (2004). A freestanding proofreading domain is required for protein synthesis quality control in Archaea. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 10260-10265.

- Kössel, H. and RajBhandary, U.L. (1968). Studies on polynucleotides. LXXXVI. Enzymic hydrolysis of N-acylaminoacyl-transfer RNA. *J. Mol. Biol.* **35**, 539-560.
- Kotik-Kogan, O., Moor, N., Tworowski, D. and Safro, M. (2005). Structural basis for discrimination of L-phenylalanine from L-tyrosine by phenylalanyl-tRNA synthetase. *Structure* **13**, 1799-1807.
- Kurland, C. G., Jørgensen, F., Richter, A., Ehrenberg, M., Bilgin, N. and Rojas, A.-M. (1990) In: *The Ribosome – Structure, Function, and Evolution*, eds. Dahlberg, A., Hill, W. E., Garrett, R. A., Moore, P. B., Schlessinger, D. and Warner, J. R., Amer. Soc. Microbiol., Washington, DC. 513–526,
- Landes, C., Perona, J.J., Brunie, S., Rould, M.A., Zelwer, C., Steitz, T.A. and Risler, J.L. (1995). A structure-based multiple sequence alignment of all class I aminoacyl-tRNA synthetases. *Biochimie* **77**, 194-203.
- Lim, K., Tempczyk, A., Bonander, N., Toedt, J., Howard, A., Eisenstein, E. and Herzberg, O. (2003). A catalytic mechanism for D-Tyr-tRNA<sup>Tyr</sup> deacylase based on the crystal structure of *Hemophilus influenzae* HI0670. *J. Biol. Chem.* **278**, 13496-13502.
- Lin, L. and Schimmel, P. (1996). Mutational analysis suggests the same design for editing activities of two tRNA synthetases. *Biochemistry* **35**, 5596-5601.
- Lincecum, T.L., Jr., Tukalo, M., Yaremchuk, A., Mursinna, R.S., Williams, A.M., Sproat, B.S., Van Den Eynde, W., Link, A., Van Calenbergh, S., Grotli, M. *et al.* (2003). Structural and mechanistic basis of pre- and posttransfer editing by leucyl-tRNA synthetase. *Mol. Cell* **11**, 951-963.
- Ling, J., Roy, H. and Ibba, M. (2007). Mechanism of tRNA-dependent editing in translational quality control. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 72-77.
- Liu, Y., Liao, J., Zhu, B., Wang, E.D. and Ding, J. (2006). Crystal structures of the editing domain of *Escherichia coli* leucyl-tRNA synthetase and its complexes with Met and Ile reveal a lock-and-key mechanism for amino acid discrimination. *Biochem. J.* **394**, 399-407.
- Lodish, H.F. and Jacobsen, M. (1972). Regulation of hemoglobin synthesis. Equal rates of translation and termination of  $\alpha$ - and  $\beta$ -globin chains. *J. Biol. Chem.* **247**, 3622-3629.
- Loftfield, R.B. (1963). The frequency of errors in protein biosynthesis. *Biochem. J.* **89**, 82-92.
- Loftfield, R.B. and Vanderjagt, D. (1972). The frequency of errors in protein biosynthesis. *Biochem. J.* **128**, 1353-1356.

