HAL
open science

Etude cristallographique des protéines NikA et NikR impliquées dans la transport du nickel chez Escherichia coli.Quant la structure de NikA met en évidence l'existence possible d'un nouveau métallophore.

Mickael V Cherrier

To cite this version:

Mickael V Cherrier. Etude cristallographique des protéines NikA et NikR impliquées dans la transport du nickel chez Escherichia coli.Quant la structure de NikA met en évidence l'existence possible d'un nouveau métallophore.. Biochimie [q-bio.BM]. Université Joseph-Fourier - Grenoble I, 2006. Français. NNT : . tel-00079162

HAL Id: tel-00079162

https://theses.hal.science/tel-00079162

Submitted on 9 Jun 2006

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.

THESE

Présentée par

Mickaël V. CHERRIER

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE JOSEPH FOURIER – GRENOBLE I

(arrêtés ministériels du 5 juillet 1984 et du 30 mars 1992)

Spécialité : Cristallographie et RMN biologiques

Etude cristallographique des protéines NikA et NikR impliquées dans le transport du nickel chez *Escherichia coli.*

Quand la structure de NikA met en évidence l'existence possible d'un nouveau métallophore.

Soutenue le 5 mai 2006 devant le jury composé de

Président du jury	**Marc FONTECAVE**
Rapporteurs	**Philippe DUMAS**
	Isabelle SCHALK
Examinateurs	**Marie-Andrée MANDRAND-BERTHELOT**
	Michael MARONEY
	Jean-Louis PIERRE
Directeur de thèse	**Juan Carlos FONTECILLA CAMPS**

Thèse préparée au Laboratoire de Cristallographie et Cristallogenèse des Protéines (LCCP)
de l'Institut de Biologie Structurale Jean-Pierre EBEL (IBS) CEA/CNRS/UJF

REMERCIEMENTS

Le travail que j'ai réalisé pendant cette thèse, et présenté dans ce manuscrit, n'aurait jamais pu exister sans l'aide et le soutien précieux de nombreuses personnes.

Je tiens de ce fait, à remercier en tout premier lieu mon directeur de thèse Juan Carlos Fontecilla Camps, qui m'a accueilli dans son laboratoire depuis la maîtrise. Merci pour les nombreuses discussions scientifiques, très fructueuses, que nous avons eues, ainsi que votre aide et votre confiance pendant ces années. J'espère que cela pourra continuer.

Je remercie les membres du jury, Marie – Andrée Mandrand – Berthelot, Isabelle Schalk, Philippe Dumas, Marc Fontecave, Michael Maroney, Jean – Louis Pierre, d'avoir pris le temps d'examiner ma thèse, et d'être venu y assister (pour certain de loin).

Je remercie Christine Cavazza et Lydie Martin de m'avoir formé aux techniques de microbiologie et biochimie, mais aussi pour leur disponibilité et la patience dont elles ont fait preuve envers moi.

Je remercie Carlos Contreras – Martel de m'avoir formé à la cristallographie, mais aussi pour son amitié.

Je remercie Lilian Jacquamet d'avoir pris le temps (ce qui ne fût pas facile pour lui au vu de son emploi du temps surchargé) de m'avoir expliqué et formé à la S.A.X., mais aussi pour l'accès privilégié à la ligne FIP dont j'ai bénéficié grâce à lui.

Je remercie Philippe Carpentier pour ses nombreux conseils et son aide concernant la collecte et le traitement des données de diffraction.

Je remercie Jérôme Dupuy et Jean – Baptiste Reiser qui ont été des compagnons de DEA et de thèse très agréables et précieux. Merci pour votre amitié.

Je remercie Cédric Bauvois (alias nonante), Franck Borel, Jérôme Dupuy, Lilian Jacqamet, et Xavier Vernede, pour leur amitié et leur soutien qui a été très important lors des moments les plus difficiles de ma thèse.

Je remercie Constance Bochot qui n'est pas au laboratoire depuis très longtemps mais avec qui j'ai déjà eu des discussions très intéressantes sur NikA.

Je remercie plus généralement tous les membres du laboratoire (Frédéric Garzoni, Yvain Nicolet, Antoine Royant…) qui m'ont accueilli, aidé et soutenu pendant les années que j'ai passé au LCCP.

Je remercie Marie – Andrée Mandrand – Berthelot et Agnès Rodrigue notre collaboration concernant NikR mais aussi l'aide précieuse qu'elles nous fournissent sur NikA.

Je remercie David Lemaire dont le travail en spectroscopie de masse fût très précieux, mais aussi pour tout le temps qu'il a bien voulu me consacrer pour m'expliquer cette technique.

Je remercie Stéphane Ménage pour les tentatives d'extraction du métallophore et de sa caractérisation par spectroscopie de masse.

Je remercie Jacques Gaillard pour les expériences de RPE que nous avons menées ensemble.

Je remercie Jean – Pierre Simorre pour les essais de caractérisation de la structure du métallophore que nous avons tentée par RMN.

Je remercie Dominique Madern pour les expériences de dichroïsme circulaire que nous avons menées ensemble sur la protéine NikA.

Je remercie Frédérique Vellieux d'avoir retrouvé les premières données de diffraction qui avaient été enregistrées en 1994 sur NikA.

Je remercie les membres des lignes de lumières de l'ESRF sur lesquelles j'ai pu réaliser les différentes collectes de données de diffraction que j'ai effectuées pendant ma thèse (BM30A – FIP ; ID14 – EH1-4 ; ID29), ainsi que celle qui m'a permis de collecter les données SAX analysées dans ce manuscrit (BM30B-FAME).

Je remercie Chantal Robesson pour sa disponibilité et son efficacité en tant que secrétaire.

Je remercie aussi le chien de Delphine Baud qui m'a fourni (malgré lui) un des poils de sa moustache (de grande qualité !) qui m'a permis de reproduire les cristaux de NikA par micro-ensemencement.

Je remercie ma mère de m'avoir permis de faire ces études grâce à son aide et son soutien précieux.

Enfin, je remercie ma femme Maud pour son amour, son soutien, mais aussi de la patience dont elle fait preuve lorsque je suis stressé (et je l'ai été). Sans elle cette période de ma vie aurait été beaucoup plus difficile à vivre.

A mon père qui m'a transmis la passion de la science…

SOMMAIRE

ET

LISTE DES FIGURES

SOMMAIRE

ABREVIATIONS **25**

INTRODUCTION **29**

1 GENERALITES SUR LE NICKEL 31

1.1 L'atome de nickel *31*

1.2 L'historique de la découverte du nickel *32*

1.3 La répartition naturelle du nickel *32*

1.3.1 Les minerais sulfurés 33

1.3.2 Les minerais oxydés 33

1.3.3 Le cas de la Nouvelle Calédonie 34

1.4 L'utilisation industrielle du nickel *35*

1.4.1 Les aciers inoxydables 35

1.4.2 Le nickelage 36

1.4.3 Les pièces de monnaie 36

1.4.4 D'autres exemples d'utilisation du nickel 36

1.5 La toxicité du nickel *36*

1.5.1 Les sources de contact avec le nickel et ses dérivés 36

1.5.2 Les réactions allergiques 37

1.5.3 Le nickel est carcinogène 37

2 LE NICKEL CHEZ LES MICROORGANISMES 38

2.1 Le rôle et l'utilisation du nickel chez les bactéries *38*

2.1.1 L'uréase 38

2.1.2 L'hydrogénase à NiFe 41

2.1.3 La monoxyde de carbone déshydrogénase (CODH) 45

2.1.4 L'acétyl-coenzyme A synthase (ACS) 45

2.1.5 La méthyl-coenzyme M réductase 46

2.1.6 La superoxyde dismutase Ni-dépendante (SOD) 48

2.1.7 La glyoxylase Ni-dépendante 48

2.1.8 L'aci-réductone dioxygénase 49

2.1.9 D'autres enzymes à activité nickel-dépendante 50

2.2 Le transport du nickel chez les bactéries *50*

2.2.1 L'import du nickel 50

2.2.1.1 Le passage à travers la membrane externe 50

2.2.1.2 Le passage à travers la membrane interne : le système du type ABC 52

2.2.1.2.1 Le système codé par l'opéron nikABCDE 53

2.2.1.2.2 La protéine NikA 55

2.2.1.3 Le passage à travers la membrane interne : les perméases aux métaux de transition 56

2.2.1.3.1 La protéine HoxN 56

2.2.1.3.2 La protéine NixA ... 57
2.2.1.3.3 D'autres exemples de perméases ... 57
2.2.1.4 D'autres systèmes permettant l'import du nickel ... 57
2.3 La régulation de la concentration cytoplasmique du nickel ... 58
2.3.1 La régulation de l'expression du système NikABCDE ... 58
2.3.1.1 L'activation de la transcription de l'opéron par FNR ... 58
2.3.1.2 La répression de la transcription de l'opéron par NikR ... 59
2.3.1.3 La répression de la transcription de l'opéron par NarL ... 61
2.3.2 Les autres régulations de transcription contrôlées par le nickel ... 61
2.3.2.1 La protéine NikR chez d'autres bactéries ... 61
2.3.2.2 D'autres exemples de régulations par le nickel ... 61
2.3.3 Le stockage cytoplasmique du nickel ... 62
2.3.4 L'export du nickel ... 62

3 GENERALITES SUR L'IMPORT DU FER ... 63
3.1 Le fer cellulaire ... 63
3.2 L'import direct du fer par les microorganismes ... 64
3.3 L'import indirect du fer par les microorganismes ... 64
3.3.1 Les sidérophores ... 64
3.3.2 Le transport du fer par les sidérophores ... 65

4 L'INTERET DE L'ETUDE STRUCTURALE DES PROTEINES NIKA ET NIKR ... 67
4.1 La protéine NikR ... 68
4.2 La protéine NikA ... 68

MATERIEL ET METHODES ... 71

1 LA BIOLOGIE MOLECULAIRE ET LA BIOCHIMIE ... 73
1.1 La biologie moléculaire ... 73
1.1.1 Le principe de la surexpression des protéines ... 73
1.1.2 Le protocole de surexpression de la protéine NikA ... 73
1.1.2.1 La surexpression de NikA en milieu minimum ... 74
1.1.2.1.1 La transformation de la souche B834(DE3) par la construction pET22b/nika ... 74
1.1.2.1.2 L'induction de la surexpression ... 74
1.1.2.2 La surexpression de NikA séléniée en milieu minimum ... 75
1.1.2.3 La surexpression de NikA en LB et TB ... 75
1.1.2.3.1 La transformation de la souche BL21(DE3) par la construction pET22b/nika ... 75
1.1.2.3.2 L'induction de la surexpression ... 76
1.2 La biochimie ... 76
1.2.1 Le principe de la purification des protéines ... 76
1.2.2 Le protocole de purification de la protéine NikA ... 77
1.2.2.1 La purification de la protéine native avec extraction périplasmique à l'EDTA ... 78
1.2.2.1.1 L'extraction périplasmique à l'EDTA ... 78

1.2.2.1.2 La précipitation au sulfate d'ammonium 78
1.2.2.1.3 La purification par chromatographie sur colonne Q – sépharose (Pharmacia) 79
1.2.2.1.4 La purification par chromatographie sur colonne Resource – Q (Pharmacia) 79
1.2.2.2 La purification de la protéine séléniée avec extraction périplasmique à l'EDTA 79
1.2.2.3 La purification de la protéine native avec extraction périplasmique au chloroforme 79
1.2.2.3.1 L'extraction périplasmique au chloroforme 79
1.2.2.3.2 La précipitation au sulfate d'ammonium 80
1.2.2.3.3 La suite de la purification 80
1.2.3 L'estimation de la concentration de protéine 80
1.2.4 L'analyse par gel de polyacrylamide 80
1.2.5 L'estimation de la concentration de fer 81

2 LA CRISTALLOGRAPHIE AUX RAYONS X DES PROTEINES 82
2.1 Le principe et l'objectif de la cristallographie des protéines 82
2.2 La cristallogenèse 84
2.2.1 Le principe de la cristallisation 84
2.2.1.1 Les paramètres influençant la solubilité des protéines 84
2.2.1.2 Le diagramme de phase 84
2.2.2 Les techniques de cristallisation des protéines utilisées durant cette thèse 86
2.2.2.1 La diffusion de vapeur en goutte suspendue 86
2.2.2.2 L'ensemencement 87
2.2.2.2.1 Le principe de l'ensemencement 87
2.2.2.2.2 Le protocole du micro-ensemencement utilisé pour reproduire les cristaux de NikA . 88
2.3 L'enregistrement et le traitement des données de diffraction 88
2.3.1 La cryocristallographie 88
2.3.1.1 L'intérêt de la congélation des cristaux de protéine 88
2.3.1.2 Les cryoprotectants 89
2.3.2 La diffraction des rayons X par les cristaux 89
2.3.2.1 La direction des rayons X diffractés par un cristal 89
2.3.2.2 Le facteur de structure 90
2.3.3 L'enregistrement des données de diffraction 92
2.3.3.1 Le diffractomètre de laboratoire 93
2.3.3.2 Le synchrotron 94
2.3.4 Le traitement des données de diffraction 94
2.3.4.1 Le principe 94
2.3.4.2 Les programmes utilisés 95
2.3.4.3 La validation des données de diffraction 95
2.3.5 L'analyse des caractéristiques du cristal 96
2.3.5.1 Les systèmes cristallins et les groupes d'espace 96
2.3.5.2 La mosaïcité 97
2.3.5.3 Le désordre des atomes du cristal 97
2.3.5.4 L'estimation du nombre de molécules de l'unité asymétrique 97
2.4 La résolution du problème de la phase 98
2.4.1 Le problème de la phase 98

2.4.2 La diffusion anomale ... 98
2.4.2.1 L'origine de la diffusion anomale ... 98
2.4.2.2 L'utilisation de la diffusion anomale pour résoudre le problème de la phase ... 99
2.4.2.2.1 Le MAD ... 99
2.4.2.2.2 Le SAD ... 100
2.4.3 Le remplacement moléculaire ... 101
2.4.3.1 La fonction de Patterson ... 101
2.4.3.2 La fonction de rotation ... 102
2.4.3.3 La fonction de translation ... 103
2.4.3.4 Les programme utilisés ... 103
2.5 La résolution de la structure de la protéine ... 104
2.5.1 La modification de densité ... 104
2.5.1.1 L'aplatissement de solvant ... 104
2.5.1.2 La moyennation des molécules reliées par symétrie non cristallographique ... 105
2.5.1.3 L'histogramme de densité ... 105
2.5.1.4 La combinaison des phases ... 105
2.5.2 L'affinement et reconstruction de la structure ... 106
2.5.2.1 L'affinement par minimisation d'un résiduel ... 107
2.5.2.2 La dynamique moléculaire ... 107
2.5.2.3 Les cartes de densité électronique ... 108
2.5.2.3.1 Les cartes différences ... 108
2.5.2.3.2 Les cartes « omit » ... 108
2.5.2.3.3 Les cartes de différence anomale ... 109
2.5.2.4 Les critères d'évaluation du modèle ... 111
2.5.2.5 Les programmes utilisés ... 112
2.5.2.5.1 CNS ... 112
2.5.2.5.2 CCP4 – REFMAC ... 112
2.5.3 L'analyse de la structure ... 113

3 LA SPECTROSCOPIE D'ABSORPTION DES RAYONS X (SAX) ... 115
3.1 Le principe de la SAX ... 115
3.1.1 Le phénomène physique ... 115
3.1.2 La SAX des métalloprotéines ... 117
3.1.3 Le préseuil ... 118
3.1.4 Le XANES ... 118
3.1.5 L'EXAFS ... 121
3.2 Le formalisme de la SAX ... 122
3.2.1 Le coefficient d'absorption ... 122
3.2.2 La formule EXAFS ... 123
3.3 L'enregistrement des données SAX ... 124
3.4 Le traitement des données EXAFS ... 125
3.4.1 Etape 1 : La détermination de l'énergie du seuil E_0 et de la hauteur du saut d'absorption ... 125
3.4.2 Etape 2 : La normalisation du spectre d'absorption ... 127

3.4.3 Etape 3 : La conversion du coefficient d'absorption dans l'espace réciproque [$Å^{-1}$] et le calcul de la fonction EXAFS $\chi(k)$ 127
3.4.4 Etape 4 : La transformée de Fourier 128
3.4.5 Etape 5 : L'ajustement et obtention des informations structurales 128
3.4.5.1 La détermination de $f_i(k)$, $\phi_{ij}(k)$ et $\lambda(k)$ 129
3.4.5.2 La calibration et validation de S_0^2 et ΔE_0 130
3.4.5.3 La détermination de R_i, N_i et σ_i du site métallique à caractériser 131

4 LA SPECTROSCOPIE DE MASSE 132
4.1 Le principe de la spectroscopie de masse 132
4.1.1 L'électrospray 132
4.1.2 La séparation et détection des ions 133
4.2 La spectroscopie de masse de la protéine NikA 134
4.2.1 La préparation des échantillons et la réalisation de l'expérience en conditions dénaturantes 134
4.2.2 La préparation des échantillons, et la réalisation de l'expérience en conditions natives 134
4.3 L'extraction et analyse par spectroscopie de masse de la molécule présente dans NikA .. 135

5 LA RESONANCE MAGNETIQUE NUCLEAIRE (RMN) 137
5.1 Le principe de la RMN 137
5.2 La préparation des échantillons, et la réalisation de l'expérience 138

6 LA RESONANCE PARAMAGNETIQUE ELECTRONIQUE (RPE) 139
6.1 Le principe de la RPE 139
6.2 La préparation des échantillons, et la réalisation de l'expérience 139

7 LE DICHROÏSME CIRCULAIRE (CD) 140
7.1 Le principe du CD 140
7.2 La préparation des échantillons, et la réalisation de l'expérience 141

ETUDE STRUCTURALE DE LA PROTEINE NIKR 143

1 LA CRISTALLOGENESE DE LA PROTEINE NIKR 145
1.1 Les cristaux de NikR obtenus en utilisant l'isopropanol comme précipitant 146
1.2 Les cristaux de NikR obtenus en utilisant les PEG comme précipitant 148

2 L'ENREGISTREMENT ET TRAITEMENT DES JEUX DE DIFFRACTION 150
2.1 Les cristaux de NikR obtenus en utilisant l'isopropanol comme précipitant 150
2.2 Les cristaux de NikR obtenus en utilisant les PEG comme précipitant 151
2.2.1 Les cristaux cryoprotégés dans l'azote liquide 151
2.2.2 Les cristaux cryoprotégés dans le propane liquide 152
2.3 Les Tentatives de résolution du problème de phases 153

2.3.1 Le jeu de données de diffraction SAD au seuil du nickel 154
2.3.2 La recherche d'atome lourds se fixant à la structure de NikR permettant un phasage par la méthode MAD ou SAD 156
2.3.3 La production et cristallisation de la protéine NikR séléniée 156

3 LES STRUCTURES DE LA PROTEINE NIKR 156
3.1 La structure de NikR d'Escherichia coli 156
3.1.1 La structure de la protéine apo entière 157
3.1.2 Le site de fixation du nickel 158
3.1.3 La fixation à l'ADN 160
3.2 La structure de NikR de Pyrococcus horikoshii 161
3.2.1 Les conformations des structures résolues 161
3.2.2 Les sites de fixation du nickel de haute affinité 161
3.2.3 Les sites de fixation du nickel de faible affinité 163
3.2.4 Le modèle de la régulation par le nickel de la fixation du tétramère de NikR à l'ADN 163

4 LE PHASAGE DE NOS DONNEES DE DIFFRACTION PAR REMPLACEMENT MOLECULAIRE 164
4.1 Le remplacement moléculaire à partir de NikR d'Escherichia coli 164
4.2 Le remplacement moléculaire à partir de NikR de Pyrococcus horikoshii 165

ETUDE STRUCTURALE DE LA PROTEINE NIKA 167

1 L'ETUDE DE LA PROTEINE NIKA, PURIFIEE SELON LE PROTOCOLE CONTENANT DE L'EDTA ... 169
1.1 La surexpression de la protéine NikA 169
1.1.1 La protéine NikA surexprimée en milieu minimum 169
1.1.1.1 La protéine NikA native 169
1.1.1.2 La protéine NikA séléniée 170
1.1.2 La surexpression en milieu riche 170
1.2 La purification de la protéine NikA 171
1.2.1 La purification de la protéine NikA surexprimée en milieu minimum 171
1.2.1.1 La purification de la protéine NikA native 171
1.2.1.2 La purification de la protéine NikA séléniée 171
1.2.1.2.1 La colonne Q – Sépharose 171
1.2.1.2.2 La colonne Resource – Q 172
1.2.1.2.3 Le rendement de purification 173
1.2.2 La protéine NikA surexprimée en milieu riche 173
1.3 La cristallisation de la protéine NikA 173
1.3.1 Les tentatives de reproduction des cristaux de la protéine native 173
1.3.2 La reproduction des cristaux par micro-ensemencement 174
1.3.2.1 La cristallisation de la protéine native 174
1.3.2.2 La cristallisation de la protéine séléniée 175
1.4 L'enregistrement et traitement des données de diffraction 176
1.4.1 Le jeu de données de la protéine NikA native 177

1.4.2 Le jeu de données MAD de la protéine NikA 177
1.4.2.1 L'enregistrement et indexation du jeu MAD 177
1.4.2.2 La localisation des atomes de sélénium 179
1.4.2.3 Des exemples de tentatives de phasage de la structure de NikA 180
1.5 Les structures de NikA publiées par Heddle et al. 181
1.5.1 La description de la structure de NikA 182
1.5.2 Le site de fixation du nickel 183
1.5.3 La comparaison de la structure de NikA avec les protéines OppA et DppA 184
1.6 La résolution de notre structure de NikA par remplacement moléculaire 185
1.7 Pourquoi n'a-t-on pas pu résoudre la structure de NikA à partir de la protéine séléniée : un début de réponse 186
1.7.1 La vérification de la présence des atomes de sélénium par cristallographie 186
1.7.2 L'analyse du contenu en atomes lourds des cristaux de NikA séléniée par fluorescence des rayons X 187
1.8 L'étude du site de fixation de NikA 189
1.8.1 La densité électronique du contenu du site de fixation 189
1.8.1.1 Les problèmes posés par la fixation du nickel à NikA sous une forme penta-hydratée 189
1.8.1.2 La densité électronique présente dans le site de fixation de la protéine NikA 190
1.8.2 La caractérisation du métal présent dans le site de fixation de NikA 191
1.8.2.1 La réalisation de cartes de double différence anomale 191
1.8.2.1.1 L'analyse du contenu en atomes lourds des cristaux de NikA par fluorescence des rayons X 191
1.8.2.1.2 L'enregistrement des données de diffraction 192
1.8.2.1.3 Le calcul des cartes de double différence anomale 193
1.8.2.2 La spectroscopie RPE 195
1.8.3 La caractérisation du contenu du site de fixation de NikA 195
1.8.3.1 L'identification par cristallographie 195
1.8.3.1.1 Les tentatives d'identification du ligand de NikA 195
1.8.3.1.2 L'obtention d'une structure diffractant à 1,8 Å de résolution 196
1.8.3.1.3 L'identification de l'EDTA comme ligand du nickel 198
1.8.3.2 La spectroscopie de masse 200
1.8.3.3 Le contenu du site de fixation de NikA dans la structure de Heddle et al 201
1.8.4 NikA contient un complexe $FeEDTA(H_20)^-$ 202
1.8.4.1 La fixation du complexe $FeEDTA(H_20)^-$ à la protéine NikA 202
1.8.4.2 L'origine du FeEDTA 203
1.8.5 La mise en évidence de l'existence d'un métallophore spécifique du nickel 204

2 L'ETUDE DE LA PROTEINE NIKA, PURIFIEE SELON LE PROTOCOLE CONTENANT DU CHLOROFORME 206
2.1 La production de la protéine NikA, purifiée selon un protocole sans EDTA 206
2.1.1 La surexpression de la protéine en milieu riche 206
2.1.2 La purification de la protéine selon le protocole contenant du chloroforme 206
2.1.2.1 L'extraction périplasmique au chloroforme et les précipitations au sulfate d'ammonium. 206
2.1.2.2 La colonne Q – Sépharose 207

2.1.2.3 La colonne Resource – Q 208
2.1.2.4 Le rendement de la purification 208
2.2 La cristallisation de la protéine 209
2.3 L'enregistrement et le traitement des données de diffraction 210
2.4 La résolution de la structure par remplacement moléculaire 211
2.4.1 Les premières données de diffraction de NikA 211
2.4.2 Les nouvelles données de diffraction 212
2.5 La comparaison de la structure hexagonale avec les structures orthorhombiques et apo 214
2.5.1 La comparaison des structures 214
2.5.2 Le contenu du site de fixation 215
2.6 La caractérisation du métal présent dans le site de fixation de NikA 217
2.6.1 L'étude d'un cristal non mis en présence de nickel 217
2.6.1.1 L'analyse du contenu en atomes lourds des cristaux de NikA par fluorescence des rayons X 217
2.6.1.2 L'enregistrement des données de diffraction 218
2.6.1.3 Le calcul des cartes différence anomale 219
2.6.2 L'étude d'un cristal mis en présence de nickel 221
2.6.2.1 L'analyse du contenu en atomes lourds des cristaux de NikA par fluorescence des rayons X 221
2.6.2.2 L'enregistrement des données de diffraction 221
2.6.2.3 Le calcul des cartes différence anomale 222
2.7 La caractérisation du métallophore présent dans le site de fixation de NikA 223
2.7.1 La cristallographie 223
2.7.1.1 La densité électronique 223
2.7.1.2 La tentative d'amélioration de la résolution 226
2.7.2 La spectroscopie d'absorption des rayons X (SAX) 227
2.7.2.1 Le XANES 227
2.7.2.2 L'EXAFS 229
2.7.2.2.1 Le complexe de nickel-EDTA 229
2.7.2.2.2 La protéine NikA au seuil du nickel 230
2.7.3 L'extraction et l'analyse de la molécule présente dans le site de fixation de NikA 233
2.7.3.1 L'extraction du nickelophore de la protéine 233
2.7.3.1.1 La concentration d'EDTA permettant le déplacement du nickelophore 233
2.7.3.1.2 L'étude de la protéine par dichroïsme circulaire 234
2.7.3.2 La spectroscopie de masse 236
2.7.3.2.1 L'étude de la protéine NikA en conditions natives 236
2.7.3.2.2 La caractérisation de la masse du métallophore 237
2.7.3.3 La spectroscopie RMN 240
2.8 La comparaison des systèmes de transport du nickel et fer 242
2.9 La comparaison de la protéine NikA d'E. coli avec ses homologues 243

CONCLUSION ET PERSPECTIVES 249

1 LA PROTEINE NIKR 251

2 LA PROTEINE NIKA 253

BIBLIOGRAPHIE 257

ANNEXES 277

Liste des Figures

Introduction

FIGURE 1.1-1 : TABLEAU PERIODIQUE DES ELEMENTS, EN ROUGE EST INDIQUE LE NICKEL 31
FIGURE 1.2-1 : FEUILLE DE NICKEL 32
FIGURE 1.3-1 : LA PERIDOTINE, LA SERPENTINE, LA GRANIERITE 34
FIGURE 1.3-2 : MINES DE NICKEL EN NOUVELLE CALEDONIE. 34
FIGURE 2.1-1 : HYDROLYSE DE L'UREE PAR L'UREASE EN CARBAMATE ET AMMONIAQUE ; 38
FIGURE 2.1-2 : STRUCTURE CRISTALLOGRAPHIQUE DE L'UREASE DE *H. PYLORI* (CODE PDB : 1E9Z). 40
FIGURE 2.1-3 : MODELE DE L'ACTIVATION DE L'UREASE. 41
FIGURE 2.1-4 : REACTION REVERSIBLE DE REDUCTION DE L'HYDROGENE MOLECULAIRE PAR L'HYDROGENASE. ... 42
FIGURE 2.1-5 : STRUCTURE DE L'HYDROGENASE A NIFE DE *D. GIGAS* 43
FIGURE 2.1-6 : SCHEMA DE LA MATURATION DE L'HYDROGENASE A NIFE ET DES SENSEURS A HYDROGENE. 44
FIGURE 2.1-7 : REACTION CATALYSEE PAR LA MONOXYDE DE CARBONE DESHYDROGENASE (CODH). 45
FIGURE 2.1-8 : REACTION CATALYSEE PAR L'ACTIVITE ACETYL-COENZYME A SYNTHETASE. 45
FIGURE 2.1-9 : REACTION CATALYSEE PAR LA METHYL-COENZYME M REDUCTASE. 47
FIGURE 2.1-10 : STRUCTURE DU F_{430} EN COMPLEXE AVEC LE NICKEL. 47
FIGURE 2.1-11 : REACTION CATALYSEE PAR LA SUPEROXYDE DISMUTASE. 48
FIGURE 2.1-12 : REACTION CATALYSEE PAR LA GLYOXYLASE. 48
FIGURE 2.1-13 : REACTION CATALYSEE PAR L'ACI-REDUCTONE DIOXYGENASE 49
FIGURE 2.2-1 : SCHEMA DES ENVELOPPES DES BACTERIES A. GRAM (+) ; B. GRAM (-). 51
FIGURE 2.2-2 : STRUCTURE DE LA PORINE OMPA D'*E. COLI* 51
FIGURE 2.2-3 : STRUCTURE DU PREMIER SYSTEME ABC (SOUS FAMILLE 8) ENTIEREMENT CRISTALLISE : 52
FIGURE 2.2-4 : SYSTEME DE TRANSPORT DU NICKEL, DU TYPE ABC, CODE PAR L'OPERON *NIKABCDE* D'*E. COLI*..... 54
FIGURE 2.3-1 : SCHEMA DU TRANSPORT ET DE LA REGULATION DU NICKEL CHEZ *E. COLI* PAR NIKABCDE..... 59
FIGURE 3.1-1 : REACTION DE FENTON. 63
FIGURE 3.3-1 : STRUCTURE DE TROIS SIDEROPHORES : FERRICHROME, ENTEROBACTINE, PYOCHELINE. 65
FIGURE 3.3-2 : SCHEMA DE L'IMPORT ET DE L'EXPORT DES SIDEROPHORES CHEZ LES GRAM (+) ET GRAM (-)..... 66

Matériel et Méthodes

FIGURE 2.1-1 : PRINCIPE DE LA CRISTALLOGRAPHIE AUX RAYONS X 82
FIGURE 2.1-2 : SCHEMA D'UN CRISTAL DE PROTEINE. 83
FIGURE 2.2-1 : DIAGRAMME DE PHASES 85
FIGURE 2.2-2 : SCHEMA DU PRINCIPE DE LA DIFFUSION DE VAPEUR EN GOUTTE SUSPENDUE 86

FIGURE 2.3-1 : SCHEMA DECRIVANT LE PHENOMENE DE LA DIFFRACTION SELON LES LOIS DE LAUE ET DE BRAG . 90
FIGURE 2.3-2 : REPRESENTATION DANS LE PLAN COMPLEXE DU FACTEUR DE STRUCTURE $F(\vec{S})$ 91
FIGURE 2.3-3: CLICHE DE DIFFRACTION DE LA PROTEINE NIKA..... 93
FIGURE 2.4-1 : REPRESENTATION VECTORIELLE DES FACTEURS DE STRUCTURE 99
FIGURE 2.5-1 : SCHEMA DE LA PROCEDURE DE MODIFICATION DE DENSITE ELECTRONIQUE. 106
FIGURE 2.5-2 : REPRESENTATION DE L'EVOLUTION DES F" DU FER, NICKEL ET ZINC EN FONCTION DE L'ENERGIE DES RAYONS X AUXQUELS EST SOUMIS LE CRISTAL 109
FIGURE 3.1-1 : REPRESENTATION SCHEMATIQUE DE L'INTERACTION RAYONNEMENT-MATIERE. 116
FIGURE 3.1-2 : SPECTRE D'ABSORPTION NORMALISE DU NI – EDTA (SEUIL K DU NICKEL) : PRESEUIL. 118
FIGURE 3.1-3 : SPECTRE D'ABSORPTION NORMALISE DU NI – EDTA (SEUIL K DU NICKEL) : XANES. 119
FIGURE 3.1-4 : PROCESSUS DE L'ABSORPTION X. 120
FIGURE 3.1-5 : SPECTRE D'ABSORPTION NORMALISE DU NI – EDTA (SEUIL K DU NICKEL) : EXAFS. 121
FIGURE 3.3-1 : EXEMPLE D'UNE LIGNE DE LUMIERE DE L'ESRF DEDIEE A LA SAX : LIGNE ID26 125
FIGURE 3.4-1 : EXTRACTION DES INFORMATIONS STRUCTURALES DU SPECTRE D'ABSORPTION X. 126
FIGURE 4.1-1 : FORMATION D'UN NUAGE DE GOUTTELETTES CHARGEES (ELECTROSPRAY). 132
FIGURE 4.1-2 : SCHEMA DE FONCTIONNEMENT DU SPECTROMETRE DE MASSE Q-TOF MICRO. 133
FIGURE 7.1-1 : SCHEMA DU DICHROÏSME CIRCULAIRE. 140

Etude structurale de la protéine NikR

FIGURE 1.1-1 : CRISTAUX DE NIKR AYANT POUSSE DANS DES CONDITIONS A BASE D'ISOPROPANOL 146
FIGURE 1.2-1 : CRISTAUX DE NIKR AYANT POUSSE DANS DES CONDITIONS A BASE DE PEG. 149
FIGURE 2.3-1 : CARTE DE PATTERSON DIFFERENCE DANS SA SECTION DE HARKER V = 0,5, CALCULEE A PARTIR DU JEU SAD OBTENU POUR UN CRISTAL DE NIKR..... 155
FIGURE 3.1-1 : STRUCTURE DE L'HOMOTETRAMERE DE NIKR D'*E. COLI*..... 158
FIGURE 3.1-2 : SITE DE FIXATION DU NICKEL DANS LE DOMAINE C TERMINAL DE NIKR. 159
FIGURE 3.1-3 : SUPERPOSITION DE LA STRUCTURE DE NIKR ENTIERE APO ET DU DOMAINE C TERMINAL FIXANT LE NICKEL 159
FIGURE 3.2-1 : STRUCTURES DES TETRAMERES DE NIKR DE *P. HORIKOSHII* 162
FIGURE 3.2-2 : SCHEMA ILLUSTRANT LES DIFFERENTES CONFORMATIONS ADOPTEES PAR LE TETRAMERE DE NIKR DE *P. HORIKOSHII*, EN FONCTION DE LA CONCENTRATION DE NICKEL ET DE LA PRESENCE OU NON D'ADN. 163
FIGURE 4.1-1 : SUPERPOSITION DE LA STRUCTURE DE NIKR APO PUBLIEE ET DE LA STRUCTURE INCOMPLETE QUE NOUS AVONS RESOLUE 165

Etude structurale de la protéine NikA

FIGURE 1.1-1 : GEL SDS - PAGE DES PROTEINES PERIPLASMIQUES BACTERIENNES AVANT ET APRES LA SUREXPRESSION DE LA PROTEINE NIKA NATIVE. 169

FIGURE 1.1-2 : SDS - PAGE DES PROTEINES PERIPLASMIQUES BACTERIENNES AVANT ET APRES LA SUREXPRESSION DE LA PROTEINE NIKA SELENIEE. 170
FIGURE 1.2-1 : ETAPE DE PURIFICATION SUR COLONNE Q - SEPHAROSE. 172
FIGURE 1.2-2 : ETAPE DE PURIFICATION SUR COLONNE RESOURCE - Q. 172
FIGURE 1.3-1 : CRISTAUX DE LA PROTEINE NIKA NATIVE REALISES PAR MICRO-ENSEMENCEMENT..... 175
FIGURE 1.3-2 : CRISTAUX DE LA PROTEINE NIKA SELENIEE REALISES PAR MICRO-ENSEMENCEMENT..... 176
FIGURE 1.4-1 : SPECTRE DE FLUORESCENCE D'UN CRISTAL DE LA PROTEINE NIKA SELENIEE. 178
FIGURE 1.4-2 : CARTE DE PATTERSON DIFFERENCE DANS SA SECTION DE HARKER U = 0,5, CALCULEE A PARTIR DU JEU MAD OBTENU POUR UN CRISTAL DE NIKA SELENIEE 180
FIGURE 1.5-1 : SUPERPOSITION DES STRUCTURES DE NIKA APO ET NIKA AYANT FIXE LE NICKEL..... 182
FIGURE 1.5-2 : VUE STEREOSCOPIQUE DE LA SUPERPOSITION DES FORMES OUVERTE ET FERMEE DE NIKA 183
FIGURE 1.5-3 : REPRESENTATIONS DU SITE DE FIXATION DU NICKEL DANS NIKA PROPOSE PAR HEDDLE *ET AL.* .. 184
FIGURE 1.5-4 : SUPERPOSITION DE LA STRUCTURE DE NIKA APO ET DE OPPA..... 185
FIGURE 1.7-1 : CARTE DE DENSITE ELECTRONIQUE $2F_{OBS}$-F_{CALC} (EN BLEU), ET DE LA DENSITE ELECTRONIQUE F_{OBS}-F_{CALC} DE LA METHIONINE 32 DE NIKA. 187
FIGURE 1.7-2 : PICS DE FLUORESCENCE X MESUREE SUR DES CRISTAUX DE NIKA SELENIEE. 188
FIGURE 1.8-1 : VUES A 90° L'UNE DE L'AUTRE DE LA DENSITE ELECTRONIQUE F_{OBS}-F_{CALC} DE LA STRUCTURE DE NIKA CONTENANT UN NICKEL RESOLUE PAR *HEDDLE ET AL.*, AVEC UN MODELE DE NICKEL PENTA-HYDRATE ET UN NICKEL HEXA-HYDRATE CANONIQUE. 190
FIGURE 1.8-2 : VUE STEREOSCOPIQUE DE LA DENSITE ELECTRONIQUE « OMIT » PRESENTE DANS LE SITE DE FIXATION DE NIKA 191
FIGURE 1.8-3 : PICS DE FLUORESCENCE X MESUREE SUR DES CRISTAUX DE NIKA..... 192
FIGURE 1.8-4 : CARTE DE DENSITE ELECTRONIQUE DOUBLE DIFFERENCE ANOMALE METTANT EN EVIDENCE LA PRESENCE DE FERDANS LE SITE DE FIXATION DE NIKA. 193
FIGURE 1.8-5 : SPECTRE RPE ENREGISTRE SUR LA PROTEINE NIKA. 195
FIGURE 1.8-6 : EXEMPLES DE MODELES DE LIGANDS DU METAL PRESENT DANS LE SITE DE NIKA. 196
FIGURE 1.8-7 : VUE STEREOSCOPIQUE DE LA DENSITE ELECTRONIQUE CALCULEE AUTOUR DE L'ATOME DE FER. 198
FIGURE 1.8-8 : LE COMPLEXE FEEDTA$(H_2O)^-$. 199
FIGURE 1.8-9 : SPECTRE DE MASSE EN CONDITIONS NATIVES DE LA PROTEINE NIKA DIALYSEE CONTRE DE L'ACETATE D'AMMONIUM 20 MM. 200
FIGURE 1.8-10 : VUE STEREOSCOPIQUE DE LA DENSITE ELECTRONIQUE CALCULEE AUTOUR DU METAL, DANS LE CAS DE LA STRUCTURE PUBLIEE PAR *HEDDLE ET AL.* 201
FIGURE 1.8-11 : REPRESENTATION DE LA FIXATION DU COMPLEXE DE FEEDTA$(H2O)^-$ DANS LE SITE DE FIXATION DE NIKA. 202
FIGURE 2.1-1 : ETAPE DE PURIFICATION SUR COLONNE Q - SEPHAROSE. 207
FIGURE 2.1-2 : ETAPE DE PURIFICATION SUR COLONNE RESOURCE – Q. 208
FIGURE 2.2-1 : CRISTAUX DE LA PROTEINE NIKA, PURIFIEE SELON LE PROTOCOLE DE PURIFICATION CONTENANT DU CHLOROFORME. 209
FIGURE 2.4-1 : REPRESENTATION DE L'EMPILEMENT CRISTALLIN DANS LA FORME HEXAGONALE DES CRISTAUX DE NIKA, PURIFIEE SELON LE PROTOCOLE CONTENANT DU CHLOROFORME. 213

FIGURE 2.5-1 : SUPERPOSITION DES STRUCTURES DE NIKA APO, NIKA PURIFIEE SELON LE PROTOCOLE DE PURIFICATION CONTENANT DE L'EDTA ET NIKA PURIFIEE SELON LE PROTOCOLE DE PURIFICATION CONTENANT DU CHLOROFORME. 215
FIGURE 2.5-2 : DENSITE ELECTRONIQUE « OMIT », DANS LE SITE DE FIXATION DE NIKA PURIFIEE SELON LE PROTOCOLE CONTENANT DU CHLOROFORME, CALCULEE A PARTIR D'UN MODELE DE LA PROTEINE NE CONTENANT QUE LE METALDANS LE SITE DE FIXATION. 216
FIGURE 2.5-3 : MOUVEMENT DE L'HISTIDINE 416 ENTRE LES FORMES DE NIKA APO, AYANT FIXE L'EDTA ET PURIFIEE SELON LE PROTOCOLE CONTENANT DU CHLOROFORME 216
FIGURE 2.5-4 : COMPARAISON DES RESIDUS DU SITE DE FIXATION DE LA PROTEINE AYANT FIXE L'EDTA ET LA PROTEINE NOUVELLEMENT PURIFIEE SELON LE PROTOCOLE CONTENANT DU CHLOROFORME..... 217
FIGURE 2.6-1 : PICS DE FLUORESCENCE X MESUREE SUR DES CRISTAUX DE NIKA PURIFIEE SELON LE PROTOCOLE CONTENANT DU CHLOROFORME. 218
FIGURE 2.7-1 : VUE STEREOSCOPIQUE DU SITE DE FIXATION ET DES LIAISONS HYDROGENES ENTRE L'ACIDE 3-(CARBOXY)HEXANEDIOÏQUE ET NIKA. 225
FIGURE 2.7-2 : FIGURES LIGPLOT DE L'ACIDE 3-(CARBOXY)HEXANEDIOIQUE EN COMPLEXE AVEC LE NICKEL DANS LE SITE DE FIXATION DE NIKA 225
FIGURE 2.7-3 : XANES DE LA PROTEINE NIKA PURIFIEE SELON LE PROTOCOLE DE PURIFICATION CONTENANT DU CHLOROFORME, DU NICKEL, ET UN COMPLEXE DE NIEDTA EN SOLUTION, ENREGISTRE AU SEUIL DU NICKEL 228
FIGURE 2.7-4 : STRUCTURE CRISTALLOGRAPHIE D'UN COMPLEXE DE NIEDTA. 230
FIGURE 2.7-5 : TRANSFORME DE FOURIER DES DONNEES EXAFS. 231
FIGURE 2.7-6 : TEST DE CRISTALLISATION DE LA PROTEINE NIKA PURIFIEE EN PRESENCE DE CHLOROFORME..... 233
FIGURE 2.7-7 : DICHROÏSME CIRCULAIRE. 235
FIGURE 2.7-8 : SPECTRE DE MASSE EN CONDITIONS NATIVES DE LA PROTEINE NIKA PURIFIEE SELON LE PROTOCOLE CONTENANT DU CHLOROFORME, ET DIALYSEE CONTRE DE L'ACETATE D'AMMONIUM 20 MM. 236
FIGURE 2.7-9 : SPECTRE DE MASSE DU SURNAGEANT D'UNE SOLUTION DE NIKA PURIFIEE SELON LE PROTOCOLE CONTENANT DU CHLOROFORME, CHAUFFEE A 70°C. 237
FIGURE 2.7-10 : SPECTRE DE MASSE DE LA MOLECULE EXTRAITE DE NIKA PURIFIEE SELON LE PROTOCOLE CONTENANT DU CHLOROFORME PAR TRAITEMENT A L'EDTA..... 239
FIGURE 2.7-11 : COMPARAISON DES STRUCTURES DE L'ACIDE 3-(CARBOXY)HEXANEDIOÏQUE ET DE L'ACIDE N-CARBOXY N-ETHANOÏQUE PROP-2-ENOÏQUE. 240
FIGURE 2.7-12 : SPECTRE RMN 1D DU PROTON DE LA MOLECULE EXTRAITE DE NIKA PURIFIEE SELON LE PROTOCOLE CONTENANT DU CHLOROFORME, PAR UN GRADIENT DE TEMPERATURE..... 241
FIGURE 2.8-1 : STRUCTURE DE LA PROTEINE PERIPLASMIQUE FHUD D'*E. COLI* IMPLIQUEE DANS LE TRANSPORT DU FER PAR L'INTERMEDIAIRE DES SIDEROPHORES DE LA FAMILLE DES HYDROXAMATES. 243
FIGURE 2.9-1 : ALIGNEMENT DES SEQUENCES DE NIKA D'*E. COLI* ET DE SES HOMOLOGUES..... 247

