



**HAL**  
open science

# Etude structurale et fonctionnelle de la protéine à radical SAM Hyde

Roman Rohac

► **To cite this version:**

Roman Rohac. Etude structurale et fonctionnelle de la protéine à radical SAM Hyde. Biologie structurale [q-bio.BM]. Université Grenoble Alpes, 2016. Français. NNT: 2016GREAV010. tel-01524860

**HAL Id: tel-01524860**

**<https://theses.hal.science/tel-01524860>**

Submitted on 19 May 2017

**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

Pour obtenir le grade de

**DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES**

Spécialité : **Biologie Structurale et Nanobiologie**

Arrêté ministériel : 7 août 2006

Présentée par

**Roman ROHAC**

Thèse dirigée par **Yvain NICOLET** et  
codirigée par **Juan C. FONTECILLA-CAMPS**

préparée au sein de **L'Institut de Biologie Structurale  
(CEA/CNRS/UJF), Groupe Métalloprotéines**  
dans **l'École Doctorale Chimie et Sciences du Vivant**

# Etude structurale et fonctionnelle de la protéine à radical SAM HydE

Thèse soutenue publiquement le **18 mai 2016**,  
devant le jury composé de:

**Mr Olivier BERTEAU**

Chargé de recherche à l'INRA, Examineur

**Mr David PIGNOL**

Directeur de recherche au CEA Cadarache, Rapporteur

**Mr Marius REGLIER**

Directeur de recherche 1 au CNRS, Rapporteur

**Mr Fabrice THOMAS**

Professeur à l'UJF Grenoble, Examineur

**Mr Juan C., FONTECILLA-CAMPS**

Directeur de recherche au CEA Grenoble, Co-directeur de thèse

**Mr Yvain, NICOLET**

Chargé de recherche au CEA Grenoble, Directeur de thèse





*Pour Veronika , mon photon intriqué*



## Remerciements

*Je souhaite remercier Juan Carlos Fontecilla-Camps et Yvain Nicolet de m'avoir accueilli au sein de leur laboratoire et de m'avoir fait confiance avec ce sujet de thèse.*

*Je remercie Juan pour de nombreuses discussions scientifiques très enrichissantes et pour des informations fort intéressantes sur les pieuvres et les origines de l'univers et bien sûr pour toutes ces blagues, qui le caractérisent si bien.*

*Un grand merci à toi, Yvain, mon directeur de thèse, pour ton encadrement, ta disponibilité, ta patience (parce qu'il en fallait pas mal), tes encouragements et également pour ton humanité. Je sais qu'avec moi cela n'a pas été tout le temps un long fleuve tranquille mais j'ai beaucoup appris à tes côtés aussi bien professionnellement que personnellement. Pour tout cela je voudrais te témoigner ma profonde gratitude.*

*Je remercie chaleureusement les membres de mon jury, Olivier Berteau, David Pignol, Marius Réglie, Fabrice Thomas, Juan C. Fontecilla-Camps et Yvain Nicolet, d'avoir accepté d'évaluer ma thèse.*

*Je remercie Jean-Luc Ferrer pour ces centaines d'heures de temps de faisceau sur la ligne BM30A à ESRF qu'il m'a généreusement accordées. Et un grand merci à Michel Pirocchi pour tout ce temps qu'il a passé sur la ligne avec moi à veiller à ce que toutes les expériences se passent au mieux.*

*Je remercie Patricia Amara pour cet immense travail de calculs théoriques, qu'elle a effectué à partir de mes structures cristallographiques, ainsi que pour ses encouragements et ses conseils, qu'elle m'a donnés au cours de la rédaction de ce manuscrit.*

*Je remercie Jean-Marie Mouesca pour son implication dans le projet de ma thèse et pour ses grandes compétences qui nous ont permis d'éclairer certains points cruciaux sur le sujet.*

*Je remercie Olivier Berteau, Alhosna Benjdia et Pauline Ruffié pour les analyses HPLC des produits de la réaction observée dans la protéine HydE.*

*Je remercie Adrien Favier pour ces expériences en RMN aux résultats fort sympathiques qui nous ont permis d'ajouter la touche finale à notre article sur la chimie radicalaire des protéines à radical SAM.*

*Je remercie Serge Gambarelli et Giuseppe Sicoli pour les analyses en RPE à basse température de mes échantillons protéiques.*

*Je remercie Laura Zeppieri et Lydie Martin de m'avoir montré les astuces de paillasse lors de mon initiation aux manips de biochimie en milieu anaérobie.*

*Je remercie Xavier Vernède pour sa disponibilité et pour son aide technique précieuse.*

*Je tiens à remercier l'ensemble des personnes qui ont travaillé avec moi de près ou de loin au cours de ces quatre années de thèses et qui ont toutes pris le temps de m'écouter, de m'expliquer, de m'encourager et de m'aider : Widade, Hugo, Sarah, Monika, Niels, Elodie, Claudine, Oriane, Fanny, Frank, David, Jade, Antoine, Eve, Jacques, Isabelle et Chantal. Merci à vous tous pour ces bons moments.*

*Je voudrais remercier Adrien pour l'amitié qu'il m'a apportée et pour la patience qu'il a eue avec moi. Il m'est impossible de penser à ces moments autrement qu'avec un sourire sur les lèvres, "mon petit".*

*Je remercie enfin ma famille ; ma mère, mon père et ma petite sœur Daniela, pour leur soutien, leur confiance en moi, leur patience et leur amour.*

*Et je te remercie toi, Veronika pour ton amour incommensurable, pour ta patience quasi-infinie, pour toujours croire en moi et pour m'avoir soutenu même dans certains moments de doute, merci pour tout ce que tu fais pour moi et pour ce que tu m'apportes au quotidien. Ta présence auprès de moi m'a été salutaire durant des périodes difficiles (et il y en a eu pas mal). Grâce à tes encouragements j'ai pu achever la rédaction de mon manuscrit. Rien de ce qui s'y trouve n'y serait arrivé sans toi.*

# SOMMAIRE

<b>ABREVIATIONS</b>	<b>6</b>
<b>INTRODUCTION GENERALE</b>	<b>7</b>
<b>INTRODUCTION BIBLIOGRAPHIQUE</b>	<b>9</b>
<b>IB-1 HYDROGENASES</b>	<b>10</b>
IB-1.1 SIGNES D'UNE EVOLUTION CONVERGENTE	10
IB-1.2 LES HYDROGENASES A [NiFe]	12
IB-1.3 LES HYDROGENASES A [FeFe]	13
IB-1.4 CENTRES [Fe-S]	14
<b>IB-2 MATURATION DU SITE ORGANOMETALLIQUE DES HYDROGENASES A [FeFe]</b>	<b>15</b>
IB-2.1 L'IDENTIFICATION DE LA MACHINERIE HYD	15
IB-2.2 MATURASE HYDF - UNE GTPASE CONTENANT UN CENTRE [Fe-S]	16
<b>IB-3 PROTEINES A RADICAL SAM - GENERALITES</b>	<b>17</b>
IB-3.1 MATURASE HYDG - UNE ENZYME A RADICAL SAM TRIFONCTIONNELLE	19
IB-3.2 MATURASE HYDE - SOURCE POTENTIELLE DU LIGAND DITHIOMETHYLAMINE	22
<b>IB-4 MECANISME DE MATURATION</b>	<b>25</b>
<b>IB-5 PROTEINES A RADICAL SAM - INSERTION DU SOUFRE PAR LA VOIE RADICALE</b>	<b>27</b>
<b>OBJECTIFS DE LA THESE</b>	<b>30</b>
<b>IB-6 REFERENCES DE L'INTRODUCTION BIBLIOGRAPHIQUE</b>	<b>31</b>
<b>PARTIE RESULTATS</b>	<b>38</b>
<b>CHAPITRE I : RECHERCHE DU SUBSTRAT DE HYDE</b>	<b>40</b>
<b>I-1 INTRODUCTION DU CHAPITRE</b>	<b>40</b>
<b>I-2 LE SUBSTRAT DE HYDE - UN METABOLITE D'E.COLI</b>	<b>41</b>
<b>I-3 REFLEXIONS SUR LA PROVENANCE DES ATOMES DE SOUFRE DANS LE LIGAND DTMN</b>	<b>41</b>
I-3.1 IMPLICATION DES CENTRES [Fe-S] DANS LES REACTIONS D'INSERTION DE SOUFRE	41
I-3.2 ORIGINES POSSIBLES DES ATOMES DE SOUFRE DANS LE LIGAND DTMN	43
<b>I-4 METHODES D'IDENTIFICATION DU SUBSTRAT D'UNE ENZYME</b>	<b>44</b>
I-4.1 CONCEPTIONS DE TESTES FONCTIONNELS POUR MESURER L'ACTIVITE DE HYDE	44
I-4.2 APPROCHE SIMILAIRE AU « RATIONAL DRUG DESIGN »	45
I-4.2.1 Limitations de la méthode	45
<b>I-5 ANALYSE DES CARACTERES PHYSICO-CHIMIQUES DE LA CAVITE CATALYTIQUE</b>	<b>47</b>
<b>I-6 PREMIERES EXPERIENCES D'ACCOSTAGE MOLECULAIRE</b>	<b>49</b>
I-6.1 COMPLEXE AVEC L'ACIDE 1,3-PHENYLENEDIACETIQUE	50
I-6.2 TESTS DE CO-CRISTALLISATION AVEC DES PETITES MOLECULES MIMETIQUES DE FRAGMENTS FONCTIONNELS DU SUBSTRAT	50
<b>I-7 AFFINEMENT D'HYPOTHESES SUR LA NATURE DU SUBSTRAT DE HYDE</b>	<b>51</b>
<b>I-8 L-CYSTINE COMME UN SUBSTRAT POTENTIEL DE HYDE</b>	<b>51</b>
I-8.1 MECANISME PROPOSE DE CONVERSION DE CYSTINE EN AZADITHIOLATE	53
I-8.2 ASPARTATE AMMONIA-LYASE DANS L'OPERON HYD	55
I-8.3 RECHERCHE DU SITE DE FIXATION D'AMMONIUM	55
I-8.4 DERIVATISATION AVEC DU DICYANOAUATE	55
I-8.5 PREMIER TEST DE CO-CRISTALLISATION AVEC DE LA SE-L-CYSTINE	58
I-8.5.1 Identification d'un ligand à un atome de sélénium	58
I-8.5.2 Partielle reconstitution de l'agrégat [Fe-S] auxiliaire avec du séléniure	61

I-8.6	ESSAIS DE CO-CRISTALLISATION AVEC DES ANALOGUES DES ESPECES INTERMEDIAIRES HYPOTHETIQUES	62
I-8.7	SECONDE CO-CRISTALLISATION DE TmHYDE SAUVAGE AVEC DE LA L-CYSTINE	63
I-8.7.1	Dérivé de L-cystéine comme ligand de HydE	63
I-8.7.2	Interactions entre le ligand (2R,4R)-MeTDA et le site actif de HydE	65
I-8.7.3	Disposition du ligand (2R,4R)-MeTDA par rapport à la SAM	67
I-8.8	ESSAIS DE CO-CRISTALLISATION DE TmHYDE ΔFeS N-STREP AVEC LA L-CYSTINE	68
I-8.8.1	Préparation de l'échantillon TmHydE ΔFeS N-strep	68
I-8.8.2	Structure cristallographique du mutant TmHydE ΔFeS N-strep	69
I-8.8.3	Co-cristallisation de TmHydE ΔFeS N-strep avec la L-cystine	71
I-8.9	COMPLEXE AVEC UN LIGAND 3-MERCAPTOPROPIONATE MODIFIE	71
<b>I-9</b>	<b>ACCOSTAGE MOLECULAIRE <i>IN SILICO</i> AVEC LES METABOLITES POLYCARBOXYLES</b>	<b>74</b>
I-9.1	TRICARBOXYLATES COMME LIGANDS DE HYDE	74
<b>I-10</b>	<b>ESSAIS DE CO-PURIFICATION DE LA PROTEINE HYDE Tm N-STREP ΔFeS AVEC LES INTERMEDIAIRES DE LA REACTION</b>	<b>77</b>
I-10.1	INCUBATION DE LA PROTEINE TmHYDE Tm NSTREP AVEC DU LYSAT ANAEROBIQUE D'E. COLI	78
I-10.2	RECONSTITUTION <i>IN VIVO</i> DE TmHYDE Tm NSTREP ΔFeS	79
I-10.2.1	Composition de la cavité du mutant ΔFeS N-strep reconstitué <i>in vivo</i> en fermentation du glucose	80
I-10.2.2	Composition de la cavité du mutant ΔFeS reconstitué <i>in vivo</i> en respiration du fumarate — Co-purification de la SAM	82
<b>I-11</b>	<b>ETUDE DES LIGANDS DERIVES DE L-CYSTEINE</b>	<b>84</b>
I-11.1	TREMPAGE DE CRISTAUX DE HYDE DE T. MARITIMA ΔFeS N-STREP AVEC DE LA CYSTEINE ET DU PYRUVATE	84
I-11.2	ETUDE DU COMPORTEMENT DU LIGAND MeTDA EN SOLUTION	84
I-11.3	STEREOSELECTIVITE DU SITE ACTIF DE TmHYDE VIS-A-VIS DU LIGAND (2R,4R)-METDA	87
I-11.4	ELARGISSEMENT DE LA FAMILLE DE LIGANDS DE HYDE	87
I-11.4.1	Détails des interactions entre HydE et les ligands nouvellement identifiés	90
I-11.4.2	Profil minimum d'un ligand stéréospécifique de HydE	92
I-11.4.3	Site T1 - un site d'affinité pour le soufre	93
<b>I-12</b>	<b>RESUME DU CHAPITRE I</b>	<b>94</b>
<b>I-13</b>	<b>REFERENCES DU CHAPITRE I</b>	<b>95</b>

## **CHAPITRE II : ETUDE FONCTIONNELLE DE HYDE PAR L'ACTIVATION D'HYDROGENASE DANS L'EXTRAIT CELLULAIRE** **99**

<b>II-1</b>	<b>INTRODUCTION DU CHAPITRE</b>	<b>99</b>
<b>II-2</b>	<b>OPTIMISATION DES CONDITIONS DE MATURATION D'HYDROGENASE CrHYDA1</b>	<b>99</b>
<b>II-3</b>	<b>IMPORTANCE DE LA MATURASE HYDE DANS LE PROCESSUS D'ACTIVATION DE L'HYDROGENASE</b>	<b>100</b>
II-3.1	NOUVEAU MODELE DE MATURATION PROPOSE PAR L'EQUIPE DE JAMES R. SWARTZ	100
II-3.2	ESSAIS DE MATURATION D'HYDROGENASE SANS ET AVEC HYDE	102
<b>II-4</b>	<b>EFFET DE LA CYSTEINE, DES THIAZOLIDINES ET DE A-KETOACIDES SUR LA MATURATION D'HYDROGENASE</b>	<b>104</b>
II-4.1	EFFET DE LA CYSTEINE	104
II-4.2	EFFET DE DERIVES THIAZOLIDINES FORMES DANS LE MELANGE DE MATURATION	105
II-4.3	EFFET DE DERIVES THIAZOLIDINES PREFORMES VS A-KETOACIDES SEULS	106
II-4.4	EFFET DE PENICILLAMINE	106
II-4.5	ROLE DU PLP DANS LE PROCESSUS DE LA MATURATION DE L'HYDROGENASE A [FeFe]	107
<b>II-5</b>	<b>ACTIVATION D'HYDROGENASE EN UTILISANT DE LA CYSTEINE <sup>13</sup>C</b>	<b>107</b>
<b>II-6</b>	<b>UTILISATION D'EXTRAITS DIALYSES</b>	<b>107</b>
II-6.1	TESTS DE FRACTIONNEMENT DE LYSAT CLARIFIE DE E. COLI	108
<b>II-7</b>	<b>CONCLUSIONS ET PERSPECTIVES DES CHAPITRES I &amp; II</b>	<b>109</b>
<b>II-8</b>	<b>REFERENCES DU CHAPITRE II</b>	<b>113</b>

**CHAPITRE III : FORMATION D'UNE LIAISON C-S PAR UN PROCESSUS RADICALAIRE CATALYSE PAR HYDE** **119**

<b>III-1</b>	<b>INTRODUCTION DU CHAPITRE</b>	<b>119</b>
<b>III-2</b>	<b>ETUDE DE LA REACTION DE HYDE AVEC LES DERIVES 1,3-THIAZOLIDINE EN SOLUTION</b>	<b>120</b>
III-2.1	ETUDE FONCTIONNELLE DE HYDE PAR LES MESURES DU CLIVAGE DE LA SAM	120
III-2.1.1	Clivage de la SAM en présence de molécules thiolées (Betz et al.)	120
III-2.1.2	Réinterprétation des données publiées par Betz et al.	121
III-2.1.3	Clivage de la SAM en présence du dérivé (2R,4R)-MeTDA	122
III-2.1.4	Production enzymatique d'un nouveau dérivé adénosyl - la SAC	124
III-2.1.5	Impact du centre [Fe-S] auxiliaire sur la production de la SAC	126
III-2.2	SUIVI PAR SPECTROSCOPIE RMN RESOLUE DANS LE TEMPS DE LA FORMATION DE PYRUVATE LIBRE AU COURS DE LA PRODUCTION ENZYMATIQUE DE LA SAC	127
<b>III-3</b>	<b>ETUDE PAR CRISTALLOGRAPHIE AUX RAYONS X DE LA REACTIVITE DE TmHYDE ΔFeS ENVERS LE LIGAND (2R,4R)-MESETDA</b>	<b>129</b>
III-3.1	REACTIVITE DE HYDE AVEC LE LIGAND (2R,4R)-MESETDA DANS LES CRISTAUX PROTEIQUES	129
III-3.1.1	Disposition du site actif à $t_0$	130
III-3.1.2	Disposition et contenu du site actif à $t_{final}$	130
III-3.1.3	Analyse des mouvements dans le site actif au cours de la réaction	132
III-3.1.4	Absence de détection d'intermédiaires de la réaction	134
<b>III-4</b>	<b>FORMATION D'UNE LIAISON C-S(E) PAR UN PROCESSUS RADICALAIRE</b>	<b>134</b>
III-4.1	LE ROLE DU PYRUVATE DANS LA PRODUCTION ENZYMATIQUE DE LA S-ADENOSYL-L-CYSTEINE	134
III-4.2	ATTAQUE RADICALAIRE DU 5'-dA• SUR L'ATOME DE S(E) D'UN SELENOETHER/THIOETHER	135
III-4.2.1	Calculs théoriques	136
III-4.2.2	Analyse orbitale	136
III-4.2.3	Analyse orbitale de la réaction de HydE avec le ligand (2R,4R)-MeSeTDA	137
III-4.2.4	Etude de la directionnalité de la substitution homolytique ( $S_H2$ ) dans HydE	139
<b>III-5</b>	<b>NATURE DU RADICAL 5'-dA•</b>	<b>142</b>
III-5.1	INTERACTIONS ENTRE LES ATOMES C5' ET SE1 AVANT LA REACTION	142
III-5.2	ANALYSE DU MOUVEMENT DU 5'-dA AU COURS DE LA REACTION	147
III-5.3	ETUDE DU ROLE DE LA PROTEINE DANS LE MECANISME RADICALAIRE DE HYDE	148
<b>III-6</b>	<b>TERMINAISON DE LA REACTION RADICALAIRE</b>	<b>150</b>
<b>III-7</b>	<b>MECANISME DETAILLE DE LA FORMATION DE LA (SE)-SAC PAR HYDE</b>	<b>152</b>
<b>III-8</b>	<b>DISCUSSION</b>	<b>153</b>
III-8.1	IMPLICATIONS SUR LES REACTIONS D'INSERTIONS D'ATOME DE SOUFRE CATALYSEES PAR LES ENZYMES A RADICAL SAM	153
III-8.2	IMPLICATIONS SUR LE MECANISME D'INITIATION DU REARRANGEMENT RADICALAIRE DU SUBSTRAT CATALYSE PAR LES ENZYMES A RADICAL SAM	155
III-8.2.1	Régiosélectivité du clivage de la SAM	155
III-8.2.2	Contrôle des mouvements du 5'-dA dans le site actif	156
III-8.2.3	Le 5'-dA• comme un état de transition dans les enzymes à radical SAM	160
III-8.3	LE 5'-dA DANS LES ENZYMES A RADICAL SAM ET DANS LES ENZYMES A ADENOSYLCOBALAMINE	161
III-8.4	IMPLICATION POUR LA MATURASE HYDE	164
<b>III-9</b>	<b>REFERENCES DU CHAPITRE III</b>	<b>166</b>

**CHAPITRE IV : ETUDE STRUCTURALE D'INTERCONVERSION ENTRE DIFFERENTES FORMES DE CENTRES [FE-S]** **171**

<b>IV-1</b>	<b>STRUCTURE DE TmHYDE AVEC DEUX CENTRES [Fe<sub>4</sub>S<sub>4</sub>]</b>	<b>172</b>
<b>IV-2</b>	<b>ROLE DE L'AGREGAT [Fe-S] AUXILIAIRE</b>	<b>173</b>
<b>IV-3</b>	<b>PERSPECTIVES</b>	<b>174</b>
IV-3.1	ETUDE D'INTERCONVERSION DES CENTRES [Fe-S] EN SOLUTION	174
IV-3.2	CENTRES [Fe-Se] ET LE PHASAGE DE DONNEES CRISTALLOGRAPHIQUES	174
<b>IV-4</b>	<b>REFERENCES DU CHAPITRE IV</b>	<b>175</b>

<b>CHAPITRE V : VERS LA SORTIE DU PRODUIT DE HYDE</b>	<b>177</b>
<b>V-1 INTRODUCTION</b>	<b>177</b>
<b>V-2 VERS LE TRANSFERT DU PRODUIT HORS DE HYDE</b>	<b>178</b>
V-2.1 EXPOSITION DE LA PARTIE BASSE DU TONNEAU AU SOLVANT	179
V-2.2 EXTERNALISATION DU CHLORURE EN BAS DU TONNEAU	180
V-2.3 CONCLUSIONS DU CHAPITRE V: HYDE - UNE LIGNE D'ASSEMBLAGE A SENS UNIQUE	182
<b>V-3 REFERENCES DU CHAPITRE V</b>	<b>183</b>
<b>CONCLUSIONS GENERALES</b>	<b>185</b>
<b>MATERIELS ET METHODES</b>	<b>187</b>
<b>MM-1 MATERIELS</b>	<b>187</b>
MM-1.1 MATERIELS BIOLOGIQUES	187
MM-1.1.1 Gènes utilisés	187
MM-1.1.2 Vecteurs de surexpression	187
MM-1.1.3 Souches bactériennes	188
MM-1.2 MILIEUX DE CULTURE	189
<b>MM-2 BIOLOGIE MOLECULAIRE</b>	<b>189</b>
MM-2.1 PREPARATION DE CELLULES COMPLETES	190
MM-2.2 TRANSFORMATION DES SOUCHES ESCHERICHIA COLI COMPETENTES	190
MM-2.3 MUTAGENESE DIRIGEE PAR PCR	191
MM-2.4 PREPARATION D'ADN PLASMIDIQUE	192
<b>MM-3 BIOCHIMIE</b>	<b>193</b>
MM-3.1 SUREXPRESSIONS DES PROTEINES	193
MM-3.1.1 Respiration du fumarate	193
MM-3.1.2 Fermentation du glucose	193
MM-3.1.3 Protocole de Swartz - Fermentation du glucose	193
MM-3.2 PURIFICATIONS DES PROTEINES	194
MM-3.2.1 Purification de HydE sans tag de Thermotoga maritima	195
MM-3.2.2 Purification de HydE N-strep de Thermotoga maritima par chromatographie d'affinité	197
MM-3.2.3 Purification de HydE N-strep de Clostridium acetobutylicum	197
MM-3.2.4 Purification de HydE N-strep $\Delta$ FeS de Thermotoga maritima	197
MM-3.2.5 Purification de HydE N-strep $\Delta$ FeS Se de Thermotoga maritima	198
MM-3.2.6 Purification de HydE N-strep $\Delta$ FeS de Thermotoga maritima reconstituée in vivo en condition de respiration de fumarate	198
MM-3.2.7 Purification de HydE N-strep $\Delta$ FeS de Thermotoga maritima reconstituée in vivo en condition de fermentation	198
MM-3.2.8 Purification de HydA1 N-strep de Chlamydomonas reinhardtii	198
MM-3.3 RECONSTITUTION DES CENTRES FES	199
MM-3.3.1 Reconstitution chimique des apoprotéines purifiées	199
MM-3.3.2 Reconstitution des apoprotéines purifiées à l'aide d'une cystéine désulfurase	199
MM-3.3.3 Reconstitution dans l'extrait	200
MM-3.3.4 Reconstitution in vivo	200
MM-3.3.5 Reconstitution au sélénium	201
MM-3.4 ANALYSES BIOCHIMIQUES	201
MM-3.4.1 Dosage protéique par la méthode du rose Bengal	201
MM-3.4.2 Dosage du fer	202
MM-3.4.3 Dosage du soufre	203
MM-3.4.4 Electrophorèse sur gel de polyacrylamide en présence de dodécylsulfate de sodium	204
MM-3.5 PREPARATIONS DE DERIVES DE THIAZOLIDINES ET DE THIAZINANES	204
MM-3.6 ESSAIS ENZYMATIQUES D'ACTIVITE HYDROGENASE	206
MM-3.6.1 Préparation de molécules nécessaires à la biosynthèse dans l'extrait du sous-agrégat à 2Fe	206

MM-3.6.2	Préparation des lysats cellulaires	207
MM-3.6.2.1	Préparation des lysats cellulaires anaérobies clarifiés pour les expériences de maturation d'hydrogénase dans l'extrait	207
MM-3.6.2.2	Préparation de lysats cellulaires anaérobies dialysés	207
MM-3.6.2.3	Fractionnement des lysats clarifiés en fonction de la taille des molécules	207
MM-3.6.3	Protocole d'activation de l'hydrogénase à [FeFe] par maturation in vitro du sous-agrégat à 2 Fe	208
MM-3.6.4	Mesures d'activité hydrogénase	209
MM-3.6.4.1	Principe de mesures d'activité hydrogénase par dosage colorimétrique du méthylviologène réduit	209
MM-3.6.4.2	Préparation de cuvettes avec une solution de méthylviologène saturée en dihydrogène	210
MM-3.6.4.3	Mesures d'activité hydrogénase par dosage colorimétrique du méthylviologène réduit	210
<b>MM-4</b>	<b>CRISTALLISATION ET CRISTALLOGRAPHIE AUX RAYONS X</b>	<b>211</b>
MM-4.1	CO-CRISTALLISATION OU TREMPAGE DE CRISTAUX DE LA PROTEINE HYDE DE THERMOTOGA MARITIMA AVEC DES SUBSTRATS POTENTIELS	211
MM-4.2	CRISTALLOGRAPHIE AUX RAYONS X	211
MM-4.2.1	Conditions expérimentales	212
MM-4.2.2	Corrections des données	213
MM-4.2.3	Calcul des cartes de densité électronique	214
MM-4.2.4	Diffusion anormale	215
MM-4.2.4.1	Absorption des rayons X	215
MM-4.2.4.2	Facteurs de la diffusion anormale	216
MM-4.2.4.3	Brisure de la loi de Friedel	217
MM-4.2.4.4	Détermination des facteurs de diffusion atomique dispersive $f''$ et anormale $f'''$	219
MM-4.2.4.5	Cartes de densité électronique anormale	220
MM-4.2.5	Carte de densité électronique double différence	220
MM-4.2.6	Affinement restreint	221
MM-4.2.6.1	Contraintes stéréochimiques	221
<b>MM-5</b>	<b>SPECTROSCOPIE</b>	<b>223</b>
MM-5.1	SPECTROSCOPIE DE RESONANCE MAGNETIQUE NUCLEAIRE DU PROTON	223
MM-5.1.1	Principe de la RMN	223
MM-5.1.2	RMN STD $^1\text{H}$	225
MM-5.1.3	Principe d'une expérience RMN STD 1D $^1\text{H}$	225
MM-5.1.4	Préparation de l'échantillon pour une expérience de RMN STD 1D $^1\text{H}$	226
MM-5.2	SPECTROSCOPIE DE RESONANCE PARAMAGNETIQUE ELECTRONIQUE	226
MM-5.2.1	Principe de la résonance paramagnétique électronique (RPE)	226
MM-5.2.2	Expérience de RPE avec HydA1 Cr Nstrep repurifiée	228
MM-5.3	SPECTROSCOPIE HYSCORE	228
MM-5.3.1	Principe de la spectroscopie HYSCORE	228
MM-5.3.2	Expérience de spectroscopie HYSCORE avec HydA1 Cr Nstrep repurifiée	229
MM-5.4	SPECTROSCOPIE DE MASSE	229
<b>MM-6</b>	<b>REFERENCES DES MATERIELS ET METHODES</b>	<b>230</b>
<b>ANNEXES</b>		<b>233</b>

## Abréviations

ADN	Acide Désoxyribonucléique	Ni-NTA	Nickel-Acide NitriloTriacétique
Amp	Ampicilline	N-strep	Etiquette strep N-terminale
APS	Ammonium Persulfate	pb	Paire de bases
ATP	Adénosine triphosphate	PCA	Acide Perchlorique
BAG	Boîte à gants	PCR	Réaction en chaîne par polymérase
BPS	Bathophénanthroline	PDT	Propane-1,3-dithiol
BSA	Bovine Sérum Albumine	PEG	Polyéthylène glycol
CHLRE	<i>Chlamydomonas reinhardtii</i>	PFL	Pyruvate Formate Lyase
CLOAC	<i>Clostridium acetobutylicum</i>	pI	Point isoélectrique
CV	Volume de colonne	PLP	Phosphate de pyridoxal
Da	Dalton	PRS	Protéine à radical SAM
$\Delta$ [FeS]	triple mutant Cys/Ser ne pouvant pas contenir un centre [Fe-S] auxiliaire	RPE	Résonnance Paramagnétique Electronique
DMPD	Diméthyl Phénylène Diamine	rpm	Rotation par minute
DO	Densité optique	SAH	S-Adénosyl-L-homocystéine
DTH	Dithionite de sodium	SAM	S-Adénosyl-L-méthionine
DTT	Dithiothréitol	SDS-PAGE	Sodium dodécyl sulfate polyacrylamide gel electrophoresis
EDTA	Acide éthylène diamine tétracétique	TB	Terrific Broth
ESRF	European Synchrotron Radiation Facility	TEMED	N, N, N', N' – TétraMéthylEthylène Diamine
HPLC	High Pressure Liquid Chromatography	THEMA	<i>Thermotoga maritima</i>
HydA <sup>ΔEFG</sup>	Hydrogénase à [FeFe] exprimée sans les maturases	TIM	Triose Phosphate Isomérase
IPTG	Isopropyl $\beta$ -D-1-thiogalactopyranoside	Tm	Température de fusion
Isc	Iron Sulfur Cluster	Tris	Tris-hydroxyméthylaminométhane
Kan	Kanamycine	u	Unité
LAM	Lysine-2,3-AminoMutase	UV	Ultra-Violet
LB	Luria Broth	XDS	X-ray Detector Software
MWCO	Molecular Weight Cut-Off		
NADPH	Nicotinamide Adénine Dinucléotide Phosphate		

## Introduction Générale

« *Vas prendre tes leçons dans la nature, c'est là qu'est notre futur* » a dit Léonard de Vinci il y a 500 ans. En effet, la nature qui nous entoure met en place des procédés sophistiqués, façonnés petit à petit au cours des milliards d'années. C'est en observant la nature que l'on peut répondre aux besoins de la vie moderne. En tirant profit des milliards d'années d'évolution, nous avons la possibilité de faire des bonds technologiques jamais imaginés.

Par exemple l'étude de l'anatomie des ailes d'oiseaux et de certains chiroptères a inspiré les premiers aviateurs dans la conception de leurs aéronefs. C'est ainsi qu'en 1886, l'ingénieur français Clément Ader a construit la machine nommée *Éole*. Ce magnifique engin aux ailes en soie élastique mimant les ailes de la chauve souris, est devenu la machine plus lourde que l'air qui a réalisé le premier décollage motorisé dans l'histoire. Moins de trois décennies plus tard, la première ligne régulière fut ouverte connectant les villes de Tampa et S<sup>t</sup> Petersburg en Floride.

Depuis l'invention du microprocesseur dans les années 1970, on vit dans l'ère de la course à la miniaturisation. La conjoncture de Moore, énoncée en 1965 par Gordon E. Moore, prévoit une diminution par moitié de la taille des microprocesseurs environ tous les deux ans.

Et si on s'approchait d'avantage de la nature à l'échelle microscopique? En s'inspirant des enzymes qui sont des véritables usines miniatures, capables de catalyser des réactions chimiques complexes nous pourrions concevoir des nanomachines biomimétiques pour mieux s'approprier notre monde à l'échelle atomique. Le but est de construire des outils plus efficaces et moins polluants que nos moyens actuels.

On peut prendre l'exemple des hydrogénases et des protéines à radical SAM. Les premières sont capables de produire du dihydrogène moléculaire à partir de protons et d'électrons, tandis que les secondes catalysent un large spectre de réactions, dont la reproduction de certaines pose des problèmes aux chimistes. Ces enzymes présentent donc un grand potentiel pour le développement de la production d'énergie et pour la chimie verte.

La première étape pour un biomimétisme réussi est l'étude approfondie et la compréhension détaillée des mécanismes mis en œuvre par la nature.

Au cours de ma thèse je me suis intéressé à l'enzyme HydE ; une protéine à radical SAM impliquée dans le processus d'assemblage du site actif de l'hydrogénase à [FeFe].

J'ai effectué ces travaux au laboratoire du Groupe Métalloprotéines de l'Institut de Biologie Structurale à Grenoble sous la direction d'Yvain Nicolet et de Juan Carlos Fontecilla-Camps.

Le présent manuscrit retrace l'ensemble de mon travail en trois parties: l'Introduction Bibliographique, Chapitres Résultats et Matériels & Méthodes.



## Introduction Bibliographique

Le réchauffement climatique et la diminution des réserves en combustibles fossiles poussent l'humanité vers la recherche de nouvelles sources d'énergie. La notion de révolution verte est synonyme de diversification des sources d'énergie plus respectueuses de l'environnement. Dans ce contexte le dihydrogène semble pouvoir compléter des sources d'énergie renouvelables comme l'énergie solaire, l'énergie éolienne ou la combustion du biodiesel<sup>1</sup>. Le dihydrogène possède l'avantage de contenir une grande densité énergétique ; 1 kg de H<sub>2</sub> dégage presque trois fois plus d'énergie que 1 kg d'essence (120 MJ.kg<sup>-1</sup> comparé au 45 MJ.kg<sup>-1</sup> pour l'essence). De plus, le produit unique de sa combustion avec le dioxygène est de l'eau. Son utilisation pourrait donc contribuer à la diminution des émissions de gaz à effet de serre, à condition que sa production ne soit ni coûteuse en énergie, ni polluante. Le H<sub>2</sub> est un bon moyen de stockage d'énergie, car sa combustion peut libérer de l'énergie thermique, tandis que son utilisation dans les piles à combustible peut directement fournir de l'électricité<sup>2</sup>. Avant que l'on puisse accéder à une nouvelle économie basée sur l'utilisation du dihydrogène, nous devons résoudre plusieurs problèmes technologiques liés au stockage, à la distribution et à la production de cette molécule. Aujourd'hui, le facteur limitant est sans doute l'absence d'un moyen de production durable et neutre en carbone, car la majorité du H<sub>2</sub> est produite par la reformation du gaz naturel<sup>2</sup>.

L'atome d'hydrogène est très abondant sur notre planète. En effet chaque molécule d'eau contient un atome d'oxygène et deux atomes d'hydrogène. L'eau couvre 70% de la surface de la Terre. La réaction de craquage de l'eau se présente donc comme une stratégie attractive pour la production de dihydrogène à grande échelle. Cependant, l'électrolyse de l'eau est un procédé gourmand en énergie et son couplage avec la production d'électricité par des panneaux photovoltaïques ou des turbines éoliennes manque d'efficacité<sup>3</sup>. Des catalyseurs métalliques à base de platine ont prouvé leur efficacité, mais le coût lié à leur production empêche leur utilisation à l'échelle industrielle<sup>4</sup>.

Il semble cependant que la nature ait déjà trouvé des catalyseurs efficaces, car dans de nombreux organismes, la réduction des protons et l'oxydation du dihydrogène jouent un rôle primordial dans le métabolisme bioénergétique. Ces catalyseurs biologiques sont des enzymes appelées hydrogénases<sup>5-7</sup>. Dans certains organismes, ces enzymes sont impliquées dans une production de dihydrogène couplée à l'oxydation de l'eau via le photosystème II et la chaîne photosynthétique de transport d'électrons<sup>8-10</sup>. C'est une façon de se débarrasser des électrons à bas potentiel. Ces organismes arrivent ainsi à produire du H<sub>2</sub> à partir de l'eau en utilisant de l'énergie solaire.

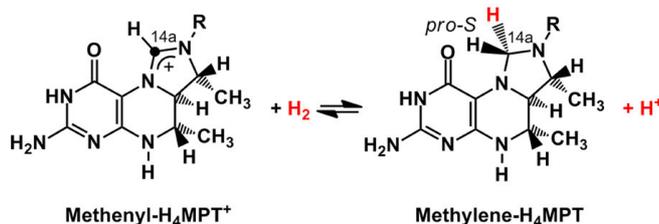
## IB-1 Hydrogénases

Les hydrogénases sont impliquées dans les réactions d'oxydoréduction réversibles avec le dihydrogène moléculaire. Leur rôle physiologique est soit de coupler l'oxydation du H<sub>2</sub> aux processus anaboliques, soit de coupler la production du H<sub>2</sub> à l'élimination de l'excès de pouvoir réducteur, produit par la fermentation ou la photosynthèse<sup>9,11</sup>. On peut trouver ces enzymes chez de nombreux micro-organismes, comme les bactéries, les archées et même certains eucaryotes<sup>8</sup>. A ce jour, nous connaissons trois classes d'hydrogénases phylogénétiquement distinctes: les hydrogénases à [NiFe], les hydrogénases à [FeFe] et les hydrogénases à [Fe].

Les hydrogénases à [FeFe] et à [NiFe] catalysent la réaction de conversion du dihydrogène en protons et électrons et la réaction inverse, c'est-à-dire la production de dihydrogène. Ces transformations sont effectuées en deux étapes:



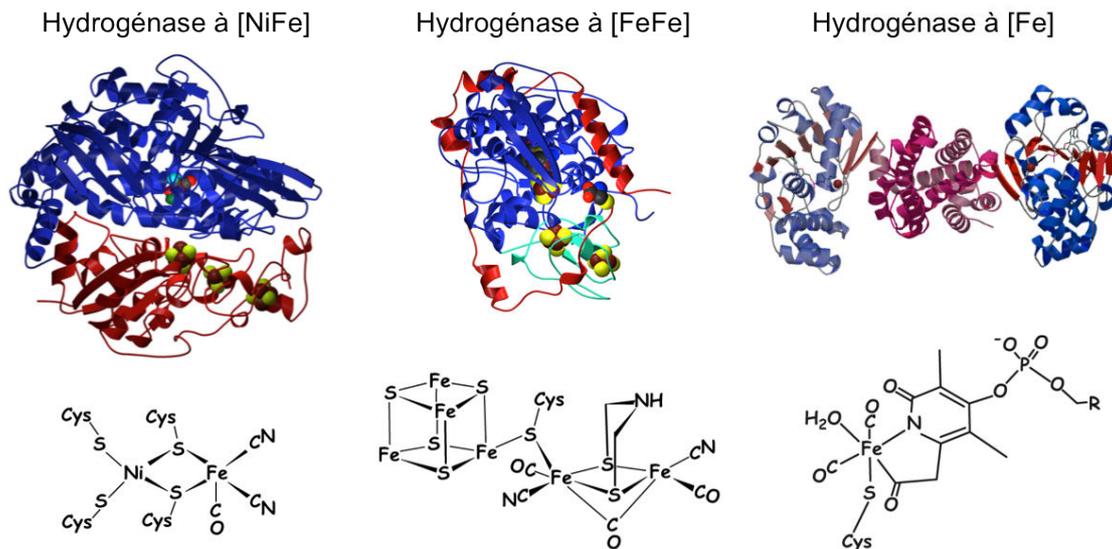
Les hydrogénases à [Fe] constituent un groupe très restreint présent uniquement chez les archées méthanogènes hydrogénotrophiques. Elles sont impliquées dans le processus de la transformation du CO<sub>2</sub> et du H<sub>2</sub> en CH<sub>4</sub>. Elles catalysent la première demi-réaction (Equation 1), ou plus précisément, la réduction réversible du substrat méthényltetrahydrométhanoptérine ou méthényl-H<sub>4</sub>MPT<sup>+</sup> avec de l'hydrure issu du clivage hétérolytique d'une molécule de H<sub>2</sub> [Figure-IB-1]<sup>12,13</sup>.



**Figure-IB-1** : Réaction du méthényl-H<sub>4</sub>MPT<sup>+</sup> avec le dihydrogène catalysée par les hydrogénase à [Fe] (référence<sup>6</sup>)

### IB-1.1 Signes d'une évolution convergente

Ces trois classes d'hydrogénases possèdent des structures tridimensionnelles différentes, mais leurs sites actifs présentent certaines caractéristiques communes, signe d'une évolution convergente. En effet, dans le site actif de toutes ces enzymes nous pouvons trouver des ligands diatomiques CN<sup>-</sup> et/ou CO liés à un atome de Fe [Figure-IB-2]. Il a été proposé que ces ligands stabilisent la forme bas-spin et l'état redox I/II de cet atome de Fe, une propriété essentielle pour l'activation de dihydrogène par ces enzymes<sup>14-16</sup>.

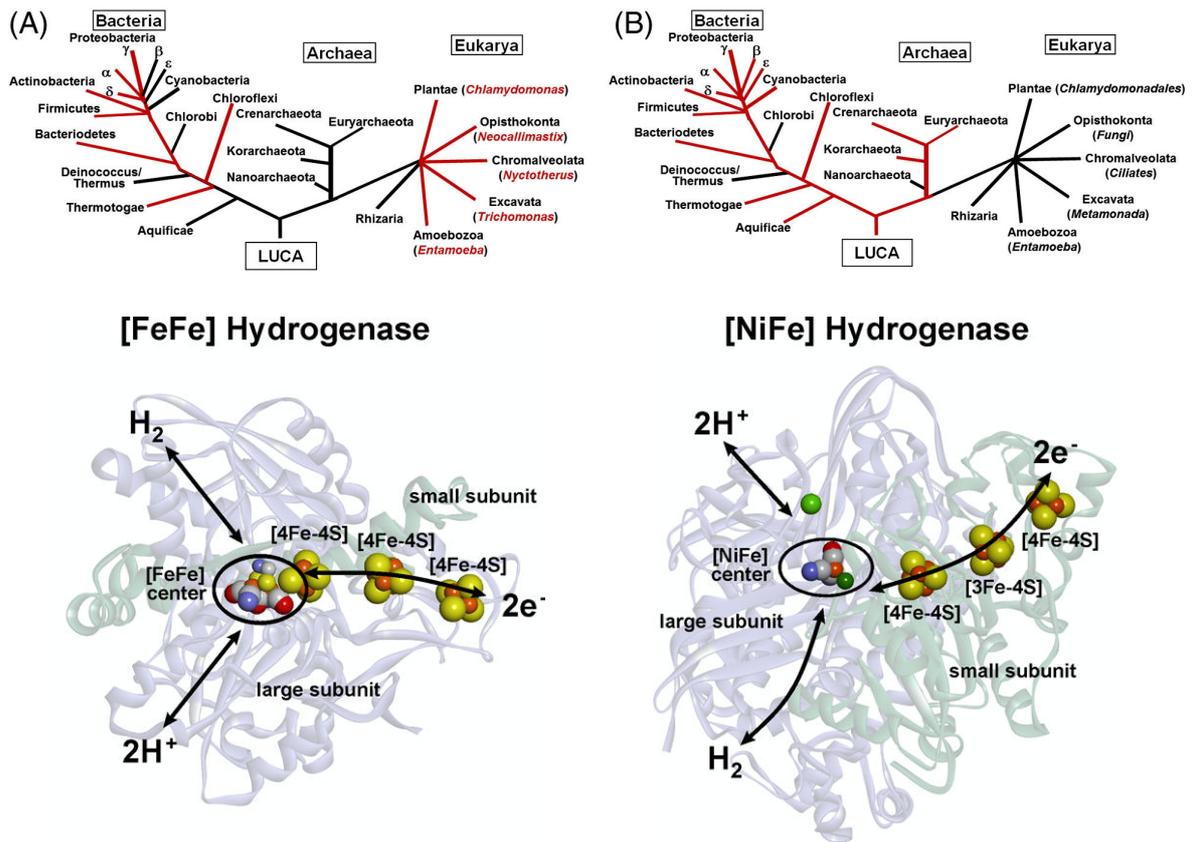


**Figure-IB-2** : Structures cristallographiques et schémas des sites organométalliques des représentants de trois classes d'hydrogénases

De gauche à droite: Structure de l'hydrogénase à [NiFe] de *Desulfovibrio gigas* (code PDB: 1FRV) ; Structure de l'hydrogénase à [FeFe] de *Desulfovibrio desulfuricans* (code PDB: 1HFE) ; Structure de l'hydrogénase à [Fe] de *Methanocaldococcus jannaschii* (code PDB : 3F47) (Image produite par Yvain Nicolet)

Les similarités entre les hydrogénases à [NiFe] et à [FeFe] sont encore plus nombreuses. Toutes les deux utilisent des centres organométalliques binucléaires avec des ligands CO et CN<sup>-</sup>. Les études par cristallographie aux rayons X sur des cristaux protéiques congelés sous pression de xénon ou krypton ont mis en évidence la présence de canaux hydrophobes reliant la surface de la protéine au site actif, servant ainsi d'accès privilégié du dihydrogène au site catalytique de ces enzymes<sup>5,17-20</sup> [Figure-IB-3]. L'analyse des structures tridimensionnelles de ces protéines, combinée aux expériences de mutagenèse dirigée a révélé un réseau de liaisons hydrogène reliant la surface de la protéine au site actif<sup>18,21</sup>. Ce réseau est un chemin de transfert de protons qui interviennent dans le processus d'oxydation réversible de la molécule de H<sub>2</sub>. La comparaison des structures de représentants de ces deux grandes classes d'hydrogénases montre la présence de plusieurs centres [Fe-S] qui relient, eux aussi, la surface de la protéine au site organométallique<sup>18,22,23</sup>. Il s'agit d'un chemin de transfert d'électrons [Figure-IB-3].

La théorie d'évolution convergente est confortée par le fait que les hydrogénases à [FeFe] n'ont pas la même occurrence dans les règnes du vivant que les hydrogénases à [NiFe] également par le fait que leurs centres organométalliques respectifs sont synthétisés par des mécanismes complètement différents<sup>5-7</sup> [Figure-IB-3]. Il semblerait donc que les hydrogénases à [NiFe] et à [FeFe] ont trouvé, de façon indépendante, des solutions très similaires au problème de l'activation biologique de la molécule de dihydrogène<sup>17</sup>.



**Figure-IB-3** : Evolution convergente des hydrogénases à [FeFe] (A) et à [NiFe] (B)

En haut, la distribution taxonomique des deux grandes classes d'hydrogénase. Lignées avec au moins un exemple d'hydrogénase sont colorées en rouge. En bas, le schéma d'accès de dihydrogène au site actif et les deux chaînes de transfert de protons et d'écrons (Images adaptées de références <sup>7</sup> et <sup>6</sup>)

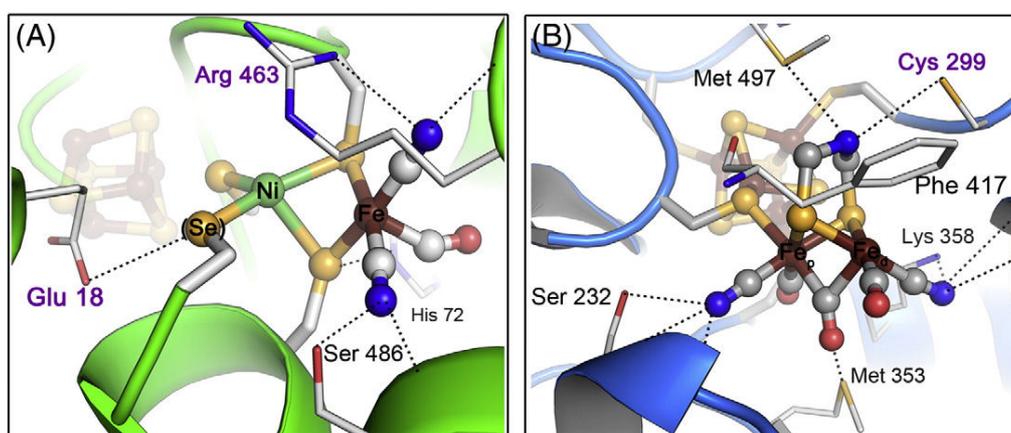
### IB-1.2 Les hydrogénases à [NiFe]

Les hydrogénase à [NiFe] sont présentes dans une grande variété d'organismes, comme les archées anaérobiques strictes et facultatives, les eubactéries et les cyanobactéries, mais sont absentes chez les eucaryotes<sup>7-9</sup> [Figure-IB-3-A]. Elles sont souvent associées au métabolisme énergétique où elles catalysent l'oxydation du dihydrogène pour générer des électrons. Dans leur site actif se trouve un centre hétérobimétallique composé d'un ion Ni coordonné par les thiolates de quatre résidus cystéine. Deux de ces cystéinyles relient le Ni à un atome de Fe sans propriétés redox<sup>6,7,24</sup>. Ce dernier est, en plus, coordonné par deux ions cyanures et une molécule de monoxyde de carbone [Figure-IB-4-A].

Le mécanisme d'assemblage de l'agrégat organométallique des hydrogénases à [NiFe] a été largement étudié et est désormais relativement bien compris<sup>7,25</sup>. L'expression de six gènes (*hypA-hypF*) est nécessaire pour la biosynthèse et l'insertion du centre catalytique [NiFe(CN)<sub>2</sub>CO] dans la grande sous-unité de l'hydrogénase.

### IB-1.3 Les hydrogénases à [FeFe]

Les hydrogénases à [FeFe] sont largement répandues chez les micro-organismes anaérobies où elles permettent le couplage de la réduction de protons avec l'élimination des électrons issus de la fermentation. On les trouve également chez plusieurs eucaryotes inférieurs, comme les algues unicellulaires ou les protistes. Cependant, contrairement aux hydrogénases à [NiFe] elles sont absentes chez les cyanobactéries et les archées. Certaines de ces enzymes sont très actives à température ambiante ; chaque molécule de la protéine peut produire de l'ordre de  $10^4$  à  $10^5$  molécules de  $H_2$  par seconde<sup>26</sup>. Leur site actif, appelé agrégat H, est composé d'un centre  $[Fe_4S_4]$  cubane relié par un groupement thiolate d'un résidu cystéine à une structure organométallique, appelé sous-agrégat à deux Fe. Ce dernier est composé de deux atomes de Fe portant chacun un ligand  $CN^-$  et un ligand CO. Les deux atomes de Fe sont, en plus, connectés entre eux par un ligand CO pontant à l'état oxydé et une petite molécule dithiolées pontante<sup>18,27</sup> [Figure-IB-4-B]. La nature exacte de cette molécule a longtemps fait débat<sup>18,27-30</sup>. La détermination de la structure tridimensionnelle de l'enzyme a mis en évidence un dithiolate composé de cinq atomes<sup>18</sup>. Puisque la résolution du jeu de données cristallographiques ne permettait pas de faire la distinction entre les atomes de C, N ou O, Yvain Nicolet et Juan Fontecilla-Camps ont initialement modélisé la molécule comme un 1,3-propanedithiolate ou PDT<sup>18</sup>. Plus tard, ils ont revu leur interprétation du modèle structural et ils ont proposé, que la molécule pontante soit une dithiométhylamine dont l'amine interviendrait en tant que base dans la polarisation de la molécule de dihydrogène et dans le transfert de protons<sup>27</sup>. Récemment, une étude de Berggren et de ses collaborateurs a confirmé que la seconde interprétation était la bonne<sup>31</sup>.

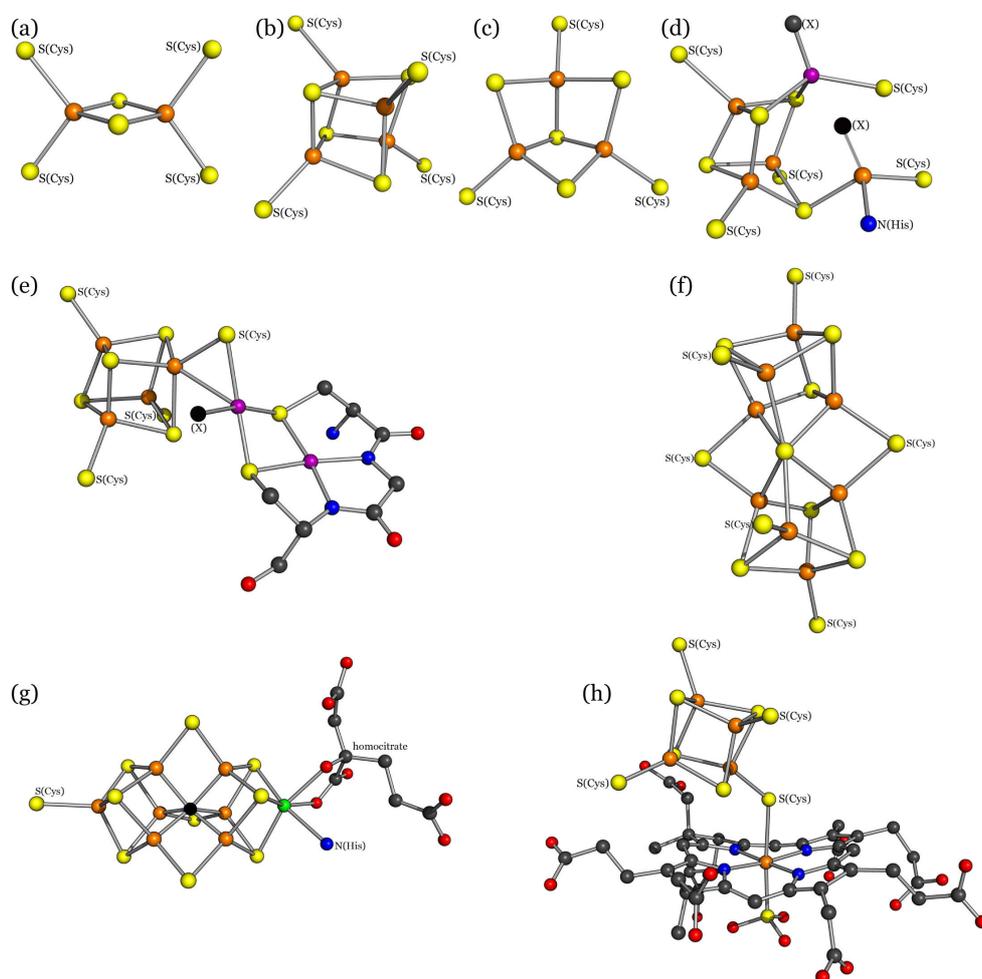


**Figure-IB-4** : Détails des structures cristallographiques des centres organométalliques de l'hydrogénase à [NiFe] et de l'hydrogénase à [FeFe]

**A** - Site actif de l'hydrogénase à [NiFe] de *Desulfovibrio gigas* (code PDB: 2FRV)      **B** - Site actif de l'hydrogénase à [FeFe] *Cpl* de *Clostridium pasteurianum* (code PDB: 3C8Y)

## IB-1.4 Centres [Fe-S]

Les centres [Fe-S] sont des groupements prosthétiques omniprésents dans les trois règnes du vivant. Ils présentent une versatilité structurale et fonctionnelle importante<sup>32</sup> [Figure-IB-5]. Ils sont impliqués dans les processus de transfert d'électron, d'activation de substrat, de maintien de stabilité structurale de protéines, de régulation de l'expression de gènes ou dans les processus d'insertion de soufre dans les molécules biologiques<sup>32-36</sup>. La forme la plus simple est un centre  $[\text{Fe}_1\text{S}_0]^{3+/2+}$  qu'on trouve dans la rubrédoxine. Les formes les plus courantes sont des centres  $[\text{Fe}_2\text{S}_2]^{2+/1+}$ ,  $[\text{Fe}_3\text{S}_4]^{1+/0}$ ,  $[\text{Fe}_4\text{S}_4]^{2+/1+}$  et les centres  $[\text{Fe}_4\text{S}_4]^{3+/2+}$  à haut potentiel redox<sup>37</sup>. La forme la plus complexe est sans doute le cofacteur FeMo-co de la nitrogénase. C'est un centre  $[\text{Fe}_7\text{CS}_9\text{Mo-homocitrates}]$  abritant un carbure central lié à six atomes de Fe<sup>38,39</sup> [Figure-IB-5-g]. Les centres  $[\text{Fe}_4\text{S}_4]^{2+/1+}$  présents dans les hydrogénases et certaines de leurs maturases servent de condensateurs biologiques, capables d'accepter ou de fournir un électron en changeant d'état d'oxydation.



**Figure-IB-5** : Différentes formes des agrégats [Fe-S]

(a) Centre  $[\text{Fe}_2\text{S}_2]$  ; (b) Centre  $[\text{Fe}_4\text{S}_4]$  ; (c) Centre  $[\text{Fe}_3\text{S}_4]$  ; (d) "C-cluster" de la CO-déshydrogénase ; (e) "A-cluster" de l'acétyl-coenzyme A synthase (f) "P-cluster" de la nitrogénase (g) "FeMo-co" de la nitrogénase (h) Centre [Fe-S] de la sulfite réductase. La coloration des atomes est standard. Vert : molybdène ; Violet : nickel.

## IB-2 Maturation du site organométallique des hydrogénases à [FeFe]

La maturation post-translacionnelle des métalloprotéines est une étape essentielle pour leur activité. Ceci a été bien démontré pour de nombreuses protéines à [Fe-S]<sup>33,40</sup> ou dans le cas de la nitrogénase<sup>41</sup>. Les mécanismes de maturation sont très bien régulés, car les métaux de transitions sous forme libre sont toxiques pour la cellule. La complexité du site organométallique de l'hydrogénase à [FeFe] suggérait l'existence de protéines additionnelles impliquées dans l'assemblage et l'insertion d'une telle structure. En effet l'expression du gène codant pour l'hydrogénase à [FeFe] dans un organisme, à la base dépourvu de cette enzyme, conduit à l'accumulation d'une protéine soluble, mais incapable d'activer le dihydrogène<sup>42-44</sup>. Les dosages en fer et en soufre combinés aux analyses spectroscopiques montrent que la protéine ainsi exprimée contient déjà le centre [Fe<sub>4</sub>S<sub>4</sub>] appartenant à l'agrégat H<sup>45-47</sup>. Ce centre [Fe<sub>4</sub>S<sub>4</sub>] est très probablement assemblé par les protéines impliquées dans la biosynthèse des centres [Fe-S] : les machineries ISC et SUF<sup>46,48,49</sup>. Donc les maturases spécifiques de l'activation de l'hydrogénase à [FeFe] doivent synthétiser le sous-agrégat à deux Fe.

### IB-2.1 L'identification de la machinerie Hyd

La découverte de la machinerie Hyd date de 2004, quand Matthew C. Posewitz et ses collaborateurs<sup>10</sup> ont généré des bibliothèques de mutants d'insertion dans le but d'inactiver les gènes qui codent pour les protéines responsables de la maturation de l'hydrogénase à [FeFe] dans l'algue verte unicellulaire *Chlamydomonas reinhardtii* - la protéine CrHydA1. Ils ont pu ainsi identifier trois gènes *hydE*, *hydF* et *hydG*, dont l'expression est indispensable à l'activation de l'hydrogénase<sup>44</sup>. L'organisation de ces gènes dans le génome des organismes possédant l'hydrogénase à [FeFe] se divise en deux groupes. Dans certains organismes ces gènes sont dispersés; dans d'autres ils sont regroupés en véritable opéron *Hyd*<sup>25</sup>. L'analyse de ces opérons a révélé que le gène *aspA*, codant pour l'enzyme aspartate ammonia lyase, qui catalyse la transformation réversible de l'aspartate en fumarate et ammonium, est souvent associée au trio de maturases<sup>50</sup>.

L'analyse des séquences en acides aminés des trois maturases a montré que HydE, HydF et HydG possèdent plusieurs résidus cysteines strictement conservés - une caractéristique des protéines à centres [Fe-S]. La fixation des centres [Fe-S] par les maturases a été caractérisée dans de nombreuses études<sup>10,44,51,52</sup>. La séquence de la protéine HydF contient les motifs A et B de Walker - la signature des GTP-ase<sup>10,51</sup>, tandis que les protéines HydE et HydG appartiennent à la superfamille des enzymes à radical SAM<sup>9,33,39,40</sup>.

La coexpression hétérologues, dans *E. coli* (un organisme qui ne contient naturellement pas d'hydrogénase à [FeFe]), de l'enzyme CrHydA1 avec les maturases HydE, HydF et HydG produit une enzyme CrHydA1<sup>EFG</sup> active<sup>10,44,45</sup>. De même, un extrait cellulaire contenant les trois maturases co-exprimées ensemble est capable d'activer l'hydrogénase CrHydA1<sup>ΔEFG</sup> exprimée séparément<sup>45</sup>. Plus important encore, la protéine HydF<sup>EFG</sup>, co-exprimée avec les protéines HydE et HydG, est capable, après purification, de maturer une hydrogénase inactive<sup>48</sup>. Une telle activation nécessite une reconstitution préalable du centre

[Fe<sub>4</sub>S<sub>4</sub>] de l'hydrogénase<sup>45,48</sup>. Toutes ces données ont convergé vers le modèle, dans lequel les trois maturases assemblent un sous-agrégat à deux Fe sur la maturase HydF, qui le transfère ensuite à l'hydrogénase possédant déjà un centre [Fe<sub>4</sub>S<sub>4</sub>]<sup>48</sup>. La protéine HydF agit donc comme une protéine d'échafaudage. Ces expériences montrent en plus, que les substrats nécessaires à la production de l'agrégat H sont présents parmi les métabolites d'*E. coli*<sup>44,45</sup>.

### IB-2.2 Maturase HydF - une GTPase contenant un centre [Fe-S]

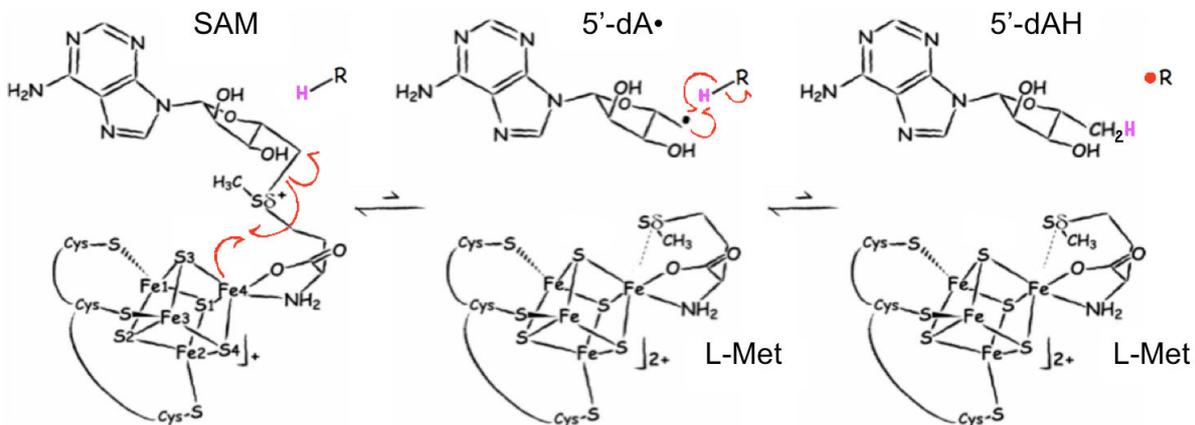
La maturase HydF est une protéine à centre [Fe-S] capable d'hydrolyser le GTP en GDP<sup>51,53</sup>. Elle possède un motif CXHX<sub>46-53</sub>CX<sub>2</sub>C conservé, impliqué dans la fixation d'un centre [Fe-S]<sup>10</sup>. La forme exacte de cet agrégat a longtemps été sujette à polémiques<sup>48,51,53-58</sup>. Aujourd'hui, il est admis que dans la protéine HydF<sup>ΔEG</sup> exprimée sans les deux autres maturases, la forme physiologique de ce centre est un agrégat [Fe<sub>4</sub>S<sub>4</sub>]<sup>7</sup>. Le quatrième ligand de cet agrégat reste inconnu. Dans une première étude le ligand du Fe unique a été identifié comme de l'imidazole issu de la purification de la protéine<sup>51</sup>. Plus tard il a été proposé que le ligand physiologique du Fe unique dans la protéine HydF<sup>ΔEG</sup> pourrait être la chaîne latérale d'un résidu histidine ou un ligand hydroxyle facilement échangeable, mais les contradictions entre les données expérimentales persistent<sup>56,57,59,60</sup>. Lorsque la protéine HydF est exprimée avec les maturases HydE et HydG, elle abrite un centre [Fe<sub>4</sub>S<sub>4</sub>] et un centre [Fe<sub>2</sub>S<sub>2</sub>]<sup>53</sup>. Les expériences de FTIR sur la protéine HydF<sup>EG</sup> ont révélé la signature des ligands CN<sup>-</sup> et CO<sup>55</sup>. Il a été proposé que HydF<sup>EG</sup> abritait un centre [Fe<sub>4</sub>S<sub>4</sub>] lié à un sous-agrégat à deux Fe similaire à l'agrégat H de l'hydrogénase<sup>54</sup>. Les expériences de mutagenèse dirigée montrent que l'intégrité du domaine GTP-ase et du site de fixation du centre [Fe-S] sont essentielles à l'assemblage du sous-agrégat à deux Fe<sup>44</sup>. Le rôle de l'activité GTP-ase dans l'assemblage du sous-agrégat à deux Fe n'est cependant toujours pas compris. Il a été prouvé que l'hydrolyse de GTP n'est pas indispensable à l'étape du transfert du sous-agrégat à deux Fe à partir de la maturase HydF<sup>EG</sup> vers HydA1<sup>ΔEFG</sup><sup>45,53</sup>. Cette activité est relativement faible par rapport à d'autres enzymes GTP-ase. Contrairement à la présence de l'hydrogénase, celle des maturases HydE ou HydG fait augmenter l'hydrolyse du GTP en GDP d'environ 50%<sup>53</sup>. Ces résultats suggèrent que la fixation et l'hydrolyse du GTP sont associées à l'interaction entre HydF et les deux autres maturases et donc à la synthèse du sous-agrégat à deux Fe plutôt qu'à son transfert dans l'hydrogénase.

La seule structure tridimensionnelle connue est celle de HydF de *Thermotoga neapolitana*, résolue en 2011<sup>59</sup>. La maturase contient trois domaines distincts; le domaine N-terminal qui contient le motif de fixation de GTP, le domaine C-terminal qui contient le motif de fixation du centre [Fe-S] et un domaine de dimérisation se situant entre les deux premiers. Dans le cristal protéique, la maturase forme un dimère de dimère dépourvu de GTP et de centres [Fe<sub>4</sub>S<sub>4</sub>], il est donc difficile d'en extraire des informations sur le processus d'assemblage du sous-agrégat à deux Fe.

Il est intéressant de noter que plusieurs machineries de maturation de métalloprotéines possèdent des enzymes avec une fonction NTP-ase<sup>53</sup>. Cependant leur rôle exact dans les processus de maturation des centres organométalliques reste à établir.

### IB-3 Protéines à radical SAM - Généralités

L'analyse des séquences en acides aminés a révélé que les maturases HydE et HydG font partie de la superfamille des enzymes à radical *S*-adénosyl-L-méthionine (SAM) identifiée en 2001<sup>10,52,61</sup>. Comme les protéines à adénosylcobalamine (un sous-groupe d'enzymes à B<sub>12</sub>) ces enzymes utilisent un radical alkyl hautement oxydant - le radical 5'-déoxyadénosyl ou 5'-dA• - pour arracher un atome d'hydrogène sur le substrat et déclencher ainsi sa conversion en produit. En arrachant un H• spécifique à partir de leur substrats, elles arrivent à catalyser des réactions qui seraient difficiles pour une chimie organique à deux électrons, car sur des carbones non activés et donc très peu réactifs dans les conditions physiologiques. Contrairement aux protéines à adénosylcobalamine, elles se servent d'un centre [Fe<sub>4</sub>S<sub>4</sub>] de bas potentiel redox pour initier le clivage réductif homolytique d'une molécule de la SAM et former la L-méthionine et le radical 5'-dA•<sup>62-64</sup> [Figure-IB-6].

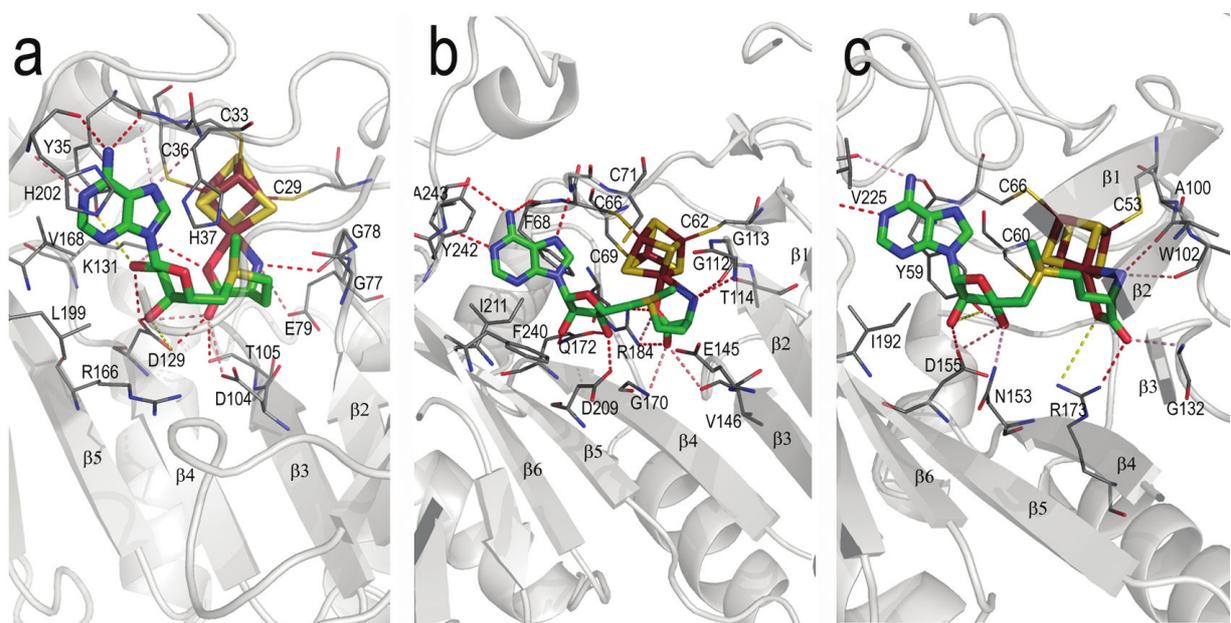


**Figure-IB-6 : Mécanisme de clivage homolytique réductif de la SAM**

Dans un premier temps le centre [Fe<sub>4</sub>S<sub>4</sub>]<sup>2+</sup> 'SAM radical' est réduit par un réducteur physiologique tel le système NADPH - flavodoxine - flavodoxine réductase. Le transfert mono-électronique du centre [Fe<sub>4</sub>S<sub>4</sub>]<sup>1+</sup> à la SAM, par l'intermédiaire de l'atome de fer unique, déclenche la coupure homolytique de la liaison S-C5' du sulfonium de la SAM, conduisant à la production d'une molécule de L-méthionine et d'un radical 5'-déoxyadénosyle (5'-dA•) hautement réactif. Ce radical arrache alors un atome d'hydrogène du substrat pour former la 5'-déoxyadénosine (5'-dAH) et un radical substrat R•. Ce dernier va ensuite évoluer en produit.

Les protéines à radical SAM interviennent notamment dans les processus de biosynthèse des vitamines, cofacteurs, dans l'activation des protéines à radical glycylyl, dans la réparation de l'ADN, dans la modification de l'ARN ou dans l'assemblage des sites actifs organométalliques<sup>65-68</sup>. Ces protéines partagent une homologie de séquences en acides aminés très limitée. Leur motif caractéristique est la séquence conservée Cx<sub>3</sub>CxφC où φ représente un résidu tyrosine, phénylalanine, histidine ou tryptophane<sup>61,69</sup>. Les trois cystéines sont impliquées dans la fixation du centre [Fe<sub>4</sub>S<sub>4</sub>] 'SAM radical'. Ce dernier étant redox actif est sensible à l'oxygène. Le résidu φ protège l'agrégat [Fe-S] du solvant<sup>70</sup>. Les expériences ENDOR et Mössbauer ont montré qu'avant le clivage de la SAM, les ligands du Fe unique ou Fe<sub>u</sub> du centre [Fe<sub>4</sub>S<sub>4</sub>] différenciés sont le groupement amine et un atome d'oxygène du groupement carboxylate de la partie L-méthionine de la SAM<sup>71-73</sup>. Ces ligands fixent le Fe<sub>u</sub> de façon bidentate [Figure-IB-6 & Figure-IB-7]. Après le clivage, l'atome de soufre de la L-

méthionine devient le sixième ligand du  $Fe_u^{62}$ . Ces interactions ont été confirmées par de nombreuses structures tridimensionnelles déterminées par la cristallographie aux rayons X<sup>74-81</sup>. Le groupement amine de la partie L-méthionine de la SAM interagit également avec une région riche en glycine - le motif GGE<sup>61,69</sup>. Le groupement carboxylate présente des interactions moins conservées. Les deux groupements hydroxyles du ribose peuvent interagir avec des résidus chargés ou polaires de nature variée<sup>70</sup>. Les interactions entre la protéine et la partie adénine de la SAM sont bien conservées et facilement identifiables dans les séquences en acides aminés des protéines à radical SAM. L'adénine se fixe dans une poche composée de 3-5 acides aminés hydrophobes et y forme entre 3-4 liaisons hydrogène conservées<sup>70</sup>. Ces liaisons font intervenir des résidus d'un motif GxIGx<sub>2</sub>E et la chaîne principale du résidu  $\phi$  appartenant au motif Cx<sub>3</sub>Cx $\phi$ C<sup>69,70</sup>. Les interactions entre la molécule de la SAM et les éléments de la structure protéiques sont illustrées sur la [Figure-IB-7](#) ci-dessous.

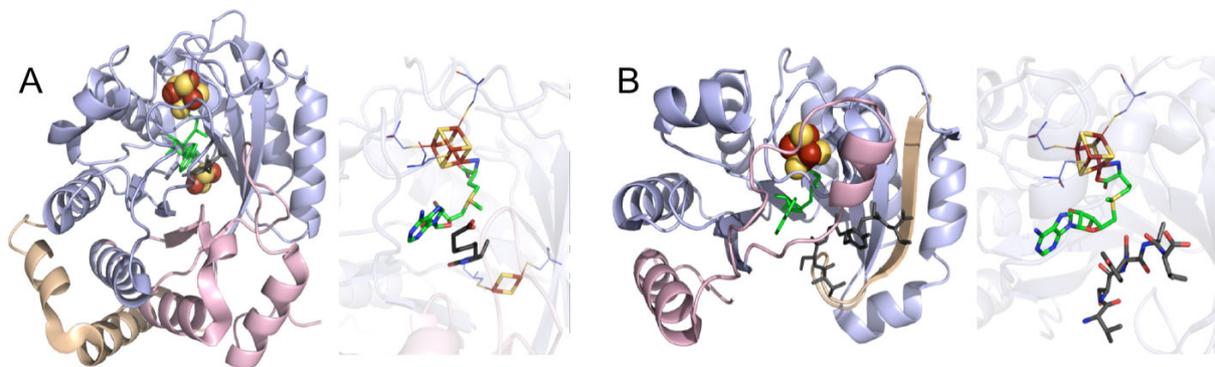


**Figure-IB-7** : Détails de la fixation de la SAM dans les sites actifs de trois protéines à radical SAM différentes

**a** - pyruvate-formate lyase activating enzyme ou PFL-AE (code PDB : 3CB8). Cette enzyme catalyse le transfert du radical sur la pyruvate-formate lyase, qui est une protéine à radical glycy.**b** - Coproporphyrinogen III oxydase ou HemN (code PDB : 10LT) catalyse la décarboxylation de propionates en vinyles dans la biosynthèse du hème **b** **c** - Biotine synthase ou BioB (code PDB : 1R30) catalyse la réaction d'insertion de soufre dans la déthiobiotine. (Image adaptée de Vey *et al.* 2011)

La majorité des protéines à radical SAM possèdent un repliement composé de 6 brins  $\beta$  et de 6 hélices  $\alpha$  arrangé de façon similaire au tonneau  $(\beta/\alpha)_8$  de la triose-phosphate isomérase ou TIM<sup>69</sup>. Il s'agit d'un repliement en tonneau TIM partiel ou '3/4'. Certaines enzymes à radical SAM possèdent deux brins  $\beta$  et deux hélices  $\alpha$  supplémentaires et présentent un repliement en tonneau  $(\beta/\alpha)_8$  complet. Les parties de structure supplémentaires contiennent des éléments spécifiques à la fonction de l'enzyme (centres [Fe-S] auxiliaires, sites de fixations du substrat ou d'un cofacteur additionnel et/ou surfaces d'oligomérisation)<sup>70,82</sup>. Peu de temps après l'identification de la superfamille de protéines, Yvain Nicolet et Catherine L. Drennan ont remarqué une corrélation intéressante entre le

repliement de l'enzyme et la taille du substrat<sup>69</sup>. En effet, les enzymes possédant un repliement en tonneau ( $\beta/\alpha$ )<sub>8</sub> complet présente une cavité à l'intérieur de la structure et reconnaissent des substrats de petites tailles. En absence des deux brins  $\beta$  et hélices  $\alpha$  supplémentaires, la protéine possède un repliement plus ouvert, qui lui permet d'accueillir des substrats bien plus volumineux, comme les parties de protéines ou d'acides nucléiques<sup>70</sup> [Figure-IB-8]. Les enzymes à radical SAM arrivent alors à catalyser plus de 70 réactions différentes sur une grande variété de substrats. Elles sont souvent impliquées dans le clivage ou la formation des liaisons C-C, C-N, C-S, C-P lors des réarrangements complexes ou des réactions d'oxydation en milieu anoxique<sup>65</sup>.



**Figure-IB-8** : Comparaison des structures de protéines à radical SAM BioB et PFL-AE

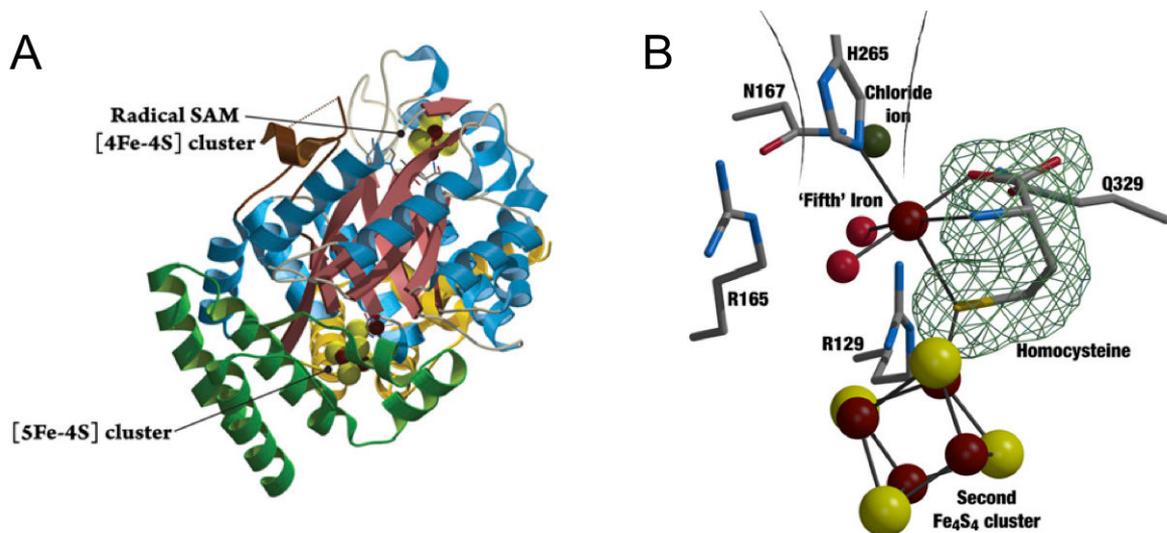
**A** - structure globale de la biotine synthase BioB (code PDB: 1R30) et le détail de son site actif avec le substrat. Le domaine N-terminal est en beige, la partie du domaine SAM radical avec un repliement en tonneau ( $\beta/\alpha$ )<sub>6</sub> en bleu et le domaine C-terminal, qui dans cette structure complète le tonneau ( $\beta/\alpha$ )<sub>8</sub> complet est en rose. **B** - structure globale de la PFL-AE (code PDB: 3CB8) et le détail du site actif avec le substrat. Le domaine N-terminal est en beige, le domaine SAM radical avec un repliement en tonneau ( $\beta/\alpha$ )<sub>6</sub> en bleu et le domaine C-terminal en rose. Dans les deux structures le cofacteur est en vert et le substrat en noir. Le repliement en tonneau TIM complet de la BioB crée une cavité interne qui peut accueillir des substrats de faibles tailles moléculaires, comme ici un centre [Fe<sub>2</sub>S<sub>2</sub>] et la déthiobiotine. Le repliement plus ouvert en tonneau TIM '3/4' fait que la PFL-AE peut avoir pour substrat une protéine entière, ici représentée par un peptide qui mime la partie de la protéine cible de l'attaque radicalaire. (Image adaptée de Vey *et al.* 2011)

Dans la plupart des cas, la SAM est utilisée comme co-substrat, car elle est consommée et se transforme en L-méthionine et 5'-deoxyadenosine ou 5'-dAH<sup>65,83-85</sup>. Dans certaines réactions de réarrangement par contre, la molécule de SAM sert de cofacteur, car elle est régénérée à la fin de chaque cycle<sup>86-88</sup>.

### IB-3.1 Maturase HydG - une enzyme à radical SAM trifonctionnelle

La protéine HydG possède, en plus de son centre [Fe<sub>4</sub>S<sub>4</sub>] 'SAM radical', un second centre [Fe-S] indispensable pour la maturation de l'hydrogénase<sup>44,52,89</sup>. Cet agrégat auxiliaire se fixe sur le motif Cx<sub>2</sub>Cx<sub>22</sub>C strictement conservé. La maturase HydG présente une homologie de séquence en acides aminés significative avec une autre protéine à radical SAM - ThiH<sup>90</sup>. Cette dernière est impliquée dans la biosynthèse de thiamine pyrophosphate et catalyse le clivage de la L-tyrosine en *p*-crésol et déhydroglycine ou DHG<sup>65,91</sup>. Il a donc été

proposé que HydG utilise comme substrat l'acide aminé libre la L-tyrosine<sup>90</sup>. Par la suite il a été prouvé que HydG catalyse le clivage homolytique de la liaison C $\alpha$ -C $\beta$  de cet acide aminé en lui arrachant un atome d'hydrogène sur son groupement amine<sup>74,76,90,92</sup>. La maturase produit les ligands diatomiques CO et CN<sup>-</sup> à partir de la partie DHG du substrat<sup>93,94</sup>. Les expériences de maturation de l'hydrogénase à [FeFe] avec de la tyrosine marquée ont révélé que cette molécule fait augmenter l'activation de l'enzyme et que les ligands CO et CN<sup>-</sup> de son agrégat H proviennent de la L-tyrosine<sup>95,96</sup>. Les expériences de spectroscopie FTIR ont montré que le second centre [Fe-S] sert de précurseur pour la production d'un synthon [Fe(CO)<sub>x</sub>(CN<sup>-</sup>)<sub>y</sub>] (avec x = 1 ou 2 et y = 1) qui se retrouvera dans le site catalytique de l'hydrogénase à [FeFe]<sup>97</sup>. La protéine possède un repliement en tonneau ( $\beta/\alpha$ )<sub>8</sub> complet, avec un domaine C-terminal attaché du côté extérieur en bas du tonneau<sup>74,98,99</sup> [Figure-IB-9-A]. C'est ce domaine qui abrite le site de fixation du centre [Fe-S] auxiliaire. Dans une des structures cristallographiques disponibles ce site contient un centre [Fe<sub>4</sub>S<sub>4</sub>] lié par une L-homocystéine libre à un cinquième Fe, lui-même coordonné par la chaîne latérale d'un résidu histidine conservé et par deux molécules d'eau<sup>99,100</sup> [Figure-IB-9-B].

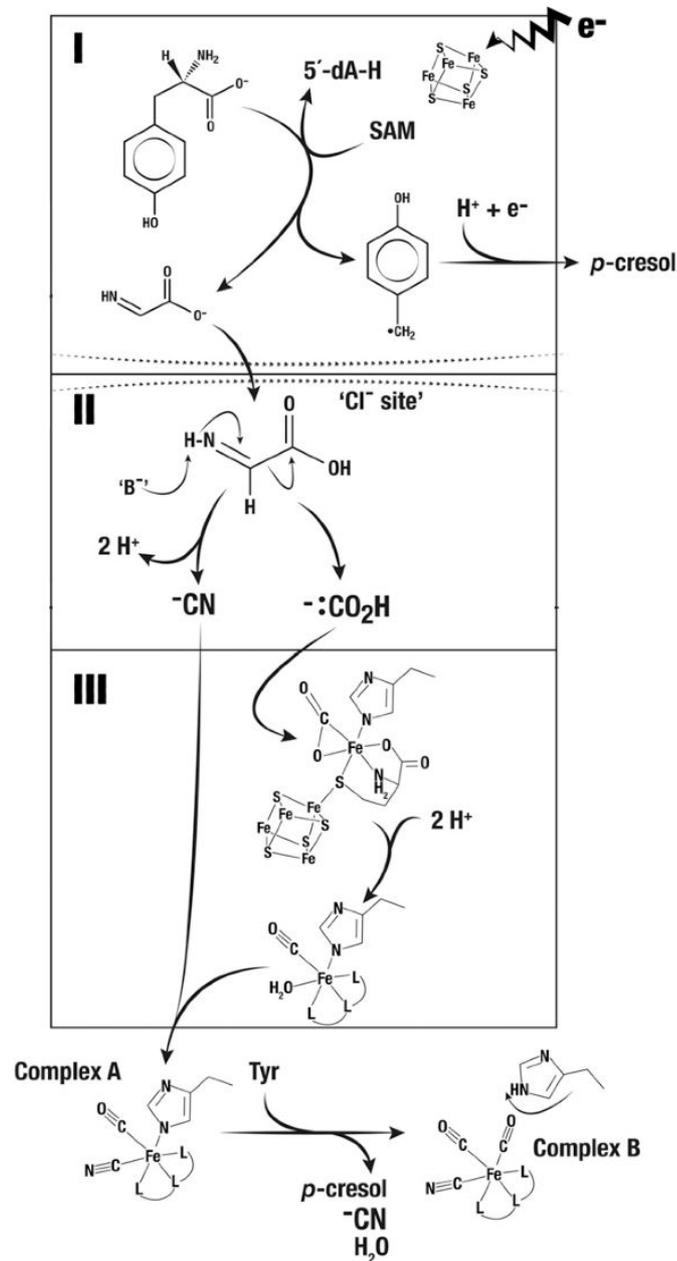


**Figure-IB-9** : Structure globale (A) et détail de l'environnement (B) de l'agrégat auxiliaire [Fe<sub>5</sub>S<sub>4</sub>] de la protéine à radical SAM HydG de *Thermoanaerobacter italicus*

A - Les deux centres [Fe-S] conservés se situent de part et d'autre de l'axe du tonneau ( $\beta/\alpha$ )<sub>8</sub> du domaine SAM radical, ici en rose et bleu. En jaune, le domaine N-terminale et en vert le domaine C-terminal responsable de la production du CO. La boucle, qui couvre le site actif du clivage de la L-tyrosine est partiellement désordonnée et elle est colorée en brun. (code PDB: 4WCX). B - Structure de l'agrégat auxiliaire [Fe<sub>5</sub>S<sub>4</sub>]-L-homocystéine. Une réinterprétation de données cristallographiques enregistrées par Dinis *et al.* effectuée par Pagnier *et al.* révèle la présence d'une molécule de L-homocystéine assurant la fixation du cinquième Fe sur un centre [Fe<sub>4</sub>S<sub>4</sub>] cubane. En vert, carte de densité électronique "omit" contournée à 7  $\sigma$ . (Images adaptées de Pagnier *et al.* 2015)

Les expériences de mutagenèse dirigée montrent que l'intégrité du centre [Fe-S] auxiliaire n'est pas indispensable pour la formation du ligand CN<sup>-</sup> à partir de la DHG, mais elle est cruciale pour la production du CO<sup>89,92,100</sup>. Les auteurs ont proposé que la maturase HydG est une enzyme trifonctionnelle; dans un premier temps elle catalyse, dans la partie supérieure du tonneau ( $\beta/\alpha$ )<sub>8</sub>, le clivage homolytique de la L-tyrosine, pour obtenir du *p*-crésol et de la DHG. Cette dernière migre le long de l'axe du tonneau vers le site de fixation du

second centre [Fe-S]. Au cours de cette migration elle est transformée en ligand  $\text{CN}^-$  et en  $^-:\text{CO}_2\text{H}$ . Cette dernière espèce sera ensuite transformée sur le cinquième Fe en ligand CO, ce qui sera accompagné par la libération d'une molécule de  $\text{H}_2\text{O}$  ou de  $\text{HO}^-$  <sup>100</sup>.



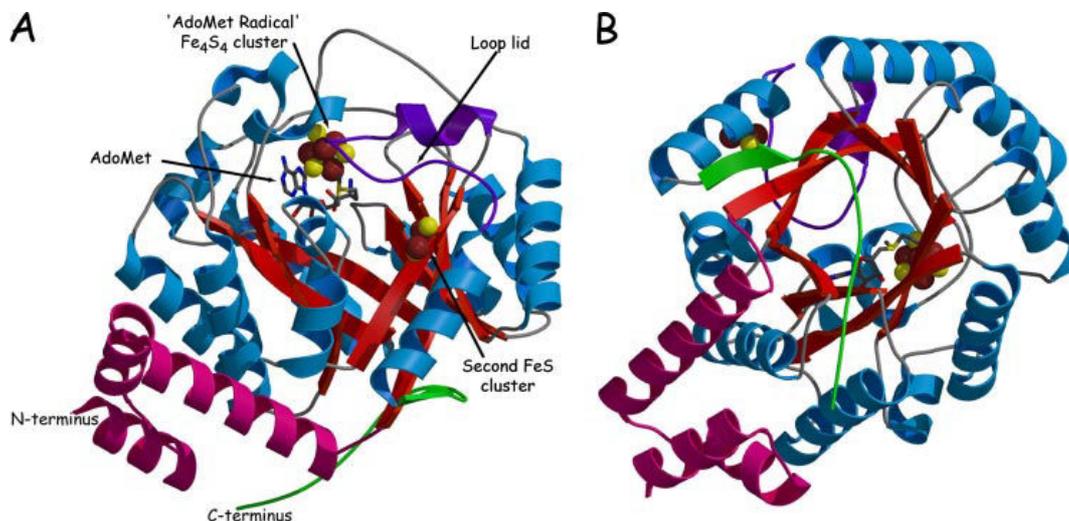
**Figure-IB-10 :** Schéma du fonctionnement des trois sites catalytiques dans l'enzyme à radical SAM HydG

I - Le clivage de la SAM et celui de la L-tyrosine a lieu dans la partie supérieure de la cavité du tonneau ( $\beta/\alpha$ )<sub>8</sub>. Cette réaction produit de la 5'-dAH, la L-méthionine, la déhydroglycine et le radical *p*-crésyl. Ce dernier va arracher un atome d'hydrogène d'une source inconnue et se transformer en *p*-crésol et diffuser hors du site actif. II - DHG franchit une couronne de résidus hydrophobes et arrive au second site actif où elle va subir une réaction d'élimination pour former le cyanure et l'espèce  $^-:\text{CO}_2\text{H}$ . Cette dernière va réagir sur le Fe<sub>u</sub> du centre auxiliaire [Fe<sub>5</sub>S<sub>4</sub>] et se transformer en ligand CO. Les deux ligands  $\text{CN}^-$  et CO peuvent alors lier le Fe<sub>u</sub> et produire le Complexe A. La formation du Complexe B est sans doute précédée par le clivage d'une seconde molécule de L-tyrosine et s'accompagne de libération d'une molécule de  $\text{CN}^-$ . Les expériences de spectroscopie FTIR 'stopped flow' ont révélé que le Complexe A est le premier à être synthétisé et qu'avec le temps il y a l'accumulation du le Complexe B. (Figure adaptée de Pagnier *et al.* 2015)

### IB-3.2 Maturase HydE - source potentielle du ligand dithiométhylamine

La protéine HydE de *Thermotoga maritima* contient deux sites de fixation de centre [Fe-S]. Le premier site est constitué du motif  $Cx_3Cx\phi C$ , caractéristique des protéines à radical SAM et le second site est constitué du motif semi-conservé  $Cx_7Cx_2C$ , présent uniquement chez un sous-groupe de maturases HydE [voir l'alignement multiséquence en Annexe]. Les expériences de spectroscopie UV-visible et RPE et les structures cristallographiques montrent que le site semi-conservé peut accueillir une large variété de centres [Fe-S] différents<sup>50,52,62,75</sup>. Par exemple le dosage de fer montre qu'en solution l'agrégat devrait exister sous la forme d'un centre  $[Fe_4S_4]$ <sup>50,52</sup>. Dans les cristaux<sup>62,75</sup> cet agrégat prend la forme d'un centre  $[Fe_2S_2]$  ou d'un centre  $[Fe_2S_2]$  avec deux persulfures sur les Cys<sup>319</sup> et Cys<sup>322</sup>. Contrairement au site de fixation d'agrégat 'SAM radical', l'agrégat auxiliaire n'est pas indispensable à la maturation de l'hydrogénase à  $[FeFe]$ <sup>75,101</sup>. Le substrat et le produit de l'enzyme HydE ne sont pas connus, mais puisque HydF est une protéine d'échafaudage et HydG fournit les atomes de Fe et les deux ligands diatomiques CO et CN, par élimination HydE devrait fabriquer la dithiométhylamine ou son précurseur<sup>102</sup>.

Les structures cristallographiques de la protéine *TmHydE* montrent que le premier site se situe dans la partie supérieure du tonneau  $(\beta/\alpha)_8$  complet et contient le centre  $[Fe_4S_4]$  'SAM radical'<sup>75</sup>. Le centre [Fe-S] auxiliaire qui se lie aux trois cystéines Cys<sup>311</sup>, Cys<sup>319</sup> et Cys<sup>322</sup>, se situe à la surface de la protéine, où il est exposé au solvant [Figure-IB-11].

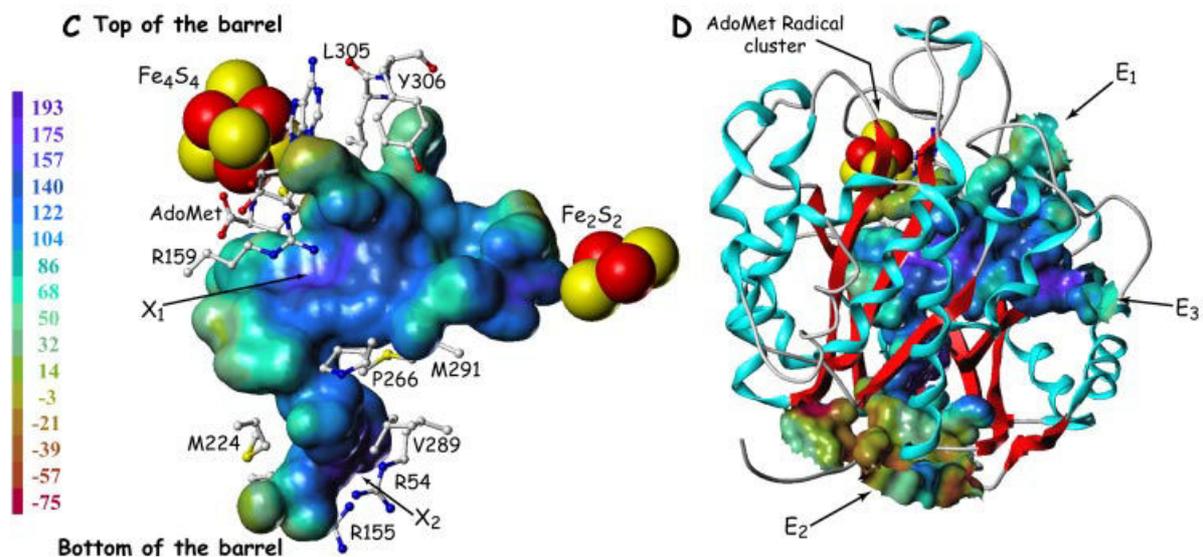


**Figure-IB-11** : Structure cristallographique de la maturase HydE de *Thermotoga maritima*

(A) et (B) sont des vues orthogonales de la structure de *TmHydE*. Le tonneau  $(\beta/\alpha)_8$  est représenté en rouge et bleu. Le domaine N-terminal est coloré en magenta. La boucle Ser297-Lys315, qui couvre la partie supérieure du tonneau est colorée en violet. De l'autre côté, la partie basse du tonneau est protégée du solvant par un brin  $\beta$  C-terminal, ici en vert.

A l'intérieur de la structure on trouve une vaste cavité au potentiel électrostatique positif, remplie de solvant. La partie supérieure du tonneau est couverte par une boucle connectant le brin  $\beta$  8 et l'hélice  $\alpha$  8. Cette boucle contient plusieurs résidus hautement ou strictement conservés, comme la Tyr<sup>303</sup>, Tyr<sup>306</sup>, Lys<sup>309</sup> ou Leu<sup>305</sup>. Cette boucle a son équivalent dans d'autres protéines à radical SAM qui partagent la même topologie que HydE et joue sans doute un rôle dans le contrôle d'accès et de protection du site catalytique. Dans HydE, l'autre extrémité du tonneau est également protégée par un brin  $\beta$  C-terminal [Figure-IB-11].

La cavité de HydE est séparée en deux volumes distincts par le biais d'une couronne de résidus hydrophobes conservés. Dans la partie supérieure se situent le site de fixation du centre [Fe<sub>4</sub>S<sub>4</sub>] 'SAM radical' et le site de la fixation de la SAM. Ainsi nous pouvons dire que cette partie supérieure de la cavité est le site d'initiation de la réaction radicalaire sur le substrat. Lorsque le site de fixation de centre [Fe-S] auxiliaire n'abrite pas d'agrégat, la cavité est connectée au solvant par trois canaux E<sub>1</sub>, E<sub>2</sub> et E<sub>3</sub> [voir Figure-IB-12 ci-dessous].

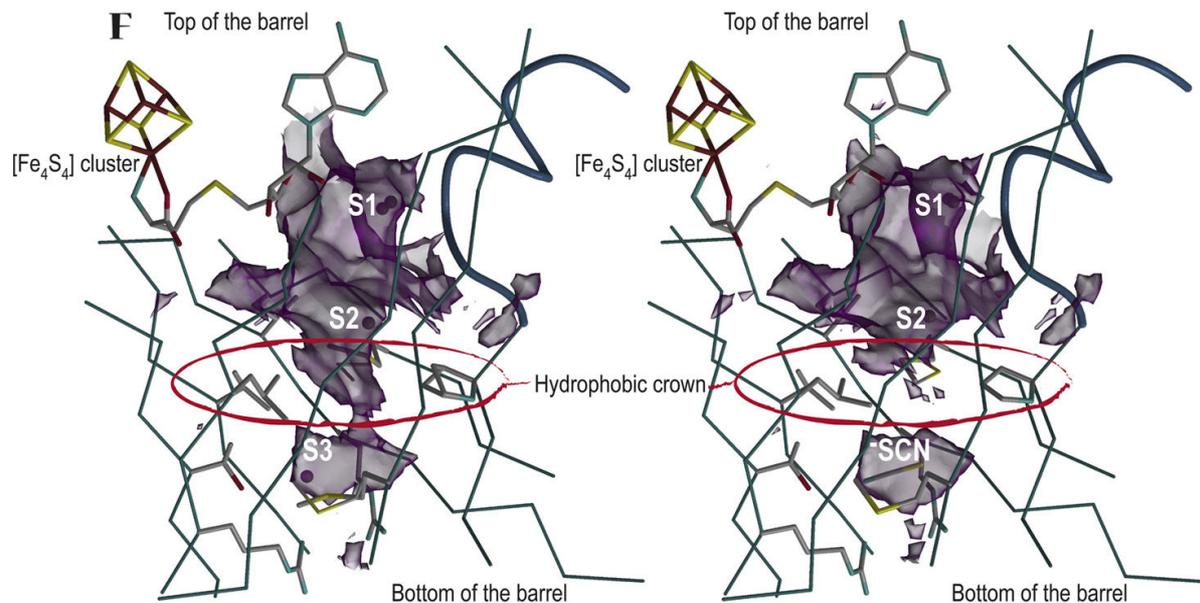


**Figure-IB-12** : Visualisation de la cavité de *TmHydE*

**C** - La cavité a été calculée en utilisant une sonde de diamètre de 1,2 Å. Elle est colorée en fonction de son potentiel électrostatique (voir la barre de potentiel électrostatique en kcal.mol<sup>-1</sup>.electron<sup>-1</sup> à gauche de l'image). Les résidus clés et les deux centres [Fe-S] sont également représentés. La cavité a un potentiel global électropositif. Deux pics de potentiels X<sub>1</sub> et X<sub>2</sub> se distinguent clairement. Ils correspondent aux sites de fixation d'anions S1 et S3, respectivement. **D** - Lorsque la cavité est calculée en absence du second centre [Fe-S] en utilisant un sonde de diamètre de 1 Å, trois canaux la reliant à la surface de la protéine peuvent être visualisés. (Images adaptées de Nicolet *et al.* 2008)

Les expériences de trempage de cristaux protéiques avec les sels de bromure ont mis en évidence trois sites de fixations d'anions disposés linéairement le long de l'axe du tonneau ( $\beta/\alpha$ )<sub>8</sub>. Les sites S1 et S2 se situent dans la partie supérieure de la cavité à proximité du site de fixation de la SAM. Le site S1 interagit avec l'hydroxyle de la chaîne latérale du résidu Tyr<sup>306</sup> et avec l'atome d'azote du résidu Ala<sup>270</sup>. Le premier est strictement conservé, tandis que le second fait partie d'un motif hautement conservé P<sup>266</sup>ATTA<sup>270</sup>. Ce site S1 peut contenir un ion chlorure en position alternative et a une affinité pour le groupement carboxylate. Le site S2

interagit avec la chaîne latérale de l'Arg<sup>159</sup>. Cette dernière fait partie d'un motif hautement conservé R<sup>155</sup>YLLRXE<sup>161</sup>. Ce même motif porte également le résidu strictement conservé Glu<sup>161</sup>, qui interagit avec les deux groupements hydroxyles de la partie ribose de la SAM. Le dernier site de fixation d'anion S3 se situe au fond de la cavité. Les expériences de trempage ont montré que ce site a une forte affinité pour le thiocyanate. La forme mésomérique sélectionnée par ce site est un ion S=C=N<sup>-</sup>. Il est intéressant de noter que lorsque le thiocyanate est fixé dans le site S3, la couronne de résidus hydrophobes se referme et sépare la partie supérieure de la cavité de la partie inférieure [Figure-IB-13].



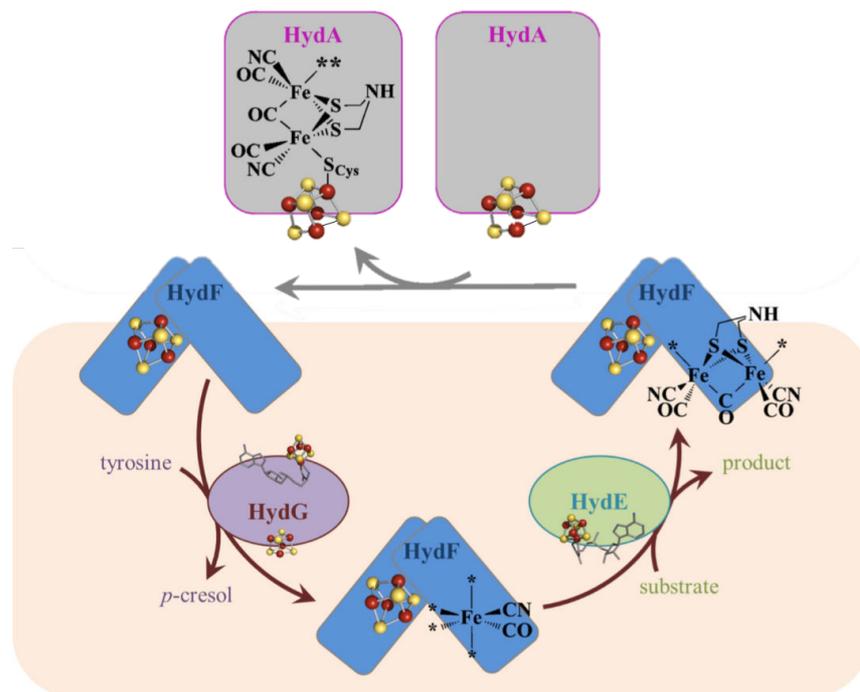
**Figure-IB-13** : L'effet de la fixation d'ion thiocyanate dans le site S3 sur la forme de la cavité dans *TmHydE*

A gauche, la forme que la cavité adopte lorsque les trois sites de fixations d'anions S1, S2 et S3 sont occupés par des ions chlorure. A droite, la fixation d'un ion thiocyanate dans le site S3, se situant au fond de la cavité, entraîne une fermeture de la couronne de résidus hydrophobes et la séparation de la cavité en deux volumes distincts. (Images adaptées de Nicolet *et al.* 2012)

En se basant sur ces observations, les auteurs de la structure ont proposé que le substrat de la maturase HydE pourrait être une molécule de faible poids moléculaire possédant deux groupements carboxylate<sup>101</sup>. Le mécanisme le plus plausible pour le fonctionnement de l'enzyme est que l'initiation de la réaction radicalaire s'effectue, comme pour HydG, dans la partie supérieure de la cavité. Ensuite le produit, ou un de ces précurseurs, franchit la ceinture de résidus hydrophobes, qui se referme et isole la nouvelle molécule dans la partie inférieure de la cavité. La molécule peut y séjourner, à l'abri du solvant, avant d'être expulsée par le canal E<sub>2</sub> hors de la protéine<sup>75,102</sup>.

## IB-4 Mécanisme de maturation

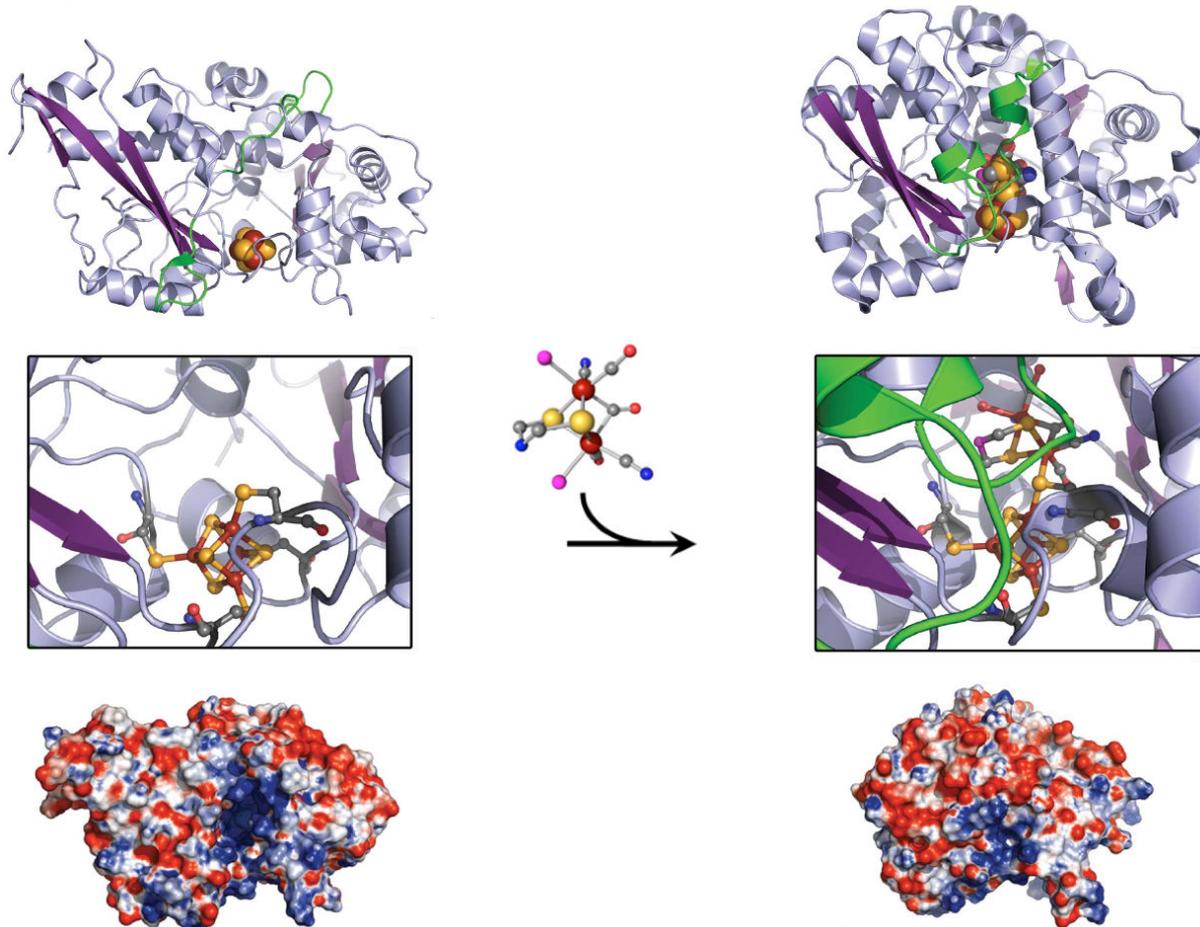
Après l'expression de l'hydrogénase à [FeFe], la maturation de l'agrégat H commence par l'assemblage du centre [Fe<sub>4</sub>S<sub>4</sub>] par les machineries d'assemblage des centres [Fe-S]<sup>46-48,53,103</sup>. L'assemblage du sous-agrégat à deux Fe se fait sur la maturase HydF<sup>45,48</sup>, qui elle aussi contient un centre [Fe<sub>4</sub>S<sub>4</sub>] préassemblé<sup>104</sup>. La maturase HydG produit des synthons [Fe(CO)<sub>2</sub>(CN<sup>-</sup>)] à partir du cinquième atome de Fe de son agrégat auxiliaire et à partir de la L-tyrosine<sup>97,99,100</sup>. Les interactions entre HydF et les deux autres maturases HydE et HydG ont été quantifiées par les expériences de résonance plasmonique de surface<sup>105</sup>. Il a été montré que HydE se lie à HydF avec plus d'affinité que ne le fait HydG. Cependant l'interaction entre HydF et HydE ne déplace pas HydG. L'activité GTP-ase de HydF, qui augmente en présence des deux autres maturases, déplace l'équilibre vers la séparation des complexes HydF-HydE/HydG. La maturase HydE est sans doute impliquée dans la production du ligand dithiométhylamine<sup>102</sup>. Dans certains organismes, comme par exemple chez *Chlamydomonas reinhardtii*, les protéines HydE et HydF sont fusionnées en un seul gène *hydEF*<sup>10</sup>. L'affinité relativement élevée de HydE pour HydF et la fusion des deux gènes dans certains organismes témoignent d'une collaboration étroite entre les deux maturase lors de l'étape de l'assemblage du sous-agrégat à deux Fe [Figure-IB-14].



**Figure-IB-14** : Schéma du mécanisme proposé pour l'assemblage et l'insertion du sous-agrégat à deux atomes de Fe dans le site actif de l'hydrogénase à [FeFe]

La maturase HydG produit des synthons [Fe(CO)<sub>x</sub>(CN)<sub>-y</sub>] à partir du cinquième atome de Fe de son agrégat auxiliaire et à partir des produits de dégradation de la L-tyrosine. Les synthons sont ensuite transférés dans HydF, qui contient déjà un centre [Fe<sub>4</sub>S<sub>4</sub>]. HydE produit le ligand dithiométhylamine ou son précurseur. Le sous-agrégat à deux atomes de Fe est ensuite transféré dans HydA. Les astérisques indiquent des ligands de nature inconnue. (Image adaptée de Broderick *et al.* 2014)

Le sous-agrégat à deux atomes de Fe assemblé sur la maturase HydF<sup>48,53-55</sup> est ensuite inséré dans HydA<sup>31,46</sup>. La présence de petites molécules exogènes, comme par exemple le GTP et SAM, n'est pas nécessaire pour le transfert<sup>48</sup>. Il a été démontré que l'insertion des molécules biomimétiques du sous-agrégat à deux Fe peut avoir lieu directement dans la solution en absence de la maturase HydF<sup>31,106</sup>. L'insertion du sous-agrégat à deux Fe se fait via un canal chargé positivement. Après le transfert, l'effondrement de ce canal va refermer le site actif de l'hydrogénase à [FeFe]<sup>46</sup> [Figure-IB-15]. Il a été proposé que ce même mode d'insertion est utilisé pour le transfert du cofacteur FeMo-co dans le site actif de la nitrogénase<sup>46</sup>.



**Figure-IB-15** : Changement de la conformation de HydA après l'insertion du sous-agrégat à deux Fe

A gauche de l'image: La structure tridimensionnelle de *Cr*HydA<sup>ΔEFG</sup> exprimée en absence de ses maturases montre que l'hydrogénase contient la partie [Fe<sub>4</sub>S<sub>4</sub>] de l'agrégat H (code PDB: 3LX4). Cette structure révèle également la présence d'un large canal électropositif par lequel le sous-agrégat à deux atomes de Fe peut être inséré après son assemblage sur HydF. A droite de l'image: La structure de *Cpl* HydA<sup>EFG</sup> exprimée en présence des trois maturases (code PDB: 3C8Y) montre qu'après l'assemblage de l'agrégat H complet, le canal s'effondre et isole ainsi le site actif du solvant. L'effondrement du canal est causé par le réarrangement de deux boucles strictement conservées, ici colorées en vert. (Images adaptées de Mulder *et al.* 2010)

Contrairement à l'hydrogénase à [NiFe], la maturation de l'hydrogénase à [FeFe] est un processus très conservé<sup>58</sup>. Les maturases bactériennes peuvent insérer le sous-agrégat à deux Fe dans l'hydrogénase à [FeFe] issue d'un organisme eucaryote. Une promiscuité entre les maturases issus de différents organismes existe aussi<sup>44,75</sup>. Ceci n'est pas le cas pour la maturation des hydrogénases à [NiFe] qui présentent une très forte spécificité. En effet, dans

ce cas, la machinerie issue d'un organisme n'est capable de maturer que l'hydrogénase issue du même organisme.

En 2012 une controverse sur le processus de maturation du site actif de l'hydrogénase à [FeFe] a été lancée par James R. Swartz et son équipe. Suite à l'activation partielle de l'enzyme en absence de HydE ces chercheurs ont émis des doutes sur l'importance de cette protéine dans l'assemblage du sous-agrégat à deux Fe<sup>58</sup>. Ce point sera d'avantage discuté dans le **Chapitre II** [voir paragraphe II-3].

L'étape qui reste la moins comprise est la fabrication de la molécule dithiolate pontant les deux atomes de Fe – réaction à priori catalysée par HydE. Ce ligand contient deux atomes de soufre, il est donc naturel de considérer des réactions d'insertion de soufre dans les molécules biologiques.

### **IB-5 Protéines à radical SAM - Insertion du soufre par la voie radicalaire**

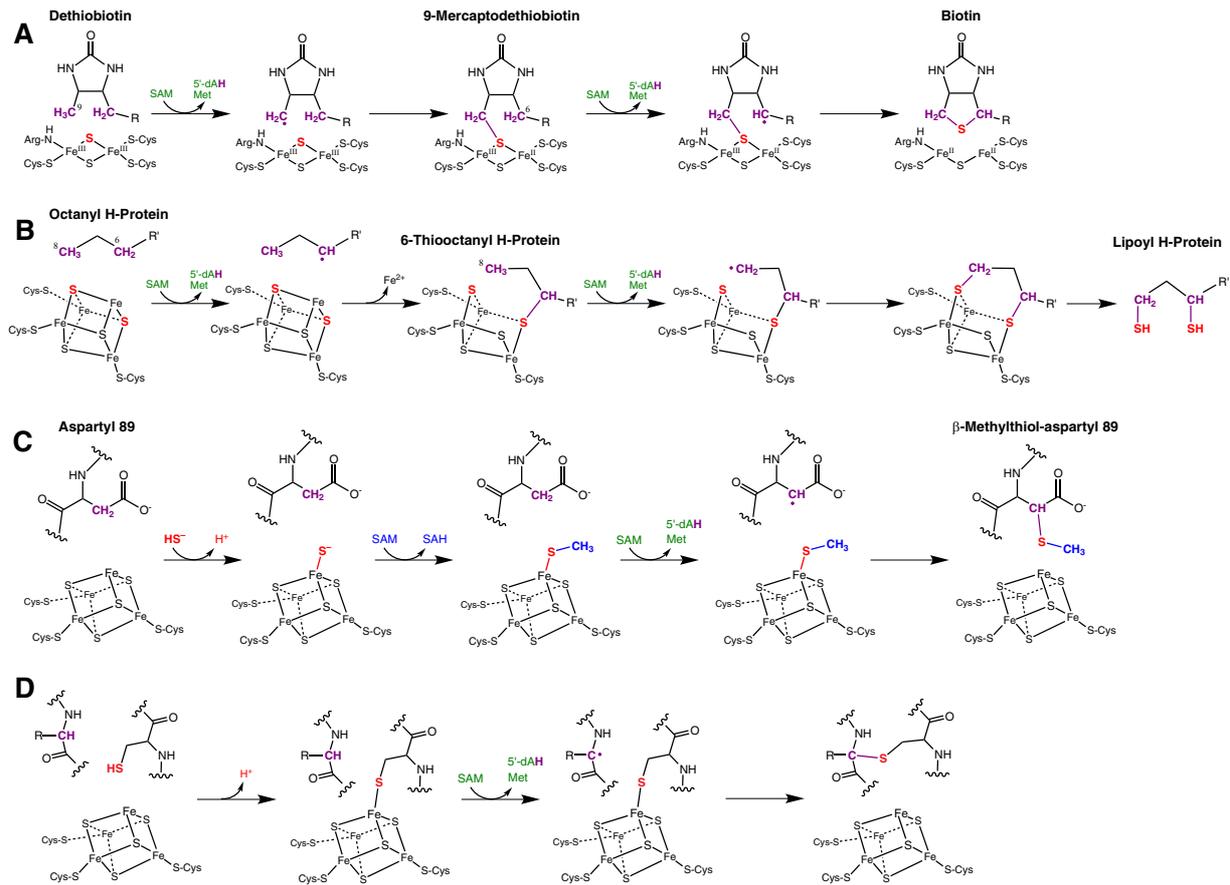
Les organothiolates anioniques sont de bons groupes partants et de bons nucléophiles. Les réactions d'insertion de soufre sur les carbones électrophiliques ou qui peuvent le devenir au cours de la réaction enzymatique sont nombreuses en biologie et passent essentiellement par substitution nucléophile ou par addition de Michael. Cependant l'insertion de soufre sur des carbones aliphatiques ou aromatiques riches en électrons n'est pas possible avec ces réactions. Il existe un groupe de protéines à radical SAM qui catalysent des réactions d'insertion de soufre par voie radicalaire. Ces enzymes ont toutes la particularité de posséder un second agrégat [Fe-S], qui sert soit pour activer l'atome de soufre d'un des substrats soit directement comme source d'atomes de soufre à insérer<sup>107,108</sup> [Figure-IB-16].

Le groupe d'enzymes qui utilisent un centre [Fe<sub>4</sub>S<sub>4</sub>]/ [Fe<sub>2</sub>S<sub>2</sub>] auxiliaire pour activer un thiol du substrat est représenté par les protéines à radical SAM MiaB, MtaB, RimO, AlbA, SkfB, QhpD et le duo d'enzymes TrnC et TrnD<sup>107</sup>. Les protéines MiaB, MtaB et RimO forme un petit sous-groupe de méthylthiotransférases. MiaB et MtaB sont impliquées dans les méthylthioliations post-transcriptionnelles des ARN de transfert<sup>109</sup>. MiaB catalyse l'insertion d'un groupement méthylthiol sur le C2 de la 6-isopentenyladenosine, tandis que son homologue MtaB catalyse la même insertion sur le C2 de la 6-threonyladenosine<sup>109</sup>. Quant à RimO, cette protéine catalyse l'insertion du groupement méthylthiol sur la position C $\gamma$  du résidu Asp89 de la protéine ribosomale S12<sup>110</sup> [Figure-IB-16-C].

Les protéines AlbA, SkfB, TrnC et TrnD sont impliquées dans la formation d'une liaison thioéther à l'intérieur des peptides cycliques antibiotiques - les sactibiotiques<sup>111-113</sup>. Puisque ces peptides sont trop petit pour posséder une structure tridimensionnelle stable, une ou plusieurs liaisons thioéther entre le soufre d'une cystéine et l'atome C $\alpha$  d'un autre résidu sont introduites pour restreindre l'espace conformationnel du peptide [Figure-IB-16-D]. Par exemple, la subtilosine A est générée à partir du précurseur SboA par l'enzyme AlbA<sup>112</sup>.

La protéine à radical SAM QhpD est essentielle pour la formation de liaisons thioéther intramoléculaires dans la petite sous-unité de la quinohémostéine amine déshydrogénase QhpC<sup>114</sup>.

Le second groupe, qui utilise un centre [Fe-S] auxiliaire comme source d'atomes de soufre, contient les protéines BioB et LipA. L'enzyme BioB catalyse une réaction d'insertion d'un atome de soufre entre les carbones aliphatiques C9 et C6 de la déthiobiotine lors de la dernière étape de la synthèse de la vitamine B<sub>8</sub> - la biotine<sup>77,115-117</sup> [Figure-IB-16-A]. LipA catalyse une double insertion d'atomes de soufre sur deux carbones aliphatiques primaires et secondaires (C6 et C8) pour transformer l'octanoate en cofacteur lipote<sup>118</sup> [Figure-IB-16-B]. Ces enzymes consomment deux molécules de SAM par produit formé. Dans BioB le donneur sacrificiel de soufre est un centre [Fe<sub>2</sub>S<sub>2</sub>], tandis que LipA utilise un agrégat [Fe<sub>4</sub>S<sub>4</sub>].



**Figure-IB-16 : Mécanismes proposés pour la formation de liaison C-S par processus radicalaires catalysée par certaines enzymes à radical SAM**

Pour tous ces enzymes, le processus de la formation d'une liaison entre l'atome de soufre (en rouge) et l'atome de carbone (en violet) commence par le clivage réductif homolytique de la SAM (en vert). L'activation de cet atome de carbone aliphatique se fait par l'arrachage d'un atome d'hydrogène effectué par le radical 5'-dA•. L'atome de soufre à insérer est coloré en rouge.

**A** - Synthèse de la biotine par la protéine BioB commence par l'arrachage régiospécifique d'un atome d'hydrogène sur l'atome C9 de la déthiobiotine, suivi de la formation d'une liaison C-S entre cet atome de carbone et un atome de soufre de l'agrégat auxiliaire. Un second clivage de la SAM et l'arrachage d'un atome d'hydrogène de façon stéréospécifique sur le C6 précèdent la formation de la liaison thioéther C9-S-C6.

**B** - Insertions de deux groupements thiols sur l'octanoate attaché à une protéine porteuse, catalysée par LipA. Dans A et B deux molécules de SAM sont consommées par produit libéré et le centre [Fe-S] secondaire ayant servi d'accepteur d'électrons et de source sacrificielle d'atomes de soufre est dégradé à la fin de la réaction.

**C** - Le mécanisme proposé pour l'insertion d'un groupement méthylthiol sur l'atome Cγ du résidu Asp89 de la protéine ribosomale S12 catalysée par l'enzyme RimO. Une première molécule de la SAM est consommée dans le processus de méthylation (en bleu) de l'atome de soufre à insérer.

**D** - Formation d'une liaison thioéther intramoléculaire entre un atome de soufre d'une cystéine et un atome de Cα d'un autre résidu du substrat catalysée par les enzymes AlbA, SkfB, QhpD et TrnC & D. Dans C et D l'agrégat auxiliaire active le thiolate du substrat et sert de piège à électrons. Il peut fonctionner de façon catalytique. (Figure adaptée de Jarrett 2015)

Tous ces enzymes utilisent le clivage réductif homolytique de la SAM pour générer le radical 5'-dA qui va ensuite arracher un atome d'hydrogène sur le substrat et former ainsi un radical alkyl sur ce dernier<sup>107</sup>. Il a été proposé que ce radical va par la suite attaquer l'atome de soufre qui fait partie du centre [Fe-S] ou qui est activé par l'interaction avec le Fe<sub>u</sub> d'un agrégat [Fe-S] différencié. Dans les deux cas, le centre [Fe-S] sert d'accepteur d'électron après la formation de la nouvelle liaison C-S<sup>107</sup>. Son second rôle serait le maintien de l'atome de soufre à insérer sous une forme où il peut être attaqué par le radical alkyl. A pH physiologique la plupart de thiols se trouveraient sous forme protonée. Le radical alkyl du substrat arracherait alors l'atome d'hydrogène du thiol, ce qui conduirait à la formation d'un radical thiyl stable et non pas à la création d'une nouvelle liaison C-S. Par contre, en liant le sulfure ou le thiolate au Fe<sub>u</sub>, le soufre ne peut plus être protoné et le centre [Fe-S] remplit en plus le rôle de l'accepteur d'électron.

Dans le cas des protéines à radical SAM qui utilisent le centre [Fe-S] auxiliaire comme la source d'atome de soufre (BioB et LipA), le mécanisme d'insertion et la formation d'espèces intermédiaires commencent à être bien compris. En revanche, dans le cas du second groupe, l'interaction entre le thiolate du substrat et le Fe<sub>u</sub> de l'agrégat auxiliaire a pu être confirmée expérimentalement uniquement pour les méthylthiotransférases RimO et MiaB<sup>110</sup>. La résolution de structure tridimensionnelle d'une des enzymes qui produisent des sactibiotiques sera sans doute nécessaire pour confirmer l'activation directe du thiolate de la cystéine du peptide substrat par sa fixation sur le Fe<sub>u</sub> du second centre [Fe-S]<sup>82</sup>.

Certaines de ces enzymes à radical SAM sont très proches de la maturase HydE. Les similarités et les différences entre HydE et la protéine BioB seront commentées au début du *Chapitre I* de la partie Résultats.

## Objectifs de la thèse

Les objectifs de ma thèse étaient d'identifier, dans un premier temps, le substrat de la protéine HydE - une protéine à radical SAM impliquée dans l'assemblage du site actif organométallique des hydrogénases à [FeFe], puis de mener une étude structurale par cristallographie aux rayons X du mécanisme réactionnel de la maturase HydE. Un des buts de ces travaux était d'apporter des informations supplémentaires sur le processus d'assemblage du site actif des hydrogénases à [FeFe]. Puisque au moment quand j'ai commencé ma thèse la composition exacte de ce site n'était pas connue, nous espérions que l'identification du substrat de HydE allait nous permettre de déterminer la nature de l'atome centrale de la molécule pontant les deux atomes de Fe dans le site actif des hydrogénases à [FeFe]. La bonne connaissance du site actif et du mécanisme catalytique sont indispensables pour produire des analogues synthétiques biomimétiques ayant une activité spécifique comparable à celle de l'enzyme. L'identification du substrat de HydE nous avancerait également vers la mise en place d'un système de maturation de l'hydrogénase *in vitro* complet. Un tel système serait particulièrement utile pour étudier la séquence de biosynthèse du sous-agrégat à deux Fe dans HydF et de son insertion dans le site actif de l'hydrogénase.

La compréhension du processus de maturation mènerait à une meilleure reconstitution de l'hydrogénase recombinante ce qui rendrait son étude plus facile, notamment en ce qui est de l'identification de mutants tolérants à l'oxygène - ceci dans le but de produire des systèmes de biogénération de dihydrogène à grande échelle.

Outre l'apport aux connaissances sur l'hydrogénase, l'étude de son système de maturation a aussi un intérêt pour la meilleure compréhension du mécanisme des protéines à radical SAM. En effet, dans notre laboratoire nous avons réussi à produire des cristaux de la protéine HydE de *Thermotoga maritima*, qui diffractent proche de la résolution atomique. Une fois le substrat de l'enzyme identifié, il devrait être possible de disséquer la réaction radicalaire en piégeant certains des états intermédiaires dans le cristal. Ceci nous renseignerait sur la chimie radicalaire d'une de nombreuses réactions catalysées par cette classe d'enzymes.

## IB-1 Références de l'Introduction Bibliographique

- 1 Artero, V. *et al.* From Enzyme Maturation to Synthetic Chemistry: The Case of Hydrogenases. *Accounts of Chemical Research* **48**, doi:10.1021/acs.accounts.5b00157 (2015).
- 2 Turner, J. A. Sustainable hydrogen production. *Science (New York, N.Y.)* **305**, 972-974, doi:10.1126/science.1103197 (2004).
- 3 Ni, M., Leung, M. & Leung, D. Energy and exergy analysis of hydrogen production by a proton exchange membrane (PEM) electrolyzer plant. *Energy Conversion and Management* **49**, 2748-2756, doi:10.1016/j.enconman.2008.03.018 (2008).
- 4 Gordon, R. B., Bertram, M. & Graedel, T. E. Metal stocks and sustainability. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 1209-1214, doi:10.1073/pnas.0509498103 (2006).
- 5 Fontecilla-Camps\*, J. C., Volbeda, A., Cavazza, C. & Nicolet, Y. Structure/Function Relationships of [NiFe]- and [FeFe]-Hydrogenases. *Chemical Reviews* **107**, 4273-4303, doi:10.1021/cr050195z (2007).
- 6 Lubitz, W., Ogata, H., Rüdiger, O. & Reijerse, E. Hydrogenases. *Chemical reviews* **114**, 4081-4148, doi:10.1021/cr4005814 (2014).
- 7 Peters, J. W. *et al.* [FeFe]- and [NiFe]-hydrogenase diversity, mechanism, and maturation. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, doi:10.1016/j.bbamcr.2014.11.021 (2014).
- 8 Vignais, P. & Billoud, B. Occurrence, Classification, and Biological Function of Hydrogenases: An Overview. *Occurrence, Classification, and Biological Function of Hydrogenases: An Overview* **107**, doi:10.1021/cr050196r (2007).
- 9 Vignais, P. M., Billoud, B. & Meyer, J. Classification and phylogeny of hydrogenases. *Classification and phylogeny of hydrogenases*, doi:10.1111/j.1574-6976.2001.tb00587.x (2001).
- 10 Posewitz, M. C. *et al.* Discovery of Two Novel Radical S-Adenosylmethionine Proteins Required for the Assembly of an Active [Fe] Hydrogenase. *Journal of Biological Chemistry* **279**, 25711-25720, doi:10.1074/jbc.M403206200 (2004).
- 11 Peters, J. W. Structure and mechanism of iron-only hydrogenases. *Current Opinion in Structural Biology* **9**, 670-676, doi:10.1016/s0959-440x(99)00028-7 (1999).
- 12 Shima, S. & Thauer, R. K. A third type of hydrogenase catalyzing H<sub>2</sub> activation. *Chemical record (New York, N.Y.)* **7**, 37-46, doi:10.1002/tcr.20111 (2007).
- 13 Shima, S. *et al.* Reconstitution of [Fe]-hydrogenase using model complexes. *Nature Chemistry* **7**, 995-1002, doi:10.1038/nchem.2382 (2015).
- 14 Nicolet, Y., Lemon, B. J., Fontecilla-Camps, J. C. & Peters, J. W. A novel FeS cluster in Fe-only hydrogenases. *Trends in biochemical sciences* **25**, 138-143, doi:10.1016/s0968-0004(99)01536-4 (2000).
- 15 Darensbourg, M. Y., Lyon, E. J., Zhao, X. & Georgakaki, I. P. The organometallic active site of [Fe] hydrogenase: models and entatic states. *Proceedings of the National Academy of Sciences* **100**, 3683-3688, doi:10.1073/pnas.0536955100 (2003).
- 16 DuBois, D. L. Development of molecular electrocatalysts for energy storage. *Inorganic Chemistry*, doi:10.1021/ic4026969 (2014).
- 17 Fontecilla-Camps, J. C., Amara, P., Cavazza, C., Nicolet, Y. & Volbeda, A. Structure-function relationships of anaerobic gas-processing metalloenzymes. *Nature* **460**, 814-822, doi:10.1038/nature08299 (2009).
- 18 Nicolet, Y., Piras, C., Legrand, P., Hatchikian, C. E. & Fontecilla-Camps, J. C. Desulfovibrio desulfuricans iron hydrogenase: the structure shows unusual coordination to an active site Fe binuclear center. *Structure* **7**, doi:10.1016/s0969-2126(99)80005-7 (1998).

- 19 Montet, Y. *et al.* Gas access to the active site of Ni-Fe hydrogenases probed by X-ray crystallography and molecular dynamics. *Nature structural biology* **4**, 523-526 (1997).
- 20 Kalms, J. *et al.* Krypton Derivatization of an O<sub>2</sub> - Tolerant Membrane - Bound [NiFe] Hydrogenase Reveals a Hydrophobic Tunnel Network for Gas Transport. *Angewandte Chemie International Edition*, doi:10.1002/anie.201508976 (2016).
- 21 Cornish, A. J., Gärtner, K., Yang, H., Peters, J. W. & Hegg, E. L. Mechanism of proton transfer in [FeFe]-hydrogenase from *Clostridium pasteurianum*. *The Journal of biological chemistry* **286**, 38341-38347, doi:10.1074/jbc.M111.254664 (2011).
- 22 Ogata, H., Nishikawa, K. & Lubitz, W. Hydrogens detected by subatomic resolution protein crystallography in a [NiFe] hydrogenase. *Nature* **520**, 571-574, doi:10.1038/nature14110 (2015).
- 23 Volbeda, A. *et al.* X-ray crystallographic and computational studies of the O<sub>2</sub>-tolerant [NiFe]-hydrogenase 1 from *Escherichia coli*. *Proceedings of the National Academy of Sciences* **109**, 5305-5310, doi:10.1073/pnas.1119806109 (2012).
- 24 Volbeda, A. *et al.* Structure of the [NiFe] Hydrogenase Active Site: Evidence for Biologically Uncommon Fe Ligands  $\perp$ . *Journal of the American Chemical Society* **118**, 12989-12996, doi:10.1021/ja962270g (1996).
- 25 Böck, A., King, P. W., Blokesch, M. & Posewitz, M. C. Maturation of Hydrogenases. *Advances in microbial physiology* **51**, 1-225, doi:10.1016/s0065-2911(06)51001-x (2006).
- 26 Hatchikian, E. C., Forget, N., Fernandez, V. M., Williams, R. & Cammack, R. Further characterization of the [Fe]-hydrogenase from *Desulfovibrio desulfuricans* ATCC 7757. *European journal of biochemistry / FEBS* **209**, 357-365, doi:10.1111/j.1432-1033.1992.tb17297.x (1992).
- 27 Nicolet, Y. *et al.* Crystallographic and FTIR Spectroscopic Evidence of Changes in Fe Coordination Upon Reduction of the Active Site of the Fe-Only Hydrogenase from *Desulfovibriodesulfuricans*. *Journal of the American Chemical Society* **123**, 1596-1601, doi:10.1021/ja0020963 (2001).
- 28 Pandey, A. S., Harris, T. V., Giles, L. J., Peters, J. W. & Szilagyi, R. K. Dithiomethylether as a ligand in the hydrogenase h-cluster. *Journal of the American Chemical Society* **130**, 4533-4540, doi:10.1021/ja711187e (2008).
- 29 Grigoropoulos, A. & Szilagyi, R. K. Evaluation of biosynthetic pathways for the unique dithiolate ligand of the FeFe hydrogenase H-cluster. *JBIC Journal of Biological Inorganic Chemistry* **15**, 1177-1182, doi:10.1007/s00775-010-0698-y (2010).
- 30 Grigoropoulos, A. & Szilagyi, R. K. In silico evaluation of proposed biosynthetic pathways for the unique dithiolate ligand of the H-cluster of [FeFe]-hydrogenase. *Journal of computational chemistry* **32**, 3194-3206, doi:10.1002/jcc.21901 (2011).
- 31 Berggren, G. *et al.* Biomimetic assembly and activation of [FeFe]-hydrogenases. *Nature* **499**, 66-69, doi:10.1038/nature12239 (2013).
- 32 Johnson, D. C., Dean, D. R., Smith, A. D. & Johnson, M. K. STRUCTURE, FUNCTION, AND FORMATION OF BIOLOGICAL IRON-SULFUR CLUSTERS. *Biochemistry* **74**, 247-281, doi:10.1146/annurev.biochem.74.082803.133518 (2005).
- 33 Lill, R. Function and biogenesis of iron-sulphur proteins. *Nature* **460**, 831-838, doi:10.1038/nature08301 (2009).
- 34 Paul, V. & Lill, R. Biogenesis of cytosolic and nuclear iron-sulfur proteins and their role in genome stability. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1853**, 1528-1539, doi:10.1016/j.bbamcr.2014.12.018 (2015).
- 35 Fontecave, M. Iron-sulfur clusters: ever-expanding roles. *Nature Chemical Biology* **2**, 171-174, doi:10.1038/nchembio0406-171 (2006).
- 36 Py, B. & Barras, F. Genetic approaches of the Fe-S cluster biogenesis process in bacteria: Historical account, methodological aspects and future challenges. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1853**, 1429-1435, doi:10.1016/j.bbamcr.2014.12.024 (2015).
- 37 Drennan, C. L. & Peters, J. W. Surprising cofactors in metalloenzymes. *Current opinion in structural biology* **13**, 220-226 (2003).