- Ludmerer, S.W. and Schimmel, P. (1987). Gene for yeast glutamine tRNA synthetase encodes a large amino-terminal extension and provides a strong confirmation of the signature sequence for a group of the aminoacyl-tRNA synthetases. *J. Biol. Chem.* **262**, 10801-10806.
- Luo, S., and Levine, R.L. (2009). Methionine in proteins defends against oxidative stress. *FASEB J.* **23**, 464-472.
- Martinis, S.A. and Schimmel, P. (1992). Enzymatic aminoacylation of sequence-specific RNA minihelices and hybrid duplexes with methionine. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 65-69.
- Menninger, J.R. (1976). Peptidyl transfer RNA dissociates during protein synthesis from ribosomes of *Escherichia coli*. *J. Biol. Chem.* **251**, 3392-3398.
- Menninger, J.R. (1979). Accumulation of peptidyl tRNA is lethal to *Escherichia coli*. *J. Bacteriol.* **137**, 694-696.
- Milligan, J.F. and Uhlenbeck, O.C. (1989). Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol.* **180**, 51-62.
- Moras, D. (1992). Structural and functional relationships between aminoacyl-tRNA synthetases. *Trends Biochem. Sci.* **17**, 159-164.
- Musier-Forsyth, K. and Schimmel, P. (1993). Aminoacylation of RNA oligonucleotides: minimalist structures and origin of specificity. *FASEB J.* **7**, 282-289.
- Netzer, N., Goodenbour, J.M., David, A., Dittmar, K.A., Jones, R.B., Schneider, J.R., Boone, D., Eves, E.M., Rosner, M.R., Gibbs, J.S., *et al.* (2009). Innate immune and chemically triggered oxidative stress modifies translational fidelity. *Nature* **462**, 522-526.
- Newberry, K.J., Hou, Y.M. and Perona, J.J. (2002). Structural origins of amino acid selection without editing by cysteinyl-tRNA synthetase. *EMBO J.* **21**, 2778-2787.
- Ninio, J. (1975). Kinetic amplification of enzyme discrimination. *Biochimie* **57**, 587-595.
- Nissen, P., Kjeldgaard, M. and Nyborg, J. (2000). Macromolecular mimicry. *EMBO J.* **19**, 489-495.
- Nissen, P., Kjeldgaard, M., Thirup, S., Polekhina, G., Reshetnikova, L., Clark, B.F. and Nyborg, J. (1995). Crystal structure of the ternary complex of Phe-tRNA<sup>Phe</sup>, EF-Tu, and a GTP analog. *Science* **270**, 1464-1472.

- Nordin, B.E. and Schimmel, P. (2002). Plasticity of recognition of the 3'-end of mischarged tRNA by class I aminoacyl-tRNA synthetases. *J. Biol. Chem.* **277**, 20510-20517.
- Nozawa, K., O'Donoghue, P., Gundllapalli, S., Araiso, Y., Ishitani, R., Umehara, T., Söll, D. and Nureki, O. (2009). Pyrrolysyl-tRNA synthetase-tRNA<sup>Pyl</sup> structure reveals the molecular basis of orthogonality. *Nature* **457**, 1163-1167.
- O'Mahony, G., Sundgren, A., Svensson, S. and Grøtli, M. (2007). A practical synthesis of 2'-aminoacylamino-2'-deoxyadenosines. *Tetrahedron* **63**, 6901-6908.
- Ogle, J.M. and Ramakrishnan, V. (2005). Structural insights into translational fidelity. *Annu. Rev. Biochem.* **74**, 129-177.
- Orgel, L.E. (1963). The maintenance of the accuracy of protein synthesis and its relevance to ageing. *Proc. Natl. Acad. Sci. U.S.A.* **49**, 517-521.
- Orgel, L.E. (1973). Ageing of clones of mammalian cells. *Nature* **243**, 441-445.
- Ostheimer, G.J., Williams-Carrier, R., Belcher, S., Osborne, E., Gierke, J. and Barkan, A. (2003). Group II intron splicing factors derived by diversification of an ancient RNA-binding domain. *EMBO J.* **22**, 3919-3929.
- Pack, G.R., Wong, L. and Lamm, G., (1998). pK<sub>a</sub> of cytosine on the third strand of triplex DNA : preliminary Poisson-Boltzman calculations. *Int. J. Q. Chem.* **70**, 1177-1184.
- Palmiter, R.D. (1974). Magnesium precipitation of ribonucleoprotein complexes. Expedient techniques for the isolation of undergraded polysomes and messenger ribonucleic acid. *Biochemistry* **13**, 3606-3615.
- Pauling., L. (1957) The probability of errors in the process of synthesis of protein molecules. In : Festschrift für Pr. Dr. Arthur Stoll, Birkhäuser-Verlag, Basel 597-602
- Phizicky, E.M. (2005). Have tRNA, will travel. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 11127-11128.
- Pisarev, A.V., Hellen, C.U. and Pestova, T.V. (2007). Recycling of eukaryotic posttermination ribosomal complexes. *Cell* **131**, 286-299.
- Plateau, P. and Blanquet, S. (1982). Zinc-dependent synthesis of various dinucleoside 5',5'''-P<sup>1</sup>,P<sup>3</sup>-tri- or 5',5'''-P<sup>1</sup>,P<sup>4</sup>-tetrphosphates by *Escherichia coli* lysyl-tRNA synthetase. *Biochemistry* **21**, 5273-5279.