Liste des tableaux

Matériel et Méthodes

TABLEAU 1.1-1 : COMPOSITION DU MILIEU NYM 74
TABLEAU 2.3-1 : SYSTEMES CRISTALLINS ET EXEMPLES DE GROUPES D'ESPACE POSSIBLES EN BIOLOGIE. 96
TABLEAU 2.5-1 : RESUME DE LA NATURE DES ATOMES IDENTIFIES DANS LES CARTES DE DIFFERENCE ET DE DOUBLE DIFFERENCE ANOMALE. 111

Etude structurale de la protéine NikR

TABLEAU 2.2-1 : STATISTIQUES DES DONNEES DE DIFFRACTION D'UN JEU NATIF DE LA PROTEINE NIKR CONGELEE DANS L'AZOTE LIQUIDE. 152
TABLEAU 2.2-2 : STATISTIQUES DES DONNEES DE DIFFRACTION D'UN JEU NATIF DE LA PROTEINE NIKR CONGELEE DANS LE PROPANE LIQUIDE 153
TABLEAU 2.3-1 : STATISTIQUES DES DONNEES DE DIFFRACTION D'UN JEU SAD AU SEUL DU NICKEL DE LA PROTEINE NIKR CONGELEE DANS L'AZOTE LIQUIDE. 154

Etude structurale de la protéine NikA

TABLEAU 1.4-1 : STATISTIQUES DES DONNEES DE DIFFRACTION D'UN JEU DE LA PROTEINE NIKA NATIVE 177
TABLEAU 1.4-2 : STATISTIQUES DES DONNEES DE DIFFRACTION DES TROIS LONGUEURS D'ONDE D'UN JEU MAD DE LA PROTEINE NIKA SELENIEE. 179
TABLEAU 1.6-1 : STATISTIQUES D'AFFINEMENT DE LA STRUCTURE DE NIKA A 2,2 Å DE RESOLUTION. 186
TABLEAU 1.8-1 : STATISTIQUES DES DONNEES DE DIFFRACTION ET D'AFFINEMENT DE LA STRUCTURE AYANT PERMIS LE CALCUL DES CARTES DE DOUBLE DIFFERENCE ANOMALE. 194
TABLEAU 1.8-2 : STATISTIQUES DES DONNEES DE DIFFRACTION ET D'AFFINEMENT DE LA STRUCTURE DE LA PROTEINE NIKA NATIVE. 197
TABLEAU 2.3-1 : STATISTIQUES DES DONNEES DE DIFFRACTION D'UN JEU DE LA PROTEINE NIKA NATIVE, PURIFIEE SELON LE PROTOCOLE DE PURIFICATION CONTENANT DU CHLOROFORME. 211
TABLEAU 2.4-1 : STATISTIQUES D'AFFINEMENT DE LA STRUCTURE DE NIKA A 2,5 Å DE RESOLUTION. 212
TABLEAU 2.6-1 : STATISTIQUES DES DONNEES DE DIFFRACTION AYANT PERMIS LE CALCUL DES CARTES DE DIFFERENCE ANOMALE A PARTIR DE LA PROTEINE NIKA NATIVE PURIFIEE SELON LE PROTOCOLE CONTENANT DU CHLOROFORME 219
TABLEAU 2.6-2 : ESTIMATION DE LA PROPORTION DE ZINC, NICKEL ET FER PRESENTS DANS LES TROIS MOLECULES DE L'UNITE ASYMETRIQUE. 220

TABLEAU 2.6-3 : STATISTIQUES DES DONNEES DE DIFFRACTION AYANT PERMIS LE CALCUL DES CARTES DE DIFFERENCE ANOMALE A PARTIR DE LA PROTEINE NIKA NATIVE PURIFIEE SELON LE PROTOCOLE CONTENANT DU CHLOROFORME, ET CRISTALLISEE EN PRESENCE D'UN EXCES DE NICKEL 222

TABLEAU 2.6-4 : ESTIMATION DE LA PROPORTION DE NICKEL ET FER PRESENTS DANS LES TROIS MOLECULES DE L'UNITE ASYMETRIQUE D'UN CRISTAL AYANT POUSSE EN PRESENCE DE NICKEL. 223

TABLEAU 2.7-1 : TRANSITIONS 1S → 3D ET 1S → $4P_Z$ VISIBLES DANS LE SPECTRE SAX EN FONCTION DE LA GEOMETRIE DU NICKEL 229

TABLEAU 2.7-2 : PARAMETRES STRUCTURAUX ET PHYSIQUES CALCULES A PARTIR DES DONNEES EXAFS ENREGISTREES SUR UN COMPLEXE DE NIEDTA AU SEUIL DU NICKEL. 229

TABLEAU 2.7-3 : PARAMETRES STRUCTURAUX ET PHYSIQUES CALCULES A PARTIR DES DONNEES EXAFS ENREGISTREES SUR LA PROTEINE NIKA PURIFIEE EN PRESENCE DE CHLOROFORME, AU SEUIL DU NICKEL.. 232

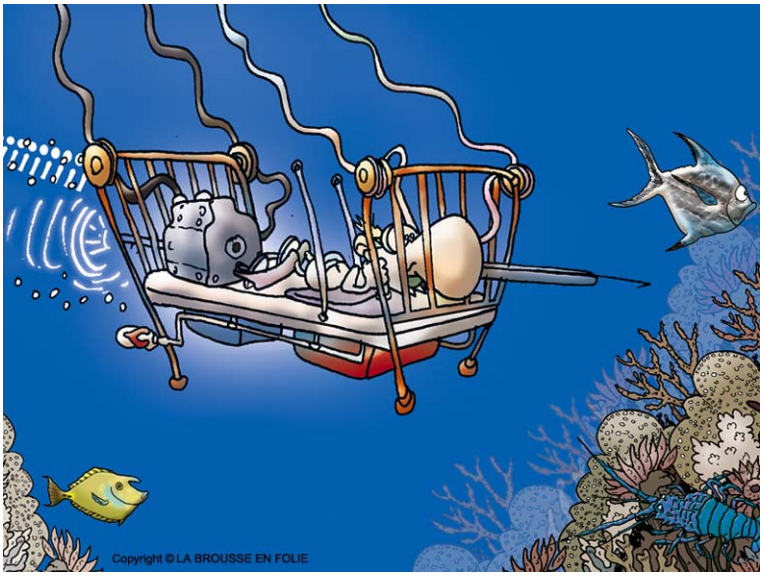
TABLEAU 2.9-1 : PROTEINES ANNOTEES COMME HOMOLOGUES A LA PROTEINE NIKA D'*E. COLI* ET PRESENTANT UNE IDENTITE DE SEQUENCE SIGNIFICATIVE 244

ABREVIATIONS

Å	Angström
α, β, γ	angles associés aux vecteurs unités de la maille cristalline
βOG	n-octyl-β-D-glucopyranoside
λ	longueur d'onde
a, b, c	vecteurs unités de la maille cristalline
ACS	Acétyl-coenzyme A synthase
apo	se dit d'une protéine n'ayant pas fixé un ou plusieurs métaux
B	facteur d'agitation thermique
CD	*Circular Dichroism*
CoA	Coenzyme A
CoB-SH	Coenzyme B
CODH	*Carbon Monoxide DeHydrogenase*
CoM	Coenzyme M
Da / kDa	Dalton / kiloDalton
D.O.	Densité Optique
EDTA	acide EthylèneDiamine-TétrAcétique
ESRF	*European Synchrotron Radiation Facility*
eV / keV	électron Volt / kiloélectron Volt
EXAFS	*Extended X-ray Absorption Fine Structure*
FIP	*French beamline for Investigation of Proteins*
FNR	*Fumarate Nitrate Regulator*
Glx	Glyoxylase
Gram (-)	Bactérie gram négative
Gram (+)	Bactérie gram positive
G-SH	Glutathion
hν	longueur d'onde d'énergie hν
HEPES	acide 4-(2-hydroxyéthyl)piperazine-1-éthanesulfonique
HMQC	*Heteronuclear Multiple-Quantum Coherence*
holo	se dit d'une protéine ayant fixé de métaux
HpNikR	NikR d'*Helicobacter pylori*
I/σ	rapport signal sur bruit moyen
IEF	*IsoElectric Focusing*

K_d	constante de dissociation
K_t	constante de transport
LB	milieu de culture Luria-Bertani
M / mM / µM / nM / pM	Molaire (mol/L) / milliMolaire / microMolaire / nanoMolaire / picoMolaire
MAD	*Multiwavelength Anomalous Dispersion*
Mes	acide 4-MorpholinEthaneSulfonique
NiCoTs	*Nickel / Cobalt Transporters*
NYM	milieu de culture *New Glycerol Minimal Medium*
PAGE	*PolyAcrylamide Gel Electrophoresis*
Pb	Paire de bases
PDB	*Protein Data Bank*
PEG	Polyéthylène Glycol
PEG 5 000 mme	Polyéthylène Glycol 5 000 monométhyléther
ppm	partie par million
RMN	Résonance Magnétique Nucléaire
rmsd	*root mean scare deviation* (déviation moyenne)
RPE	Résonance Paramagnétique Electronique
SAD	*Singlewavelength Anomalous Dispersion*
SAX	Spectroscopie d'Absorption des rayons X
SDS	Sodium Dodécyl Sulfate
SOD	SuperOxyde Dismutase
TB	milieu de culture Terrific Broth
Tris	Tris-hydroxyméthyl-aminométhane
u.a.	unite arbitraire
XANES	*X-ray Absorption Near Edge Structure*

INTRODUCTION



1 Généralités sur le nickel

1.1 L'atome de nickel

Le nickel, de symbole Ni, est un élément chimique appartenant au groupe des métaux de transition (Figure 1.1-1). Son numéro atomique est 28 et il présente une masse atomique moyenne de 58,6934 g.mol^{-1}.

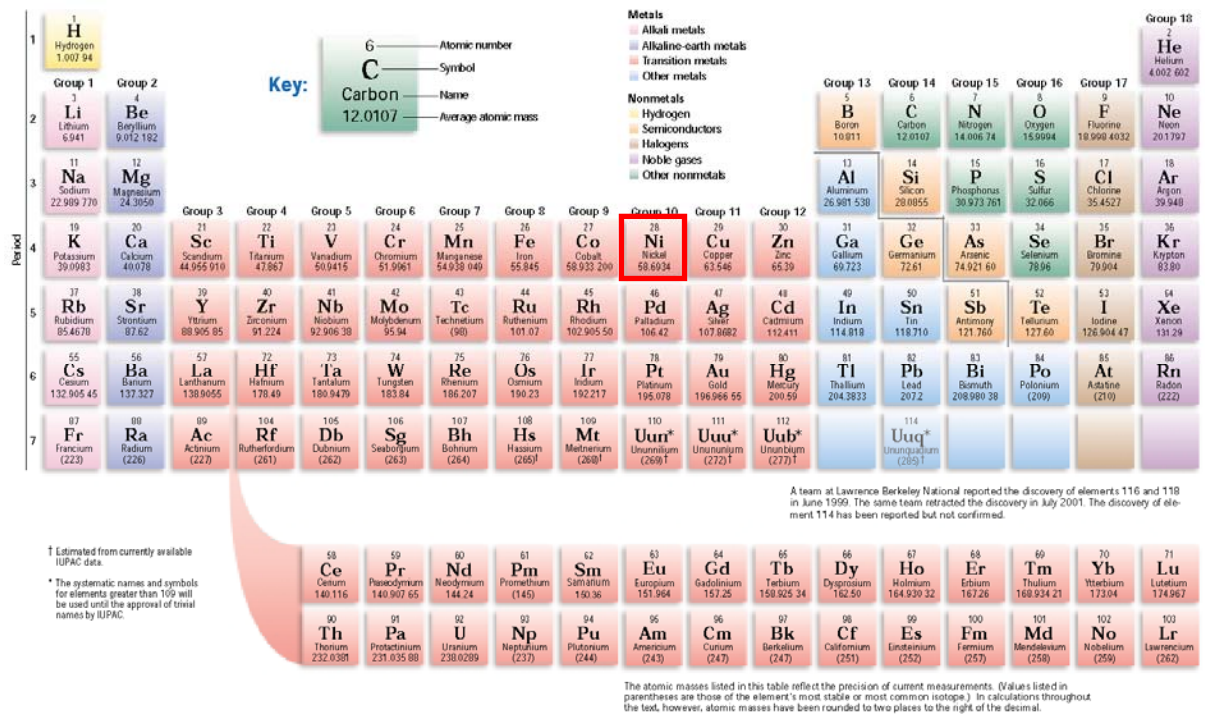


Figure 1.1-1 : Tableau périodique des éléments, en rouge est indiqué le nickel (http://www.sciencegeek.net/tables/Textbook.pdf).

Il existe 5 isotopes naturels (non radioactifs) du nickel : ^{58}Ni, ^{60}Ni, ^{61}Ni, ^{62}Ni, et ^{64}Ni, ayant une abondance naturelle respective de 68,08%, 26,22%, 1,14%, 3,63%, 0,93%. Les radioisotopes sont : ^{56}Ni, ^{57}Ni, ^{59}Ni, ^{63}Ni, ^{65}Ni et ^{66}Ni.

L'état d'oxydation de nickel le plus courant est Ni(II), mais il est également présent dans les états, plus rares, Ni(0), Ni(I), Ni(III) et Ni(IV).

1.2 L'historique de la découverte du nickel

Le nickel est utilisé depuis très longtemps par l'homme. En effet on retrouve des traces de son utilisation remontant jusqu'à 3500 ans av. JC (Wikipédia). Mais du fait qu'il a été très longtemps confondu avec l'argent, sa considération en tant qu'élément à part entière ne fut que très tardive.

C'est en 1751 que le baron Axel Frederik Cronstedt, en voulant extraire du cuivre du Kupfernickel ou « cuivre du diable », un minerai que l'on utilisait pour colorer le verre en vert, obtint un métal blanc (Figure 1.2-1) : le nickel venait d'être découvert.

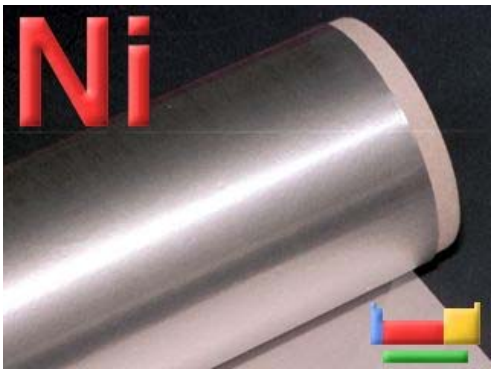


Figure 1.2-1 : Feuille de nickel (http://www.webelements.com).

1.3 La répartition naturelle du nickel

Le nickel est le 24ème élément le plus abondant sur l'écorce terrestre (Sunderman F.W. J.r, 2004). Il est naturellement diffusé dans l'environnement par l'activité volcanique et l'érosion des roches et des sols (environ 150 000 tonnes par an), auxquelles on doit ajouter 180 000 tonnes par an de nickel en provenance des activités humaines (consommation de sources d'énergie fossiles, productions industrielles …).

Dans l'environnement, les teneurs moyennes du nickel sont d'environ :

- 80 ppm dans l'écorce terrestre (chimie) ;
- 2 ppm dans l'eau de mer (soit environ 30 nM) (Eitinger and Mandrand-Berthelot, 2000) ;
- 0,3 ppm dans l'eau douce (soit environ 5 nM) (Eitinger and Mandrand-Berthelot, 2000).

C'est sous forme de minerai que l'on trouve naturellement le nickel, le plus souvent associé au fer (Fe), au cuivre (Cu), au chrome (Cr), ou au cobalt (Co). Ces minerais peuvent être classés en deux grandes familles : les minerais sulfurés et oxydés (chimie).

1.3.1 Les minerais sulfurés

Ils représentent environ 65% de la production mondiale, dans des mines le plus souvent souterraines. Ce type de minerais est principalement exploité en Australie de l'ouest, dans les pays de l'est de l'Europe, au Canada, en Chine, en Afrique australe et en Finlande.
Le nickel s'y trouve sous la forme de pentlandite $(Ni,Fe)_9S_8$, le plus souvent associée à de la pyrrhotite (Fe_7S_8), de la pyrite (FeS_2) et de la chalcopyrite ($CuFeS_2$), et représente entre 0,7 et 3% du minerai. Mais on y trouve aussi du cuivre (environ 1%), des platinoïdes (Pt), du cobalt, de l'argent (Ag), et de l'or (Au), qui sont alors récupérés.

1.3.2 Les minerais oxydés

Ils représentent environ 35% de la production mondiale et sont exploités, contrairement aux minerais sulfurés, dans des mines à ciel ouvert.
Ces minerais ne contiennent pas de cuivre ni de métaux précieux, mais peuvent renfermer du cobalt. On doit en distinguer deux types :

- Les minerais silicatés (également appelés saprolites), comme par exemple la graniérite, sont principalement exploités en Nouvelle Calédonie et en Colombie. La graniérite est une serpentine ($3MgO,2SiO_2,2H_2O$) dans laquelle le magnésium a été partiellement remplacé par du nickel (Figure 1.3-1). Elle contient alors entre 2,3 et 3% de nickel. On y trouve également 10 à 30% de fer et de cobalt ;
- Les latérites nickélifères (également appelés limonites) sont exploitées en Nouvelle Calédonie, en Australie, à Cuba et en Grèce. Dans ce cas, le nickel se substitue au fer de la goethite (FeOOH). Ces minerais sont relativement pauvres en nickel : entre 1 à 1,5% seulement. En plus du fer (40 à 50%), on trouve également du cobalt (0,1 à 0,2%) et du chrome (2 à 5%).

1.3.3 Le cas de la Nouvelle Calédonie

La Nouvelle Calédonie présente la particularité d'être recouverte pour 1/3 par des roches ultramafiques (péridotites et serpentines : Figure 1.3-1), on y trouve donc une forte proportion de graniérite (Figure 1.3-1).

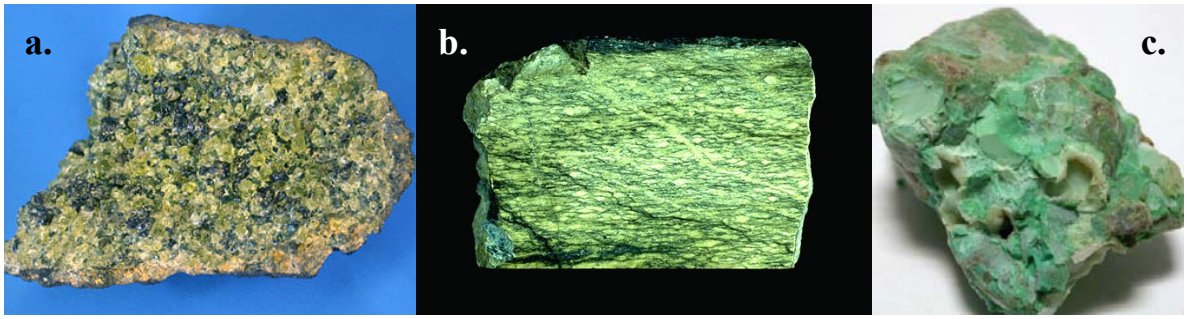


Figure 1.3-1 : La péridotite (a.) sous l'effet de la chaleur et de l'hydratation permet d'obtenir la serpentine (b.) qui donne, par remplacement de son magnésium par du nickel, la graniérite (c.) (http://www.musee.ensmp.fr/).

La Nouvelle Calédonie possède de ce fait 17% des réserves mondiales de nickel (chimie). Cette particularité fut découverte au XIXème siècle par Jules Garnier, qui déposa alors un brevet pour l'exploitation industrielle du nickel calédonien (Wikipédia). A l'heure actuelle, le territoire produit 6% de l'exploitation mondiale (Figure 1.3-2).

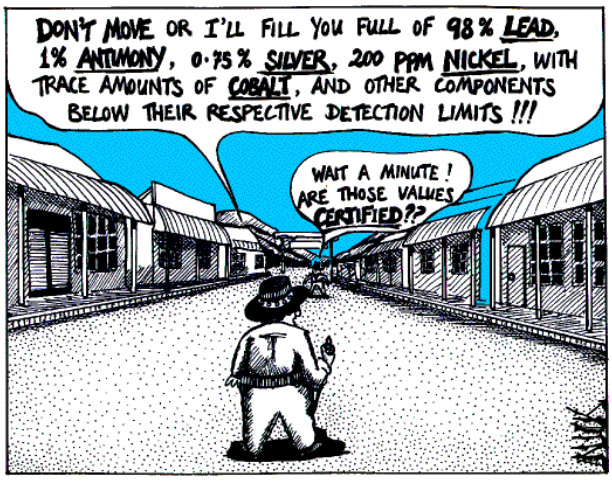
Figure 1.3-2 : Mines de nickel en Nouvelle Calédonie.

L'écosystème calédonien a dû s'adapter à ces sols particuliers présentant une haute « pollution » en métaux lourd (le taux de nickel étant en moyenne de 20 000 ppm), ainsi qu'à une déficience en éléments élémentaires tels que le carbone (C ; 0,2%), l'azote (N ; 0,01%),

ou le phosphore (P). Ces sols sont donc d'excellents modèles pour l'étude de l'adaptation d'organismes, tels que les bactéries et les plantes, à ce type de pollution (Hery *et al.*, 2003; Hery *et al.*, 2005).

1.4 L'utilisation industrielle du nickel

Le nickel est utilisé dans de nombreux domaines industriels concernant la fabrication d'aciers inoxydables jusqu'à celle des pièces de monnaie (chimie; Wikipédia).



ANALYTICAL CHEMISTS IN THE WILD WEST

copyright Nick Kim
http://strangematter.sci.waikato.ac.nz/

1.4.1 Les aciers inoxydables

L'acier est un alliage à base de fer et d'une proportion faible, mais variable, de carbone : entre 0,05 et 2% de la masse, en fonction des propriétés finales souhaitées. Mais au vu des propriétés redox du fer, l'acier a une tendance à s'oxyder. C'est pourquoi on ajoute du nickel, en plus du chrome, afin de le rendre inoxydable.

En fonction de la quantité de nickel ajouté, l'acier acquiert des propriétés caractéristiques :

- L'acier dit de construction voit sa résistance mécanique augmenter par l'ajout de Ni ;
- L'ajout de 9% de nickel permet à l'acier d'être non fragile à froid ;
- L'Invar possède, grâce à l'ajout de 36% de nickel, un coefficient de dilatation nul. C'est donc un matériau de choix lorsque que l'on veut être insensible aux écarts de température, comme par exemple pour la réalisation d'étalons secondaires de mesure, de thermostats à dilatation de solide, ou en horlogerie.

1.4.2 Le nickelage

Les objets, que l'on appelle chromés, sont en réalité essentiellement nickelés : l'acier est recouvert d'une couche de 20 à 30 µm de Ni, elle-même recouverte d'une fine pellicule de chrome de 0,2 à 0,3 µm, cette dernière uniquement destinée à en faciliter l'entretien.
Ces pièces chromées sont, pour une grande proportion, retrouvées dans les automobiles, ce qui explique que cette industrie utilise 6 à 8% du nickel consommé en Occident.

1.4.3 Les pièces de monnaie

Le nickel est souvent utilisé dans les pièces de monnaie, soit pur, soit sous forme d'alliage.
C'est en 1881 que fut créée la première pièce de monnaie entièrement réalisée en nickel : la pièce de 5 cents (USA). En Anglais, cette dernière est d'ailleurs encore appelée « nickel ».
En Europe, les pièces de 1 et 2 euros sont élaborées à partir d'alliages contenant du nickel.

1.4.4 D'autres exemples d'utilisation du nickel

De nombreux autres domaines utilisent le nickel, comme par exemple :

- Les batteries nickel-cadmium ;
- Les catalyseurs d'hydrogénation (Ni-Al) ;
- De nombreux alliages, comme les alliages Ni-Ti qui possèdent une mémoire de forme ;
- …

1.5 La toxicité du nickel

1.5.1 Les sources de contact avec le nickel et ses dérivés

Etant donnée l'importante utilisation par l'industrie du nickel, il existe une contamination non négligeable de l'environnement par ce dernier et ses dérivés. L'homme est de ce fait relativement exposé, et les sources de contact assez nombreuses.
La contamination la plus importante concerne les employés travaillant dans la métallurgie du nickel par inhalation et ingestion (IARC, 1990).
Le contact par voie cutanée concerne une population plus importante et moins spécialisé que dans le cas précédent, puisqu'elle se fait par le contact des pièces de monnaie et des objets d'usage courant contenant du nickel (Hostynek and Maibach, 2002).

L'activité urbaine, et notamment la circulation routière, provoque une augmentation de la quantité de nickel et de ses dérivés dans l'atmosphère : 120 à 170 ng/m^3, au regard des 6 à 17 ng/m^3 que l'on mesure en dehors des villes (Norseth and Piscator, 1979).
On peut par ailleurs rencontrer le nickel dans le domaine médical. Néanmoins la majorité des cas décrits sont relativement anciens. Deux cas de figure peuvent se présenter :

- Le contact par certaines prothèses contenant le métal (McGregor *et al.*, 2000) ;
- L'administration de médicaments contenant du nickel (Leach and Sunderman, 1985; Leach and Sunderman, 1987).

Chez l'homme, deux types principaux de problèmes médicaux peuvent apparaître, en fonction du type de contamination et de son importance.

1.5.2 Les réactions allergiques

Le nickel peut provoquer des réactions allergiques à la suite de contacts cutanés prolongés (bijoux, rivets, boutons de vêtements…). Elles se manifestent sous la forme de dermatites de contact, qui sont des inflammations locales de la peau, accompagnées de démangeaisons, de rougeurs, et de petites vésicules ou de papules (petits cercles rouges à centre blanc donnant des démangeaisons) sur la peau (CCHST; Cavani, 2005).

1.5.3 Le nickel est carcinogène

Les premiers cas de cancer des voies respiratoires potentiellement imputés au nickel ont été identifiés en 1933 ((Kasprzak *et al.*, 2003) pour une revue). Dans les années suivantes, un nombre important de cas on été décrits chez des personnes professionnellement exposées au nickel (Baader, 1937; NSA, 1975; Sunderman F.W. J.r, 2004).
Depuis, il a été clairement démontré que le nickel était carcinogène grâce à des études épidémiologique chez l'homme et des essais sur l'animal (IARC, 1990; Denkhaus and Salnikow, 2002; Sunderman F.W. J.r, 2004). Il a été montré que l'exposition chronique des travailleurs à des poussières et fumées contenant du nickel, provoquait une augmentation de la mortalité par tumeurs malignes des cavités nasales et des poumons.
Pendant longtemps, on a pensé que seules les formes insolubles du nickel (par exemple : Ni_3S_2, ou NiO) étaient carcinogènes. Mais des études plus récentes (Grimsrud *et al.*, 2003) ont montré que les formes solubles et les aérosols peuvent également induire des cancers des voies respiratoires (par exemple les sulfates de Ni(II)).

2 Le nickel chez les microorganismes

Le nickel est par ailleurs un cofacteur essentiel à l'activité enzymatique de protéines appartenant au monde procaryote mais aussi eucaryote. En effet, les enzymes possédant une activité dépendante du nickel sont retrouvées chez les bactéries, les archaebactéries, les champignons, les algues, les plantes supérieures, et quelques animaux (Hausinger, 1993). Dans le cadre de cette thèse, nous ne nous intéresserons qu'au cas des microorganismes. Dans la majorité des cas, c'est le Ni^{2+} (le nickel ayant l'état d'oxydation II) qu'ils utilisent.

2.1 Le rôle et l'utilisation du nickel chez les bactéries

C'est en 1965 que l'on mit pour la première fois en évidence le caractère indispensable du nickel pour deux souches d'*Hydrogenomonas* H1 et H16 (Bartha and Ordal, 1965). Par la suite, un nombre limité de protéines possédant une activité dépendante du nickel (une petite dizaine) ont été identifiées ((Maroney, 1999; Mulrooney and Hausinger, 2003) pour une revue). Néanmoins, certaines peuvent avoir un rôle central dans le métabolisme de l'organisme (comme par exemple l'uréase chez *Helicobacter pylori*). En plus de ces enzymes, il existe un nombre important de protéines impliquées dans l'incorporation du nickel dans les enzymes (par exemple : des protéines chaperonnes, ou les protéines impliquées dans le transport du nickel).

2.1.1 L'uréase

L'uréase catalyse l'hydrolyse de l'urée en ammoniaque et en carbamate. Ce dernier est alors spontanément décomposé en acide carbonique et ammoniaque (Figure 2.1-1 ; (Mobley *et al.*, 1995; Hausinger and Karplus, 2001) pour des revues).

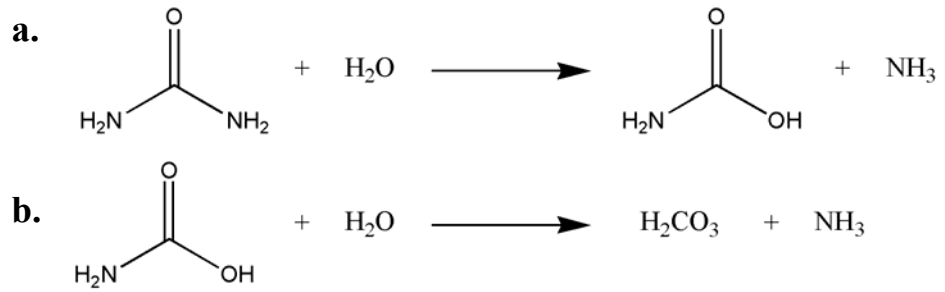


Figure 2.1-1 : a. Hydrolyse de l'urée par l'uréase en carbamate et ammoniaque ;
b. Décomposition du carbamate en acide carbonique et ammoniaque.

Il s'agit de la première enzyme identifiée comme ayant une activité dépendante du nickel (Dixon *et al.*, 1975). Toutes les uréases bactériennes, d'algues, de champignons et de plantes identifiées jusqu'à présent contiennent du nickel, qui semble avoir pour rôle la polarisation de la fonction carbonyle de l'urée (Wages *et al.*, 1993).

En fonction de l'organisme, l'uréase peut avoir deux rôles métaboliques :

- Dans la majorité des microorganismes, l'ammoniaque produit par les uréases sert de source d'azote, leur permettant ainsi de vivre sur des milieux contenant des précurseurs d'urée comme l'arginine ou les purines (Mobley and Hausinger, 1989) ;
- L'uréase peut aussi tenir le rôle de facteur de virulence. En effet, grâce à sa capacité à hydrolyser l'urée en ion ammonium, elle permet à *H. pylori* de survivre dans le système digestif humain. Bien qu'elle se développe dans la muqueuse de l'estomac, *H. pylori* est extrêmement sensible aux pH acides. Il lui est donc indispensable de neutraliser son environnement. Pour ce faire, elle utilise ainsi l'uréase, qui représente environ 6% des protéines solubles de la bactérie (Mobley, 1996). Cette bactérie est connue comme impliquée dans la genèse du cancer de l'estomac, et d'autres pathologies gastro-duodénales telles les gastrites chroniques, les ulcères duodénaux (B. J. Marshall et J. R. Warren : Prix Nobel de médecine 2005). *Proteus mirabilis* est un autre exemple de pathogène humain (urinaire cette fois) qui possède une uréase servant de facteur de virulence.

Du point de vue structural, les uréases peuvent être composées d'une, deux ou trois sous-unités, la structure quaternaire ainsi créée possédant deux atomes de nickel (Figure 2.1-2), coordinés par quatre histidines, une lysine et trois molécules d'eau.

- Dans le cas de *Klebsiella aerogenes* et *Bacillus pasteurii*, trois sous-unités existent : UreA (~ 11 kDa), UreB (~ 12-14 kDa), et UreC (~ 60-61 kDa). L'uréase ainsi formée va alors adopter une structure en trimère de trimères : $(UreABC)_3$ (Pearson *et al.*, 1998) ;
- Dans le cas de *H. pylori*, la sous-unité UreA (~ 26,5 kDa) remplace deux des sous-unités d'un trimère. L'uréase correspond donc à un dimère UreAB, qui, là aussi, va trimériser pour donner un complexe $(UreAB)_3$ (Ha *et al.*, 2001) ;

- Enfin, les uréases eucaryotes n'ont qu'une seule sous-unité (~ 90 kDa) correspondant à la fusion des sous-unités UreA, UreB, et UreC. Elles peuvent alors trimériser ou hexamériser (Takishima *et al.*, 1988).

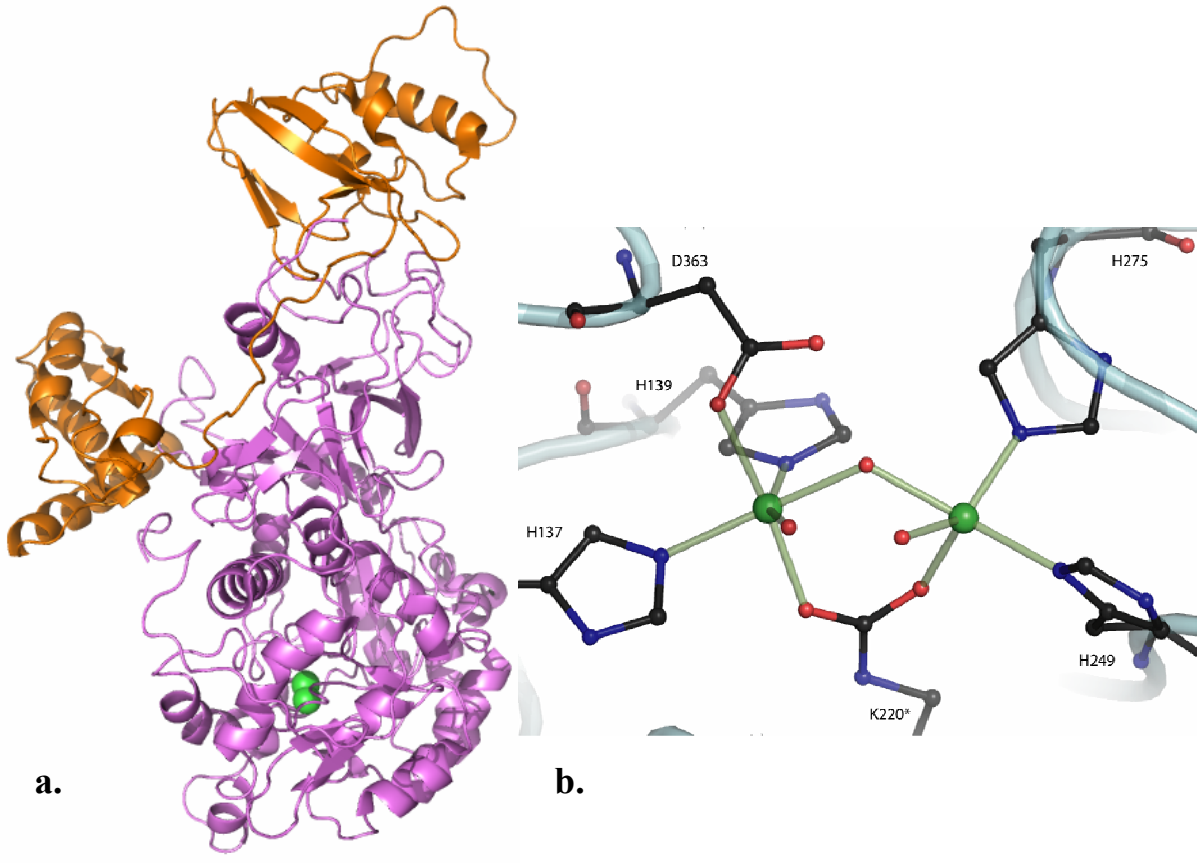


Figure 2.1-2 : a. Structure cristallographique de l'uréase de *H. pylori* (code PDB : 1E9Z). En orange : UreA ; En violet : UreB ; En vert : les deux atomes de nickel. b. Site actif de l'uréase de *B. pasteurii* (code PDB : 2UBP). K220* est une lysine carbamylée. En vert les deux atomes de nickel.

Dans tous les cas, les structures primaire et tertiaire sont conservées, et notamment le site actif possédant les deux nickels.

Des études cristallographiques ont montré que, dans le cas de *K. aerogenes*, la structure de la protéine holo (avec les deux nickels) et celle de l'apo protéine (sans les nickels) étaient identiques (Jabri and Karplus, 1996).

Un modèle de l'activation de l'uréase est présenté dans la Figure 2.1-3 (Mulrooney and Hausinger, 2003). Le trimère de la forme apo protéine de l'uréase $(UreABC)_3$ est associé à trois complexes multiprotéiques UreDFG, composés des protéines accessoires UreD, UreF et UreG (Park and Hausinger, 1995; Moncrief and Hausinger, 1997). Le nickel est alors délivré

par la protéine chaperonne UreE qui le fixe avec une stœchiométrie dépendant de l'organisme : par exemple, dans le cas de *K. aerogenes*, le dimère UreE fixe six nickels (Lee *et al.*, 1993), contre seulement un nickel par dimère dans le cas de *B. pasteurii* (Ciurli *et al.*, 2002).

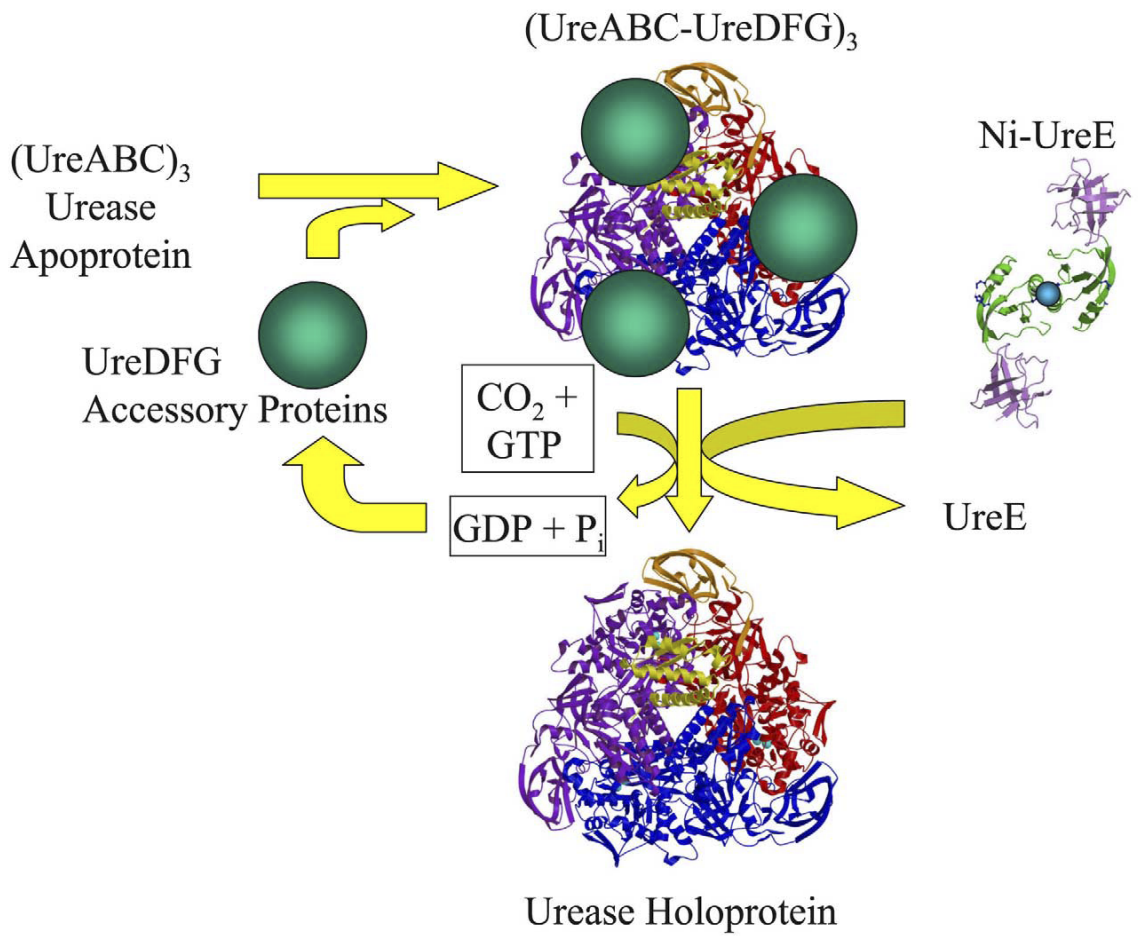


Figure 2.1-3 : Modèle de l'activation de l'uréase (Mulrooney and Hausinger, 2003).
La structure de l'uréase est celle de *K. aerogenes* (code PDB : 1FWJ) composée de trois sous-unités ici représentées en orange, rouge et jaune. Les deux autres uréases du trimère sont colorées en bleu et violet.
La structure de la métallo-chaperonne UreE est celle de *K. aerogenes* (code PDB : 1CFZ). Les deux domaines sont colorés en mauve et vert.
Le nickel est représenté par une sphère bleue.