- 38 Spatzal, T. *et al.* Evidence for Interstitial Carbon in Nitrogenase FeMo Cofactor. *Science* **334**, 940940, doi:10.1126/science.1214025 (2011).
- 39 Lancaster, K. M. *et al.* X-ray emission spectroscopy evidences a central carbon in the nitrogenase iron-molybdenum cofactor. *Science (New York, N.Y.)* **334**, 974-977, doi:10.1126/science.1206445 (2011).
- 40 Johnson, D. C., Dean, D. R., Smith, A. D. & Johnson, M. K. Structure, function, and formation of biological iron-sulfur clusters. *Annual review of biochemistry* **74**, 247-281, doi:10.1146/annurev.biochem.74.082803.133518 (2005).
- 41 Rubio, L. M. & Ludden, P. W. Maturation of nitrogenase: a biochemical puzzle. *Journal of bacteriology* **187**, 405-414, doi:10.1128/jb.187.2.405-414.2005 (2005).
- 42 Voordouw, G., Hagen, W. R., Krüse-Wolters, K. M., van Berkel-Arts, A. & Veeger, C. Purification and characterization of *Desulfovibrio vulgaris* (Hildenborough) hydrogenase expressed in *Escherichia coli*. *European journal of biochemistry / FEBS* **162**, 31-36 (1987).
- 43 Atta, M. & Meyer, J. Characterization of the gene encoding the [Fe]-hydrogenase from *Megasphaera elsdenii*. *Biochimica et biophysica acta* **1476**, 368-371 (2000).
- 44 King, P. W., Posewitz, M. C., Ghirardi, M. L. & Seibert, M. Functional Studies of [FeFe] Hydrogenase Maturation in an *Escherichia coli* Biosynthetic System. *Journal of Bacteriology* **188**, 2163-2172, doi:10.1128/jb.188.6.2163-2172.2006 (2006).
- 45 McGlynn, S. E. *et al.* In vitro activation of [FeFe] hydrogenase: new insights into hydrogenase maturation. *JBIC Journal of Biological Inorganic Chemistry* **12**, 443-447, doi:10.1007/s00775-007-0224-z (2007).
- 46 Mulder, D. W. *et al.* Stepwise [FeFe]-hydrogenase H-cluster assembly revealed in the structure of HydA  $\Delta$  EFG. *Nature* **465**, 248-251, doi:10.1038/nature08993 (2010).
- 47 Mulder, D. W. *et al.* Activation of HydA  $\Delta$  EFG Requires a Preformed [4Fe-4S] Cluster. *Biochemistry* **48**, 6240-6248, doi:10.1021/bi9000563 (2009).
- 48 McGlynn, S. E. *et al.* HydF as a scaffold protein in [FeFe] hydrogenase H-cluster biosynthesis. *FEBS Letters* **582**, 2183-2187, doi:10.1016/j.febslet.2008.04.063 (2008).
- 49 Frazzon, J. & Dean, D. R. Formation of iron-sulfur clusters in bacteria: an emerging field in bioinorganic chemistry. *Current Opinion in Chemical Biology* **7**, 166-173, doi:10.1016/s1367-5931(03)00021-8 (2003).
- 50 Betz, J. N. *et al.* [FeFe]-Hydrogenase Maturation: Insights into the Role HydE Plays in Dithiomethylamine Biosynthesis. *Biochemistry*, 1574462297, doi:10.1021/bi501205e (2015).
- 51 Brazzolotto, X. *et al.* The [Fe-Fe]-Hydrogenase Maturation Protein HydF from *Thermotoga maritima* Is a GTPase with an Iron-Sulfur Cluster. *Journal of Biological Chemistry* **281**, 769-774, doi:10.1074/jbc.M510310200 (2006).
- 52 Rubach, J. K., Brazzolotto, X., Gaillard, J. & Fontecave, M. Biochemical characterization of the HydE and HydG iron-only hydrogenase maturation enzymes from *Thermotoga maritima*. *FEBS Letters* **579**, 5055-5060, doi:10.1016/j.febslet.2005.07.092 (2005).
- 53 Shepard, E. M. *et al.* Synthesis of the 2Fe subcluster of the [FeFe]-hydrogenase H cluster on the HydF scaffold. *Proceedings of the National Academy of Sciences*, doi:10.1073/pnas.1001937107 (2010).
- 54 Czech, I. *et al.* The [FeFe]-hydrogenase maturation protein HydF contains a H-cluster like [4Fe4S]-2Fe site. *FEBS Letters* **585**, 225-230, doi:10.1016/j.febslet.2010.11.052 (2010).
- 55 Czech, I., Silakov, A., Lubitz, W. & Happe, T. The [FeFe]-hydrogenase maturase HydF from *Clostridium acetobutylicum* contains a CO and CN- ligated iron cofactor. *FEBS Letters* **584**, 638-642, doi:10.1016/j.febslet.2009.12.016 (2010).
- 56 Berto, P. *et al.* The [4Fe-4S]-cluster coordination of [FeFe]-hydrogenase maturation protein HydF as revealed by EPR and HYSCORE spectroscopies. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1817**, 2149-2157, doi:10.1016/j.bbabi.2012.09.004 (2012).

- 57 Albertini, M. *et al.* Probing the Solvent Accessibility of the [4Fe-4S] Cluster of the Hydrogenase Maturation Protein HydF from *Thermotoga neapolitana* by HYSCORE and 3p-ESEEM. *The Journal of Physical Chemistry B* **119**, 13680-13689, doi:10.1021/acs.jpcc.5b03110 (2015).
- 58 Kuchenreuther, J. M., Britt, D. R. & Swartz, J. R. New Insights into [FeFe] Hydrogenase Activation and Maturation Function. *PLoS ONE* **7**, doi:10.1371/journal.pone.0045850 (2012).
- 59 Cendron, L. *et al.* Crystal structure of HydF scaffold protein provides insights into [FeFe]-hydrogenase maturation. *The Journal of biological chemistry* **286**, 43944-43950, doi:10.1074/jbc.M111.281956 (2011).
- 60 Berggren, G. *et al.* An EPR/HYSCORE, Mössbauer, and resonance Raman study of the hydrogenase maturation enzyme HydF: a model for N-coordination to [4Fe-4S] clusters. *JBIC Journal of Biological Inorganic Chemistry*, doi:10.1007/s00775-013-1062-9 (2013).
- 61 Sofia, H. J., Chen, G., Hetzler, B. G., Reyes-Spindola, J. F. & Miller, N. E. Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic Acids Research*, doi:10.1093/nar/29.5.1097 (2001).
- 62 Nicolet, Y., Amara, P., Mouesca, J.-M. M. & Fontecilla-Camps, J. C. Unexpected electron transfer mechanism upon AdoMet cleavage in radical SAM proteins. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 14867-14871, doi:10.1073/pnas.0904385106 (2009).
- 63 Wang, S. C. & Frey, P. A. S-adenosylmethionine as an oxidant: the radical SAM superfamily. *Trends in Biochemical Sciences* **32**, 101-110, doi:10.1016/j.tibs.2007.01.002 (2007).
- 64 Wang, S. C. & Frey, P. A. Binding energy in the one-electron reductive cleavage of S-adenosylmethionine in lysine 2,3-aminomutase, a radical SAM enzyme. *Biochemistry* **46**, 12889-12895, doi:10.1021/bi701745h (2007).
- 65 Broderick, J. B., Duffus, B. R., Duschene, K. S. & Shepard, E. M. Radical S-Adenosylmethionine Enzymes. *Chemical Reviews* **114**, 4229-4317, doi:10.1021/cr4004709 (2014).
- 66 Hu, Y. & Ribbe, M. W. Maturation of nitrogenase cofactor—the role of a class E radical SAM methyltransferase NifB. *Current Opinion in Chemical Biology*, doi:10.1016/j.cbpa.2016.02.016 (2016).
- 67 Mehta, A. P. *et al.* Radical SAM Enzymes in Cofactor Biosynthesis: a treasure trove of complex organic radical rearrangement reactions. *Journal of Biological Chemistry*, doi:10.1074/jbc.R114.623793 (2014).
- 68 Mehta, A. P., Abdelwahed, S. H. & Begley, T. P. Molybdopterin biosynthesis—Mechanistic studies on a novel MoaA catalyzed insertion of a purine carbon into the ribose of GTP. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1854**, 1073-1077, doi:10.1016/j.bbapap.2015.04.008 (2015).
- 69 Nicolet, Y. & Drennan, C. L. AdoMet radical proteins—from structure to evolution—alignment of divergent protein sequences reveals strong secondary structure element conservation. *Nucleic Acids Research* **32**, 4015-4025, doi:10.1093/nar/gkh728 (2004).
- 70 Vey, J. L. & Drennan, C. L. Structural Insights into Radical Generation by the Radical SAM Superfamily. *Chemical Reviews* **111**, 2487-2506, doi:10.1021/cr9002616 (2011).
- 71 Walsby, C. J. *et al.* Electron-nuclear double resonance spectroscopic evidence that S-adenosylmethionine binds in contact with the catalytically active [4Fe-4S](+) cluster of pyruvate formate-lyase activating enzyme. *Journal of the American Chemical Society* **124**, 3143-3151 (2002).
- 72 Walsby, C. J., Ortillo, D., Broderick, W. E., Broderick, J. B. & Hoffman, B. M. An anchoring role for FeS clusters: chelation of the amino acid moiety of S-adenosylmethionine to the unique iron site of the [4Fe-4S] cluster of pyruvate formate-lyase activating enzyme. *Journal of the American Chemical Society* **124**, 11270-11271, doi:10.1021/ja027078v (2002).

- 73 Walsby, C. J. *et al.* Spectroscopic Approaches to Elucidating Novel Iron–Sulfur Chemistry in the “Radical-SAM” Protein Superfamily. *Inorganic Chemistry* **44**, 727-741, doi:10.1021/ic0484811 (2005).
- 74 Nicolet, Y. *et al.* Crystal structure of HydG from Carboxydotherrmus hydrogenoformans: a trifunctional [FeFe]-hydrogenase maturase. *Chembiochem : a European journal of chemical biology* **16**, 397-402, doi:10.1002/cbic.201402661 (2015).
- 75 Nicolet, Y. *et al.* X-ray Structure of the [FeFe]-Hydrogenase Maturase HydE from *Thermotoga maritima*. *Journal of Biological Chemistry* **283**, 18861-18872, doi:10.1074/jbc.M801161200 (2008).
- 76 Nicolet, Y., Zeppieri, L., Amara, P. & Fontecilla - Camps, J. C. Crystal Structure of Tryptophan Lyase (NosL): Evidence for Radical Formation at the Amino Group of Tryptophan. *Angewandte Chemie International Edition* **53**, 11840-11844, doi:10.1002/anie.201407320 (2014).
- 77 Berkovitch, F., Nicolet, Y., Wan, J. T., Jarrett, J. T. & Drennan, C. L. Crystal Structure of Biotin Synthase, an S-Adenosylmethionine-Dependent Radical Enzyme. *Science* **303**, 76-79, doi:10.1126/science.1088493 (2004).
- 78 Vey, J. L., Yang, J., Li, M. & Broderick, W. E. Structural basis for glycy radical formation by pyruvate formate-lyase activating enzyme. *Proceedings of the ...*, doi:10.1073/pnas.0806640105 (2008).
- 79 Goldman, P. J. *et al.* X-ray structure of an AdoMet radical activase reveals an anaerobic solution for formylglycine posttranslational modification. *Proceedings of the National Academy of Sciences* **110**, 8519-8524, doi:10.1073/pnas.1302417110 (2013).
- 80 Goldman, P. J., Grove, T. L., Booker, S. J. & Drennan, C. L. X-ray analysis of butirosin biosynthetic enzyme BtrN redefines structural motifs for AdoMet radical chemistry. *Proceedings of the National Academy of Sciences* **110**, 15949-15954, doi:10.1073/pnas.1312228110 (2013).
- 81 Quitterer, F., List, A., Eisenreich, W., Bacher, A. & Groll, M. Crystal Structure of Methylornithine Synthase (PylB): Insights into the Pyrrolysine Biosynthesis. *Angewandte Chemie International Edition* **51**, 1339-1342, doi:10.1002/anie.201106765 (2012).
- 82 Grell, T. A. J., Goldman, P. J. & Drennan, C. L. SPASM and Twitch Domains in AdoMet Radical Enzyme Structures. *Journal of Biological Chemistry*, doi:10.1074/jbc.R114.581249 (2014).
- 83 Yokoyama, K., Numakura, M., Kudo, F., Ohmori, D. & Eguchi, T. Characterization and mechanistic study of a radical SAM dehydrogenase in the biosynthesis of butirosin. *Journal of the American Chemical Society* **129**, 15147-15155, doi:10.1021/ja072481t (2007).
- 84 Ruszczycky, M. W., Choi, S.-h. H., Mansoorabadi, S. O. & Liu, H.-w. W. Mechanistic studies of the radical S-adenosyl-L-methionine enzyme DesII: EPR characterization of a radical intermediate generated during its catalyzed dehydrogenation of TDP-D-quinovose. *Journal of the American Chemical Society* **133**, 7292-7295, doi:10.1021/ja201212f (2011).
- 85 Ruszczycky, M. W., Ogasawara, Y. & Liu, H.-W. W. Radical SAM enzymes in the biosynthesis of sugar-containing natural products. *Biochimica et biophysica acta* **1824**, 1231-1244, doi:10.1016/j.bbapap.2011.11.006 (2012).
- 86 Lees, N. S. *et al.* How an enzyme tames reactive intermediates: positioning of the active-site components of lysine 2, 3-aminomutase during enzymatic turnover as determined by ENDOR spectroscopy. *Journal of the American Chemical Society* **128**, 10145-10154, doi:10.1021/ja061282r (2006).
- 87 McCarty, R. M., Krebs, C. & Bandarian, V. Spectroscopic, steady-state kinetic, and mechanistic characterization of the radical SAM enzyme QueE, which catalyzes a complex cyclization reaction in the biosynthesis of 7-deazapurines. *Biochemistry* **52**, 188-198, doi:10.1021/bi301156w (2012).
- 88 Yang, L. *et al.* A Radical Transfer Pathway in Spore Photoproduct Lyase. *Biochemistry* **52**, 1155023668, doi:10.1021/bi3016247 (2013).
- 89 Nicolet, Y., Martin, L., Tron, C. & Fontecilla-Camps, J. C. A glycy radical as the precursor in the synthesis of carbon monoxide and cyanide by the [FeFe]-hydrogenase maturase HydG. *FEBS Letters* **584**, 4197-4202, doi:10.1016/j.febslet.2010.09.008 (2010).

- 90 Pilet, E. *et al.* The role of the maturase HydG in [FeFe]-hydrogenase active site synthesis and assembly. *FEBS Letters* **583**, 506-511, doi:10.1016/j.febslet.2009.01.004 (2008).
- 91 Kriek, M., Martins, F., Challand, M. R., Croft, A. & Roach, P. L. Thiamine Biosynthesis in *Escherichia coli*: Identification of the Intermediate and By - Product Derived from Tyrosine. *Angewandte Chemie International Edition* **46**, 9223-9226, doi:10.1002/anie.200702554 (2007).
- 92 Driesener, R. C. *et al.* Biochemical and Kinetic Characterization of Radical S-Adenosyl-l-methionine Enzyme HydG. *Biochemistry* **52**, 8696-8707, doi:10.1021/bi401143s (2013).
- 93 Shepard, E. M. *et al.* [FeFe]-hydrogenase maturation: HydG-catalyzed synthesis of carbon monoxide. *Journal of the American Chemical Society* **132**, 9247-9249, doi:10.1021/ja1012273 (2010).
- 94 Driesener, R. C. *et al.* [FeFe] - Hydrogenase Cyanide Ligands Derived From S - Adenosylmethionine - Dependent Cleavage of Tyrosine. *Angewandte Chemie International Edition* **49**, 1687-1690, doi:10.1002/anie.200907047 (2010).
- 95 Kuchenreuther, J. M., Stapleton, J. A. & Swartz, J. R. Tyrosine, Cysteine, and S-Adenosyl Methionine Stimulate In Vitro [FeFe] Hydrogenase Activation. *PLOS ONE* **4**, doi:10.1371/journal.pone.0007565 (2009).
- 96 Kuchenreuther, J. M., George, S. J., Grady-Smith, C. S., Cramer, S. P. & Swartz, J. R. Cell-free H-cluster Synthesis and [FeFe] Hydrogenase Activation: All Five CO and CN- Ligands Derive from Tyrosine. *PLOS ONE* **6**, doi:10.1371/journal.pone.0020346 (2011).
- 97 Kuchenreuther, J. M. *et al.* The HydG enzyme generates an Fe(CO)<sub>2</sub>(CN) synthon in assembly of the FeFe hydrogenase H-cluster. *Science (New York, N.Y.)* **343**, 424-427, doi:10.1126/science.1246572 (2014).
- 98 Tron, C. *et al.* Further Characterization of the [FeFe]-Hydrogenase Maturase HydG. *European Journal of Inorganic Chemistry* **2011**, doi:10.1002/ejic.201001101 (2011).
- 99 Dinis, P. *et al.* X-ray crystallographic and EPR spectroscopic analysis of HydG, a maturase in [FeFe]-hydrogenase H-cluster assembly. *Proceedings of the National Academy of Sciences* **112**, 1362-1367, doi:10.1073/pnas.1417252112 (2015).
- 100 Pagnier, A., Martin, L., Zeppieri, L., Nicolet, Y. & Fontecilla-Camps, J. C. CO and CN- syntheses by [FeFe]-hydrogenase maturase HydG are catalytically differentiated events. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 104-109, doi:10.1073/pnas.1515842113 (2015).
- 101 Nicolet, Y., Fontecilla-Camps, J. C. & Fontecave, M. Maturation of [FeFe]-hydrogenases: Structures and mechanisms. *International Journal of Hydrogen Energy* **35**, 1075010760, doi:10.1016/j.ijhydene.2010.02.056 (2010).
- 102 Nicolet, Y. & Fontecilla-Camps, J. C. Structure-Function Relationships in [FeFe]-Hydrogenase Active Site Maturation. *Journal of Biological Chemistry* **287**, 13532-13540, doi:10.1074/jbc.R111.310797 (2012).
- 103 Shepard, E. M., Boyd, E. S., Broderick, J. B. & Peters, J. W. Biosynthesis of complex iron-sulfur enzymes. *Current Opinion in Chemical Biology* **15**, 319-327, doi:10.1016/j.cbpa.2011.02.012 (2011).
- 104 Broderick, J. B. *et al.* H-Cluster assembly during maturation of the [FeFe]-hydrogenase. *JBIC Journal of Biological Inorganic Chemistry* **19**, 747-757, doi:10.1007/s00775-014-1168-8 (2014).
- 105 Vallese, F. *et al.* Biochemical Analysis of the Interactions between the Proteins Involved in the [FeFe]-Hydrogenase Maturation Process. *Journal of Biological Chemistry* **287**, 36544-36555, doi:10.1074/jbc.M112.388900 (2012).
- 106 Esselborn, J. *et al.* Spontaneous activation of [FeFe]-hydrogenases by an inorganic [2Fe] active site mimic. *Nature chemical biology* **9**, 607-609, doi:10.1038/nchembio.1311 (2013).
- 107 Jarrett, J. T. The Biosynthesis of Thiol- and Thioether-Containing Cofactors and Secondary Metabolites Catalyzed by Radical S-Adenosylmethionine Enzymes. *Journal of Biological Chemistry*, doi:10.1074/jbc.R114.599308 (2014).

- 108 Lanz, N. D. & Booker, S. J. Auxiliary iron–sulfur cofactors in radical SAM enzymes. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1853**, doi:10.1016/j.bbamcr.2015.01.002 (2015).
- 109 Arragain, S. *et al.* Post-translational modification of ribosomal proteins: structural and functional characterization of RimO from *Thermotoga maritima*, a radical S-adenosylmethionine methylthiotransferase. *The Journal of biological chemistry* **285**, 5792-5801, doi:10.1074/jbc.M109.065516 (2010).
- 110 Forouhar, F. *et al.* Two Fe-S clusters catalyze sulfur insertion by radical-SAM methylthiotransferases. *Nature Chemical Biology* **9**, 333-338, doi:10.1038/nchembio.1229 (2013).
- 111 Flühe, L. & Marahiel, M. A. Radical S-adenosylmethionine enzyme catalyzed thioether bond formation in sactipeptide biosynthesis. *Current Opinion in Chemical Biology* **17**, 605-612, doi:10.1016/j.cbpa.2013.06.031 (2013).
- 112 Flühe, L. *et al.* The radical SAM enzyme Alba catalyzes thioether bond formation in subtilisin A. *Nature Chemical Biology* **8**, 350-357, doi:10.1038/nchembio.798 (2012).
- 113 Flühe, L. *et al.* Two [4Fe-4S] Clusters Containing Radical SAM Enzyme SkfB Catalyze Thioether Bond Formation during the Maturation of the Sporulation Killing Factor. *Journal of the American Chemical Society* **135**, 959-962, doi:10.1021/ja310542g (2013).
- 114 Nakai, T. *et al.* The Radical S-Adenosyl-L-methionine Enzyme QhpD Catalyzes Sequential Formation of Intra-protein Sulfur-to-Methylene Carbon Thioether Bonds. *Journal of Biological Chemistry*, doi:10.1074/jbc.M115.638320 (2015).
- 115 Fugate, C. J. *et al.* 9-Mercaptodethiobiotin Is Generated as a Ligand to the [2Fe–2S]<sup>+</sup> Cluster during the Reaction Catalyzed by Biotin Synthase from *Escherichia coli*. *Journal of the American Chemical Society* **134**, 90429045, doi:10.1021/ja3012963 (2012).
- 116 Fugate, C. J. & Jarrett, J. T. Biotin synthase: Insights into radical-mediated carbon–sulfur bond formation. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1824**, 1213-1222, doi:10.1016/j.bbapap.2012.01.010 (2012).
- 117 Jarrett, J. T. The novel structure and chemistry of iron–sulfur clusters in the adenosylmethionine-dependent radical enzyme biotin synthase. *Archives of Biochemistry and Biophysics* **433**, 312-321, doi:10.1016/j.abb.2004.10.003 (2005).
- 118 Harmer, J. E. *et al.* Structures of lipoyl synthase reveal a compact active site for controlling sequential sulfur insertion reactions. *The Biochemical journal* **464**, 123-133, doi:10.1042/bj20140895 (2014).

## **Partie Résultats**

## Répartition des chapitres décrivant les résultats

Le **Chapitre I** de la *Partie Résultats* sera consacré aux travaux de recherche du substrat de HydE, basée sur l'étude structurale de l'enzyme par cristallographie aux rayons X. Ces expériences d'analyses structurales reposaient sur la co-purification d'un complexe de HydE avec un éventuel intermédiaire de la réaction, mais surtout sur la co-cristallisation de l'enzyme avec des ligands potentiels, ces derniers sélectionnés sur la base de résultats d'accostage moléculaire *in silico*. Nous avons donc utilisé une approche similaire à la méthode, qui est connue sous son nom anglais - le "rational drug design".

Les ligands ainsi identifiés ont ensuite été utilisés dans une analyse fonctionnelle, qui reposait sur la capacité de ces ligands à augmenter l'activité de l'hydrogénase dans un système de maturation dans l'extrait cellulaire d'*Escherichia coli* contenant des maturases HydE, HydG et HydF. Les résultats de cette analyse seront exposés dans le **Chapitre II**.

Le **Chapitre III** correspond à l'étude exhaustive d'une réaction radicalaire catalysée par la protéine HydE sur un groupe de substrats nouvellement identifiés. Plusieurs approches complémentaires ont été utilisées; une étude de clivage de la SAM *in vitro* par HydE en présence des ligands avec une caractérisation des produits par chromatographie en phase liquide couplée à la spectrométrie de masse, une étude cinétique de la réaction dans les cristaux protéiques et une étude cinétique par RMN du  $^{13}\text{C}$  d'accumulation d'un des produit. Les résultats de ces expériences ont été utilisés pour effectuer des calculs théoriques et proposer un modèle de clivage de la SAM et d'attaque radicalaire catalysés par les enzymes à radical SAM. Ces travaux ont aussi contribué à la compréhension du mécanisme de création d'une liaison carbone - soufre par un processus radicalaire. Ces travaux ont été acceptés pour publication dans le journal *Nature Chemistry* (l'article Rohac *et al.* 2016 en Annexe).

En se servant de la protéine HydE comme d'une métalloprotéine modèle, nous avons obtenue des formes particulières de centres [Fe-S] qui correspondent à des états intermédiaires d'assemblage ou de dégradation aérobie de ces agrégats. Ceci nous a permis de proposer un modèle détaillé d'interconversion entre différentes formes de centres [Fe-S]. Cette étude cristallographique a fait l'objet d'une publication le journal *PNAS*<sup>1</sup> et sera brièvement commentée dans le mini **Chapitre IV** (l'article Nicolet *et al.* 2013 en Annexe).

Les centaines de structures de la protéine HydE analysées au cours de mes travaux de thèse nous ont conduit à l'identification d'une nouvelle structure de HydE qui présente un changement conformationnel unique du brin- $\beta$  C-terminal. Le mini **Chapitre V** est consacré à l'analyse de cette structure et de son impact sur le fonctionnement de cette enzyme et sur celui d'un sous-groupe de protéines à radical SAM, qui partagent la même topologie que la maturase HydE.

# Chapitre I : Recherche du substrat de HydE

## I-1 Introduction du Chapitre

L'objectif principal de ma thèse était de faire progresser notre compréhension de la maturation de l'hydrogénase à [FeFe], notamment par l'identification du substrat de la maturase HydE. Comme nous l'avons vu précédemment, quelques années après la découverte des trois maturases par Posewitz et ses collaborateurs<sup>2</sup>, d'importants progrès ont été réalisés dans l'identification des rôles physiologiques des maturases HydG et HydF. Il a été prouvé que HydG produit des ligands diatomiques CO et CN<sup>-</sup> à partir de la L-Tyrosine<sup>3-5</sup> et que HydF agit comme une protéine échafaudage<sup>6,7</sup> sur laquelle le sous-agrégat à deux Fe est synthétisé avant d'être inséré dans l'hydrogénase à [FeFe]<sup>8-11</sup>. En procédant par élimination, la protéine HydE devrait être impliquée dans la synthèse du ligand DTMN<sup>12</sup>. Une première caractérisation spectroscopique et biochimique de la protéine HydE issue de *T. maritima* a confirmé qu'il s'agissait bien d'une protéine à radical SAM, capable de cliver ce cofacteur<sup>13</sup>. La structure cristallographique de l'enzyme<sup>14</sup> reste jusqu'à aujourd'hui celle réalisée à la plus haute résolution pour une protéine à radical SAM [I-8.8.2]. Malgré ces avancées, le substrat, le produit exact, ainsi que la réaction catalysée par cette enzyme restent, à ce jour, inconnus. Cela dit, la bonne connaissance de la structure de l'enzyme et les expériences de maturation de l'hydrogénase dans l'extrait effectuées par les équipes de Maria L. Ghirardi, de John W. Peters et de James R. Swartz nous ont tout de même livré des informations sur la nature du substrat de HydE.

Dans ce chapitre nous allons voir, dans un premier temps, comment nous avons utilisé l'approche dite "rational drug design" pour identifier des ligands de la protéine HydE. Au final cette approche nous a conduit à l'identification de deux grands groupes de ligands; le premier groupe est composé de molécules polycarboxylées et le second contient des dérivés 1-thia-3-aza-cyclopentanes (thiazolidines) et des dérivés 1-thia-3-aza-cyclohexanes (thiazinanes). Seuls les membres du second groupe contenant un cycle thiazolidine présentent une affinité élevée pour le site actif de l'enzyme étudiée et forment des complexes très stables avec *TmHydE*. Ces ligands ont tous la particularité de présenter un atome de soufre, engagé dans une liaison thioéther, à l'atome C5' de la SAM, qui est sensé de porter l'attaque radicalaire.

Nous allons également commenter nos résultats sur nos tentatives de co-purification d'un complexe de HydE avec son substrat, son produit ou un éventuel intermédiaire de sa réaction physiologique.

## I-2 Le substrat de HydE - un métabolite d'*E.coli*

La coexpression de l'hydrogénase à [FeFe] avec ses trois maturases dans *E. coli*, un organisme dépourvu de ces protéines, permet l'obtention d'une hydrogénase active<sup>2,6,15-17</sup>. Il est donc fort probable que le(s) substrat(s) de HydE soit un ou plusieurs métabolites ubiquitaires – comme c'est le cas pour HydG qui utilise la tyrosine pour synthétiser les ligands CO et CN<sup>-</sup>. Cette hypothèse est corroborée par le fait, qu'aucune protéine dédiée à la production du substrat de HydE n'a été identifiée, ni par l'invalidation des gènes par la génération de mutants par insertion, ni par l'inspection des portions du génome adjacentes aux gènes des trois maturases (opéron *Hyd*). Les expériences de génération de mutants d'insertion laissent supposer que le substrat de HydE intervient dans plusieurs voies métaboliques ou bien qu'il soit indispensable à la survie de la cellule. En effet, si un seul gène était responsable de la synthèse de ce métabolite, sa délétion aurait permis son identification comme faisant partie du système de maturation de l'hydrogénase à [FeFe]. Dans le cas contraire, on peut imaginer que plusieurs voies métaboliques conduisent à la biosynthèse de la molécule en question. Alternativement, on peut imaginer que c'est bien un seul gène qui est responsable de la synthèse de cette molécule, mais que l'incapacité de la cellule de produire cette molécule conduit à sa mort. Ainsi il serait impossible d'identifier ce gène par cette méthode, parce que son inactivation serait létale pour la cellule avant même qu'un test d'activité hydrogénase ne soit réalisable.

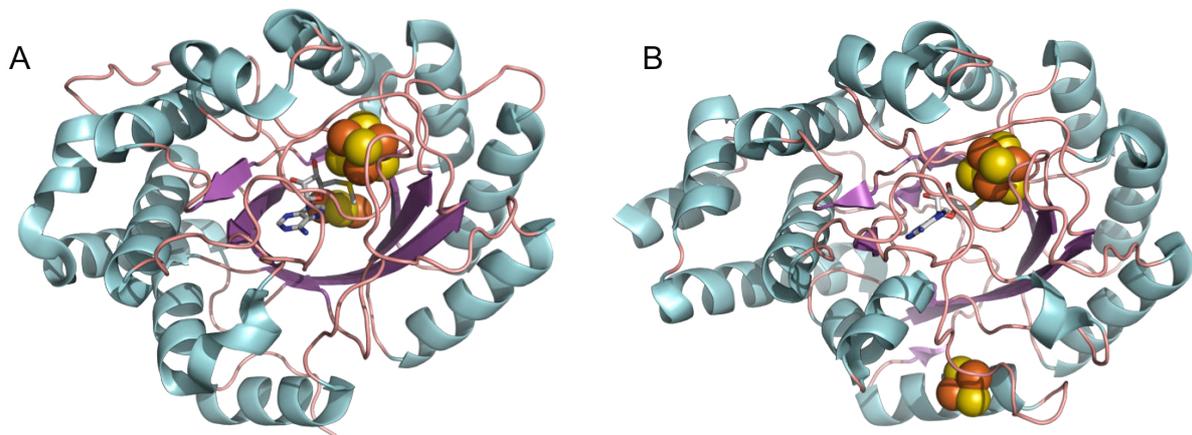
## I-3 Réflexions sur la provenance des atomes de soufre dans le ligand DTMN

Au début de ma thèse nous avons commencé notre recherche du substrat de HydE par une analyse de la provenance des atomes de soufre du ligand DTMN. Ce dernier contient deux atomes de soufre, chacun d'eux engagé dans une liaison C-S. En identifiant les mécanismes plausibles de création de ces liaisons on peut restreindre la recherche du substrat sur certains groupes de métabolites plutôt que sur d'autres.

### I-3.1 Implication des centres [Fe-S] dans les réactions d'insertion de soufre

Comme cela a été mentionné dans l'Introduction Bibliographique de ce manuscrit, la protéine HydE présente une forte ressemblance avec une autre protéine à radical SAM – la biotine synthase BioB. Cette dernière constitue, avec LipA, un petit sous-groupe de protéines à radical SAM, qui catalysent une réaction d'insertion de soufre sur un carbone aliphatique dans un ligand organique<sup>18</sup> [voir IB-5]. BioB catalyse l'insertion d'un atome de soufre lors de la dernière étape de la synthèse de la vitamine B<sub>8</sub>. LipA catalyse une double insertion de soufre sur l'octanoate pour le transformer en cofacteur lipoate. Ces deux protéines utilisent pour cela un centre [FeS] auxiliaire comme donneur sacrificiel de soufre. La protéine HydE de *T. maritima* possède elle aussi un second centre [FeS]. On pourrait donc être amené à conclure que ces trois enzymes catalysent des réactions d'insertion de soufre à partir d'un centre [FeS] sur des substrats différents. Cependant, il y a une différence majeure au niveau de l'emplacement de son agrégat auxiliaire. Dans BioB et dans LipA le second agrégat [FeS]

se trouve à l'intérieur de la cavité centrale, à côté du site de fixation du substrat et à l'opposé du centre  $[\text{Fe}_4\text{S}_4]$  lié à la SAM<sup>19,20</sup>. Contrairement à ces deux enzymes, dans HydE le second centre  $[\text{FeS}]$  se trouve à la surface de la protéine et il est exposé au solvant [Figure-I-17]. Ce second agrégat se situe à 21Å de l'agrégat lié à la SAM et à 17Å de l'atome C5' de la SAM. Il est donc peu probable que ce second centre  $[\text{FeS}]$ , situé loin du site d'activation radicalaire du substrat et à l'extérieur de la cavité centrale, puisse servir de source directe de soufre. D'autant plus que les comparaisons de séquences en acides aminées de protéines HydE issues de différents organismes montrent l'existence de deux groupes distincts : un groupe possédant trois résidus cystéinyles impliqués dans la fixation de ce second agrégat et un second groupe sans ces trois cystéines<sup>2,14,21</sup> [voir l'alignement multiséquence en Annexe]. Un autre argument contre l'hypothèse d'une insertion de soufre à partir d'un agrégat  $[\text{FeS}]$  auxiliaire repose sur le fait que la mutation de ses trois cystéines ligands en alanines dans HydE de *C. acetobutylicum*, n'affecte pas l'activation de l'hydrogénase *in vivo*<sup>14</sup>. D'ailleurs, le remplacement de HydE de *C. acetobutylicum* par son homologue de *B. thetaiotaomicron*, une enzyme qui ne possède pas les trois cystéines ligands du second agrégat  $[\text{FeS}]$ , n'affecte pas non plus la production d'hydrogénase active<sup>14</sup>. Ainsi ce second agrégat ne serait pas indispensable à la fonction de HydE.



**Figure-I-17 : Comparaison de l'emplacement du centre  $[\text{Fe-S}]$  auxiliaire dans BioB et HydE**

Structures cristallographiques de : A - *EcBioB* et B - *TmHydE*. Seules les structures secondaires de la chaîne principale sont représentées: les feuillets  $\beta$  en violet, les hélices en cyan et les boucles en rose. Les centres  $[\text{Fe-S}]$  sont représentés par les sphères et la molécule de SAH/SAM par des bâtons gris.

### I-3.2 Origines possibles des atomes de soufre dans le ligand DTMN

Les trois hypothèses les plus plausibles sur la provenance des atomes de soufre dans le ligand DTMN sont les suivantes :

- (i) HydE produit le ligand DTMN à partir d'un substrat qui contient déjà les atomes de soufre.
- (ii) HydE produit le ligand DTMN à partir d'un substrat dépourvu de soufre et à partir d'une source de soufre extérieure, comme par exemple une cystéine désulfurase.
- (iii) HydE produit un intermédiaire dans la synthèse du DTMN dépourvu de soufre. Cette molécule va ensuite réagir sur un agrégat  $[\text{Fe}_2\text{S}_2]\text{-CO}_x\text{-CN}_y$  contenu dans HydF.

Une dernière hypothèse selon laquelle les ligands CO et  $\text{CN}^-$  arriveraient après la préformation d'un centre  $[\text{Fe}_2\text{S}_2]^{10}$  nécessiterait un échange de Fe sur cet agrégat, car les récentes expériences de marquages montrent que le synthon  $\text{Fe}(\text{CO})_2(\text{CN})$ , produit par HydG, se retrouve dans le site actif de l'hydrogénase<sup>22</sup>. Un échange du Fe d'un centre  $[\text{Fe}_2\text{S}_2]$  paraît alors peu probable.

L'ensemble des données expérimentales dont on dispose sur les maturases ne permet pas de privilégier de façon précise une hypothèse par rapport aux deux autres, car certaines observations expérimentales n'ont pas qu'une seule interprétation possible. Par exemple, la co-purification de HydF avec HydG et HydE<sup>7</sup> montre que HydF interagit avec les deux autres maturases. Les mesures d'affinité par résonance plasmonique de surface ont prouvé que l'interaction entre HydF et HydE est plus forte que celle entre HydF et HydG<sup>23</sup>. Chez certains organismes, comme par exemple *C. reinhardtii*, HydE et HydF existent sous forme d'un seul gène *hyDEF*<sup>2</sup>. Ceci suggérerait que HydE et HydF forment un complexe fonctionnel pendant l'assemblage du sous-agrégat à deux Fe<sup>15</sup>. Cette observation ne contredit aucune des trois hypothèses mentionnées plus haut. On peut en effet imaginer que HydE interagisse étroitement avec HydF afin d'y insérer le ligand DTMN, une molécule facilement hydrolysable, en même temps que HydG délivre les synthons  $\text{Fe}(\text{CO})_2(\text{CN})$ . Alternativement, on peut imaginer que HydE interagisse étroitement avec HydF, parce qu'elle doit acheminer un précurseur du DTMN dans HydF, où il va réagir sur un agrégat  $[\text{Fe}_2\text{S}_2]\text{-CO}_x\text{-CN}_y$  préformé. Une collaboration mécanistique entre HydE et HydF est suggérée par l'augmentation de l'activité GTPase de HydF lors de son incubation avec HydE<sup>7</sup>.

## I-4 Méthodes d'identification du substrat d'une enzyme

Les expériences initiales d'inactivation génétiques (en anglais « gene knockout ») effectuées par Posewitz *et al.* n'ont pas permis d'extraire plus d'informations sur la nature du substrat de HydE, hormis le fait qu'il s'agissait d'un métabolite présent chez *E. coli*. L'utilisation d'autres méthodes d'identification de substrat d'une enzyme s'imposaient alors. Les techniques que nous avons utilisées ont été sélectionnées en fonction de la pertinence de l'information à obtenir, de notre savoir-faire, de l'équipement de notre laboratoire, de la possibilité d'établir des collaborations avec d'autres groupes de recherche et aussi de la volonté de ne pas forcément répéter, du moins dans un premier temps, les expériences de nos concurrents.

### I-4.1 Conceptions de tests fonctionnels pour mesurer l'activité de HydE

En absence d'information sur la réaction catalysée, ou sur la nature exacte du produit, il est difficile de concevoir des tests fonctionnels directement avec l'enzyme purifiée. Comme on le verra au début du **Chapitre III**, l'identification du substrat en mesurant l'augmentation du clivage de la SAM ne semble pas être une méthode fiable, car elle peut conduire à de faux positifs [voir paragraphe III-2.1.2]. De plus, ces expériences se font en présence d'une concentration élevée de réducteurs forts, ce qui peut, dans certains cas, interférer avec la catalyse enzymatique. Une alternative est donc d'utiliser, comme détaillé dans le **Chapitre II**, le taux d'activation de l'hydrogénase à [FeFe] pour mesurer, de façon indirecte, l'activité de HydE. Pour cela, il est indispensable de disposer d'un système de maturation d'hydrogénase *in vitro* ou au moins acellulaire, dit « cell-free », dans lequel la concentration du substrat de HydE serait le facteur limitant. La recherche du substrat pourrait alors se faire en construisant un plan factoriel de mélange de composés à partir d'une librairie de molécules dont la présence a été confirmée dans *E. coli*. Il est également envisageable de procéder à des fractionnements successifs de l'extrait d'*E. coli* pour identifier les fractions capables d'activer l'hydrogénase à [FeFe] en présence de ses maturases [voir II-6.1]. Cette approche paraît plus adaptée que l'utilisation de librairie de composés chimiques, car elle serait moins chère et en plus elle permettrait l'identification éventuelle d'autres facteurs, que des petites molécules, qui ont un impact sur l'activation de l'hydrogénase. En effet, jusqu'à présent l'activation de l'hydrogénase dans l'extrait d'*E. coli* nécessite l'utilisation d'un fort excès d'extraits avec les maturases<sup>6,17,24</sup>, ce qui prouve que le système fonctionne loin de ses conditions optimales. Deux possibilités peuvent expliquer ce phénomène: soit l'action des maturases entraîne une rapide déplétion d'un métabolite de l'extrait – potentiellement le substrat de HydE, ou bien, comme cela a été proposé par Kuchenreuther et ses collaborateurs la fraction protéique de l'extrait contiendrait un élément dont la présence augmenterait l'efficacité de la synthèse du sous-agrégat à deux Fe ou de son transfert<sup>24</sup>. Ces aspects seront commentés plus en détails dans le **Chapitre II**.

Puisque d'autres groupes de recherche avaient beaucoup d'avance dans le développement du système de maturation dans l'extrait, nous avons privilégié, dans un premier temps, d'autres approches, complémentaires à celles utilisées par nos concurrents. D'autant plus que la mise en place d'un test de maturation dans l'extrait capable de cribler des centaines de molécules et la mise en place du fractionnement de l'extrait nécessite l'utilisation d'un système robotique, dont nous ne disposons pas.

#### **I-4.2 Approche similaire au « rational drug design »**

Puisque notre laboratoire a été le premier à avoir accédé à la structure cristalline de HydE nous avons décidé d'utiliser notre bonne connaissance de la structure de l'enzyme pour en déduire des informations qui pourraient nous mener vers l'identification du substrat. Une première présélection des molécules issues de la base de données ZINC<sup>25</sup> a été effectuée, en se basant sur les informations extraites de la structure. Ces molécules ont ensuite été utilisées pour des tests d'accostage moléculaire *in silico*, ceci dans le but d'identifier de potentiels ligands de HydE. Grâce à la facilité avec laquelle on peut obtenir des cristaux diffractant à haute résolution il est possible de confirmer rapidement la fixation de tels ligands dans le site actif de HydE. Ces molécules pourraient constituer des fragments du substrat ou du produit ou encore mimer les états intermédiaires et donc apporter des informations sur la réaction catalysée par l'enzyme. Cette approche s'appelle « rational drug design » et elle est basée sur la relation réciproque entre la structure d'une enzyme et sa fonction<sup>26,27</sup>. L'alternance d'expériences d'accostage moléculaire *in silico* et de co-cristallisation avec de potentiels ligands aboutie, en théorie, à l'identification de molécules plus complexes qui ont une forte affinité pour la protéine cible<sup>28</sup>.

##### **I-4.2.1 Limitations de la méthode**

Il est cependant important de comprendre les limitations de l'accostage moléculaire *in silico* basé sur une structure cristalline de la protéine étudiée. Puisqu'il est difficile de prendre en compte toutes les positions alternatives de chaînes latérales des résidus du site actif, on choisit une seule position par résidu pour construire un modèle optimisé de la protéine. En choisissant une seule orientation d'une chaîne latérale il y a un risque que cette orientation ne soit pas compatible avec la fixation du substrat. Il est également possible que le repliement de la protéine, qui est sélectionné par l'empilement cristallin, ne soit pas compatible avec la fixation du substrat. On peut envisager qu'un réarrangement local de sa structure soit nécessaire pour permettre la reconnaissance du substrat par l'enzyme. Nous pouvons prendre pour exemple la maturase HydG, dont plusieurs conditions de cristallisation sont connues, mais aucune expérience de co-cristallisation ou de trempage de cristaux avec son substrat physiologique - la L-tyrosine - n'a produit une structure du complexe<sup>29,30</sup>.

Il faut aussi considérer la conservation de molécules d'eau dans le modèle optimisé de HydE. En effet, en absence de ligands la cavité est occupée par trois ions chlorure et plusieurs molécules d'eau [voir paragraphe IB-3.2]. Ces molécules peuvent être divisées en deux catégories. La première est constituée de celles qui se trouverait dans le site actif même en présence du substrat, parce qu'elles sont nécessaires pour sa fixation ou bien parce qu'elles sont impliquées dans le mécanisme catalytique, comme par exemple dans une réaction d'hydrolyse. L'autre catégorie comprend les molécules d'eau qui miment les groupements polaires du substrat.

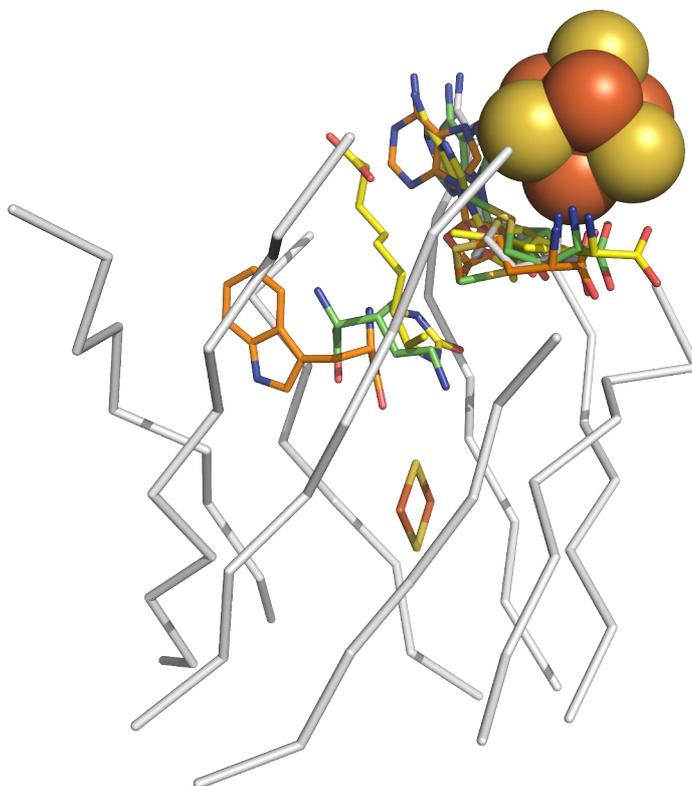
Tandis qu'il est préférable de garder dans le modèle optimisé les molécules d'eau de la première catégorie il est indispensable d'éliminer celles appartenant à la seconde catégorie, car ces dernières empêcheraient l'accostage du substrat. Il est extrêmement difficile d'attribuer correctement les molécules d'eau du site actif à une de ces deux catégories, surtout quand on ne connaît ni le substrat, ni la réaction catalysée.

Le dernier problème constitue le principe même du classement de ligands en fonction de l'enthalpie libre de fixation. Cette méthode a la tendance de sélectionner les ligands qui ont le plus d'interactions possibles avec la protéine et qui se retrouvent par la suite avec des scores les plus élevés. Cette approche s'est révélée adaptée pour la conception d'inhibiteurs. Par contre, l'identification de substrats physiologiques s'est montrée beaucoup plus difficile. D'autant plus que les enzymes présentent des affinités élevées surtout vis-à-vis d'un état de transition particulier, plutôt que vis-à-vis de l'état fondamental (substrat) qu'on trouve dans la base de données<sup>31</sup>.

Conscients de ces limitations nous avons procédé à une analyse détaillée des propriétés physico-chimiques de la cavité de HydE afin d'en extraire des informations utiles pour la conception d'une expérience d'accostage moléculaire *in silico*. Nous avons fait ce choix stratégique en connaissance de cause.

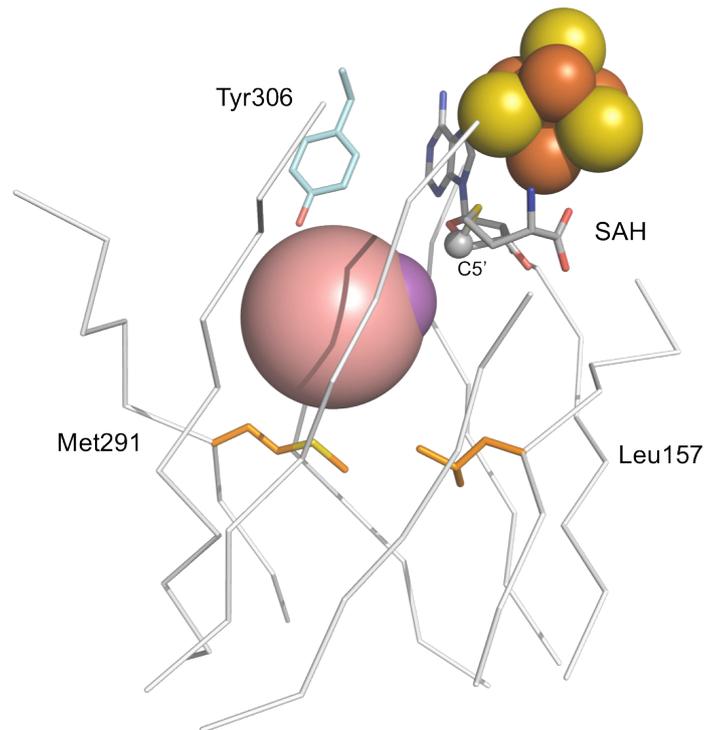
## I-5 Analyse des caractères physico-chimiques de la cavité catalytique

Comme mentionné dans le paragraphe IB-3.2 de l'Introduction Bibliographique, la détermination de la structure cristalline de l'enzyme a mis en évidence la présence d'une large cavité, qui se situe le long de l'axe du tonneau ( $\beta\alpha$ )<sub>8</sub>. La face interne de la cavité est tapissée de résidus conservés chez les protéines HydE issues de différents organismes. Ces résidus sont cependant différents de ceux que l'on trouve dans BioB, ce qui montre leur implication dans la réaction catalysée spécifiquement par HydE<sup>14,32</sup>. Une couronne de résidus hydrophobes, hautement conservés, divise la cavité en deux parties. La partie supérieure contient les sites de fixation de la SAM et de l'agrégat [Fe<sub>4</sub>S<sub>4</sub>] conservé et il est donc probable que cette partie soit l'endroit où l'initiation de la réaction radicalaire a lieu<sup>14</sup>. En comparant la structure de HydE à d'autres structures de protéines à radical SAM de topologie similaire, qui contiennent le substrat ou le produit de leur réaction [Figure-I-18], comme BioB<sup>19</sup>, PylB<sup>33</sup> ou encore NosL<sup>34</sup> et en prenant en compte la conservation de résidus dans la séquence de HydE<sup>14</sup> il est possible de définir une région de fixation de substrat [Figure-I-19].



**Figure-I-18** : Comparaison de l'emplacement du ligand par rapport à la position du cofacteur SAM dans le site actif des protéines à radical SAM : BioB, PylB et NosL

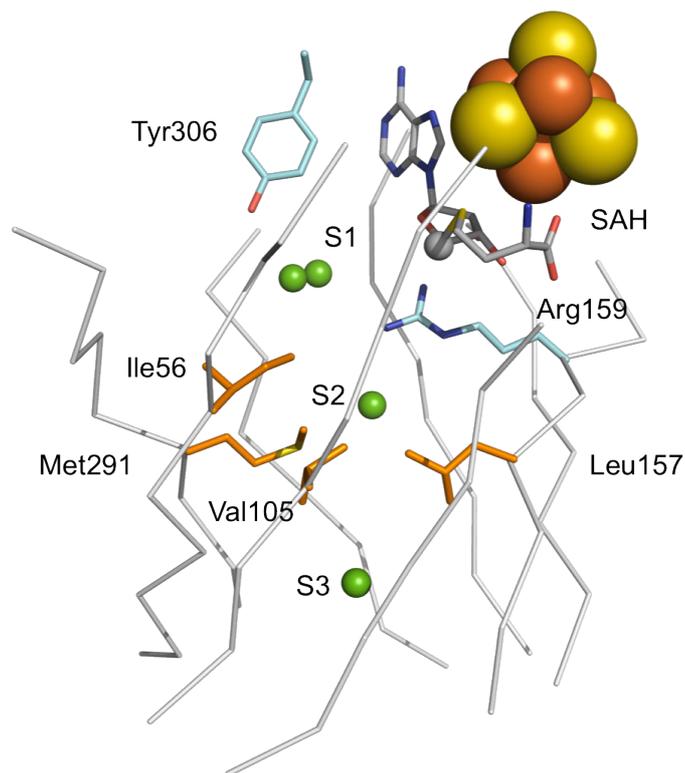
Le tonneau ( $\beta\alpha$ )<sub>8</sub> de *TmHydE* est représenté par des brins blancs. Le centre [Fe<sub>4</sub>S<sub>4</sub>] SAM radical, commun à ces quatre structures, est représenté par des sphères, tandis que le centre [Fe<sub>2</sub>S<sub>2</sub>], conservé uniquement chez la protéine BioB, est représenté par des bâtons et se situe au milieu du tonneau. Les ligands sont coloriés de la façon suivante: en orange - L-Trp et SAH dans NosL, en vert - la (3R)-3-méthyle-D-ornithine et SAM dans PylB et en jaune - la déthiobiotine et SAM dans BioB.



**Figure-I-19 : Région probable de fixation du substrat dans la protéine HydE**

Le tonneau  $(\beta\alpha)_8$  de *TmHydE* est représenté par des brins blancs. La SAH est représenté par des bâtons gris et l'atome C5' par une sphère grise. Le centre  $[\text{Fe}_4\text{S}_4]$  SAM radical est représenté par des sphères. En rose: l'enveloppe minimale de la zone du fixation du substrat et en magenta, la zone probable d'attaque radicalaire du substrat par le C5' de la SAM. La position des chaînes latérales des résidus Met291 et Leu157, permet de visualiser l'emplacement de la ceinture de résidus hydrophobes.

Le volume de cette région, ainsi que sa composition en acides aminés, définissent des contraintes en matière de taille et de charge du substrat. Les protéines à radical SAM présentant un repliement en tonneau  $(\beta\alpha)_8$  complet utilisent comme substrat des métabolites de faible poids moléculaire<sup>32</sup>, il devrait donc en être autant pour la maturase HydE. Dans la majorité de structures de protéines à radical SAM avec leur substrats on remarque que la distance entre l'atome C5' de la SAM et l'atome du substrat qui va céder son atome d'hydrogène varie entre 4,2 Å et 4,5 Å<sup>35,36</sup>. Ceci permet alors de définir une zone dans HydE, où devrait se trouver l'atome qui sera la cible de l'attaque radicalaire [Figure-I-19]. Il nous paraît raisonnable de supposer que le site de fixation du substrat sera délimité d'une part par la Tyr306, un résidu strictement conservé, et de l'autre part par une couronne hydrophobe, qui contient des résidus strictement (Leu157 & Met291) et hautement (Ile56 & Val105) conservés<sup>14</sup> [voir Annexes]. La surface de la cavité dans cette zone présente un potentiel positif et la structure montre la présence de deux site de fixation de chlorure, le site S1 qui interagit avec le groupement -OH de la Tyr306 et le site S2 qui interagit avec le guanidinium de l'Arg159 hautement conservée [Figure-I-20]. Il en a été conclu que le substrat pourrait donc avoir deux groupements chargés négativement. Quand le substrat de l'enzyme est présent dans la cavité, le site S1 est très probablement occupé par un groupement carboxylate. En effet, comme on peut le voir sur la Figure-I-20, l'ion chlorure se fixe dans ce site en deux conformations alternatives, ce qui mimerait la position des atomes d'oxygène d'un groupement carboxylate [voir IB-3.2].



**Figure-I-20 : Composition de la cavité de HydE de *T. maritima***

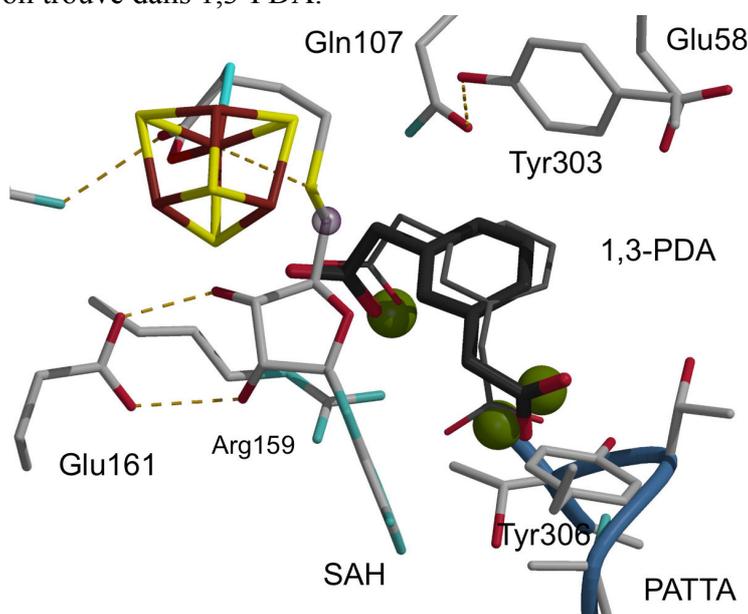
Le tonneau  $(\beta\alpha)_8$  de *TmHydE* est représenté par des brins blancs. La SAH est représenté par des bâtons gris et l'atome C5' par une sphère grise. Le centre  $[Fe_4S_4]$  SAM radical est représenté par des sphères. En cyan: les résidus conservés Tyr306 et Arg159 et en orange sont représentés les résidus conservés de la ceinture hydrophobe. Les sphères vertes représentent l'emplacement des ions chlorures dans les trois sites de fixation d'anions S1, S2 et S3.

## I-6 Premières expériences d'accostage moléculaire

Avant mon arrivée au laboratoire, Yvain Nicolet a effectué des expériences d'accostage moléculaire *in silico* avec un groupe de 20 000 molécules sélectionnées dans la base ZINC<sup>25</sup> sur les critères de la charge (charge nette entre -2 et +1 dans la gamme de pH 7.0  $\pm$ 2) et sur les critères de la taille (masse comprise entre 20Da et 150Da)<sup>14</sup>. Cette première tentative n'a pas permis d'identifier de potentiels ligands de HydE. Par contre, la comparaison des meilleurs ligands virtuels a confirmé que le substrat de HydE pourrait être une molécule anionique, possédant au moins un groupement carboxylate. En effet, les 50 molécules avec le plus haut score ont toutes une charge nette comprise entre -2 et 0 et parmi les 20 meilleurs ligands virtuels, 16 possèdent un groupement carboxylate qui se loge dans le site de fixation d'anion S1. Quand le ligand potentiel portait un groupement chargé positivement, ce dernier se situait systématiquement entre les chaînes latérales des acides amines strictement conservés Glu58 et Gln107.

### I-6.1 Complexe avec l'acide 1,3-phénylènediacétique

Une seconde expérience d'accostage moléculaire *in silico* avec un groupe de 64 000 molécules, sélectionnées sur la base de critères similaires aux ceux de l'expérience précédente (charge nette comprise entre -2 et 0 dans la gamme de pH  $7.0 \pm 2$ , masse inférieure à 200 Da), a permis d'identifier un seul ligand potentiel - l'acide 1,3-phénylènediacétique - dont la fixation dans le site actif de HydE a été confirmée expérimentalement par sa co-cristallisation avec l'enzyme [Figure-I-21]. Comme cette molécule n'est pas un métabolite naturel, il n'est pas possible qu'il s'agisse du substrat de HydE. Les deux groupements carboxylate de ce ligand occupent les deux sites de fixation d'anions S1 et S2. Ce résultat suggère que le substrat naturel de l'enzyme pourrait contenir deux groupements carboxylates avec le même espacement que l'on trouve dans 1,3-PDA.



**Figure-I-21** : Comparaison de structures du site actif de HydE issues d'accostage moléculaire *in silico* et de co-cristallisation avec de l'acide 1,3-phénylènediacétique

La conformation de la molécule de l'acide 1,3-phénylènediacétique obtenue par la co-cristallisation est représentée par des bâtons épais et celle obtenue par l'accostage moléculaire *in silico* par des bâtons fins. Les ions chlorures dans les sites de fixations d'anions S1 et S2 sont représentés par des sphères vertes.

### I-6.2 Tests de co-cristallisation avec des petites molécules mimétiques de fragments fonctionnels du substrat

Dans le but d'acquérir plus d'informations sur les différents groupements chimiques que le substrat pourrait contenir la protéine a été co-cristallisée avec des petites molécules, tels que l'acétate, l'acétamide et le formamide. Leur fixation dans HydE aurait pu non seulement nous renseigner sur la présence de sites fonctionnels potentiels dans la cavité, mais nous aurais aussi permis de reconstituer, au moins en partie, le substrat bloc par bloc. Malheureusement cette approche n'a pas été fructueuse. L'affinité des fragments pour la cavité de HydE n'est pas suffisante pour permettre leur fixation, et ce, même à des concentrations de l'ordre de 100 mM.

## I-7 Affinement d'hypothèses sur la nature du substrat de HydE

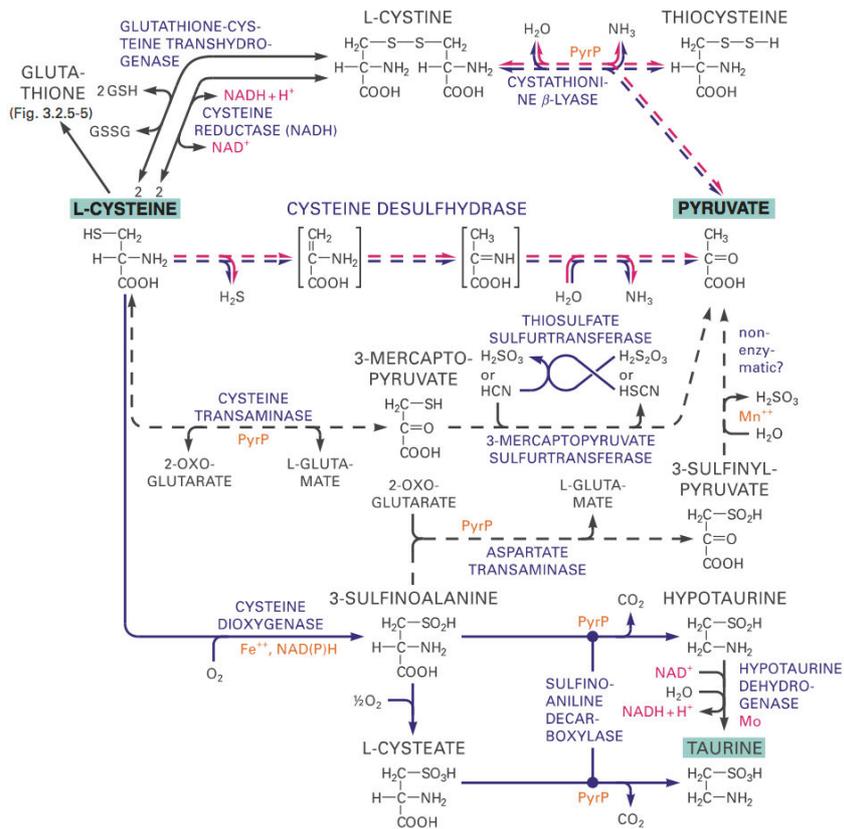
Comme détaillé dans les paragraphes précédents, les données expérimentales suggèrent que le substrat de l'enzyme serait un métabolite ubiquitaire. La structure en tonneau ( $\beta\alpha$ )<sub>8</sub> complet de l'enzyme laisse penser, que cette molécule devrait être de taille relativement petite<sup>14,32</sup>. Les modélisations suggèrent que le substrat contiendrait deux groupements carboxylates espacés de façon à ce que chacun d'eux puisse se loger dans un des deux sites de fixation d'anions S1 et S2, situés dans la partie supérieure de la cavité centrale de HydE.

Dans la partie inférieure de la cavité se trouve le site de fixation d'anions S3. Ce dernier possède une forte affinité pour le thiocyanate [Figure-I-20][voir IB-3.2]. Les changements structuraux engendrés par la fixation du thiocyanate nous ont fait penser que HydE pourrait produire soit le ligand DTMN entier soit un précurseur soufré. La fixation de thiocyanate dans le site S3 entraîne un déplacement des chaînes latérales de la Leu157 et de la Pro266, ce qui a pour effet la fermeture de la couronne de résidus hydrophobes [Figure-IB-13]. Le thiocyanate se retrouve alors isolé des molécules d'eau qui occupent la partie supérieure de la cavité. Puisque le thiocyanate mime une partie du ligand DTMN, on peut imaginer que ce dernier (ou un précurseur lui ressemblant) aurait le même comportement. C'est-à-dire qu'une fois synthétisé, il se retrouverait au fond de la cavité, isolé du solvant. Ainsi il serait à l'abri d'une éventuelle hydrolyse avant son transfert dans HydF<sup>12</sup>.

En proposant que le produit de HydE soit porteur d'atomes de soufre, nous écartons l'hypothèse (iii) sur l'origine possible d'atomes de soufre dans le ligand DTMN [I-3.2]. Donc, comme mentionné plus haut, les seules options qui restaient sont les suivantes : soit le substrat de HydE ne contient pas de soufre et il réagit, au cours de la réaction dans HydE, avec un donneur externe de soufre ou bien c'est le substrat lui-même qui apporte les atomes de soufre nécessaires à la fabrication du ligand DTMN. La première possibilité nous paraît moins probable, parce qu'à ce jour aucune protéine dédiée à l'insertion du soufre dans HydE n'a été identifiée. Ainsi, et par élimination, nous avons privilégié l'hypothèse que le substrat de HydE contient déjà les atomes de soufre.

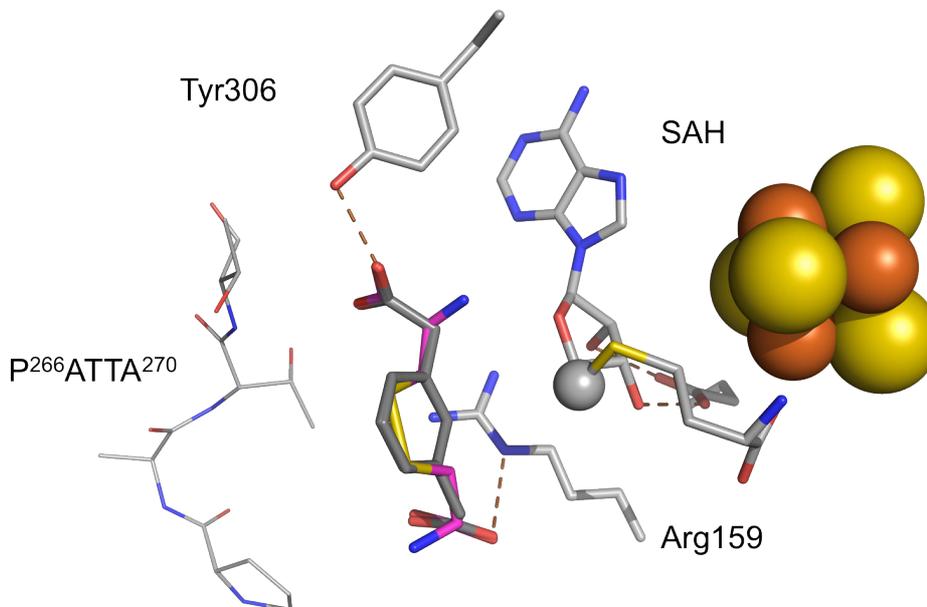
## I-8 L-Cystine comme un substrat potentiel de HydE

Nous avons ensuite essayé de trouver des molécules, qui seraient en accord avec les hypothèses que nous avons émises sur la nature du substrat. Notre meilleur candidat était la forme oxydée de la L-cystéine – la L-cystine. C'est un métabolite naturel [Figure-I-22], de petite taille qui contient tous les atomes nécessaires pour la fabrication du ligand DTMN. Surtout, cette molécule contient, à l'image du ligand dithiométhylamine, deux atomes de soufre, chacun d'eux lié à un atome de carbone. La superposition d'une de ses conformations possibles avec la structure de 1,3-PDA liée dans le site actif de HydE montre que les deux groupements carboxylates de la cystine ont le même espacement [Figure-I-23]. Ainsi la cystine devrait pouvoir se loger dans le site actif de HydE de la même manière que 1,3-PDA – c'est-à-dire, chaque groupement carboxylate occuperait un des deux sites de fixation d'anions S1 et S2.



**Figure-I-22 : Voies de conversions et de dégradation de la L-cystéine**

Image adaptée de la publication "Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology", Second Edition, ouvrage édité par Gerhard Michal & Dietmar Schomburg, publié par John Wiley & Sons, Inc., Hoboken, New Jersey, 2012;<sup>37</sup>



**Figure-I-23 : Modélisation de la L-cystine dans le site actif de HydE**

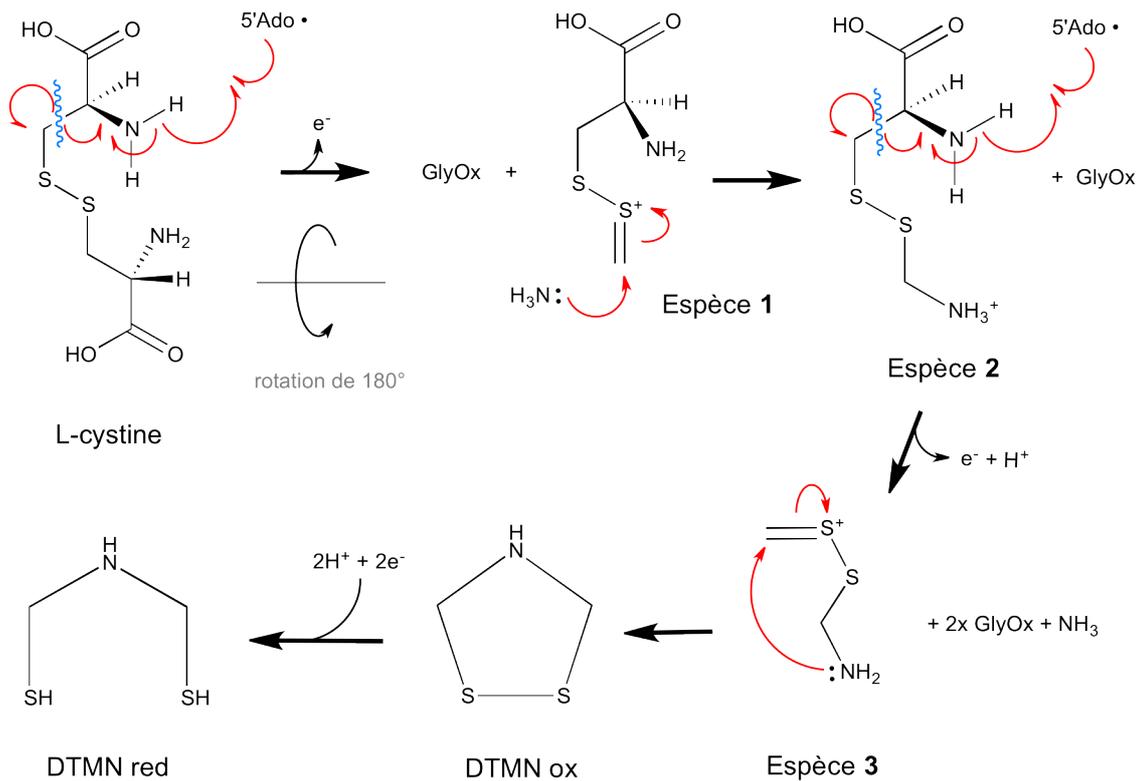
Molécule de L-cystine en magenta, placée dans le site actif de HydE de façon à ce que ces deux carboxylates se superposent avec ceux du 1,3-phénylènediacétate, ce dernier représenté par des bâtons épais gris foncé.

Il est également intéressant de mentionner que l'ajout de la L-cystéine dans le mélange de maturation dans l'extrait cellulaire d'*E. coli* augmente l'activité de l'hydrogénase à [FeFe]<sup>17</sup>. La cystéine est une des sources de soufre les plus répandues dans la cellule. Elle est mobilisée par la cystéine désulfurase lors de la fabrication de centre [FeS]. Les bactéries et les plantes, organismes chez lesquelles on peut trouver l'hydrogénase à [FeFe], sont capables de convertir la L-serine en L-cystéine via l'acétylserine. L'acétyle est directement échangé avec le H<sub>2</sub>S, qui lui, provient de la réduction de sulfate inorganique<sup>37</sup>. Ainsi la cystéine constitue la première molécule organique soufrée fabriquée par la cellule. Il a été conclu, que l'augmentation de l'activité hydrogénase après l'ajout de la cystéine peut s'expliquer par une meilleure reconstitution de protéines à [Fe-S], dont les maturases. Cependant, puisque la cystéine intervient à un certain niveau de la biosynthèse de toutes les molécules soufrées, il est possible d'interpréter le résultat expérimental différemment ; plus précisément, on pourrait envisager que la cystéine sert aussi de précurseur dans la chaîne de la biosynthèse du substrat de HydE, qui pourrait être potentiellement la L-cystine.

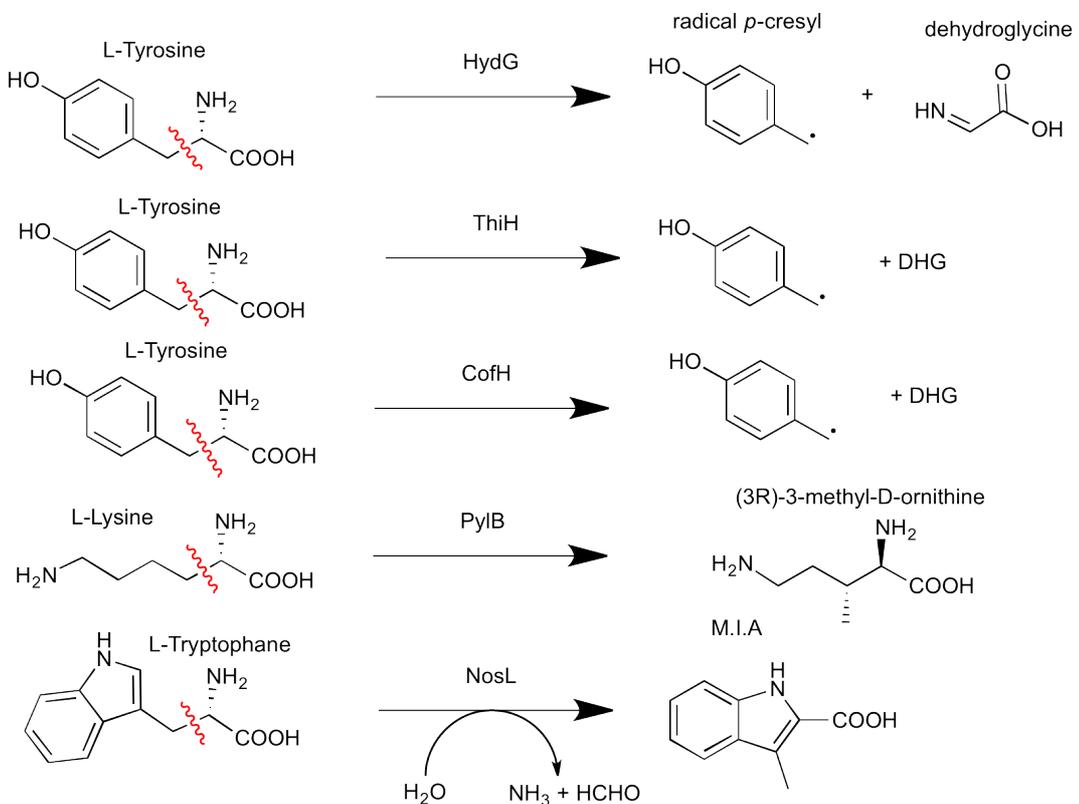
### I-8.1 Mécanisme proposé de conversion de cystine en azadithiolate

Nous avons imaginé un mécanisme dans lequel HydE transformerait la cystine en ligand DTMN en faisant deux fois la même réaction. Dans ce mécanisme l'enzyme arracherait à deux reprises un atome d'hydrogène à partir de l'amine de la cystine pour déclencher un clivage homolytique des deux liaisons C<sub>α</sub>-C<sub>β</sub> du substrat [Figure-I-24].

Il est intéressant de noter, qu'il existe un groupe d'enzymes à radical SAM, qui possèdent une structure en tonneau (β/α)<sub>8</sub> complet et qui catalysent le clivage de la liaison C<sub>α</sub>-C<sub>β</sub> d'un acide aminé<sup>5,29,33,34,38,39</sup> [Figure-I-25]. La maturase HydG, par exemple, fait partie de ce groupe. Pour certaines de ces enzymes il a été prouvé qu'elles initient la réaction radicalaire en arrachant un atome d'hydrogène de l'amine d'un acide aminé libre<sup>34</sup>. Ceci avait renforcé notre hypothèse selon laquelle la cystine pourrait être le substrat de HydE.



**Figure-I-24** : Mécanisme hypothétique proposé pour la conversion de le L-cystine en DTMN potentiellement catalysée par l'enzyme HydE.



**Figure-I-25** : Réactions catalysées par les enzymes à radical SAM qui initient la conversion du substrat en produit par un clivage de la liaison C<sub>α</sub>-C<sub>β</sub>

## I-8.2 Aspartate ammonia-lyase dans l'opéron Hyd

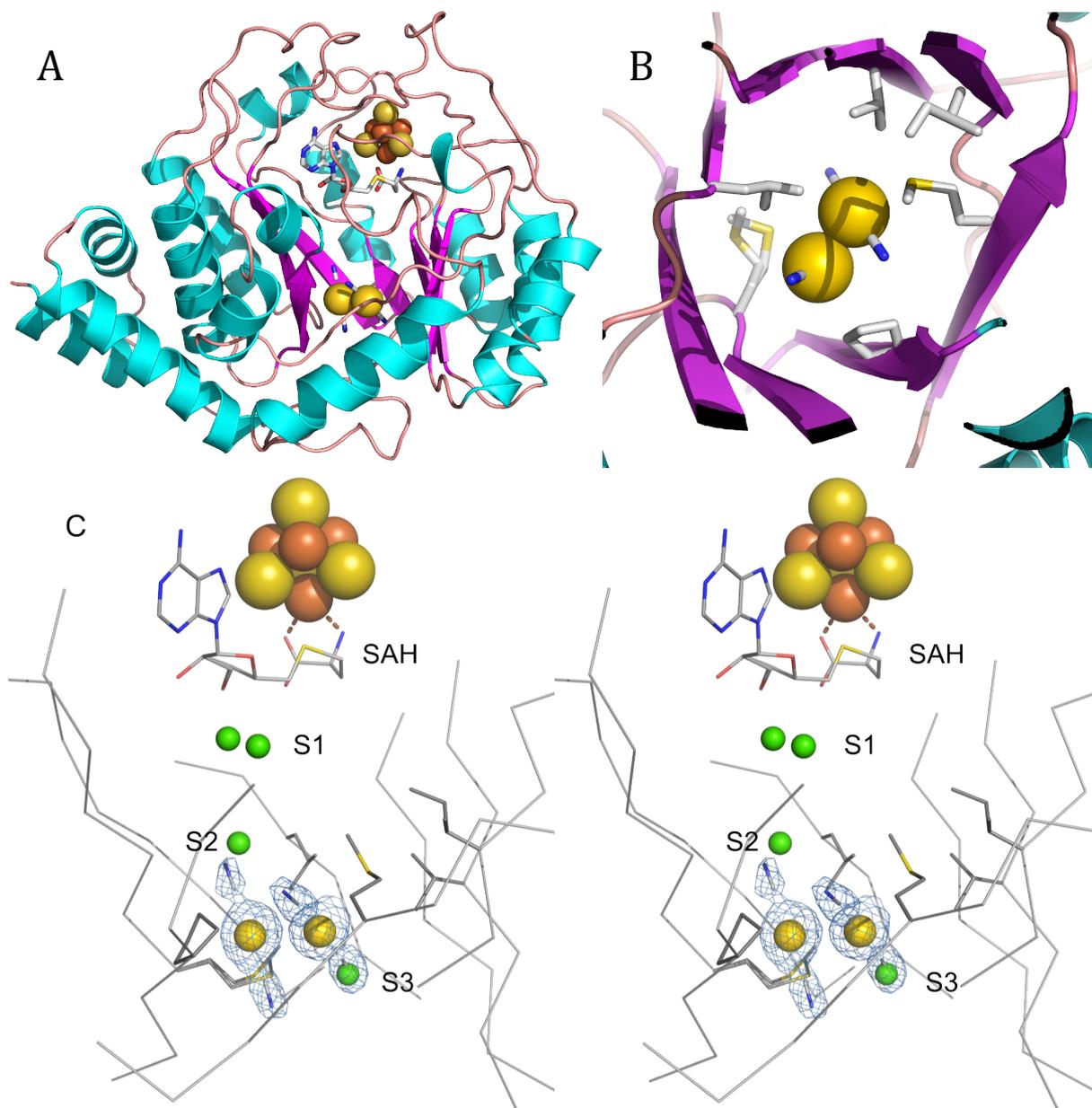
Le mécanisme que nous avons proposé expliquerait également la présence du gène codant pour l'aspartate ammonia-lyase (AspA) dans l'opéron Hyd de certains organismes [voir IB-2.1]. Le clivage homolytique d'une des deux liaisons C<sub>α</sub>-C<sub>β</sub> de la L-cystine produirait une déhydroglycine et une espèce radicalaire sur un méthyl. Ce radical pourrait être oxydé pour donner l'espèce **1**. La déhydroglycine, étant très facilement hydrolysable, se transformerait en glyoxylate, libérant un ion ammonium. Ainsi deux clivages des deux liaisons C<sub>α</sub>-C<sub>β</sub> de la L-cystine produiraient deux ions ammonium. Cependant un seul ammonium est nécessaire pour la fabrication du produit DTMN. Le surplus d'ammonium, libéré lors du clivage de la seconde liaison C<sub>α</sub>-C<sub>β</sub> de la L-cystine, serait transformé par AspA en L-aspartate selon la réaction suivante : fumarate + NH<sub>4</sub><sup>+</sup> ⇌ L-aspartate. Nous avons aussi imaginé que l'ammonium libéré dans la première partie de la réaction pourrait se loger dans le site de fixation de cation entre Q107 et E58 [I-6 & I-8.3], identifié par l'accostage moléculaire *in silico*, avant d'interagir avec l'espèce **1** pour produire l'espèce **2**.

## I-8.3 Recherche du site de fixation d'ammonium

Lors de la première caractérisation de la structure de HydE de *T. maritima* la présence de trois sites de fixation d'anions a été confirmée en utilisant des sels de bromure ou d'iodure<sup>14</sup>. De façon similaire nous voulions tester l'existence potentielle d'un site de fixation d'ions ammonium par la co-cristallisation de la protéine avec divers sels contenant un cation monovalent lourd. Nous espérons observer un pic dans la carte de différence anormale là où le cation se fixerait dans la structure de HydE. Cette fois-ci cette approche n'a pas mis en évidence l'existence d'un tel site de fixation de cation.

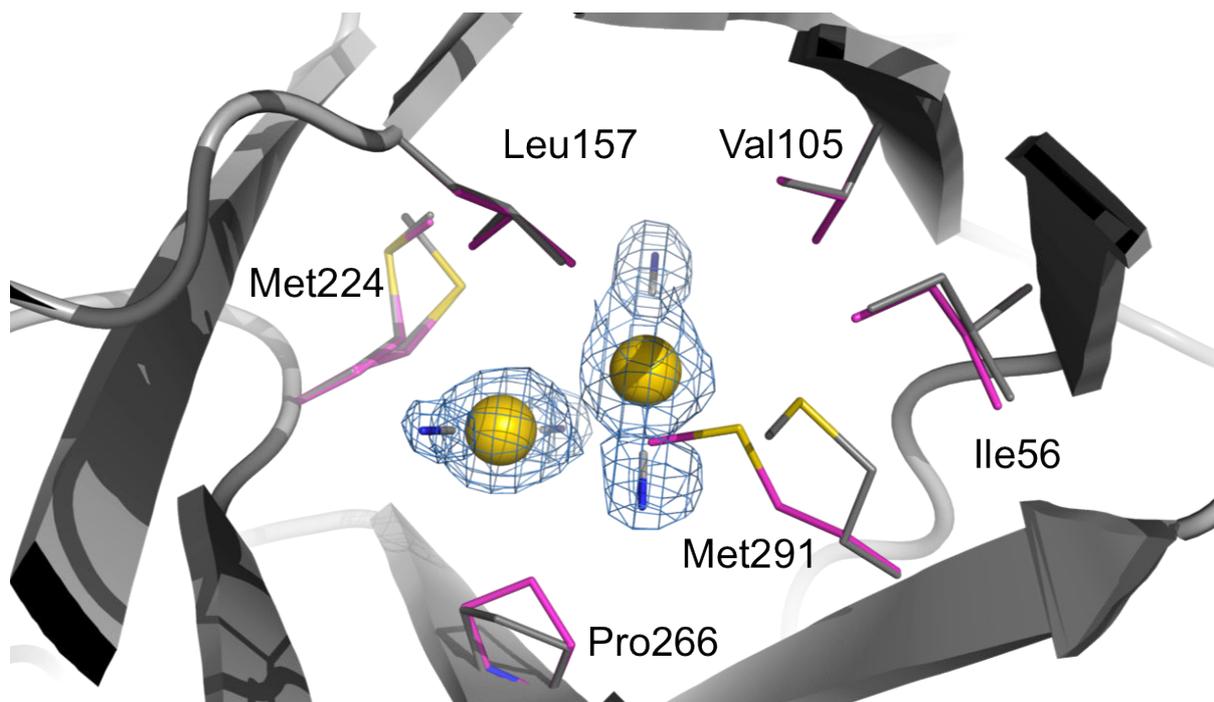
## I-8.4 Dérivatisation avec du dicyanoaurate

Suite à des échecs avec les sels de rubidium et de césium nous avons effectué quelques testes avec des cations multichargés lourds et avec des complexes de métaux de transition. L'un d'eux, le dicyanoaurate de potassium, nous a permis d'obtenir deux structures anecdotiques avec deux molécules de dicyanoaurate logées au centre de la couronne hydrophobe dans la cavité de HydE [Figure-I-26 A&B]. Chaque ion dicyanoaurate possède une topologie linéaire avec les deux cyanures de part et d'autre de l'atome d'or. La distance entre les deux atomes d'or est de 3.06 Å, ce qui correspondrait à une interaction de type Van der Waals. Ces atomes d'or sont encastrés dans un environnement hydrophobe, tandis que les cyanures pointent vers les parties de la cavité occupées par des résidus polaires ou par du solvant. Un des ions cyanure occupe le site de fixation d'anion S3 et un second ion se situe proche du site S2 [Figure-I-26-C]. En raison de la présence des ions dicyanoaurates dans la cavité de HydE on observe un déplacement des chaînes latérales des résidus Ile56, Leu157, Pro266 et Met291. Ceci se traduit par une légère augmentation du diamètre de la couronne hydrophobe [Figure-I-27].



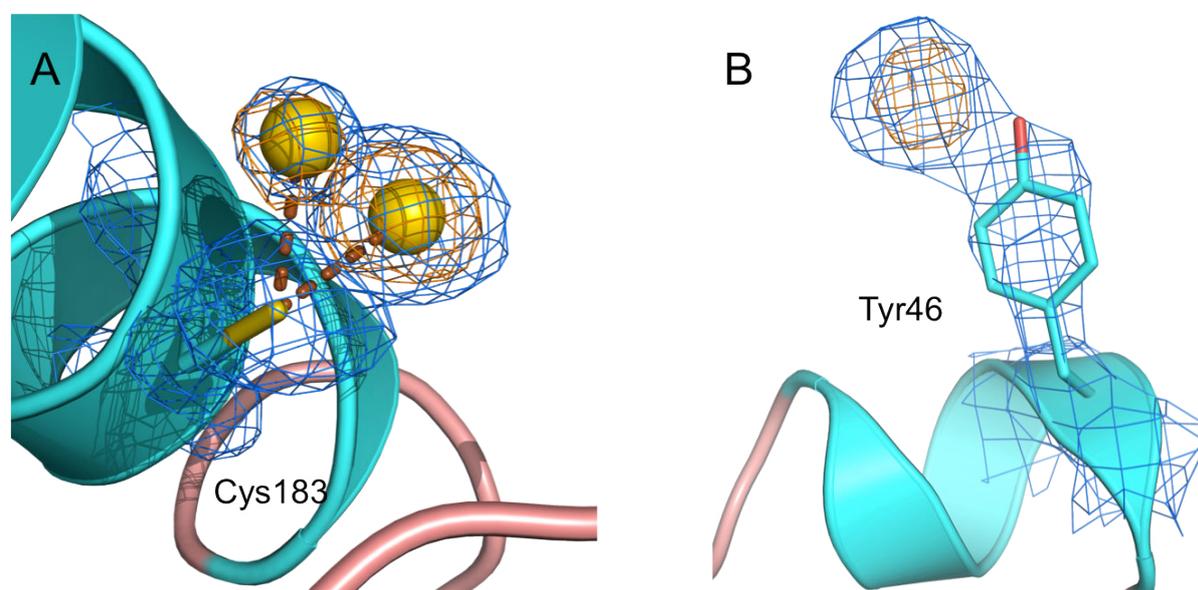
**Figure-I-26** : Structure cristallographique de *TmHydE* dérivatisée avec des ions dicyanoaurates

**A** - Structure générale du complexe de *TmHydE* avec deux ions dicyanoaurate. **B** - Coupe du tonneau ( $\beta\alpha$ )<sub>8</sub> dans le plan de la ceinture de résidus hydrophobes dans HydE. **C** - Vue stéréoscopique de deux ions dicyanoaurates encastrés dans la ceinture de résidus hydrophobes (en gris foncé). Les ions chlorures, tels qu'observés dans la structure native, sont représentés par des sphères vertes. En bleu, la carte de densité électronique  $2F_o-F_c$  contournée à  $1\sigma$  et visualisée à  $2\text{ \AA}$  autour des ions dicyanoaurates (résolution  $1,46\text{ \AA}$ ).



**Figure-I-27** : Variation du diamètre de la ceinture de résidus hydrophobes suite à la fixation de deux ions dicyanoaurates

Orientation des chaînes latérales des résidus hydrophobes dans la structure native (en magenta) et dans la structure du complexe (en gris foncé) avec deux ions dicyanoaurates encastrés dans la ceinture de résidus hydrophobes. En bleu, la carte de densité électronique  $2Fo-Fc$  contournée à  $1 \sigma$  (résolution 1,46 Å).



**Figure-I-28** : Dérivatisation de la protéine HydE avec des ions aurates

**A** - Deux ions aurates liés au thiol de la Cys183 et modélisés en positions alternatives avec un taux d'occupation de 0.4 et 0.15. En bleu, carte de densité électronique  $2Fo-Fc$  contournée à  $1 \sigma$  et en orange, carte anormale contournée à  $5 \sigma$  (résolution 1,46 Å). **B** - Ion aurate lié au phénol de la Tyr46 et modélisés avec un taux d'occupation de 0.3. En bleu, carte de densité électronique  $2Fo-Fc$  contournée à  $1 \sigma$  et en orange, carte anormale contournée à  $5 \sigma$  (résolution 2,59 Å).

Dans une de nos structures on peut voir deux ions d'or liés chacun à un groupement phénol d'une tyrosine [Figure-I-28-A]. Dans la seconde structure on trouve un ion d'or en positions alternatives sur la Cys183 [Figure-I-28-B]. De similaires interactions entre un ion aurate (I) ou aurate (III) et la chaîne latérale d'un acide aminé ont été répertoriées dans la pdb [codes pdb: 3P66, 3H7G, 4A65, 3K3F, 1W1D, 1A8D]. La faible résolution de nos données cristallographiques (résolution 2,59 Å) et/ou la basse occupation de ces ions dans nos structures ne permettent pas de trancher si les ions aurates en interaction avec la chaîne latérale d'une tyrosine ou cystéine ont gardé un de leurs ligands cyanure.

La disposition de la paire d'ions dicyanoaurates dans la structure de HydE est très proche de celle observée dans la structure de la  $\beta$ -lactamase de classe D de *Pseudomonas aeruginosa* après la dérivation de l'enzyme avec du dicyanoaurate de potassium<sup>40</sup> (pdb: 1E3U). Dans les deux structures, les deux ions dicyanoaurates sont logés dans un environnement hydrophobe. Puisque dans cette expérience l'or est sous la forme d'un ion dicyanoaurate très stable il est difficile d'utiliser ces données pour en déduire des informations physiologiquement pertinentes. Il est très probable que la paire d'ions dicyanoaurate se fixe dans HydE parce qu'au milieu de sa cavité on trouve une partie à un caractère hydrophobe, tandis que le bas et le haut de la cavité présentent un potentiel électrostatique positif.

### **I-8.5 Premier test de co-cristallisation avec de la Se-L-cystine**

Pour tester notre hypothèse sur la cystine comme substrat de HydE, nous avons, dans un premier temps, décidé de co-cristalliser la forme sauvage de la protéine avec une solution de L-cystine à 25mM de concentration finale. Dans la gamme de pH où HydE de *T. maritima* cristallise la solubilité de la cystine reste très faible et ne dépasse pas 0,7 mM<sup>41</sup>. Pour palier à ce problème, nous avons d'abord préparé une solution concentrée de cystine en la solubilisant avec une titration à la soude. Ensuite, cette solution a été mélangée à l'échantillon protéique tamponné. Cette procédure, ainsi que le protocole de préparation des cristaux, sont détaillés dans le paragraphe MM-4.1 et dans l'article Rohac *et al.* 2016 en Annexe.

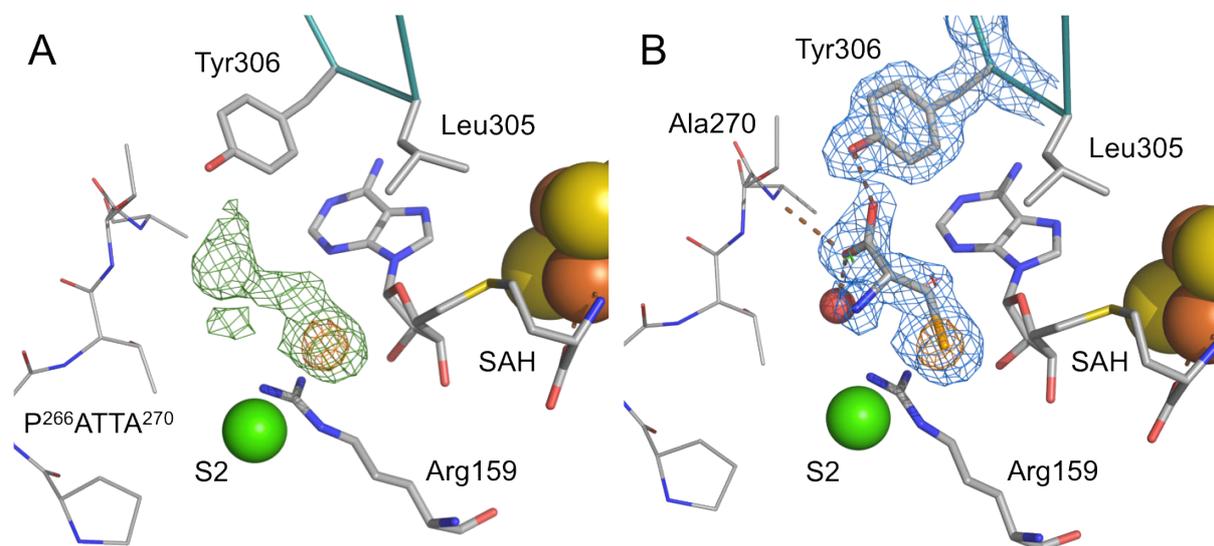
Pour le premier test de cocrystallisation nous avons préféré utiliser un analogue sélénié de la cystine – la séléno-L-cystine. Le sélénium possède 34 électrons, contrairement au soufre qui n'en possède que 16. Ainsi la séléno-L-cystine serait plus facilement identifiable dans les cartes de densité électronique que son analogue soufré. D'autant plus que le signal de dispersion anormale spécifique du sélénium au seuil d'absorption K permettrait l'identification de cet atome même si le taux d'occupation du ligand dans la cavité de HydE est faible, à condition que le facteur d'agitation thermique de l'atome en question ne soit pas très élevé.

#### **I-8.5.1 Identification d'un ligand à un atome de sélénium**

Les principes de l'enregistrement et du traitement de données cristallographiques, ainsi que de la résolution de structure par remplacement moléculaire sont mentionnés dans les paragraphes [paragraphe MM-4.1], respectivement. La résolution du jeu de données cristallographiques collecté atteignait 1,83 Å. L'inspection de la carte de densité électronique

de Fourier différence *F<sub>o</sub>-F<sub>c</sub>* a montré la présence d'une molécule de petite taille dans le site actif de l'enzyme [Figure-I-29-A]. Un seul pic fort dans la carte de différence anormale au niveau du site actif a confirmé que le nouveau ligand contenait un seul atome de sélénium. Cette molécule n'était donc de toute évidence pas la séléno-L-cystéine. L'intensité de ce pic atteignait 9,5  $\sigma$  vs. 10,0  $\sigma$  pour les atomes de Fe du centre [FeS] SAM radical. Puisque les données ont été enregistrées au seuil du sélénium, ceci suggérait que l'occupation du ligand dans la cavité de HydE était inférieure à 100%.

En examinant les détails de la densité électronique du ligand nous avons constaté la présence d'un groupement carboxylate dans le site S1. Il y avait aussi une connexion très claire entre ce groupement carboxylate et l'atome de sélénium. Cette densité ressemblait fortement à une molécule de séléno-L-cystéine (la forme réduite de la séléno-L-cystéine) [Figure-I-29-B]. Pour comprendre comment cette molécule avait pu se former nous nous sommes penchés sur la composition de l'échantillon protéique utilisé pour cette première expérience de co-cristallisation.

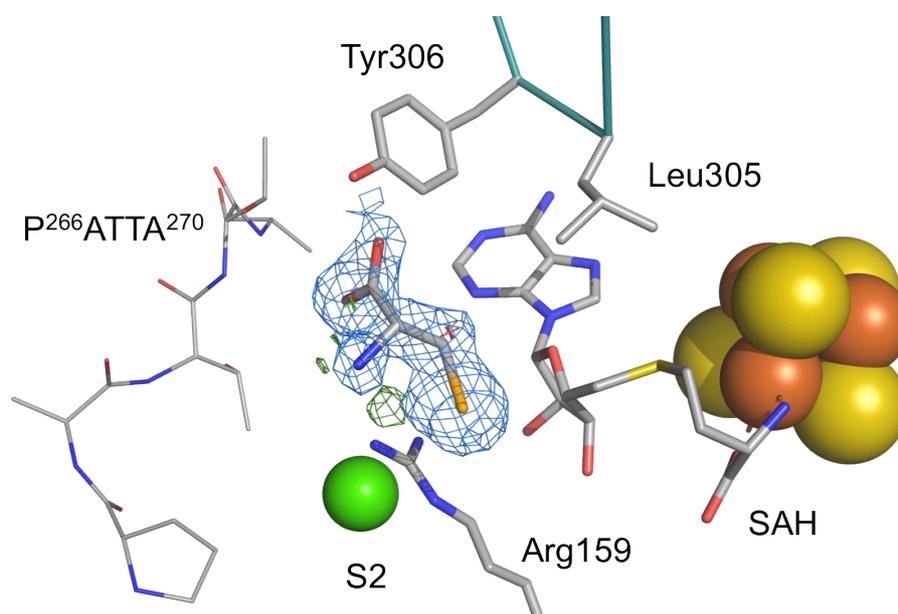


**Figure-I-29** : Composition du site actif de HydE suite a la co-cristallisation avec la séléno-L-cystéine

**A** - En vert foncé, carte de densité électronique *F<sub>o</sub>-F<sub>c</sub>* contournée à 3  $\sigma$  et en orange, carte anormale contournée à 5  $\sigma$  (résolution 1,83 Å). Le chlorure dans le site S2 est représenté par une sphère vert clair. **B** - Une molécule de cystéine modélisée dans le site actif de HydE. En bleu, carte de densité électronique *2F<sub>o</sub>-F<sub>c</sub>* contournée à 1  $\sigma$ . Le chlorure dans le site S1 est représenté par une croix verte et la molécule d'eau proche du C $\beta$  de la cystéine libre est représentée par une croix rouge.

En fait, cet échantillon a été préparé avant mon arrivée au laboratoire et il contenait la forme native de la protéine de *TmHydE*, possédant les deux centres [FeS] reconstitués chimiquement à la fin de la purification. La solution présentait un aspect très foncé, parce qu'elle n'avait pas été débarrassée de l'excès de Fe et de DTT et donc elle contenait des agrégats [FeS] en solution. Nous avons donc initialement pensé que la séléno-L-cystéine s'était réduite en présence de DTT pour former la séléno-L-cystéine et que c'était cette dernière que nous avons observée dans la cavité de HydE.

La modélisation du ligand comme une molécule de sélénio-L-cystéine avec un taux d'occupation de 60% expliquait raisonnablement bien la nouvelle densité électronique dans le site actif de HydE. Cependant nous avons pu remarquer quelques densités résiduelles en proximité du nouveau ligand. La présence d'une densité résiduelle à côté de l'atome C $\beta$  a été expliquée comme provenant d'une molécules d'eau avec un taux d'occupation de 40%. Cette molécule d'eau est présente dans toutes les structures dites « natives », c'est-à-dire contenant le cofacteur SAM ou SAH mais pas de ligand (pdb codes : 3CIW, 3CIX, 3IIZ). Une autre densité résiduelle, plus faible, se situait entre l'atome de Se et l'amine de la sélénocystéine et se propageait vers le site S2 où nous avons modélisé un ion chlorure [Figure-I-30]. Cette densité n'a pas pu être toute de suite modélisée, sans doute à cause de la résolution limitée de données cristallographiques et la faible occupation du ligand [voir paragraphe I-8.7.1].



**Figure-I-30** : Sélénio-L-cystéine modélisée dans le site actif de HydE avec un taux d'occupation de 0.6

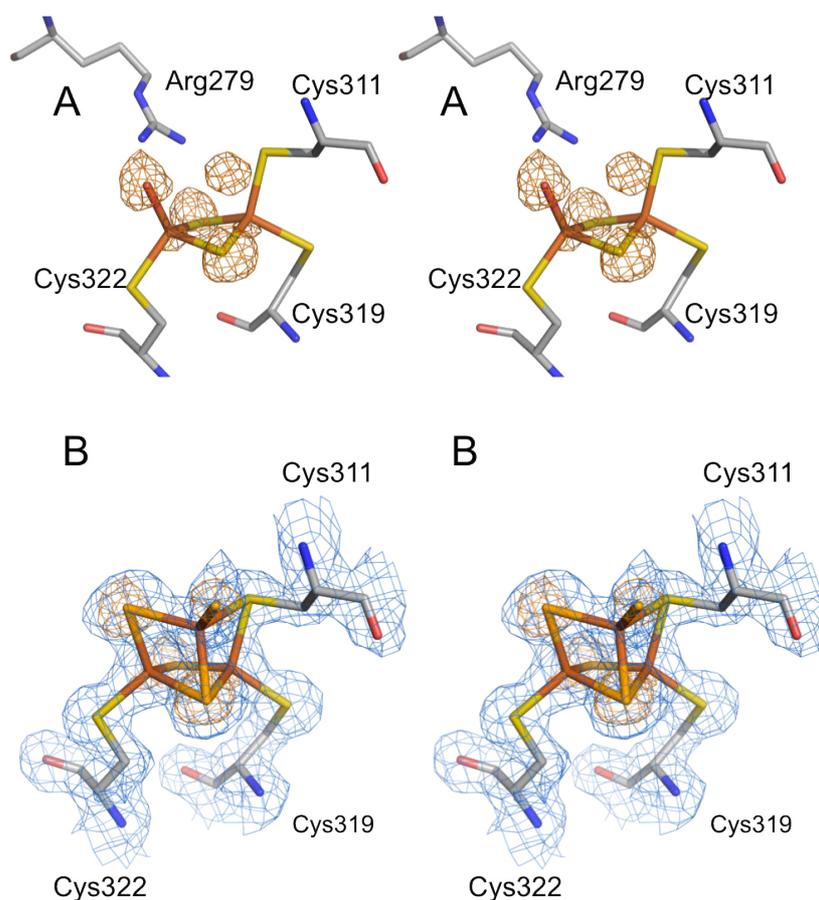
En bleu, carte de densité électronique  $2Fo-Fc$  contournée à  $1 \sigma$  et en vert foncé, carte de densité électronique  $Fo-Fc$  contournée à  $3 \sigma$  (résolution 1,83 Å). Le chlorure dans le site S2 est représenté par une sphère vert clair et celui dans le site S1 par une croix verte. La molécule d'eau qui se trouve près du C $\beta$  de la sélénio-L-cystéine est modélisée avec un taux d'occupation de 0.4 et elle est représentée par une croix rouge.

La fixation d'une sélénocystéine dans le site actif de HydE était quelque peu surprenante, car de nombreuses tentatives de co-cristallisation de l'enzyme avec la cystéine n'ont pas permis l'obtention d'un tel complexe. Cette observation sera discutée plus en détails dans [I-11.4.2 & III-2.1.2].

### I-8.5.2 Partielle reconstitution de l'agrégat [Fe-S] auxiliaire avec du sélénium

La présence du Fe libre et de l'excès de DTT a eu un impact sur la forme de l'agrégat auxiliaire. La carte de Fourier différence montre sept pics forts dans la zone où se situe habituellement le second centre [Fe-S], suggérant qu'il y a sept atomes lourds à cet endroit. La carte de différence anormale a montré quatre pics très forts à des positions où on s'attendrait de trouver des atomes de soufre [Figure-I-31-A]. Puisque ce jeu de données cristallographiques a été enregistré à la longueur d'onde du seuil K du sélénium, quatre de ces atomes lourds sont en fait des atomes de sélénium. Ceci constitue une preuve de la libération de sélénium à partir de la sélénocystéine et de son incorporation dans l'agrégat auxiliaire de *TmHydE*. La libération de sulfures par un processus de  $\beta$ -élimination à partir de la cystéine en présence de Fe a déjà été observée auparavant<sup>42,43</sup>. Les auteurs ont remarqué que la présence de traces d'ions de Fe était indispensable pour que cette dégradation de la cystéine ait lieu.

En se servant du signal anomal du sélénium, l'agrégat auxiliaire a pu être modélisé comme un centre [Fe<sub>3</sub>Se<sub>3</sub>-Se] [Figure-I-31-B]. Pour plus d'informations sur l'interconversion entre les différentes formes d'agrégats [Fe-S] voir *Chapitre IV*.



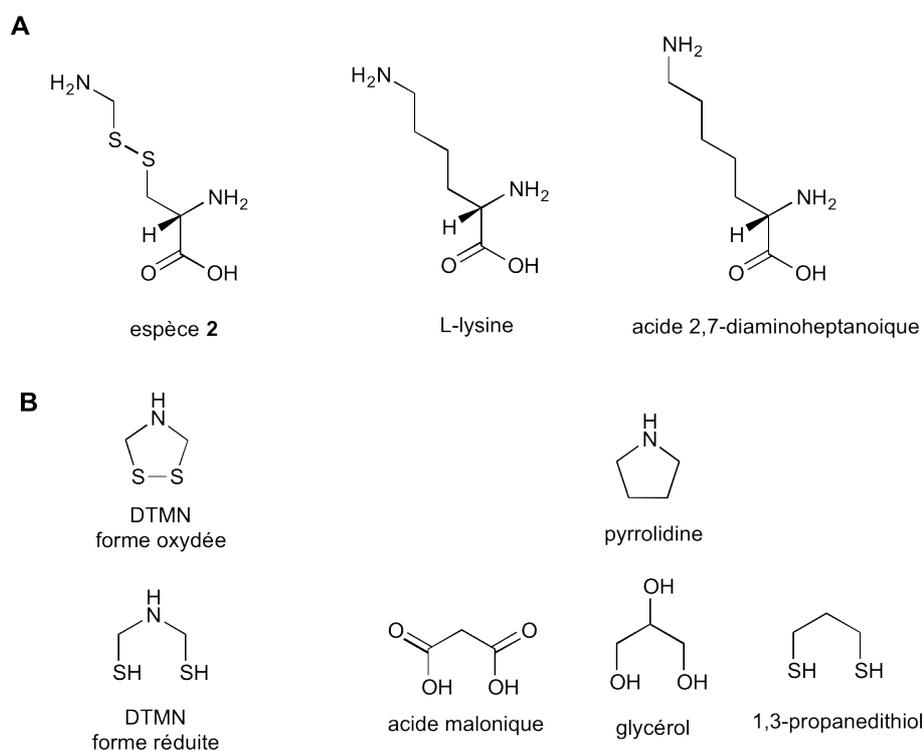
**Figure-I-31** : Vue stéréoscopique de l'environnement du site de fixation du centre [Fe-S] auxiliaire après l'incubation de la protéine avec une solution contenant de la séléno-L-cystine en milieu réducteur

**A** - Superposition de la structure contrôle du second centre [Fe-S] sous sa forme Fe<sub>2</sub>S<sub>2</sub> (pdb: 3CIW) avec la carte de densité électronique anormale contourée à 5  $\sigma$  issue de l'incubation de la protéine avec une solution contenant de la séléno-L-cystine. **B** - En bleu, carte de densité électronique 2Fo-Fc contourée à 1  $\sigma$  et en orange, carte de densité électronique anormale contourée à 5  $\sigma$  (résolution 1,83 Å). Dans ces cartes est modélisé l'agrégat [Fe<sub>3</sub>Se<sub>3</sub>-Se], reconstitué avec du sélénium issu de la dégradation de la sélénocystéine.

## I-8.6 Essais de co-cristallisation avec des analogues des espèces intermédiaires hypothétiques

Le mécanisme que nous proposons pour la conversion de la cystine en azadithiolate par HydE passe par l'espèce **2**. Le produit hypothétique DTMN est quant à lui facilement hydrolysable. Dans le but de tester la fixation de ces espèces dans le site actif de HydE nous avons utilisé leurs analogues stables et commercialement accessibles [Figure-I-32-A&B]. L'espèce **2** a une structure proche de la L-lysine. Si on tient compte de la distance, qui sépare les deux soufres engagés dans le pont disulfure, le plus proche analogue de l'espèce **2** pourrait être le 2,7-diaminoheptanoate [Figure-I-32-A].

Le produit final, avant sa réduction en azadithiolate ressemble à l'hétérocycle de la pyrrolidine. Après sa réduction, les analogues mimant plus ou moins bien la topologie de l'azadithiolate sont le malonate, le glycérol et le 1,3-propanedithiol [Figure-I-32-B]. Ce dernier avait même été initialement proposé comme le vrai constituant du site actif de l'hydrogénase à [FeFe]<sup>44</sup> [voir IB-1.3].



**Figure-I-32** : Schéma d'espèces théoriques et de leur analogues commercialement accessibles

**A** - A gauche, l'intermédiaire de la réaction proposée, "espèce 2" et à droite ses analogues stables et commercialement accessibles. **B** - A gauche, formes oxydée et réduite du produit hypothétique de la réaction proposée et à droite leurs analogues stables et commercialement accessibles.

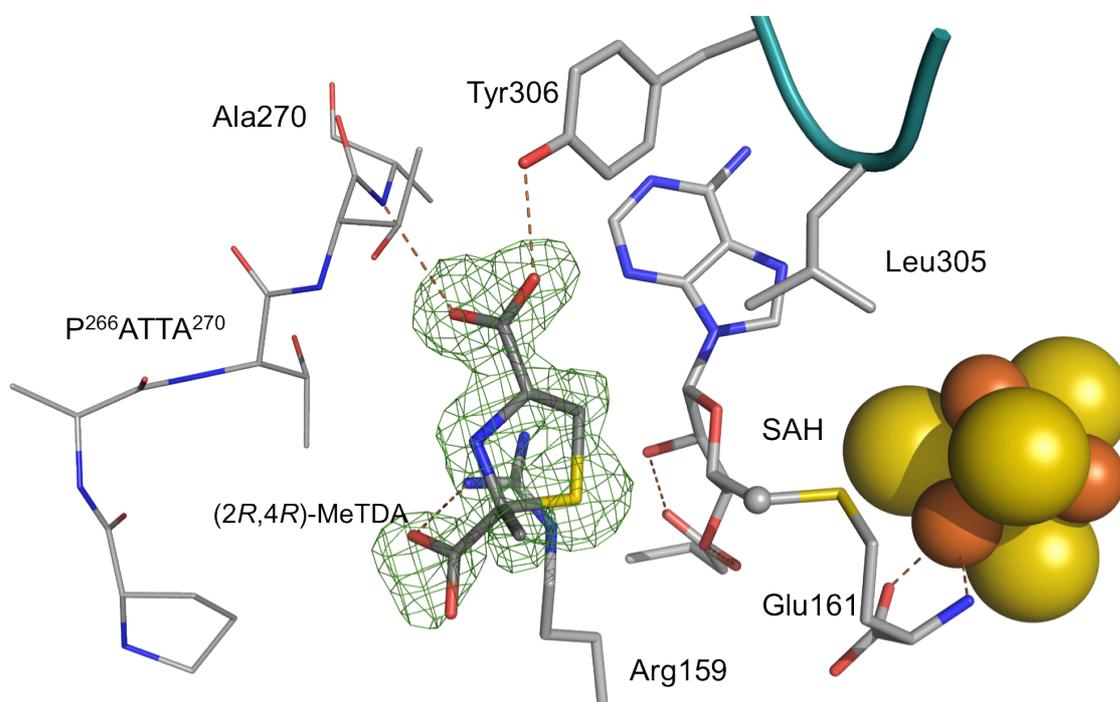
Malheureusement, aucun de nos essais de co-cristallisations de HydE avec ces molécules n'a abouti à la formation d'un complexe. Ainsi notre mécanisme se retrouvait en difficulté. Il fallait donc impérativement révéifier la fixation de la cystine dans le site actif de HydE, cette fois-ci en utilisant un échantillon dépourvu de DTT, afin d'éviter la réduction de la cystine en cystéine.

## I-8.7 Seconde co-cristallisation de TmHydE sauvage avec de la L-cystine

Nous avons alors préparé un nouvel échantillon de *TmHydE* sauvage, qui a été débarrassé de l'excès de Fe et de DTT [voir MM-3.2.4]. La procédure de préparation d'une solution concentrée de L-cystine, ainsi que la préparation des cristaux sont identiques à celles utilisées lors de l'expérience précédente.

### I-8.7.1 Dérivé de L-cystéine comme ligand de HydE

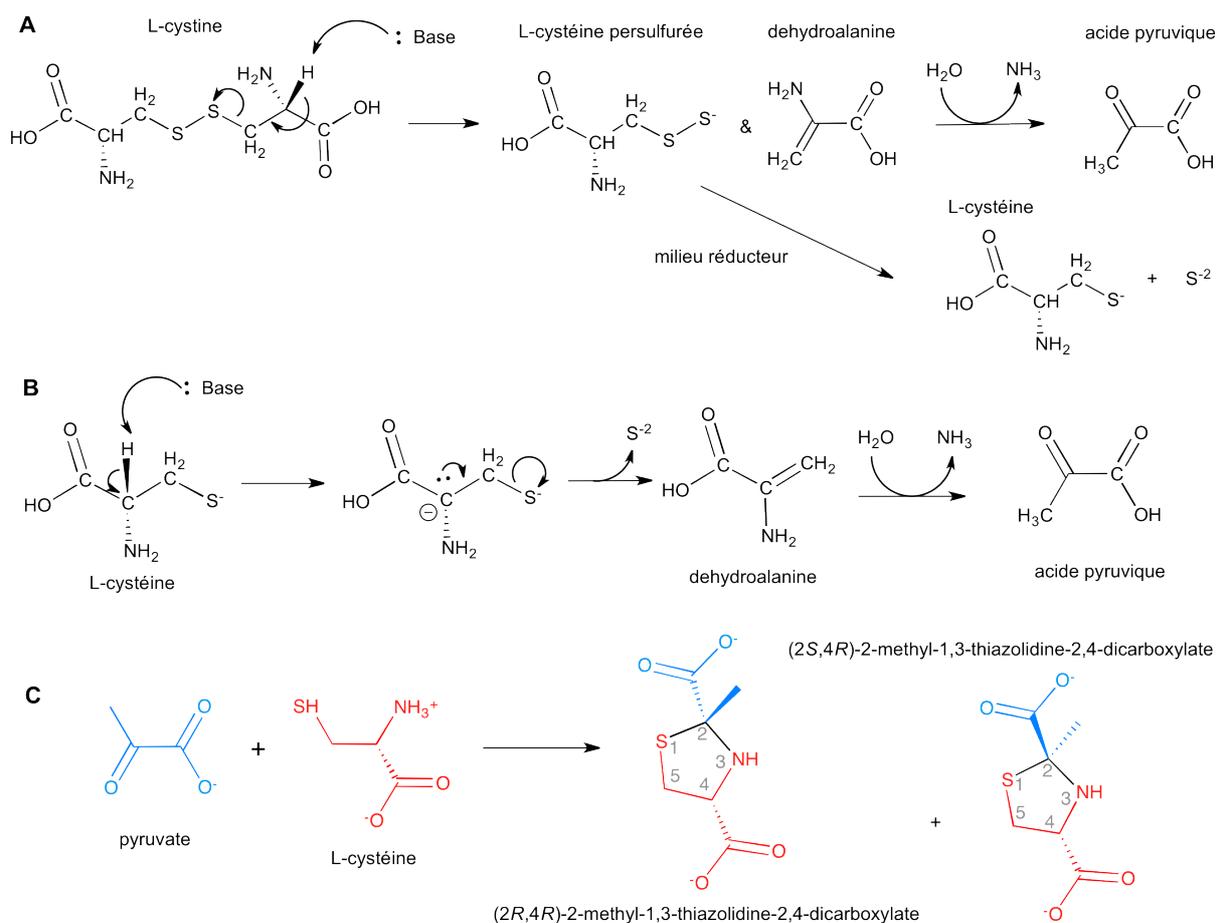
La résolution du nouveau jeu de données cristallographiques obtenu à partir de la protéine sauvage était de 1,55 Å. Cette résolution relativement élevée et l'occupation du ligand qui approchait 80 %, ont eu pour conséquence que la densité électronique dans la carte de Fourier différence adoptait clairement la forme d'une molécule cyclique avec trois extensions. Deux de ces extensions avaient la forme de deux groupements carboxylates, chacun logé dans un des sites S1 et S2. Ce ligand a pu être modélisé comme le (2*R*,4*R*)-2-méthyle-1,3-thiazolidine-2,4-dicarboxylate [Figure-I-33], qu'on appellera par convenance (2*R*,4*R*)-MeTDA (« Me » désigne la présence du méthyle, « T » est pour le cycle thiazolidine et le suffixe « DA » pour dicarboxylic acid en anglais - désignant la présence de deux groupements carboxylates).



**Figure-I-33** : Structure cristallographique du complexe de *TmHydE* avec le dérivé thiazolidine (2*R*,4*R*)-MeTDA

En vert foncé, carte de densité électronique *Fo-Fc* "omit" contourée à 3  $\sigma$ , calculée à partir du modèle final contenant le ligand (2*R*,4*R*)-MeTDA, modélisé avec un tau d'occupation de 0 et représenté en gris foncé au centre de l'image (résolution 1,55 Å).

Nous avons cherché à comprendre d'où venait cette molécule dans le site actif de HydE. Puisque la protéine a été co-cristallisée en présence de la *S*-adényl-L-cystéine (SAH), la forme inactive du cofacteur SAM<sup>45</sup>, et en absence de réducteurs forts, il semblait peu plausible que la protéine ait été impliquée dans la transformation de la L-Cystine en (2*R*,4*R*)-MeTDA selon un mécanisme radicalaire. Par contre, il est connu depuis les années 1930 que la cystine subit, en milieu alcalin et en présence de métaux de transition, une dégradation par  $\beta$ -élimination<sup>46,47</sup>. Elle se décompose en L-cystéine, pyruvate et sulfure d'hydrogène [Figure-I-34-A]. La cystéine peut elle aussi subir une dégradation dans les mêmes conditions pour libérer le sulfure d'hydrogène et le pyruvate<sup>42</sup> [Figure-I-34-B].

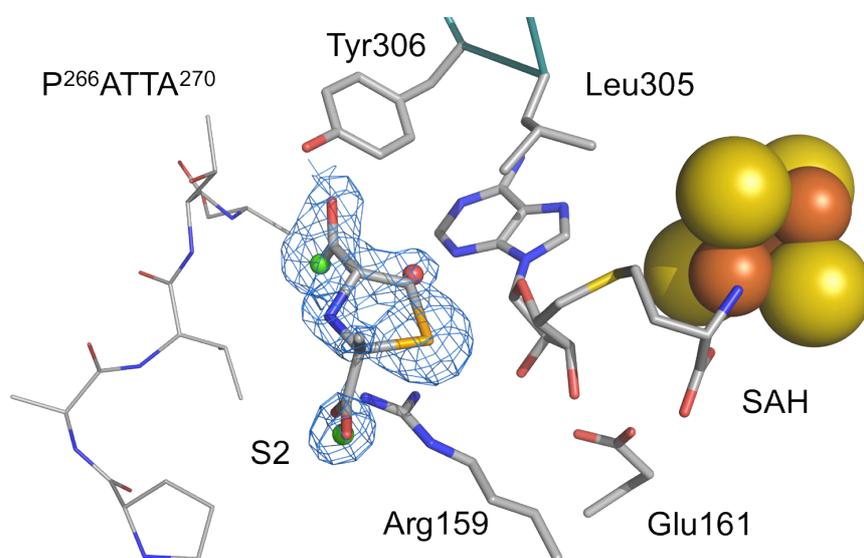


**Figure-I-34**

**A** - Schéma de la dégradation de la L-cystine par  $\beta$ -élimination en milieu alcalin. **B** - Schéma de la dégradation de la L-cystéine par  $\beta$ -élimination. **C** - Schéma de la cyclocondensation de la L-cystéine avec du pyruvate. En gris, la numérotation des atomes du cycle thiazolidine.

Il a aussi été démontré que la cystéine libre réagit avec les carbonyles des aldéhydes ou des cétones ce qui conduit à la formation de dérivés thiazolidines correspondants<sup>48,49</sup>. Il est fort probable que notre solution de co-cristallisation contenait des traces d'ions ferreux ou ferriques, provenant de la dégradation partielle des agrégats [FeS] de HydE de *T. maritima*. Nous en avons donc conclu, que la L-cystéine s'était partiellement décomposée en pyruvate, L-cystéine et sulfure, ce qui a conduit à la formation du dérivé (2*R*,4*R*)-MeTDA et probablement aussi (2*S*,4*R*)-MeTDA [Figure-I-34-C] [I-11.2 & I-11.3].

Dans le cas de la sélénocystéine, la libération de séléniure expliquait la présence de sélénium dans le centre [FeS] auxiliaire de HydE lors du premier essai de co-cristallisation [I-8.5.2]. D'ailleurs, une réinterprétation de données cristallographiques, issues de cette première expérience, a permis de modéliser le dérivé thiazolidine sélénié (2*R*,4*R*)-SeMeTDA (le préfixe « Se » témoigne de la présence du sélénium dans le ligand) dans le site actif avec un taux d'occupation de 60% [Figure-I-35]. Contrairement à l'ancien modèle avec la sélénocystéine, ce nouveau modèle explique parfaitement la densité électronique résiduelle entre le Se et l'amine [Figure-I-30][I-8.5.1]. Nous pouvons donc conclure, que notre première structure avec le ligand à un atome de sélénium contenait le dérivé thiazolidine (2*R*,4*R*)-SeMeTDA. Ce dérivé provenait de la condensation de la sélénocystéine avec du pyruvate, le dernier issu de la dégradation de la sélénocystéine en présence de Fe.



**Figure-I-35** : Réinterprétation de données cristallographiques issues de la première expérience de co-cristallisation de la protéine *TmHydE* avec de la séléno-L-cystéine

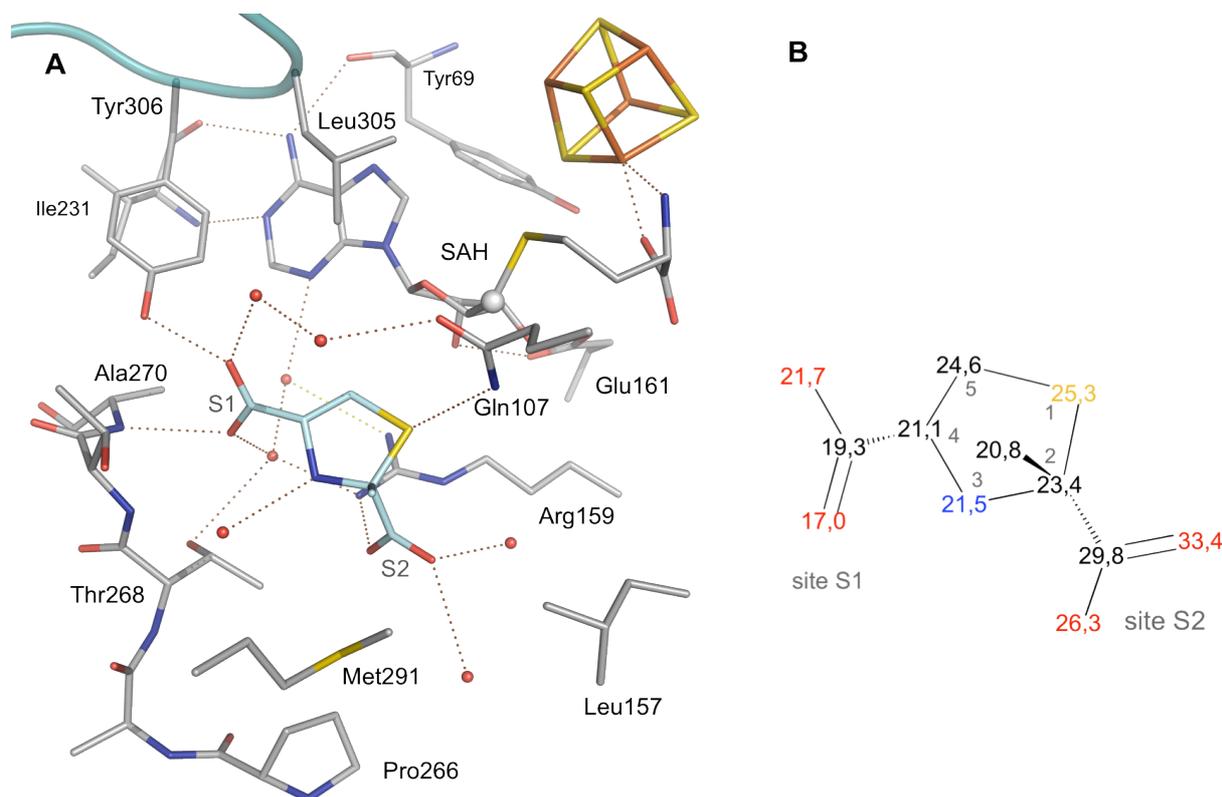
Le ligand (2*R*,4*R*)-SeMeTDA est modélisé dans le site actif de HydE avec un taux d'occupation de 0.6. En bleu, carte de densité électronique  $2F_o-F_c$  contourée à  $1 \sigma$  (résolution 1,83 Å). Les ions chlorures dans les sites S1 et S2 sont modélisés avec un taux d'occupation de 0.4 et ils sont représentés par des sphères vertes. La molécule d'eau qui se trouve près du C5 du cycle thiazolidine est modélisée avec un taux d'occupation de 0.4 et elle est représentée par une sphère rouge. Il est à noter que la nouvelle carte de densité électronique a été obtenue en utilisant l'affinement "jelly body" dans refmac5, ce qui a sensiblement amélioré la visibilité des détails dans le site actif.

### **I-8.7.2 Interactions entre le ligand (2*R*,4*R*)-MeTDA et le site actif de HydE**

Le ligand (2*R*,4*R*)-MeTDA nouvellement identifié se loge dans la partie supérieure de la cavité dans HydE entre les résidus Tyr306 et Arg159. Cette zone correspond à l'endroit où le substrat physiologique devrait se fixer [I-5].

Les deux groupements carboxylate du ligand occupent les sites anioniques S1 et S2 précédemment identifiés. Le groupement carboxylate, qui provient de la partie cystéine, se fixe dans le site S1. Il crée des liaisons hydrogènes avec l'atome d'oxygène du phénol de la Tyr306 et avec l'atome d'azote de l'Ala270, cette dernière appartenant à la séquence PATTA

hautement conservée. Le second groupement carboxylate, issu du pyruvate, se fixe dans le site S2 où il interagit avec le groupement guanidinium de l'Arg159 [Figure-I-36-A]. Les atomes de ce carboxylate présentent des facteurs d'agitation thermique plus élevés que ceux du carboxylate dans le site S1 [Figure-I-36-B]. Ceci est dû à l'absence d'une seconde interaction entre le carboxylate dans S2 et un résidu de la protéine et confirme la différence d'affinité entre les deux sites anioniques pour le groupement carboxylate, que nous avons proposé [I-7].

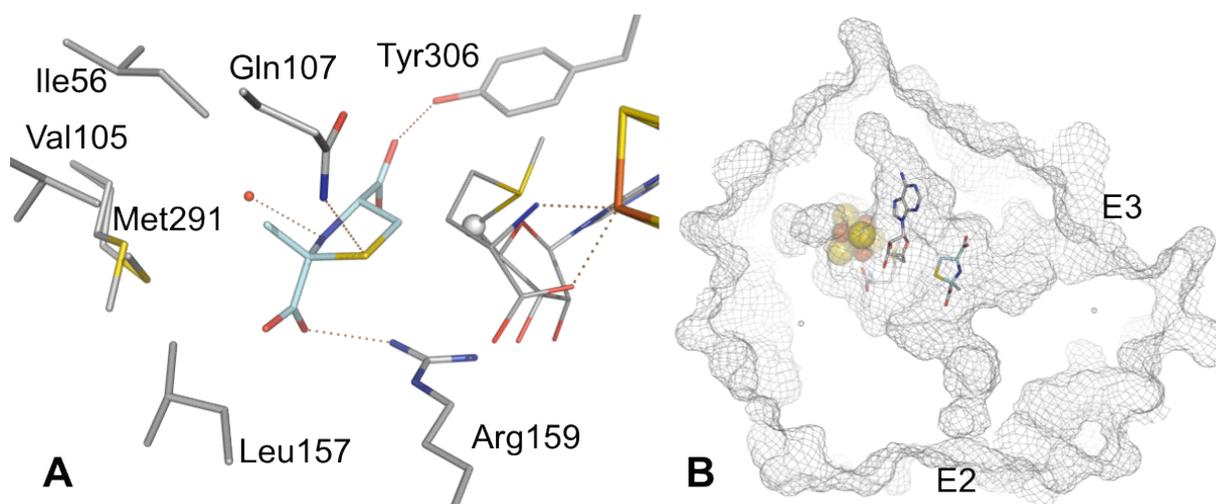


**Figure-I-36** : Détails d'interaction entre le ligand (2R,4R)-MeTDA et la protéine *TmHydE*

**A** - Vue de la cavité de *TmHydE* (en gris clair) avec le ligand (2R,4R)-MeTDA (en cyan) fixé à proximité de la SAH. La première et la seconde couche de solvation du ligand sont représentées par des sphères rouges. **B** - Schéma du ligand (2R,4R)-MeTDA avec les facteurs d'agitation thermique notés pour chaque atome (en Å<sup>2</sup>). Le facteur d'agitation thermique moyen du ligand est de 27,9 Å<sup>2</sup>. Celui de la SAH par exemple est de 18,9 Å<sup>2</sup>. Les atomes du cycle 1,3-thiazolidine sont numérotés en gris.

Le méthyl du dérivé (2R,4R)-MeTDA, provenant de la partie pyruvate, pointe vers une zone à caractère hydrophobe, composée des résidus Met291, Ile56 et plus loin Leu157 et Val105. L'atome de soufre du thioéther est pris en sandwich entre les résidus strictement conservés Gln107 et Arg159. L'amine du cycle thiazolidine forme une liaison hydrogène avec une molécule d'eau [Figure-I-37-A].

On peut noter, que le ligand n'occupe qu'une fraction du volume disponible dans le site actif de la maturase HydE et que de nombreuses molécules d'eau remplissent le reste de cet espace [Figure-I-37-B].



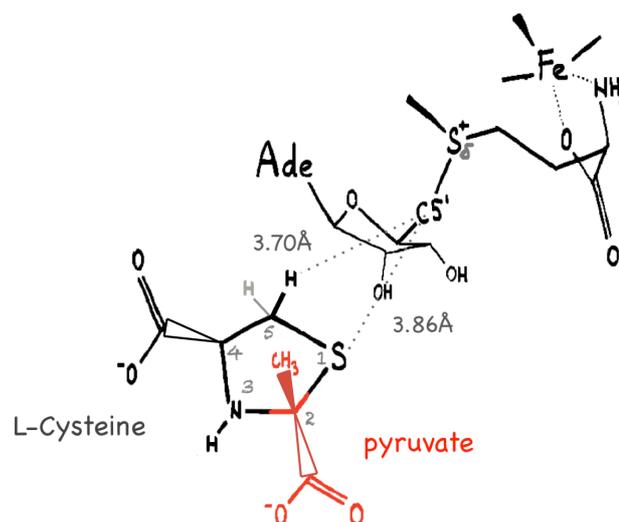
**Figure-I-37** : Détails d'interaction entre le ligand (2*R*,4*R*)-MeTDA et la protéine *TmHydE*

**A** - Disposition du ligand (2*R*,4*R*)-MeTDA (en cyan) par rapport à la ceinture de résidus hydrophobes. **B** - Visualisation de l'enveloppe de la protéine *TmHydE*, ainsi que du volume que le ligand occupe dans la cavité interne de l'enzyme. E2 et E3 sont deux pores de sortie visibles sous ce point de vue. L'image a été créée avec le logiciel PyMol en utilisant une sonde avec un diamètre de 0,9 Å.

### I-8.7.3 Disposition du ligand (2*R*,4*R*)-MeTDA par rapport à la SAM

Les atomes S1 et C5 du cycle thiazolidine se trouvent dans la zone où l'attaque radicalaire, portée par l'atome C5' du radical 5'-déoxyadénosyl, devrait avoir lieu [I-5]. Le C5' se situe à 3,86 Å et à 4,25 Å des atomes S1 et C5, respectivement [Figure-I-38]. La distance entre l'atome C5' et un des atomes d'hydrogène du C5, qui semble être bien placé pour être arraché, est de 3,7 Å.

Les conséquences que cette orientation du ligand vis-à-vis de l'atome C5' ont sur l'activité de l'enzyme seront commentées plus en détails dans le *Chapitre III*. Ce dernier est entièrement consacré à l'étude de la réaction radicalaire avec le ligand (2*R*,4*R*)-MeTDA observée en solution et dans les cristaux de HydE.



**Figure-I-38** : Disposition du ligand (2*R*,4*R*)-MeTDA par rapport à la SAM dans *TmHydE*

## I-8.8 Essais de co-cristallisation de *TmHydE* $\Delta$ FeS N-strep avec la L-cystine

### I-8.8.1 Préparation de l'échantillon *TmHydE* $\Delta$ FeS N-strep

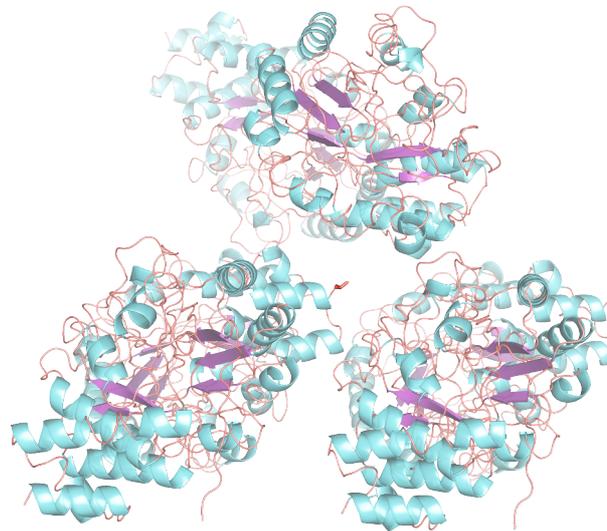
Afin de pouvoir tester la fixation de la cystine il fallait éviter sa dégradation lors de la préparation de cristaux. Pour cela on devait s'assurer que l'échantillon protéique n'allait pas libérer de Fe dans la solution. Au cours de nombreuses expériences de cristallisation avec la protéine *TmHydE* nous avons remarqué que les deux centres [Fe-S] de la protéine sauvage n'ont pas la même stabilité<sup>1</sup> [**Chapitre IV**]. En effet, le centre [Fe<sub>4</sub>S<sub>4</sub>] SAM radical lié soit au cofacteur SAM, soit à l'analogue de ce dernier - la SAH, est stable dans le temps. D'un autre côté, le centre [Fe-S] auxiliaire se dégrade en libérant du Fe dans la solution [**Chapitre IV**]. Le dosage en Fe et en S de la protéine sauvage chimiquement reconstituée montre 8 atomes de Fe et 8 atomes de soufre par monomère. Ceci laisse penser qu'à la fin de la reconstitution la protéine en solution possède deux centres [Fe<sub>4</sub>S<sub>4</sub>] [voir IB-3.2]. Le repliement de la forme avec les deux centres [Fe<sub>4</sub>S<sub>4</sub>] est resté pendant longtemps inconnu. Dans la plupart de nos structures nous retrouvons le centre [Fe-S] SAM radical sous la forme d'un [Fe<sub>4</sub>S<sub>4</sub>], tandis que l'agrégat auxiliaire est le plus souvent sous la forme d'un centre [Fe<sub>2</sub>S<sub>2</sub>] partiellement occupé (pdb : 3CIX) [**Figure-IB-12**]. Nous en avons conclu que ce second centre [Fe-S] se dégradait au cours du temps en libérant du Fe et du S dans la solution. De plus, la dégradation partielle du centre [Fe-S] auxiliaire était une source d'hétérogénéité structurale [I-8.8.2]. Enfin, comme il a été montré que ce centre [Fe-S] auxiliaire ne joue pas de rôle clé dans le processus de maturation de l'hydrogénase<sup>14</sup> [voir les paragraphes I-3.1 & II-3.2], nous avons décidé de le supprimer pour faciliter nos études structurales.

Nous avons donc préparé un triple mutant C311S, C319S et C322S, possédant une étiquette strep-tag N-terminale. Dans ce second échantillon, appelé ici  $\Delta$ FeS N-strep, les trois cystéines responsables de la fixation de l'agrégat auxiliaire de *TmHydE* ont été donc mutées en serines [voir paragraphe MM-2.3]. La suppression de ce second centre [FeS] augmente l'homogénéité de l'échantillon et améliore la résolution de données cristallographiques [I-8.8.2]. L'échantillon a été débarrassé de l'excès de Fe et de DTT [voir paragraphe MM-3.2.4]. Les procédures de préparation d'une solution stock de L-cystine et de préparation de cristaux sont identiques à celles utilisées lors des expériences précédentes.

### I-8.8.2 Structure cristallographique du mutant *TmHydE* $\Delta$ FeS N-strep

Dans ces conditions, la protéine *TmHydE* Nstrep  $\Delta$ FeS cristallise dans le même groupe d'espace ( $P2_12_12_1$ ) que la protéine sauvage, dépourvue de l'étiquette strep et de la triple mutation C311S-C319S-C322S. La structure globale reste la même ; cela dit, deux différences mineures sont cependant visibles : *i.*) la disparition de la densité correspondante à la Met1 (numérotation de la séquence de la protéine sauvage) et *ii.*) un léger réarrangement de la boucle 311-319.

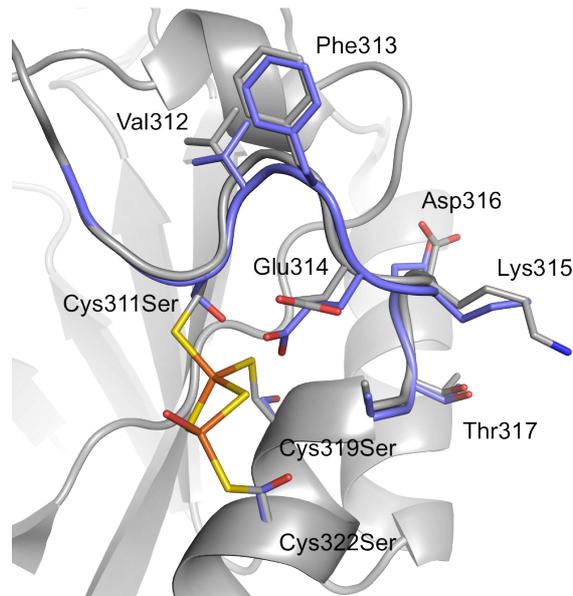
La disparition du résidu Met1 est due à l'insertion d'une étiquette strep entre ce résidu et la Thr2. Les 10 résidus de l'étiquette ne sont pas visibles - l'étiquette strep N-terminale doit donc être désordonnée et évoluer librement dans le canal de solvant qui entoure la partie N-terminale de la protéine [Figure-I-39].