Powers, R., Mirkovic, N., Goldsmith-Fischman, S., Acton, T.B., Chiang, Y., Huang, Y.J., Ma, L., Rajan, P.K., Cort, J.R., Kennedy, M.A. *et al.* (2005). Solution structure of *Archaeoglobus fulgidis* peptidyl-tRNA hydrolase (Pth2) provides evidence for an extensive conserved family of Pth2 enzymes in archaea, bacteria, and eukaryotes. *Protein Sci.* **14**, 2849-2861.

Pulavarti, S.V., Jain, A., Pathak, P.P., Mahmood, A. and Arora, A. (2008). Solution structure and dynamics of peptidyl-tRNA hydrolase from *Mycobacterium tuberculosis* H37Rv. *J. Mol. Biol.* **378**, 165-177.

Rao, A.R. and Varshney, U. (2001). Specific interaction between the ribosome recycling factor and the elongation factor G from *Mycobacterium tuberculosis* mediates peptidyl-tRNA release and ribosome recycling in *Escherichia coli*. *EMBO J.* **20**, 2977-2986.

Redko, Y., Li de Lasierra-Gallay, I. and Condon, C. (2007). When all's zed and done: the structure and function of RNase Z in prokaryotes. *Nat. Rev. Microbiol.* **5**, 278-286.

Reynolds, N.M., Ling, J., Roy, H., Banerjee, R., Repasky, S.E., Hamel, P., and Ibba, M. (2010). Cell-specific differences in the requirements for translation quality control. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 4063-4068.

Ribas de Pouplana, L. and Schimmel, P. (2001). Aminoacyl-tRNA synthetases: potential markers of genetic code development. *Trends Biochem. Sci.* **26**, 591-596.

Rock, F.L., Mao, W., Yaremchuk, A., Tukalo, M., Crepin, T., Zhou, H., Zhang, Y.K., Hernandez, V., Akama, T., Baker, S.J. *et al.* (2007). An antifungal agent inhibits an aminoacyl-tRNA synthetase by trapping tRNA in the editing site. *Science* **316**, 1759-1761.

Rosas-Sandoval, G., Ambrogelly, A., Rinehart, J., Wei, D., Cruz-Vera, L.R., Graham, D.E., Stetter, K.O., Guarneros, G. and Söll, D. (2002). Orthologs of a novel archaeal and of the bacterial peptidyl-tRNA hydrolase are nonessential in yeast. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 16707-16712.

Roy, H. and Ibba, M. (2006). Molecular biology: sticky end in protein synthesis. *Nature* **443**, 41-42.

Roy, H., Ling, J., Alfonzo, J. and Ibba, M. (2005). Loss of editing activity during the evolution of mitochondrial phenylalanyl-tRNA synthetase. *J. Biol. Chem.* **280**, 38186-38192.

Roy, H., Ling, J., Irnov, M. and Ibba, M. (2004). Post-transfer editing in vitro and in vivo by the beta subunit of phenylalanyl-tRNA synthetase. *EMBO J.* **23**, 4639-4648.