2.1.2 L'hydrogénase à NiFe

Les hydrogénases ont été mises en évidence en 1931 (Stephenson, 1931). Elles sont présentes dans de nombreux organismes, tels que les procaryotes aérobies, anaérobies stricts

ou facultatifs (Adams *et al.*, 1981), et certains eucaryotes comme par exemple des algues unicellulaires (Healey, 1970) ou des protozoaires (Payne *et al.*, 1993).
En fonction de l'organisme, elles peuvent permettre la production ou la consommation d'hydrogène moléculaire (Figure 2.1-4).

$$H_2 \rightleftharpoons 2H^+ + 2e^-$$

Figure 2.1-4 : Réaction réversible de réduction de l'hydrogène moléculaire par l'hydrogénase.

Dans le cas où les organismes utilisent l'hydrogène dans le sens de sa consommation, la production de deux électrons de bas potentiel permet par la suite la réduction de différents substrats (comme le CO_2) et la production d'énergie nécessaire à la synthèse d'ATP.
Dans le cas de l'utilisation de l'hydrogénase comme producteur d'hydrogène moléculaire, celle-ci permet d'éliminer l'excès de pouvoir réducteur généré par le métabolisme. On trouve ce cas de figure pour certains microorganismes anaérobies utilisant la fermentation de substrats organiques comme source de réducteurs riches en énergie pour la respiration anaérobie (pour revue (Pelmont, 1993)).
Il existe trois classes d'hydrogénases :

- Les hydrogénases à nickel-fer (NiFe), qui constituent la classe la plus représentée et la plus étudiée. On trouve aussi des hydrogénases à nickel-fer-sélénium (NiFeSe) chez certains organismes comme par exemple *Desulfomicrobium baculatum* (Eidsness *et al.*, 1989) ;
- Les hydrogénases à fer (Fe) ;
- Les hydrogénases sans FeS : elles étaient précédemment connues comme ne contenant pas de métal, pourtant, Lyon *et al*. ont montré qu'elles possédaient du fer (Lyon *et al.*, 2004). Elles sont donc maintenant appelées « sans FeS » car elles ne contiennent pas de centres FeS, contrairement aux hydrogénases à fer.

Plusieurs structures d'hydrogénases à NiFe (et une à NiFeSe) ont été résolues par cristallographie, et notamment la première d'entre elles, résolue au laboratoire, qui provient de *Desulfovibrio gigas* (Volbeda *et al.*, 1995; Volbeda *et al.*, 1996). Elles possèdent un ion nickel par molécule, qui sert de site de fixation du substrat, coordiné par quatre cystéines, ou trois cystéines et une sélénocystéine dans le cas des hydrogénases à NiFeSe.

Elles possèdent en général au moins deux sous-unités formant un hétérodimère (Figure 2.1-5). La petite sous-unité (SS : *Small Subunit*) et la grande sous-unité (LS : Large Subunit) ont respectivement des masses de 23 à 38 kDa, et de 46 à 72 kDa. En plus du centre NiFe, les hydrogénases contiennent des atomes de Fe arrangés en un nombre variable d'agrégats [4Fe4S] ou [3Fe4S].

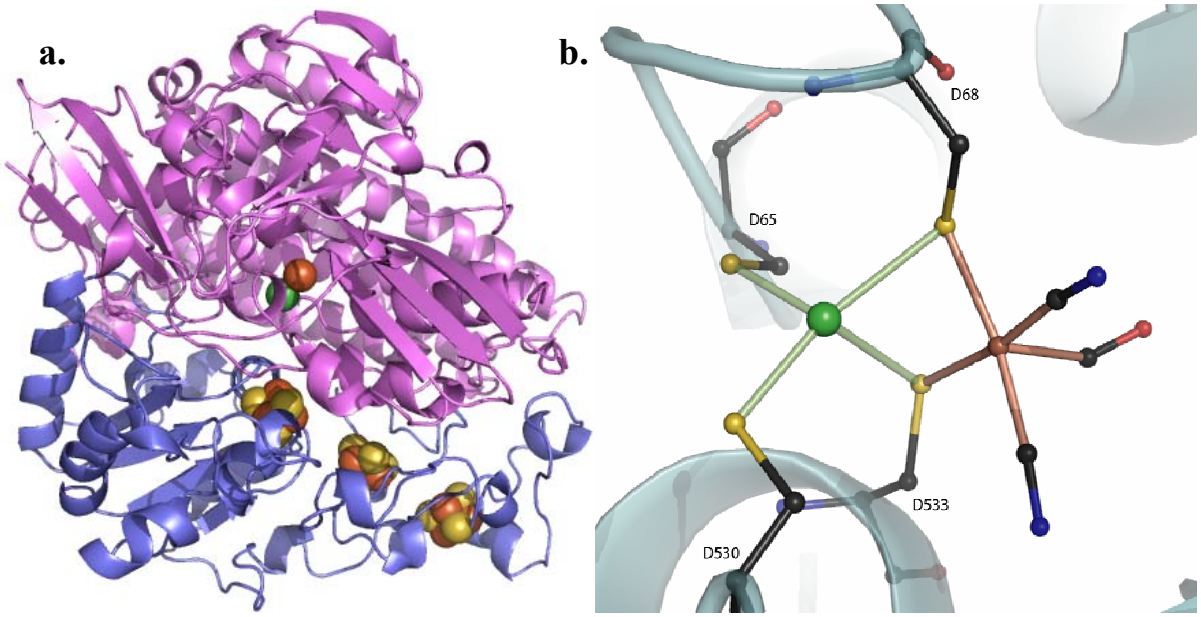


Figure 2.1-5 : Structure de l'hydrogénase à NiFe de *D. gigas* (code PDB : 2FRV).
a. Structure cristallographique de l'hydrogénase à NiFe de *D. gigas.* La petite sous-unité est représentée en bleu. La grande sous-unité est représentée en violet. A l'interface des deux sont représentés les agrégats [4Fe4S] et [3Fe4S]. Le nickel et le fer sont respectivement représentés par des sphères vertes et rouges.
b. Site actif de l'hydrogénase à NiFe de *D. gigas.*

A l'image de la maturation de l'uréase, celle de l'hydrogénase fait intervenir de nombreuses protéines. La situation est même plus complexe puisque dans ce cas deux métaux doivent être incorporés : le fer et le nickel (Mulrooney and Hausinger, 2003).

Les protéines chaperonnes HypC et HypD (Figure 2.1-6) forment avec le fer un complexe qui va être reconnu par le dimère composé des protéines HypE et HypF (ou leur équivalents) (Drapal and Bock, 1998; Blokesch *et al.*, 2001). Sous l'action de ces deux protéines et d'une hydrolyse d'ATP, le fer va acquérir un CO et deux CN (ici représentés par des traits), constitutifs du site actif. La grande sous-unité de l'hydrogénase est dans un premier temps synthétisée avec une extension en C terminal (Pre-LS) (Rodrigue *et al.*, 1996). La fixation de HypC est supposée provoquer un changement de conformation de Pre-LS, ce qui permet le

chargement du fer et ses trois ligands. Le nickel est ensuite incorporé par le complexe composé des protéines HypA et HypB ou leurs homologues. Cette étape est couplée à une hydrolyse de GTP (Maier *et al.*, 1993; Fu *et al.*, 1995; Maier *et al.*, 1995). Le départ de HypC permet alors à une protéase spécifique nickel dépendante d'enlever le domaine C terminal (ici est représentée celle d'*Escherichia coli*. Code PDB : 1CFZ (Fritsche *et al.*, 1999)), et ainsi d'obtenir la grande sous unité maturée (Magalan and Bock, 2000). L'ajout de la petite sous-unité (SS) permet d'obtenir une hydrogénase activée.

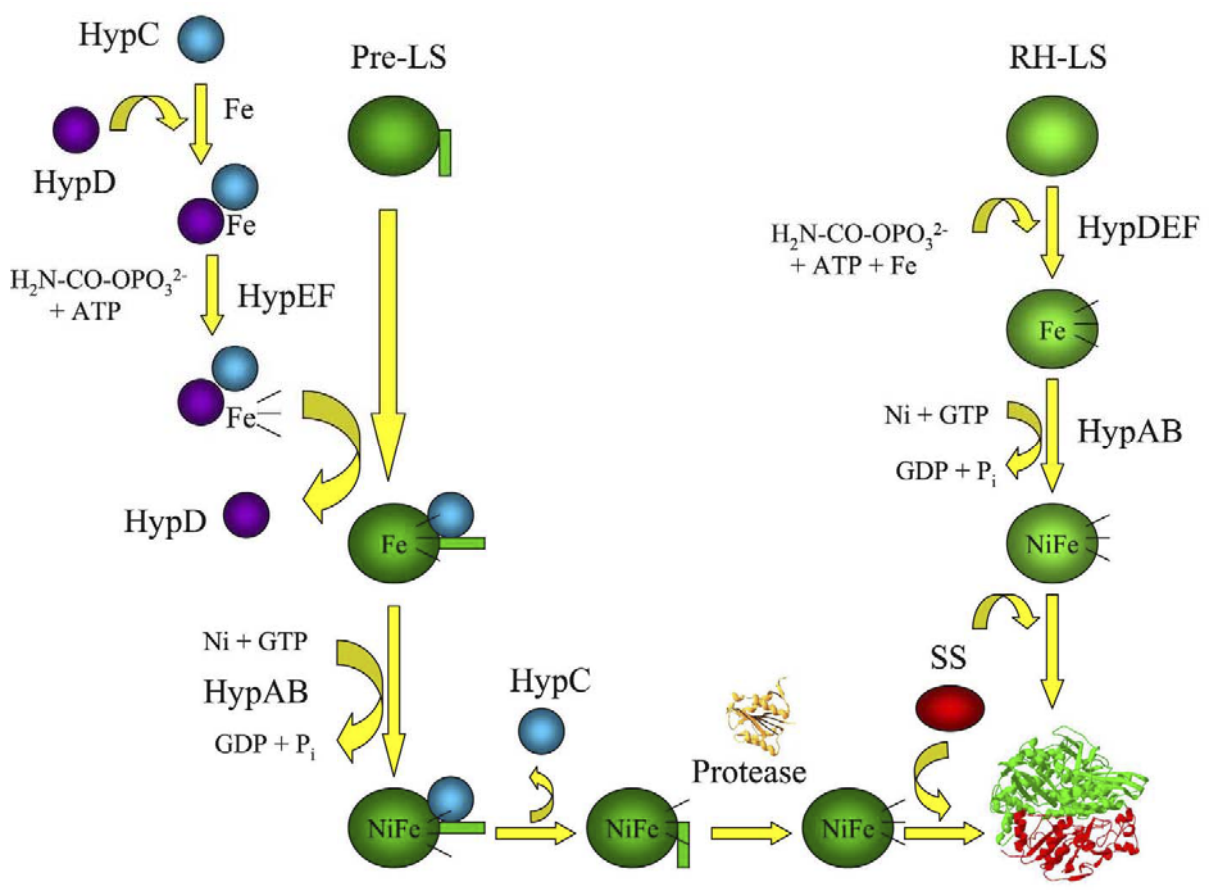


Figure 2.1-6 : Schéma de la maturation de l'hydrogénase à NiFe et des senseurs à hydrogène (Mulrooney and Hausinger, 2003).

Les capteurs à hydrogène, que l'on appelle aussi hydrogénases régulatrices, sont des hydrogénases capables de fixer l'hydrogène moléculaire, mais qui ne présentent par contre qu'une faible activité d'oxydation ou de production de l'hydrogène moléculaire. Elles possèdent cependant une activité de remplacement de l'hydrogène par du deutérium (H/D) (Vignais *et al.*, 2000). Leur activation est plus simple que pour les hydrogénases (Figure 2.1-6). La grande sous-unité est synthétisée sans l'extension N terminale (RH-LS). Le fer et

ses trois ligands sont incorporés à RH-LS en une seule étape par le complexe HypDEF, couplé à une hydrolyse d'ATP (Colbeau *et al.*, 1998). Une seconde étape couplée à l'hydrolyse de GTP permet l'ajout du nickel au site actif grâce au complexe HypAB. Il suffit alors d'ajouter la petite sous-unité pour avoir un capteur à hydrogène activé.

Dans le cas d'*E. coli*, il existe trois hydrogénases clairement identifiées codées par les gènes *hyaAB*, *hybOC*, et *hycGE*. Un quatrième groupe de gènes a été identifié, mais la protéine produite n'a pas d'activité hydrogénase détectable (Skibinski *et al.*, 2002).

2.1.3 La monoxyde de carbone déshydrogénase (CODH)

L'activité CODH est indispensable aux organismes se servant du monoxyde de carbone comme source d'énergie. Elle catalyse la réaction réversible d'oxydation du monoxyde de carbone en dioxyde de carbone (Figure 2.1-7).

$$CO + H_2O \rightleftharpoons CO_2 + 2e^- + 2H^+$$

Figure 2.1-7 : Réaction catalysée par la monoxyde de carbone déshydrogénase (CODH).

Les structures de la CODH de *Carboxythermus hydrogenaformans* (Dobbek *et al.*, 2001) et de *Rhodospirillum rubrum* (Drennan *et al.*, 2001) sont connues. Dans les deux cas, il s'agit d'un homodimère contenant cinq centres [4Fe-4S] : deux agrégats B classiques, un agrégat D partagé par les deux sous-unités, et les deux sites catalytiques C qui consistent chacun en un centre [1Ni-3Fe-4S] associé à un site de fer mononucléaire. Le pontage entre le fer seul et le nickel semble se réaliser par l'intermédiaire, soit d'un soufre provenant d'une cystéine, soit d'un sulfide.

La maturation de la protéine nécessite l'intervention de protéines chaperonnes : CooS, CooC, et CooJ. La réaction est couplée à une hydrolyse d'ATP et il semble que ce soit CooJ qui délivre le nickel.

2.1.4 L'acétyl-coenzyme A synthase (ACS)

L'acétyl-coenzyme A synthase (ACS) catalyse la synthèse d'un groupement acétyle à partir d'un groupement méthyle (Figure 2.1-8). La réaction inverse est également possible.

$$CoA\text{-}S^- + CH_3\text{-}Co(III) + CO \rightleftharpoons Co(I) + H_3C-\overset{\overset{\displaystyle O}{\|}}{C}-S-CoA$$

Figure 2.1-8 : Réaction catalysée par l'activité Acétyl-coenzyme A synthétase.
Protéine fer-soufre corrinoïde : Co(I) – FeSP.

L'ACS possède un site actif contenant deux nickels adoptant tous les deux une géométrie plan carré. Le premier est fixé à la protéine par l'intermédiaire des soufres de deux cystéines et de deux azotes de la chaîne principale. La coordination du second nickel implique également les soufres deux cystéines précédemment évoquées, ainsi que celui d'une troisième, également impliquée dans la fixation d'un agrégat [4Fe4S] (Dobbek *et al.*, 2001; Drennan *et al.*, 2001; Darnault *et al.*, 2003), ce dernier devant servir d'accepteur des électrons produits par la réaction. La coordination plan carré est complétée par un quatrième ligand non identifié appartenant au solvant.

Dans le cas des microorganismes méthanogènes et acétogènes (Ferry, 1995; Ragsdale and Kumar, 1996; Lindahl, 2002), l'ACS n'est pas indépendante, mais associée à la CODH, formant ainsi un complexe multiprotéique.

Dans les organismes dégradant l'acétate, ou méthanogènes, tels que *Methanosarcina barkeri*, ou *Methanosarcina thermophila*, la réaction catalysée par l'activité ACS du complexe multiprotéique va dans le sens du transfert et la coupure de l'acétyle de l'acétylcoenzyme A en CH_3 et CO (Figure 2.1-8). Le CO ainsi produit est pris en charge par l'activité CODH (Figure 2.1-7) pour donner du CO_2. Par la suite, le méthyle transféré à la protéine fer-soufre corrinoïde est transformé en méthane. Dans ce cas, le complexe multiprotéique est $\alpha_2\beta_2\gamma\delta\varepsilon$ (avec α la CODH, β l'ACS, γ et δ deux corrinoïdes, et ε une protéine inconnue).

Par contre, chez les acétogènes tels que *Clostridium thermoaceticum*, c'est la situation inverse qui est rencontrée. Le CO_2 est réduit en CO par l'activité CODH (Figure 2.1-7) de la protéine, puis l'activité ACS catalyse la formation d'acétyl-coenzyme A (CH_3CO-S-CoA). Dans ce cas, le complexe multiprotéique est un hétérotétramère $\alpha_2\beta_2$ (avec α l'ACS et β la CODH), les corrinoïdes étant alors des protéines libres.

Dans les deux cas, on a suspecté l'existence d'un tunnel pour le CO, entre les deux sites contenant les nickels des activités CODH et ACDS (Seravalli and Ragsdale, 2000). La structure de la protéine de *Moorella thermoacetica* confirme cette hypothèse en montrant un réseau de canaux.

2.1.5 La méthyl-coenzyme M réductase

Les archaebactéries méthanogènes et les organismes utilisant comme source de nutriments l'acétate, le méthanol, la méthylamine, le formate et le dioxyde de carbone, produisent du coenzyme M méthylé (CH_3-S-CoM). La méthyl-coenzyme M réductase

catalyse la formation de méthane et d'un CoB-S-S-CoM à partir de coenzyme B et de méthyl coenzyme M (Figure 2.1-9). Par la suite, le CoB-S-S-CoM est réduit par l'hétérodisulfite réductase.

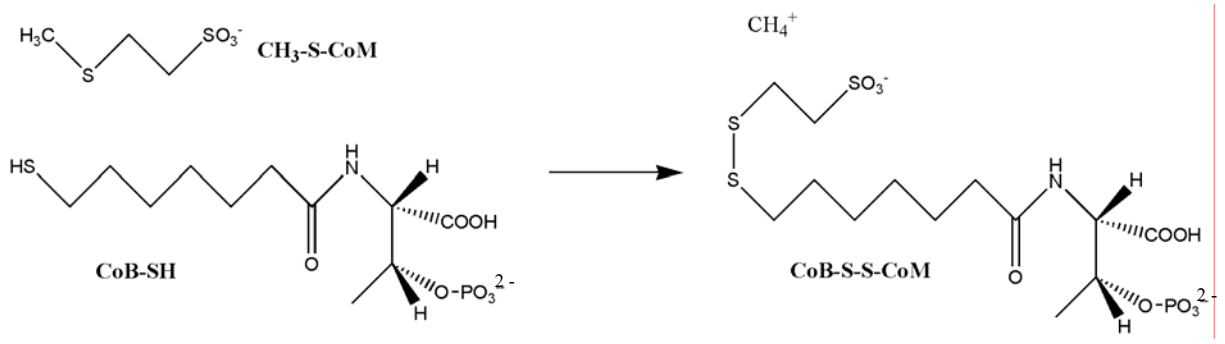


Figure 2.1-9 : Réaction catalysée par la méthyl-coenzyme M réductase.

Les structures cristallines de la protéine (Ermler *et al.*, 1997; Grabarse *et al.*, 2000; Grabarse *et al.*, 2001) montrent qu'elle est du type $\alpha_2\beta_2\gamma_2$ ($\alpha \sim 60$ kDa, $\beta \sim 46$ kDa, $\gamma \sim 28$ kDA). Le nickel est fixé dans le site actif de la protéine par l'intermédiaire d'un tétrapyrrole, le F_{430} (Figure 2.1-10). Ce dernier est fortement associé à la protéine mais de façon non covalente. Dans la protéine, le nickel est fixé par les quatre azotes du F_{430} (formant ainsi un plan carré), plus une arginine comme ligand axial. Le sixième ligand axial est donc libre pour interagir avec le CH_3-S-CoM. La forme active de la protéine est celle contenant un Ni(I).

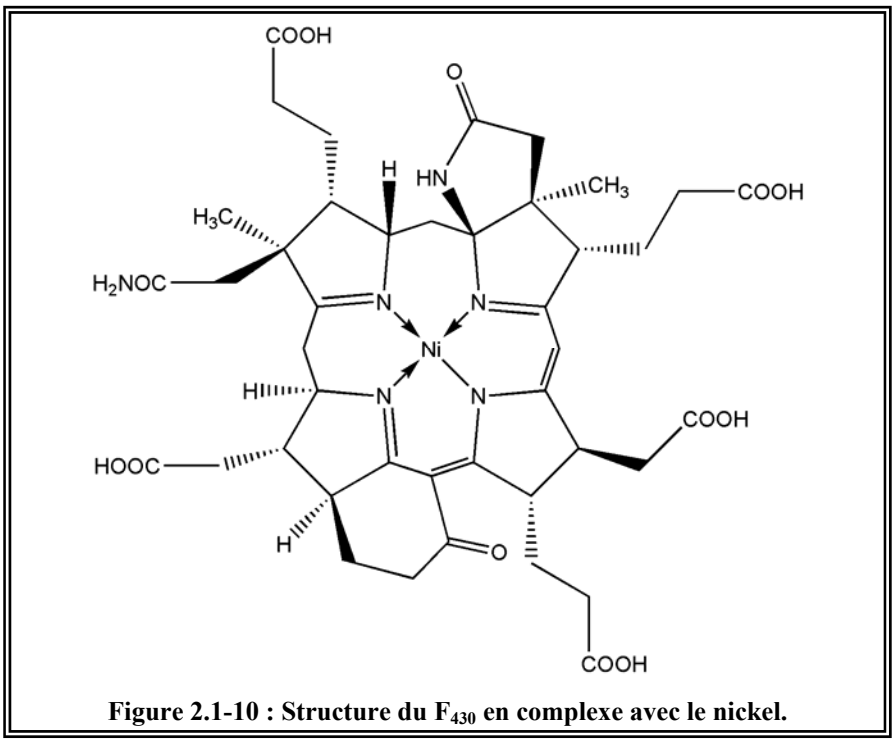


Figure 2.1-10 : Structure du F_{430} en complexe avec le nickel.

2.1.6 La superoxyde dismutase Ni-dépendante (SOD)

La SOD a pour rôle de protéger la cellule de la toxicité des radicaux superoxydes grâce à leur réduction en peroxyde d'hydrogène (Figure 2.1-11).

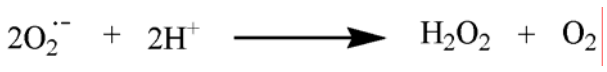


Figure 2.1-11 : Réaction catalysée par la superoxyde dismutase.

Il existe différentes familles de SOD qui contiennent, soit un centre dinucléaire de Cu ou de Zn, soit un centre mononucléaire de Fe, de Mn ou de Ni.

Les SOD nickel-dépendantes se retrouvent dans plusieurs souches de *Streptomyces* (Youn *et al.*, 1996; Youn *et al.*, 1996). Il est à noter qu'elles ont été retrouvées chez tous les *Streptomyces* isolés en milieu clinique et à partir de sols (Leclere *et al.*, 1999).

La première structure (Wuerges *et al.*, 2004) d'une SOD nickel-dépendante, montre que la forme biologique est un hexamère, avec un ion Ni(III) par molécule.

2.1.7 La glyoxylase Ni-dépendante

De nombreuses voies métaboliques produisent du méthylglyoxale. Cette espèce est toxique pour les cellules puisqu'elle peut réagir de façon covalente avec les protéines ou l'ADN. Un des systèmes permettant d'éliminer le méthylglyoxal utilise la glyoxylase I (Glx I) et la glyoxylase II (Glx II). Ce système présent dans de très nombreux organismes concernant la bactérie jusqu'à l'homme fonctionne grâce à une molécule de glutathion (G-SH) (Figure 2.1-12).

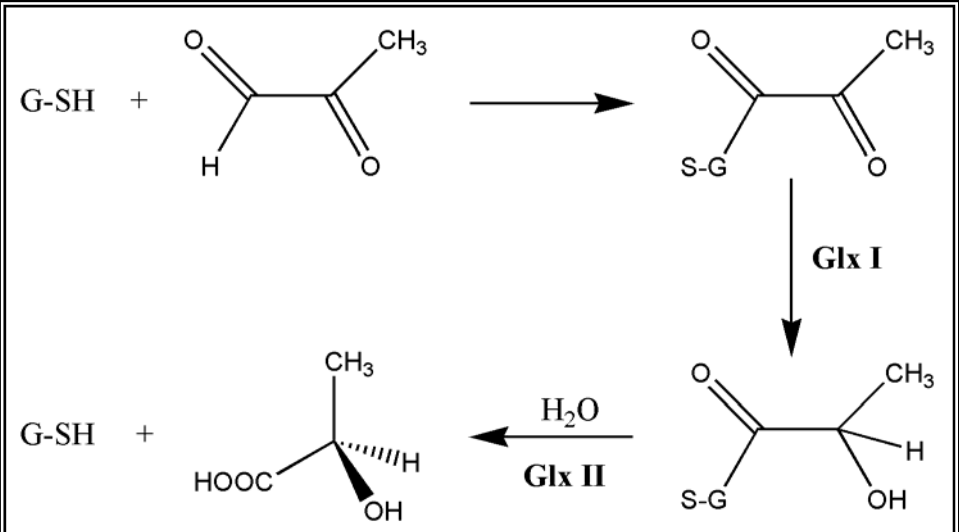


Figure 2.1-12 : Réaction catalysée par la glyoxylase.

La Glx I est classiquement une enzyme à activité zinc-dépendante. Il a été de ce fait surprenant de découvrir que celle d'*E. coli* fixait préférentiellement le nickel et avait un grand pourcentage d'identité avec la glyoxylase à zinc (Clugston *et al.*, 1998). La structure de la protéine à nickel (He *et al.*, 2000) montra que sa coordination octaédrique est similaire à celle du zinc de l'enzyme humaine : deux glutamates, deux histidines et deux molécules d'eau.
Par ailleurs, *Yersinia pestis* possède également un Glx I à activité nickel-dépendante (Clugston and Honek, 2000).

2.1.8 L'aci-réductone dioxygénase

De nombreux organismes recyclent la méthylthioadénosine produite durant la biosynthèse de polyamine à partir de la S-adénosylméthionine, en reformant l'acide aminé original. Cette voie métabolique passe par l'aci-réductone (1,2-dihydroxy-3-kéto-5-méthylthiopentane). L'aci-réductone dioxygénase va catalyser deux types de réactions en fonction du métal présent dans la structure (Figure 2.1-13) (Dai *et al.*, 1999).

- La protéine contenant du fer a une activité dite productive. Elle catalyse la formation d'acide formique et d'α-ketoacide, précurseur de la méthionine ;
- La protéine contenant du nickel a une activité dite non productive. Elle catalyse alors la formation de formiate, de monoxyde de carbone et de méthylthiopropionate.

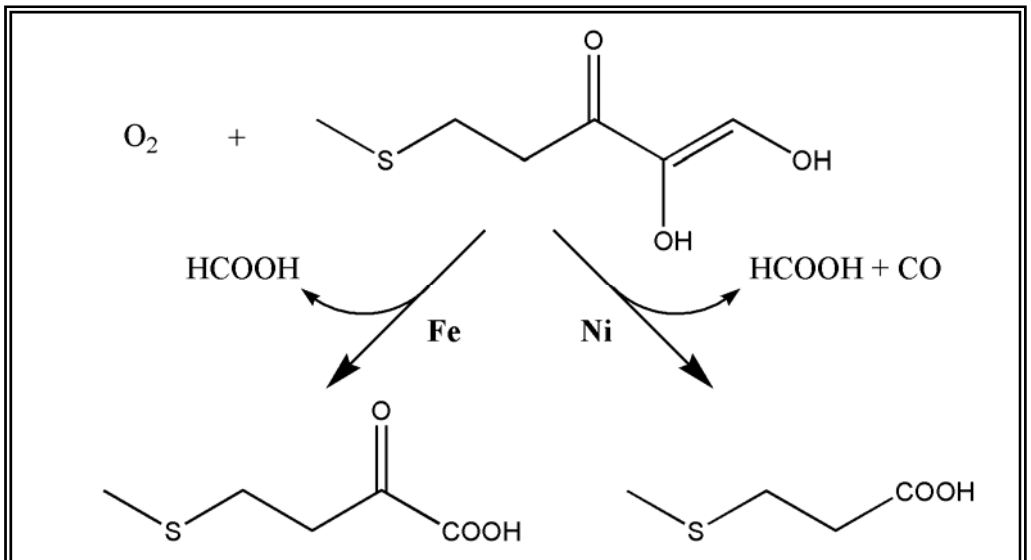


Figure 2.1-13 : Réaction catalysée par l'aci-réductone dioxygénase

Des études biophysiques montrent que le nickel est lié à la protéine par trois chaînes latérales d'histidines et trois autres donneurs oxygène ou azote (Al-Mjeni *et al.*, 2002).

2.1.9 D'autres enzymes à activité nickel-dépendante

Les méthylène-urées sont des molécules de formule générale NH_2-(CO-NH-CH_2-NH)$_i$-CO-NH_2 utilisées comme engrais. La méthylène-diuréase, isolée des *Brukholderia*, est une enzyme fixant un équivalent de nickel capable de dégrader les méthylène-urées où i = 1. Elle est composée de deux sous-unités, dont la plus grande ressemble aux uréases. De plus, les résidus impliqués dans la fixation du nickel chez les uréases, sont conservés chez la méthylène-diuréase. Ces résultats n'ont été reportés que par *Mulrooney et al.* dans sa revue (Mulrooney and Hausinger, 2003) provenant d'une communication personnelle : il convient donc d'utiliser ces résultats avec précautions.

De nombreuses métalloenzymes peuvent fixer un nickel à la place du métal physiologique. Il faut donc être prudent avant de classer une protéine comme ayant une activité dépendante du nickel.

2.2 Le transport du nickel chez les bactéries

Nous venons de voir que des processus clés du métabolisme de plusieurs bactéries dépendent de la présence de nickel dans leur cytoplasme. Il est donc indispensable pour ces dernières de disposer de systèmes de transport du métal ((Eitinger and Mandrand-Berthelot, 2000; Mulrooney and Hausinger, 2003) pour des revues).

2.2.1 L'import du nickel

2.2.1.1 Le passage à travers la membrane externe

Il existe principalement deux familles de bactéries :

- Les Gram positives (Gram (+)), possèdent, en plus de leur membrane plasmique, une paroi épaisse de peptidoglycane. L'espace entre ces deux couches est appelé périplasme (Figure 2.2-1a) ;
- Les Gram négatives (Gram (-)), possèdent deux membranes plasmiques (une interne et une externe). Là encore l'espace entre ces deux dernières est appelé périplasme. Ces bactéries possèdent une paroi de peptidoglycane moins épaisse que les Gram (+) (Figure 2.2-1b).

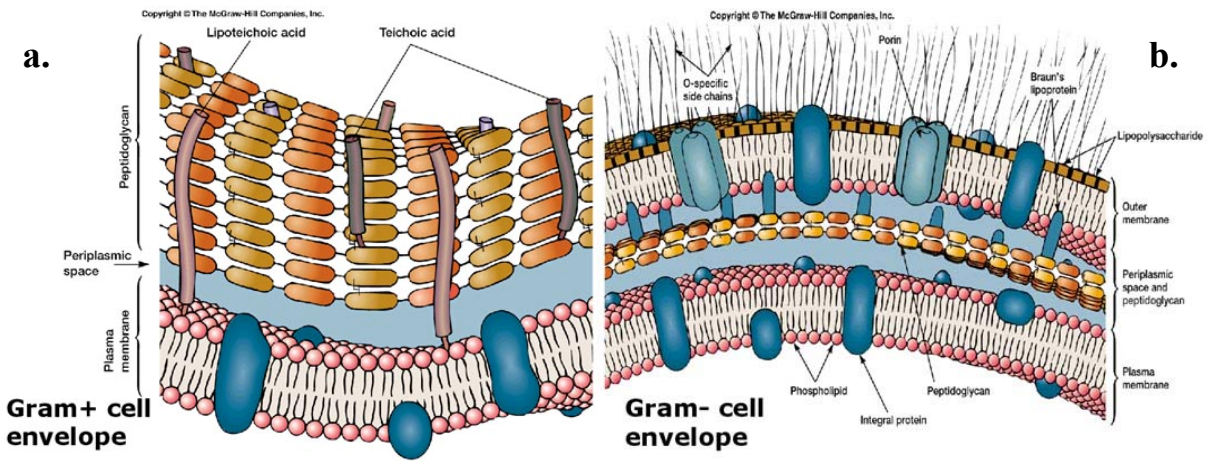


Figure 2.2-1 : Schéma des enveloppes des bactéries a. Gram (+) ; b. Gram (-) (http://faculty.ircc.edu/).

Dans le cas des bactéries Gram (-) la première étape du processus d'import du nickel est le passage de la membrane externe. Cette membrane jouant le rôle principal de barrière perméable à certains types de molécules, on peut se demander quel est le système permettant le passage du nickel. Bien que ce dernier n'ait pas encore été identifié, on peut suspecter, dans le cas d' *E. coli*, l'implication de porines (Figure 2.2-2).

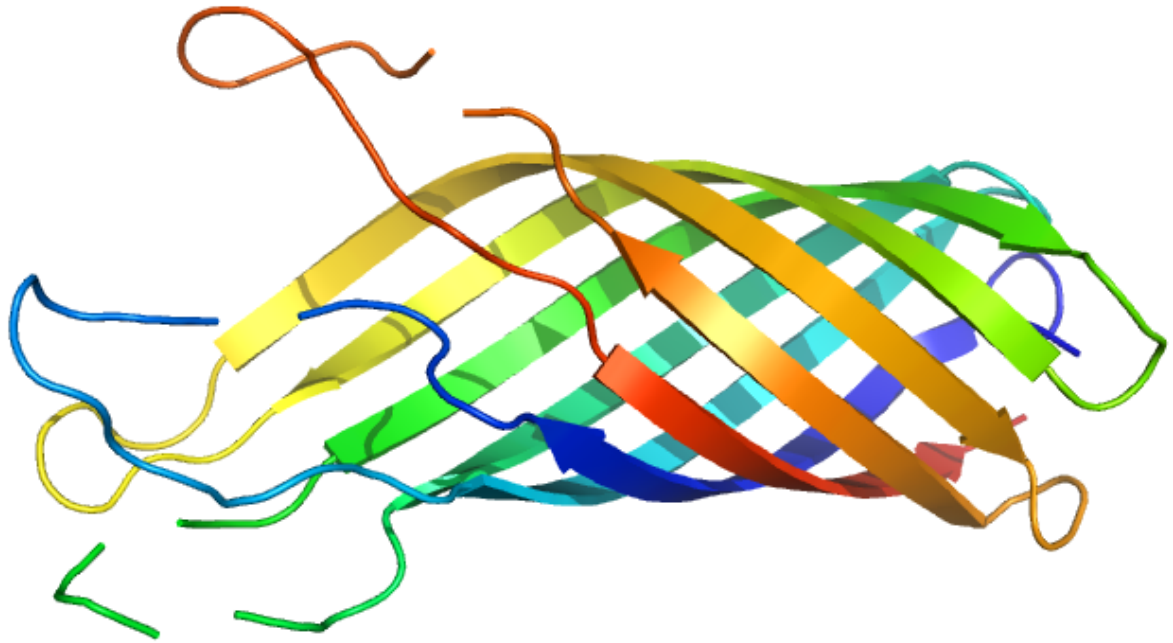


Figure 2.2-2 : Structure de la porine OmpA d'*E. coli* (code PDB : 1BXW) (Pautsch and Schulz, 1998).

2.2.1.2 Le passage à travers la membrane interne : le système du type ABC

Les transporteurs ABC (*ATP Binding Casette*) sont composés de cinq domaines (Figure 2.2-3) :

- Un domaine périplasmique ;
- Deux domaines transmembranaires formant un pore à travers la membrane interne ;
- Deux domaines cytoplasmiques associés à la membrane plasmique capables d'hydrolyser l'ATP, apportant ainsi l'énergie nécessaire au transport.

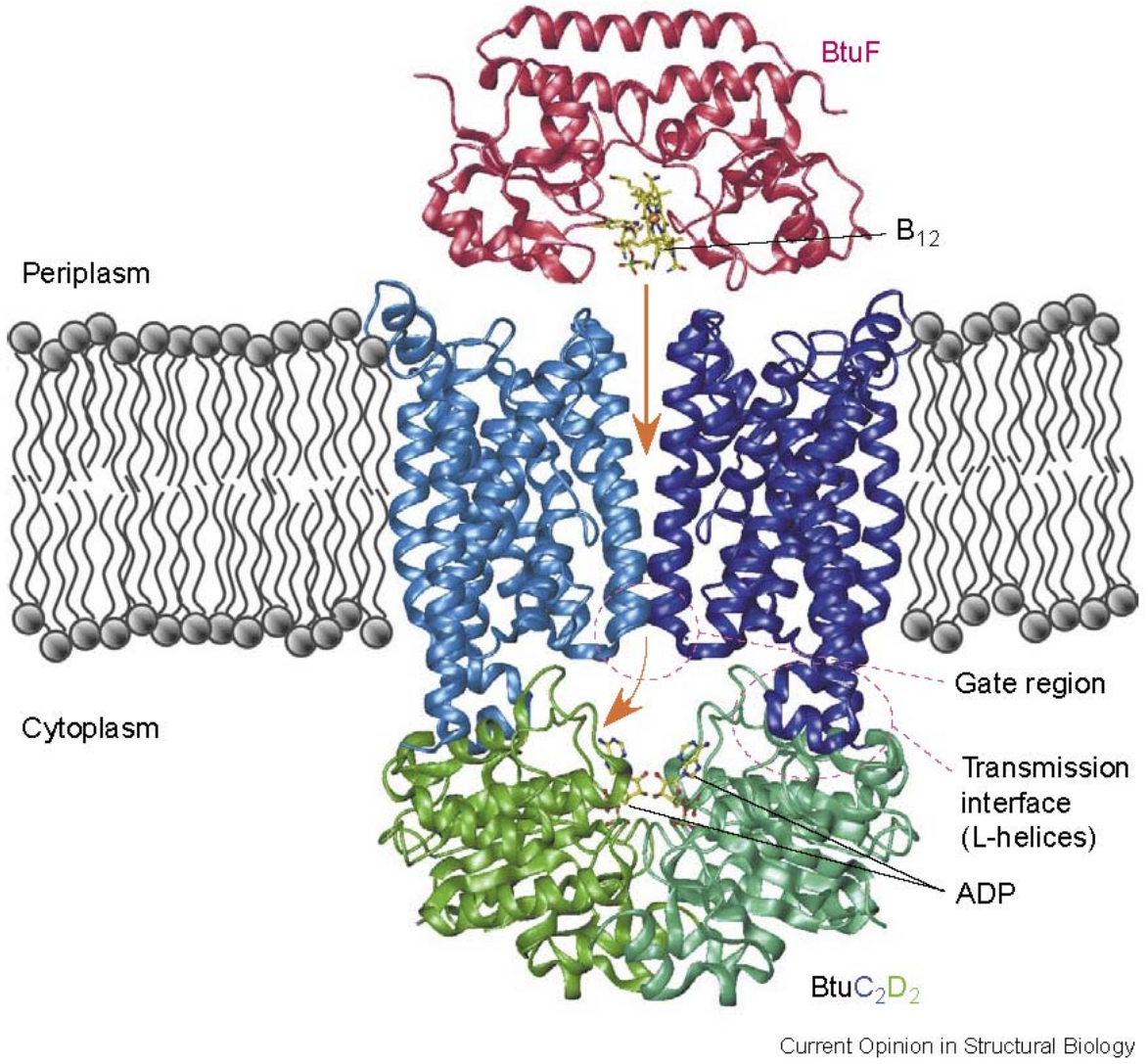


Figure 2.2-3 : Structure du premier système ABC (sous famille 8) entièrement cristallisé : le transporteur de la vitamine B_{12} d'*E. coli* (Locher, 2004).

Ils peuvent être classés en dix sous-familles sur la base d'une comparaison de séquence des domaines cytoplasmiques (Linton and Higgins, 1998).

2.2.1.2.1 Le système codé par l'opéron nikABCDE

Wu *et al.* ont décrit et caractérisé, en 1986, un mutant d'*E. coli* (Nik^-) ayant perdu son activité hydrogénase (Wu and Mandrand-Berthelot, 1986). En cultivant ce mutant en milieu riche, il a été surprenant de découvrir qu'il présentait un taux de Ni^{2+} cellulaire très faible (moins de 1% du taux mesuré dans la souche sauvage) (Wu *et al.*, 1989). Finalement, des expériences d'import de nickel radioactif ($^{63}Ni^{2+}$, émettant un rayonnement β^- de basse énergie (67 keV), avec une demi-vie de 100 ans, pour donner un atome stable de ^{63}Cu), ont montré que le mutant Nik^- était déficient en un système de transport du nickel codé par le locus *nik* (Navarro *et al.*, 1993).

Le clonage et le séquençage du locus *nik* (Wu *et al.*, 1991) mirent en évidence qu'il correspond à un opéron *nikABCDE* codant pour cinq protéines (NikA, NikB, NikC, NikD et NikE) possédant des similarités de séquence avec les protéines impliquées dans le transport de dipeptides et d'oligopeptides de plusieurs bactéries Gram (-) et Gram (+), comme par exemple *Salmonella typhimurium*, *Bacillus subtilis* ou *Streptococcus pneumoniae* (Navarro *et al.*, 1993).

Il s'avère que le système NikABCDE appartient à la sous-famille 2 des transporteurs ABC (Linton and Higgins, 1998), et se compose, de ce fait, de cinq protéines (Figure 2.2-4) :

- La protéine NikA (502 résidus) constitue la partie périplasmique du système de transport que nous décrirons plus en détail dans le chapitre 2.2.1.2.2 de l'introduction ;
- NikB (314 résidus) et NikC (277 résidus) sont deux protéines hydrophobes associées à la membrane, composées de six régions transmembranaires chacune. Elles formeraient un pore permettant le passage du nickel à l'intérieur des cellules ;
- Enfin, NikD (253 résidus) et NikE (268 résidus) sont deux protéines cytoplasmiques associées à la membrane plasmique qui possèdent les motifs conservés A et B de Walker (Walker A : GXXGXGKST ; Walker B : hhhhDEPT, h étant un résidu hydrophobique) (Linton and Higgins, 1998). Ces derniers sont impliqués dans la fixation et l'hydrolyse de l'ATP, apportant ainsi l'énergie indispensable au transport.