**Figure-I-39** : Canaux de solvant dans la structure cristallographique du mutant *TmHydE*  $\Delta$ FeS N-strep

Le résidu Thr2 (numérotation de la séquence de la protéine sauvage) est le premier résidu visible dans les cartes de densité électronique. L'étiquette strep N-terminale, se situant en amont de la Thr2, évolue sans aucun doute librement dans un large canal de solvant.

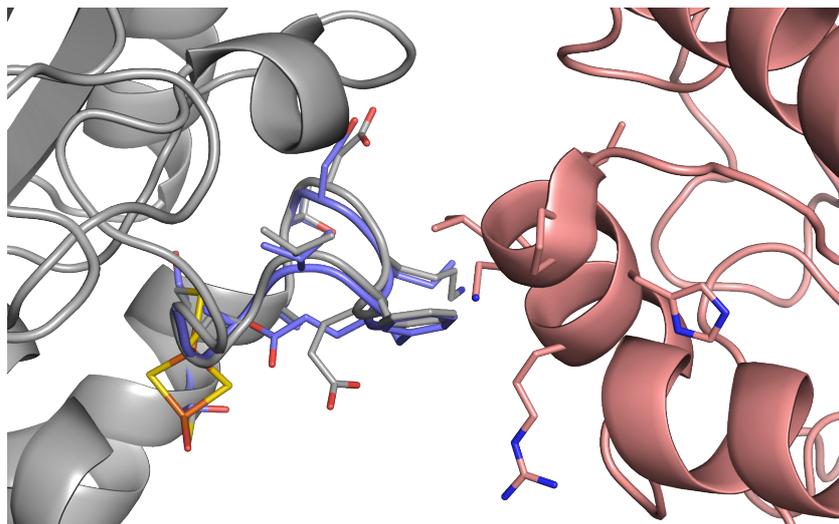
En ce qui concerne la boucle 311-319, elle contient les mutations C311S et C319S. Cette boucle se situe à la surface de la protéine et fait partie du site de chélation du second centre [FeS]. La comparaison de différentes structures de la protéine sauvage montre que cette zone présente une variabilité structurale en fonction de la forme du second agrégat [FeS] [Chapitre IV]. Chez le mutant, la mutation des deux résidus cystéines qui délimitent cette boucle déclenche sa translation vers l'endroit où, chez le sauvage, se trouve le second centre FeS. [Figure-I-40].



**Figure-I-40** : Déplacement de la boucle 311-319 suite à l'insertion de la triple mutation C311S, C319S, C322S

Superposition de la structure contrôle (en gris) avec un centre  $[Fe_2S_2]$  auxiliaire et de la structure du triple mutant C311S, C319S, C322S. Dans le cas du mutant (en violet), seuls les résidus de la boucle 311-319 et la Ser322 sont visualisés.

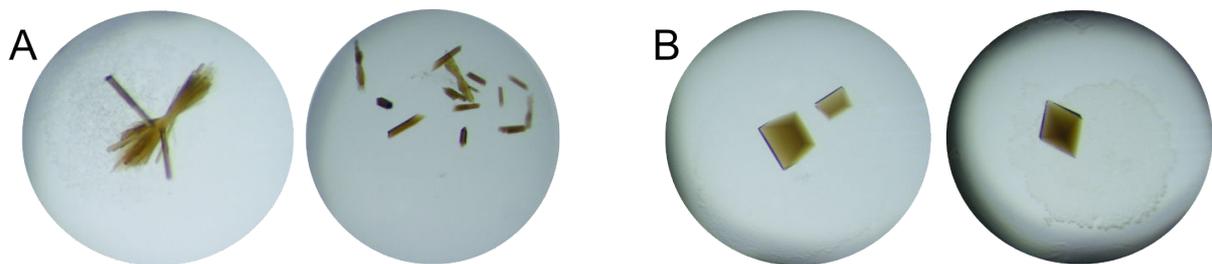
De façon macroscopique, cette mutation a aussi modifiée l'aspect des cristaux. Ceux de la protéine sauvage ont une forme en baguette, tandis que le mutant forme des cristaux en pyramide. Bien que le site de fixation du second agrégat  $[Fe-S]$  se trouve complètement exposé au solvant, certains résidus de la boucle 311-319 sont directement impliqués dans des contacts cristallins [Figure-I-41].



**Figure-I-41** : Contacts entre deux monomères de l'empilement cristallin impliquant interaction avec les résidus de la boucle 311-319

En gris et rose sont visualisés deux monomères de l'empilement cristallin dans la structure contrôle avec un centre  $[Fe_2S_2]$  auxiliaire et en violet, les résidus de la boucle 311-319 et la Ser322 du triple mutant C311S, C319S, C322S.

La cristallisation de l'échantillon de *TmHydE* sauvage, qui présente une inhomogénéité au niveau du second centre [Fe-S] et donc au niveau de la boucle 311-319, doit entraîner une croissance cristalline qui se traduit par la formation de cristaux en forme de baguettes fines [Figure-I-42-A]. Le mutant  $\Delta\text{FeS}$  N-strep, plus homogène, car ne pouvant contenir que le centre [Fe-S] SAM radical, forme des cristaux plus prismatiques [Figure-I-42-B]. Ainsi, dans le cas du mutant, un volume diffractant du cristal plus important, combiné avec une mosaïcité plus faible, se traduit par une augmentation de résolution de données cristallographique. Tandis qu'avec la forme sauvage il était rare de descendre en dessous de 1.40 Å, certains cristaux du triple mutant ont diffracté au-delà de 1.10 Å.



**Figure-I-42** : Aspect macroscopiques des cristaux de *TmHydE*

**A** - Faciès des cristaux de la forme native de la protéine *TmHydE*. **B** - Faciès des cristaux du mutant *TmHydE*  $\Delta\text{Fes}$  N-strep. Les deux formes A & B appartiennent au même groupe d'espace  $P2_12_12_1$ .

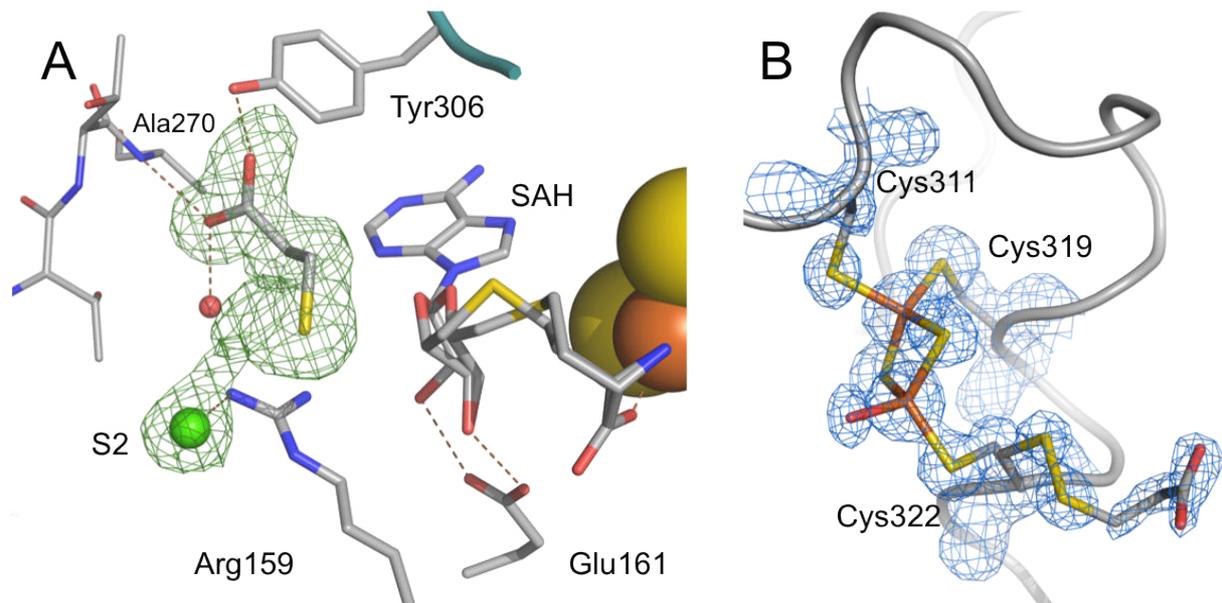
### **I-8.8.3 Co-cristallisation de *TmHydE* $\Delta\text{Fes}$ N-strep avec la L-cystine**

La co-cristallisation du mutant  $\Delta\text{Fes}$  N-strep avec la cystine n'a pas révélé la présence du ligand (2*R*,4*R*)-MeTDA dans le site actif. Il n'y avait pas de précipitation de protéine apparente lors de la cristallogénèse, donc pas de libération de Fe dans la solution. En l'absence d'ions de Fe libres la cystine ne subit pas ou très peu de dégradation. Cette expérience en particulier apporte deux conclusions : d'une part elle confirme notre hypothèse sur la dégradation de la cystine en présence de métaux de transition et d'autre part elle laisse penser que la L-cystine n'est pas ligand de HydE de *T. maritima* ; en tout cas pas pour la conformation de la protéine sélectionnée par l'empilement de molécules observé dans nos cristaux.

### **I-8.9 Complexe avec un ligand 3-mercaptopropionate modifié**

Parallèlement aux essais de co-cristallisation de l'enzyme avec la L-cystine nous avons testé des analogues de cette dernière, tels que la cystathionine et l'acide dithiopropionique. En utilisant ce dernier, nous avons réussi à obtenir une structure qui montre une densité électronique nouvelle à l'intérieur du site actif [Figure-I-43-A].

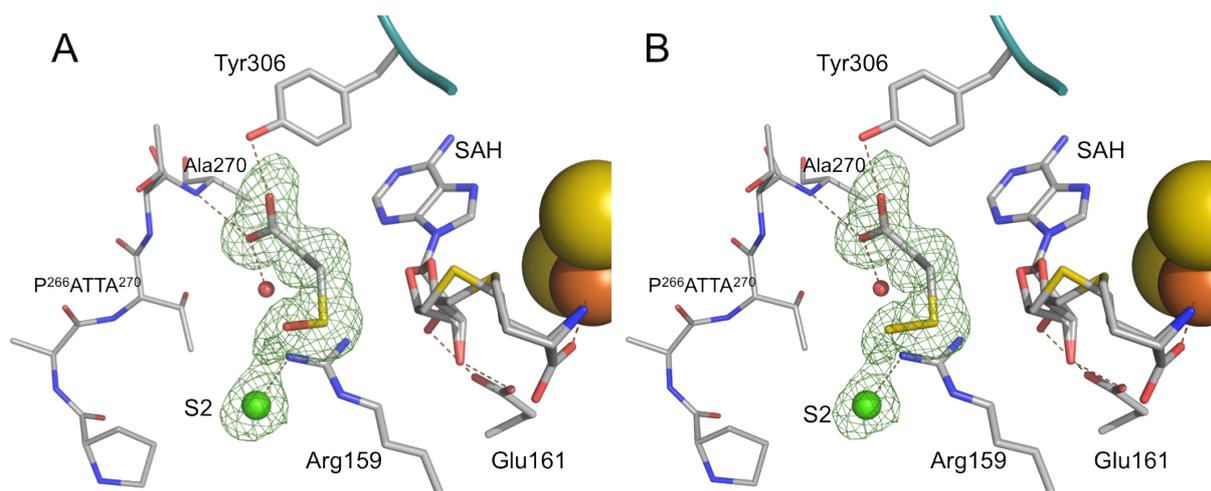
Une seconde molécule, faiblement occupée, semble être fixée par un pont disulfure à la Cys322 [Figure-I-43-B]. Ceci pourrait être la conséquence d'une attaque du thiol de la Cys322 sur le pont disulfure du dithiopropionate, ce qui conduirait à la libération du 3-mercaptopropionate réduit et à la formation d'un pont disulfure entre la Cys322 et une entité 3-mercaptopropionate libre. Il est à noter que l'échantillon protéique utilisé pour cette expérience était le même que celui de la protéine sauvage et contenant potentiellement du Fe libre en solution, issu de la dégradation partielle de l'agrégat [Fe-S] auxiliaire.



**Figure-I-43:** Détails de la structure du complexe de *TmHydE* avec un dérivé de 3-mercaptopropionate

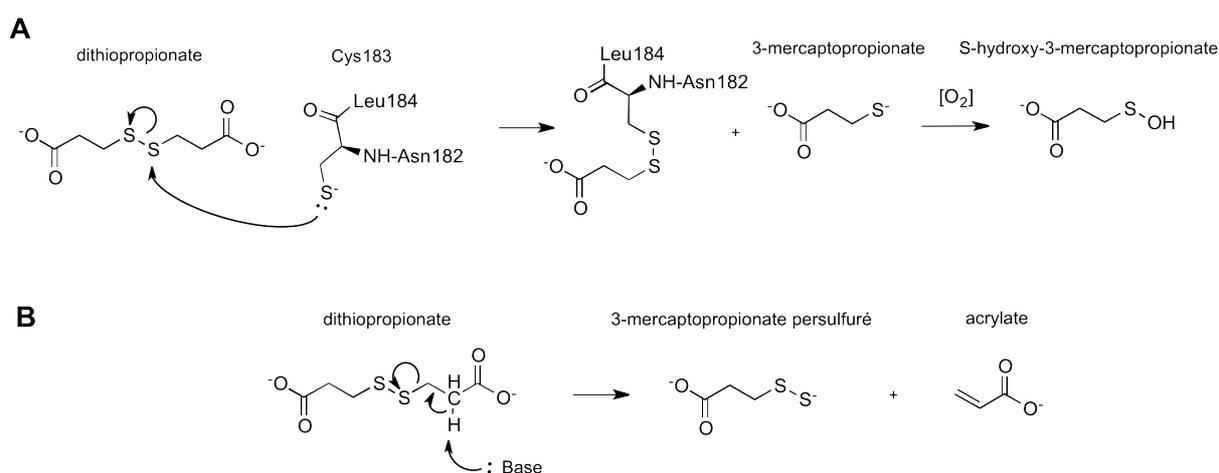
**A** - Environnement du site actif de HydE après l'incubation avec du dithiopropionate. En vert, la carte de densité électronique *Fo-Fc* "omit" contournée à  $3\sigma$ . Une molécule de 3-mercaptopropionate est modélisée dans cette carte. **B** - Environnement du site de fixation de l'agrégat [Fe-S] auxiliaire après l'incubation avec du dithiopropionate. En bleu, carte de densité électronique *2Fo-Fc* contournée à  $1\sigma$  (résolution 1,46 Å). A droite en bas de l'image, une molécule de 3-mercaptopropionate, modélisée avec un facteur d'occupation de 0,4, est liée par un pont disulfure à la chaîne latérale de la Cys322.

Nous proposons que la molécule dans le site actif pourrait être le S-hydroxy-3-mercaptopropionate [Figure-I-44-A] ou plus probablement encore, du 3-mercaptopropionate persulfuré [Figure-I-44-B]. Nous pensons que la formation du dernier composé résulterait d'une  $\beta$ -élimination d'un acrylate à partir du dithiopropionate utilisé dans la co-cristallisation [Figure-I-45-B]. La formation du premier dérivé pourrait être précédée par la réduction du dithiopropionate via une attaque du thiol de la Cys322 sur le pont disulfure de la molécule initiale. Ceci conduirait à la libération du 3-mercaptopropionate réduit et à la formation d'un pont disulfure entre la Cys322 et le 3-mercaptopropionate. La molécule libérée pourrait alors s'oxyder en présence de traces d'oxygène dans la boîte à gants pour former le S-hydroxy-3-mercaptopropionate [Figure-I-45-A]. La protéine elle-même témoigne de présence d'attaques oxydantes, car la Cys183, exposée au solvant, a été entièrement transformée en S-hydroxy-L-cystéine. Cependant, nous ne pouvons pas exclure que l'oxydation de ce résidu s'était produite pendant l'expression aérobie de la protéine.



**Figure-I-44: Site actif de *TmHydE* avec un dérivé de 3-mercaptopropionate**

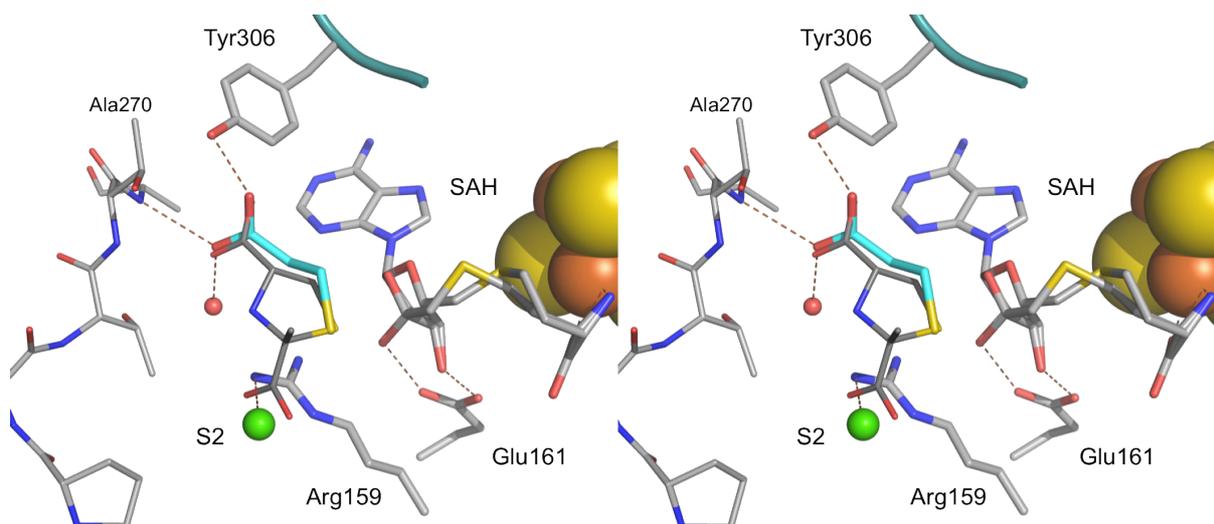
**A** - Environnement du site actif de HydE après l'incubation avec du dithioproprionate. En vert, la carte de densité électronique *Fo-Fc* "omit" contourée à  $3\sigma$ . Une molécule de S-hydroxy-3-mercaptopropionate est modélisée dans cette carte. **B** - Même données cristallographiques que l'image A, cette fois-ci avec une molécule de 3-mercaptopropionate persulfurée, modélisée dans la carte de densité électronique *Fo-Fc* "omit".



**Figure-I-45 : Schéma de conversion de dithioproprionate en dérivés de 3-mercaptopropionate**

**A** - Schéma de la production du S-hydroxy-3-mercaptopropionate à partir du dithioproprionate.  
**B** - Schéma de la production du 3-mercaptopropionate persulfuré à partir du dithioproprionate.

Dans les deux cas, la présence d'un atome supplémentaire, plus agité ou partiellement occupé, expliquerait raisonnablement bien la densité résiduelle proche de l'atome de soufre du dérivé 3-mercaptopropionate [Figure-I-44]. La fixation du ligand 3-mercaptopropionate oxydé n'apporte pas de preuve qui soutienne notre hypothèse sur une éventuelle fixation de la cystine dans le site actif de HydE. En effet le ligand ressemble plus à la cystéine ou une S-hydroxycystéine. Le fait que le groupement carboxylate de ce ligand se fixe dans le site de fixation d'anion S1 illustre bien l'affinité du site S1 pour ce groupement chimique. Il est intéressant de noter que son atome de soufre et son groupement carboxylate occupent exactement la même position que l'atome de soufre du ligand (2*R*,4*R*)-MeTDA [Figure-I-46].



**Figure-I-46** : Vue stéréoscopique de la superposition de l'emplacement du dérivé 3-mercaptopropionate et du dérivé (2*R*,4*R*)-MeTDA dans le site actif de *TmHydE*

La molécule de 3-mercaptopropionate est représentée par des bâtons de couleur cyan, tandis que le dérivé (2*R*,4*R*)-MeTDA est représenté par des bâtons fins de couleur gris foncé.

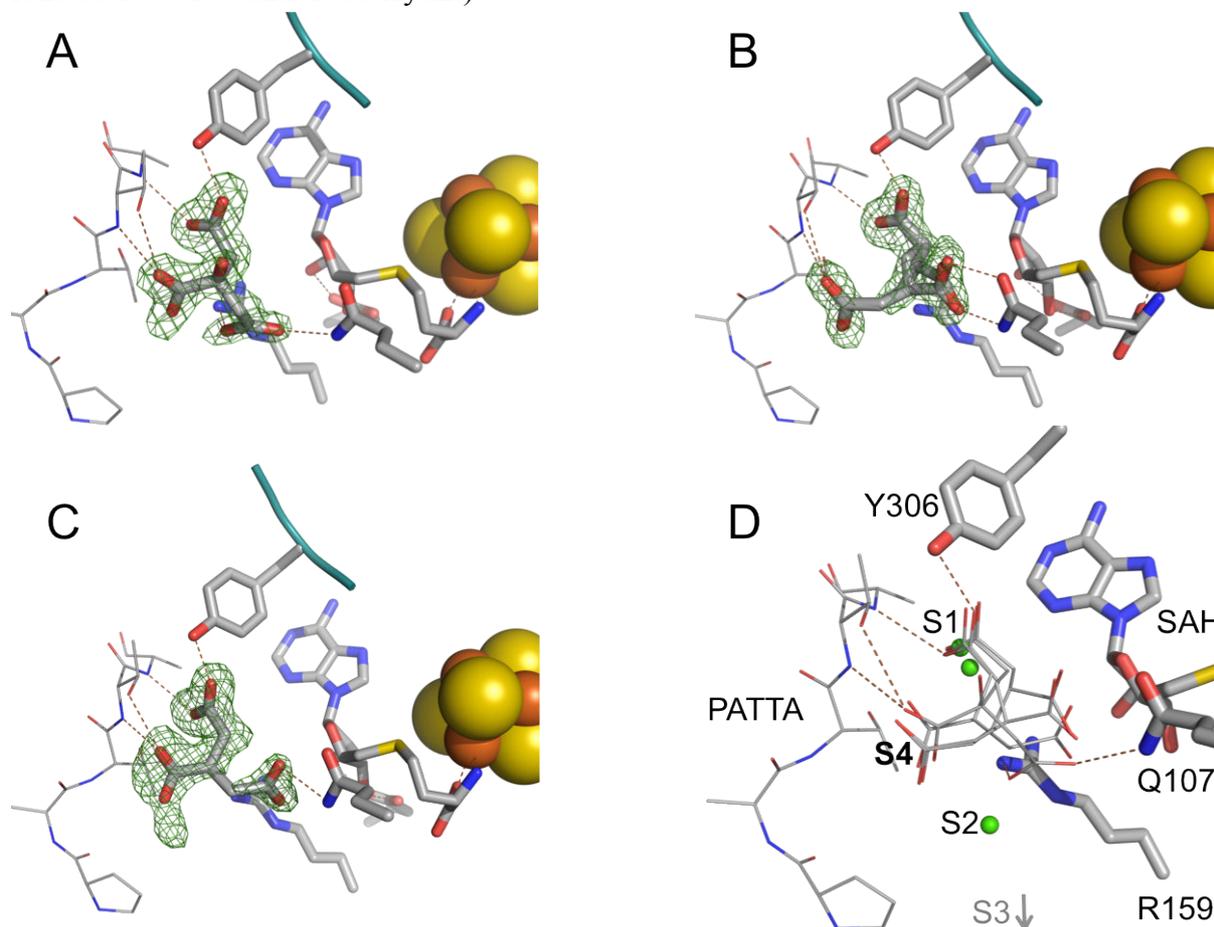
## I-9 Accostage moléculaire *in silico* avec les métabolites polycarboxylés

Les résultats des testes décrits dans les paragraphes précédents suggèrent que la cystine n'est pas le substrat de HydE. Nous n'avons pas réussi à prouver sa fixation dans la cavité de HydE ni la fixation d'un analogue d'un intermédiaire de la réaction proposée [Figure-I-32 A&B] [I-8.6 & I-8.8.3]. Nous avons alors décidé de refaire les expériences d'accostage moléculaire *in silico*, cette fois-ci en élargissant notre recherche en incluant les métabolites polycarboxylés de *E. coli* à partir de la base KEGG<sup>50</sup> ne contenant pas forcément d'atomes de soufre. Cette expérience a été réalisée par Patricia Amara de notre laboratoire. De cette façon nous espérons trouver des substrats potentiels de HydE qui satisfassent l'hypothèse (ii) sur l'origine d'atomes de soufre dans le ligand DTMN; c'est-à-dire, que HydE produit le ligand DTMN à partir d'un substrat dépourvu de soufre et d'une source de soufre extérieure, comme par exemple une cystéine désulfurase persulfurée [I-3.2]. A l'issue de ces calculs, une des molécules avec le score le plus fort était le citrate. Nous avons donc essayé de co-cristalliser HydE avec cette molécule et plusieurs de ses analogues [MM-4.1].

### I-9.1 Tricarboxylates comme ligands de HydE

Nous avons obtenu des structures cristallographiques de complexes de *TmHydE* avec les trois molécules tricarboxylées suivants : le citrate [Figure-I-47-A], le 1,2,4-butanetricarboxylate [Figure-I-47-B] et le tricarballylate [Figure-I-47-C]. L'analyse des structures de ces complexes montre que les trois ligands logent un de leurs groupements carboxylates dans le site de fixation d'anions S1 où ils interagissent avec le groupement hydroxyle de la Tyr306 et avec l'atome d'azote de l'Ala270, cette dernière appartenant au motif PATTA. Le second groupement carboxylate se fixe à proximité du site de fixation d'anions S2, proche de l'Arg159 et interagit avec la Gln107. Cette position présente le plus de

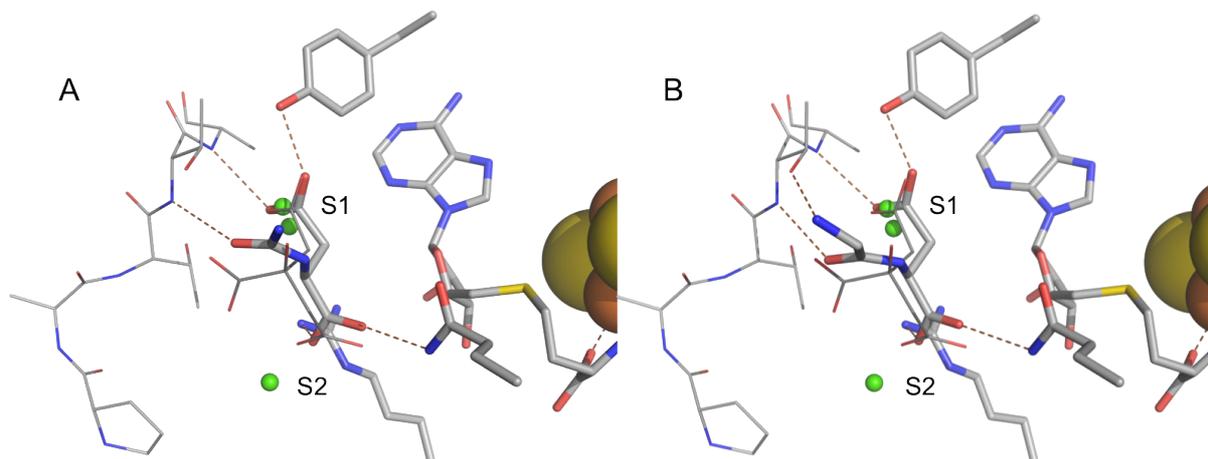
divergence entre les trois structures. Il est intéressant de noter, que les trois ligands tricarboxylés placent leur troisième carboxylate à proximité du motif PATTA où il forme une liaison hydrogène avec l'atome d'azote et le groupement hydroxyle de la Thr269 [Figure-I-47-D]. Cette zone diffuse qui présente une certaine affinité pour les groupements au potentiel électrostatique négatif sera appelée le site S4. (Rappelons que le site S3 se situe au fond de la cavité centrale de HydE.)



**Figure-I-47 : Structures de complexes de *TmHydE* avec trois molécules tricarboxylées**

Structures de complexes de *TmHydE* obtenues en présence de : A - citrate (résolution 1,74 Å), B - tricarballoylate (résolution 1,40 Å) et de C - 1,2,4-butanetricarboxylate (résolution 1,64 Å). En vert foncé, carte de densité électronique *F<sub>o</sub>-F<sub>c</sub>* "omit" contournée à 3  $\sigma$  et visualisée avec un rayon de 3 Å autour des ligands polycarboxylés. D - Superposition des trois structures représentées dans A, B et C.

Avant mon arrivée au laboratoire de nombreux essais de co-cristallisation de HydE avec des composés dicarboxylés ont été effectués. Certaines de ces molécules ressemblaient beaucoup à un des trois nouveaux ligands identifiés [Annexe-II]. Parmi ces molécules, les seules qu'on a pu retrouver dans la cavité de HydE étaient le N-carbamoyl-L-aspartate [Figure-I-48-A] et le dipeptide Glycyl-aspartate [Figure-I-48-B]. Comme pour les tricarboxylates, ces deux molécules placent un groupement carboxylate dans le site S1 et l'autre proche du site S2. Le carbonyle appartenant au groupement carbamoyl et celui de la liaison peptidique du dipeptide GlyAsp se trouvent à proximité du site S4. Pour compléter notre analyse, on notera que ni le L-aspartate, ni le L-glutamate ne se fixent dans le site actif.



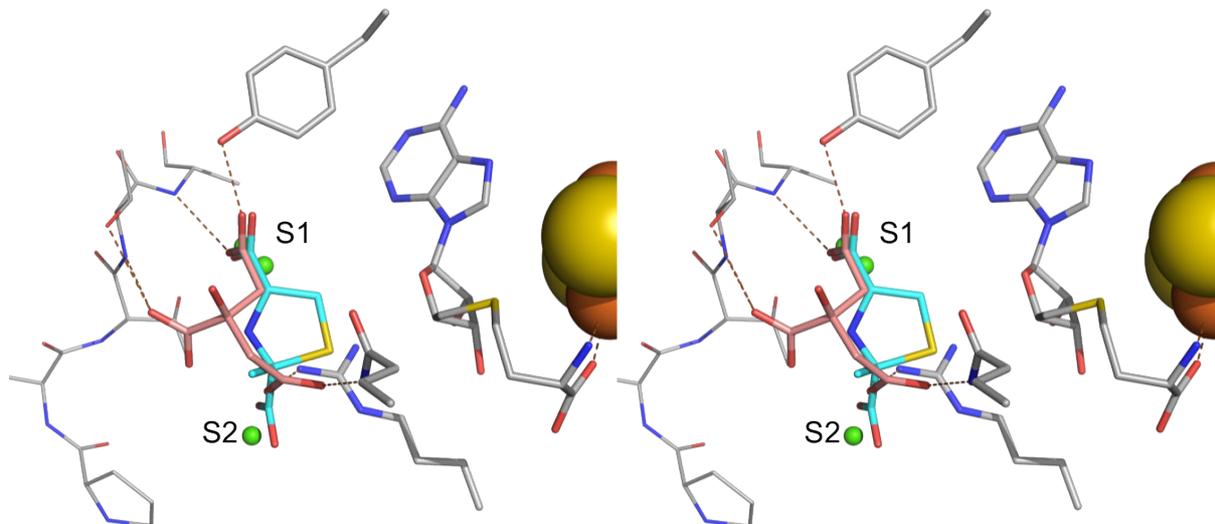
**Figure-I-48** : Structures des complexes de HydE avec le N-carbamoyl-L-aspartate et le glycyL-aspartate

Structures de complexes de *TmHydE* obtenues en présence de A - N-carbamoyl-L-aspartate (résolution 1,30 Å) et B - dipeptide glycyL-aspartate (résolution 1,43 Å). La molécule de citrate est représentée par des bâtons fins de couleur gris foncé.

L'absence de fixation de molécules telles que le glutarate, l'adipate ou encore le pimelate [voir Annexes] suggère que la présence seule de deux groupements carboxylates correctement espacés pour se loger dans les sites S1 et S2 ne garantit pas la fixation de la molécule dans la cavité de HydE. L'ajout d'un troisième groupement fonctionnel, pour lequel la cavité présente une affinité, ne serait ce que faible, augmente l'affinité globale de la molécule pour le site actif et permet la formation d'un complexe suffisamment stable pour être observé par cristallographie aux rayons X.

Cela dit, même si l'affinité de ces cinq ligands pour le site actif n'a pas été proprement quantifiée, elle est sans doute faible, parce que lorsqu'on lave les cristaux contenant un de ces ligands dans une solution de cryoprotection sans ligands, la molécule est rapidement éliminée du site actif. De plus, le 1,2,4-butanetricarboxylate se fixe dans le site actif en deux positions alternatives [Figure-I-47-B] et le tricarballylate présente des facteurs d'agitation thermique élevés sur deux de ses groupements carboxylate (carboxylate près du site S2: 33 Å<sup>2</sup>, carboxylate dans le site S4: 30 Å<sup>2</sup>, tandis que carboxylate dans le site S1: 22 Å<sup>2</sup>). On en conclut que la fixation de ces ligands, dont un n'est pas un produit naturel, n'est pas physiologique. Cette fixation est sans doute due à la co-cristallisation de l'enzyme avec de très fortes concentrations de ligands (25 mM final). Le potentiel électrostatique positif de la cavité combiné avec la présence d'au moins deux sites qui ont une affinité pour les groupements carboxylates facilite la formation de complexe avec des molécules polycarboxylées.

Lorsqu'on compare le mode de fixation des ligands polycarboxylés mentionnés plus haut avec le mode de fixation du dérivé (2*R*,4*R*)-MeTDA on s'aperçoit que seul ce dernier arrive à placer son second groupement carboxylate entièrement dans le site S2 [Figure-I-49]. C'est peut-être cette interaction avec la protéine, combinée à la présence d'un thioéther, qui sont à l'origine de l'affinité plus élevée que ce ligand a pour le site actif de HydE [I-11.3 & I-11.4.2]. En effet, comme on le verra plus loin, les traces du ligand (2*R*,4*R*)-MeTDA sont visibles dans la cavité de HydE même au bout de 24 heures d'un lavage du cristal dans une solution de cryoprotection qui ne contient pas de ligand [Figure-I-57].



**Figure-I-49** : Vue stéréoscopique de la comparaison de fixation de ligands (2*R*,4*R*)-MeTDA et citrate dans HydE

La molécule de citrate est représentée par des bâtons de couleur saumon et le dérivé (2*R*,4*R*)-MeTDA par des bâtons de couleur grise. La position des ions chlorures dans les sites S1 et S2, telle qu'observée dans la structure native, est représentée par des sphères vertes.

## I-10 Essais de co-purification de la protéine HydE Tm N-strep ΔFeS avec les intermédiaires de la réaction

Certaines enzymes ont des affinités relativement élevées pour leurs substrats à cause de la faible concentration de ce dernier dans les conditions physiologiques. Il arrive alors de co-purifier la protéine avec un ligand et de retrouver ce complexe ensuite dans le cristal. Même si ce cas de figure est rare, cette approche « quick-and-dirty » est une des premières expériences à faire pour tenter d'identifier le substrat ou un intermédiaire de réaction de l'enzyme. Quitterer et ses collègues<sup>33</sup> ont pu, de cette manière, résoudre la structure du complexe de la protéine à radical SAM PylB co-purifiée avec la SAM et son produit – la (3*R*)-3-méthyle-D-ornithine. Cependant, dans le cas de *Tm*HydE il y a eu plusieurs raisons qui rendaient cette expérience compliquée à mettre en œuvre au début de mes travaux de thèse.

La purification de la protéine sauvage se fait par chromatographie échangeuse d'ions [voir paragraphe MM-3.2.6 et MM-3.2.7]. Dans cette technique l'usage de fortes concentrations de NaCl peut induire la perte de ligand. Les ions chlorure peuvent en effet entrer en compétition avec les groupements carboxylate d'un ligand liés aux sites de fixations d'anions S1 et S2. De ce fait, une autre méthode de purification s'impose afin de garder le complexe protéine-ligand intact jusqu'à la cristallisation.

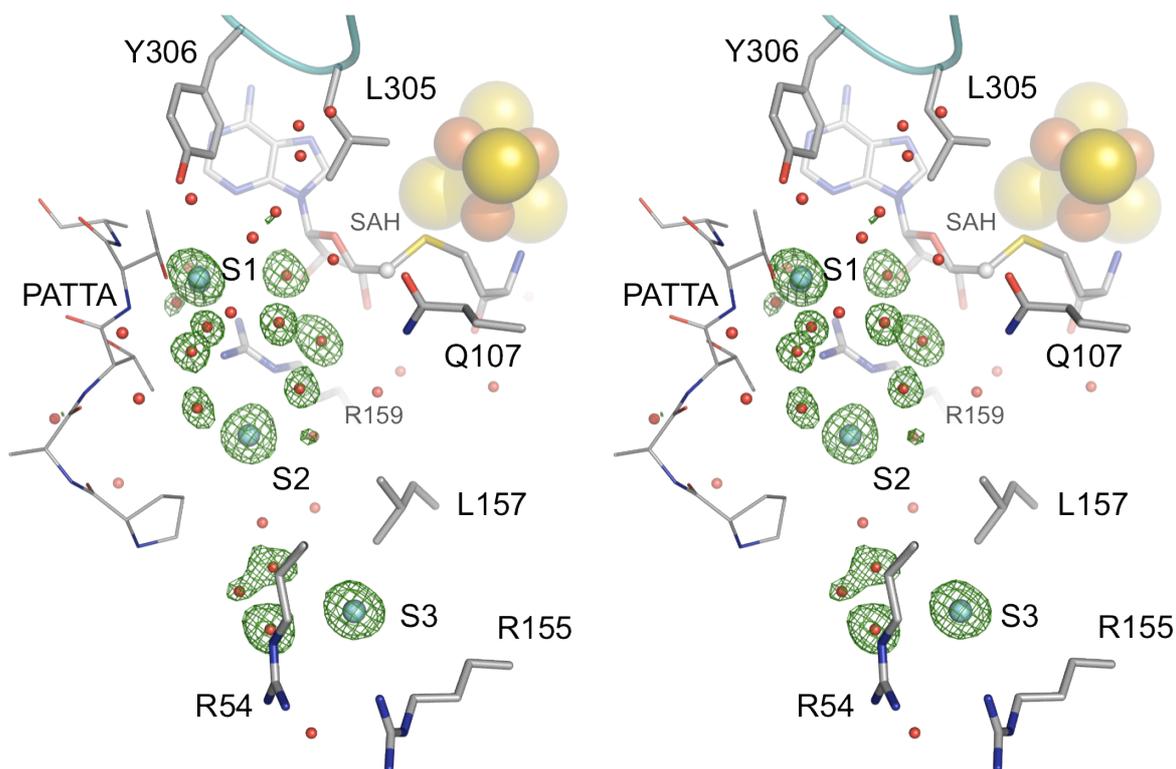
Les essais de cristallisation réalisés peu de temps après la purification de la protéine ont donné des amas fortement polycristallins, provoqués sans doute par l'inhomogénéité du second centre [Fe-S] [I-8.8.2]. Le même échantillon protéique, incubé pendant plusieurs semaines en boîte à gants et co-cristallisé avec de la SAH, a fourni une structure dite « native » de l'enzyme. On a pu remarquer que le second centre [Fe-S] était sous sa forme persulfurée. Ceci constitue à nos yeux la preuve d'une dégradation d'un centre [Fe<sub>4</sub>S<sub>4</sub>] initialement présent dans le site de fixation du second centre [Fe-S] [*Chapitre IV*].

Nous avons décidé de refaire ces expériences avec le mutant  $\Delta$ FeS N-strep parce que premièrement, la présence de l'étiquette strep du côté N-terminal nous permettait de nous affranchir de l'usage de forte concentration de NaCl et deuxièmement, la mutation des trois cystéines impliquées dans la fixation du second centre [Fe-S] élimine le problème d'inhomogénéité de l'échantillon lié à ce centre. Le dernier point était l'utilisation d'une concentration élevée de Fe (1mM) et de DTT (5mM) pendant la reconstitution de la protéine. Ces molécules ne sont pas anodines dans l'extrait et peuvent avoir pour conséquence la modification du profil métabolique de l'extrait. Ainsi nous avons préféré séparer la reconstitution chimique des centres [Fe-S] de l'incubation de l'holoenzyme HydE avec l'extrait anaérobique de *E. coli*.

### **I-10.1 Incubation de la protéine TmHydE Tm Nstrep avec du lysat anaérobique d'*E. coli***

Nous avons tout d'abord reconstitué l'enzyme chimiquement dans l'extrait et nous l'avons ensuite purifié par chromatographie d'affinité [voir MM-3.2.2]. Cette enzyme pure et possédant son centre [Fe<sub>4</sub>S<sub>4</sub>] 'SAM radical' a été incubée en présence d'un excès de SAH. La protéine a, par la suite, été immobilisée sur une colonne de Streptactine, préalablement équilibré dans un tampon 50mM Tris pH 8 sans NaCl. HydE a été ensuite incubée dans la colonne avec un lysat anaérobique clarifié. Enfin, elle a été purifiée et cristallisée en utilisant les protocoles standards modifiés : le NaCl a été exclu de toutes les solutions et la protéine a été cristallisée sans l'ajout de la SAH.

La structure de l'enzyme ainsi traitée présentait dans son site actif une densité allongée dans la carte de Fourier différence. Grâce à la très haute résolution de données cristallographique (1,15Å) au cours de l'affinement cette densité a pu être modélisée comme plusieurs molécules d'eau partiellement occupées. Dans ces conditions la protéine n'a donc pas pu être co-purifiée avec quelconque substrat ou produit. Par contre la structure obtenue a été utilisée comme référence structurale de la protéine sauvage à haute résolution, qui répertorie toutes les molécules d'eau présentes dans la cavité à son état natif [[Figure-I-50](#)].



**Figure-I-50** : Vue stéréoscopique de la cavité de *TmHydE* sous sa forme native à une résolution de 1,15 Å.

En vert foncé, carte de densité électronique *Fo-Fc* "omit" contourée à  $3\sigma$  et visualisée avec un rayon de 3 Å autour des molécules d'eau et des ions chlorure. Dans les structures de complexes ces molécules se trouvent déplacés par les ligands ou bien présentent souvent des densités électroniques irrégulières.

### I-10.2 Reconstitution *in vivo* de *TmHydE* *TmNstrep* $\Delta$ FeS

La concentration du substrat de HydE dans le cytoplasme de *E. coli* n'est pas connue. Il est donc tout-à-fait possible que lors de l'expérience d'incubation de la protéine *TmHydE* avec le lysat anaérobie, un grand excès de protéine immobilisée sur la résine a été incubé avec une faible quantité de ligand. Dans le but de se rapprocher des conditions physiologiques nous avons procédé à une reconstitution *in vivo* de l'enzyme. Comme nous l'avons déjà évoqué [IB-2.1 & I-2], l'activation de l'hydrogénase à [FeFe] dans l'extrait d'*E. coli* contenant les maturases HydE, F & G montre que le substrat de HydE est présent dans cette bactérie<sup>2,6,15-17,51</sup>. Cette molécule doit donc être un métabolite commun, dont la synthèse ne nécessite pas d'intervention d'une enzyme spécifique des organismes possédant une hydrogénase.

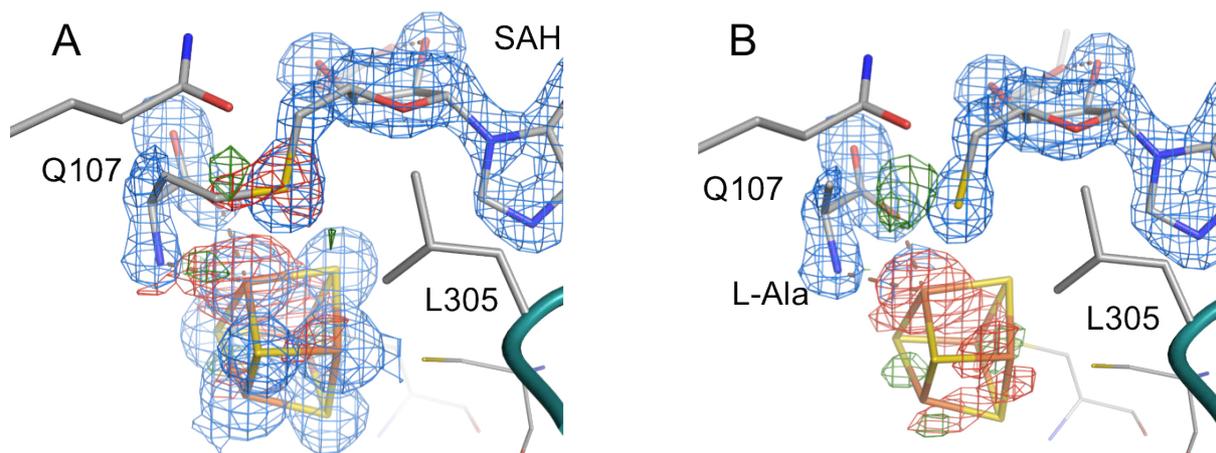
*E. coli*, appartenant à la famille des *Enterobacteriaceae*, est un organisme anaérobie facultatif. En absence d'oxygène il est capable de produire de l'ATP par fermentation acide mixte ou par respiration anaérobie. Les expériences réalisées par l'équipe de James R. Swartz montrent que la maturation d'hydrogénase est possible dans les deux cas<sup>52-55</sup>. Ceci veut dire que le substrat de HydE doit être produit par les deux métabolismes. Cependant, comme les concentrations des métabolites peuvent varier en fonction du métabolisme il nous a paru plus prudent de tester les deux.

Nous avons donc repris et adapté deux protocoles de maturation *in vivo* des protéines à centre [Fe-S], mis au point par l'équipe de James R. Swartz<sup>53</sup> et nous avons exprimé la protéine *TmHydE* *Tm Nstrep*  $\Delta$ FeS en anaérobiose sous respiration de fumarate [voir MM-3.1.1] et fermentation de glucose [voir MM-3.1.2].

Nous avons aussi utilisé une souche d'*E. coli* porteuse du plasmide *pRSRF Isc MetK*, qui, comme détaillé dans [MM-1.1.3 & MM-3.6.2], sert à améliorer le rendement de reconstitution des centre [Fe-S] et à augmenter la concentration intracellulaire de SAM<sup>56-59</sup>. Afin d'éviter une déplétion du ligand dans le cytoplasme suite à une forte surexpression de HydE nous avons réduit la concentration d'IPTG lors de l'induction de 500 à 10  $\mu$ M. Cette concentration a été préalablement déterminée par des tests d'induction. Elle correspond à la concentration d'IPTG limite à laquelle on peut encore observer une bande de HydE sur un gel d'expression. L'inconvénient de cette approche est qu'elle affecte aussi la surexpression de la machinerie Isc et de la protéine MetK. Cependant, puisque ces protéines sont des enzymes et des transporteurs, même une faible surexpression devrait, à condition de l'abondance de leurs substrats, augmenter la concentration intracellulaire de SAM et de centres [FeS]. Ainsi, si leur niveau d'expression est du même ordre de grandeur que celui de la protéine HydE, cette dernière devrait, en théorie, posséder une quantité stœchiométrique de SAM et d'agrégat Fe<sub>4</sub>S<sub>4</sub>. A l'issue de l'expérience d'incubation avec le lysat la protéine a été purifiée dans un tampon dépourvu de NaCl et elle a été cristallisée sans l'ajout de SAH ou de SAM.

#### ***I-10.2.1 Composition de la cavité du mutant $\Delta$ FeS N-strep reconstitué *in vivo* en fermentation du glucose***

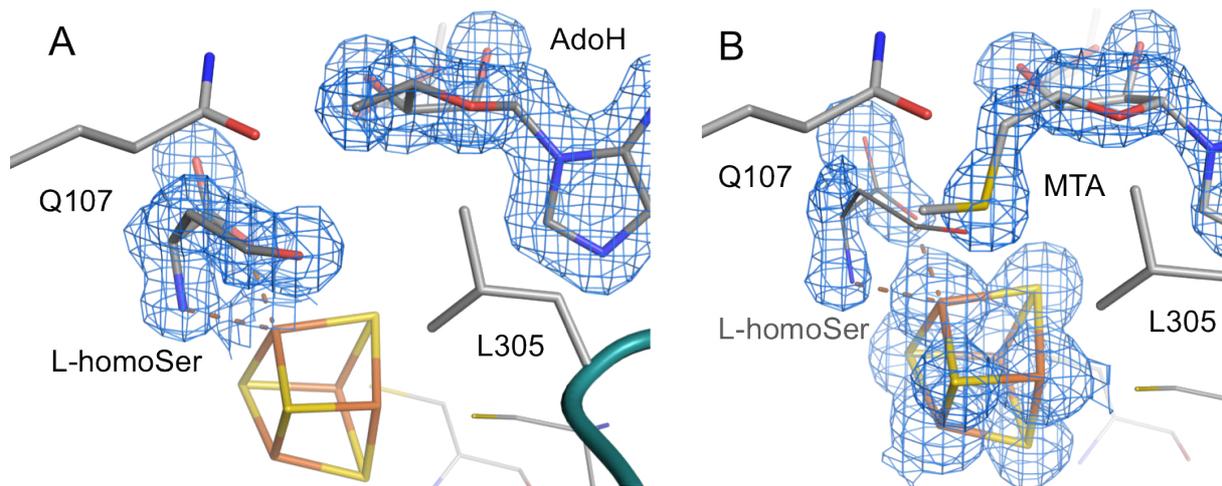
La structure cristallographique du mutant  $\Delta$ FeS N-strep reconstitué *in vivo* en condition de fermentation du glucose ne montre pas de densité électronique dans les sites S1, S2 et S3 qui ne pourrait pas s'expliquer par une molécule d'eau ou de chlorure qu'on observe dans la structure native de référence. Cependant le cofacteur SAM et le centre Fe<sub>4</sub>S<sub>4</sub> SAM radical présente des modifications apparentes. Dans la carte de Fourier différence on peut voir une densité négative sur l'atome de Fe unique du centre [Fe-S] et sur l'atome de soufre de la SAH. On remarque également une rupture de continuité de carte électronique entre le C <sub>$\beta$</sub>  et le S <sub>$\delta$</sub>  du cofacteur [Figure-I-51-A]. Des cartes très similaires ont été obtenues lors d'un essai de co-cristallisation de la protéine sauvage avec de la SAM commerciale. Dans les deux structures la densité électronique pourrait être raisonnablement bien expliquée comme de l'alanine et de la 5'-thioadénosine [Figure-I-51-B].



**Figure-I-51** : Environnement du centre [Fe-S] SAM radical après sa reconstitution *in vivo* en condition de fermentation du glucose

En bleu, carte de densité électronique  $2Fo-Fc$  contourée à  $1 \sigma$  et en vert foncé, carte  $Fo-Fc$  contourée à  $3 \sigma$  (résolution 1,48 Å). **A** - Résultat d'affinement de données cristallographiques avec une molécule de SAH modélisée avec un taux d'occupation de 100%. **B** - Les mêmes données cristallographiques avec une molécule de L-alanine et 5'-thioadénosine dans le modèle. On notera que la densité résiduelle positive (en vert), proche de l'atome de soufre du composé 5'-thioadénosine modélisé, ne correspond à la position d'aucun atome du cofacteur SAM ou SAH.

La SAM est connue pour sa stabilité limitée en milieu neutre ou basique<sup>60-62</sup>. C'est d'ailleurs la raison pour laquelle elle est vendue dans un tampon 0.005M H<sub>2</sub>SO<sub>4</sub> pH 4. La SAM peut subir de façon simultanée plusieurs dégradations différentes. Parmi les produits de sa dégradation hydrolytique on trouve de l'homoserine et de la méthylethioadénosine<sup>61</sup>. Nous disposons d'une structure avec de la 5'-deoxyadénosine et une homoserine liée au Fe unique du centre [Fe-S] SAM radical [Figure-I-52-A]. La comparaison de cette structure de référence avec notre modèle montre clairement que l'atome de l'oxygène de l'homoserine occuperait le même espace que l'atome de soufre de la 5'-thioadénosine modélisée [Figure-I-52-B]. Il est donc possible que dans notre structure la chaîne latérale de l'homoserine soit déplacée et que l'agitation des atomes C<sub>γ</sub> et O<sub>δ</sub> crée une densité qui ressemble à celle de l'alanine. Une agitation plus élevée du méthyle d'une méthylethioadénosine pourrait expliquer la densité résiduelle proche de l'atome de soufre de la 5'-thioadénosine modélisée [Figure-I-51-B]. Les similitudes entre notre modèle et la structure obtenue par co-cristallisation de l'enzyme avec la SAM commerciale suggèrent que la protéine reconstituée *in vivo* en condition de fermentation du glucose contient un mélange de produits de dégradation de son cofacteur SAM.



**Figure-I-52** : Comparaison entre les données cristallographiques issue de l'expérience de reconstitution *in vivo* en condition de fermentation du glucose avec la structure du complexe avec AdoH et L-homoserine

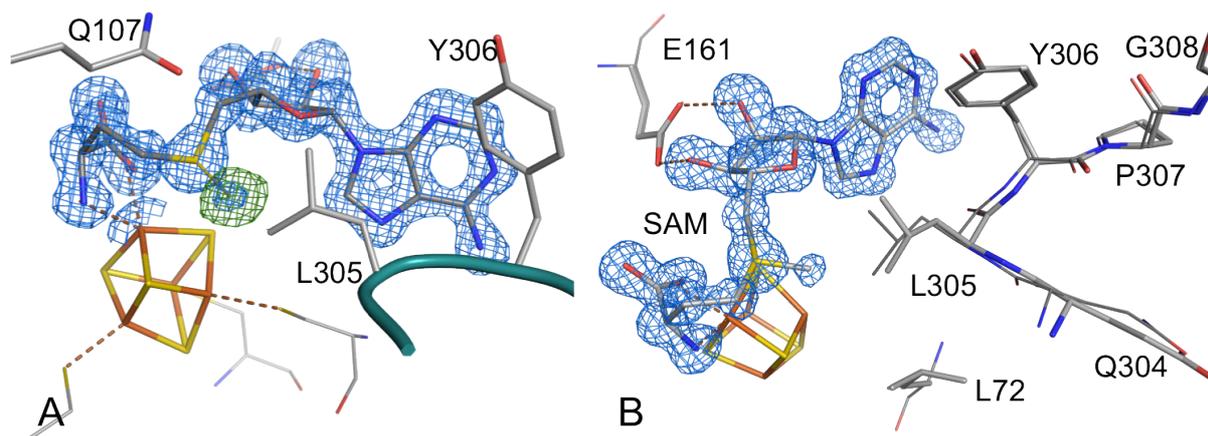
En bleu, carte de densité électronique  $2Fo-Fc$  contournée à  $1\sigma$ . A - Détails de la structure du complexe de TmHydE avec AdoH et une molécule de L-homoserine fixée sur le centre [Fe-S] 'SAM radical' (résolution 1,45 Å). B - Molécules de L-homoserine et MTA modélisées dans la carte de densité électronique correspondant à l'expérience de reconstitution *in vivo* (résolution 1,48 Å).

On estime que l'occupation de l'homoserine dans notre structure est aux alentours de 80%, ce qui correspond à l'occupation de l'atome de Fe unique, auquel l'homoserine est liée. Dans certaines molécules de HydE l'absence de liaisons entre le cofacteur et le Fe unique engendre la perte partielle de cet atome. Cela a déjà été observé lors de la co-cristallisation de l'enzyme chimiquement reconstituée avec de la 5'-deoxyadenosine seule. Dans la structure ainsi obtenue, le Fe unique, ne pouvant pas être coordonné par le cofacteur, est progressivement perdu pour former un  $Fe_3S_4$  [Chapitre IV]. Nous en concluons, que la protéine reconstituée *in vivo* en condition de fermentation du glucose contient un mélange de produits de dégradation de la SAM en quantité sous-stœchiométrique et que de ce fait le centre  $[Fe_4S_4]$  'SAM radical' perd progressivement son Fe unique. Il semblerait donc que cette condition ne soit pas appropriée pour l'isolation d'un complexe de HydE avec un intermédiaire de sa réaction physiologique.

#### **I-10.2.2 Composition de la cavité du mutant $\Delta FeS$ reconstitué *in vivo* en respiration du fumarate — Co-purification de la SAM**

La structure cristallographique du mutant  $\Delta FeS$  N-strep reconstitué *in vivo* en condition de respiration anaérobie n'a révélé aucune nouvelle densité électronique dans la région du site actif. On en conclue que soit la concentration du substrat de HydE dans le cytoplasme de *E. coli* est faible soit que l'affinité de cette molécule pour le site actif de HydE n'est pas suffisante pour permettre la co-purification du complexe. Cependant, dans cette nouvelle structure le centre [Fe-S] semble être entièrement reconstitué et le cofacteur, quant à lui, présente une densité résiduelle à l'endroit où devrait se trouver le méthyle de la SAM [Figure-I-53-A]. La présence du cofacteur intact dans notre échantillon protéique est corroborée par le fait, que les atomes de la chaîne latérale du résidu Leu305, qui se situe en

face du méthyle de la SAM, ont des facteurs d'agitation élevés (SAM -  $S_{\delta}$ : 26 Å<sup>2</sup> & CH<sub>3</sub>: 29 Å<sup>2</sup> ; Leu305 -  $C_{\delta 1}$ : 39 Å<sup>2</sup> &  $C_{\alpha}$ : 24 Å<sup>2</sup>). On observe le même phénomène dans la structure de référence 3IIZ, qui contient un mélange SAM/SAH<sup>14</sup>. En effet, en présence de la SAM, la chaîne latérale de la Leu305 est légèrement poussée par le méthyle du cofacteur. Ce déplacement affecte également la position de la boucle 304-308 [Figure-I-53-B]. Grâce à ces changements structuraux subtils, mais bien visibles, il est facile de distinguer la SAM de son analogue SAH fixé sur le centre [Fe<sub>4</sub>S<sub>4</sub>] 'SAM radical' de HydE.



**Figure-I-53** : Environnement du centre [Fe-S] 'SAM radical' après sa reconstitution *in vivo* en condition de respiration de fumarate

En bleu, carte de densité électronique  $2Fo-Fc$  contournée à  $1 \sigma$  et en vert foncé, carte de densité électronique  $Fo-Fc$  "omit" contournée à  $3 \sigma$  (résolution 1,15 Å). A - Apparition d'une nouvelle densité correspondant au méthyl de la SAM. B - Déplacement de la boucle 304-308 (en bâtons épais) en présence du méthyl de la SAM. La conformation de la boucle en absence du méthyl, c'est-à-dire en présence de la SAH, est représentée pqr des bâtons fins. On remarquera la bonne superposition de structure globale illustrée par la conformation des résidus Leu72 et Glu161.

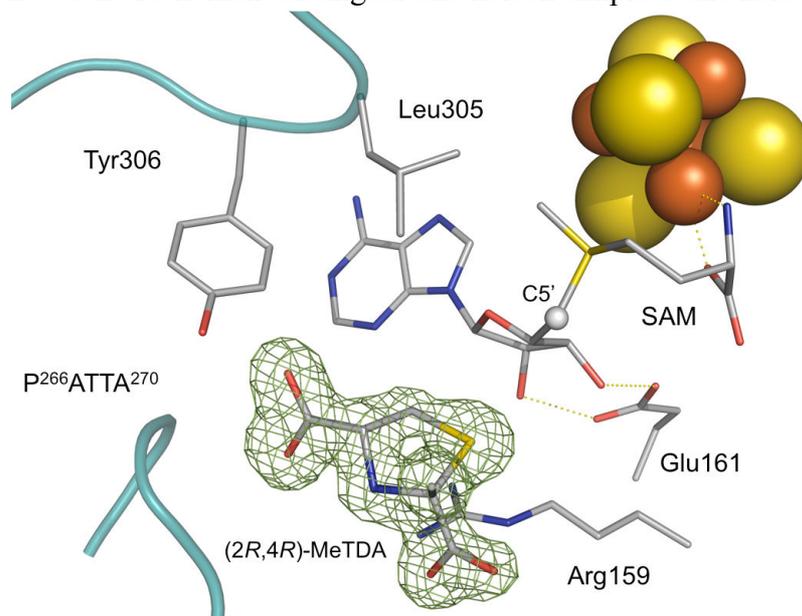
La protéine reconstituée *in vivo* en condition de respiration de fumarate contient un centre [Fe<sub>4</sub>S<sub>4</sub>] entièrement reconstitué et est co-purifiée avec son cofacteur physiologique SAM intact. Ainsi notre méthode de reconstitution *in vivo* des centres [Fe-S] d'une métalloprotéine modérément surexprimée en condition anaérobie et en présence du plasmide *pRSRF Isc MetK* semble être efficace. L'équipe de Joan Broderick (MSU, USA) a récemment observée une co-purification à hauteur de 25% de SAM avec la protéine HydE de *C. acetobutylicum*, fortement surexprimée en condition d'anaérobie<sup>63</sup>.

L'échantillon de HydE  $\Delta FeS$  reconstitué *in vivo* en condition de respiration anaérobie a joué un rôle crucial dans mes travaux de thèse, car il produit des cristaux qui contiennent des quantités quasi-stœchiométriques du cofacteur SAM intact. C'est cet échantillon qui a rendu possible l'étude de la réaction radicalaire dans les cristaux de HydE [Chapitre III-3 & III-5]. Une co-cristallisation avec de la SAM commerciale n'aurait pas pu être utilisée, car elle produit souvent des structures cristallographiques avec un mélange de SAH et SAM ou/et des produits de dégradation de cette dernière [voir plus haut].

## I-11 Etude des ligands dérivés de L-cystéine

### I-11.1 Trempage de cristaux de HydE de *T. maritima* $\Delta$ FeS N-strep avec de la cystéine et du pyruvate

Parmi tous les ligands de HydE, identifiés lors de nos efforts pour trouver le substrat de cette enzyme, le dérivé (2*R*,4*R*)-MeTDA semblait avoir le plus d'affinité pour le site actif de HydE. Nous avons alors décidé de nous intéresser de plus près à ce ligand. Le trempage de cristaux du mutant  $\Delta$ FeS N-strep avec une solution de cryoprotection contenant un mélange équimolaire de cystéine et de pyruvate a généré les mêmes cartes de densité électronique dans le site actif de HydE [Figure-I-54] que celles observées pour la co-cristallisation de la protéine sauvage avec la cystéine [Figure-I-33]. Ce jeu de données atteignait une résolution de 1.15 Å. Ce nouveau jeu proche de la résolution atomique a formellement confirmé la nature du ligand comme (2*R*,4*R*)-MeTDA car, à cette résolution, on peut distinguer un atome d'azote d'un atome d'oxygène. On peut également mesurer les distances interatomiques avec une grande précision et donc facilement distinguer une liaison simple d'une liaison double.



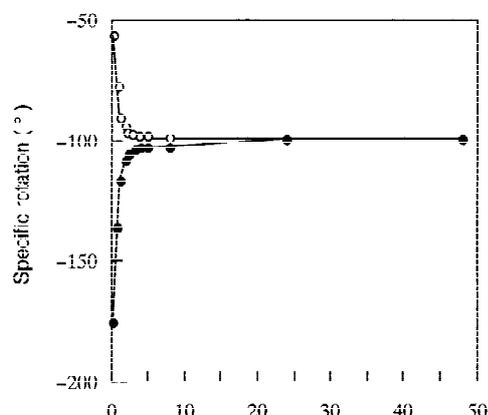
**Figure-I-54** : Structure cristallographique du complexe de *TmHydE* avec le dérivé thiazolidine (2*R*,4*R*)-MeTDA

La structure cristallographique du complexe de *TmHydE* avec le ligand (2*R*,4*R*)-MeTDA a été obtenue par co-cristallisation de la protéine avec un mélange équimolaire de L-cystéine et de pyruvate. En vert foncé, carte de densité électronique *Fo-Fc* "omit" contournée à 3  $\sigma$  (résolution 1,15 Å).

### I-11.2 Etude du comportement du ligand MeTDA en solution

L'étude RMN  $^1\text{H}$  en solution de la condensation de la L-cystéine avec le pyruvate, mélangés avec un rapport 1:1, révèle la production de deux diastéréoisomères (2*S*,4*R*)-2-Méthyle-1,3-thiazolidine-2,4-dicarboxylate et (2*R*,4*R*)-2-Méthyle-1,3-thiazolidine-2,4-dicarboxylate [Figure-I-56-A&B]. La conversion est, comme rapporté dans la

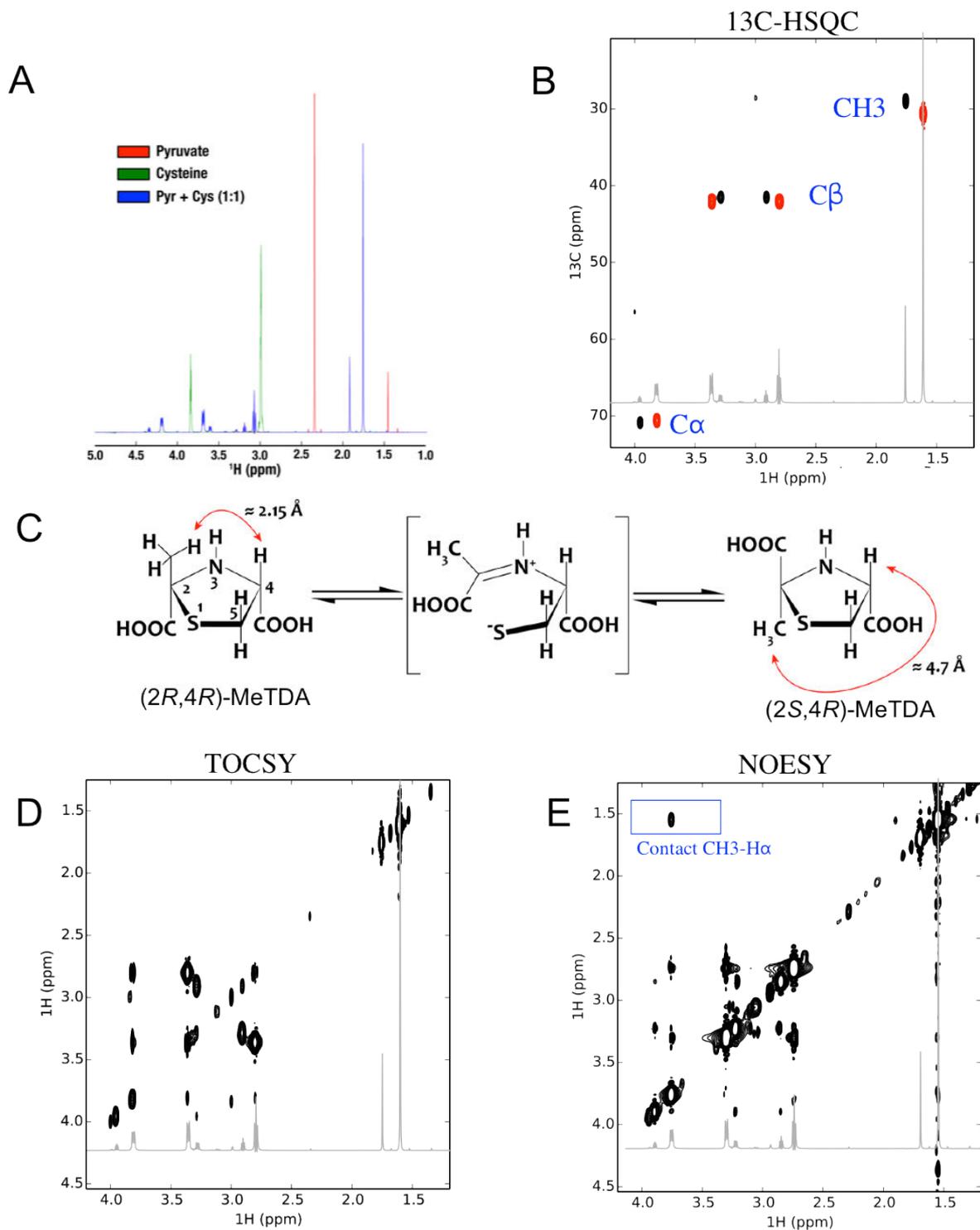
littérature<sup>49,64,65</sup>, rapide et totale. Les deux épimères ne peuvent pas être séparés par recristallisation à partir d'une solution aqueuse bien que l'épimère (2*S*,4*R*) soit moins soluble que l'épimère (2*R*,4*R*)<sup>66</sup>. En effet, il a été démontré que les 1,3-thiazolidines portant une substitution sur le C2 sont soumises dans l'eau à une épimérisation relativement rapide<sup>49,65,67,68</sup> [Figure-I-55]. Cette réaction procède par une ouverture du cycle thiazolidine entre les atomes S1 et C2<sup>49,64,69</sup> [Figure-I-56-C]. Il est cependant possible d'effectuer une séparation en utilisant les amines organiques<sup>65</sup>. La synthèse asymétrique est également utilisée, car le rapport final entre les deux épimères est fortement dépendant du pH de la solution. Ainsi en contrôlant le pH initial il est possible de synthétiser préférentiellement un seul épimère sous contrôle cinétique<sup>49,69,70</sup>.



**Figure-I-55** : Epimérisation des dérivés (2*S*,4*R*)- et (2*R*,4*R*)-2,4-thiazolidine-dicarboxylate

Représentation en cercles noirs du dérivé (2*R*,4*R*)- pur solubilisé dans l'eau et en cercles blancs du dérivé (2*S*,4*R*)- pur solubilisé dans l'eau. L'axe des abscisses représente le temps en heures. (Image adaptée de Shiraiwa et al, Chirality, 1999).

L'attribution stéréospécifique des résonances a été effectuée en exploitant l'effet Overhauser nucléaire<sup>71,72</sup> ou «NOE», observé par des expériences 2D de RMN du <sup>1</sup>H [Figure-I-56-E][MM-5.1]. Seul l'épimère (2*R*,4*R*) peut présenter un pic dans la carte NOESY correspondant au transfert de polarisation entre le proton sur le C4 (l'ancien proton H<sub>α</sub> de la L-cystéine) et les protons du méthyle, qui lui est attaché sur le C2 [Figure-I-56-C,D&E]. La distance élevée qui sépare ces protons dans l'épimère (2*S*,4*R*) empêche l'apparition de NOE [Figure-I-56-C].

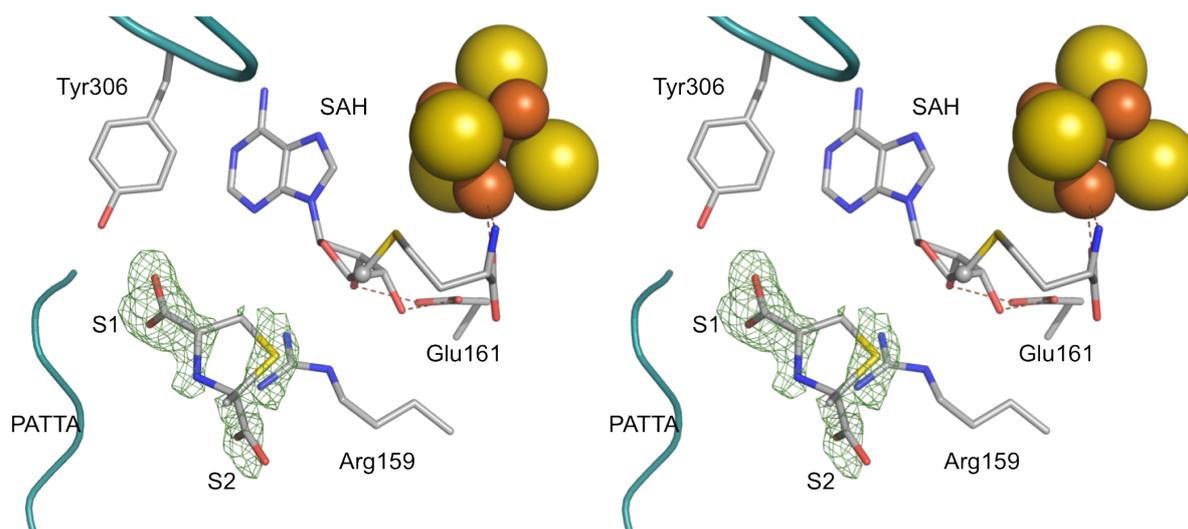


**Figure-I-56** : Etude spectroscopique par RMN de  $^1\text{H}$  et  $^{13}\text{C}$  du dérivé thiazolidine (4R)-MeTDA

**A** - Spectre 1D de RMN du  $^1\text{H}$  montrant une conversion complète d'un mélange équimolaire de pyruvate et de L-cystéine en dérivés thiazolidines (2S,4R)-MeTDA et (2R,4R)-MeTDA. **B** - Spectre HSQC  $^1\text{H}$ - $^{13}\text{C}$  utilisé, avec le spectre NOESY en **E**, pour l'attribution stéréospécifique de différents pics : l'épimère (2S,4R)-MeTDA en noir et l'épimère (2R,4R)-MeTDA en rouge. **C** - Mécanisme proposé pour l'épimérisation du dérivé (4R)-MeTDA sur le carbone C2. Les distances minimales entre le méthyle et l'hydrogène sur le C4 (C $\alpha$  de la partie L-cystéine) sont indiquées en rouge pour chaque épimère. **D** - Spectre 2D TOCSY de RMN du  $^1\text{H}$  de dérivés thiazolidines (2S,4R)-MeTDA et (2R,4R)-MeTDA. **E** - Spectre 2D NOESY de RMN du  $^1\text{H}$  utilisé pour l'attribution stéréospécifique des deux épimères grâce à la présence d'une résonance méthyle-H $\alpha$  possible uniquement chez l'épimère (2R,4R)-MeTDA. Dans l'autre épimère la longue distance qui sépare ces protons empêche l'apparition d'un tel pic.

### I-11.3 Stéréosélectivité du site actif de TmHydE vis-à-vis du ligand (2R,4R)-MeTDA

Dans notre préparation le rapport entre les deux épimères est de 2:8. L'épimère (2R,4R) étant le plus stable thermodynamiquement il devient majoritaire avec le temps<sup>65</sup>. Nous n'avons pas cherché à séparer les deux épimères, car dans la structure cristallographique du complexe avec HydE on ne détecte que l'épimère (2R,4R). Ainsi, de toute évidence, la protéine possède une stéréospécificité vis-à-vis du dérivé MeTDA. La fixation de l'épimère (2S,4R) est fortement défavorisée, car le méthyle du ligand pointerait vers le site de fixation d'anions S2 et le groupement carboxylate sur le C2 devrait s'accommoder dans la zone hydrophobe, mentionnée dans le paragraphe I-8.7.2. On remarque que non seulement la protéine sélectionne un seul épimère du mélange, mais elle possède aussi une affinité relativement élevée pour cette molécule. Un teste de lavage de cristaux contenant le ligand (2R,4R)-MeTDA démontre que même après 24 heures de lavage le ligand, bien que partiellement occupé, est toujours présent dans le site actif [Figure-I-57].

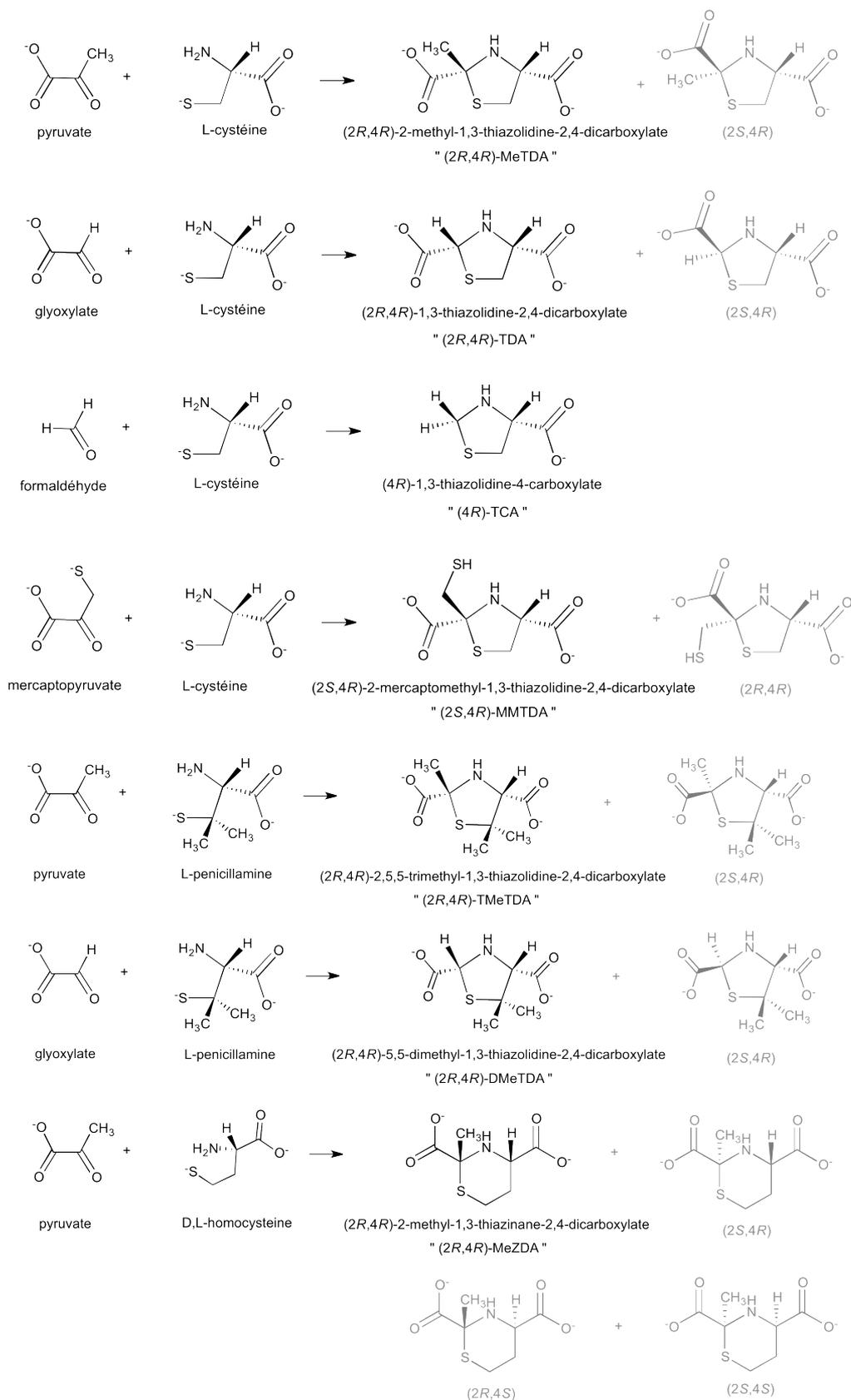


**Figure-I-57** : Résultat du lavage du ligand (2R,4R)-MeTDA du site actif de TmHydE

Vue stéréoscopique de l'environnement du site actif de TmHydE après l'incubation d'un cristal de la protéine avec le dérivé (2R,4R)-MeTDA suivie d'un lavage du cristal pendant 24 heures dans une solution de cryoprotection qui ne contenait pas de ligand. En vert foncé, carte de densité électronique  $F_o-F_c$  "omit" contournée à  $3\sigma$  (résolution 1,40 Å).

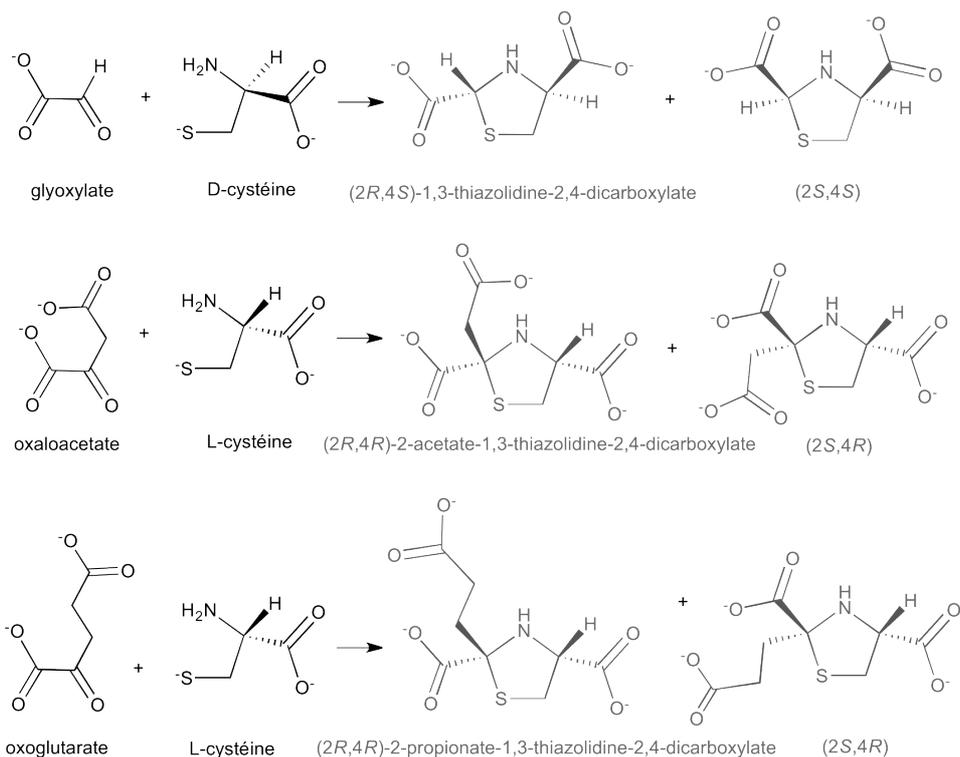
### I-11.4 Elargissement de la famille de ligands de HydE

Une bonne affinité et stéréospécificité d'une protéine pour un épimère d'une molécule sont souvent caractéristiques d'un couple enzyme-substrat ou enzyme-inhibiteur. Nous avons décidé d'utiliser le cycle thiazolidine comme un échafaudage qu'on allait fonctionnaliser pour étudier la fixation de ces dérivés dans HydE. Le but étant d'identifier de nouveaux ligands de haute affinité, parmi lesquels pourrait se trouver le substrat de HydE. Nous avons procédé à une co-cristallisation de HydE avec différents produits de cyclo-condensation de la cystéine (ou de ces analogues) avec des aldéhydes ou des cétones. La [Figure-I-58] décrit la formation de molécules dont un des épimères se fixe dans la cavité de HydE, tandis que la [Figure-I-59] contient des molécules qui ne sont pas ligands de HydE.

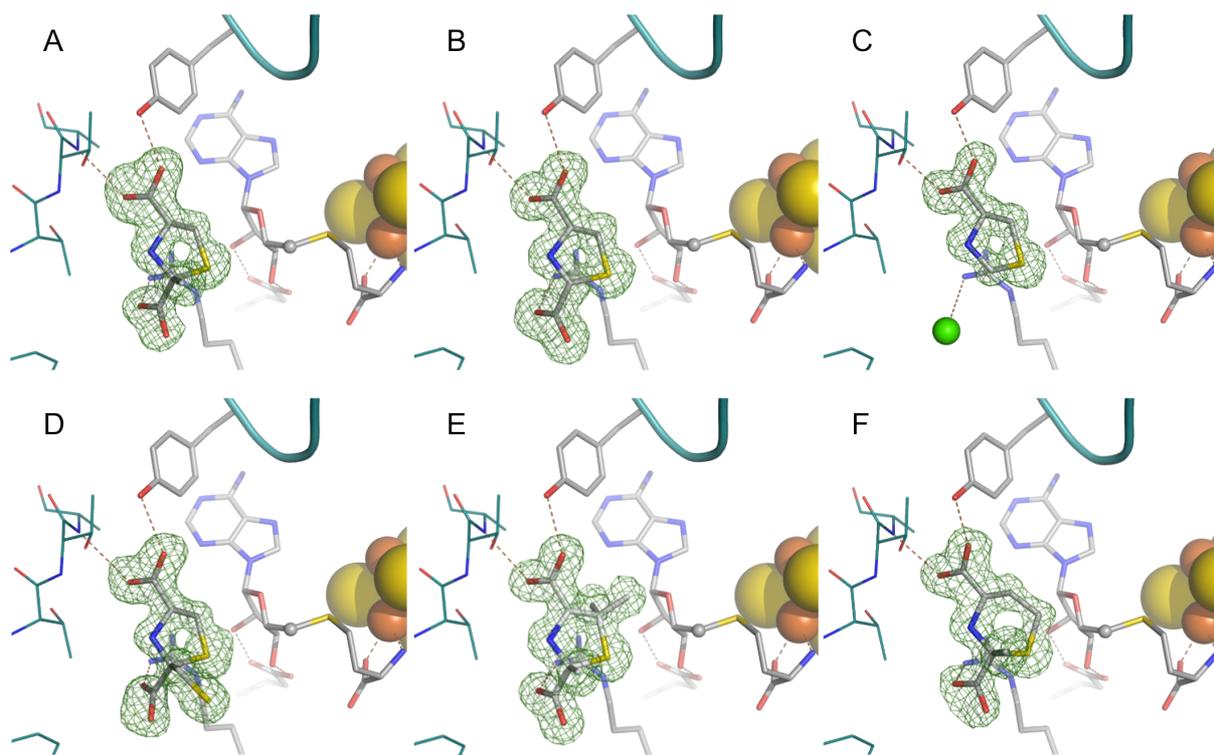


**Figure-I-58** : Schéma de réactions de formation de différents ligands de *TmHydE*

Réactions de condensation conduisant à des dérivés thiazolidines ou thiazinanes dont un des épimères (en noir) s'est avéré d'être ligand de *TmHydE*. Les épimères pour lesquels la formation du complexe n'a pas été observée sont colorés en gris.



**Figure-I-59** : Schéma de dérivés thiazolidines, qui ne sont pas ligands de *TmHydE*



**Figure-I-60** : Structures cristallographiques de complexes avec les dérivés thiazolidines ou thiazinanes

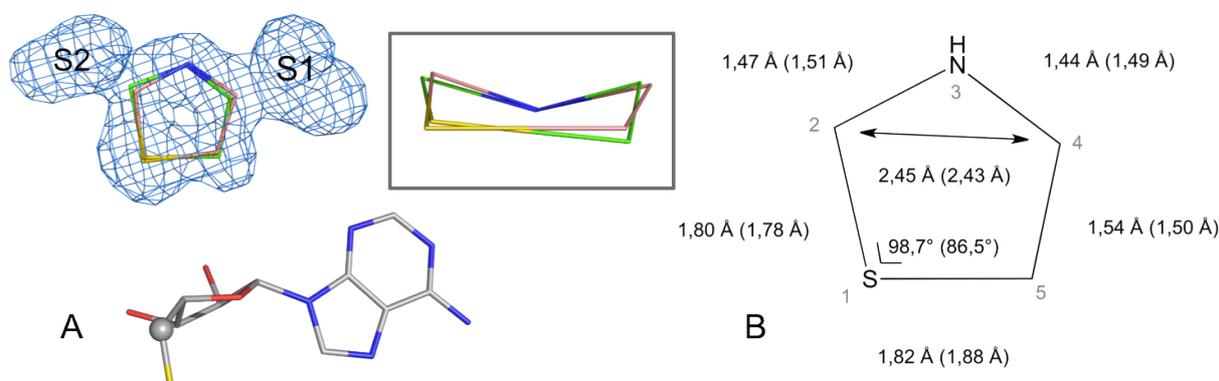
Environnement du site actif de *TmHydE* après la co-cristallisation avec des dérivés: A - (2*R*,4*R*)-MeTDA (résolution 1,15 Å) ; B - (2*R*,4*R*)-TDA (résolution 1,15 Å) ; C - (4*R*)-TCA (résolution 1,18 Å) ; D - (2*S*,4*R*)-MMTDA (résolution 1,12 Å) ; E - (2*R*,4*R*)-TMeTDA (résolution 1,35 Å) ; F - (2*R*,4*R*)-MeZDA (résolution 1,36 Å). Le dérivé (2*R*,4*R*)-1,3-thiazinane-2,4-dicarboxylate, qui ne diffère du dérivé (2*R*,4*R*)-2-méthyl-1,3-thiazinane-2,4-dicarboxylate que par l'absence d'un groupement méthyl sur le C2, n'est pas représenté sur cette figure. En vert foncé, carte de densité électronique  $F_o - F_c$  "omit" contourée à 3  $\sigma$ .

La [Figure-I-60](#) représente les structures cristallographiques de ligands dans le site actif de *TmHydE* pour lesquelles nous avons réussi à obtenir des structures de complexes.

Comme cela a été discuté dans le paragraphe précédent, notre façon de synthétiser ces dérivés produit un mélange d'epimères en position C2. Dans nos structures de complexes avec HydE on ne trouve qu'un seul epimère pour tous ces ligands. La protéine est donc stéréospécifique pour chacune de ces molécules.

#### I-11.4.1 Détails des interactions entre HydE et les ligands nouvellement identifiés

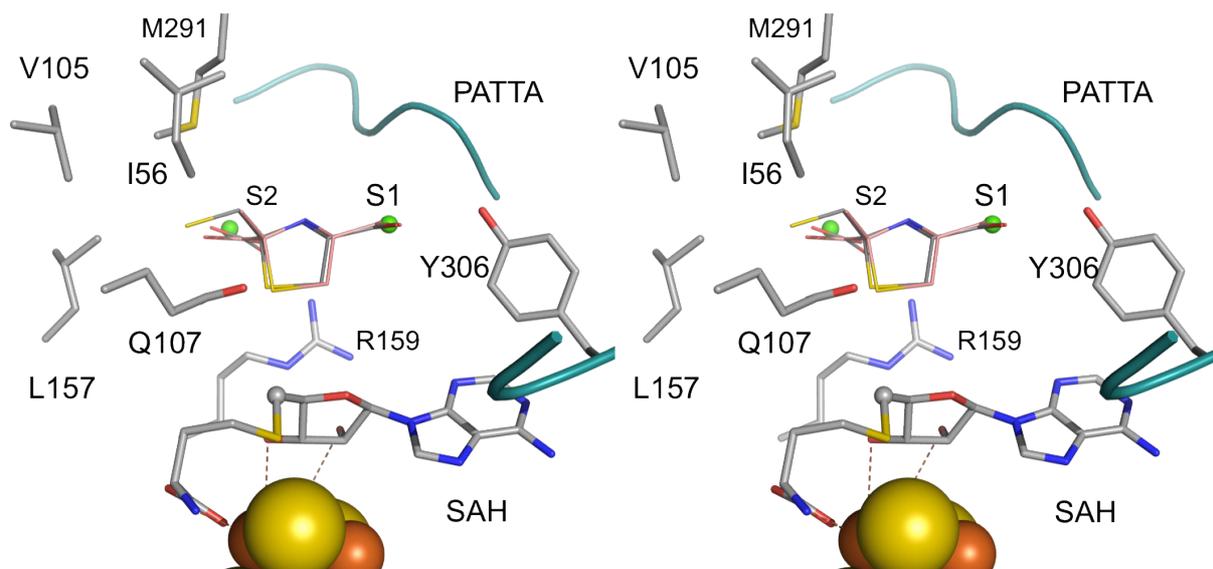
La comparaison de ces structures dévoile de nouveaux éléments sur les propriétés physico-chimiques du site actif et sur la sélection de ligand par l'enzyme. Les deux groupements carboxylates du ligand (2*R*,4*R*)-TDA, logés dans les sites de fixation d'anions S1 et S2 ont un impact sur le plissement du cycle thiazolidine [[Figure-I-61-A](#)]. En présence du second groupement carboxylate, le cycle subit une contrainte mécanique. Ceci se traduit par un changement d'angles et de distances entre les atomes du cycle thiazolidine. Le plus notable des changements est celui de l'angle entre les atomes C2-S1-C5. Il passe de 86,5° à 98,7° [[Figure-I-61-B](#)]. L'écartement entre les atomes C2 et C4, porteurs des deux groupements carboxylates en question, reste inchangé. Sur la [Figure-I-61-A](#) on peut voir que, si le second groupement carboxylate est là, ces deux atomes se décalent légèrement vers le site S2.



**Figure-I-61** : Changements du plissement du cycle thiazolidine induits par le carboxylate sur le C2

**A** - En rose, le cycle thiazolidine tel qu'il se trouve dans le ligand (4*R*)-TCA, qui ne possède pas de groupement carboxylate sur le C2. En vert, le cycle thiazolidine du ligand (2*R*,4*R*)-TDA. La partie adénosyl de la SAH est représentée en bas de l'image. En bleu, carte de densité électronique 2*Fo*-*Fc* contourée à 1  $\sigma$  correspondant au ligand (2*R*,4*R*)-TDA. Encadrée, vue orthogonale de la comparaison des cycles thiazolidine appartenant à ces deux ligands. **B** - Variation du plissement du cycle thiazolidine en présence du second carboxylate sur le C2. Les valeurs d'angles et de longueurs de liaisons pour le ligand (4*R*)-TCA sont entre parenthèses. La numérotation des atomes du cycle thiazolidine est en gris.

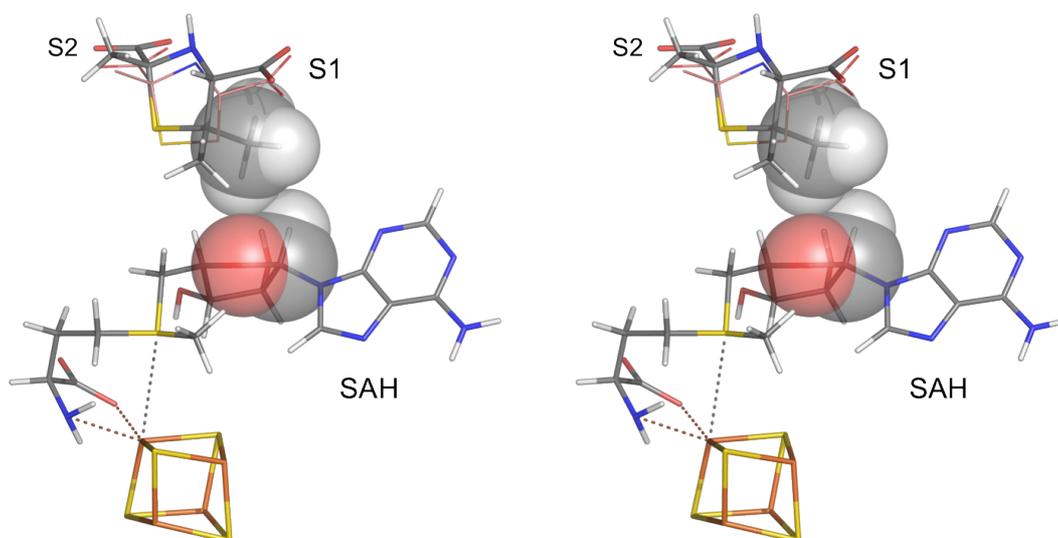
Les seules fonctionnalisations du C2 appartenant au cycle thiazolidine ou thiazinane que nous avons pu observé dans nos structures de complexes sont: un groupement carboxylate du côté du site du S2 et un groupement méthyle ou mercaptométhyle du côté de la Ile56 [[Figure-I-62](#)]. Un groupement plus volumineux, un carboxylate par exemple, ne pourrait que difficilement se trouver dans cette zone à cause de la proximité des chaînes latérales des résides Ile56 et Met291 [I-8.7.2].



**Figure-I-62** : Superposition de complexes avec les ligands (2*R*,4*R*)-MeTDA et (2*S*,4*R*)-MMTDA

Vue stéréoscopique d'une superposition de complexes avec les ligands (2*R*,4*R*)-MeTDA (en rose) et (2*S*,4*R*)-MMTDA (en gris).

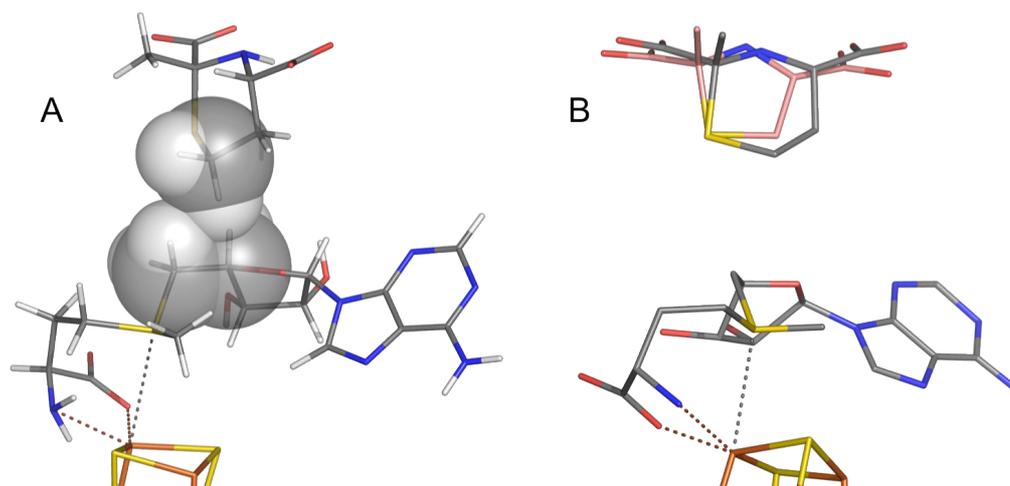
Le ligand (2*R*,4*R*)-2,5,5-triméthyle-1,3-thiazolidine-2,4-dicarboxylate ou (2*R*,4*R*)-TMeTDA a son cycle thiazolidine poussé plus loin du cofacteur SAH à cause d'un clash stérique entre les atomes O4' et C1' de ce dernier et les deux méthyles sur le C5 du ligand [Figure-I-63]. Ceci ne modifie pas de façon notable son affinité vis-à-vis du site actif par rapport au dérivé (2*R*,4*R*)-MeTDA, qui ne possède pas ces méthyles.



**Figure-I-63** : Vue stéréoscopique de l'interaction entre le ligand (2*R*,4*R*)-TMeTDA et la SAH

Le ligand (2*R*,4*R*)-TMeTDA (en haut de l'image en gris) est repoussé plus loin de la SAH que le ligand (2*R*,4*R*)-MeTDA (en rose) à cause de la présence de deux méthyles sur le C5.

Pour ce qui est du ligand (2*R*,4*R*)-2-méthyle-1,3-thiazinane-2,4-dicarboxylate ou (2*R*,4*R*)-MeZDA, on remarque qu'il se fixe de façon à placer l'atome de soufre du cycle thiazinane exactement à la même position que l'atome de soufre du cycle thiazolidine du dérivée (2*R*,4*R*)-MeTDA [Figure-I-64-B]. Ce détail est d'autant plus surprenant qu'aucun autre atome ne se retrouve dans une même position dans les deux structures. Le ligand (2*R*,4*R*)-2-méthyle-1,3-thiazinane-2,4-dicarboxylate a une affinité plus faible pour le site actif, car il part facilement lors du lavage des cristaux. Sans doute ceci est observé parce que son hétérocycle à six, plus volumineux qu'un cycle thiazolidine, crée un clash stérique au niveau du C6 avec les atomes C5' et C4' de la molécule SAH [Figure-I-64-A].



**Figure-I-64** : Détail de l'interaction entre le ligand (2*R*,4*R*)- MeZDA et la SAH dans le site actif de *TmHydE*

**A** - Contacts entre le C6 du cycle thiazinane du ligand (2*R*,4*R*)-MeZDA avec les atomes C5' et C4' de la SAH. **B** - Comparaison du positionnement du ligand (2*R*,4*R*)-MeZDA (en gris foncé) et (2*R*,4*R*)-MeTDA (en rose) dans le site actif de *TmHydE*.

#### **I-11.4.2 Profil minimum d'un ligand stéréospécifique de HydE**

Le motif commun qu'on retrouve dans les ligands thiazolidines est un cycle thiazolidine possédant un groupement carboxylate en position 4*R* (ceci est vrai également pour les dérivés thiazinanes). Ce motif est l'unité minimale de fixation d'un dérivé thiazolidine dans la cavité de HydE. Il correspond au ligand (4*R*)-TCA [Figure-I-60-C], issu de la cyclocondensation de la L-cystéine avec le formaldéhyde. Cependant nos expériences de lavage de cristaux montrent que la présence du second groupement carboxylate augmente considérablement l'affinité de la molécule pour le site actif de HydE.

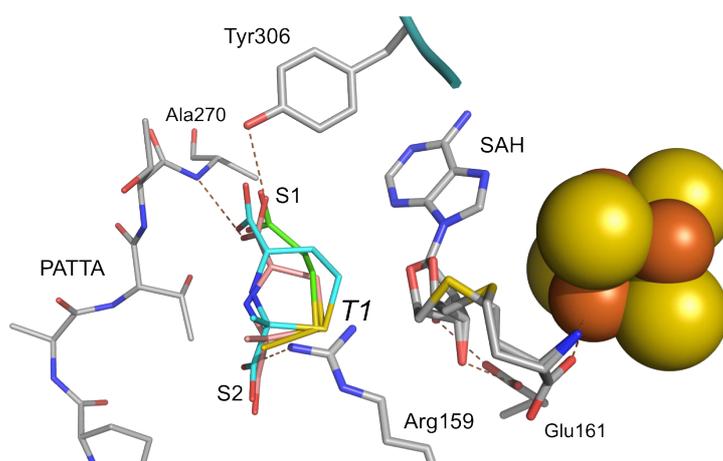
Rappelons juste que nous n'avons jamais réussi à obtenir un complexe avec la L-cystéine libre ni avec la L-proline. Le p*K*<sub>a</sub> de l'amine secondaire du dérivé (2*R*,4*R*)-TDA a été déterminé expérimentalement et se situe au alentours de 5,86<sup>73</sup>. Dans nos structures, cette amine forme une liaison hydrogène avec une molécule d'eau [Figure-I-37]. A pH physiologique cette amine est déprotonée, contrairement à l'amine primaire d'une cystéine ou d'une proline, dont le p*K*<sub>a</sub> est de 10,78 et 10,60, respectivement. Ainsi il semblerait que ce qui

contrôle la fixation d'un ligand dérivé de la L-cystéine est la transformation de son thiol en thioéther par fermeture du cycle thiazolidine et l'absence d'une charge positive sur son amine.

Aucun complexe n'a été identifié avec les dérivés produits à partir de la D-cystéine ou D-homocystéine. Cette observation souligne l'importance de la présence de l'epimère naturel de la cystéine dans la composition des ligands. La présence d'une fonction carboxylate sur le cycle thiazolidine en position 2*S* ne permettrait pas un positionnement correct de ce groupement dans le site de fixation d'anions S1 et, par la suite, du soufre du thioéther à l'endroit où on le retrouve systématiquement - en proximité du carbone C5' du cofacteur SAM ou SAH.

#### I-11.4.3 Site T1 - un site d'affinité pour le soufre

La conservation de la position de l'atome de soufre dans toutes nos structures de complexes avec des ligands soufrés suggère la présence d'un site qui a une affinité élevée pour un atome de ce type; en particulier si ce dernier est engagé dans une liaison thioéther [Figure-I-65]. Cette zone, avec une affinité élevée pour le soufre, qui se situe face au C5' de la SAH ou de la SAM sera appelée "le site T1".



**Figure-I-65** : Localisation d'un site d'affinité pour le soufre dans le site actif de *TmHydE*

Superposition de structures de complexes de *TmHydE* avec les ligands (2*R*,4*R*)-MeTDA (en rose), (2*R*,4*R*)-MeZDA (en cyan) et 3-mercaptopyrivate persulfuré (en vert). Le seul atome dont l'emplacement reste le même pour ces ligands est un soufre dans un site d'affinité pour le soufre, le site *T1*, nouvellement identifié.

## I-12 Résumé du Chapitre I

Les études présentées dans ce chapitre ont permis l'identification de plusieurs nouveaux ligands de la protéine HydE. Ces ligands possèdent tous au moins un groupement carboxylate, ce dernier se loge systématiquement dans le site de fixation d'anions S1. Ceci illustre bien le potentiel électrostatique positif de la cavité de l'enzyme, mais surtout l'affinité élevée du site S1 pour le groupement carboxylate. Parmi ces molécules nous avons identifié un groupe de ligands spécifiques de la maturase HydE, qui contiennent un cycle thiazolidine avec au moins un groupement carboxylate. Ces ligands sont formés par la cyclocondensation spontanée de la L-cystéine et d'un aldéhyde ou d'un  $\alpha$ -ketoacide de faible poids moléculaire. La protéine HydE reconnaît ces ligands de façon stéréospécifique - aucun dérivé thiazolidine fait à partir de la D-cystéine n'a été trouvé dans le site actif de l'enzyme et un seul épimère en C2 est observé fixé dans ce site. Lorsque ces ligands possèdent deux groupements carboxylates, un sur le C2 et l'autre sur le C4, ces deux groupements fixent les sites S1 et S2. Ces interactions, combinées avec la présence d'un atome de soufre engagé dans une liaison thioéther du cycle thiazolidine, augmentent de façon considérable l'affinité des ligands pour le site actif de la protéine. La comparaison entre une L-proline et le ligand (4R)-TCA suggère que cet atome de soufre est indispensable à la formation de complexes stables entre ces ligands et la protéine *TmHydE*. La particularité de l'interaction entre HydE et ces dérivés thiazolidines est que l'atome de soufre du thioéther du cycle thiazolidine se fixe dans le site de fixation de soufre T1. Ainsi, cet atome de soufre se trouve à proximité de l'atome C5' du cofacteur SAM. En fait, il devient, après un des atomes d'hydrogènes liés sur le C5 du cycle thiazolidine, le second atome du ligand le plus proche de l'atome C5'. Ce dernier va porter l'attaque radicalaire initiale sur le substrat après le clivage de la SAM.

Puisque HydE fixe les dérivés thiazolidine-2,4-dicarboxylates de façon stéréospécifique et avec une forte affinité, nous nous sommes demandé si parmi ces ligands se trouvait aussi le substrat physiologique de HydE. L'étude d'activation d'hydrogénase à [FeFe], qui devait répondre à cette question, est décrite dans le chapitre suivant.

## I-13 Références du Chapitre I

- 1 Nicolet, Y., Rohac, R., Martin, L. & Fontecilla-Camps, J. C. X-ray snapshots of possible intermediates in the time course of synthesis and degradation of protein-bound Fe<sub>4</sub>S<sub>4</sub> clusters. *Proceedings of the National Academy of Sciences* **110**, 7188-7192, doi:10.1073/pnas.1302388110 (2013).
- 2 Posewitz, M. C. *et al.* Discovery of Two Novel Radical S-Adenosylmethionine Proteins Required for the Assembly of an Active [Fe] Hydrogenase. *Journal of Biological Chemistry* **279**, 25711-25720, doi:10.1074/jbc.M403206200 (2004).
- 3 Shepard, E. M. *et al.* [FeFe]-Hydrogenase Maturation: HydG-Catalyzed Synthesis of Carbon Monoxide. *Journal of the American Chemical Society* **132**, 9247-9249, doi:10.1021/ja1012273 (2010).
- 4 Driesener, R. C. *et al.* [FeFe] - Hydrogenase Cyanide Ligands Derived From S - Adenosylmethionine - Dependent Cleavage of Tyrosine. *Angewandte Chemie International Edition* **49**, 1687-1690, doi:10.1002/anie.200907047 (2010).
- 5 Pilet, E. *et al.* The role of the maturase HydG in [FeFe]-hydrogenase active site synthesis and assembly. *FEBS Letters* **583**, 506-511, doi:10.1016/j.febslet.2009.01.004 (2008).
- 6 McGlynn, S. E. *et al.* In vitro activation of [FeFe] hydrogenase: new insights into hydrogenase maturation. *JBIC Journal of Biological Inorganic Chemistry* **12**, 443-447, doi:10.1007/s00775-007-0224-z (2007).
- 7 McGlynn, S. E. *et al.* HydF as a scaffold protein in [FeFe] hydrogenase H-cluster biosynthesis. *FEBS Letters* **582**, 2183-2187, doi:10.1016/j.febslet.2008.04.063 (2008).
- 8 Czech, I. *et al.* The [FeFe]-hydrogenase maturation protein HydF contains a H-cluster like [4Fe4S]-2Fe site. *FEBS Letters* **585**, 225-230, doi:10.1016/j.febslet.2010.11.052 (2010).
- 9 Czech, I., Silakov, A., Lubitz, W. & Happe, T. The [FeFe]-hydrogenase maturase HydF from *Clostridium acetobutylicum* contains a CO and CN- ligated iron cofactor. *FEBS letters* **584**, 638-642, doi:10.1016/j.febslet.2009.12.016 (2010).
- 10 Shepard, E. M. *et al.* Synthesis of the 2Fe subcluster of the [FeFe]-hydrogenase H cluster on the HydF scaffold. *Proceedings of the National Academy of Sciences*, doi:10.1073/pnas.1001937107 (2010).
- 11 Berggren, G. *et al.* Biomimetic assembly and activation of [FeFe]-hydrogenases. *Nature* **499**, 66-69, doi:10.1038/nature12239 (2013).
- 12 Nicolet, Y. & Fontecilla-Camps, J. C. Structure-Function Relationships in [FeFe]-Hydrogenase Active Site Maturation. *Journal of Biological Chemistry* **287**, 13532-13540, doi:10.1074/jbc.R111.310797 (2012).
- 13 Rubach, J. K., Brazzolotto, X., Gaillard, J. & Fontecave, M. Biochemical characterization of the HydE and HydG iron-only hydrogenase maturation enzymes from *Thermotoga maritima*. *FEBS Letters* **579**, 5055-5060, doi:10.1016/j.febslet.2005.07.092 (2005).
- 14 Nicolet, Y. *et al.* X-ray Structure of the [FeFe]-Hydrogenase Maturase HydE from *Thermotoga maritima*. *Journal of Biological Chemistry* **283**, 18861-18872, doi:10.1074/jbc.M801161200 (2008).
- 15 King, P. W., Posewitz, M. C., Ghirardi, M. L. & Seibert, M. Functional Studies of [FeFe] Hydrogenase Maturation in an *Escherichia coli* Biosynthetic System. *Journal of Bacteriology* **188**, 2163-2172, doi:10.1128/jb.188.6.2163-2172.2006 (2006).
- 16 Mulder, D. W. *et al.* Activation of HydAΔEFG Requires a Preformed [4Fe-4S] Cluster. *Biochemistry* **48**, 6240-6248, doi:10.1021/bi9000563 (2009).
- 17 Kuchenreuther, J. M., Stapleton, J. A. & Swartz, J. R. Tyrosine, Cysteine, and S-Adenosyl Methionine Stimulate In Vitro [FeFe] Hydrogenase Activation. *PLoS ONE* **4**, doi:10.1371/journal.pone.0007565 (2009).

- 18 Lanz, N. D. & Booker, S. J. Auxiliary iron–sulfur cofactors in radical SAM enzymes. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1853**, doi:10.1016/j.bbamcr.2015.01.002 (2015).
- 19 Berkovitch, F., Nicolet, Y., Wan, J. T., Jarrett, J. T. & Drennan, C. L. Crystal Structure of Biotin Synthase, an S-Adenosylmethionine-Dependent Radical Enzyme. *Science* **303**, 76-79, doi:10.1126/science.1088493 (2004).
- 20 Harmer, J. E. *et al.* Structures of lipoyl synthase reveal a compact active site for controlling sequential sulfur insertion reactions. *The Biochemical journal* **464**, 123-133, doi:10.1042/bj20140895 (2014).
- 21 Böck, A., King, P. W., Blokesch, M. & Posewitz, M. C. Maturation of Hydrogenases. *Advances in microbial physiology* **51**, 1-225, doi:10.1016/s0065-2911(06)51001-x (2006).
- 22 Kuchenreuther, J. M. *et al.* The HydG enzyme generates an Fe(CO)<sub>2</sub>(CN) synthon in assembly of the FeFe hydrogenase H-cluster. *Science (New York, N.Y.)* **343**, 424-427, doi:10.1126/science.1246572 (2014).
- 23 Vallese, F. *et al.* Biochemical Analysis of the Interactions between the Proteins Involved in the [FeFe]-Hydrogenase Maturation Process. *Journal of Biological Chemistry* **287**, 36544-36555, doi:10.1074/jbc.M112.388900 (2012).
- 24 Kuchenreuther, J. M., Britt, D. R. & Swartz, J. R. New Insights into [FeFe] Hydrogenase Activation and Maturase Function. *PLoS ONE* **7**, doi:10.1371/journal.pone.0045850 (2012).
- 25 Irwin, J. J. & Shoichet, B. K. ZINC--a free database of commercially available compounds for virtual screening. *Journal of chemical information and modeling* **45**, 177-182, doi:10.1021/ci049714+ (2005).
- 26 Kuntz, I. D. Structure-based strategies for drug design and discovery. *Science (New York, N.Y.)* **257**, 1078-1082 (1992).
- 27 Verlinde, C. L. & Hol, W. G. Structure-based drug design: progress, results and challenges. *Structure (London, England : 1993)* **2**, 577-587 (1994).
- 28 Nienaber, V. L. *et al.* Discovering novel ligands for macromolecules using X-ray crystallographic screening. *Nature Biotechnology* **18**, 1105-1108, doi:10.1038/80319 (2000).
- 29 Nicolet, Y. *et al.* Crystal structure of HydG from Carboxydotherrmus hydrogenoformans: a trifunctional [FeFe]-hydrogenase maturase. *Chembiochem : a European journal of chemical biology* **16**, 397-402, doi:10.1002/cbic.201402661 (2015).
- 30 Dinis, P. *et al.* X-ray crystallographic and EPR spectroscopic analysis of HydG, a maturase in [FeFe]-hydrogenase H-cluster assembly. *Proceedings of the National Academy of Sciences* **112**, 1362-1367, doi:10.1073/pnas.1417252112 (2015).
- 31 Hermann, J. C. *et al.* Structure-based activity prediction for an enzyme of unknown function. *Nature* **448**, 775-779, doi:10.1038/nature05981 (2007).
- 32 Nicolet, Y. & Drennan, C. L. AdoMet radical proteins—from structure to evolution—alignment of divergent protein sequences reveals strong secondary structure element conservation. *Nucleic Acids Research* **32**, 4015-4025, doi:10.1093/nar/gkh728 (2004).
- 33 Qwitterer, F., List, A., Eisenreich, W., Bacher, A. & Groll, M. Crystal Structure of Methylornithine Synthase (PylB): Insights into the Pyrrolysine Biosynthesis. *Angewandte Chemie International Edition* **51**, 1339-1342, doi:10.1002/anie.201106765 (2012).
- 34 Nicolet, Y., Zeppieri, L., Amara, P. & Fontecilla - Camps, J. C. Crystal Structure of Tryptophan Lyase (NosL): Evidence for Radical Formation at the Amino Group of Tryptophan. *Angewandte Chemie International Edition* **53**, 11840-11844, doi:10.1002/anie.201407320 (2014).
- 35 Fenwick, M. K. *et al.* Non-canonical active site architecture of the radical SAM thiamin pyrimidine synthase. *Nature Communications* **6**, 6480, doi:10.1038/ncomms7480 (2015).
- 36 Horitani, M. *et al.* Why Nature Uses Radical SAM Enzymes so Widely: Electron Nuclear Double Resonance Studies of Lysine 2,3-Aminomutase Show the 5' -dAdo· “Free Radical” Is Never Free. *Journal of the American Chemical Society*, 2022414680, doi:10.1021/jacs.5b00498 (2015).

- 37 Michal, G. & Schomburg, D. Biological Pathways: An atlas of biochemistry and molecular biology. *Biological Pathways: An atlas of biochemistry and molecular biology*, doi:10.1007/s002130050786 (2012).
- 38 Kriek, M., Martins, F., Challand, M. R., Croft, A. & Roach, P. L. Thiamine Biosynthesis in *Escherichia coli*: Identification of the Intermediate and By - Product Derived from Tyrosine. *Angewandte Chemie International Edition* **46**, 9223-9226, doi:10.1002/anie.200702554 (2007).
- 39 Philmus, B., Decamps, L., Berteau, O. & Begley, T. P. Biosynthetic Versatility and Coordinated Action of 5' -Deoxyadenosyl Radicals in Deazaflavin Biosynthesis. *Journal of the American Chemical Society* **137**, 1740106492, doi:10.1021/ja513287k (2015).
- 40 Maveyraud, L. *et al.* Insights into class D beta-lactamases are revealed by the crystal structure of the OXA10 enzyme from *Pseudomonas aeruginosa*. *Structure (London, England : 1993)* **8**, 1289-1298, doi:10.1016/s0969-2126(00)00534-7 (2000).
- 41 Carta, R. & Tola, G. Solubilities of L-cystine, L-tyrosine, L-leucine, and glycine in aqueous solutions at various pHs and NaCl concentrations. *Solubilities of L-cystine, L-tyrosine, L-leucine, and glycine in aqueous solutions at various pHs and NaCl concentrations*, doi:10.1021/je9501853 (1996).
- 42 Kredich, N. M., Foote, L. J. & Keenan, B. S. The stoichiometry and kinetics of the inducible cysteine desulfhydrase from *Salmonella typhimurium*. *The Journal of biological chemistry* **248**, 6187-6196 (1973).
- 43 Pagnier, A., Nicolet, Y. & Fontecilla-Camps, J. C. IscS from *Archaeoglobus fulgidus* has no desulfurase activity but may provide a cysteine ligand for [Fe<sub>2</sub>S<sub>2</sub>] cluster assembly. *Biochimica et biophysica acta* **1853**, 1457-1463, doi:10.1016/j.bbamcr.2014.10.015 (2015).
- 44 Nicolet, Y., Piras, C., Legrand, P., Hatchikian, C. E. & Fontecilla-Camps, J. C. Desulfovibrio desulfuricans iron hydrogenase: the structure shows unusual coordination to an active site Fe binuclear center. *Structure* **7**, doi:10.1016/s0969-2126(99)80005-7 (1998).
- 45 Challand, M. R. *et al.* Product inhibition in the radical S-adenosylmethionine family. *FEBS letters* **583**, 1358-1362, doi:10.1016/j.febslet.2009.03.044 (2009).
- 46 Robbins, F. M. & Fioriti, J. A. Alkaline Degradation of Cystine, Glutathione and Sulphur-containing Proteins. *Nature* **200** (1963).
- 47 Clarke, H. T. & Inouye, J. M. Some observations on the action of alkali upon cystine and cysteine. *Journal of Biological Chemistry* (1930).
- 48 Schubert, M. P. Compounds of Thiol acids with Aldehydes. *JBC* (1936).
- 49 Kallen, R. G. The mechanism of reactions involving Schiff base intermediates. Thiazolidine formation from L-cysteine and formaldehyde. *Journal of the American Chemical Society* **93**, 6236-6248, doi:10.1021/ja00752a040 (1971).
- 50 Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research* **28**, 27-30, doi:10.1093/nar/28.1.27 (2000).
- 51 Boyer, M. E., Stapleton, J. A., Kuchenreuther, J. M., Wang, C. w. & Swartz, J. R. Cell - free synthesis and maturation of [FeFe] hydrogenases. *Biotechnology and Bioengineering* **99**, 59-67, doi:10.1002/bit.21511 (2008).
- 52 Stapleton, J. A. & Swartz, J. R. Development of an In Vitro Compartmentalization Screen for High-Throughput Directed Evolution of [FeFe] Hydrogenases. *PLoS ONE* **5**, doi:10.1371/journal.pone.0015275 (2010).
- 53 Kuchenreuther, J. M. *et al.* High-Yield Expression of Heterologous [FeFe] Hydrogenases in *Escherichia coli*. *PLoS ONE* **5**, doi:10.1371/journal.pone.0015491 (2010).
- 54 Kuchenreuther, J. M., George, S. J., Grady-Smith, C. S., Cramer, S. P. & Swartz, J. R. Cell-free H-cluster Synthesis and [FeFe] Hydrogenase Activation: All Five CO and CN- Ligands Derive from Tyrosine. *PLoS ONE* **6**, doi:10.1371/journal.pone.0020346 (2011).
- 55 Kuchenreuther, J. M. ENGINEERING IN VIVO AND IN VITRO SYSTEMS FOR EXPRESSION AND ACTIVATION OF [FeFe] HYDROGENASES. *ENGINEERING IN VIVO AND IN VITRO SYSTEMS FOR EXPRESSION AND ACTIVATION OF [FeFe] HYDROGENASES* (2010).

- 56 Nakamura, M., Saeki, K. & Takahashi, Y. Hyperproduction of recombinant ferredoxins in *Escherichia coli* by coexpression of the ORF1-ORF2-iscS-iscU-iscA-hscB-hscA-fdx-ORF3 gene cluster. *Journal of Biochemistry* (1999).
- 57 Cicchillo, R. M. *et al.* *Escherichia coli* lipoyl synthase binds two distinct [4Fe-4S] clusters per polypeptide. *Biochemistry* **43**, 11770-11781, doi:10.1021/bi0488505 (2004).
- 58 Hänzelmann, P. & Schindelin, H. Crystal structure of the S-adenosylmethionine-dependent enzyme MoaA and its implications for molybdenum cofactor deficiency in humans. *Proceedings of the National ...*, doi:10.1073/pnas.0404624101 (2004).
- 59 Hänzelmann, P., Hernández, H. L. & Menzel, C. Characterization of MOCS1A, an oxygen-sensitive iron-sulfur protein involved in human molybdenum cofactor biosynthesis. *Journal of Biological ...*, doi:10.1074/jbc.M313398200 (2004).
- 60 Desiderio, C. Evaluation of chemical and diastereoisomeric stability of S-adenosylmethionine in aqueous solution by capillary electrophoresis. *Journal of Pharmaceutical and Biomedical Analysis*.
- 61 Parks, L. W. & Schlenk, F. The stability and hydrolysis of S-adenosyl-methionine; isolation of S-ribosylmethionine. *Journal of Biological Chemistry* (1958).
- 62 Cantoni, G. L. S-ADENOSYLMETHIONINE; A NEW INTERMEDIATE FORMED ENZYMATIALLY FROM L-METHIONINE AND ADENOSINETRIPHOSPHATE. *Journal of Biological Chemistry* (1953).
- 63 Betz, J. N. *et al.* [FeFe]-Hydrogenase Maturation: Insights into the Role HydE Plays in Dithiomethylamine Biosynthesis. *Biochemistry*, 1574462297, doi:10.1021/bi501205e (2015).
- 64 Refouvelet, B., Robert, J.-F., Couquelet, J. & Tronche, P. Thiazolidine-2,4-dicarboxylic acid and its esters: Synthesis, in solution behaviour and regioselective cyclocondensation. *Journal of Heterocyclic Chemistry* **31** (1994).
- 65 Shiraiwa, T. *et al.* Preparation of (2S,4R) - 2,4 - thiazolidinedicarboxylic acid by separation of diastereoisomeric salts with organic amines. *Chirality* **11**, 326-329, doi:10.1002/(sici)1520-636x(1999)11:4<326::aid-chir11>3.0.co;2-2 (1999).
- 66 Shiraiwa, T., Katayama, T., Kaito, T., Tanigawa, H. & Kurokawa, H. Facile Preparation of Four Stereoisomers of 2,4-Thiazolidinedicarboxylic Acid. *Bulletin of the Chemical Society of Japan* **71**, doi:10.1246/bcsj.71.1911 (1998).
- 67 Nagasawa, H. T., Goon, D. J. W. & Shirota, F. N. Epimerization at C-2 of 2-substituted thiazolidine-4-carboxylic acids. *Journal of Heterocyclic Chemistry* **18**, 1047-1051, doi:10.1002/jhet.5570180538 (1981).
- 68 Chiarino, D., Ferrario, F., Pellacini, F. & Sala, A. Synthesis and in Solution Behaviour of New 2-Substituted-4-thiazolidinescarboxylic Acid Derivatives. *Journal of Heterocyclic Chemistry* **26** (1989).
- 69 Baker, P. W., Bais, R. & Rofe, A. M. Formation of the l-cysteine-glyoxylate adduct is the mechanism by which l-cysteine decreases oxalate production from glycollate in rat hepatocytes. *Biochemical Journal* **302**, 753-757, doi:10.1042/bj3020753 (1994).
- 70 Bringmann, G. *et al.* A "chemical" concept for the therapy of glyoxylate-induced oxalurias (1). *Life Sciences* **50**, 1597-1605, doi:10.1016/0024-3205(92)90445-u (1992).
- 71 Wagner, G. & Wüthrich, K. Truncated driven nuclear overhauser effect (TOE). A new technique for studies of selective  $^1\text{H}$   $\rightarrow$   $^1\text{H}$  overhauser effects in the presence of spin diffusion. *Journal of Magnetic Resonance (1969)* **33**, doi:10.1016/0022-2364(79)90180-x (1979).
- 72 Kumar, A., Ernst, R. R. & Wüthrich, K. A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochemical and biophysical research communications* **95**, 1-6 (1980).
- 73 Butvin, P., MitkovÁ, M., Sv TlÍK, J. & HavrÁNek, E. Complex Formation of Thiazolidine-2,4-dicarboxylic Acid with Selected Divalent and Trivalent Metal Ions. *Chem.Pap.* (2002).

## Chapitre II : Etude fonctionnelle de HydE par l'activation d'hydrogénase dans l'extrait cellulaire

### II-1 Introduction du chapitre

Avec l'accumulation de nouveaux ligands identifiés par l'accostage moléculaire *in silico* et par la cristallographie au rayons X, il devenait nécessaire de se doter d'un test d'activité enzymatique de HydE fiable. Puisque les mesures de clivage de la SAM et de production de la 5'-déoxyadénosine peuvent conduire à des conclusions erronées [voir paragraphe III-2.1.2], nous avons décidé d'utiliser un système de maturation d'hydrogénase à [FeFe] inspiré des travaux de James R. Swartz<sup>24,51-54</sup>. Ce système est composé de l'hydrogénase HydA1 N-strep de *Chlamydomonas reinhardtii* purifiée, de trois extraits cellulaires d' *E. coli* clarifiés, contenant chacun une des trois maturases de *Clostridium acetobutylicum* surexprimée et d'une série de petites molécules exogènes (DTT, DTH, Fe<sup>2+</sup>, S<sup>2-</sup>, L-Cys, L-Tyr, GTP, PLP, SAM) [voir MM-3.6.3]. Dans tous les extraits cellulaires la SAM synthase MetK et la machinerie ISC ont été également surexprimées. En ajoutant à ce système nos ligands un par un, nous avons pu quantifier leur l'impacts sur l'activation d'hydrogénase et par extrapolation sur l'activité de ses maturases.

L'activité spécifique de l'hydrogénase a été mesurée par un essai colorimétrique de consommation de dihydrogène en présence de méthylviologène oxydé [voir MM-3.6.4]. Dans un premier temps nous avons utilisé des extraits d' *E. coli* clarifiés contenant ou pas des maturases HydE, HydF et HydG. Plus tard, nous avons utilisé sur des extraits clarifiés dialysés [voir MM-3.6.2.3], ceci dans le but de mieux contrôler la concentration des petites molécules lors de la maturation de l'hydrogénase.

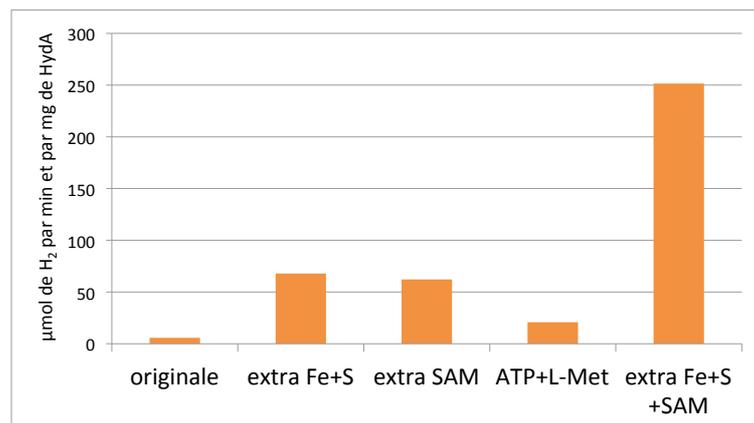
### II-2 Optimisation des conditions de maturation d'hydrogénase CrHydA1

L'avantage d'utiliser l'hydrogénase à [FeFe] de *Chlamydomonas reinhardtii* est que celle-ci ne contient pas les centres [Fe-S] auxiliaires et donc le seul centre à assembler est l'agrégat H. Le dosage du contenu en atomes de Fe de CrHydA1 N-strep avant son activation a montré entre 3,3 et 3,5 atomes de Fe par molécule de l'enzyme purifiée. L'hydrogénase purifiée possède donc un centre [Fe<sub>4</sub>S<sub>4</sub>] prêt à accueillir le sous-agrégat à deux atomes de Fe, comme rapporté dans la littérature<sup>74</sup>.

La première expérience de maturation a produit de l'hydrogénase active, cependant l'activité spécifique de l'enzyme était très faible - de l'ordre de 6 μmol de H<sub>2</sub> consommé par minute et par mg de l'enzyme. L'hydrogénase HydA1 isolée à partir de *C. reinhardtii* montre un taux d'évolution de H<sub>2</sub> de l'ordre de 650–950 μmol·min<sup>-1</sup>·mg<sup>-1</sup><sup>75,76</sup>. Nous avons donc cherché à optimiser notre système de maturation en augmentant les concentrations de molécules qui ont été identifiées comme ayant un effet bénéfique sur l'activation de l'hydrogénase. Une augmentation de 50% de la concentration finale en SAM ou Fe<sup>2+</sup> et S<sup>2-</sup> a conduit à une augmentation d'activité hydrogénase de façon à obtenir plus de 60 μmol·min<sup>-1</sup>

$1 \cdot \text{mg}^{-1}$  [Figure-II-66]. Parallèlement, l'ajout de 2 mM d'ATP et de L-méthionine a fait tripler l'activité hydrogénase par rapport à la condition initiale qui était à optimiser. Ces molécules sont les substrats de la SAM synthase, qui est surexprimée dans notre système [voir MM-1.1.3]. Ainsi la présence de ces molécules augmente certainement la concentration de la SAM dans notre mélange de maturation. Suite à ces résultats nous avons augmenté les concentrations de SAM, de  $\text{Fe}^{2+}$  et de  $\text{S}^{2-}$  dans le protocole standard de 50%. Dans cette condition nous avons observé un effet de synergie, car nous sommes arrivés à des activités spécifiques de l'ordre de  $250 \mu\text{mol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$  au bout de 18 heures de maturation. Ceci reste faible par rapport à l'activité spécifique de l'enzyme sauvage. Par contre d'après des données non publiées de Jon M. Kuchenreuther, la présence d'une étiquette strep du côté N-terminale de l'hydrogénase CrHydA1 fait baisser son activité spécifique d'à peu près de 50%. On peut également noter que les mesures d'activités hydrogénase qui figurent dans la littérature ont été faites à  $37^\circ\text{C}$ , tandis que nos mesures ont été réalisées, pour des raisons pratiques, à  $25^\circ\text{C}$ .

Il est important de souligner que l'activité spécifique de l'échantillon standard peut varier fortement d'une expérience à l'autre. Swartz et ses collaborateurs ont eux aussi observé cette variabilité élevée<sup>55</sup>. Sa cause a été attribuée à l'utilisation d'extraits cellulaires, qui ont subi plusieurs cycles de congélation et de décongélation.



**Figure-II-66** : L'optimisation de l'activation de l'hydrogénase CrHydA1 *in vitro*

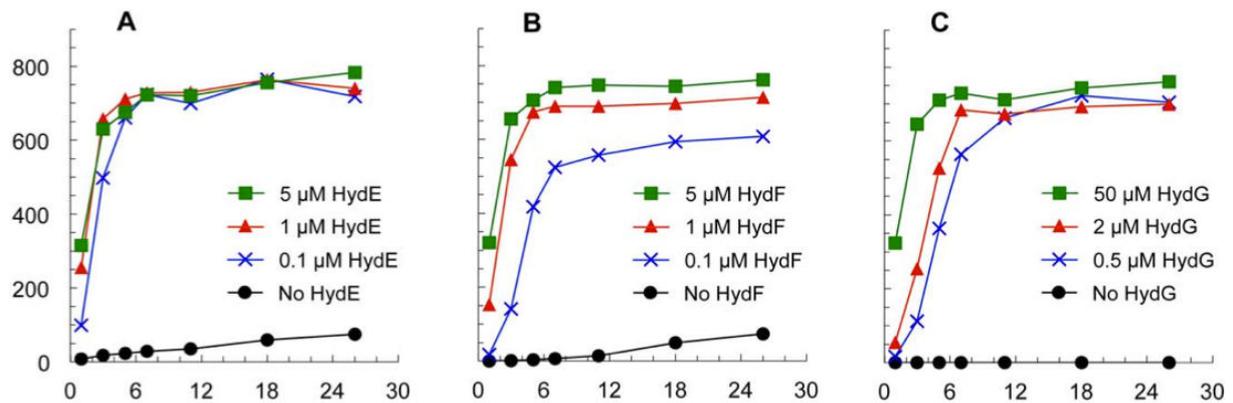
L'augmentation de l'activation anaérobie de l'hydrogénase CrHydA1 en présence de molécules exogènes dont les concentrations ont été augmentées de 50 % par rapport au protocole standard (condition "originale"), nombre d'expériences n=1.

## II-3 Importance de la maturase HydE dans le processus d'activation de l'hydrogénase

### II-3.1 Nouveau modèle de maturation proposé par l'équipe de James R. Swartz

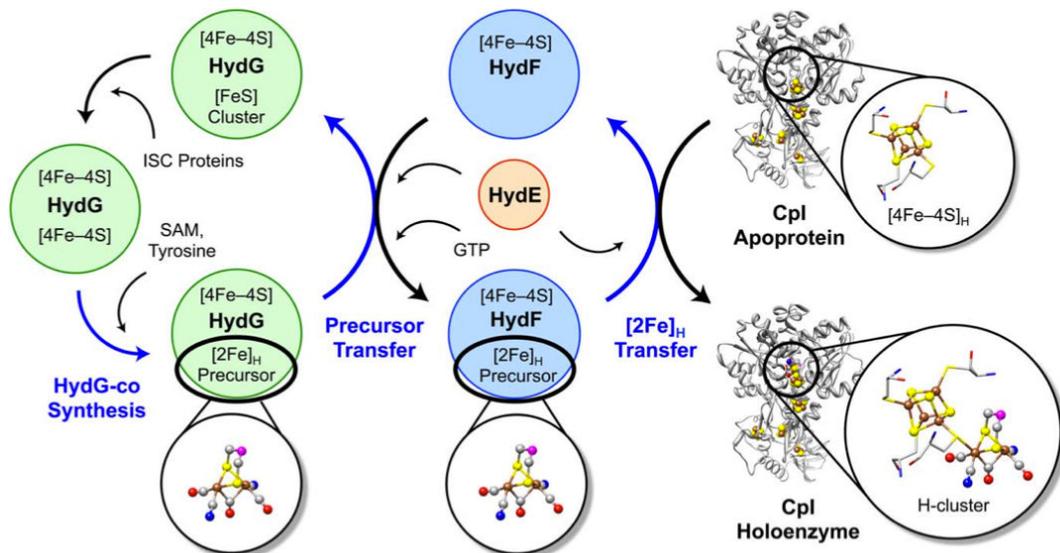
En 2012 l'équipe de James R. Swartz avait publié un article qui défiait le modèle de maturation de l'hydrogénase généralement admis<sup>24</sup>. D'après leurs données la maturation de l'enzyme en absence de HydE ou de HydF produisait une hydrogénase dont l'activité spécifique représentait environ 9% de l'activité maximale [Figure-II-67]. Ils en ont conclu

que ni HydE ni HydF n'étaient indispensables à l'activation de l'hydrogénase. Le rôle de HydF en tant qu'une protéine chaperonne ne peut être que difficilement mis en doute<sup>7-10</sup>. Les auteurs ont cependant proposé un nouveau mécanisme dans lequel HydG fabriquerait un précurseur du sous-agrégat à de Fe complet, qui était ensuite transféré à HydF avant d'être acheminé dans le site actif de l'hydrogénase [Figure-II-68]. Le rôle de HydE dans ce mécanisme n'est pas clair, car les auteurs disent qu'elle pourrait assister HydF pour assurer le transfert du sous-agrégat vers l'hydrogénase.



**Figure-II-67** : Effets de variation de concentration des maturases HydE, HydF ou HydG sur l'activation de l'hydrogénase CrHydA1 *in vitro* observés par Kuchenreuther et collègues

Activation *in vitro* d'hydrogénase CrHydA1 en présence de lysat anaérobie d'*E. coli*, d'une série de petites molécules exogènes et de maturases purifiées de concentrations variables (n=3). (Images adaptées de Kuchenreuther et al. PLoS ONE 2012) A - 5 μM de HydF, 50 μM de HydG et variable concentration de HydE. B - 5 μM de HydE, 50 μM de HydG et variable concentration de HydF. C - 5 μM de HydE, 5 μM de HydF et variable concentration de HydG.



**Figure-II-68** : Mécanisme d'activation d'hydrogénase à [FeFe] proposé par Kuchenreuther et collègues

D'après les auteurs, HydG synthétiserait un précurseur du sous-agrégat à deux Fe à partir de son second centre [Fe-S] et de la L-tyrosine. Les auteurs proposent que cette dernière pourrait même servir de précurseur pour la production du ligand DTMN. HydE utiliserait ensuite des électrons issus de la chimie radicalaire pour faciliter le transfert du précurseur du sous-agrégat à deux Fe vers HydF. Le second centre [Fe-S] de HydG devrait être reconstitué par la machinerie Isc après chaque cycle de maturation.

Un tel mécanisme nous paraît peu probable pour plusieurs raisons:

i.) La conservation des trois maturases dans les séquences génomiques des organismes qui possèdent une hydrogénase à [FeFe], combinée avec l'utilisation des mutants d'insertion au niveau du gène de *hydE*<sup>2</sup>, ainsi que les expériences de coexpression des trois maturases avec l'hydrogénase montrent clairement que les protéines sont nécessaires à l'activation de l'enzyme *in vivo*<sup>6,14,51</sup>. Paul W. King a montré que si on ne co-exprime que deux des trois maturases avec l'hydrogénase, l'activité hydrogénase d'extraits cellulaires passe de 100% à 0.4% - ce qui correspondait à l'activité basale de l'extrait<sup>15</sup>.

ii.) Les protéines à radical SAM sont connues pour être impliquées dans des réactions de réarrangements plus ou moins complexes impliquant, entre autres, des coupures ou créations de liaisons C-C, C-N, C-P et C-S<sup>77</sup> [voir paragraphe IB-3 et IB-5]. A ce jour, aucune protéine à radical SAM n'a été décrite comme ayant pour seul rôle de transférer un cofacteur d'une protéine à une autre. Nos données structurales vont à l'encontre d'un tel rôle pour HydE [voir Chapitre III].