- Rozenki, J., Crain, P.F. and McCloskey, J.A. (1999). The RNA Modification Database: 1999 update. *Nucleic Acids Res.* **27**, 196-197.
- Ruan, B. and Söll, D. (2005). The bacterial YbaK protein is a Cys-tRNA<sup>Pro</sup> and Cys-tRNA<sup>Cys</sup> deacylase. *J. Biol. Chem.* **280**, 25887-25891.
- Sankaranarayanan, R., Dock-Bregeon, A.C., Rees, B., Bovee, M., Caillet, J., Romby, P., Francklyn, C.S. and Moras, D. (2000). Zinc ion mediated amino acid discrimination by threonyl-tRNA synthetase. *Nat. Struct. Biol.* **7**, 461-465.
- Sankaranarayanan, R., Dock-Bregeon, A.C., Romby, P., Caillet, J., Springer, M., Rees, B., Ehresmann, C., Ehresmann, B. and Moras, D. (1999). The structure of threonyl-tRNA synthetase-tRNA<sup>Thr</sup> complex enlightens its repressor activity and reveals an essential zinc ion in the active site. *Cell* **97**, 371-381.
- Sasaki, H.M., Sekine, S., Sengoku, T., Fukunaga, R., Hattori, M., Utsunomiya, Y., Kuroishi, C., Kuramitsu, S., Shirouzu, M. and Yokoyama, S. (2006). Structural and mutational studies of the amino acid-editing domain from archaeal/eukaryal phenylalanyl-tRNA synthetase. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 14744-14749.
- Sauerwald, A., Zhu, W., Major, T.A., Roy, H., Palioura, S., Jahn, D., Whitman, W.B., Yates, J.R., 3rd, Ibba, M. and Söll, D. (2005). RNA-dependent cysteine biosynthesis in archaea. *Science* **307**, 1969-1972.
- Schlutzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F. and Yonath, A. (2000). Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. *Cell* **102**, 615-623.
- Schmidt, E. and Schimmel, P. (1994). Mutational isolation of a sieve for editing in a transfer RNA synthetase. *Science* **264**, 265-267.
- Schmidt, E. and Schimmel, P. (1995). Residues in a class I tRNA synthetase which determine selectivity of amino acid recognition in the context of tRNA. *Biochemistry* **34**, 11204-11210.
- Schmitt, E., Fromant, M., Plateau, P., Mechulam, Y. and Blanquet, S. (1997a). Crystallization and preliminary X-ray analysis of *Escherichia coli* peptidyl-tRNA hydrolase. *Proteins* **28**, 135-136.
- Schmitt, E., Mechulam, Y., Fromant, M., Plateau, P. and Blanquet, S. (1997b). Crystal structure at 1.2 Å resolution and active site mapping of *Escherichia coli* peptidyl-tRNA hydrolase. *EMBO J.* **16**, 4760-4769.
- Schulman, L.H. and Pelka, H. (1975). The structural basis for the resistance of *Escherichia coli* formylmethionyl transfer ribonucleic acid to cleavage by



*Escherichia coli* peptidyl transfer ribonucleic acid hydrolase. *J. Biol. Chem.* **250**, 542-547.

Selvaraj, M., Roy, S., Singh, N.S., Sangeetha, R., Varshney, U. and Vijayan, M. (2007). Structural plasticity and enzyme action: crystal structures of *Mycobacterium tuberculosis* peptidyl-tRNA hydrolase. *J. Mol. Biol.* **372**, 186-193.

Sharma, N., Furter, R., Kast, P. and Tirrell, D.A. (2000). Efficient introduction of aryl bromide functionality into proteins in vivo. *FEBS Lett.* **467**, 37-40.

Sheoran, A. and First, E.A. (2008). Activation of D-tyrosine by *Bacillus stearothermophilus* tyrosyl-tRNA synthetase: 2. Cooperative binding of ATP is limited to the initial turnover of the enzyme. *J. Biol. Chem.* **283**, 12971-12980.

Sheoran, A., Sharma, G. and First, E.A. (2008). Activation of D-tyrosine by *Bacillus stearothermophilus* tyrosyl-tRNA synthetase: 1. Pre-steady-state kinetic analysis reveals the mechanistic basis for the recognition of D-tyrosine. *J. Biol. Chem.* **283**, 12960-12970.

Shi, H. and Moore, P.B. (2000). The crystal structure of yeast phenylalanine tRNA at 1.93 Å resolution: a classic structure revisited. *RNA* **6**, 1091-1105.