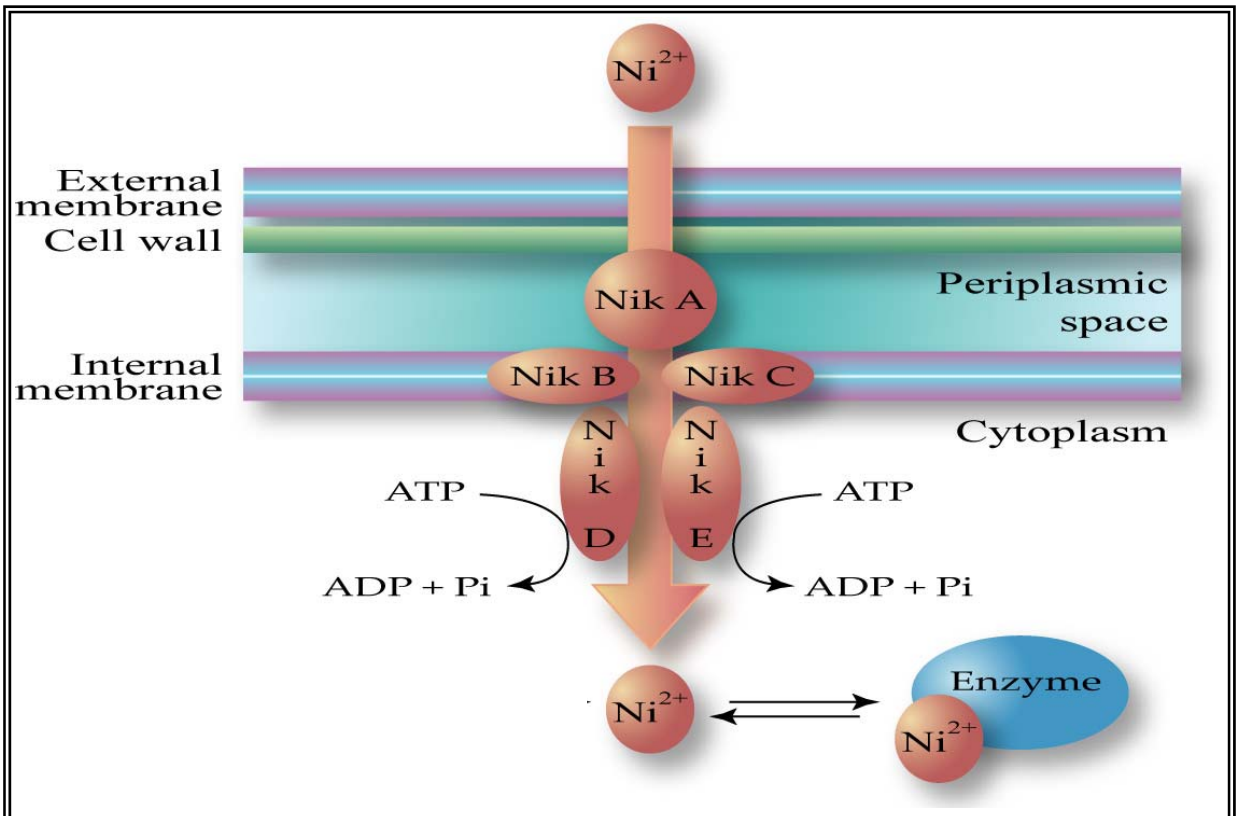


Figure 2.2-4 : Système de transport du nickel, du type ABC, codé par l'opéron *nikABCDE* d'*E. coli*.

Des transporteurs du nickel appartenant à la famille des transporteurs ABC du même type que celui décrit plus haut se retrouvent dans d'autres organismes, assurant seuls ce transport, ou le partageant avec des perméases aux métaux de transition (chapitre 2.2.1.3 de l'introduction).

A titre d'exemple, *Yersinia pseudotuberculosis* (une bactérie dont le génome est proche de celui de *Y. pestis*) possède l'opéron *yntABCDE*, codant pour un transporteur de nickel de type ABC (Sebbane *et al.*, 2002). Le système a été identifié chez d'autres pathogènes humains tels *Vibrio parahaemolyticus* (Park *et al.*, 2000), ou *Brucella suis* (Jubier-Maurin *et al.*, 2001).

Le système de transport du nickel décrit chez *H. pylori* est une perméase aux métaux de transition (chapitre 2.2.1.3 de l'introduction). En créant une souche mutante présentant une déficience de cette fonction, l'activité uréase est réduite de 42%, mais toujours présente (Bauerfeind *et al.*, 1996). Ce résultat permet de supposer l'existence d'un autre système de transport du nickel. La bactérie possède un opéron *abcABCD* codant pour quatre protéines dont l'une d'entre elles (abcD) présente une similarité de séquence avec NikD (Hendricks and Mobley, 1997). Mais à l'heure actuelle, il est impossible d'affirmer qu'il s'agisse d'un transporteur ABC du nickel.

2.2.1.2.2 La protéine NikA

NikA (code Swiss-Prot : P33590) est une protéine périplasmique de 502 résidus (soit 56,273 kDa) dans un premier temps synthétisée avec une séquence signal en N terminal, permettant son exportation dans le périplasme (la protéine est alors composé de 524 résidus soit 58,719 kDa). Ces 22 résidus sont ensuite clivés par une protéase lorsque la protéine traverse la membrane (Navarro *et al.*, 1993).

NikA est décrite comme membre d'une famille contenant des protéines capables de fixer des substrats tels des dipeptides et des oligopeptides, des hèmes ou des phéromones, sur la base des identités de séquences (Wu and Mandrand-Berthelot, 1995). Ces protéines se retrouvant chez les protéobactéries et chez les Gram (+), elles doivent avoir un ancêtre commun antérieur à la séparation des deux lignées (Eitinger and Mandrand-Berthelot, 2000).

Le nombre de copies de NikA dans le périplasme est estimé à environ 23 000, ce qui correspond au niveau d'expression des autres protéines périplasmiques (Maier *et al.*, 1995).

Des expériences de *quenching* intrinsèque du tryptophane (mesure de la modification de la fluorescence du tryptophane en fonction de son environnement), et de dialyse à l'équilibre ont montré qu'elle est capable de fixer un ion nickel (Ni^{2+}) par molécule, avec un K_d de l'ordre de 0,1 µM, mais qu'elle présente aussi une affinité pour d'autres cations divalents comme le cobalt (Co^{2+}), le cuivre (Cu^{2+}), ou le fer (Fe^{2+}) (de Pina *et al.*, 1995). Dans le cas du Co^{2+}, le K_d est estimé à environ 1 µM, soit dix fois plus que pour le Ni^{2+}. Par contre, des expériences plus récentes de microcalorimétrie estiment le K_d du Ni^{2+} à 11 ± 1,7 µM, contre 246 ± 29 µM pour le Co^{2+} (Heddle *et al.*, 2003). La situation n'est pas claire, et il faudrait donc comprendre ce qui provoque les différences dans la mesure du K_d, mais aussi pourquoi ces valeurs sont élevées, alors que cette protéine semble jouer un rôle clé dans la sélectivité du système NikABCDE vis-à-vis du nickel.

Des études d'absorption des rayons X semblent indiquer que le nickel est maintenu dans son site par six ou sept ligands : cinq ou six ligands N – ou O – donneurs, et un ligand S – donneur (Allan *et al.*, 1998).

NikA présente une identité de séquence protéique de 23% (calculée avec le programme Blastp : 114 résidus identiques sur les 476 résidus de OppA (Madden *et al.*, 1996)) avec OppA, dont la structure est connue (Tame *et al.*, 1994) (code Swiss-Prot : P06202 ; code PDB : 1B05) ; cette protéine périplasmique, que l'on trouve chez *S. typhimurium*, transporte vers l'intérieur de la bactérie des oligopeptides de deux à cinq résidus, mais sans spécificité de séquence. NikA et OppA ont la particularité d'avoir une taille presque deux fois plus

importantes que les autres transporteurs du périplasme (Eitinger and Mandrand-Berthelot, 2000).

Un autre aspect remarquable de NikA est son rôle dans la réponse chimiotactique de la bactérie vis-à-vis du nickel : en effet, en présence de trop fortes concentrations de Ni^{2+}, NikA interagit avec la protéine Tar, ce qui provoque la "fuite" de la bactérie devant l'excès de nickel (Charon *et al.*, 1994; de Pina *et al.*, 1995). La protéine Tar est aussi une protéine du périplasme, également capable d'interagir avec MalE (Gardina *et al.*, 1997) (une molécule périplasmique liant le maltose) mais dans ce cas, la bactérie est "attirée" par les grandes concentrations de maltose.

2.2.1.3 Le passage à travers la membrane interne : les perméases aux métaux de transition

Le second système de transport du nickel que l'on peut trouver chez les bactéries est constitué d'une protéine formant un pore à travers la membrane interne permettant ainsi le passage du nickel et du cobalt, avec une sélectivité différente en fonction de l'organisme. De cette façon, ces protéines appartiennent à la famille des transporteurs de nickel/cobalt (NiCoTs). Dans de nombreux microorganismes tels que par exemple *Ralstonia eutropha* ou *H. pylori*, la perméase présente une spécificité stricte pour le nickel.

2.2.1.3.1 La protéine HoxN

La première perméase spécifique du nickel a été mise en évidence chez *R. eutropha* et est appelée HoxN (Eberz *et al.*, 1989; Eitinger and Friedrich, 1991; Wolfram *et al.*, 1995). Il s'agit d'une protéine transmembranaire de la membrane interne.

Sa constante de transport (K_t) vis-à-vis du nickel a été estimée à environ 20 nM (Eitinger and Friedrich, 1994; Wolfram *et al.*, 1995). Le système peut importer le nickel de manière à obtenir une concentration dix à quinze fois supérieure à celle du milieu extérieur (Eitinger and Friedrich, 1994; Eitinger *et al.*, 1997). HoxN est donc un système de transport du nickel hautement spécifique (elle est incapable de prendre en charge le cobalt) mais de faible capacité.

Les études menées sur HoxN suggèrent qu'elle est constituée de huit segments transmembranaires et que ses extrémités C et N terminales sont cytoplasmiques (Eitinger and Friedrich, 1994; Eitinger *et al.*, 1997).

2.2.1.3.2 La protéine NixA

Dans le cas d'*H. pylori*, pour qui le nickel est indispensable à l'activité de son uréase, enzyme lui permettant de survire dans l'estomac des mammifère, un homologue de HoxN existe : il s'agit de NixA (Mobley *et al.*, 1995).

Cet homologue est très similaire à HoxN, puisqu'il s'agit aussi d'une protéine composée de huit segments transmembranaires, fonctionnant en tant que monomère (Fulkerson and Mobley, 2000). Son K_t est estimé à 11 nM (Mobley *et al.*, 1995).

2.2.1.3.3 D'autres exemples de perméases

Les perméases capables de transporter sélectivement le nickel se retrouvent chez une grande variété de bactéries Gram (+) et Gram (-), comme par exemple chez *Yersinia enterocolitica, Y. pseudotuberculosis*, *Y. pestis* (Sebbane *et al.*, 2002), ou *Bradyrhizobium japonicum* (Fu *et al.*, 1994) (56% d'identité avec HoxN). Mais dans certains cas, comme par exemple NhlF de *Rhodococcus rhodochrous*, la perméase est capable de transporter le nickel mais aussi le cobalt, avec une légère préférence pour ce dernier (Degen *et al.*, 1999).

Deux autres familles de protéines ont été identifiées comme des perméases potentielles spécifiques du nickel : HupE/UreJ et UreH (Eitinger *et al.*, 2005). Les gènes *hupE* et *ureJ* se trouvent chez certaines bactéries, le plus souvent associés au *cluster* des gènes de l'hydrogénase à [NiFe] et de l'uréase respectivement. En ce qui concerne le gène de *ureH*, il se trouve, chez les cyanobactéries marines, associé au gène de la superoxyde dismutase. Ces observations suggèrent l'implication de ces deux familles de protéines dans le transport du nickel. Cette hypothèse a été confirmée grâce à des expériences d'accumulation de nickel radioactif, réalisées à partir de HupE et UreH de *Rhodopseudomonas palustris*, et UreJ de *Cupriavidus necator*.

2.2.1.4 D'autres systèmes permettant l'import du nickel

Tous les systèmes de transport spécifiques du nickel connus jusqu'à présent, se classent, soit dans la famille des transporteurs ABC, soit dans la famille des perméases.

Le nickel peut toutefois aussi être pris en charge par les systèmes de transport du magnésium. En conditions artificielles, c'est-à-dire en présence d'un excès de nickel, ce dernier entre en compétition avec le Mg^{2+}, et est importé dans la bactérie (Eitinger and Friedrich, 1997). On peut considérer que le transport du magnésium contribue faiblement à l'import du nickel (Snavely *et al.*, 1991).

Il est également à noter qu'il vient d'être identifié chez plusieurs organismes (bactéries, archaebactéries, protéobactéries et cyanobactéries) une nouvelle famille de protéines (NikM, NikQ et NikO étant toujours présentes, alors que NikN, NikL et NikK sont optionnelles) qui correspond à un nouveau type de transporteur ABC spécifique du nickel (Rodionov *et al.*, 2006). Il est toutefois surprenant de constater que, dans la majorité des cas, la protéine NikK, qui fixe le nickel à l'extérieur de la bactérie, est absente (ce qui correspond à la protéine périplasmique chez les Gram (-)), comme c'est le cas pour *Rhodobacter capsulatus*.
Enfin, les études récentes de Rowe *et al.* suggèrent l'existence chez *E. coli* d'un second système de transport du nickel (en plus de NikABCDE) non encore identifié, mais déjà envisagé par *Navarro et al.* (Navarro *et al.*, 1993; Rowe *et al.*, 2005).

2.3 La régulation de la concentration cytoplasmique du nickel

Malgré la nécessité pour les bactéries de disposer de nickel cytoplasmique, il peut aussi se révéler toxique pour les cellules, puisqu'une concentration intracellulaire trop importante peut entraîner divers évènements :

- La catalyse, par le métal, de la formation d'espèces oxygénées hautement réactives, ou ROS pour *Reactive Oxygen Species* ;
- La fixation non spécifique du nickel à des protéines ou aux membranes (Chivers and Sauer, 2000) ;
- Le déplacement de métaux de leurs sites actifs (Chivers and Sauer, 2000).

Il est donc indispensable pour la bactérie d'avoir des systèmes de contrôle et de régulation de la concentration intracellulaire de nickel.

2.3.1 La régulation de l'expression du système NikABCDE

2.3.1.1 L'activation de la transcription de l'opéron par FNR

L'activation de l'opéron *nikABCDE* est assurée par la protéine FNR (*Fumarate Nitrate Regulator*) qui permet le contrôle de la transcription de différentes protéines impliquées dans les activités métaboliques de la fermentation et de la respiration anaérobie (Wu and Mandrand-Berthelot, 1986).
FNR assure son rôle biologique en se fixant à un site situé en amont (du +1 de transcription) de l'opéron *nikABCDE* (Eitinger and Mandrand-Berthelot, 2000) (Figure 2.3-1).

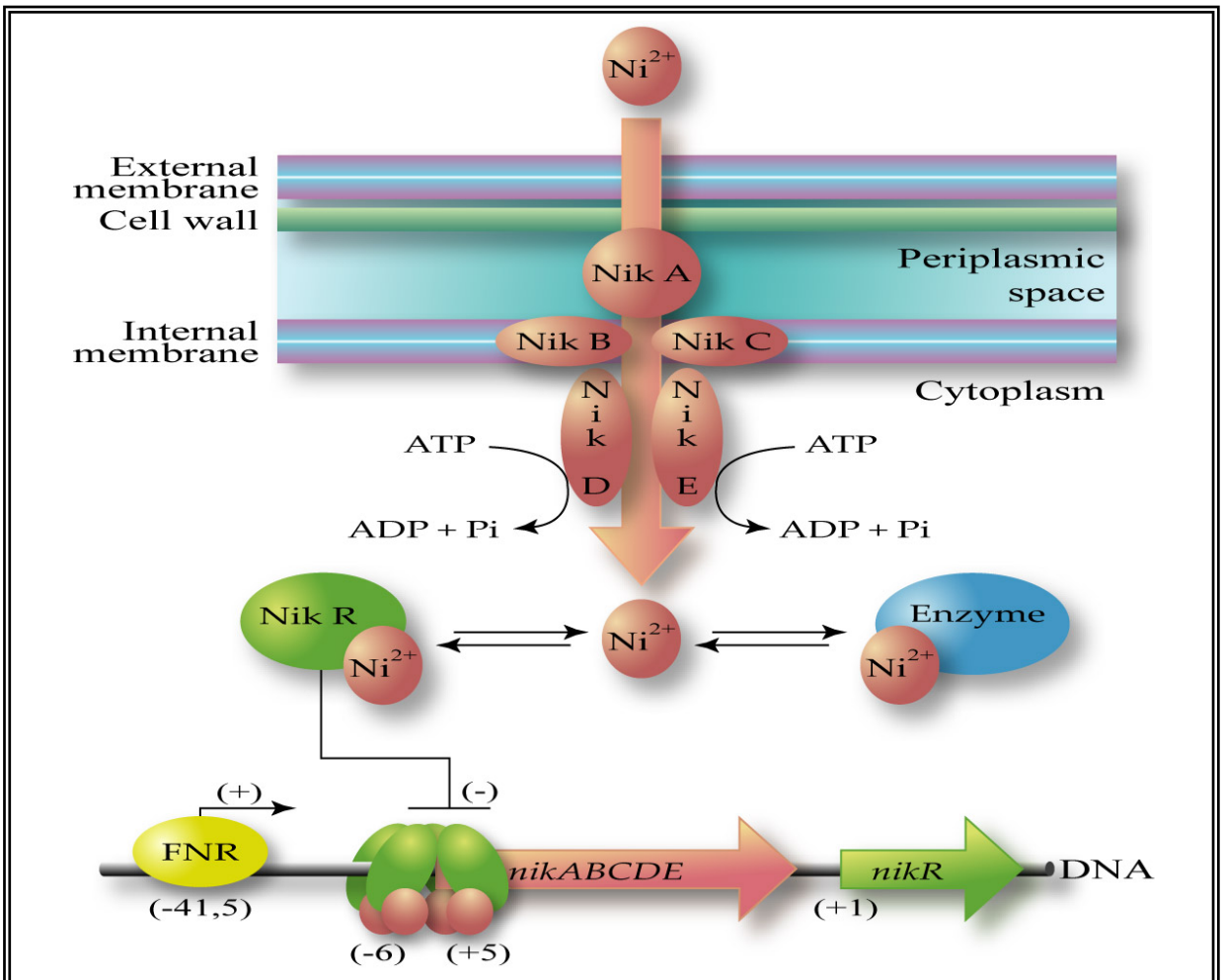


Figure 2.3-1 : Schéma du transport et de la régulation du nickel chez *E. coli* par NikABCDE

2.3.1.2 La répression de la transcription de l'opéron par NikR

La protéine NikR (code Swiss-Prot : P0A6Z6 ; (De Pina *et al.*, 1999)) est le répresseur de l'opéron *nikABCDE*. Elle est constituée de 133 résidus, soit 15,094 kDa ((Chivers and Sauer, 2000; Dosanjh and Michel, 2006) pour une revue), et se compose de deux domaines, séparés par une région charnière :

- Le domaine N terminal, allant jusqu'au résidu 48, qui correspond au site de fixation à l'ADN ;
- La région centrale, correspondant aux acides aminés 49 à 55, qui constitue une région inter-domaine flexible ;
- Enfin, le domaine C terminal, constitué des résidus 56 à 133, où se situe le site de fixation du nickel.

Elle peut être classée dans la famille des brin – hélice – hélice (βαα), qui contient déjà quatre répresseurs procaryotes se fixant également à l'ADN sous forme dimérique : Arc, CopG,

MetJ, et Mnt (Knight *et al.*, 1989). Lors de la dimérisation de ces dernières, chaque monomère (qui adopte la conformation βαα) interagit avec le second de façon à ce que leurs brins β constituent un feuillet β antiparallèle (Chivers and Sauer, 1999). Ce dimère se fixe alors à l'ADN.

En présence d'un excès de nickel, NikR se fixe à l'ADN au niveau d'une séquence GTATGA-16N-TCATAC chevauchant le nucléotide +1 de l'opéron *nikABCDE* (Figure 2.3-1). Cette fixation réprime alors la transcription des gènes, en masquant à la polymérase le site d'initiation de la transcription (le nucléotide +1).

Le gène *nikR* se situe juste en aval de l'opéron *nikABCDE*, par conséquent, la transcription peut se faire à partir de deux promoteurs différents (Figure 2.3-1) :

- Son propre promoteur, qui n'est pas sous contrôle de la protéine FNR, permet de conserver NikR à un niveau basal, indépendamment de la concentration intracellulaire du nickel ;
- Le promoteur de l'opéron *nikABCDE*, qui permet la transcription du gène *nikR* à la suite des gènes de l'opéron.

Contrairement aux protéines décrites précédemment, NikR fonctionne en tant que tétramère (Chivers and Sauer, 2002; Schreiter *et al.*, 2003). Deux sites de fixation du nickel ont été identifiés par monomère (Chivers and Sauer, 2002) :

- Le premier est un site de forte affinité (avec un K_d de l'ordre de 1 pM). Son occupation permet une interaction faible avec l'ADN de l'ordre du nanomolaire ;
- Le deuxième site est de faible affinité (avec un K_d de l'ordre de 30 μM). Il permet d'étendre la zone d'interaction de la protéine avec l'ADN (passant d'environ 40 pb à environ 65 pb), augmentant son affinité à 15 pM.

Des études récemment publiées semblent montrer que l'activité de NikR est inhibée par un ou plusieurs des composants du système d'assemblage des hydrogénases (Rowe *et al.*, 2005), permettant ainsi à la bactérie d'assurer la diponibilité du nickel nécéssaire à l'activité des hydrogénases.

2.3.1.3 La répression de la transcription de l'opéron par NarL

Des études récentes tendraient à monter que NikR n'est pas le seul répresseur de l'opéron *nikABCDE*, mais que la protéine NarL joue également ce rôle (Rowe *et al.*, 2005). En effet, en présence de nitrate, l'expression des hydrogénases se voit réprimée en faveur de l'expression de la nitrate réductase (Unden and Bongaerts, 1997). Dans ces conditions, la protéine NarX va phosphoryler NarL, augmentant ainsi son affinité pour l'ADN au niveau d'un site en amont de *nikABCDE* distinct de celui de NikR ou de FNR (Maris *et al.*, 2002). Par contre, la répression de l'opéron *nikABCDE* par NarL n'est pas observée en absence de nitrate. Ces résultats doivent néanmoins être confirmés.

2.3.2 Les autres régulations de transcription contrôlées par le nickel

2.3.2.1 La protéine NikR chez d'autres bactéries

Des homologues de NikR se retrouvent dans de nombreux organismes.
Il a notamment été montré que *H. pylori* possédait une protéine NikR (HpNikR) impliquée dans plusieurs régulations (van Vliet *et al.*, 2002; Contreras *et al.*, 2003) :

- Elle permet d'activer l'expression de l'uréase en fonction de la quantité de nickel disponible dans le milieu (van Vliet *et al.*, 2002) ;
- Elle active ou réprime l'expression de plusieurs protéines en fonction du nickel, telles que par exemple celles impliquées dans l'import de ce dernier (Contreras *et al.*, 2003) ;
- Elle est aussi impliquée dans la réponse globale de la bactérie lorsqu'elle se trouve en conditions acides modérées (Bury-Mone *et al.*, 2004; van Vliet *et al.*, 2004; van Vliet *et al.*, 2004).

2.3.2.2 D'autres exemples de régulations par le nickel

Il existe de nombreuses protéines dont la transcription est régulée par la concentration de nickel, comme par exemple :

- La transcription des gènes de l'hydrogénase chez *B. japonicum* (Kim and Maier, 1990; Kim *et al.*, 1991) ;
- La transcription des gènes de la SOD chez plusieurs *Streptomyces* (Youn *et al.*, 1996; Youn *et al.*, 1996). Chez *S. griseus*, les protéines SrnQ et SrnR sont capables de former un hétéro-octamère qui, après fixation de nickel, réprime la

transcription de la SOD à fer et zinc (*sodF*) (Kim *et al.*, 2003). Plus récemment, la protéine Nur (qui est une protéine homologue de Fur) a été identifiée chez *S. coelicolor*. Elle est capable, sous l'action de la fixation de nickel, de réprimer la transcription des gènes de la SOD à fer et zinc (*sodF*), ainsi que ceux impliqués dans l'import du nickel. Elle est aussi indirectement impliquée dans l'activation de la transcription du gène de la SOD à nickel (*sodN*) (Ahn *et al.*, 2006) ;

- …

2.3.3 Le stockage cytoplasmique du nickel

Une seule protéine a été décrite comme ayant potentiellement pour rôle le stockage du nickel dans le cytoplasme : Hpn de *H. pylori* (Ge *et al.*, 2006). Il s'agit d'une protéine de soixante résidus, riche en histidines (28 histidines, soit 46,7% de la protéine), dont l'expression est régulée positivement par HpNikR, et qui représente environ 2% des protéines de la bactérie. Elle est présente en solution sous la forme d'un multimère. Chaque protéine semble pouvoir fixer cinq ions nickel avec un K_d de 7,1 µM. Cette fixation est réversible, puisque l'EDTA est capable de retirer le nickel de la protéine.
Les travaux de Ge *et al.* suggèrent également qu'il existe chez *H. pylori* d'autres protéines ayant la même fonction que Hpn (Ge *et al.*, 2006).

2.3.4 L'export du nickel

Une autre stratégie de contrôle de la concentration cytoplasmique du nickel consiste à exporter l'excès de nickel importé.
Récemment, la protéine YohM a été mise en évidence chez *E. coli* (Rodrigue et al., 2005). Il s'agit d'un système d'efflux permettant à la bactérie d'exporter son excès de nickel et de cobalt. La protéine semble être associée à la membrane interne grâce à six segments transmembranaires. Elle dispose d'une boucle riche en histidine qui est sûrement impliquée dans la fixation du métal.
Chez les *Ralstonia*, on trouve deux types de systèmes d'export du nickel : Cnr et Ncc (Silver and Phung, 1996).

3 Généralités sur l'import du fer

Au cours de ma thèse, nous nous sommes aperçus que le transport du nickel présente certaines similarités avec celui du fer. C'est la raison pour laquelle cette introduction possède une partie consacrée à des généralités sur le fonctionnement de l'import du fer chez les bactéries. Les informations qui y sont présentes serviront à l'interprétation de nos résultats.

3.1 Le fer cellulaire

Le fer est l'un des éléments les plus importants pour les cellules procaryotes et eucaryotes, car il est impliqué dans un large éventail de voies métaboliques, et de signalisations. Parmi celles-ci, on trouve par exemple les chaînes respiratoires, l'activation d'oxygène, la réduction de peroxydes, la biosynthèse d'acides aminés et de nucléosides, ou la photosynthèse ((Wandersman and Delepelaire, 2004) pour une revue).

Les deux états redox, Fe^{2+} (le fer ferreux) et Fe^{3+} (le fer ferrique), sont utilisés par les cellules. Le potentiel redox varie en fonction de la coordination et de l'environnement protéique du fer. Le fer est l'un des éléments chimiques les plus abondants. Mais en conditions physiologiques (en présence d'oxygène et à pH neutre), le métal est rapidement oxydé du Fe^{2+} au Fe^{3+}, puis devient insoluble sous la forme d'hydroxyde. Les organismes, et notamment les bactéries, ont donc dû développer des systèmes d'import hautement spécifiques et efficaces.

Lorsque le fer est de nouveau réduit en Fe^{2+} la réaction de Fenton devient alors possible (Figure 3.1-1).

$$Fe^{2+} + H_2O_2 \longrightarrow Fe^{3+} + OH^- + OH^\bullet$$

Figure 3.1-1 : Réaction de Fenton.

Le fer ferreux réagit avec le peroxyde d'hydrogène (H_2O_2) ou des lipides peroxydés, pour générer du fer ferrique, des ions hydroxydes (OH^-), et des radicaux hydroxyles ($OH^\bullet$) qui sont très réactifs (comme par exemple des radicaux lipidiques comme $LO^\bullet$ et $LOO^\bullet$). L'ensemble de ces radicaux représente alors un risque pour l'intégrité des lipides membranaires, des protéines et des acides nucléiques. Il est tout de même à noter que l'existence de ce mécanisme in vivo est controversée car il semble nécessiter des conditions non physiologiques, et notamment un pH trop acide. Malgré cela, une surcharge cellulaire en fer, aussi bien qu'un déficit, a toutes les chances de conduire la cellule à une mort certaine : c'est

pourquoi la concentration en fer doit être soigneusement contrôlée et limitée. Le fer est alors séquestré dans la cellule par les lactoferrines, les transferrines, et les ferritines, ou fixé aux protoporphyrines des hémoprotéines. De ce fait, il n'y a virtuellement pas de fer libre dans la cellule.

3.2 L'import direct du fer par les microorganismes

Le Fe^{3+} étant insoluble, son incorporation directe par la bactérie est impossible. Par contre en conditions réductrices ou anaérobies, c'est le Fe^{2+} soluble qui est majoritaire.
Dans le cas des bactéries Gram (-) il peut donc diffuser librement, ou contre un potentiel de protons, à travers les porines de la membrane externe. Il est ensuite pris en charge par un système de transport de type ABC que l'on retrouve dans de nombreuses espèces (Kammler *et al.*, 1993). Dans le cas des bactéries pathogènes *Salmonella enterica* et *H. pylori* (qui colonisent respectivement l'intestin et l'estomac des mammifères, deux milieux dont les conditions permettent l'existence de Fe^{3+}), le transport est assuré par le système Feo (Velayudhan *et al.*, 2000; Boyer *et al.*, 2002).

3.3 L'import indirect du fer par les microorganismes

3.3.1 Les sidérophores

La quantité de Fe^{3+} sous forme de complexes dans l'environnement étant importante mais non disponible (le fer ne se trouvant jamais sous forme libre), les microorganismes ont développé des moyens de le solubiliser et le séquestrer : les sidérophores. Il s'agit en effet de molécules de faible poids moléculaire, capable de fixer le Fe^{3+} en formant un complexe hexacoordonné avec une constante d'affinité très élevée, de l'ordre de 10^{30} M^{-1} (Faraldo-Gomez and Sansom, 2003; Wandersman and Delepelaire, 2004). Plus de 500 sidérophores appartenant à différentes classes ont été jusqu'à présent décrits, la plupart chez les bactéries Gram (+) et Gram (-) (Winkelmann, 2001).
Ils sont produits par des protéines ressemblant à celles impliquées dans la biosynthèse des antibiotiques. Ces gènes sont régulés par la concentration cytoplasmique du fer, et se trouvent souvent en *cluster* avec les gènes des systèmes d'import du métal. On peut les classer en fonction de la nature chimique du groupe chélatant :

- Hydroxamate : comme par exemple le ferrichrome (740 Da) produit par les champignons. *E. coli* ainsi que d'autres bactéries sont capables de l'utiliser (Figure 3.3-1a) ;

- Cathécholate (présentant un cycle catéchol) : comme par exemple l'entérobactine (669 Da) produite par *E. coli* et *S. typhumurium*. Il est également utilisé par toutes les entérobactéries connues (Figure 3.3-1b) ;
- Hydroxyacide : comme par exemple la pyochéline (324 Da) produite par *Pseudomonas aeruginosa* (Figure 3.3-1c).

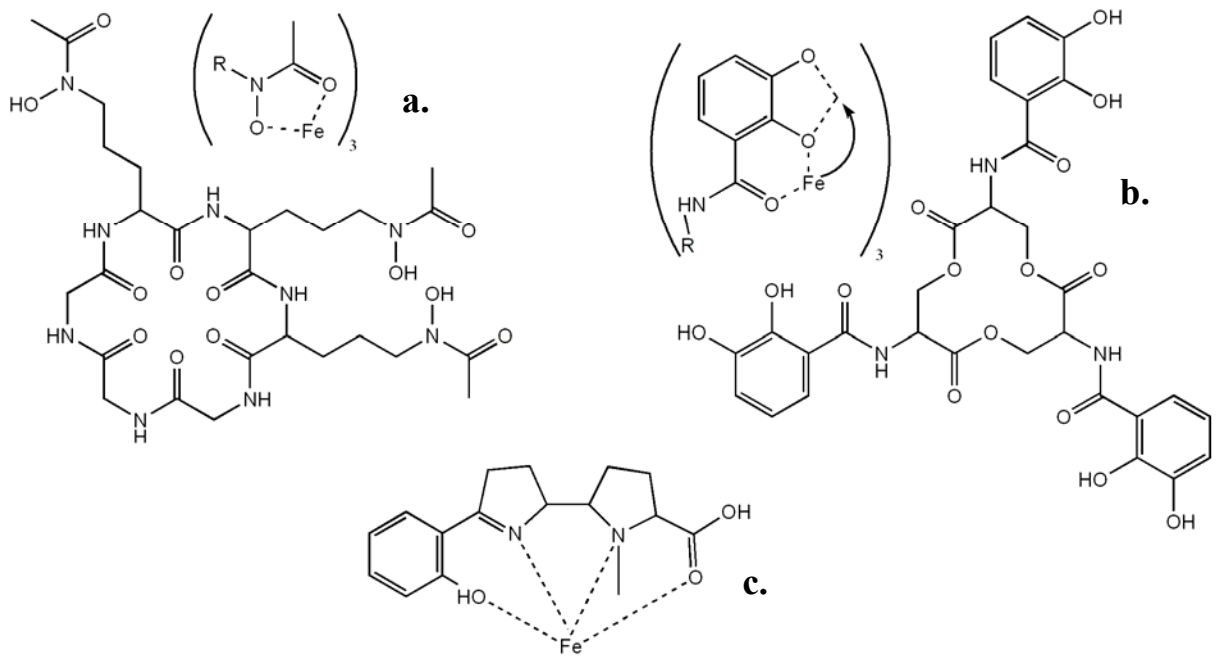


Figure 3.3-1 : Structure de trois sidérophores. a. ferrichrome ; b. entérobactine ; c. pyochéline.

Mais cela ne permet pas de classer exhaustivement les sidérophores. En effet, certains contiennent plus d'un de ces groupements fonctionnels, ou même d'autres fonctions.
La mycobactine des mycobactéries se distingue des autres sidérophores, car elle est ancrée à la membrane externe grâce à une longue chaîne aliphatique (Faraldo-Gomez and Sansom, 2003).

3.3.2 Le transport du fer par les sidérophores

Après synthèse par la machinerie cellulaire, les sidérophores sont exportés dans le milieu extérieur pour regagner ensuite le cytoplasme de la bactérie, complexés au Fe^{3+} puisé dans le milieu (Figure 3.3-2).

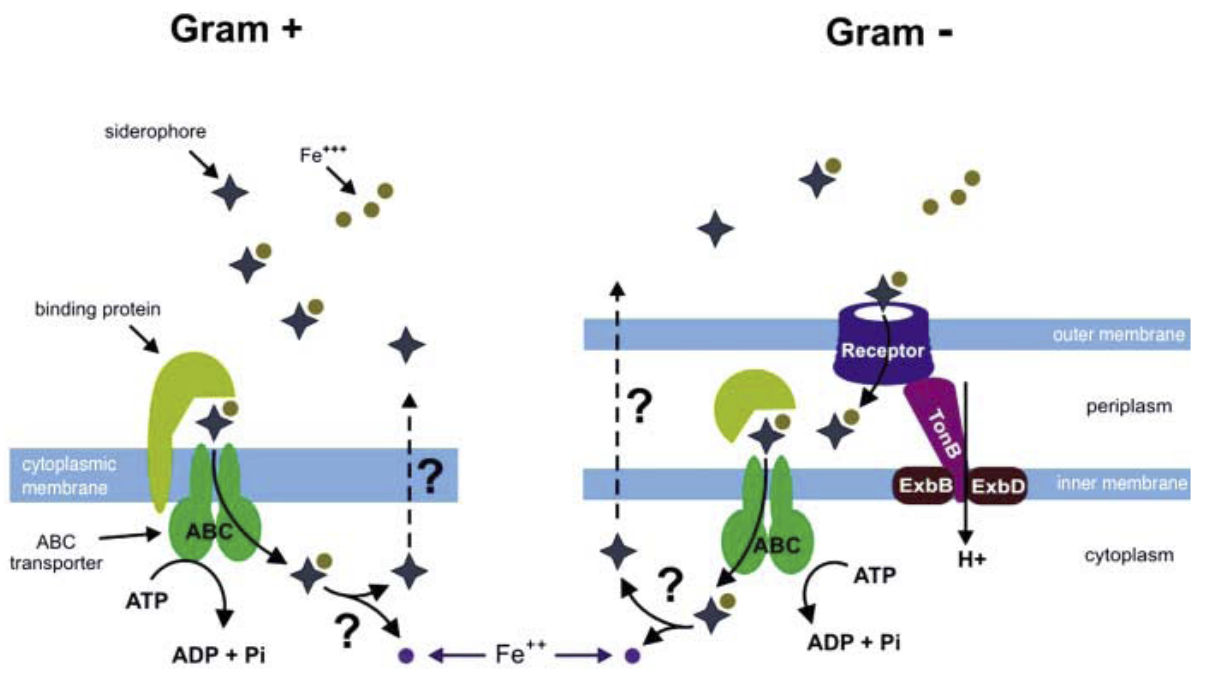


Figure 3.3-2 : Schéma de l'import et de l'export des sidérophores chez les Gram (+) et Gram (-) (Wandersman and Delepelaire, 2004).

Dans le cas des Gram (-), le complexe fer – sidérophore est reconnu par un récepteur de la membrane externe composé de deux domaines, dont un tonneau β associé à la membrane externe (Bonhivers *et al.*, 2001; Ferguson *et al.*, 2002). Le récepteur ayant une haute affinité pour le complexe fer – sidérophore, l'énergie nécessaire à son relargage dans le périplasme est apportée par le système TonB (ou son homologue) (Schalk *et al.*, 2004). Il se compose de la protéine membranaire ExbB, et de deux protéines périplasmiques ancrées à la membrane plasmique ExbD et TonB (Klebba *et al.*, 1993; Postle, 1993; Braun, 1995; Moeck and Coulton, 1998). Grâce à la force motrice du proton, TonB apporte au récepteur avec lequel il est en interaction, l'énergie nécessaire à l'import du complexe de fer.

A partir de ce moment, la situation est la même dans le cas des bactéries Gram (-) et Gram (+) : le complexe fer – sidérophore est pris en charge par des transporteurs du type ABC (Koster, 2001).

Dans la majorité des cas, le fer est importé avec le sidérophore, et parfois il passe par deux sidérophores avant d'atteindre le cytoplasme. Une fois dans celui-ci, le fer est séparé du sidérophore, soit par hydrolyse de ce dernier, soit par réduction du Fe^{3+} en Fe^{2+}. Pourtant l'existence de cette dernière voie métabolique est controversée (Muller *et al.*, 1998; Wandersman and Delepelaire, 2004).

4 L'intérêt de l'étude structurale des protéines NikA et NikR

Nous venons de voir que le nickel était un élément essentiel pour les microorganismes, le plus souvent pour assurer leurs activités uréase et/ou hydrogénase (Maier, 2005), et notamment dans le cas des pathogènes (ASPC) tels que :

- *H. pylori* : Il s'agit de la première bactérie impliquée dans la genèse du cancer de l'estomac : deuxième cause de cancers dans le monde (600 000 nouveaux cas/an dont 9 000 en France) et première cause dans les pays en voie de développement ;
- *Y. enterocolitica* (Sebbane *et al.*, 2002) : une bactérie de la même famille que *Y. pestis* : bactérie responsable de la peste bubonique (mortelle dans 70% des cas en absence de traitement) et de la peste pulmonaire (mortelle dans 100% des cas en absence de traitement) ;
- *S. enterica* : responsable de la salmonellose, une maladie provoquant des diarrhées, des crampes abdominales et de la fièvre ;
- *E. coli* et *Campylobacter jejuni* : provoquent toutes deux des diarrhées et des gastro-entérites ;
- *Bordetella bronchiseptica* : entraîne des infections respiratoires semblables à la coqueluche ;
- *Ureaplasma urealyticum* : provoque des urétrites non gonococciques ;
- …

Une bonne compréhension de tous les processus impliqués dans l'utilisation du nickel par les microorganismes peut ainsi présenter un intérêt médical.
L'objectif de cette thèse a donc été la résolution de la structure de deux protéines impliquées dans l'import du nickel par le transporteur ABC de la bactérie modèle *E. coli*, par la technique de la cristallographie aux rayons X des protéines :

- La protéine NikR qui est le répresseur de l'opéron *nikABCDE* ;
- La protéine NikA qui est la partie périplasmique du système NikABCDE.

4.1 La protéine NikR

Au moment où a commencé cette thèse, aucune structure tridimensionnelle de NikR (chapitre 2.3.1.2 de l'introduction) n'était disponible, que ce soit d'*E. coli* ou d'aucun autre organisme.

Un groupe américain avait obtenu des cristaux du domaine N terminal (le domaine responsable de la fixation à l'ADN) diffractant à 3,5 Å, mais les données de diffraction n'ont pu permettre la résolution de la structure (Chivers and Sauer, 1999).

Il était cependant intéressant d'obtenir la structure de la protéine entière, pour différentes raisons :

- On aurait pu, dans un premier temps, confirmer ou infirmer l'existence du repliement de type brin – hélice – hélice pour le domaine N terminal ;
- La structure aurait aussi pu nous permettre de comprendre comment s'organisent les sites de fixation du nickel, ainsi que d'appréhender les changements structuraux provoqués par l'occupation d'un ou plusieurs de ses sites, puisque ce doivent être eux qui induisent la dimérisation, et donc la fonction biologique de NikR ;
- Un autre point intéressant aurait été de comparer la structure de NikR fixée à l'ADN avec celle des autres protéines adoptant une conformation brin – hélice – hélice, car dans le cas de ces dernières, les deux séquences répétées auxquelles se fixe le dimère, ne sont séparées que par quatre à six résidus ;
- La connaissance de la structure de NikR d'*E. coli*, aurait finalement permis d'aider la compréhension et l'étude des homologues de la protéine des autres organismes, comme par exemple chez *H. pylori*.

4.2 La protéine NikA

Le premier intérêt de l'étude de la structure de NikA (chapitre 2.2.1.2.2 de l'introduction) est de comprendre la fixation du nickel au sein de la molécule : comment la protéine sélectionne préférentiellement le Ni^{2+} par rapport aux autres métaux ? L'étude de cette fixation permettrait peut-être de comprendre les différences que l'on trouve dans la littérature pour la valeur du K_d de NikA concernant le nickel, et ainsi de clarifier la situation. De plus, elle permettrait aussi de comprendre pourquoi la protéine semble si peu spécifique : la valeur de sa constante de dissociation étant élevée.