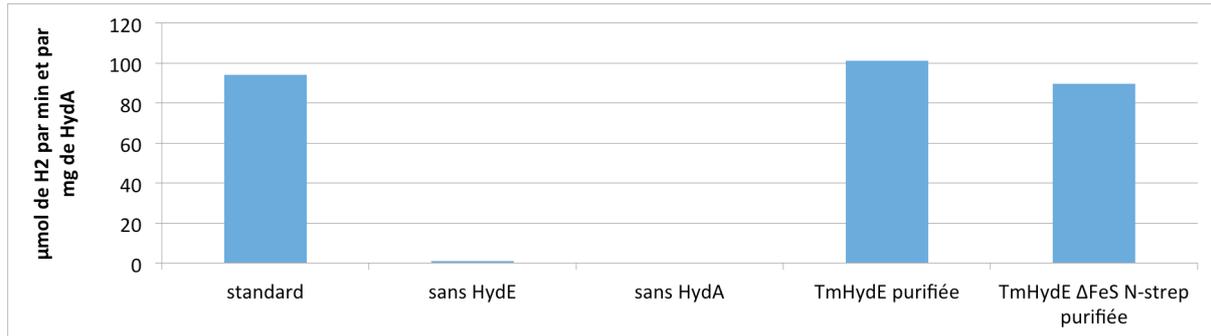
iii.) D'après le nouveau modèle de Kuchenreuther le ligand dithiométhylamine serait fabriqué par une protéine inconnue et inséré directement dans HydG. Cependant la structure de HydG avec ses deux centres [Fe-S] suggère que la protéine synthétiserait un complexe à un atome de Fe chélaté par des ligands organiques<sup>30</sup> (pdb: 4WCX) [voir paragraphe IB-3.1].

iv.) Les avancées au niveau de la réaction enzymatique catalysée par la maturase HydG montrent que lors des expériences *in vitro* cette enzyme libère en solution des synthons Fe-(CO)<sub>2</sub>-(CN)<sup>-</sup> fait à partir du cinquième atome de Fe de l'agrégat auxiliaire. Le centre [Fe<sub>4</sub>S<sub>4</sub>] auxiliaire ne semble pas se convertir en une autre forme<sup>78-81</sup> [voir paragraphe IB-4 et Conclusions du chapitre].

v.) Pour finir, rappelons qu'aucune autre protéine spécifique des organismes possédant une hydrogénase à [FeFe] qui pourrait être à l'origine du ligand azadithiolate n'a été identifiée. Il est également improbable qu'un tel ligand puisse être synthétisé par une des protéines d' *E. coli*.

### II-3.2 Essais de maturation d'hydrogénase sans et avec HydE

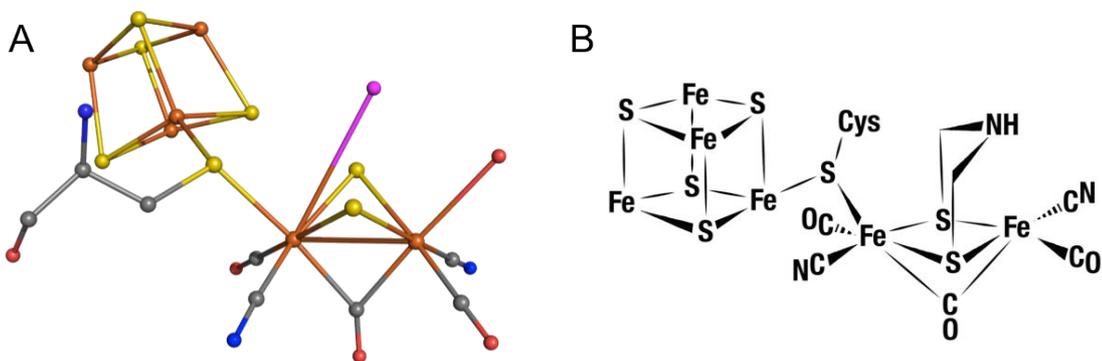
De notre côté nous avons pu constater que *CaHydE* peut être remplacée par *TmHydE* purifiée ou encore par *TmHydE* ΔFeS N-strep purifiée [Figure-II-69]. Ceci démontre que l'agrégat auxiliaire [Fe-S] n'est pas nécessaire à la maturation de l'hydrogénase et que le mécanisme de maturation est un processus conservé parmi les espèces possédant une hydrogénase à [FeFe] [I-3.1]. Par contre, en excluant cette maturase du milieu réactionnel l'activité hydrogénase spécifique de l'échantillon passe de 100% à 1%, tandis que l'activité basale dans nos expériences et dans celles de Kuchenreuther est de 0. Nous sommes loin des 9% observés par Kuchenreuther, nous observons également une très faible activité hydrogénase en absence de HydE.



**Figure-II-69** : Activation *in vitro* de hydrogénase à [FeFe] en présence ou en absence de la maturase HydE

Contrairement à Kuchenreuther *et al.*, nous pensons que HydE est indispensable à la maturation de l'agrégat H. Cependant nous pensons aussi, que, sans cette maturase et en présence d'un fort excès de sulfures, de fer et de HydG, un agrégat incomplet pourrait se former. Il serait composé d'un centre [Fe<sub>2</sub>S<sub>2</sub>] portant des ligands diatomiques CO et CN<sup>-</sup>. Cet agrégat ressemblerait à celui modélisé dans la première structure d'une hydrogénase à [FeFe]<sup>82</sup>, publié en 1998. Au lieu du ligand DTMN, John W. Peters *et al.* ont modélisé une molécule d'eau ou un ammoniac qui occupe la place de l'atome d'azote du ligand azadithiolate [Figure-II-70-A&B]. Bien que des travaux postérieurs sur la même enzyme ont montré que ce modèle était erroné<sup>83</sup>, il est possible qu'une telle structure puisse être assemblée en absence de HydE. La molécule d'ammoniaque ou d'hydroxyle pourrait agir comme une base dans le mécanisme du clivage hétérolytique du dihydrogène. Il est connu que la maturase HydF, dotée de l'agrégat H, possède une faible activité hydrogénase<sup>54</sup>. Il est toujours possible qu'à l'image de la maturase qui abrite le cofacteur complet, mais entouré par un environnement protéique différent, l'hydrogénase, possédant un cofacteur incomplet entouré d'un environnement protéique adapté, pourrait elle aussi présenter une activité hydrogénase résiduelle.

En se basant sur nos connaissances de la structure de HydE, sur les généralités sur les protéines à radical SAM et sur les expériences de maturation de l'hydrogénase, nous pensons que la maturase HydE est indispensable à la biosynthèse d'un agrégat H complet et donc à l'activation post-translationnelle de l'hydrogénase à [FeFe].



**Figure-II-70** : Modèles de l'agrégat H de l'hydrogénase à [FeFe]

**A** - Première structure cristallographique de l'agrégat H de l'hydrogénase à [FeFe] proposée par Peters *et al.* en 1998. **B** - Schéma de l'agrégat H sous sa forme active appelée "H<sub>ox</sub>" proposé par Nicolet *et al.* en 2001<sup>84</sup>.

## II-4 Effet de la cystéine, des thiazolidines et de $\alpha$ -ketoacides sur la maturation d'hydrogénase

Le système de maturation d'hydrogénase avec des extraits cellulaires clarifiés a ensuite été utilisé pour étudier l'impact de certains dérivés thiazolidines - ligands de *TmHydE* - identifiés par les expériences de co-cristallisation. Parmi ces molécules il y a eu des produits naturels faisant partie de la voie métabolique de la L-cystéine, comme par exemple le (2*R*,4*R*)-MeTDA<sup>85,86</sup> ou encore le (2*R*,4*R*)-TDA<sup>69</sup>. Il était donc possible qu'une de ces molécules soit le substrat naturel de HydE. Rappelons juste que, dans toutes ces expériences, les deux épimères sur le carbone C2 du dérivé thiazolidine en question sont formés. Puisque la L-cystéine est l'unité commune de tous les dérivés thiazolidines nous avons voulu, dans un premier temps, vérifier son impact sur la maturation de l'hydrogénase.

### II-4.1 Effet de la cystéine

L'exclusion de la cystéine du milieu réactionnel engendre une forte baisse d'activité spécifique de l'hydrogénase, qui dans cette condition n'atteint que 7% de l'activité de l'échantillon standard [Figure-II-71]. La cystéine est connue pour augmenter la maturation. Cependant, comme discuté dans le paragraphe I-8, il n'est pas clair si cet effet est dû à l'augmentation de la concentration de sulfures et donc à une meilleure reconstitution des centre [Fe-S] ou si la cystéine entre dans la voie métabolique de biosynthèse du substrat de HydE. Dans le but d'éclaircir ce point nous avons mesuré les activités hydrogénase d'un échantillon qui contenait 4 mM de L-cystéine (échantillon "extra L-Cys") et d'un autre avec 2 mM de cystéine et 2 mM supplémentaires de sulfures (échantillon "extra S"). L'activité du premier ne différait pas de celle de l'échantillon standard, qui contient systématiquement 2 mM de cystéine. Par contre l'échantillon enrichi en sulfure présentait une augmentation d'activité de plus de 90% [Figure-II-71]. Puisque les extraits cellulaires utilisés pour ces tests contiennent tous la protéine IscS surexprimée, il semblerait que la cystéine est impliquée dans un processus bénéfique pour la maturation autre que la production de sulfures (voir plus loin dans le texte).

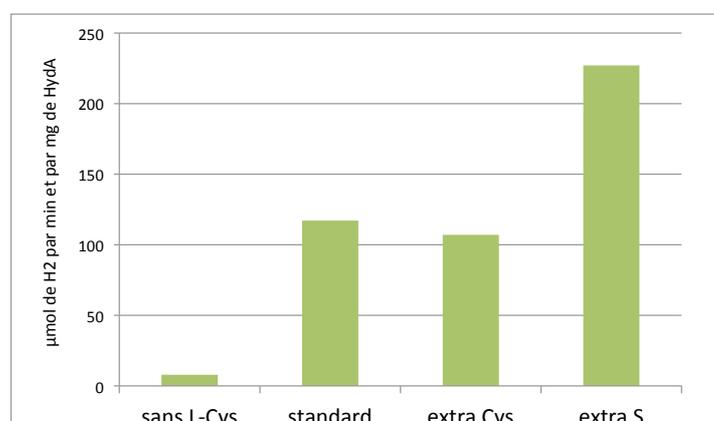
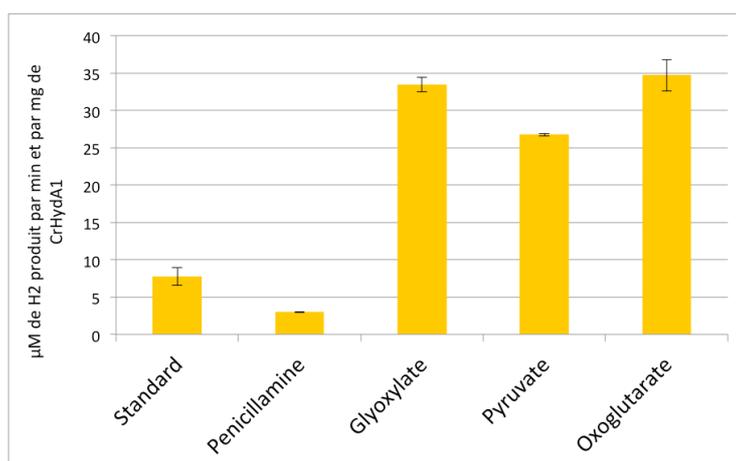


Figure-II-71 : Effets de l'ajout de L-cystéine ou de sulfures sur l'activation *in vitro* de hydrogénase à [FeFe]

## II-4.2 Effet de dérivés thiazolidines formés dans le mélange de maturation

Nous avons ainsi pensé que la L-cystéine pourrait, sous forme d'une thiazolidine, faire partie du substrat de HydE. Pour tester cette hypothèse nous avons quantifié l'impact de l'ajout de trois  $\alpha$ -ketoacides - glyoxylate, pyruvate et oxoglutarate - dans le mélange de maturation d'hydrogénase. Puisque le milieu réactionnel standard contenait 2 mM de cystéine nous espérons former les dérivés thiazolidines par cyclocondensation des  $\alpha$ -ketoacides avec la cystéine directement dans le mélange de maturation [Figure-II-72]. Dans cette série d'essais de maturation nous avons inclus une condition avec 2 mM de L-pénicillamine, comme un analogue de la cystéine. Nous voulions vérifier si la pénicillamine pourrait réagir avec le pyruvate, naturellement présent dans les extraits cellulaires, et former le dérivé (2*R*,4*R*)-TMeTDA. Ce dernier pourrait éventuellement agir comme inhibiteur de HydE à cause du clash stérique entre les deux méthyles sur le C5 du cycle thiazolidine et le O4' de la SAM [I-11.4.1]. Nos premiers résultats montrent que l'ajout de pénicillamine faisait baisser l'activité de l'hydrogénase d'environ 60%. Ceci allait dans le sens de notre hypothèse sur l'inhibition de HydE par le dérivé (2*R*,4*R*)-TMeTDA. De l'autre côté, les trois  $\alpha$ -ketoacides ont tous fait augmenter l'activité hydrogénase de l'extrait de plus de 300% par rapport à la condition standard. Ceci était cependant quelque peu étonnant, parce que nos essais de co-cristallisation ont bien montré que les dérivés (2*R*,4*R*)-TDA et (2*R*,4*R*)-MeTDA forment des complexes très stables avec la protéine *TmHydE*, mais aucun complexe n'a été observé pour le dérivé (2*R*,4*R*)-2-propionate-1,3-thiazoldinie-2,4-dicarboxylate [Figure-I-43], sans doute à cause de la présence du groupement propionate sur le C2. Trop volumineux, ce groupement engendrerait un clash stérique avec la chaîne latérale de la Ile56 et empêcherait ainsi la fixation du dérivé (2*R*,4*R*)-2-propionate-1,3-thiazoldinie-2,4-dicarboxylate dans le site actif de HydE [I-8.7.2]. Puisque les trois  $\alpha$ -ketoacides ont été ajoutés au mélange réactionnel séparément de la L-cystéine, il était possible que leur effet sur la maturation de l'hydrogénase fût lié à une autre réaction que celle de la formation de dérivés thiazolidines.

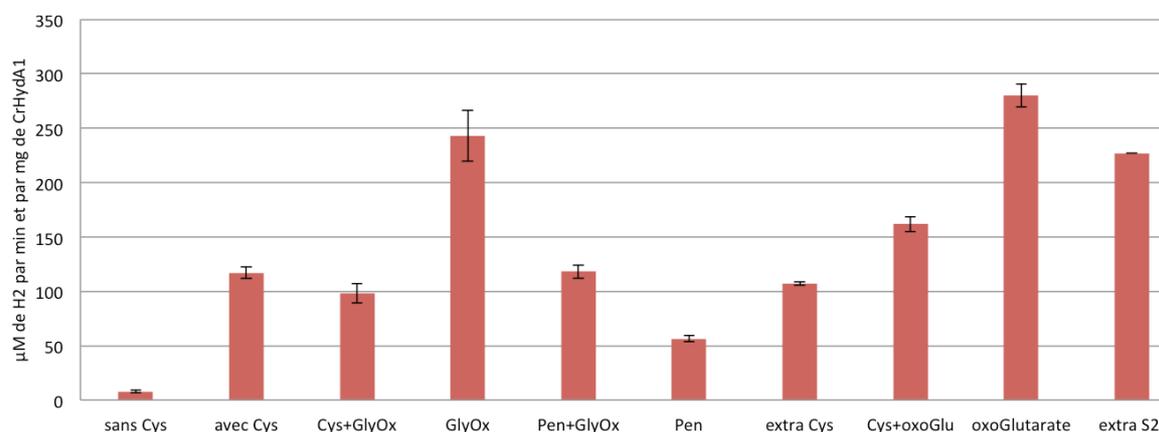


**Figure-II-72** : Effets de l'ajout de pénicillamine ou de  $\alpha$ -ketoacides sur l'activation de hydrogénase à [FeFe]

Augmentation de l'activation anaérobie de l'hydrogénase CrHydA1 en présence d'un des trois  $\alpha$ -ketoacides testés. L'ajout de pénicillamine a eu un effet inhibiteur sur l'activation d'hydrogénase. Les mesures d'activité hydrogénase ont été effectuées après une période de maturation anaérobie de 18 heures (nombre d'expériences n=2-4). On remarquera cependant la faible activité de tous les échantillons de cette série.

### II-4.3 Effet de dérivés thiazolidines préformés vs $\alpha$ -ketoacides seuls

Nous avons donc fait une nouvelle série d'expériences avec les dérivés thiazolidines (2*R*,4*R*)-TDA et (2*R*,4*R*)-2-propionate-1,3-thiazolidine-2,4-dicarboxylate [Figure-II-73], préformés en mélangeant de la cystéine avec un équivalent de glyoxylate ou oxoglutarate avant d'être ajoutés au mélange de maturation. Le dérivé (2*R*,4*R*)-DMeTDA [Figure-I-42] a été préparé à partir d'un mélange équimolaire de la L-pénicillamine et du glyoxylate.



**Figure-II-73** : Effets de l'ajout de dérivés thiazolidines préformés vs.  $\alpha$ -ketoacides seuls sur l'activation de l'hydrogénase à [FeFe]

Seul la première condition "sans Cys" ne contient pas de L-cystéine exogène. La condition "avec Cys" est l'échantillon standard. Les mesures d'activité hydrogénase ont été effectuées après une période de maturation anaérobie de 18 heures (n=2).

Etonnamment il n'y a pas eu d'importants changements d'activité hydrogénase ni pour le dérivé (2*R*,4*R*)-TDA, ni pour le présumé inhibiteur (2*R*,4*R*)-DMeTDA. Seul le dérivé (2*R*,4*R*)-2-propionate-1,3-thiazolidine-2,4-dicarboxylate présentait une augmentation d'activité notable de +38%. Ces données indiquaient que nos dérivés thiazolidines qu'on retrouvait dans la cavité de HydE n'étaient pas impliqués dans le processus d'assemblage du site actif de l'hydrogénase à [FeFe].

Par contre l'ajout de glyoxylate ou d'oxoglutarate seuls augmentait l'activité hydrogénase de +108% et +139% respectivement [Figure-II-73]. Ceci suggère que les  $\alpha$ -ketoacides jouent un rôle à un certain niveau de la maturation de l'enzyme.

### II-4.4 Effet de pénicillamine

La pénicillamine a fait baisser l'activité hydrogénase d'environ 50%. Par contre son dérivé thiazolidine (2*R*,4*R*)-DMeTDA n'affecte en rien l'activité hydrogénase [Figure-II-73]. Il est donc clair que c'est la pénicillamine elle-même qui présente un effet inhibiteur de la maturation d'hydrogénase. Or il a été démontré que la L-pénicillamine agit comme inhibiteur des enzymes à PLP<sup>87</sup>. L'exclusion du cofacteur PLP du mélange réactionnel engendre une chute d'activité hydrogénase de plus de 90%<sup>24</sup>. Tous ces résultats pris en compte suggèrent qu'une enzyme à PLP constitue un élément important de la maturation de l'hydrogénase à [FeFe].

#### II-4.5 Rôle du PLP dans le processus de la maturation de l'hydrogénase à [FeFe]

L'analyse des gènes associés avec l'opéron *hydAEEFG* montre la présence de deux clusters de gènes impliqués dans la synthèse de la cystéine<sup>63</sup>. Puisque ces gènes codent pour des enzymes PLP-dépendantes, Betz et collaborateurs ont proposé, que ces enzymes seraient impliquées dans le processus de la maturation de l'hydrogénase, ce qui expliquerait l'effet stimulateur de l'activation de l'hydrogénase en utilisant les extraits cellulaires.

Il se pourrait aussi que l'enzyme PLP-dépendante indispensable à la maturation des hydrogénase à [FeFe] soit une transaminase. Cette dernière pourrait avoir un lien avec l'effet activateur observé pour certains  $\alpha$ -ketoacides testés dans notre laboratoire. D'autres études seront nécessaires pour comprendre le rôle du PLP dans le processus de maturation.

#### II-5 Activation d'hydrogénase en utilisant de la cystéine <sup>13</sup>C

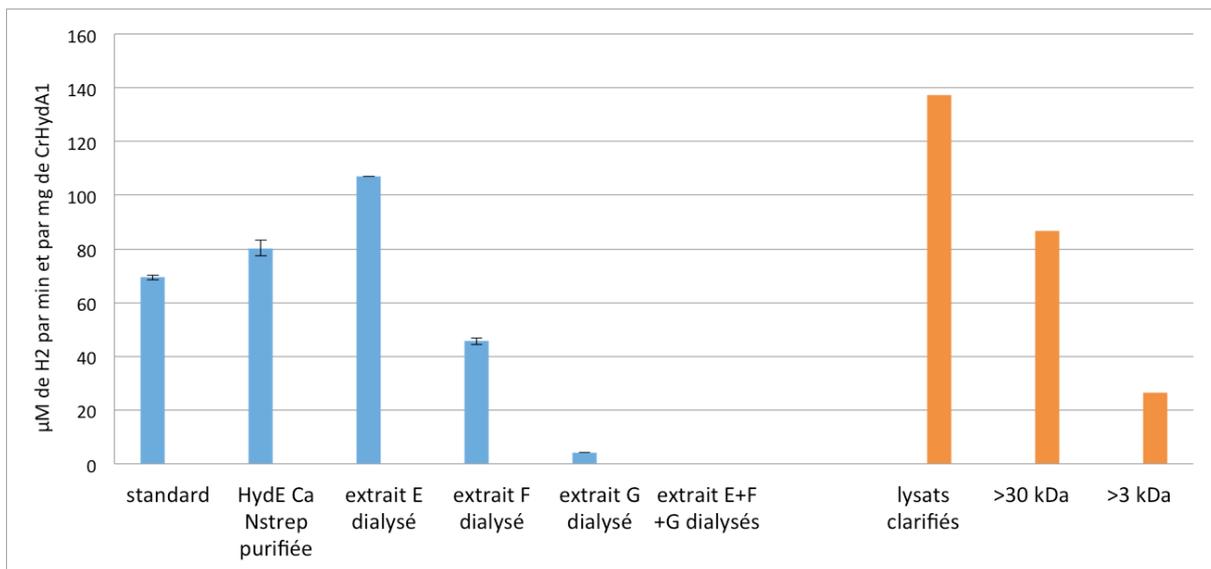
Parallèlement aux premiers tests d'activation d'hydrogénase à [FeFe] avec les  $\alpha$ -ketoacides [II-4.2] nous avons effectué une expérience de maturation d'hydrogénase en présence de la cystéine marquée au <sup>13</sup>C uniquement sur le carbone C <sub>$\beta$</sub> . En purifiant l'hydrogénase ainsi activée nous voulions, à l'aide de la spectroscopie HYSCORE, vérifier si le C <sub>$\beta$</sub>  de la cystéine se retrouvait dans le ligand DTMN du site actif de cette enzyme. Son analyse par résonance paramagnétique électronique (RPE) a montré que cette dernière se trouvait sous sa forme inhibée par le CO - l'état H<sub>ox</sub>-CO. Le spectre HYSCORE ne montrait aucun couplage hyperfin visible entre un éventuel <sup>13</sup>C du ligand DTMN et le sous-agrégat [Fe<sub>2</sub>S<sub>2</sub>]. Ceci suggère que l'atome C <sub>$\beta$</sub>  de la L-cystéine n'est pas la source d'atomes de carbone du ligand DTMN.

#### II-6 Utilisation d'extraits dialysés

Dans la dernière série de tests, nous avons essayé de reproduire les expériences de maturations avec les extraits cellulaires clarifiés et débarrassés de molécules de faible poids moléculaire<sup>24</sup>. Les trois lysats clarifiés et dialysés contenant chacun une des trois maturases surexprimée étaient actifs, parce que quand un seul lysat dialysé a été utilisé avec deux lysats non-dialysés, une activité hydrogénase a pu être mesurée [Figure-II-74]. Par contre lorsque les trois lysats dialysés ont été utilisés ensemble, le mélange de maturation ne présentait plus aucune activité hydrogénase. Ce résultat est en contradiction avec celui de Kuchenreuther et ses collègues. Ils affirment pouvoir activer l'hydrogénase à [FeFe] rien qu'en utilisant des extraits dialysés combinés aux petites molécules exogènes. Cependant nous émettons des doutes sur l'efficacité de leur procédure de préparation d'extraits débarrassés de molécules de faible poids moléculaire. Tandis que nous avons dialysé nos extraits dans deux bains de dialyse successifs [voir MM-3.6.2.2], Kuchenreuther et ses collègues ont changé le tampon des extraits clarifiés en utilisant une colonne de chromatographie d'exclusion stérique<sup>24</sup>. Leur système est sans doute moins performant que le notre, car ils arrivent à activer l'hydrogénase même sans l'ajout de tyrosine exogène. Ceci voudrait dire qu'il y a toujours de la tyrosine libre dans leurs extraits dialysés et donc qu'ils n'ont pas réussi à enlever efficacement toutes les petites molécules.

## II-6.1 Tests de fractionnement de lysat clarifié de *E. coli*

Afin d'identifier l'élément qu'on a perdu pendant la dialyse nous avons effectué un premier test de fractionnement du lysat clarifié frais de *E. coli*. Pour cela nous avons préparé trois fractions en utilisant différentes membranes d'ultra-concentration [voir MM-3.6.2.3]. L'ajout d'une fraction de lysat clarifié de *E. coli* dans le mélange de maturation constitué d'extraits clarifiés dialysés permet de récupérer entièrement l'activité hydrogénase [Figure-II-74]. L'ajout de lysat contenant des molécules de tailles inférieures à 30 kDa permet de récupérer 61% d'activité hydrogénase. Par contre l'ajout de lysat contenant des molécules de tailles inférieures à 3 kDa ne permet de récupérer que 19% d'activité hydrogénase. Ces résultats préliminaires sont difficiles à être interprétés. Il est possible que notre système de préparation de fractions ne soit pas très adapté pour la récupération de molécules de faible poids moléculaires. En effet on peut imaginer que la précipitation de protéines sur la membrane pendant l'étape d'ultra-concentration crée un film de précipitât sur la membrane qui capterait certains métabolites de l'extrait. Ainsi ces derniers se retrouveraient dilués dans la dernière fraction récupérée. D'autres essais qui utiliseront un système différent de fractionnement d'extrait anaérobie seront sans aucun doute nécessaires.



**Figure-II-74** : Test d'activation d'hydrogénase en utilisant des lysats cellulaires clarifiés et dialysé

La série en bleu représente les résultats du test d'activation anaérobie d'hydrogénase en utilisant soit trois lysats clarifiés, chacun contenant une maturase Hyd surexprimée (échantillon standard), ou bien deux lysats clarifiés et un lysat dialysé (échantillons "extrait E dialysé", "extrait F dialysé" et "extrait G dialysé") ou bien trois lysats dialysés (échantillon "extrait E+F+G dialysés") (n=2). La série en orange représente un premier essai de fractionnement du lysat en fonction de la taille des molécules.

## II-7 Conclusions et Perspectives des Chapitres I & II

Les travaux de recherche du substrat de la protéine HydE, présentés dans le *Chapitre I*, ont abouti à l'identification d'un groupe de ligands dérivés de la thiazolidine. Ces ligands sont formés par la cyclocondensation spontanée de la L-cystéine et d'un aldéhyde ou d'un  $\alpha$ -ketoacide de faible poids moléculaire. La protéine reconnaît ces molécules de façon stéréospécifique. Si les ligands possèdent deux groupements carboxylates, un sur C2 et l'autre sur C4, l'affinité du site actif pour ce ligand est suffisamment élevée pour former un complexe très stable. Ces ligands possèdent tous un atome de soufre engagé dans une liaison thioéther. Cet atome se place à proximité de l'atome C5' du cofacteur SAM. Il se fixe dans le site *T1*, identifié au cours de ces travaux, qui présente une haute affinité pour un soufre. Puisque nos essais de maturation d'hydrogénase à [FeFe] en utilisant des extraits anaérobiques clarifiés d'*E. coli* ont montré que ces ligands en particulier ne sont pas impliqués dans le processus de maturation, il semble que ces dérivés thiazolidine ne sont pas des substrats physiologiques de HydE. Cependant, comme HydE fixe ces molécules de façon stéréospécifique et avec une forte affinité, il est possible que le substrat physiologique de l'enzyme ressemble en partie à un de ces ligands.

L'identification du site *T1*, combinée avec l'affinité du site S3 pour le thiocyanate et avec la présence systématique d'un groupement carboxylate dans le site S1 suggère que le ou les substrats physiologiques de HydE devraient posséder un atome de soufre et au moins un groupement carboxylate. Cette hypothèse est en accord avec les résultats récemment obtenus par Betz et collaborateurs, qui ont observé un arrachage d'un atome d'hydrogène échangeable à partir de petites molécules thiolées<sup>63</sup>. Bien que nous ne sommes pas d'accord avec toutes les conclusions de l'étude en question (pour plus de détails consulter le paragraphe III-2.1.2), il reste claire que la protéine HydE possède une affinité pour les molécules contenant un atome de soufre et ayant une charge nette ou partielle négative. L'ensemble d'informations recueillit par l'analyse des ligands de HydE nous laisse penser que cette enzyme est en effet impliquée dans la biosynthèse d'une molécule soufrée, probablement le ligand DTMN ou son précurseur.

La présence de nombreuses molécules d'eau, qui entourent les ligands dans la cavité de HydE, indique que le substrat physiologique de cette protéine pourrait être plus volumineux que les ligands identifiés et/ou qu'une de ces molécule d'eau serait indispensable à la réaction catalysée par l'enzyme.

Nos efforts d'isoler un complexe de HydE avec un des intermédiaires de sa réaction à partir du cytoplasme n'ont pas été fructueux; sans doute à cause de la faible concentration du substrat dans le cytoplasme ou à la stabilité limitée d'un tel complexe. Les limitations de notre approche ont été que nous n'avions pas testé un autre organisme qu'*E. coli* et que la protéine a due être purifiée après qu'elle a été en présence du substrat. Or la concentration du substrat dans le cytoplasme varie probablement d'une espèce à l'autre. Il serait donc intéressant d'essayer plusieurs espèces différentes. Pour contourner le problème de la séparation du complexe enzyme-substrat pendant l'étape de la purification il est possible d'incuber la protéine pure avec du lysat clarifié ou d'une fraction de ce dernier juste avant la préparation de la goutte de cristallisation. Cette méthode a été utilisée pour identifier le chélateur naturel

du nickel dans le transporteur NikA de *S. aureus*<sup>85</sup>. Les auteurs de cette étude ont remarqué que la nature du chélateur dépendait de la souche cellulaire utilisée pour la préparation d'extraits cellulaires.

Nos tentatives de co-purifier *TmHydE* avec un intermédiaire de sa réaction nous ont cependant permis d'isoler un complexe très stable de l'enzyme avec son cofacteur naturel intact - la SAM - en quantité quasi-stœchiométrique. Cette forte affinité du site actif pour la SAM suggère que HydE utilise cette molécule comme cofacteur plutôt que co-substrat. Cette hypothèse avait été proposée par Yvain Nicolet sur la base de l'affinité élevée du site actif de HydE pour les produits du clivage de la SAM, la L-Met et 5'-dAH<sup>88</sup>. Ainsi, après le réarrangement radicalaire du substrat, un intermédiaire radicalaire de la réaction arracherait un atome d'hydrogène de la 5'-dAH pour former le produit final et pour régénérer le cofacteur SAM. Ce mécanisme a été observé pour la LAM<sup>89</sup>, la SPL<sup>90-92</sup>, et la protéine QueE<sup>93</sup>; trois autres membres de la superfamille de protéines à radical SAM<sup>77</sup>.

Alternativement, un  $k_{on}$  élevé de la fixation de la SAM pourrait vouloir dire que deux molécules de la SAM sont nécessaires pour un turnover complet de l'enzyme<sup>63</sup>. Ce cas a été observée pour les protéines à radical SAM BioB<sup>94,95</sup> et LipA<sup>96,97</sup>, toutes les deux impliquées dans les réactions d'insertion de soufre à partir d'un agrégat [Fe-S] auxiliaire<sup>18,95</sup>. Puisque HydE est censée produire le ligand DTMN, une molécule avec une symétrie d'ordre 2, l'utilisation de deux molécules de SAM par turnover ne serait pas surprenante.

Cela dit, une troisième possibilité existe; nous pouvons imaginer que HydE utilise la SAM comme cofacteur dans la première partie de la réaction et comme co-substrat dans la seconde partie. Ainsi, durant la production du ligand DTMN une seule molécule de SAM serait consommée, mais deux événements d'arrachage d'atomes d'hydrogène du substrat auraient lieu. Ce dernier cas reste cependant pour le moment sans précédent.

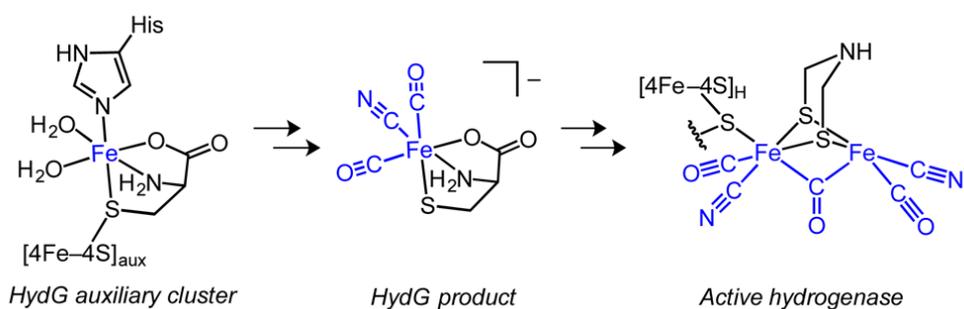
Au cours de ma thèse nous avons également essayé d'identifier le substrat de l'enzyme par les techniques de résonance magnétique nucléaire du proton. En 1999, M. Mayer et B. Meyer ont proposé une méthode de RMN <sup>1</sup>H adaptée à l'identification d'interactions transitoires entre une protéine et un ligand<sup>98,99</sup>. Le principe de la méthode STD, en anglais Saturation Transfer Difference, est décrit plus en détail dans [voir paragraphe MM-5.1.3 & MM-5.1.4]. L'avantage de cette méthode est qu'elle peut permettre l'identification d'un ligand parmi un cocktail de métabolites et qu'elle peut également fournir des informations sur l'affinité du ligand pour la protéine cible. Dans le cas idéal il est possible d'identifier des résidus sur la protéine et des groupements chimiques sur le ligand impliqués dans l'interaction entre les deux partenaires. Ceci a fait de la STD une des méthodes les plus utilisées pour l'identification de ligands d'une protéine. Nous avons obtenu quelques résultats préliminaires, qui ne sont pas présentés dans ce manuscrit, mais la méthode reste à être optimisée pour une protéine contenant des centres [Fe-S]. En effet, les centres paramagnétiques, tels que le centre [Fe<sub>4</sub>S<sub>4</sub>]<sup>+</sup> ou le Fe<sup>3+</sup>, mais aussi le Fe<sup>2+</sup> dans la plupart de cas, créent des artefacts qui rendent l'interprétation des spectres très difficile.

Pour continuer la recherche du substrat de HydE d'autres techniques restent à envisager. Avec la mise en service de lignes de lumière automatiques, comme par exemple la ligne ID30A à ESRF (*European Synchrotron Radiation Facility*) à Grenoble, la cristallographie à haut débit devient une méthode accessible. L'obstacle majeur de cette technique constitue la préparation d'un très grand nombre de cristaux, traités avec différents ligands et congelés dans des cryo-boucles. L'équipe de Alexei S. Soares de Brookhaven National Laboratory aux Etats-Unis a trouvé une astuce pour contourner cet obstacle<sup>100</sup>. Ses membres utilisent les ultrasons pour éjecter une gouttelette de 2,5 nL d'une solution de ligand pour co-cristalliser la protéine en présence de ce ligand directement sur des microgrilles MiTeGen. Ceci leur permet de cribler une seule molécule à la fois et d'éviter donc les mélanges de ligands. D'autres systèmes de congélation automatiques de cristaux ont été développés<sup>101-103</sup>, cependant, aucun de ces systèmes n'a encore été implémenté dans une boîte à gants et donc cette approche se limite, pour le moment, à l'étude cristallographique de protéines insensibles à l'oxygène.

Bien que nos essais de maturation d'hydrogénase dans l'extrait n'aient pas fourni de preuves que les ligands dérivés thiazolidines interviennent dans le processus de maturation nous avons pu identifier un groupe d' $\alpha$ -ketoacides dont l'ajout dans le mélange réactionnel augmente de façon significative l'activité hydrogénase de l'extrait. Il reste à élucider par quel moyen ces molécules influencent ce processus.

L'utilisation de la protéine *TmHydE*  $\Delta$ FeS purifiée dans nos expériences d'activation de l'hydrogénase *CrHydA1* dans les extraits cellulaires contenant des maturases *CaHydG* et *CaHydF* montre qu'il s'agit d'un processus très conservé. En effet, l'enzyme de *T. maritima* est l'une des hydrogénases les plus complexes, tandis que celle *C. reinhardtii* est la plus simple; dépourvue de centres [Fe-S] auxiliaires, cette dernière ne possède que l'agrégat H. Il est non seulement possible de croiser les maturases issues d'une espèce avec l'hydrogénase issue d'une autre espèce, mais également utiliser des maturases issues d'espèces différentes<sup>14</sup>.

Nos essais de marquage de la L-cystéine ont montré que cette dernière, tout en restant indispensable à notre système d'activation d'hydrogénase, ne semble pas être la source des atomes de carbone du ligand DTMN. Récemment, Daniel LM Suess et ses collègues ont proposé que la L-cystéine pourrait se lier au second centre [Fe<sub>4</sub>S<sub>4</sub>] de HydG et servir non seulement de chélateur du cinquième Fe, mais constituer une partie du synthon [(Cys)Fe(CO)<sub>2</sub>(CN)]<sup>-</sup> produit par HydG<sup>78,79</sup>. Les résultats obtenus par Adrien Pagnier dans notre laboratoire et la réinterprétation de données cristallographiques de la structure de HydG obtenues par Dinis et collaborateurs suggéraient que le chélateur du cinquième Fe de l'agrégat auxiliaire serait plutôt la L-homocystéine [voir IB-3.1]. Cependant au cours de la rédaction de ce manuscrit Suess et collaborateurs ont réussi à prouver que la L-cystéine, et non pas l'homocystéine, et le cinquième Fe sont indispensables à la formation de l'agrégat auxiliaire complet dans HydG. Dans une série de publications les auteurs ont montré que la forme active de cet agrégat est donc un centre [Fe<sub>4</sub>S<sub>4</sub>]<sub>aux</sub>[(Cys)Fe] et le synthon produit par HydG est un complexe [(Cys)Fe(CO)<sub>2</sub>(CN)]<sup>-</sup><sup>78,79,104</sup> [Figure-II-75]. Cette observation confirme notre hypothèse que la L-cystéine était impliquée dans un processus lié à la maturation de l'hydrogénase autre que la reconstitution des centres [Fe-S]. Cependant nous ne pouvons pas exclure que cet acide aminé soit également directement impliqué dans la biosynthèse du substrat de HydE.



**Figure-II-75** : Nouveau modèle du fonctionnement de l'agrégat [Fe-S]<sub>aux</sub> de HydG

Source de l'image: Suess *et al.* Inorganic Chemistry 2016 <sup>78</sup>

Nous n'avons pas encore été en mesure de produire un système de maturation d'hydrogénase qui fonctionne avec des extraits clarifiés dialysés. Ceci nous laisse penser qu'un métabolite de l'extrait clarifié, essentiel à la maturation, est perdu pendant la dialyse. Ce métabolite pourrait être le substrat de HydE. Pour confirmer cette hypothèse il faudrait procéder à un fractionnement du lysat clarifié, en utilisant des méthodes plus adaptées que celle d'ultra-concentration sur des membranes de différentes tailles de pores.

Il est à noter que jusqu'à présent, aucun système de maturation d'hydrogénase à [FeFe], qui utilisait les trois maturases Hyd, n'a pu fonctionner sans l'ajout d'un lysat anaérobie (dialysé ou pas) d'*E. coli*. Bien que la présence de métabolites de faible poids moléculaire dans ce lysat soit discutable [II-6], celle de protéines d'*E. coli* ne l'est pas. En d'autres termes, il est tout à fait possible, qu'une protéine ubiquitaire non-identifiée d'*E. coli* soit indispensable à la création d'un système de maturation d'hydrogénase *in vitro* composé de protéines purifiées. L'effet du PLP et des  $\alpha$ -ketoacides sur l'activation de l'hydrogénase dans l'extrait nous fait penser qu'une transaminase PLP-dépendante pourrait être impliquée à un certain niveau du processus de maturation du site actif d'hydrogénase à [FeFe].

Les résultats obtenus dans notre laboratoire et par l'équipe de James R. Swartz suggèrent qu'une enzyme à PLP constitue en effet un élément essentiel de la maturation des hydrogénases à [FeFe]. L'identification de cette enzyme pourrait sans doute apporter des éléments sur la nature du substrat de HydE et aussi ouvrir la voie à la création d'un système de maturation de l'hydrogénase à [FeFe] *in vitro* complet.

## II-8 Références du Chapitre II

- 1 Nicolet, Y., Rohac, R., Martin, L. & Fontecilla-Camps, J. C. X-ray snapshots of possible intermediates in the time course of synthesis and degradation of protein-bound Fe<sub>4</sub>S<sub>4</sub> clusters. *Proceedings of the National Academy of Sciences* **110**, 7188-7192, doi:10.1073/pnas.1302388110 (2013).
- 2 Posewitz, M. C. *et al.* Discovery of Two Novel Radical S-Adenosylmethionine Proteins Required for the Assembly of an Active [Fe] Hydrogenase. *Journal of Biological Chemistry* **279**, 25711-25720, doi:10.1074/jbc.M403206200 (2004).
- 3 Shepard, E. M. *et al.* [FeFe]-Hydrogenase Maturation: HydG-Catalyzed Synthesis of Carbon Monoxide. *Journal of the American Chemical Society* **132**, 9247-9249, doi:10.1021/ja1012273 (2010).
- 4 Driesener, R. C. *et al.* [FeFe] - Hydrogenase Cyanide Ligands Derived From S - Adenosylmethionine - Dependent Cleavage of Tyrosine. *Angewandte Chemie International Edition* **49**, 1687-1690, doi:10.1002/anie.200907047 (2010).
- 5 Pilet, E. *et al.* The role of the maturase HydG in [FeFe]-hydrogenase active site synthesis and assembly. *FEBS Letters* **583**, 506-511, doi:10.1016/j.febslet.2009.01.004 (2008).
- 6 McGlynn, S. E. *et al.* In vitro activation of [FeFe] hydrogenase: new insights into hydrogenase maturation. *JBIC Journal of Biological Inorganic Chemistry* **12**, 443-447, doi:10.1007/s00775-007-0224-z (2007).
- 7 McGlynn, S. E. *et al.* HydF as a scaffold protein in [FeFe] hydrogenase H-cluster biosynthesis. *FEBS Letters* **582**, 2183-2187, doi:10.1016/j.febslet.2008.04.063 (2008).
- 8 Czech, I. *et al.* The [FeFe]-hydrogenase maturation protein HydF contains a H-cluster like [4Fe4S]-2Fe site. *FEBS Letters* **585**, 225-230, doi:10.1016/j.febslet.2010.11.052 (2010).
- 9 Czech, I., Silakov, A., Lubitz, W. & Happe, T. The [FeFe]-hydrogenase maturase HydF from *Clostridium acetobutylicum* contains a CO and CN- ligated iron cofactor. *FEBS letters* **584**, 638-642, doi:10.1016/j.febslet.2009.12.016 (2010).
- 10 Shepard, E. M. *et al.* Synthesis of the 2Fe subcluster of the [FeFe]-hydrogenase H cluster on the HydF scaffold. *Proceedings of the National Academy of Sciences*, doi:10.1073/pnas.1001937107 (2010).
- 11 Berggren, G. *et al.* Biomimetic assembly and activation of [FeFe]-hydrogenases. *Nature* **499**, 66-69, doi:10.1038/nature12239 (2013).
- 12 Nicolet, Y. & Fontecilla-Camps, J. C. Structure-Function Relationships in [FeFe]-Hydrogenase Active Site Maturation. *Journal of Biological Chemistry* **287**, 13532-13540, doi:10.1074/jbc.R111.310797 (2012).
- 13 Rubach, J. K., Brazzolotto, X., Gaillard, J. & Fontecave, M. Biochemical characterization of the HydE and HydG iron-only hydrogenase maturation enzymes from *Thermotoga maritima*. *FEBS Letters* **579**, 5055-5060, doi:10.1016/j.febslet.2005.07.092 (2005).
- 14 Nicolet, Y. *et al.* X-ray Structure of the [FeFe]-Hydrogenase Maturase HydE from *Thermotoga maritima*. *Journal of Biological Chemistry* **283**, 18861-18872, doi:10.1074/jbc.M801161200 (2008).
- 15 King, P. W., Posewitz, M. C., Ghirardi, M. L. & Seibert, M. Functional Studies of [FeFe] Hydrogenase Maturation in an *Escherichia coli* Biosynthetic System. *Journal of Bacteriology* **188**, 2163-2172, doi:10.1128/jb.188.6.2163-2172.2006 (2006).
- 16 Mulder, D. W. *et al.* Activation of HydAΔEFG Requires a Preformed [4Fe-4S] Cluster. *Biochemistry* **48**, 6240-6248, doi:10.1021/bi9000563 (2009).
- 17 Kuchenreuther, J. M., Stapleton, J. A. & Swartz, J. R. Tyrosine, Cysteine, and S-Adenosyl Methionine Stimulate In Vitro [FeFe] Hydrogenase Activation. *PLoS ONE* **4**, doi:10.1371/journal.pone.0007565 (2009).

- 18 Lanz, N. D. & Booker, S. J. Auxiliary iron–sulfur cofactors in radical SAM enzymes. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1853**, doi:10.1016/j.bbamcr.2015.01.002 (2015).
- 19 Berkovitch, F., Nicolet, Y., Wan, J. T., Jarrett, J. T. & Drennan, C. L. Crystal Structure of Biotin Synthase, an S-Adenosylmethionine-Dependent Radical Enzyme. *Science* **303**, 76-79, doi:10.1126/science.1088493 (2004).
- 20 Harmer, J. E. *et al.* Structures of lipoyl synthase reveal a compact active site for controlling sequential sulfur insertion reactions. *The Biochemical journal* **464**, 123-133, doi:10.1042/bj20140895 (2014).
- 21 Böck, A., King, P. W., Blokesch, M. & Posewitz, M. C. Maturation of Hydrogenases. *Advances in microbial physiology* **51**, 1-225, doi:10.1016/s0065-2911(06)51001-x (2006).
- 22 Kuchenreuther, J. M. *et al.* The HydG enzyme generates an Fe(CO)<sub>2</sub>(CN) synthon in assembly of the FeFe hydrogenase H-cluster. *Science (New York, N.Y.)* **343**, 424-427, doi:10.1126/science.1246572 (2014).
- 23 Vallese, F. *et al.* Biochemical Analysis of the Interactions between the Proteins Involved in the [FeFe]-Hydrogenase Maturation Process. *Journal of Biological Chemistry* **287**, 36544-36555, doi:10.1074/jbc.M112.388900 (2012).
- 24 Kuchenreuther, J. M., Britt, D. R. & Swartz, J. R. New Insights into [FeFe] Hydrogenase Activation and Maturase Function. *PLoS ONE* **7**, doi:10.1371/journal.pone.0045850 (2012).
- 25 Irwin, J. J. & Shoichet, B. K. ZINC--a free database of commercially available compounds for virtual screening. *Journal of chemical information and modeling* **45**, 177-182, doi:10.1021/ci049714+ (2005).
- 26 Kuntz, I. D. Structure-based strategies for drug design and discovery. *Science (New York, N.Y.)* **257**, 1078-1082 (1992).
- 27 Verlinde, C. L. & Hol, W. G. Structure-based drug design: progress, results and challenges. *Structure (London, England : 1993)* **2**, 577-587 (1994).
- 28 Nienaber, V. L. *et al.* Discovering novel ligands for macromolecules using X-ray crystallographic screening. *Nature Biotechnology* **18**, 1105-1108, doi:10.1038/80319 (2000).
- 29 Nicolet, Y. *et al.* Crystal structure of HydG from *Carboxydotherrmus hydrogenoformans*: a trifunctional [FeFe]-hydrogenase maturase. *Chembiochem : a European journal of chemical biology* **16**, 397-402, doi:10.1002/cbic.201402661 (2015).
- 30 Dinis, P. *et al.* X-ray crystallographic and EPR spectroscopic analysis of HydG, a maturase in [FeFe]-hydrogenase H-cluster assembly. *Proceedings of the National Academy of Sciences* **112**, 1362-1367, doi:10.1073/pnas.1417252112 (2015).
- 31 Hermann, J. C. *et al.* Structure-based activity prediction for an enzyme of unknown function. *Nature* **448**, 775-779, doi:10.1038/nature05981 (2007).
- 32 Nicolet, Y. & Drennan, C. L. AdoMet radical proteins—from structure to evolution—alignment of divergent protein sequences reveals strong secondary structure element conservation. *Nucleic Acids Research* **32**, 4015-4025, doi:10.1093/nar/gkh728 (2004).
- 33 Quitterer, F., List, A., Eisenreich, W., Bacher, A. & Groll, M. Crystal Structure of Methylornithine Synthase (PylB): Insights into the Pyrrolysine Biosynthesis. *Angewandte Chemie International Edition* **51**, 1339-1342, doi:10.1002/anie.201106765 (2012).
- 34 Nicolet, Y., Zeppieri, L., Amara, P. & Fontecilla - Camps, J. C. Crystal Structure of Tryptophan Lyase (NosL): Evidence for Radical Formation at the Amino Group of Tryptophan. *Angewandte Chemie International Edition* **53**, 11840-11844, doi:10.1002/anie.201407320 (2014).
- 35 Fenwick, M. K. *et al.* Non-canonical active site architecture of the radical SAM thiamin pyrimidine synthase. *Nature Communications* **6**, 6480, doi:10.1038/ncomms7480 (2015).
- 36 Horitani, M. *et al.* Why Nature Uses Radical SAM Enzymes so Widely: Electron Nuclear Double Resonance Studies of Lysine 2,3-Aminomutase Show the 5' -dAdo• “Free Radical” Is Never Free. *Journal of the American Chemical Society*, 2022414680, doi:10.1021/jacs.5b00498 (2015).

- 37 Michal, G. & Schomburg, D. Biological Pathways: An atlas of biochemistry and molecular biology. *Biological Pathways: An atlas of biochemistry and molecular biology*, doi:10.1007/s002130050786 (2012).
- 38 Kriek, M., Martins, F., Challand, M. R., Croft, A. & Roach, P. L. Thiamine Biosynthesis in *Escherichia coli*: Identification of the Intermediate and By - Product Derived from Tyrosine. *Angewandte Chemie International Edition* **46**, 9223-9226, doi:10.1002/anie.200702554 (2007).
- 39 Philmus, B., Decamps, L., Berteau, O. & Begley, T. P. Biosynthetic Versatility and Coordinated Action of 5' -Deoxyadenosyl Radicals in Deazaflavin Biosynthesis. *Journal of the American Chemical Society* **137**, 1740106492, doi:10.1021/ja513287k (2015).
- 40 Maveyraud, L. *et al.* Insights into class D beta-lactamases are revealed by the crystal structure of the OXA10 enzyme from *Pseudomonas aeruginosa*. *Structure (London, England : 1993)* **8**, 1289-1298, doi:10.1016/s0969-2126(00)00534-7 (2000).
- 41 Carta, R. & Tola, G. Solubilities of L-cystine, L-tyrosine, L-leucine, and glycine in aqueous solutions at various pHs and NaCl concentrations. *Solubilities of L-cystine, L-tyrosine, L-leucine, and glycine in aqueous solutions at various pHs and NaCl concentrations*, doi:10.1021/je9501853 (1996).
- 42 Kredich, N. M., Foote, L. J. & Keenan, B. S. The stoichiometry and kinetics of the inducible cysteine desulfhydrase from *Salmonella typhimurium*. *The Journal of biological chemistry* **248**, 6187-6196 (1973).
- 43 Pagnier, A., Nicolet, Y. & Fontecilla-Camps, J. C. IscS from *Archaeoglobus fulgidus* has no desulfurase activity but may provide a cysteine ligand for [Fe<sub>2</sub>S<sub>2</sub>] cluster assembly. *Biochimica et biophysica acta* **1853**, 1457-1463, doi:10.1016/j.bbamcr.2014.10.015 (2015).
- 44 Nicolet, Y., Piras, C., Legrand, P., Hatchikian, C. E. & Fontecilla-Camps, J. C. Desulfovibrio desulfuricans iron hydrogenase: the structure shows unusual coordination to an active site Fe binuclear center. *Structure* **7**, doi:10.1016/s0969-2126(99)80005-7 (1998).
- 45 Challand, M. R. *et al.* Product inhibition in the radical S-adenosylmethionine family. *FEBS letters* **583**, 1358-1362, doi:10.1016/j.febslet.2009.03.044 (2009).
- 46 Robbins, F. M. & Fioriti, J. A. Alkaline Degradation of Cystine, Glutathionine and Sulphur-containing Proteins. *Nature* **200** (1963).
- 47 Clarke, H. T. & Inouye, J. M. Some observations on the action of alkali upon cystine and cysteine. *Journal of Biological Chemistry* (1930).
- 48 Schubert, M. P. Compounds of Thiol acids with Aldehydes. *JBC* (1936).
- 49 Kallen, R. G. The mechanism of reactions involving Schiff base intermediates. Thiazolidine formation from L-cysteine and formaldehyde. *Journal of the American Chemical Society* **93**, 6236-6248, doi:10.1021/ja00752a040 (1971).
- 50 Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research* **28**, 27-30, doi:10.1093/nar/28.1.27 (2000).
- 51 Boyer, M. E., Stapleton, J. A., Kuchenreuther, J. M., Wang, C. w. & Swartz, J. R. Cell - free synthesis and maturation of [FeFe] hydrogenases. *Biotechnology and Bioengineering* **99**, 59-67, doi:10.1002/bit.21511 (2008).
- 52 Stapleton, J. A. & Swartz, J. R. Development of an In Vitro Compartmentalization Screen for High-Throughput Directed Evolution of [FeFe] Hydrogenases. *PLoS ONE* **5**, doi:10.1371/journal.pone.0015275 (2010).
- 53 Kuchenreuther, J. M. *et al.* High-Yield Expression of Heterologous [FeFe] Hydrogenases in *Escherichia coli*. *PLoS ONE* **5**, doi:10.1371/journal.pone.0015491 (2010).
- 54 Kuchenreuther, J. M., George, S. J., Grady-Smith, C. S., Cramer, S. P. & Swartz, J. R. Cell-free H-cluster Synthesis and [FeFe] Hydrogenase Activation: All Five CO and CN- Ligands Derive from Tyrosine. *PLoS ONE* **6**, doi:10.1371/journal.pone.0020346 (2011).
- 55 Kuchenreuther, J. M. ENGINEERING IN VIVO AND IN VITRO SYSTEMS FOR EXPRESSION AND ACTIVATION OF [FeFe] HYDROGENASES. *ENGINEERING IN VIVO AND IN VITRO SYSTEMS FOR EXPRESSION AND ACTIVATION OF [FeFe] HYDROGENASES* (2010).

- 56 Nakamura, M., Saeki, K. & Takahashi, Y. Hyperproduction of recombinant ferredoxins in *Escherichia coli* by coexpression of the ORF1-ORF2-iscS-iscU-iscA-hscB-hscA-fdx-ORF3 gene cluster. *Journal of Biochemistry* (1999).
- 57 Cicchillo, R. M. *et al.* *Escherichia coli* lipoyl synthase binds two distinct [4Fe-4S] clusters per polypeptide. *Biochemistry* **43**, 11770-11781, doi:10.1021/bi0488505 (2004).
- 58 Hänzelmann, P. & Schindelin, H. Crystal structure of the S-adenosylmethionine-dependent enzyme MoaA and its implications for molybdenum cofactor deficiency in humans. *Proceedings of the National ...*, doi:10.1073/pnas.0404624101 (2004).
- 59 Hänzelmann, P., Hernández, H. L. & Menzel, C. Characterization of MOCS1A, an oxygen-sensitive iron-sulfur protein involved in human molybdenum cofactor biosynthesis. *Journal of Biological ...*, doi:10.1074/jbc.M313398200 (2004).
- 60 Desiderio, C. Evaluation of chemical and diastereoisomeric stability of S-adenosylmethionine in aqueous solution by capillary electrophoresis. *Journal of Pharmaceutical and Biomedical Analysis*.
- 61 Parks, L. W. & Schlenk, F. The stability and hydrolysis of S-adenosyl-methionine; isolation of S-ribosylmethionine. *Journal of Biological Chemistry* (1958).
- 62 Cantoni, G. L. S-ADENOSYLMETHIONINE; A NEW INTERMEDIATE FORMED ENZYMATIALLY FROM L-METHIONINE AND ADENOSINETRIPHOSPHATE. *Journal of Biological Chemistry* (1953).
- 63 Betz, J. N. *et al.* [FeFe]-Hydrogenase Maturation: Insights into the Role HydE Plays in Dithiomethylamine Biosynthesis. *Biochemistry*, 1574462297, doi:10.1021/bi501205e (2015).
- 64 Refouvelet, B., Robert, J.-F., Couquelet, J. & Tronche, P. Thiazolidine-2,4-dicarboxylic acid and its esters: Synthesis, in solution behaviour and regioselective cyclocondensation. *Journal of Heterocyclic Chemistry* **31** (1994).
- 65 Shiraiwa, T. *et al.* Preparation of (2S,4R) - 2,4 - thiazolidinedicarboxylic acid by separation of diastereoisomeric salts with organic amines. *Chirality* **11**, 326-329, doi:10.1002/(sici)1520-636x(1999)11:4<326::aid-chir11>3.0.co;2-2 (1999).
- 66 Shiraiwa, T., Katayama, T., Kaito, T., Tanigawa, H. & Kurokawa, H. Facile Preparation of Four Stereoisomers of 2,4-Thiazolidinedicarboxylic Acid. *Bulletin of the Chemical Society of Japan* **71**, doi:10.1246/bcsj.71.1911 (1998).
- 67 Nagasawa, H. T., Goon, D. J. W. & Shirota, F. N. Epimerization at C-2 of 2-substituted thiazolidine-4-carboxylic acids. *Journal of Heterocyclic Chemistry* **18**, 1047-1051, doi:10.1002/jhet.5570180538 (1981).
- 68 Chiarino, D., Ferrario, F., Pellacini, F. & Sala, A. Synthesis and in Solution Behaviour of New 2-Substituted-4-thiazolidinescarboxylic Acid Derivatives. *Journal of Heterocyclic Chemistry* **26** (1989).
- 69 Baker, P. W., Bais, R. & Rofe, A. M. Formation of the l-cysteine-glyoxylate adduct is the mechanism by which l-cysteine decreases oxalate production from glycollate in rat hepatocytes. *Biochemical Journal* **302**, 753-757, doi:10.1042/bj3020753 (1994).
- 70 Bringmann, G. *et al.* A "chemical" concept for the therapy of glyoxylate-induced oxalurias (1). *Life Sciences* **50**, 1597-1605, doi:10.1016/0024-3205(92)90445-u (1992).
- 71 Wagner, G. & Wüthrich, K. Truncated driven nuclear overhauser effect (TOE). A new technique for studies of selective  $^1\text{H}$   $^1\text{H}$  overhauser effects in the presence of spin diffusion. *Journal of Magnetic Resonance (1969)* **33**, doi:10.1016/0022-2364(79)90180-x (1979).
- 72 Kumar, A., Ernst, R. R. & Wüthrich, K. A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochemical and biophysical research communications* **95**, 1-6 (1980).
- 73 Butvin, P., Mitková, M., Sv TlÍK, J. & HavrÁnek, E. Complex Formation of Thiazolidine-2,4-dicarboxylic Acid with Selected Divalent and Trivalent Metal Ions. *Chem.Pap.* (2002).
- 74 Mulder, D. W. *et al.* Stepwise [FeFe]-hydrogenase H-cluster assembly revealed in the structure of HydAΔEFG. *Nature* **465**, 248-251, doi:10.1038/nature08993 (2010).

- 75 Kamp, C. *et al.* Isolation and first EPR characterization of the [FeFe]-hydrogenases from green algae. *Biochimica et biophysica acta* **1777**, 410-416, doi:10.1016/j.bbabi.2008.02.002 (2008).
- 76 Happe, T. & Naber, J. D. Isolation, characterization and N-terminal amino acid sequence of hydrogenase from the green alga *Chlamydomonas reinhardtii*. *European journal of biochemistry / FEBS* **214**, 475-481, doi:10.1111/j.1432-1033.1993.tb17944.x (1993).
- 77 Broderick, J. B., Duffus, B. R., Duschene, K. S. & Shepard, E. M. Radical S-Adenosylmethionine Enzymes. *Chemical Reviews* **114**, 4229-4317, doi:10.1021/cr4004709 (2014).
- 78 Suess, D. L. M., Kuchenreuther, J. M., Paz, L., Swartz, J. R. & Britt, D. R. Biosynthesis of the [FeFe] Hydrogenase H Cluster: A Central Role for the Radical SAM Enzyme HydG. *Inorganic Chemistry*, doi:10.1021/acs.inorgchem.5b02274 (2015).
- 79 Suess, D. L. M. *et al.* Cysteine as a ligand platform in the biosynthesis of the FeFe hydrogenase H cluster. *Proceedings of the National Academy of Sciences* **112**, 201508440, doi:10.1073/pnas.1508440112 (2015).
- 80 Suess, D. L. M. *et al.* The Radical SAM Enzyme HydG Requires Cysteine and a Dangler Iron for Generating an Organometallic Precursor to the [FeFe]-Hydrogenase H-Cluster. *Journal of the American Chemical Society*, doi:10.1021/jacs.5b12512 (2016).
- 81 Pagnier, A., Martin, L., Zeppieri, L., Nicolet, Y. & Fontecilla-Camps, J. C. CO and CN- syntheses by [FeFe]-hydrogenase maturase HydG are catalytically differentiated events. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 104-109, doi:10.1073/pnas.1515842113 (2015).
- 82 Peters, J. W., Lanzilotta, W. N., Lemon, B. J. & Seefeldt, L. C. X-ray crystal structure of the Fe-only hydrogenase (CpI) from *Clostridium pasteurianum* to 1.8 angstrom resolution. *Science (New York, N.Y.)* **282**, 1853-1858, doi:10.1126/science.282.5395.1853 (1998).
- 83 Nicolet, Y., Lemon, B. J., Fontecilla-Camps, J. C. & Peters, J. W. A novel FeS cluster in Fe-only hydrogenases. *Trends in biochemical sciences* **25**, 138-143, doi:10.1016/s0968-0004(99)01536-4 (2000).
- 84 Nicolet, Y. *et al.* Crystallographic and FTIR Spectroscopic Evidence of Changes in Fe Coordination Upon Reduction of the Active Site of the Fe-Only Hydrogenase from *Desulfovibrio desulfuricans*. *Journal of the American Chemical Society* **123**, 1596-1601, doi:10.1021/ja0020963 (2001).
- 85 Lebrette, H. *et al.* Novel insights into nickel import in *Staphylococcus aureus* : the positive role of free histidine and structural characterization of a new thiazolidine-type nickel chelator. *Metallomics* **7**, doi:10.1039/c4mt00295d (2015).
- 86 Daßler, T., Maier, T., Winterhalter, C. & Böck, A. Identification of a major facilitator protein from *Escherichia coli* involved in efflux of metabolites of the cysteine pathway. *Molecular Microbiology* **36**, 1101-1112, doi:10.1046/j.1365-2958.2000.01924.x (2000).
- 87 Muramatsu, H. *et al.* Discovery and characterization of d-phenylserine deaminase from *Arthrobacter* sp. TKS1. *Applied Microbiology and Biotechnology* **90**, 159-172, doi:10.1007/s00253-010-3028-y (2010).
- 88 Nicolet, Y., Amara, P., Mouesca, J.-M. M. & Fontecilla-Camps, J. C. Unexpected electron transfer mechanism upon AdoMet cleavage in radical SAM proteins. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 14867-14871, doi:10.1073/pnas.0904385106 (2009).
- 89 Frey, P. A. Travels with Carbon-Centered Radicals. 5' -Deoxyadenosine and 5' -Deoxyadenosine-5' -yl in Radical Enzymology. *Accounts of Chemical Research*, doi:10.1021/ar400194k (2014).
- 90 Himo, F. C-C bond formation and cleavage in radical enzymes, a theoretical perspective. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1707**, doi:10.1016/j.bbabi.2004.04.006 (2005).
- 91 Buis, J. M., Cheek, J., Kalliri, E. & Broderick, J. B. Characterization of an active spore photoproduct lyase, a DNA repair enzyme in the radical S-adenosylmethionine superfamily. *The Journal of biological chemistry* **281**, 25994-26003, doi:10.1074/jbc.M603931200 (2006).

- 92 Yang, L. *et al.* A Radical Transfer Pathway in Spore Photoproduct Lyase. *Biochemistry* **52**, 1155023668, doi:10.1021/bi3016247 (2013).
- 93 McCarty, R. M., Krebs, C. & Bandarian, V. Spectroscopic, steady-state kinetic, and mechanistic characterization of the radical SAM enzyme QueE, which catalyzes a complex cyclization reaction in the biosynthesis of 7-deazapurines. *Biochemistry* **52**, 188-198, doi:10.1021/bi301156w (2012).
- 94 Fugate, C. J. & Jarrett, J. T. Biotin synthase: insights into radical-mediated carbon-sulfur bond formation. *Biochimica et biophysica acta* **1824**, 1213-1222, doi:10.1016/j.bbapap.2012.01.010 (2012).
- 95 Jarrett, J. T. The Biosynthesis of Thiol- and Thioether-Containing Cofactors and Secondary Metabolites Catalyzed by Radical S-Adenosylmethionine Enzymes. *Journal of Biological Chemistry*, doi:10.1074/jbc.R114.599308 (2014).
- 96 Lanz, N. D. *et al.* Evidence for a Catalytically and Kinetically Competent Enzyme-Substrate Cross-Linked Intermediate in Catalysis by Lipoyl Synthase. *Biochemistry* **53**, 4557-4572, doi:10.1021/bi500432r (2014).
- 97 Lanz, N. D. *et al.* Characterization of a Radical Intermediate in Lipoyl Cofactor Biosynthesis. *Journal of the American Chemical Society* **137**, 13216-13219, doi:10.1021/jacs.5b04387 (2015).
- 98 Mayer, M. & Meyer, B. Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angewandte Chemie International Edition* **38**, 1784-1788, doi:10.1002/(sici)1521-3773(19990614)38:12<1784::ai-anie1784>3.0.co;2-q (1999).
- 99 Mayer, M. & Meyer, B. Group Epitope Mapping by Saturation Transfer Difference NMR To Identify Segments of a Ligand in Direct Contact with a Protein Receptor. *Journal of the American Chemical Society* **123**, doi:10.1021/ja0100120 (2001).
- 100 Yin, X. *et al.* Hitting the target: fragment screening with acoustic in situ co-crystallization of proteins plus fragment libraries on pin-mounted data-collection micromeshes. *Acta Crystallographica Section D: Biological Crystallography* **70**, doi:10.1107/s1399004713034603 (2014).
- 101 Khajepour, M. Y. *et al.* REACH: Robotic Equipment for Automated Crystal Harvesting using a six-axis robot arm and a micro-gripper. *Acta Crystallographica Section D: Biological Crystallography* **69**, 381-387, doi:10.1107/s0907444912048019 (2013).
- 102 Deller, M. C. & Rupp, B. Approaches to automated protein crystal harvesting. *Acta Crystallographica Section F: Structural Biology Communications* **70**, 133-155, doi:10.1107/s2053230x14000387 (2014).
- 103 Cipriani, F. *et al.* CrystalDirect: a new method for automated crystal harvesting based on laser-induced photoablation of thin films. *Acta crystallographica. Section D, Biological crystallography* **68**, 1393-1399, doi:10.1107/s0907444912031459 (2012).
- 104 Suess, D. L. M. *et al.* The Radical SAM Enzyme HydG Requires Cysteine and a Dangler Iron for Generating an Organometallic Precursor to the [FeFe]-Hydrogenase H-Cluster. *Journal of the American Chemical Society*, doi:10.1021/jacs.5b12512 (2016).

## Chapitre III : Formation d'une liaison C-S par un processus radicalaire catalysé par HydE

### III-1 Introduction du Chapitre

Les travaux présentés dans le *Chapitre I* ont conduit à l'identification d'un groupe de ligands dérivés de la thiazolidine, dont certains possèdent une forte affinité pour le site actif de HydE. Bien que les essais de maturation de l'hydrogénase décrits dans le *Chapitre II* n'aient pas prouvé que parmi ces ligands se trouve le substrat physiologique de HydE, nous avons quand même décidé de nous intéresser de plus près à ces molécules. En effet, elles ont la particularité de se fixer dans HydE de façon à ce que l'atome de soufre du thioéther se situe à proximité de l'atome C5' du cofacteur SAM. Les tests de clivage de la SAM ont montré que certains de ces ligands peuvent subir des transformations qui dépendent de la présence de l'enzyme HydE, de la SAM et d'un réducteur fort. A la fin de cette transformation nous avons pu observer l'apparition d'une nouvelle liaison C-S entre l'atome C5' de la SAM et l'atome de soufre, qui lui provient du cycle 1,3-thiazolidine. La particularité de cette réaction est qu'elle ne fait pas intervenir un centre [Fe-S] auxiliaire, comme le font d'autres protéines à radical SAM, qui catalysent les réactions d'insertion de soufre sur un atome de carbone aliphatique (voir paragraphe IB-5).

La découverte de la réactivité de HydE envers les ligands thiazolidines, combinée avec notre capacité à obtenir des cristaux de la protéine qui diffractent à haute résolution ont constitué l'opportunité idéale pour étudier les événements propres aux protéines à radical SAM, tels que le clivage du cofacteur SAM et la génération du radical 5'-dA•.

Ce chapitre est donc consacré à l'étude de la réaction radicalaire observée en solution et dans les cristaux de HydE. La première partie de ce chapitre décrit les résultats de mesure du clivage de la SAM en présence de dérivés thiazolidines. Ensuite nous décrivons la caractérisation en solution des produits de la réaction de HydE avec le ligand (2*R*,4*R*)-MeTDA ((2*R*,4*R*)-2-méthyl-1,3-thiazolidine-2,4-dicarboxylate). Les produits de réaction ont été étudiés en utilisant, dans un premier temps, la chromatographie en phase liquide couplée à la spectroscopie de masse, et dans un second temps, par la spectroscopie RMN du proton et du <sup>13</sup>C. A notre connaissance, il s'agit de la première expérience de suivi en temps réel d'une réaction catalysée par une protéine à radical SAM par RMN. En parallèle nous avons réalisé une étude structurale de cette même réaction dans les cristaux de HydE. Les structures cristallographiques à haute résolution ainsi obtenues ont par la suite été utilisées pour des calculs théoriques dans le but d'étudier l'évolution de la structure électronique du complexe HydE - [Fe<sub>4</sub>S<sub>4</sub>] - SAM - (2*R*,4*R*)-MeTDA au cours de la réaction radicalaire.

L'ensemble des données recueillies apporte des informations sur le clivage de la SAM et la gestion d'espèces hautement réactives par la protéine, mais également sur le mécanisme de la formation d'une liaison C-S par un processus radicalaire et la libération des produits secondaires de réaction.

## III-2 Etude de la réaction de HydE avec les dérivés 1,3-thiazolidine en solution

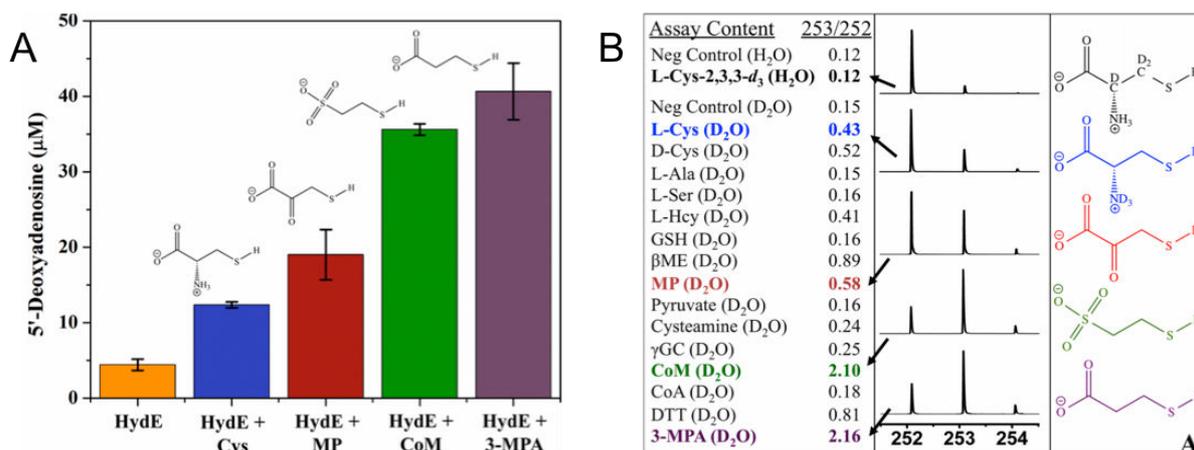
### III-2.1 Etude fonctionnelle de HydE par les mesures du clivage de la SAM

#### III-2.1.1 Clivage de la SAM en présence de molécules thiolées (Betz et al.)

Avant d'exposer nos propres résultats sur les expérience de clivage de la SAM, il me semblait judicieux de commenter les résultats de l'étude de Jeremiah N. Betz et ses collaborateurs<sup>1</sup>, publiée vers la fin de mes travaux de thèse. Les auteurs ont effectué une quantification de la production du 5'-dAH en présence de différents substrats potentiels de HydE.

Ils ont émis l'hypothèse, que le substrat physiologique de HydE devrait posséder un groupement thiol. Cette hypothèse est basée sur leurs mesures de production *in vitro* de la 5'-deoxyadenosine en présence de molécule contenant une fonction thiol. Lorsqu'ils ont effectué la réaction en substituant l'eau par du D<sub>2</sub>O, ils ont pu détecter une incorporation de deutérium dans la 5'-deoxyadenosine [Figure-III-76]. Ils en ont conclu que HydE arrachait un atome d'hydrogène échangeable sur un groupement thiol.

Betz et collaborateurs ont ensuite proposé un mécanisme de production du ligand azadithiolate à partir de la L-cystéine. Dans ce mécanisme l'arrachage d'un atome d'hydrogène du thiol produirait du thioformaldéhyde et un radical glycyl. Deux molécules de thioformaldéhyde ainsi générées réagiraient sur un ion ammonium pour former le produit dithiométhylamine à l'image de la production de ce ligand par la voie synthétique<sup>2</sup>.



**Figure-III-76** : Résultats d'incorporation de deutérium dans AdoH publiés par Betz et collaborateurs

**A** - Stimulation de la production du AdoH par HydE à partir de la SAM suite à l'ajout de L-cystéine, mercaptopyruvate, coenzyme M ou 3-mercaptopropionate. **B** - Résultats d'incorporation du deutérium dans AdoH par HydE en présence de substrats potentiels dans D<sub>2</sub>O. Les chiffres représentent le rapport de populations des molécules AdoH deutérées et des molécules AdoH non-deutérées. Les spectres de masse pour certains substrats, qui ont servi pour le calcul de ce rapport, sont représentés au milieu de l'image. Les structures des substrats sélectionnés sont représentées à droite de l'image. (Images adaptées de Betz et al, Biochemistry, 2015 <sup>1</sup>).

### III-2.1.2 Réinterprétation des données publiées par Betz et al.

Bien que le mécanisme proposé par Betz et ses collaborateurs soit très intéressant, plusieurs éléments le contredisent : (i) notre étude montre que la cystéine libre ne forme pas de complexe avec *TmHydE* ; (ii) certains des substrats identifiés par JN Betz, sont quatre fois plus actifs que la L-cystéine [Figure-III-76-A] ; (iii) la D-cystéine présente un taux de clivage de la SAM légèrement plus élevé que celui avec la L-cystéine [Figure-III-76-B] ; la protéine ne serait donc pas stéréospécifique vis-à-vis de la cystéine ; (iv) parmi les meilleurs substrats identifiés il y a des molécules non naturelles comme le 3-mercaptopropionate, le  $\beta$ -mercaptoéthanol et le DTT qui présentent des activités plus importantes que la L-cystéine [Figure-III-76-B] ; (v) pour finir, l'arrachage d'un atome d'hydrogène sur un thiol conduirait à la production d'un radical thiyl relativement stable<sup>3</sup>, ce qui est peu compatible avec une énergie suffisante pour subséquemment cliver une liaison C-C. Ce radical thiyl est d'ailleurs utilisé par certaines enzymes à radicaux tyrosyl et glycyl, comme par exemple la ribonucléotide réductase aérobie (classe I RNR) et la ribonucléotide réductase anaérobie (classe III RNR), qui s'en servent comme radical "exécuteur" pour attaquer le substrat, tandis que les radicaux tyrosyl et glycyl servent de radicaux de "stockage"<sup>4-6</sup>. Dans le cas de la ribonucléotide réductase de classe II, une enzyme adénosylcobalamine-dépendante, il a été proposé que le radical thiyl était généré directement par un arrachage de l'atome d'hydrogène à partir du thiol du résidu Cys<sup>408</sup> par le radical 5'-dA• issu du clivage homolytique du cofacteur adénosylcobalamine<sup>7,8</sup>. Ce dernier exemple montre que l'attaque du radical 5'-dA• sur le thiol de la L-cystéine ne conduit pas au clivage de la liaison C $\alpha$ -C $\beta$ .

En se basant sur nos connaissances de la structure de HydE et sur les expériences effectuées par d'autres groupes nous pensons que Betz et collaborateurs ont quantifié une réaction parasite. L'augmentation du clivage de la SAM avec le DTT a déjà été observée par Rubach et collaborateurs, non seulement pour HydE, mais également pour la maturase HydG<sup>9</sup>. Tandis que dans le cas de HydE il y a deux fois plus de 5'-déoxyadénosine produite avec le DTT que sans cette molécule, pour HydG cela représente une augmentation de plus de huit fois [Figure-III-77-A&B]. Les expériences avec le DTT montrent que les thiols sont de bons accepteurs de radicaux<sup>10</sup>, ce qui fait sans doute déplacer l'équilibre de la réaction du clivage de la SAM vers la production de la 5'-deoxyadénosine. Il est donc possible qu'une telle réactivité envers les thiols soit observée dans d'autres protéines à radical SAM.

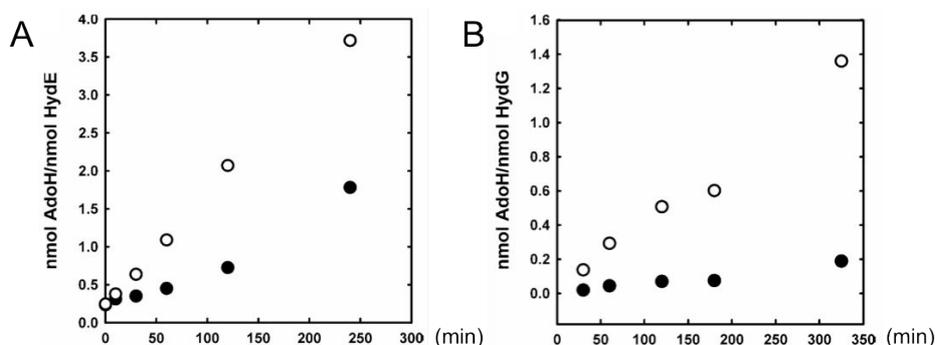


Figure-III-77 : Mesures de la production d'AdoH par HydE et HydG publiées par Rubach et collaborateurs

Production d'AdoH à partir de la SAM par A - HydE et B - HydG et en absence (cercles noirs) ou en présence (cercles blancs) de 5 mM de DTT. (Images adaptées de Rubach et collaborateurs<sup>9</sup> FEBS Letters 2005).

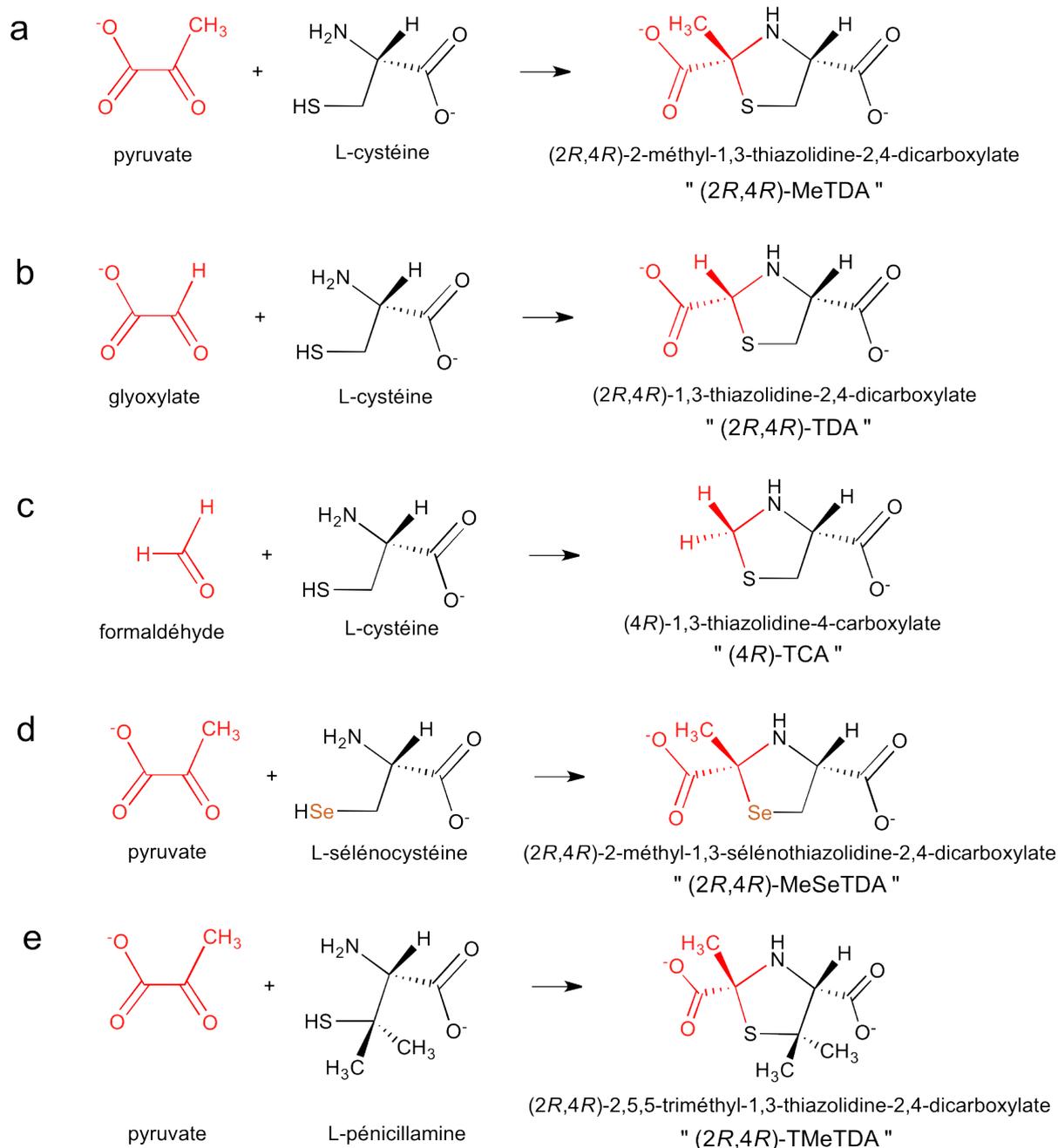
Comme mentionné précédemment [voir [Figure-IB-12](#) & paragraphe [I-12](#)], la cavité de *TmHydE* a un potentiel électrostatique positif et possède différents sites de fixation d'anions. La présence d'un groupement négativement chargé (le plus souvent un carboxylate) chez la plupart de substrats identifiés par Betz *et al.* va sans aucun doute guider ces molécules dans le site actif de HydE. On remarquera que plus la molécule est petite et plus sa charge nette est négative, plus elle produit de 5'-deoxyadenosine [[Figure-III-76-A&B](#)]. Ces ligands n'ont pas une affinité suffisante pour former de complexes stables avec l'enzyme, mais à cause de leur petite taille et leur charge négative, ils vont passer un certain temps dans la cavité. Leur groupement thiol va probablement se positionner dans le site T1, qui a une affinité élevée pour le soufre [[I-11.4.3](#)]. Ils vont donc s'orienter de la même façon que la molécule du 3-mercaptopropionate modifiée que nous avons observée dans le site actif de l'enzyme [[I-8.9](#)]. Il est à noter que le 3-mercaptopropionate est le meilleur substrat identifié par Betz [[Figure-III-76-A](#)]. De cette façon ces substrats vont être orientés de telle manière que leur groupement thiol sera à proximité directe du carbone C5' de la SAM, ce qui leur permettra, soit de déclencher le clivage du cofacteur, soit de réagir avec le radical 5'-deoxyadénosyl déjà formé [[III-5.3](#)].

Nous pensons que les résultats de Betz *et al.* ne permettent pas de conclure de façon définitive que le substrat de HydE contient un groupement thiol qui sera cible de l'attaque du radical 5'-dA•. Cependant ces expériences montrent que dans certaines conditions une telle réaction est favorable. Les observations de Betz *et al.* confirment nos conclusions sur l'affinité du site actif de HydE pour les molécules soufrées et ils montrent que les observations que nous avons pu faire dans les cristaux protéiques sont extrapolables au comportement de HydE en solution. Le fait que l'enzyme soit en mesure d'arracher un atome d'hydrogène sur le thiol de la  $\gamma$ -L-glutamyl-L-cystéine illustre bien que le substrat physiologique pourrait être plus volumineux que nos ligands dérivés de thiazolidine.

### **III-2.1.3 Clivage de la SAM en présence du dérivé (2R,4R)-MeTDA**

Dans la même période que Betz et ses collaborateurs ont publié leur étude nous avons vérifié la réactivité de l'enzyme vis-à-vis du dérivé (2R,4R)-MeTDA [[Figure-III-78-a](#)]. Ce ligand a été choisi, plutôt qu'un autre, parce qu'il possède une affinité élevée pour le site actif de HydE [[I-11.3](#)] et historiquement c'était le premier dérivé thiazolidine identifié comme ligand de HydE [[I-8.7.1](#)]. Nous avons quantifié la production de la 5'-déoxyadénosine en présence du DTT, de la L-cystéine, du (2R,4R)-MeTDA et du (2R,4R)-MeTDA + DTT. Ces essais enzymatiques *in vitro* [pour plus de détails voir article Rohac *et al.* 2016 en Annexe] ont été effectués par nos collaborateurs Alhosna Benjdia, Pauline Ruffié et Olivier Bertheau de l'équipe ChemSyBio de l'Institut Micalis. Nous avons pu confirmer les observations de Betz *et al.* sur la stimulation de la production de la 5'-deoxyadenosine par la L-cystéine et le DTT [[Figure-III-79-A](#)]. Nous pensons que le DTT est un meilleur substrat de HydE, sans doute à cause de l'absence d'une charge positive sur ce substrat. En plus le DTT possède deux fonctions thiols. Il a donc deux fois plus d'atomes hydrogènes cibles potentielles d'attaque radicalaire. Ce premier test a également montré qu'il n'y a pas d'augmentation de production

de la 5'-déoxyadénosine pour le (2*R*,4*R*)-MeTDA et donc pas d'arrachage d'un atome d'hydrogène à partir de ce ligand. Le radical 5'-déoxyadényl semble donc incapable d'arracher un atome d'hydrogène de notre dérivé de thiazolidine. Dans le cas d'un mélange (2*R*,4*R*)-MeTDA + DTT, nous avons observé une diminution d'environ 50% de la production de 5'-déoxyadénosine par rapport au DTT seul [Figure-III-79-A]. Ceci témoigne très clairement d'une compétition directe entre le dérivé thiazolidine et le DTT pour le même site de fixation dans le site actif de HydE tout à fait compatible avec nos résultats de cristallographie.



**Figure-III-78** : Ligands dérivés de la thiazolidine utilisés dans l'étude présentée dans le *Chapitre III*

Rappel de formation et de structures de ligands dérivés de la thiazolidine, utilisés dans ce chapitre. Notons que pour les ligands **a**, **b**, **d** et **e**, seuls les épimères 2*R* sont représentés, car comme nous l'avons établi dans le *Chapitre I*, l'épimère 2*S* ne se fixe pas dans le site catalytique de *TmHydE* [voir paragraphe I-11.4].

### III-2.1.4 Production enzymatique d'un nouveau dérivé adénosyl - la SAC

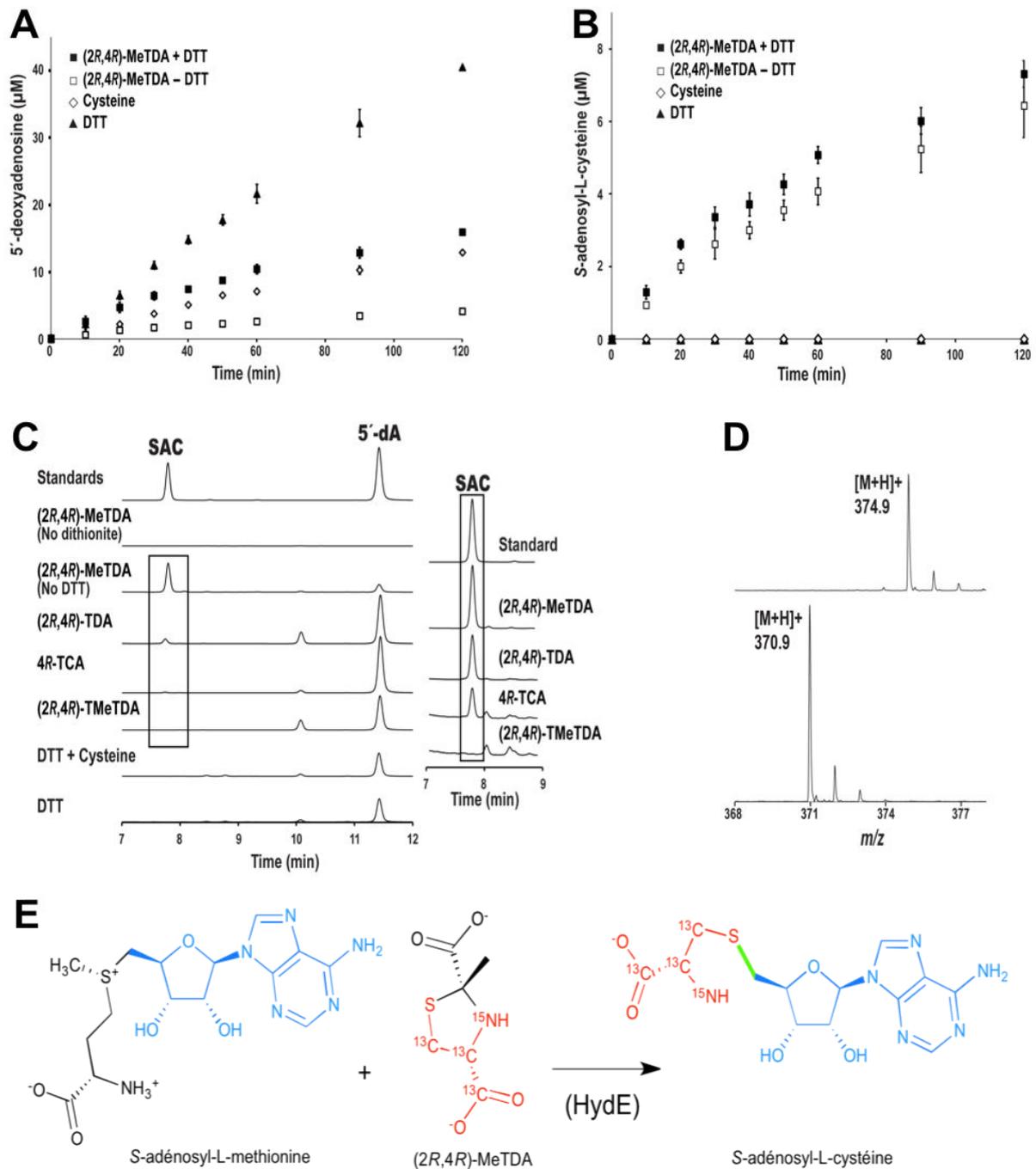
L'analyse du milieu réactionnel à la fin de la réaction de HydE avec le dérivé (2*R*,4*R*)-MeTDA par chromatographie en phase liquide (HPLC) met en évidence la formation d'un composé de nature inconnue. Cette nouvelle molécule possède un spectre d'absorption UV-visible similaire à celui des dérivés adénosyl et sa production dans le temps dépend de la présence de HydE, du dérivé (2*R*,4*R*)-MeTDA et d'un réducteur fort - ici le dithionite [Figure-III-79-B&C].

Une analyse par spectroscopie de masse utilisant la méthode MALDI-TOF nous a permis d'identifier cette molécule; il s'agit de la *S*-adénosyl-L-cystéine ou SAC ( $m/z = 370,98$ ). Lorsque le même protocole d'analyse par HPLC-MS est appliqué à la SAC commerciale ( $M=370,38$  Da), cette dernière présente le même temps de rétention et les mêmes propriétés spectroscopiques que le produit nouvellement identifié.

Lorsque nous avons utilisé un dérivé (2*R*,4*R*)-MeTDA, qui avait été préparé à partir du pyruvate et de la L-cystéine doublement marquée au  $^{13}\text{C}$  et  $^{15}\text{N}$ , la masse de ce produit a présenté un déplacement de +4 Da [Figure-III-79-D]. Ce déplacement correspond certainement à la présence de trois atomes de  $^{13}\text{C}$  et d'un atome de  $^{15}\text{N}$  issus de la L-cystéine isotopiquement marquée. Ce résultat suggère que le produit SAC est formé par la condensation de la partie 5'-dA de la SAM avec la partie L-cystéine du ligand (2*R*,4*R*)-MeTDA [Figure-III-79-E]. Cette conclusion est confirmée par la production de la sélénio-SAC lorsque HydE réagit avec le dérivé sélénié (2*R*,4*R*)-MeSeTDA [Figure-III-83-A].

Comme on peut le voir sur la Figure-III-79-C, la production de la *S*-adénosyl-L-cystéine a été observée également pour les ligands (2*R*,4*R*)-TDA (glyoxylate + L-cystéine) [Figure-III-78-b] et (4*R*)-TDA (formaldéhyde + L-cystéine) [Figure-III-78-c]. Ceci montre que le type d'aldéhyde ou de cétone qui a été utilisé pour la préparation du dérivé 1,3-thiazolidine en question n'est pas déterminant pour la réactivité de ce dernier avec HydE. Par contre, bien que le dérivé (2*R*,4*R*)-TMeTDA (pyruvate + L-pénicillamine) [Figure-III-78-e] ait une meilleure affinité pour le site actif que le ligand (4*R*)-TDA [I-11.4.2 & I-11.4.3], aucun produit dérivé d'adénosyl nouvellement formé n'a été mis en évidence pour ce ligand [Figure-III-79-C]. Ainsi il semblerait que parmi les ligands récemment identifiés, seuls ceux obtenus par la condensation de la L-cystéine avec un aldéhyde ou une cétone de faible taille sont substrats de HydE.

Les raisons pour lesquelles le dérivé (2*R*,4*R*)-TMeTDA n'est pas un substrat de HydE seront exposées dans la partie du manuscrit qui décrit l'étude de la réactivité des dérivés 1,3-thiazolidines dans les cristaux de HydE [voir paragraphe III-5.1].



**Figure-III-79** : Etude *in vitro* de l'activité de HydE en présence de dérivés 1,3-thiazolidine

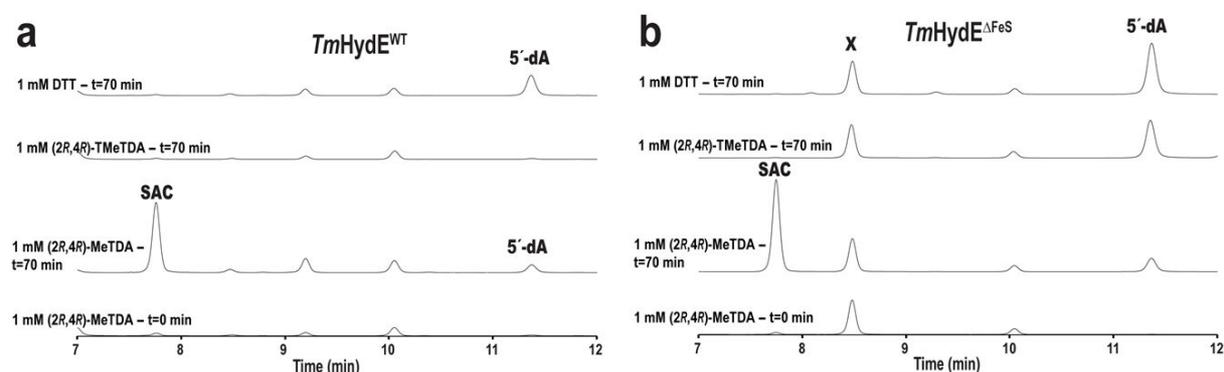
**A** - Production d'AdoH par *CaHydE* en présence de 1 mM de DTT (▲), 1 mM de L-cystéine (◇), 1 mM de (2R,4R)-MeTDA (□) ou 1 mM de (2R,4R)-MeTDA avec 1 mM DTT (■). **B** - Production de SAC par *CaHydE* en présence de 1 mM de DTT (▲), 1 mM de L-cystéine (◇), 1 mM de (2R,4R)-MeTDA (□) ou 1 mM de (2R,4R)-MeTDA avec 1 mM DTT (■). **C** - Profils d'élué HPLC en phase reverse des standards 5'-dAH et SAC comparés avec ceux obtenus lors d'incubation de *CaHydE* avec différents substrats potentiels. A droite, détail des profils d'élué correspondants au temps d'élué de la SAC obtenus pour la réaction de HydE avec les dérivés 1,3-thiazolidine testés. **D** - Analyse MS MALDI-TOF de la SAC formée au cours de la réaction de *CaHydE* avec le dérivé (2R,4R)-MeTDA (spectre du bas) et avec le dérivé (2R,4R)-MeTDA isotopiquement marqué, obtenu à partir de la L-cystéine-<sup>13</sup>C<sub>3</sub>,<sup>15</sup>N (spectre du haut). **E** - Schéma simplifié de la production de la SAC formée à partir de la partie adénosyl de la SAM (en bleu) et de la partie L-cystéine-<sup>13</sup>C<sub>3</sub>,<sup>15</sup>N du dérivé (2R,4R)-MeTDA isotopiquement marqué (en rouge). En vert est représentée la liaison C-S nouvellement formée entre le carbone C5' de l'adénosyl et l'atome S1 de ce qui était avant le cycle 1,3-thiazolidine du dérivé (2R,4R)-MeTDA.

### III-2.1.5 Impact du centre [Fe-S] auxiliaire sur la production de la SAC

Les mesures de clivage de la SAM, présentées dans les paragraphes précédents, ont été faites avec la forme native de la protéine de *CaHydE*, qui contenait deux agrégats [Fe<sub>4</sub>S<sub>4</sub>] [voir Annexe Rohac *et al.* 2016]. Dans une seconde série d'essais nous avons fait réagir la forme native de *TmHydE* (contenant les deux centres [Fe<sub>4</sub>S<sub>4</sub>]) et le triple mutant *TmHydE*  $\Delta FeS$  (ne possédant que le centre [Fe-S] SAM radical) avec le ligand (2*R*,4*R*)-MeTDA. Nous avons également inclus le dérivé (2*R*,4*R*)-TMeTDA en tant que contrôle négatif.

La figure [Figure-III-80](#) montre que la SAC est produite en présence de (2*R*,4*R*)-MeTDA par les deux échantillons indépendamment de la présence ou de l'absence de l'agrégat [Fe-S] auxiliaire. Cette fois-ci non plus nous n'avons pas observé de production de SAC à partir du composé (2*R*,4*R*)-TMeTDA. Les bosses sur les profils d'éluion témoignent de la présence de molécules autres que SAC et 5'-dAH dans le milieu réactionnel. Puisque ces mêmes bosses sont visibles à  $t_0$  et que leurs aires n'évoluent pas avec le temps nous avons conclu qu'il s'agit de contaminants qu'on trouve fréquemment dans la SAM ou qui ont été co-purifiés avec HydE (voir par exemple le composé X dans le cas de *TmHydE*  $\Delta FeS$ ).

Il est à noter qu'aucune production de SAC n'a été observée lorsque HydE a été remplacée par AlbA, une protéine à radical SAM impliquée dans la formation de liaison C-S [voir [Figure-IB-16-D](#)].



**Figure-III-80** : Profils d'éluion de HPLC en phase reverse de réaction de *TmHydE* ou *TmHydE*  $\Delta FeS$

Mesures de production de la SAC par a -*TmHydE* (avec l'agrégat [Fe-S] auxiliaire) et b - *TmHydE*  $\Delta FeS$  (sans l'agrégat [Fe-S] auxiliaire), en présence de DTT, du ligand (2*R*,4*R*)-MeTDA ou du ligand (2*R*,4*R*)-TMeTDA. Le protocole expérimental est le même que pour la première expérience effectuée avec la protéine *CaHydE*. Les profils d'éluion visualisés correspondent à  $t_0$  (t=0min) pour le ligand (2*R*,4*R*)-MeTDA et à  $t_{final}$  (t=70min) pour les trois ligands testés. L'accumulation dans le temps de la SAC est observée uniquement pour le ligand (2*R*,4*R*)-MeTDA et cela indépendamment de la présence de l'agrégat [Fe-S] auxiliaire.

Les expériences décrites ci-dessus ont prouvé que la production de la SAC par HydE à partir de la SAM et de certains ligands dérivés de thiazolidine ne dépend pas de la présence de l'agrégat [Fe-S] auxiliaire et que cette réaction, qui n'a pas de précédent chez les protéines à radical SAM, serait propre à HydE.

### III-2.2 Suivi par spectroscopie RMN résolue dans le temps de la formation de pyruvate libre au cours de la production enzymatique de la SAC

Dans les paragraphes précédents nous avons montré que la molécule de SAC, produite par HydE, est composée de la partie 5'-dA issue de la SAM et de la partie L-cystéine appartenant au ligand (2R,4R)-MeTDA [Figure-III-79-E]. Puisque dans les protéines à radical SAM le clivage du cofacteur produit le radical 5'-dA• et la L-méthionine, nous pouvons supposer que cette dernière fait partie des produits de la réaction de HydE-SAM avec le ligand (2R,4R)-MeTDA [Figure-III-83-A] (voir plus loin dans le texte). Nous avons voulu éclaircir le devenir de la partie pyruvate du ligand (2R,4R)-MeTDA au cours de la réaction. Dans ce but nous avons fait réagir HydE avec un ligand marqué au  $^{13}\text{C}$  sur le méthyl. Cette molécule a été préparée en mélangeant de la L-cystéine avec un équivalent de pyruvate marqué au  $^{13}\text{C}$  sur le carbone C3 (son groupement méthyl) [Figure-III-81-a]. Nous avons ensuite enregistré, à des intervalles réguliers, des spectres RMN 1D du proton filtrés au  $^{13}\text{C}$  pour suivre sélectivement l'évolution dans le temps de résonances des noyaux d'hydrogènes couplés à des atomes  $^{13}\text{C}$  [Figure-III-81-b3] [voir Annexe Rohac *et al.* 2016]. De cette manière, seuls les signaux d'atomes d'hydrogène qui sont liés à un atome  $^{13}\text{C}$  sont visibles; les protons de la protéine et des autres constituant du milieu réactionnel ne le sont pas.

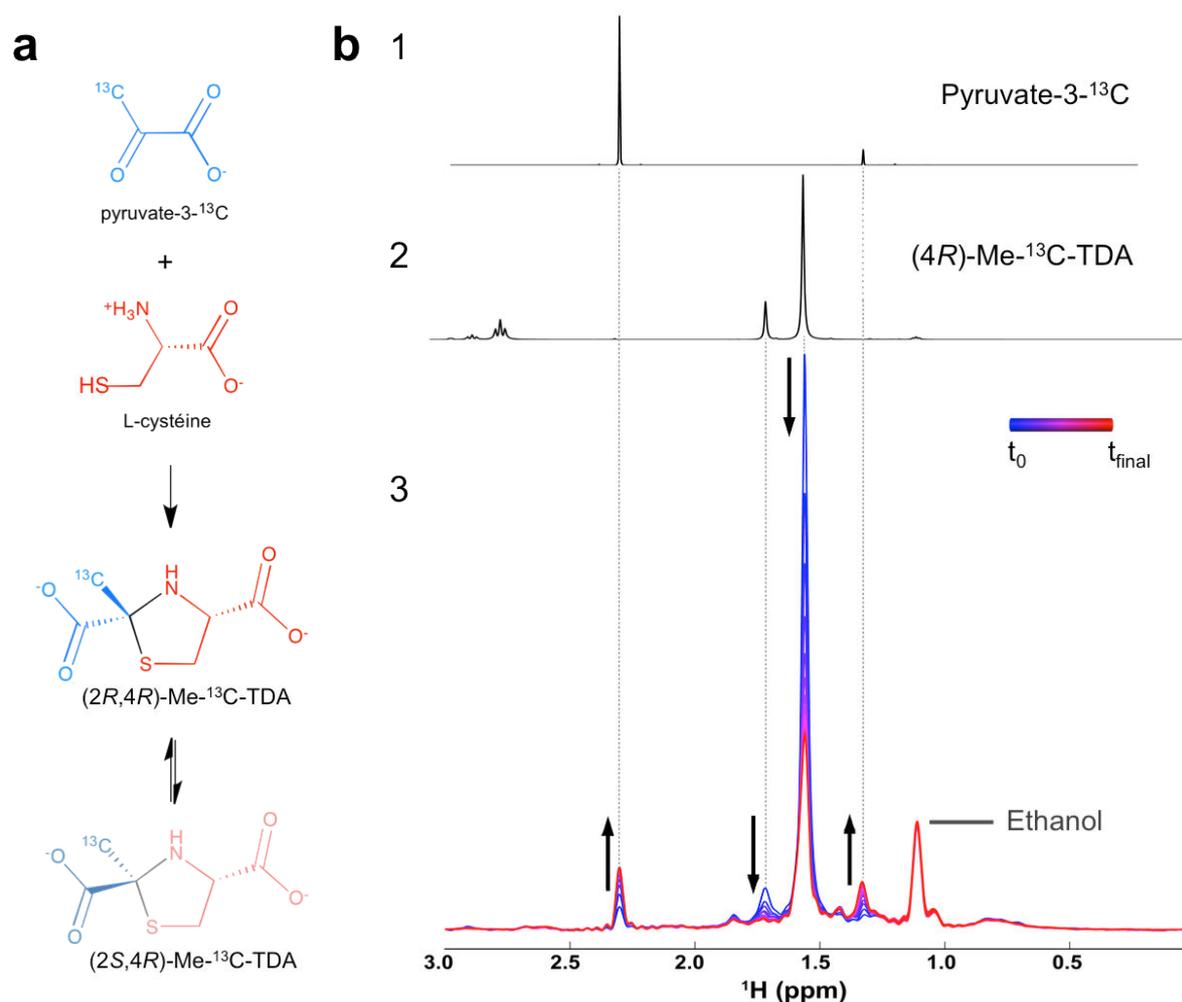
La Figure-III-81-b1 montre le spectre de référence du pyruvate avec les résonances des noyaux d'hydrogènes liés de façon covalente au méthyl  $^{13}\text{C}$  du pyruvate. Le pic majoritaire ( $\delta$  2.31 (s)) correspond aux hydrogènes du méthyl de la forme  $\alpha$ -cétoacide du pyruvate, tandis que le pic minoritaire ( $\delta$  1.33 (s)) correspond aux hydrogènes du méthyl de la forme hydratée du pyruvate<sup>11</sup>. Après l'ajout de L-cystéine (un équivalent de pyruvate total) les deux pics disparaissent pour laisser place à deux autres pics, dont les résonances correspondent aux groupements méthyl des deux épimères sur le C2 du dérivé (4R)-MeTDA [Figure-III-81-b1 & Figure-I-56-A,B&E]. Les protons du méthyl  $^{13}\text{C}$  de l'épimère (2S,4R)-MeTDA résonnent à  $\delta$  1.71 (s), tandis que ceux du méthyl  $^{13}\text{C}$  de l'épimère (2R,4R)-MeTDA résonnent à  $\delta$  1.56 (s).

Le suivi RMN de la cinétique de la réaction catalysée par HydE montre une consommation du ligand (2R,4R)-MeTDA au cours de la réaction [Figure-III-81-b3]. Cette consommation se traduit par une diminution des résonances spécifiques des groupements méthyl des deux épimères (2S,4R)- et (2R,4R)-MeTDA. Comme nous l'avons vu dans le **Chapitre I**, seul le dernier épimère est capable de former un complexe stable avec HydE. Cependant, puisque un échange rapide entre les deux épimères existe en solution [Figure-I-55], la diminution de la concentration d'un d'eux aura une répercussion sur la concentration de l'autre.

Il est intéressant de noter que parallèlement à la consommation du substrat nous observons une apparition graduelle de deux pics dont les valeurs de résonances correspondent exactement aux groupements méthyl de la forme  $\alpha$ -cétoacide et de la forme hydratée du pyruvate [Figure-III-81-b3]. Ceci suggère qu'au cours de la synthèse enzymatique de SAC il y a formation de pyruvate à partir du substrat. Au cours de cette expérience nous avons eu un problème de contamination de l'échantillon protéique par de l'éthanol [Figure-III-81-b3]. Le filtre  $^{13}\text{C}$  appliqué sur les résonances des protons a fait baisser le pic du méthyl de l'éthanol à environ 1% de la concentration de l'éthanol dans le milieu réactionnel (abondance naturelle du

$^{13}\text{C}$  est de 1.109%). Nous avons utilisé la résonance de ce méthyl comme un standard interne. L'aire de cette résonance n'évolue pas au cours de l'expérience. Ceci confirme qu'il n'y a pas de détérioration de l'homogénéité du champ magnétique dans l'échantillon. Ceci prouve également que les variations qu'on observe pour les résonances des méthyls du substrat ou du produit sont bien dues à la réaction enzymatique en cours.

En conclusion, l'analyse HPLC-MS des produits de la réaction observée dans HydE a prouvé que la partie L-cystéine du dérivé thiazolidine se retrouve dans la SAC, tandis que l'étude RMN présentée ci-dessus a montré que la partie pyruvate du substrat est libérée en solution sous la forme de pyruvate libre. Ainsi les produit de la réaction de HydE avec le dérivé (2*R*,4*R*)-MeTDA sont la SAC, la L-méthionine et le pyruvate.



**Figure-III-81** : Etude par RMN  $^1\text{H}$  de la libération de pyruvate au cours de la production enzymatique de la SAC

**a** - Préparation des ligands (4*R*)-Me- $^{13}\text{C}$  -TDA marqués au  $^{13}\text{C}$  sur le méthyl à partir d'un mélange équimolaire de L-cystéine et de pyruvate marqué au  $^{13}\text{C}$  uniquement sur le carbone C3 (son groupement méthyl). Rappelons que de cette manière les deux épimères sur le C2 sont formés et qu'un échange entre les deux existe en solution. Seul l'épimère (2*R*,4*R*)-Me- $^{13}\text{C}$  -TDA peut se fixer dans le site actif de HydE. **b1**: Spectre RMN 1D  $^1\text{H}$  filtré au  $^{13}\text{C}$  de référence du pyruvate-3- $^{13}\text{C}$ . **b2**: Spectre RMN 1D  $^1\text{H}$  filtré au  $^{13}\text{C}$  de référence des deux épimères sur le C2 du dérivé thiazolidine (4*R*)-Me- $^{13}\text{C}$  -TDA. **b3**: Superposition de spectres RMN 1D du proton filtrés au  $^{13}\text{C}$  enregistrés toutes les deux minutes au cours de la réaction de CaHydE avec le ligand (2*R*,4*R*)-Me- $^{13}\text{C}$  -TDA. Les flèches verticales illustrent les tendances d'évolution dans le temps de résonances observées. Il est à noter que ces mesures ont été effectuées en anaérobiose stricte [voir Annexe Rohac *et al.* 2016].

### III-3 Etude par cristallographie aux rayons X de la réactivité de *TmHydE* $\Delta FeS$ envers le ligand (2*R*,4*R*)-MeSeTDA

Les études de suivi de réactions enzymatiques dans les cristaux sont assez rares. Une des raisons possibles est que certaines réactions nécessitent, pour arriver à leur complétude, des changements du repliement de la protéine, ce qui risque de perturber l'empilement cristallin et par la suite de détruire le cristal. Une autre raison est la difficulté à synchroniser la réaction dans toutes les molécules de protéine dans le cristal. Dans le cas de HydE, nous avons vu que le ligand se trouve dans une large cavité remplie de molécules d'eau. Le triple mutant *TmHydE*  $\Delta FeS$ , ne pouvant pas contenir l'agrégat [Fe-S] auxiliaire, possède deux canaux d'accès  $E_1$  et  $E_3$  [Figure-I-37-B], qui permettent la diffusion du substrats dans le site actif lors du trempage de cristaux dans les solution de cryoprotection contenant des ligands. Dans le *Chapitre I* nous avons vu que ce triple mutant, lorsque reconstitué *in vivo*, se co-purifie avec une quantité stœchiométrique de la SAM. Cet échantillon produit facilement des cristaux qui diffractent à des résolutions au-delà de 1,2 Å. Puisque nous avons établi que le second centre [Fe-S] n'est pas impliqué dans la production enzymatique de la SAC par HydE les cristaux du triple mutant *TmHydE*  $\Delta FeS$  contenant de la SAM et diffractant à haute résolution s'avèrent particulièrement adaptés pour une étude cristallographique de la réactivité de HydE avec les ligands dérivés de thiazolidine. Aussi, le fait que l'initiation de la réaction dans les cristaux est très lente permet, dans une certaine mesure, sa synchronisation.

#### III-3.1 Réactivité de HydE avec le ligand (2*R*,4*R*)-MeSeTDA dans les cristaux protéiques

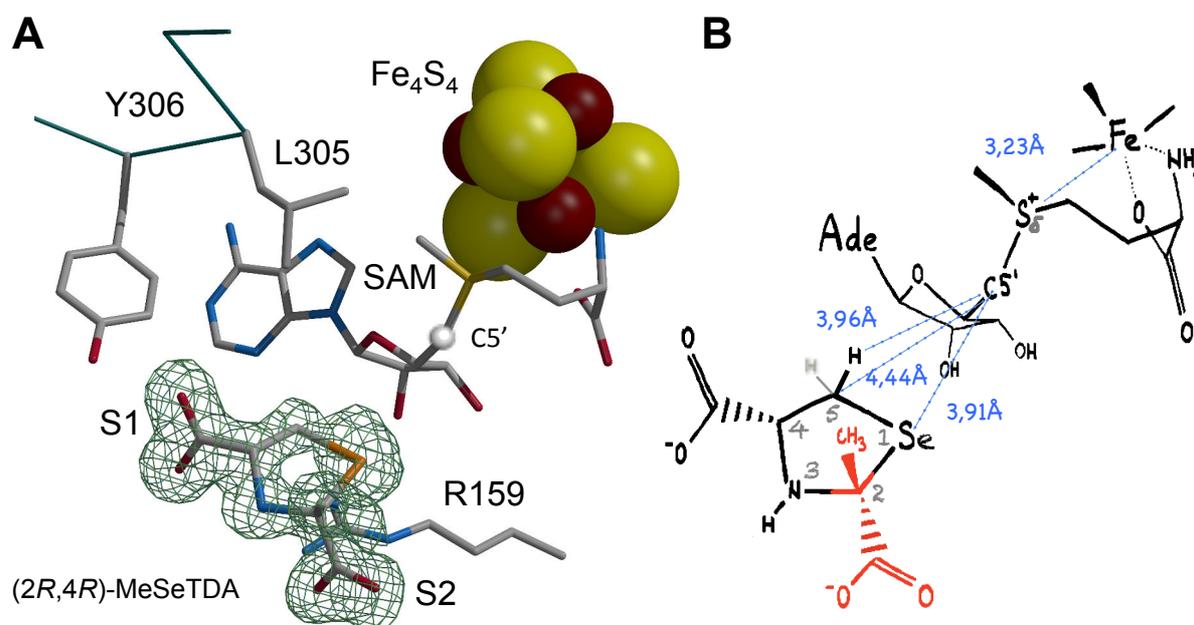
Pour cette étude nous avons choisi le ligand (2*R*,4*R*)-MeSeTDA, parce que le sélénium peut être facilement suivi dans les cartes de densité électronique à cause de son nombre d'électrons élevé et de son signal anomal fort. Les expériences de contrôle avec le substrat soufré ont montré que les deux chalcogènes ont le même comportement lors de la réaction enzymatique dans HydE. De ce fait, dans l'analyse structurale par cristallographie aux rayons X, tout ce qui s'applique au substrat sélénié peut s'appliquer aussi au substrat soufré. Seuls les paramètres cinétiques de la réaction sont légèrement différents dans le sens où la réaction avec le substrat sélénié semble être plus rapide.

Pour suivre la réaction enzymatique nous avons, dans un premier temps, produit des cristaux de *TmHydE*  $\Delta FeS$  contenant de la SAM. Ces cristaux ont ensuite été incubés dans une solution de cryoprotection contenant le ligand avant d'être lavés de tout excès de ligand par trempage dans une solution de cryoprotection sans ligands. La réaction a été déclenchée par le trempage de cristaux dans une solution de cryoprotection contenant un excès de dithionite. Les cristaux qui ont subi ce traitement ont ensuite été congelés à l'intérieur de la boîte à gants à différents temps [MM-4.1]. De cette manière nous avons obtenus une série de structures cristallographiques qui correspondent aux différents états de la réaction. Le réarrangement des atomes de la SAM ou du substrat a été suivi en calculant des cartes de densité électronique double différence  $F_o(t_x) - F_o(t_0)$ . Notons que  $F_o(t_0)$  est le facteur de structure correspondant au jeu de données du cristal congelé à  $t=0$  et  $F_o(t_x)$  est celui du cristal

congelé à  $t=x$  minutes. Le nom de chaque cristal, la résolution du jeu de données, ainsi que la composition du site actif sont détaillés dans le [Tableau-III-2](#).

### III-3.1.1 Disposition du site actif à $t_0$

La résolution du jeu de données cristallographique **X2**, obtenu à partir d'un cristal contenant la SAM intacte et le ligand (2*R*,4*R*)-MeSeTDA [[Figure-III-82-A](#)], est de 1,27 Å [[Tableau-III-2](#)]. Les atomes Se1 et C5 du cycle thiazolidine se trouvent dans la zone où l'attaque radicalaire, portée par l'atome C5' du radical 5'-déoxyadénosyl, devrait avoir lieu. Le C5' se situe à 3,91 Å et à 4,44 Å des atomes Se1 et C5, respectivement. L'atome d'hydrogène du substrat le plus proche du carbone C5' est l'hydrogène *pro-R* porté par l'atome C5 du cycle thiazolidine. Il se situe à 3,96 Å de l'atome C5' de la SAM [[Tableau-III-01](#)][[Figure-III-82-B](#)]. L'angle entre les atomes C5'-Se1-C2 est de 165,4°, tandis que celui entre les atomes C5'-Se1-C5 est de 92,4°. Le sulfonium de la SAM se trouve à 3,23 Å du Fe unique de l'agrégat [Fe<sub>4</sub>S<sub>4</sub>] 'SAM radical'.



**Figure-III-82** : Disposition du site actif de *TmHydE*  $\Delta$ FeS à  $t_0$  avec la SAM et le ligand (2*R*,4*R*)-MeSeTDA

**A** - En vert foncé, carte de densité électronique *Fo-Fc* "omit" contourée à 3  $\sigma$  et visualisée à 2 Å autour du ligand (2*R*,4*R*)-MeSeTDA (cristal: **X2**, résolution du jeu: 1,27 Å). **B** - Schéma de la disposition du substrat par rapport à la SAM. La partie pyruvate du substrat est colorée en rouge. La numérotation des atomes du cycle thiazolidine est représentée par des chiffres gris.

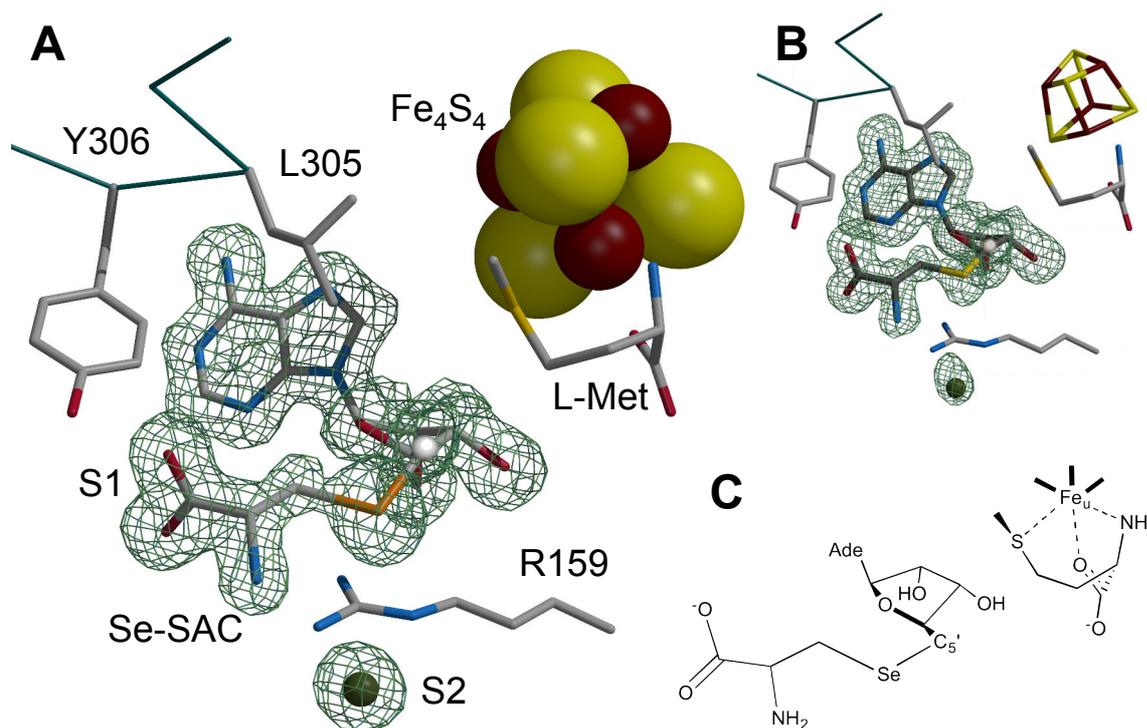
### III-3.1.2 Disposition et contenu du site actif à $t_{final}$

La résolution du jeu de données cristallographique **X4** est de 1,32 Å [[Tableau-III-2](#)]. Ce jeu a été obtenu en congelant un cristal après plus de 120 minutes depuis la réduction de l'agrégat [Fe<sub>4</sub>S<sub>4</sub>] 'SAM radical'. La carte de densité électronique *Fo-Fc* "omit" montre clairement qu'une réaction dans les cristaux protéiques et qu'une molécule que l'on a pu

aisément modéliser comme de la Se-SAC se trouve maintenant dans l'espace qui était occupé précédemment par la SAM et le ligand (2*R*,4*R*)-MeSeTDA [Figure-III-83-A].

La carte de densité électronique, obtenue par la congélation d'un cristal au bout de deux heures après le début de la réaction avec le ligand (2*R*,4*R*)-MeTDA (cristal **X5**), superposée avec le modèle de la SAC obtenu par la co-cristallisation de HydE avec la SAC commerciale (cristal **X6**) montre que celle-ci commerciale adopte exactement la même conformation que la molécule produite par l'enzyme [Figure-III-83-B]. Cette observation confirme nos conclusions de l'analyse du produit de réaction; c'est-à-dire que HydE produit de la SAC à partir de la partie 5'-dA de la SAM et de la partie L-cystéine du ligand (2*R*,4*R*)-MeTDA [Figure-III-83-C].

Ces données cristallographiques confirment le clivage de la SAM dans le cristal protéique et la production de la L-méthionine et de la Se-SAC. La création d'une nouvelle liaison C-Se est bien visible. A la fin de la réaction, la partie pyruvate n'est plus présente dans les cartes de densités électroniques et le site S2, qui au début de la réaction abritait un groupement carboxylate, est occupé par un ion chlorure. Ceci est en accord avec les résultats d'analyse par RMN avec un suivi dans le temps, qui prouve qu'au cours de la formation de la SAC il y a bien libération du pyruvate [Figure-III-81-b1].



**Figure-III-83** : Disposition du site actif de *TmHydE*  $\Delta$ FeS à  $t_{\text{final}}$  contenant le produit de la réaction

**A** - En vert foncé, carte de densité électronique *Fo-Fc* "omit" contournée à 3  $\sigma$  et visualisée à 2 Å autour de la séléno-SAC et du site S2. (cristal: **X4**, résolution du jeu: 1,32 Å). L'atome C5' est représenté par une sphère blanche, tandis que le chlorure dans le site S2 par une sphère verte. **B** - En vert foncé, carte de densité électronique *Fo-Fc* "omit" contournée à 4  $\sigma$  et visualisée à 3,5 Å autour de la SAC, obtenue par la réaction enzymatique de HydE avec le substrat (2*R*,4*R*)-MeTDA (cristal: **X5**, résolution du jeu: 1,17 Å). En gris foncé, le modèle de la SAC placé à l'intérieur de cette densité électronique, qui correspond à la conformation de la SAC issue de co-cristallisation de *TmHydE* $\Delta$ FeS avec la SAC commerciale (cristal: **X6**, résolution du jeu: 1,49 Å) **C** - Schéma de la conformation de la Se-SAC produite par voie enzymatique dans le site actif de HydE.

### III-3.1.3 Analyse des mouvements dans le site actif au cours de la réaction

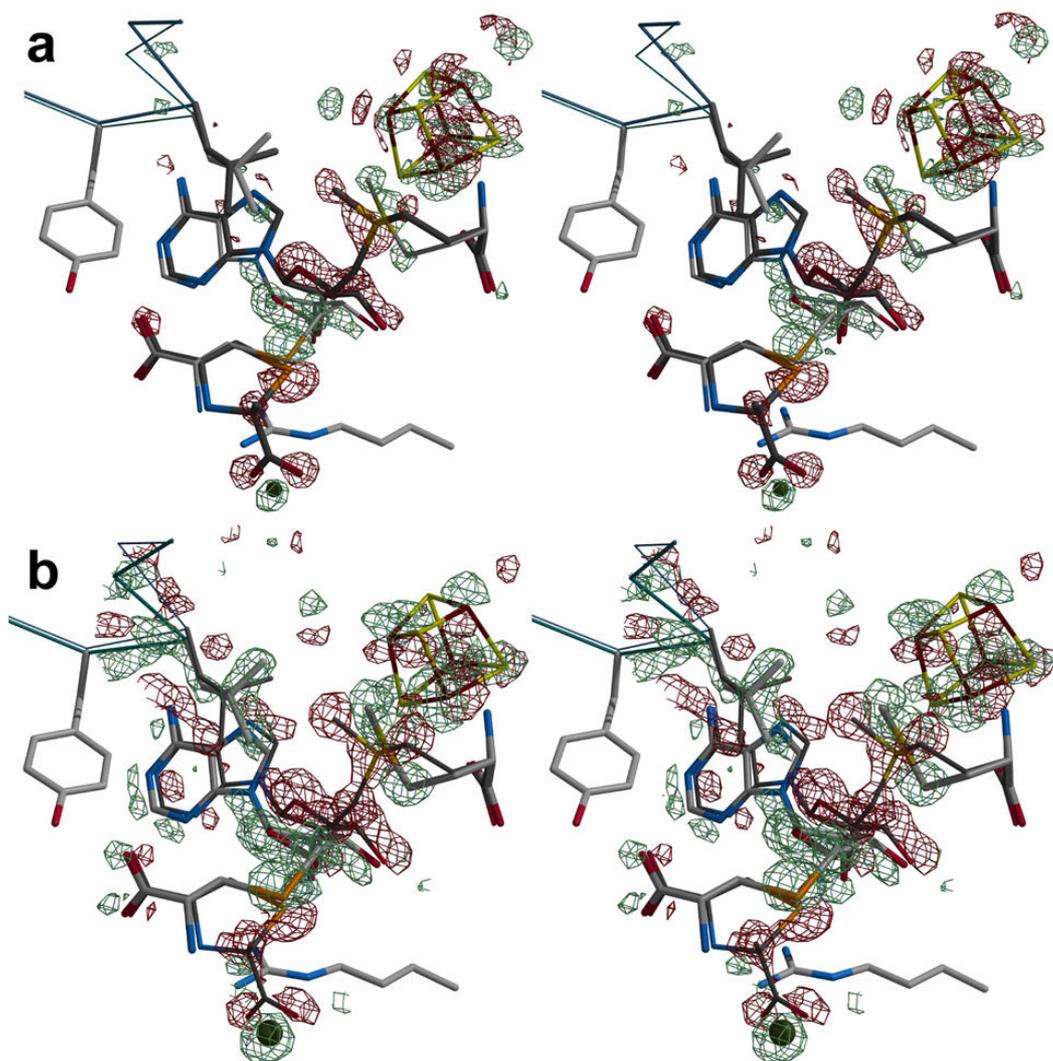
Les cartes de densité électronique double différence  $F_o(t_x)-F_o(t_0)$  (avec  $x = 45$  ou  $120$  minutes) illustrent que tous les changements structuraux au cours de la réaction se situent dans une sphère avec un rayon de  $12 \text{ \AA}$  autour de l'atome C5' de la SAM [Figure-III-84-A&B]. Cette sphère enveloppe le centre  $[\text{Fe}_4\text{S}_4]$  'SAM radical', le cosubstrat SAM, le ligand (2*R*,4*R*)-MeSeTDA et la boucle  $\text{Y}^{303}\text{NLY}^{306}$ . Cette dernière contient le résidu Leu305, dont la chaîne latérale interagit avec le méthyl de la SAM<sup>12,13</sup> [voir paragraphe I-10.2.2]. La production enzymatique de la SAC n'affecte pas le reste de la structure de la protéine.

Les réarrangements les plus notables se situent sur un axe qui relie l'atome C2 du substrat à l'atome de Fe unique de l'agrégat  $[\text{Fe}_4\text{S}_4]$ . Ces réarrangements peuvent être divisés dans les mouvements suivants:

- i)* Le clivage de la SAM, qui se traduit par l'éloignement mutuel des atomes C5' et  $\text{S}_8$  de la SAM. La distance entre ces deux atomes passe de  $1,81 \text{ \AA}$  à  $3,91 \text{ \AA}$  [Tableau-III-01].
- ii)* En même temps on assiste à la formation d'une nouvelle liaison C-Se entre les atomes C5' de la SAM et Se1 du ligand. La distance entre ces deux atomes passe de  $3,91 \text{ \AA}$  à  $2,08 \text{ \AA}$ .
- iii)* Nous observons également le clivage de la liaison Se1-C2 [ $d_{\text{initiale}}(\text{Se1-C2}) = 1,99 \text{ \AA}$ ] et la rupture du cycle thiazolidine.
- iv)* Le dernier changement important est la libération du fragment pyruvate du ligand (2*R*,4*R*)-MeSeTDA, suivi du remplacement du groupement carboxylate, initialement présent dans le site S2, par un ion chlorure [Figure-III-83-A&B et Figure-III-84-A&B].

Un changement plus subtil, cependant bien visible, est le réarrangement de la chaîne latérale du résidu Leu305 [Figure-III-84-B], qui, après le clivage de la SAM, adopte la même conformation que celle dans la structure avec le 5'-dAH et L-Met (code PDB: 3IIX).

Puisque la réaction en solution et dans les cristaux protéiques semble être particulièrement lente nous avons tenté de piéger des espèces intermédiaires de cette réaction en congelant les cristaux à des temps très court (2-5 min), moyennement longs (15-45 min) et très longs après le début de la réaction (120-210 min). Les temps très courts ne présentent aucune modification de la structure. La composition du site actif est identique à celle du  $t_0$ , malgré la présence du dithionite en excès. Les temps moyennement longs présentent les mêmes réarrangements que les temps très longs [Figure-III-84-A]. En effet, toutes les cartes de densité électronique double différence  $F_o(t_x)-F_o(t_0)$  montrent que les temps intermédiaires contiennent un mélange de deux états : l'état  $t_0$  [X2, Tableau-III-2, Figure-III-82-A] et l'état  $t_{\text{final}}$  [X4, Tableau-III-2, Figure-III-83-A]. Les proportions entre ces deux états passent, au cours de la réaction, progressivement de 1:0 à 0:1. Ceci suggère qu'avec cette approche il est impossible de piéger d'autres états intermédiaires que les substrats SAM + (2*R*,4*R*)-MeSeTDA ou les produits Se-SAC + L-Met. Plus important encore, puisque les temps de congélation court ne montrent aucun clivage, tandis que les temps plus longs contiennent directement des produits de la réaction, ceci laisse penser que l'étape limitante de cette réaction est le clivage de la SAM et la génération du radical 5'-dA•.



**Figure-III-84** : Vues stéréoscopiques de cartes de densité électronique double différence  $F_o(t_x)-F_o(t_0)$  montrant un suivi de la conversion des substrats (2R,4R)-MeSeTDA et SAM en produits Se-SAC et L-Met

L'orientation des structures est la même que dans les images Figure-III-06 & -07. En rouge et vert, carte de densité électronique double différence  $F_o(t_x)-F_o(t_0)$  contournée à  $\pm 5 \sigma$  et visualisée dans un rayon de 12 Å autour de l'atome C5'. Cette carte permet de visualiser les différences dans la position d'atomes entre deux jeux de données cristallographiques différents, à condition que ces jeux soient issus de cristaux isomorphes. La densité rouge (négative) correspond à la matière qui était présente dans une position donnée à  $t_0$ , c'est-à-dire au début de la réaction, mais qui n'est plus à cette position à un temps avancé au cours de la réaction -  $t_x$ . De l'autre côté, la densité verte (positive) correspond à la matière qui est présente dans une position donnée à  $t_x$ , mais qui n'était pas dans cette position au temps de congélation initial ( $t_0$ ). Donc si un atome est entouré par les densités rouge et verte, cela signifie que cet atome subit, au cours de la réaction, une translation de l'endroit avec la densité rouge vers l'endroit avec la densité verte. Cette carte nous permet de suivre l'accumulation progressive des produits Se-SAC et L-méthionine au détriment des substrats SAM et (2R,4R)-MeSeTDA. Elle permet également d'analyser les mouvements qui accompagnent cette conversion. Les phases utilisées pour la transformée de Fourier inverse ont été calculées à partir du modèle affiné correspondant à la structure à  $t_0$  (cristal: X2) [MM-4.2.5] - ici représenté par des bâtons gris foncé. L'utilisation de modèles correspondant à des temps de congélation intermédiaires conduit à des cartes très similaires. La composition du site actif à  $t_{2h00}$  (cristal: X4) est représentée par des bâtons gris clair. L'image **a** correspond à la carte de densité électronique double différence avec les coefficients  $F_o(t_{45min})-F_o(t_0)$  et **b** à celle avec les coefficients  $F_o(t_{2h00})-F_o(t_0)$ . Les deux images montrent les mêmes mouvements, à la seule différence, que l'intensité du signal est plus forte dans le cas du temps de congélation plus avancé, c'est-à-dire à  $t_{2h00}$  - **b**. Ceci est dû simplement au fait, que à  $t_{45min}$  une partie de substrats n'a pas encore été transformée en produits. Les changements structuraux bien visibles à  $t_{45min}$  sont les réarrangements importants comme le clivage de la SAM et la création de la liaison C5'-Se. A  $t_{2h00}$  nous pouvons voir en plus des changements plus subtils, comme le déplacement de la chaîne latérale du résidu Leu305 ou encore une légère rotation de cycle de l'adénine.

### **III-3.1.4 Absence de détection d'intermédiaires de la réaction**

Les structures cristallographiques présentées dans les paragraphes précédents complètent nos analyses du clivage de la SAM en solution et celles par RMN. Ces structures montrent clairement qu'à la fin de la réaction, HydE contient les produits L-méthionine et la SAC (ou Se-SAC lorsque un substrat sélénié a été utilisé) [Figure-III-83-A&B]. Le suivi dans le temps par les expériences de RMN du devenir de la partie pyruvate du ligand montre que ce fragment est effectivement libéré à la fin de la réaction sous forme de pyruvate [Figure-III-81-b3]. L'apparition des résonances, autres que celles des deux formes du pyruvate, n'a pas été détectée au cours de la réaction suivi par RMN. Cela dit, il faut tenir compte du fait que la SAC, ses précurseurs et d'éventuels intermédiaires restent liés à la protéine au cours de la réaction. Les phénomènes de relaxation des signaux RMN de la protéine et des molécules qu'elle contient empêchent l'observation de ces signaux dans l'expérience que nous avons effectuée. En d'autres termes, nous avons regardé uniquement les signaux provenant des molécules libres en solution.

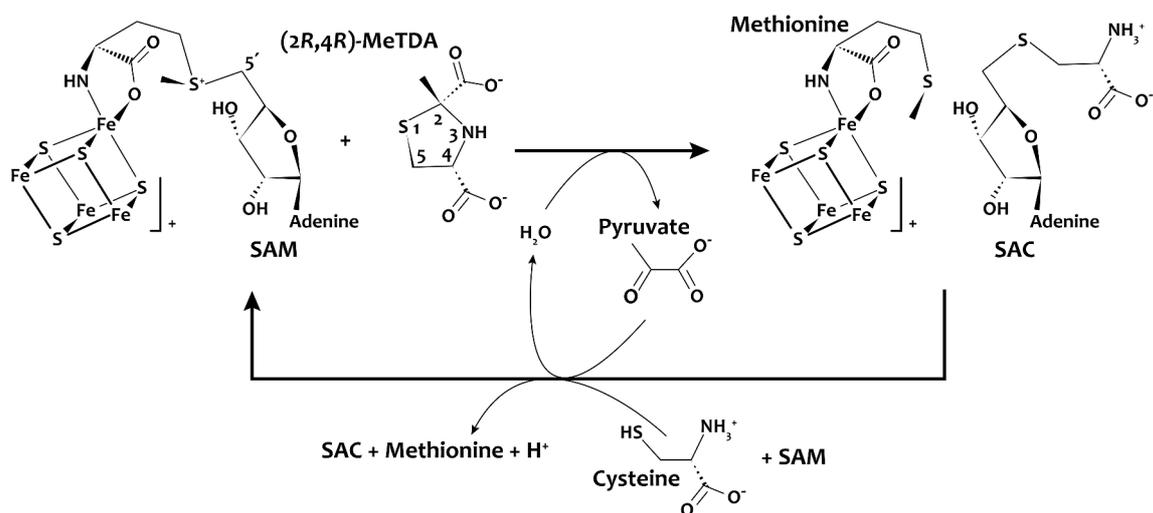
Les expériences de cinétique dans les cristaux n'ont pas abouti à l'isolation d'espèces intermédiaires. Nous avons également tenté de piéger des espèces radicalaires en solution par congélation à de différents temps et procéder à une analyse par RPE. Cette approche est une méthode de choix pour la détection et la caractérisation d'espèces à un électron non apparié. Elle a permis d'identifier de certaines espèces radicalaires produites par les protéines à radical SAM<sup>14</sup>. Cependant, cette approche n'a pas été fructueuse dans notre cas. Ces observations signifient que ces données expérimentales ne peuvent fournir plus de détails sur le mécanisme de la conversion des substrats SAM + (2*R*,4*R*)-Me(Se)TDA en produits (Se)-SAC + L-méthionine + pyruvate. En effet, l'accumulation des produits en solution et dans les cristaux semble être très lente - de l'ordre de quelques dizaines de minutes par turnover. Puisqu'aucune espèce intermédiaire n'a été détectée on peut en déduire que la synchronisation de la réaction dans les molécules de HydE est faible. Ainsi la concentration des espèces intermédiaires plus ou moins stables reste, au cours de la réaction, en dessous de la limite de détection des techniques utilisées. Si en plus ces espèces ont une durée de vie très courte, leur détection devient en pratique impossible.