Shi, J.P., Martinis, S.A. and Schimmel, P. (1992). RNA tetraloops as minimalist substrates for aminoacylation. *Biochemistry* **31**, 4931-4936.

Shiloach, J., Bauer, S., de Groot, N. and Lapidot, Y. (1975a). The influence of the peptide chain length on the activity of peptidyl-tRNA hydrolase from *E. coli*. *Nucleic Acids Res.* **2**, 1941-1950.

Shiloach, J., Bauer, S. and Lapidot, Y. (1975b). Proceedings: Role of the peptide moiety in the interaction between peptidyl-tRNA and peptidyl-tRNA hydrolase. *Isr. J. Med. Sci.* **11**, 1195-1196.

Shiloach, J., Lapidot, Y. and de Groot, N. (1975c). The specificity of peptidyl-tRNA hydrolase from *E. coli*. *FEBS Lett.* **57**, 130-133.

Shimizu, K., Kuroishi, C., Sugahara, M. and Kunishima, N. (2008). Structure of peptidyl-tRNA hydrolase 2 from *Pyrococcus horikoshii* OT3: insight into the functional role of its dimeric state. *Acta Crystallogr. D Biol. Crystallogr.* **64**, 444-453.

Shimizu, S., Juan, E.C., Sato, Y., Miyashita, Y., Hoque, M.M., Suzuki, K., Sagara, T., Tsunoda, M., Sekiguchi, T., Dock-Bregeon, A.C. *et al.* (2009). Two complementary enzymes for threonylation of tRNA in crenarchaeota: crystal structure of *Aeropyrum pernix* threonyl-tRNA synthetase lacking a cis-editing domain. *J. Mol. Biol.* **394**, 286-296.

- Sickmann, A., Reinders, J., Wagner, Y., Joppich, C., Zahedi, R., Meyer, H.E., Schonfisch, B., Perschil, I., Chacinska, A., Guiard, B. *et al.* (2003). The proteome of *Saccharomyces cerevisiae* mitochondria. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 13207-13212.
- Singh, N.S., Ahmad, R., Sangeetha, R. and Varshney, U. (2008). Recycling of ribosomal complexes stalled at the step of elongation in *Escherichia coli*. *J. Mol. Biol.* **380**, 451-464.
- Singh, N.S., Das, G., Seshadri, A., Sangeetha, R. and Varshney, U. (2005). Evidence for a role of initiation factor 3 in recycling of ribosomal complexes stalled on mRNAs in *Escherichia coli*. *Nucleic Acids Res.* **33**, 5591-5601.
- Singh, N.S. and Varshney, U. (2004). A physiological connection between tmRNA and peptidyl-tRNA hydrolase functions in *Escherichia coli*. *Nucleic Acids Res.* **32**, 6028-6037.
- Snoussi, K., Nonin-Lecomte, S. and Leroy, J.L. (2001). The RNA i-motif. *J. Mol. Biol.* **309**, 139-153.
- Sokabe, M., Okada, A., Yao, M., Nakashima, T. and Tanaka, I. (2005). Molecular basis of alanine discrimination in editing site. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 11669-11674.
- Sokabe, M., Ose, T., Nakamura, A., Tokunaga, K., Nureki, O., Yao, M. and Tanaka, I. (2009). The structure of alanyl-tRNA synthetase with editing domain. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 11028-11033.
- Soutourina, J., Blanquet, S. and Plateau, P. (2000a). D-tyrosyl-tRNA<sup>Tyr</sup> metabolism in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **275**, 11626-11630.
- Soutourina, J., Plateau, P. and Blanquet, S. (2000b). Metabolism of D-aminoacyl-tRNAs in *Escherichia coli* and *Saccharomyces cerevisiae* cells. *J. Biol. Chem.* **275**, 32535-32542.
- Soutourina, J., Plateau, P., Delort, F., Peirottes, A. and Blanquet, S. (1999). Functional characterization of the D-Tyr-tRNA<sup>Tyr</sup> deacylase from *Escherichia coli*. *J. Biol. Chem.* **274**, 19109-19114.
- Soutourina, O., Soutourina, J., Blanquet, S. and Plateau, P. (2004). Formation of D-tyrosyl-tRNA<sup>Tyr</sup> accounts for the toxicity of D-tyrosine toward *Escherichia coli*. *J. Biol. Chem.* **279**, 42560-42565.
- Splan, K.E., Ignatov, M.E. and Musier-Forsyth, K. (2008). Transfer RNA modulates the editing mechanism used by class II prolyl-tRNA synthetase. *J. Biol. Chem.* **283**, 7128-7134.