Il serait également intéressant de comparer cette fixation à celle des peptides liés à OppA : si les deux protéines possèdent bien un repliement similaire, quels sont les changements structuraux du site actif permettant, dans un cas, la reconnaissance et la fixation d'un ion, et dans un autre cas, la fixation de peptides sans spécificité de séquence.
On pourrait également essayer de comprendre comment MalE et NikA peuvent induire deux réponses opposées en interagissant avec la même protéine. Pour cela, il faudrait comparer les structures de NikA et MalE complexées à Tar, mais il faudrait obtenir des cristaux des complexes, ce qui implique la réalisation d'une nouvelle étude cristallographique.
Pour ces raisons, en 1994, la détermination de la structure de NikA avait été tentée au L.C.C.P. (Charon *et al.*, 1994). A l'époque, l'équipe de L.F. Wu, fournissait la protéine, à partir de laquelle deux formes de cristaux ont pu être obtenues :

- Les premiers mesuraient 0,4 x 0,4 x 0,6 mm^3 et diffractaient à environ 3,0 Å de résolution. Le traitement des données a permis de déterminer leur appartenance aux groupes d'espace $P6_2$ ou $P6_4$ (ce qui reste à vérifier), ainsi que les paramètres de maille suivants : a = b = 160,3 Å, c = 138,4 Å ;
- La deuxième forme cristalline a été obtenue grâce à un traitement à l'EDTA 1 mM. Les cristaux, d'une taille de 0,05 x 0,40 x 0,40 mm^3, ont diffracté à 5,6 Å, et appartenaient au groupe d'espace $P2_12_12_1$, avec les paramètres de maille : a = 192,8 Å, b = 69,7 Å, c = 149,9 Å.

Afin de pouvoir résoudre la structure de la protéine, et éviter de devoir déterminer les phases par la méthode MIR (car la fixation d'atomes lourds dans les cristaux n'a pas fonctionné, ce qui constitue une étape indispensable à la réalisation de cette méthode), un remplacement moléculaire avec OppA a été tenté, mais n'a malheureusement donné aucun résultat exploitable. De plus, le phasage de la structure par méthode MAD en utilisant la diffusion anomale du nickel n'a pas pu être mis en œuvre, du fait de la faible intensité de cette dernière dans les cristaux de NikA.
Cette thèse avait donc pour buts, de reproduire les cristaux de NikA, de tenter d'en améliorer la qualité et, dans un second temps, de cristalliser la protéine séléniée afin de déterminer les phases, non plus par méthode MIR, mais par méthode MAD.

MATERIEL

ET

METHODES

1 La biologie moléculaire et la biochimie

1.1 La biologie moléculaire

Toute expérience de cristallographie des protéines débute toujours par l'obtention d'une quantité suffisante de protéine pure en solution afin de pouvoir par la suite tenter de la cristalliser.

1.1.1 Le principe de la surexpression des protéines

La solution la plus simple, lorsqu'elle est envisageable, est la surexpression de la molécule d'intérêt par une bactérie (le plus souvent on utilise *E. coli*). Cette méthode se réalise en plusieurs étapes :

- Dans un premier temps, on doit introduire le gène codant pour la protéine d'intérêt dans un plasmide, au niveau de la cassette de clonage (région nucléotidique possédant de nombreux sites de restriction), sous contrôle d'un promoteur fort et inductible. Le plasmide est un ADN circulaire, qui peut être considéré comme un mini-chromosome capable de se répliquer de façon autonome. Il doit posséder, pour surproduire une protéine : une origine de réplication bactérienne ; une origine de transcription spécifique d'une ADN polymérase (dont le gène est présent sur le plasmide ou dans le génome bactérien), suivie d'une cassette de clonage ; un site de fixation ribosomal ; et un gène de résistance à un antibiotique qui permettra la sélection des bactéries ayant intégré la construction (plasmide + gène d'intérêt) ;
- La construction ainsi obtenue est introduite par transformation bactérienne (choc thermique, électroporation …) dans une souche compétente : souche bactérienne capable d'intégrer le plasmide grâce à la porosité de sa paroi de peptidoglycane ;
- Enfin, la surexpression est provoquée par induction de la transcription du gène.

1.1.2 Le protocole de surexpression de la protéine NikA

Le plasmide utilisé nous a été fourni par l'équipe de L. F. Wu, et correspond à celui qui a été utilisé lors des études précédentes de cristallisation réalisées au laboratoire (Charon *et al.*, 1994). Il s'agit d'une construction où la séquence codante du gène de NikA (avec la séquence signal correspondant à l'exportation dans le périplasme) a donc été clonée dans le

vecteur d'expression pET22b (NOVAGEN) sous le contrôle du promoteur T7. Le plasmide possède un gène de résistance à l'ampicilline, et la protéine est surexprimée chez *E. coli*.

1.1.2.1 La surexpression de NikA en milieu minimum

1.1.2.1.1 La transformation de la souche B834(DE3) par la construction pET22b/nika

Dans un premier temps, on ajoute 1 µL de la construction pET22b/*nik*a à 200 µL de bactéries de la souche compétente d'*E. Coli* B834(DE3), et on incube 30 minutes dans la glace. La souche d'*E. coli* B834 (DE3) présente la particularité d'être auxotrophe à la méthionine, et possède une copie du gène codant pour la polymérase T7, sous contrôle du promoteur *lac UV5* inductible à l'IPTG. Par la suite, la transformation se fait grâce à un choc thermique : 45 secondes à 42°C. Les 200 µL de bactéries sont alors mises en préculture, en ajoutant 800 µL de LB, à 37°C, sous agitation pendant 60 minutes. On commence la sélection des bactéries ayant intégré la construction en mettant les cellules en préculture dans un milieu sélectif LB et ampicilline (100 mg/L) à 37°C sous agitation pendant 6 heures ; puis cette sélection se poursuit dans un volume de 200 mL de milieu minimum NYM (Tableau 1.1-1), méthionine (300 µM) et ampicilline (100 mg/L) à 37°C sous agitation toute la nuit (soit environ 15 heures).

1.1.2.1.2 L'induction de la surexpression

Dans un second temps, cette préculture est utilisée pour inoculer 2 L de milieu minimum NYM (Tableau 1.1-1), plus 0,3 mM de méthionine et 100 mg/L d'ampicilline.

Solutions stock autoclavées		**Solutions stock filtrées**	
Composé	**Concentration finale**	**Composé**	**Concentration finale**
$(NH_4)SO_4$	7,5 mM	Ca^{2+}	1 µg/mL
NaCl	8,5 mM	Fe^{2+}	1 µg/mL
KH_2PO_4	55 mM	Oligoéléments	1 ng/mL
K_2HPO_4	100 mM	Thiamine	10 µg/mL
$MgSO_4$	1 mM	Biotine	10 µg/mL
Glycérol	4,8%		

Tableau 1.1-1 : Composition du milieu NYM (modifié à partir du milieu NMM (Budisa *et al.*, 1995))

La surexpression de NikA est induite par ajout d'IPTG (0,5 mM) lorsque la D.O. à 600 nm de la culture atteint une valeur comprise entre 0,6 et 0,8 (après environ 3 heures 30 minutes). L'expression est alors réalisée pendant 15 heures. Finalement, les bactéries sont récupérées par centrifugation : 8000 rpm pendant 15 minutes.

1.1.2.2 La surexpression de NikA séléniée en milieu minimum

Dans le cas de la surexpression de la protéine séléniée, le protocole est identique au précédent, sauf en ce qui concerne la méthionine :

- Après l'étape de préculture dans les 200 mL de NYM, on centrifuge les bactéries et on ne récupère que le culot. Cette étape permet d'éliminer la grande majorité de la méthionine qui avait été ajoutée à l'étape précédente. A l'étape suivante, on ajoute alors aux 2 L de milieu minimum NYM, 0,05 mM de séléno – méthionine et 100 mg/L d'ampicilline;
- On induit alors les 2 L de culture avec 0,5 mM d'IPTG, lorsque la D.O. (à 600 nm) de la culture a atteint environ 0,4. On ajoute également 0,3 mM de séléno – méthionine et 100 mg/L d'ampicilline.

1.1.2.3 La surexpression de NikA en LB et TB

1.1.2.3.1 La transformation de la souche BL21(DE3) par la construction pET22b/nika

Le protocole de transformation de la souche d'*E. coli* BL21(DE3) par la construction pET22b/*nika* est très similaire à celui utilisé pour la transformation la souche B834(DE3), auxotrophe à la méthionine, par la même construction. La différence principale est l'absence d'utilisation du milieu NYM.

On ajoute 1 µL de la construction pET22b/*nik*a à 200 µL de bactéries de la souche compétente d'*E. Coli* BL21(DE3), et on incube 30 minutes dans la glace. Puis la transformation se fait grâce à un choc thermique : 45 secondes à 42°C. Les 200 µL de bactéries sont alors mises en préculture, en ajoutant 800 µL de LB, à 37°C sous agitation pendant 60 minutes. On sélectionne les bactéries ayant intégré la construction en étalant les cellules sur une boîte de Pétri contenant un milieu sélectif LB et ampicilline (100 mg/L). Les bactéries sont ensuite placées à 37°C pendant 12 heures.

Par la suite, on sélectionne un clone qui nous sert à ensemencer 2 mL de LB et ampicilline (100 mg/L). On prélève 500 µL de cette culture, à la quelle on ajoute 500 µL de glycérol afin de réaliser un stock de bactéries ayant intégré le plasmide, que l'on stocke à - 80°C.

1.1.2.3.2 L'induction de la surexpression

On réalise, à l'aide du stock produit à l'étape précédente, une préculture nous permettant d'inoculer au 1/20ème les milieux de culture LB ou TB plus ampicilline (100 mg/L). La surexpression de NikA est induite par ajout d'IPTG (0,5 mM) lorsque la D.O. à 600 nm de la culture atteint entre 0,6 et 0,8 (après environ 2 heures 30 minutes dans le cas du LB et environ 1 heure 30 minutes dans le cas du TB). Les bactéries sont alors laissées 3 heures à 37°C sous agitation, avant de les récupérer par centrifugation : 8000 rpm pendant 15 minutes.

1.2 La biochimie

1.2.1 Le principe de la purification des protéines

Pour pouvoir réaliser des essais de cristallisation, il est impératif de disposer d'une solution de protéine très pure, car tout contaminant présent en quantité trop importante dans la solution peut, le plus souvent, provoquer des problèmes lors de la cristallisation (comme par exemple provoquer l'apparition de cristaux mosaïques), et même jusqu'à empêcher l'apparition de cristaux. C'est pourquoi, en général on cherche à avoir une pureté de la solution protéique d'intérêt de l'ordre de 98 à 99%.

Une des méthodes les plus utilisées est la chromatographie. Elle permet de séparer des protéines en fonction de leurs propriétés physico-chimiques ou de leur affinité pour une molécule. Les protéines, lors de leur passage sur une résine contenue dans une colonne, vont être retenues avec plus ou moins d'affinité, et séparées par ajout d'un éluant : les molécules ayant le moins d'affinité pour la colonne, se détachant avant celles qui sont liées plus fortement.

L'étape de purification permet alors de séparer le mélange de protéines, en fonction de leurs propriétés :

- Les protéines vont pouvoir être séparées en fonction de leur taille et de leur forme. Ces propriétés sont utilisées notamment dans la technique de filtration sur gel, où le gel utilisé est poreux : les petites molécules vont y pénétrer, et

ainsi parcourir une distance plus importante que les protéines de taille plus grande, qui, elles, vont en être exclues et sortir de la colonne plus rapidement ;

- La charge des molécules va être mise à profit dans des techniques telles que la chromatographie par échange d'ions. Il existe deux types de résines, celles chargées positivement permettant la séparation des anions, et celles chargées négativement adaptées à la séparation des cations ;
- La solubilité vis-à-vis d'un solvant est exploitée par les techniques de précipitation. Le plus souvent la précipitation se fait par ajout de sulfate d'ammonium ;
- L'hydrophobicité de la protéine peut aussi être utilisée dans des techniques comme la chromatographie sur colonne hydrophobe ;
- Enfin on peut se servir de la spécificité de la protéine pour une molécule pour la purifier, comme en chromatographie d'affinité.

Un impératif supplémentaire imposé par la biologie, et de façon encore plus importante par la biologie structurale, est le maintien de l'intégrité de la protéine. Toutes les molécules doivent être fonctionnellement et structuralement identiques en fin de purification, pour pouvoir résoudre la structure de la protéine : l'absence d'isomorphisme entre les protéines pouvant empêcherer la cristallisation, et dans le cas où des cristaux seraient obtenus, ils risquent d'être de très mauvaise qualité. Dans ce cas, comme la cristallographie biologique fournit une image moyenne de toutes les molécules du cristal, la qualité de la structure serait également affectée. Les précautions les plus simples à mettre en place pour maintenir l'intégrité structurale de la protéine peuvent être par exemple, de réaliser le plus rapidement les différentes étapes de la purification, et de les effectuer le plus possible à 4°C.

1.2.2 Le protocole de purification de la protéine NikA

A l'époque des premiers essais de cristallisation et de résolution de la structure de NikA, au laboratoire, par Charon *et al.* (Charon *et al.*, 1994), la protéine était fournie par l'équipe de L.F. Wu. La purification se faisait selon le protocole suivant :

- Extraction périplasmique au chloroforme dans un premier temps, (cette étape a été remplacée par une extraction périplasmique réalisée à l'EDTA à cause de problèmes de protéolyse de NikA) ;
- Précipitation avec 80% (pourcentage de saturation) de sulfate d'ammonium ;

- Chromatographie sur une colonne hydrophobe : TSK-phénol ;
- Chromatographie sur une colonne échangeuse d'anions : DEAE ;
- Enfin une chromatographie sur une seconde colonne échangeuse d'anions : mono-Q.

Comme nous l'avons déjà vu précédemment, pour résoudre la structure de NikA, nous souhaitions réaliser une expérience du type MAD sur la protéine séléniée. Mais le rendement de l'ancienne purification (décrite précédemment), était trop faible (environ 7 mg de protéine pour 100 g de bactérie) pour pouvoir envisager la production de NikA séléniée. En effet la toxicité de la séléno-méthionine vis-à-vis des bactéries provoque une diminution du rendement de la surexpression. La première étape a donc été de mettre au point, sur la protéine native, un nouveau protocole de purification.

1.2.2.1 La purification de la protéine native avec extraction périplasmique à l'EDTA

1.2.2.1.1 L'extraction périplasmique à l'EDTA

L'extraction périplasmique permet, à partir de bactéries non lysées, de séparer facilement les protéines du périplasme (dont fait partie NikA), des autres protéines de la cellule, en cassant uniquement la membrane externe. Les bactéries sont resuspendues dans un tampon Tris base – EDTA 50 mM pH 9 (van der Westen *et al.*, 1978), avec une pastille contenant un cocktail d'inhibiteurs de protéases. Le tout est alors placé à 37°C sous agitation douce pendant 30 à 45 minutes. Pour terminer l'extraction, on centrifuge la solution à 13000 rpm, à 4°C pendant 60 minutes.

1.2.2.1.2 La précipitation au sulfate d'ammonium

Après la centrifugation de l'étape précédente, on récupère le surnageant et on ajoute 40% de sulfate d'ammonium. On laisse alors sous agitation pendant 20 minutes. On centrifuge à 18000 rpm, à 4°C, pendant 60 minutes pour récupérer les protéines précipitées dont fait partie NikA. Le surnageant est récupéré, et dialysé trois fois afin de mettre la protéine en solution dans du Tris base 40 mM pH 7,5, pour l'étape suivante.

1.2.2.1.3 La purification par chromatographie sur colonne Q – sépharose (Pharmacia)

Il s'agit d'une colonne échangeuse d'anions. Elle est équilibrée avec un tampon 40 mM Tris pH 7,5. A ce pH, NikA est chargée négativement (pI = 5,3 ; déterminé par séparation sur gel IEF), et par conséquent peut être retenue par la colonne. L'élution se fait alors par un gradient continu de 0 à 450 mM de NaCl en 10 volumes de colonne.

1.2.2.1.4 La purification par chromatographie sur colonne Resource – Q (Pharmacia)

La colonne Resource – Q est aussi une échangeuse d'anions, mais elle est plus résolutive que la Q – Sépharose et va donc permettre d'améliorer la pureté de la solution protéique. Mais avant d'injecter l'échantillon sur la colonne, on doit lui enlever toute trace de sels, ce que l'on réalise grâce une dialyse contre du tampon Tris base 40 mM pH 7,5. La colonne est équilibrée avec du tampon 40 mM Tris pH 7,5 et les protéines sont éluées par un gradient continu de 0 à 250 mM de NaCl en 25 volumes de colonne. La purification se termine alors par quatre lavages de la protéine sur centriprep 30 kDa, puis une concentration sur centricon avec un seuil de coupure 10 KDa.

1.2.2.2 La purification de la protéine séléniée avec extraction périplasmique à l'EDTA

La séléno-méthionine étant plus sensible à l'oxydation que la méthionine, on ajoute à tous les tampons de la purification 5 mM de DTT, afin de maintenir l'intégrité de la protéine en la gardant en milieu réducteur. C'est le seul élément qui ait été modifié par rapport au protocole de purification de la protéine native pour l'adapter à la protéine séléniée.

1.2.2.3 La purification de la protéine native avec extraction périplasmique au chloroforme

1.2.2.3.1 L'extraction périplasmique au chloroforme

Le protocole est décrit par Ames *et al.* (Ames *et al.*, 1984). Dans un premier temps, on resuspend les bactéries à l'aide d'un volume minimum de Tris base 10 mM pH 7,5. On ajoute ensuite un volume de chloroforme déterminé en fonction de la masse de bactéries utilisées (environ 2,5 mL de chloroforme par gramme de bactéries), et une pastille contenant un

cocktail d'inhibiteurs de protéases (sans EDTA). On laisse la solution sous agitation pendant 3 heures en moyenne.

Pour stopper l'extraction, on ajoute, à la solution bactérienne, un volume dix fois supérieur au volume de chloroforme, en Tris base 10 mM pH 7,5.

On centrifuge la solution obtenue pendant 30 minutes à 7200 tours/min, et on récupère le surnageant.

1.2.2.3.2 La précipitation au sulfate d'ammonium

On réalise ensuite une précipitation au sulfate d'ammonium à 40%, comme dans le protocole débutant par une extraction périplasmique à l'EDTA (chapitre 1.2.2.1.1 du matériel et méthodes). Mais au vu des volumes importants que l'on doit manipuler dans ce cas, on ajoute une étape de précipitation au sulfate d'ammonium à 80%, ayant pour objectif de concentrer les protéines. Après centrifugation, le précipité est resuspendu dans 50 mL de Tris base 40 mM pH 7,5.

1.2.2.3.3 La suite de la purification

La protéine est alors dialysée contre du Tris base 40 mM pH 7,5, et le reste de la purification est similaire à celle décrite dans le chapitre 1.2.2.1.1 du matériel et méthodes.

1.2.3 L'estimation de la concentration de protéine

La concentration de protéine a été estimée en mesurant la densité optique de la solution à 280 nm. On utilisa les valeurs du coefficient d'extinction molaire suivantes pour réaliser le calcul :

- NikA : 56,273 kDa ; $\varepsilon = 72,83\ mM^{-1}cm^{-1}$ (de Pina *et al.*, 1995) ;
- NikR : 15,094 kDa ; $\varepsilon = 4,720\ mM^{-1}cm^{-1}$ (valeur calculée).

1.2.4 L'analyse par gel de polyacrylamide

L'analyse de la pureté des solutions protéiques se fit par électrophorèse sur gel de polyacrylamide en conditions dénaturantes (SDS –PAGE).

L'électrophorèse est une technique permettant la migration de particules chargées sous l'effet d'un champ électrique. Dans notre cas, ce sont des protéines dénaturées par l'action du SDS (sodium dodécyl sulfate CH_3-$(CH_2)_2$-OSO_3^- Na^+), et de β-mercaptoéthanol (permettant la réduction des ponts disulfures).

La migration se fait dans un gel de polyacrylamide, résultant de la polymérisation de l'acrylamide (CH_2=CH-CO-NH_2) et du méthylène bis acrylamide ((CH_2=CH-CONH)$_2CH_2$). Ceci permet alors de séparer les protéines en fonction de leur masse : plus la molécule est petite, plus elle va migrer rapidement dans le gel. La taille des mailles du gel est fonction de la quantité d'acrylamide polymérisé, et plus elle est élevée, plus on sépare les petites molécules. Dans notre cas, les gels ont été réalisés avec 12% de polyacrylamide.
La visualisation des protéines se fait par coloration au bleu de Coomassie.

1.2.5 L'estimation de la concentration de fer

La concentration du fer dans certains échantillons de protéines a été estimée par la méthode colorimétrique de Fish utilisant le disulfonate de bathophénanthroline après dénaturation de la protéine par l'acide perchlorique (Fish, 1988). Les concentrations sont estimées par comparaison avec une gamme étalon de fer. Les résultats sont mesurés par la D.O. des échantillons à 535 nm.

2 La cristallographie aux rayons X des protéines

2.1 Le principe et l'objectif de la cristallographie des protéines

La cristallographie aux rayons X est une technique physique permettant d'observer la structure des protéines, des ADN et des ARN, c'est-à-dire l'organisation tridimensionnelle de biomolécules. En biologie, la structure est toujours étroitement liée à la fonction, par conséquent, ces études permettent de mieux comprendre le fonctionnement moléculaire des systèmes vivants.

La technique s'apparente au principe de fonctionnement d'un microscope optique : lorsque la lumière traverse l'objet observé, les rayons lumineux sont diffusés, puis reconcentrés par une série de lentilles, ce qui permet de reconstituer l'image (Figure 2.1-1a).

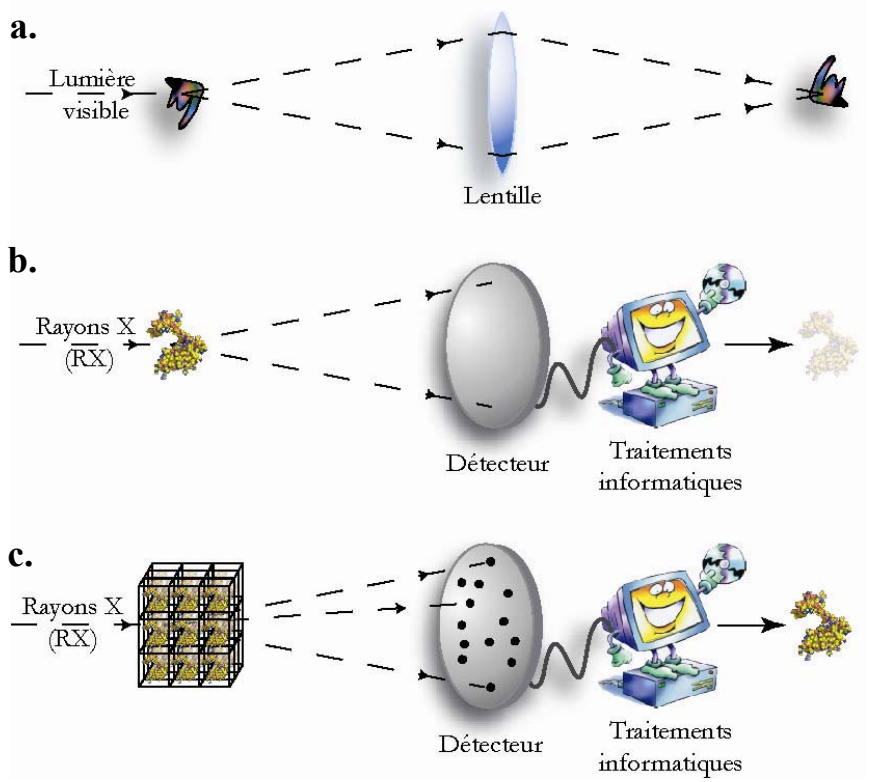


Figure 2.1-1 : Principe de la cristallographie aux rayons X

Le même principe peut s'appliquer à l'observation d'atomes, mais dans ce cas, il faut utiliser des ondes électromagnétiques dont la longueur d'onde est de l'ordre de grandeur de l'objet observé. Pour visualiser l'arrangement des atomes dans une molécule, on utilise donc les RX, dont la longueur d'onde est de l'ordre de l'Angström (10^{-10} m). Malheureusement, il n'existe

pas à l'heure actuelle de lentille capable de refocaliser les RX, et l'on ne peut donc pas appliquer directement le principe du microscope pour la visualisation d'atomes. La solution est alors de remplacer les lentilles par un traitement informatique (basé sur la transformée de Fourier). Les rayons diffusés par la molécule sont enregistrés par un détecteur. La transformée de Fourier de ces données génère l'image des nuages d'électrons qui correspondent aux positions atomiques (Figure 2.1-1b).

L'utilisation d'une seule molécule est impossible, du fait de la faible intensité des rayons diffusés. La solution est donc d'utiliser, à la place d'une molécule unique, un cristal, c'est-à-dire un empilement moléculaire régulier et périodique dans les trois dimensions de l'espace (Figure 2.1-2). L'utilisation du cristal a pour effet l'apparition d'un phénomène de diffraction « discrète » : les rayons X ne sont plus diffusés dans toutes les directions, mais seulement dans un nombre de directions limité, à cause des phénomènes d'interférences constructives et destructives du rayonnement. Ceci permet alors une amplification du signal de diffusion dans les directions d'interférences constructives. Les clichés enregistrés sur le détecteur sont donc appelés clichés de diffraction (Figure 2.1-1c).

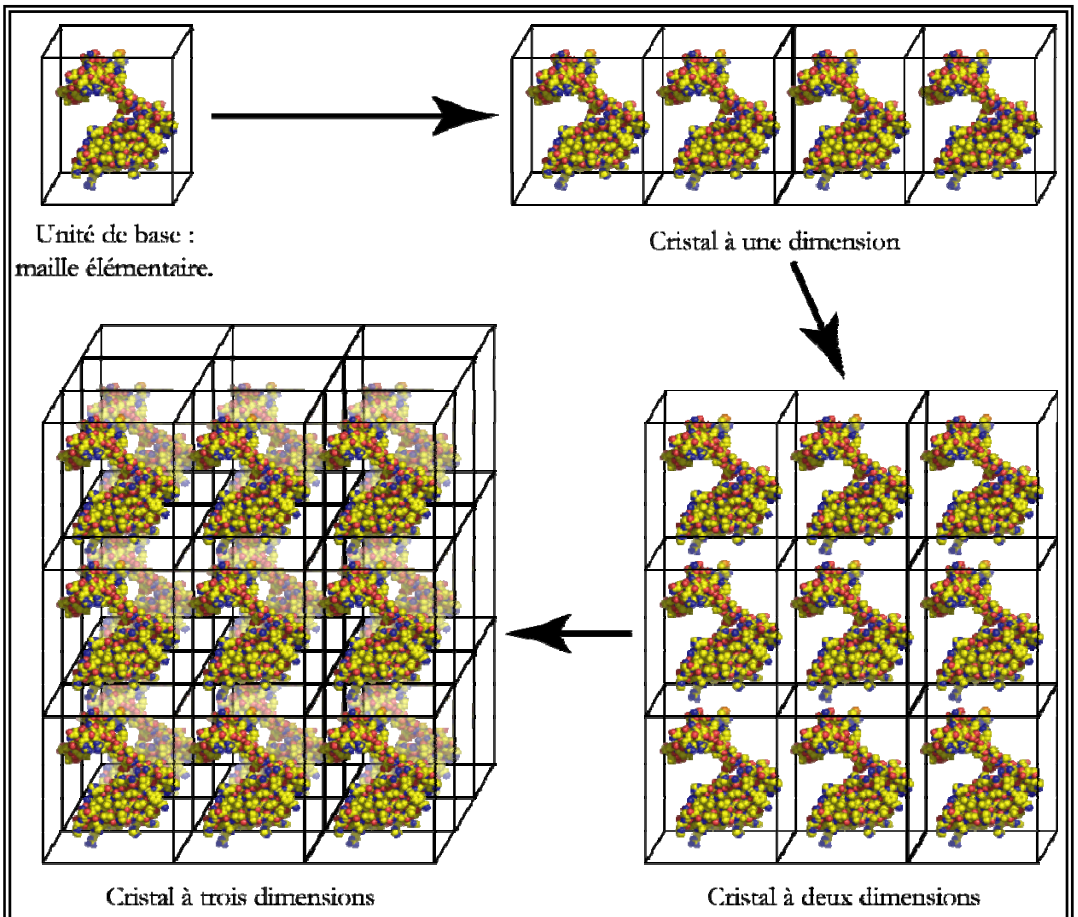


Figure 2.1-2 : Schéma d'un cristal de protéine.

2.2 La cristallogenèse

2.2.1 Le principe de la cristallisation

2.2.1.1 Les paramètres influençant la solubilité des protéines

La cristallisation d'une protéine est le passage de cette dernière d'un état liquide vers un état solide organisé (chapitre 1.2 (Drenth)), et s'effectue par une diminution de sa solubilité en fonction de facteurs tels que :

- La nature et la concentration des agents précipitants. On peut les séparer en trois familles majoritaires :
  - ⇒ Les polymères qui agissent par exclusion du solvant au voisinage de la protéine (le PEG est le plus utilisé) ;
  - ⇒ Les sels non chaotropiques (non dénaturants pour les protéines) qui modifient la force ionique et augmentent les interactions hydrophobes entre les molécules de protéine en excluant le solvant (le sulfate d'ammonium, les acétates ou les chlorures…) ;
  - ⇒ Les solvants organiques qui diminuent la constante diélectrique du milieu (MPD, isopropanol, dioxane…).
- La nature et le pH du tampon, comme par exemple le Mes, l'HEPES, ou le Tris ;
- La nature et la concentration d'additifs, comme par exemple des détergents ;
- La concentration de la protéine ;
- La température ;
- La pureté et l'homogénéité de la protéine.

2.2.1.2 Le diagramme de phase

Le diagramme de phases, qui est obtenu en faisant varier la concentration en agent précipitant (ou tout autre paramètre influençant la cristallisation : concentration en sel, pH, température…) en fonction de la concentration de protéine, met en évidence l'existence de différentes zones (Figure 2.2-1) :

- La zone de solubilité de la protéine (sous-saturation), où la protéine reste en solution. Dans ces conditions aucun cristal ne peut être obtenu ;

- La zone de précipitation (qui appartient à la zone de sursaturation), où les conditions provoquent l'apparition d'un précipité amorphe ou cristallin, ou de germes cristallins ;
- La zone métastable (qui appartient également à la zone de sursaturation), où les contraintes thermodynamiques sont favorables à l'apparition d'un précipité ou d'un germe cristallin (trop petit pour être visible), mais la cinétique est trop lente pour permettre la formation d'un cristal. C'est la zone la plus adaptée à la réalisation des techniques d'ensemencement.

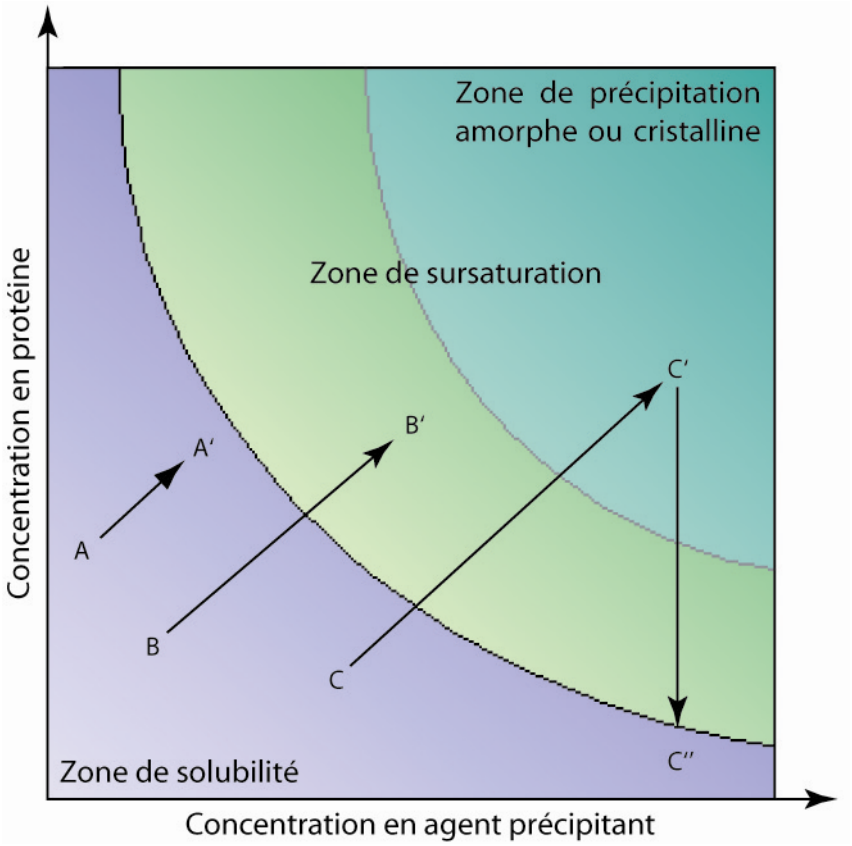


Figure 2.2-1 : Diagramme de phases : chemins expérimentaux en technique de goutte suspendue à partir des conditions initiales A, B et C.

La zone de solubilité et celle de sursaturation sont séparées par la courbe de saturation : condition d'équilibre où un cristal ajouté à la solution ne pourra ni croître ni se dissoudre.

Lors d'une expérience de cristallisation, on va essayer de mettre la protéine dans des conditions permettant l'apparition d'un minimum de germes cristallins en un laps de temps raisonnable, c'est à dire de quelques heures à quelques semaines.

On peut également noter que dans une expérience de cristallisation, toutes les solutions qui entrent dans la composition de la condition de cristallisation sont passées sur filtre 0,22 µm, afin d'éliminer au maximum tous les éléments qui pourraient provoquer une modification de la cristallisation (nucléation hétérogène, non cristallisation…).

2.2.2 Les techniques de cristallisation des protéines utilisées durant cette thèse

2.2.2.1 La diffusion de vapeur en goutte suspendue

La technique de cristallisation de diffusion de vapeur en goutte suspendue est l'une des plus utilisées (Wlodawer and Hodgson, 1975; McPherson *et al.*, 1995).

Son principe est le suivant : on dépose une goutte contenant la protéine purifiée et une concentration c_1 d'agent précipitant, sur une lamelle siliconée (pour éviter que la goutte ne s'étale), qui est alors retournée sur un puits hermétiquement fermé par de la graisse (silicone), contenant une solution réservoir d'agent précipitant à la concentration c_2 supérieure à c_1. Cette différence de concentration, entre la goutte et le réservoir, va progressivement disparaître grâce à des échanges par diffusion de vapeur ; et comme le volume du réservoir est très supérieur à celui de la goutte, cette dernière va voir son volume diminuer et sa concentration en agent précipitant augmenter jusqu'à s'approcher de c_2 (Figure 2.2-2).

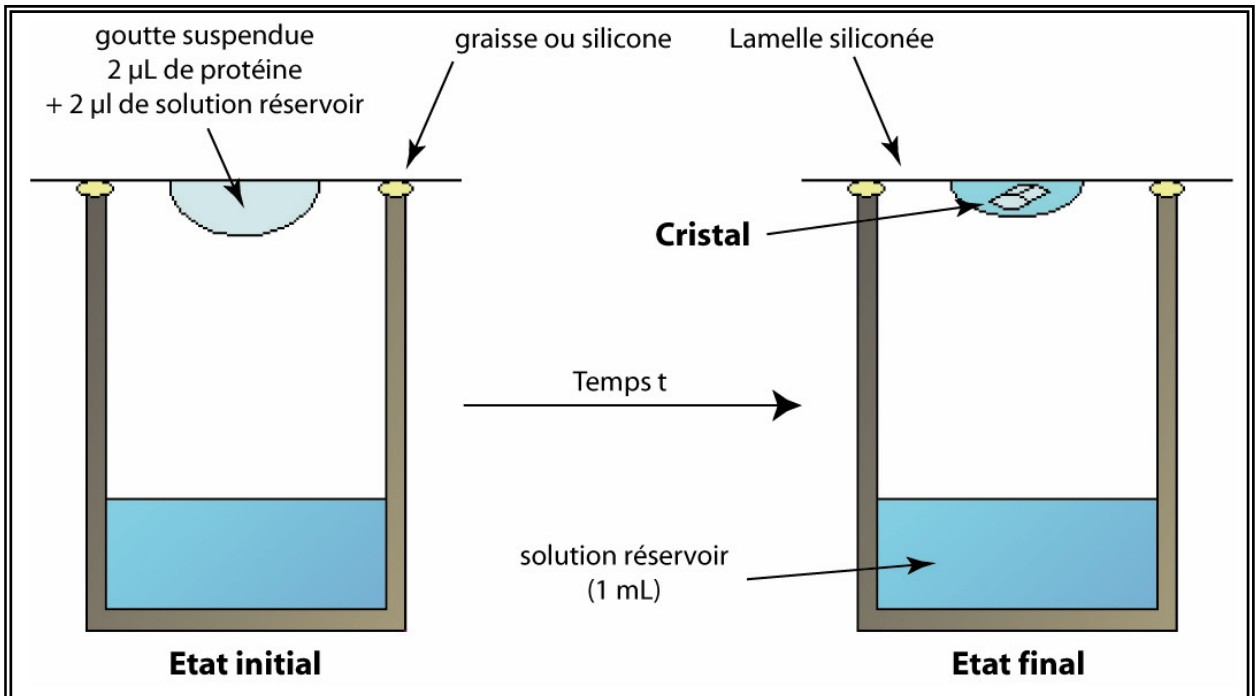


Figure 2.2-2 : Schéma du principe de la diffusion de vapeur en goutte suspendue

Il existe aussi des boîtes de cristallisation (Nextal), où l'on visse la lamelle sur le réservoir, et où l'étanchéité est assurée par un joint.

En diminuant de volume, les concentrations de protéine et d'agent précipitant dans la goutte vont augmenter, et si les conditions sont remplies, un phénomène de nucléation (homogène ou hétérogène) va se produire, conduisant à l'apparition des cristaux.

La Figure 2.2-1 montre, sur le diagramme de phases, les différents chemins rencontrés lors d'une expérience de cristallisation par diffusion de vapeur en goutte suspendue :

- La première étape est toujours une augmentation des concentrations de la protéine et de l'agent précipitant (chemins A→A', B→B', et C→C') ;
- Dans le cas de la condition A, après l'équilibration de la goutte, la protéine est toujours dans la zone de solubilité, par conséquent aucun cristal ne peut être obtenu ;
- Pour la condition B, après équilibration, la protéine se trouve dans la zone métastable. Là non plus, aucun cristal ne pourra apparaître dans un délai raisonnable ;
- Pour la condition C, après équilibration, on se trouve dans la zone de précipitation, ce qui va provoquer l'apparition d'un précipité (cristallin ou amorphe) ou de germes cristallins. La concentration de la protéine va alors diminuer jusqu'à atteindre la courbe de solubilité (C''), conjointement à la croissance du cristal.

2.2.2.2 L'ensemencement

2.2.2.2.1 Le principe de l'ensemencement

Le principe de l'ensemencement est d'introduire des germes cristallins ou un monocristal dans une goutte de cristallisation prééquilibrée pour se trouver dans la zone métastable, ce qui permet de solutionner le problème de nucléation, et ainsi obtenir des cristaux de taille raisonnable. Deux solutions sont alors possibles :

- Soit réaliser un macro-ensemencement, en ajoutant à la goutte de cristallisation un cristal, provenant d'une autre goutte, que l'on veut faire grossir ;
- Soit effectuer un micro-ensemencement, en amenant dans la goutte de cristallisation des germes cristallins, à l'aide d'un cil, qui permettront l'apparition et la croissance de cristaux.

Cette technique impose toutefois l'obtention préalable de cristaux, même s'ils sont de mauvaise qualité : le micro-ensemencement pouvant éventuellement l'améliorer.

2.2.2.2.2 Le protocole du micro-ensemencement utilisé pour reproduire les cristaux de NikA

Le micro-ensemencement est réalisé par diffusion de vapeur en goutte suspendue. Les gouttes sont composées de 2 µL de solution de protéine à 11 mg/mL additionnée de 2 µL de solution réservoir. On les laisse s'équilibrer pendant 24 heures, puis on amène dans la goutte, un germe cristallin, que l'on obtient en touchant un cristal préalablement formé avec, dans notre cas, un poil de moustache de chien.

Dans le cas de nucléation trop importante, il est possible de laver le poil avant de le passer dans la goutte à ensemencer, afin d'enlever l'excès de germes. La solution de lavage a la même composition que celle du réservoir, mais avec une concentration moindre de précipitant. Pour la déterminer, on effectue des tests de stabilité des cristaux en fonction de la concentration de précipitant : on réalise une gamme de ce dernier, et on place un cristal par condition. On choisit finalement comme solution de lavage, une condition un peu plus concentrée que celle qui a commencé à abîmer le cristal pour que, lors du lavage, la majorité des germes soit dissoute, et que l'on n'amène dans la goutte de cristallisation que quelques uns d'entre eux.

2.3 L'enregistrement et le traitement des données de diffraction

2.3.1 La cryocristallographie

2.3.1.1 L'intérêt de la congélation des cristaux de protéine

Les sources intenses, comme les synchrotrons, peuvent provoquer, lors de l'enregistrement de données de diffraction, un dommage important dû au rayonnement X : il peut y avoir formation de radicaux libres qui diffusent et endommagent les molécules du cristal (réduction des métaux, coupure des ponts disulfures…), entraînant une diminution importante de la qualité de diffraction. La solution a été apportée par l'apparition de la cryocristallographie : le cristal est amené à une température inférieure à 100K par refroidissement rapide dans de l'azote gazeux ou liquide, ce qui limite considérablement la propagation du dommage radiatif, permettant ainsi une bonne qualité de diffraction pendant une longue durée. La congélation peut également se faire dans du propane liquide. En effet,

l'azote a une faible chaleur latente de vaporisation. Cela entraîne un phénomène de caléfaction lorsque le cristal est plongé dans l'azote (apparition d'une couche de vapeur d'azote entre le cristal et le liquide), ce qui peut empêcher une congélation rapide optimale surtout si le cristal est de grande taille. Dans le cas du propane, ce phénomène n'apparaît pas et la congélation est de meilleure qualité.

Pendant toute l'expérience d'enregistrement des données de diffraction, le cristal est maintenu dans un flux d'azote gazeur à 100K. L'utilisation des techniques de congélation sont devenues aujourd'hui quasi-systématiques.

2.3.1.2 Les cryoprotectants

Si l'on refroidit un cristal tel quel, c'est à dire en présence uniquement de sa solution de cristallisation, l'eau que contiennent le cristal (de 30 à 80% en fonction de l'empilement des protéines dans le cristal) et la solution de cristallisation, peut cristalliser, ce qui va l'abîmer de façon très importante (effets mécaniques liés à l'augmentation du volume de la maille cristalline), et l'on va en plus visualiser la diffraction due à la glace, ce qui va gêner l'indexation et l'intégration des taches de diffraction. Pour que l'eau ne forme pas de glace et se fige de façon amorphe, on ajoute au cristal un cryoprotectant, comme par exemple le PEG 400, le glycérol, le MPD… Le choix du cryoprotectant impose généralement plusieurs tests pour déterminer sa nature et sa concentration optimales, car souvent, son ajout perturbe l'équilibre et peut endommager le cristal. La solution cryoprotectante peut être ajoutée, en fonction du cristal, soit directement dans la goutte, soit par trempages successifs par paliers, soit par trempage rapide.