## **III-4 Formation d'une liaison C-S(e) par un processus radicalaire**

### **III-4.1 Le rôle du pyruvate dans la production enzymatique de la S-adénosyl-L-cystéine**

La libération du pyruvate à la fin de la production de la SAC est un phénomène surprenant. Le pyruvate est nécessaire pour la transformation de la L-cystéine en ligand (2*R*,4*R*)-MeTDA, qui forme un complexe stable avec la protéine HydE. En absence de pyruvate, la cystéine est un substrat de l'enzyme; HydE arrive à lui arracher un atome d'hydrogène à partir de son groupement thiol [III-2.1.1], mais elle ne forme pas de complexe stable avec la protéine. Par contre en présence de pyruvate, il n'y a plus d'hydrogène sur le thiol, car le dernier a été transformé en thioéther. Ceci a un impact drastique sur la réaction: cette fois-ci une nouvelle liaison C-S est créée, tandis que le pyruvate est recyclé. Ainsi le

pyruvate agit comme un véritable cofacteur, transformant le thiol de la L-cystéine en thioéther, permettant ainsi la condensation de la partie adénosyl de la SAM avec la L-cystéine. A la fin de la réaction, le pyruvate libéré peut interagir avec une nouvelle molécule de la L-cystéine pour refaire un autre cycle [Figure-III-85]. Pour avoir la preuve expérimentale que le pyruvate puisse fonctionner de façon catalytique il faudrait refaire nos essais de suivi d'accumulation de la SAC [III-2.1.4], cette fois-ci en présence d'un excès de L-cystéine et d'une quantité sous-stoechiométrique de pyruvate.

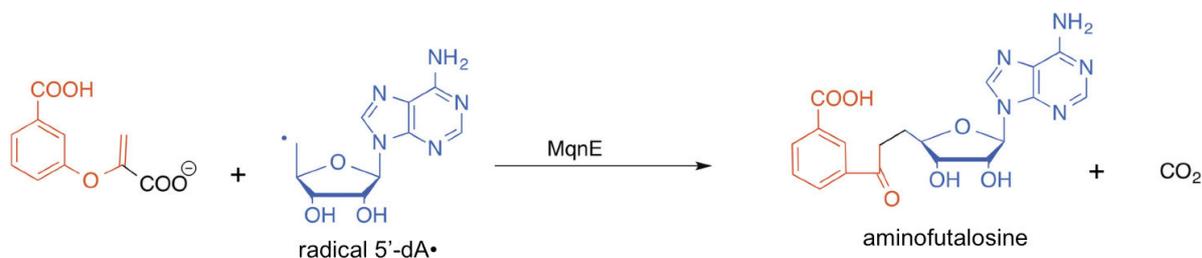


**Figure-III-85** : Pyruvate comme cofacteur dans la synthèse de la SAC à partir de la SAM et de la L-cystéine

Nos données expérimentales prouvent que les trois produits libérés au cours de la réaction dans HydE sont la SAC, la L-méthionine et le pyruvate. Le pyruvate est aussi le précurseur du substrat (2R,4R)-MeTDA qu'il forme par la réaction de condensation avec de la L-cystéine. Cet acide aminé sera intégré dans le produit SAC, tandis que le pyruvate sera régénéré. Le pyruvate pourrait donc agir de façon catalytique et servir comme un véritable cofacteur dans la production de la SAC à partir de la L-cystéine et de la partie adénosyl de la SAM.

#### III-4.2 Attaque radicalaire du 5'-dA• sur l'atome de S(e) d'un sélénoéther/thioéther

Comme nous l'avons vu dans les paragraphes précédents, HydE est capable d'arracher un atome d'hydrogène à partir d'un thiol d'une petite molécule, comme par exemple la cystéine, ou bien de réagir directement sur le thioéther d'un dérivé 1,3-thiazolidine. A ce jour et à notre connaissance la seule protéine à radical SAM qui n'est pas impliquée dans l'abstraction d'un atome d'hydrogène est l'aminofutalosine synthase MqnE<sup>15</sup>. Cette enzyme catalyse une addition directe du radical 5'-dA• sur une double liaison de l'acide 3-[(1-carboxyvinyl)oxy]benzoïque dans la voie de la biosynthèse de la ménaquinone (vitamine K<sub>2</sub>) chez *Streptomyces coelicolor*<sup>16</sup> [Figure-III-86]. Une addition radicalaire similaire a été observée pour l'activase de la pyruvate formate lyase - la PFL-AE, qui normalement arrache l'atome d'hydrogène pro-S sur un résidu glycine de la pyruvate formate lyase. Lorsque ce résidu est remplacé par une déhydroalanine, la PFL-AE catalyse une addition du radical 5'-dA• sur la double liaison de la déhydroalanine, formant ainsi une liaison C5'-C<sub>β</sub><sup>17</sup>. En synthèse de polymères, les radicaux alkyls sont souvent utilisés pour la formation des liaisons C-C qui procède par une addition du radical alkyl sur la double liaison d'un système non saturé<sup>18</sup>.



**Figure-III-86** : Réaction catalysée par l'aminofutalosine synthase MqnE

La protéine à radical SAM MqnE catalyse une addition directe du radical 5'-dA• sur une double liaison de l'acide 3-[(1-carboxovinyl)oxy]benzoïque (à gauche de l'image) pour former l'aminofutalosine et le  $\text{CO}_2$ .

D'ailleurs cette chimie est utilisée par certains protéines à radical SAM, comme par exemple PqqA<sup>19</sup>, Cfr<sup>20-22</sup> et RlmN<sup>20,23,24</sup>, qui sont impliquées dans la formation de liaisons C-C lors de la biosynthèse de la pyrroloquinoline quinone et dans les réactions de méthylation de l'ARN ribosomique 23S, respectivement. Ces enzymes utilisent le radical 5'-dA• pour activer un atome de carbone aliphatique qui sera par la suite rajouté sur une double liaison du substrat.

#### III-4.2.1 Calculs théoriques

Afin de compléter nos données expérimentales la chimie théorique a été utilisée pour décortiquer les différentes étapes de la réaction observée dans HydE. Dans un premier temps des modèles du site actif simples ont été utilisés et des calculs de mécanique quantique (QM) ont été effectués. Ensuite, une approche hybride combinant la mécanique quantique et la mécanique moléculaire (QM/MM) a été utilisée pour étudier des modèles plus complexes, qui inclurait l'effet de la protéine – ceci dans le but de comprendre l'impact de la matrice protéique sur la réaction catalysée par l'enzyme [voir article Rohac *et al.* 2016 en Annexe]. Ces calculs ont été effectués par Patricia Amara de notre laboratoire, en collaboration avec Jean-Marie Mouesca de l'Institut Nanosciences et Cryogénie au CEA à Grenoble.

Les deux ligands (2*R*,4*R*)-MeSeTDA et (2*R*,4*R*)-MeTDA ont été étudiés. Dans les paragraphes suivants nous allons parler d'avantage du ligand (2*R*,4*R*)-MeSeTDA, parce que les analyses avec ce ligand séléné sont plus claires. Cependant, comme nous le verrons plus loin dans le texte, les calculs avec le ligand soufré aboutissent à des résultats similaires [Figure-III-87-b&c et Figure-III-94-a&b].

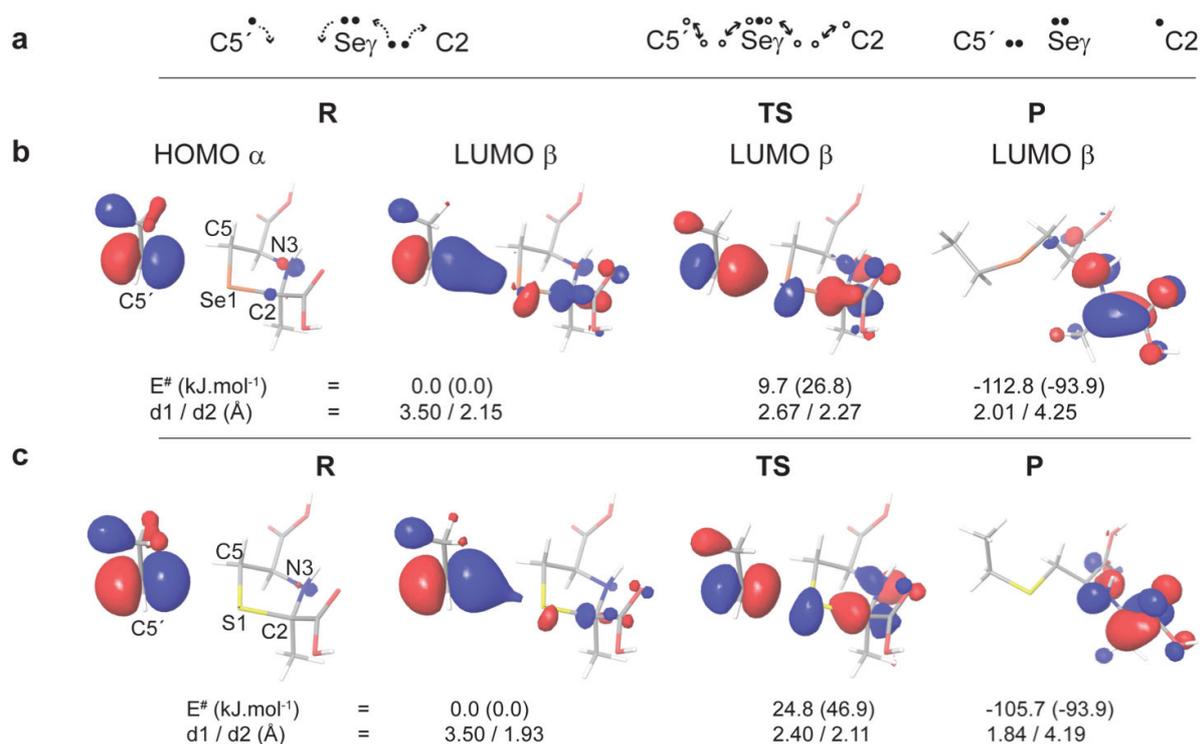
#### III-4.2.2 Analyse orbitaire

Dans l'analyse orbitaire le caractère radicalaire est représenté par l'orbitale HOMO  $\alpha$  (acronyme anglais de *Highest Occupied Molecular Orbital*). Cependant, dans certains cas, cette orbitale peut être altérée par d'autres orbitales qui ont un niveau d'énergie proche<sup>25,26</sup>.

Dans ce cas il est préférable de visualiser le caractère radicalaire à l'aide de l'orbitale vide LUMO  $\beta$  (acronyme anglais de *Lowest Unoccupied Molecular Orbital*). Cette dernière ne contient pas d'électron, elle est donc moins affectée par d'autres orbitales et reflète plus clairement l'évolution du caractère radicalaire au cours de la réaction [pour plus d'explications voir article Rohac et al. 2016 Schème S1 en Annexe]. Dans ce manuscrit, dans certains cas nous présenterons l'orbitale HOMO  $\alpha$ , pour montrer l'endroit où l'électron non apparié se trouve, dans d'autres cas nous présenterons l'orbitale vide LUMO  $\beta$  pour illustrer où le caractère radicalaire va apparaître.

### **III-4.2.3 Analyse orbitale de la réaction de HydE avec le ligand (2R,4R)-MeSeTDA**

L'étude de l'évolution de la structure électronique du petit modèle représentant l'attaque du ligand dérivé de thiazolidine et de la formation de la Se-SAC par le radical C5' montre que : (i) la formation de la liaison C5'-Se1 et le clivage de la liaison Se1-C2 sont concomitants et conduisent à la formation d'un radical centré sur le carbone C2. Ce radical est délocalisé sur les atomes d'azote et de carbone adjacents au C2. En effet, comme on peut le voir sur la [Figure-III-87-b](#), les atomes C5', Se1, C2, N3 et les atomes du groupement carboxylate attaché sur le C2 contribuent à l'orbitale LUMO  $\beta$  de l'état de transition TS, suggérant que la formation de la liaison C5'-Se1 et le clivage de la liaison Se1-C2 se font en une seule étape ; (ii) la barrière énergétique de cette étape est faible - de l'ordre de 9,7 kJ.mol<sup>-1</sup> pour le ligand sélénié et 24,8 kJ.mol<sup>-1</sup> dans le cas du ligand soufré [[Figure-III-87-b&c](#) et [Tableau-III-02](#)]. La réaction est plus favorable avec le ligand sélénié, parce que le sélénium, ayant des orbitales plus diffuses, permet un meilleur recouvrement orbitalaire entre le radical sur le C5' et le ligand. Ce phénomène est bien visible sur la comparaison de la LUMO  $\beta$  du réactif R sélénié [[Figure-III-87-b](#)] et la même orbitale du réactif R soufré [[Figure-III-87-c](#)]. D'autres études ont observé la même différence de réactivité entre les thioéthers et sélénoéthers vis-à-vis des radicaux alkyls<sup>27</sup> ; (iii) la formation de la liaison C5'-Se1 et le clivage concomitant de la liaison Se1-C2 est en fait une substitution homolytique (S<sub>H</sub>2) qui, comme le suggère l'orbitale LUMO  $\beta$  du réactif R, fait intervenir un des doublets non liants du Se1 (ou S1) et l'électron radicalaire issu du clivage de la SAM [[Figure-III-87-a&b](#)] (voir plus loin dans le texte). L'autre doublet non liant du Se1 (ou S1) est passif et adopte une conformation perpendiculaire au plan défini par les atomes C5', Se1/S1 et C2.

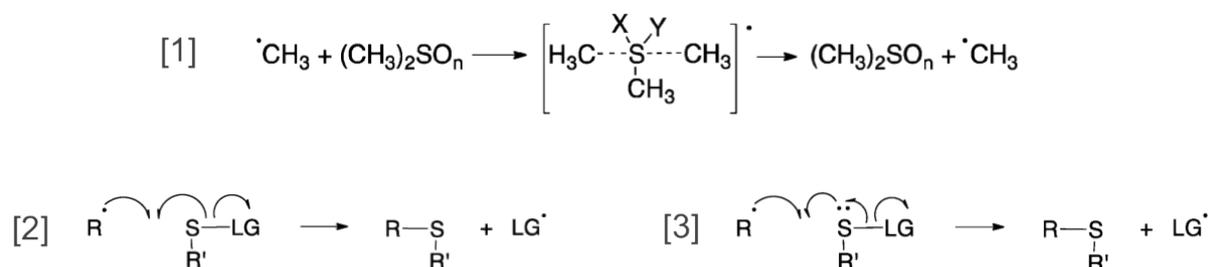


**Figure-III-87** : Petit modèle de l'attaque du radical C5'• sur le ligand (2*R*,4*R*)-Me(Se)TDA

**a** - Schéma de la substitution homolytique ( $S_H2$ ) sur un chalcogène (ici un atome de sélénium) faisant intervenir des électrons d'un de ses doublets non liants. Les électrons 'entiers' sont représentés par des cercles pleins, tandis que le réarrangement électronique correspondant à la délocalisation des électrons dans l'état de transition TS est représenté par des cercles vides interconnectés. **b** - Modèle QM minimal de l'attaque initiale du radical C5' (ici modélisé comme un  $CH_3CH_2\bullet$ ) sur le sélénoéther du ligand (2*R*,4*R*)-MeSeTDA. Les énergies de l'état de transition TS et du produit P, relatives à celle du réactif R, sont indiquées en kJ.mol<sup>-1</sup> pour les deux méthodes de calculs B3LYP et (B)HandHLYP). Les distances entre les atomes C5'-Se1 (d1) et entre les atomes Se1-C2 (d2) sont indiquées en Å sous format d1/d2. **c** - Modèle QM minimal similaire à celui dans (**b**) excepté la nature du chalcogène, cible de l'attaque radicalaire (ici un thioéther).

Le dernier point est en accord avec l'étude d'Aitken et de ses collaborateurs<sup>28</sup>, qui ont montré, à l'aide de calculs théoriques basés sur un modèle d'un radical méthyl attaquant l'atome de soufre du sulfure de diméthyle [Figure-III-88-[1]], qu'une attaque d'un radical alkyl sur un atome de soufre d'un thioéther ou d'un sulfoxyde est dominée par une interaction entre l'orbitale SOMO du radical et le doublet non liant du soufre et non par l'interaction  $SOMO \rightarrow \sigma^*$ , comme on le pensait jusqu'à récemment [Figure-III-88-[2]]. Ainsi, contrairement aux additions radicalaires sur une double liaison (dominées par l'interaction  $SOMO \rightarrow \pi^*$ ) où le radical alkyl est considéré comme *nucléophile*, les radicaux alkyls impliqués dans des réactions  $S_H2$  sur des chalcogènes (dominées par l'interaction  $SOMO \rightarrow$  doublet non liant) pourraient être considérés comme de nature *électrophile*. Nos résultats sont globalement en accord avec les conclusions de l'étude en question. La barrière énergétique de l'état de transition du modèle d'Aitken est plus élevée que celle de notre modèle. Cette différence d'énergie peut être expliquée par la différence de stabilité des espèces radicalaires formées à la fin de la réaction. Tandis que dans le modèle d'Aitken on forme un radical méthyl de haute énergie - un radical méthyl, l'espèce formée dans HydE présente une délocalisation stabilisante du caractère radicalaire. D'autres études ont montré que la barrière

énergétique dans les réactions de substitution homolytique est fortement dépendante de la nature du radical partant, tandis que la nature du radical alkyl qui porte l'attaque initiale semble avoir relativement peu d'impact sur cette barrière<sup>27,29</sup>.



**Figure-III-88** : Attaque d'un radical alkyl sur un atome de soufre (Aitken et al. 2012<sup>28</sup>)

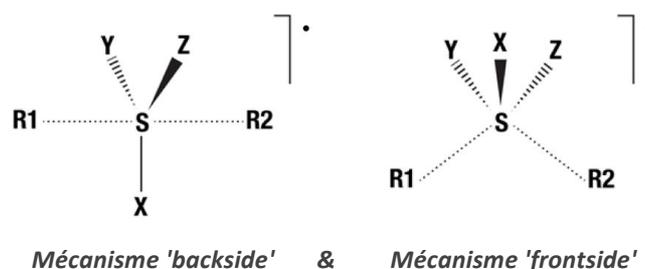
[1] - Modèle utilisé par Aitken et ses collaborateurs pour étudier une attaque d'un radical alkyl sur un atome de soufre d'un thioéther (n=0), d'un sulfoxyde (n=1) ou d'un sulfone (n=2). X et Y représentent soit des atomes d'oxygènes ou des doublets non liants. Dans le cas d'un sulfure de diméthyle (n=0) la barrière de l'état de transition est de 54,9 (82,4) kJ.mol<sup>-1</sup>, tandis que celle de notre modèle avec le ligand soufré est de 24,8 (46,9) kJ.mol<sup>-1</sup> pour les calculs B3LYP (BHandHLYP). [2] - L'ancien modèle d'une attaque du radical alkyl sur le soufre du thioéther - réaction (S<sub>H2</sub>) dominée par l'interaction (SOMO → σ\*). 'LG' désigne le groupe partant qui portera le radical final. [3] - Nouveau modèle d'une attaque du radical alkyl sur le soufre du thioéther, découlant des travaux d'Aitken et al. - réaction (S<sub>H2</sub>) dominée par l'interaction (SOMO → doublet non liant).

En résumé, nos calculs avec le petit modèle montrent que la formation de la liaison C5'-Se1 et le clivage de la liaison Se1-C2 se font de façon concomitante et que cette réaction de substitution homolytique (S<sub>H2</sub>) fait intervenir l'interaction entre l'orbitale occupée par l'électron non apparié appartenant au radical 5'dA• et le doublet non liant du chalcogène, ce dernier étant la cible de l'attaque radicalaire.

#### III-4.2.4 Etude de la directionnalité de la substitution homolytique (S<sub>H2</sub>) dans HydE

La réaction observée dans HydE est bien connue de la chimie organique; elle ressemble à la désulfuration réductive des thioéthers ou des N,S-acétals catalysée par des composés organostanniques<sup>30,31</sup>. En solution, les réactions S<sub>H2</sub> favorisent le plus souvent le clivage d'une liaison par mécanisme 'backside' et non mécanisme 'frontside'<sup>30,32</sup> [Figure-III-89]. En effet, Il a été prouvé que les réactions S<sub>H2</sub> sur des thioéthers ou sulfoxydes nécessitent un arrangement colinéaire entre l'atome de soufre, l'atome qui va former la nouvelle liaison et l'atome du groupe partant et que la réaction passe par une inversion de la configuration sur le soufre<sup>33,34</sup>. Le même phénomène a été observé pour la synthèse enzymatique de la SAC par HydE. Avant la réaction l'angle entre les atomes C5'-Se1-C2 est de 165,4°, tandis que celui entre les atomes C5'-Se1-C5 est de 92,4°. Donc les atomes C5', Se1 et C2, qui sont impliqués dans la formation et le clivage des liaisons, sont presque parfaitement alignés [Figure-III-82-B].

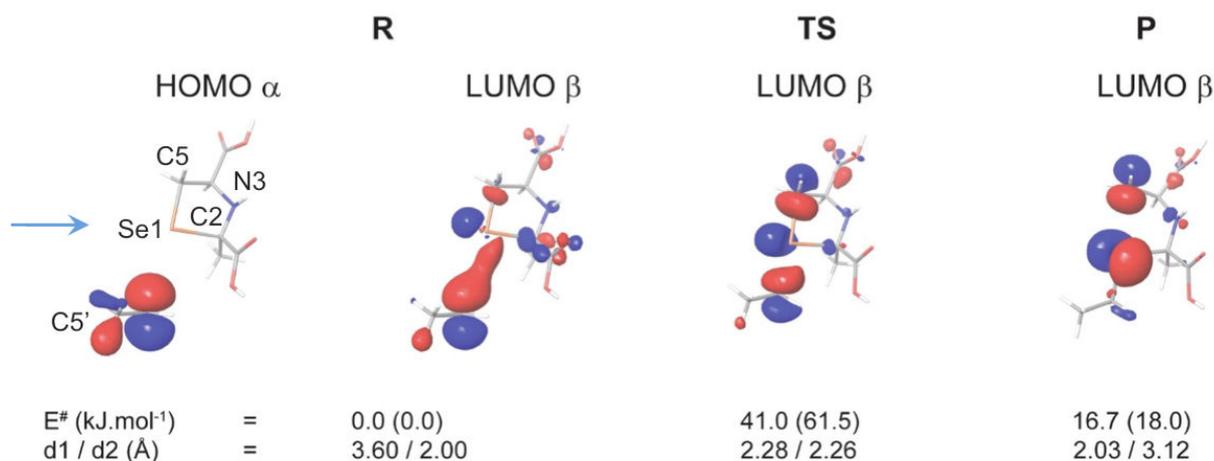
A l'aide d'un petit modèle QM, nous avons simulé une attaque du radical C5'• sur le Se1, cette fois-ci dans l'axe de la liaison Se1-C5 pour déterminer si le clivage de la liaison Se1-C2 dépend de la configuration initiale des réactants ou bien si le clivage de cette liaison est inhérent à la formation de la liaison C5'-Se1 [pour plus de détails sur les calculs quantiques voir article Rohac *et al.* 2016 en Annexe].



**Figure-III-89** : Réactions de substitution homolytique (S<sub>H2</sub>)

Principe de la formation et du clivage homolytique d'une liaison par mécanisme 'backside' et 'frontside' lors d'une réaction de substitution homolytique (S<sub>H2</sub>). R1 représente le groupe qui effectue l'attaque radicalaire sur l'atome S, tandis que R2 représente le groupe partant. A gauche, la nouvelle liaison formée R1-S est dans l'axe avec la liaison S-R2, qui subira un clivage homolytique. Ceci est le mécanisme 'backside', qui ressemble à une substitution nucléophile S<sub>N2</sub> et conduit, lui aussi, à une inversion de stéréochimie sur l'atome S (inversion de Walden). Ce mécanisme est parfois appelé 'T-shaped transition', parce que le substitut X du thioéther est perpendiculaire à l'axe R1-S-R2. A droite, la nouvelle liaison formée R1-S et la liaison clivée S-R2 se trouvent du même côté (mécanisme 'frontside' dans la littérature) - la stéréochimie reste donc inchangée<sup>32</sup>.

L'orbitale LUMO β du réactant R, c'est-à-dire du radical C5'• en approche du ligand (2*R*,4*R*)-MeSeTDA, montre qu'il y a une dégénérescence énergétique entre l'orbitale du radical C5'•, une orbitale du doublet non liant du Se1, les deux orbitales du C5 (clivage 'backside') et du C2 (clivage 'frontside') [Figure-III-90]. La présence d'un caractère radicalaire aussi sur le C2 ('frontside') s'explique par le fait, que le groupe partant résultant (radical sur l'atome de carbone C2) est plus stable que le radical sur le C5, à cause de sa délocalisation sur l'atome d'azote et le groupement carboxylate adjacents au C2 [voir paragraphe III-4.2.3]. Lorsqu'on approche le radical C5' de l'atome Se1 en maintenant cette direction, l'alignement entre les atomes C5', Se1 et C5 fait augmenter le recouvrement orbitalaire entre ces atomes et induit le clivage de la liaison Se1-C5. Ces résultats sont en accord avec l'étude théorique de réactions de substitutions homolytiques déclenchées par une attaque d'un radical alkyl sur un atome de sélénium d'un sélénioéther, réalisée par Lyons et Schiesser. Ces auteurs ont démontré qu'une telle attaque procède de façon concomitante (pas d'intermédiaire radicalaire sur le chalcogène) et que l'état de transition présente une quasi-colinéarité entre le radical attaquant et le radical partant<sup>27</sup>.



**Figure-III-90** : Impact de la directionnalité pour l'attaque radicalaire du C5'• sur le Se1

Modèle QM minimal de l'attaque du radical C5' sur l'atome Se1. Ce modèle particulier a été construit de façon à simuler l'attaque du C5'• dans l'axe de la liaison Se1-C5. Le mode de fixation du ligand (2*R*,4*R*)-MeSeTDA dans le site actif de HydE empêche un tel alignement. L'attaque dans la protéine se fait dans l'axe de la liaison Se1-C2. Cette direction est représentée par une flèche bleu clair à gauche de l'image. Dans la protéine et dans le modèle présenté ici, le recouvrement orbitalaire causé par un alignement d'atomes quasi-parfait induit le clivage de la liaison qui est dans l'axe de la liaison C5'-Se1 nouvellement formée (mécanisme *'backside'*). Puisque cette fois-ci ce sont les atomes C5', Se1 et C5 qui sont alignés, la liaison clivée de manière homolytique est celle entre Se1 et C5. Ceci conduit à la formation d'une espèce hypothétique avec le caractère radicalaire, qui se trouve en partie sur l'atome de carbone C5.

En conclusion, nos calculs montrent que l'attaque du radical 5'dA• sur le chalcogène du ligand (2*R*,4*R*)-Me(Se)TDA procèdent par le clivage de la liaison Se-C qui est dans l'axe de la nouvelle liaison C5'-Se1 formée (mécanisme *'backside'*). La matrice protéique impose une directionnalité de l'approche du radical C5'• dans l'axe de la liaison Se1-C2. Cette sélection d'orientation est, *in fine*, le paramètre qui impose la régiosélectivité. La barrière énergétique du second modèle est plus élevée que celle du premier (41,0 vs. 9,7 kJ.mol<sup>-1</sup>), sans doute à cause de la différence de stabilité du radical partant, résultant du clivage de la liaison Se1-C5 vs. Se1-C2 [voir [Tableau-III-02](#)]. Ce paramètre sera donc prépondérant en absence de contrôle par la matrice protéique lors d'une réaction en solution.

### III-5 Nature du radical 5'-dA•

#### III-5.1 Interactions entre les atomes C5' et Se1 avant la réaction

Nos données cristallographiques montrent très clairement qu'au cours de la conversion des réactifs en produits, les atomes qui effectuent les déplacements les plus importants au sein du site actif sont ceux qui appartiennent au cosubstrat SAM; plus précisément à sa partie ribose et L-méthionine [Figure-III-84-a & Figure-III-93]. En comparant la conformation que le 5'-dA adopte à la fin de la réaction dans HydE (cristal: X5) avec la conformation obtenue par la co-cristallisation de l'enzyme avec le 5'-dAH + L-méthionine en absence du substrat (code PDB: 3IIX) [Figure-III-91] on voit que les deux conformations sont quasi-superposables. L'atome C5', qui portera le radical, occupe la même position dans les deux structures ( $\pm 0,3$  Å), ce qui suggère que le radical 5'-dA•, généré par le clivage de la SAM, n'est pas capable de couvrir une distance plus longue dans la cavité de HydE. Lorsque la protéine co-cristallisée avec les produits du clivage de la SAM est mise en présence du substrat (2R,4R)-MeSeTDA (cristal: X3), le 5'-dA s'accommode dans une position intermédiaire entre celle dans la SAM intacte (cristal: X2) et celle dans la Se-SAC (cristal: X5). Ceci est dû à un clash stérique entre le méthyl du 5'-dAH et l'atome Se1 du cycle thiazolidine. D'ailleurs le substrat lui-même se retrouve légèrement repoussé par le 5'-dA plus loin de centre [Fe<sub>4</sub>S<sub>4</sub>] [Figure-III-91 ; X2 vs. X3]. Cette observation implique qu'avant la réaction, les atomes C5' et Se1, qui formeront la nouvelle liaison C-S(e), sont quasiment en contact van der Waals [Tableau-III-01 & Figure-III-92-a].

En analysant les distances interatomiques au sein du site actif de HydE on constate qu'au cours de la réaction, tous les mouvements préservent les contacts de van der Waals [Figure-III-92-a,b&c]. Les atomes qui interviennent dans la formation ou dans le clivage des liaisons atomiques se déplacent le long d'un axe entre le Fe unique et l'atome de carbone C2 du substrat [Figure-III-84-a]. Tous ces atomes se trouvent constamment en contact van der Waals et donc la position de chaque atome est influencée par la position des atomes voisins.

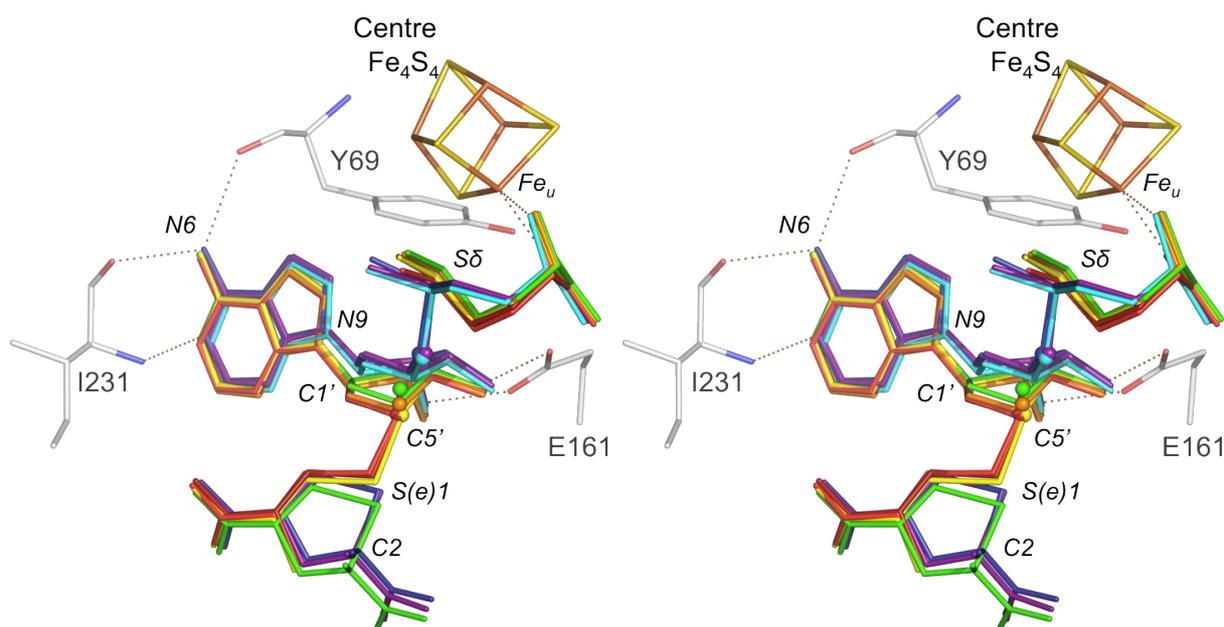
Nos essais de réactivité de HydE envers le ligand (2R,4R)-TMeTDA n'ont conduit ni à la formation en solution d'un nouveau dérivé adénosyl ni au clivage de la SAM dans les cristaux. Comme discuté dans le paragraphe I-11.4.1, ce ligand présente un clash stérique important entre son méthyl pro-S sur l'atome C5 du cycle thiazolidine et les atomes O4' et C1' de la SAM [Figure-III-92-c & Figure-I-63]. Ce contact empêche le rapprochement du 5'-dA vers l'atome S1, ce qui rend la formation de la liaison C5'-S1 impossible.

L'atome du ligand (2R,4R)-TMeTDA le plus proche de l'atome C5' de la SAM est un des hydrogènes du méthyl pro-R sur l'atome C5 du cycle thiazolidine. La distance entre ces atomes est aux alentours de 3 Å [Figure-III-92-c]. Cet atome d'hydrogène est donc suffisamment proche du C5' pour être arraché par le radical 5'-dA•. Cependant on remarque que l'angle H-C5'-S<sub>8</sub> est d'environ 115°, contrairement à l'alignement quasi-parfait des atomes impliqués dans la formation de la SAC (circa 165°). L'absence de réactivité de HydE envers le ligand (2R,4R)-TMeTDA illustre à quel point le mouvement effectué par le 5'-dA au cours de la réaction et le positionnement correct de l'atome cible d'attaque radicalaire sont déterminants pour que la réaction puisse avoir lieu.

**Tableau-III-01** : Tableau récapitulatif contenant les distances entre les atomes impliqués dans les principaux réarrangements lors de la conversion des ligands (2R,4R)-Me(Se)TDA en produits (Se-)SAC

En gras et italique, les codes des cristaux ou les codes PDB des structures utilisées dans ce chapitre. Le nom de la couleur entre parenthèse permet de relier les distances à leur structure correspondante dans la Figure-III-14 et dans le Tableau-III-2. (\*) désigne les distances mesurées entre le Se1 et un atome de carbone. Les distances interatomiques peuvent être directement comparées à des sommes de rayons van der Waals des atomes correspondants; C = 1.70 Å; S = 1.80 Å; Se = 1.90 Å; H = 1.20 Å [référence <sup>35</sup>]. Cette comparaison permet d'identifier les atomes qui sont en interaction de type van der Waals même avant l'initiation de la réaction. Les distances ont été mesurées à partir de nos structures de haute résolution [Tableu-III-2] dont la superposition est représentée sur la Figure-III-14. Pour les détails concernant la structure X9 voir Figure-III-15-d.

Bond	<b>X1</b> <i>Bleu</i>	<b>X2</b> <i>Pourpre</i>	<b>3IIZ</b> <i>Cyan</i>	<b>X3</b> <i>Vert</i>	<b>3IIX</b> <i>Orange</i>	<b>X4</b> <i>Jaune</i>	<b>X5</b> <i>Crimson</i>	<b>X6</b> <i>Rouge</i>	<b>X9</b> -
Fe-S $\delta$	3.25 Å	3.23 Å	3.25 Å	2.63 Å	2.67 Å	2.88 Å	2.82 Å	2.78 Å	3.27 Å
S $\delta$ -C5'	1.86 Å	1.81 Å	1.83 Å	3.50 Å	3.80 Å	3.91 Å	3.91 Å	3.81 Å	1.80 Å
C5'-S1	3.85 Å	3.91 Å*	-	3.31 Å*	-	2.08 Å*	1.83 Å	1.94 Å	4.24 Å
S1-C2	1.80 Å	1.99 Å*	-	1.93 Å*	-	-	-	-	1.85 Å



**Figure-III-91** : Vue stéréoscopique des conformations possibles du 5'-dA pendant la réaction dans HydE

Détails du site actif avec en blanc les résidus conservés en interaction avec la SAM ou avec le 5'-dAH. Les sphères colorées représentent l'atome C5' pour chaque structure superposée. Le code couleur est le suivant: **SAM+(2R,4R)-MeTDA to** (cristal X1; bleu), **SAM+(2R,4R)-MeSeTDA to** (cristal X2; pourpre), **la SAM seule** (code PDB: 3IIZ; cyan), **5'-dAH+L-Méthionine+(2R,4R)-MeSeTDA** (cristal X3; vert), **5'-dAH+L-Méthionine** (code PDB: 3IIX; orange), **Se-SAC+L-Met ; t<sub>final</sub> d'une cinétique dans les cristaux protéiques** (cristal X4; jaune), **SAC+L-Met ; t<sub>final</sub> d'une cinétique dans les cristaux protéiques** (cristal X5; rouge carmin) and **SAC+L-Méthionine co-cristallisation** (cristal X6; rouge). Les statistiques de collectes de données cristallographiques et d'affinement pour chaque jeu/cristal sont détaillées dans le Tableau-III-XXX. Le résumé de distances entre les atomes impliqués dans la réaction catalysée par HydE se trouve dans le Tableau-III-01.

**Tableau-III-2 (partie 1/2) | Statistiques de traitement de données cristallographiques et d'affinement de modèles**

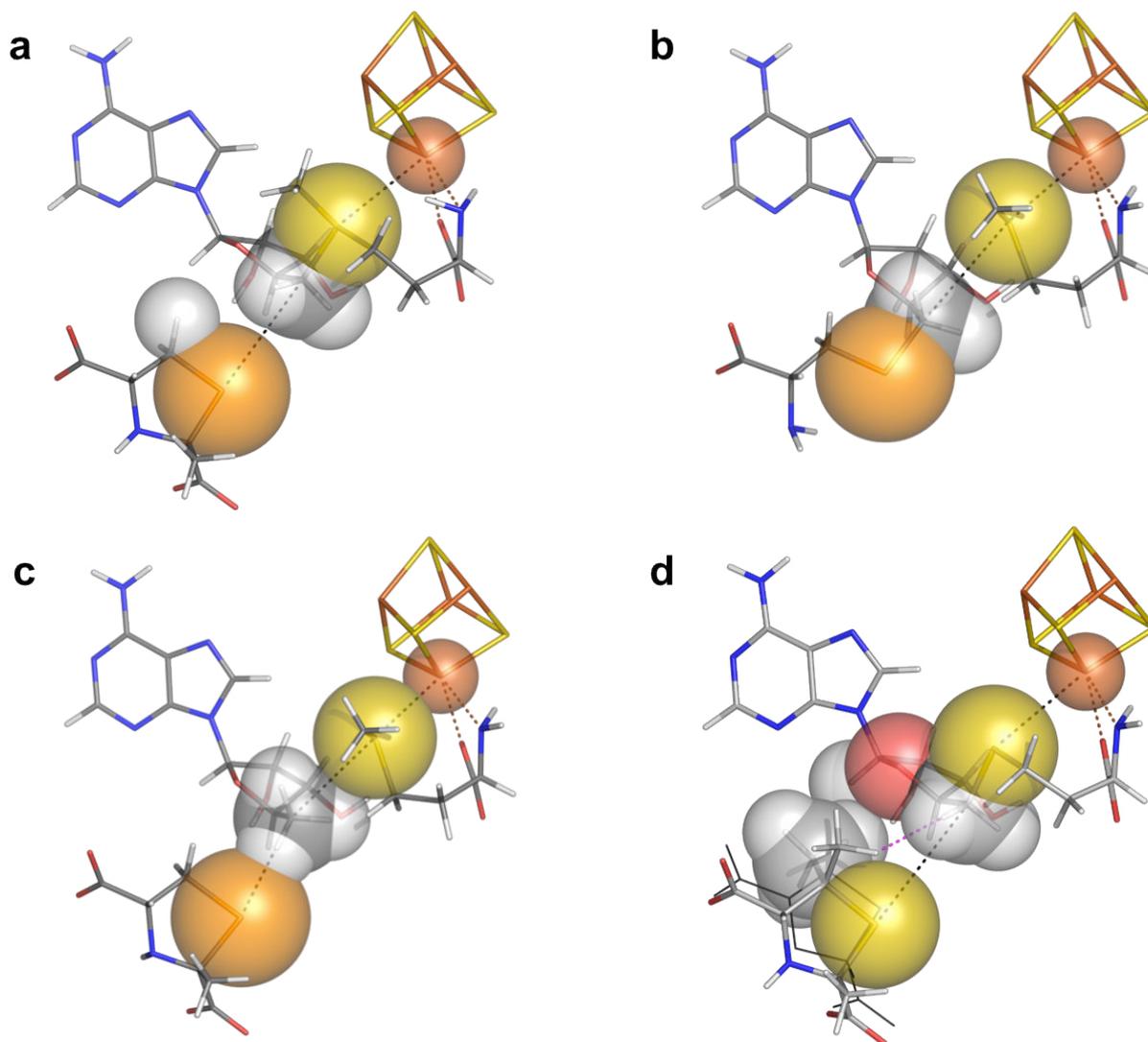
Nom du cristal	X1	X2	X3	X4	X5
Condition	t <sub>0</sub> (cinétique); SAM & (2 <i>R</i> ,4 <i>R</i> )-MeTDA	t <sub>0</sub> (cinétique); SAM & (2 <i>R</i> ,4 <i>R</i> )-MeSeTDA	5'dA + L-Met & (2 <i>R</i> ,4 <i>R</i> )-MeSeTDA	t <sub>final</sub> (cinétique); L-Met & Se-SAC	t <sub>final</sub> (cinétique); L-Met & SAC
<b>Enregistrement de données</b>					
Maille (Å)	51.17 80.02 86.36	51.07 79.10 86.26	50.80 78.87 86.25	51.12 79.46 86.57	50.99 78.97 86.27
Groupe d'espace	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Longueur d'onde (Å)	0.979844	0.980250	0.979129	0.979677	0.979247
Résolution (Å)	1.45 (1.54-1.45)	1.27 (1.35-1.27)	1.35 (1.43-1.35)	1.32 (1.40-1.32)	1.17 (1.24-1.17)
Rsym, %	4.7 (81.0)	4.9 (62.7)	7.2 (93.9)	5.5 (66.3)	9.7 (174.4)
I/σ	14.86 (1.69)	13.97 (2.01)	15.27 (1.98)	12.61 (1.80)	19.08 (2.21)
CC(1/2)	99.9 (67.0)	99.9 (68.0)	99.9 (74.5)	99.9 (70.1)	99.9 (80.9)
Complétude (%)	99.5 (98.6)	99.1 (98.8)	99.8 (99.4)	98.8 (97.6)	99.6 (98.4)
Multiplicité	3.4 (3.3)	2.9 (2.8)	5.9 (5.8)	3.2 (3.1)	20.0 (18.9)
<b>Affinement de modèle</b>					
R <sub>cris</sub> /R <sub>free</sub> <sup>21</sup>	16.16/18.80	10.77/13.53	11.90/13.52	11.29/14.67	10.97/13.23
No. de réflexions utilisées/Test	60277/3183	87752/4519	72790/3833	78733/4132	112915/5529
<b>Déviations de la géométrie idéale</b>					
Liaisons atomiques	0.0231	0.0264	0.0243	0.0252	0.0269
Angles	2.3962	2.5138	2.2769	2.4474	2.4886
No. d'atomes affinés	3729	3577	3638	3593	3631
Protéine	3003	2868	2892	2887	2930
Molécules d'eau	486	469	506	469	427
Atomes de Fe	4	4	4	4	4
Autres	236	236	236	233	270
<b>Diagramme de Ramachandran</b>					
Préférentiel/acceptable/défavorisé	95.2% 4.8% 0%	96.1% 3.6% 0.3%	95.6% 4.4% 0%	95.7% 4.3% 0%	95.8% 4.2% 0%
<b>Valeur du facteur B (Å<sup>2</sup>)</b>					
Protéine	20.57	16.36	15.58	16.68	16.20
Ligand	23.58	23.75	20.31	-	-
SAM / 5'-dA / (Se)-SAC	20.64 (SAM)	15.46 (SAM)	11.60 (5'-dA)	15.93 (Se-SAC)	16.71 (SAC)
Agrégat [Fe <sub>4</sub> S <sub>4</sub> ] / [Fe <sub>4</sub> Se <sub>4</sub> ] <sup>*</sup>	18.22	13.32	13.23 <sup>*</sup>	13.48	11.94

Seuls le résidu Val<sup>348</sup> C-terminal et les résidés appartenant à l'étiquette strep N-terminale (MWSHPQFEKN) ne sont pas modélisés, car ils sont systématiquement absents des cartes de densité électronique.

**Tableau-III-2 (partie 1/2) | Statistiques de traitement de données cristallographiques et d'affinement de modèles**

Nom du cristal	X1	X2	X3	X4	X5
Condition	$t_0$ (cinétique); SAM & (2R,4R)-MeTDA	$t_0$ (cinétique); SAM & (2R,4R)-MeSeTDA	5'dA + L-Met & (2R,4R)-MeSeTDA	$t_{\text{final}}$ (cinétique); L-Met & Se-SAC	$t_{\text{final}}$ (cinétique); L-Met & SAC
<b>Enregistrement de données</b>					
Maille (Å)	51.17 80.02 86.36	51.07 79.10 86.26	50.80 78.87 86.25	51.12 79.46 86.57	50.99 78.97 86.27
Groupe d'espace	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Longueur d'onde (Å)	0.979844	0.980250	0.979129	0.979677	0.979247
Résolution (Å)	1.45 (1.54-1.45)	1.27 (1.35-1.27)	1.35 (1.43-1.35)	1.32 (1.40-1.32)	1.17 (1.24-1.17)
Rsym, %	4.7 (81.0)	4.9 (62.7)	7.2 (93.9)	5.5 (66.3)	9.7 (174.4)
I/ $\sigma$	14.86 (1.69)	13.97 (2.01)	15.27 (1.98)	12.61 (1.80)	19.08 (2.21)
CC(1/2)	99.9 (67.0)	99.9 (68.0)	99.9 (74.5)	99.9 (70.1)	99.9 (80.9)
Complétude (%)	99.5 (98.6)	99.1 (98.8)	99.8 (99.4)	98.8 (97.6)	99.6 (98.4)
Multiplicité	3.4 (3.3)	2.9 (2.8)	5.9 (5.8)	3.2 (3.1)	20.0 (18.9)
<b>Affinement de modèle</b>					
$R_{\text{cris}}/R_{\text{free}}$	16.16/18.80	10.77/13.53	11.90/13.52	11.29/14.67	10.97/13.23
No. de réflexions utilisées/Test	60277/3183	87752/4519	72790/3633	78733/4132	112915/5529
<b>Déviations de la géométrie idéale</b>					
Liaisons atomiques	0.0231	0.0264	0.0243	0.0252	0.0269
Angles	2.3962	2.5138	2.2769	2.4474	2.4886
No. d'atomes affinés	3729	3577	3638	3593	3631
Protéine	3003	2868	2892	2887	2930
Molécules d'eau	486	469	506	469	427
Atomes de Fe	4	4	4	4	4
Autres	236	236	236	233	270
<b>Diagramme de Ramachandran</b>					
Préférentiel/acceptable/défavorisé	95.2% 4.8% 0%	96.1% 3.6% 0.3%	95.6% 4.4% 0%	95.7% 4.3% 0%	95.8% 4.2% 0%
<b>Valeur du facteur B (Å<sup>2</sup>)</b>					
Protéine	20.57	16.36	15.58	16.68	16.20
Ligand	23.58	23.75	20.31	-	-
SAM / 5'-dA / (Se)-SAC	20.64 (SAM)	15.46 (SAM)	11.60 (5'-dA)	15.93 (Se-SAC)	16.71 (SAC)
Agrégat [Fe <sub>4</sub> S <sub>4</sub> ] / [Fe <sub>4</sub> Se <sub>4</sub> ]*	18.22	13.32	13.23*	13.48	11.94

Seuls le résidu Val<sup>348</sup> C-terminal et les résidés appartenant à l'étiquette strep N-terminale (MWSHPQFEKN) ne sont pas modélisés, car ils sont systématiquement absents des cartes de densité électronique.

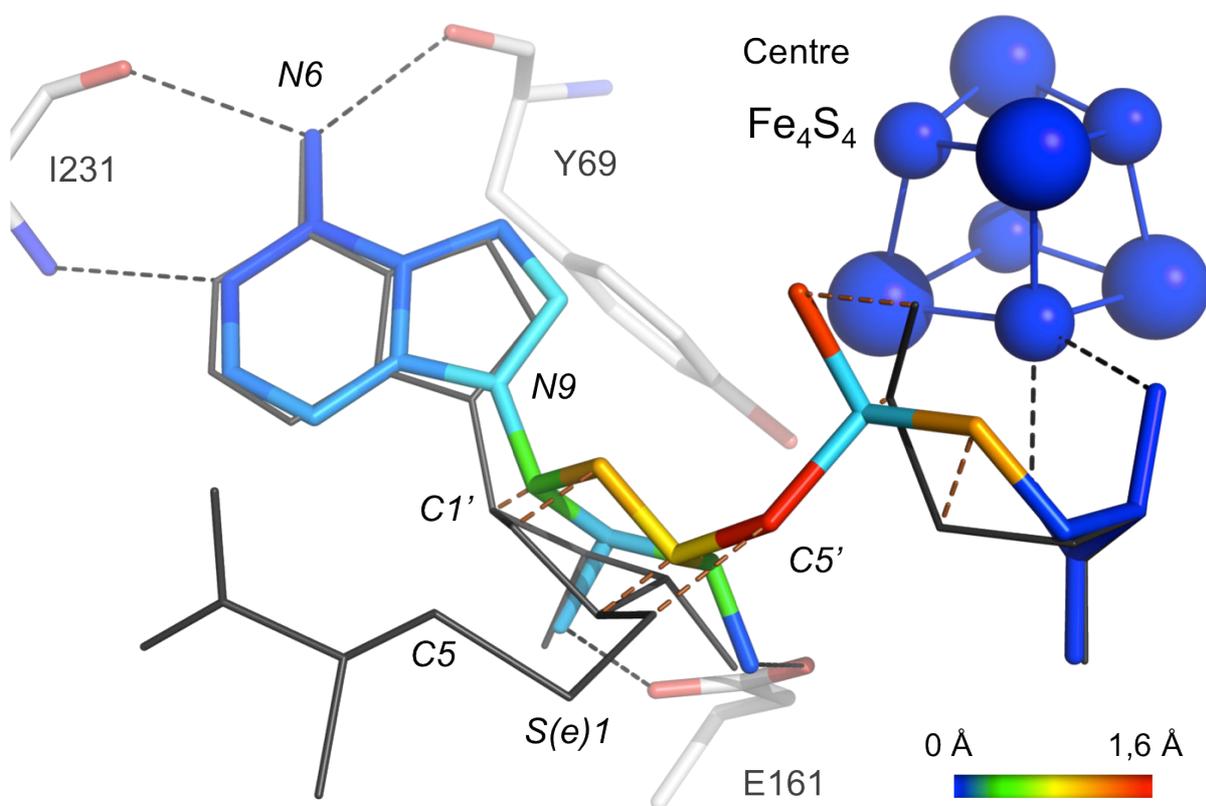


**Figure-III-92** : Analyse des contacts van der Waals au sein du site actif de *TmHydE*

Composition du site actif de *TmHydE*  $\Delta$ FeS telle qu'observée dans: **a** - la structure avec SAM + (2*R*,4*R*)-MeSeTDA à  $t_0$  (cristal X2), **b** - la structure avec Se-SAC+L-Met ;  $t_{\text{final}}$  d'une cinétique dans les cristaux protéiques (cristal X4), **c** - la co-cristallisation avec 5'-dAH + L-Méthionine + (2*R*,4*R*)-MeSeTDA (cristal X3), et **d** - la co-cristallisation avec SAM + (2*R*,4*R*)-TMeTDA (cristal X9). Code couleur utilisé: hydrogène en blanc, carbone en gris clair, nitrogène en bleu, oxygène en rouge, soufre en jaune, fer en ocre rouge et sélénium en orange. La comparaison des conformations des réactifs (**a**) avec celle des produits (**b**) montre très clairement que l'atome qui exerce le mouvement le plus important au cours de la formation de la SAC est l'atome C5' de la SAM. Ce mouvement s'effectue le long de l'axe qui relie le Fe unique à l'atome C2 du substrat. Après le clivage de la SAM, la rupture de la liaison S<sub>8</sub>-C5' augmente l'écartement entre ces deux atomes à cause de la création d'un nouveau contact van der Waals entre eux. Ceci est illustré par la structure (**c**) contenant les produits du clivage de la SAM (5'-dAH + L-Méthionine) et le ligand (2*R*,4*R*)-MeSeTDA intact. Le contact van der Waals entre les atomes S<sub>8</sub> et C5' augmente la compression des réactifs dans le site actif ce qui se traduit par un léger allongement de la distance entre le Fe unique et le Se1 du cycle thiazolidine. L'image (**d**) met en l'évidence la présence d'un clash stérique entre la SAM et le ligand (2*R*,4*R*)-TMeTDA. La distance entre l'atome O4' de la SAM et les méthyls pro-*R* et pro-*S* est de 3,79 et 3,57 Å, respectivement. De plus la distance entre le méthyl pro-*S* et l'atome C1' est de 3,70 Å. Cette interaction empêche l'attaque du radical 5'-dA• sur l'atome de soufre du ligand. La ligne pointillée de couleur magenta relie l'atome C5' avec l'atome d'hydrogène du ligand le plus proche. La distance entre ces atomes n'est que de 3,04 Å. [Pour une vue stéréoscopique sous un autre angle du site actif de la structure X9 voir la Figure-I-63.]

### III-5.2 Analyse du mouvement du 5'-dA au cours de la réaction

La plage conformationnelle de l'espèce 5'-dA dans HydE, visible sur la superposition de différentes structures [Figure-III-91], peut être expliquée par un simple mouvement de rotation d'environ  $6^\circ$  autour de l'atome N6 de l'adénine, accompagné d'une seconde rotation de  $26^\circ$  autour de la liaison N9-C1'. Puisque N6 est l'atome du 5'-dA le plus éloigné du C5', ces deux rotations légères s'amplifient progressivement plus on avance vers l'atome C5'. Ainsi un changement minimal de la conformation du 5'-dA maximise le déplacement de l'atome C5' vers le substrat [Figure-III-93]. Le radical alkyl sur le C5' est donc supposé de réagir avec sa cible au cours de ce déplacement de moins de 1,8 Å.



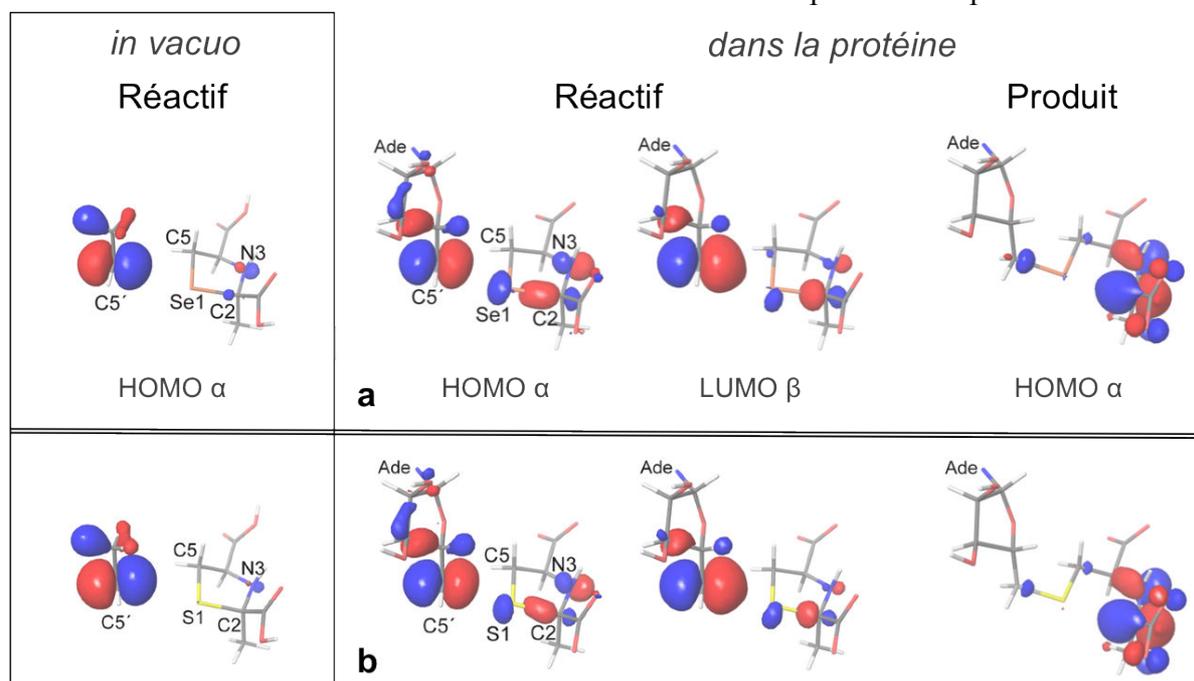
**Figure-III-93** : Visualisation des distances parcourues par chaque atome de la SAM au cours de la réaction

Les distances parcourues par chaque atome de la SAM après son clivage ont été mesurées à partir de la superposition de la structure SAM+(2*R*,4*R*)-MeTDA  $t_0$  (cristal X1) avec la structure SAC -  $t_{\text{final}}$  d'une cinétique dans les cristaux protéiques (cristal X5). Chaque distance a été ensuite reportée sur son atome de la SAM intacte correspondant en utilisant un code couleur (voir en bas à droite de l'image). Les conformations de molécules de la SAC et de la L-méthionine, telles qu'observées dans la structure X5 sont représentées par des bâtons fins de couleur gris foncé. Les lignes pointillées brunes connectent la position initiale et finale des atomes qui effectuent des mouvements les plus importants. Les résidus E161, I231, Y69 qui interagissent avec la partie 5'-dA de la SAM sont représentés par des bâtons épais de couleur blanche. Ces interactions sont mises en évidence par des lignes pointillées de couleur noire. Le centre [Fe<sub>4</sub>S<sub>4</sub>] 'SAM radical', représenté par des sphères, ne subit qu'un changement de conformation très subtil. Il se manifeste au niveau du Fe unique, qui se déplace d'environ 0,1 Å. Le déplacement du Fe unique est dû à la transition entre une coordination pseudo tétraédrique de ce Fe à une coordination quasi-octaédrique. Cette distorsion du centre [Fe<sub>4</sub>S<sub>4</sub>] 'SAM radical', qui résulte du clivage de la SAM, a été bien documentée dans la littérature<sup>13</sup>. La très faible amplitude de ce mouvement empêche sa visualisation sur le modèle ci-dessus.

### III-5.3 Etude du rôle de la protéine dans le mécanisme radicalaire de HydE

Après l'étude de l'évolution au cours de la réaction de la structure électronique en utilisant les petits modèles QM nous avons voulu déterminer de quelle manière la protéine influençait la réaction radicalaire dans HydE.

Le modèle QM/MM, qui prend compte de la contribution de la protéine, montre une délocalisation partielle du caractère radicalaire du 5'-dA• sur les atomes Se1 et C2 du substrat [Figure-III-94]. Ceci est en contraste avec le modèle QM minimal, où l'électron non apparié est presque entièrement localisé sur l'atome C5' [Figure-III-94 - encadré & Figure-III-87-b]. La délocalisation du caractère radicalaire dans le modèle QM/MM est due au fait que la matrice protéique restreint les mouvements des réactifs dans le site actif et maintient les atomes en contact van der Waals tout au long de la réaction. Ceci a pour effet d'augmenter le recouvrement orbitalaire entre les atomes C5' et Se1 durant la première étape de la réaction.



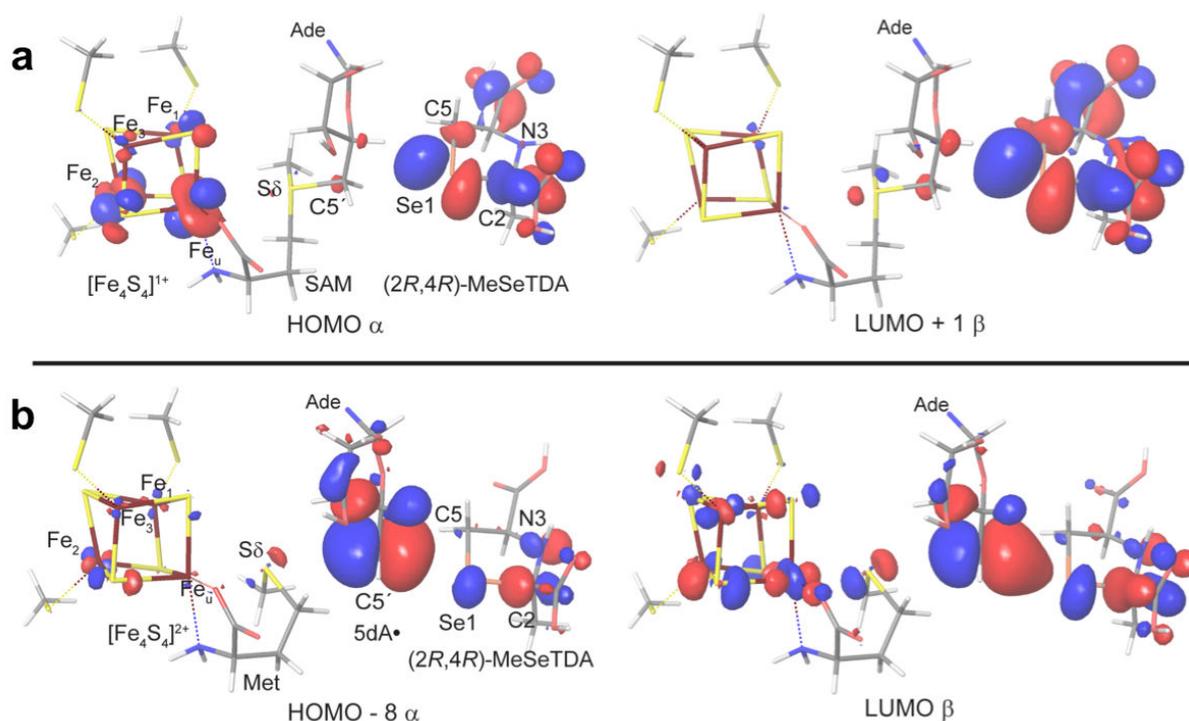
**Figure-III-94** : Impact de la matrice protéique sur l'attaque du radical C5' sur le substrat

**Encadré**: l'orbitale HOMO $\alpha$  du réactif calculé à partir du modèle QM minimal avec le substrat (2R,4R)-MeSeTDA en haut et le substrat (2R,4R)-MeTDA en bas [pour une description plus complète voir la Figure-III-87-b&c]. **A droite de l'encadré**: le modèle QM/MM, qui contient la protéine dans la partie MM et le radical 5'-dA• avec le substrat sélénié (en haut) et le substrat soufrée (en bas) dans la partie QM. Seule certaines orbitales de la partie QM sont représentées. La comparaison de l'orbitale HOMO  $\alpha$  du réactif du modèle QM minimal avec la même orbitale dans le modèle QM/MM montre que lorsque la partie protéique est incluse dans les calculs, le caractère radicalaire n'est plus localisé uniquement sur le carbone C5', mais il est aussi partiellement délocalisé sur les atomes Se1 et C2 du substrat.

L'inclusion de la protéine dans notre modèle théorique nous a confirmé que la barrière énergétique de la réaction concertée (formation de la liaison C5'-Se1 et clivage de la liaison Se1-C2) est basse et que cette réaction est favorable et facile (9,7 kJ.mol<sup>-1</sup> dans le vide et 8,2 kJ.mol<sup>-1</sup> dans le modèle avec la protéine) [Tableau-III-02]. Cette barrière est très faible par rapport à la barrière correspondante à l'étape de l'initiation de la réaction radicalaire dans les protéines à radical SAM, c'est-à-dire au clivage de la

SAM et à la génération du radical 5'-dA•. Dans HyDE<sup>13</sup> la barrière énergétique du clivage de la liaison C5'-S<sub>δ</sub> en absence du substrat est de 54,0 kJ.mol<sup>-1</sup>. Il est vrai que les recherches sur la LAM et la RNR III ont montré que, la barrière énergétique associée au clivage de la SAM a la tendance de s'abaisser une fois que le substrat se fixe dans le site actif<sup>36-40</sup>. Cependant, cette barrière reste sans doute plus élevée que celle associée à la formation de la SAC. Le clivage de la SAM est donc l'étape limitante de la réaction. Ceci explique le manque de synchronisation de la progression de la réaction observée dans les cristaux protéiques de HyDE.

Lorsque le centre [Fe<sub>4</sub>S<sub>4</sub>]<sup>1+</sup> 'SAM radical' est inclus dans la partie QM du modèle QM/MM, nous pouvons voir, que même avant le clivage de la SAM, il y a une dégénérescence orbitale entre le substrat et le centre [Fe-S] réduit [Figure-III-95-a]. De plus, la modélisation du radical 5'-dA• et du centre [Fe<sub>4</sub>S<sub>4</sub>]<sup>2+</sup> oxydé dans la partie QM montre que dans l'enzyme, la matrice protéique maintient en contact les trois espèces [Fe<sub>4</sub>S<sub>4</sub>]<sup>2+</sup>, 5'-dA• et (2R,4R)-MeSeTDA, ce qui maximise le recouvrement orbitalaire entre les réactifs [Figure-III-95-b]. Ces résultats suggèrent que le radical 5'-dA• serait en faite un état de transition et non un intermédiaire isolable.



**Figure-III-95 : Inclusion du centre [Fe<sub>4</sub>S<sub>4</sub>] dans la partie QM de la modélisation (clivage de la SAM)**

a - Etat initial du site actif de HyDE avant le clivage de la SAM. Le centre [Fe<sub>4</sub>S<sub>4</sub>] 'SAM radical' est dans son état réduit et la SAM est modélisée avec une charge positive sur l'atome S<sub>δ</sub>. L'orbitale HOMO α montre la présence d'une partie du caractère radicalaire sur le substrat même avant le clivage réductif de la SAM. Par exemple, la population de spin sur l'atome Se1 du substrat est de 0,19. Puisque avant le clivage du cosubstrat l'électron non apparié est supposé d'être entièrement localisé sur le centre [Fe<sub>4</sub>S<sub>4</sub>]<sup>1+</sup>, l'observation de la délocalisation du caractère radicalaire sur le substrat, confirmée par le LUMO + 1β, démontre que le substrat est enclin à une oxydation radicalaire à un stade très précoce de la réaction. b - Dans la protéine, après le clivage de la SAM, le caractère radicalaire est surtout localisé sur l'atome C5' ; voir HOMO - 8α. Le centre [Fe<sub>4</sub>S<sub>4</sub>] se trouve dans son état oxydé. On remarque que dans LUMO β apparaissent, en plus de l'atome C5', l'agrégat [Fe<sub>4</sub>S<sub>4</sub>]<sup>2+</sup>, l'atome S<sub>δ</sub> et le substrat (2R,4R)-MeSeTDA. Ceci suggère qu'il y a une dégénérescence orbitale entre ces quatre espèces et que le radical 5'-dA• serait un état de transition plutôt qu'une espèce intermédiaire isolable.

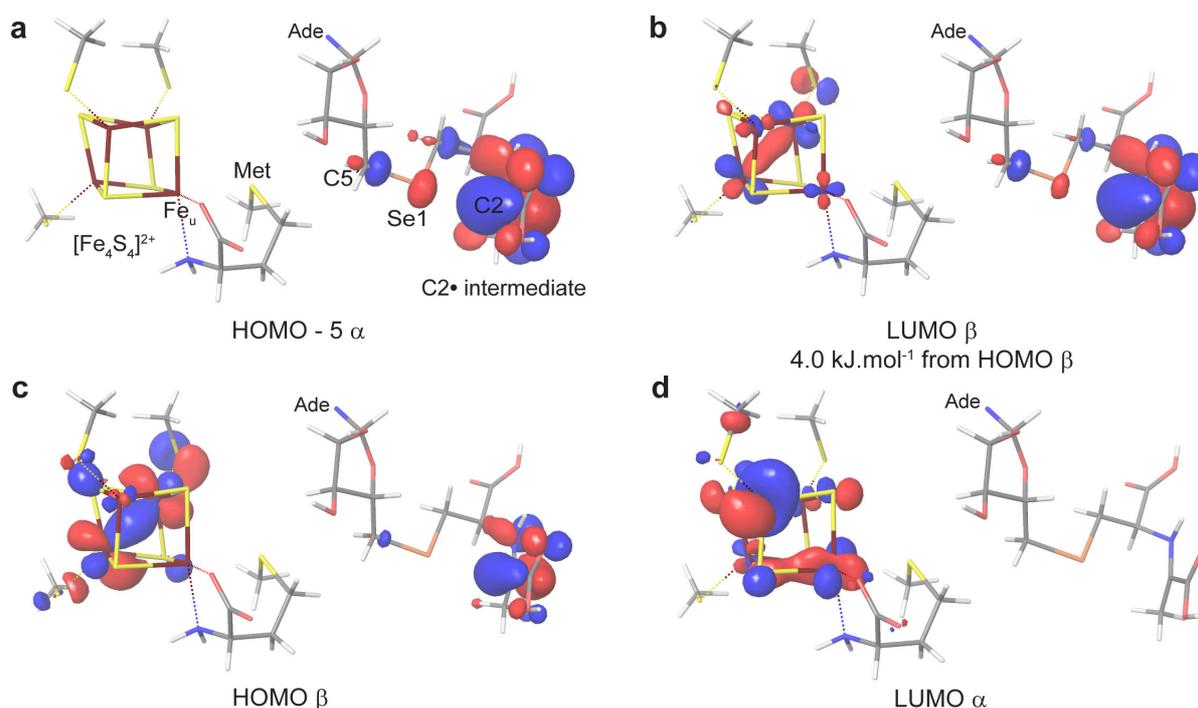
### III-6 Terminaison de la réaction radicalaire

Nos données expérimentales ont convergé vers un consensus que HydE catalyse la formation des produits (Se)-SAC et L-méthionine à partir des substrats SAM et (2*R*,4*R*)-Me(Se)TDA [III-2 & III-3.1.2]. Le troisième substrat, comme montré par les analyses cinétiques de RMN  $^1\text{H}$   $^{13}\text{C}$ , est le pyruvate [III-4.1]. Or les calculs théoriques indiquent qu'une espèce radicalaire sur le C2, plus ou moins stable, serait formée au cours de la réaction [III-4.2.3]. Cette réaction fait penser à une désulfuration radicalaire<sup>30</sup>. En synthèse organique, les radicaux générés par des de telles désulfurations radicalaires peuvent se propager en solution et régénérer l'espèce radicalaire initiale<sup>30,41</sup>. Par contre dans la protéine les espèces chimiques ne peuvent pas diffuser librement.

En se basant sur nos données structurales nous pouvons constater qu'à proximité du radical sur l'atome C2 il n'y a pas d'atome d'hydrogène, qui pourrait être transféré sur le C2 et propager ainsi la réaction radicalaire. Cependant il a été démontré que certaines protéines à radical SAM catalysent des réactions d'oxydation à deux électrons<sup>36,42-44</sup>. La protéine DesII, par exemple, termine l'oxydation du substrat en transférant l'électron libre sur le centre  $[\text{Fe}_4\text{S}_4]^{2+}$  'SAM radical'. De cette manière, cette enzyme arrive à catalyser plusieurs turnovers en recyclant à chaque fois la forme réduite du centre  $[\text{Fe}_4\text{S}_4]$  'SAM radical' avec un électron issu de l'intermédiaire radicalaire<sup>45</sup>.

Dans HydE, le radical centré sur l'atome C2 se situerait à un peu moins de 12 Å du centre  $[\text{Fe}_4\text{S}_4]^{2+}$  'SAM radical'. Cette distance est estimée à partir de la disposition des substrats avant la réaction et en prenant compte du déplacement de l'atome C2 suite au clivage de la liaison Se1-C2. (A  $t_0$ , l'atome C2 se situe à 10,4 Å du centre  $[\text{Fe-S}]$  'SAM radical' [Figure-III-82-B].) Un transfert direct de l'électron libre à partir du C2• vers le centre  $[\text{Fe}_4\text{S}_4]^{2+}$  'SAM radical' n'a pas été mesuré expérimentalement.

Cependant, un modèle QM/MM du site actif à la fin du clivage de la liaison Se1-C2, où le radical sur le C2 et l'agrégat  $[\text{Fe-S}]^{2+}$  oxydé ont été inclus dans la partie QM, montre que les niveaux d'énergie des orbitales de ces deux espèces sont très proches [Figure-III-96]. Ceci suggère qu'après un léger réarrangement électronique ou conformationnel des constituants du site actif l'électron pourrait être transféré sur le centre  $[\text{Fe-S}]$  et terminer ainsi la réaction radicalaire en régénérant l'agrégat  $[\text{Fe}_4\text{S}_4]^{1+}$  réduite. Le véritable produit de la réaction radicalaire serait une forme dérivé kétimine de la SAC, qui serait hydrolysée pour libérer le pyruvate, ainsi que cela est montré par nos expériences de RMN [voir le paragraphe suivant].



**Figure-III-96** : Inclusion du centre  $[\text{Fe}_4\text{S}_4]$  oxydé dans la partie QM de la modélisation (radical sur le C2)

La combinaison de la SOMO  $\alpha$  (a) - ici le HOMO-5  $\alpha$  - et la LUMO  $\beta$  (b) montre qu'à la fin de la formation de la liaison C5'-Se1 et du clivage de la liaison Se1-C2 le caractère radicalaire est localisé sur l'atome C2. La LUMO  $\beta$  (b) et la HOMO  $\beta$  (c) présentent une quasi-dégénérescence énergétique ( $4\text{kJ.mol}^{-1}$ ), ce qui a pour effet une délocalisation partielle de ces deux orbitales entre le centre  $[\text{Fe}_4\text{S}_4]$  'SAM radical' et l'atome C2 du produit portant l'électron non apparié. Cette situation est similaire à celle observée avant le clivage de la SAM [Figure-III-19-a]. Notons la présence du Fe unique dans la LUMO  $\beta$  (b), ce qui montre une communication orbitale entre ce Fe et l'atome C2. Cette communication suggère que l'électron non apparié pourrait être transféré du C2 au Fe et réduire le centre  $[\text{Fe}_4\text{S}_4]^{2+}$ . La localisation de la LUMO  $\alpha$  sur le centre [Fe-S] (d) favorise aussi ce transfert. Elle montre en effet que c'est l'agrégat  $[\text{Fe}_4\text{S}_4]$  qui sera l'accepteur électronique final.

Modèle	Figure		R	TS	P
(2 <i>R</i> ,4 <i>R</i> )-MeSeTDA	III-10-b	E d1 / d2	0.0 (0.0) 3.50 / 2.15	9.7 (26.8) 2.67 / 2.27	-112.8 (-93.9) 2.01 / 4.25
(2 <i>R</i> ,4 <i>R</i> )-MeSeTDA*	III-13	E d1 / d2'	0.0 (0.0) 3.60 / 2.00	41.0 (61.5) 2.28 / 2.26	16.7 (18.0) 2.03 / 3.12
(2 <i>R</i> ,4 <i>R</i> )-MeTDA	III-10-c	E d1 / d2	0.0 (0.0) 3.50 / 1.93	24.8 (46.9) 2.40 / 2.11	-105.7 (-93.9) 1.84 / 4.19
<b>(2<i>R</i>,4<i>R</i>)-MeSeTDA</b>	III-17-a	E d1 / d2	0.0 (0.0) 3.00 / 2.15	8.2 (31.0) 2.57 / 2.32	-41.2 (-43.5) 2.03 / 2.93
<b>(2<i>R</i>,4<i>R</i>)-MeTDA</b>	III-17-b	E d1 / d2	0.0 (0.0) 2.88 / 1.94	27.3 (54.6) 2.38 / 2.22	-50.8 / -48.7 1.85 / 2.96

**Tableau-III-02** : Tableau récapitulatif contenant les caractéristiques des modèles QM et QM/MM (en gras) en termes de barrières énergétiques et de distance interatomiques

Les énergies de l'état de transition TS et du produit P, relatives à celle du réactif R, sont indiquées en  $\text{kJ.mol}^{-1}$  pour les deux méthodes de calculs B3LYP et (B)HandHLYP). L'évolution des distances entre les atomes C5'-Se1 (d1) et entre les atomes Se1-C2 (d2) est indiquée en Å sous format d1/d2. La seule exception constitue la Figure-III-14 (\*), qui représente un modèle de l'attaque du radical C5' dans l'axe de la liaison Se1-C5. Dans ce modèle, d2' désigne la distance entre les atomes Se1 et C5.

### III-7 Mécanisme détaillé de la formation de la (Se)-SAC par HydE

La transformation enzymatique des dérivés thiazolidines par la protéine HydE, dont les différents aspects ont été étudiés dans les paragraphes précédents, peut être décrite comme une condensation radicalaire de la L-cystéine avec le 5'-dA. La molécule de pyruvate sert de cofacteur dans cette réaction et elle est régénérée à la fin du cycle [Figure-III-81-b3].

Les données expérimentales combinées aux calculs théoriques indiquent qu'au cours de la formation enzymatique de la (Se)-SAC la protéine maintient les réactifs en contact van der Waals [Figure-III-92-a,b&c], ce qui a pour effet de maximiser le recouvrement orbitalaire entre les atomes impliqués dans la formation et clivage de liaisons [Figure-III-94]. L'alignement presque parfait entre les atomes  $Fe_4---S_8-C5'---S(e)1-C2$ , maintenu tout au long de la réaction [Figure-III-84], et la dégénérescence orbitalaire le long de cet axe nous font penser que le réarrangement électronique du système se fait en une seule étape [Figure-III-95]. En d'autres termes, le clivage réductif de la SAM, la formation d'une nouvelle liaison C-S(e) et le clivage de la liaison S(e)-C2 procèdent par un mécanisme concerté. Ceci suggère, que dans le cas étudié, le radical 5'-dA• appartient à un état de transition d'un système conjugué et n'est donc pas une espèce isolable.

Après le clivage de la liaison S(e)-C2 [Figure-III-96], la réduction du centre  $[Fe_4S_4]^{2+}$  avec l'électron libre du  $C2\bullet$  conduirait à la formation d'une kétimine secondaire (base de Schiff) entre les atomes C2 et N3 [Figure-III-97]. Cette liaison pourrait être facilement hydrolysée par une des nombreuses molécules d'eau qui entourent le ligand dans le site actif de HydE [Figure-I-36]. La conséquence directe de cette hydrolyse serait la formation de la (Se)-SAC et du pyruvate libre [Figure-III-85 & Figure-III-97]; l'accumulation avec le temps de ces deux produits a été prouvée expérimentalement [Figure-III-80-b & Figure-III-81-b3].

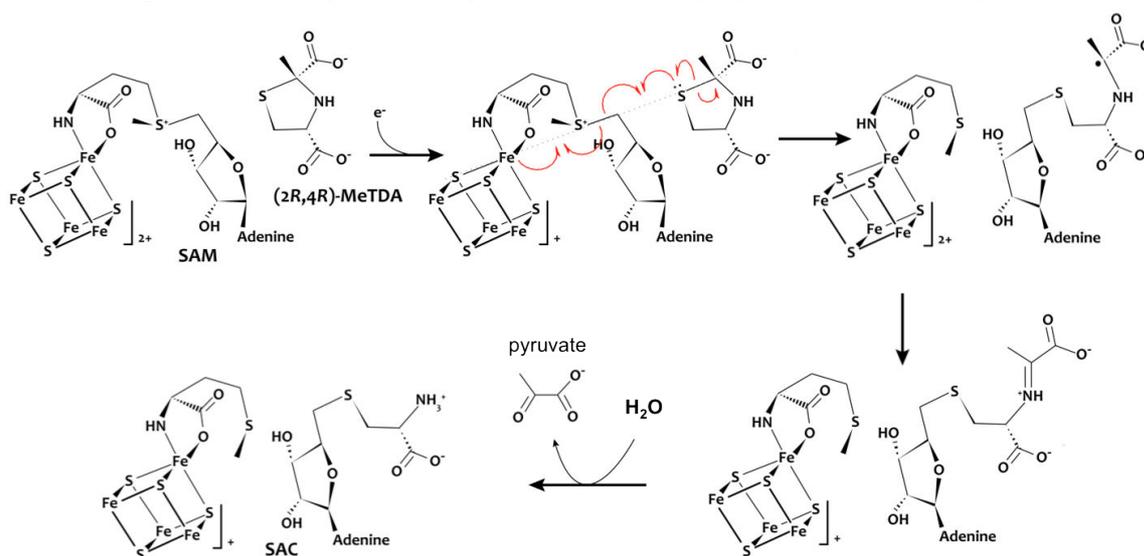


Figure-III-97 : Mécanisme détaillé de la formation enzymatique de la SAC par HydE

La production de la (Se)-SAC à partir de la SAM et du ligand (2R,4R)-Me(Se)TDA procède, dans un premier temps, par un mécanisme concerté. Pendant cette étape, le clivage de la SAM, la formation de la liaison C5'-S(e)1 et le clivage de la liaison S(e)1-C2 se font de façon concomitante. Cette partie de la réaction fait intervenir le doublet non liant de l'atome S(e)1 engagé dans une liaison thioéther (ou sélénoéther). L'oxydation radicalaire de l'espèce intermédiaire avec le radical sur l'atome C2 permet de régénérer le centre  $[Fe_4S_4]$  dans son état réduit et a pour conséquence la formation d'une base de Schiff entre le carbone C2 et l'azote N3. Cette liaison peut être facilement hydrolysée, ce qui conduira à la formation du produit final (Se)-SAC et à la libération du cofacteur pyruvate.

### III-8 Discussion

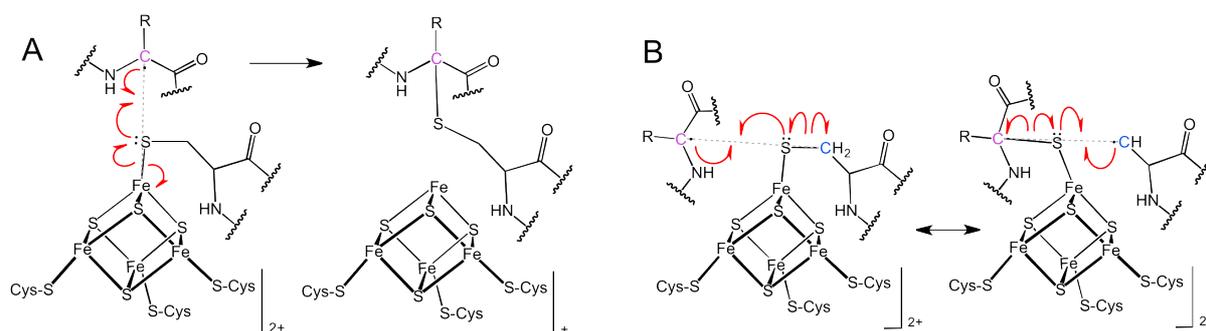
Nous venons de voir que la configuration des réactifs dans le site actif de HydE est telle que le clivage de la SAM et l'attaque du radical 5'-dA• sur l'atome de chalcogène du substrat peuvent se faire en une seule étape. L'alignement et les distances entre les atomes qui sont impliqués dans la création et le clivage des liaisons atomiques sont presque identiques à ceux des structures optimisées des modèles utilisés pour les calculs QM de substitutions homolytiques sur des atomes de S, Se et Te, reporté dans la littérature<sup>27-29</sup>. Dans ces modèles nous pouvons remarquer une légère déviation de la linéarité entre le chalcogène et les deux atomes de carbone qui l'entourent. Cette même déviation (l'angle C5'-S(e)1-C2 est d'environ 164°) est reproduite par HydE. En d'autres termes, la configuration des réactifs dans HydE est très proche de la configuration idéale pour une substitution homolytique sur un atome de chalcogène. Les déplacements des atomes sont minimes et ils se limitent à un mouvement de translation le long d'un axe. La compacité du site actif, contrôlée par la matrice protéique, maximalise le recouvrement orbitalaire, ce qui se traduit par un abaissement de la barrière énergétique de l'état de transition. Grâce à un tel arrangement du site actif, la protéine HydE arrive à faire une réaction de désulfuration oxydative en milieu anoxique.

#### III-8.1 Implications sur les réactions d'insertions d'atome de soufre catalysées par les enzymes à radical SAM

La réaction observée dans HydE est relativement éloignée des réactions d'arrachage d'un atome d'hydrogène, habituellement catalysées par les enzymes à radical SAM. Cependant il est intéressant de la comparer aux réactions de création de la liaison C-S catalysées par certaines protéines à radical SAM, comme par exemple BioB, LipA, RimO, MiaB, AlbA ou encore SkfB. Elles ont toutes un centre [Fe-S] auxiliaire, qu'elles utilisent pour activer un thiol du substrat ou directement comme source d'atome de soufre, qui sera inséré sur le substrat [voir IB-5]. Ces enzymes utilisent le radical 5'-dA• pour générer un radical alkyl sur le substrat qui va ensuite interagir avec l'atome du soufre, qui est lié au centre [Fe-S] auxiliaire. Un des ions ferriques de l'agrégat [Fe-S] sert de piège à l'électron non-apparié et sa réduction va terminer la réaction radicalaire. Dans HydE, nous observons la génération et l'attaque concertées d'un radical alkyl sur un atome de soufre qui n'est pas lié à un centre [Fe-S], mais qui est engagé dans une liaison thioéther. L'oxydation du substrat se fait par la libération d'un groupe partant radicalaire, qui sera vite oxydé par le centre [Fe<sub>4</sub>S<sub>4</sub>]<sup>2+</sup> 'SAM radical'.

Nous pensons que les propriétés réactives du thioéther vis-à-vis des radicaux alkyl pourraient être extrapolées à l'atome de S activé par un centre [Fe-S]. Cela s'appliquerait plus particulièrement à des cas où l'atome de soufre est déjà lié à un atome de carbone, comme dans les méthylthiotransférases RimO et MiaB, les enzymes impliquées dans la production des sactibiotiques AlbA et SkfB ou encore lors de la seconde étape d'insertion de l'atome de soufre dans BioB. Dans ces cas, le radical alkyl va attaquer un atome de soufre lié au Fe

différencié et à un atome de carbone ; il est donc important que la protéine s'assure que la liaison qui sera clivée après l'attaque du radical sera celle entre l'atome de S et le Fe et pas celle entre le soufre et le carbone [Figure-III-98-A]. Dans le cas contraire cela ne conduirait pas à la formation d'une liaison thioéther, mais à la création d'un second radical alkyl [Figure-III-98-B]. Ainsi il est fort probable que la directionnalité de l'attaque radicalaire, le recouvrement orbitalaire optimal et le clivage de liaison par un mécanisme 'backside' soient des éléments clés pour diriger la réaction vers la formation d'une liaison thioéther et la réduction concomitante d'un ion ferrique plutôt que vers la formation d'une nouvelle liaison C-S et le départ d'un groupe partant radicalaire organique.



**Figure-III-98 : Possibles configurations du site actif de la protéine Alba**

**A** - Le correct alignement de l'atome radicalaire C $\alpha$  (en violet) avec la liaison S $\gamma$ -Fe $_{\text{u}}$  conduirait à la formation d'une liaison thioéther entre les deux résidus et à la réduction du centre [Fe $_4$ S $_4$ ]. **B** - Alternativement, si l'attaque se fait dans l'axe de la liaison S $\gamma$ -C $\beta$  du résidu cystéine, la substitution homolytique S $_{\text{H}2}$  sur le soufre produirait d'un côté un résidu R lié au Fe unique par un thiolate connecté à l'atome C $\alpha$  et de l'autre côté un radical alanyl. Puisque cette réaction produirait un second radical alkyl à proximité du thioéther, il est possible qu'elle soit réversible.

Il est connu que certaines protéines à radical SAM, comme par exemple la LAM ou la SPL, régénèrent la SAM à la fin de chaque cycle. Cette étape de reformation du cofacteur est censée passer par la régénération du radical 5'-dA•, qui va ensuite attaquer le soufre du thioéther appartenant à la L-méthionine liée au Fe unique de l'agrégat [Fe $_4$ S $_4$ ] 'SAM radical'<sup>6,36</sup>. Ceci conduirait à la formation du sulfonium de la SAM et à la réduction du centre [Fe $_4$ S $_4$ ]. Or, cette attaque du radical alkyl sur l'atome de soufre du thioéther ressemble beaucoup à la réaction caractérisée dans HydE. A l'exception près que dans ce cas il n'y a pas de clivage de liaison, et l'électron libre se retrouve sur l'agrégat [Fe $_4$ S $_4$ ] et non pas sous la forme d'un radical organique. Il sera intéressant de comparer, lors d'une prochaine étude, nos résultats sur HydE aux modèles de la régénération de la SAM pour voir comment la directionnalité d'attaque oriente la réaction vers la reformation de la liaison C5'-S $_{\delta}$  et empêche une désulfuration accidentelle. Cette étude devrait également répondre à la question si la génération du radical 5'-dA•, la reformation de la liaison S $_{\delta}$ -C5' et la réduction du centre [Fe $_4$ S $_4$ ] 'SAM radical' se font en une étape.

## III-8.2 Implications sur le mécanisme d'initiation du réarrangement radicalaire du substrat catalysé par les enzymes à radical SAM

### III-8.2.1 Régiosélectivité du clivage de la SAM

Le bon alignement entre les atomes impliqués dans le réarrangement des liaisons atomiques est sans doute aussi responsable de la régiosélectivité du clivage de la SAM. Plusieurs auteurs ont souligné le fait que le clivage de chacune des trois liaisons C-S<sub>δ</sub> de la SAM devrait être possible [Figure-III-99]<sup>46,47</sup>. Cependant, la quasi-totalité des enzymes à radical SAM clivent la liaison S<sub>δ</sub>-C5' pour produire le radical 5'-dA•. Les seules exceptions connues à ce jour sont la protéine Dph2 et la glycérol-déhydratase activase GD-AE<sup>47</sup>. Ces enzymes clivent la liaison C<sub>γ</sub>-S<sub>δ</sub> et produisent un radical 3-amino-3-carboxypropyl. Dans Dhp2, ce dernier sera par la suite attaché sur une des doubles liaisons d'un résidu histidine lors de la maturation du facteur d'élongation 2<sup>48</sup>. Dans GD-AE, le radical sert à arracher un atome d'hydrogène sur un résidu glycine pour former un radical glycyloxy<sup>49,50</sup>. Ces protéines ne possèdent pas le motif CX<sub>3</sub>CX<sub>2</sub>C et les résidus cystéines impliqués dans la fixation du centre [Fe<sub>4</sub>S<sub>4</sub>] sont dispersés le long de la séquence. Il est donc fort probable qu'elles fixent la SAM de façon à maintenir un alignement entre le Fe<sub>u</sub>, les atomes C<sub>γ</sub> et S<sub>δ</sub> de la SAM et le carbone cible de l'attaque radicalaire. Cela contrasterait avec l'alignement Fe<sub>u</sub>--S<sub>δ</sub>-C5'--C/N qu'on trouve habituellement dans les structures des protéines appartenant à cette superfamille [Figure-III-101]. En se basant sur les structures cristallographiques disponibles, l'angle Fe<sub>u</sub>--S<sub>δ</sub>-C5' varie entre 139° et 161° et la distance Fe<sub>u</sub>--S<sub>δ</sub> varie entre 3,1 et 3,6 Å<sup>51</sup>. Il a été montré que dans cette configuration il y a une interaction entre les orbitales du centre [Fe<sub>4</sub>S<sub>4</sub>] et l'orbitale σ\* de la liaison S<sub>δ</sub>-C5'<sup>13,52</sup>. Il sera intéressant de voir si la structure cristallographique de l'holoenzyme Dph2 ou GD-AE confirmera cet alignement et si les calculs théoriques montreront une interaction entre les orbitales de l'agrégat [Fe-S] et l'orbitale σ\* de la liaison S<sub>δ</sub>-C<sub>γ</sub>.

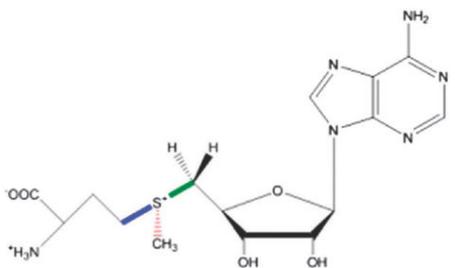
S-adénosyl-L-méthionine	Liaison clivée	Radical formé	Exemples
	S — 5'C	5'-dA	HydG, BioB, NosL, RimO, AlbA, QueE, BtrN, anSME, SPL, PFL-AE etc.
	S — γC	ACP	Dph2 & GD-AE
	S — CH <sub>3</sub>	Méthyl •CH <sub>3</sub>	Pas connu

Figure-III-99 : Trois clivages homolytiques réductives de la SAM (adapté de Hutchison Metallomics 2012 )

La quasi-totalité des protéines à radical SAM clivent la liaison S<sub>δ</sub>-C5' (en vert), bien que la liaison S<sub>δ</sub>-C<sub>γ</sub> (en bleu) devrait voir la même solidité et produire une espèce radicalaire de stabilité comparable au 5'-dA•. Le clivage homolytique de la liaison S<sub>δ</sub>-C5' conduit à la formation du radical 5'-dA• et de la L-méthionine, tandis que le clivage de la liaison S<sub>δ</sub>-C<sub>γ</sub> conduit à la formation du radical ACP et de la S-méthyl-5'-thioadénosine (MTA), et le clivage de la liaison S<sub>δ</sub>-C<sub>ε</sub> (en rouge) conduit à la formation du radical méthyl et de la SAH.

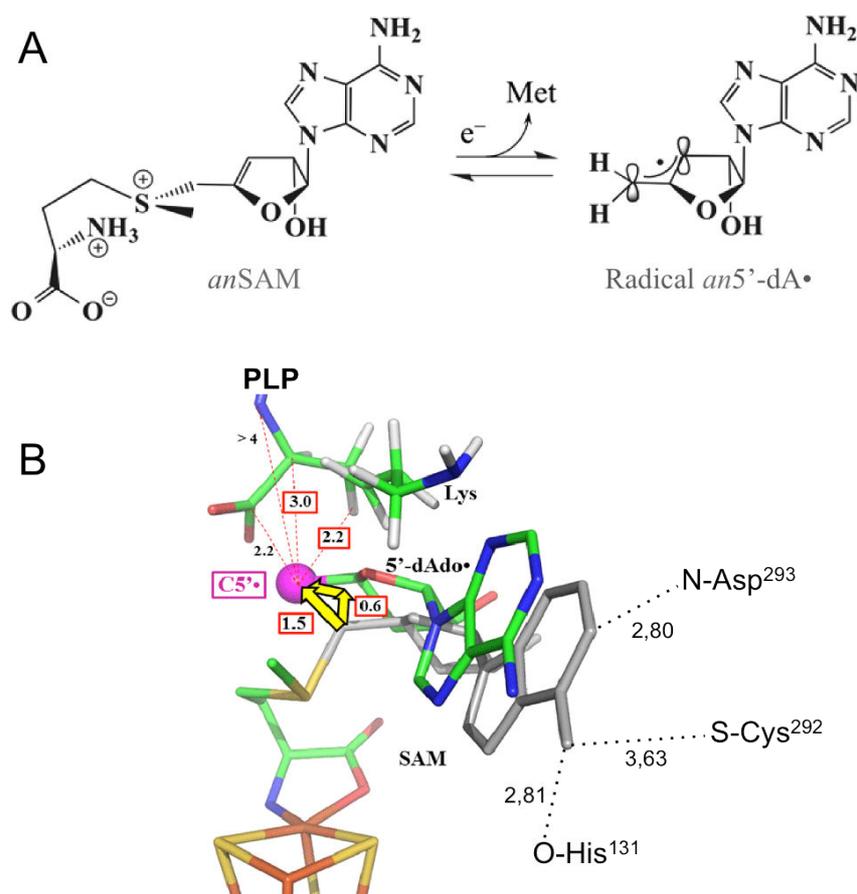
### III-8.2.2 *Contrôle des mouvements du 5'-dA dans le site actif*

Dans le paragraphe III-5.1 nous avons vu que dans HydE les contacts van der Waals entre les réactifs sont maintenus tout au long de la réaction. Nous avons aussi vu que la protéine contrôle le mouvement de la partie 5'-dA grâce à de nombreuses liaisons hydrogènes de façon à minimiser le mouvement total du 5'-dA tout en maximisant le déplacement de l'atome C5' vers sa cible [Figure-III-93].

Ces observations sont globalement en accord avec les conclusions d'une étude récente effectuée par Horitani et ses collaborateurs, qui ont utilisé la technique ENDOR pour suivre la transformation de la L-lysine par la LAM dans le but d'en déduire la variation des distances entre les atomes au cours de la réaction<sup>53</sup>. Les auteurs de l'étude ont utilisé un substrat sélectivement marqué sur certaines positions par substitution isotopique <sup>13</sup>C ou <sup>2</sup>H et un analogue de la SAM, la 3',4'-anhydre-S-adénosyl-L-méthionine (*anSAM*), capable de générer un radical alkyl stable dans le temps grâce à la délocalisation du caractère radicalaire sur une double liaison<sup>54,55</sup> [Figure-III-100-A]. La structure cristallographique de la LAM<sup>56</sup> ainsi que les premières expériences ENDOR sur l'enzyme<sup>57</sup> avaient montré qu'avant le clivage de la SAM les atomes C5' et le carbone cible du substrat sont presque en contact van der Waals. Les auteurs de la nouvelle étude ont proposé, à partir des contraintes sur les distances entre le C5' et les atomes du substrat, déduites des constantes de couplages hyperfins, que le clivage de la SAM fait augmenter la distance entre C5' et S<sub>δ</sub> à approximativement la somme de leur rayons van der Waals respectifs. Cette distance va donc passer de 1,9 Å à environ 3,2 Å, ce qui va "projeter" le carbone C5' vers le substrat. Les auteurs ont conclu que le radical alkyl du 5'-dA• n'est jamais libre et que son évolution dans l'espace est étroitement contrôlée par des contacts van der Waals. Chez les protéines à radical SAM la distance que l'atome C5' doit parcourir est donc minimisée, ce qui leur permet d'éviter tout mouvement non productif, ainsi que des réactions parasites qui en résulteraient.

Horitani et collaborateurs ont proposé un modèle du mouvement que la partie 5'-dA de la SAM doit exécuter après le clivage du cofacteur [Figure-III-100-B]. A cause de l'absence de données expérimentales sur la position des autres atomes du 5'-dA leur modèle est basé uniquement sur la variation de la distance entre le C5' et certains atomes du substrat. Un élément de ce modèle est en contradiction avec nos résultats, car il prévoit une translation de l'hétérocycle adénine. En effet, nous n'avons pas observé un tel mouvement dans notre étude cristallographique de la réaction catalysée par HydE sur les dérivés thiazolidines. Au lieu de cela nous avons clairement mis en évidence un léger mouvement de rotation d'environ 6° autour de l'azote N6 [Figure-III-91]. La comparaison de structures cristallographiques et d'alignements de séquences des enzymes à radical SAM montre que les interactions entre l'adénine et la protéine sont fortement conservées chez la superfamille [voir IB-3.1]. Notamment, il s'agit d'interactions avec les atomes N6, N1 et parfois C2 de l'adénine. Pour illustrer le degrés de conservation de ces interactions rappelons que l'atome N6 interagit avec le carbonyle du résidu Φ (Φ = Tyr, Phe, His, Trp) qui fait partie de la séquence CX<sub>3</sub>CXΦC - motif caractéristique de la superfamille de protéines à radical SAM<sup>58</sup>. Les atomes N6, N1 et C2 se trouvent à l'autre extrémité du 5'-dA par rapport à l'atome C5' [Figure-III-93 & Figure-III-101]. Il semble donc que dans les protéines à radical SAM le cycle adénine soit maintenu

en place par des interactions conservées qui l'empêcherait d'effectuer un mouvement de translation. L'adénine serait alors contrainte à effectuer des mouvements de rotation de faible amplitude (limités par les contacts entre les atomes N1 et C2 et la protéine) dont l'axe se situerait à proximité de son amine. Nous en concluons que le mouvement de rotation autour de l'azote N6 observée dans HydE est un élément commun aux protéines à radical SAM. Cette rotation, combinée à celle autour de la liaison N-osidique de la SAM, est un moyen efficace pour la protéine, en effectuant des minimaux mouvements, de maximiser le déplacement de l'atome C5' vers le substrat [Figure-III-93].



**Figure-III-100** : Utilisation de *anSAM* pour étudier les mouvements de l'atome C5' au cours de la réaction dans la LAM (Horitani *et al.* 2015)

A - Clivage réductif homolytique de la liaison Sδ-C5' de la 3',4'-anhydre-S-adenosyl-L-méthionine (*anSAM*) conduit à la formation d'un radical allylique *an5'-dA•* stable<sup>54,55</sup>. La modélisation de constantes de couplage hyperfin a montré que le spin de l'électron non apparié est distribué de façon homogène entre les deux carbones C5' et C3' du ribose. La densité de spin  $\rho$  pour chacun des atomes C5' et C3' est estimée à  $0,59 \pm 0,02$ <sup>55</sup>. L'atome de carbone se trouvant au milieu d'un système radicalaire allylique (ici l'atome C4') a une densité de spin légèrement négative; ceci s'expliquerait par la polarisation du spin des électrons liant cet atome de carbone à ces voisins<sup>59</sup>. B - Modèle du clivage de la *anSAM* et des mouvements qui l'accompagnent proposés par Horitani et ses collaborateurs. Ce modèle prévoit une translation du 5'-dA et ne tiens donc pas compte des interactions entre la partie adénine et la protéine. En plus, l'absence du groupement OH sur l'atome C3' de l'*anSAM* empêche l'interaction que cet hydroxyle a avec le carboxylate de la SAM dans la structure de l'enzyme avec la forme naturelle du cofacteur<sup>56</sup>. Il est donc possible que la fixation et le comportement de la partie ribose de l'*anSAM* dans le site actif soient affectés par cette modification. (Images adaptées de Horitani *et al.* 2015)

Il est intéressant de noter que la conservation des résidus en interaction avec le cofacteur se limite à des résidus qui interagissent avec la partie L-méthionine et adénosine<sup>58,60</sup>, tandis que la nature de résidus en interaction avec les deux groupements alcools du ribose est plus variée [Tableau-III-03]. Il est très tentant d'en conclure que dans les protéines à radical SAM l'amplitude de la rotation autour du N6 sera la même que dans HydE, par contre l'amplitude de la rotation autour de la liaison N9-C1' pourrait dépendre de la nature d'interaction entre les groupement alcools du ribose et la protéine. De cette manière chaque protéine à radical SAM pourrait contrôler de façon précise la distance que l'atome C5' doit parcourir après le clivage de la SAM. Il reste à vérifier s'il existe une réelle coévolution entre la position du substrat dans le site actif et le mode de fixation du ribose.

Le rôle du ribose dans le contrôle du mouvement du radical 5'dA a été étudié dans le cas des enzymes à adénosylcobalamine. Il a été proposé que les deux groupements alcools servent de "poignées polaires" que la protéine utilise à la fois pour faciliter le clivage homolytique de la liaison C5'-Co et pour assister la diffusion contrôlée du radical vers le substrat<sup>61</sup>. La deuxième partie de cette hypothèse pourrait s'appliquer aussi aux protéines à radical SAM. La présence de groupements fonctionnels par lesquels la protéine peut diriger le mouvement de la partie 5'-dA peut expliquer pourquoi les protéines à radical SAM clivent de façon préférentiel la liaison S<sub>δ</sub>-C5' et non pas la liaison S<sub>δ</sub>-C<sub>γ</sub>. Un radical ACP possède moins de groupements fonctionnels et une flexibilité plus élevée que le radical 5'-dA, il est donc plus difficile de le guider vers le substrat et d'éviter des réactions parasites. Par ailleurs, sa génération s'accompagne de la formation d'une molécule de MTA [Figure-III-99], qui est un inhibiteur compétitif de ces enzymes.

Dans la protéine *TmHydE* le résidu Glu<sup>161</sup> qui interagit avec le ribose de la SAM fait partie du motif R<sup>155</sup>YLLRXE<sup>161</sup>T qui caractérise les maturases HydE [voir l'alignement multiséquence en Annexe]. Bien qu'il est vrai que les résidus en interaction avec le ribose ne sont pas conservés parmi les séquences appartenant aux différents membres de la superfamille des protéines à radical SAM, les structures disponibles montrent que le plus souvent c'est effectivement un résidu glutamate qui interagit avec les deux groupement alcools du ribose de la même manière que dans la structure de *TmHydE*. Il est donc très probable que dans les enzymes qui possèdent un tel résidu glutamate, le 5'-dA va exécuter le même mouvement que dans HydE et plus important, que l'atome C5' va parcourir sensiblement la même distance vers sa cible.

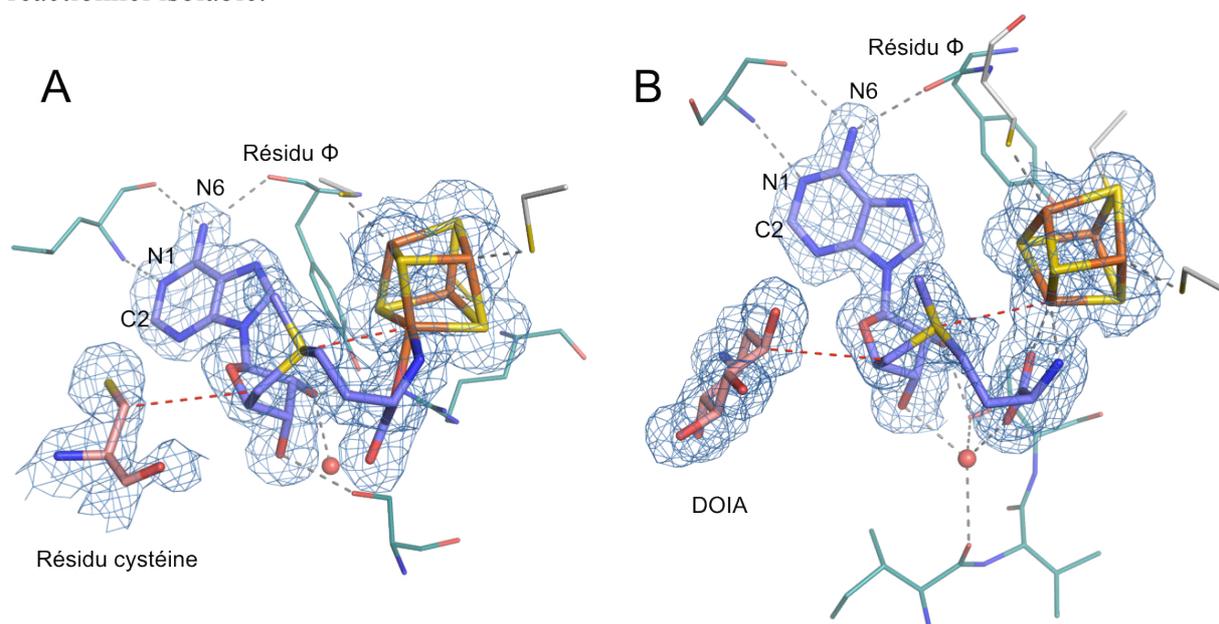
Protéine	code PDB	Résolution (Å)	Distance C5'-C/N (Å)	Interaction ribose-protéine
BioB	1R30 <sup>62</sup>	3,40	3,70 *	Asn <sup>153</sup> -Leu <sup>154</sup> -Asp <sup>155</sup>
NosL	4R33	1,78	3,74 *	<u>Glu<sup>204</sup></u>
BtrN	4M7T <sup>63</sup>	1,56	3,77	Ile <sup>113</sup> -Val <sup>114</sup> -Thr <sup>115</sup>
LAM	2A5H <sup>56</sup>	2,10	3,80	<u>Gln<sup>258</sup></u>
QueE	4NJI <sup>64</sup>	2,20	3,83	<u>Ser<sup>133</sup>-Pro<sup>134</sup>-Lys<sup>135</sup></u>
SPL	4FHD <sup>65</sup>	2,00	3,88 *	Ser <sup>195</sup> , Ala <sup>234</sup> , Lys <sup>174</sup>
PFL-AE	3CB8 <sup>66</sup>	2,77	3,93 *	<u>Asp<sup>129</sup>, Lys<sup>131</sup>, Tyr<sup>35</sup></u>
anSME	4K38 <sup>67</sup>	1,83	4,08	<u>Ser<sup>122</sup>, Arg<sup>134</sup></u>
ThiC	4S25 <sup>51</sup>	1,45	4,17 *	<u>Glu<sup>489</sup></u>
MoaA	2FB3 <sup>68</sup>	2,35	> 5,30 * §	<u>Ser<sup>126</sup>-Leu<sup>127</sup>-Asp<sup>128</sup>, Tyr<sup>30</sup></u>
RlmN	3RFA <sup>20</sup>	2,05	5,92 #	<u>Ser<sup>233</sup>-Leu<sup>234</sup>-His<sup>235</sup></u>
PylB	3T7V <sup>69</sup>	1,50	4,30 * [produit]	<u>Glu<sup>171</sup></u>
LipA	4UOP <sup>70</sup>	1,62	2x [Fe <sub>4</sub> S <sub>4</sub> ] + L-Met	<u>Glu<sup>171</sup></u>
RimO	4JCO <sup>71</sup>	3,30	2x [Fe <sub>4</sub> S <sub>4</sub> ] + polypersulfures	Gln <sup>256</sup> , Arg <sup>268</sup> ?
HydG	4WCX <sup>72</sup>	1,59	[Fe <sub>4</sub> S <sub>4</sub> ] + SAM + [Fe <sub>5</sub> S <sub>4</sub> ]-L-homoCys	<u>Glu<sup>194</sup></u>
HydG	4RTB	2,79	[Fe <sub>4</sub> S <sub>4</sub> ] + SAM *	<u>Glu<sup>189</sup></u>
HemN	1OLT <sup>73</sup>	2,07	[Fe <sub>4</sub> S <sub>4</sub> ] + 2x SAM	<u>Gln<sup>172</sup>, Asp<sup>209</sup></u>
HydE	3IIZ <sup>13</sup>	1,62	[Fe <sub>4</sub> S <sub>4</sub> ] + SAM + [Fe <sub>2</sub> S <sub>4</sub> ]	<u>Glu<sup>161</sup></u>

**Tableau-III-03** : Analyse des distances interatomiques entre le carbone C5' de la SAM et le C ou N cible du substrat dans le site actif de protéines à radical SAM, dont la structure cristallographique est disponible

La distance entre l'atome C5' de la SAM (ou de la SAH) et l'atome de C ou N du substrat qui sera cible d'arrachage d'un atome d'hydrogène a été mesurée pour chaque protéine dont la structure présentait un site actif avec le centre [Fe<sub>4</sub>S<sub>4</sub>] 'SAM radical', le cofacteur SAM/SAH et au moins un des substrats. Dans le cas contraire, la composition de la cavité est détaillée. Dans le cas de PylB, au lieu de contenir le substrat L-lysine, la cavité contient la SAM avec le produit (3R)-3-méthyl-D-ornithine. Il a été proposé qu'un radical organique sur le substrat pourrait arracher un atome d'hydrogène sur le C5' du 5'-dAH pour régénérer le cofacteur et terminer la transformation du substrat en produit<sup>69</sup>. Dans MoaA (S), la distance entre C5' et le carbone C3' du substrat est inhabituellement longue. Par contre les expériences de RPE ont montré que dans la structure cristallographique le substrat GTP ne fixe pas le second agrégat [Fe-S] de façon physiologique<sup>74</sup>. Dans un modèle, qui prends en compte les données RPE, l'atome C3' du GTP pourrait se positionner différemment par rapport à l'atome C5' de la SAM. Dans le cas de la protéine RlmN (#), un réarrangement important de la structure est attendu après la fixation du second substrat qui est l'adénosine 2503 de l'ARN ribosomique 23S. Il est à noter que la comparaison de centaines de nos structures de *TmHydE* a révélé que la SAH peut se fixer sur le centre [Fe<sub>4</sub>S<sub>4</sub>] 'SAM radical' exactement de la même manière que la SAM, cependant elle peut présenter une conformation alternative, sans doute a cause de l'absence du méthyl et donc de l'absence d'une contrainte géométrique et stérique. Dans cette conformation la distance entre le Sδ et le Fe unique est allongée, mais surtout, l'alignement entre les atomes C5'-Sδ-Fe n'est pas optimale. Puisque certaines structures cristallographiques reportées dans ce tableau ont été obtenues avec la SAH, il est possible que la configuration du site actif ne soit pas exactement la même que celle de l'holoenzyme avec le substrat et le cofacteur intact. Toute structure où l'identité du cofacteur n'a pas pu être confirmée de façon non ambiguë grâce à la présence du méthyl de la SAM dans les cartes de densité électronique est marquée par (\*). La dernière colonne du tableau contient la liste des résidus qui interagissent avec les deux groupements alcools de la partie ribose de la SAM (résidus en noir). Les résidus qui dominent cette interaction sont soulignés. Les résidus en gris n'interagissent pas directement avec la SAM, mais assurent le bon repliement de la boucle qui porte les résidus dont les chaînes latérales forment des liaisons hydrogène avec le ribose. Dans le cas de RimO, la structure ne contenait pas de cofacteur; les interactions ont donc été déduites (d'où le point d'interrogation) à partir de l'alignement de sa structure avec celles des autres protéines à radical SAM et en tenant compte de la conservation de résidus dans les séquence de protéines RimO issues de divers organismes.

### III-8.2.3 Le 5'-dA• comme un état de transition dans les enzymes à radical SAM

L'analyse de structures cristallographiques des complexes de protéines à radical SAM avec leurs substrats révèle que le carbone C5' de la SAM se trouve systématiquement face à l'atome du substrat, à partir duquel il va arracher un atome d'hydrogène. La distance qui sépare ces deux atomes est comprise entre 3,7 et 4,2 Å [Tableau-III-03]. Ainsi la compacité du site actif avant le clivage de la SAM, observée dans *TmHydE* et la LAM, serait une propriété commune de ces enzymes. Une seconde caractéristique générale est l'alignement des atomes impliqués dans le clivage et la formation des liaisons [Figure-III-101]. Ces deux observations soulèvent la question de l'existence du radical 5'-dA• comme un intermédiaire réactionnel isolable.



**Figure-III-101** : Détail du site actif de la structure cristallographique de protéines anSME et BtrN qui illustre la compacité du site actif et l'alignement entre les atomes impliqués dans le clivage et la formation de liaisons

**A** - Le détail du site actif de anSME avec son substrat - un résidu de cystéine de l'arylsulfatase<sup>67</sup>. **B** - Le détail du site actif de BtrN avec son substrat - 2-déoxy-scyloinosamine (DOIA)<sup>63</sup>. En traits pointillés rouges est représenté, de gauche en droite, l'alignement entre les atomes C(cible)--C5'-S<sub>8</sub>--Fe<sub>u</sub>.

Depuis la découverte de la LAM, il y a 46 ans, ce radical furtif n'a jamais été détecté dans les protéines à radical SAM. Son existence dans les protéines à adenosylcobalamine reste très probable (voir le paragraphe suivant). En remplaçant la SAM par la anSAM, le turnover de la LAM est divisé par 400 et un radical 3',4'-anhydro-5'-déoxyadénosyl est détecté<sup>75</sup>. Son observation pourrait être considérée comme un argument pour l'existence d'un radical 5'-dA• éphémère dans les protéines à radical SAM. Cependant la stabilisation résonante du radical allylique an5'-dA• ne reflète pas exactement le caractère radicalaire du 5'-dA•. En effet les données RPE et les calculs théoriques montrent que le spin de l'électron non apparié est distribué de façon homogène entre les carbones C5' et C3'<sup>55</sup> [voir la légende de la Figure-III-100-A]. Cette distribution contraste avec celle du radical 5'-dA•, calculée à partir des données cristallographiques de HydE [Figure-III-95]. La stabilisation du radical

*an*5'-dA• peut fortement modifier le profil énergétique de la réaction et créer un intermédiaire qui n'existerait pas autrement.

En étudiant la réaction catalysée par HydE sur les dérivés thiazolidine nous avons pu accumuler des indices qui suggèrent que dans le cas de HydE, le clivage de la SAM et l'addition radicalaire de l'atome C5' sur le chalcogène se font en une seule étape. Certains de ces indices sont de nature expérimentale, comme par exemple la compacité du site actif, l'alignement quasi-parfait des atomes concernés par la réaction ou encore l'absence de détection d'états intermédiaires. D'autres proviennent de calculs théoriques, comme la différence importante entre la barrière de clivage de la SAM et la barrière de la formation de la liaison C5'-S(e)1 ou encore la dégénérescence orbitale le long de l'axe  $Fe_u---C2$  qui se traduit par la perturbation de la structure électronique du substrat en présence du centre  $[Fe_4S_4]$  réduit avant même le clivage réductif de la SAM [Figure-III-95-a].

Aucune étude comparable à la notre n'a été effectuée avec un arrachage d'un atome d'hydrogène. Cependant la configuration du site actif de différentes protéines à radical SAM (compacité et alignement des atomes) nous fait penser que les principes structuraux et les mécanismes électroniques observés dans HydE devraient gouverner également la réaction d'arrachage de l'atome d'hydrogène catalysée par ces enzymes. Nous proposons que l'arrachage d'un atome d'hydrogène procéderait de façon concomitante au clivage de la SAM. Ainsi, dans les protéines à radical SAM, le radical 5'-dA• ne serait pas une espèce intermédiaire isolable, mais plutôt un élément de l'état de transition.

### III-8.3 Le 5'-dA dans les enzymes à radical SAM et dans les enzymes à adénosylcobalamine

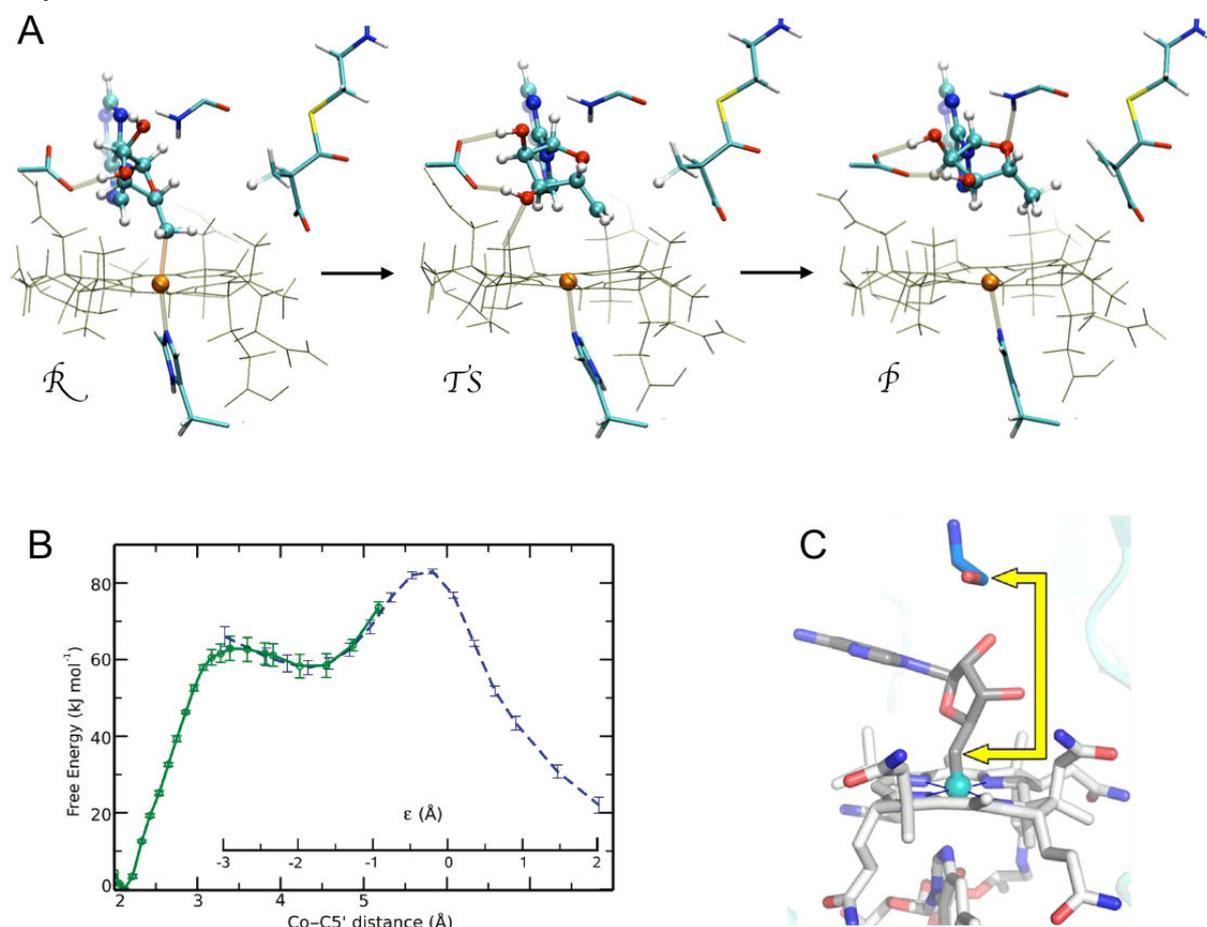
La formation du radical 5'-dA• dans les protéines à adénosylcobalamine a longtemps été discutée. Bien que ce radical n'ait jamais été directement détecté, plusieurs indices expérimentaux suggèrent que dans les enzymes à  $B_{12}$  ce radical probablement éphémère chez les enzymes à radical SAM pourrait exister comme espèce isolable.

En effectuant des mesures ENDOR sur l'éthanolamine ammonia-lyase (EAL), une enzyme bactérienne impliquée dans la déamination des aminoalcools vicinaux, Lobrutto et collaborateurs ont montré que son substrat se fixe à environ 12 Å de l'atome de cobalt et que l'atome C5' effectue un déplacement sur une distance de près de 7 Å vers sa cible<sup>76</sup>. Ces observations ont été confirmées par d'autres groupes<sup>77,78</sup>. Une étude similaire avec la lysine 5,6-aminomutase (une enzyme  $B_{12}$ -dependante) a révélé un déplacement d'environ 6 Å de l'atome C5' en direction du carbone cible<sup>79</sup>. Ces distances sont effectivement beaucoup trop longues pour que la réaction de clivage de la liaison Co-C5' puisse se faire de façon concomitante à l'arrachage de l'atome d'hydrogène.

Sans doute la preuve la plus parlante de l'existence du radical 5'dA• dans ces enzymes a été apportée par Chen et collaborateurs, qui ont étudié l'arrachage d'un atome d'hydrogène à partir d'un résidu cystéine catalysé par la ribonucléotide réductase (RNR) de classe II [pour plus de détails sur la réaction catalysée par cette enzyme voir paragraphe III-2.1.2]. Ces chercheurs ont utilisé de la (5'-S)-[5'-<sup>2</sup>H<sub>1</sub>]-adénosylcobalamine pour étudier l'épimérisation du proton et du deutérium sur l'atome C5' après le clivage et la reformation de la liaison Co-C5'.

Ils ont pu montrer que cette épimérisation est beaucoup plus rapide que le reste de la réaction catalysée par l'enzyme et plus important encore, cette épimérisation a lieu également quand le résidu cystéine cible est muté en alanine. Ainsi il semblerait que le thiol cible d'arrachage de l'atome d'hydrogène n'est pas indispensable au clivage de la liaison Co-C5'. Les auteurs ont conclu que l'enzyme arrive à effectuer des clivages réversibles de cette liaison Co-C5' et générer de façon réversible le radical 5'-dA•<sup>80</sup>. Cette épimérisation rapide avait déjà été observée dans la méthylmalonyl-CoA mutase<sup>81</sup>. Notons qu'il serait extrêmement intéressant de voir le résultat d'une expérience analogue dans les protéines à radical SAM, car elle pourrait confirmer ou pas l'existence du radical 5'-dA• dans ces enzymes.

Récemment, une étude théorique, basée sur les structures cristallographiques de la méthylmalonyl-CoA mutase, a démontré que le radical 5'-dA• est effectivement une entité distincte<sup>61</sup>. Les trajectoires d'atomes au cours de la réaction, calculées par la dynamique moléculaire QM/MM, et l'évolution du profil énergétique montrent que le 5'-dA• doit exister pendant un laps de temps court [Figure-III-102-A & B]. D'autres sources ont estimé la durée moyenne d'existence de ce radical à environ 2 ns<sup>82</sup>.



**Figure-III-102 : Le radical 5'-dA• comme un intermédiaire de la réaction dans les enzymes AdoCbl**

**A** - Mouvements observés lors de l'initiation de la réaction radicalaire dans la méthylmalonyl-CoA mutase (MCM), obtenus par la dynamique moléculaire QM/MM. Les lettres R, TS et P signifient réactant, état de transition et produit, respectivement. **B** - Profil de l'énergie libre (en kJ.mol<sup>-1</sup>) du clivage homolytique de la liaison Co-C5' (en vert) et de l'arrachage d'un atome d'hydrogène (en bleu) correspondants à la réaction représentée dans A. Le minimum local à 400-450 pm; bien que peu profond, mais clairement visible, indique que le radical 5'-dA• est une espèce intermédiaire et que la réaction procède en deux étapes distinctes. **C** - Site actif de l'éthanolamine ammonia-lyase (EAL) avec le substrat (en bleu) fixé à 6,6 Å du carbone C5' et à 8,6 Å du cobalt. Images A & B adaptées de Bucher JCS 2012, image C adaptée de Horitani JACS 2015.

La comparaison des structures disponibles des enzymes à adénosylcobalamine montre que, contrairement à ce qu'on observe dans les protéines à radical SAM, il n'y a pas de réel consensus en ce qui concerne les distances entre le C5' et l'atome cible avant le début de la réaction. Dans la MCM par exemple, cette distance se situe entre 3,6 et 4,2 Å avec un angle Co-C5'---C<sub>cible</sub> qui varie entre 100° et 127° [code PDB: 4REQ & 2XIQ]. Dans la EAL, cette distance est de 6,56 Å. Par contre l'alignement des atomes Co-C5'---C<sub>cible</sub> est très marquant; environ 173° [code PDB: 3ABS] [Figure-III-102-C]. Ainsi dans la MCM l'attaque du radical 5'-dA• ne se fait pas du tout dans l'axe du clivage de la liaison Co-C5' et dans la EAL la distance qui sépare le C5' de l'atome cible rend impossible d'effectuer le clivage de la liaison Co-C5' et l'arrachage de l'atome d'hydrogène de façon concomitante. Puisque l'absence d'alignement d'atomes et la distance élevée qui les sépare ont été observées dans d'autres structures des enzymes à adénosylcobalamine<sup>83</sup> il est hautement probable que dans le mécanisme de ces enzymes le radical 5'-dA• soit une espèce intermédiaire distincte. Il semblerait donc que ce radical puisse exister dans l'environnement protéique, mais sa génération est fortement couplée avec l'arrachage d'un atome d'hydrogène<sup>84,85</sup> ou avec la reformation de la liaison Co-C5'<sup>80</sup>.

Ces observations pourraient contribuer à l'explication, avancée par Horitani et ses collaborateurs, que même s'il y a plus de 100 000 protéines à radical SAM connues à travers tout le domaine du vivant, on ne connaît que 12 protéines à adénosylcobalamine et aucune ne se trouve chez les plantes<sup>53</sup>. Les protéines à radical SAM ont évolué pour pouvoir accueillir une grande variété de substrats, tout en gardant un contrôle étroit sur l'étape de l'initiation de la réaction radicalaire. Grâce au mode de fixation de la SAM sur le centre [Fe<sub>4</sub>S<sub>4</sub>] et à la topographie de leur domaine TIM, la taille des substrats peut varier d'un acide aminé jusqu'à une protéine entière. Le substrat se lie à proximité du ribose du 5'-dA avec le C5' qui est face à l'atome à partir duquel il va arracher un atome d'hydrogène. En maintenant la compacité du site actif et un alignement des atomes impliqués dans la formation et le clivage de liaisons, ces enzymes arrivent à effectuer le clivage de la SAM et l'arrachage de l'atome d'hydrogène du substrat en une seule étape. Le déplacement de l'atome C5' sur une distance relativement courte est conditionné par des contacts van der Waals, ce qui minimise les risques de réactions parasites.

De l'autre côté, les protéines à adénosylcobalamine présentent des contraintes géométriques supplémentaires qui sont dues au mode de fixation du 5'-dA sur la cobalamine. Ces contraintes empêchent que l'atome cible se fixe à proximité du C5' ou que l'alignement des atomes soit optimal pour que la génération du radical 5'-dA• et l'initiation du réarrangement radicalaire du substrat puisse se faire en une seule étape. Ceci oblige l'enzyme à contrôler la diffusion du radical 5'-dA• sur des distances plus longues, ce qui peut expliquer pourquoi ces enzymes sont si peu répandues.

### III-8.4 Implication pour la maturase HydE

Dans le *Chapitre II* nous avons établi que les dérivés de thiazolidine qui sont ligands de HydE n'augmentent pas l'activation de l'hydrogénase à [FeFe] dans l'extrait. Ainsi, bien que la protéine HydE arrive à catalyser une transformation de ces ligands, ces derniers ne sont pas ses substrats physiologiques. Pour le moment il nous est difficile de lier de façon concrète la réaction observée dans les cristaux de HydE et confirmée en solution au rôle que cette enzyme joue dans le processus de la maturation de l'hydrogénase à [FeFe]. Cependant, l'observation de la formation d'une nouvelle liaison thioéther, combinée avec les résultats de Betz et de ces collaborateurs montre quels types de réactions le site actif de HydE peut catalyser. Le rôle central d'un atome de soufre dans ces réactions suggère que l'affinité du site *T* [voir paragraphe I-11.4.3] pour le soufre (et le sélénium) peut être une propriété physiologiquement pertinente.

Dans le but de collecter le maximum d'informations sur la réactivité de HydE nous avons voulu vérifier s'il était possible de catalyser la réaction vers un autre produit en utilisant les dérivés thiazolidines fonctionnalisés. Nous avons donc tenté d'utiliser un ligand thiazolidine (2*R*,4*R*)-BrMeTDA, fonctionnalisé au bromure sur le méthyl, obtenu en mélangeant de la L-cystéine avec du 3-bromopyruvate. Nous pensons que cette molécule pourrait se fixer dans TmHydE de la même manière que le ligand (2*R*,4*R*)-MMTDA [Figure-I-60-d] et que la présence du bromure sur le substrat permettrait de 'quencher' le radical sur le C2. Nous espérons éviter ainsi la libération du pyruvate et donc synthétiser une nouvelle molécule, dérivée de la SAC. Ces essais n'ont pas abouti, car lors de la préparation du dérivé en question nous avons systématiquement observé une perte de bromure, qui se manifestait par une rapide coloration de l'échantillon.

Nous avons effectué plusieurs séries d'essais de cinétiques dans les cristaux avec soit le 1,3-phénylènediacétate, le citrate ou le N-carbamoyl-L-aspartate. Dans le *Chapitre I* nous avons montré que ces trois molécules sont ligands de la protéine TmHydE [Figure-I-21, Figure-I-47-A et Figure-I-48-A, respectivement]. Nous espérons observer un clivage de la SAM concomitant à l'arrachage d'un atome d'hydrogène à partir d'un de ces ligands. Ceci nous aurait permis d'étudier en détails une réaction catalysée par la quasi-totalité des enzymes à radical SAM. Cependant, la faible affinité de ces molécules pour le site actif fait qu'elles sont rapidement éluées du site catalytique avant que la réaction ait le temps de commencer. Lorsqu'on élimine l'étape de lavage, c'est-à-dire qu'on garde un excès de ligand tout au long de la cinétique, nous observons, dans certains cristaux, une faible quantité de clivage de la SAM. Cependant les cartes de densité électronique ont été difficiles à interpréter. D'autres essais, notamment les mesures du clivage de la SAM en présence de ligands deutérés, seront nécessaires pour déterminer si un de ces ligands est aussi le substrat de HydE.

Suite à cet échec nous avons fait l'expérience avec le DTT, qui certes ne forme pas de complexe avec l'enzyme, mais qui est bien substrat de HydE, comme cela a été prouvé par Betz et ses collaborateurs [Figure-III-76]. Les cartes de densité électronique étaient une fois de plus difficiles à interpréter, car elles ne ressemblaient pas aux produits habituels du clivage de la SAM - la L-méthionine et le 5'-dAH, mais plutôt aux produits de sa dégradation hydrolytique en absence de substrat.

Ainsi, malgré nos efforts, nous n'avons pas pu utiliser les cristaux de *TmHydE* pour déceler les réactions radicalaires physiologiques catalysées par cet enzyme.

## CONCLUSIONS

La présente étude de la réaction de la protéine HydE avec le ligand (2*R*,4*R*)-(Se)MeTDA contient la première observation d'une réaction radicalaire dans les cristaux protéiques d'une enzyme à radical SAM et également un premier suivi en temps réel par la RMN du  $^{13}\text{C}$  et  $^1\text{H}$  de l'accumulation d'un des produits de la réaction catalysée par ces enzymes. La réaction observée dans HydE est la formation d'une nouvelle liaison thioéther entre le carbone C5' de la SAM et un soufre (ou sélénium) d'un cycle (séléno)thiazolidine du substrat. A notre connaissance il s'agit ici du premier exemple d'une enzyme à radical SAM qui fait réagir son atome C5' directement sur un thioéther ou sélénoéther.

L'analyse cristallographique a révélé que dans la cavité de HydE l'atome C5' n'exécute qu'un mouvement très restreint dans la direction de sa cible. Les calculs théoriques ont montré que grâce à la compacité du site actif et à l'alignement des atomes concernés, la formation du radical 5'-dA• et son attaque du thioéther procèdent de façon concomitante. En d'autres termes, le clivage de la liaison S $\delta$ -C5' de la SAM, la formation de la liaison C5'-S(e)1 et le clivage de la liaison et S(e)1-C2 du cycle thiazolidine se font en une seule étape.

La configuration du site actif de différentes protéines à radical SAM (compacité et alignement des atomes) suggère que les principes structuraux et les mécanismes électroniques observés dans HydE devraient gouverner également la réaction d'arrachage de l'atome d'hydrogène catalysée par ces enzymes. Nous proposons que l'arrachage d'un atome d'hydrogène procéderait de façon concomitante au clivage de la SAM. Ainsi, dans les protéines à radical SAM, le radical 5'-dA• ne serait pas une espèce intermédiaire isolable, mais plutôt un élément de l'état de transition. Les protéines à radical SAM auraient donc résolu le problème du contrôle d'une espèce hautement réactive en fusionnant en une seule étape l'évènement de la genèse de cette espèce radicalaire avec l'évènement de son attaque sur le substrat. Ceci, combiné à leur repliement et topologie qui leur permet de s'adapter à une grande variété de substrats de tailles différentes peut expliquer pourquoi ces enzymes sont tant répandues dans la nature.

Nous pensons que les éléments tels que le bon alignement des atomes, un recouvrement orbitalaire optimal avec un des doublets non liants du soufre et clivage et formation de liaisons par mécanisme '*backside*' seront retrouvés aussi dans les réactions d'insertion de l'atome de soufre dans les molécules biologiques catalysées par certaines protéines à radical SAM [voir paragraphe III-8.1].

Nos données apportent également des informations sur la possibilité d'insertion d'un atome de soufre sans avoir à recourir à des centres [Fe-S] et sur la possibilité d'utiliser les réactions de substitution homolytique sur un atome de soufre ou de sélénium pour produire de nouveaux radicaux organiques centrés sur un atome de carbone<sup>86</sup>.

### III-9 Références du Chapitre III

- 1 Betz, J. N. *et al.* [FeFe]-Hydrogenase Maturation: Insights into the Role HydE Plays in Dithiomethylamine Biosynthesis. *Biochemistry*, 1574462297, doi:10.1021/bi501205e (2015).
- 2 Li, H. & Rauchfuss, T. B. Iron Carbonyl Sulfides, Formaldehyde, and Amines Condense To Give the Proposed Azadithiolate Cofactor of the Fe-Only Hydrogenases. *Journal of the American Chemical Society* **124**, 726-727, doi:10.1021/ja016964n (2002).
- 3 Dénès, F., Pichowicz, M., Povie, G. & Renaud, P. Thiyl radicals in organic synthesis. *Chemical Reviews* **114**, 2587-2693, doi:10.1021/cr400441m (2014).
- 4 Shibata, N. & Toraya, T. Molecular Architectures and Functions of Radical Enzymes and Their (Re)activating Proteins. *Journal of Biochemistry* (2015).
- 5 Stubbe, J. & van Der Donk, W. A. Protein Radicals in Enzyme Catalysis. *Chemical Reviews* **98**, 705-762 (1998).
- 6 Frey, P. A., Hegeman, A. D. & Ruzicka, F. J. The Radical SAM Superfamily. *Critical Reviews in Biochemistry and Molecular Biology* **43**, 63-88, doi:10.1080/10409230701829169 (2008).
- 7 Gerfen, G. J., Licht, S., Willems, P. J., Hoffman, B. M. & Stubbe, J. Electron Paramagnetic Resonance Investigations of a Kinetically Competent Intermediate Formed in Ribonucleotide Reduction: Evidence for a Thiyl Radical-Cob(II)alamin Interaction. *Journal of the American Chemical Society* **118**, 8192-8197, doi:10.1021/ja960363s (1996).
- 8 Sintchak, M. D., Arjara, G., Kellogg, B. A., Stubbe, J. & Drennan, C. L. The crystal structure of class II ribonucleotide reductase reveals how an allosterically regulated monomer mimics a dimer. *Nature structural biology* **9**, 293-300, doi:10.1038/nsb774 (2002).
- 9 Rubach, J. K., Brazzolotto, X., Gaillard, J. & Fontecave, M. Biochemical characterization of the HydE and HydG iron-only hydrogenase maturation enzymes from *Thermotoga maritima*. *FEBS Letters* **579**, 5055-5060, doi:10.1016/j.febslet.2005.07.092 (2005).
- 10 Zhang, N., Schuchmann, H. & Sonntag, C. The reaction of superoxide radical anion with dithiothreitol: a chain process. *The Journal of Physical Chemistry* **95**, 4718-4722, doi:10.1021/j100165a024 (1991).
- 11 Griffiths, V. S. & Socrates, G. N.m.r. study of the hydration of pyruvic acid. *Transactions of the Faraday Society* **63**, 673-677, doi:10.1039/tf9676300673 (1967).
- 12 Nicolet, Y. *et al.* X-ray Structure of the [FeFe]-Hydrogenase Maturase HydE from *Thermotoga maritima*. *Journal of Biological Chemistry* **283**, 18861-18872, doi:10.1074/jbc.M801161200 (2008).
- 13 Nicolet, Y., Amara, P., Mouesca, J.-M. M. & Fontecilla-Camps, J. C. Unexpected electron transfer mechanism upon AdoMet cleavage in radical SAM proteins. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 14867-14871, doi:10.1073/pnas.0904385106 (2009).
- 14 Stich, T. A., Myers, W. K. & Britt, D. R. Paramagnetic Intermediates Generated by Radical S-Adenosylmethionine (SAM) Enzymes. *Accounts of Chemical Research* **47**, 2235-2243, doi:10.1021/ar400235n (2014).
- 15 Mehta, A. P. *et al.* Radical SAM Enzymes in Cofactor Biosynthesis: a treasure trove of complex organic radical rearrangement reactions. *Journal of Biological Chemistry*, doi:10.1074/jbc.R114.623793 (2014).
- 16 Mahanta, N., Fedoseyenko, D., Dairi, T. & Begley, T. P. Menaquinone Biosynthesis: Formation of Aminofutalosine Requires a Unique Radical SAM Enzyme. *Journal of the American Chemical Society* **135**, 15318-15321, doi:10.1021/ja408594p (2013).
- 17 Wagner, A. F., Demand, J., Schilling, G., Pils, T. & Knappe, J. A dehydroalanyl residue can capture the 5'-deoxyadenosyl radical generated from S-adenosylmethionine by pyruvate formate-lyase-activating enzyme. *Biochemical and biophysical research communications* **254**, 306-310, doi:10.1006/bbrc.1998.9931 (1999).