- Sprinzi, M., Hartmann, T., Weber, J., Blank, J. and Zeidler, R. (1989). Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.* **17 Suppl**, r1-172.
- Sprinzi, M., Horn, C., Brown, M., Ioudovitch, A. and Steinberg, S. (1998). Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.* **26**, 148-153.
- Srinivasan, G., James, C.M. and Krzycki, J.A. (2002). Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science* **296**, 1459-1462.
- Steinmetz, L.M., Scharfe, C., Deutschbauer, A.M., Mokranjac, D., Herman, Z.S., Jones, T., Chu, A.M., Giaever, G., Prokisch, H., Oefner, P.J. and Davis, R.W. (2002). Systematic screen for human disease genes in yeast. *Nat. Genet.* **31**, 400-404.
- SternJohn, J., Hati, S., Siliciano, P.G. and Musier-Forsyth, K. (2007). Restoring species-specific posttransfer editing activity to a synthetase with a defunct editing domain. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 2127-2132.
- Swairjo, M.A. and Schimmel, P.R. (2005). Breaking sieve for steric exclusion of a noncognate amino acid from active site of a tRNA synthetase. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 988-993.
- Takaku, H. and Nashimoto, M. (2008). *Escherichia coli* tRNase Z can shut down growth probably by removing amino acids from aminoacyl-tRNAs. *Genes Cells* **13**, 1087-1097.
- Takarada, T., Takahata, Y., Iemata, M., Hinoi, E., Uno, K., Hirai, T., Yamamoto, T. and Yoneda, Y. (2009). Interference with cellular differentiation by D-serine through antagonism at N-methyl-D-aspartate receptors composed of NR1 and NR3A subunits in chondrocytes. *J. Cell. Physiol.* **220**, 756-764.
- Takayama, T., Ogawa, T., Hidaka, M., Shimizu, Y., Ueda, T. and Masaki, H. (2005). Esterification of *Escherichia coli* tRNAs with D-histidine and D-lysine by aminoacyl-tRNA synthetases. *Biosci. Biotechnol. Biochem.* **69**, 1040-1041.
- Thanedar, S., Kumar, N.V. and Varshney, U. (2000). The fate of the initiator tRNAs is sensitive to the critical balance between interacting proteins. *J. Biol. Chem.* **275**, 20361-20367.
- Torres-Larios, A., Swinger, K.K., Pan, T. and Mondragon, A. (2006). Structure of ribonuclease P--a universal ribozyme. *Curr. Opin. Struct. Biol.* **16**, 327-335.