2.3.2 La diffraction des rayons X par les cristaux

2.3.2.1 La direction des rayons X diffractés par un cristal

Comme nous l'avons expliqué plus haut, lorsqu'un cristal est soumis à un rayonnement X, une partie du faisceau de rayons X va être diffusée. Dans le cadre de la théorie ondulatoire, ces derniers peuvent êtres représentés par des ondes qui vont alors interagir par interférences constructives et destructives, ce qui constitue le phénomène de la diffraction : des directions particulières de l'espace vont être privilégiées en fonction de la dimension et de la géométrie de la maille cristalline.

Mathématiquement, les équations de Laue traduisent ce phénomène en reliant le vecteur de diffusion $\vec{S}$ ($\vec{S} = \vec{s}_0 - \vec{s}$, avec $\vec{s}_0$ la direction du faisceau incident, et $\vec{s}$ la direction du

faisceau diffracté) aux vecteurs définissant la maille du cristal ($\vec{a}$, $\vec{b}$ et $\vec{c}$). On a alors un faisceau diffracté lorsque :

$$\vec{S}.\vec{a} = h \qquad \vec{S}.\vec{b} = k \qquad \vec{S}.\vec{c} = l \qquad \text{avec h, k, et l entiers.}$$

Ces équations permettent de définir une représentation mathématique de l'ensemble des directions de diffraction : l'espace réciproque, qui a pour base :

$$\vec{a}^* = (\vec{b}\wedge\vec{c})/V \qquad \vec{b}^* = (\vec{c}\wedge\vec{a})/V \qquad \vec{c}^* = (\vec{a}\wedge\vec{b})/V$$

avec V le volume de la maille cristalline.

Le vecteur de diffusion peut ainsi s'écrire : $\vec{S} = h.\vec{a}^* + k.\vec{b}^* + l.\vec{c}^*$; h, k et l correspondant alors aux indices des réflexions (Figure 2.3-1a).

Une deuxième représentation mathématique similaire de la diffraction est la loi de Bragg. Dans ce cas, on a diffraction pour l'angle 2θ (angle entre les faisceaux incident et diffracté) vérifiant : $2d_{hkl} \cdot \sin\theta = n \cdot \lambda$, avec d_{hkl} la distance interréticulaire de la famille de plans sur laquelle se « réfléchit » le faisceau direct, n un entier, et λ la longueur d'onde des rayons X. L'égalité $\frac{2\sin\theta}{\lambda} = \frac{1}{d_{hkl}}$ montre alors la relation réciproque entre l'espace direct représenté par d_{hkl} et la résolution des ondes diffractées déterminée par l'angle θ (Figure 2.3-1b).

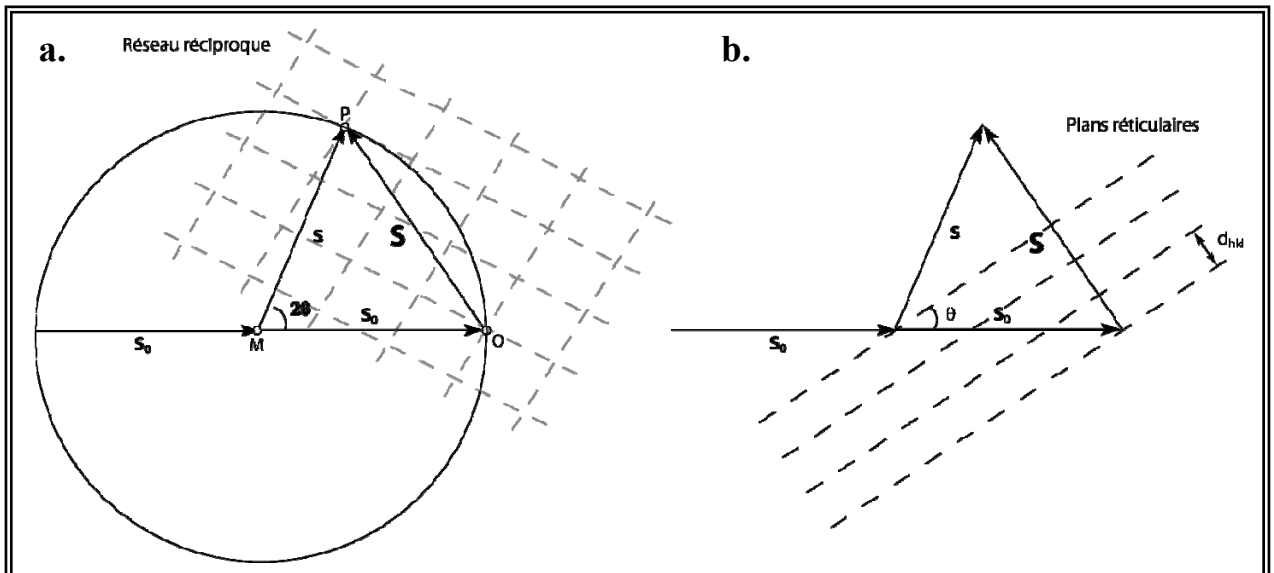


Figure 2.3-1 : Schéma décrivant le phénomène de la diffraction selon les lois de Laue (a.) et de Bragg (b.)

2.3.2.2 Le facteur de structure

L'onde diffusée par un objet, ici le cristal, est la somme de la diffusion de tous les atomes de l'objet et peut donc être représentée mathématiquement comme la transformée de

Fourier de cet objet. En cristallographie, un cristal est représenté comme le produit de convolution d'une fonction motif (la densité électronique des atomes contenus dans la maille du cristal) par une fonction réseau qui répète ce motif par translation dans le cristal (définie par les paramètres de maille). Or, la transformée de Fourier échange produit ordinaire et produit de convolution : la transformée de Fourier d'un produit de convolution est égale au produit des transformées de Fourier. Les ondes diffractées correspondent donc à la transformée de Fourier de la densité électronique du motif, échantillonnée aux nœuds du réseau réciproque. Ces réflexions peuvent être représentées par un nombre complexe $F(\vec{S})$ appelé facteur de structure :

$$F(\vec{S}) = \iiint_V \rho(\vec{r}) e^{i2\pi\vec{r}\vec{S}} dV$$

avec $\vec{r}$ et $\rho(\vec{r})$ respectivement le vecteur position dans l'espace réel, et la densité électronique (Figure 2.3-2).

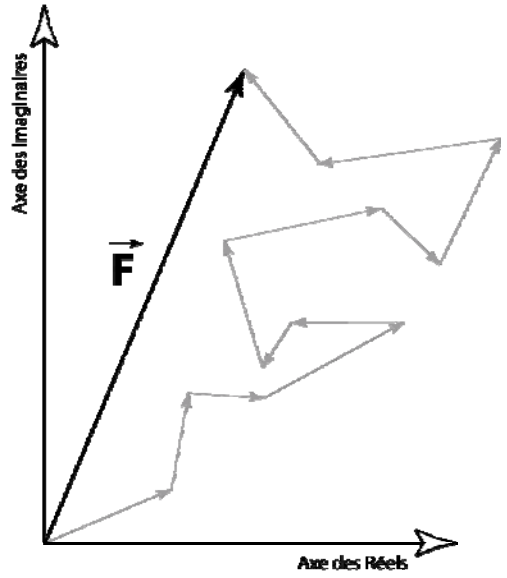


Figure 2.3-2 : Représentation dans le plan complexe du facteur de structure $F(\vec{S})$ comme la somme de la contribution du facteur de diffusion f_j de chaque atome de l'unité cristalline du cristal.

Etant donné qu'en cristallographie des protéines on n'atteint généralement pas la résolution suffisante pour visualiser les orbitales atomiques, on considère la densité électronique de chaque atome comme étant sphérique et centro-symétrique. Pour chaque type d'atomes, un facteur de diffusion atomique f_j est calculé. Cette approximation permet alors de réécrire le facteur de structure sous la forme d'une série de Fourier :

$$\vec{F}_{hkl} = \sum_{j=1}^{n} q_j f_j e^{-B_j\left(\frac{\sin\theta}{\lambda}\right)^2} e^{i2\pi(hx_j+ky_j+lz_j)} = \left|F_{hkl}\right| e^{i\varphi_{hkl}}$$

avec q_j, B_j et x_j, y_j, z_j, respectivement l'occupation, le facteur d'agitation thermique et les trois coordonnées de l'atome j. Le facteur d'agitation thermique peut être défini comme un facteur anisotropique dans le cas où le nombre de réflexions mesurées est suffisant. Cela permet de mieux définir les mouvements de l'atome, et ainsi de mieux caractériser le facteur de structure.

On voit ici que le facteur de structure peut être défini par une amplitude $\left|F_{hkl}\right|$ et une phase φ_{hkl}.

En appliquant la transformée de Fourier inverse, nous pouvons retrouver l'image de la densité électronique du motif diffractant :

$$\rho(\vec{r}) = \frac{1}{V}\sum_h \sum_k \sum_l \vec{F}_{hkl} e^{-i2\pi(hx_j+ky_j+lh_j)}$$

2.3.3 L'enregistrement des données de diffraction

Lors d'une expérience de diffraction des rayons X par un cristal, on enregistre à l'aide d'un détecteur bidimensionnel l'intensité des faisceaux diffractés qui est proportionnelle au facteur de structure (HERCULES) :

$$I_{hkl} = \frac{\lambda^3}{\omega \cdot V^2} \times \left(\frac{e^2}{mc^2}\right)^2 \times V_{cr} \times I_0 \times L \times P \times T_r \times \left|\vec{F}_{hkl}\right|^2 \propto \left|\vec{F}_{hkl}\right|^2$$

avec :

- I_{hkl} et $\vec{F}_{hkl}$ respectivement l'intensité mesurée et le facteur de structure de la réflexion hkl ;
- λ la longueur d'onde du faisceau de rayons X, ω la vélocité de la rotation du cristal et V le volume de la maille cristalline ;
- e la charge de l'électron, et m sa masse ;
- c la vitesse de la lumière ;
- V_{cr} le volume du cristal ;
- I_0 l'intensité du faisceau incident de rayons X ;
- L le facteur de Lorentz ;
- P le facteur de polarisation du faisceau de rayons X ;

- et T_r le facteur de transmission du faisceau de rayons X.

Les images ainsi enregistrées sont appelées clichés de diffraction (Figure 2.3-3), et correspondent à l'exposition du cristal au faisceau monochromatique de rayons X, associé à la rotation de l'échantillon protéique d'en général 0,5 à 1° selon un axe parallèle au plan du détecteur, et perpendiculaire aux RX. Cette méthode permet de rapidement mettre en conditions de diffraction une grande partie des réflexions de l'espace réciproque. Afin d'obtenir un jeu complet de diffraction, il convient d'enregistrer un nombre d'images de diffraction correspondant à des oscillations successives, permettant de d'enregistrer la totalité des taches de diffraction non reliées par symétries. Le temps d'exposition dépend de l'intensité du faisceau incident et du pouvoir de diffraction du cristal.

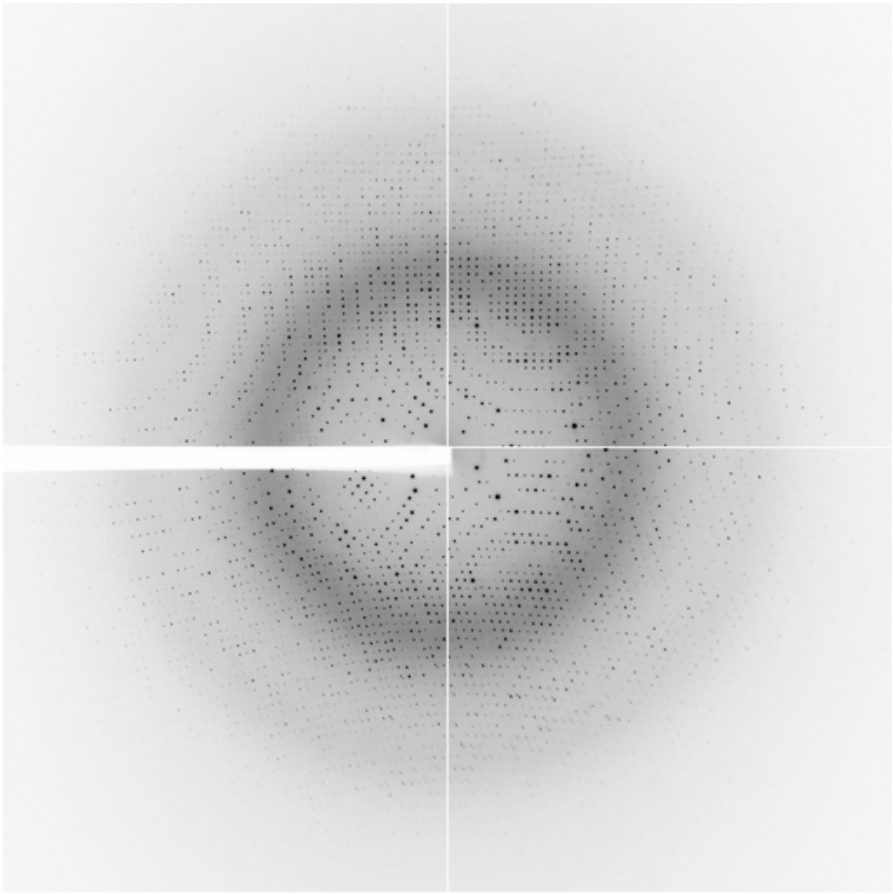


Figure 2.3-3: Cliché de diffraction de la protéine NikA.

2.3.3.1 Le diffractomètre de laboratoire

Les sources classiques de laboratoire telles que les anodes tournantes à cible de cuivre, produisent un rayonnement X monochromatique, de longueur d'onde λ = 1,54 Å, non

polarisé. La durée d'exposition du cristal aux rayons X permettant d'enregistrer un cliché de diffraction est dans ce cas comprise entre 15 min et 1 heure, ce qui implique que l'enregistrement d'un jeu complet de diffraction prend plusieurs heures, ce qui est très long.

2.3.3.2 Le synchrotron

Les synchrotrons sont de grands instruments dont la fonction est la production de rayons X à partir d'électrons accélérés à une vitesse proche de celle de la lumière. Le rayonnement produit est beaucoup plus intense que les sources de laboratoire, blanc (le faisceau est polychromatique et continu dans une gamme de longueur d'onde autour d'environ 1,5 Å), peu divergent, et polarisé. Dans ce cas, la durée d'exposition du cristal aux rayons X permettant d'enregistrer un cliché de diffraction est comprise entre 0,1 s et quelques minutes au maximum, réduisant ainsi considérablement la durée de l'expérience.
Toutes les données présentées dans cette thèse ont été enregistrées sur les différentes lignes de lumière de l'ESRF (*European Synchrotron Radiation Facility*) dédiées à la biocristallographie (BM30A, ID29, ID14eh1-4…).

2.3.4 Le traitement des données de diffraction

2.3.4.1 Le principe

Une fois que l'on a enregistré un jeu complet de diffraction, on effectue un traitement de ces données qui se décompose en trois étapes :

- L'indexation : cela consiste en la détermination de l'orientation et des paramètres de la maille du cristal ($|\vec{a}|$, $|\vec{b}|$ et $|\vec{c}|$, le module des trois vecteurs directeur de la maille, ainsi qu'α, β et γ, les angles qu'ils forment entre eux), afin de pouvoir redonner à chaque réflexion ses indices hkl et ensuite intégrer leur intensité pour chaque image séparément ;
- L'intégration des intensités : la dispersion spatiale des raies de diffraction, due essentiellement à la mosaïcité du cristal, à la largeur spectrale et à la divergence du faisceau incident, a pour conséquence d'étaler une tache sur un diamètre de moins d'un à quelques millimètres. Les programmes de traitement permettent d'établir les profils moyens de ces taches et de calculer l'intensité de chaque réflexion par intégration numérique suivant ces profils calculés. Il est aussi nécessaire de corriger les effets d'absorption et de polarisation du faisceau incident ;

- La mise à l'échelle : lorsque une réflexion et/ou ses symétriques sont mesurées plus d'une fois sur deux images différentes, leur intensité doit être théoriquement la même. Mais en pratique ce n'est pas le cas à cause de différents paramètres comme la dégradation du cristal soumis au flux de rayons X ou la forme du cristal, la décroissance du rayonnement synchrotron… On effectue alors une mise à l'échelle globale des intensités, en déterminant pour chaque image un facteur de mise à l'échelle k linéaire et un facteur d'agitation thermique B en fonction de la résolution. Par la suite, les intensités de ces réflexions équivalentes sont moyennées et le module des facteurs de structure est calculé : c'est la réduction des données qui permet d'obtenir l'ensemble des réflexions uniques.

2.3.4.2 Les programmes utilisés

Plusieurs programmes existent pour réaliser le traitement de données de diffraction, les résultats présentés dans cette thèse ont tous été obtenus par une indexation et une mise à l'échelle par le programme XDS (Kabsch, 1993). Il est composé de sept sous-programmes qui s'enchaînent automatiquement. Il a l'avantage de déterminer des profils tridimensionnels des réflexions. Ceci lui permet de mieux traiter que les autres programmes des cas où les cristaux sont mosaïques et pour lesquels l'intensité d'une tache peut donc être répartie sur plusieurs images.

2.3.4.3 La validation des données de diffraction

La qualité d'un jeu de données est déterminée par différents critères statistiques calculés au cours du traitement. Les paramètres les plus significatifs sont :

- La résolution : elle déterminera la finesse des détails dans l'image de la densité électronique. La limite haute de la résolution des données mesurées peut être imposée par la configuration de l'expérience (taille du détecteur, distance cristal-détecteur, longueur d'onde…) ou par la qualité de diffraction du cristal. Lors de cette thèse, cette limite a été fixée lorsqu'était atteint, pour la dernière tranche de résolution, l'un de ces deux critères : $I/\sigma(I) \leq 2$ et $R_{sym} \geq 30\%$;
- Le rapport signal sur bruit ($I/\sigma(I)$) : il permet d'avoir une estimation de l'intensité moyenne des réflexions mesurées en fonction de la résolution (avec I, l'intensité des réflexions, et $\sigma(I)$, l'écart type des intensités des réflexions, ce qui correspond au bruit de fond) ;

- La complétude : elle donne le pourcentage de réflexions mesurées par rapport au nombre total de taches mesurables, à la résolution donnée. Elle doit être la plus grande possible (supérieure à 90%) ;
- La redondance : elle établit le nombre moyen de fois qu'une réflexion et ses symétriques sont mesurés. Plus sa valeur est grande, meilleure sera l'estimation de l'intensité moyenne d'une réflexion unique en diminuant l'influence de l'erreur de mesure commise ;
- Le facteur R_{sym} : $R_{sym} = \frac{\sum_{hkl} \left| I_{hkl} - \langle I \rangle \right|}{\sum_{hkl} I_{hkl}}$ qui permet de comparer l'intensité de chaque réflexion équivalente par la symétrie du cristal (I_{hkl}) à la valeur moyenne (<I>) de ces réflexions. Plus ce facteur est bas, meilleure est la cohérence du jeu de données ;
- D'autres paramètres peuvent également être pris en compte pour évaluer la qualité des données de diffraction comme par exemple le graphique de Wilson qui, grâce à l'étude de la distribution des intensités en fonction de la résolution, permet d'estimer un facteur de mise à l'échelle et un facteur de température, mais aussi d'évaluer la présence d'une macle dans les cristaux (Yeates, 1997).

2.3.5 L'analyse des caractéristiques du cristal

2.3.5.1 Les systèmes cristallins et les groupes d'espace

Système cristallin	Critères d'appartenance	Exemples de groupe d'espace
Triclinique	$a \neq b \neq c$; $\alpha \neq \beta \neq \gamma$	P1
Monoclinique	$a \neq b \neq c$; $\alpha = \gamma = 90°$ $\beta \neq 90°$	$P2$, $P2_1$, $C2$
Orthorhombique	$a \neq b \neq c$; $\alpha = \beta = \gamma = 90°$	$P222$, $P2_12_12_1$, $P2_12_12$ …
Quadratique	$a = b \neq c$; $\alpha = \beta = \gamma = 90°$	$P4$, $P422$, $I4_122$ …
Rhombohédrique	$a = b = c$; $\alpha = \beta = \gamma \neq 90°$	$P3$, $P312$, $R32$ …
Hexagonal	$a = b \neq c$; $\alpha = \beta = 90°$ $\gamma = 120°$	$P6$, $P6_2$, $P6_522$ …
Cubique	$a = b = c$; $\alpha = \beta = \gamma = 90°$	$P23$, $P4_132$, $I432$ …

Tableau 2.3-1 : Systèmes cristallins et exemples de groupes d'espace possibles en biologie.

Les paramètres cristallins déterminés lors de l'étape d'indexation permettent de classer le cristal parmi les 7 systèmes cristallins et les 64 groupes d'espace possibles en biologie (parmi les 230 possibles en excluant les symétries inversant la chiralité) (Tableau 2.3-1).

2.3.5.2 La mosaïcité

La mosaïcité caractérise la qualité de l'empilement cristallin. En effet, les cristaux de protéine présentent souvent des défauts dont les origines sont multiples :

- Lors de la croissance cristalline, du fait du nombre limité de contacts assurant la cohésion cristalline, des perturbations mécaniques ou thermiques peuvent engendrer des erreurs dans l'empilement cristallin ;
- La congélation, si elle est mal maîtrisée, peut également faire augmenter la mosaïcité.

Typiquement, elle varie de 0,2 à 1,5°, et une forte mosaïcité risque d'entraîner des problèmes de superposition de réflexions, ou de causer des difficultés lors de la mesure des réflexions de faibles intensités.

2.3.5.3 Le désordre des atomes du cristal

Les protéines présentent souvent dans leur structure des zones flexibles et agitées. De plus, ce désordre est dépendant de la température et de la compacité du cristal (les cristaux de protéine contenant de 30 à 80% de solvant). Il se traduit par un affaiblissement global des intensités diffractées, qui peut être pris en compte dans le facteur de structure par un facteur d'agitation thermique global B. Le facteur B moyen est exprimé en $Å^2$, et reflète aussi bien le désordre statique que le désordre dynamique. Généralement, les atomes dans les cristaux de protéines possèdent des facteurs B moyens compris entre 10 et 70 $Å^2$.

2.3.5.4 L'estimation du nombre de molécules de l'unité asymétrique

L'unité asymétrique est le plus petit volume de la maille contenant des points non reliés entre eux par la symétrie cristallographique. Elle peut contenir plusieurs molécules qui sont alors reliées entre elles par des opérateurs de rotation-translation définissant une symétrie non cristallographique.

B. W. Matthews a montré que, pour la plupart des cristaux protéiques, le rapport (V_m) entre le volume de la maille (V) et le poids moléculaire contenu dans celle-ci était compris entre 2 et 4 $Å^3$/Da (Matthews, 1968; Kantardjieff and Rupp, 2003).

$$V_m = \frac{V}{M \cdot n \cdot Z_a}$$

avec V le volume de la maille, n le nombre de positions équivalentes, Z_a le nombre de molécules dans l'unité asymétrique et M la masse moléculaire de la protéine. La quantité de solvant ($S_{\%}$) peut quant à elle être estimée par la relation :

$$S_{\%} = 1 - \frac{1{,}23}{V_m}$$

1,23 g/cm^3 étant la valeur de la masse spécifique moyenne des protéines (0,807 Da/Å^3).
La méthode de Matthews permet donc, par comparaison, d'estimer la quantité de solvant dans le cristal ($S_{\%}$) ainsi que le nombre de molécules de protéine contenues dans l'unité asymétrique (Z_a), à partir du poids moléculaire de la protéine cristallisée et des paramètres de l'arrangement cristallin.

2.4 La résolution du problème de la phase

2.4.1 Le problème de la phase

Les détecteurs de rayons X ne sont capables de mesurer que l'intensité des rayons diffractés par le cristal, et la différence de phase qu'il existe entre eux est perdue, ce qui rend impossible le calcul de la densité électronique. Alors que dans le cas des petites molécules, ce problème est surmonté grâce aux méthodes directes, dans le cas des molécules biologiques il faut, en général, passer par des méthodes nécessitant pour la plupart l'enregistrement des données de diffractions provenant d'un nouveau cristal. Ces méthodes indirectes peuvent être classées en deux groupes :

- Celles nécessitant l'ajout au cristal d'un atome lourd, dont on pourra ou non utiliser les propriétés de diffuseur anomal ;
- Et le remplacement moléculaire, nécessitant l'utilisation d'un modèle de structure similaire et connue.

2.4.2 La diffusion anomale

2.4.2.1 L'origine de la diffusion anomale

Jusqu'à présent, nous avons considéré que le faisceau de rayons X interagissait avec les électrons en respectant la théorie de la diffusion cohérente (ou diffusion de Thomson), qui suppose que les électrons sont libres. Dans ce cas, les facteurs de structure respectent la loi de Friedel : $|F_{hkl}| = |F_{-h-k-l}|$ et $\varphi_{hkl} = -\varphi_{-h-k-l}$ (Blundell and Johnson, 1976) (Figure 2.4-1).

Par contre, lorsque l'énergie des rayons X est légèrement supérieure ou égale au seuil d'absorption d'un atome de la structure, on ne peut plus considérer les électrons de ce dernier comme libres, et leurs entrées en résonance avec les photons vont conduire à un phénomène de diffusion anomale. Dans ce cas, le facteur de diffusion de l'atome s'écrira :

$^{\lambda}f = {}^{\circ}f + {}^{\lambda}f' + i\,{}^{\lambda}f''$ avec $^{\circ}f$ le facteur de diffusion en l'absence de diffusion anomale, indépendant de λ et diminuant avec la résolution ; $^{\lambda}f'$ la partie anomale réelle (dispersive), introduite par les électrons de cœur, dépendant de λ, mais insensible à la résolution ; $^{\lambda}f''$ la partie anomale imaginaire due à l'absorption des rayons X par les électrons qui entrent en résonance, et réémettent le photon avec un déphasage de π/2. De la même façon que $^{\lambda}f'$, il est indépendant de la résolution et varie avec λ.

Cette modification va rompre la loi de Friedel, et dans ce cas $|F_{hkl}| \neq |F_{-h-k-l}|$ et $\varphi_{hkl} \neq \varphi_{-h-k-l}$ (Figure 2.4-1).

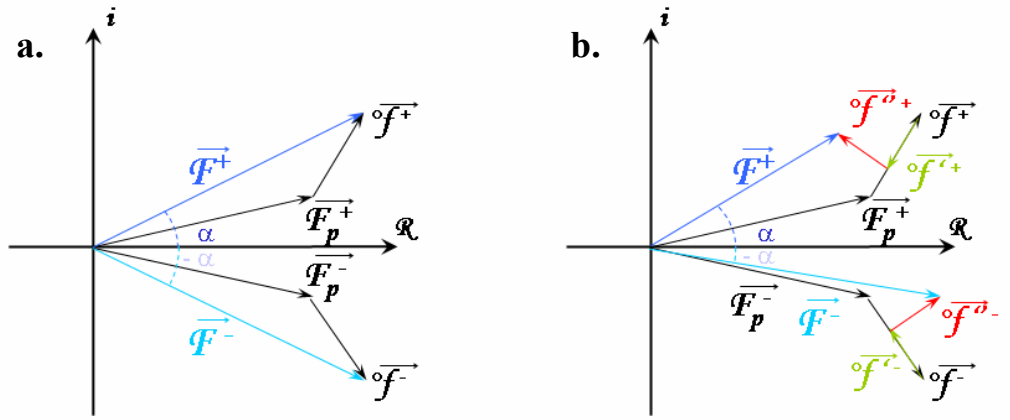


Figure 2.4-1 : Représentation vectorielle des facteurs de structure de la protéine ($\vec{F}_P$) contenant un diffuseur anomal ($^{0}\vec{f}$, $^{0}\vec{f}'$ et $^{0}\vec{f}''$). $\vec{F}$ correspondant à la protéine, plus le diffuseur anomal.

a. à une longueur d'onde induisant une absence de diffusion anomale.

b. à une longueur d'onde induisant une présence de diffusion anomale.

2.4.2.2 L'utilisation de la diffusion anomale pour résoudre le problème de la phase

2.4.2.2.1 Le MAD

Hendrickson *et al.* (Hendrickson *et al.*, 1985) ont montré qu'il est possible de séparer le signal anomal du signal normal selon l'équation :

$$^{\lambda}F_T(\vec{h}) = {}^{\circ}F_T(\vec{h}) + {}^{\lambda}F_T(\vec{h}) = {}^{\circ}F_T(\vec{h}) + \left(\frac{^{\lambda}f'}{^{\circ}f} + i\frac{^{\lambda}f''}{^{\circ}f} \right)$$

Ce qui permet d'obtenir l'équation de phase :

$$\left|^{\lambda}F(\pm h)\right|^2 = \left|^{0}F_T\right|^2 + a(\lambda)\left|^{0}F_A\right|^2 + b(\lambda)\left|^{0}F_T\right|\left|^{0}F_A\right|\cos(^{0}\varphi_T - {}^{0}\varphi_A) \pm c(\lambda)\left|^{0}F_T\right|\left|^{0}F_A\right|\sin(^{0}\varphi_T - {}^{0}\varphi_A)$$

$$a(\lambda) = \frac{f'^2 + f''^2}{^{0}f^2}$$

$$\text{avec : } b(\lambda) = 2\frac{f'}{^{0}f}$$

$$c(\lambda) = 2\frac{f''}{^{0}f}$$

avec $\left|^{0}F_T\right|$, $^{0}\varphi_T$, $\left|^{0}F_A\right|$ et $^{0}\varphi_A$, respectivement le module du facteur de structure total et sa phase, et le module du facteur de structure du diffuseur anomal et sa phase.

On a donc quatre inconnues ($\left|^{0}F_T\right|$, $\left|^{0}F_A\right|$, $\left|^{0}F_T\right|\left|^{0}F_A\right|\cos(^{0}\varphi_T - {}^{0}\varphi_A)$ et $\left|^{0}F_T\right|\left|^{0}F_A\right|\sin(^{0}\varphi_T - {}^{0}\varphi_A)$), et la diffusion anomale à une longueur d'onde λ, va apporter les $|F^+|$ et les $|F^-|$, soit deux équations et la contrainte : $\cos^2(°\phi_T - °\phi_A) + \sin^2(°\phi_T - °\phi_A) = 1$.

Pour réaliser une expérience MAD, la première étape est donc d'ajouter aux cristaux un diffuseur anomal, sous la forme d'un atome lourd (comme par exemple le sélénium avec des séléno-méthionines). Par contre, dans le cas des métalloprotéines, l'atome lourd est déjà présent. Grâce à deux longueurs d'ondes, il sera possible de déterminer mathématiquement les phases ϕ_T et ϕ_A, mais généralement on enregistre pour un MAD, afin d'augmenter le nombre de mesures, trois longueurs d'ondes : au minimum de $^{\lambda}f'$, au maximum de $^{\lambda}f''$, et une troisième longueur d'onde loin du seuil. Après enregistrement des trois jeux de données de diffraction du jeu MAD, il est nécéssaire de déterminer la position des atomes lourds dans la structure. Ceci ce fait par l'étude de la fonction de Patterson différence $\left(\left|F_{A\ hkl}\right| - \left|F_{A\ -h-k-l}\right|\right)^2$ (chapitre 2.4.3.1 du matériel et méthodes) et plus particulièrement de ses sections de Harker.

2.4.2.2.2 Le SAD

La technique du SAD est basée sur les mêmes données théoriques que le MAD, mais ici on enregistre un jeu de diffraction à une seule longueur d'onde : le plus souvent au maximum de $^{\lambda}f''$. La structure est alors obtenue après des processus de modification de

densité comme, par exemple, l'aplatissement de solvant, ce qui permet de résoudre l'indétermination qui reste dans la résolution de l'équation des phases avec l'utilisation d'un seul « dérivé ».
Cette technique est particulièrement adaptée aux protéines formant des cristaux avec une forte proportion de solvant et des symétries non cristallographiques.

2.4.3 Le remplacement moléculaire

La technique de phasage par remplacement moléculaire est employée lorsque l'on dispose d'un modèle d'une protéine supposée structuralement proche de la protéine d'intérêt. La comparaison des séquences en acides aminés est une bonne indication de l'homologie structurale. Deux molécules ayant plus de 30% d'identité séquentielle ont toutes les chances d'avoir un repliement similaire. Afin d'améliorer artificiellement la corrélation entre le modèle et la protéine d'intérêt, il peut aussi être préférable de modifier le modèle en supprimant les chaînes latérales ou certaines parties (boucles, domaines) lorsqu'on le suppose trop éloigné de la structure de la protéine inconnue.
On va donc chercher les six paramètres de la transformation géométrique (trois pour la rotation et trois pour la translation) permettant de passer du modèle à notre protéine. Les procédures du remplacement moléculaire ont été initiées en 1962 (Rossman and Blow, 1962) et, depuis lors, largement améliorées par le développement des techniques de calcul numérique (Navaza, 1994; Read, 2001).
Le positionnement du modèle dans la maille cible est découpé en deux étapes clés : l'étape de rotation pour déterminer l'orientation du modèle et l'étape de translation pour déterminer la position du modèle dans la maille. Le découplage de ces deux étapes permet une recherche de la solution de remplacement moléculaire sur deux espaces indépendants, de trois paramètres chacun, à la place d'un espace à six paramètres. Le nombre de calculs de corrélation nécessaires à la recherche est ainsi diminué d'au moins trois ordres de grandeur.

2.4.3.1 La fonction de Patterson

La fonction de Patterson est la fonction d'autocorrélation de la densité diffusante et peut aussi être considérée comme la transformée de Fourier des intensités diffractées :

$$\mathrm{P}(\vec{u}) = \int_{\mathrm{V}} \rho(\vec{r}).\rho(\vec{r}+\vec{u})\mathrm{d}\vec{r} = \sum_{\mathrm{hkl}} \left|\mathrm{F}(\vec{S})\right|^2 e^{\mathrm{i}2\pi\vec{S}.\vec{u}} = \rho(\vec{r}) * \rho(-\vec{r})$$

avec $\rho(\vec{r})$ la densité électronique au point $\vec{r} = x\vec{a} + y\vec{b} + z\vec{c}$; et $F(\vec{S})$ le facteur de structure du vecteur de diffusion $\vec{S}$.

De cette façon, la fonction de Patterson présente des maxima aux points correspondant aux distances interatomiques. En ces points, la valeur de la fonction correspond au produit des densités électroniques des deux atomes. Elle est ainsi représentative de la distribution des vecteurs entre atomes d'une même molécule (vecteurs intramoléculaires) et entre atomes de molécules distinctes reliées par symétrie (vecteurs intermoléculaires).

Les éléments de symétrie, imposés par le groupe d'espace du cristal, impliquent une concentration particulière de pics au niveau de régions de la fonction de Patterson appelées sections de Harker. Ces pics correspondent alors aux vecteurs interatomiques impliquant des atomes reliés par symétrie.

La fonction de Patterson du modèle est directement calculée à partir de sa structure atomique alors que la fonction de Patterson de la structure recherchée peut être obtenue à partir des facteurs de structure mesurés.

2.4.3.2 La fonction de rotation

Cette première étape consiste à la recherche de l'orientation du modèle initial. La méthode de Rossman et Blow (Rossman and Blow, 1962) consiste à déterminer les trois angles d'Euler ($\theta_1, \theta_2, \theta_3$) pour lesquels la fonction de corrélation entre la fonction de Patterson du modèle orienté, et celle issue du jeu de vecteurs observés, est maximale.

$$\mathrm{RF}(\theta_1, \theta_2, \theta_3) = \int_V \mathrm{P_{obs}}(\vec{u}).\mathrm{U}(\vec{u}).\mathrm{P_{calc}}(\mathbf{R}.\vec{u})\mathrm{d}\vec{u}$$

avec $P_{obs}(\vec{u})$ et $P_{calc}(R \cdot \vec{u})$ respectivement la fonction de Patterson correspondant aux facteurs de structure observés en $\vec{u}$, et la fonction de Patterson correspondant au modèle, calculée en $R \cdot \vec{u}$. Elle est équivalente à la fonction de Patterson calculée en $\vec{u}$ ayant subi la rotation par la matrice de rotation **R** correspondant aux angles $\theta_1, \theta_2, \theta_3$; $U(\vec{u})$, et est une fonction masque limitant le volume d'intégration.

Ce dernier point est très important pour obtenir une solution correcte à la fonction de rotation. En effet, il convient de choisir un volume d'intégration de telle sorte que le maximum de vecteurs intramoléculaires soient inclus, tout en excluant un maximum de vecteurs intermoléculaires. Il est typiquement de l'ordre du rayon du modèle utilisé pour la recherche en rotation.

2.4.3.3 La fonction de translation

Après l'obtention de l'orientation du modèle, ce dernier doit être positionné en translation dans la nouvelle maille. La méthode la plus rapide numériquement est celle développée par Crowther et Blow (Crowther and Blow, 1967) qui utilise les vecteurs entre molécules symétriques de la fonction de Patterson. Si le modèle est placé de façon arbitraire dans la maille, la position des symétriques est tout aussi arbitraire et détermine une distribution singulière des vecteurs intermoléculaires. Ainsi, la fonction de translation définie par Crowther et Blow comme la corrélation entre les distributions de ces vecteurs, et ceux issus des facteurs de structure observés, sera maximale pour un vecteur $\vec{t}$ correspondant à un vecteur entre deux positions symétriques réelles.

$$\mathrm{TF}(\vec{t}) = \int_{V} \mathrm{P}_{\mathrm{obs}}(\vec{u}).\mathrm{P}_{\mathrm{calc}}\left(\vec{u},\vec{t}\right)\mathrm{d}\vec{u}$$

avec les mêmes conventions que pour la fonction de rotation.

2.4.3.4 Les programme utilisés

Il existe de nombreux programmes de remplacement moléculaire, tous automatisés. Au cours de cette thèse nous avons utilisé principalement deux d'entre eux :

- AMoRe (*Automatic Molecular Replacement*) (Navaza, 1994) : Il présente l'avantage d'être automatisé et très rapide, mais aussi, en plus de la fonction de translation de Crowther et Blow, il présente plusieurs fonctions de translation basées sur des approches physiques différentes, qui permettent parfois de trouver la solution en translation là où la méthode précédente échoue. Pour vérifier l'accord entre le modèle placé dans la nouvelle maille et les données expérimentales, deux principaux facteurs sont calculés : la corrélation entre les intensités calculées et mesurées, et un facteur R ;
- Phaser (Read, 2001) : Il s'agit du dernier-né des programmes de remplacement moléculaire. De ce fait, il intègre de nombreux développements de ces dernières années. Et notamment il applique la technique du maximum de vraisemblance pour déterminer la rotation et de la translation à appliquer au modèle. Phaser s'est montré très efficace quant au phasage de structures pour lesquelles les autres programmes avaient échoué. L'accord entre le modèle et la protéine d'intérêt est caractérisé par un « Z-score » :
  - ⇒ Inférieur à 5 : aucune solution n'a été trouvé ;

⇒ Entre 5 et 6 : la solution est incertaine ;

⇒ Entre 6 et 7 : la solution est dite possible ;

⇒ Entre 7 et 8 : la solution est dite probable ;

⇒ Supérieur à 8 : la solution est définitive.

2.5 La résolution de la structure de la protéine

2.5.1 La modification de densité

Les phases obtenues par les méthodes décrites précédemment sont entachées d'erreurs expérimentales dues par exemple aux erreurs dans le modèle d'atomes lourds, ou le défaut d'isomorphisme. De plus, les phases peuvent dans certains cas être calculées à une résolution différente du jeu natif. De ce fait, les premières cartes de densité peuvent contenir un nombre important d'erreurs, et même ne plus être interprétables. On peut alors tenter d'améliorer les phases expérimentales en appliquant certaines contraintes à la densité électronique. Ces dernières introduisent des relations entre les différents facteurs de structure, et donc aussi sur les phases, ce qui peut permettre leur amélioration. Il existe plusieurs procédures visant à améliorer la qualité des phases expérimentales, mais les deux plus efficaces sont l'aplatissement de solvant et la moyennation des densités électroniques reliées par symétrie non cristallographique.

2.5.1.1 L'aplatissement de solvant

Cette procédure est basée sur le fait que la plupart du solvant n'est pas ordonnée dans le cristal, et ne participe pas à la diffraction, sauf à basse résolution. La densité électronique correspondant aux zones de solvant devrait donc être plate, ce qui n'est pas le cas, du fait des erreurs de phase qui introduisent des fluctuations de la densité.

La méthode décrite par Wang (Wang, 1985) consiste à identifier les zones de solvant des zones de protéine, sur la base de la valeur locale moyenne de la densité électronique (plus élevée dans la protéine que dans le solvant), en tenant compte du pourcentage de solvant du cristal pour définir le seuil de coupure. Cela permet de déterminer un masque ayant pour valeur 1 dans les zones de protéine, et 0 pour les zones de solvant. Il suffit alors d'appliquer ce masque, puis de recalculer des phases à partir de cette nouvelle carte de densité électronique nivelée. Ensuite, il est possible de refaire quelques cycles de détermination de masque et nivellement jusqu'à convergence du facteur R_{free}. Ce facteur R_{free} est le même que celui utilisé pour suivre l'affinement (chapitre 2.5.2.4 du matériel et méthodes). Il mesure

l'accord entre des F_{obs} mis de côté, et n'intervenant pas dans les calculs de carte de densité électronique, et les F_{calc} restaurés par la procédure, qui introduisent des relations entre les différents facteurs de structure, et donc des contraintes sur le modèle comme sur les phases. Cette méthode sera d'autant plus puissante que le pourcentage de solvant dans le cristal est élevé.

2.5.1.2 La moyennation des molécules reliées par symétrie non cristallographique

Dans le cas où il existe dans l'unité asymétrique des molécules reliées entre elles par des symétries non cristallographiques, on peut remplacer la densité électronique de chaque motif par la densité électronique moyenne correspondante. Cette procédure est extrêmement puissante pour introduire des contraintes dans les facteurs de structure calculés, et donc améliorer drastiquement les phases, ou même procéder à une extension de phases sur des facteurs de structure à plus haute résolution. Cette procédure utilise, comme la précédente, un masque de la protéine afin de déterminer la transformation à appliquer pendant la moyennation.

2.5.1.3 L'histogramme de densité

Cette technique permet d'imposer une limitation de la densité électronique. A une résolution donnée, l'histogramme de répartition de la densité électronique des protéines est très similaire. On fait alors correspondre l'histogramme de la protéine étudiée à un histogramme théorique mis au point (Zhang and Main, 1990). Afin de les faire coïncider, on partage les deux histogrammes en N intervalles d'aire égale (le même nombre d'électrons) qui sont mis à l'échelle deux à deux.

2.5.1.4 La combinaison des phases

La combinaison des phases par modification de la densité électronique est un processus itératif (Figure 2.5-1). A chaque cycle, de nouvelles phases calculées sont obtenues et combinées avec les phases expérimentales de façon à éviter le biais que peut introduire cette procédure. Cela est effectué jusqu'à convergence du facteur R_{free}, et permet d'augmenter en résolution.