- 18 Moad, G., Rizzardo, E. & Thang, S. H. Radical addition–fragmentation chemistry in polymer synthesis. *Polymer* **49**, 1079-1131, doi:10.1016/j.polymer.2007.11.020 (2008).
- 19 Barr, I. *et al.* The pyrroloquinoline quinone (PQQ) biosynthetic pathway: Demonstration of de novo carbon-carbon cross-linking within the peptide substrate (PqqA) in the presence of the Radical SAM enzyme (PqqE) and its peptide chaperone (PqqD). *The Journal of biological chemistry*, doi:10.1074/jbc.C115.699918 (2016).
- 20 Boal, A. K. *et al.* Structural basis for methyl transfer by a radical SAM enzyme. *Science (New York, N.Y.)* **332**, 1089-1092, doi:10.1126/science.1205358 (2011).
- 21 Grove, T. L., Radle, M. I., Krebs, C. & Booker, S. J. Cfr and RlmN contain a single [4Fe-4S] cluster, which directs two distinct reactivities for S-adenosylmethionine: methyl transfer by SN2 displacement and radical generation. *Journal of the American Chemical Society* **133**, 19586-19589, doi:10.1021/ja207327v (2011).
- 22 Grove, T. L. *et al.* A substrate radical intermediate in catalysis by the antibiotic resistance protein Cfr. *Nature chemical biology* **9**, 422-427, doi:10.1038/nchembio.1251 (2013).
- 23 Grove, T. L. *et al.* A Radically Different Mechanism for S-Adenosylmethionine–Dependent Methyltransferases. *Science* **332**, 604-607, doi:10.1126/science.1200877 (2011).
- 24 Silakov, A. *et al.* Characterization of a cross-linked protein-nucleic acid substrate radical in the reaction catalyzed by RlmN. *Journal of the American Chemical Society* **136**, 8221-8228, doi:10.1021/ja410560p (2014).
- 25 Onofrio, N. & Mouesca, J.-M. Valence Bond/Broken Symmetry Analysis of the Exchange Coupling Constant in Copper(II) Dimers. Ferromagnetic Contribution Exalted through Combined Ligand Topology and (Singlet) Covalent-Ionic Mixing. *The Journal of Physical Chemistry A* **114**, 6149-6156, doi:10.1021/jp912114d (2010).
- 26 Onofrio, N. & Mouesca, J.-M. Analysis of the Singlet–Triplet Splitting Computed by the Density Functional Theory–Broken-Symmetry Method: Is It an Exchange Coupling Constant? *Inorganic Chemistry* **50**, 5577-5586, doi:10.1021/ic200198f (2011).
- 27 Lyons, J. E. & Schiesser, C. H. An ab initio study of some free radical homolytic substitution reactions at selenium. *Journal of Organometallic Chemistry* **437**, 165-176, doi:10.1016/0022-328x(92)83441-j (1992).
- 28 Aitken, H. M., Hancock, A. N. & Schiesser, C. H. Understanding (the lack of) homolytic substitution chemistry of sulfones. *Chemical Communications* **48**, 8326-8328, doi:10.1039/c2cc33856d (2012).
- 29 Horvat, S. M. & Schiesser, C. H. An ab initio and DFT study of homolytic substitution reactions of acyl radicals at sulfur, selenium, and tellurium. *New Journal of Chemistry* **34**, 1692, doi:10.1039/c0nj00125b (2010).
- 30 Dénès, F., Schiesser, C. H. & Renaud, P. Thiols, thioethers, and related compounds as sources of C-centred radicals. *Chemical Society Reviews* **42**, 7900-7942, doi:10.1039/c3cs60143a (2013).
- 31 Gutierrez, C. G. & Summerhays, L. R. Organotin-mediated selective desulfurization: tri-n-butyltin hydride reduction of unsymmetric sulfides. *The Journal of Organic Chemistry* **49**, 5206-5213, doi:10.1021/jo00200a039 (1984).
- 32 Schiesser, C. H. Taming the free radical shrew – learning to control homolytic reactions at higher heteroatoms. *Chemical Communications* **0**, 4055-4065, doi:10.1039/b608150a (2006).
- 33 Kampmeier, J. A., Jordan, R. B., Liu, M. S., Yamanaka, H. & Bishop, D. J. ACS Symposium Series. *ACS Symposium Series*, 275-289, doi:10.1021/bk-1978-0069.ch016 (1978).
- 34 Beckwith, A. L. J. & Boate, D. R. Stereochemistry of intramolecular homolytic substitution at the sulphur atom of a chiral sulphoxide. *Journal of the Chemical Society, Chemical Communications* **0**, 189-190, doi:10.1039/c39860000189 (1986).
- 35 Bondi, A. van der Waals Volumes and Radii. *The Journal of Physical Chemistry* **68**, 441-451, doi:10.1021/j100785a001 (1964).

- 36 Broderick, J. B., Duffus, B. R., Duschene, K. S. & Shepard, E. M. Radical S-Adenosylmethionine Enzymes. *Chemical Reviews* **114**, 4229-4317, doi:10.1021/cr4004709 (2014).
- 37 Wu, W. *et al.* Lysine 2,3-Aminomutase and trans-4,5-Dehydrolysine: Characterization of an Allylic Analogue of a Substrate-Based Radical in the Catalytic Mechanism†. *Biochemistry* **39**, doi:10.1021/bi000658p (2000).
- 38 Mulliez, E., Padovani, D., Atta, M., Alcouffe, C. & Fontecave, M. Activation of class III ribonucleotide reductase by flavodoxin: a protein radical-driven electron transfer to the iron-sulfur center. *Biochemistry* **40**, 3730-3736 (2001).
- 39 Frey, M., Rothe, M., Wagner, A. F. & Knappe, J. Adenosylmethionine-dependent synthesis of the glycy radical in pyruvate formate-lyase by abstraction of the glycine C-2 pro-S hydrogen atom. Studies of [2H]glycine-substituted enzyme and peptides homologous to the glycine 734 site. *The Journal of biological chemistry* **269**, 12432-12437 (1994).
- 40 Wang, S. C. & Frey, P. A. Binding energy in the one-electron reductive cleavage of S-adenosylmethionine in lysine 2,3-aminomutase, a radical SAM enzyme. *Biochemistry* **46**, 12889-12895, doi:10.1021/bi701745h (2007).
- 41 Padwa\*, A., Beall†, S. L., Heidelbaugh, T. M., Liu, B. & Sheehan, S. M. A One-Pot Bicycloannulation Method for the Synthesis of Tetrahydroisoquinoline Systems. *The Journal of Organic Chemistry* **65**, 2684-2695, doi:10.1021/jo991742h (2000).
- 42 Grell, T. A. J., Goldman, P. J. & Drennan, C. L. SPASM and Twitch Domains in AdoMet Radical Enzyme Structures. *Journal of Biological Chemistry*, doi:10.1074/jbc.R114.581249 (2014).
- 43 Ruszczycky, M. W., Ogasawara, Y. & Liu, H.-W. W. Radical SAM enzymes in the biosynthesis of sugar-containing natural products. *Biochimica et biophysica acta* **1824**, 1231-1244, doi:10.1016/j.bbapap.2011.11.006 (2012).
- 44 Yokoyama, K., Ohmori, D., Kudo, F. & Eguchi, T. Mechanistic study on the reaction of a radical SAM dehydrogenase BtrN by electron paramagnetic resonance spectroscopy. *Biochemistry* **47**, 8950-8960, doi:10.1021/bi800509x (2008).
- 45 Ruszczycky, M. W., Choi, S.-H. H. & Liu, H.-W. W. Stoichiometry of the redox neutral deamination and oxidative dehydrogenation reactions catalyzed by the radical SAM enzyme DesII. *Journal of the American Chemical Society* **132**, 2359-2369, doi:10.1021/ja909451a (2010).
- 46 Kampmeier, J. A. Regioselectivity in the Homolytic Cleavage of S-Adenosylmethionine. *Biochemistry* **49**, 10770-10772, doi:10.1021/bi101509u (2010).
- 47 Hutcheson, R. U. & Broderick, J. B. Radical SAM enzymes in methylation and methylthiolation. *Metallomics* **4**, doi:10.1039/c2mt20136d (2012).
- 48 Zhang, Y. *et al.* Diphthamide biosynthesis requires an organic radical generated by an iron-sulphur enzyme. *Nature* **465**, 891-896, doi:10.1038/nature09138 (2010).
- 49 Demick, J. M. & Lanzilotta, W. N. Radical SAM Activation of the B12-Independent Glycerol Dehydratase Results in Formation of 5' -Deoxy-5' -(methylthio)adenosine and Not 5' -Deoxyadenosine. *Biochemistry* **50**, 440-442, doi:10.1021/bi101255e (2011).
- 50 Raynaud, C., Sarçabal, P., Meynial-Salles, I., Croux, C. & Soucaille, P. Molecular characterization of the 1,3-propanediol (1,3-PD) operon of *Clostridium butyricum*. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 5010-5015, doi:10.1073/pnas.0734105100 (2003).
- 51 Fenwick, M. K. *et al.* Non-canonical active site architecture of the radical SAM thiamin pyrimidine synthase. *Nature Communications* **6**, 6480, doi:10.1038/ncomms7480 (2015).
- 52 Dey, A. *et al.* S K-edge XAS and DFT Calculations on SAM Dependent Pyruvate Formate-Lyase Activating Enzyme: Nature of Interaction between the Fe<sub>4</sub>S<sub>4</sub> Cluster and SAM and its Role in Reactivity. *Journal of the American Chemical Society* **133**, 18656-18662, doi:10.1021/ja203780t (2011).

- 53 Horitani, M. *et al.* Why Nature Uses Radical SAM Enzymes so Widely: Electron Nuclear Double Resonance Studies of Lysine 2,3-Aminomutase Show the 5' -dAdo• “Free Radical” Is Never Free. *Journal of the American Chemical Society*, 2022414680, doi:10.1021/jacs.5b00498 (2015).
- 54 Magnusson, O., Reed, G. H. & Frey, P. A. Spectroscopic Evidence for the Participation of an Allylic Analogue of the 5'-Deoxyadenosyl Radical in the Reaction of Lysine 2,3-Aminomutase. *Journal of the American Chemical Society* **121**, 9764-9765, doi:10.1021/ja9925507 (1999).
- 55 Magnusson, O. T., Reed, G. H. & Frey, P. A. Characterization of an allylic analogue of the 5'-deoxyadenosyl radical: an intermediate in the reaction of lysine 2,3-aminomutase. *Biochemistry* **40**, 7773-7782 (2001).
- 56 Lepore, B. W., Ruzicka, F. J., Frey, P. A. & Ringe, D. The x-ray crystal structure of lysine-2,3-aminomutase from *Clostridium subterminale*. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 13819-13824, doi:10.1073/pnas.0505726102 (2005).
- 57 Lees, N. S. *et al.* How an enzyme tames reactive intermediates: positioning of the active-site components of lysine 2, 3-aminomutase during enzymatic turnover as determined by ENDOR spectroscopy. *Journal of the American Chemical Society* **128**, 10145-10154, doi:10.1021/ja061282r (2006).
- 58 Vey, J. L. & Drennan, C. L. Structural Insights into Radical Generation by the Radical SAM Superfamily. *Chemical Reviews* **111**, 2487-2506, doi:10.1021/cr9002616 (2011).
- 59 Heller, C. & Cole, T. Electron Magnetic Resonance of an Oriented Allyl-Type Free Radical. *The Journal of Chemical Physics* **37**, 243, doi:10.1063/1.1701312 (1962).
- 60 Nicolet, Y. & Drennan, C. L. AdoMet radical proteins—from structure to evolution—alignment of divergent protein sequences reveals strong secondary structure element conservation. *Nucleic Acids Research* **32**, 4015-4025, doi:10.1093/nar/gkh728 (2004).
- 61 Bucher, D., Sandala, G. M., Durbeej, B., Radom, L. & Smith, D. M. The elusive 5'-deoxyadenosyl radical in coenzyme-B12-mediated reactions. *Journal of the American Chemical Society* **134**, 1591-1599, doi:10.1021/ja207809b (2012).
- 62 Berkovitch, F., Nicolet, Y., Wan, J. T., Jarrett, J. T. & Drennan, C. L. Crystal Structure of Biotin Synthase, an S-Adenosylmethionine-Dependent Radical Enzyme. *Science* **303**, 76-79, doi:10.1126/science.1088493 (2004).
- 63 Goldman, P. J., Grove, T. L., Booker, S. J. & Drennan, C. L. X-ray analysis of butirosin biosynthetic enzyme BtrN redefines structural motifs for AdoMet radical chemistry. *Proceedings of the National Academy of Sciences* **110**, 15949-15954, doi:10.1073/pnas.1312228110 (2013).
- 64 Dowling, D. P. *et al.* Radical SAM enzyme QueE defines a new minimal core fold and metal-dependent mechanism. *Nature Chemical Biology*, doi:10.1038/nchembio.1426 (2014).
- 65 Benjdia, A., Heil, K., Barends, T. R. M., Carell, T. & Schlichting, I. Structural insights into recognition and repair of UV-DNA damage by Spore Photoproduct Lyase, a radical SAM enzyme. *Nucleic Acids Research* **40**, 9308-9318, doi:10.1093/nar/gks603 (2012).
- 66 Vey, J. L., Yang, J., Li, M. & Broderick, W. E. Structural basis for glycy radical formation by pyruvate formate-lyase activating enzyme. *Proceedings of the ...*, doi:10.1073/pnas.0806640105 (2008).
- 67 Goldman, P. J. *et al.* X-ray structure of an AdoMet radical activase reveals an anaerobic solution for formylglycine posttranslational modification. *Proceedings of the National Academy of Sciences* **110**, 8519-8524, doi:10.1073/pnas.1302417110 (2013).
- 68 Hänzelmann, P. & Schindelin, H. Binding of 5'-GTP to the C-terminal FeS cluster of the radical S-adenosylmethionine enzyme MoeA provides insights into its mechanism. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 6829-6834, doi:10.1073/pnas.0510711103 (2006).
- 69 Qwitterer, F., List, A., Eisenreich, W., Bacher, A. & Groll, M. Crystal Structure of Methylornithine Synthase (PylB): Insights into the Pyrrolysine Biosynthesis. *Angewandte Chemie International Edition* **51**, 1339-1342, doi:10.1002/anie.201106765 (2012).

- 70 Harmer, J. E. *et al.* Structures of lipoyl synthase reveal a compact active site for controlling sequential sulfur insertion reactions. *The Biochemical journal* **464**, 123-133, doi:10.1042/bj20140895 (2014).
- 71 Forouhar, F. *et al.* Two Fe-S clusters catalyze sulfur insertion by radical-SAM methylthiotransferases. *Nature Chemical Biology* **9**, 333-338, doi:10.1038/nchembio.1229 (2013).
- 72 Dinis, P. *et al.* X-ray crystallographic and EPR spectroscopic analysis of HydG, a maturase in [FeFe]-hydrogenase H-cluster assembly. *Proceedings of the National Academy of Sciences* **112**, 1362-1367, doi:10.1073/pnas.1417252112 (2015).
- 73 Layer, G., Moser, J., Heinz, D. W., Jahn, D. & Schubert, W.-D. D. Crystal structure of coproporphyrinogen III oxidase reveals cofactor geometry of Radical SAM enzymes. *The EMBO journal* **22**, 6214-6224, doi:10.1093/emboj/cdg598 (2003).
- 74 Lees, N. S. *et al.* ENDOR spectroscopy shows that guanine N1 binds to [4Fe-4S] cluster II of the S-adenosylmethionine-dependent enzyme MoaA: mechanistic implications. *Journal of the American Chemical Society* **131**, 9184-9185, doi:10.1021/ja903978u (2009).
- 75 Frey, P. A. Travels with Carbon-Centered Radicals. 5' -Deoxyadenosine and 5' -Deoxyadenosine-5' -yl in Radical Enzymology. *Accounts of Chemical Research*, doi:10.1021/ar400194k (2014).
- 76 LoBrutto, R. *et al.* 5 -Deoxyadenosine Contacts the Substrate Radical Intermediate in the Active Site of Ethanolamine Ammonia-lyase: 2H and 13C Electron Nuclear Double Resonance Studies. *Biochemistry* **40**, doi:10.1021/bi001865s (2001).
- 77 Bender, G., Poyner, R. R. & Reed, G. H. Identification of the substrate radical intermediate derived from ethanolamine during catalysis by ethanolamine ammonia-lyase. *Biochemistry* **47**, 11360-11366, doi:10.1021/bi801316v (2008).
- 78 Canfield, J. M. & Warncke, K. Geometry of Reactant Centers in the CoII-Substrate Radical Pair State of Coenzyme B12-Dependent Ethanolamine Deaminase Determined by Using Orientation-Selection-ESEEM Spectroscopy. *The Journal of Physical Chemistry B* **106**, 8831-8841, doi:10.1021/jp0207634 (2002).
- 79 Maity, A. N. *et al.* Evidence for Conformational Movement and Radical Mechanism in the Reaction of 4-Thia-l-lysine with Lysine 5,6-Aminomutase. *The Journal of Physical Chemistry B* **113**, 12161-12163, doi:10.1021/jp905357a (2009).
- 80 Chen, D., Abend, A., Stubbe, J. & Frey, P. A. Epimerization at carbon-5' of (5'R)-[5'-2H]adenosylcobalamin by ribonucleoside triphosphate reductase: cysteine 408-independent cleavage of the Co-C5' bond. *Biochemistry* **42**, 4578-4584, doi:10.1021/bi030018x (2003).
- 81 Gaudemer, A. *et al.* Reversible cleavage of the cobalt-carbon bond to coenzyme B12 catalysed by methylmalonyl-CoA mutase from *Propionibacterium shermanii*. The use of coenzyme B12 stereospecifically deuterated in position 5'. *European journal of biochemistry / FEBS* **119**, 279-285 (1981).
- 82 Buckel, W., Friedrich, P. & Golding, B. T. Hydrogen Bonds Guide the Short - Lived 5' -Deoxyadenosyl Radical to the Place of Action. *Angewandte Chemie International Edition*, doi:10.1002/anie.201205299 (2012).
- 83 Reitzer, R. *et al.* Glutamate mutase from *Clostridium cochlearium*: the structure of a coenzyme B12-dependent enzyme provides new mechanistic insights. *Structure* **7**, 891-902, doi:10.1016/s0969-2126(99)80116-6 (1999).
- 84 Toraya, T. Radical catalysis in coenzyme B12-dependent isomerization (eliminating) reactions. *Chemical Reviews* **103**, 2095-2127, doi:10.1021/cr020428b (2003).
- 85 Jones, A. R., Rentergent, J., Scrutton, N. S. & Hay, S. Probing reversible chemistry in coenzyme B12 - dependent ethanolamine ammonia lyase with kinetic isotope effects. *Chemistry (Weinheim an der Bergstrasse, Germany)* **21**, 8826-8831, doi:10.1002/chem.201500958 (2015).
- 86 Crich, D. Homolytic Substitution at the Sulfur Atom as a Tool for Organic Synthesis. *Helvetica Chimica Acta* **89**, 2167-2182, doi:10.1002/hlca.200690204 (2006).

## Chapitre IV : Etude structurale d'interconversion entre différentes formes de centres [Fe-S]

Les centres [Fe-S], du fait de leur capacité à accepter un donneur des électrons sont habituellement sensibles à l'attaque par le dioxygène et subissent des dégradations oxydatives qui peuvent conduire jusqu'à leur destruction totale. Ils peuvent donc être utilisés par la cellule comme des capteurs qui détectent des variations de la concentration en O<sub>2</sub>. Un exemple bien étudié est la protéine FNR (*fumarate nitrate reductase regulator*) qui régule le passage entre le métabolisme anaérobie vers le métabolisme aérobie et vice-versa<sup>1,2</sup>. Quand la concentration de O<sub>2</sub> est inférieure à 0,5 % la protéine FNR, qui contient un centre [Fe<sub>4</sub>S<sub>4</sub>], forme un dimère qui interagit avec l'ADN et influence ainsi l'expression d'un grand nombre de gènes. Quand la concentration de O<sub>2</sub> commence à monter, l'agrégat [Fe<sub>4</sub>S<sub>4</sub>] se dégrade en [Fe<sub>2</sub>S<sub>2</sub>], le dimère de FNR se dissocie, ce qui empêche son interaction avec l'ADN<sup>3</sup>. L'exact mécanisme de la dégradation de son centre [Fe<sub>4</sub>S<sub>4</sub>] ainsi que les changements structuraux qui en découlent faisaient débat au moment quand j'effectuais mes travaux de thèse. Afin d'apporter des éléments à ce sujet nous avons décidé d'étudier les phénomènes de dégradation et de synthèse d'un centre [Fe-S] en prenant comme modèle la protéine HydE de *Thermotoga maritima*.

La protéine *TmHydE* contient, à part le centre [Fe<sub>4</sub>S<sub>4</sub>] 'SAM radical', un centre [Fe-S] auxiliaire. Ce second agrégat se lie sur les trois cystéines Cys<sup>311</sup>, Cys<sup>319</sup> et Cys<sup>322</sup>, qui appartiennent au motif semi-conservé CX<sub>7</sub>CX<sub>2</sub>C [voir IB-3.2 & Annexes]. Ce site de fixation se situe à la surface de la protéine [Figure-IB-12 & Figure-I-17]. Le centre [Fe-S] y est exposée au solvant et peut donc facilement subir des attaques oxydantes. Nous avons utilisé cette propriété pour effectuer une étude cristallographique de l'interconversion entre différentes formes de cet agrégat [Fe-S] auxiliaire. De nombreux essais de cristallisation de cette protéine dans de diverses conditions oxydoréductrices, avec ou sans une source de fer et/ou de soufre, nous ont fourni toute une gamme de structures, dans laquelle certaines formes de ce second centre [Fe-S] correspondaient à de différentes étapes d'assemblage ou de dégradation d'un centre [Fe<sub>4</sub>S<sub>4</sub>]. Dans nos structures, le second centre [Fe-S] existe majoritairement sous la forme d'un [Fe<sub>2</sub>S<sub>2</sub>] [code PDB : 3CIX] ou d'un [Fe<sub>2</sub>S<sub>2</sub>] persulfuré [code PDB : 3IIZ]. Un agrégat [Fe<sub>3</sub>S<sub>3</sub>] [code PDB : 4JYE] avait été observé suite au trempage des cristaux contenant un centre [Fe<sub>2</sub>S<sub>2</sub>] persulfuré dans une solution avec du DTT et Fe<sup>3+</sup>. La dernière structure qui est venue compléter le bestiaire de ces différentes formes a été obtenue au cours de ma thèse et elle contient un agrégat auxiliaire [Fe<sub>4</sub>S<sub>4</sub>] [code PDB : 4JY8]. Nous avons donc pu caractériser des étapes cruciales de l'assemblage et de la dégradation de ce centre [Fe-S] dans les cristaux de *TmHydE*. Par exemple l'observation d'un centre [Fe<sub>2</sub>S<sub>2</sub>] avec des persulfures sur les cystéines Cys<sup>319</sup> et Cys<sup>322</sup> a confirmé l'hypothèse émise par Zhang et ses collaborateurs<sup>4</sup>, que l'oxydation de l'agrégat [Fe<sub>4</sub>S<sub>4</sub>] de la FNR par O<sub>2</sub> passe, dans un premier temps, par la libération d'ions de Fe, tandis que les atomes de soufre se retrouvent sous la forme de persulfures. De cette façon, le soufre pourrait être recyclé sur place dans un processus de la réparation du centre [Fe<sub>4</sub>S<sub>4</sub>], ce qui constituerait une économie pour la cellule.

Le recyclage des atomes de soufre est bien illustré par la formation de l'agrégat  $[\text{Fe}_3\text{S}_3]$  à partir d'un centre  $[\text{Fe}_2\text{S}_2]$  persulfuré après que ce dernier a été exposé au DTT et au  $\text{Fe}^{3+}$ .

Notre étude a également révélé d'importants changements structuraux qui accompagnaient la conversion de la forme  $[\text{Fe}_4\text{S}_4]$  en la forme  $[\text{Fe}_2\text{S}_2]$  avec (ou sans) des persulfures. De tels changements de conformation peuvent faire partie d'un mécanisme de régulation de l'activité de certaines protéines et de transmission de signal et être, par exemple, e à l'origine de la transition dimère - monomère dans la protéine FNR<sup>5</sup>.

Les résultats de cette étude ont fait l'objet d'une publication dans le journal PNAS<sup>6</sup>. L'article en question se trouve en Annexe de ce manuscrit. La contribution de mes travaux à ce projet, qui a été entamé bien avant mon arrivée au laboratoire, était la résolution de structures de *TmHydE* possédant des centres  $[\text{Fe-Se}]$  [Figure-I-15-B], mais surtout la caractérisation structurale de la forme de *TmHydE* abritant deux centres  $[\text{Fe}_4\text{S}_4]$ . Cette dernière est brièvement commentée ci-dessous.

#### IV-1 Structure de *TmHydE* avec deux centres $[\text{Fe}_4\text{S}_4]$

La protéine *TmHydE* sauvage reconstituée chimiquement soit dans l'extrait brut au début de la purification soit une fois pure contient 8 atomes de Fe et 8 atomes de soufre<sup>7,8</sup>. Ceci correspond à un agrégat  $[\text{Fe}_4\text{S}_4]$  'SAM radical', responsable du clivage de la SAM, et un agrégat  $[\text{Fe}_4\text{S}_4]$  auxiliaire. Cette forme de la protéine, qui a longtemps résisté à nos essais de cristallisation, a finalement pu être cristallisée dans un nouveau groupe d'espace ( $P6_522$ ) en présence de NaCl. La structure montre un important déplacement de la boucle 299-317, qui connecte le brin  $\beta 8$  à l'hélice  $\alpha 8$  et qui, dans toutes les autres structures de *HydE*, couvre la partie supérieure du tonneau  $(\beta\alpha)_8$  [Figure-IV-103-A]. Dans la nouvelle structure, cette boucle effectue un mouvement important en direction du site de fixation de l'agrégat auxiliaire [Figure-IV-103-C]. Le mouvement atteint une amplitude maximale de 13,6 Å au niveau de  $C_\alpha$  du résidu Leu<sup>305</sup>. De ce fait, la cavité centrale de *HydE* se retrouve complètement exposée au solvant. Certains résidus strictement conservés situés sur cette boucle, comme les deux tyrosines Tyr<sup>303</sup> & Tyr<sup>306</sup>, qui, avant, tapissait la face intérieure de la cavité, se retrouvent externalisés [Figure-IV-103-B]. Dans cette configuration le groupement OH de la Tyr<sup>306</sup> fixe le quatrième Fe de l'agrégat auxiliaire  $[\text{Fe}_4\text{S}_4]$ . Il est connu que les chaînes latérales de résidus aspartate, glutamate ou serine peuvent être ligands du Fe unique<sup>9-11</sup>, mais l'interaction avec une tyrosine est sans précédent. Le déplacement de la boucle 299-317 affecte également la position de la Cys<sup>311</sup>; son atome de carbone  $C_\alpha$  subit une translation de plus de 6 Å. La comparaison des structures avec différentes formes d'agrégat  $[\text{Fe-S}]$  auxiliaire montre qu'une des trois cystéines impliquées dans sa fixation ne peut pas occuper le même espace en présence d'un centre  $[\text{Fe}_4\text{S}_4]$  qu'elle occupe en présence d'un centre  $[\text{Fe}_2\text{S}_2]$  [Figure-IV-103-D]. L'interconversion entre ces deux formes s'accompagne donc obligatoirement d'un réarrangement conformationnel important. Ce résultat illustre comment la transition entre un centre  $[\text{Fe}_4\text{S}_4]$  cubane vers un centre  $[\text{Fe}_2\text{S}_2]$  rhombique induit des changements de conformation pouvant affecter le repliement d'une vaste zone à la surface d'une protéine, ce qui dans le cas de la FNR, conduit jusqu'à la monomérisation de la protéine.

**Figure-IV-103** : Etude structurale des changements conformationnels induits par la conversion du centre [Fe<sub>2</sub>S<sub>2</sub>] auxiliaire en centre [Fe<sub>4</sub>S<sub>4</sub>]

Repliement global de la protéine *TmHydE* en présence d'un centre auxiliaire sous la forme [Fe<sub>2</sub>S<sub>2</sub>] en (A) et d'un centre auxiliaire sous la forme [Fe<sub>4</sub>S<sub>4</sub>] en (B)

oxydants réagiraient préférentiellement avec l'agrégat exposé au solvant avant de s'attaquer au centre  $[\text{Fe}_4\text{S}_4]$  'SAM radical', qui lui est essentiel à l'activité de l'enzyme.

Alternativement nous pouvons imaginer que le centre  $[\text{Fe-S}]$  auxiliaire pourrait servir pour fixer un éventuel substrat ou son précurseur et augmenter de cette façon sa concentration à proximité du site actif dans le but d'accélérer la synthèse du produit. En effet, sous sa forme rhombique, un de ces atomes de Fe possède un ligand OH exposé au solvant, qui pourrait facilement s'échanger avec un groupement du substrat. La comparaison de séquences de protéines HydE montre une corrélation entre la présence des trois résidus cystéines impliqués dans la fixation de l'agrégat auxiliaire et la présence de certains résidus à proximité de ce site. Le plus marquant est le résidu Asp/Glu<sup>314</sup>, qui est strictement conservé dans les séquences possédant le motif CX<sub>7</sub>CX<sub>2</sub>C. Il est possible que ce résidu soit impliqué dans l'assemblage de l'agrégat auxiliaire ou même dans la fixation d'un substrat sur ce dernier.

### **IV-3 Perspectives**

#### **IV-3.1 Etude d'interconversion des centres $[\text{Fe-S}]$ en solution**

Nous avons envisagé d'enregistrer les spectres Raman résonant pour chaque cristal de *TmHydE* contenant une forme particulière de l'agrégat  $[\text{Fe-S}]$  auxiliaire. En faisant la soustraction du signal dû au centre  $[\text{Fe}_4\text{S}_4]$  'SAM radical' il devrait être possible d'attribuer une signature Raman résonant à chaque intermédiaire de la synthèse ou le dégradation des centres  $[\text{Fe-S}]$ . Il serait alors possible d'effectuer une étude similaire à la notre, mais cette fois-ci en solution, et relier directement le spectre Raman résonant de l'échantillon à la structure du centre  $[\text{Fe-S}]$  qu'il contient. Ceci devrait permettre le suivi en temps réel d'interconversion entre différentes formes d'agrégats  $[\text{Fe-S}]$  en solution pour toutes les protéines à  $[\text{Fe-S}]$  de structure cristallographique inconnue.

#### **IV-3.2 Centres $[\text{Fe-Se}]$ et le phasage de données cristallographiques**

Les expériences de reconstitution des agrégats  $[\text{Fe-S}]$  avec du sélénium ont prouvé que cette technique peut être utilisée pour la production de cristaux de protéines à  $[\text{Fe-S}]$  dérivatisés au sélénium [voir paragraphe MM-3.3.5]. Les avantages d'utiliser un centre  $[\text{Fe-Se}]$  plutôt qu'un centre  $[\text{Fe-S}]$  pour le phasage expérimental de jeux de données cristallographiques sont nombreuses. Le sélénium possède un signal anomal plus élevé que celui du Fe à cause de sa raie blanche. Les lignes de lumière aux synchrotrons ont souvent un flux photonique maximal à l'énergie correspondante au seuil K du sélénium. L'enregistrement des données au seuil K du Fe augmente le bruit de fond dû à la diffusion des rayons X par des atomes légers, car cette dernière augmente avec la longueur d'onde. De plus, dans le cas des cristaux ayant un pouvoir diffractant élevé, des pertes de données de haute résolution peuvent se produire à cause de la géométrie du détecteur. En faisant la collecte au seuil du sélénium, il est toujours possible d'utiliser la contribution résiduelle du Fe lors du phasage. Pour finir, dans le cas où la résolution de données était faible il devrait être possible de se servir de

l'agrégat [Fe-Se] comme d'un 'superatome'. Dans le cas des protéines à [Fe-S] cette méthode de dérivation pourrait compléter ou même remplacer la méthode de marquage par la L-séléno-méthionine, plus onéreuse.

#### IV-4 Références du Chapitre IV

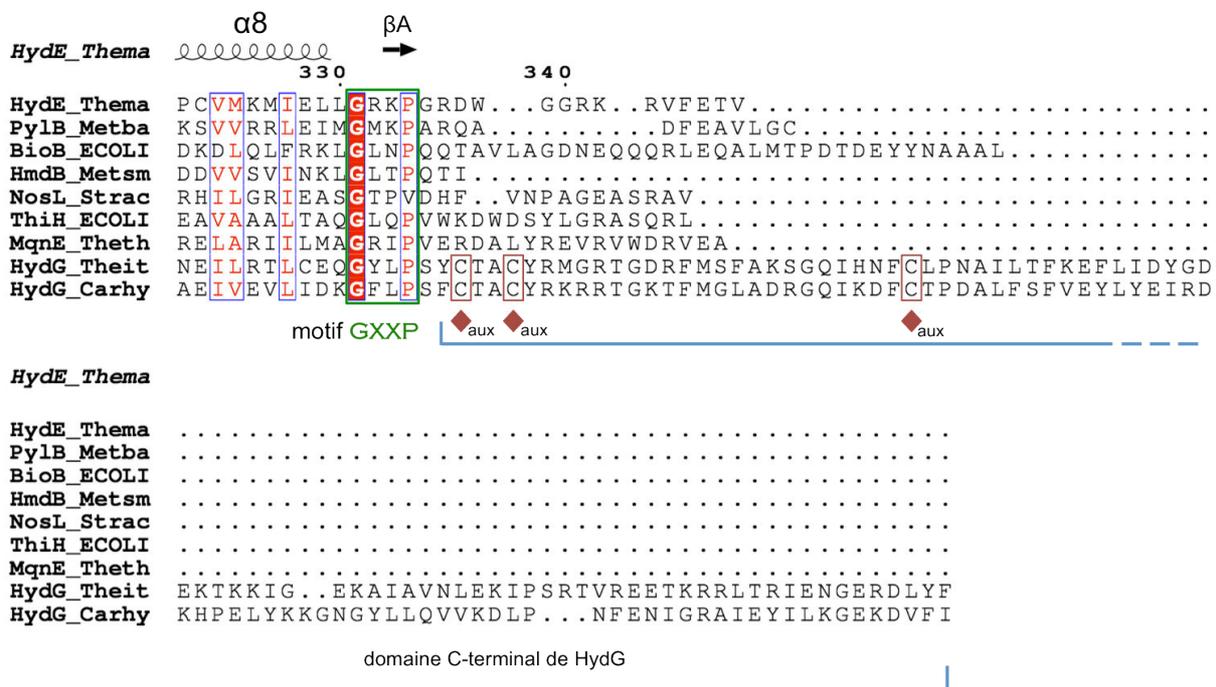
- 1 Crack, J. C., Green, J., Thomson, A. J. & Le Brun, N. E. Iron-sulfur cluster sensor-regulators. *Current Opinion in Chemical Biology* **16**, 35-44, doi:10.1016/j.cbpa.2012.02.009 (2012).
- 2 Fleischhacker, A. S. & Kiley, P. J. Iron-containing transcription factors and their roles as sensors. *Current Opinion in Chemical Biology* **15**, 335-341, doi:10.1016/j.cbpa.2011.01.006 (2011).
- 3 Alexeeva, S. & Hellingwerf, K. J. Quantitative assessment of oxygen availability: perceived aerobiosis and its effect on flux distribution in the respiratory chain of *Escherichia coli*. *Journal of Bacteriology*, doi:10.1128/jb.184.5.1402-1406.2002 (2002).
- 4 Zhang, B. *et al.* Reversible cycling between cysteine persulfide-ligated [2Fe-2S] and cysteine-ligated [4Fe-4S] clusters in the FNR regulatory protein. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 15734-15739, doi:10.1073/pnas.1208787109 (2012).
- 5 Lazazzera, B. A., Beinert, H. & Khoroshilova, N. DNA binding and dimerization of the Fe- S-containing FNR protein from *Escherichia coli* are regulated by oxygen. *Journal of Biological ...* (1996).
- 6 Nicolet, Y., Rohac, R., Martin, L. & Fontecilla-Camps, J. C. X-ray snapshots of possible intermediates in the time course of synthesis and degradation of protein-bound Fe<sub>4</sub>S<sub>4</sub> clusters. *Proceedings of the National Academy of Sciences* **110**, 7188-7192, doi:10.1073/pnas.1302388110 (2013).
- 7 Rubach, J. K., Brazzolotto, X., Gaillard, J. & Fontecave, M. Biochemical characterization of the HydE and HydG iron-only hydrogenase maturation enzymes from *Thermatoga maritima*. *FEBS Letters* **579**, 5055-5060, doi:10.1016/j.febslet.2005.07.092 (2005).
- 8 Betz, J. N. *et al.* [FeFe]-Hydrogenase Maturation: Insights into the Role HydE Plays in Dithiomethylamine Biosynthesis. *Biochemistry*, 1574462297, doi:10.1021/bi501205e (2015).
- 9 Calzolari, L. *et al.* 1H NMR Investigation of the Electronic and Molecular Structure of the Four-Iron Cluster Ferredoxin from the Hyperthermophile *Pyrococcus furiosus*. Identification of Asp 14 as a Cluster Ligand in Each of the Four Redox States. *Biochemistry* **34**, 11373-11384, doi:10.1021/bi00036a010 (1995).
- 10 Brereton, P. S., Duderstadt, R. E., Staples, C. R., Johnson, M. K. & Adams, M. W. Effect of serinate ligation at each of the iron sites of the [Fe<sub>4</sub>S<sub>4</sub>] cluster of *Pyrococcus furiosus* ferredoxin on the redox, spectroscopic, and biological properties. *Biochemistry* **38**, 10594-10605, doi:10.1021/bi990671d (1999).
- 11 Johnson, D. C., Dean, D. R., Smith, A. D. & Johnson, M. K. Structure, function, and formation of biological iron-sulfur clusters. *Annual review of biochemistry* **74**, 247-281, doi:10.1146/annurev.biochem.74.082803.133518 (2005).



# Chapitre V : Vers la sortie du produit de HydE

## V-1 Introduction

Comme nous l'avons vu dans l'Introduction Bibliographique, les membres de la superfamilles de protéines à radical SAM adoptent un repliement en tonneau  $(\beta\alpha)_6$  partiel ou  $(\beta\alpha)_8$  complet. Les motifs caractéristiques des cette superfamille sont un motif CX<sub>3</sub>CX $\phi$ C responsable de la fixation de l'agrégat [Fe<sub>4</sub>S<sub>4</sub>], essentiel à l'initiation de la réaction radicalaire, et puis des motifs impliqués de la fixation du cofacteur SAM. Il a été proposée que les protéines avec un repliement de type tonneau  $(\beta\alpha)_6$  ou  $(\beta\alpha)_8$  qui partagent la même structure tridimensionnelle possèdent des résidus glycines et prolines conservés sur certaines positions<sup>1</sup>. Ces résidus marquent les transitions entre les différentes structures secondaires et verrouillent ainsi le repliement global de la protéine. Certaines protéines à radical SAM qui partagent un repliement en tonneau  $(\beta\alpha)_8$  complet possèdent une motif GXXP qui se trouve juste après l'hélice  $\alpha_8$  [Figure-V-104]. Ce motif sépare donc le corps en tonneau  $(\beta\alpha)_8$  d'un éventuelle extension ou domaine C-terminale. Le rôle exact de ce motif semi-conservé n'est pas clair, mais il a été proposé qu'il puisse servir de charnière pour permettre un mouvement du domaine C-terminale relatif au tonneau  $(\beta\alpha)_8$ . Ce mini-chapitre contient une étude cristallographique d'une structure de la protéine *TmHydE* native dont le repliement particulier pourrait apporter des éléments sur ce motif GXXP.

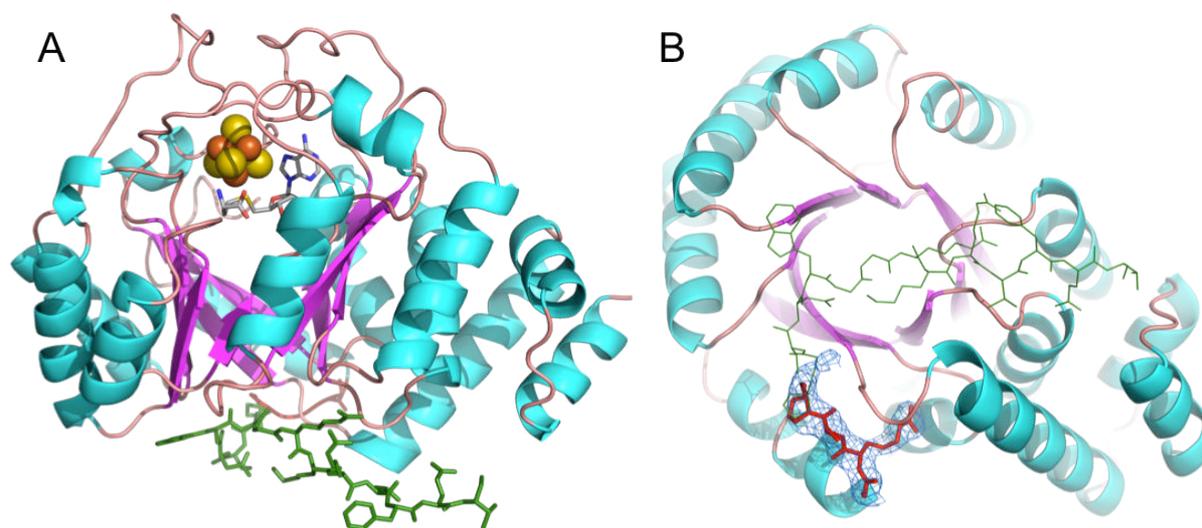


**Figure-V-104** : Détails d'un alignement de séquences des protéines à radical SAM qui possèdent un repliement en tonneau  $(\beta\alpha)_8$  complet

Les éléments de la structure secondaire et la numérotation des résidus et sont celles de *TmHydE*. Les losanges bruns marquent la position des résidus cystéines du motif CX<sub>2</sub>CX<sub>22</sub>C, conservé dans HydG et impliqué dans la fixation du centre [Fe-S] auxiliaire. En vert le motif semi-conservé GXXP.

## V-2 Vers le transfert du produit hors de HydE

Bien que la cystine a résisté à toutes nos tentatives de prouver sa fixation à l'intérieur de HydE, lors d'un de nos nombreux tests de co-cristallisation de *TmHydE* native avec cette molécule nous avons obtenu une structure de l'enzyme présentant un important changement de conformation du brin- $\beta$  C-terminal. Ce brin- $\beta$  se situe juste après l'hélice  $\alpha 8$  et il est délimité du côté de cette dernière par un motif GXXP et de l'autre côté par le dernier résidu de la séquence [Figure-V-104]. Dans cette nouvelle structure ce brin disparaît entièrement des cartes de densité électronique, mais la structure globale de l'enzyme reste inchangée [Figure-V-105-A]. C'est la seule structure (parmi les centaines étudiées au cours de mes travaux de thèse) qui présente ce changement conformationnel.



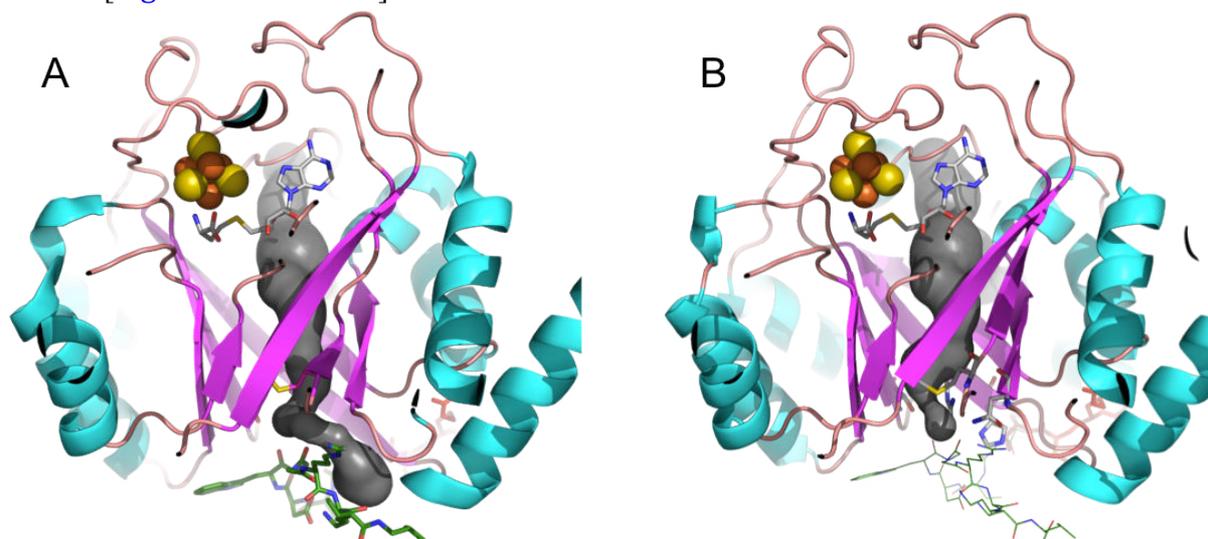
**Figure-V-105** : Vues orthoscopiques de la nouvelle structure avec un brin C-terminal désordonné

Le brin C-terminal en question est coloré en vert et le motif GXXP en rouge. Le reste de la protéine est coloré en fonction de la structure secondaire. B - En bleu, carte de densité électronique  $2Fo-Fc$  contournée à  $1\sigma$  et visualisée à  $3\text{ \AA}$  autour du motif GXXP (résolution  $2,58\text{ \AA}$ ).

L'inspection de la structure générale montre que le repliement de la protéine reste globalement inchangé ( $rmsd = 0.316\text{ \AA}$  pour 303  $C\alpha$  superposés sur 346  $C\alpha$  possibles). La protéine cristallise dans le même groupe d'espace  $P2_12_12_1$  que la structure de référence. Les paramètres de la maille ne sont que peu affectés par le désordre du côté du brin C-terminal. Le changement du volume de la maille est de l'ordre de 2%. Les données cristallographiques présentent une résolution relativement basse de  $2,58\text{ \AA}$  sans doute à cause d'un léger réarrangement de la chaîne principale dans la zone de la fixation du second agrégat FeS, entre les résidus 307 et 323. Le changement de conformation de cette région peut s'expliquer par la présence de différents degrés de reconstitution de ce second centre [Fe-S] [I-8.8.2]. Le désordre du côté du brin C-terminal ne semble pas avoir d'impact sur l'empilement cristallin, car le repliement de la boucle 78-88 qui, dans la structure de référence interagit avec ce brin, n'est pas affecté par l'absence de cette interaction. D'ailleurs dans nombre de nos structures à haute résolution, les résidus du brin C-terminal sont partiellement désordonnés et présentent des facteurs d'agitation thermique élevés par rapport au reste de la structure.

## V-2.1 Exposition de la partie basse du tonneau au solvant

En examinant la structure plus en détail, nous observons que toute la chaîne principale se situant en aval de la Pro334, c'est-à-dire à partir de la Gly335, jusqu'au dernier résidu dans la séquence de HydE de *T. maritima*, la Val348, disparaît entièrement des cartes de densité électronique [Figure-V-105-B]. Le résidu Pro334 appartient au motif GXXP, situé juste après l'hélice  $\alpha 8$ . On peut trouver ce motif chez certaines protéines à radical SAM présentant un repliement en tonneau  $(\beta\alpha)_8$  complet. Il a été proposé que ce motif GXXP puisse servir de charnière pour exposer la partie basse du tonneau au solvant. Ceci devrait faciliter l'évacuation du produit en dehors de la cavité catalytique de l'enzyme<sup>2,3</sup>. Dans la structure native de *TmHydE* (pdb: 3CIW), le brin C-terminal recouvre de l'extérieur la partie basse du tonneau  $(\beta\alpha)_8$ . Il protège ainsi partiellement les résidus conservés Arg54, Thr103, Mrt224 et le site S3 de leur exposition au solvant. Quand le brin C-terminal est en place, il y a un canal long et étroit, le canal E<sub>2</sub>, qui relie le site S3 avec le solvant [voir Figure-IB-12]. En absence de ce brin, la longueur du canal est raccourcie, car les résidus His52, Arg54, Thr103 et Asn264, qui forment le fond de la cavité de HydE, se retrouvent complètement exposés au solvant [Figure-V-106-A&B].



**Figure-V-106** : Comparaison des longueurs du canal E2 entre la structure native (A) et la structure avec le brin C-terminal désordonné (B)

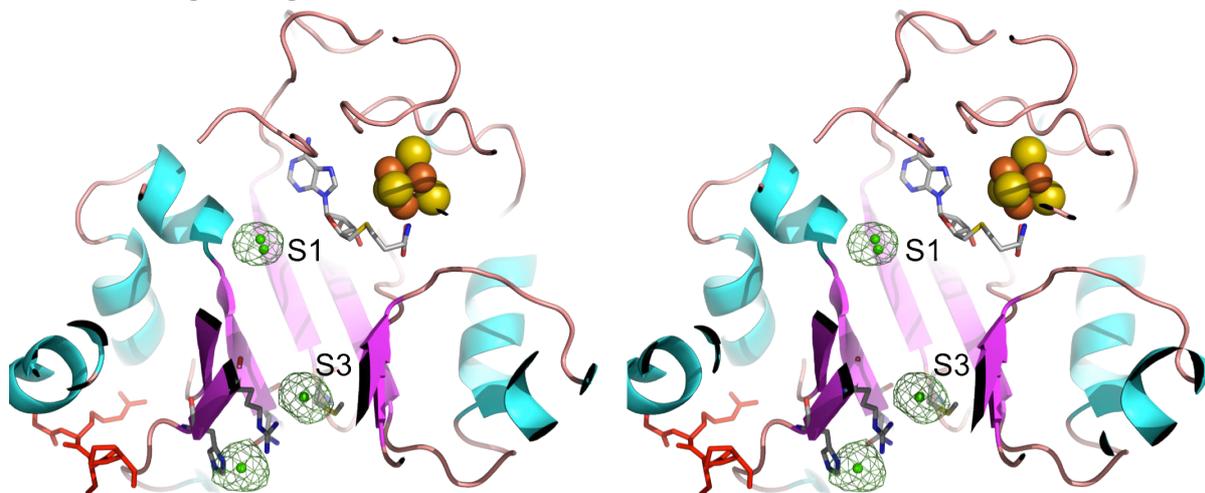
La représentation de la surface de la cavité a été générée par le logiciel Caver 3 dans PyMol en utilisant une sonde avec un diamètre de 0.9 Å. Dans la structure de référence (A) le brin C-terminal est représenté par des bâtons épais de couleur verte. Dans la nouvelle structure (B) ce brin est désordonné, il est donc visualisé par des bâtons fins.

Malheureusement la résolution limitée de données cristallographiques ne nous permet pas d'examiner le réarrangement des molécules d'eau dans la cavité de HydE ou dans de la première couche de solvatation, induit par le changement de conformation du brin C-terminal.

L'absence d'un éventuel clivage de la chaîne peptidique après le motif GXXP, qui pourrait expliquer la disparition du brin C-terminal des cartes de densité électronique a été vérifiée en déposant les cristaux qui ont donné la structure en question sur un gel SDS-PAGE. Par contre, nous n'avons pas réalisé de tests plus précis, comme par exemple une analyse de l'échantillon protéique par spectrométrie de masse.

## V-2.2 Externalisation du chlorure en bas du tonneau

Dans la nouvelle structure avec le brin- $\beta$  C-terminal désordonné, nous remarquons l'apparition d'une densité électronique sphérique dans la carte de Fourier différence en bas du tonneau, du côté extérieur. Cette densité a été modélisée comme un ion de chlorure coordonné par les chaînes latérales des résidus Arg54, His52 et Thr103. A l'intérieur de la cavité de HydE on ne trouve plus que deux ions chlorure : un qui occupe le site S1 en haut du tonneau et le second, modélisé avec un taux d'occupation de 60%, qui occupe le site S3 en bas du tonneau [Figure-V-107]. Ce dernier est coordonné, comme dans la structure de référence, par les résidus Arg54, Arg155 et Thr134.

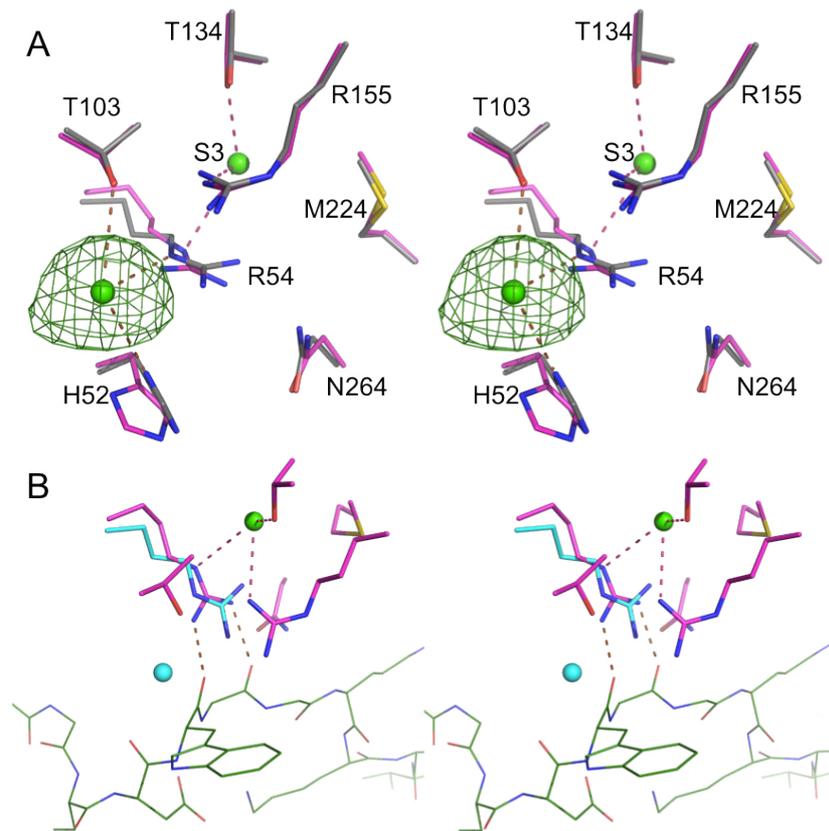


**Figure-V-107** : Vue stéréoscopique d'une coupe le long de l'axe de la cavité de la nouvelle structure

En vert foncé, carte de densité électronique  $F_o - F_c$  "omit" contournée à  $3\sigma$  (résolution 2,58 Å). Les ions chlorure fixés dans les sites S1, S3 et du côté extérieur de la partie basse du tonneau sont représentés par des sphères vertes.

Etrangement, la sphère de coordination du chlorure 'externe' est similaire à celle du chlorure qui occupe le site S3. Les deux ions partagent la chaîne latérale du résidu Arg54 et chacun possède un ligand hydroxyle de la thréonine 103 ou 134. De plus, nous remarquons que dans la structure où le brin- $\beta$  C-terminal n'est pas visible, la chaîne latérale de l'Arg54 pivote de  $180^\circ$  par rapport à sa conformation dans la structure de référence. De ce fait, l'atome d'hydrogène lié à l'azote  $N\epsilon$  de ce résidu pointe maintenant vers l'extérieur de la cavité [Figure-V-108-A]. Ceci diminue sans doute l'affinité du site S3 pour l'ion chlorure et en même temps stabilise la fixation d'un autre ion chlorure du côté extérieur de la cavité.

Il est intéressant de noter, que dans la structure native, ce site extérieur est très proche de la Gly339, un résidu qu'on retrouve souvent dans les séquences d'acides aminés des membres de la famille HydE<sup>4</sup> [voir l'alignement multiséquence en Annexe de ce manuscrit]. Ce résidu appartient au brin- $\beta$  C-terminal et dans toutes nos autres structures il interagit avec l'Arg54 [Figure-V-108-B]. Dans la nouvelle structure le changement de la conformation de la chaîne latérale de cette arginine rend l'interaction impossible à cause d'un clash stérique. La rupture de cette interaction pourrait être à l'origine du réarrangement de ce brin C-terminal.



**Figure-V-108** : Vues stéréoscopiques de l'environnement du site S3 dans la structure native et dans la structure avec le brin C-terminal désordonné

**A** - En magenta, configuration de chaînes latérales des résidus telle qu'observée dans la structure native. En gris, les mêmes résidus dans la structure avec le brin- $\beta$  C-terminal désordonné. En vert foncé, carte de densité électronique *Fo-Fc* "omit" contourée à  $3\sigma$  et visualisée dans un rayon de 4 Å autour d'ion chlorure qui se trouve du côté extérieur de la partie basse du tonneau. **B** - La structure native en magenta et le brin- $\beta$  C-terminal en vert. En cyan, la chaîne latérale de l'arginine 54 avec l'ion chlorure externalisé, tels qu'observés dans la nouvelle structure.

Nous avons donc conclu que la nouvelle structure montre une conformation que la protéine pourrait adopter lors du transfert de son produit hors de sa cavité. Le transfert semble être initié par un basculement de la chaîne latérale de l'Arg54 ce qui engendre un réarrangement de la chaîne principale après le motif GXXP. Ceci diminue l'affinité du site S3 pour le produit de HydE négativement chargé et facilite son externalisation. Le changement de la conformation du brin C-terminal expose la partie basse du tonneau au solvant. Les résidus formant le fond de la cavité de HydE sont alors accessibles pour une éventuelle interaction avec la maturase HydF. Cette dernière pourrait alors accueillir le produit de HydE [voir paragraphes IB-4 & I-3.2].

### V-2.3 Conclusions du Chapitre V: HydE - une ligne d'assemblage à sens unique

De nombreux essais de co-cristallisation de HydE avec d'éventuels ligands nous ont permis d'obtenir, hors les structures de complexes, une structure où la protéine adaptait un repliement particulier. Dans cette structure la chaîne principale se retrouve désordonnée du côté du brin C-terminal, tout de suite après la séquence GXXP. Ceci expose au solvant les résidus qui forment la partie basse de la cavité interne de HydE. On remarque aussi des changements de conformation de chaînes latérales de certains de ces résidus, comme par exemple l'Arg54, ce qui a pour effet l'externalisation du chlorure du site S3. Nous proposons que cette structure corresponde à un état où la protéine serait prête à transférer un éventuel produit hors de sa cavité catalytique. Les résidus strictement conservés, qui se trouvent au fond de la cavité et qui par le mouvement du brin- $\beta$  C-terminal se retrouvent exposés au solvant et pourraient éventuellement interagir avec HydF pour assurer le transfert du produit.

La nouvelle structure de HydE serait en accord avec de récents progrès dans notre compréhension de la structure et du mécanisme catalytique de HydG<sup>5-10</sup>. Cette maturase, qui possède elle aussi un repliement en tonneau ( $\beta\alpha$ )<sub>8</sub> complet, présente une directionalité dans le devenir de son substrat. La réaction radicalaire initiale a lieu dans la partie supérieure de la cavité de HydG. L'intermédiaire de la réaction descend le long de cette cavité vers sa partie inférieure, où il sera transformé en produit final. Comme cela avait été proposée dans la littérature<sup>2</sup>, la structure de HydE présentée dans ce mini-chapitre suggère qu'il est fort probable que cette enzyme présente une directionalité comparable à celle de HydG.

Ainsi dans HydE, le substrat, fixé entre le site S1 et S2, subirait une attaque radicalaire dans la partie supérieure de la cavité. Ensuite l'intermédiaire de la réaction franchirait la ceinture de résidus hydrophobes et irait se loger dans le site S3 au fond de la cavité où il se retrouverait à l'abri de solvant. Il serait alors transféré à HydF par un basculement de la chaîne latérale de l'Arg54 et un résultant réarrangement du brin C-terminal, comme on l'a observé dans la structure présentée ci-dessus. A ce stade, il n'est pas clair quant au rôle exact de la ceinture de résidus hydrophobes. Elle peut servir de sas pour évacuer le produit final hors HydE. Ou bien on peut imaginer que HydE doit effectuer deux turnover sur deux molécules de substrat pour fabriquer une molécule de produit. Dans ce cas, la ceinture pourrait servir à protéger l'intermédiaire de la réaction logé dans le site S3 contre l'hydrolyse pendant que le second turnover a lieu dans la partie supérieure de la cavité.

Nos données cristallographiques suggèrent que l'hypothèse que le motif GXXP puisse agir comme une charnière dans certaines protéines à radical SAM a été correcte. Il est possible que ce motif semi-conservé soit un vestige d'un ancêtre commun d'un groupe de protéines à radical SAM possèdent une topologie en tonneau ( $\beta\alpha$ )<sub>8</sub> complet. Nous pouvons imaginer que dans cette protéine ancestrale un changement de conformation du brin- $\beta$  en aval du motif GXXP était à l'origine de la libération d'un des produits de la réaction de cette enzyme. Quoi qu'il en soit, la structure présentée ci-dessus semble confirmer la proposition que HydE et HydG, et potentiellement d'autres protéines à radical SAM avec la même topologie, puissent être considérées comme des « nano-usines », qui se servent de leurs cavités catalytiques comme de lignes d'assemblage de produits à sens unique<sup>11</sup>.

### V-3 Références du Chapitre V

- 1 Nicolet, Y. & Drennan, C. L. AdoMet radical proteins—from structure to evolution—alignment of divergent protein sequences reveals strong secondary structure element conservation. *Nucleic Acids Research* **32**, 4015-4025, doi:10.1093/nar/gkh728 (2004).
- 2 Nicolet, Y. *et al.* X-ray Structure of the [FeFe]-Hydrogenase Maturase HydE from *Thermotoga maritima*. *Journal of Biological Chemistry* **283**, 18861-18872, doi:10.1074/jbc.M801161200 (2008).
- 3 Nicolet, Y., Martin, L., Tron, C. & Fontecilla-Camps, J. C. A glycyI free radical as the precursor in the synthesis of carbon monoxide and cyanide by the [FeFe]-hydrogenase maturase HydG. *FEBS Letters* **584**, 4197-4202, doi:10.1016/j.febslet.2010.09.008 (2010).
- 4 Posewitz, M. C. *et al.* Discovery of Two Novel Radical S-Adenosylmethionine Proteins Required for the Assembly of an Active [Fe] Hydrogenase. *Journal of Biological Chemistry* **279**, 25711-25720, doi:10.1074/jbc.M403206200 (2004).
- 5 Dinis, P. *et al.* X-ray crystallographic and EPR spectroscopic analysis of HydG, a maturase in [FeFe]-hydrogenase H-cluster assembly. *Proceedings of the National Academy of Sciences* **112**, 1362-1367, doi:10.1073/pnas.1417252112 (2015).
- 6 Nicolet, Y. *et al.* Crystal structure of HydG from *Carboxydotherrmus hydrogenoformans*: a trifunctional [FeFe]-hydrogenase maturase. *Chembiochem : a European journal of chemical biology* **16**, 397-402, doi:10.1002/cbic.201402661 (2015).
- 7 Suess, D. L. M. *et al.* Cysteine as a ligand platform in the biosynthesis of the FeFe hydrogenase H cluster. *Proceedings of the National Academy of Sciences* **112**, 201508440, doi:10.1073/pnas.1508440112 (2015).
- 8 Suess, D. L. M., Kuchenreuther, J. M., Paz, L., Swartz, J. R. & Britt, D. R. Biosynthesis of the [FeFe] Hydrogenase H Cluster: A Central Role for the Radical SAM Enzyme HydG. *Inorganic Chemistry*, doi:10.1021/acs.inorgchem.5b02274 (2015).
- 9 Suess, D. L. M. *et al.* The Radical SAM Enzyme HydG Requires Cysteine and a Dangler Iron for Generating an Organometallic Precursor to the [FeFe]-Hydrogenase H-Cluster. *Journal of the American Chemical Society*, doi:10.1021/jacs.5b12512 (2016).
- 10 Pagnier, A., Martin, L., Zeppieri, L., Nicolet, Y. & Fontecilla-Camps, J. C. CO and CN- syntheses by [FeFe]-hydrogenase maturase HydG are catalytically differentiated events. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 104-109, doi:10.1073/pnas.1515842113 (2015).
- 11 Nicolet, Y. & Fontecilla-Camps, J. C. Structure-Function Relationships in [FeFe]-Hydrogenase Active Site Maturation. *Journal of Biological Chemistry* **287**, 13532-13540, doi:10.1074/jbc.R111.310797 (2012).



## Conclusions Générales

Les travaux réalisés au cours de ma thèse ont abouti à une première étude cristallographique d'une réaction radicalaire dans une protéine à radical SAM, observé dans l'enzyme HydE [l'article Rohac *et al.* Nature Chem. 2016 en Annexe]. Dans un premier temps nous avons réussi à identifier un groupe de molécules, dérivées de la cystéine, contenant un cycle thiazolidine avec un ou deux groupements carboxylates, qui ont une très bonne affinité pour le site actif de HydE. Certains de ces ligands se sont montrés d'être des substrats de l'enzyme. Grâce à ces substrats nous avons pu mettre en évidence un nouveau mécanisme d'attaque radicalaire dans les protéines à radical SAM. En effet, la quasi-totalité de ces enzymes activent leur substrat en lui arrachant un atome d'hydrogène. Dans HydE nous avons observé une attaque directe du radical 5'-déoxyadénosyl sur l'atome soufre du thioéther appartenant au cycle thiazolidine. Ceci résulte en la formation d'une nouvelle liaison C-S et en la rupture de la liaison C-S du thioéther qui se trouve dans l'axe de cette nouvelle liaison créée. Les expériences avec un analogue sélénié du substrat conduisent à la formation d'une nouvelle liaison C-Se. Cette réaction constitue un exemple pas comme les autres d'une insertion d'un atome de soufre ou de sélénium catalysée par une enzyme à radical SAM.

Les calculs de QM et de QM/MM, basés sur nos données cristallographiques à haute résolution, montrent que cette attaque radicalaire procède par un mécanisme de substitution homolytique  $S_H2$ . Nous en avons conclu que le clivage de la SAM, la formation de la nouvelle liaison C-S et la rupture de l'autre se font de façon concomitante. Ce mécanisme et la configuration du site actif de différentes protéines à radical SAM (compacité et alignement des atomes) suggèrent, que dans le cas de cette superfamille de protéines le radical 5'-déoxyadénosyl serait plutôt un état de transition et donc pas une espèce intermédiaire isolable.

La réaction observée dans HydE ne nous permet pas de prédire la nature exacte du substrat physiologique de l'enzyme. Cependant, l'identification d'un site d'affinité pour l'atome de soufre à proximité de l'atome C5' de la SAM et le fait que l'enzyme arrive à attaquer un groupement thiol ou thioéther indique clairement que cette protéine est impliquée dans la production d'une molécule soufrée.

Pendant la dernière année de ma thèse j'ai pu mettre en place un système de maturation de l'hydrogénase à [FeFe] dans l'extrait, adapté des travaux de James R. Swartz. Nous avons utilisé cet outil pour vérifier l'impact des ligands de la protéine HydE sur la maturation de l'hydrogénase à [FeFe] *CrHydA1*. Bien que les substrats de HydE n'aient pas eu d'effet sur l'activité hydrogénase de l'extrait, nous avons pu identifier un group d'alphacéto acides, qui semblent augmenter le taux de maturation de l'enzyme. La prochaine étape est d'élucider par quel mécanisme ces molécules influence le processus de maturation.

Le système de maturation de l'hydrogénase maintenant implémenté aux techniques de notre laboratoire pourra être utilisé pour l'identification de fractions de l'extrait cellulaire nécessaires pour l'activation de l'enzyme à condition de trouver un système de fractionnement de l'extrait approprié.



## Matériels et Méthodes

### MM-1 Matériels

Cette section de Matériels et Méthodes résume les matériels biologiques utilisés dans les expériences tout au long de ma thèse, ainsi que les milieux de culture et leurs procédés de préparation.

#### MM-1.1 Matériels biologiques

Dans les matériels biologiques, nous considérons les gènes utilisés, des vecteurs de surexpression et de souches bactériennes.

##### MM-1.1.1 Gènes utilisés

- *hydE* (gène : TM\_1269 dans Unip) de *Thermotoga maritima* (THEMA - souche ATCC 43589 / MSB8 / DSM 3109 / JCM 10099) : 1047 bp
- *hydE* (gène : CA\_C1631 dans Unip) de *Clostridium acetobutylicum* (CLOAC - souche ATCC 824 / DSM 792 / JCM 1419 / LMG 5710 / VKM B-1787) : 1053 bp
- *hydF* (gène : CA\_C1651 dans Unip) de *Clostridium acetobutylicum* (CLOAC - souche ATCC 824 / DSM 792 / JCM 1419 / LMG 5710 / VKM B-1787) : 1236 bp
- *hydG* (gène : CA\_C1356 dans Unip) de *Clostridium acetobutylicum* (CLOAC - strain ATCC 824 / DSM 792 / JCM 1419 / LMG 5710 / VKM B-1787) : 1419 bp
- *hydA1* (gène : 5718949 dans Genbank) de *Chlamydomonas reinhardtii* : 4664 bp

##### MM-1.1.2 Vecteurs de surexpression

Nom du plasmide	Protéine (mutations éventuelles sur le gène d'intérêt)	Taille du plasmide et résistance
pT7-7 <i>hydE Tm</i>	HydE de THEMA sauvage	2743 pb pT7-7 Ampicilline
pET-15b <i>hydE Ca</i> N-strep	HydE de CLOAC Sauvage strep N-terminale	5708 pb pET-15b Ampicilline
pET-15b <i>hydE Ca</i>	HydE de CLOAC avec une étiquette	5708 pb pET-15b Ampicilline
pET-15b <i>hydF Ca</i>	HydF de CLOAC sauvage	5708 pb pET-15b Ampicilline
pET-15b <i>hydG Ca</i>	HydG de CLOAC sauvage	5708 pb pET-15b Ampicilline
pET-22b <i>hydA1 Cr</i> N-strep	HydA1 de CHLRE avec une étiquette strep N-terminale	5493 pb pET-22b Ampicilline

Table MM-1 : Vecteurs d'expression contenant les gènes de protéines d'intérêt

### ***MM-1.1.3 Souches bactériennes***

#### **- Souche d'expression *E. coli* BL21(DE3)**

Génotype : F<sup>-</sup> *ompT gal dcm lon hsdS<sub>B</sub>(r<sub>B</sub><sup>-</sup>, m<sub>B</sub><sup>-</sup>) λ(DE3) [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]*

Cette souche est utilisée dans le but de produire une protéine dont le gène a été cloné en aval du promoteur T7. La transcription du gène d'intérêt est inductible à l'IPTG.

#### **- Souche d'expression *E. coli* BL21 Codonplus<sup>TM</sup> (DE3) RIL**

Génotype : F<sup>-</sup> *ompT hsdS<sub>B</sub>(r<sub>B</sub><sup>-</sup>, m<sub>B</sub><sup>-</sup>) dcm<sup>+</sup> Tet<sup>r</sup> gal λ(DE3) endA Hte [argU ileY leuW Cam<sup>r</sup>]*

Bien que le code génétique soit universel pour l'ensemble des êtres vivants sur Terre, la fréquence d'utilisation de certains codons peut différer de façon importante entre les espèces. Ainsi l'expression hétérologue d'un gène dans un organisme hôte phylogénétiquement éloigné de l'organisme d'origine peut avoir pour conséquence un faible rendement d'expression de la protéine d'intérêt. En effet, nous avons remarqué que l'expression de HydE de THEMA dans la souche d'expression *E. coli* BL21(DE3) ne donne pas de quantités de HydE observables sur un gel SDS-PAGE dénaturant. Pour palier à ce problème deux solutions se proposent :

Utiliser un gène synthétique dont la séquence a été optimisée de façon à contenir que les codons fréquemment utilisés chez *E. coli*

Utiliser une souche d'expression portant les gènes codant pour les ARN de transfert des codons rares chez *E. coli*

Nous avons opté pour la seconde solution en exprimant HydE de THEMA dans la souche d'expression *E. coli* BL21 Codonplus<sup>TM</sup> (DE3) RIL. Ces cellules contiennent un vecteur portant une résistance au chloramphénicol et des copies supplémentaires pour les gènes des ARN de transfert *argU*, *ileY*, et *leuW*, reconnaissant respectivement les codons AGA/AGG, AUA et CUA.

#### **- Souche d'expression *E. coli* BL21(DE3) pRSF *Isc metK* Duet-1**

La surexpression de protéines à centres Fe-S produit souvent une population de protéines hétérogène avec des agrégats Fe-S partiellement reconstitués. Afin d'optimiser la maturation de ces protéines à centres Fe-S nous avons utilisé une souche d'expression *E. coli* BL21(DE3) pRSF *Isc metK* Duet-1. Cette souche contient le vecteur pRSF, dans lequel ont été insérés, sous le contrôle du promoteur T7, les gènes de l'opéron *isc* (dépourvu du régulateur IscR) ainsi que le gène codant pour la SAM synthétase MetK, issus d'*E. coli*. L'induction entraîne une accumulation non-réprimée de la machinerie bactérienne de maturation des centres Fe-S. La surexpression de la SAM synthétase a pour effet l'augmentation de la concentration de la SAM dans le cytoplasme. Ceci devrait aboutir à l'augmentation de la population des enzymes à radical SAM avec le cofacteur lié au site actif.

#### **- Souche d'expression *E. coli* BL21 Codonplus<sup>TM</sup> (DE3) RIL pRSF *Isc metK* Duet-1**

Cette souche d'expression est utilisée pour la surproduction d'enzymes à radical SAM dont le gène contient des codons rares chez *E. coli* (voir les paragraphes précédents).

- Cellules compétentes *E. coli* DH5 $\alpha$

Génotype : F<sup>-</sup> *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG*  $\Phi$ 80*lacZ* $\Delta$ M15  $\Delta$ (*lacZYA-argF*) U169, *hsdR17*(r<sub>K</sub><sup>-</sup> m<sub>K</sub><sup>+</sup>),  $\lambda$ -

Cette souche, dont l'endonucléase EndA a été désactivée, permet la surproduction et l'accumulation de plasmides bactériens. D'autre part, la mutation sur le gène *recA*, rend les cellules déficientes pour la recombinaison homologe, et donc évite les risques de mutations du plasmide au cours de sa préparation. Cette souche est utilisée pour constituer un stock d'ADN plasmidique.

## MM-1.2 Milieux de culture

Le milieu *LB* ou Luria Bertani (Sigma) est composé de 10 g.L<sup>-1</sup> de peptone tryptique de caséine, 5 g.L<sup>-1</sup> d'extrait autolytique de levure et 5 g.L<sup>-1</sup> de NaCl, pH 7. Sa stérilisation se fait à l'autoclave pendant 15 min à 121°C. Une fois stérile et refroidi il a été utilisé pour l'ensemble de pré-cultures des souches transformées par les plasmides, pour certaines cultures d'expression et également pour les cultures en milieu solide (milieu LB-Agar).

Le milieu *TB* modifié (Fluka) a été préparé en re-suspendant 47.6 g de poudre et 8 mL de glycérol dans 1 L d'eau. Ce milieu a été utilisé pour certaines cultures d'expression en milieu aérobie.

Pour l'expression de protéines en milieu anaérobie stricte deux milieux tamponnés différents ont été préparés pour tester l'impact du type de métabolisme de la cellule sur la qualité de maturation des centres Fe-S *in vivo*. Les deux milieux sont basés sur le milieu *TB*, tamponné avec 100 mM de MOPS (Sigma Aldrich) pH 7.1. Ils diffèrent par le rajout de 25 mM de glucose ou de fumarate 30 minutes avant l'induction en anaérobie stricte, ce qui va orienter le métabolisme de cellules vers la fermentation acide mixte ou la respiration du fumarate, respectivement.

## MM-2 Biologie moléculaire

Une étude de caractérisation *in vitro* d'une macromolécule biologique, par des méthodes biochimiques ou biophysiques, commence en général par l'obtention d'une grande quantité de cette macromolécule à étudier, avec un haut degré de pureté (souvent >95%) et d'homogénéité (présence de cofacteurs ou différents états d'oligomérisation).

Dans le cas des protéines bactériennes, telles que les protéines de l'opéron *Hyd* issues de *THEMA* et *CLOAC*, le moyen de production fréquemment utilisé est le système de surexpression de protéines recombinantes pET chez *E. coli*. Le développement de ce système découle des travaux de Monod et Jacob sur la compréhension de la régulation de l'expression des gènes de l'opéron *lac*. Le système d'expression pET est composé de deux parties :

- un vecteur d'expression, tel qu'un plasmide bactérien artificiel
- une souche d'*E. coli*, qui a été génétiquement modifiée, de façon à insérer le gène codant pour la T7 RNA polymérase dans le chromosome de la bactérie en aval d'un promoteur et d'un opérateur *lac*.

Le gène de la protéine d'intérêt est cloné dans le vecteur d'expression en aval d'un promoteur T7 et d'un opérateur *lac*. Ce vecteur est ensuite introduit dans la souche d'*E. coli* génétiquement modifiée. L'induction de l'expression du gène d'intérêt s'effectue par l'ajout d'IPTG, un analogue non hydrolysable d'allolactose, dans le milieu de culture cellulaire. L'IPTG lie le répresseur de l'opéron lactose, ce qui rend ce dernier accessible à la fixation de la machinerie de transcription. L'utilisation d'un promoteur fort, comme le promoteur viral T7, d'un opérateur inductible et d'un vecteur d'expression à grand nombre de copies par cellule permet une rapide accumulation de la protéine d'intérêt dans la cellule au moment désiré. Souvent en fin de culture, la protéine ainsi exprimée constitue l'élément majoritaire du protéome de la cellule.

### **MM-2.1 Préparation de cellules complètes**

La transformation est un procédé dans lequel les cellules vont incorporer une molécule d'ADN de leur environnement dans leur génome. Toutes nos souches de cellules compétentes ont été transformées avec la méthode du choc thermique. Pour produire les cellules compétentes 10 mL de culture en phase exponentielle (densité optique à 600 nm égale à 0.4) sont culottés (10 min à 5000 rpm, 4°C) et les cellules sont ensuite mises dans la glace pendant 10 minutes. Le culot est repris dans 5 mL de solution stérile de CaCl<sub>2</sub> à 100 mM. Après avoir passé 30 minutes dans la glace les cellules sont culottées à nouveau, puis le culot est re-séché pour être finalement repris dans 700 µL de solution stérile de CaCl<sub>2</sub> à 100 mM contenant 10 % de glycérol. Les cellules compétentes ainsi préparées sont immédiatement congelées dans l'azote liquide et stockées à -80°C.

### **MM-2.2 Transformation des souches *Escherichia coli* compétentes**

Pour réaliser une transformation, les cellules compétentes sont d'abord pré-incubées en présence du plasmide dans un milieu riche en chlorure de calcium à 4°C (dans un bain de glace). Les quantités de plasmide typiquement utilisées varient entre 0.5 et 1.5 µL à 80–320 ng.µL<sup>-1</sup> de plasmide pour 100 µL de cellules compétentes. Au bout de 30 minutes, ces cellules sont incubées pendant une minute dans un bain-marie préchauffé à 42°C et ensuite refroidies dans de la glace pendant 2 minutes. 250 µL de LB sont ensuite ajoutés aux cellules transformées et ces dernières sont d'abord incubées à 37°C pendant une heure avant d'être étalées sur une boîte de Petri avec du LB-Agar contenant l'antibiotique approprié. L'apparition de colonies au bout de 12h d'incubation à 37°C est un signe de l'incorporation réussie du plasmide par les cellules. Les cellules transformées sont ensuite prélevées soit à partir d'une seule colonie (un seul clone) ou à partir de plusieurs colonies et elles sont

incubées à 37°C pendant 12h pour permettre leur prolifération. Cette culture est par la suite utilisée pour créer un stock glycérol ou, dans le cas d'*E. coli* DH5 $\alpha$ , pour purifier le plasmide à l'aide d'un kit de purification d'ADN plasmidique.

### **MM-2.3 Mutagénèse dirigée par PCR**

L'insertion de l'étiquette strep en N-terminal de la séquence de HydE de THEMA ainsi que la préparation du plasmide contenant une triple mutation C311S-C319S-C322S a été effectuée par l'amplification en chaîne par l'ADN polymérase (PCR). Cette technique permet la biosynthèse de fragments d'ADN ou des plasmides entiers grâce à 2 oligonucléotides (amorces simple brin), une matrice d'ADN et l'enzyme ADN polymérase thermorésistante issue d'un organisme thermophile. La réaction consiste en une répétition de 3 étapes qui forment un cycle de température : une étape de dénaturation (séparation des 2 brins de l'ADN matrice), une étape d'hybridation (appariement des amorces à leur séquence complémentaire sur la matrice d'ADN) et une étape d'élongation (synthèse de l'ADN complémentaire par la polymérase). En utilisant les amorces qui possèdent une séquence en nucléotide spécifiquement modifiée par rapport au brin matrice, il est possible d'introduire une ou plusieurs mutations ponctuelles dans le gène d'intérêt.

#### *Amplification du plasmide pT7-7 hydE Tm N-strep*

L'étiquette strep a été insérée entre les résidus Met1 et Thr2 pour générer une séquence N-terminale modifiée : MWSHPQFEKAST. La mutagenèse est réalisée dans le tampon Phusion<sup>TM</sup> HF avec 100 ng de plasmide (pT7-7 hydE Tm) [ref. Rubach FEBS 2005], 15 pmol d'amorces [décrites dans [Table MM-2](#)], 1 mM de dNTP et 1 U de Phusion<sup>TM</sup> DNA Polymerase (Finnzymes). L'échantillon est alors chauffé 5 min à 98°C, puis il subit 25 cycles de 30 s à 98°C / 30 s à la température d'hybridation des amorces / 1 min 30 s à 72°C, pour finir avec une dernière étape de 10 min à 72°C. A la différence des brins d'ADN synthétiques, les brins parentaux sont issus d'un système d'expression bactérien et présentent donc des méthylations au sein de leur séquence. A la fin de l'amplification par PCR le produit est incubé pendant 1h30 à 37°C avec 30 U de DpnI dans le tampon NEBuffer 4 (NEB). La digestion à l'endonucléase DpnI, qui reconnaît spécifiquement la séquence GA(Me)TC de l'ADN méthylé, permet donc de dégrader l'ADN parental non muté. Le plasmide double brin linéaire que l'on obtient après cette digestion est transformé dans une souche de *E. coli* DH5 $\alpha$  pour une purification d'ADN plasmidique (voir paragraphe suivant).

Nom du plasmide	Mutations éventuelles sur le gène d'intérêt	Taille du plasmide et résistance	Séquence des amorces	Conditions de PCR	Matrice : Plasmide d'origine
pT7-7 hydE Tm N-strep	HydE de THEMA avec une étiquette strep N-terminale	2743 pb pT7-7 Ampicilline	Hybridation de cette amorce sur le brin non codant : 5'-CACCCGCAGTTCGAAAAGGCAAGCACC GGTAGAGAAATTCTGGAAAAA-3' Hybridation de cette amorce sur le brin codant : 5'-CTTTTCGAACTGCGGGTGGCTCCACATG GTATATCTCCTTCTTAAAGTTAAACAAA A-3'	30'' à 98°C 30'' à 76°C 1'30" à 72 °C 25 cycles	pT7-7 hydE Tm
pT7-7 hydE Tm N-strep ΔFeS	HydE de THEMA avec une étiquette strep N-terminale et une triple mutation C311S, C319S et C322S	2743 pb pT7-7 Ampicilline	Hybridation de cette amorce sur le brin non codant : 5'-GAGAAGGACACTGCGTCTATTCCCTCCG TGATGAAGATGATA-3' Hybridation de cette amorce sur le brin codant : 5'-CGCAGTGTCTTCTCGAAAACGGAGATC TTTCCAGGATA-3'	30'' à 98°C 30'' à 70°C 1'30" à 72 °C 25 cycles	pT7-7 hydE Tm N-strep

**Table MM-2 : Amorces de PCR utilisées pour les mutagenèses dirigées sur les gènes hydE et vecteurs d'expression résultant**

#### Amplification du plasmide pT7-7 hydE Tm N-strep ΔFeS

Afin de supprimer le centre Fe-S semi-conservé de HydE nous avons remplacé les trois résidus cystéines impliqués dans sa chélation par des serines. Le plasmide pT7-7 hydE Tm N-strep, dont le protocole de préparation est décrit dans le paragraphe précédent, a été utilisé comme matrice pour produire le triple mutant C311S-C319S-C322S. Les séquences des amorces et les conditions de l'amplification par PCR sont résumées dans le [Table MM-2](#). La présence de mutations et de l'insertion de l'étiquette strep a été confirmée par le séquençage du gène entier.

#### **MM-2.4 Préparation d'ADN plasmidique**

Les cellules compétentes *E. coli* DH5α sont transformées et étalées sur une boîte de Pétri LB-Agar contenant l'antibiotique approprié comme décrit dans le paragraphe précédent. Une colonie est sélectionnée pour inoculer 5 mL de milieu LB qui sont ensuite mis en culture sur la nuit à 37°C. Les cellules en phase stationnaire sont culotées (5 min, 5 000 rpm à 4°C) et le kit QIAprep Miniprep (Qiagen) est utilisé pour purifier les plasmides. Les cellules subissent d'abord une lyse alcaline, en présence de RNase, ainsi qu'une précipitation de l'ADN génomique par l'acide acétique. L'ADN génomique est alors éliminé par centrifugation, et le surnageant est déposé sur une colonne comportant une membrane constituée de gel de silice, permettant une adsorption sélective de l'ADN plasmidique, en présence d'une forte

concentration en sel. Après un lavage à l'éthanol pour éliminer les sels, l'ARN et les protéines, l'ADN plasmidique est élué avec un tampon de faible force ionique (eau ou tampon d'éluion du kit : Tris 10 mM, pH 8,5). La concentration L'ADN plasmidique purifié est déterminée à partir de l'absorbance à 260 nm et sa pureté est estimée à partir du rapport d'absorbances à 260 nm et à 280 nm. Le plasmide est alors stocké à  $-20^{\circ}\text{C}$ .

## **MM-3 Biochimie**

### **MM-3.1 Surexpressions des protéines**

#### ***MM-3.1.1 Respiration du fumarate***

De notre côté nous avons utilisé le plasmide pRSFiscMetKDuet-1 décrit dans [MM-1.1.2]. Le stock glycérol de cellules co-transformées par le plasmide pRSFiscMetKDuet-1 et par un plasmide porteur du gène codant pour la protéine d'intérêt [MM-2.3] est repris dans le milieu TB tamponné avec 50 mM MOPS pH 7.1 contenant les antibiotiques appropriés. La culture est maintenue à  $25^{\circ}\text{C}$  et agitée à 200 rpm jusqu'à une  $\text{DO}_{600\text{ nm}}=0.2-0.3$ . La température est ensuite diminuée à  $20^{\circ}\text{C}$  et quand la concentration de cellules atteint une  $\text{DO}_{600\text{ nm}}=0.4$ , la culture est transférée en boîte à gants, pré-incubée à  $20^{\circ}\text{C}$ . L'oxygène dissout dans la solution peut être rapidement enlevé par bullage de la solution à l'aide d'une petite pompe à aquarium à l'intérieur de la boîte à gant (BAG). Sinon, il va être progressivement consommé par les cellules. Trente minutes après le passage en milieu anaérobie, sont ajoutés : 25 mM de sodium fumarate (SIGMA), un mélange de précurseurs de cofacteurs MEM Vitamines 100x (SIGMA M6895) (25 mL pour 2 L de culture), 400  $\mu\text{M}$  de L-Cystéine et 100  $\mu\text{M}$  de citrate de  $\text{Fe}^{3+}$ . Quand la  $\text{DO}_{600\text{ nm}}$  atteint 0.8, l'expression est induite avec 10  $\mu\text{M}$  d'IPTG à  $20^{\circ}\text{C}$ , 160 rpm et se poursuit sur la nuit (12-16h). La culture est ensuite transvasée dans les pots étanches et centrifugée pendant 10 min à 5000 rpm à  $4^{\circ}\text{C}$ .

#### ***MM-3.1.2 Fermentation du glucose***

Alternativement, on peut faire croître les cellules par fermentation acide mixte, plutôt que par la respiration du fumarate<sup>1-3</sup>. Toutes les étapes d'expression anaérobie sont les mêmes pour les deux métabolismes, sauf que dans le cas de la fermentation on utilise du milieu LB tamponné avec 100 mM MOPS pH 7.4, contenant 25 mM de D-glucose et par conséquent l'ajout de fumarate est éliminé du protocole.

#### ***MM-3.1.3 Protocole de Swartz - Fermentation du glucose***

Ce protocole est dérivé des travaux de James R. Swartz et de son étudiant Jon M. Kuchenreuther et a été utilisé surtout pour la production de lysats clarifiés contenant des maturases HydE, HydG et HydF reconstituées *in vivo*. Ces lysats ont été par la suite utilisés

lors de tests de maturation de l'hydrogénase à [FeFe] HydA1 de *Chlamydomonas reinhardtii* [MM-1.1.1]. Les cellules sont d'abord cultivées en aérobie dans le milieu SBC stérile (0,9 L 1,1x LB autoclavé, 100 mL 1 M MOPS pH 7.4 filtré, 10 mL D-glucose à 25 % w/v filtré), avec les antibiotiques appropriés, à 30°C, 200 rpm jusqu'à une  $DO_{600\text{ nm}}=0.2$ . 200 mg de citrate de Fer, 350 mg de L-cystéine et 10 mL de MEM Vitamines 100x sont alors ajoutés par litre de culture et la température est diminuée à 20°C. Une fois que la température du milieu a atteint les 20°C, la culture est transférée en boîte à gants, pré-incubée à la même température. L'induction se fait avec 500  $\mu\text{M}$  d'IPTG au moins trente minutes après le passage en milieu anaérobie stricte et une fois que la culture a atteint une  $DO_{600\text{ nm}}=0.8$ . L'induction se poursuit sur la nuit à 20°C, 160 rpm.

### MM-3.2 Purifications des protéines

Les protéines étudiées dans ces travaux et ayant des centres FeS ont été purifiées avec un système GE ÄKTA purifier, placé dans une boîte à gant JACOMEX à l'abri de l'oxygène. L'atmosphère de la boîte à gants se compose 100% d'azote et moins de 5 ppm  $\text{O}_2$ .

Le protéine HydE et THEMA ont été purifiées en une étape par chromatographie par échange d'ions. Le principe de cette méthode consiste à choisir, en fonction du pI de la protéine d'intérêt, un pH, auquel cette protéine aura une charge nette opposée à la charge de la phase stationnaire. Dans ces conditions, la protéine va se lier à la phase stationnaire par interactions électrostatiques. L'élution s'effectue par une augmentation progressive de la force ionique, ce qui va créer une compétition entre la protéine chargée et un ion de la même charge. Cette compétition va provoquer le décrochage de la protéine de la résine et son élution.

Les protéines HydE, et HydA de THEMA portant d'une étiquette *strep* ont été purifiées par chromatographie d'affinité. L'étiquette *strep* est composée d'une séquence de huit résidus (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys) ayant une forte affinité pour la streptavidine. Dans le cas de toutes nos protéines, cette étiquette se trouve en extrémité N-terminale et elle est suivie d'un court linker (Ala-Ser). A l'aide d'une résine fonctionnalisée avec de la streptavidine modifiée on peut séparer les protéines portant une étiquette *strep* du lysat cellulaire. Après une étape de lavage, l'élution se fait avec une faible concentration de déthiobiotine – une petite molécule qui se lie de façon compétitive à la streptavidine et déplace donc la protéine d'intérêt.

Les deux purifications standards sont décrites en détail dans les sections [MM-3.2.1 et MM-3.2.2]. Les diverses préparations protéiques ont été purifiées en suivant les grandes lignes de ces protocoles. En fonction de la pureté de l'échantillon une étape de purification par chromatographie d'exclusion stérique a été rajoutée. Les autres préparations d'échantillons protéiques sont brièvement décrites dans les sections [MM-3.2.3 - MM-3.2.8 et toutes les variations par rapport aux protocoles standards sont résumées dans le [Table MM-3](#)].

Nom de l'échantillon	Souche Antibiotiques	Milieu Expression	Tampon Sonication
HydE <i>Tm</i> no tag	<i>E. coli</i> BL21 Codonplus™ (DE3) RIL pT7-7 hydE <i>Tm</i> Cam & Amp	LB 3h, 200 rpm, 37°C (+Fe <sup>3+</sup> , L-Cys)	Tris 100 mM à pH 7.6, 0.5% Triton X-100 90%, 2 min, (2s :10s)
HydE <i>Tm</i> N-strep	<i>E. coli</i> BL21 Codonplus™ (DE3) RIL pT7-7 hydE <i>Tm</i> N-strep Cam & Amp	LB 3h, 200 rpm, 37°C (+Fe <sup>3+</sup> , L-Cys)	100 mM Tris, 150 mM NaCl, pH 8.0 90%, 2 min, (2s :10s)
HydE <i>Tm</i> N-strep ΔFeS	<i>E. coli</i> BL21 Codonplus™ (DE3) RIL pT7-7 hydE <i>Tm</i> N-strep Cam & Amp	LB 3h, 200 rpm, 37°C (+Fe <sup>3+</sup> , L-Cys)	100 mM Tris, 150 mM NaCl, pH 8.0 90%, 2 min, (2s :10s)
HydE <i>Tm</i> N-strep ΔFeS Se	<i>E. coli</i> BL21 Codonplus™ (DE3) RIL pT7-7 hydE <i>Tm</i> N-strep Cam & Amp	LB 3h, 200 rpm, 37°C	100 mM Tris, 150 mM NaCl, pH 8.0, 5 mM EDTA, 5 mM Tiron 90%, 3 min, (2s :10s)
HydE <i>Tm</i> N-strep ΔFeS recon in vivo fumarate	<i>E. coli</i> BL21 Codonplus™ (DE3) RIL pT7-7 hydE <i>Tm</i> N-strep pRSF <i>Isc metK</i> Duet-1 Cam, Kana & Amp	TB, 50 mM MOPS pH 7.1, 25 mM fumarate (+Fe <sup>3+</sup> , L-Cys, MEM Vit.) 16h, 160 rpm, 20°C induction ≤10 μM IPTG	100 mM Tris, 50 mM NaCl, pH 8.0 35%, 3 min, (2s :8s)
HydE <i>Tm</i> Nstrep ΔFeS recon. in vivo fermentation	<i>E. coli</i> BL21 Codonplus™ (DE3) RIL pT7-7 hydE <i>Tm</i> N-strep pRSF <i>Isc metK</i> Duet-1 Cam, Kana & Amp	TB, 50 mM MOPS pH 7.1, 25 mM glucose (+Fe <sup>3+</sup> , L-Cys, MEM Vit.) 16h, 160 rpm, 20°C induction <10 μM IPTG	100 mM Tris, 50 mM NaCl, pH 8.0 35%, 3 min, (2s :8s)
HydE <i>Ca</i> N-strep	<i>E. coli</i> BL21 Codonplus™ (DE3) RIL pRSF <i>Isc metK</i> Duet-1 pET-15b hydE <i>Ca</i> N-strep Cam, Kana & Amp	LB, 0.1 M MOPS pH 7.4, 2.5% D-glucose (+Fe <sup>3+</sup> , L-Cys, MEM Vit.) 16h, 160 rpm, 20°C	50 mM Tris pH 7.5, 300 mM KCl et 10% glycérol 35%, 3 min, (2s :8s)
HydA1 <i>Cr</i> N-strep	<i>E. coli</i> BL21 (DE3) pRSF <i>Isc metK</i> Duet-1 pET-22b hydA1 <i>Cr</i> N-strep Kana & Amp	LB, 0.1 M MOPS pH 7.4, 2.5% D-glucose (+Fe <sup>3+</sup> , L-Cys, MEM Vit.) 14h, 160 rpm, 20°C	100 mM Tris pH 8.0, 300 mM NaCl, 5% glycérol 35%, 3 min, (2s :8s)

**Table MM-3 :** Conditions de purification des différents échantillons protéiques

### MM-3.2.1 Purification de HydE sans tag de *Thermotoga maritima*

Une préculture est réalisée à partir du stock glycérol de cellules *E. coli* BL21 Codonplus<sup>TM</sup> (DE3) RIL porteuses du plasmide pT7-7 hydE *Tm* contenant le gène codant pour HydE de THEMA. Un étaleur stérile est utilisé pour prélever un faible volume de cellules toujours congelées. Ces cellules sont resuspendues dans 5 mL de milieu LB, Cam 34  $\mu\text{g.mL}^{-1}$ , Amp 50  $\mu\text{g.mL}^{-1}$  et incubées pendant la nuit (12h) à 37°C, avec une agitation de 200 rpm. Le lendemain, cette préculture est diluée dans 200 mL de milieu LB, Cam 34  $\mu\text{g.mL}^{-1}$ , Amp 100  $\mu\text{g.mL}^{-1}$  et incubée jusqu'à une  $\text{DO}_{600\text{ nm}} \approx 0.2$ , suite à quoi elle est à nouveau diluée dans 4 L de milieu LB, Cam 34  $\mu\text{g.mL}^{-1}$ , Amp 100  $\mu\text{g.mL}^{-1}$ . A ce moment le milieu peut être enrichi en soufre et en fer par ajout de 400  $\mu\text{M}$  de L-Cystéine et 100  $\mu\text{M}$  de citrate de  $\text{Fe}^{3+}$ . Cette culture d'expression est incubée à 37°C, avec une agitation de 200 rpm, jusqu'à ce qu'elle n'atteigne une  $\text{DO}_{600\text{ nm}} = 0.6-0.8$ . L'expression est alors induite à l'aide d'1 mM d'IPTG final. Au bout de 3h d'incubation à 37°C, 200 rpm la culture est centrifugée à 7800 g, 4°C, 20 min. Le culot est récupéré, puis transféré dans deux tubes Falcons de 50 mL et re-centrifugé à 3200 g, 4°C, 10 min avant d'être congelé dans de l'azote liquide et stocké à -80°C.

Le culot cellulaire obtenu après la surexpression est re-suspendu dans du tampon Tris 100 mM à pH 7.6, 0.5% Triton X-100, dégazé à l'intérieur de la boîte à gants ; 4 mL de ce tampon sont utilisés par un gramme de culot. Une tablette d'inhibiteurs de protéases 'Complete EDTA-free' (Roche) est rajoutée par cinq grammes de culot. Les cellules sont lysées par sonication pendant 2 minutes à 90% d'amplitude (2s de pulse et 10s de pause). Alternativement, un mode de sonication plus doux semble être suffisant pour induire la lyse des membranes en utilisant 35% d'amplitude (2s de pulse et 8s de pause) pendant 3 minutes. Dans les deux cas l'énergie totale déposée ne dépasse pas 10 kJ. Au cours de la sonication la solution est refroidie à l'aide d'un bloc métallique, préalablement incubée à -20°C. La protéine est ensuite reconstituée dans le lysat brut comme décrit dans [MM-3.3]. En fonction de la composition souhaitée d'agrégats [Fe-S], il est possible de purifier la protéine sous la forme apo et faire une reconstitution à la fin de la purification. Dans le cas d'une reconstitution dans le lysat, une première étape de purification – optionnelle – peut se faire par thermoprécipitation. Le lysat reconstitué est chauffé au bain-marie à 60°C pendant 30 min. Cette solution est ensuite transvasée dans des tubes étanches et ultracentrifugée à 165000 g pendant 45 minutes à 4°C dans le but d'éliminer, par sédimentation, les débris cellulaires et les protéines précipitées. Ce lysat clarifié est filtré sur une membrane de porosité de 20  $\mu\text{m}$ , puis il est injecté (0,5 mL/min) sur une colonne Resource Q, équilibrée avec 3 volumes de colonne (CV) de tampon A (100 mM Tris pH 7,6). L'élution se fait avec un gradient du tampon A vers le tampon B (100 mM Tris pH 7,6, 1 M NaCl). La température de la boîte à gants est maintenue à 25°C, tout le long de la purification. Lors de la première caractérisation de cette enzyme par Rubach *et al.* une étape de purification sur une colonne HiTrap<sup>TM</sup> Blue HP V = 5 mL résine Blue Sepharose (GE Healthcare) avait été rajouté<sup>4</sup>. Nous avons supprimé cette étape, car elle augmentait la pureté de l'échantillon seulement de façon marginal, et, en plus, elle diminuait fortement le rendement total.

Les fractions sont analysées par SDS-PAGE [MM-3.4.4] et celles contenant la protéine d'intérêt pure sont rassemblées et injectées sur une colonne de filtration sur gel HiLoad 16/60 Superdex 75 (GE Healthcare Life Sciences) équilibrée dans 50 mM Tris pH 7,6. Les fractions dessalées sont concentrées sur une cellule Amicon équipée d'une membrane

30 kDa (Millipore). La concentration en protéine est alors estimée (après une dilution au  $1/10^6$ ) par la méthode colorimétrique utilisant le rose bengal [MM-3.4.1]. Les teneurs en fer et en soufre sont déterminées par les méthodes de Fish et Beinert, respectivement [MM-3.4.2 et MM-3.4.3]. Les aliquotes contenant l'holoprotéine concentrée sont congelés et stockés dans de l'azote liquide.

### ***MM-3.2.2 Purification de HydE N-strep de Thermotoga maritima par chromatographie d'affinité***

La protéine a été exprimée à partir d'un stock glycérol de cellules BL21(DE3) RIL pT7-7 *NstrepTmhydE*. L'expression, la lyse de cellules et la reconstitution des centres [Fe-S] dans l'extrait ont été faites comme décrit dans [MM-3.3]. La seule modification est la composition du tampon dans lequel les cellules ont été suspendues avant la lyse par sonication ; 100 mM Tris, 150 mM NaCl, pH 8.0. Après la reconstitution, le lysat brut a été clarifié par ultracentrifugation à 165 000 g, 4°C pendant 30 minutes.

La purification a été faite sur une colonne de Strep-Tactine® (fournisseur IBA), préalablement équilibrée avec du tampon 100 mM Tris, 150 mM NaCl, 5 mM DTT, pH 8.0. Le lavage se fait avec 3 volume de colonne (CV) de tampon 100 mM Tris, 150 mM NaCl, pH 8.0. La protéine est éluée avec 100 mM Tris, 150 mM NaCl, pH 8.0 et 2,5 mM de déthiobiotine. Les fractions sont rassemblées et concentrées avec une cellule Amicon sur une membrane de 30 kDa. Le tampon est ensuite échangé grâce à une colonne GE Healthcare HiPrep 26/10 Desalting®, équilibrée dans 100 mM Tris, 150 mM NaCl, pH 8.0. L'holoprotéine est reconcentrée et les concentrations de protéine, de fer et de soufre sont déterminées comme décrit dans les sections MM-3.4.1 ; MM-3.4.2 et MM-3.4.3 respectivement. Les aliquotes sont congelées et entreposées dans de l'azote liquide.

### ***MM-3.2.3 Purification de HydE N-strep de Clostridium acetobutylicum***

L'expression de la protéine HydE *Ca* N-strep se fait dans des cellules *E. coli* BL21 Codonplus™ (DE3) RIL pRSF *Isc metK* Duet-1 contenant le plasmide pET-15b hydE *Ca* N-strep d'après le protocole de Swartz [MM-3.1.3]. Le culot est repris dans 50 mM Tris pH 7.5, 300 mM KCl et 10% glycérol. Après la lyse de cellules par sonication, la purification de HydE *Ca* N-strep se fait alors en une étape par chromatographie d'affinité sur une colonne de Strep-Tactine. Toutes les étapes se font en anaérobie stricte.

### ***MM-3.2.4 Purification de HydE N-strep ΔFeS de Thermotoga maritima***

L'expression de la protéine HydE *Tm* N-strep ΔFeS se fait dans les cellules *E. coli* BL21 Codonplus™ (DE3) RIL contenant le plasmide pT7-7 hydE *Tm* N-strep ΔFeS. Après une induction aérobie sur 3h à 37°C en milieu TB enrichi en Fe<sup>3+</sup> et L-cystéine les cellules sont culotées et les membranes sont cassées par sonication en boîte à gants. Les centres [Fe-S]

sont reconstitués dans l'extrait [MM-3.3.3] et une première étape de purification est faite par thermoprécipitation à 60°C pendant 20 min. Après l'ultracentrifugation (165000 g pendant 60 minutes à 4°) la protéine HydE *Tm* N-strep  $\Delta$ FeS reconstituée chimiquement est purifiée en une étape par chromatographie d'affinité sur une colonne de Strep-Tactine.

#### **MM-3.2.5 Purification de HydE N-strep $\Delta$ FeS Se de *Thermotoga maritima***

L'expression de la protéine HydE *Tm* N-strep  $\Delta$ FeS Se a été fait dans des cellules *E. coli* BL21 Codonplus<sup>TM</sup> (DE3) RIL contenant le plasmide pT7-7 hydE *Tm* N-strep  $\Delta$ FeS. L'induction se fait en aérobie sur 3h à 37°C en milieu TB. Les détails concernant l'étape de sonication de cellules et de chélation de cation de fer en aérobie, ainsi que la technique de purification de protéine, de reconstitution des centres [Fe-Se] et de dessalage de la protéine reconstituée sont résumées dans le paragraphe [MM-3.3.5].

#### **MM-3.2.6 Purification de HydE N-strep $\Delta$ FeS de *Thermotoga maritima* reconstituée *in vivo* en condition de respiration de fumarate**

La production d'un culot de cellules *E. coli* BL21 Codonplus<sup>TM</sup> (DE3) RIL pRSF *Isc metK* Duet-1 pT7-7 hydE *Tm* N-strep  $\Delta$ FeS exprimant la protéine HydE N-strep  $\Delta$ FeS de *Thermotoga maritima* reconstituée *in vivo* en condition de respiration du fumarate est décrite dans le paragraphe [MM-3.1.1]. Les cellules sont reprises dans 100 mM Tris pH 8, 50 mM NaCl (8 mL par gramme de cellules), puis elles sont lysées par sonication (35% d'amplitude, 2s de pulse et 8s de pause, total 3 min) et le lysat est clarifié par ultracentrifugation (100 000 g, 30 min, 4°C). La purification se fait par chromatographie d'affinité sur une colonne de Strep-Tactine, suivi d'une étape de gel filtration sur une colonne HiLoad 16/60 Superdex 75 (GE Healthcare Life Sciences). Toutes les étapes se font en anaérobie stricte.

#### **MM-3.2.7 Purification de HydE N-strep $\Delta$ FeS de *Thermotoga maritima* reconstituée *in vivo* en condition de fermentation**

La production de cet échantillon est similaire à la production de la protéine reconstituée *in vivo* décrite dans le paragraphe précédent. La seule différence constitue le métabolisme de cellules lors de l'expression de la protéine, comme mentionné dans le paragraphe [MM-3.1.2].

#### **MM-3.2.8 Purification de HydA1 N-strep de *Chlamydomonas reinhardtii***

L'expression de HydA1 *Cr* N-strep dans des cellules *E. coli* BL21 (DE3) pRSF *Isc metK* Duet-1 contenant le plasmide pET-22b hydA1 *Cr* N-strep se fait selon le protocole de Swartz [MM-3.1.3]. Le culot de cellules est repris dans 100 mM Tris pH 8.0, 300 mM NaCl, 5% glycérol. Après la lyse de cellules par sonication, le lysat est clarifié par ultracentrifugation et la protéine est ensuite purifiée à partir du surnageant par

chromatographie d'affinité sur une colonne de Strep-Tactine. Toutes les étapes se font en anaérobiose stricte.

### **MM-3.3      Reconstitution des centres FeS**

Au cours de nombreuses purifications de la protéine HydE nous avons observé que la qualité de reconstitution des centres [Fe-S] dépendait fortement de la méthode utilisée. Nos structures avec les différents types d'agrégats démontrent l'impact de la forme de l'agrégat sur le repliement de la protéine et par la suite sur son activité. Les paragraphes suivants contiennent les détails des différentes techniques de reconstitution des centres [Fe-S] utilisées au cours de mes travaux de thèse.

#### ***MM-3.3.1      Reconstitution chimique des apoprotéines purifiées***

Les apoprotéines purifiées sont reconstituées chimiquement à l'abri de l'oxygène dans une boîte à gants. La protéine est concentrée jusqu'à au moins 25  $\mu$ M dans du tampon Tris 50 mM pH 7,6. Nous avons remarqué qu'il était préférable de maintenir le pH au-dessus de la valeur du pKa du groupement thiol des résidus cystéine pour permettre sa déprotonation. La solution est réduite avec l'ajout de 5 mM de DTT et incubée pendant 5 minutes. Ensuite, on ajoute progressivement du  $\text{Fe}^{2+}$  sous forme d'une solution concentrée de Sel de Mohr  $[(\text{NH}_4)_2\text{Fe}(\text{SO}_4) \cdot 6\text{H}_2\text{O}]$ . La quantité de fer est ajustée pour correspondre à environ 125% de la quantité de fer théorique maximale contenue dans la protéine. Par exemple pour une protéine qui devrait contenir deux centres  $\text{Fe}_4\text{S}_4$  on ajoute 10 excès molaires de  $\text{Fe}^{2+}$ . Après une période d'incubation de 5 minutes, une quantité de  $\text{Na}_2\text{S}$ , équimolaire à celle du  $\text{Fe}^{2+}$ , est ajoutée goutte-à-goutte au mélange. L'ajout lent et progressive du soufre est essentiel pour favoriser la formation de centres [Fe-S] chélatés par la protéine, plutôt que des agrégats [Fe-S] dans la solution. La réaction de reconstitution se poursuit à 20°C pendant deux heures sous agitation modérée. L'échantillon est ensuite filtré sur une membrane de porosité de 20  $\mu$ m et dessalé à l'aide d'une colonne HiPrep Desalting V=50 mL, équilibrée dans du tampon Tris 50 mM pH 7.5. Cette étape d'échange de tampon permet d'éliminer les composés de faible poids moléculaire. La protéine chimiquement reconstituée à la fin de la purification est concentrée sur une cellule Amicon équipée d'une membrane 30 kDa (Millipore), aliquotée et congelée dans l'azote liquide.

#### ***MM-3.3.2      Reconstitution des apoprotéines purifiées à l'aide d'une cystéine désulfurase***

Alternativement, on peut utiliser comme source de soufre la L-Cystéine en présence d'une cystéine désulfurase, comme par exemple l'enzyme NifS d'*Azotobacter vinelandii*. La purification de la protéine NifS est détaillée dans la publication<sup>5</sup> de Zheng *et al.* 1993. Cependant, il est important de noter, que pendant l'expression de cette protéine une solution de Vitamines MEM (SIGMA) contenant de l'hydrochlorure de pyridoxal – un précurseur du PLP - a été ajouté au milieu de culture et, par conséquent, la protéine NifS est copurifiée avec

son cofacteur PLP, ce qui lui confère une couleur jaune. Pour la reconstitution à l'aide d'une cystéine désulfurase, une faible quantité de NifS pure, typiquement de l'ordre de 225 nmol pour 1 mmol de L-Cystéine, est ajoutée à la solution de protéine purifiée, contenant 5 mM de DTT. Ensuite la L-Cystéine et le fer sont ajoutés en alternance, deux équivalents molaires toutes les cinq minutes, pour arriver à la fin à un excès molaire de 10 par rapport à la concentration de l'apoprotéine. Le reste de la réaction de reconstitution est identique à celle décrite dans la section précédente. La concentration de NifS présente dans la solution est de l'ordre de 2,5 ‰ par rapport à la concentration de la protéine d'intérêt. Dans le cas de HydE, cette faible contamination de l'holoprotéine pure par NifS n'a pas gêné sa cristallisation.

### ***MM-3.3.3 Reconstitution dans l'extrait***

Dans le cas où l'apoprotéine pure présente une stabilité limitée, il est possible de procéder à une reconstitution avant l'étape de purification. Cette reconstitution dans l'extrait s'effectue juste après la lyse de cellules par sonication. Il est préférable d'ajuster le pH du lysat brut à l'aide d'une solution de KOH 2N pour arriver à un pH > 7.5. Le lysat brut est incubé pendant 5 minutes en présence de 5 mM de DTT, puis on ajoute progressivement une solution concentrée de sel de Mohr pour arriver à 1 mM de Fe<sup>2+</sup> final. Au bout de 5 minutes une solution concentrée de Na<sub>2</sub>S est ajoutée goutte-à-goutte jusqu'à atteindre 1 mM de S<sup>2-</sup> final. La réaction de reconstitution se poursuit à 20°C pendant une heure sous agitation modérée.

L'introduction de cette méthode dans le protocole de purification peut s'avérer bénéfique pour une obtention rapide d'holoenzyme pure. En effet, les métalloprotéines reconstituées présentent souvent une stabilité accrue par rapport à leur forme apo. Ainsi la protéine peut mieux supporter les étapes de purification et de concentration, ce qui a pour effet l'augmentation du rendement de la purification. Dans le cas d'une protéine dont le gène a été issu d'un organisme thermophile, on peut alors rajouter une étape de purification par thermoprécipitation. Les essais de thermoprécipitation réalisés par Elodie Marinoni avant mon arrivée au laboratoire avait montré que *TmHydE* résistait à 60-65°C pendant 20 min. D'autre part, le fort excès de fer ajouté au lysat brut a pour conséquence la précipitation de certaines protéines de la cellule. Cela fait diminuer la quantité de protéines contaminants dans le lysat, ce qui est particulièrement utile dans le cas d'une purification par chromatographie d'échange d'ions. Cependant, cet excès de fer dans le lysat peut aussi entraîner une précipitation de la protéine d'intérêt. Dans ce cas une reconstitution plus douce s'impose.

### ***MM-3.3.4 Reconstitution in vivo***

Avant une reconstitution chimique, les métalloprotéines purifiées présentent souvent des quantités sous-stœchiométriques de centres Fe-S. Ceci est sans doute dû à la sensibilité de ces centres à l'oxygène, à la déplétion du fer du milieu de culture et à une trop forte concentration de la protéine surexprimée par rapport aux protéines des machineries d'assemblage de centre Fe-S présentes dans la cellule. Dans le but d'obtenir la protéine d'intérêt avec une quantité stœchiométrique d'agrégat Fe-S nous avons utilisé une méthode

d'expression anaérobie, inspirée des travaux de James R. Swartz<sup>1,2,6</sup>. Pour palier au problème de déséquilibre entre la concentration de la protéine surexprimée et celle des protéines de machineries d'assemblage de centre Fe-S l'équipe de Swartz a utilisé une souche d'expression *E. coli* dont le gène codant pour la protéine IscR a été remplacée par un gène codant pour la résistance à la kanamycine<sup>2</sup>. La protéine IscR est un facteur de transcription qui régule, parmi d'autres, l'expression de l'opéron *isc*<sup>7</sup>. Quand cette protéine est dépourvue de centres FeS, la synthèse des machineries *isc* et *suf* est activée, ce qui a pour effet l'accumulation de centre Fe-S dans la cellule<sup>8</sup>.

### **MM-3.3.5     *Reconstitution au sélénium***

Afin d'obtenir une complète substitution du soufre par le sélénium dans les agrégats [Fe-S] il est indispensable d'enlever, dans un premier temps, tous les centres [Fe-S] résiduels. Pour cela, la préparation d'un tel échantillon se passe à l'extérieur de la boîte à gant jusqu'à l'étape de purification par chromatographie d'affinité. Les cellules sont resuspendues dans du tampon 100 mM Tris pH 8.0, 50 mM NaCl, 5 mM EDTA, 5 mM Tiron (4 mL par gramme de cellules). L'EDTA est un chélateur d'ions divalents, tandis que le Tiron présente une affinité élevée pour le Fe<sup>3+</sup>. Leur présence dans le tampon de lyse, ainsi que celle d'oxygène dissout devrait fortement déstabiliser les centres [Fe-S] liés à la protéine. Alternativement une combinaison de DTH-EDTA ou ferricyanure-Tiron peut être utilisée pour réduire ou oxyder le fer, respectivement, et ensuite le séquestrer par un chélateur plus approprié. On peut également considérer l'ajout de DTT pour la réduction d'éventuel persulfures [Kennedy JBC 1988]. Avant la lyse de cellules une tablette d'inhibiteurs de protéases Complete EDTA-free (Roche) est ajoutée par cinq grammes de culot. Après l'étape de sonication à l'extérieur de la boîte à gants, dont l'équivalent anaérobie est décrit dans [MM-3.2.1], le lysat est centrifugé à 27 000 g, 40 min, 4°C et transféré en BAG. La purification se fait sur une résine Streptactine II [MM-3.2.2], l'éluion se fait avec les tampons dépourvus de chélateurs de métaux. La protéine pure est concentrée jusqu'à 25 µM et les agrégats Fe-Se sont reconstitués avec du sel de Mohr, DTT et Se-Cystine, à l'aide de NifS, comme décrit dans [MM-3.3.2]. La protéine est ensuite filtrée et dessalée sur une colonne HiPrep Desalting, après quoi, l'holoprotéine pure abritant les centres [Fe-Se] est concentrée et les aliquotes congelées et stockées dans de l'azote liquide.

### **MM-3.4     *Analyses biochimiques***

#### **MM-3.4.1     *Dosage protéique par la méthode du rose Bengal***

Dans les travaux présentés dans ce manuscrit, nous avons estimé la concentration protéique de nos échantillons en utilisant la méthode de dosage colorimétrique au rose Bengal<sup>9</sup>. Par rapport à la méthode la plus répandue – le dosage de Bradford<sup>10</sup>, le dosage au rose Bengal a l'avantage d'être moins sensible à la teneur en métaux dans l'échantillon et il est donc plus adapté au dosage de protéines contenant des agrégats Fe-S.

Une solution stock de colorant rose Bengal à 10 mg/mL est préparée dans de l'eau. Une série de volumes croissant d'échantillon à analyser est diluée dans 2 mL de tampon  $\text{KH}_2\text{PO}_4$  10 mM pH 6,8. 100  $\mu\text{L}$  de solution stock de rose Bengal sont ajoutés dans chaque tube. Entre deux et cinq minutes plus tard, 100  $\mu\text{L}$  d'acide acétique glacial à 50 % sont ajoutés, le tube est passé au vortex et l'absorbance à 560 nm est lue immédiatement après, car l'échantillon précipite au cours du temps. Le blanc se fait avec une solution ne contenant pas d'échantillon protéique. Après l'acidification du milieu le coefficient d'extinction molaire du colorant libre chute de façon considérable et donc l'absorbance est proportionnelle à la concentration de protéine dans le tube. Une gamme étalon, avec des concentrations en protéine connues est nécessaire pour pouvoir relier une absorbance donnée à une valeur de concentration précise. Cette gamme étalon est typiquement préparée dans 2 mL de tampon  $\text{KH}_2\text{PO}_4$  10 mM pH 6,8 avec des concentrations de Bovine Sérum Albumine (BSA) allant de 5 à 60  $\mu\text{g/mL}$ . En divisant le coefficient de pente de l'absorbance en fonction du volume de l'échantillon à analyser par le coefficient de pente de l'absorbance en fonction de la concentration de BSA et en tenant compte du facteur de dilution nous obtenons la concentration de l'échantillon en mg/mL.

Au cours du temps, plusieurs gammes étalons ont été produites, avec une bonne reproductibilité. Ainsi, un coefficient de pente moyen a été calculé avec une valeur de 0,025 mL/ $\mu\text{g}$  [Figure MM-109] et c'est celui-ci qui a été utilisé pour le calcul de concentration protéique dans les échantillons.

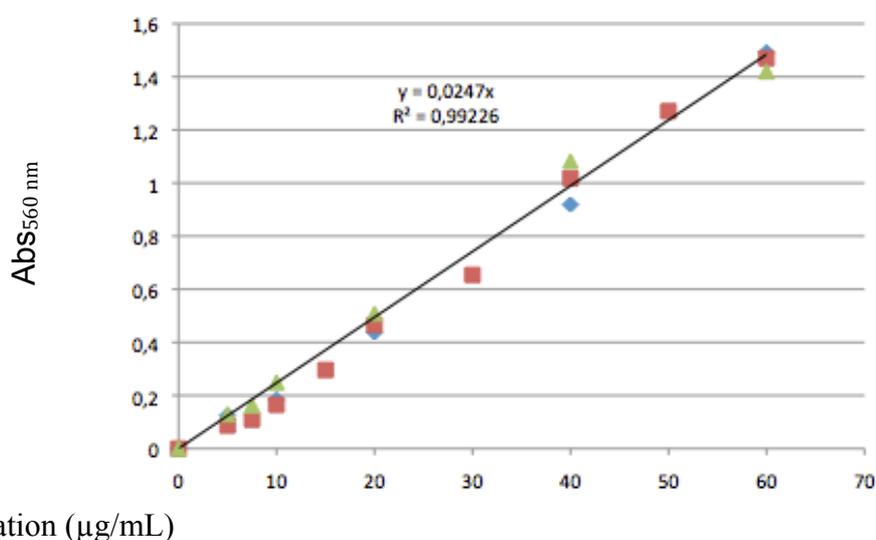


Figure MM-109 : Courbe étalon pour le dosage au rose Bengal (n=3)

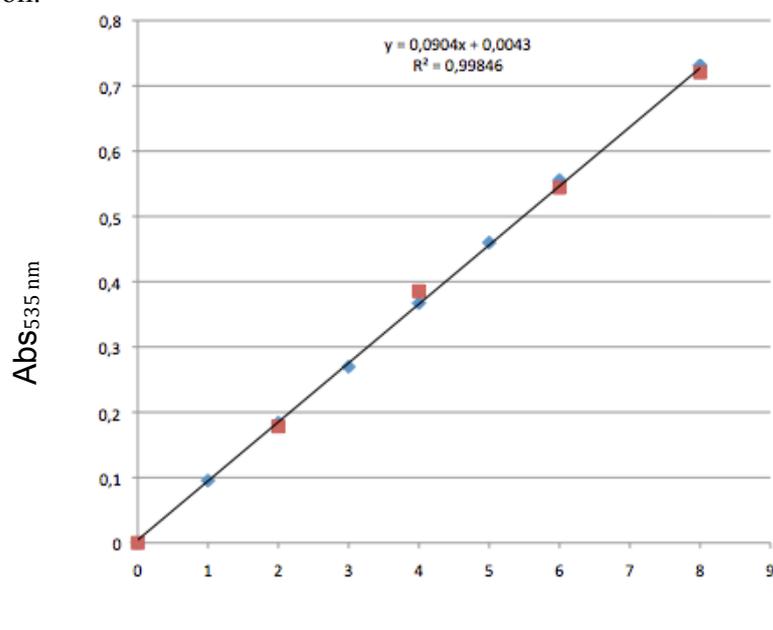
### MM-3.4.2 Dosage du fer

La teneur en Fer des échantillons protéiques est déterminée par la méthode de Fish<sup>11</sup>. Une gamme étalon est préparée à partir d'une solution de sel de Mohr  $(\text{NH}_4)_2(\text{SO}_4)_2$  à 10 mM dans du HCl 1N, diluée au 1/20<sup>e</sup> dans de l'eau. Des standards contenant respectivement 0, 2, 4, 6, 8 et 10 nmoles de fer sont préparés dans un volume final de 65  $\mu\text{L}$ . Dans le même volume, la concentration de l'échantillon protéique est ajustée de façon à ce que la quantité de fer dans l'échantillon soit comprise dans la gamme étalon. L'ensemble des échantillons et des standards est dénaturé par incubation avec 45  $\mu\text{L}$  d'acide perchlorique à 1M pendant 1 h. Le

précipité de protéines est éliminé par une étape de centrifugation à 13 000 rpm pendant 5 minutes. Après cela, 90  $\mu\text{L}$  sont prélevés et mélangés dans l'ordre avec :

- 72  $\mu\text{L}$  de bathophénaントroline (1.7 mg/mL dans l'eau)
- 36  $\mu\text{L}$  d'ascorbate de sodium (38 mg/mL dans l'eau) préparée le jour du dosage
- 27  $\mu\text{L}$  d'acétate d'ammonium (solution saturée diluée aux 1/3).

Après 30 minutes d'incubation à température ambiante les tubes sont à nouveau centrifugés à 13 000 rpm pendant 5 minutes. L'ascorbate de sodium va réduire le  $\text{Fe}^{3+}$  issu de la dégradation des centres FeS et la bathophénaントroline va former un complexe coloré avec le  $\text{Fe}^{2+}$ . L'absorbance à 535 nm [Figure MM-110] est mesurée sur un spectromètre Varian 50 Bio dans des cuves en quartz en ayant fait au préalable le zéro sur le standard ne contenant pas de Fer. Connaissant la concentration de la protéine, préalablement déterminée par le dosage de rose Bengal, la courbe étalon permet d'obtenir la concentration en fer moyenne dans l'échantillon.



*Fer (nmole)*

**Figure MM-110 :** Courbe étalon pour le dosage du fer (n=2)

### **MM-3.4.3 Dosage du soufre**

Les sulfures et le  $\text{S}^0$ , contenus dans la protéine, sont dosés par la méthode colorimétrique de Beinert<sup>12</sup>. Une gamme étalon est préparée à partir d'une solution de sulfure de sodium ( $\text{Na}_2\text{S}$ ) à 200  $\mu\text{M}$ . Des standards entre 0 et 20 nmol de sulfure sont préparés dans un volume d'eau de 100  $\mu\text{L}$ . La protéine est diluée de la même manière que pour le dosage de Fe, c'est-à-dire pour que la quantité de soufre contenue dans les 100  $\mu\text{L}$  d'échantillon à analyser soit comprise dans la gamme étalon. Tous les tubes sont d'abord traités avec 300  $\mu\text{L}$  d'acétate de zinc à 1% (soit une solution à 100 mg/mL diluée au 1/10<sup>e</sup>) suivi immédiatement de 15  $\mu\text{L}$  de soude à 3 M. L'acétate de zinc va dénaturer la protéine et induire une libération

du soufre. L'ajout de soude va provoquer la précipitation immédiate d'hydroxyde de zinc et en même temps, l'augmentation de pH va empêcher la perte de soufre sous forme de H<sub>2</sub>S. Le mélange est incubé pendant 5 à 10 minutes à température ambiante. Ensuite on rajoute dans chaque tube simultanément 75 µL de diméthyl phénylène diamine (DMPD) 0,01% (soit une solution à 1 mg/mL dans HCl 5M) et 16 µL d'une solution de chlorure de fer (FeCl<sub>3</sub>) à 23 mM (soit 37,5 mg dans 10 mL HCl 1,2 M). La formation lente du bleu de méthylène se fait par réaction des sulfures avec le DMPD en milieu oxydant. Les standards sont incubés pendant 30 minutes à température ambiante, tandis que les échantillons à analyser sont incubés pendant 3h à 4°C. Après une étape de centrifugation à 10 000 rpm pendant 5 minutes, on mesure l'absorbance à 670 nm. La quantité de soufre par monomère de protéine est obtenue en rapportant l'absorbance à celle de la droite standard.

#### ***MM-3.4.4 Electrophorèse sur gel de polyacrylamide en présence de dodécylsulfate de sodium***

Au cours de différentes étapes de préparation de l'échantillon la composition des fractions protéiques a été suivie par électrophorèse sur gel de polyacrylamide en conditions dénaturantes (acronyme anglais SDS-PAGE). Cette technique permet de séparer des protéines selon leur poids moléculaire. De faibles volumes d'échantillons à analyser (2-5 µL) sont mélangés avec 5 µL de tampon de charge 5x, contenant un détergent fort - le dodécylsulfate de sodium (SDS) et du β-mercaptoéthanol. Ce mélange est ensuite complété avec de l'eau pour arriver à 25 µL final. Les échantillons ainsi préparés sont ensuite chauffés à 95°C pendant une dizaine de minutes, pour permettre la dénaturation de protéines, avant d'être déposés sur le gel de polyacrylamide. Le rôle du β-mercaptoéthanol est de réduire d'éventuels ponts disulfures, tandis que le SDS empêche le repliement de la protéine après l'étape de dénaturation et lui confère une charge nette négative. La séparation en fonction du poids des molécules s'effectue grâce à la migration de protéines chargées négativement dans un champ électrique à travers une matrice de copolymère d'acrylamide et de bis-acrylamide. Les protéines ayant un poids moléculaire plus faible migreront plus rapidement. Pour visualiser les bandes de migration sur le gel nous avons utilisé du bleu de Coomassie. Ce colorant s'adsorbe aux protéines via interactions hydrophobes avec les résidus aromatiques et interactions ioniques avec les résidus basiques. D'autres méthodes de révélation des bandes de migration utilisent comme colorant le nitrate d'argent, du zinc ou des sondes fluorescentes.

#### **MM-3.5 Préparations de dérivés de thiazolidines et de thiazinanes**

Le [Table MM-4](#) résume la préparation de ligands contenant un cycle thiazolidine ou thiazinane. Les solutions stock à 250 mM de ces différentes composées ont été préparées en mélangeant une solution aqueuse d'une *molécule I* contenant des groupements thiol et amine primaire avec une poudre d'une *molécule II* contenant un groupement carbonyle ou cétone. Ceci aboutit à la formation de racémiques 2R et 2S de dérivés thiazolidine ou thiazinane correspondants. Le pH final de toutes les solutions stock est ajusté à des valeurs entre 7 et 9 avec une solution de NaOH/HCl 2N. Les solutions stock ont été préparées, de préférence, à

l'intérieur de la boîte à gants avec de l'eau dégazée. Dans le cas contraire, elles ont été bullées avec de l'azote une fois entrées en boîte à gants. En fonction de la stabilité des composés en milieu aqueux, ils ont été soit préparés le jour de l'expérience ou bien dans le cas de certains composés thiazolidines et thiazinanes stables, ils ont été stockée en boîte à gants à 20°C.

acronyme du composé	<i>molécule I</i> (-S <sup>-</sup> & -NH <sub>2</sub> )	<i>molécule II</i> (RCOH/RCOR')	composé cyclique correspondant
(2 <i>RS</i> ,4 <i>R</i> )-MeTDA	L-cystéine	pyruvate	(2 <i>R</i> , <i>S</i> -4 <i>R</i> )-2-méthyle-1,3-thiazolidine-2,4-dicarboxylate
(2 <i>RS</i> ,4 <i>R</i> )-TDA	L-cystéine	glyoxylate	(2 <i>R</i> , <i>S</i> -4 <i>R</i> )-1,3-thiazolidine-2,4-dicarboxylate
(4 <i>R</i> )-TDA	L-cystéine	formaldéhyde	(4 <i>R</i> )-1,3-thiazolidine-4-carboxylate
(2 <i>RS</i> ,4 <i>R</i> )-DMeTDA	L-pénicillamine	glyoxylate	(2 <i>R</i> , <i>S</i> -4 <i>R</i> )-5,5-diméthyle-1,3-thiazolidine-2,4-dicarboxylate
(2 <i>RS</i> ,4 <i>R</i> )-TMeTDA	L-pénicillamine	pyruvate	(2 <i>R</i> , <i>S</i> -4 <i>R</i> )-2,5,5-triméthyle-1,3-thiazolidine-2,4-dicarboxylate
(2 <i>RS</i> ,4 <i>S</i> )-MeTDA	D-cystéine	pyruvate	(2 <i>R</i> , <i>S</i> -4 <i>S</i> )-2-méthyle-1,3-thiazolidine-2,4-dicarboxylate
(2 <i>RS</i> ,4 <i>S</i> )-TDA	D-cystéine	glyoxylate	(2 <i>R</i> , <i>S</i> -4 <i>S</i> )-1,3-thiazolidine-2,4-dicarboxylate
(4 <i>S</i> )-TDA	D-cystéine	formaldéhyde	(4 <i>S</i> )-1,3-thiazolidine-4-carboxylate
(2 <i>RS</i> ,4 <i>RS</i> )-MeZDA	DL-homocystéine	pyruvate	(2 <i>R</i> , <i>S</i> -4 <i>R</i> , <i>S</i> )-2-méthyle-1,3-thiazinane-2,4-dicarboxylate
(2 <i>RS</i> ,4 <i>RS</i> )-ZDA	DL-homocystéine	glyoxylate	(2 <i>R</i> , <i>S</i> -4 <i>R</i> , <i>S</i> )-1,3-thiazinane-2,4-dicarboxylate
(2 <i>RS</i> ,4 <i>R</i> )-MMTDA	L-cystéine	mercapto-pyruvate	(2 <i>R</i> , <i>S</i> -4 <i>R</i> )-2-mercaptométhyle-1,3-thiazolidine-2,4-dicarboxylate
(2 <i>RS</i> ,4 <i>R</i> )-BrMTDA	L-cystéine	bromo-pyruvate	(2 <i>R</i> , <i>S</i> -4 <i>R</i> )-2-bromométhyle-1,3-thiazolidine-2,4-dicarboxylate
(2 <i>RS</i> ,4 <i>R</i> )-AcTDA	L-cystéine	oxaloacétate	(2 <i>R</i> , <i>S</i> -4 <i>R</i> )-2-acétate-1,3-thiazolidine-2,4-dicarboxylate
(2 <i>RS</i> ,4 <i>R</i> )-PrTDA	L-cystéine	oxoglutarate	(2 <i>R</i> , <i>S</i> -4 <i>R</i> )-2-propionate-1,3-thiazolidine-2,4-dicarboxylate

**Table MM-4 :** Préparation de ligands contenant un cycle thiazolidine ou thiazinane

Les solutions stock de L-Cystine (Sigma) ou de Séléno-L-Cystine (Sigma) ont été préparées en solubilisant les poudres de ces composés en déprotonant leur groupements amino par titration avec une solution de soude à 50%. L'analogue sélénié du composé (2*RS*,4*R*)-MeTDA a été préparé à partir d'une solution aqueuse de Séléno-L-Cystine. Le pont di-séléniure est d'abord réduit par l'addition de 5 équivalents de DTT. Le composé (2*RS*,4*R*)-MeSeTDA est alors formé par l'ajout de 2 équivalents de poudre de pyruvate de sodium.

## MM-3.6 Essais enzymatiques d'activité hydrogénase

### MM-3.6.1 Préparation de molécules nécessaires à la biosynthèse dans l'extrait du sous-agrégat à 2Fe

Le [Table MM-5](#) ci-dessous résume brièvement les protocoles de préparation de solutions stock de molécules utilisées dans une expérience standard de maturation de l'hydrogénase à [FeFe] dans l'extrait, telle que décrite par James R. Swartz<sup>13</sup>. Les composés utilisés lors des testes de maturation et d'activité hydrogénase ont été préparés à l'intérieur de la boîte à gants, ensuite ils ont été aliquotés en petits volumes et congelés et stockés dans de l'azote liquide. Ainsi, pour les testes d'activité, une aliquote de chaque molécule est décongelé et utilisé qu'une seule fois. S'il n'est pas précisé autrement, le pH final des solutions stock est ajusté à des valeurs entre 7.0 et 9.0 avec une solution de NaOH/HCl 2N.

molécule	préparation	concentr. finale	stockage
HCl 1N	2020 µL d'acide chlorhydrique à 37% dilué dans de l'eau, 25 mL de volume final	1M	BAG 25°C
KOH 2N	2.8 g de hydroxyde de potassium dissout dans 25 mL d'eau	2 M	BAG 25°C
HEPES 20x	23,8 g d'acide 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique dissout dans 100 mL d'eau. pH est ajusté à 8.2 avec des pastilles de KOH et la solution est filtrée	1 M	BAG 25°C
KCl 20x	7.46 g de chlorure de potassium dissout dans 100 mL d'eau, la solution est filtrée	1 M	BAG 25°C
Fe <sup>2+</sup>	98.04 mg de Fe(NH <sub>4</sub> ) <sub>2</sub> (SO <sub>4</sub> ) <sub>2</sub> dissout dans 1000 µL d'eau. 1N HCl est ajouté pour une concentration finale de 0.075 N	250 mM	N <sub>2</sub> liq
S <sup>2-</sup>	39.02 mg de sulfure de sodium dissout dans 10 mL d'eau	50 mM	N <sub>2</sub> liq
PLP	61.8 mg de phosphate de pyridoxal dissout dans 1ml d'eau. le pH est ajusté a 6.5-7.5 avec KOH 2N.	250 mM	N <sub>2</sub> liq
SAM	solution commerciale (New England Biolabs)	32 mM	N <sub>2</sub> liq
GTP	261.6 mg de guanosine-5'-triphosphate dissout dans 1000 µL d'eau. le pH est ajusté à 6.5-7.5 avec KOH 2N.	500 mM	N <sub>2</sub> liq
DTT	77.1 mg de dithiothréitol dissout dans 1000 µL d'eau	500 mM	N <sub>2</sub> liq
L-Cystéine	30.3 mg de L-cystéine dissout dans 1000 µL d'eau	250 mM	N <sub>2</sub> liq
L-Tyrosine	45.3 mg de L-tyrosine de dissout dans 1000 µL d'eau	250 mM	N <sub>2</sub> liq
DTH	17.4 mg de dithionite de sodium dissout dans 400 µL d'eau le jour de l'expérience	250 mM	frais
H <sub>2</sub> SO <sub>4</sub> 1N	Faire la dilution appropriée	0,5 M	BAG 25°C

**Table MM-5 : Protocoles de préparation de solutions stock de molécules utilisées dans une expérience standard de maturation de l'hydrogénase à [FeFe] dans l'extrait**

## **MM-3.6.2 Préparation des lysats cellulaires**

### **MM-3.6.2.1 Préparation des lysats cellulaires anaérobies clarifiés pour les expériences de maturation d'hydrogénase dans l'extrait**

Le protocole général d'expression hétérologue en condition d'anaérobie et sous fermentation acide mixte de métalloprotéines sensibles à l'oxygène est détaillé dans le paragraphe [MM-3.1.2]. Après l'expression, la culture est transférée dans des pots étanches et les cellules sont récupérées par deux centrifugations successives (à 7 800 g, 4°C, 20 min et 3 200 g, 4°C, 10 min). 4 mL de tampon de lyse (100 mM HEPES pH 8.2, 50 mM KCl) sont utilisés pour resuspendre 1 g de culot et une tablette d'inhibiteurs de protéases Complete EDTA-free est rajoutée par 5 g de cellules. La lyse se fait par sonication (35% d'amplitude pendant 3 minutes - 2s de pulse et 8s de pause). Si nécessaire, le pH du lysat brut est ajusté avec du tampon HEPES pH 8.2 20x pour se situer entre 7.5 et 7.8. Le lysat brut est alors transféré dans des tubes étanches et il est clarifié par l'ultracentrifugation à 165 000 g, 4°C pendant 30 minutes. Le lysat clarifié est aliquoté par 500 µL dans des tubes Nunc de 1.8 mL, qui sont ensuite rapidement congelés et stockés dans de l'azote liquide.