- Tsui, W.C. and Fersht, A.R. (1981). Probing the principles of amino acid selection using the alanyl-tRNA synthetase from *Escherichia coli*. *Nucleic Acids Res.* **9**, 4627-4637.
- Tu, G.F., Reid, G.E., Zhang, J.G., Moritz, R.L. and Simpson, R.J. (1995). C-terminal extension of truncated recombinant proteins in *Escherichia coli* with a 10Sa RNA decapeptide. *J. Biol. Chem.* **270**, 9322-9326.
- van Dijk, M., van Dijk, A.D., Hsu, V., Boelens, R. and Bonvin, A.M. (2006). Information-driven protein-DNA docking using HADDOCK: it is a matter of flexibility. *Nucleic Acids Res.* **34**, 3317-3325.
- Vidales, F.J., Bernabeu, C. and Ballesta, J.P. (1979). Peptidyl transfer ribonucleic acid hydrolase activity of proteinase K. *Biochemistry* **18**, 4155-4158.
- Watson, J.D. and Crick, F.H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**, 737-738.
- Williams, A.M. and Martinis, S.A. (2006). Mutational unmasking of a tRNA-dependent pathway for preventing genetic code ambiguity. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 3586-3591.
- Wimberly, B.T., Brodersen, D.E., Clemons, W.M., Jr., Morgan-Warren, R.J., Carter, A.P., Vornrhein, C., Hartsch, T. and Ramakrishnan, V. (2000). Structure of the 30S ribosomal subunit. *Nature* **407**, 327-339.
- Winn, M.D., Murshudov, G.N. and Papiz, M.Z. (2003). Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol.* **374**, 300-321.
- Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L. and Sykes, B.D. (1995). <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* **6**, 135-140.
- Wishart, D.S. and Sykes, B.D. (1994). The <sup>13</sup>C chemical-shift index: a simple method for the identification of protein secondary structure using <sup>13</sup>C chemical-shift data. *J. Biomol. NMR* **4**, 171-180.
- Wolosker, H., Blackshaw, S. and Snyder, S.H. (1999). Serine racemase: a glial enzyme synthesizing D-serine to regulate glutamate-N-methyl-D-aspartate neurotransmission. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13409-13414.
- Wong, F.C., Beuning, P.J., Nagan, M., Shiba, K. and Musier-Forsyth, K. (2002). Functional role of the prokaryotic proline-tRNA synthetase insertion domain in amino acid editing. *Biochemistry* **41**, 7108-7115.

Wong, F.C., Beuning, P.J., Silvers, C. and Musier-Forsyth, K. (2003). An isolated class II aminoacyl-tRNA synthetase insertion domain is functional in amino acid editing. *J. Biol. Chem.* **278**, 52857-52864.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York

Wydaŭ, S., van der Rest, G., Aubard, C., Plateau, P. and Blanquet, S. (2009). Widespread distribution of cell defense against D-aminoacyl-tRNAs. *J. Biol. Chem.* **284**, 14096-14104.

Xiong, Y. and Steitz, T.A. (2004). Mechanism of transfer RNA maturation by CCA-adding enzyme without using an oligonucleotide template. *Nature* **430**, 640-645.

Yusupov, M.M., Yusupova, G.Z., Baucom, A., Lieberman, K., Earnest, T.N., Cate, J.H. and Noller, H.F. (2001). Crystal structure of the ribosome at 5.5 Å resolution. *Science* **292**, 883-896.

Zaher, H.S. and Green, R. (2009a). Fidelity at the molecular level: lessons from protein synthesis. *Cell* **136**, 746-762.

Zaher, H.S. and Green, R. (2009b). Quality control by the ribosome following peptide bond formation. *Nature* **457**, 161-166.

Zhai, Y. and Martinis, S.A. (2005). Two conserved threonines collaborate in the *Escherichia coli* leucyl-tRNA synthetase amino acid editing mechanism. *Biochemistry* **44**, 15437-15443.

Zhai, Y., Nawaz, M.H., Lee, K.W., Kirkbride, E., Briggs, J.M. and Martinis, S.A. (2007). Modulation of substrate specificity within the amino acid editing site of leucyl-tRNA synthetase. *Biochemistry* **46**, 3331-3337.

Zhang, H., Huang, K., Li, Z., Banerjee, L., Fisher, K.E., Grishin, N.V., Eisenstein, E. and Herzberg, O. (2000). Crystal structure of YbaK protein from *Haemophilus influenzae* (HI1434) at 1.8 Å resolution: functional implications. *Proteins* **40**, 86-97.

Zhao, M.W., Zhu, B., Hao, R., Xu, M.G., Eriani, G. and Wang, E.D. (2005). Leucyl-tRNA synthetase from the ancestral bacterium *Aquifex aeolicus* contains relics of synthetase evolution. *EMBO J.* **24**, 1430-1439.

Zhouravleva, G., Frolova, L., Le Goff, X., Le Guellec, R., Inge-Vechtomy, S., Kisselev, L. and Philippe, M. (1995). Termination of translation in eukaryotes is governed by two interacting polypeptide chain release factors, eRF1 and eRF3. *EMBO J.* **14**, 4065-4072.