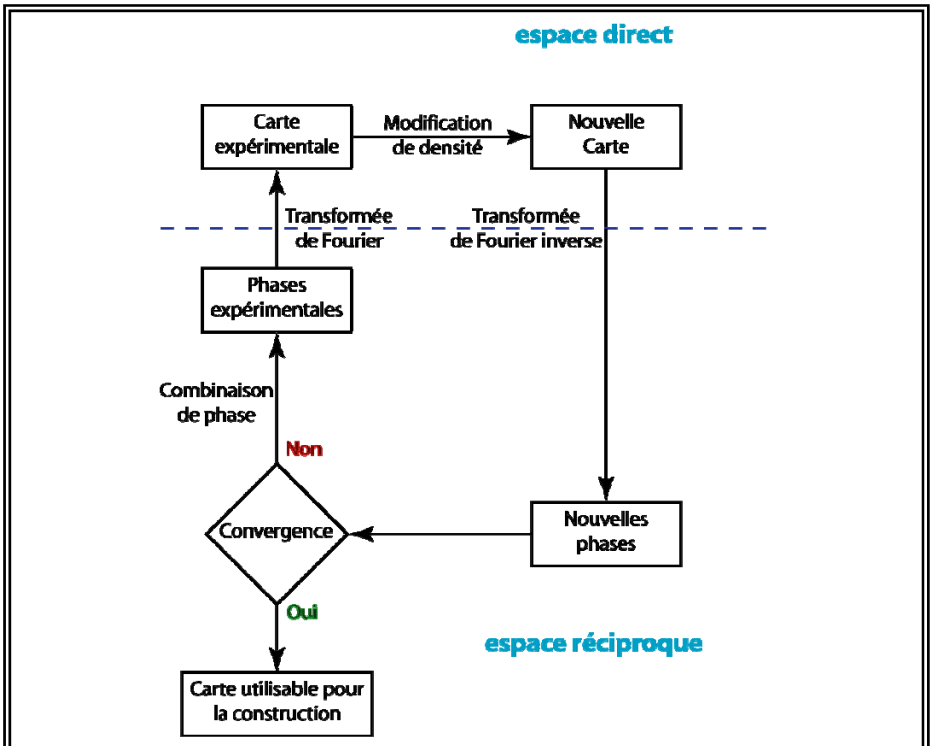


Figure 2.5-1 : Schéma de la procédure de modification de densité électronique.

2.5.2 L'affinement et reconstruction de la structure

Nous avons donc maintenant à notre disposition une carte de densité électronique interprétable et le modèle qui en découle. A ce stade, l'accord entre les facteurs de structures observés et calculés à partir du modèle n'est en général pas satisfaisant. Il convient donc de l'améliorer par des techniques de reconstruction manuelle avec les programmes de visualisation, et des techniques d'affinement numérique des paramètres du modèle (position et agitation de chaque atome). En pratique, on procède en alternant les étapes de reconstruction manuelle et d'affinement numérique, afin de :

- modifier manuellement le modèle lorsque celui-ci est trop éloigné de la structure recherchée, ce qui peut empêcher les algorithmes d'affinement de converger ;
- compenser par l'affinement numérique les erreurs systématiques induites par la reconstruction manuelle.

Il existe deux méthodes d'affinement très répandues : les méthodes de minimisation d'un résiduel (moindres carrés, vraisemblance) ou de dynamique moléculaire.

2.5.2.1 L'affinement par minimisation d'un résiduel

Cette méthode d'affinement permet de minimiser un terme qui tient compte des différences entre les valeurs des différents facteurs calculées à partir du modèle et les valeurs observées de ces mêmes facteurs. La fonction à minimiser correspond alors à l'énergie potentielle du modèle plus une contribution des rayons X :

$$S = W_{xray}F_{xray} + \sum_i W_i \sum_j \frac{1}{\sigma_j^2}(P_{id\,j} - P_{obs\,j})^2$$

avec P_{obsj} et P_{idj} qui sont respectivement les paramètres structuraux observés, et leur valeur calculée à partir du modèle. Il s'agit de contraintes stéréochimiques (longueurs, angles des liaisons interatomiques, répulsions de Van der Waals entre atomes non liés), cristallographiques (similarité par symétrie non cristallographique) ou encore contraintes sur la variation lors de l'affinement des paramètres de position ou d'agitation. Les facteurs W_i et σ_j sont des facteurs de pondération de chaque terme. Enfin F_{xray} rend compte du désaccord entre les facteurs de structure calculés et observés. Il s'agit du terme à minimiser, et il s'écrit différemment en fonction de la procédure de minimisation utilisée.

Il existe deux procédures permettant de minimiser cette fonction :

- La méthode des moindres carrés : elle présente l'avantage d'une convergence rapide. Toutefois, cette méthode a pour désavantage de compenser les erreurs de données mesurées par des erreurs dans le modèle affiné, et de souvent donner des solutions correspondant à un minimum secondaire ;
- Le maximum de vraisemblance : le concept central de cette méthode est la définition de la vraisemblance du modèle (des facteurs de structure calculés) comme étant la probabilité de réaliser la mesure des F_{obs} si l'hypothèse du modèle est juste. L'affinement se fait alors par la recherche du maximum du logarithme de la probabilité. La fonction de probabilité est estimée à partir du jeu de réflexions utilisé pour le calcul du R_{free}. Il permet d'atteindre plus facilement le minimum des facteurs R et R_{free}.

2.5.2.2 La dynamique moléculaire

Dans le cas de la dynamique moléculaire, on ne minimise pas l'énergie potentielle du modèle, mais on calcule en chaque atome la force dérivée de ce potentiel (qui est donc fonction des facteurs de structure), et le déplacement appliqué à chaque atome est celui qui résulte de cette force et d'une vitesse initiale supposée. Le potentiel qui en résulte est minimisé par moindre carré, ou maximum de vraisemblance.

Dans le cas où la structure du modèle est très éloignée de la vraie structure, la fonction d'énergie peut être bloquée dans un minimum local au lieu d'atteindre le vrai minimum lors de la procédure de minimisation. Afin de surmonter cette difficulté, on peut employer une méthode de dynamique moléculaire : le recuit simulé. Le principe est d'appliquer à chaque atome de la protéine les lois de la mécanique newtonienne. Chaque atome se voit assigné une énergie cinétique, lui permettant de sortir du minimum dans lequel elle était bloquée en franchissant les barrières d'énergie. En diminuant par paliers la température, on amène le modèle dans un nouveau minimum d'énergie plus bas. L'inconvénient de cette méthode est le calcul assez long.

2.5.2.3 Les cartes de densité électronique

2.5.2.3.1 Les cartes différences

Lorsque l'on calcule des cartes de densité électronique à partir d'un modèle affiné, ce dernier peut biaiser les phases à cause de son inexactitude. Pour rendre compte de ces erreurs, on utilise le plus souvent en cristallographie les cartes de densité électronique $2mF_{obs}$-DF_{calc} et mF_{obs}-DF_{calc} que l'on appellera par la suite respectivement $2F_{obs}$-F_{calc} et F_{obs}-F_{calc}. Les coefficients m (figure de mérite), qui est une estimation de la précision de la phase, et D, sont des poids appliqués aux termes de phase et aux modules des facteurs de structure calculés. Ils prennent en compte les erreurs apportées par l'incomplétude du modèle et l'inexactitude des positions des atomes, et permettent d'améliorer la qualité des cartes différences.

La première est utilisée pour guider la construction du modèle. En effet, elle montre une densité électronique pour les erreurs du modèle (apportée par F_{obs}-F_{calc} dans $2F_{obs}$-F_{calc}), mais aussi les régions correctement placées (apportées par F_{obs} dans 2 F_{obs}-F_{calc}).

La seconde met en évidence les erreurs du modèle : les pics positifs correspondant à des zones où il manque des électrons ; et les pics négatifs représentant la situation inverse, c'est-à-dire qu'ils marquent une zone où le modèle présente un excès d'électrons.

2.5.2.3.2 Les cartes « omit »

Si l'on veut étudier une région de la protéine en s'affranchissant totalement du biais introduit par le modèle, on peut calculer une carte « omit » de densité électronique. Pour ce faire il suffit de retirer du modèle la partie que l'on veut étudier. Le reste du modèle apportant l'information de phase pour cette région, la densité électronique ainsi calculée ne présente pas

de biais. En pratique, on étudie une carte différence F_{obs}-F_{calc} après quelques cycles d'affinement pour enlever le biais.

2.5.2.3.3 Les cartes de différence anomale

La première étape de toute expérience visant à réaliser des cartes de différence anomale est l'enregistrement d'un ou plusieurs jeux complets de données de diffraction à des longueurs d'ondes choisies, en fonction des diffuseurs anomaux présents dans le cristal de protéine. Dans le cadre de cette thèse, ces expériences avaient pour objectif de déterminer la nature du métal présent dans le site de fixation de NikA.

L'analyse de la composition, en diffuseurs anomaux, des cristaux de NikA par l'étude du spectre de fluorescence émis par ce dernier lorsqu'il est soumis à un rayonnement X, a montré qu'ils contenaient dans la plupart des cas, du fer, du nickel et du zinc. L'enregistrement du spectre de fluorescence a été réalisé grâce au détecteur de fluorescence Roentec® de la ligne BM30A (FIP) de l'ESRF. L'origine de la fluorescence d'un métal et son analyse seront détaillées dans le chapitre 3 du matériel et méthodes.

Afin de déterminer lequel de ces trois métaux est présent dans le site de fixation de NikA, nous avons enregistré quatre jeux de données en choisissant les longueurs d'onde d'enregistrement en fonction des seuils d'absorption des trois atomes lourds (Figure 2.5-2).

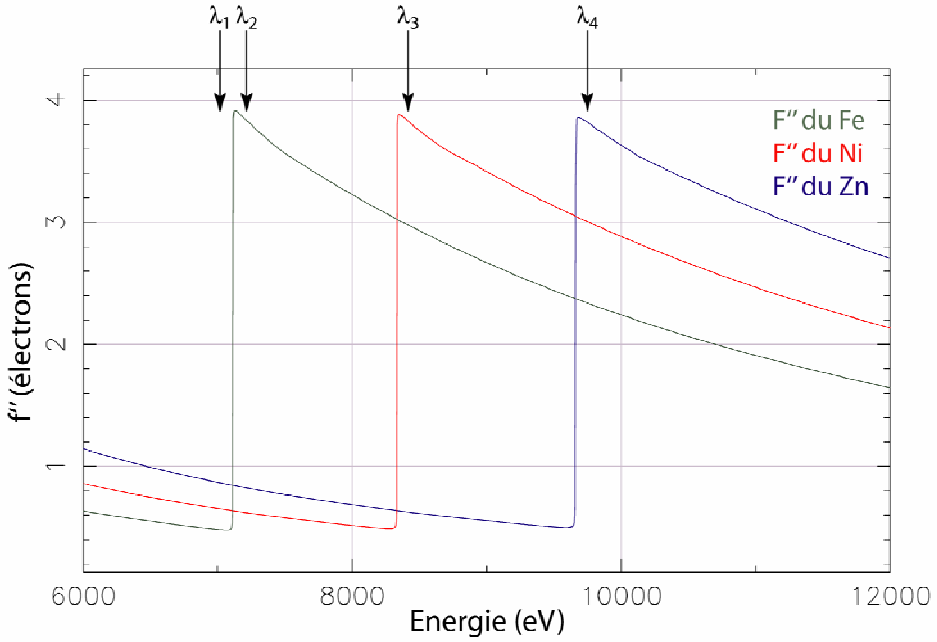


Figure 2.5-2 : Représentation de l'évolution des f'' du fer, nickel et zinc en fonction de l'énergie des rayons X auxquels est soumis le cristal. La brusque augmentation de la valeur de f '' correspond à l'énergie du seuil d'absorption du métal.

Afin d'appréhender la contribution anomale de chaque atome en fonction de l'énergie des rayons X, on analyse l'évolution de f '' uniquement, car l'évolution de f ' n'est significative qu'au seuil de fluorescence :

- λ_1 : à cette énergie, inférieure au seuil d'absorption du fer, les valeurs des f '' des trois atomes considérés sont faibles. De ce fait, la contribution au facteur de diffusion de l'atome (chapitre 2.4.2.1 du matériel et méthodes) sera négligeable. A cette longueur d'onde, aucun des trois atomes lourds considérés n'aura de contribution anomale au facteur de structure ;
- λ_2 : à cette énergie, supérieure au seuil d'absorption du fer, les valeurs des f '' du nickel et du zinc restent faibles, mais celle du fer est maintenant élevée. Le facteur de diffusion du fer s'en trouve donc modifié, et le facteur de structure aura donc une contribution anomale ;
- λ_3 : à cette énergie, supérieure au seuil d'absorption du nickel, la valeur de f '' du zinc reste faible, mais celles du fer et du nickel sont maintenant élevée. Ils ont donc tous les deux, à cette longueur d'onde, une contribution anomale au facteur de structure ;
- λ_4 : enfin à cette énergie, supérieure au seuil d'absorption du zinc, les valeurs de f '' des trois atomes sont élevées. Ils se comportent donc tous les trois comme des diffuseurs anomaux.

Les cartes de différence anomale permettent de visualiser et localiser la diffusion anomale d'un atome. Pour ce faire, on calcule des cartes F^+-F^- (F^+ et F^- étant les paires de Friedel non équivalentes dans le cas d'une diffusion anomale : chapitre 2.4.2.2 du matériel et méthodes).
Les cartes de double différence anomale se calculent en soustrayant deux cartes de différence anomale provenant de deux jeux de données enregistrés à deux longueurs d'onde supérieure et inférieure au seuil de fluorescence du diffuseur anomal considéré. Théoriquement, la seule différence entre les deux cartes est l'intensité du pic du diffuseur anomal. En les soustrayant, on réduit le bruit de fond de la carte, ce qui facilite l'identification du pic de diffusion anomale.
De cette façon, en calculant des cartes de différence et de double différence anomale à partir des quatre longueurs d'ondes décrites précédemment, on identifie et localise les métaux présents dans la structure de NikA.

Carte différence anomale		Carte double différence anomale	
Longueur d'onde	**Atome identifié**	**Longueur d'onde**	**Atome identifié**
λ_1	∅	λ_2 - λ_1	Fe
λ_2	Fe	λ_3 - λ_2	Ni
λ_3	Fe / Ni	λ_4 - λ_3	Zn
λ_4	Fe / Ni / Zn		

Tableau 2.5-1 : Résumé de la nature des atomes identifiés dans les cartes de différence et de double différence anomale.

2.5.2.4 Les critères d'évaluation du modèle

Afin d'évaluer et de valider le modèle affiné, plusieurs critères doivent être surveillés :

- Le facteur R : il donne une idée de la qualité du modèle en mesurant le manque d'accord entre les facteurs de structure observés et les facteurs de structure calculés : $R = \frac{\sum_{hkl} \left\| F_{obs} \right| - \left| F_{calc} \right\|}{\sum_{hkl} \left| F_{obs} \right|}$. Mais les méthodes d'affinement tentant de minimiser au maximum ce facteur, il peut arriver que les erreurs du modèle compensent les erreurs des mesures lorsque le système observable-mesures est peu surdéterminé. Une estimation moins biaisée de la qualité du modèle peut être définie par le facteur R_{free} calculé de la même façon que le facteur R mais sur un ensemble de réflexions choisies aléatoirement (5 à 10%) et non inclues dans l'affinement (Brunger, 1992a). Ce facteur est fortement corrélé à la précision des phases, et permet de détecter de multiples types d'erreurs dans la structure. L'écart entre les facteurs R et R_{free} est finalement un moyen de qualifier l'état de surdétermination du système : plus le rapport entre le nombre d'observations et le nombre de paramètres à affiner est faible, plus cet écart est important (Tickle *et al.*, 1998). Ainsi, dans le cas d'un suraffinement du modèle lorsque le nombre d'observations n'est pas suffisant, le R_{free} aura tendance à augmenter pendant que le facteur R diminue (Kleywegt and Brunger, 1996) ;
- Les déviations moyennes des longueurs et angles de liaisons entre atomes du modèle estiment la qualité stéréochimique de celui-ci. La qualité stéréochimique d'une structure est considérée comme bonne quand ses déviations ne sont pas supérieures aux variations physiques des longueurs et

des angles de liaison déjà observés pour les petites molécules (environ 0.02 Å pour les longueurs et 2 ° pour les angles) (Engh and Huber, 1991) ;

- Le diagramme de Ramachandran : il permet lui aussi de juger la qualité stéréochimique du modèle affiné (Ramakrishnan and Ramachandran, 1965). En raison de l'encombrement stérique des chaînes latérales des acides aminés, les liaisons peptidiques ne peuvent adopter qu'un nombre limité de conformations. Pour chaque résidu, ces conformations sont caractérisées par les angles ϕ et ψ de la liaison peptidique et sont représentées sur un graphe : le diagramme de Ramachandran, qui présente les régions énergétiquement favorables, les régions permises et celles non autorisées. Les chaînes latérales sont elles aussi soumises à des conformations énergétiquement favorables qui seront fortement représentées parmi les différentes conformations rencontrées dans les structures de protéines. Le même type de diagramme permet de valider leur stéréochimie ;
- Enfin, les contacts atomiques, la longueur et l'angle des liaisons hydrogènes potentiels permettent, lorsque ceux-ci ne sont pas corrects, d'identifier les régions du modèle erronées.

2.5.2.5 Les programmes utilisés

2.5.2.5.1 CNS

Le programme CNS (*Crystallography & NMR Systeme*) (Brunger *et al.*, 1998) a été développé à partir de son prédécesseur XPLOR (Brunger, 1992b). Il fonctionne sous la forme de modules que l'on enchaîne selon notre objectif, permettant d'effectuer toutes les étapes cristallographiques, du phasage à l'affinement de la structure. Il propose entre autres une technique d'affinement par recuit simulé, efficace, mais coûteuse en temps de calcul.

2.5.2.5.2 CCP4 – REFMAC

La suite CCP4 (CCP4, 1994) regroupe un grand nombre de programmes de cristallographie développés indépendamment. A l'origine, elle fonctionnait grâce à l'exécution de scripts, mais elle dispose depuis quelques années d'une interface graphique (CCP4i) simplifiant son utilisation. De nombreux programmes de cette suite furent utilisés pendant cette thèse, et notamment le programme REFMAC (Murshudov *et al.*, 1997).

Il a été spécialement développé pour affiner les structures par les méthodes de minimisation. Tout d'abord, à partir des coordonnées atomiques du modèle, il détermine les contraintes structurales (liaisons atomiques, angles de liaisons…), calcule les facteurs de structure, en module et en phase, et les met à l'échelle par rapport aux facteurs de structure observés. Puis vient l'étape d'affinement proprement dite, qui peut se faire soit par la méthode des moindres carrés, ou de maximum de vraisemblance. REFMAC minimise la fonction choisie par la méthode du gradient conjugué : dans l'espace des paramètres affinés, le minimum de la fonction est atteint en faisant varier les valeurs des paramètres afin de se déplacer dans la direction opposée au gradient de cette fonction. Plusieurs cycles sont nécessaires pour atteindre un minimum local. Grâce à différentes options, le programme propose soit d'affiner la structure en corps rigides (modèle divisé en morceaux considérés comme rigides), ou d'affiner l'ensemble des coordonnées avec ou sans le facteur d'agitation de chaque atome. Le programme à la fin de cette procédure calcule les nouvelles coordonnées du modèle affiné et ses facteurs de structures associés.
L'avantage de REFMAC par rapport à CNS, est sa rapidité de calcul.

2.5.3 L'analyse de la structure

Afin d'analyser et d'interpréter les structures affinées, plusieurs programmes de visualisation on été utilisés :

- Turbo-Frodo (Roussel and Cambillaud, 1989) : il s'agit du programme le plus ancien des trois utilisés pendant cette thèse. Il présente d'avantage d'être relativement simple d'utilisation et efficace lors de l'étape de construction. Il est également très complet du point de vue des fonctions proposées. Par contre, il présente certaines limitations et problèmes, comme par exemple une mauvaises gestion de la géométrie des prolines, ou le passage obligatoire par le programme Mapman pour convertir le format des cartes de densité électronique ;
- Pymol (DeLano, 2002) : il s'agit d'un programme récemment développé, et régulièrement mis à jour, optimisé pour l'analyse et surtout la création de figures de la structure étudiée. C'est pour cette raison que toutes les figures présentées dans cette thèse ont été créées avec Pymol. Par contre, les étapes de construction et d'affinement des structures y sont impossibles ;

- Coot (Emsley and Cowtan, 2004) : c'est le plus récent des trois programmes présentés. Comme dans le cas de Pymol, il est régulièrement mis à jour, et encore en développement. Il se révèle extrêmement performant et simple d'utilisation pour la construction et l'affinement du modèle. En effet, il vient récemment d'intégrer la suite CCP4 (dans sa version 6.0) et y est associé à REFMAC.

L'utilisation conjointe de Coot et Pymol, permet de disposer d'un système de reconstruction (coot) et d'analyse (coot et pymol) de structures protéiques extrêmement performant, du fait de leur efficacité et facilité d'utilisation.

3 La Spectroscopie d'Absorption des rayons X (SAX)

La Spectroscopie d'Absorption des Rayons X (SAX) est une technique qui permet d'apporter des informations électroniques et structurales sur un élément considéré, en sondant sélectivement ses états électroniques et son environnement local (à quelques Å). Elle permet alors d'accéder aux informations électroniques que sont l'état de spin et le degré d'oxydation de l'élément sondé, et aux informations structurales que sont la géométrie autour de l'élément sondé, la distance entre ce même élément et ses voisins, ainsi que le nombre et la nature de ces derniers.

Un spectre d'absorption des rayons X présente trois domaines que nous détaillerons dans la suite de ce chapitre :

- Le préseuil ;
- Le XANES (X-ray Absorption Near Edge Structure) ;
- L'EXAFS (Extended X-ray Absorption Fine Structure).

L'EXAFS a été observé pour la première fois par Fricke et Hertz (Fricke, 1920; Hertz, 1920) en 1920. En 1932, pour expliquer les oscillations EXAFS, Kronig (Kronig, 1932) a proposé une première approche théorique basée sur un ordre à courte distance. La théorie de Kronig était incomplète, mais le principe physique de base était correct. Il a fallu attendre les années 1970 pour que Sayers *et al.* (Sayers *et al.*, 1970; Sayers *et al.*, 1971) proposent la première expression mathématique de l'EXAFS qui est celle actuellement utilisée.

Le développement des sources de rayonnement synchrotron de deuxième et troisième génération a permis un développement important de la SAX dans de nombreux domaines d'application différents tels que : la chimie inorganique, la chimie des lanthanides et actinides, la physique des matériaux et des surfaces, la biogéochimie, et bien sûr le domaine qui nous intéresse ici, la biologie.

3.1 Le principe de la SAX

3.1.1 Le phénomène physique

Le phénomène physique observé en SAX trouve son origine dans le processus général d'interaction entre le rayonnement et la matière. Lorsqu'un photon d'intensité I_0 interagit avec la matière, il s'en suit deux phénomènes importants (Figure 3.1-1) :

- L'absorption de ce photon et la photoémission d'un électron. La photoémission est toujours suivie d'une étape de relaxation au cours de laquelle l'atome ionisé

perd, en une ou plusieurs étapes, l'excès d'énergie qu'il contient, par excitation d'un photon (fluorescence d'intensité I_f) et par éjection d'un électron (processus Auger) ;

- Et/ou la diffusion élastique (de Thomson) ou inélastique (de Compton) du photon incident.

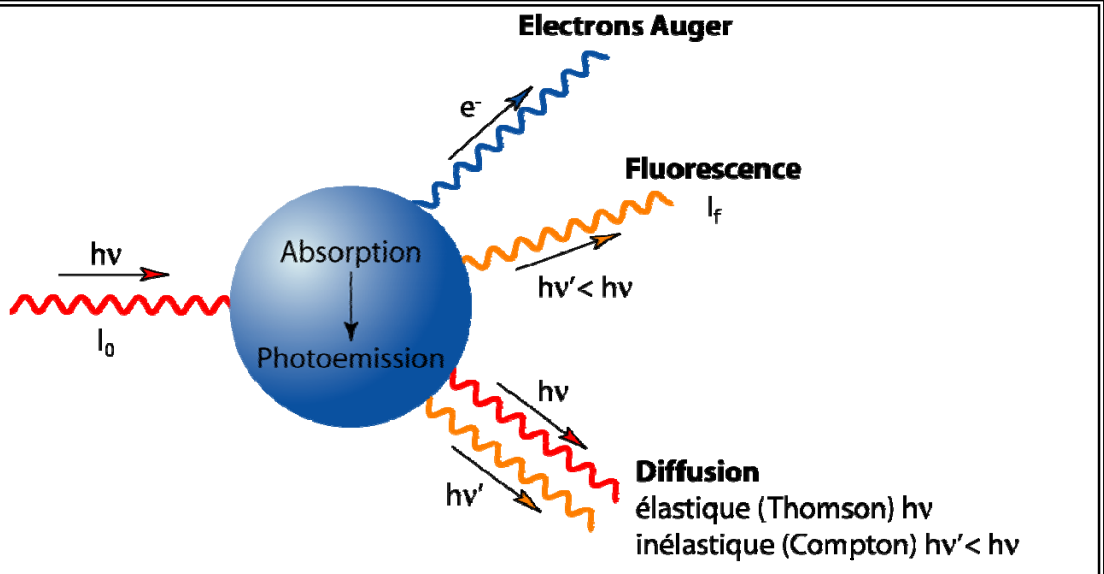


Figure 3.1-1 : Représentation schématique de l'interaction rayonnement-matière.

La SAX étudie le spectre d'absorption des rayons X, l'absorption et la photoémission. L'ion métallique (ou l'élément) étudié, absorbe un photon X entraînant l'excitation puis l'éjection d'un photoélectron de cœur. Cet électron va alors sonder, suivant l'énergie incidente du photon X, les états liés puis l'environnement de l'ion métallique, donnant lieu aux différentes structures dans le spectre d'absorption. On peut alors considérer en toute rigueur la SAX comme un phénomène à un électron. Par la suite, le photoélectron sera simplement appelé électron.

Expérimentalement, le spectre d'absorption des rayons X est obtenu en mesurant l'absorbance μx en fonction de l'énergie incidente du rayonnement. Suivant la concentration de l'ion métallique dans l'échantillon, le spectre sera enregistré en mode fluorescence (pour de faibles concentrations) ou en mode transmission. En mode transmission, l'absorbance est donnée par la loi de Beer-Lambert :

$$\mu x = \ln\left(\frac{I_0}{I_1}\right)$$

avec μ le coefficient d'absorption, x l'épaisseur de l'échantillon et I_0 et I_1 respectivement l'intensité des faisceaux incident et transmis.

En mode fluorescence, l'absorbance est donnée par la relation :

$$\mu x \approx \frac{I_f}{I_0}$$

avec I_f l'intensité de la fluorescence.

Pour un atome donné, le coefficient d'absorption présente des augmentations brutales à certaines valeurs de l'énergie, appelées seuils d'absorption. Dès que l'énergie du photon incident atteint ou dépasse l'énergie du seuil d'absorption E_0, c'est à dire l'énergie d'ionisation (ou énergie de Fermi) un électron est éjecté de l'atome absorbeur. Dans un gaz monoatomique, au delà du seuil d'absorption, la décroissance du coefficient d'absorption est monotone. En revanche, pour tout autre composé, l'environnement de l'atome étudié va induire des structures dans la décroissance du coefficient d'absorption après le seuil.

3.1.2 La SAX des métalloprotéines

Dans l'étude des métalloprotéines, l'intérêt structural est essentiellement focalisé sur les sites métalliques et leur environnement proche. Cette méthode a connu un réel développement depuis l'apparition de sources synchrotrons de plus en plus puissantes. En effet, compte tenu des faibles concentrations de l'ion métallique dans une étude de protéines (de l'ordre du millimolaire), le flux de photons devenait alors suffisant pour que la durée d'enregistrement d'un spectre ne soit plus rédhibitoire pour de telles expériences ((Scott, 1988; Garner, 1991) pour revues).

Les études des métalloprotéines par SAX ont débuté dans les années 70. L'une des premières a été réalisée sur la rubrédoxine en 1975 (Shulman *et al.*, 1975). La SAX s'est alors révélée comme parfaitement adaptée à l'étude des sites métalliques des métalloprotéines et de leur réactivité, puisqu'elle permet d'obtenir des informations structurales en solution, telles que la géométrie et le nombre de coordination de l'ion métallique, la distance entre ce même ion métallique et ses voisins, ainsi que le nombre et la nature de ses voisins. De plus, elle s'est révélée être une technique complémentaire de la cristallographie puisqu'elle permet d'atteindre des informations auxquelles cette dernière ne permet pas d'accéder, telles la géométrie et les distances entre le métal et ses ligands.

3.1.3 Le préseuil

La zone du préseuil est définie pour des énergies inférieures à environ 10 eV en dessous de l'énergie du seuil d'absorption E_0 (Figure 3.1-2a). L'électron éjecté va alors sonder les états liés inoccupés de l'ion métallique. En particulier, pour un ion métallique de transition de la première ligne du tableau périodique, ces états liés sont les états à caractère dominant 3d (Figure 3.1-2b), notés par la suite ψ(3d). Aux transitions vers ces états ψ(3d) correspondent des pics appelés prépics dans le domaine du préseuil. Le préseuil donnera alors des informations électroniques telles que l'état de spin et le degré d'oxydation de l'ion métallique, puis des informations telles que la géométrie et le nombre de coordination de cet ion.

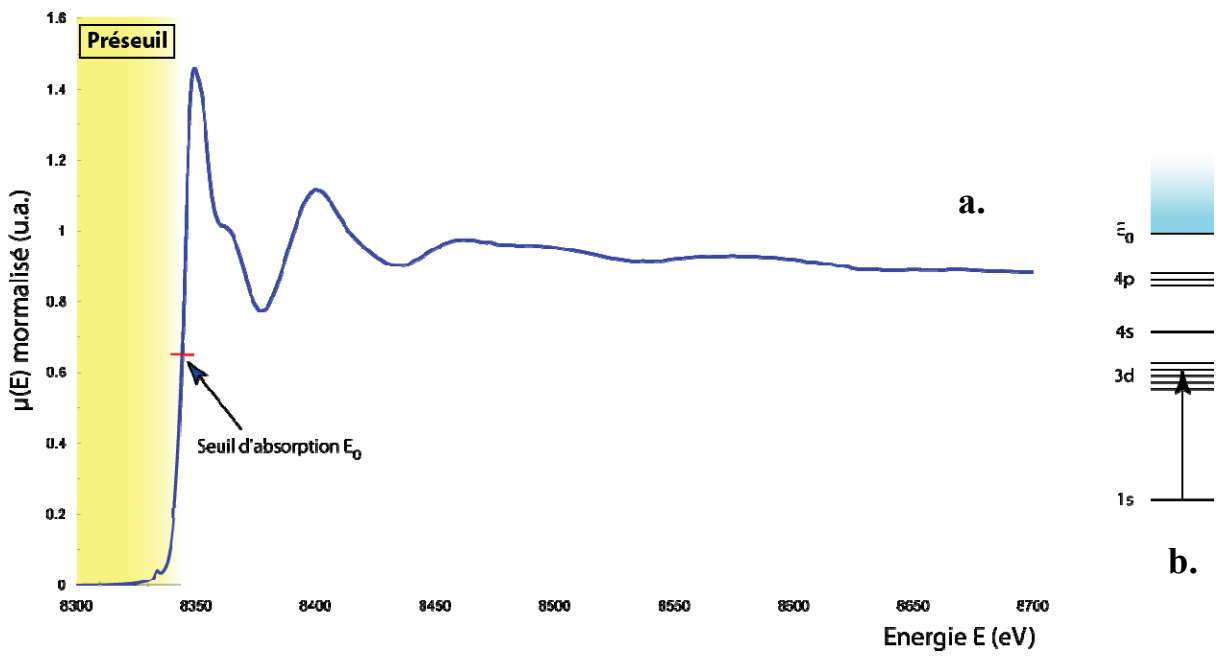


Figure 3.1-2 : Spectre d'absorption normalisé du Ni – EDTA (seuil K du nickel).
a. localisation du préseuil.
b. diagramme d'énergie de l'atome de nickel, et représentation de l'excitation d'un électron 1s par un rayonnement X dont la longueur d'onde appartient au préseuil.

3.1.4 Le XANES

Le XANES s'étend de quelques eV en dessous de E_0 jusqu'à environ 50 eV au delà de E_0 (Figure 3.1-3a).Tant que l'énergie du photon est inférieure à l'énergie du seuil E_0, l'électron excité va sonder les états liés de l'ion métallique, et d'après les règles de sélection, les états à caractère dominant 4p, notés par la suite ψ(4p). La transition la plus intense dans la montée du seuil, due à la transition 1s → ψ(4p), est souvent appelée "raie blanche". Suivant la symétrie

locale autour de l'ion métallique, les orbitales ψ(4p) pourront être dégénérées et les transitions vers les orbitales $\psi(4p_x)$, $\psi(4p_y)$, $\psi(4p_z)$ pourront être observées (Figure 3.1-3b).

Cette méthode permettra donc d'accéder également à des informations électroniques telles que l'état de spin et le degré d'oxydation de l'ion métallique ainsi qu'à des informations telles que la géométrie et le nombre de coordination de cet ion. A partir de la valeur de E_0, il est en effet possible d'obtenir le degré d'oxydation de l'ion métallique, par comparaison avec des composés de degré d'oxydation connu ayant une même symétrie, impliquant le même type de liaisons.

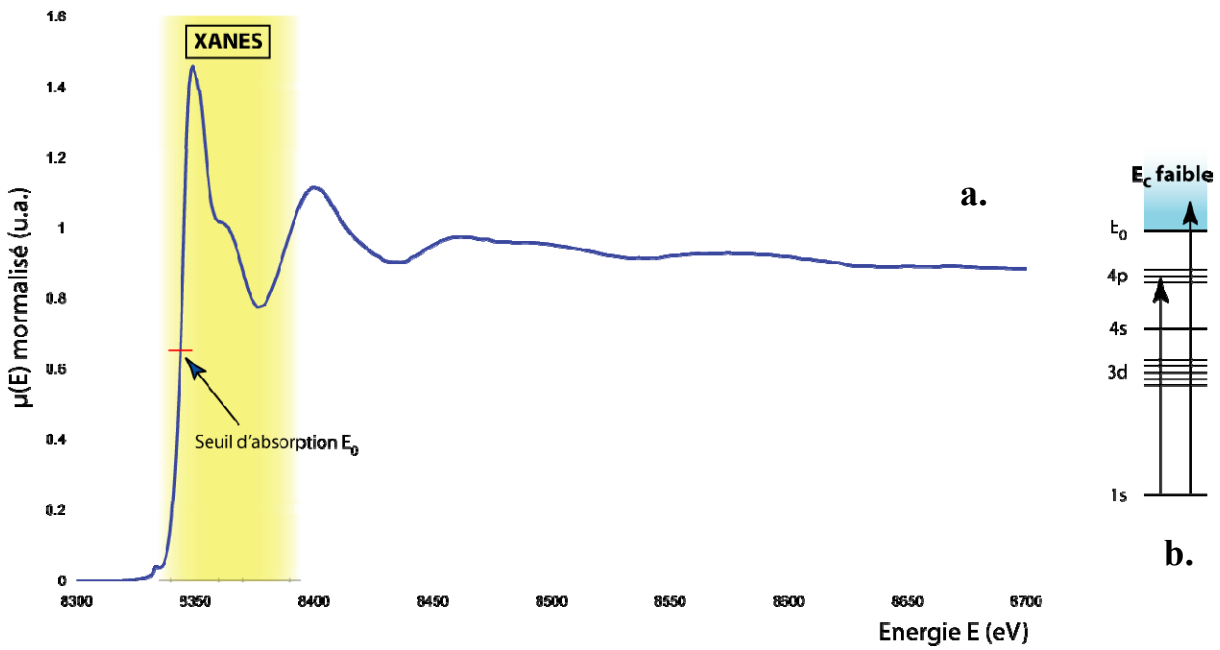


Figure 3.1-3 : Spectre d'absorption normalisé du Ni – EDTA (seuil K du nickel).
a. localisation du XANES.
b. diagramme d'énergie de l'atome de nickel, et représentation de l'excitation d'un électron 1s par un rayonnement X dont la longueur d'onde appartient au XANES.

Lorsque l'énergie du photon devient supérieure à l'énergie du seuil E_0, l'électron est éjecté dans les états du continuum avec une faible énergie cinétique E_c. L'onde électromagnétique de l'électron éjecté de l'ion métallique va se propager et interagir avec le potentiel des atomes voisins (Figure 3.1-4a). Ces atomes vont alors rétrodiffuser l'onde incidente (Figure 3.1-4b). Les interférences constructives et destructives entre l'onde incidente et les ondes rétrodiffusées vont donner lieu aux structures dans la partie XANES. Mais, pour de faibles valeurs de l'énergie cinétique, l'électron possède un long temps de vie, ce qui correspond à un libre parcours moyen de quelques dizaines d'Angströms. L'électron peut être alors rétrodiffusé

par des atomes éloignés de l'ion métallique ou même rétrodiffusé plusieurs fois par des atomes proches (Figure 3.1-4c), ce phénomène est appelé diffusion multiple. Elle est extrêmement dépendante de la géométrie de l'édifice moléculaire, et importante pour des arrangements atomiques particuliers tels que les groupes imidazole des histidines. Par conséquent, les structures de la diffusion multiple peuvent fournir des informations structurales, notamment la géométrie, et même la nature des atomes voisins de l'ion métallique, complémentaires de celles obtenues par EXAFS.

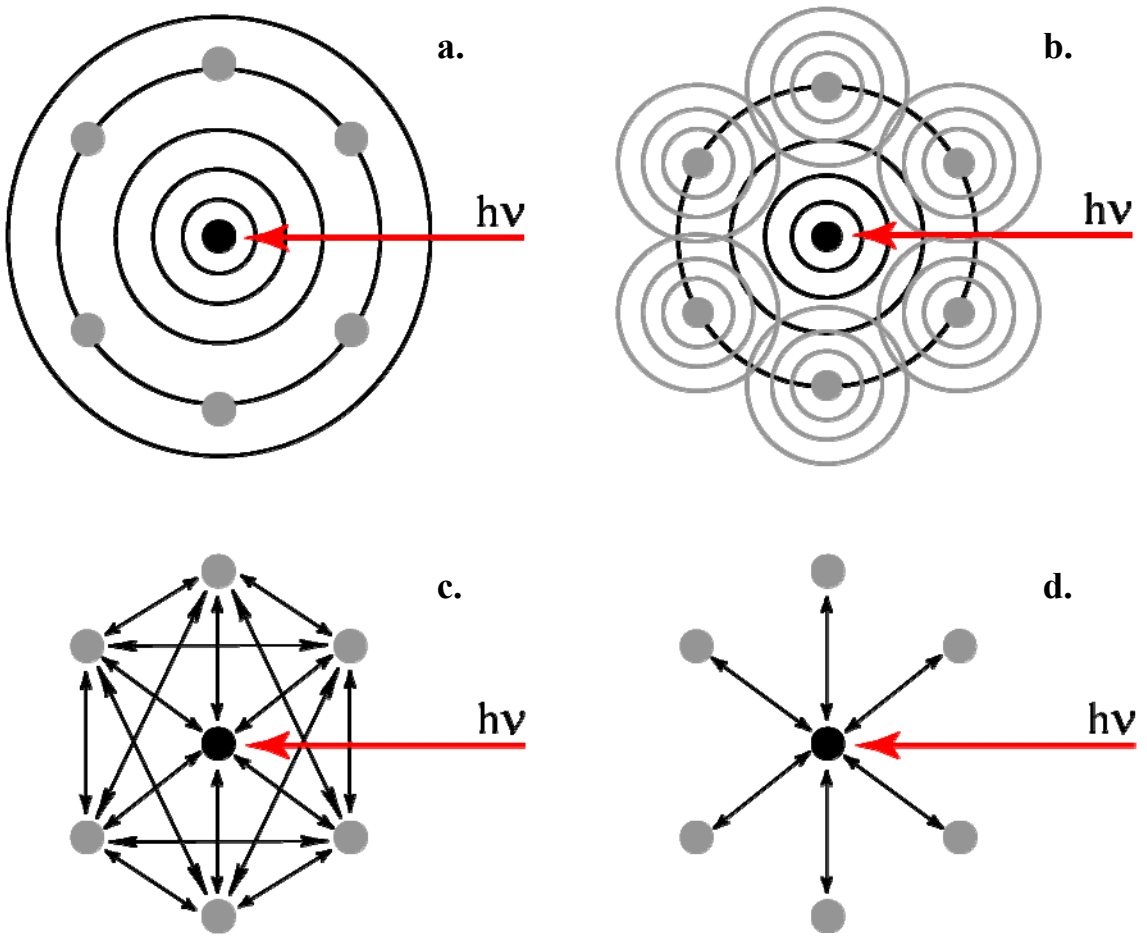


Figure 3.1-4 : Processus de l'absorption X.

a. diffusion (représentée par les cercles concentriques noirs).

b. rétrodiffusion (représentée par les cercles concentriques gris).

c. diffusion multiple (représentée par les doubles flèches).

d. diffusion simple (représentée par les doubles flèches).

3.1.5 L'EXAFS

L'EXAFS s'étend à partir de 50 eV au delà de E_0 (Figure 3.1-5a). L'énergie de l'électron étant supérieure à l'énergie du seuil E_0, l'électron est éjecté dans les états du continuum avec une énergie cinétique E_c élevée (Figure 3.1-5b).