### **MM-3.6.2.2 Préparation de lysats cellulaires anaérobies dialysés**

Les composés de faible poids moléculaire, naturellement présent dans le lysat, ont été enlevés par deux bains de dialyse successifs en boîte à gants. 5 ml de lysat clarifié sont d'abord versés dans un boudin de dialyse (MWCO=3kDa). Le boudin est incubé sous agitation à 20°C dans 500 mL de tampon anaérobie 100 mM HEPES pH 8.2, 50 mM KCl. Au bout de trois heures le boudin est transféré dans un second bain de dialyse, contenant 500 mL de tampon neuf et y est incubé pendant encore quatre heures. Après la dialyse, le lysat est concentré par ultrafiltration sur une membrane de 30kDa pour revenir au volume initial de 5 mL. Les lysats dialysés sont ensuite aliquotés par 1500 µL, puis congelés et stockés dans de l'azote liquide. Les lysats contenant les différentes maturases ont été dialysés dans des bains de dialyse séparés.

### **MM-3.6.2.3 Fractionnement des lysats clarifiés en fonction de la taille des molécules**

Pour séparer les fractions protéiques en fonction de la taille des protéines, le lysat dialysé a été fractionné par des étapes d'ultrafiltrations successives utilisant des membranes avec la taille des pores de plus en plus petites. Dans un premier temps, 3ml de lysat dialysé a été filtré avec une membrane de MWCO=100 kDa. Une partie du filtrat a été aliquoté, congelé et stockée dans l'azote, tandis que le reste du filtrat a été refiltré avec une membrane de MWCO=30 kDa. De nouveau, une partie du filtrat est aliquoté et congelé et le reste est encore une fois filtré, cette fois-ci sur une membrane de MWCO=10 kDa. Ce dernier filtrat est alors lui aussi aliquoté puis rapidement congelé et stocké dans de l'azote liquide.

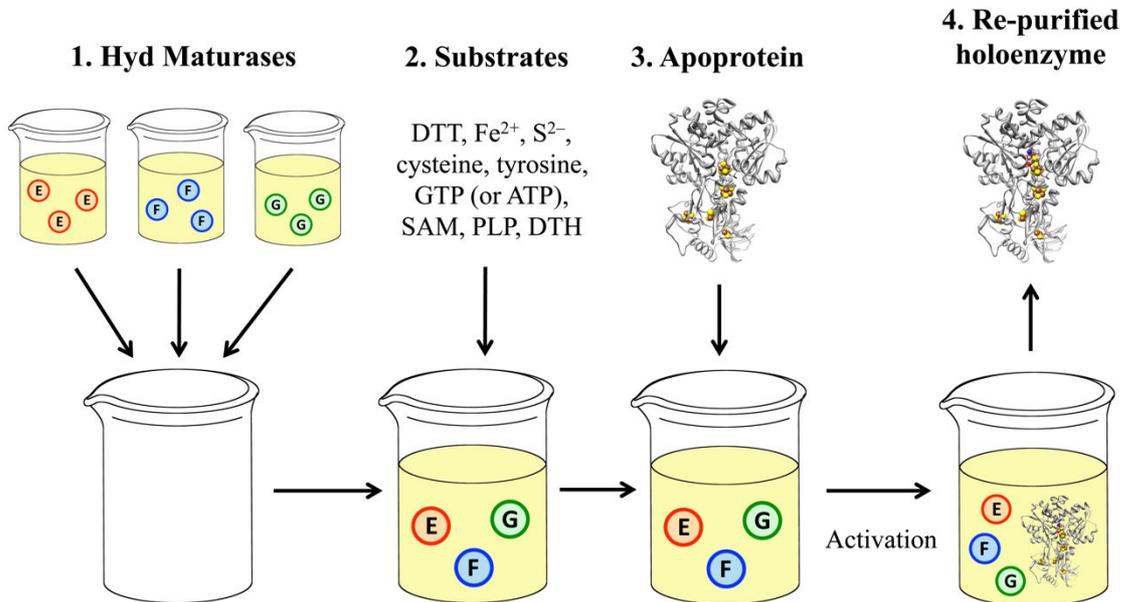
### MM-3.6.3 Protocole d'activation de l'hydrogénase à [FeFe] par maturation *in vitro* du sous-agrégat à 2 Fe

La maturation *in vitro* se fait en boîte à gants sous une faible agitation à l'aide d'un barreau aimanté. Les lysats contenant les maturases sont d'abord décongelés en boîte à gants. Les aliquotes avec les substrats et d'autres composés chimiques, indispensables pour la maturation de l'hydrogénase, sont transférés en boîte à gants et décongelés à 4°C peu de temps avant leur utilisation. Une solution stock de DTH à 250 mM est préparée en solubilisant 17.4 mg de poudre de dithionite de sodium, stocké à l'abri d'oxygène, dans 400 µL d'eau anaérobie. Pour chaque millilitre de l'apoenzyme HydA1 Cr N-strep purifié à 8 mg.L<sup>-1</sup> sont ajoutés 0,5 mL de Isc-MetK<sup>lys</sup>, 1 mL de HydE<sup>lys</sup>, 1 mL de HydF<sup>lys</sup> et 5 mL de HydG<sup>lys</sup> [Figure MM-111]. Le mélange est ensuite réduit avec 100 µL de DTT. Après 5 minutes, le Fe<sup>2+</sup> est ajouté et la solution est incubée pendant 5 à 10 minutes. On ajoute alors le S<sup>2-</sup> en quatre fois 200 µL ; chaque injection est espacée de 5 minutes de la précédente. Entre temps on prépare un mélange de solutions stock de GTP/SAM/PLP/L-Tyrosine avec un rapport de 2/4,688/0,2/0,4. Après la reconstitution des centres [FeS] on ajoute 400 µL de L-Cystéine suivi de 2,5 mL du tampon HEPES 20x et de 7,288 mL du mélange GTP/SAM/PLP/L-Tyrosine. La catalyse enzymatique des protéines à radical SAM est alors déclenchée par l'ajout de 400 µL de solution de DTH, fraîchement préparée. On ajoute alors 1 mL de l'apoenzyme HydA1 Cr N-strep. La maturation *in vitro* se poursuit alors à une température constante de 25°C pendant 16 à 18h. Le Table MM-6 résume les concentrations et les volumes de tous les composants du système de maturation.

Composant	Concentration initiale	Concentration finale	Volume ajouté
HydE <sup>lys</sup>	1x	0.10x	5 mL
HydF <sup>lys</sup>	1x	0.10x	5 mL
HydG <sup>lys</sup>	1x	0.50x	25 mL
Isc-MetK <sup>lys</sup>	1x	0.05x	2,5 mL
DTT	500 mM	1 mM	100 µL
Fe <sup>2+</sup>	250 mM	1.6 mM	320 µL
S <sup>2-</sup>	50 mM	0.8 mM	800 µL
L-Cystéine	250 mM	2 mM	400 µL
HEPES 20x	1000 mM	50 mM	2,5 mL
GTP	500 mM	20 mM	2 mL
SAM	32 mM	3 mM	4,688 mL
PLP	250 mM	1 mM	200 µL
L-Tyrosine	250 mM	2 mM	400 µL
apo HydA1 Cr N-strep purifiée	180 µM	3,6 µM	1 mL
HEPES	20x	1x	2,5 mL
DTH	250 mM	2 mM	400 µL

Table MM-6 : Composition du mélange réactionnel pour la maturation de l'hydrogénase à [FeFe]

Lors des testes de maturation en présence de potentiels substrats de HydE, le mélange de maturation complet est dispatché, juste après l'ajout de l'hydrogénase, dans des tubes Eppendorf par un volume de 100  $\mu$ L. Dans chaque tube on injecte alors 5  $\mu$ L de solution stock de ligands [Table MM-4] dont on veut tester l'impact sur la maturation de l'hydrogénase.



**Figure MM-111 :** Préparation du milieu réactionnel acellulaire pour la maturation de l'hydrogénase

(1) Le milieu réactionnel est préparé en mélangeant les différents lysats contenant les maturases HydE, HydF et HydG. (2) Les substrats de faibles poids moléculaires sont alors ajoutés au mélange réactionnel, (3) suivi par l'hydrogénase apo. (4) Après l'incubation du milieu réactionnel en condition anaérobie pendant 12–18 heures, l'hydrogénase active peut être repurifiée du mélange et caractérisée. (Figure adaptée de Kuchenreuther *et al.* 2012)

### MM-3.6.4 Mesures d'activité hydrogénase

#### MM-3.6.4.1 Principe de mesures d'activité hydrogénase par dosage colorimétrique du méthylviologène réduit

L'avancement de la maturation de l'hydrogénase est déterminé en mesurant l'activité hydrogénase du mélange de maturation *in vitro*. Cette expérience nécessite une arrivée de gaz dihydrogène en boîte à gants, soit sous forme d'une rampe à gaz ou sous forme d'une bouteille de dihydrogène moléculaire mise à l'intérieur de la boîte à gants. La mesure d'activité hydrogénase s'effectue en boîte à gants Jacomex équipée d'un porte cuvette optique, reliée par une fibre optique à un spectromètre équipé d'une lampe à arc de deutérium de gamme spectrale entre 180 et 450 nm et d'une seconde lampe à incandescence halogène de gamme spectrale supérieure à 450 nm. Un faible volume d'extrait contenant l'hydrogénase mature est dilué dans une solution de méthylviologène saturée en dihydrogène. Le méthylviologène est un composé électrochrome, c'est-à-dire il change de couleur suite au changement de son état

redox. Transparent quand oxydé, il présente une couleur bleue intense avec une bande d'absorbance large, centrée sur 600 nm, après sa réduction avec des électrons, générés par l'hydrogénase à partir de H<sub>2</sub>. L'absorbance peut être donc directement reliée à la consommation de H<sub>2</sub> par l'hydrogénase et donc à son activité spécifique.

#### MM-3.6.4.2 Préparation de cuvettes avec une solution de méthylviologène saturée en dihydrogène

La veille de l'expérience une série de cuvettes en plastique avec les bouchons en caoutchouc et une série de aiguilles sont dégazés dans le sas et rentrés en boîte à gants. 100 mL de tampon 50 mM Tris pH 8.5 sont dégazés par bullage de l'azote. Le jour de l'expérience une solution de dichlorure de méthyl viologène hydrate (abréviation MV) à 16 mg.mL<sup>-1</sup> est préparée dans de l'eau anaérobie. Dans chaque cuvette sont ajoutés 1 mL de tampon 50 mM Tris pH 8.5 et 20 µL de MV et la cuvette est fermée hermétiquement avec un bouchon. Le bouchon est ensuite percé avec une aiguille longue, qui descend jusqu'au fond de la cuvette. Une autre aiguille, plus courte, est insérée dans le bouchon de façon à ce que la pointe de l'aiguille sort à peine de l'autre côté du bouchon. La première aiguille sert du port d'entrée du dihydrogène, tandis que la seconde aiguille fait l'office de l'échappement. Afin d'éviter d'éventuelles projection de MV et la contamination de la boîte à gant, un morceau de papier absorbant est inséré dans la tête de l'aiguille d'échappement. Les cuvettes sont alors bullées avec du dihydrogène pendant 10 min pour permettre la saturation en gaz de la solution. Avant la fermeture de l'arrivée du H<sub>2</sub> on retire d'abord l'aiguille d'échappement et ensuite l'aiguille d'arrivée pour s'assurer que le gaz qui remplit l'espace entre le liquide et le bouchon n'ait pas le temps de s'échanger avec l'atmosphère de la boîte à gants.

#### MM-3.6.4.3 Mesures d'activité hydrogénase par dosage colorimétrique du méthylviologène réduit

Entre 2 et 10 µL d'extrait de maturation sont dilués dans un tube Eppendorf contenant 200 µL de tampon 50 mM Tris pH 8.5. Le tube est ensuite centrifugé sur une petite centrifugeuse de paillasse pendant 15s, dans le but de faire sédimenter d'éventuel précipité de protéines. Une seringue Hamilton d'un volume total de 100 µL est alors utilisée pour prélever 100 µL de surnageant et pour l'injecter à travers le bouchon dans la cuvette saturée en dihydrogène. Le tube est rapidement agité pour homogénéiser la solution. L'absorbance à 604 et celle à 578 nm sont suivies dans le temps en ayant fait au préalable le zéro sur la cuvette tout de suite après l'injection de l'extrait dilué. Le volume initial d'extrait de maturation est ajusté en fonction de l'activité de l'enzyme.

L'activité est calculée à partir de la pente initiale de la mesure d'absorbance en fonction du temps en utilisant la régression linéaire. Dans le cas de ces échantillons très dilués, le coefficient de pente initiale est directement proportionnel à la quantité d'hydrogénase active<sup>6</sup>. Un coefficient d'extinction molaire de 9.78 AU/(mM.cm)<sup>14</sup> et une épaisseur de cuvette de 1 cm ont été utilisés pour calculer la vitesse de consommation de H<sub>2</sub>.

## **MM-4 Cristallisation et Cristallographie aux rayons X**

### **MM-4.1 Co-cristallisation ou trempage de cristaux de la protéine Hyde de *Thermotoga maritima* avec des substrats potentiels**

Dans le cas du trempage, 10  $\mu\text{L}$  d'échantillon protéique à 10 mg/mL dans 20 mM Tris pH 8, 200 mM NaCl sont incubés avec 20  $\mu\text{L}$  de tampon (50 mM Tris pH 8.0, 30 mM CHAPS). Les gouttes suspendues sont préparées en condition anaérobie en BAG à 20°C en mélangeant 3  $\mu\text{L}$  de cette solution avec 1  $\mu\text{L}$  d'une solution contenant 22-28% PEG 4000, 100 mM Tris pH 8.0, 200 mM  $\text{Li}_2\text{SO}_4$ . Les cristaux obtenus sont incubés pendant 30-120 minutes dans une solution cryoprotectrice contenant 30% PEG 4000, 20% glycérol, 80 mM CHAPS, 50 mM Tris pH 8.0, 100 mM  $\text{Li}_2\text{SO}_4$ , 20 mM de substrats potentiels. Les cristaux sont alors lavés dans une solution cryoprotectrice nouvellement préparée qui ne contient pas de ligand (30% PEG 4000, 20% glycérol, 80 mM CHAPS, 50 mM Tris pH 8.0, 100 mM  $\text{Li}_2\text{SO}_4$ ) pendant 30 minutes avant congélation anaérobie. Pour les expériences de cinétiques enzymatiques, les cristaux sont lavés successivement pendant 30 s dans 2 gouttes de 10  $\mu\text{L}$  de solution cryoprotectrice sans ligand. Au cours de ces transferts, l'utilisation de boucles de congélation permet de diminuer le volume de solution apportée avec les cristaux. Les cristaux sont finalement trempés dans une solution cryoprotectrice nouvellement préparée (30% PEG 4000, 20% glycérol, 80 mM CHAPS, 50 mM Tris pH 8.0, 100 mM  $\text{Li}_2\text{SO}_4$ , 12 mM  $\text{Na}_2\text{S}_2\text{O}_4$ ) et congelés dans du propane liquide à différents temps en condition anaérobie en BAG<sup>15</sup>.

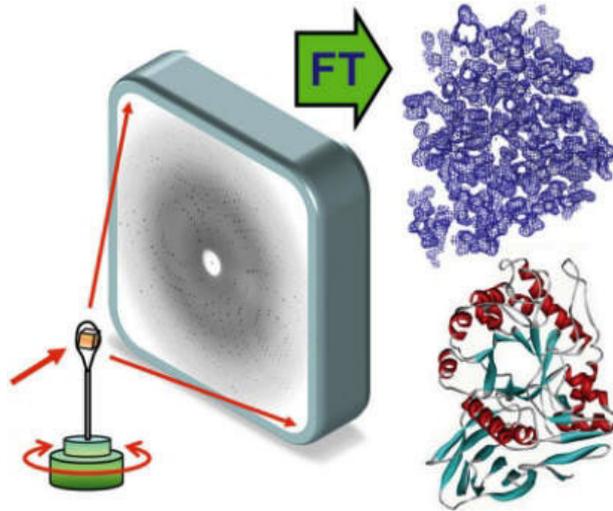
Dans le cas de la co-cristallisation, l'échantillon protéique est incubé avec 10 excès molaires de S-adénosyl-L-méthionine (SAH) et 20-100 mM de ligand. 10  $\mu\text{L}$  d'échantillon protéique sont incubés avec 10  $\mu\text{L}$  de solution contenant 100 mM Tris pH 8.0, 40 excès molaires de SAH et 60 mM CHAPS avant cristallisation. Les gouttes suspendues sont préparées en mélangeant 2  $\mu\text{L}$  de cette solution avec 1  $\mu\text{L}$  de solution contenant 23-24% PEG 4000, 100 mM Tris pH 8.0, 200 mM  $\text{Li}_2\text{SO}_4$  et 1  $\mu\text{L}$  de solution contenant 80 mM de ligand. La solution de cryoprotection contient 30% PEG4000, 20% glycérol, 80 mM CHAPS, 50 mM Tris pH 8.0, 100 mM  $\text{Li}_2\text{SO}_4$ , 20-100 mM de ligand.

### **MM-4.2 Cristallographie aux rayons X**

Cette partie de Matériels et Méthodes est dédiée à la brève présentation des différentes étapes nécessaires pour la détermination structurale à l'aide des rayons X (RX).

Le principe de la résolution de structure tridimensionnelle des protéines par cristallographie aux RX est schématisé dans la [Figure MM-112](#). Le cristal de la protéine cryoprotégé est exposé aux rayons X qui sont ensuite diffusés par les atomes du cristal et enregistrés par un détecteur 2D. Les clichés de diffraction sont des signatures dans l'espace réciproque de la transformée de Fourier de la macromolécule cristallisée convoluée avec la fonction réseau du cristal. Ainsi, pour retrouver la structure de la protéine nous faisons le

procédé inverse, la transformée de cette transformée de Fourier. Nous nous servons des intensités des pics de diffraction et des phases (d'un modèle ou déterminées expérimentalement) pour construire la carte de densité électronique représentée par le grillage bleu sur la [Figure MM-112](#). Un modèle atomique de la structure, représenté ici par des rubans rouges et bleus, est ensuite construit pour expliquer la carte de densité électronique.

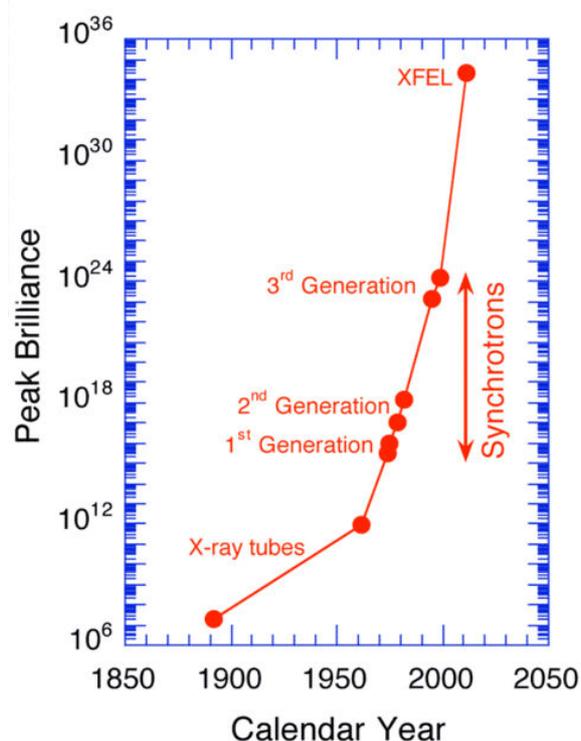


**Figure MM-112** : Le principe de détermination de la structure d'une protéine

Les paragraphes suivants vont expliquer un peu plus en détails les étapes de la cristallographie aux rayons X, spécifiques à mon travail de thèse, notamment les conditions expérimentales, les corrections des données collectées, le calcul des cartes de densité électronique, les cartes de double différence, la diffraction anormale et l'affinement restreint.

#### **MM-4.2.1**    *Conditions expérimentales*

Une fois la protéine cristallisée et congelée, on procède à l'expérience de diffraction aux rayons X, afin d'acquérir les clichés de diffraction nous permettant de déterminer sa structure tridimensionnelle. Lors de ma thèse, j'ai effectué de nombreuses collectes sur les lignes de lumières dédiées à la biocristallographie à l'ESRF (European Synchrotron Radiation Facility) à Grenoble, notamment BM30, ID14 et ID23-2. Les rayons X y sont produits grâce aux électrons accélérés à la vitesse proche de celle de la lumière. L'anneau de l'ESRF produit des faisceaux X monochromatiques ou des faisceaux blancs avec des énergies allant de 10 à 120 keV. La brillance est un facteur important de qualité d'un faisceau de photons. En pratique, plus un faisceau est brillant, plus les photons du faisceau sont concentrés en un point. La brillance des lignes à l'ESRF est de  $10^{21}$  photons/s/mm<sup>2</sup>/mr<sup>2</sup>/0.1%bw, ce qui le positionne parmi les synchrotrons de 3<sup>ème</sup> génération [[Figure MM-113](#)].



**Figure MM-113 :** La comparaison de la brillance de différentes sources de rayons X au cours des années

(source : <https://www.psi.ch/swissfel/why-swissfel>)

L'expérience de diffraction consiste en alignant le cristal de la protéine congelée avec le faisceau X et le détecteur 2D collectant les photons diffractés. Les détecteurs utilisés lors de mes expériences étaient la camera CCD et le détecteur Pilatus. Le support du cristal est monté sur un axe de rotation permettant au cristal de tourner selon un axe parallèle au plan du détecteur 2D. De manière générale, lors de l'expérience, on collecte un cliché de diffraction par 0.1 ou 1 degré de rotation du cristal. Le pas de rotation dépend de la taille de la maille du cristal. L'étendu de la rotation est choisie de manière à couvrir le maximum de plans diffractant non reliés par symétrie.

#### **MM-4.2.2**    *Corrections des données*

Après l'acquisition, nous procédons à l'étape de la « réduction de données ». Il s'agit de l'étape d'indexation des pics de diffraction, de la mise à l'échelle et de la conversion au format reconnu par les logiciels de calcul de structure. Ceci est fait par un lot de trois logiciels : XDS<sup>16</sup>, XSCALE et XDSCONV<sup>17</sup>.

Le logiciel XDS recherche à attribuer une maille et une matrice d'orientation permettant d'expliquer la position des taches de diffraction des clichés enregistrés. A l'issue de cette analyse, XDS propose les groupes d'espace du cristal les plus probables. Les positions des pics étant maintenant connues, le logiciel intègre le signal mesuré pour chaque

pic afin d'en estimer l'intensité (étape d'intégration). La comparaison du changement d'intensités des pics équivalents (par symétrie du cristal) sur différents clichés permet de déterminer la qualité du jeu de donnée et va servir à la mise à l'échelle des données. En théorie, il ne devrait pas y avoir de variations. Cependant, en réalité plusieurs phénomènes entravant cette règle entrent en jeu. L'intensité des pics est proportionnelle au volume diffractant. De cette manière, les dimensions du cristal peuvent être à l'origine de la variation d'intensité des pics en fonction de l'orientation du cristal. Ensuite, le cristal de la protéine peut subir des dégradations au cours de la collecte, causées par le rayonnement ionisant. La détérioration structurale peut amener une baisse d'intensité des pics de diffraction. Une des particularités de l'ESRF est la diminution d'intensité des rayons X au cours du temps. Pour cette raison, dans le cas où l'expérience s'étend sur plusieurs heures, les plans de diffractions équivalents ne vont pas recevoir la même intensité.

Les corrections d'intensité permettent ensuite d'attribuer à chaque pic hkl (équivalent au pic -h-k-l par la loi de Friedel), une amplitude du facteur de structure native FP et un écart-type de FP, SIGFP. Cette amplitude correspond à la racine carrée de l'intensité enregistrée. Les clichés de diffractions sont donc convertis après cette étape en un tableau (fichier format texte) (HKL) contenant une liste des pics de diffraction avec comme attributs les valeurs de FP et SIGFP. Chaque ligne correspond à un facteur de structure donné, avec ses coordonnées HKL, son module FP et son écart type SIGFP.

### ***MM-4.2.3    Calcul des cartes de densité électronique***

La densité électronique d'une structure de protéine est calculée par transformée de Fourier des facteurs de structures. Ces derniers correspondent à des nombres complexes composé d'un module et d'une phase et contiennent, chacun, une partie de l'information complète de la structure protéique. Cependant, la diffraction ne nous permet d'acquérir que l'intensité des pics de diffraction. L'information de phase (déphasage) qui est la plus informative sur la position des différents atomes de la structure est perdue. La grande problématique de la cristallographie est, ce qui a été appelé : « le problème de la phase ».

Pour répondre à ce problème, plusieurs méthodes ont été développées. Il existe notamment différentes techniques de phasage expérimental, soit par remplacement isomorphe avec des sels d'atomes lourds, soit par l'utilisation du phénomène de diffusion anomal, soit par remplacement moléculaire. Ces méthodes sont maintenant très bien décrites dans tous les bons livres de cristallographie et ne seront pas développées dans mon manuscrit. Dans notre cas très particulier, les cristaux de la protéine HydE étant isostructuraux (à une exception près - voir ***Chapitre IV***), nous n'avons même pas eu besoin d'utiliser la méthode du remplacement moléculaire et nous avons pu effectuer des affinements directs en utilisant un modèle de HydE simplifié dans lequel nous avons supprimé les atomes contenus dans la cavité du site actif.

Pour inverser le processus de la diffraction nous pouvons procéder par l'inversion de son formalisme mathématique, la transformation de Fourier (FT). En exprimant le facteur de structure  $F(h)$  dans l'espace réciproque  $R^*$  comme une intégrale de Fourier de la densité électronique  $\rho(xyz)$  dans l'espace réel  $R$ , nous obtenons:

$$F(h) = V \iiint_{x=y=z=0}^1 \rho(xyz) \cdot \exp(2\pi i h x) dx dy dz$$

Maintenant, en effectuant la transformation de Fourier inverse ( $FR^{-1}$ ), nous obtenons l'expression de la densité électronique  $\rho(xyz)$ :

$$\rho(xyz) = \frac{1}{V} \sum_{h=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} F(hkl) \cdot \exp(-2\pi i(hx + ky + lz))$$

On peut utiliser cette dernière expression et faire de nouveau une transformation de Fourier pour retrouver le facteur de structure exprimé par une somme discrète :

$$F(hkl) = V \sum_{x=0}^a \sum_{y=0}^b \sum_{z=0}^c \rho(xyz) \cdot \exp(2\pi i(hx + ky + lz))$$

Cette expression est utilisée pour transformer la carte de densité électronique en facteurs de structure de l'espace réciproque.

Les équations de transformations sont utilisées dans la construction du modèle et son affinement par les logiciels COOT et Refmac5, respectivement.

#### **MM-4.2.4 Diffusion anormale**

Nous avons utilisé la diffusion anormale d'un atome lourd pour visualiser sa position en dépit de son taux d'occupation très faible rendant impossible son identification sur une carte de densité électronique de différence normale. Afin d'expliquer cet outil, nous allons préciser d'abord le phénomène d'absorption des rayons X, pour ensuite présenter les facteurs de la diffusion anormale, la brisure de la loi de Friedel et le calcul des cartes de la diffusion anormale.

##### **MM-4.2.4.1 Absorption des rayons X**

Tous les matériaux absorbent les rayons X. En règle générale, l'absorption diminue rapidement lorsque l'énergie des rayons X augmente (ou leur longueur d'onde diminue). Pour une longueur d'onde donnée, le coefficient d'absorption  $\mu$  est supérieur pour les éléments plus lourds par rapport aux éléments légers.

Le profil d'absorption des rayons X par les éléments lourds n'est pas lisse mais présente des variations abruptes ressemblant à des marches, appelées les seuils d'absorption. Il s'agit des longueurs d'onde spécifiques des rayons X, pour lesquelles les électrons entrent en résonance. Un transfert d'énergie se produit entre le rayonnement et les électrons, conduisant à l'arrachage d'un ou de plusieurs électrons de l'atome. C'est l'effet photoélectrique. L'énergie des rayons X étant très élevée, ce sont les électrons des niveaux proches du noyau qui sont affectés, notamment les niveaux K, L et M. Les électrons du niveau K sont les électrons plus fortement liés dans l'atome. L'extraction de l'un des deux électrons du niveau K nécessite l'absorption des rayons X les plus énergétiques. Pour cette raison, le seuil d'absorption K pour un élément donné est le seuil le plus énergétique. Les électrons des niveaux L et M sont moins fortement liés, ils absorbent le rayonnement moins énergétique formant ainsi les seuils d'absorption L et M.

L'atome ionisé revient à son état énergétique de départ par la relaxation d'un électron d'un niveau supérieur. L'excès d'énergie ainsi créée produit dans la majorité des cas l'émission des rayons X caractéristiques. Le spectre de fluorescence X ainsi créée, reflète la structure du spectre d'absorption. La fluorescence X est pratiquée lors de la collection des données pour identifier de façon précise le seuil d'absorption.

#### MM-4.2.4.2 Facteurs de la diffusion anormale

Lorsque l'énergie des rayons X est égale ou supérieure à l'énergie du seuil d'absorption d'un atome du cristal, on ne peut plus considérer les électrons de cet atome comme libres. Le fait qu'ils entrent en résonance en interagissant avec les rayons X va conduire au phénomène de diffusion anormale. La théorie de diffusion cohérente (diffusion de Thomson) ne sera plus respectée.

Le comportement des électrons sous le rayonnement peut être décrit par le modèle de l'oscillateur harmonique amorti. La fréquence de résonance  $\omega_0$  correspond à la fréquence d'absorption de l'atome lourd. Le facteur de diffusion atomique  $f$  pour un électron lié devient  $\frac{\omega^2}{\omega^2 - \omega_0^2 - ik\omega}$ . Cette expansion de l'équation de la diffusion atomique résonnante amène aux deux importantes conséquences.

Premièrement, le facteur de diffusion atomique contient maintenant des termes dépendants de la longueur d'onde du faisceau  $f'(\lambda)$  et  $f''(\lambda)$  en plus du terme indépendant de la longueur d'onde mais dépendant de l'angle de diffusion  $f_s^0$  dont la valeur peut être estimée par l'équation de *Cromer-Mann*. L'expression du facteur de diffusion atomique avec les termes de diffusion anormale devient :

$$f_{(s,\lambda)} = f_s^0 + f'_{(\lambda)} + i.f''_{(\lambda)}$$

Le terme réel de la diffusion anormale constitue la contribution dispersive  $f'(\lambda)$  et le terme imaginaire correspond à la contribution anormale  $f''(\lambda)$ . Les deux termes de diffusion anormale sont indépendants de l'angle de diffusion. Le fait d'observer les contributions additionnelles au facteur de diffusion atomique pour différentes longueurs d'onde veut dire qu'autour du seuil d'absorption, nous allons mesurer, pour les mêmes pics, différentes intensités diffractées en fonction de la longueur d'onde des rayons X incidents. Cela conduit aux *différences dispersives*.

La seconde conséquence de la diffusion anormale est la présence du terme imaginaire (à l'origine du changement de phase de  $+90^\circ$  dans la contribution de la diffusion anormale  $f''$ ) qui rompt la centrosymétrie du motif de diffraction. Cela a pour conséquence les *différences anormales* (ou de Bijvoet), dans la valeur des facteurs de structure  $F_{hkl}$  et son pair  $F_{-h-k-l}$  qui ne sont plus égaux.

Les deux contributions au facteur de diffusion atomique, dispersive  $f'(\lambda)$  et anormale  $f''(\lambda)$ , sont utilisées dans les expériences de diffraction afin de positionner les atomes lourds dans les cartes de densité électronique et dans le processus de phasage des structures cristallines pour déterminer la structure de nouveaux protéines.

#### MM-4.2.4.3 Brisure de la loi de Friedel

En absence de la diffusion anormale, les paires conjuguées de facteurs de structure  $F_{hkl}$  et  $F_{-h-k-l}$  ont exactement la même valeur ( $|F_{hkl}|=|F_{-h-k-l}|$ ) et les angles de phase opposés ( $\varphi_{hkl}=-\varphi_{-h-k-l}$ ). Pour cette raison en absence de la diffusion anormale, les valeurs des intensités diffractées sont égales. Cette égalité est appelée la loi de Friedel et les réflexions conjuguées sont les paires de Friedel. En considérant les facteurs de structure :

$$F_{hkl} = F_{hkl}(\cos \varphi_{hkl} + i \sin \varphi_{hkl}) = A_{hkl} + iB_{hkl}$$

$$F_{\bar{h}\bar{k}\bar{l}} = F_{\bar{h}\bar{k}\bar{l}}(\cos(-\varphi_{hkl}) + i \sin(-\varphi_{hkl})) = F_{hkl}(\cos \varphi_{hkl} - i \sin \varphi_{hkl}) = A_{hkl} - iB_{hkl}$$

L'amplitude de la réflexion est obtenue comme produit de F et de son complexe conjugué  $F^*$  :

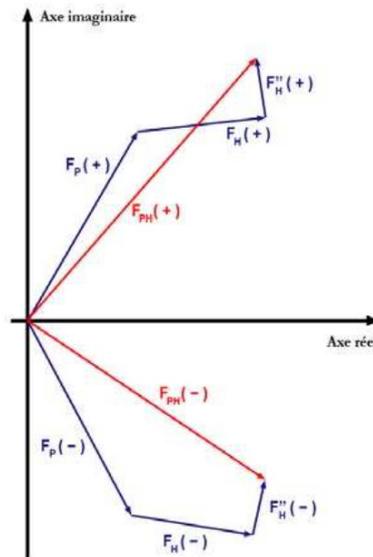
$$I_{hkl} \approx F_{hkl}^2 = (A_{hkl} + iB_{hkl})(A_{hkl} + iB_{hkl}) = A_{hkl}^2 + iB_{hkl}^2$$

$$I_{\bar{h}\bar{k}\bar{l}} \approx F_{\bar{h}\bar{k}\bar{l}}^2 = (A_{hkl} - iB_{hkl})(A_{hkl} - iB_{hkl}) = A_{hkl}^2 + iB_{hkl}^2$$

Les intensités  $I_{hkl}$  et  $I_{-h-k-l}$  de la paire de Friedel sont égales.

Maintenant considérons la situation avec un atome dispersif dans la structure de la protéine. La figure d'Argand correspondante à cette situation est présentée dans la [Figure MM-114](#). Le vecteur  $F_P$  représente la somme de toutes les contributions des atomes diffractant normalement,  $F_H$  et  $F''_H$  représentent la contributions réelle et imaginaire du diffuseur anormal et  $F_{PH}$  est la somme de toutes les contributions. La [Figure MM-114](#) illustre qu'en présence de

la diffusion anormale les facteurs de structures conjugués  $F_{PH(+)}$  et  $F_{PH(-)}$  ne sont plus égaux à cause du terme imaginaire de la contribution anormale  $F''_H$ .



**Figure MM-114 :** Diagramme d'Argand illustrant la brisure de la loi de Friedel

La brisure de la loi de Friedel nous permet de définir *la différence anormale*,  $D_{ANO}$ . Elle est définie comme la différence des amplitudes des facteurs de structure d'une paire de Friedel :

$$D_{ANO} = |F(+)| - |F(-)| = |F(hkl)| - |F(\overline{hkl})|$$

Le rapport de Bijvoet permet d'estimer l'intensité des effets anormaux :

$$\frac{\langle |D_{ANO}| \rangle}{\langle |\overline{F}| \rangle} = \sqrt{\frac{2N_A f''}{N_P f_P}}$$

Avec :

$\langle x \rangle \equiv \sqrt{\overline{x^2}}$  est la racine carrée de la valeur quadratique moyenne de la grandeur  $x$

$|\overline{F}| = \frac{1}{2} (|F(+)| + |F(-)|) = \frac{1}{2} (|F(hkl)| + |F(\overline{hkl})|)$  (est la valeur moyenne des amplitudes d'une paire de Friedel)

$N_P$  est le nombre d'atomes de la protéine dans la maille possédant le facteur de diffusion moyen  $f_P$

$N_A$  est le nombre de diffuseurs anormaux dans la maille dont le composant imaginaire du facteur de diffusion atomique est  $f''$

Dans le cas où l'effet anormal est dû à un seul type de diffuseur anormal, le rapport de Bijvoet est proportionnel au composant imaginaire du facteur de diffusion atomique  $f''$ . Afin d'obtenir la différence anormale la plus importante, on se place à l'énergie du faisceau correspondant à  $f''$  maximale.

#### MM-4.2.4.4 Détermination des facteurs de diffusion atomique dispersive $f'$ et anormale $f''$

Afin d'optimiser le signal dispersif et anormal de la diffusion anormale, nous avons besoin de mesurer la contribution du signal anormal à travers le scan du seuil d'absorption. Généralement, la mesure se fait par l'analyse de fluorescence X. Lors d'un scan du seuil d'absorption, on positionne le détecteur de fluorescence à côté du cristal de manière perpendiculaire à l'arrivée du faisceau. L'énergie du faisceau est augmentée par des pas de 0.2 à 1 eV de manière à passer à travers le seuil d'absorption de l'élément diffractant anormalement. L'intensité du signal de fluorescence X varie proportionnellement à l'absorption. Le coefficient d'absorption normalisée  $\mu(\lambda)$  est ensuite déduit à partir du spectre de fluorescence.

La contribution imaginaire à la diffusion anormale  $f''(\lambda)$  est proportionnelle au coefficient d'absorption  $\mu$ , correspondant avec la longueur d'onde  $\lambda$  :

$$f''(\lambda) = \frac{mc}{4e^2\hbar} E(\lambda)\mu(\lambda)$$

$m$  correspond à la masse de l'électron,  $c$  est la vitesse de la lumière,  $e$  est la charge de l'électron,  $\hbar$  est la constante de Dirac et  $E$  l'énergie du photon incident. La valeur de  $f''(\lambda)$  est donc déterminée directement par l'expérience de fluorescence. La partie réelle de la diffusion anormale  $f'(\lambda)$  est obtenue indirectement à partir de la partie imaginaire, grâce à la transformation de Kramers-Kronig :

$$f'(\lambda) = \frac{2}{\pi} \int_0^{\infty} E' \frac{f''(E')}{E^2 - E'^2} dE'$$

La valeur de la partie réelle de la diffusion anormale  $f'(\lambda)$  peut être obtenue aussi par la simple dérivation du spectre mesuré de  $f''(\lambda)$ . La Figure MM-115 illustre l'évolution des valeurs des coefficients  $f'$  et  $f''$  en fonction de l'énergie des rayons X pour le sélénium, le fer et le soufre.

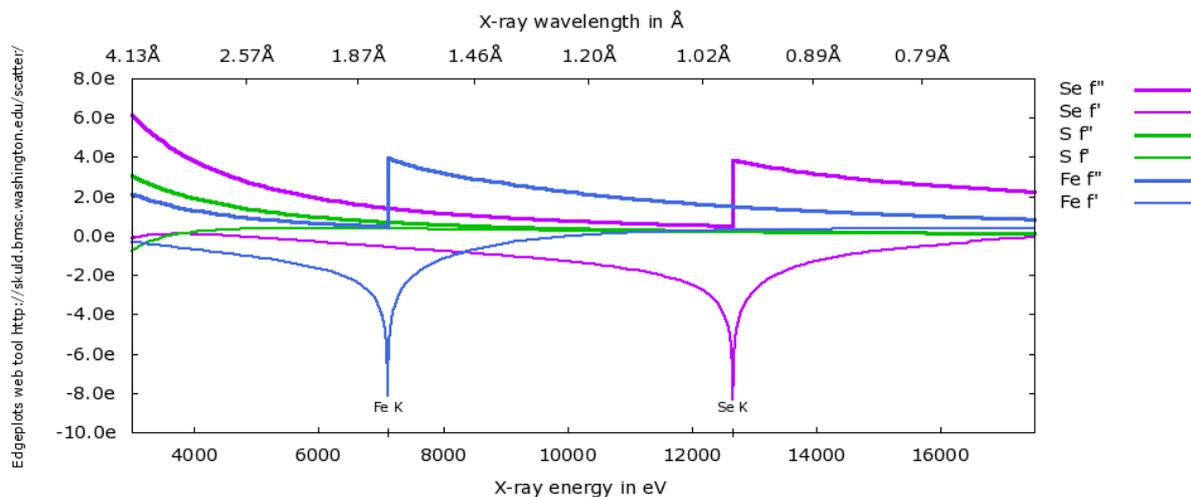


Figure MM-115: Coefficients de dispersion anormale  $f'$  et  $f''$  pour les atomes de Se, Fe et S.

#### MM-4.2.4.5 Cartes de densité électronique anormale

Les cartes de densité électronique anormale nous ont servi à positionner le diffuseur anormal. En pratique, pour calculer ces cartes, nous faisons l'acquisition d'un set de données de diffraction à la longueur d'onde du seuil d'absorption de l'atome lourd. Nous utilisons ensuite ces données expérimentales et un modèle de la protéine étudiée qui nous donne la phase. Ces deux informations sont ensuite entrées dans le programme de calcul. Nous utilisons un des programmes du volet CCP4, dont le nom est FFT, pour *Fast Fourier Transform*. Les données et les paramètres d'entrée et leurs origines, sont précisés ci-dessous :

Modèle préexistant :

*PHIC* : l'angle de phase

*FOM* : le facteur de qualité

Données expérimentales :

*DANO* : la différence du signal anormal ( $D_{ANO}$ )

*SIGDANO* : écart-type de *DANO*

A l'issu du calcul nous obtenons les cartes avec les positions du diffuseur anormal. Le détail du calcul des densités électroniques est présenté dans MM-4.2.3.

#### **MM-4.2.5 Carte de densité électronique double différence**

La carte de double différence est utilisée pour voir les différences dans les positions atomiques entre deux cristaux isomorphes (ayant le même groupe d'espace), à condition que les mailles n'ont que de très petites variations (<1.0%). Cet outil est utilisé pour améliorer l'interprétabilité de la densité électronique.

Pour construire la carte de densité électronique, nous utilisons la différence d'amplitudes des facteurs de structure pour les deux jeux :

$$\Delta F = F_{hkl(1\zeta)} - F_{hkl(2\zeta)}$$

La phase utilisée dans le calcul est la phase du modèle d'un des jeux de données. Les deux modèles ont été testés comme source de phase. Nous n'avons pas remarqué de différences majeures. La carte de densité électronique double différence a été calculée avec le logiciel FFT.

Un exemple de carte de densité électronique double différence peut être trouvé dans le paragraphe III-3.1.1 du *Chapitre III* de la partie Résultats. Les zones rouges sont les zones avec la densité négative. En considérant l'équation ci-dessus, les zones qui apparaissent en rouge coïncident avec la densité du jeu de données 2 et ne se trouvent pas dans le jeu de données 1. Les volumes délimités par le grillage vert représentent la densité électronique positive. C'est la densité propre au jeu de données 1. Les parties communes entre le jeu de données 1 et 2 ne sont pas représentées.

En pratique, le calcul des cartes de densité électronique double différence nous a été très utile lors des comparaisons de structures de *TmHydE* au début de la réaction à  $t_0$  et aux différents moments au cours de la réaction à  $t_x$ . Ces cartes représentaient les mouvements atomiques.

#### **MM-4.2.6 Affinement restreint**

La résolution des clichés de diffraction des macromolécules est typiquement de 3-2 Å ce qui correspond à environ 1.2-5 observations par atome. Même si c'est proche de la résolution atomique, cette résolution est insuffisante pour modéliser les petites molécules nécessitant environ 80 observations par atome, c'est-à-dire 0.8 Å de résolution. Avec la résolution standard d'une macromolécule, la complexité de la maille de la protéine rend le rapport entre le nombre d'observations connues et le nombre de paramètres à affiner trop faible pour un affinement sans contraintes. Le modèle divergerait ou aboutirait à une solution sans aucun sens chimique. Dans le but d'améliorer le modèle et diminuer l'incertitude sur les positions atomiques, nous attribuons aux atomes les contraintes stéréochimiques<sup>18</sup>. L'affinement restreint est ensuite effectué uniquement sur les structures possédant ces descriptions géométriques.

##### MM-4.2.6.1 Contraintes stéréochimiques

Les contraintes stéréochimiques limitent la mobilité des atomes lors de l'affinement du modèle. Parmi les exemples typiques des contraintes, on peut trouver la longueur et les angles des liaisons interatomiques, la planarité, la chiralité et les angles de torsion.

Pour utiliser ces contraintes, nous avons besoin de connaître la valeur idéale et l'écart-type pour chaque paramètre. Nous pouvons ainsi attribuer des poids ou des pénalités pour chaque écartement par rapport à la valeur idéale du paramètre donné. Le but de l'affinement est de minimiser la valeur globale des pénalités définies selon l'équation suivante :

$$\text{Pénalité} = \sum \text{poids} (\text{Valeur observée} - \text{Valeur idéale})^2$$

Les contraintes stéréochimiques des nombreuses protéines se trouvent dans les bases de données, donc nous pouvons facilement retrouver un set de contraintes applicable pour notre modèle. Elles ont été générées à partir des structures des acides aminés isolés. Ces données sont déjà intégrées dans les logiciels d'affinement de structures cristallographiques. Cependant, ce n'est pas le cas pour les ligands.

Afin de pouvoir inclure les nouvelles structures de ligands dans le modèle d'affinement, nous avons procédé de la manière suivante: d'abord, nous avons choisi les structures à haute résolution qui contenaient le ligand et nous avons construit cette molécule dans les cartes de densité électroniques atome par atome. Nous avons pris en compte le pH de la condition de cristallisation pour modéliser l'état de protonation de ses groupements amines, carboxylates et thiols éventuellement présents. Ensuite ce modèle initial a été utilisé pour le calcul d'une structure idéalisée de la molécule. La géométrie de cette structure a été optimisée par les logiciels tels que *ProDRG*<sup>19</sup> de la suite de programmes CCP4 ou le program *elbow* de la suite *Phenix*. Ces logiciels génèrent ensuite les contraintes stéréochimiques pour la nouvelle structure. Ces contraintes sont alors utilisées dans le logiciel d'affinement et la structure du complexe protéine-ligand peut être modélisée par la méthode d'affinement restreint.

## MM-5 Spectroscopie

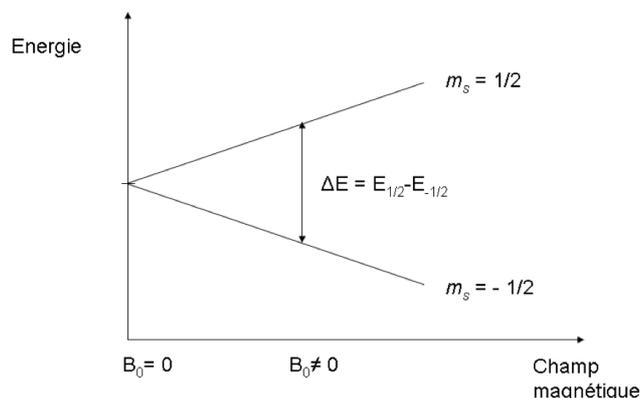
### MM-5.1 Spectroscopie de résonance magnétique nucléaire du proton

Ce travail a été réalisé en collaboration avec Adrien Favier et Eric Condamine, Plateforme Nationale et Européenne de RMN à l'IBS.

#### MM-5.1.1 Principe de la RMN

La spectroscopie de résonance magnétique nucléaire est une technique qui permet d'avoir des informations sur l'environnement d'un atome en étudiant l'interaction entre son spin nucléaire ( $I$ ), qui correspond au moment magnétique du noyau, et un rayonnement électromagnétique dans la bande de fréquence des ondes radio. Cette méthode est applicable à tout atome avec un spin nucléaire non nul, c'est-à-dire à tout atome dont le noyau n'a pas les deux nombres de protons et de neutrons paires. En effet un noyau atomique est composé de neutrons et protons, eux mêmes composés de trois quarks fermioniques. Dans la physique des particules, un fermion possède un spin demi-entier, tandis qu'un boson possède un spin entier. Comme chaque particule composée de nombre impaire de fermions est elle aussi un fermion, le proton et le neutron possèdent chacun un spin  $\frac{1}{2}$ . Si un atome a les deux nombres de protons et de neutrons simultanément pairs, le spin nucléaire résultant est nul ( $^{12}\text{C}$ ,  $^{16}\text{O}$ ). Si par contre les deux nombres de protons et de neutrons sont impairs, le moment magnétique nucléaire est un nombre entier ( $^2\text{H}$ ,  $^{14}\text{N}$ ). Dans le dernier cas, qui est le plus étudié en RMN du liquide, les noyaux avec un nombre de masse impaire, c'est-à-dire dont la somme de neutrons et de protons donne un nombre impair, possèdent un spin nucléaire d'une valeur demi-entier ( $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ,  $^{17}\text{O}$ ).

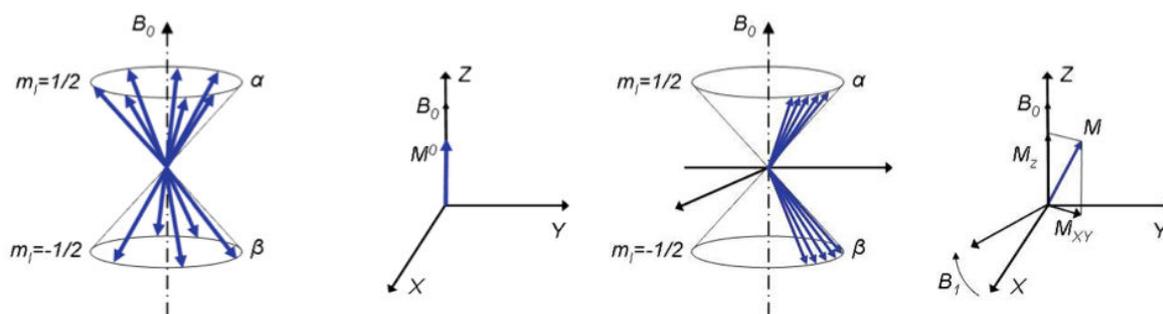
En absence de champ magnétique il n'y a pas d'orientation préférentielle du moment magnétique du noyau. Par contre, quand un atome est soumis à un champ magnétique statique, les niveaux d'énergie du spin nucléaire ne sont plus dégénérés (effet de Zeeman) et le spin nucléaire peut alors avoir  $2I + 1$  orientations possibles. Dans le cas du proton isolé ( $I=\frac{1}{2}$ ), il va soit s'aligner avec le champ magnétique externe de façon parallèle ( $m=+\frac{1}{2}$ ) ; c'est le niveau énergétique le plus stable, ou de façon antiparallèle ( $m=-\frac{1}{2}$ ) [Figure MM-116]. Dans un ensemble de protons, on aura une population majoritaire de protons dont le moment magnétique nucléaire sera parallèle au champ magnétique externe et une population minoritaire dont le spin prendra une orientation antiparallèle. Cette répartition s'appelle la distribution de Boltzmann. La somme de ces moments magnétiques nous donne l'aimantation longitudinale macroscopique de l'échantillon.



**Figure MM-116 :** Effet Zeeman – Séparation des niveaux d'énergie des états de spin en présence d'un champ magnétique externe

Dans une expérience de RMN 1D  $^1\text{H}$  classique, un échantillon est d'abord placé dans un champ magnétique statique. En présence de ce champ inducteur, les moments magnétiques de spins entament un mouvement de précession à la fréquence de Larmor autour de l'axe du champ. Pour pouvoir observer l'aimantation de l'échantillon, infiniment plus petite que  $B_0$ , il faut la faire basculer dans le plan transversal au moyen d'un champ radiofréquence oscillant. Pour cela, le champ  $B_1$  doit remplir une condition géométrique - être perpendiculaire au  $B_0$  et une condition énergétique - avoir la fréquence de Larmor du noyau observé, autrement dit, le photon incident doit avoir l'énergie qui correspond à la différence d'énergies entre les deux états d'énergie quantifiés  $m = +\frac{1}{2}$  et  $m = -\frac{1}{2}$ . Une fois l'aimantation basculée de  $90^\circ$ , on arrête l'impulsion radiofréquence et on mesure la disparition de l'aimantation transverse de chaque proton (le déclin de précession libre) qui obéit à différents phénomènes de relaxation. En absence de l'impulsion radiofréquence, l'échantillon va progressivement récupérer l'aimantation longitudinale, car soumis au champ magnétique, les spins ont la tendance de retrouver la distribution de Boltzmann [Figure MM-117]. Le déclin de précession libre enregistré est une somme de sinusoides, ainsi sa transformée de Fourier nous donne un spectre à une dimension avec les fréquences de Larmor de chaque proton de l'échantillon.

Dans le cas d'une population de protons non isolés, les courants magnétiques créés par le mouvement des électrons (blindage électronique) modifient les champs magnétiques locaux perçus par les protons ce qui engendre des différences dans les fréquences de Larmor de protons en question (déplacement chimique  $\delta$ ). De plus, les interactions avec les spins nucléaires des atomes avoisinants (couplage scalaire  $J$ ), véhiculées à travers les liaisons atomiques, rajoutent des niveaux d'énergie du spin nucléaire supplémentaires, ce qui a pour effet la démultiplication de raies de résonances des protons affectées. Les valeurs de déplacement chimique et du couplage  $J$  nous renseignent alors sur la nature chimique et la structure de la molécule étudiée.



**Figure MM-117 :** Distribution selon la loi de Boltzmann d'une population de spins demi-entier sous l'effet d'un champ magnétique statique  $B_0$  ; résultante des moments magnétiques nucléaires ou l'aimantation macroscopique de l'échantillon

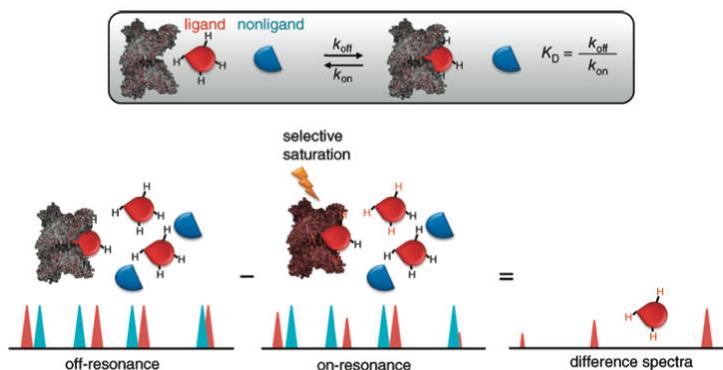
### MM-5.1.2 RMN STD $^1H$

La RMN STD  $^1H$  (en anglais *Saturation-Transfert Difference*) est une méthode de RMN du liquide développée par Meyer B. et Mayer M. en 1999 qui permet de confirmer la formation d'un complexe transitoire entre une protéine et un ligand<sup>20,21</sup>. Cette méthode est particulièrement adaptée pour identifier des ligands d'une protéine à partir d'un mélange de molécules. De plus, il est possible de déterminer par quels atomes le ligand interagit avec le récepteur, car les atomes qui sont en contact direct avec la protéine présentent un plus fort signal STD.

### MM-5.1.3 Principe d'une expérience RMN STD 1D $^1H$

Une protéine consiste d'un grand nombre de protons, qui sont couplée à travers les liaisons atomiques - couplage J et à travers l'espace par les interactions dipôle - dipôle. En solution, cette dernière interaction se manifeste sous forme d'une relaxation nucléaire transversale (l'effet Overhauser nucléaire). Sous son effet, la saturation sélective d'une fréquence de la protéine va se propager rapidement sur la totalité de résonances (diffusion de spin). Si une formation d'un complexe ligand-protéine a lieu, le transfert de polarisation de la protéine vers le ligand va entraîner une saturation progressive des résonances du ligand.

Dans une expérience STD standard on enregistre d'abord un spectre 1D  $^1H$  précédé d'un pulse de saturation hors de la zone de résonance de la protéine et du ligand (spectre "off-résonance"). Dans un second temps, on enregistre un spectre 1D  $^1H$  précédé d'un pulse de saturation sélective de la protéine en dehors de la zone spectrale du ligands (spectre "on-résonance"). La soustraction des deux spectres donne le spectre de différence de transfert de saturation ou spectre STD. Dans ce spectre, seuls sont visibles les signaux des ligands qui interagissent avec la protéine [Figure MM-118].



**Figure MM-118 :** Expérience de RMN STD. La fixation du ligand à la protéine permet un transfert de polarisation de la protéine vers le ligand.

#### **MM-5.1.4 Préparation de l'échantillon pour une expérience de RMN STD 1D $^1H$**

Une colonne GE Healthcare HiPrep 26/10 Desalting, équilibrée dans du tampon 50 mM phosphate pH 8.0, 150 mM NaCl, a été utilisée pour mettre la protéine HydE Tm Nstrep  $\Delta FeS$  dans un tampon dépourvu de protons autre que ceux de l'eau. La protéine a été ensuite reconcentrée à 14 mg/mL, soit l'équivalent de 350  $\mu M$ . Un excès de 25x de S-adénosyl-L-homocystéine a été ajouté à la protéine. Après une incubation de 2h, la protéine a été filtrée et dialysée dans deux bains successifs pendant 12h contre un tampon ne contenant pas d'analogue de cofacteur. Différents échantillons ont été préparés dans 5% de  $D_2O$  contenant de la protéine à 300  $\mu M$  et un des potentiels ligands à 5 mM. 200  $\mu M$  de cette solution ont été introduits au fond d'un tube RMN de J.Young et le tube a été sellé avec son bouchon en téflon avant d'être évacué de la boîte à gant. Des tubes contrôle sans la protéine ont été réalisés avec un unique ligand ou une combinaison de ligands.

### **MM-5.2 Spectroscopie de résonance paramagnétique électronique**

Ce travail a été réalisé en collaboration avec Serge Gambarelli et Guisepppe Sicoli, du Service de Chimie Inorganique et Biologique, Institut de Nanosciences et Cryogénie, CEA, Grenoble.

#### **MM-5.2.1 Principe de la résonance paramagnétique électronique (RPE)**

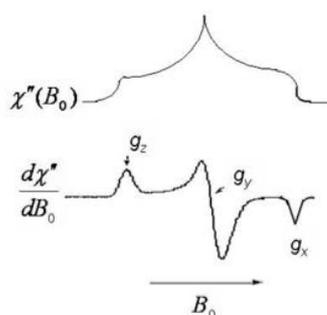
La spectroscopie RPE permet d'étudier des espèces paramagnétiques - avec un spin électronique non nul, c'est-à-dire celles qui possèdent un ou plusieurs électrons non appariés. En biologie, cette technique est utilisée pour caractériser des metalloprotéines présentant des centres organométalliques dans un état redox paramagnétiquement actif ou encore des radicaux organiques libres. Dans une expérience standard de RPE en bande X ( $\sim 9GHz$ ), l'absorption résonante d'une radiation micro-onde a une longueur d'onde donnée est mesurée en faisant varier l'intensité du champ magnétique externe.

L'électron est une particule élémentaire fermionique, avec donc un spin électronique  $S=1/2$ , qui obéit au principe d'exclusion de Pauli. Ainsi dans un atome ou une molécule possédant les couches électroniques remplies (composé diamagnétique), les électrons s'apparient deux à deux tout en gardant les états de spin opposés ( $m_s = +1/2$  pour l'un et  $m_s = -1/2$  pour l'autre). Le moment magnétique électronique de l'ensemble est alors nul. Par contre, la présence d'un électron célibataire dans le cas de radicaux libres et de certains ions de métaux de transition, ou la présence de plusieurs électrons célibataires, dans le cas de centres métalliques plus complexes, confère à l'ensemble un moment magnétique électronique  $S$  non nul. Quand ce système est soumis à un champ magnétique extérieur, les niveaux d'énergie d'un spin  $S$  se séparent, sous l'effet Zeeman, en  $(2S + 1)$  états. Cette levée de dégénérescence d'états de spin électronique est identique à la levée de dégénérescence d'états de spin nucléaire exploitée dans la RMN [MM-5.1]. Il est possible d'induire des transitions entre les niveaux d'états de spin en utilisant un rayonnement micro-onde si l'énergie du photon incident correspond à la différence d'énergie entre les états de spin - condition de résonance.

$$h\nu = \Delta E = E_{m=+1/2} - E_{m=-1/2} = g\mu_B B_0$$

La condition de résonance paramagnétique électronique se résume dans l'équation ci-dessus avec la constante de Planck  $h$ ,  $\nu$  est la fréquence du photon absorbé,  $g$  est le facteur de Landé (dans le cas d'un électron libre sa valeur est  $g_e = 2.0023$ ),  $\mu_B$  est le magnéton de Bohr et  $B_0$  est l'intensité du champ magnétique externe.

Le facteur  $g$  est l'équivalent du déplacement chimique  $\delta$  en RMN. Sa valeur dépend de la nature du centre paramagnétique et de son environnement moléculaire. L'acquisition des données se fait à des températures cryogéniques proches de la température d'hélium liquide, pour minimiser le phénomène de relaxation. Dans un liquide rapidement congelé, l'orientation des molécules est aléatoire et aucun mouvement n'est possible, le spectre enregistré est donc un "spectre de poudre". De ce fait le facteur  $g$  est anisotrope, il prend la forme d'une grandeur tensorielle définie par les composantes  $g_x$ ,  $g_y$  et  $g_z$ . Ces trois composantes correspondent aux points d'inflexions et à l'extremum du spectre d'absorption enregistré au cours du balayage des valeurs du champ magnétique [Figure MM-119].



**Figure MM-119 :** Signal d'absorption  $\chi''$  et sa dérivée  $d\chi''/dB_0$  obtenus en faisant varier le champ  $B_0$ . La dérivée du signal d'absorption permet de déterminer les grandeurs tensorielles  $g_x$ ,  $g_y$  et  $g_z$  qui sont caractéristiques du centre paramagnétique étudié.

Les données enregistrées représentent la dérivée du signal d'absorption. Ainsi après une intégration des données brutes on obtient un spectre d'absorption en fonction de l'intensité du champ magnétique. Une seconde intégration de ce spectre nous donne la quantité de spin dans l'échantillon.

### ***MM-5.2.2 Expérience de RPE avec HydA1 Cr Nstrep repurifiée***

A la fin de la maturation dans l'extrait, le mélange réactionnel est clarifié par centrifugation à 20 000 g pendant 10 minutes, puis le surnageant est filtré et chargé sur une colonne de streptactine improvisée, équilibrée dans 100 mM Tris, 300 mM NaCl, pH 8.0, 5% glycérol. Cette colonne est fabriquée à partir d'une pipette de 10 mL, dont la partie conique a été enlevée et remplacée par un filtre centricon. L'avantage de cette colonne improvisée par rapport à une colonne à flux gravitationnel classique est dans l'augmentation de la hauteur de la résine sans avoir à augmenter son volume. En effet, nous avons observée que la protéine HydA1 Cr N-strep a la tendance de migrer lentement avec le flux lors du lavage. Une plus grande hauteur de la résine permet d'optimiser sa séparation des protéines contaminants tout en limitant les pertes de la protéine d'intérêt. L'élution suit un protocole standard de purification par chromatographie d'affinité sur une colonne de Strep-Tactine [MM-3.2.2].

Une fois l'holoenzyme HydA1 Cr N-strep purifiée, elle est concentrée dans un centricon de MWCO de 30 kDa à 80  $\mu$ M. Un volume de 100  $\mu$ L de cette solution est ensuite injecté au fond d'un tube RPE à l'aide d'une seringue Hamilton et d'une longue aiguille. Le tube RPE est congelé à l'intérieur de la boîte à gants en utilisant un bloc d'aluminium disposant d'un puits dont le diamètre est identique à celui du tube RPE. Le bloc d'aluminium est d'abord refroidi dans l'azote liquide à l'extérieur de la boîte à gants. Il est ensuite rapidement transféré à l'intérieur via le sas. Le tube RPE est inséré dans le puits et l'ensemble est évacué de la boîte à gants. A la fin de l'opération, le tube est transféré et stocké dans l'azote liquide.

## **MM-5.3 Spectroscopie HYSCORE**

### ***MM-5.3.1 Principe de la spectroscopie HYSCORE***

La spectroscopie HYSCORE (en anglais HYperfine Sublevel CORRElation) est une technique de RPE pulsée à deux dimensions, qui permet de sonder l'environnement immédiat d'un centre paramagnétique en visualisant les couplages hyperfins entre certains noyaux et le spin étudié. Les couplages hyperfins  $a_n$  résultent des interactions entre un spin électronique et le moment gyromagnétique faible des noyaux possédant un spin nucléaire non nul. Son analogue en RMN serait le couplage scalaire J.

Un spectre HYSORE est composé de deux cadrans. Sur la diagonale du quadrant nommé (+,+), on trouve les noyaux dits « distants », caractérisés par des constantes hyperfines très faibles. En dehors des diagonales des quadrants (-,+) et (+,+), on trouve respectivement les couplages forts ( $|a|/2 > v_n$ ) et faibles ( $|a|/2 < v_n$ ). Par exemple, la présence d'un noyau de  $^{14}\text{N}$  ( $I = 1$ ) couplé à un électron non apparié du centre [Fe-S] se manifestera par l'apparition d'un couplage fort.

### **MM-5.3.2**    *Expérience de spectroscopie HYSORE avec HydA1 Cr Nstrep repurifiée*

Idem que pour l'expérience de RPE [MM-5.2].

### **MM-5.4**    **Spectroscopie de masse**

Ce travail a été réalisé en collaboration avec Olivier Berteau et Alhosna Benjdia, de l'Institut Micalis du centre Inra de Jouy-en-Josas. Les analyses des produits de la réaction catalysée par HydE sont effectuées sur un spectromètre de masse à temps de vol Voyager DE PRO (Appera, France) couplé d'un réflectron, c'est-à-dire un miroir électrostatique.

Lors d'une analyse par un spectromètre MALDI-TOF (*Matrix-Assisted Laser Desorption/Ionisation Time-of-Flight mass spectrometry*), les analytes sont d'abord cristallisés avec de l'acide  $\alpha$ -cyano-4-hydroxycinnamique sur une plaque d'acier. Les composés sont ensuite ionisés et éjectés de la matrice à l'aide d'un pulse de laser. Le détecteur est un analyseur de temps de vol équipé d'un miroir à ions. A l'entrée de l'analyseur tous les ions reçoivent la même énergie cinétique ( $E_c = 1/2 mv^2$ ), ainsi le temps qu'un ion donné va mettre à traverser l'analyseur est directement proportionnel à son rapport  $m/z$ . L'ajout d'un miroir électrostatique au bout de l'analyseur permet de doubler le trajet à parcourir et à améliorer donc la séparation des molécules.

## MM-6 Références des Matériels et Méthodes

- 1 Kuchenreuther, J. M., George, S. J., Grady-Smith, C. S., Cramer, S. P. & Swartz, J. R. Cell-free H-cluster Synthesis and [FeFe] Hydrogenase Activation: All Five CO and CN- Ligands Derive from Tyrosine. *PLoS ONE* **6**, doi:10.1371/journal.pone.0020346 (2011).
- 2 Kuchenreuther, J. M. *et al.* High-Yield Expression of Heterologous [FeFe] Hydrogenases in Escherichia coli. *PLoS ONE* **5**, doi:10.1371/journal.pone.0015491 (2010).
- 3 Kuchenreuther, J. M. ENGINEERING IN VIVO AND IN VITRO SYSTEMS FOR EXPRESSION AND ACTIVATION OF [FeFe] HYDROGENASES. *ENGINEERING IN VIVO AND IN VITRO SYSTEMS FOR EXPRESSION AND ACTIVATION OF [FeFe] HYDROGENASES* (2010).
- 4 Rubach, J. K., Brazzolotto, X., Gaillard, J. & Fontecave, M. Biochemical characterization of the HydE and HydG iron-only hydrogenase maturation enzymes from *Thermatoga maritima*. *FEBS Letters* **579**, 5055-5060, doi:10.1016/j.febslet.2005.07.092 (2005).
- 5 Zheng, L., White, R. H., Cash, V. L., Jack, R. F. & Dean, D. R. Cysteine desulfurase activity indicates a role for NIFS in metallocluster biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 2754-2758 (1993).
- 6 Boyer, M. E., Stapleton, J. A., Kuchenreuther, J. M., Wang, C. w. & Swartz, J. R. Cell-free synthesis and maturation of [FeFe] hydrogenases. *Biotechnology and Bioengineering* **99**, 59-67, doi:10.1002/bit.21511 (2008).
- 7 Schwartz, C. J. *et al.* IscR, an Fe-S cluster-containing transcription factor, represses expression of Escherichia coli genes encoding Fe-S cluster assembly proteins. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 14895-14900, doi:10.1073/pnas.251550898 (2001).
- 8 Giel, J. L., Rodionov, D., Liu, M., Blattner, F. R. & Kiley, P. J. IscR-dependent gene expression links iron-sulphur cluster assembly to the control of O<sub>2</sub>-regulated genes in Escherichia coli. *Molecular Microbiology* **60**, 1058-1075, doi:10.1111/j.1365-2958.2006.05160.x (2006).
- 9 Elliott, J. I. & Brewer, J. M. The inactivation of yeast enolase by 2,3-butanedione. *Archives of Biochemistry and Biophysics* **190**, 351-357, doi:10.1016/0003-9861(78)90285-0 (1978).
- 10 Bradford, M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Analytical Chemistry* **72**, doi:10.1021/ac50005a713 (1976).
- 11 Fish, W. W. Rapid colorimetric micromethod for the quantitation of complexed iron in biological samples. *Methods in Enzymology* **158**, 357-364 (1988).
- 12 Beinert, H. Semi-micro methods for analysis of labile sulfide and of labile sulfide plus sulfane sulfur in unusually stable iron-sulfur proteins. *Analytical Biochemistry* **131**, 373-378, doi:10.1016/0003-2697(83)90186-0 (1983).
- 13 Kuchenreuther, J. M., Britt, D. R. & Swartz, J. R. New Insights into [FeFe] Hydrogenase Activation and Maturase Function. *PLoS ONE* **7**, doi:10.1371/journal.pone.0045850 (2012).
- 14 Park, D. H., Laivenieks, M., Guettler, M. V., Jain, M. K. & Zeikus, J. G. Microbial utilization of electrically reduced neutral red as the sole electron donor for growth and metabolite production. *Applied and environmental microbiology* **65**, 2912-2917 (1999).
- 15 Warkentin, M., Berejnov, V., Hussein, N. S. & Thorne, R. E. Hyperquenching for protein cryocrystallography. *Journal of Applied Crystallography* **39**, 805-811, doi:10.1107/s0021889806037484 (2006).
- 16 Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography* **67**, 235-242, doi:10.1107/s0907444910045749 (2011).
- 17 Kabsch, W. XDS. *Acta Crystallographica Section D: Biological Crystallography* **66**, 125-132, doi:10.1107/s0907444909047337 (2010).
- 18 Evans, P. R. An introduction to stereochemical restraints. *Acta Crystallographica Section D: Biological Crystallography* **63**, 58-61, doi:10.1107/s090744490604604x (2006).

- 19 Schüttelkopf, A. W. & van Aalten, D. M. F. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta crystallographica. Section D, Biological crystallography* **60**, 1355-1363, doi:10.1107/s0907444904011679 (2004).
- 20 Mayer, M. & Meyer, B. Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angewandte Chemie International Edition* **38**, 1784-1788, doi:10.1002/(sici)1521-3773(19990614)38:12<1784::ai@-anie1784>3.0.co;2-q (1999).
- 21 Mayer, M. & Meyer, B. Group Epitope Mapping by Saturation Transfer Difference NMR To Identify Segments of a Ligand in Direct Contact with a Protein Receptor. *Journal of the American Chemical Society* **123**, doi:10.1021/ja0100120 (2001).



## **ANNEXES**







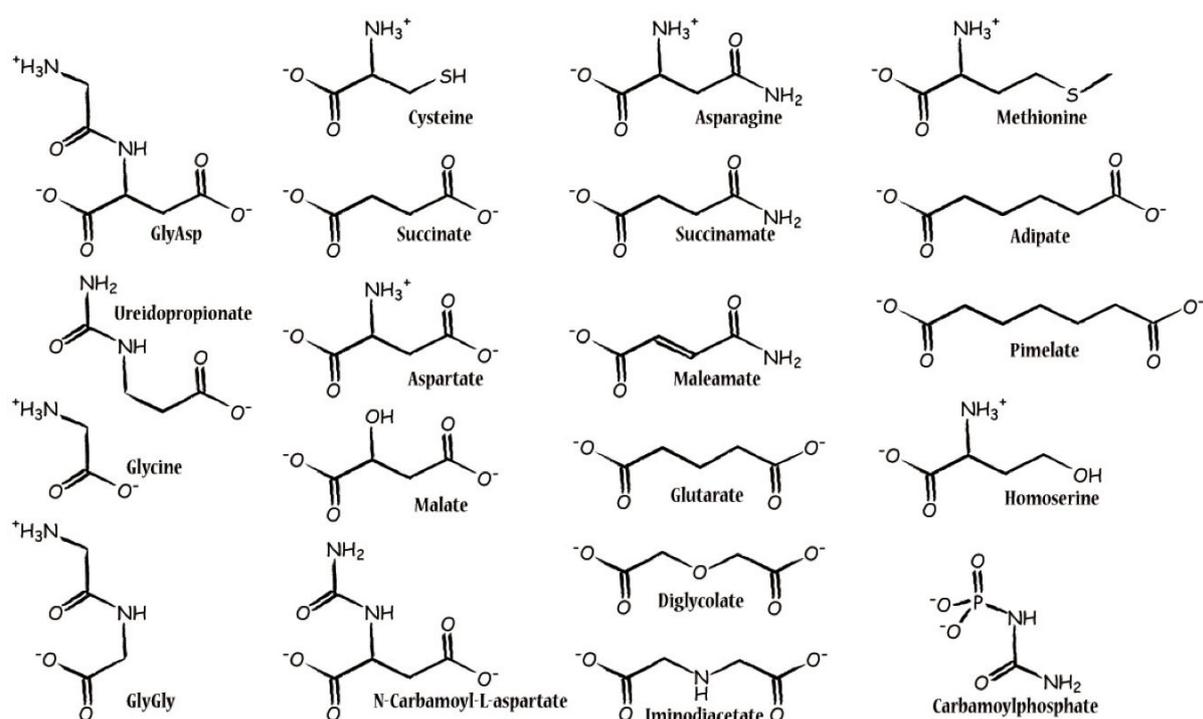




**Figure A-120:** Alignement de séquences de protéines HydE issues d'organismes différents (pages precedentes)

Les éléments de structure secondaire et la numérotation des résidus sont ceux de la forme native de *TmHydE* (pdb: 3CIW). Les losanges pleins de couleur brune marquent la position des résidus cystéines impliqués dans la fixation du centre [Fe-S] SAM radical, qui est conservé dans les protéines à radical SAM. Les losanges bruns vides marquent la position des résidus cystéines impliqués dans la fixation du centre [Fe-S] auxiliaire de HydE. Cet agrégat n'est que partiellement conservé dans le groupe de protéines HydE. Les résidus strictement conservés sont surlignés en rouge et ceux qui présentent une conservation partielle ou une conservation de propriétés physicochimiques sont en lettre rouge. Les résidus qui ont été mentionnés dans ce manuscrit sont marqués d'un triangle noir. Cet alignement a été obtenu à partir de 146 séquences de HydE issues d'une grande variété d'organismes. L'alignement a été effectué avec le logiciel ClustalW. Les 77 séquences visualisées ont été choisies avec soins dans le but de minimaliser le nombre de séquences dans l'affichage tout en préservant la diversité des séquences.

Sur la page précédente se trouvent également deux vues orthogonales d'une visualisation de la conservation de résidus à travers les différentes séquences de HydE alignées. Le pourcentage de conservation de chaque résidu est projetée sur la structure de *TmHydE* native (pdb: 3CIW). Les positions des résidus les moins conservés (<18%) sont colorées en bleu foncé et celles des résidus les mieux conservés (>95%) sont colorées en rouge vif (voir spectre en bas de l'image).



**Figure A-121:** Schéma de différents ligands potentiels de HydE testés lors des premiers essais de cristallisation



# Carbon-Sulfur Bond-Forming Reaction Catalyzed by the Radical SAM Enzyme HydE

Roman Rohac<sup>1</sup>, Patricia Amara<sup>1</sup>, Alhosna Benjdia<sup>2,3</sup>, Lydie Martin<sup>1</sup>, Pauline Ruffié<sup>2,3</sup>, Adrien Favier<sup>4</sup>, Olivier Berteau<sup>2,3</sup>, Jean-Marie Mouesca<sup>5,6</sup>, Juan C. Fontecilla-Camps<sup>1\*</sup> and Yvain Nicolet<sup>1\*</sup>

<sup>1</sup> Metalloproteins Unit, Institut de Biologie Structurale, CEA, CNRS, UGA, 71, Avenue des Martyrs, 38044, Grenoble Cedex 9, France

<sup>2</sup> INRA, ChemSyBio, UMR 1319 Micalis, F-78350 Jouy-en-Josas, France

<sup>3</sup> AgroParisTech, ChemSyBio, UMR Micalis, F-78350 Jouy-en-Josas, France

<sup>4</sup> Biomolecular NMR Spectroscopy Group, Institut de Biologie Structurale, CEA, CNRS, UGA, 71, Avenue des Martyrs, 38044, Grenoble Cedex 9, France

<sup>5</sup> Univ. Grenoble Alpes, INAC, SCIB/LRM, F-38000 Grenoble, France

<sup>6</sup> CEA, INAC, SCIB/LRM, F-38054 Grenoble, France

\*email: yvain.nicolet@ibs.fr, juan.fontecilla@ibs.fr

## Abstract

Carbon-sulfur bond formation at aliphatic positions is a challenging reaction effectively performed by radical *S*-adenosyl-L-methionine (SAM) enzymes. We identified 1,3-thiazolidines as ligands of the radical SAM enzyme HydE involved in the assembly of the active site of [FeFe]-hydrogenase and we demonstrate that these ligands are also substrates. We used X-ray crystallography, *in vitro* assays and NMR spectroscopy to follow a radical-based reaction that is best described as the formation of a C-centered radical, which concomitantly attacks a thioether S. Theoretical calculations based on our high-resolution structures allowed us to follow the evolution of the electronic structure to form *S*-adenosyl-L-cysteine. Our results suggest that, at least in our case, the widely proposed and highly reactive 5'-deoxyadenosyl radical species, which triggers the reaction in radical SAM enzymes is not an isolable intermediate.

Free radical chemistry provides high-energy intermediates and thus can participate in pathways not easily accessible to standard two-electron chemistry. In Nature, free radicals are more commonly present than initially expected and living organisms have adapted to both prevent their deleterious effects and take advantage of their powerful chemistry. Over the last twenty years, radical *S*-adenosyl-L-methionine (SAM) enzymes have emerged as a new superfamily<sup>1</sup> containing over 110,000 amino acid sequences and catalyzing more than 70 different reactions. Radical-based chemistry is known to be used in the three domains of life<sup>2</sup>. As a general mechanism, radical SAM enzymes use a reduced [Fe<sub>4</sub>S<sub>4</sub>]<sup>1+</sup> cluster to perform the reductive homolytic cleavage of the C5'-Sδ bond of SAM, generating a highly reactive 5'-deoxyadenosyl radical species (5'-dA•), which is then used to abstract a hydrogen atom from substrate. This produces a second radical species that leads to product formation. Radical SAM enzymes such as biotin synthase (BioB)<sup>3</sup>, lipoic acid synthase (LipA)<sup>4</sup> and the methylthiotransferases RimO<sup>5</sup> and MiaB<sup>6</sup> produce a C-centered radical species at specific positions on aliphatic carbon atoms to make a C-S bond. These enzymes possess a second FeS cluster, which carries the sulfur atom to be inserted. In addition, this FeS cluster is essential to trap the unpaired electron produced by the radical-based reaction in order to terminate it<sup>7</sup>. Another radical SAM enzyme, HydE, most likely involved in the synthesis of the azadithiolate bridging molecule of the active site of [FeFe]-hydrogenase has also been

proposed to synthesize C-S bond formation<sup>8-10</sup>. Although some HydE proteins have an accessory FeS cluster, the latter is neither conserved nor required for function<sup>11</sup>.

While searching for the natural substrate(s) of HydE using X-ray crystallography, we came across an active site-bound 1,3-thiazolidine species. This molecule is not the physiological substrate of HydE. However, it is a great model for dissecting the mechanism of a carbon-centered radical reaction during C-S bond formation, without the requirement of an additional FeS cluster. Here, we have combined X-ray crystallography, *in vitro* enzymatic assays, NMR spectroscopy and theoretical calculations, to characterize this reaction. We find it corresponds to the pyruvate-mediated synthesis of *S*-adenosyl-L-cysteine from cysteine and SAM (Scheme 1), and we demonstrate it is similar to the radical-based desulfurization carried out using tin-derivatives in organic synthesis<sup>12,13</sup>. Our findings shed new light on both the chemistry performed by radical SAM enzymes in general and radical-based C-S bond formation in particular.

## Results

**Identification of (2*R*,4*R*)-2-Methyl-1,3-thiazolidine-2,4-dicarboxylic acid as a HydE ligand.** It has been shown that co-expression in *Escherichia coli* (*E. coli*) of HydE with the other maturases HydF and HydG and HydA leads to active [FeFe]-hydrogenase<sup>14</sup>. Consequently, while searching for the HydE substrate we concentrated in the sulfur-containing metabolites present in *E. coli*. Among these, cystine, the disulfide-bound cysteine dimer of formula (SCH<sub>2</sub>CH(NH<sub>2</sub>)CO<sub>2</sub>H)<sub>2</sub>, was particularly appealing because, like the azadithiolate hydrogenase ligand, it contains two sulfur atoms. We were able to collect 1.47 Å resolution diffraction data from a crystal of SAM-bound *Thermotoga maritima* (*Tm*) HydE, grown in the presence of 25 mM cystine. The corresponding electron density map displayed an unprecedented peak in the active site that we could model as a (2*R*,4*R*)-2-methyl-1,3-thiazolidine-2,4-dicarboxylic acid ((2*R*,4*R*)-MeTDA). Although this finding may initially seem unexpected it is known that i) in alkaline aqueous solutions containing traces of transition metals, cystine can decompose into pyruvate, cysteine and hydrogen sulfide<sup>15,16</sup> and ii) cysteine readily reacts with carbonyls from aldehydes or ketones to produce the corresponding 1,3-thiazolidines<sup>17-20</sup> (Fig. 1). Because it is likely that our crystallization solutions contained traces of ferrous or ferric iron coming from the HydE FeS clusters, we hypothesized that the two reactions described above led to the formation of the two diastereoisomers (2*S*,4*R*)-MeTDA and (2*R*,4*R*)-MeTDA. Subsequently, we confirmed this condensation using <sup>1</sup>H-NMR spectroscopy and, as expected from the literature<sup>18-20</sup>, we observed the complete conversion of an aqueous 1:1 mixture of L-cysteine and pyruvate into (2*S*,4*R*)-MeTDA and (2*R*,4*R*)-MeTDA in a 2:8 ratio (Figs. 1d and 1e). Soaking of SAM-bound HydE crystals with a solution containing such a 1:1 mixture of L-cysteine and pyruvate generated the same electron density as observed for the cystine-soaked crystal (crystal **XI**; Table 1), thus confirming our initial assignment of the unknown ligand as (2*R*,4*R*)-MeTDA (Fig. 1a) and also indicating that the protein stereo-specifically selects only this epimer for binding. This was very convenient because as previously shown<sup>21-23</sup>, 2-substituted 1,3-thiazolidines undergo epimerization at C2 in water, rendering purification of one particular epimer difficult (Supplementary Fig. S1a).

(2*R*,4*R*)-MeTDA sits in the large active site cavity of HydE close to SAM, where the physiological substrate is expected to bind (Fig. 1b)<sup>11</sup>. It only fills a small portion of the accessible volume of the very large HydE inner cavity, which is otherwise occupied by several water molecules. Its two carboxylate moieties are located in the previously identified anion-binding **S1** and **S2** sites, close to the highly conserved residue R159<sup>11</sup>. The carboxylate from the cysteine fragment sits in the **S1** site but does not contact the guanidinium moiety of R159 (Fig. 1b). Instead, it establishes hydrogen bonds with the main chain nitrogen atom of

A270 and the phenoxy oxygen of the strictly conserved Y306. The S1 and C5 atoms of the thiazolidine ring are respectively at 3.9 Å and 4.2 Å from the SAM C5' atom (Scheme 1). Consequently, the closest C5-bound hydrogen atom should be at about 3.7 Å from C5', a distance suitable for potential H-atom abstraction. The (C5', S1, C2) angle is about 165° indicating that these three atoms are roughly aligned. Conversely, the (C5', S1, C5) angle is about 90°. In addition, the methyl group of (2*R*,4*R*)-MeTDA sits in a hydrophobic pocket facing the respectively highly and strictly conserved I56 and M291. The combination of this hydrophobic pocket and the anion-binding site **S2** (Fig. 1b) is thus likely responsible for the selection by the protein of the 2*R*-epimer instead of its 2*S* counterpart. This diastereoisomeric specificity of the protein cavity allowed us to use the 1:1 mixture of L-cysteine and pyruvate for ligand preparation in all our subsequent studies (see Methods). In order to probe the factors responsible for ligand binding, we co-crystallized HydE with several analogs of (2*R*,4*R*)-MeTDA (Table 1 and Supplementary Methods). All our thiazolidine derivatives bind to the protein (Supplementary Fig. S3). Neither the methyl moiety nor the carboxylate group at C2 are required for ligand binding, as shown by HydE complexes with (2*R*,4*R*)-1,3-thiazolidine-2,4-dicarboxylic acid ((2*R*,4*R*)-TDA) and 4*R*-1,3-thiazolidine-4-carboxylic acid (4*R*-TCA), respectively (crystals **X7** and **X8** in Table 1; Supplementary Figs. S3a and S3b). Both molecules bind the protein as (2*R*,4*R*)-MeTDA does. Again, the enzyme selects only the (2*R*,4*R*)-TDA epimer through the binding of the C2 carboxylate at site **S2**. Replacing the H atoms at C5 with two methyl groups does not prevent (2*R*,4*R*)-2,5,5-trimethyl-1,3-thiazolidine-2,4-dicarboxylic acid ((2*R*,4*R*)-TMeTDA) from binding either (crystal **X9**; Table 1; Supplementary Fig. S3c). The same 2*R*- toward 2*S*- epimer selection is observed. However, steric clashes between the additional methyl groups at C5 and atom O4' from the ribose moiety of *S*-adenosyl-L-homocysteine push the ligand into the cavity by about 0.5 Å when compared to (2*R*,4*R*)-MeTDA (crystals **XI** and **X9**; Table 1). Conversely, neither L-proline (sulfur to carbon substitution compared to 4*R*-TCA) nor L-cysteine bind to HydE crystals even at high concentrations (100 mM), suggesting the requirement of a thioether for stable binding. This latter observation does not support a previous functional analysis on HydE where L-cysteine was modeled bound close to SAM<sup>24</sup>.

**Reactivity of HydE toward (2*R*,4*R*)-MeTDA in solution.** Betz and co-workers have recently reported that small thiol-containing molecules such as dithiothreitol (DTT), L- or D-cysteine and coenzyme M are substrates of HydE. They all stimulate SAM cleavage to methionine and 5'-deoxyadenosine (5'-dA) through the abstraction of an exchangeable hydrogen atom, presumably from their thiol moiety<sup>24</sup>. The discrepancy between their measured activity with L-cysteine and the absence of its binding in the *Tm*HydE crystals is puzzling. Perhaps cysteine, and other thiols like DTT diffuse through the wide active site cavity and transiently interact with 5'-dA• rather than binding to a specific site. Although 1,3-thiazolidines are good HydE ligands and they contain the required atoms for the synthesis of at least parts of the azadithiolate ligand, a few preliminary experiments have shown they do not enhance [FeFe]-hydrogenase activation. Nevertheless, we decided to take advantage of the observed binding of 1,3-thiazolidines to HydE to carry out reactivity studies including their comparison with free thiol-containing molecules. We first determined and compared the rates of 5'-dA produced by HydE from cysteine, DTT and (2*R*,4*R*)-MeTDA (Fig. 2a). As previously described<sup>24</sup>, both DTT and cysteine stimulate 5'-dA production. However, cysteine is not as good a substrate as DTT confirming it is not a specific ligand of HydE. Intriguingly, although a good specific ligand, (2*R*,4*R*)-MeTDA does not significantly enhance 5'-dA formation either; instead, it reduces it by about a half in the presence of DTT when compared to DTT alone (Fig. 2a). This suggests that, as might have been expected, DTT and (2*R*,4*R*)-MeTDA compete for the same region of the active site cavity.

Almost no hydrogen abstraction by 5'-dA• occurred with (2*R*,4*R*)-MeTDA. We instead detected an unknown compound with UV-visible properties similar to adenosyl derivatives and which production over time depended on the presence of HydE and sodium dithionite (DTH) (Figs. 2b and c). This compound was identified as *S*-adenosyl-L-cysteine (SAC) (M=370.38) from MALDI-TOF MS analyses using either (2*R*,4*R*)-MeTDA or a <sup>13</sup>C<sub>3</sub>, <sup>15</sup>N-labeled L-cysteine derived (2*R*,4*R*)-MeTDA (*m/z* of 370.98 and 374.90 respectively) (Fig. 2d). Furthermore, a commercial SAC authentic gave the same UV-visible spectrum and retention time by HPLC analysis (Fig 2c). The mass shift of 4 Da is consistent with three <sup>13</sup>C atoms and one <sup>15</sup>N atom originating from the isotopically labeled cysteine. This, in turn, strongly supports SAC as the HydE product resulting from the condensation of the 5'-dA moiety of SAM and the cysteinyl moiety of (2*R*,4*R*)-MeTDA (Scheme 1). Both (2*R*,4*R*)-TDA and 4*R*-TCA also produced SAC (Fig. 2c). Conversely, when using (2*R*,4*R*)-TMeTDA (Supplementary Fig. S3c), no additional adenosyl derivative was detected indicating that the observed steric clashes may impede 5'-dA• from approaching the ligand, thus preventing this molecule from being a substrate for HydE.

**Reactivity of HydE in the crystals.** In order to further characterize this novel reaction, we took advantage of the robustness and the high-resolution diffraction quality (1.5 to 1.2 Å resolution) of our *TmHydE* crystals. Furthermore, the relative sluggishness of the reaction allowed us to record time-resolved snapshots (Table 1 and Supplementary Table 1). Thus, we investigated SAC formation directly in our crystals (Supplementary methods). For practical reasons, we used selenocysteine instead of cysteine in order to produce Se-substituted (2*R*,4*R*)-MeTDA ((2*R*,4*R*)-MeSeTDA), as a suitable alternative substrate. Indeed, selenium has a strong anomalous signal and its use instead of sulfur improved both the crystallographic observations and theoretical calculations (see below). As a control, we performed the same experiment using (2*R*,4*R*)-MeTDA and although the use of both chalcogens led to the same result selenium-based reactions were faster (Table 1, crystals **X4** and **X5**; Fig. 3 and Supplementary Fig. S5). The resulting structure when using (2*R*,4*R*)-MeTDA is identical to a control obtained when co-crystallizing *TmHydE* with methionine and authentic SAC (crystal **X6**; Table 1; Supplementary Fig. S5).

Using an excess of 12 mM DTH as reducing agent, we triggered the radical-based reaction directly in our crystals prior to flash cooling at different times (see Methods and Supplementary methods). We thus obtained a series of X-ray structures corresponding to the succession of cross-faded events during the reaction that we followed using double difference  $F_{O_{tx}}-F_{O_{t0}}$  electron density maps ( $F_{O_{tx}}$  corresponds to structure factors of a dataset at a particular time and  $F_{O_{t0}}$  to those of the starting structure; Fig. 3). We first observed that all the structural changes are restricted to a sphere of about 12 Å radius around C5'. It contains the [Fe<sub>4</sub>S<sub>4</sub>] cluster, SAM, (2*R*,4*R*)-MeSeTDA and the Y(303)NLY(306) loop that includes residue L305, previously described as interacting with the methyl group of SAM<sup>11</sup>. The rest of the protein remains unaffected by the reaction. These structural changes correspond to: i) the SAM C5'-Sδ bond breakage (initial/final distance: 1.81/3.91 Å); ii) the formation of a C5'-Se1 bond (3.91/2.08 Å) between 5'-dA and (2*R*,4*R*)-MeSeTDA; iii) the cleavage of the Se1-C2 bond of the thiazolidine ring and iv) the disappearance of the (2*R*,4*R*)-MeSeTDA pyruvate fragment with the concomitant binding of a chloride ion at anion-binding site **S2** (Figs. 3b, c and d). All the double difference  $F_{O_{tx}}-F_{O_{t0}}$  electron density maps generated at intermediate times (Figs. 3c and d) show a mixture of only two states: t=0 (crystal **X2**; Table 1; Fig. 3a) and t<sub>final</sub> (crystal **X4**; Table 1; Fig. 3b) with ratios varying from 1:0 to 0:1 with time. This suggests that the limiting step of the reaction is SAM cleavage and 5'-dA• generation.

**C-S bond formation.** Our results indicate that instead of abstracting an H-atom the radical SAM enzyme HydE reacts with 1,3-thiazolidines directly on their chalcogen atom. The only other radical SAM reaction that does not involve H abstraction described so far is the addition of 5'-dA• to a double bond, catalyzed by MqnE, a protein implicated in the biosynthesis of menaquinone (vitamin K2)<sup>25</sup>. In order to gain more information about the C-S bond formation catalyzed by HydE we investigated the steps after SAM cleavage to 5'-dA• and methionine, using quantum mechanical (QM) and hybrid quantum mechanical/molecular mechanical (QM/MM) calculations including the protein matrix effect (see Supplementary methods for details). Small models of the 5'-dA• attack to Se (Fig. 4a, top) and S (Supplementary Fig. S7) show that (i) the formation of the C5'-Se1(S1) bond and the cleavage of the Se1(S1)-C2 bond are concomitant and lead to the formation of a C2-centered radical, delocalized over the adjacent N and carboxylate C atoms (Fig. 4a); (ii) the energy cost is small (Supplementary Table S2); the reaction is facilitated in the case of the (2*R*,4*R*)-MeSeTDA because the Se orbitals are more diffuse than those of S, resulting in a better overlap with the neighboring radical on C5' and (iii) this homolytic substitution (S<sub>H</sub>2) involves the interaction between one of the Se1 (S1) lone pairs and the radical electron (HOMO  $\alpha$ )<sup>26</sup> as reflected by the corresponding lowest unoccupied molecular orbital (electron hole, LUMO  $\beta$ ) of the reactant (Fig. 4a (top) and Supplementary methods for the choice of following spin/charge transfers via LUMO  $\beta$  representations). The other lone pair orbital is passive being perpendicular to the plane defined by C5, Se1 (or S1) and C2. This reaction is reminiscent of the reductive desulfurization of thioethers or N,S-acetals catalyzed by tin hydride derivatives<sup>13</sup>. S<sub>H</sub>2 reactions in solution usually favor a backside reaction over a frontside one (Supplementary Fig. S8). This is what we see with HydE. We also tested computationally the possibility of cleaving the Se1(S1)-C5 bond with a small model where the radical C5' attacks along the direction of this bond. The LUMO  $\beta$  of the C5'• approaching (2*R*,4*R*)-MeSeTDA shows that there is an energy match between the orbital of the C5'•, one of the orbitals of the Se lone pair and *both* orbitals of the C5 (frontside reaction) and C2 (backside reaction) atoms (Supplementary Fig. S9). The resulting C2-centered radical is expected to be much more stable than the one on C5 due to extra stabilizations by the neighboring carboxylate and amine. However, when C5' approaches further, the alignment between C5', Se1 and the C5 atoms results in a maximum overlap and the Se1-C5 bond is cleaved. This shows that, in the protein, the binding mode of (2*R*,4*R*)-Me(Se)TDA imposes both the radical attack along the Se1(S1)-C2 direction and the cleavage of this bond following a backside mechanism.

**Reaction on (2*R*,4*R*)-Me(Se)TDA.** Our crystallographic data show that the largest structural changes during the reaction take place at the 5'-dA moiety, which moves toward Se1 as this atom approaches C5'. Interestingly, the 5'-dA moiety moves to the position previously described when no substrate was present in the active site<sup>27</sup>, suggesting that 5'-dA• may not be able to move any further (Fig. 3e). Conversely, when substituting SAM with 5'-dA and methionine in the presence of (2*R*,4*R*)-MeSeTDA (crystal X3; Table 1), steric hindrance between H<sub>3</sub>C5' and Se1 prevents such large movement. Consequently, the 5'-dA sits in between its position in SAM and in SAC (SeAC). The 5'-dA movement can be best described as a rotation of about 6° centered about the N6 atom of the adenine moiety along with a rotation of 26° about its N9-C1' bond (Fig. 3e). As N6 is the atom farthest from C5', these displacements, while minimizing the motion of 5'-dA• as a whole, maximize the movement of C5' toward (2*R*,4*R*)-Me(Se)TDA. The radical is thus expected to react with its substrate after a remarkably short journey of less than 1.8 Å. Our calculations including the protein matrix (Fig. 4a, bottom and Supplementary Table S2) confirmed that the reaction of 5'-dA• on (2*R*,4*R*)-MeSeTDA is facile (8.2 kJ.mol<sup>-1</sup>). This barrier is small compared to the energy

cost for the first step of the reaction (SAM cleavage to 5'-dA•) that we previously estimated at 54.0 kJ.mol<sup>-1</sup> without substrate<sup>27</sup>. In the presence of substrate, this latter barrier is expected to be lowered, as it was estimated in LAM<sup>28</sup>. In addition, we observe that the radical on C5' is partially delocalized over Se1 and C2, in contrast to what is observed in the models without the protein environment (see LUMO  $\beta$  and HOMO  $\alpha$  of the QM/MM (QM) reactant in Figs. 4a (bottom and top), respectively). Calculations including the [Fe<sub>4</sub>S<sub>4</sub>]<sup>1+</sup> cluster in the QM part (see Supplementary methods) show that even before SAM cleavage there is nearly orbital degeneracy between the reduced cluster and (2*R*,4*R*)-MeSeTDA (Fig. 4b). Modeling 5'-dA•, including the oxidized [Fe<sub>4</sub>S<sub>4</sub>]<sup>2+</sup> in the QM part, shows that contrary to small models, the protein matrix pushes together 5'-dA•, (2*R*,4*R*)-MeSeTDA and the [Fe<sub>4</sub>S<sub>4</sub>] cluster (Fig. 4c), strongly supporting 5'-dA• as part of a conjugated transition state rather than as an isolable intermediate.

**Radical reaction termination.** Our calculations indicate that the reaction leads to an intermediate species with a free radical located at the C2 atom. The first step is reminiscent of a radical-based desulfurization<sup>13</sup>. Yet, unlike reductive desulfurization, which usually takes place in solution and regenerates the radical propagator<sup>12,13</sup>, our reaction takes place in the buried active site of an enzyme and cannot propagate. Indeed, there is no obvious candidate to donate a hydrogen atom to the C2-centered radical intermediate. Conversely, radical SAM enzymes are known to catalyze oxidations<sup>29-31</sup>. As C2 sits less than 12 Å away from the oxidized [Fe<sub>4</sub>S<sub>4</sub>]<sup>2+</sup> cluster, its unpaired electron could be transferred back to the cluster as in the DesII enzyme, which uses a catalytic electron to perform multiple turn-overs<sup>29</sup>. We did not succeed in performing their experiment with HyDE. However, modeling the C2-centered radical with the oxidized [Fe<sub>4</sub>S<sub>4</sub>]<sup>2+</sup> included in the QM part reveals an orbital energy match between both species, suggesting that minor electronic or conformational changes will help transferring the electron back to the cluster (Supplementary Fig. S10).

Both our *in vitro* assays and our crystallographic data confirmed that SAC is the final reaction product. In the crystals, SAC synthesis was coupled to the disappearance of the pyruvate fragment of (2*R*,4*R*)-MeTDA from the electron density maps. We also followed the fate of the pyruvate fragment in a time-resolved NMR spectroscopy experiment using a methyl-<sup>13</sup>C-labeled (2*R*,4*R*)-MeTDA produced from a 1:1 mixture of L-cysteine and <sup>13</sup>C-labeled pyruvate at C3 atom (methyl group) (see Supplementary methods). By triggering the reaction directly in an anaerobic NMR tube, we monitored the consumption of (2*R*,4*R*)-MeTDA by the disappearance of the specific signal from <sup>13</sup>C-labeled methyl carbon protons ( $\delta$  1.56 (s) for (2*R*,4*R*)-MeTDA and  $\delta$  1.71 (s) for (2*S*,4*R*)-MeTDA; Fig. 5a middle and bottom). Simultaneously, we observed the appearance of another signal, which we assigned to pyruvate protons ( $\delta$  2.31 (s) for the  $\alpha$ -ketoacid form;  $\delta$  1.33 (s) for the hydrated form; Fig 5a top and bottom). We thus confirmed that pyruvate is released at the end of the reaction both in our HyDE crystals and in solution.

## Discussion

We can now use our data to propose a plausible mechanism for the conversion of (2*R*,4*R*)-MeTDA into SAC by HyDE (Fig. 5b). Overall, the reaction can be best described as a radical-based pyruvate-mediated synthesis of SAC from cysteine and SAM. The position of the sulfur atom, along with the alignment of the Fe---S $\delta$ -C5'---S1-C2 atoms, favor a concerted mechanism combining electron transfer from the cluster, SAM cleavage, C5'-S1 bond formation and S1-C2 bond breakage. Indeed, these striking atomic arrangements maximize compactness through van der Waals contacts and result in the overlap of these moieties' orbitals. These conditions, necessary for a concerted mechanism, generate near orbital degeneracy along the path from Fe<sub>u</sub> to C2 via S $\delta$ , C5' and Se1 (Figs. 4c and 3e). A subsequent radical oxidation would lead to the formation of a secondary ketimine (Schiff base) between

C2 and N1 (Fig. 5b). Subsequently, water molecules close to (2*R*,4*R*)-MeTDA would hydrolyze this ketimine producing SAC and pyruvate (Fig. 5b). Thanks to its active site pocket structure, HydE can carry out an anoxic oxidative desulfurization that significantly differs from related radical-based chain reactions usually performed during organic synthesis.

We can conclude from both our structural data and theoretical calculations that during the reaction, (2*R*,4*R*)-MeTDA is optimally oriented in the HydE cavity for the 5'-dA• attack on S1. All known X-ray structures of radical SAM enzymes show numerous and conserved hydrogen bonds and van der Waals contacts between the protein and both the ribose and the adenine moieties. These interactions should restrict large movements during SAM cleavage, as shown by a recent ENDOR study on lysine 2,3-aminomutase (LAM). Indeed, Horitani and co-workers reported that “5'-dA• is not 'free' but most probably tightly shepherded by van der Waals contacts”<sup>32</sup>. Furthermore, in all the known structures of radical SAM enzymes in complex with a productive substrate (PDB ID: 1R30, 4K38, 3CB8, 2A5H, 4NJI and 4R33) there is a short distance between C5' and the C (or N) atom which binds the target hydrogen atom (from 3.6 to 4.2 Å)<sup>3,33-37</sup>. Thus, as expected, their active sites are optimized to correctly orientate their substrate leaving no room for unproductive molecular motions. This observation is relevant when discussing the possible existence of a free 5'-dA• in SAM-dependent enzymes. Although it has never been observed in either radical SAM proteins or adenosylcobalamine (B<sub>12</sub>) enzymes, its existence as a transient species is highly probable in the latter<sup>38-40</sup>. Furthermore, the use of 3',4'-anhydro-*S*-adenosyl-*L*-methionine (*anSAM*) instead of SAM allowed for the observation of a stable 5'-anhydrodeoxyadenosyl radical species supporting the existence of 5'-dA• as a fleeting intermediate<sup>40</sup>. However, such an allylic-radical stabilization may not reflect what really happens when using SAM. Combined with the crystallographic results, our calculations on HydE using 1,3-thiazolidines will contribute to this debate. Indeed, both the structural and electronic principles that govern the reaction we observed in HydE should apply to the other radical SAM enzymes. Thus our observation that 5'-dA• is not an isolable species may also be a general feature for these enzymes.

It will be informative now to compare our reaction to C-S bond formation by radical SAM enzymes of known structure, namely BioB<sup>3</sup>, LipA<sup>4</sup> and RimO<sup>5</sup>. These enzymes use 5'-dA• to generate a second C-centered radical, which, in turn, reacts with a sulfur atom. Thiols are not suitable substrates because the favored radical reaction is H-atom abstraction leading to a stable thiyl radical. Instead, these enzymes use sulfur atoms, which are either bound to or are part of FeS clusters. Transition metals such as ferric ions can trap the unpaired electron and, consequently, are maybe a better option than even good organic leaving groups<sup>7,41</sup>. Indeed, FeS clusters can be employed in a large number of sulfur-insertion reactions while using a rather simple sulfur source. However, iron-bound sulfide ions are not easily accessible to organic chemists and the use of thioethers may become a good alternative. In this respect, our data shed useful light onto the reactive properties of thioethers toward C-centered radicals (*i.e.* correct orientation, optimal orbital overlap and a backside mechanism).

We have not yet succeeded in finding the natural substrate of HydE. However, our identification of 1,3-thiazolidines as ligands and serendipitous substrates, combined with the work of Betz and co-workers, gives a comprehensive description of the events that can take place in the vast cavity of this protein<sup>24</sup>. The site where the 1,3-thiazolidine sulfur atom binds may have a physiologically relevant affinity for this element. This would explain the reaction of HydE with numerous thiol- or thioether-containing molecules<sup>24</sup>.

## Conclusion.

Radical chemistry is a very promising alternative to carry out reactions that are difficult otherwise. Among them C-S bond formation is a really challenging problem especially when

it involves inert aliphatic carbon positions. Some radical SAM enzymes overcome this obstacle using the unique combination of a  $[\text{Fe}_4\text{S}_4]$  cluster and SAM to abstract an H-atom at such carbon positions to form a reactive  $\text{C}\cdot$ . Taking advantage of our identification of (2*R*,4*R*)-MeTDA as a substrate for the  $[\text{FeFe}]$ -hydrogenase maturase HydE, we have obtained an accurate description of a C-S bond formation in a radical SAM enzyme. To the best of our knowledge, this is the first time a radical SAM enzyme is described reacting directly on a sulfur (selenium) atom instead of abstracting an H-atom. The studied reaction is related to tin-derivative-based reductive desulfurization on N,S-acetals<sup>13</sup>. The observed tight control of C5' motions is a feature likely to apply to all radical SAM proteins. Radical SAM enzymes may have solved the problem of controlling the highly reactive 5'-dA $\cdot$  by coupling SAM cleavage to the catalyzed reaction, resulting in a single concerted step. Using alternative substrates in HydE will certainly open the way to a better understanding of the radical-based chemistry that uses C-centered radicals. It will also make it possible to discriminate between the aspects of the reaction that intrinsically depend on basic chemistry principles and those that are modulated by the protein matrix. Finally, although we found that 1,3-thiazolidines did not enhance  $[\text{FeFe}]$ -hydrogenase activation, the remarkable fit of (2*R*,4*R*)-MeTDA to the HydE active site cavity suggests that the chemistry described here could be relevant to the  $[\text{FeFe}]$ -hydrogenase maturation process. This certainly needs to be further examined and we hope our work will contribute to this endeavor.

## Methods

**Protein sample preparation.** Three protein samples were used for this work. All the crystallographic analyses were performed using the HydE protein from *T. maritima*. For the *in vitro* functional analyses we used HydE from the mesophilic organism *C. acetobutylicum* in order to increase SAM stability and to facilitate kinetic measurements. This protein is very similar to *TmHydE*<sup>11</sup>. The accessory FeS cluster of *TmHydE* provokes a local heterogeneity in the X-ray structures due to the different compositions this cluster can have<sup>42</sup>. Because it has no reported function in HydE<sup>11</sup>, we decided to remove it by mutating its three cysteine ligands to serine, leading to the *TmHydE* <sup>$\Delta\text{FeS}$</sup>  sample (see Supplementary Methods). As a control, SAC production was assayed using *TmHydE* with and without its second cluster for the different 1,3-thiazolidine ligands identified in this study and no significant differences were found. As standard procedures, *TmHydE* containing its second FeS cluster and *TmHydE* <sup>$\Delta\text{FeS}$</sup>  were overexpressed aerobically and purified under anaerobic conditions as previously described<sup>42</sup>. The only difference was that FeS cluster reconstitution was performed prior to purification, by addition of 5 mM DTT, 1 mM  $\text{Na}_2\text{S}$  and 1 mM  $\text{FeCl}_3$  right after cell disruption, under anaerobic conditions. The crude extract was further incubated 2 hours prior to ultracentrifugation and subsequent purification. For kinetic experiments in the crystals, *TmHydE* <sup>$\Delta\text{FeS}$</sup>  was expressed and purified under anaerobic conditions in order to avoid radical SAM FeS cluster *in vitro* reconstitution and SAM addition (see Supplementary Methods). The complete expression and purification protocol for *CaHydE* is described in detail in the Supplementary Methods.

**Ligand preparation.** L-cystine was solubilized by drop-by-drop addition of 50% NaOH to obtain a 100 mM stock solution. For the standard preparation of the (2*R*,4*R*)-MeTDA ligand stock solution, 60.6 mg L-cysteine were solubilized in 2 mL water (25 mM stock) and this solution was subsequently used to solubilize 1 molar equivalent of pyruvic acid powder. The pH was then adjusted with 2N NaOH/HCl to a 7.5-8.5 range leading to a mixture of 2*R* and 2*S* epimers at C2 with ratio of about 8:2 (2*R*:2*S*) as measured by NMR. The same procedure was used for the preparation of the other ligands: 1:1 mixture of L-cysteine and glyoxylate for (2*R*,4*R*)-TDA; 1:1 mixture of L-cysteine and formaldehyde for 4*R*-TCA; 1:1 mixture of L-penicillamine and pyruvate for (2*R*,4*R*)-TMeTDA (25 mM stocks). The selenated analogue

(2*R*,4*R*)-SeMeTDA was obtained by a similar procedure, except that 66.8 mg L-selenocystine were first solubilized in 1 mL water by addition of 50% NaOH, and then reduced to L-selenocysteine with 5 DTT molar equivalents. All the ligand stock solutions were sparged with N<sub>2</sub> inside the glove box to eliminate any dissolved O<sub>2</sub> prior to their addition to the *TmHydE* protein solution. For the initial *in vitro* assays, a 1.5 molar excess pyruvate was added to L-cysteine in order to prevent any trace of free L-cysteine from reacting with HydE. Conversely, a 1:1 ratio of L-cysteine and <sup>13</sup>C-labeled pyruvate (12.1 and 11.1 mg respectively in 1 mL 50 mM phosphate buffer pH 8) was used for the time-resolved NMR spectroscopy analysis. The absence of both pyruvate and L-cysteine was controlled by <sup>1</sup>H-NMR prior to perform the reaction.

***In vitro* assays.** *CaHydE* (100 μM) after FeS anaerobic reconstitution was incubated in the presence of (2*R*,4*R*)-MeTDA (1 mM), SAM (1 mM) and sodium dithionite as one electron donor (2 mM) in Tris pH 8 (50 mM), KCl (300 mM) buffer pH 8. Reaction using (2*R*,4*R*)-TDA, 4*R*-TCA or (2*R*,4*R*)-TMeTDA ligands were performed under the same conditions with the only exception that Tris buffer was replaced by 50 mM phosphate buffer pH 8. The same conditions were used for assaying *TmHydE* and *TmHydE*<sup>ΔFeS</sup> proteins. Samples (3 μL) were quenched anaerobically at different times with 27 μL trifluoroacetic acid solution (0.1% TFA). HPLC analyses were performed on an Agilent 1200 series infinity with a reversed phase column (LiChroCART RP-18e 5 μm) and the following gradient with A (H<sub>2</sub>O, 0.1% TFA) and solvent B (80 % CH<sub>3</sub>CN, 0.1% TFA): 0-1 min 100% A/0% B; 1-15 min, a linear gradient up to 80% A/20% B; 15-50 min, a linear gradient to 40% A/60% B; 50-55 min, a linear gradient to 100% A/0%B. Flow rate was of 1 ml/min and detection was made at 257 nm.

**NMR-spectroscopy.** (2*R*,4*R*)-MeTDA characterization and kinetic measurements by NMR-spectroscopy are described in the Supplementary methods.

**X-ray crystallography analyses.** Crystallization of *TmHydE*<sup>ΔFeS</sup> was performed as previously described for *TmHydE* with minor modifications<sup>11</sup>. As standard procedure, the different ligands (final concentration 25-50 mM) were mixed with pure *TmHydE* (or *TmHydE*<sup>ΔFeS</sup>) protein prior to setting the crystallization trials. Crystal compositions are succinctly described in Table 1. A more detailed description of crystal preparation and treatment, in particular for kinetic studies, is found in the Supplementary Methods. Anaerobic flash cooling of the crystals was carried out inside the glove box as previously described<sup>43</sup>. Data were collected at the BM30A beamline of the European Synchrotron Radiation Facility and were processed with the XDS package<sup>44</sup>. Manual model building was performed using COOT<sup>45</sup> and model refinements were carried out using REFMAC from the CCP4 Package<sup>46</sup>. Special care was taken to use the same set of R<sub>free</sub>-labeled reflections and to define the proper geometrical restraints to be used for the different ligands and cofactors during refinement (see Supplementary Methods).

**Theoretical calculations.** The Schrödinger suite<sup>47</sup> was used for all calculations (see Supplementary methods).

## References

1. Sofia, H. J., Chen, G., Hetzler, B. G., Reyes-Spindola, J. F. & Miller, N. E. Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic Acids Res.* **29**, 1097–1106 (2001).
2. Broderick, J. B., Duffus, B. R., Duschene, K. S. & Shepard, E. M. Radical S-adenosylmethionine enzymes. *Chem. Rev.* **114**, 4229–4317 (2014).
3. Berkovitch, F., Nicolet, Y., Wan, J. T., Jarrett, J. T. & Drennan, C. L. Crystal structure of biotin synthase, an S-adenosylmethionine-dependent radical enzyme. *Science* **303**, 76–79

(2004).

4. Lanz, N. D. *et al.* Evidence for a catalytically and kinetically competent enzyme-substrate cross-linked intermediate in catalysis by lipoyl synthase. *Biochemistry* **53**, 4557–4572 (2014).
5. Forouhar, F. *et al.* Two Fe-S clusters catalyze sulfur insertion by radical-SAM methylthiotransferases. *Nat. Chem. Biol.* **9**, 333–338 (2013).
6. Hernandez, H. L. *et al.* MiaB, a bifunctional radical *S*-adenosylmethionine enzyme involved in the thiolation and methylation of tRNA, contains two essential [4Fe-4S] clusters. *Biochemistry* **46**, 5140–5147 (2007).
7. Jarrett, J. T. The Biosynthesis of Thiol- and Thioether-containing Cofactors and Secondary Metabolites Catalyzed by Radical *S*-Adenosylmethionine Enzymes. *J. Biol. Chem.* **290**, 3972–3979 (2015).
8. Nicolet, Y. & Fontecilla-Camps, J. C. Structure-Function Relationships in [FeFe]-Hydrogenase Active Site Maturation. *J. Biol. Chem.* **287**, 13532–13540 (2012).
9. Mulder, D. W. *et al.* Insights into [FeFe]-hydrogenase structure, mechanism, and maturation. *Struct. Lond. Engl. 1993* **19**, 1038–1052 (2011).
10. Broderick, J. B. *et al.* H-cluster assembly during maturation of the [FeFe]-hydrogenase. *J. Biol. Inorg. Chem. JBIC Publ. Soc. Biol. Inorg. Chem.* **19**, 747–757 (2014).
11. Nicolet, Y. *et al.* X-ray structure of the [FeFe]-hydrogenase maturase HydE from *Thermotoga maritima*. *J. Biol. Chem.* **283**, 18861–18872 (2008).
12. Padwa, A., Beall, L. S., Heidelbaugh, T. M., Liu, B. & Sheehan, S. M. A one-pot bicycloannulation method for the synthesis of tetrahydroisoquinoline systems. *J. Org. Chem.* **65**, 2684–2695 (2000).
13. Denes, F., Schiesser, C. H. & Renaud, P. Thiols, thioethers, and related compounds as sources of C-centred radicals. *Chem. Soc. Rev.* **42**, 7900–7942 (2013).
14. Posewitz, M. C. *et al.* Discovery of two novel radical *S*-adenosylmethionine proteins required for the assembly of an active [Fe] hydrogenase. *J. Biol. Chem.* **279**, 25711–25720 (2004).
15. Robbins, F. & Fioriti, J. Alkaline degradation of cystine, glutathione and sulphur-containing proteins. *Nature* **200**, 577–& (1963).
16. Kredich, N., Foote, L. & Keenan, B. The stoichiometry and kinetics of the inducible cysteine desulfhydrase from *Salmonella typhimurium*. *J. Biol. Chem.* **248**, 6187–6196 (1973).
17. Schubert, M. P. Compounds of thiol acids with aldehydes. *J. Biol. Chem.* **114**, 341–350 (1936).
18. Kallen, R. Mechanism of reactions involving Schiff base intermediates - Thiazolidine formation from L-cysteine and formaldehyde. *J. Am. Chem. Soc.* **93**, 6236–& (1971).
19. Refouvelet, B., Robert, J., Couquelet, J. & Tronche, P. Thiazolidine-2,4-dicarboxylic and its esters - synthesis, in solution behavior and regioselective cyclocondensation. *J. Heterocycl. Chem.* **31**, 77–80 (1994).
20. Shiraiwa, T., Katayama, T., Ishikawa, J., Asai, T. & Kurokawa, H. Preparation of optically active 2-thiazolidinecarboxylic acid by asymmetric transformation. *Chem. Pharm. Bull. (Tokyo)* **47**, 1180–1183 (1999).
21. Nagasawa, H., Goon, D. & Shirota, F. Epimerization at C-2 of 2-substituted thiazolidine-4-carboxylic acid. *J. Heterocycl. Chem.* **18**, 1047–1051 (1981).
22. Chiarino, D., Ferrario, F., Pellacini, F. & Sala, A. Synthesis and in solution behaviour of new 2-substituted-4-thiazolidinecarboxylic acid derivatives. *J. Heterocycl. Chem.* **26**, 589–593 (1989).
23. Shiraiwa, T., Katayama, T., Kaito, T., Tanigawa, H. & Kurokawa, H. Facile preparation of four stereoisomers of 2,4-thiazolidinedicarboxylic acid. *Bull. Chem. Soc. Jpn.* **71**, 1911–1914 (1998).

24. Betz, J. N. *et al.* [FeFe]-Hydrogenase Maturation: Insights into the Role HydE Plays in Dithiomethylamine Biosynthesis. *Biochemistry* **54**, 1807–1818 (2015).
25. Mahanta, N., Fedoseyenko, D., Dairi, T. & Begley, T. P. Menaquinone biosynthesis: formation of aminofutalosine requires a unique radical SAM enzyme. *J. Am. Chem. Soc.* **135**, 15318–15321 (2013).
26. Aitken, H. M., Hancock, A. N. & Schiesser, C. H. Understanding (the lack of) homolytic substitution chemistry of sulfones. *Chem. Commun.* **48**, 8326–8328 (2012).
27. Nicolet, Y., Amara, P., Mouesca, J.-M. & Fontecilla-Camps, J. C. Unexpected electron transfer mechanism upon AdoMet cleavage in radical SAM proteins. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 14867–14871 (2009).
28. Wang, S. C. & Frey, P. A. Binding energy in the one-electron reductive cleavage of *S*-adenosylmethionine in lysine 2,3-aminomutase, a radical SAM enzyme. *Biochemistry* **46**, 12889–12895 (2007).
29. Yokoyama, K., Numakura, M., Kudo, F., Ohmori, D. & Eguchi, T. Characterization and mechanistic study of a radical SAM dehydrogenase in the biosynthesis of butirosin. *J. Am. Chem. Soc.* **129**, 15147–15155 (2007).
30. Ruzsyczky, M. W., Ogasawara, Y. & Liu, H.-W. Radical SAM enzymes in the biosynthesis of sugar-containing natural products. *Biochim. Biophys. Acta* **1824**, 1231–1244 (2012).
31. Grell, T. A. J., Goldman, P. J. & Drennan, C. L. SPASM and twitch domains in *S*-adenosylmethionine (SAM) radical enzymes. *J. Biol. Chem.* **290**, 3964–3971 (2015).
32. Horitani, M. *et al.* Why Nature Uses Radical SAM Enzymes so Widely: Electron Nuclear Double Resonance Studies of Lysine 2,3-Aminomutase Show the 5'-dAdo\* 'Free Radical' Is Never Free. *J. Am. Chem. Soc.* **137**, 7111–7121 (2015).
33. Goldman, P. J. *et al.* X-ray structure of an AdoMet radical activase reveals an anaerobic solution for formylglycine posttranslational modification. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 8519–8524 (2013).
34. Vey, J. L. *et al.* Structural basis for glycyl radical formation by pyruvate formate-lyase activating enzyme. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 16137–16141 (2008).
35. Lepore, B. W., Ruzicka, F. J., Frey, P. A. & Ringe, D. The x-ray crystal structure of lysine-2,3-aminomutase from *Clostridium subterminale*. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 13819–13824 (2005).
36. Dowling, D. P. *et al.* Radical SAM enzyme QueE defines a new minimal core fold and metal-dependent mechanism. *Nat. Chem. Biol.* **10**, 106–112 (2014).
37. Nicolet, Y., Zepieri, L., Amara, P. & Fontecilla-Camps, J. C. Crystal structure of tryptophan lyase (NosL): evidence for radical formation at the amino group of tryptophan. *Angew. Chem. Int. Ed* **53**, 11840–11844 (2014).
38. Gruber, K., Reitzer, R. & Kratky, C. Radical shuttling in a protein: Ribose pseudorotation controls alkyl-radical transfer in the coenzyme B-12 dependent enzyme glutamate mutase. *Angew. Chem. Int. Ed.* **40**, 3377–+ (2001).
39. Warncke, K. Characterization of the product radical structure in the Co(II)-product radical pair state of coenzyme B12-dependent ethanolamine deaminase by using three-pulse 2H ESEEM spectroscopy. *Biochemistry* **44**, 3184–3193 (2005).
40. Frey, P. A. Travels with Carbon-Centered Radicals. 5'-Deoxyadenosine and 5'-Deoxyadenosine-5'-yl in Radical Enzymology. *Acc. Chem. Res.* **47**, 540–549 (2014).
41. Fugate, C. J. *et al.* 9-Mercaptodethiobiotin is generated as a ligand to the [2Fe-2S]<sup>+</sup> cluster during the reaction catalyzed by biotin synthase from *Escherichia coli*. *J. Am. Chem. Soc.* **134**, 9042–9045 (2012).
42. Nicolet, Y., Rohac, R., Martin, L. & Fontecilla-Camps, J. C. X-ray snapshots of possible intermediates in the time course of synthesis and degradation of protein-bound Fe<sub>4</sub>S<sub>4</sub>

- clusters. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 7188–7192 (2013).
43. Vernede, X. & Fontecilla-Camps, J. C. A method to stabilize reduced and or gas-treated protein crystals by flash-cooling under a controlled atmosphere. *J. Appl. Crystallogr.* **32**, 505–509 (1999).
44. Kabsch, W. XDS. *Acta Crystallogr. Sect. -Biol. Crystallogr.* **66**, 125–132 (2010).
45. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. -Biol. Crystallogr.* **66**, 486–501 (2010).
46. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. -Biol. Crystallogr.* **67**, 235–242 (2011).
47. a) Maestro version 9.6 b) QSite version 6.1 c) Jaguar version 8.2, Schrödinger, LLC, New York, NY, 2013.
48. Onofrio, N. & Mouesca, J.-M. Valence bond/broken symmetry analysis of the exchange coupling constant in copper(II) dimers. Ferromagnetic contribution exalted through combined ligand topology and (singlet) covalent-ionic mixing. *J. Phys. Chem. A* **114**, 6149–6156 (2010).
49. Onofrio, N. & Mouesca, J.-M. Analysis of the singlet-triplet splitting computed by the density functional theory-broken-symmetry method: is it an exchange coupling constant? *Inorg. Chem.* **50**, 5577–5586 (2011).

### Acknowledgements

The authors thank M. Pirrochi for assistance with data collection at BM30A, P. Gans for fruitful discussions and the Commissariat à l’Energie Atomique (CEA), the Centre National de la Recherche Scientifique, and the Université Grenoble-Alpes for institutional support. We appreciate the help from the staff of the computing facility provided by the Commissariat à l’Energie Atomique (CEA/DSV/GIPSI), Saclay, and (CCRT), Bruyères-le-Châtel. R.R. is the recipient of an ‘Irtelis’ Fellowship from the CEA. This work was also supported by Contract ANR-12-BSV8-0013 from the French Agence Nationale pour la Recherche and used the platforms of the Grenoble Instruct center (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) within the Grenoble Partnership for Structural Biology (PSB).

### Author contribution

Y.N. and J.C.F.-C. designed and conducted the project. R.R. and Y.N. performed the structural work. L.M. R.R. and Y.N. performed the protein sample preparation. A.B. P.R. and O.B. performed the functional work. A.F. and R.R. performed the NMR spectroscopy measurements. P.A. and J.M.M. performed the theoretical calculations. R.R., P.A., J.M.M., A.B., O.B., J.C.F.-C. and Y.N. analyzed data. R.R., P.A., O.B., J.M.M., J.C.F.-C. and Y.N. wrote the manuscript.

### Additional information

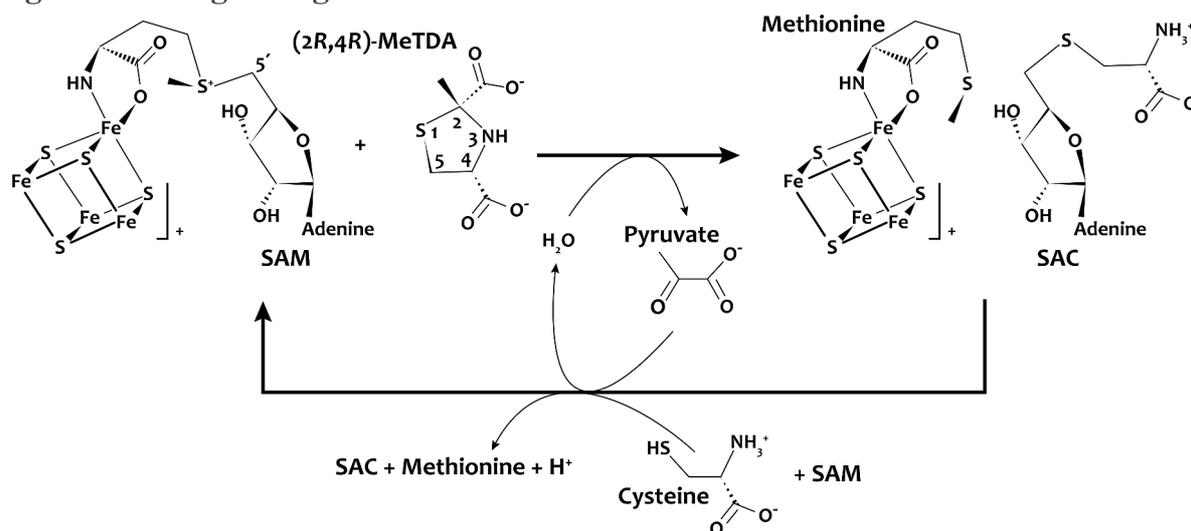
The atomic coordinates and structure factors will be deposited in the Protein Data Bank, [www.pdb.org](http://www.pdb.org) (PDB ID codes XXXX). Supplementary information is available in the online version of the paper. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing interests. Correspondence and requests for materials should be addressed to Y.N. or J.C.F.-C.

## Tables

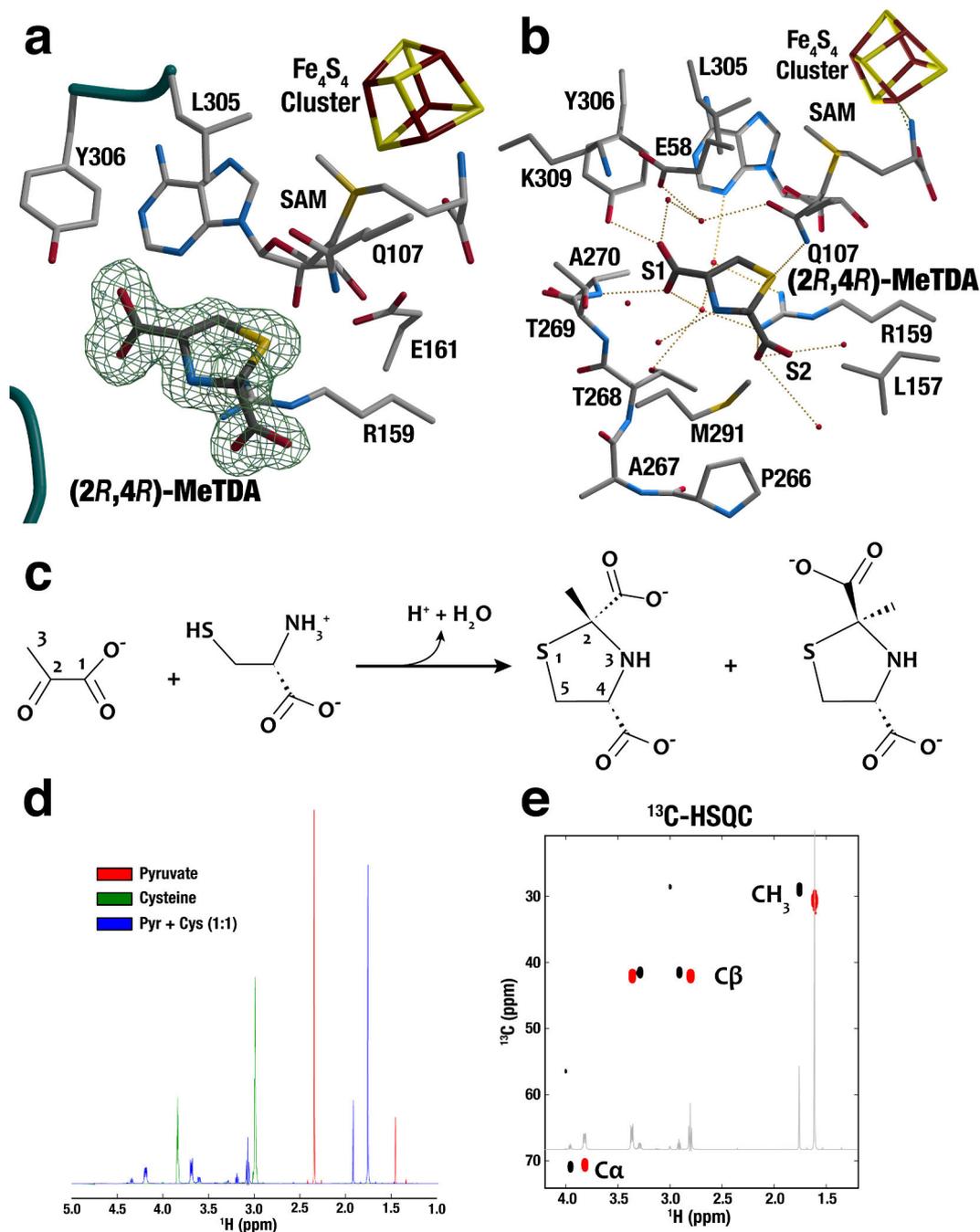
**Table 1 | Summary of the crystal structures presented in this work.** A detailed description of each crystal preparation and treatment is presented in the Supplementary methods.

Crystal name	PDB code	Resolution	Ligand – specific crystal treatment
X1		1.45 Å	(2 <i>R</i> ,4 <i>R</i> )-MeTDA – $t_0$ (kinetics)
X2		1.28 Å	(2 <i>R</i> ,4 <i>R</i> )-MeSeTDA – $t_0$ (kinetics)
X3		1.35 Å	(2 <i>R</i> ,4 <i>R</i> )-MeSeTDA + 5'-dA + Met – none
X4		1.32 Å	SeAC – $t_{\text{final}}$ (kinetics)
X5		1.17 Å	SAC – $t_{\text{final}}$ (kinetics)
X6		1.48 Å	SAC + Met – none
X7		1.47 Å	(2 <i>R</i> ,4 <i>R</i> )-TDA + SAH – none
X8		1.18 Å	4 <i>R</i> -TCA + SAH – none
X9		1.35 Å	(2 <i>R</i> ,4 <i>R</i> )-TMeTDA + SAH – none

## Figures and Figure legends.

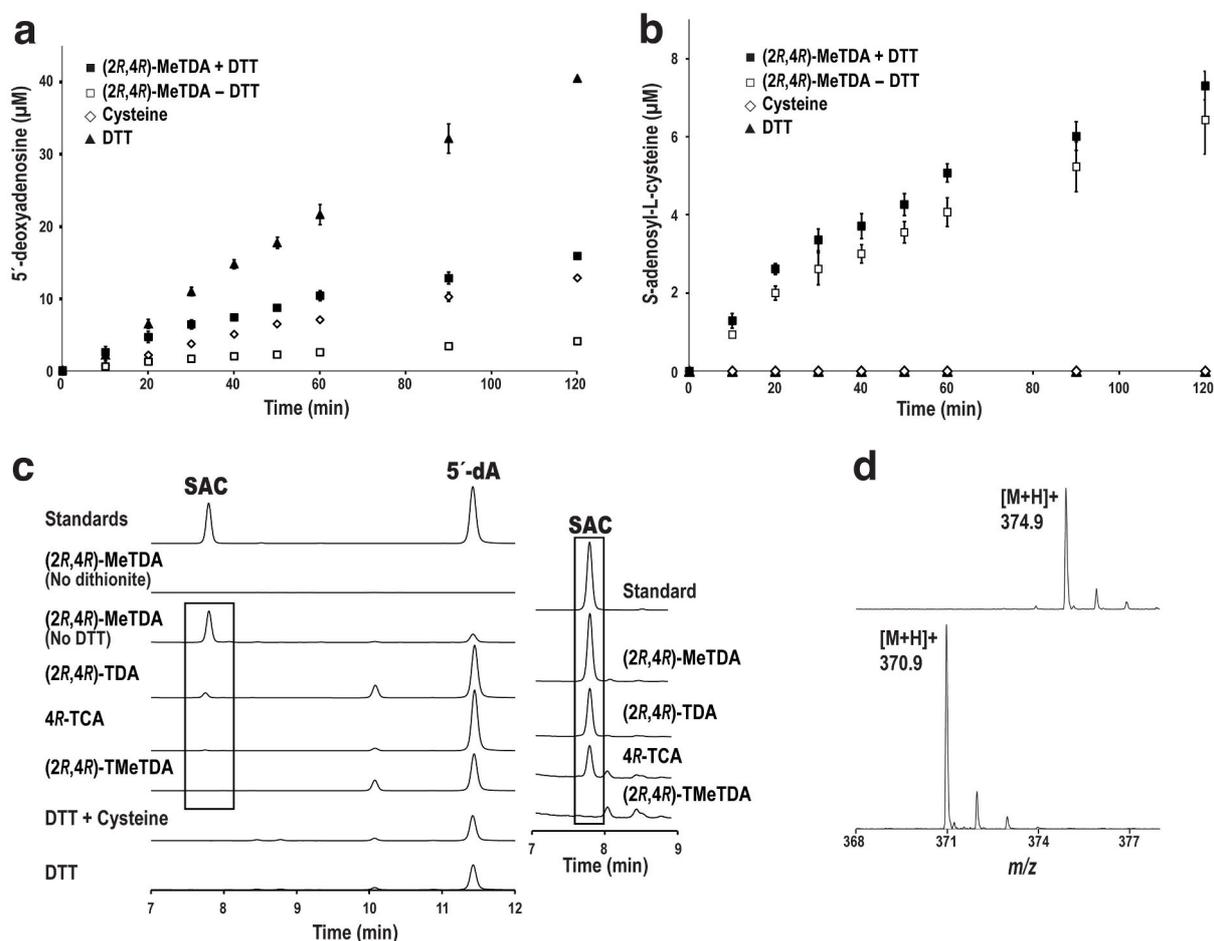


**Scheme 1 | Pyruvate-mediated synthesis of *S*-adenosyl-L-cysteine from cysteine and *S*-adenosyl-L-methionine.**

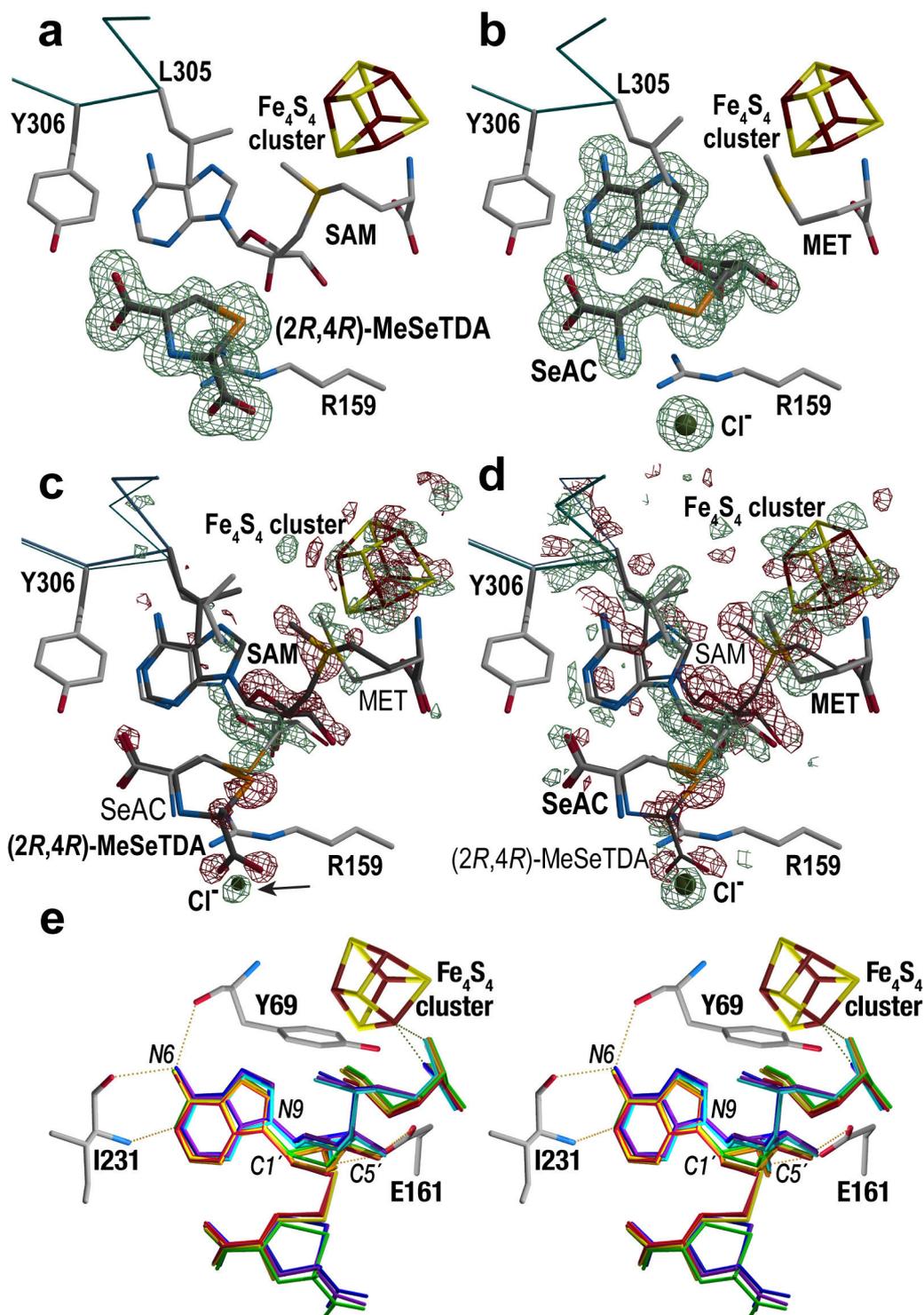


**Figure 1 | (2R,4R)-MeTDA bound to HydE and its formation.** **a.** Omit *Fo-Fc* electron density map for (2R,4R)-MeTDA contoured at  $3\sigma$  and depicted as a green mesh with a cover-radius of 2.0 Å; Data are from crystal *XI* (Table 1) containing SAM and (2R,4R)-MeTDA. SAM, the  $[\text{Fe}_4\text{S}_4]$  cluster and five neighboring residues are also depicted. The color codes are grey, blue, red, yellow and brown for C, N, O, S and Fe atoms, respectively. **b.** View of the cavity around (2R,4R)-MeTDA in HydE including water molecules from the first and second solvation spheres. Hydrogen bonds are depicted as dashed golden lines. The previously identified anion-binding **S1** and **S2** sites are also shown<sup>11</sup>. A stereo view of panel **b** is depicted in Supplementary Fig. S2. **c.** Reaction for the formation of (2R,4R)-MeTDA and (2S,4R)-MeTDA from a 1:1 mixture of L-cysteine and pyruvate<sup>16</sup>. **d.**  $^1\text{H}$ -NMR data showing the complete conversion of L-cysteine (green) and pyruvate (red) into (2R,4R)-MeTDA and (2S,4R)-MeTDA when mixed in a 1:1 ratio (blue). **e.**  $^{13}\text{C}$ -edited HSQC spectrum used for the assignment of the different peaks to either (2R,4R)-MeTDA (red) or (2S,4R)-MeTDA (black)

epimers. The corresponding  $^1\text{H-NMR}$  spectrum is depicted in grey. The data used for the complete attribution of the different NMR peaks are presented in Supplementary Fig. S1.



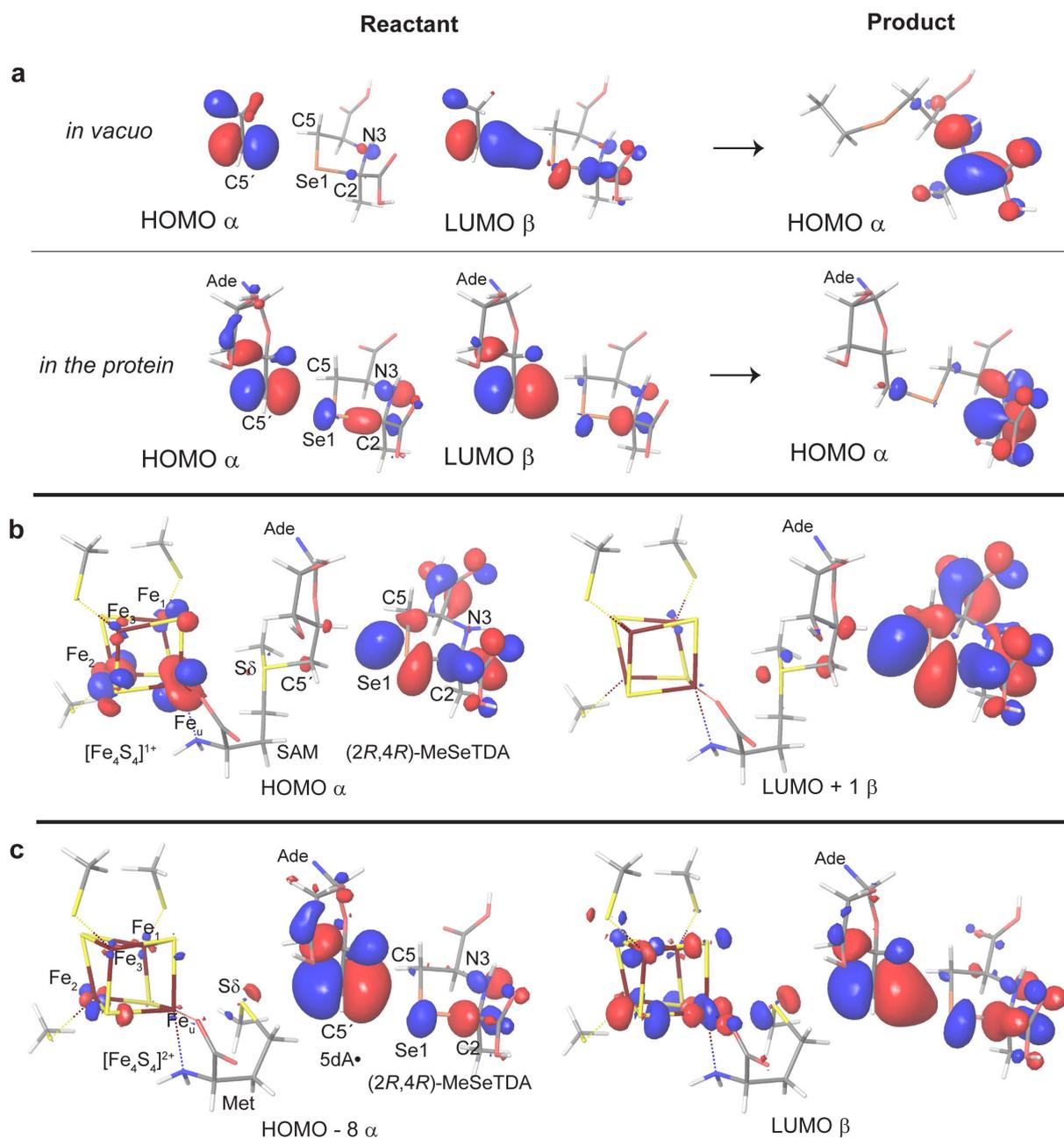
**Figure 2 | *In vitro* activity of HyDE.** **a.** Time course for the production of 5'-dA by *CaHyDE* in the presence of 1 mM DTT (▲), 1 mM Cys (◇), 1 mM (2R,4R)-MeTDA (□) or 1 mM (2R,4R)-MeTDA and DTT (■). **b.** Time course for the production of SAC by HyDE in the presence of 1 mM DTT (▲), 1 mM Cys (◇), 1 mM (2R,4R)-MeTDA (□) or 1 mM (2R,4R)-MeTDA and DTT (■). **c.** Reverse phase HPLC analysis of SAC and 5'-dA standards compared to reaction profiles obtained with *CaHyDE* incubated with various potential substrates. On the right is presented a close-up of the region in the rectangle corresponding to SAC retention time for all 1,3-thiazolidine ligands. It shows that all but (2R,4R)-TMeTDA lead to SAC production. **d.** MALDI TOF MS analysis of enzymatically synthesized SAC by HyDE incubated in the presence of L-cysteine derived (2R,4R)-MeTDA (lower trace) or  $^{13}\text{C},^{15}\text{N}$ -labeled L-cysteine isotopolog-derived (2R,4R)-MeTDA (upper trace). In order to facilitate the *in vitro* assay measurements, HyDE enzyme from *Clostridium acetobutylicum* (*Ca*) was used (Methods). Similar experiments using *TmHyDE* with and without its second non-essential cluster are presented in Supplementary Fig. S4. They confirm that both enzymes produce SAC from 1,3-thiazolidines and that the second cluster does not play any role in this reaction.



**Figure 3 | HydE reactivity with (2R,4R)-MeSeTDA.** **a** and **b** are *Fo-Fc* omit electron density maps depicted as green meshes (contoured at  $3.0 \sigma$ ; and a cover map of  $2.0 \text{ \AA}$ ). They correspond to the initial (**a**) and final (**b**) structures (crystals **X2** –  $1.28 \text{ \AA}$  resolution and **X4** –  $1.32 \text{ \AA}$  resolution; Table 1) obtained during the reaction catalyzed by HydE using (2R,4R)-MeSeTDA as substrate. The selenium derivatives of (2R,4R)-MeTDA ((2R,4R)-MeSeTDA in **a**) and SAC (SeAC in **b**) are modeled in their corresponding electron densities. The orientation and depiction are the same as in Fig. 1. Selenium atoms are shown in orange and the chloride ion in **b** is shown as a green sphere. **c** and **d** correspond to *Fo*( $t_{45\text{min}}$ )-*Fo*( $t_0$ ) and *Fo*( $t_{2\text{h}00}$ )-*Fo*( $t_0$ ) difference maps, respectively (see Methods). They are contoured at  $\pm 5 \sigma$  (green/red mesh). The positive peaks in green and the negative peaks in red represent features

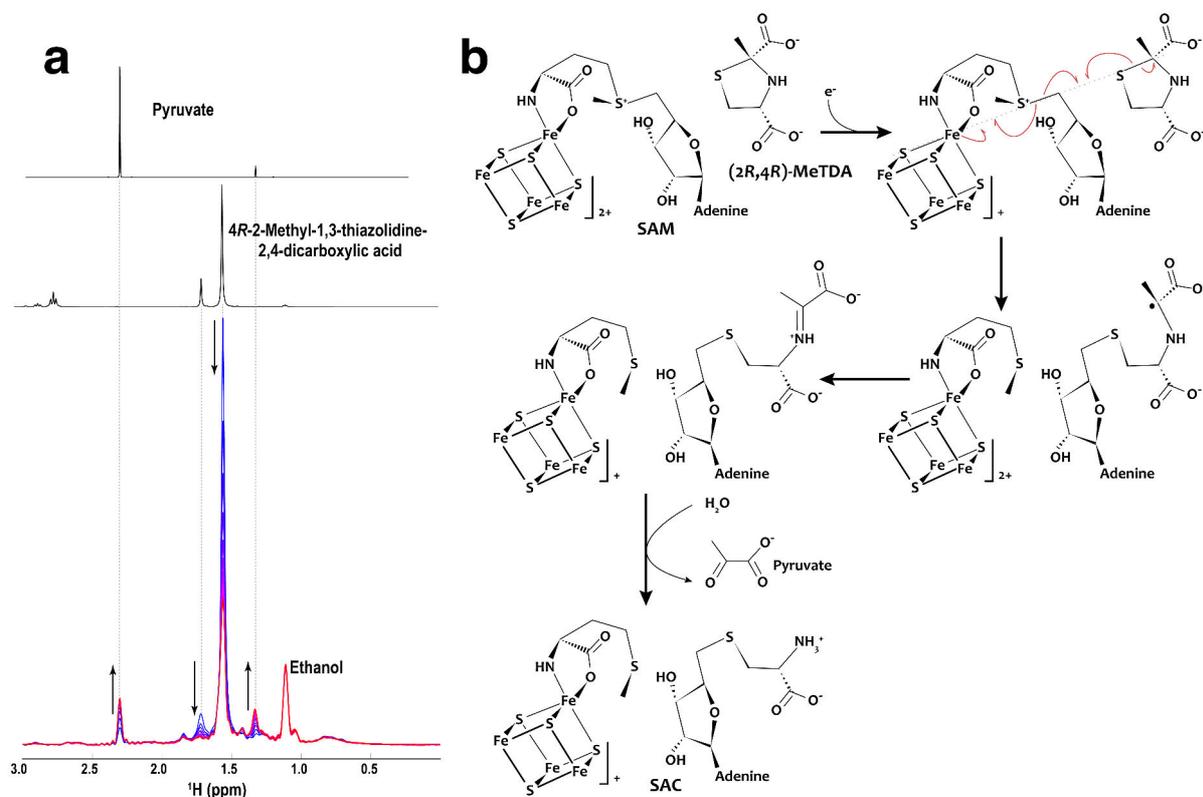
that respectively appear or disappear as a function of time. SAM cleavage and 5'-dA motions toward the selenium atom of (2*R*,4*R*)-MeSeTDA are clearly seen. Likewise, the disappearance of the pyruvate fragment is clearly indicated by the presence of negative peaks at the positions of the oxygen atoms of its carboxylate group and the concomitant occupation of the same site by a chloride ion, which is illustrated by the increasing height of a positive peak in the electron density maps (arrow in **c**). Stereo views of **c** and **d** panels are presented in the Supplementary Fig. S6. The use of selenium instead of sulfur led to crystals displaying intermediate time-lapses of better quality. However, the same SAC production was observed when starting with (2*R*,4*R*)-MeTDA instead of (2*R*,4*R*)-MeSeTDA (Supplementary Fig. S5).

**e.** Stereo view of the motions around 5'-dA during the reaction catalyzed by HydE. The hydrogen bonds between N6, O2' and O3' from 5'-dA (or SAM) and the protein matrix restrict the cofactor movements toward the (2*R*,4*R*)-MeTDA substrate. The following structures are depicted: SAM+(2*R*,4*R*)-MeTDA (crystal **X1**; blue), SAM+(2*R*,4*R*)-MeSeTDA (crystal **X2**; purple), SAM only (PDB code 3IIZ<sup>27</sup>; cyan), 5'-dA+Methionine+(2*R*,4*R*)-MeSeTDA (crystal **X3**; green), 5'-dA+Methionine (PDB code 3IIX<sup>27</sup>; orange), SeAC (crystal **X4**;  $t_{\text{final}}$  of the kinetic; yellow), SAC (crystal **X5**;  $t_{\text{final}}$  of the kinetic; crimson) and SAC+Methionine (crystal **X6**; red). The relevant distances extracted from these structures are presented in Supplementary Table S3.



**Figure 4 | Models of the radical C5' attack on (2R,4R)-MeSeTDA (first step in SeAC formation, see Scheme 1).** **a.** (Top) Small QM model of the reaction where the 5'dA• is modeled as CH<sub>3</sub>CH<sub>2</sub>• and (bottom) QM/MM model of the same reaction (only the QM part is represented here; see Supplementary methods for details). The HOMO  $\alpha$  and LUMO  $\beta$  are represented for the reactant (**R**) and the HOMO  $\alpha$  for the product (**P**) (see Supplementary Scheme S1 for the signification of HOMO  $\alpha$  and LUMO  $\beta$ <sup>48,49</sup>). [+] and [-] molecular orbital surfaces are depicted in red and blue, respectively (isosurface value = 0.05). Standard atom color code is used. **b.** HOMO  $\alpha$  of the initial state: before SAM cleavage (S $\delta$  is positively charged) the [Fe<sub>4</sub>S<sub>4</sub>] cluster is formally reduced. The LUMO  $\alpha$ , almost identical to the LUMO+1  $\beta$ , is nearly degenerate with the HOMO  $\alpha$  (1.8 kJ.mol<sup>-1</sup>). As a consequence, part of (2R,4R)-MeSeTDA appears in the HOMO  $\alpha$ , revealing the slight radical character of (2R,4R)-MeSeTDA (spin population of 0.19 on Se1), further confirmed by the (2R,4R)-MeSeTDA centered LUMO+1  $\beta$ . Because the departing electron is formally expected to be located in the [Fe<sub>4</sub>S<sub>4</sub>] cluster, this observation illustrates the (2R,4R)-MeSeTDA proneness to radical oxidation which will occur after SAM cleavage. **c.** After SAM cleavage the radical is

mostly located on C5' as shown by the HOMO-8  $\alpha$  (versus LUMO  $\beta$ ) and the  $[\text{Fe}_4\text{S}_4]$  cluster is oxidized. Notice that in addition to C5' the  $[\text{Fe}_4\text{S}_4]$  cluster the S $\delta$  and (2*R*,4*R*)-MeSeTDA appear in the LUMO  $\beta$ ; this indicates an orbital match between the four species. [+] and [-] molecular orbital surfaces are depicted in red and blue, respectively (isosurface value = 0.03).



**Figure 5 | Time-resolved NMR spectroscopy and putative mechanism for the formation of SAC from (2*R*,4*R*)-MeTDA as catalyzed by HydE. a.** Time-resolved NMR-spectroscopy showing the disappearance of (2*R*,4*R*)-MeTDA over time and the concomitant formation of pyruvate. Top and middle figures correspond to the reference  $^1\text{H}$ -NMR spectra for pyruvate and 4*R*-2-methyl-1,3-thiazolidine-2,4-dicarboxylic acid, respectively. (Bottom) Overlay of spectra measured every 2 min (blue to red traces) for the reaction catalyzed by *CaHydE* (see Supplementary methods for details). A  $^{13}\text{C}$ -filter was used to detect only the protons covalently bound to a  $^{13}\text{C}$  nucleus allowing to eliminate the proton signals arising from the unlabeled protein, SAM, SAC and buffer. **b.** According to the observation that in our crystals, even at short time-lapses, the pyruvate fragment gradually disappears in favor of chloride ion binding (Figs. 3c and d), we can conclude that the reaction limiting step corresponds to the one electron transfer from the  $[\text{Fe}_4\text{S}_4]^+$  cluster to SAM. The putative C2• radical re-oxidation followed by the hydrolysis of the resulting ketimine should be faster and irreversible steps.



## Carbon-Sulfur Bond-Forming Reaction Catalyzed by the Radical SAM Enzyme HydE

Roman Rohac<sup>1</sup>, Patricia Amara<sup>1</sup>, Alhosna Benjdia<sup>2,3</sup>, Lydie Martin<sup>1</sup>, Pauline Ruffié<sup>2,3</sup>, Adrien Favier<sup>4</sup>, Olivier Berteau<sup>2,3</sup>, Jean-Marie Mouesca<sup>5,6</sup>, Juan C. Fontecilla-Camps<sup>1\*</sup> and Yvain Nicolet<sup>1\*</sup>

<sup>1</sup> Metalloproteins Unit, Institut de Biologie Structurale, CEA, CNRS, UGA, 71, Avenue des Martyrs, 38044, Grenoble Cedex 9, France

<sup>2</sup> INRA, ChemSyBio, UMR 1319 Micalis, F-78350 Jouy-en-Josas, France

<sup>3</sup> AgroParisTech, ChemSyBio, UMR Micalis, F-78350 Jouy-en-Josas, France

<sup>4</sup> Biomolecular NMR Spectroscopy Group, Institut de Biologie Structurale, CEA, CNRS, UGA, 71, Avenue des Martyrs, 38044, Grenoble Cedex 9, France

<sup>5</sup> Univ. Grenoble Alpes, INAC, SCIB/LRM, F-38000 Grenoble, France

<sup>6</sup> CEA, INAC, SCIB/LRM, F-38054 Grenoble, France

### Table of Contents

### Page

General Comments

I. Methods

II. Results

III. References

### General Comments

#### *Provenance of chemicals*

All the following chemicals: S-(5'-Adenosyl)-L-cysteine (SAC); 5'-deoxyadenosine; L-Methionine; sodium pyruvate; sodium glyoxylate; formaldehyde; L-penicillamine; L-Cysteine and L-Cystine were purchased from Sigma-Aldrich. Seleno-L-cystine and Tiron were purchased from Fluka and the CHAPS detergent was purchased from Calbiochem. <sup>13</sup>C-labeled pyruvate (3-<sup>13</sup>C) and <sup>13</sup>C,<sup>15</sup>N-labeled cysteine were purchased from Eurisotop.

### I. Methods

#### **Removal of the semi-conserved accessory FeS cluster in HydE from *Thermotoga maritima* (TmHydE).**

In order to improve the structural homogeneity of the protein samples for crystallographic analyses, the semi-conserved second FeS cluster was removed by mutation of the three C-terminal cysteine ligands to serine. This Cys to Ser (C311S-C319S-C321S) triple mutant of the accessory FeS cluster ligands was obtained following the QuikChange® site-directed mutagenesis kit strategy as previously described<sup>1</sup>. Our pStrepTmHydE plasmid served as template and the following primers:

*Forward:* 5'-GAGAAGGACACTGCGTCTATTCCCTCCGTGATGAAGATGATA-3';  
*Reverse:* 5'-CGCAGTGCCTTCTCGAAAACGGAGATCTTTCCAGGATA-3'  
were used to produce the pStrep*TmHydE*<sup>ΔFeS</sup> construct. The correctness of the cloned DNA sequence was confirmed by sequencing the entire gene.

### Protein expression and purification.

The sample used for enzymatic kinetic experiments in crystals (datasets **X1**, **X2**, **X4** & **X5**, Table S1) was prepared without *in vitro* FeS cluster reconstitution. The *TmHydE*<sup>ΔFeS</sup> gene was co-expressed with the *isc* operon and *metK* (*S*-adenosyl-L-methionine synthase) using an *E. coli* BL21-CodonPlus(DE3)-RIL strain containing both pStrep*TmHydE*<sup>ΔFeS</sup> and pRSF-ISC-MetK plasmids<sup>2</sup>. Cells were grown at 37°C in buffered TB medium supplemented with 50 mM MOPS/KOH pH 7.1, 0.5% D-glucose, ampicillin (100 μg.ml<sup>-1</sup>), kanamycine (50 μg.ml<sup>-1</sup>) and chloramphenicol (35 μg.ml<sup>-1</sup>) until the culture reached an OD<sup>600</sup> of about 0.4. The temperature was then lowered to 20°C and after 30 min, the culture was transferred to 2 L sealed bottles inside an anaerobic glove box at 20°C. From this point on, all the procedures were performed anaerobically. After 15 min, the bottles were opened and incubated under mild shaking until the OD<sup>600</sup> reached 0.7. We considered that by then the medium had equilibrated with the glove box atmosphere and the culture was under strict anaerobic conditions. The medium was then enriched with ammonium ferric citrate (40 μM), L-Cysteine (400 μM), MEM Vitamin solution 100X (SIGMA M6895) (12 mL.L<sup>-1</sup>) and sodium fumarate (25 mM). Protein expression was induced at an OD<sup>600</sup> slightly above 0.8 with the addition of 10 μM isopropyl-β-D-1-thiogalactopyranoside (IPTG) and cells were grown overnight at 20°C before harvesting. The cell pellets were re-suspended in anaerobic low salt buffer (100 mM Tris pH 8.0; 50 mM NaCl) supplemented with EDTA-free protease inhibitor cocktail and then disrupted by sonication on ice inside the glove box (2 s 35% amplitude, 10 s pause, 2.5 min total time). Crude extract was subsequently cleared by ultracentrifugation in sealed tubes at 100,000 × g during 30 min at 4°C. The clarified supernatant was filtered, and loaded onto a streptavidin-agarose column (Iba®) equilibrated with anaerobic low-salt buffer. After an extensive wash step, the protein was eluted with a 1 mM desthiobiotin solution. The protein fractions were concentrated by ultrafiltration (30 kDa molecular weight cut-off) to 8 mg.mL<sup>-1</sup> and subsequently applied to a Superdex S-75 column previously equilibrated with anaerobic low-salt buffer (100 mM Tris pH 8.0; 50 mM NaCl). The protein fractions corresponding to the monomeric form of *TmHydE* were pooled and concentrated to 12 mg.mL<sup>-1</sup> by ultrafiltration.

Protein expression of the selenium-substituted cluster-containing sample (datasets **X3**, **X6** & **X9**, Table S1) was carried out aerobically. *E. coli* BL21-CodonPlus(DE3)-RIL cells transformed with the pStrep*TmHydE*<sup>ΔFeS</sup> plasmid were grown in LB medium containing ampicillin (100 μg.ml<sup>-1</sup>) and chloramphenicol (35 μg.ml<sup>-1</sup>) at 37°C. Neither iron nor L-cysteine were added to the growth medium. When the culture reached an OD<sup>600</sup> of about 0.8, protein expression was induced with 0.5 mM IPTG and the cells were grown for an additional 3.5 hours before harvesting, flash freezing and storage at -80°C. The pelleted cells were washed, then re-suspended in iron chelation buffer containing 100 mM Tris pH 8.0, 50 mM NaCl, 5 mM EDTA and 5 mM Tiron, supplemented with a protease inhibitor cocktail. Cells were then disrupted by sonication on ice (2 s 40% amplitude, 8 s pause, 3 min total time). The crude extract was subsequently cleared by ultracentrifugation at 180,000 × g during 1 h at 4°C. The clarified supernatant was transferred to a glove box, filtered and loaded onto a streptavidin-agarose column equilibrated with low-salt buffer

containing 100 mM Tris pH 8.0 and 50 mM NaCl. Before elution, the column was first washed with 1.5 column volumes (CV) of iron chelation buffer and then with 3 CV of anaerobic low-salt buffer. The protein was eluted with 1 mM desthiobiotin. The purified apo-protein was concentrated by ultrafiltration (30 kDa molecular weight cut-off) to 30  $\mu$ M and incubated with 5 mM DTT for 15 min. Then, 10 molar equivalents of seleno-L-cystine and 0.01 molar equivalents of NifS from *Azotobacter vinelandii*<sup>3</sup> were added and the solution was incubated, under mild stirring, for 30 min. Subsequently, 10 molar equivalents of  $(\text{NH}_4)_2\text{Fe}(\text{SO}_4)_2$  were added drop-by-drop to the mixture. After 30 min incubation, the dark solution was heated to 70°C for 20 min, then filtered and loaded onto a GE Healthcare HiPrep 16/60 Sephacryl S-100 HR column, buffered with anaerobic low-salt buffer. The *TmHydE* <sup>$\Delta$ FeS</sup> fractions were pooled and the Se-reconstituted desalted protein (Se-sub) was concentrated to 12 mg.mL<sup>-1</sup>.

The HydE protein from *Clostridium acetobutylicum* (*CaHydE*) used for *in vitro* assays was purified according to the following procedure. *E. coli* cells BL21(DE3) cells transformed with the pStrep*CaHydE* plasmid coding for the following protein sequence:

MWSHPQFEKASNI I KLINKAEVTHDLTKDELVTLLKDDTHNEEIYKAADRVREKYVGEVHHL  
RGLIEFSNICKRNCMYCGLRRDNKNIKRYRLEPDEI IHLAKSAKNYGYQTVVLSQGEDDYTT  
VEKMKYIVSEIKKLNMAITLSIGEKTFEEYEEYRKSADRYLIRIETTDKELYEKLDPKMSH  
ENRINCLKNLRKLGYEYVSGCLVGLPNQTIESLADDILFFKEIDADMIGVGPFI PNEDTPLG  
EEKGGEFFMSVKVTALIRLLLPDINIPATTAMESLYPNGRSIALTSGANVVMNVTEGEYRK  
LYALYPGKICVNDTPGHCRQCISLKINKINRKVSATKGFRRKSYKESIG, were grown at  
37°C in LB medium with kanamycine (50  $\mu$ g.mL<sup>-1</sup>) until the culture reached an OD<sup>600</sup>  
of about 0.6. The temperature was then decreased to 21°C and the medium was  
complemented with ammonium ferric citrate (250  $\mu$ M). Protein expression was  
induced at a OD<sup>600</sup> 0.8 by the addition of 1mM IPTG and cells were grown overnight  
at 21°C before harvesting. The cell pellets were re-suspended in 50 mM Tris, 300  
mM KCl pH 8.0 (Tris buffer) supplemented with EDTA-free protease inhibitor cocktail  
and disrupted by sonication on ice (5 s 65% amplitude, 30 s pause, 3 min total time).  
The crude extract was subsequently cleared by ultracentrifugation (150,000  $\times$  g,  
1h30) at 4°C. The supernatant was loaded onto a Strep-Tactin column (Iba®)  
equilibrated with Tris buffer. After an extensive wash step, the protein was eluted  
with the same buffer containing 3 mM desthiobiotin. The protein fractions were  
concentrated by ultrafiltration (30 kDa molecular weight cut-off) to ~8 mg.mL<sup>-1</sup> and  
frozen. *CaHydE* was anaerobically reconstituted by adding a 10-fold excess of  
 $\text{Fe}(\text{NH}_4)_2(\text{SO}_4)_2$  and  $\text{Na}_2\text{S}$  at 12°C. The enzyme preparation was further purified by  
size exclusion chromatography before enzymatic assays performed in Tris buffer pH  
8.

#### Mass spectrometry analysis.

MALDI-TOF mass spectrometry analysis of 5'-dA and SAC was performed on a Voyager DE PRO (Applera, France) in the reflector mode with  $\alpha$ -cyano-4-hydroxycinnamic acid as a matrix.

#### Crystallization and crystal treatments.

Crystals that produced datasets **X1** to **X9** (Supplementary Table S1) were obtained as previously described<sup>4</sup>; minor modifications were applied to the protocol as follows:

- As the *in vivo* reconstituted protein used for crystals **X1**, **X2**, **X4** & **X5** purifies with its  $[\text{Fe}_4\text{S}_4]$  cluster and SAM, no cofactor was added to the crystallization solution. 10  $\mu$ l of *TmHydE* <sup>$\Delta$ FeS</sup> protein sample was incubated with 20  $\mu$ l of 50 mM Tris pH 8.0;

30mM CHAPS. Hanging drops were prepared by mixing 3  $\mu$ l of this solution with 1  $\mu$ l of a solution containing 22-26% PEG 4000; 100 mM Tris pH 8.0; 200 mM Li<sub>2</sub>SO<sub>4</sub>. Crystals were soaked for 2 h in a cryoprotecting solution containing: 30% PEG 4000; 20% glycerol, 80 mM CHAPS; 50 mM Tris pH 8.0; 100 mM Li<sub>2</sub>SO<sub>4</sub>; 20 mM (2*R*,4*R*)-MeTDA (crystals **X1** and **X5**) or (2*R*,4*R*)-MeSeTDA (crystals **X2** and **X4**), respectively. Crystals were subsequently washed in a fresh cryoprotecting solution that did not contain the ligands, composed of 30% PEG 4000; 20% glycerol, 80 mM CHAPS; 50 mM Tris pH 8.0; 100 mM Li<sub>2</sub>SO<sub>4</sub> during 30 min prior to anaerobic flash-cooling (crystals **X1** and **X2**). For the enzymatic kinetic experiments, crystals used to collect datasets **X4** and **X5** were subsequently washed twice during 30 s in two successive 10  $\mu$ L drops of cryoprotecting solution without (2*R*,4*R*)-MeTDA. The transfers were performed using a cryo-loop to minimize the volume of solution brought along with the crystal. The crystals were finally soaked in a fresh cryoprotecting solution containing: 30% PEG 4000; 20% glycerol, 80 mM CHAPS; 50 mM Tris pH 8.0; 100 mM Li<sub>2</sub>SO<sub>4</sub>; 12 mM Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub> and flash frozen at different times (2 hours for crystals **X4** and **X5**).

- **X3** was obtained by co-crystallization of Se-substituted *TmHydE* <sup>$\Delta$ FeS</sup> sample with 10 molar excess of 5'-dA and 10 mM L-Methionine, by the same procedure that led to the 3IIX dataset<sup>5</sup>. Crystallization conditions: 100 mM Tris, pH 8.0; 200 mM Li<sub>2</sub>SO<sub>4</sub>; 26-28% PEG 4000; 80 mM CHAPS. The crystals were soaked for 30 min in a cryoprotecting solution containing: 30% PEG 4000; 20% glycerol, 80 mM CHAPS; 50 mM Tris pH 8.0; 100 mM Li<sub>2</sub>SO<sub>4</sub>; 20 mM 4*R*-MeSeTDA (mixture of the two epimers). Anaerobic flash cooling of the crystals was carried out inside the glove box as previously described<sup>6</sup>.

- A similar procedure was used to obtain the **X6** crystal, except that 10 molar excess *S*-adenosyl-L-cysteine (SAC) replaced 5'-dA and the cryoprotecting solution contained no additional ligands. Crystallization conditions: 100 mM Tris, pH 8.0; 200 mM Li<sub>2</sub>SO<sub>4</sub>; 24% PEG 4000; 80 mM CHAPS. Cryoprotecting solution: 30% PEG 4000; 20% glycerol, 80 mM CHAPS; 50 mM Tris pH 8.0; 100 mM Li<sub>2</sub>SO<sub>4</sub>.

- **X7** and **X8** crystals were grown by co-crystallization of *TmHydE*<sup>WT</sup> sample with 10 molar excess of *S*-adenosyl-L-homocysteine (SAH) and 20 mM of 4*R*-1,3-thiazolidine-2,4-dicarboxylic acid, and 4*R*-1,3-thiazolidine-4-carboxylic acid, respectively. 10  $\mu$ l of protein sample was incubated with 10  $\mu$ l of 100 mM Tris pH 8.0 buffer containing 40 molar excess of SAH and 60mM CHAPS prior to crystallization. Hanging drops were prepared by mixing 2  $\mu$ l of this solution with 1  $\mu$ l of a solution containing 23 % PEG 4000; 100 mM Tris pH 8.0; 200 mM Li<sub>2</sub>SO<sub>4</sub> and 1  $\mu$ l of solution containing 80 mM of ligand. The cryoprotecting solution contained 30% PEG 4000; 20% glycerol, 80 mM CHAPS; 50 mM Tris pH 8.0; 100 mM Li<sub>2</sub>SO<sub>4</sub>; 20 mM of ligand (4*R*-1,3-thiazolidine-2,4-dicarboxylic acid, and 4*R*-1,3-thiazolidine-4-carboxylic acid, respectively).

- **X9** was obtained by co-crystallization of a Se-substituted *TmHydE* <sup>$\Delta$ FeS</sup> sample with 10 molar excess of *S*-adenosyl-L-homocysteine (SAH) and 25 mM of 4*R*-2,5,5-trimethyl-1,3-thiazolidine-2,4-dicarboxylic acid (4*R*-TMeTDA; 2*R* and 2*S* epimer mixture). 10  $\mu$ l of protein sample were incubated with 10  $\mu$ l of 100 mM Tris pH 8.0 buffer containing 40 molar excess of SAH and 60 mM CHAPS prior to crystallization. Hanging drops were prepared by mixing 2  $\mu$ l of this solution with 1  $\mu$ l of a solution containing 24% PEG 4000; 100 mM Tris pH 8.0; 200 mM Li<sub>2</sub>SO<sub>4</sub> and 1  $\mu$ l of solution containing 100 mM of 4*R*-TMeTDA ligand. The cryoprotecting solution contained 30% PEG 4000; 20% glycerol, 80 mM CHAPS; 50 mM Tris pH 8.0; 100 mM Li<sub>2</sub>SO<sub>4</sub>; 25 mM 4*R*-TMeTDA.

### Model refinement and geometry restraints.

The X-ray crystal structures were refined using REFMAC from the CCP4 Package<sup>7</sup>. Geometric restraints for [Fe<sub>4</sub>Se<sub>4</sub>] cluster were manually adjusted based on their standard library description and on a 1.18 Å resolution dataset of a seleno-substituted *TmHydE* crystal previously solved in our laboratory. The geometric restraints for ligands that lack a standard library description were generated with ProDRG<sup>8</sup> from the CCP4 program suite. The initial three-dimensional 1,3-thiazolidine models were based on our high-resolution structures and a geometry optimization step was performed in ProDRG, prior to restraints generation. Although SAM has a standard library topology file, a new file was generated using ProDRG because some of its atomic distances in the standard library are erroneous. As an example, it gives a 1.66 Å bond length for the C-S bond instead of the more canonical 1.82 Å bond length observed in the CSD<sup>9</sup>. In **X4**, the *S*-(5'-adenosyl)-L-selenocysteine ligand was modelled as a selenocysteine linked to 5'-deoxyadenosine. A modified selenocysteine topology file with lower restraints for atomic distances was used for the refinement of this structure. All data processing and refinements statistics are summarized in Supplementary Table S1. Anomalous scattering electron density peaks were used to track the thiazolidine-derived selenium atom in the structure, even at low occupancy.

### Quantum calculations.

All QM calculations were performed using the Schrödinger suite<sup>10</sup>. Our starting model for the 5'-dA• attack on (2*R*,4*R*)-MeSeTDA was the X-ray structure containing 5'-dA and (2*R*,4*R*)-MeSeTDA (**X3** in Table 1). We first constructed a quantum minimal model for the radical, namely CH<sub>3</sub>CH<sub>2</sub>•, using the coordinates of the C4' and C5' atoms in 5'-dA; hydrogen atoms were also constructed and optimized. Both (2*R*,4*R*)-MeSeTDA carboxylate groups were protonated in order to mimic the local charge compensation arising from hydrogen bonding to the protein. The (2*R*,4*R*)-MeSeTDA nitrogen atom N3 was singly protonated considering its pK<sub>a</sub> of 5.86<sup>11</sup>. The DFT method implemented in Jaguar<sup>10</sup> was employed with the B3LYP hybrid functional and the lacvp\*\* basis set. Results were compared to the BHandHLYP functional<sup>12</sup> that, contrary to B3LYP, partially takes non-dynamical electron correlation into account. Note: in all reaction path calculations, the geometries of the reactant, transition state and product are very similar with both functionals; the main difference resides in the transition state energy relative to that of the reactant; besides with the BHandHLYP the activation energy is 2 to 4 times higher than the one obtained with B3LYP. However, the reaction remains possible.

The C5'-Se1-C2 radical attack angle observed in the protein (166° in the crystal structure) was kept to its initial value during the scan of the C5'-Se1 distance from 3.29 to 1.89 Å, geometry-optimizing all atoms. We then optimized the reactants and product and a transition state for the C5'-Se1 bond formation and concomitant Se1-C2 bond cleavage, (Fig. 4a, top). Vibrational frequency calculations were performed to verify that all the minima had positive frequencies and that the transition state had only one imaginary frequency. Using the intrinsic reaction coordinate method, we verified that the pre-reactive and post-reactive complexes corresponded to the C5'• radical attack and the C5'-Se1 bond formation, respectively. Single energy point calculations were done with the lacv3p\*\* basis set for Se and the cc-pVTZ(-f) basis set for all other atoms.

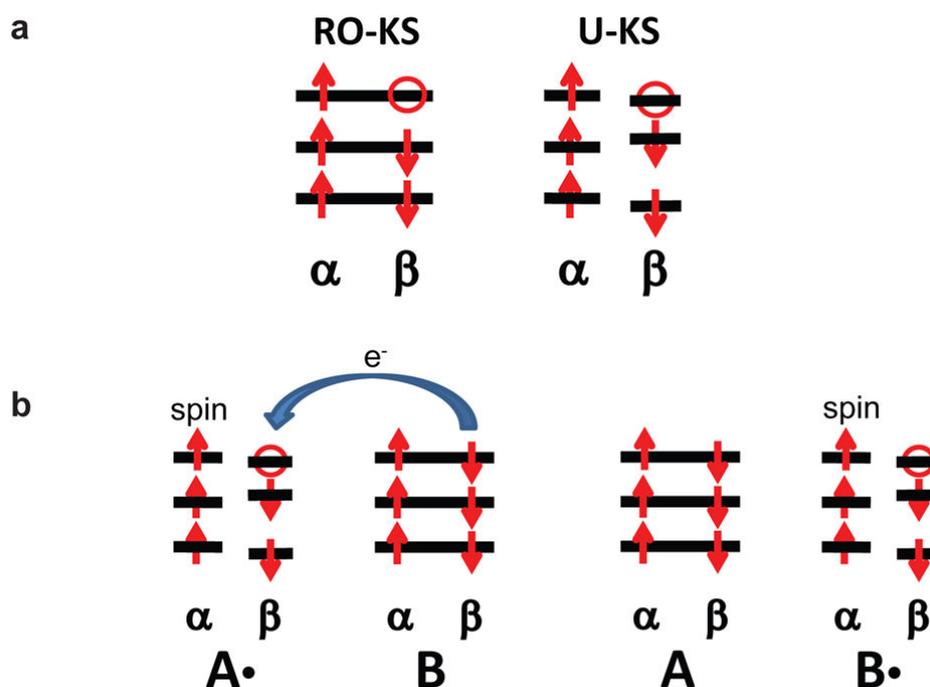
To see the effect of the attack directionality, we placed the  $\text{CH}_3\text{CH}_2\bullet$  fragment so that the  $\text{C5}'\text{-Se1-C5}$  angle was  $166^\circ$  (Supplementary Fig. S9); the protocol described above was used to scan the  $\text{C5}'\text{-Se}$  distance.

To include the protein matrix influence, we performed calculations using the QSite program<sup>10</sup>. Thus, hybrid quantum mechanical/molecular mechanical (QM/MM) calculations allowed us to treat  $5'\text{dA}\bullet$  and  $(2R,4R)\text{-MeSeTDA}$  quantum mechanically while the rest of the system was treated with the molecular mechanical OPLS2005 force field<sup>13</sup>. Both residue side chains missing in the crystal structure and hydrogen atoms were built and the MM part was optimized. The whole system was then geometry-optimized, constraining the atoms of residues lying over 10 Å away from the atoms belonging to the QM part. Starting from  $5'\text{dA}\bullet$  and  $(2R,4R)\text{-MeSeTDA}$ , we then scanned the  $\text{C5}'\text{-Se1}$  distance with the protocol described above (Fig. 4a bottom).

Finally, to evaluate the role of the iron sulfur cluster in the reaction (Figs. 4b, 4c and Supplementary Fig. S10), we geometry-optimized three models where the QM part consisted, on the one hand, of the 3 cysteine side chains coordinating the cluster, the glutamate side chain interacting with the ribose moiety, the arginine side chain interacting with the 1,3-thiazolidine carboxylate group bound to C2 and, on the other hand, either i) the reduced  $[\text{Fe}_4\text{S}_4]^{+1}$ , SAM and  $(2R,4R)\text{-MeSeTDA}$ , ii) the oxidized  $[\text{Fe}_4\text{S}_4]^{+2}$ , Met,  $5'\text{dA}\bullet$  and  $(2R,4R)\text{-MeSeTDA}$  or iii) the oxidized  $[\text{Fe}_4\text{S}_4]^{+2}$ , Met, C2-centered radical intermediate. The rest of the protein was treated with molecular mechanics. Link atoms<sup>14</sup> were placed at the QM ( $\text{C}\beta$ ) / MM ( $\text{C}\alpha$ ) junction of Cys 63, Cys 67, Cys 70, Glu 161 and Arg 159.

In  $[\text{Fe}_4\text{S}_4]$  clusters, the iron ions exhibit local 5/2 (ferric) and 2 (ferrous) high spin configurations. We used the broken symmetry (BS) method to model both an oxidized and a reduced  $[\text{Fe}_4\text{S}_4]$  cluster<sup>15</sup>. The oxidized cluster state consists of two mixed-valence high-spin ( $S=9/2$ ) iron (formally  $\text{Fe}^{\text{II}}\text{-Fe}^{\text{III}}$ ) pairs antiferromagnetically coupled, yielding a cluster spin multiplicity  $M_s = 0$ <sup>16</sup>. We tested the three possible oxidized states, namely, BS12 (initial spin counts:  $\text{Fe}_1 = 5/2$ ,  $\text{Fe}_2 = 4/2$ ,  $\text{Fe}_3 = -5/2$ ,  $\text{Fe}_4 = -4/2$ ), BS13 ( $\text{Fe}_1 = 5/2$ ,  $\text{Fe}_2 = -5/2$ ,  $\text{Fe}_3 = 4/2$ ,  $\text{Fe}_4 = -4/2$ ), and BS23 ( $\text{Fe}_1 = 5/2$ ,  $\text{Fe}_2 = -4/2$ ,  $\text{Fe}_3 = -5/2$ ,  $\text{Fe}_4 = 4/2$ ). Here, we present the results for BS12, which was the most stable state for the oxidized cluster. The reduced cluster state consists of one high-spin ( $S=9/2$ ) mixed-valence iron (formally  $\text{Fe}^{\text{II}}\text{-Fe}^{\text{III}}$ ) pair and a high-spin ( $S=4$ ) ferrous pair ( $\text{Fe}^{\text{II}}\text{-Fe}^{\text{II}}$ ) antiferromagnetically coupled, yielding a total cluster spin of  $S = 1/2$ . We tested the three possible reduced states, namely, BS12 ( $\text{Fe}_1 = 5/2$ ,  $\text{Fe}_2 = 4/2$ ,  $\text{Fe}_3 = -4/2$ ,  $\text{Fe}_4 = -4/2$ ), BS13 ( $\text{Fe}_1 = 5/2$ ,  $\text{Fe}_2 = -4/2$ ,  $\text{Fe}_3 = 4/2$ ,  $\text{Fe}_4 = -4/2$ ), and BS23 ( $\text{Fe}_1 = -4/2$ ,  $\text{Fe}_2 = 5/2$ ,  $\text{Fe}_3 = 4/2$ ,  $\text{Fe}_4 = -4/2$ ). The three other states, where the unique iron is ferric, were not tested as they had been previously excluded<sup>5</sup>. Here, we present the results for BS13, which was the most stable state for the reduced cluster. Models presented in Figs. 4b, 4c and Supplementary Fig. S10 were geometry-optimized using a QM/MM potential with (i) the BLYP functional to slightly exaggerate covalency and ease orbital interpretation<sup>5</sup> and (ii) the lacvp\*\* basis set. Single energy point calculations were done with the lacv3p\*\* basis set for the metals and the cc-pVTZ(-f) basis set for all other atoms. Orbital analyses for all the states were done by extracting the QM atoms for the sake of clarity; both carboxylate groups were protonated to take into account H-bonding with the protein matrix. This allowed us to characterize the radical and its transfer to  $(2R,4R)\text{-MeSeTDA}$  (see below). All these calculations were repeated with sulfur-containing  $(2R,4R)\text{-MeTDA}$  and the results are very similar to those obtained for  $(2R,4R)\text{-MeSeTDA}$  (Supplementary Fig. S7).

**Orbital analysis:** In a simplified *restricted* open-shell Kohn-Sham (RO-KS) scheme where all electrons are paired but the highest occupied molecular orbital (HOMO)  $\alpha$  (which is therefore also here the radical singly occupied  $\alpha$  orbital: SOMO  $\alpha$ ), one can follow the radical electron either with the HOMO  $\alpha$  or with its empty spin  $\beta$  lowest occupied molecular orbital (LUMO) because they are orbitally identical (Supplementary Scheme S1a, left). When performing *unrestricted* Kohn-Sham (U-KS) QM/MM calculations, as we did in this study, spin polarization breaks the  $\alpha$  versus  $\beta$  orbital symmetry (Supplementary Scheme S1a, right). In our experience<sup>17,18</sup>, orbital mixing within the  $\alpha$  spin orbital set may blur part of the orbital interaction within the radical SOMO  $\alpha$  (which is no longer necessarily the HOMO  $\alpha$ : see for example Fig. 4c or Supplementary Fig. S10a); in contrast, the LUMO  $\beta$  “hole” is less affected (being empty) and reflects the evolution of the radical character as the reaction proceeds. We therefore used both HOMO  $\alpha$  and LUMO  $\beta$  combined to visually follow the oppositely oriented radical *versus* charge transfers (Supplementary Scheme S1b). We also chose to present in some cases the HOMO  $\alpha$  (to show where the unpaired electron actually is) and/or the LUMO  $\beta$  (to anticipate where the radical character will appear).

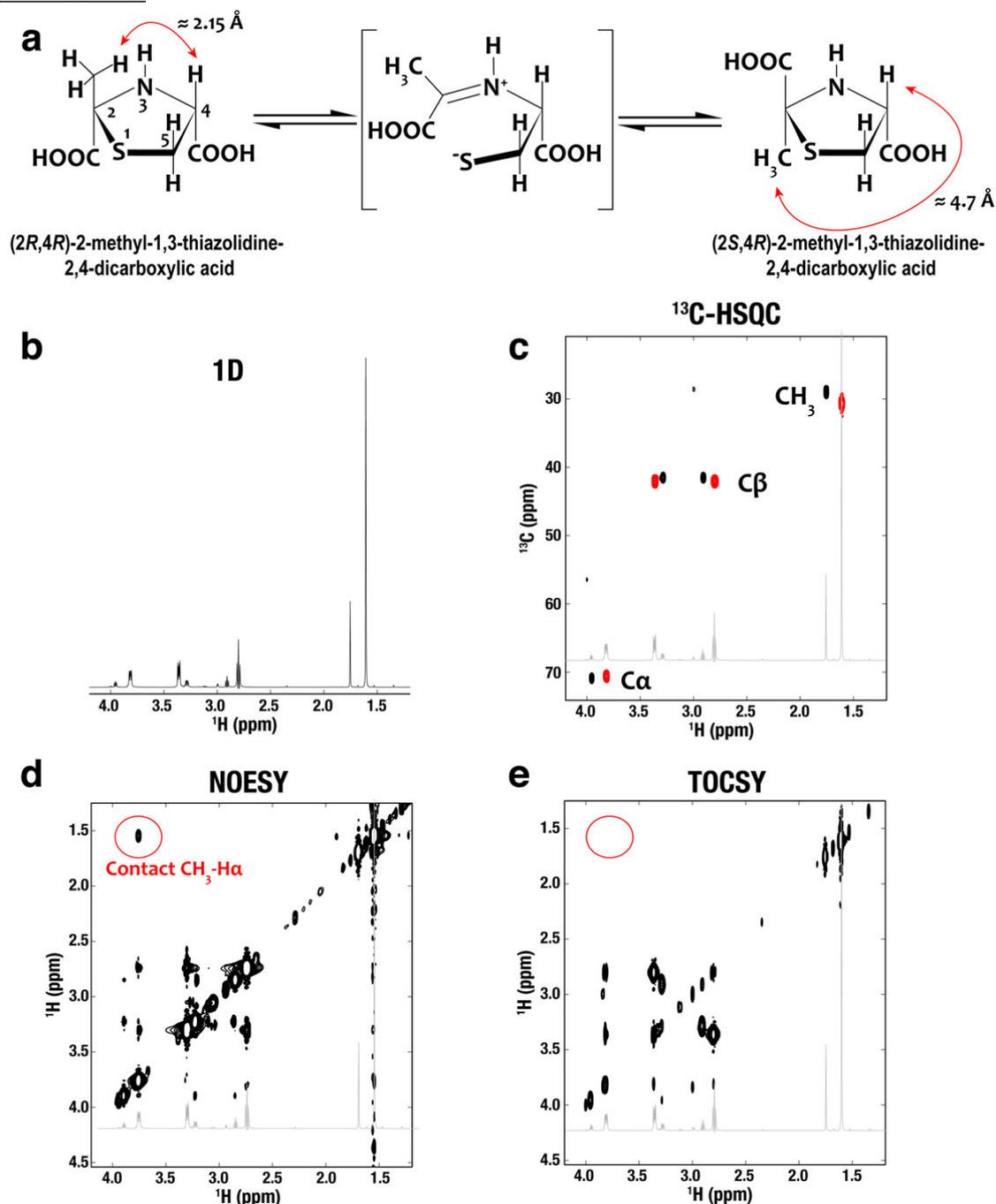


**Supplementary Scheme S1 | Representation of radical versus charge transfers.**  
**a.** Left: schematic restricted-open shell Kohn-Sham (RO-KS) orbital level diagram with all doubly and one singly occupied molecular orbitals. The HOMO  $\alpha$  (highest occupied molecular orbital) and LUMO  $\beta$  (lowest unoccupied molecular orbital) are identical except for the spin. Right: unrestricted Kohn-Sham (U-KS) scheme where spin polarization (within each spin set) breaks the orbital symmetry between  $\alpha$  and  $\beta$  spin sets. The HOMO  $\alpha$  and LUMO  $\beta$  are no longer identical although they remain similar. **b.** The ( $\beta$  spin) electron transfer from donor B to (radical) acceptor A• (left) results in a parallel radical spin transfer from A• to B (now B•: right). Both charge and spin transfers can be followed either with the HOMO  $\alpha$  or the LUMO  $\beta$ .

### NMR spectroscopy analyses

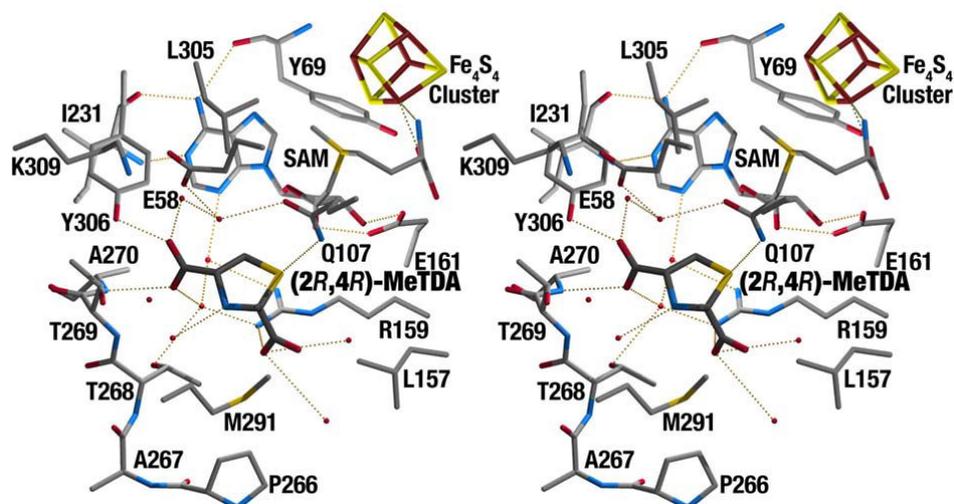
The NMR experiments were performed at 300 K on a Bruker Avance III 600 MHz spectrometer equipped with a cryoprobe. The 4*R*-2-methyl-1,3-thiazolidine-2,4-dicarboxylic acid mixture (originating from a 1:1 mixture of L-cysteine and pyruvate) was characterized by 1D, 2D TOCSY and NOESY experiments performed using an excitation-sculpting scheme for water suppression and a gradient selected sensitivity enhanced <sup>13</sup>C-HSQC experiment. 540 μl of 100 mM 4*R*-MeTDA mixture stock solution in water, pH 8.0 were mixed with 60 μl of D<sub>2</sub>O. For the enzymatic kinetic experiment, 615 μl of 200 μM anaerobically reconstituted CaHydE; 50 mM phosphate buffer; 300 mM KCl, pH 8.0 were mixed inside a glove box with 70 μl D<sub>2</sub>O and 7 μl 100 μM <sup>13</sup>C-methyl labelled (2*R*,4*R*)-MeTDA ligand stock solution. The reaction was followed using a 1D version of the methyl SOFAST pulse sequence to optimize the experiment sensitivity. In order to prevent any deleterious oxygen access, the reaction mixture was anaerobically introduced into a gas-tight J. Young NMR tube and rapidly taken out the glove box for immediate NMR analysis. The reaction was triggered by addition of 7 μl fresh 100 μM Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub> just before the NMR measurements.

## II. Results

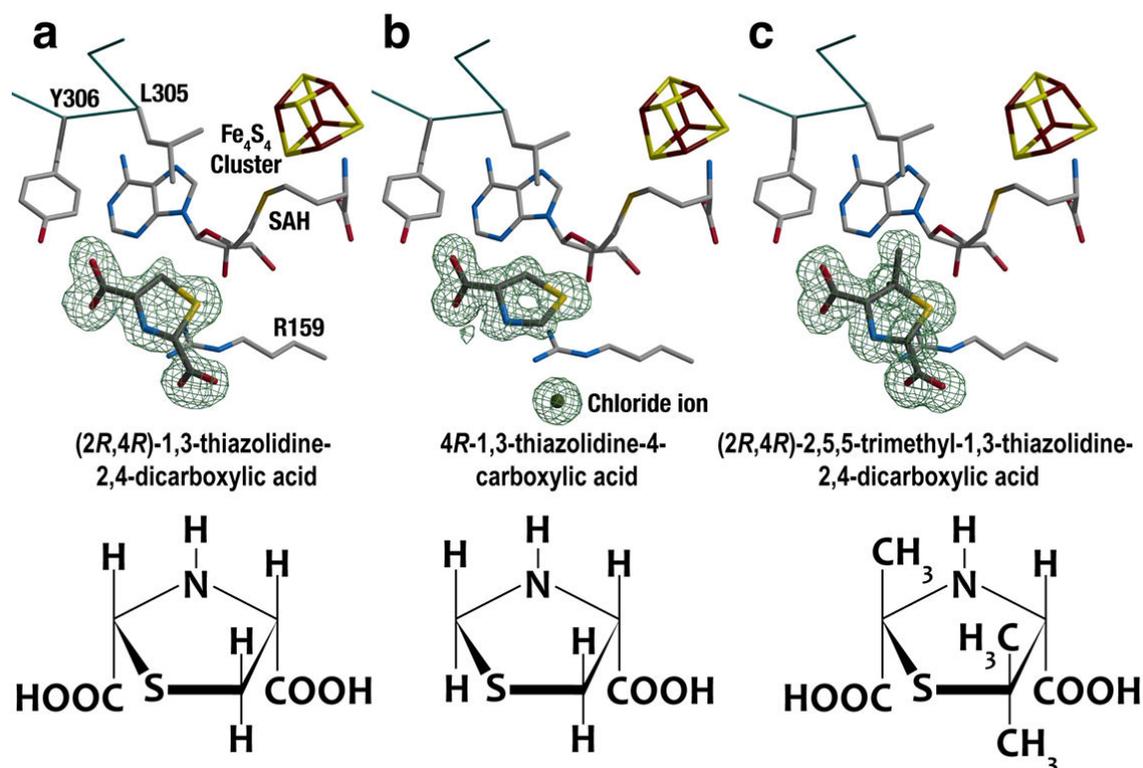


**Supplementary Figure S1 | Characterization of (2*R*,4*R*)-MeTDA and (2*S*,4*R*)-MeTDA from a 1:1 mixture of L-cysteine and pyruvate. **a.** Putative mechanism for the epimerization of 4*R*-2-methyl-1,3-thiazolidine-2,4-dicarboxylic acid at C2<sup>19</sup>. The minimal distances for methyl proton-H $\alpha$  for each epimer are indicated in red. **b.** 1D <sup>1</sup>H-NMR spectrum from a 1:1 mixture of L-cysteine and pyruvate (see Fig. 1d). Panels **c**, **d** and **e.** <sup>13</sup>C-edited HSQC, 2D-NOESY and 2D-TOCSY spectra of the same mixture as in **b.** The 1D <sup>1</sup>H-NMR spectrum is added for clarity. The H $\alpha$ , H $\beta$  and methyl protons were assigned from their bonded <sup>13</sup>C chemical shift (**c**). The highlighted cross peak observed only in the NOESY spectrum for the main**

conformation corresponds to a short methyl proton-H $\alpha$  contact and was thus assigned to the (2*R*,4*R*)-2-methyl-1,3-thiazolidine-2,4-dicarboxylic acid epimer (a).

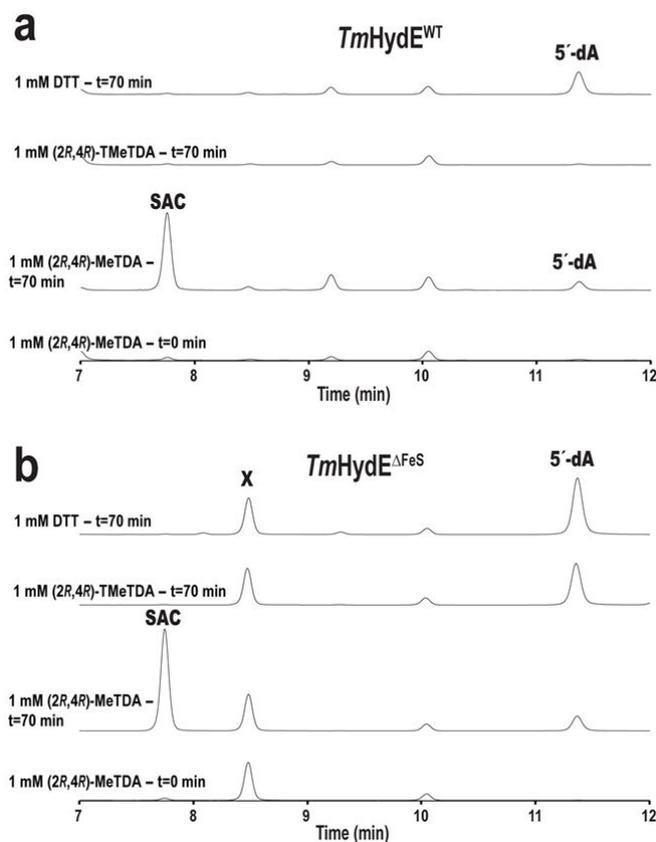


**Supplementary Figure S2 | Stereo view of the cavity surrounding (2*R*,4*R*)-MeTDA.** The orientation is the same as in Fig. 1 in the main text.

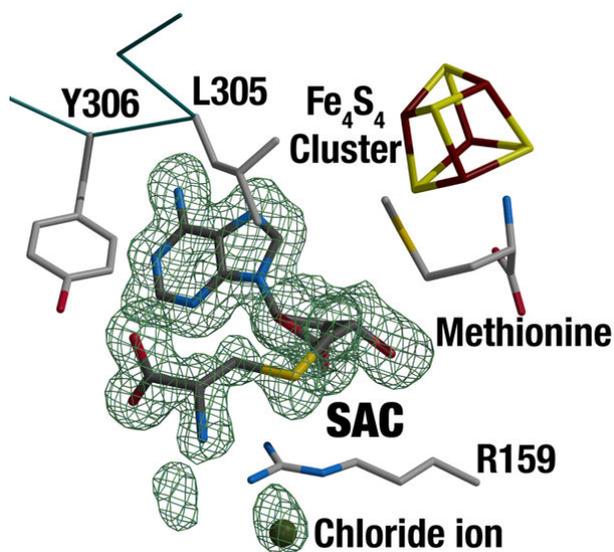


**Supplementary Figure S3 | *Fo-Fc* omit maps for various 1,3-thiazolidine analog complexes in HydE.** The orientation and representation are the same as in Fig. 1. The omit maps contoured at 3  $\sigma$  (cover radius 2.0 Å) are depicted as green meshes **a.** crystal **X7** corresponding to a complex of HydE with (2*R*,4*R*)-1,3-thiazolidine-2,4-dicarboxylic acid (1:1 L-cysteine + glyoxylate mixture). **b.** crystal **X8** corresponding to

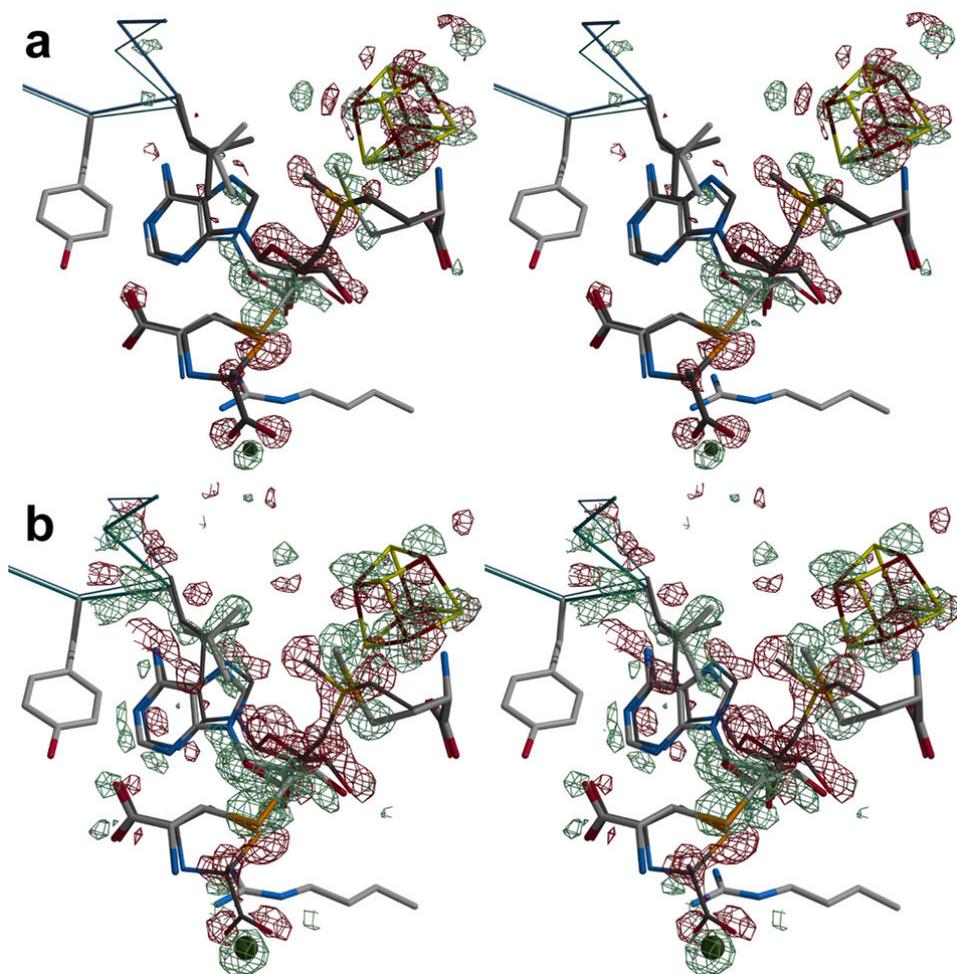
a complex of HydE with 4*R*-1,3-thiazolidine-4-carboxylic acid (1:1 L-cysteine + formaldehyde mixture) and **d.** crystal **X9** corresponding to a complex of HydE with (2*R*,4*R*)-2,5,5-trimethyl-1,3-thiazolidine-2,4-dicarboxylic acid (1:1 L-penicillamine + pyruvate mixture).



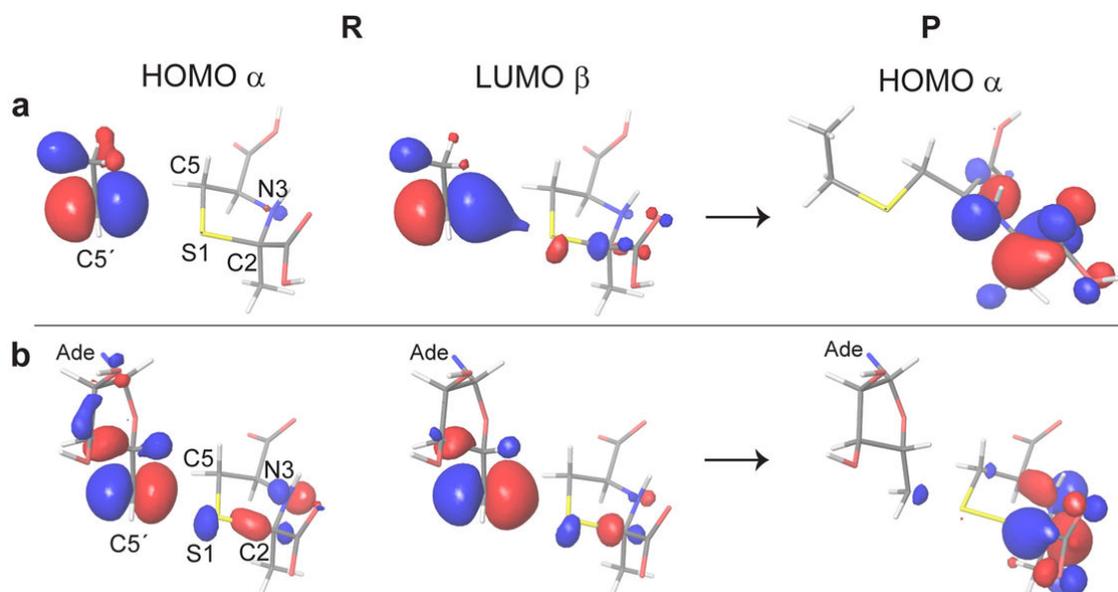
**Supplementary Figure S4 | Reverse phase HPLC reactions profiles. a and b.** SAC production from (2*R*,4*R*)-MeTDA by *TmHydE* with or without its additional [Fe<sub>4</sub>S<sub>4</sub>] cluster (*TmHydE*<sup>WT</sup> and *TmHydE*<sup>ΔFeS</sup>), respectively. The conditions are the same as in the *in vitro* analyses performed using *CaHydE* protein. Traces corresponding to t=0 and t<sub>final</sub> (70 min) are depicted. The non-substrate ligand (2*R*,4*R*)-TMeTDA was also tested, as a control. As *CaHydE*, *TmHydE* produces SAC *in vitro* from 1,3-thiazolidines and the second cluster does not play any role during the reaction.



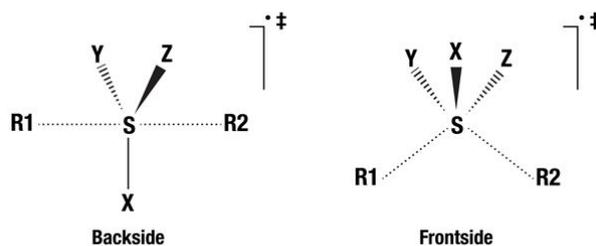
**Supplementary Figure S5 | *Fo-Fc* omit map for the final product in HydE fitted with the *S*-adenosyl-*L*-cysteine (SAC) complex structure.** The omit map contoured at  $4\sigma$  (cover radius 3.5 Å) is depicted as a green mesh. It corresponds to crystal **X5** flash cooled after a 2 h reaction started with (2*R*,4*R*)-MeTDA (Supplementary Table S1). The stick model corresponds to the *Tm*HydE – SAC complex (crystal **X6**) superimposed to crystal **X5**. The excellent fit between the model and the electron density map strongly supports SAC as the final product of the reaction in HydE crystals. This also corresponds to what we observed in solution.



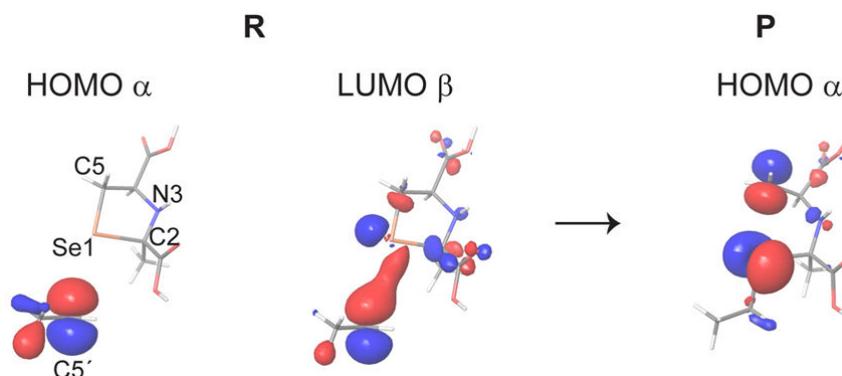
**Supplementary Figure S6 | Stereo view of the  $F_o$  double-difference electron density maps depicting the conversion of (2*R*,4*R*)-MeSeTDA to SeAC.** The orientation is the same as in Fig. 3 of the main text. The  $F_{o1}-F_{o2}$  difference maps depict the differences between two datasets (1 and 2 and not between a dataset and a model as in  $F_o-F_c$  omit maps). The  $F_{o1}-F_{o2}$  difference map is very well suited to follow even small structural changes resulting from a reaction taking place in isomorphous crystals, like in our case (Supplementary Table S1). Using these maps we can follow both the gradual conversion of (2*R*,4*R*)-MeTDA into SeAC and the atomic motions associated with this reaction. All datasets were first scaled to each other using SCALEIT from the CCP4 suite<sup>7</sup>. Phases for Fourier transform calculations were calculated using the initial model (crystal **X2**); phases calculated from any of the other models from the kinetic series produced very similar maps. **a** corresponds to the  $t=45 \text{ min} - t_0$   $F_o$  double-difference map and **b** to the  $t=2 \text{ h} - t_0$  double-difference map. Green and red meshes represent positive and negative electron density peaks, respectively.



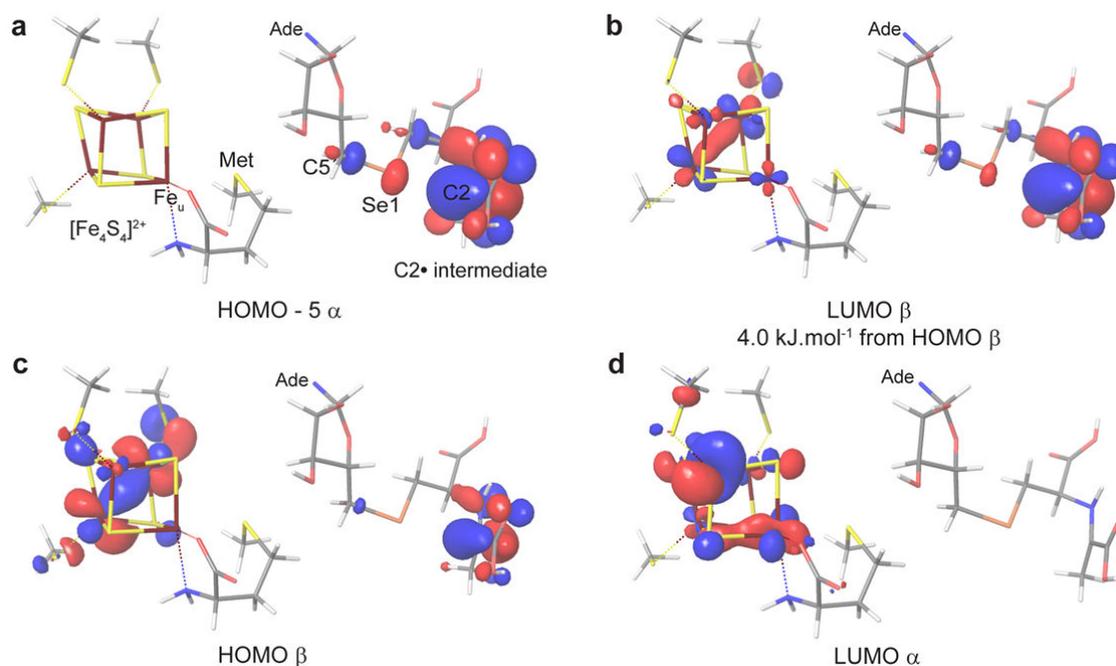
**Supplementary Figure S7 | Models of the radical C5' attack on (2*R*,4*R*)-MeTDA (first step in S-containing SAC formation).** **a.** Small QM model of the reaction where the 5'dA• is modeled as CH<sub>3</sub>CH<sub>2</sub>• and **b.** QM/MM model of the same reaction (only the QM part is shown here). The HOMO  $\alpha$  and LUMO  $\beta$  are represented for the reactant and the HOMO  $\alpha$  for the product (see Supplementary Scheme S1 for the combined use of HOMO  $\alpha$  and LUMO  $\beta$  to follow the reaction). [+] and [-] molecular orbital surfaces are depicted in red and blue, respectively (isosurface value = 0.05). Standard atom color code is used.



**Supplementary Figure S8 | Backside versus frontside mechanisms for S<sub>H</sub>2 reactions.** R1 and R2 represent the attacking and leaving groups, respectively. In the backside mechanism (left), the direction of the forming bond between R1 and S lines up with that of the disrupting bond between S and R2. This mechanism is reminiscent of the nucleophilic substitution S<sub>N</sub>2 and leads to an inversion of the stereochemistry. Conversely, in a frontside mechanism (right), the bond that is being formed between R1 and S is on the same side as the bond being broken between S and R2 and the stereochemistry is kept<sup>20</sup>.



**Supplementary Figure S9 | Model of the radical C5' attack on (2*R*,4*R*)-MeSeTDA along the Se-C5 direction.** The HOMO  $\alpha$  and LUMO  $\beta$  are represented for the reactant (**R**) and the HOMO  $\alpha$  for the product (**P**) (see Supplementary Scheme S1 for the combined use of HOMO  $\alpha$  and LUMO  $\beta$  to follow the reaction). [+] and [-] molecular orbital surfaces are depicted in red and blue, respectively (isosurface value = 0.05). Standard atom color code is used.



**Supplementary Figure S10 | Inclusion of the  $[\text{Fe}_4\text{S}_4]$  cluster in the QM part.** **a.** The combination of SOMO  $\alpha$  (which is here the HOMO-5  $\alpha$ ) and **b.** LUMO  $\beta$  show that the radical character is located on C2. Moreover, the LUMO  $\beta$  (**b**) and HOMO  $\beta$  (**c**) are nearly degenerate ( $4 \text{ kJ.mol}^{-1}$ ), causing the partial delocalization of both orbitals between the  $[\text{Fe}_4\text{S}_4]$  cluster and C2. This is reminiscent of what was observed before SAM cleavage (Fig. 4b). Notice the presence of the unique  $\text{Fe}_u$  iron in the LUMO  $\beta$ , which reflects the fact that C2 and  $\text{Fe}_u$  orbitally communicate, suggesting one electron transfer from C2 back to the  $[\text{Fe}_4\text{S}_4]$  cluster. This is also supported by the LUMO  $\alpha$  (**d**), located on the  $[\text{Fe}_4\text{S}_4]$  cluster, which will be the final electron acceptor (see Supplementary Scheme S1 for the combined use of SOMO  $\alpha$  and LUMO  $\beta$  to follow the reaction). [+] and [-] molecular orbital surfaces are depicted in red and blue, respectively (isosurface value = 0.03). Standard atom color code is used.

Supplementary Table S1 (part 1) | Statistics for data processing and X-ray structure refinement.

Dataset PDB code	X1			X2			X3			X4			X5			X6		
Crystal content	t <sub>0</sub> (kinetics); SAM & (2 <i>R</i> ,4 <i>R</i> )-MeTDA			t <sub>0</sub> (kinetics); SAM & (2 <i>R</i> ,4 <i>R</i> )-MeSeTDA			5'-dA + L-Met & (2 <i>R</i> ,4 <i>R</i> )-MeSeTDA			t <sub>final</sub> (kinetics); L-Met & Se-SAC			t <sub>final</sub> (kinetics); L-Met & SAC			SAC & L-Met		
<b>Data collection</b>																		
Cell parameters (Å)	51.17	80.02	86.36	51.07	79.10	86.26	50.80	78.87	86.25	51.12	79.46	86.57	50.99	78.97	86.27	50.89	79.34	86.20
Space group	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>			<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>			<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>			<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>			<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>			<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>		
Wavelength (Å)	0.979844			0.980250			0.979129			0.979677			0.979247			0.979107		
Resolution (Å)	1.45 (1.54-1.45)			1.27 (1.35-1.27)			1.35 (1.43-1.35)			1.32 (1.40-1.32)			1.17 (1.24-1.17)			1.49 (1.58-1.49)		
Rsym, %	4.7 (81.0)			4.9 (62.7)			7.2 (93.9)			5.5 (66.3)			9.7 (174.4)			6.0 (51.1)		
I/σ	14.86 (1.69)			13.97 (2.01)			15.27 (1.98)			12.61 (1.80)			19.08 (2.21)			13.73 (2.03)		
CC(1/2)	99.9 (67.0)			99.9 (68.0)			99.9 (74.5)			99.9 (70.1)			99.9 (80.9)			99.8 (74.3)		
Completeness (%)	99.5 (98.6)			99.1 (98.8)			99.8 (99.4)			98.8 (97.6)			99.6 (98.4)			98.9 (93.9)		
Multiplicity	3.4 (3.3)			2.9 (2.8)			5.9 (5.8)			3.2 (3.1)			20.0 (18.9)			3.8 (2.7)		
<b>Refinement statistics</b>																		
R <sub>cryst</sub> /R <sub>free</sub> <sup>21</sup>	16.16/18.80			10.77/13.53			11.90/13.52			11.29/14.67			10.97/13.23			15.03/16.85		
No. of reflections	60277/3183			87752/4519			72790/3833			78733/4132			112915/5529			54620/2898		
<b>Rms deviation from ideal geometry</b>																		
Bonds	0.0231			0.0264			0.0243			0.0252			0.0269			0.0239		
Angles	2.3962			2.5138			2.2769			2.4474			2.4886			2.3266		
# refined atoms	3729			3577			3638			3593			3631			3703		
Protein	3003			2868			2892			2887			2930			2956		
Water molecules	486			469			506			469			427			503		
Iron atoms	4			4			4			4			4			4		
others	236			236			236			233			270			240		
<b>Ramachandran plots</b>																		
Preferred/allowed/outliers	95.2% 4.8% 0%			96.1% 3.6% 0.3%			95.6% 4.4% 0%			95.7% 4.3% 0%			95.8% 4.2% 0%			96.4% 3% 0.7%		
<b>B-factor values (Å<sup>2</sup>)</b>																		
Protein	20.57			16.36			15.58			16.68			16.20			18.85		
Ligand	23.58			23.75			20.31			15.76			16.71			17.77		
SAM or 5'-dA	20.64 (SAM)			15.46 (SAM)			11.60 (5'-dA)			16.00 (5'-dA)			Not Applicable			Not Applicable		
Fe <sub>4</sub> S <sub>4</sub> cluster	18.22			13.32			13.23			13.48			11.94			18.31		

Supplementary Table S1 (part 2)

Dataset	X7			X8			X9		
PDB code									
Crystal content	SAH & (2 <i>R</i> ,4 <i>R</i> )-TDA			SAH & 4 <i>R</i> -TCA			SAH & (2 <i>R</i> ,4 <i>R</i> )-TMeTDA		
<b>Data collection</b>									
Cell parameters (Å)	51.07	79.10	86.26	51.00	79.18	86.24	51.11	79.38	86.16
Space group	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>			<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>			<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>		
Wavelength (Å)	0.979688			0.979774			0.920423		
Resolution (Å)	1.47 (1.56-1.47)			1.18 (1.25-1.18)			1.35 (1.43-1.35)		
Rsym, %	8.5 (68.3)			6.7 (77.9)			4.8 (57.5)		
1/σ	8.05 (1.78)			12.65 (1.79)			12.33 (1.58)		
CC(1/2)	99.5 (66.6)			99.9 (70.0)			99.9 (67.0)		
Completeness (%)	98.4 (98.2)			99.6 (98.9)			98.7 (96.3)		
Multiplicity	2.9 (2.8)			4.5 (4.3)			2.7 (2.3)		
<b>Refinement statistics</b>									
R <sub>cryst</sub> /R <sub>free</sub> <sup>21</sup>	16.50/18.65			11.78/14.22			11.57/15.28		
No. of reflections									
Work/Test set	56772/3009			109463/5384			73521/3868		
<b>Rms deviation from ideal geometry</b>									
Bonds	0.0216			0.025			00.0247		
Angles	2.2962			2.5914			2.3603		
# refined atoms	3559			3735			3705		
Protein	2874			3045			2951		
Water molecules	448			448			511		
Iron atoms	4			6			4		
others	233			236			239		
<b>Ramachandran plots</b>									
Preferred/allowed/outlier s	95.9% 3.8% 0.3%			95.9% 4.1% 0%			95.3% 4.3% 0.3%		
<b>B-factor values (Å<sup>2</sup>)</b>									
Protein	17.63			16.14			18.22		
Ligand	20.36			16.97			16.04		
SAM, SAH or 5'-dA	15.53 (SAH)			14.00 (SAH)			15.84 (SAH)		
Fe <sub>4</sub> S <sub>4</sub> cluster	13.06			11.60			16.86		

The MWSHPQFEKN N-ter strep tag residues and the Cter residue V348 are missing in all the structures.

**Supplementary Table S2 I** For the models we used for calculations (QM/MM models in bold), the energies in  $\text{kJ}\cdot\text{mol}^{-1}$  of the transition state (**TS**) and the product (**P**), relative to the one of the reactant (**R**), are indicated for the B3LYP and, in parentheses, the BHandHLYP exchange correlation (XC) potentials. The evolution of the C5'-Se1 (d1) and the Se1-C2 (d2) distances is also given in Å.

Models	Figure		R	TS	P
(2 <i>R</i> ,4 <i>R</i> )-MeSeTDA	4a, top	E	0.0 (0.0)	9.7 (26.8)	-112.8 (-93.9)
		d1 / d2	3.50 / 2.15	2.67 / 2.27	2.01 / 4.25
(2 <i>R</i> ,4 <i>R</i> )-MeSeTDA*	S9	E	0.0 (0.0)	41.0 (61.5)	16.7 (18.0)
		d1 / d2'	3.60 / 2.00	2.28 / 2.26	2.03 / 3.12
(2 <i>R</i> ,4 <i>R</i> )-MeTDA	S7a	E	0.0 (0.0)	24.8 (46.9)	-105.7 (-93.9)
		d1 / d2	3.50 / 1.93	2.40 / 2.11	1.84 / 4.19
<b>(2<i>R</i>,4<i>R</i>)-MeSeTDA</b>	4a, bottom	E	0.0 (0.0)	8.2 (31.0)	-41.2 (-43.5)
		d1 / d2	3.00 / 2.15	2.57 / 2.32	2.03 / 2.93
<b>(2<i>R</i>,4<i>R</i>)-MeTDA</b>	S7b	E	0.0 (0.0)	27.3 (54.6)	-50.8 / -48.7
		d1 / d2	2.88 / 1.94	2.38 / 2.22	1.85 / 2.96

\*This corresponds to the attack along the Se1-C5 bond (Supplementary Fig. S9); d2' is the Se1-C5 distance (in Å).

**Supplementary Table S3 I** Main distances measured in the different crystal structures illustrate the motions associated to the conversion of (2*R*,4*R*)-Me(Se)TDA to S(Se)AC. The colors in brackets refer to the ones used to represent the particular structure in Fig. 3e of the main text. \* Indicates that these structures contain selenium from (2*R*,4*R*)-MeSeTDA rather than sulfur from (2*R*,4*R*)-MeTDA. These distances can be compared to the tabulated van der Waals radii for the corresponding atoms, which are C = 1.70 Å; S = 1.80 Å; Se = 1.90 Å; H = 1.2 Å. This comparison indicates that SAM and (2*R*,4*R*)-MeTDA are already in van der Waals contacts before the reaction.

Bond	X1 (Blue)	X2 (Purple)	3IIZ (Cyan)	X3 (Green)	3IIX (Orange)	X4 (Yellow)	X5 (Crimson)	X6 (Red)
Fe-S $\delta$	3.25 Å	3.23 Å	3.25 Å	2.63 Å	2.67 Å	2.88 Å	2.82 Å	2.78 Å
S $\delta$ -C5'	1.86 Å	1.81 Å	1.83 Å	3.50 Å	3.80 Å	3.91 Å	3.91 Å	3.81 Å
C5'-S1	3.85 Å	3.91 Å*	-	3.31 Å*	-	2.08 Å*	1.83 Å	1.94 Å
S1-C2	1.80 Å	1.99 Å*	-	1.93 Å*	-	-	-	-

### III. References

- Nicolet, Y., Rohac, R., Martin, L. & Fontecilla-Camps, J. C. X-ray snapshots of possible intermediates in the time course of synthesis and degradation of protein-bound Fe<sub>4</sub>S<sub>4</sub> clusters. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 7188–7192 (2013).
- Nicolet, Y., Zeppieri, L., Amara, P. & Fontecilla-Camps, J. C. Crystal structure of tryptophan lyase (NosL): evidence for radical formation at the amino group of tryptophan. *Angew. Chem. Int. Ed* **53**, 11840–11844 (2014).
- Zheng, L., White, R. H., Cash, V. L., Jack, R. F. & Dean, D. R. Cysteine desulfurase activity indicates a role for NIFS in metallocluster biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 2754–2758 (1993).
- Nicolet, Y. *et al.* X-ray structure of the [FeFe]-hydrogenase maturase HydE from *Thermotoga maritima*. *J. Biol. Chem.* **283**, 18861–18872 (2008).
- Nicolet, Y., Amara, P., Mouesca, J.-M. & Fontecilla-Camps, J. C. Unexpected electron transfer mechanism upon AdoMet cleavage in radical SAM proteins. *Proc.*

- Natl. Acad. Sci. U. S. A.* **106**, 14867–14871 (2009).
6. Vernede, X. & Fontecilla-Camps, J. C. A method to stabilize reduced and or gas-treated protein crystals by flash-cooling under a controlled atmosphere. *J. Appl. Crystallogr.* **32**, 505–509 (1999).
  7. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. -Biol. Crystallogr.* **67**, 235–242 (2011).
  8. Schuttelkopf, A. W. & van Aalten, D. M. F. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 1355–1363 (2004).
  9. Allen, F. H. The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallogr. B* **58**, 380–388 (2002).
  10. a) Maestro version 9.6 b) QSite version 6.1 c) Jaguar version 8.2, Schrödinger, LLC, New York, NY, 2013.
  11. Butvin, P., Mitkova, M., Svetlik, J. & Havranek, E. Complex formation of thiazolidine-2,4-dicarboxylic acid with selected divalent and trivalent metal ions. *Chem. Pap.* **56**, 174–177 (2002).
  12. Aitken, H. M., Hancock, A. N. & Schiesser, C. H. Understanding (the lack of) homolytic substitution chemistry of sulfones. *Chem. Commun.* **48**, 8326–8328 (2012).
  13. Jorgensen, W. L., Maxwell, D. S. & TiradoRives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **118**, 11225–11236 (1996).
  14. Amara, P. & Field, M. J. Evaluation of an *ab initio* quantum mechanical/molecular mechanical hybrid-potential link-atom method. *Theor. Chem. Acc.* **109**, 43–52 (2003).
  15. Noodleman, L., Peng, C., Case, D. & Mouesca, J. Orbital interactions, electron delocalization and spin coupling in iron-sulfur clusters. *Coord. Chem. Rev.* **144**, 199–244 (1995).
  16. Mouesca, J.-M. Density functional theory-broken symmetry (DFT-BS) methodology applied to electronic and magnetic properties of bioinorganic prosthetic groups. *Methods Mol. Biol.* **1122**, 269–296 (2014).
  17. Onofrio, N. & Mouesca, J.-M. Valence bond/broken symmetry analysis of the exchange coupling constant in copper(II) dimers. Ferromagnetic contribution exalted through combined ligand topology and (singlet) covalent-ionic mixing. *J. Phys. Chem. A* **114**, 6149–6156 (2010).
  18. Onofrio, N. & Mouesca, J.-M. Analysis of the singlet-triplet splitting computed by the density functional theory-broken-symmetry method: is it an exchange coupling constant? *Inorg. Chem.* **50**, 5577–5586 (2011).
  19. Nagasawa, H., Goon, D. & Shirota, F. Epimerization at C-2 of 2-substituted thiazolidine-4-carboxylic acid. *J. Heterocycl. Chem.* **18**, 1047–1051 (1981).
  20. Schiesser, C. H. Taming the free radical shrew - learning to control homolytic reactions at higher heteroatoms. *Chem. Commun.* 4055–4065 (2006).
  21. Brunger, A. T. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472–475 (1992).



# X-ray snapshots of possible intermediates in the time course of synthesis and degradation of protein-bound Fe<sub>4</sub>S<sub>4</sub> clusters

Yvain Nicolet<sup>1</sup>, Roman Rohac, Lydie Martin, and Juan C. Fontecilla-Camps<sup>1</sup>

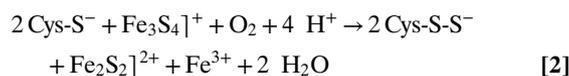
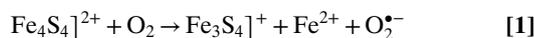
Metalloproteins Unit, Institut de Biologie Structurale J.-P. Ebel, Commissariat à l'Energie Atomique–Centre National de la Recherche Scientifique–Université Grenoble-Alpes, 38027 Grenoble, France

Edited by Richard H. Holm, Harvard University, Cambridge, MA, and approved March 25, 2013 (received for review February 5, 2013)

Fe<sub>4</sub>S<sub>4</sub> clusters are very common versatile prosthetic groups in proteins. Their redox property of being sensitive to O<sub>2</sub>-induced oxidative damage is, for instance, used by the cell to sense oxygen levels and switch between aerobic and anaerobic metabolisms, as exemplified by the fumarate, nitrate reduction regulator (FNR). Using the hydrogenase maturase HydE from *Thermotoga maritima* as a template, we obtained several unusual forms of FeS clusters, some of which are associated with important structural changes. These structures represent intermediate states relevant to both FeS cluster assembly and degradation. We observe one Fe<sub>2</sub>S<sub>2</sub> cluster bound by two cysteine persulfide residues. This observation lends structural support to a very recent Raman study, which reported that Fe<sub>4</sub>S<sub>4</sub>-to-Fe<sub>2</sub>S<sub>2</sub> cluster conversion upon oxygen exposure in FNR resulted in concomitant production of cysteine persulfide as cluster ligands. Similar persulfide ligands have been observed in vitro for several other Fe<sub>4</sub>S<sub>4</sub> cluster-containing proteins. We have also monitored FeS cluster conversion directly in our protein crystals. Our structures indicate that the Fe<sub>4</sub>S<sub>4</sub>-to-Fe<sub>2</sub>S<sub>2</sub> change requires large structural modifications, which are most likely responsible for the dimer–monomer transition in FNR.

FeFe-hydrogenase | X-ray crystallography

Iron–sulfur (FeS) clusters are among the most ubiquitous and functionally versatile prosthetic groups in nature. They can be involved in electron transfer, substrate binding and activation, be used as stress sensors to regulate gene expression or simply play a structural role (1). FeS clusters have variable compositions such as the common Fe<sub>2</sub>S<sub>2</sub>, Fe<sub>3</sub>S<sub>4</sub>, and Fe<sub>4</sub>S<sub>4</sub> centers or the more unusual FeMo-cofactor Fe<sub>7</sub>S<sub>9</sub>Mo cluster of nitrogenase (2). When Fe<sub>4</sub>S<sub>4</sub> clusters are redox active, they are also usually sensitive to oxidation and can be damaged (or disassembled) by oxygen. *Escherichia coli* uses this property to switch between aerobic and anaerobic metabolisms (3, 4). Indeed, at under about 0.5% oxygen concentration, the fumarate nitrate reduction regulator (FNR) can accommodate an Fe<sub>4</sub>S<sub>4</sub> cluster and is a dimer that can bind specific palindromic DNA sequences to either repress or activate the expression of hundreds of genes (5). When exposed to higher oxygen levels, the Fe<sub>4</sub>S<sub>4</sub> cluster is rapidly converted into an Fe<sub>2</sub>S<sub>2</sub> cluster, causing FNR to dissociate into monomers that do not bind DNA. Finally, the Fe<sub>2</sub>S<sub>2</sub> cluster slowly decomposes into two Fe<sup>3+</sup> and two S<sup>2-</sup> ions, leading to the apo protein form of FNR (6). Thus, the cell response to oxygen and its ensuing commitment to aerobic metabolism are modulated by structural changes associated with the rapid conversion of the Fe<sub>4</sub>S<sub>4</sub> cluster into an Fe<sub>2</sub>S<sub>2</sub> cluster in FNR. The tridimensional structure of an FNR has not yet been determined. A very recent study combining Raman and UV-visible absorption/circular dichroism (CD) spectroscopies and mass spectrometry (7) indicates that Fe<sub>4</sub>S<sub>4</sub>-to-Fe<sub>2</sub>S<sub>2</sub> cluster degradation does not proceed with concomitant release of one Fe<sup>2+</sup>, one Fe<sup>3+</sup>, and two S<sup>2-</sup> ions as was previously reported (8). Instead, that study argues that cluster conversion leads to an Fe<sub>2</sub>S<sub>2</sub> cluster with retention of two sulfides as cysteine persulfide ligands according to the following reactions (7):



O<sub>2</sub>-induced Fe<sub>4</sub>S<sub>4</sub> cluster degradation leading to Fe<sub>2</sub>S<sub>2</sub> cluster with cysteine persulfide ligands does not seem to be restricted to FNR and is likely to be a general process, as it is also reported for other Fe<sub>4</sub>S<sub>4</sub> cluster-containing proteins such as biotin synthase and aconitase (7, 9). Spectroscopic inference of species containing Fe<sub>2</sub>S<sub>2</sub> cluster forms with only one or two cysteine persulfide ligands suggests successive oxidations but the mechanism involved is not fully understood. Here, we have used X-ray structures of HydE from *Thermotoga maritima* (*Tm*) as templates to shed light on the process of O<sub>2</sub>-induced Fe<sub>4</sub>S<sub>4</sub> cluster disassembly in proteins. Our structural results provide direct evidence that the O<sub>2</sub>-induced oxidation of an Fe<sub>4</sub>S<sub>4</sub> cluster can proceed via cysteine persulfide production and thus establishes a sulfur-centered mechanism for cluster degradation as a plausible biological process. *TmHydE* is one of the three essential FeFe-hydrogenase active site maturases (10, 11) which belongs to the radical *S*-adenosyl-L-methionine (SAM) protein superfamily (12, 13). Besides the conserved Fe<sub>4</sub>S<sub>4</sub> cluster of radical SAM proteins, *TmHydE* has an additional, more exposed cluster-binding site defined by semi-conserved Cys311, Cys319, and Cys322, in a Cx<sub>7</sub>Cx<sub>2</sub>C motif. Cys311 is located in a loop connecting β-strand 8 to α-helix 8 of the (β<sub>α</sub>)<sub>8</sub> triose phosphate isomerase barrel of the enzyme, and Cys319 and Cys322 are located on the same side of α-helix 8 (11) (Fig. 1). HydEs are divided into two groups depending on whether the Cys ligands of the additional FeS cluster are present, but it is known that this cluster is not important for function (11). In the first *TmHydE* X-ray structure we solved, the additional cluster, located at the protein surface, was an Fe<sub>2</sub>S<sub>2</sub> rhomb (11). Conversely, solution studies showed that after in vitro reconstitution *TmHydE* binds an Fe<sub>4</sub>S<sub>4</sub> cluster at the additional site (14). Here, we report eight different structures of *TmHydE* that contain dissimilar species at their additional cluster site. These species are in agreement with several intermediate states, either postulated or previously described, in the Fe<sub>4</sub>S<sub>4</sub> cluster degradation

Author contributions: Y.N. and J.C.F.-C. designed research; Y.N., R.R., L.M., and J.C.F.-C. performed research; Y.N., R.R., L.M., and J.C.F.-C. analyzed data; and Y.N. and J.C.F.-C. wrote the paper.

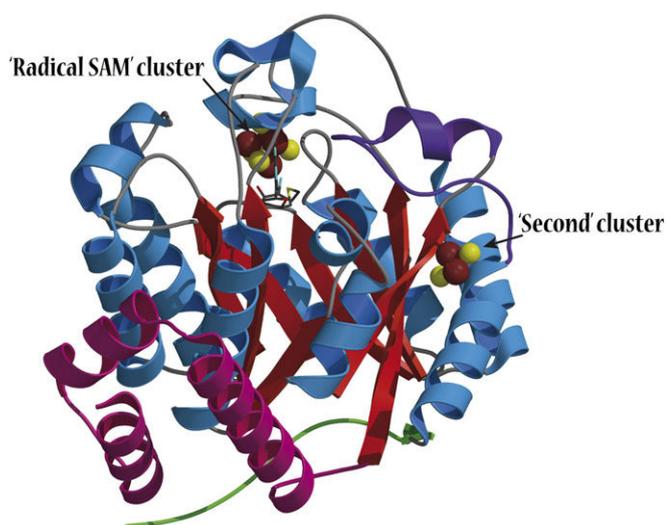
The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, [www.pdb.org](http://www.pdb.org) (PDB ID codes 3CIW, 3IIX, 4JXC, 4JY8, 4JY9, 4JYD, 4JYE and 4JYF).

<sup>1</sup>To whom correspondence may be addressed. E-mail: yvain.nicolet@ibs.fr or juan-carlos.fontecilla@ibs.fr.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1302388110/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1302388110/-DCSupplemental).



**Fig. 1.** Structure of the HydE from *T. maritima* (PDB code 3CIW). The SAM-dependent  $\text{Fe}_4\text{S}_4$  cluster is located at the top of the  $(\beta\alpha)_8$  barrel (color codes are S in yellow, iron in brown,  $\beta$ -strands in red, and  $\alpha$ -helices in blue). The SAM-independent additional  $\text{Fe}_2\text{S}_2$  cluster is located on a side of the barrel at the surface of the protein. The loop containing the cysteine ligand Cys311 is depicted in purple. The N-terminal three helices and the C-terminal stretch are depicted in pink and green, respectively.

pathway occurring, for instance, in FNR. These observed intermediate states in our structures suggest a mechanism for  $\text{Fe}_4\text{S}_4$  cluster degradation upon low-level oxygen exposure not restricted to FNR, but more general than previously thought.

## Results

As mentioned above, *in vitro* anaerobically reconstituted *TmHydE* has been reported to contain an  $\text{Fe}_4\text{S}_4$  center at its additional cluster-binding site in solution (14). However, when crystallized without DTT, the additional site coordinates an  $\text{Fe}_2\text{S}_2$  cluster bound by the semiconserved Arg279, Cys311, and persulfidated forms of Cys319 and Cys322 (species **3** in Fig. 2 and Fig. S1). The N $\eta$ 2 atom of Arg279 interacts directly with one of the iron ions ( $d = 2.3 \text{ \AA}$ ), whereas N $\eta$ 1 makes a hydrogen bond with one of the sulfide ions ( $d = 3.2 \text{ \AA}$ ). Freshly reconstituted samples normally produce crystals of this form after 1 wk in the glove box. Persulfide formation must result from either traces of  $\text{O}_2$  present in the glove box or transient exposure to air during the transfer of the frozen protein sample to it, because when intentionally oxygenated buffers are used these crystals appear more rapidly. Although in these experiments it is not possible to quantify the amount of  $\text{O}_2$  that reached our crystals, it must be extremely low as indicated by the high stability of the SAM-dependent  $\text{Fe}_4\text{S}_4$  cluster under the same conditions. Indeed, when SAM is substituted by 5'-deoxyadenosine, this cluster is slowly converted to an  $\text{Fe}_3\text{S}_4$  center in weeks without any further decay (Fig. S2). In the meantime, the additional cluster is rapidly converted into species **3**. The latter is very stable and only decays after several months, leading to **7** (Fig. 2).

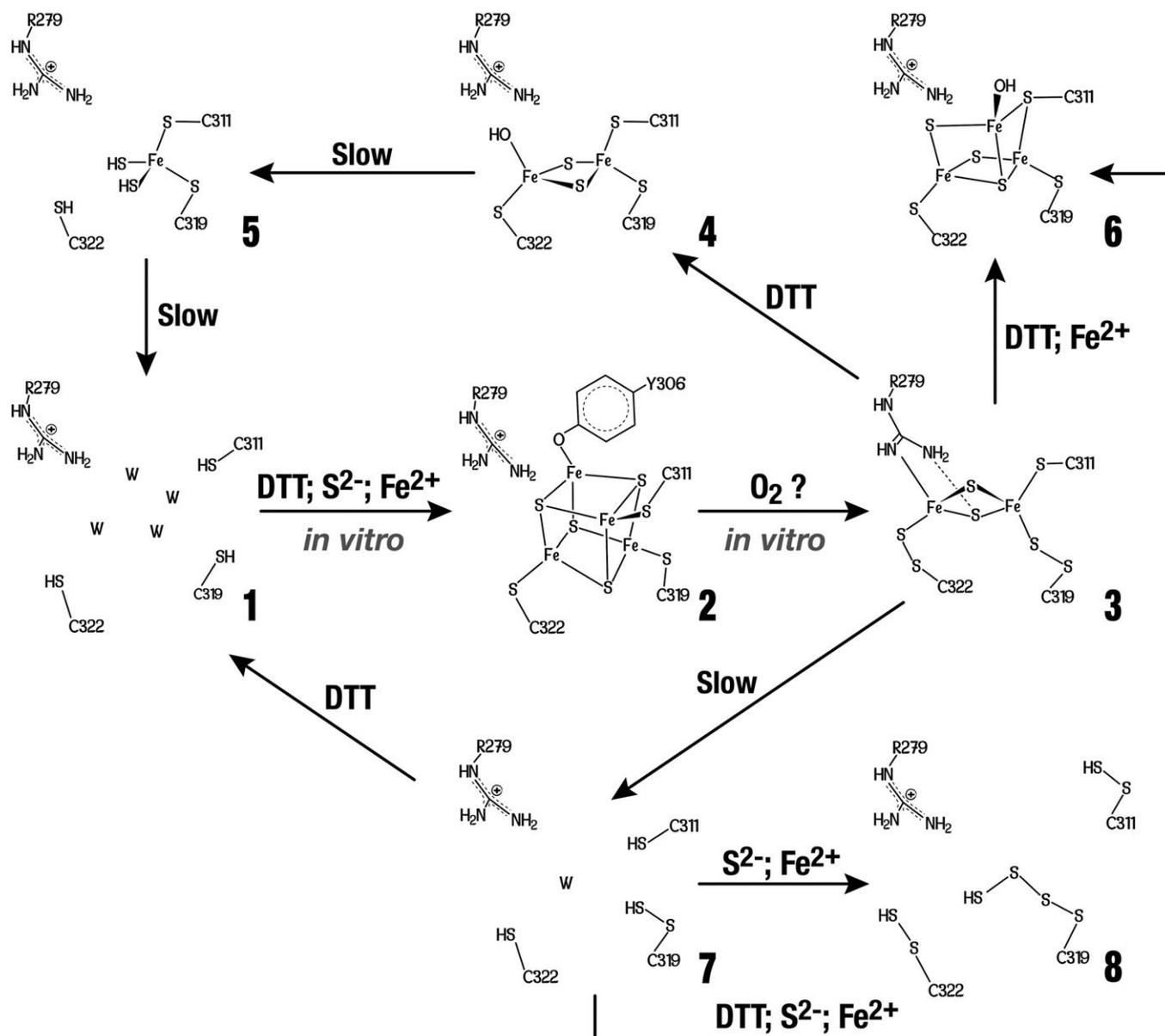
When crystals containing **3** are treated with DTT before flash cooling, a different  $\text{Fe}_2\text{S}_2$  cluster, coordinated by Cys311, Cys319, and Cys322, is observed (**4** in Fig. 2). There are no major structural changes at the protein level between **3** and **4** and the DTT treatment only causes the cluster to slightly move as a rigid body in its binding site. Conversely, the electron density indicates that the Fe fourth ligand may be  $\text{OH}^-$ , as Arg279 is clearly too far to be a ligand. We cannot exclude a mixture of  $\text{OH}^-$  and  $\text{SH}^-$  at that position because our structures unambiguously show that sulfide ions are released from the cysteine persulfide ligands upon DTT treatment. X-ray structures indicate that this cluster form

decays within weeks in the glove box. This process, which starts with the loss of one iron atom, is followed by the dissociation of the sulfide ions and finally results in the vacant-site species **1** (Fig. 2). Species **5** represents an intermediate form in the degradation of **4** (Fig. 2 and Fig. S1). X-ray structures from crystals flash-cooled at different times after DTT addition show that the slow disassembly of the additional cluster starts with the most solvent-exposed iron ion followed by the sulfide ions and, finally, by the second iron ion (**5** to **1** in Fig. 2). The slow decomposition of the  $\text{Fe}_2\text{S}_2$  form is likely to be a nonredox process and may result from slow ligand exchange with solvent.

When crystals containing **7** are treated with ferrous ions and sulfide, no iron bound to the protein is observed but Cys311 and Cys22 of the additional cluster site are modified as persulfides and Cys319 as polysulfide (**8** in Fig. 2 and Fig. S1). Treating similar crystals with ferrous ions, sulfide, and DTT leads to another unusual species (**6** in Fig. 2 and Fig. S1) that can be described as an  $\text{Fe}_2\text{S}_2$  cluster onto which an Fe-S unit is attached. This cluster resembles a standard  $\text{Fe}_3\text{S}_4$  center except that Cys311 bridges two iron ions. Modeling indicates that the structural changes required to allow Cys311 to become a classic terminal  $\text{Fe}_4\text{S}_4$  cluster ligand are prevented by crystal packing contacts (not shown). This problem was circumvented by the growth of a different crystal form of anaerobically purified *TmHydE*, treated with DTT, sulfide, and ferrous or ferric ions (Materials and Methods, Table S1). This form that was obtained by increasing the NaCl concentration in the crystallization solution contains a regular additional  $\text{Fe}_4\text{S}_4$  cluster, as previously described in solution (14), with Tyr306 as its fourth ligand (**2** in Fig. 2 and Fig. S1). Fig. 3 shows that shifts of over  $6.0 \text{ \AA}$  for the C $\alpha$  of Cys311 and  $12.4 \text{ \AA}$  for the C $\alpha$  of Tyr-306 are involved in the transition between the  $\text{Fe}_2\text{S}_2$ -containing and the  $\text{Fe}_4\text{S}_4$ -containing forms. This change requires a complete rearrangement of the loop that connects strand  $\beta_8$  to helix  $\alpha_8$  and contains Cys311 (Fig. 3). Thus, the  $\text{Fe}_4\text{S}_4$ -to- $\text{Fe}_2\text{S}_2$  cluster conversion (**2** to **3**) is a conformationally drastic one. In the reverse direction, addition of DTT and  $\text{Fe}^{2+}$  (but not of  $\text{S}^{2-}$ ) converts **3** into **6** (Fig. 2 and Fig. S1). The peculiar  $\text{Fe}_3\text{S}_3$  cluster in **6** probably results from the aborted assembly of an  $\text{Fe}_4\text{S}_4$  cluster hindered, as mentioned above, by crystal packing interactions. A major conclusion from this observation is that there is no need to add sulfide ions to obtain the observed  $\text{Fe}_3\text{S}_3$  cluster because they are already present as part of persulfide ligands in **3**.

## Discussion

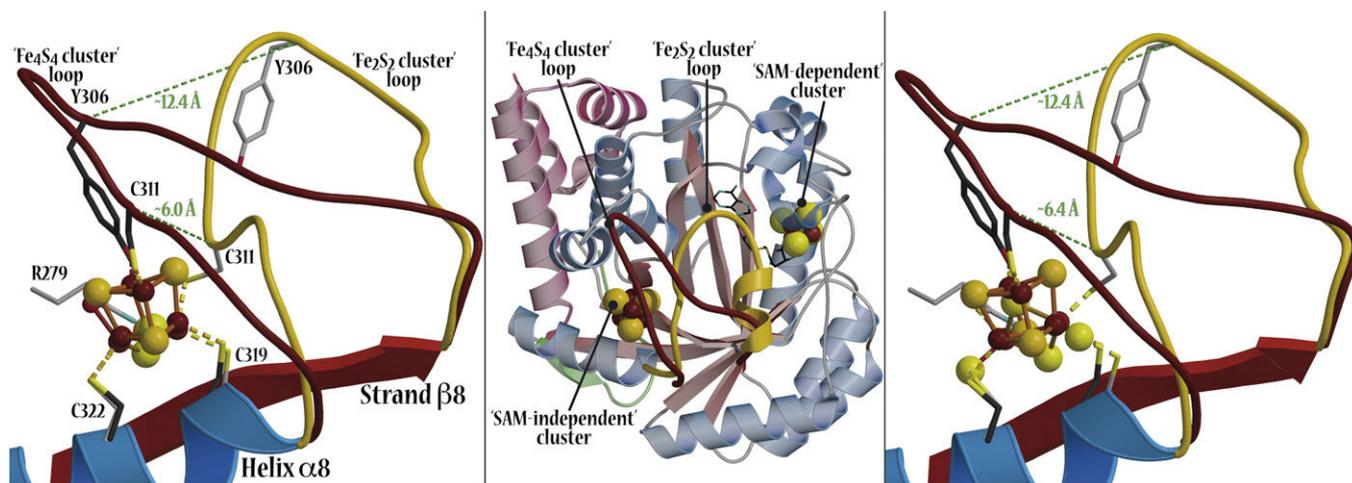
Many of our X-ray crystal structures are in full agreement with the very recent Raman resonance (RR) spectroscopic study of Zhang et al. and consequently may be relevant to FeS cluster interconversion in FNR (7). As opposed to the additional cluster, the SAM-related  $\text{Fe}_4\text{S}_4$  cluster of *TmHydE* was stable for a long time inside the glove box. This is most likely explained by the protection afforded to the latter by the  $\text{C}_x\text{C}_x\text{C}$  motif-containing loop. Conversely, the additional cluster is fully exposed to solvent and hence to traces of  $\text{O}_2$  either in the glove box or during sample transfer to it. This unanticipated situation is actually advantageous because it mimics what should happen for example to *E. coli* in the rectal segment of the intestinal tract when oxygen levels begin to rise and FNR is inactivated, switching off anaerobic metabolism. The very slow reaction of the additional FeS cluster with traces of  $\text{O}_2$  allowed us to trap intermediate species in our protein template that would have been difficult to obtain otherwise. Thus, the  $\text{Fe}_2\text{S}_2$  cluster with two cysteine persulfide ligands proposed from RR spectroscopic data in FNR is equivalent to **3** (Fig. 2 and Fig. S1). In addition, when treated with ferrous ions and DTT, this FNR  $\text{Fe}_2\text{S}_2$  species is readily converted to a regular  $\text{Fe}_4\text{S}_4$  cluster (**7**), in a process similar to the **3** to **6** transition in Fig. 2. According to reaction **1**, the first step of  $\text{O}_2$ -induced  $\text{Fe}_4\text{S}_4$  cluster degradation corresponds to iron oxidation, leading to an intermediate  $\text{Fe}_3\text{S}_4\text{]}^+$  cluster with concomitant release of one ferrous ion and a superoxide



**Fig. 2.** Postulated degradation and assembly of an FeS cluster based on our crystal structures. Unless specified, all reactions were performed in our crystals. W in species 1 and 7 indicates solvent molecules. X-ray structures of *TmHydE* containing species 1 to 8 were determined at 1.35, 2.9, 1.45, 1.6, 1.71, 1.7, 1.45, and 1.25 Å resolution, respectively (Table S2).

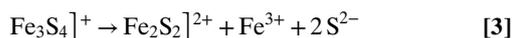
radical species (15). Conversely, species 3 with cysteine persulfide ligands demonstrates that subsequent oxidative damage proceeds preferentially via sulfur rather than iron oxidation. Furthermore, the intermediate species with only one persulfide ligand observed in FNR indicates that successive oxidation steps are possible (7). At variance with what is proposed by Zhang et al. for FNR (reaction 2), instead of the addition of a new oxygen molecule, we advance that the next step corresponds to the reaction of the superoxide anion resulting from reaction 1 with one of the cysteine ligands forming a cysteine persulfenate. Such species that would be transient in our case has been observed in the crystal structure of cysteine dioxygenase (16). The expected conversion of the cysteine persulfenate to sulfinate does not occur, maybe due to the sulfide-rich environment of the adjacent iron ion. Instead, this species is converted to cysteine persulfide with production of one hydrogen peroxide molecule and release of a second ferrous ion, as previously observed in FNR (17, 18). Reaction of a superoxide ion, instead of molecular oxygen,

with the  $\text{Fe}_3\text{S}_4]^+$  intermediate would be consistent with the observation that in FNR only 0.37 superoxide ions per damaged  $\text{Fe}_4\text{S}_4$  cluster were detected from reaction 1 (15). Substantial amounts of hydrogen peroxide were also produced (15), which could subsequently oxidize a second cysteine ligand to cysteine persulfenate. Reaction of this species with the last available hydrogen sulfide ion would generate a second persulfide, thus leading to a form equivalent to 3. Such mechanism would explain why only substoichiometric amounts of  $\text{H}_2\text{O}_2$  are produced by FNR and is consistent with the recent proposition that some FNR molecules have only one cysteine persulfide ligand (7, 15, 17). Furthermore, our mechanism is supported by the observation that both  $\text{O}_2$  and  $\text{H}_2\text{O}_2$  can provoke an  $\text{Fe}_4\text{S}_4$ -to- $\text{Fe}_2\text{S}_2$  cluster transition in FNR. However, the resulting  $\text{Fe}_2\text{S}_2$  cluster has a different circular dichroism spectrum depending on the oxidant used (17). This difference is probably due to the expected number of cysteine persulfide ligands produced: two when using  $\text{O}_2$  and one when using  $\text{H}_2\text{O}_2$ . The observation that in FNR one ferric iron is



**Fig. 3.** Structural differences between SAM-independent Fe<sub>4</sub>S<sub>4</sub> and Fe<sub>2</sub>S<sub>2</sub> cluster-containing *TmHydE*. (Left) Superposition of the additional cluster sites containing species 2 and 4 (Fig. 2). Loop containing Cys311 is depicted in brown for the Fe<sub>4</sub>S<sub>4</sub>-cluster structure and in gold for the Fe<sub>2</sub>S<sub>2</sub>-cluster structure. Dashed line indicates the shift of the C $\alpha$  atom of Cys311 and Tyr-306 between the two structures. (Center) View of the whole structure. Loop containing Cys311 covers the barrel cavity when the protein contains an Fe<sub>2</sub>S<sub>2</sub> cluster (gold) and moves away from it when the protein contains an Fe<sub>4</sub>S<sub>4</sub> cluster (brown). (Right) Superposition between species 2 and 3. Only minor differences are observed in the conformation of the Cys311-containing loop between species 3 and 4, indicating that conversion of species 2 into 3 or 4 should induce similar conformational changes.

released during the Fe<sub>3</sub>S<sub>4</sub><sup>+</sup> to Fe<sub>2</sub>S<sub>2</sub><sup>2+</sup> cluster conversion may be explained by the oxygen-independent reaction 3 (8):



This nonredox reaction may compete with sulfide oxidation and could be favored in vitro by the presence of strong ferrous or ferric ion chelators (8), suggesting an alternative nonbiological pathway for oxygen-induced FeS cluster degradation.

Addition of a reductant and extra ferrous ions converts species 3 back to 2 (or 6 in our crystals). Both our structural results and those reported by Zhang et al. (7) using RR spectroscopy raise the possibility of a persulfide-mediated Fe<sub>4</sub>S<sub>4</sub> cluster repair mechanism that, avoiding de novo sulfide synthesis by the iron-sulfur cluster (ISC) system, would be less costly to the cell. Cysteine persulfide ligands as an O<sub>2</sub>-induced degradation product of Fe<sub>4</sub>S<sub>4</sub> clusters appear to be generally distributed, being also proposed in biotin synthase and other radical SAM enzymes (7).

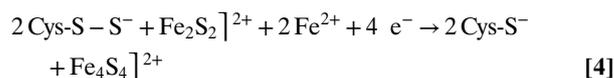
Our observation that species 4 slowly decomposes into 5 and then 1, most probably via a nonredox process (Fig. 2 and Fig. S1), agrees well with the previous observation that iron depletion itself can mediate the holo to apo transition in cytoplasmic aconitase (19). Our study also shows that the transition from the Fe<sub>4</sub>S<sub>4</sub> cube to the Fe<sub>2</sub>S<sub>2</sub> rhomb would require a major reorganization of the coordination site in *TmHydE*. The loop rearrangement resulting from this reorganization implies that an Fe<sub>4</sub>S<sub>4</sub>-to-Fe<sub>2</sub>S<sub>2</sub> cluster conversion will necessarily trigger significant protein structural changes. Equivalent changes may be responsible for O<sub>2</sub>-induced dimer dissociation in FNR (20). Similarly, the transition between holo and apo cytoplasmic aconitase has been structurally characterized and shown to involve large structural rearrangements (21, 22). Finally, an early study (9) showed that controlled oxidation of apo-aconitase resulted in the formation of both persulfides and polysulfides, a situation illustrated here by 7 and 8 (Fig. 2 and Fig. S1).

Our results may also shed light on the sequence of events during in vivo ISC-dependent FeS cluster assembly in the complex formed by the L-cysteine desulfurase IscS and the scaffold IscU. The order of arrival of iron and sulfide to the IscU assembly site has not been clearly established. This scaffold protein can bind iron, but only at lower-than-physiological temperatures (23, 24). Conversely, several

groups have reported that S<sup>0</sup>, transferred from the cysteine desulfurase, can form a Cys-S-S species in the scaffold protein, in the absence of iron (25, 26). However, it has been reported that the persulfided scaffold does not bind iron (24). The latter observation applies also to the oxidized *TmHydE* 7 species where Cys311 and persulfided Cys319 form a long S-S bond (Fig. S1). Addition of sulfide and Fe<sup>2+</sup> to 7 produces 8, a species that contains persulfided and polysulfided cysteines but does not bind iron (Fig. 2 and Fig. S1). This observation implies that free Fe<sup>2+</sup> is not able to reduce the S-S bond of the modified Cys319 in 7. This result favors in turn a concerted mechanism for Fe-S cluster assembly by ISC, where two ferrous ions bind simultaneously the first persulfide S<sup>0</sup> from the active Cys residue of IscS at the assembly site of IscU and reduce it to S<sup>2-</sup>. We have already suggested such mechanism based on the *Archeoglobus fulgidus* (IscS-IscU)<sub>2</sub> complex structure (27).

### Conclusion

The various FeS cluster structures found in our *TmHydE* crystals can be ordered in a coherent sequence that provides a solid base for understanding cluster degradation and assembly (1 through 5 in Fig. 2 and Fig. S1). The plasticity of the additional FeS cluster coordination site has proved to be well suited for a systematic characterization of different protein-bound FeS species using X-ray crystallography. Taken together, our results and the RR spectroscopic studies of Zhang et al. (7) provide a convincing structural and functional interpretation of the Fe<sub>4</sub>S<sub>4</sub>-to-Fe<sub>2</sub>S<sub>2</sub> cluster conversion upon O<sub>2</sub> exposure in FeS cluster-containing proteins. The main proposition derived from our work is the existence of FeS cluster direct repair mechanisms involving the reduction of persulfide ligands in an Fe<sub>2</sub>S<sub>2</sub> cluster similar to the 3 to 6 transition we have observed in our crystals. Without the packing constraints in the high-resolution crystal structure, 3 should have readily converted to the Fe<sub>4</sub>S<sub>4</sub> cluster-containing species 2 observed in the high-salt crystal form (Figs. 2 and 3) following reaction 4:



The requirement of very significant conformational changes in going from 3 to 2 clearly illustrates the extensive structural

rearrangement expected in the dimer–monomer transition of FNR. Finally, the resistance of the 7 species to reduction by free Fe<sup>2+</sup> favors the concerted arrival of iron and sulfur to the scaffold protein during ISC-dependent Fe<sub>2</sub>S<sub>2</sub> cluster biogenesis.

## Materials and Methods

**Protein Purification and Crystallization.** HydE was purified either aerobically, as previously described (14), or the purification was carried out anaerobically following the same protocol. Crystals were obtained as described previously (11). All of the crystallization experiments were carried out in glove boxes with oxygen concentrations of about 2–5 ppm. Great care was taken to remove oxygen from the various solutions by sparging them with the glove box atmosphere, except when aerated solutions were intentionally used. The sample leading to the Fe<sub>4</sub>S<sub>4</sub> cluster-containing structure at the additional site was obtained using a different protocol. In this case, a streptavidin tag was added at the N terminus of the amino acid sequence between Met1 and Thr2 (new N-terminus sequence: MWSPQFEKAST-). The Strep-tag-containing HydE construct was obtained following the QuikChange site-directed mutagenesis kit strategy (Stratagene) using plasmid pTmHydE (Rubach et al., ref.14), Phusion polymerase, and the following primers: Forward: 5'-CACCCGAGTTCGAAAAGCAAGCACCGGTAGAGAAATCTGG-AAAAA-3'; Reverse: 5'-CTTTTCGAAGTCCGGGTGGCTCCACATGGTATATCTC-TTCTTAAAGTTAAACAAA-3'. This produced the pStrepTmHydE construct. The correctness of the cloned DNA sequence was confirmed by sequencing the entire gene. Protein expression was carried out following the protocol previously described for the untagged protein (14). After cell disruption in the glove box, the crude extract was treated directly with 1 mM FeCl<sub>3</sub>, 5 mM DTT, and 1 mM Na<sub>2</sub>S during 60 min, with mild stirring, to reconstitute the FeS clusters. The crude extract was subsequently cleared by ultracentrifugation at 165,000 × g during 30 min at 4 °C. The clear supernatant was loaded onto a streptavidin-agarose column buffered with 100 mM Tris pH 8.0; 150 mM NaCl; and 5 mM DTT. After extensive washing of the column, the protein was eluted by adding 2.5 mM desthiobiotin, which was subsequently removed using

a GE Healthcare HiPrep 26/10 desalting column equilibrated with 100 mM Tris pH 8.0; 150 mM NaCl. HydE was then concentrated to 16 mg/mL, as determined by the Rose Bengal method (28), using an Amicon concentrator with a 10-kDa cutoff membrane. This HydE sample contained 8 iron atoms per protein molecule as determined by the method of Fish (29). Crystallization was performed as previously described (11), except that the protein sample contained 100 mM Tris pH 8.0 and 150 mM NaCl instead of 50 mM Tris pH 8.0 solution without NaCl. Large hexagonal brownish crystals were produced after 1 d. In the absence of NaCl, we obtained orthorhombic crystals like those previously reported (11). This shows that salt, and not the presence of the Strep-tag, is responsible for the change in space group. Table S1 summarizes the crystallization and cryoprotecting conditions used for each X-ray structure presented.

**Data Collection and X-Ray Structure Refinement.** Data were collected at the European Synchrotron Radiation Facility and Swiss Light Source and processed with the XDS package (30). Model refinements were carried out using REFMAC5 (31) from the Collaborative Computational Project No. 4 package (32). The set of flagged reflections for R<sub>free</sub> calculations was the same used previously (11) although the initial set was extended to 1.25 Å resolution for the datasets diffracting that far. All of the data processing and refinement statistics are presented in Table S2. Manual model building was performed using COOT (33). Identification of Fe versus S atoms was based on the observed anomalous scattering electron density peaks for the corresponding atoms.

**ACKNOWLEDGMENTS.** We thank Prof. Michael Johnson and Dr. Isabelle Artaud for fruitful discussions and the Commissariat à l'Énergie Atomique (CEA), the Centre National de la Recherche Scientifique, and the Université Grenoble-Alpes for institutional support. We also thank the staff of the European Synchrotron Radiation Facility beamlines ID14, ID23, and ID29 (France) and Swiss Light Source PXI beamline (Switzerland) for help with data collection. R.R. is the recipient of an "Irtelis" Fellowship from the CEA. This work was financed by Contract ANR-08-BLAN-0224-01 from the French Agence Nationale pour la Recherche.

- Johnson DC, Dean DR, Smith AD, Johnson MK (2005) Structure, function, and formation of biological iron-sulfur clusters. *Annu Rev Biochem* 74:247–281.
- Drennan CL, Peters JW (2003) Surprising cofactors in metalloenzymes. *Curr Opin Struct Biol* 13(2):220–226.
- Crack JC, Green J, Thomson AJ, Le Brun NE (2012) Iron-sulfur cluster sensor-regulators. *Curr Opin Chem Biol* 16(1–2):35–44.
- Fleischhacker AS, Kiley PJ (2011) Iron-containing transcription factors and their roles as sensors. *Curr Opin Chem Biol* 15(2):335–341.
- Alexeeva S, Hellingwerf KJ, Teixeira de Mattos MJ (2002) Quantitative assessment of oxygen availability: Perceived aerobiosis and its effect on flux distribution in the respiratory chain of *Escherichia coli*. *J Bacteriol* 184(5):1402–1406.
- Achebach S, Selmer T, Udden G (2005) Properties and significance of apoFNR as a second form of air-inactivated [4Fe-4S]<sub>2</sub> FNR of *Escherichia coli*. *FEBS J* 272(16):4260–4269.
- Zhang B, et al. (2012) Reversible cycling between cysteine persulfide-ligated [2Fe-2S] and cysteine-ligated [4Fe-4S] clusters in the FNR regulatory protein. *Proc Natl Acad Sci USA* 109(39):15734–15739.
- Crack JC, et al. (2008) Influence of the environment on the [4Fe-4S]<sub>2</sub> to [2Fe-2S]<sub>2</sub> cluster switch in the transcriptional regulator FNR. *J Am Chem Soc* 130(5):1749–1758.
- Kennedy MC, Beinert H (1988) The state of cluster SH and S<sup>2-</sup> of aconitase during cluster interconversions and removal. A convenient preparation of apoenzyme. *J Biol Chem* 263(17):8194–8198.
- Fontecilla-Camps JC, Volbeda A, Cavazza C, Nicolet Y (2007) Structure/function relationships of [NiFe]- and [FeFe]-hydrogenases. *Chem Rev* 107(10):4273–4303.
- Nicolet Y, et al. (2008) X-ray structure of the [FeFe]-hydrogenase maturase HydE from *Thermotoga maritima*. *J Biol Chem* 283(27):18861–18872.
- Nicolet Y, Drennan CL (2004) AdoMet radical proteins—from structure to evolution—alignment of divergent protein sequences reveals strong secondary structure element conservation. *Nucleic Acids Res* 32(13):4015–4025.
- Sofia HJ, Chen G, Hetzler BG, Reyes-Spindola JF, Miller NE (2001) Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: Functional characterization using new analysis and information visualization methods. *Nucleic Acids Res* 29(5):1097–1106.
- Rubach JK, Brazzolotto X, Gaillard J, Fontecave M (2005) Biochemical characterization of the HydE and HydG iron-only hydrogenase maturation enzymes from *Thermotoga maritima*. *FEBS Lett* 579(22):5055–5060.
- Crack JC, Green J, Cheesman MR, Le Brun NE, Thomson AJ (2007) Superoxide-mediated amplification of the oxygen-induced switch from [4Fe-4S] to [2Fe-2S] clusters in the transcriptional regulator FNR. *Proc Natl Acad Sci USA* 104(7):2092–2097.
- Simmons CR, et al. (2008) A putative Fe<sup>2+</sup>-bound persulfenate intermediate in cysteine dioxygenase. *Biochemistry* 47(44):11390–11392.
- Crack J, Green J, Thomson AJ (2004) Mechanism of oxygen sensing by the bacterial transcription factor fumarate-nitrate reduction (FNR). *J Biol Chem* 279(10):9278–9286.
- Sutton VR, Mettert EL, Beinert H, Kiley PJ (2004) Kinetic analysis of the oxidative conversion of the [4Fe-4S]<sub>2</sub><sup>2+</sup> cluster of FNR to a [2Fe-2S]<sub>2</sub><sup>2+</sup> Cluster. *J Bacteriol* 186(23):8018–8025.
- Varghese S, Tang Y, Imlay JA (2003) Contrasting sensitivities of *Escherichia coli* aconitases A and B to oxidation and iron depletion. *J Bacteriol* 185(1):221–230.
- Lazzerza BA, Beinert H, Khoroshilova N, Kennedy MC, Kiley PJ (1996) DNA binding and dimerization of the Fe-S-containing FNR protein from *Escherichia coli* are regulated by oxygen. *J Biol Chem* 271(5):2762–2768.
- Dupuy J, et al. (2006) Crystal structure of human iron regulatory protein 1 as cytosolic aconitase. *Structure* 14(1):129–139.
- Walden WE, et al. (2006) Structure of dual function iron regulatory protein 1 complexed with ferritin IRE-RNA. *Science* 314(5807):1903–1908.
- Agar JN, et al. (2000) Modular organization and identification of a mononuclear iron-binding site within the NifU protein. *J Biol Inorg Chem* 5(2):167–177.
- Nuth M, Yoon T, Cowan JA (2002) Iron-sulfur cluster biosynthesis: Characterization of iron nucleation sites for assembly of the [2Fe-2S]<sub>2</sub><sup>2+</sup> cluster core in IscU proteins. *J Am Chem Soc* 124(30):8774–8775.
- Smith AD, et al. (2001) Sulfur transfer from IscS to IscU: The first step in iron-sulfur cluster biosynthesis. *J Am Chem Soc* 123(44):11103–11104.
- Urbina HD, Silberg JJ, Hoff KG, Vickery LE (2001) Transfer of sulfur from IscS to IscU during FeS cluster assembly. *J Biol Chem* 276(48):44521–44526.
- Marinoni EN, et al. (2012) (IscS-IscU)<sub>2</sub> complex structures provide insights into FeS<sub>2</sub> biogenesis and transfer. *Angew Chem Int Ed Engl* 51(22):5439–5442.
- Elliott JI, Brewer JM (1978) The inactivation of yeast enolase by 2,3-butanedione. *Arch Biochem Biophys* 190(1):351–357.
- Fish WW (1988) Rapid colorimetric micromethod for the quantitation of complexed iron in biological samples. *Methods Enzymol* 158:357–364.
- Kabsch W (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J Appl Cryst* 26:795–800.
- Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53(Pt 3):240–255.
- Winn MD, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67:235–242.
- Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.

# Supporting Information

Nicolet et al. 10.1073/pnas.1302388110

## Conserved Radical *S*-Adenosyl-L-Methionine FeS Cluster Stability in the Glove Box

The structure presented in Fig. S2 was obtained using 5'-deoxy-adenosine instead of *S*-adenosyl-L-homocysteine (SAM) after in vitro FeS cluster reconstitution. No DTT was added either for crystallization or for crystal flash cooling. When crystals are frozen within days after crystallization, the conserved cluster consists of a mixture of Fe<sub>4</sub>S<sub>4</sub> and Fe<sub>3</sub>S<sub>4</sub> species as indicated by the progressive disappearance of the electron density corresponding to the unique iron atom (not shown). If crystals are

frozen within weeks only the Fe<sub>3</sub>S<sub>4</sub> cluster species is present. This indicates that in our crystals this cluster is relatively stable in our anaerobic chamber, both under the Fe<sub>4</sub>S<sub>4</sub> and Fe<sub>3</sub>S<sub>4</sub> forms. This result shows that the radical SAM cluster and the second cluster of HydE have very different stabilities when exposed to low-level oxygen. In this crystal, the second cluster site is occupied by a mixture of states including species 3 (major) and 4 (minor). The relatively high stability of the Fe<sub>3</sub>S<sub>4</sub> form of the radical SAM cluster most likely results from the low level of oxygen in the glove box.

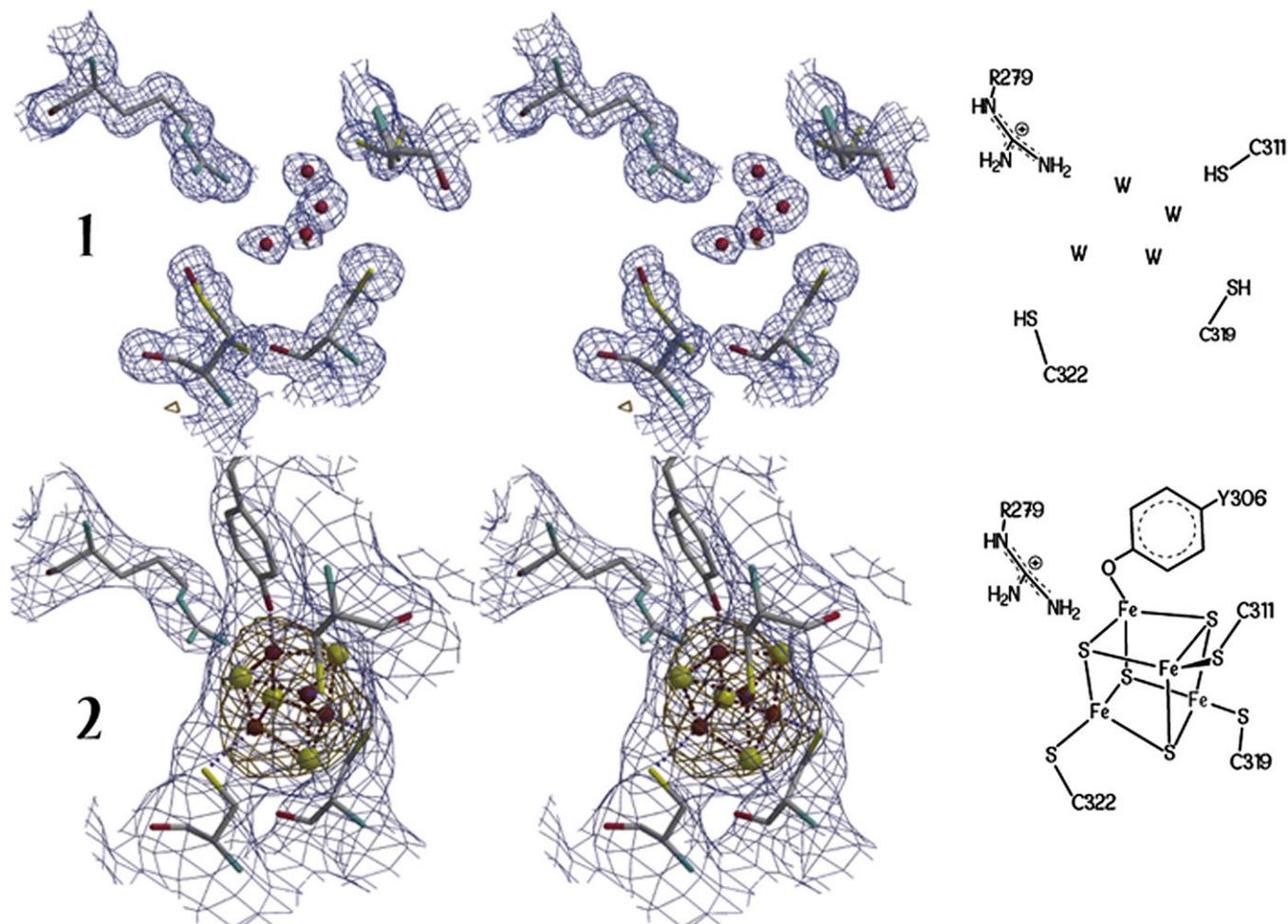
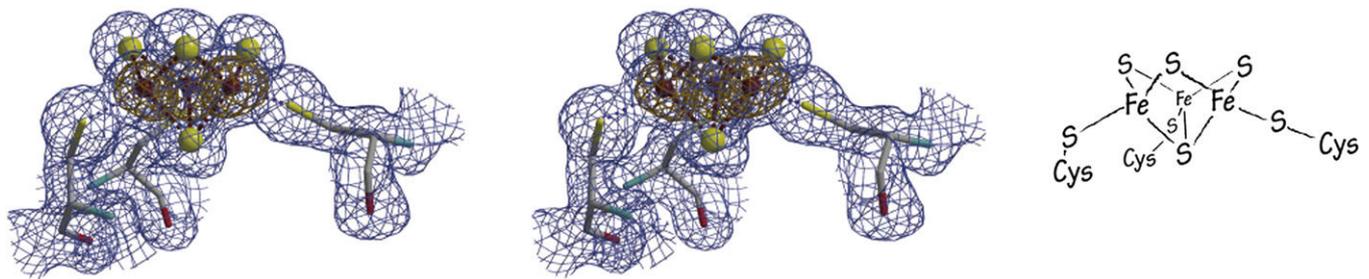


Fig. S1. (Continued)







**Fig. S2.** Radical SAM cluster of HydE in its  $\text{Fe}_3\text{S}_4$  form at 1.7 Å resolution. Stereoview of the atomic model and corresponding ( $2F_o-F_c$ ) electron density maps (shown as a blue mesh and contoured at the  $1\sigma$  level). The anomalous difference ( $\Delta_{\text{anom}}$ ) electron density map (depicted as an orange mesh, contoured at the  $3.5\sigma$  level) shows the positions of the iron atoms.

**Table S1.** Summary of the crystallization conditions and crystal treatment before flash cooling

Sample preparation conditions	Species							
	1	2	3	4	5	6	7	8
Crystallization condition								
DTT 5 mM	Yes	—	—	—	Yes	Yes	—	—
NaCl	—	Yes	—	—	—	—	—	—
Cryo condition								
DTT 5 mM	Yes	—	—	Yes	Yes	Yes	—	—
$\text{Na}_2\text{S}$ 1 mM	—	—	—	—	—	—	—	Yes
$\text{Fe}(\text{NH}_4)_2(\text{SO}_4)_2$ 1 mM	—	—	—	—	—	Yes	—	Yes
Further soak	—	—	—	—	NaBr 100 mM	—	—	—
Approximate time between crystallization and crystal flash cooling	2 mo	2 wk	2 wk	6 mo	1 mo	1 mo	4 mo	4 mo