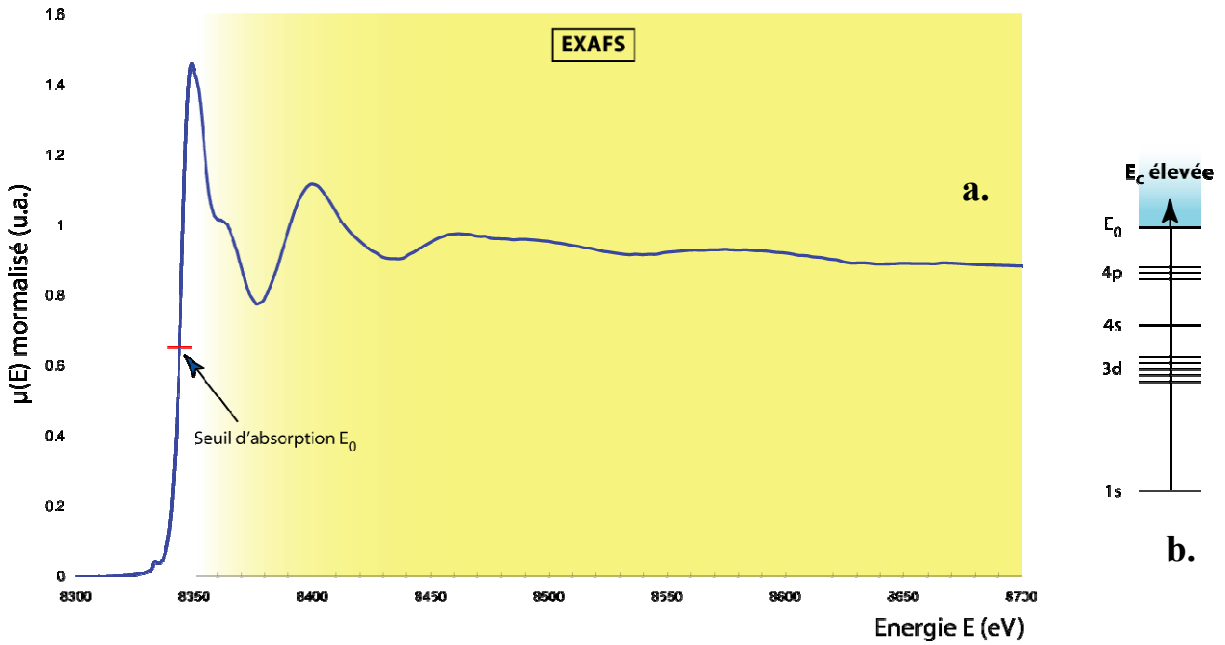


Figure 3.1-5 : Spectre d'absorption normalisé du Ni – EDTA (seuil K du nickel).
a. localisation du EXAFS.
b. diagramme d'énergie de l'atome de nickel, et représentation de l'excitation d'un électron 1s par un rayonnement X dont la longueur d'onde appartient à l'EXAFS.

De même que pour le XANES, l'onde électromagnétique de l'électron éjecté de l'ion métallique va se propager et interagir avec le potentiel des atomes voisins (Figure 3.1-4a). Ces atomes vont alors rétrodiffuser l'onde incidente (Figure 3.1-4b). Les interférences constructives et destructives entre l'onde incidente et les ondes rétrodiffusées vont donner lieu aux structures dans la partie EXAFS, plus connues sous le nom d'oscillations EXAFS. La différence avec le domaine du XANES réside dans le fait que, pour des valeurs élevées de l'énergie cinétique, la diffusion atomique est importante seulement vers l'avant et sur un nombre limité de couches de voisins. De plus, les processus de diffusion vont être limités par un effet de type Debye-Waller. Les processus de diffusion simple sont donc majoritaires (Figure 3.1-4d). Toutefois, pour des édifices moléculaires particuliers (tels que les groupes imidazole), l'influence de la diffusion multiple peut s'étendre sur le domaine énergétique de l'EXAFS.

Les oscillations EXAFS sont dépendantes de la géométrie autour de l'ion métallique, de la distance entre cet ion métallique et ses voisins, ainsi que du nombre et de la nature de ses voisins. De plus, il est possible, d'après la valeur du terme de Debye-Waller, d'obtenir des informations sur la distribution de distances autour de la distance moyenne entre l'ion métallique et ses voisins, mais il est par contre difficile de distinguer en EXAFS deux éléments possédant des numéros atomiques proches (par exemple entre O, N ou C).

3.2 Le formalisme de la SAX

3.2.1 Le coefficient d'absorption

Le coefficient d'absorption est donné, en mécanique quantique, par la règle d'or de Fermi pour un électron dans l'approximation dipolaire :

$$\mu(E) = C^{te} \times \left|\langle \psi_i | \vec{\varepsilon} \cdot \vec{r} | \psi_f \rangle\right|^2 \times \rho(E_f) \times \delta(h\nu - (E_f - E_i))$$

avec :

- ψ_i et ψ_f respectivement la fonction d'onde de l'électron dans son état initial et final ;
- $\vec{\varepsilon}.\vec{r}$ l'hamiltonien d'interaction rayonnement-matière dans l'approximation dipolaire, qui est le produit scalaire du vecteur polarisation du champ électromagnétique ($\vec{\varepsilon}$) et du vecteur position de l'électron ($\vec{r}$). Il est également appelé opérateur électrique dipolaire, et noté $\hat{O}_{dip}$;
- $\delta(h\nu - (E_f - E_i))$ la fonction de Dirac qui assure la conservation de l'énergie lors du processus d'absorption ;
- $\rho(E_f)$ la densité d'états vacants.

Le coefficient d'absorption μ(E) donne donc la probabilité de transition d'un électron dans son état fondamental (E_i, ψ_i), vers un état excité d'énergie (E_f, ψ_f) par absorption d'un photon X d'énergie hν.

Dans le cadre de cette thèse, nous nous sommes placés au seuil K, c'est-à-dire que le photon X absorbé va exciter un électron 1s. Dans ce cas, le coefficient d'absorption s'écrit :

$$\mu(E) = C^{te} \times \left|\langle \psi_{1S} | \hat{O}_{dip} | \psi_f \rangle\right|^2 \times \rho(E_f) \times \delta(h\nu - (E_f - E_{1S}))$$

Une transition dipolaire électrique est alors permise si $\langle \psi_{1S} | \hat{O}_{dip} | \psi_f \rangle \neq 0$. Il en découle des règles de sélection pour la transition dipolaire. Dans le cas des gaz monoatomiques, elles sont alors : $\Delta S = 0$, et $\Delta l = \pm 1$. Un électron s (l = 0) ne peut donc être excité que vers des orbitales atomiques p ($\Delta l = +1$). Les transitions vers les orbitales atomiques d sont interdites puisque $\Delta l = +2$. Dans le cas de l'étude d'un site métallique d'une protéine, et plus généralement d'une molécule, les orbitales atomiques de l'ion métallique s'hybrident avec celles de même symétrie des voisins de l'ion métallique pour donner des orbitales moléculaires. Les règles de sélection doivent donc être "reconsidérées" en spectroscopie moléculaire.

3.2.2 La formule EXAFS

La formule EXAFS nécessite l'application de certaines approximations :

- La transition est considérée de nature dipolaire ;
- Seul l'électron excité du niveau 1s participe à l'absorption, il n'y a pas d'interaction entre les électrons ;
- Le potentiel des atomes est de symétrie sphérique ;
- L'onde sortante est une onde sphérique ;
- La distribution des distances est une distribution gaussienne normalisée ;
- Seule la diffusion simple entre l'ion métallique et les atomes voisins est considérée, c'est à dire que le déphasage correspond à deux fois la distance entre l'ion métallique et son voisin. Cette approximation peut ne plus être valable au delà de la deuxième couche, en effet, au delà de cette dernière et pour certains arrangements atomiques, la diffusion multiple peut avoir une contribution non négligeable. Néanmoins, la formule EXAFS peut rester valable puisque introduire la diffusion multiple pour trois atomes revient à multiplier $f_i(k)$ par un terme A(k) et ajouter à $\phi_{ij}(k)$ un terme B(k).

Elle s'écrit alors :

$$k_\chi(k) = -S_0^2 \sum_i \frac{N_i}{R_i^2} |f_i(k)| e^{-2\sigma_i^2 k^2} e^{-\frac{2R_i}{\lambda(k)}} \sin\left(2kR_i + \phi_{ij}(k)\right)$$

avec :

- $k = \sqrt{\frac{2m_e \times (E - E_0)}{\hbar^2}}$ avec m_e la masse de l'électron et $\hbar$ la constante de Planck réduite ($\hbar = \frac{h}{2\pi}$) ;
- S_0^2 : terme de réduction d'amplitude due aux effets multiélectroniques ;
- N_i : nombre de voisins de type i à la distance R_i ;
- R_i : distance moyenne entre l'ion métallique et les voisins de type i ;
- $f_i(k)$: amplitude de rétrodiffusion de l'atome de type i ;
- σ_i : facteur de Debye-Waller, dans l'approximation d'une distribution gaussienne. Il comprend un terme de distribution statique (distribution des distances) et dynamique (agitation thermique) ;
- $\lambda(k)$: libre parcours moyen de l'électron excité. Ce terme est un terme phénoménologique qui tient compte des pertes inélastiques de l'électron ;
- $\phi_{ij}(k)$: déphasage entre l'ion métallique j et les voisins de type i.

La formule EXAFS contient des termes structuraux R_i, N_i, σ_i et des termes électroniques $f_i(k)$, $\phi_{ij}(k)$, $\lambda(k)$, S_0^2, liés à la nature des atomes mis en jeu. Ces termes influent différemment sur les oscillations EXAFS : $\lambda(k)$, S_0^2 et σ_i sont des termes d'amortissement alors que N_i, $f_i(k)$ sont des termes d'intensité de diffusion.
L'incertitude classique sur les paramètres structuraux en EXAFS est de :

- 0,02 Å pour R_i (du moins pour la première sphère de coordination) ;
- 20% pour N_i.

3.3 L'enregistrement des données SAX

Ce travail a été mené avec Lilian Jacquamet, du laboratoire, qui m'a ainsi formé à l'enregistrement, au traitement, et à l'interprétation les spectres d'absorption EXAFS.
L'enregistrement des données SAX s'est fait sur la ligne de lumière BM30B (FAME) de l'ESRF (Figure 3.3-1), où l'on a enregistré les spectres :

- Au seuil du nickel du composé référence : Ni(II)-EDTA 50 mM + glycérol 20% ;
- Au seuil du fer et du nickel de la protéine NikA, purifiée par le protocole contenant du chloroforme, à 75,2 mg/mL (soit 1,3 mM) + glycérol 20%.

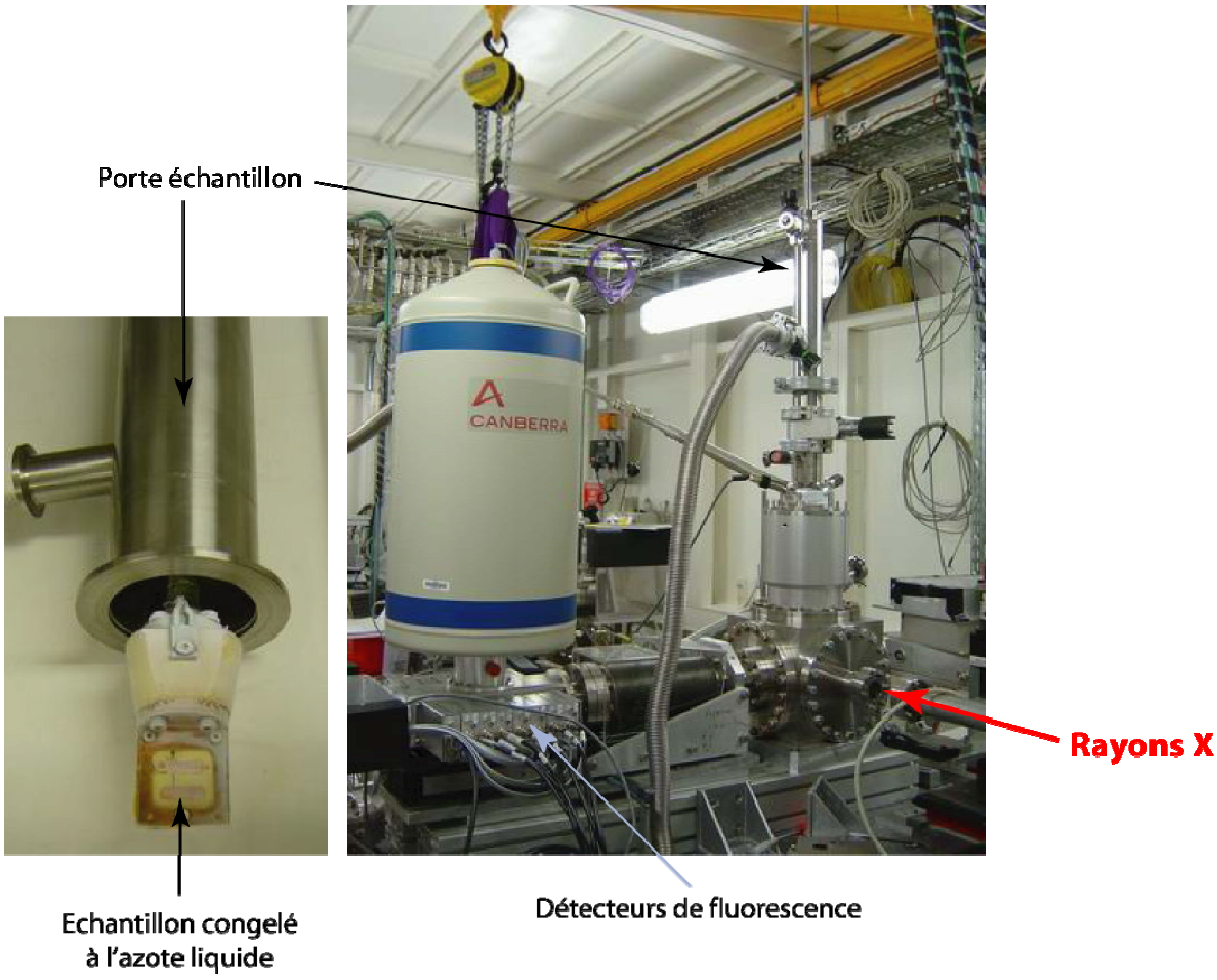


Figure 3.3-1 : Exemple d'une ligne de lumière de l'ESRF dédiée à la SAX : Ligne ID26. Sa configuration est équivalente à celle de la ligne BM30B (FAME).

3.4 Le traitement des données EXAFS

Le traitement des données EXAFS a été réalisé au cours de cette thèse à l'aide des programmes SEDEM (Aberdam, 1998), la suite de programme Ifeffit (Ifeffit; Zabinsky *et al.*, 1995). Il se décompose en cinq étapes détaillées ci-dessous et illustrées par la Figure 3.4-1.

3.4.1 Etape 1 : La détermination de l'énergie du seuil E_0 et de la hauteur du saut d'absorption

L'énergie du seuil E_0 est déterminée par le maximum de la dérivée du spectre d'absorption dans le domaine du XANES. Celle-ci est arbitraire et n'a pas de fondement physique, puisque la valeur de E_0 s'avère difficilement accessible expérimentalement compte-tenu des différentes structures de cette région. Pour les composés de référence, les valeurs de E_0 sont fixées à celle du composé à caractériser.

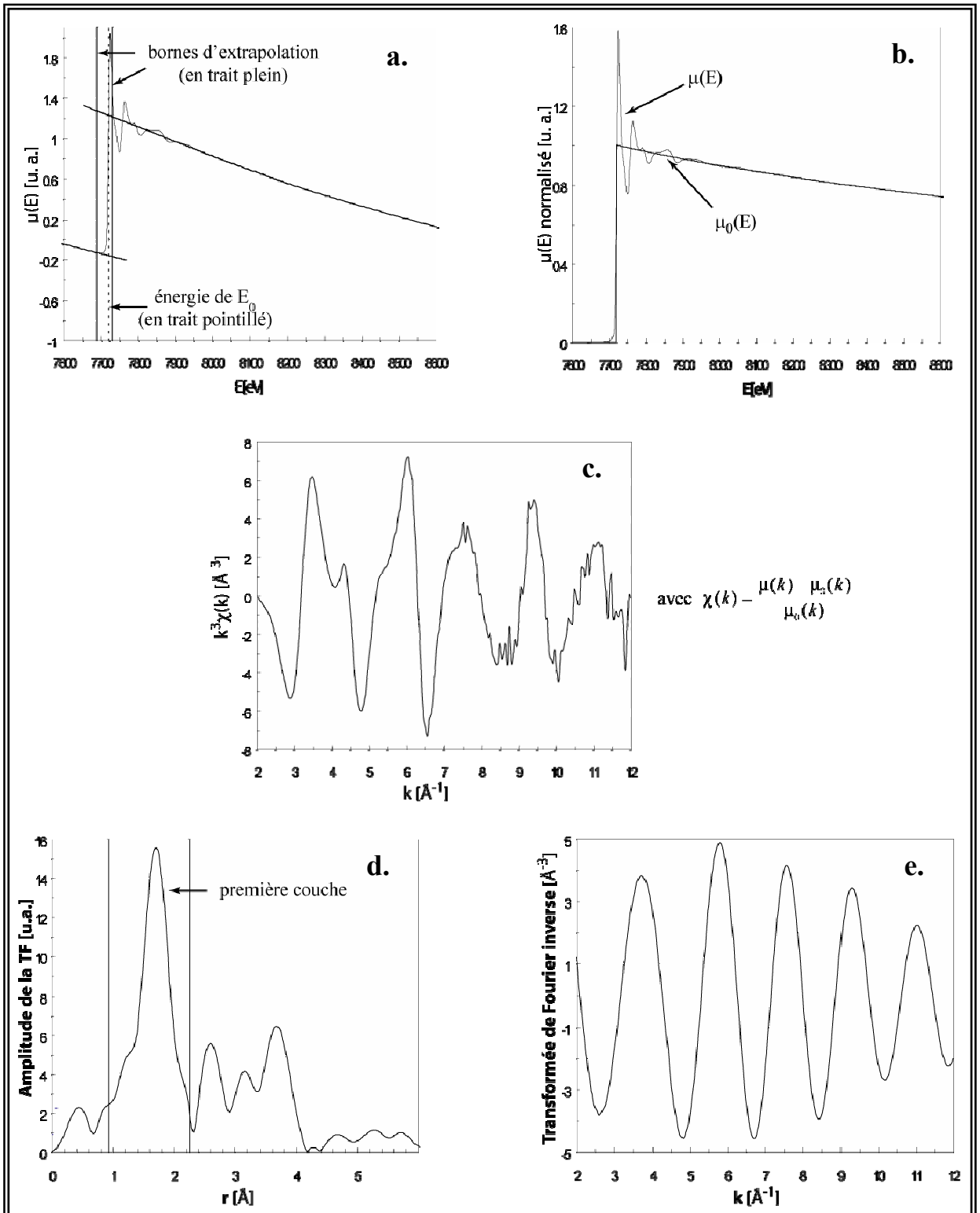


Figure 3.4-1 : Extraction des informations structurales du spectre d'absorption X.

La normalisation des spectres d'absorption, réalisée à l'étape 2, nécessite la détermination de la hauteur h du saut d'absorption à l'énergie du seuil. Cette dernière étant directement proportionnelle au nombre de moles d'ion métallique présent dans l'échantillon, on l'utilise pour normaliser les spectres d'absorption. Pour cela, la ligne de base avant le seuil est déterminée par une droite, et la ligne de base après le seuil, c'est à dire la ligne de base passant à travers les oscillations EXAFS, est déterminée à l'aide d'une fonction de splin cubique à deux segments de poids différents (Figure 3.4-1a). Une fois obtenues, ces deux lignes de base sont extrapolées jusqu'à E_0. Le saut d'absorption h est alors le saut du coefficient d'absorption, à l'énergie de E_0, entre la ligne de base avant seuil extrapolée à E_0 et la ligne de base après seuil extrapolée à E_0.

3.4.2 Etape 2 : La normalisation du spectre d'absorption

Après normalisation des spectres d'absorption par la hauteur h, les lignes de base avant et après seuil, déterminées précédemment, sont remplacées par une absorption atomique théorique $\mu_0(E)$ calculée par le programme f-Prime (Cromer and Liberman, 1970) (Figure 3.4-1b). Cette étape assure une reproductibilité dans la normalisation des spectres, et permet de réaliser une comparaison facile et significative entre différents spectres d'absorption.

3.4.3 Etape 3 : La conversion du coefficient d'absorption dans l'espace réciproque [Å^{-1}] et le calcul de la fonction EXAFS $\chi(k)$

Le spectre jusqu'alors dans l'espace des énergies (en eV) est converti dans l'espace réciproque (en Å^{-1}) par la relation :

$$k = \sqrt{\frac{2m_e \times (E - E_0)}{\hbar^2}}$$

avec k le nombre d'onde de l'électron exprimé en Å^{-1} et m_e la masse de l'électron.

Le coefficient d'absorption atomique théorique $\mu_0(E)$ est donc converti en $\mu_0(k)$. Il est alors de nouveau affiné, dans sa partie après seuil, à l'aide d'une fonction de splin cubique à deux segments de poids différents. En effet, la ligne de base, calculée précédemment dans l'espace des énergies, peut nécessiter une amélioration puisque la sensibilité dans l'espace réciproque peut s'avérer différente de celle dans l'espace des énergies. Un des critères d'obtention d'une "bonne" ligne de base $\mu_0(k)$ est d'effectuer la transformée de Fourier des oscillations EXAFS après soustraction de $\mu_0(k)$ ainsi obtenu (comme décrit dans l'étape 4).

Elle doit présenter le minimum de pics non physiques (correspondant à une distance inférieure à 1,7 Å) et avec la plus faible intensité possible. Tant que la condition du minimum de pics non physiques ne semble pas atteinte, il faudra reconstruire une nouvelle ligne de base $\mu_0(k)$.

Une fois la "meilleure" ligne de base $\mu_0(k)$ obtenue, on la soustrait au spectre d'absorption $\mu(k)$. La fonction EXAFS $\chi(k)$ normalisée est alors calculée suivant la relation :

$$\chi(k) = \frac{\mu(k) - \mu_0(k)}{\mu_0(k)}$$

En fonction du site métallique à étudier, la fonction $\chi(k)$ peut être multipliée par une puissance différente de k : 1, 2 ou 3. Cette multiplication est en fait utile pour compenser l'amortissement du signal, et s'avère souvent indispensable pour l'étude au delà de la première sphère de coordination. Dans le cadre de cette thèse, la fonction EXAFS a été multipliée par k^3 (Figure 3.4-1c).

3.4.4 Etape 4 : La transformée de Fourier

La fonction $k^3\chi(k)$ est la transformée de Fourier de l'espace réciproque [$Å^{-1}$] dans l'espace réel [Å]. Elle est calculée directement comme une intégrale de Fourier et non pas à l'aide de l'algorithme de transformée de Fourier rapide. Elle nous donne la distribution des couches atomiques autour de l'ion métallique (Figure 3.4-1d). Les distances obtenues pour chaque couche sont des distances effectives, dans le sens où elles doivent être corrigées par les différents déphasages électroniques $\phi_{ij}(k)$ entre les paires atomiques ion métallique-voisin pour correspondre à des distances réelles. Il faut typiquement ajouter environ 0,5 Å aux distances pour obtenir les distances réelles.

Une fenêtre d'apodisation, de type Kaiser "avec plateau", est utilisée pour limiter les effets de troncatures engendrés par la transformée de Fourier

3.4.5 Etape 5 : L'ajustement et obtention des informations structurales

Afin d'obtenir les informations structurales R_i, N_i et σ_i, l'ajustement de la fonction $k^3\chi(k)$ du spectre expérimental ou issue du filtre de Fourier d'une couche particulière d'atomes (ce qui correspond au pic de la fonction $k^3\chi(k)$ que l'on veut étudier, comme par exemple la première couche de la Figure 3.4-1d) est réalisé (Figure 3.4-1e). Les paramètres structuraux R_i, N_i et σ_i obtenus par ajustement doivent être physiques, c'est-à-dire doivent avoir une

signification physique. Par exemple, un ajustement donnant une distance R_i inférieure à 1,5 Å ne peut être retenu. Dans le cas du terme de Debye-Waller σ_i, il n'a un sens physique que s'il est inférieur à 10^{-2} $Å^{-1}$. Le nombre de voisins N_i n'est, quant à lui, pas ajusté mais fixé à des valeurs chimiquement possibles.

L'algorithme utilisé pour l'ajustement est un algorithme de Marquardt (Marquardt, 1963) permettant de minimiser une fonction de moindre carré.

L'examen de la formule EXAFS (chapitre 3.2.2 du matériel et méthodes), montre qu'il est nécessaire pour déterminer les paramètres structuraux R_i, N_i et σ_i de connaître les paramètres électroniques que sont S_0^2, $f_i(k)$, $\lambda(k)$ et $\phi_{ij}(k)$ ainsi que le paramètre ΔE_0, qui est le décalage en énergie entre la valeur de E_0 choisie et la fonction de déphasage $\phi_{ij}(k)$ choisie. Compte tenu du nombre important de paramètres et des corrélations entre ces différents paramètres, nous avons appliqué une méthodologie développée et décrite par Lilian Jacquamet dans sa thèse (Jacquamet, 1999). Elle implique l'utilisation de composés de référence pour déterminer les paramètres structuraux.

3.4.5.1 La détermination de $f_i(k)$, $\phi_{ij}(k)$ et $\lambda(k)$

Ces paramètres électroniques peuvent être obtenus expérimentalement ou théoriquement et sont déterminés pour une paire ion métallique-voisin M-X. Dans un site métallique de protéine, les voisins de l'ion métallique se limitent essentiellement aux atomes de carbone, d'azote, d'oxygène et de soufre (les atomes d'hydrogène n'étant pas détectés de façon classique en EXAFS). Comme il est ici difficile de distinguer deux atomes de la même ligne du tableau périodique, les paramètres $f_i(k)$, $\phi_{ij}(k)$ et $\lambda(k)$ doivent être essentiellement déterminés pour les paires M-C/N/O et M-S.

Afin de déterminer expérimentalement ces trois paramètres, il faut disposer de composés de référence, de structure connue, et possédant un environnement du métal le plus proche possible de celui du site métallique à caractériser.

Les paramètres $f_i(k)$, $\phi_{ij}(k)$ et $\lambda(k)$ peuvent être aussi obtenus à partir de calculs *ab initio* réalisés par le programme FEFF. Ce programme peut calculer ces paramètres pour une paire M-X isolée, et ce, quelle que soit la distance interatomique, ou les calculer dans une structure cristallographique. Ainsi, le nombre de paramètres à déterminer par ajustement est réduit de trois, puisque les paramètres η, β et γ sont déterminés automatiquement par le programme SEDEM à partir des valeurs de $\lambda(k)$ calculées par FEFF.

C'est cette dernière méthode que nous avons choisi d'utiliser dans le cadre des études menées pendant cette thèse. Dans ce cas la formule EXAFS devient :

$$k_\chi(k) = -S_0^2 \sum_i \frac{N_i}{R_i^2} \left| F_{eff_i}(k,R) \right| e^{-2\sigma_i^2 k^2} \sin\left(2kR_i + \phi_{ij}(k,R)\right)$$

avec : $F_{eff_i}(k) = f_i(k,R) e^{-\frac{2R_i}{\lambda(k,R)}}$.

Mais, même en utilisant les paramètres *ab initio*, les composés de référence restent indispensables à l'obtention des paramètres électroniques S_0^2 et ΔE_0.

Nous avons calculé les paramètres électroniques $f_i(k)$, $\phi_{ij}(k)$ et $\lambda(k)$ avec le programme FEFF pour deux atomes isolés M et X (C/N/O ou S) séparés par une distance R_i, puisqu'ils sont fonction de k mais aussi de R_i. Le choix de R_i est guidé par les composés de référence. Bien que le choix de la valeur de R_i soit important, elle tolère un certain taux d'erreur. En effet, il s'avère que pour une valeur de la distance R_i choisie pour le calcul, il est possible d'obtenir par ajustement la distance cristallographique de composés de référence à une distance $R_i \pm 0{,}25$ Å. Ainsi, les paramètres électroniques calculés par le programme FEFF dépendent de R_i mais cette dépendance ne rend pas pour autant le choix de ce dernier déterminant pour le résultat final.

3.4.5.2 La calibration et validation de S_0^2 et ΔE_0

Pour obtenir les paramètres S_0^2 et ΔE_0, nous avons utilisé des composés de structure cristallographique connue que nous avons considérés comme « homogènes » (MX_i), c'est-à-dire avec une première sphère de coordination constituée d'un seul type d'atomes différenciables en EXAFS. Dans le composé MX_i, la distance interatomique R_i et le nombre de voisins N_i sont connus, ce qui permet alors de calculer à l'aide du programme FEFF les paramètres $f_i(k)$, $\phi_{ij}(k)$, $\lambda(k)$ pour la paire M-X. L'ajustement permet alors de déterminer S_0^2 et ΔE_0 pour une paire M-X donnée. Si N_i est fixé à sa valeur cristallographique, R_i est par contre ajusté. En effet la distance cristallographique et la distance déterminée par EXAFS peuvent différer très légèrement de par leurs incertitudes intrinsèques. De plus, suivant les conditions expérimentales de température, les distances cristallographiques peuvent également varier. Mais en tout état de cause, l'écart entre la distance R_i déterminée par EXAFS et celle obtenue par cristallographie ne doit en principe pas dépasser 0,02 Å.

3.4.5.3 La détermination de R_i, N_i et σ_i du site métallique à caractériser

L'ajustement se fait par le programme SEDEM, dans un premier temps sur le filtre de Fourier du premier pic de la transformée de Fourier (Figure 3.4-1d). Ceci permet d'obtenir les paramètres structuraux R_i, N_i et σ_i pour chaque couche ou type d'atome. Bien entendu, l'ajustement peut être réalisé sur le spectre total $k^3\chi(k)$, mais il est souvent plus pratique de réaliser l'ajustement en premier lieu sur chaque pic de la transformée de Fourier, car le nombre de paramètres ajustés est alors moindre.

L'obtention de chaque distance par ajustement va dépendre du domaine énergétique atteint pendant l'enregistrement du spectre. En effet, la résolution en EXAFS est donnée par la relation : $\Delta R = \frac{\pi}{2 \cdot \Delta k}$. Pour une valeur ΔR de 0,16 Å, le Δk est égal à 10 $Å^{-1}$, ce qui signifie que deux couches séparées par moins de 0,16 Å ne seront pas distinguées en EXAFS avec un Δk usuel de 10 $Å^{-1}$ pour l'étude de métalloprotéines. Il apparaît donc que, plus le domaine en k du spectre sera étendu, meilleure sera la résolution. La nécessité d'enregistrer les spectres avec le plus grand domaine en énergie possible est donc évidente, mais cette condition impose d'améliorer la qualité des données. En biologie, compte tenu des concentrations protéiques, le temps d'accumulation est important et un faisceau extrêmement stable est donc nécessaire. L'augmentation du nombre de détecteurs (que l'on peut constater sur de nombreuses lignes consacrées à la SAX) a permis de diminuer le temps d'accumulation mais en tout état de cause, il reste toujours un compromis entre le temps d'accumulation pour obtenir un spectre avec le meilleur rapport signal sur bruit, et le temps de faisceau alloué.

En plus du caractère physique des valeurs des paramètres structuraux obtenus, on peut juger de la qualité des résultats grâce à la valeur du résiduel donné par le programme SEDEM. Il représente une estimation de la corrélation du modèle que l'on affine et des données expérimentales. Sa valeur doit être la plus faible possible.

4 La spectroscopie de masse

L'étude de NikA par spectroscopie de masse a été réalisée en collaboration avec David Lemaire, du Laboratoire de Spectroscopie de Masse des Protéines (LSMP) de l'IBS.
La caractérisation par spectroscopie de masse de petites molécules a été effectuée en collaboration avec Stéphane Ménage, du Laboratoire de Chimie Biologique (CB) du CEA.

4.1 Le principe de la spectroscopie de masse

La spectroscopie de masse est une technique permettant d'obtenir la masse d'un atome, ou dans notre cas, d'une molécule.

4.1.1 L'électrospray

Pour ce faire, la molécule que l'on veut étudier doit d'abord être ionisée. Ce processus se fait par l'utilisation d'une source d'ions électrospray. Il s'agit d'une technique qui a révolutionné la spectrométrie de masse car elle est dite "douce" : elle ne fragmente pas les molécules et peut s'appliquer à des composés polaires déjà chargés. Dans le cas des protéines, la nature du tampon peut imposer le nombre et les états de charge de la molécule. La solution protéique est introduite par infusion dans un capillaire porté à un haut potentiel électrique. Le champ électrique intense appliqué entre le capillaire et une contre-électrode sous un flux d'azote chauffé provoque la formation d'un nuage de gouttelettes chargées (nébulisation, Figure 4.1-1).

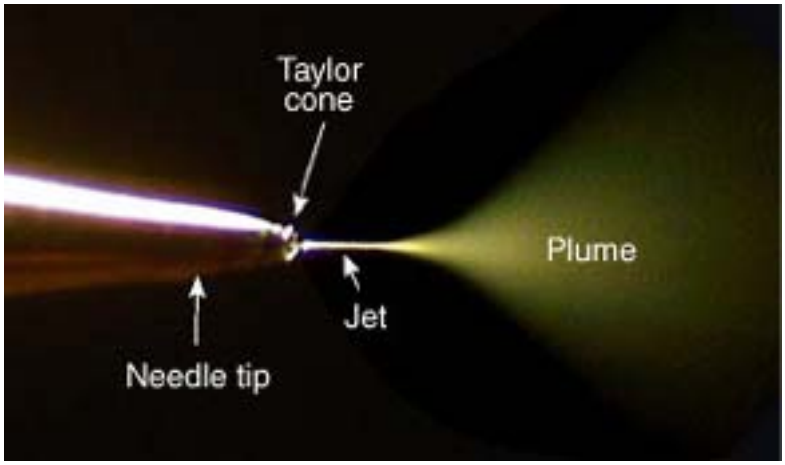


Figure 4.1-1 : Formation d'un nuage de gouttelettes chargées à la sortie du capillaire (http://www.ohsu.edu/proteomics/resources/techniques.cfm).

Pendant leur transport vers l'analyseur, ces dernières vont voir leur taille diminuer, d'une part car elles sont soumises à une atmosphère d'azote chaud et sec et d'autre part car les charges qu'elles contiennent vont s'opposer et provoquer leur éclatement (explosions « coulombiennes » successives). Les ions ainsi formés à pression atmosphérique vont être dirigés vers l'analyseur où il règne un vide poussé.
Tous les spectres de masse que nous avons enregistrés impliquaient l'utilisation de cette technique.

4.1.2 La séparation et détection des ions

L'étape suivante est la séparation des ions générés en fonction de leur rapport m/z (avec m la masse de la molécule et z sa charge) par un champ magnétique uniforme perpendiculaire au plan de leur trajectoire. Sous l'effet de la force de Lorentz, la trajectoire de la molécule adopte un rayon de courbure fonction de sa masse et de sa charge. En déterminant le point d'impact de l'ion sur un détecteur, on estime précisément sa masse (m/z, avec m la masse de l'ion ; z la charge de l'ion).
Dans le cas des expériences menées sur la protéine NikA, les spectres ont été obtenus sur un spectromètre de masse de géométrie hybride Q-TOF, le Q-TOF micro (Micromass, Manchester, U.K. ; Figure 4.1-2).

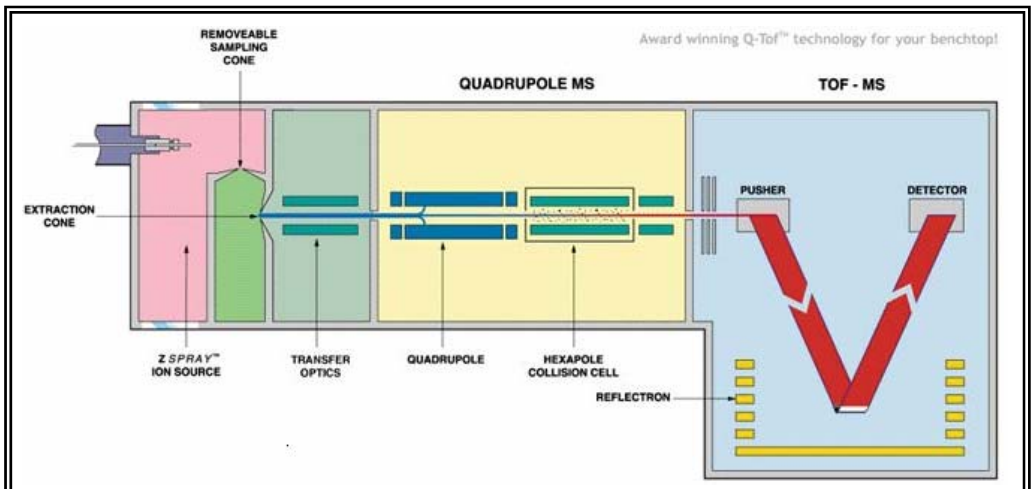


Figure 4.1-2 : Schéma de fonctionnement du spectromètre de masse Q-TOF micro (http://www.trentu.ca/wqc/organicinstrument.html).

La mesure de masse de l'échantillon se fait par l'analyseur à temps de vol (TOF). Dans son principe, c'est l'analyseur le plus simple de tous. Après une accélération initiale à l'entrée du tube de vol libre de champ, l'énergie cinétique est une fonction de la vitesse et de la masse.

L'analyse TOF consiste donc à séparer les ions en fonction de leur vitesse. Il s'agit par conséquent de la réalisation d'une mesure de temps. Les ions "volent" d'autant plus vite qu'ils sont légers. Les TOF mesurent donc le temps que mettent les ions pour franchir une distance fixe. Le temps final se mesure à l'arrivée sur le détecteur.

4.2 La spectroscopie de masse de la protéine NikA

4.2.1 La préparation des échantillons et la réalisation de l'expérience en conditions dénaturantes

Afin de vérifier l'intégrité de la protéine NikA après purification par le protocole contenant de l'EDTA et par celui contenant du chloroforme, nous avons enregistré dans les deux cas les spectres de masse, en conditions dénaturantes, et dans les mêmes conditions.
La protéine purifiée est concentrée à 700 nM dans une solution H_2O/acétonitrile (1/1, rapport volume/volume) avec 0,2% d'acide formique, afin de dénaturer et charger la protéine. L'échantillon est injecté à 5 µL/min. Le spectromètre est réglé avec une tension de capillaire de 3 kV, et les voltages du cône d'échantillon et du cône d'extraction sont respectivement de 45 et 2 V. Le spectre de masse est enregistré entre 700 et 2000 m/z. Les données sont enregistrées en mode positif, et la calibration du spectromètre est faite grâce aux états de charge de la myoglobine de cœur de cheval, dissoute dans une solution H_2O/acétonitrile (1/1, rapport volume/volume) avec 0,2% d'acide formique. Les données sont enregistrées et traitées avec le programme MasLynx 4.0 (Micromass).

4.2.2 La préparation des échantillons, et la réalisation de l'expérience en conditions natives

Afin de caractériser la masse du contenu du site de fixation de NikA purifiée par le protocole contenant de l'EDTA et par celui contenant du chloroforme, nous avons enregistré dans les deux cas les spectres de masse en conditions natives.
Dans les deux cas, l'échantillon est dans un premier temps dialysé sur centricon® (Millipore) contre une solution d'acétate d'ammonium 20 mM, dans le but d'éliminer toutes les fixations entre la protéine et des molécules du tampon. Puis, on se place à une concentration de 25 et 27 µM, respectivement pour la protéine purifiée grâce au protocole contenant de l'EDTA, et celle contenant du chloroforme. De cette façon, NikA ne doit plus fixer, à la fin de cette étape, que le contenu de son site de fixation. L'échantillon est alors injecté dans le spectromètre de masse, sans changement de tampon supplémentaire. La protéine est donc dans un état natif, la

structure est conservée, mais dans ce cas le nombre d'états de charge est moindre par rapport à la protéine dans l'expérience dénaturante.

Dans le cas de la protéine purifiée par le protocole contenant de l'EDTA, l'échantillon est injecté à 7 µL/min. Le spectromètre est réglé avec une tension de capillaire de 2,7 kV, et les voltages du cône d'échantillon et du cône d'extraction sont respectivement de 90 et 5V. Le spectre de masse est enregistré entre 2700 et 4700 m/z. Les données sont enregistrées en mode positif, et la calibration du spectromètre est faite grâce une solution de CsI 0,5 mg/mL dans H_2O/isopropanol (1/1, rapport volume/volume). Les données sont enregistrées et traitées avec le programme MasLynx 4.0 (Micromass).

Dans le cas de la protéine purifiée par le protocole contenant du chloroforme, l'échantillon est injecté à 7 µL/min. Le spectromètre est réglé avec une tension de capillaire de 3 kV, et les voltages du cône d'échantillon et du cône d'extraction sont respectivement de 150 et 2 V. Le spectre de masse est enregistré entre 2700 et 6000 m/z. Les données sont enregistrées en mode positif, et la calibration du spectromètre est faite grâce une solution de CsI 0,5 mg/mL dans H_2O/isopropanol (1/1, rapport volume/volume). Les données sont enregistrées et traitées avec le programme MasLynx 4.0 (Micromass).

4.3 L'extraction et analyse par spectroscopie de masse de la molécule présente dans NikA

L'objectif de cette expérience était de caractériser la masse de la petite molécule présente dans NikA purifiée avec le protocole de purification contenant du chloroforme.

La première analyse s'est faite à partir d'un échantillon concentré de la protéine, préparé de la façon suivante :

- Une solution de NikA 125 µM (7 mg/ml) est chauffé à 75°C pendant 20 minutes ;
- La protéine précipitée est éliminée par centrifugation ;
- Le surnageant est récupéré et évaporé jusqu'à un volume d'environ 10 µL, puis récupéré pour être analysé en spectroscopie de masse.

De cette façon, on peut admettre que la molécule se trouve à une concentration maximale de 160 µM.

La seconde analyse a été réalisée sur des échantillons préparés initialement pour une étude par RMN. Deux échantillons de NikA furent, dans un premier temps, dialysés sur centricon® (Millipore) contre une solution d'acétate d'ammonium 40 mM. Puis, chacun des deux échantillons est traité de la façon suivante afin d'extraire la molécule de la protéine :

- A 172 nmol de protéine fut ajouté cinquante fois plus d'EDTA (8,6 µmol). On laissa la préparation 20 heures à température ambiante. Puis la protéine fut retirée par concentration sur microcon® (Millipore). On récupéra le filtrat.
- On applique un gradient de température à 166 nmol de protéine (concentration : 0,7 mg/mL) permettant de la précipiter progressivement. Puis, cette dernière est retirée par filtration sur un filtre 0,22 µm à 40°C. On récupère le filtrat. Le gradient de température est réalisé selon le protocole suivant.
  - ⇒ On incube la protéine à 30°C pendant 15 min ;
  - ⇒ Entre 30 et 62°C la température est augmentée d'1°C toutes les 5 min ;
  - ⇒ On laisse la protéine à 63°C pendant 1 heure;
  - ⇒ On laisse la protéine à 75°C pendant 30 min.

Les deux filtrats récupérés sont lyophilisés et redissouts dans 400 µL de D_2O, pour être alors utilisés dans l'expérience RMN. De cette façon, on peut admettre que la molécule, que l'on a tenté d'extraire par les traitements décrits précédemment, se trouve à une concentration maximale de 415 et 430 µM respectivement pour le traitement à l'EDTA et le traitement à la chaleur. Par la suite, 10 µL furent prélevés et analysés par spectroscopie de masse.
Dans tous les cas présentés ici, pour calculer ces concentrations, il faut poser l'hypothèse que chaque protéine contenait et a libéré la molécule que l'on tente ici d'identifier.

5 La Résonance Magnétique Nucléaire (RMN)

Ce travail a été réalisé en collaboration avec Jean-Pierre Simorre, du Laboratoire de Résonance Magnétique Nucléaire (LRMN) de l'IBS.

5.1 Le principe de la RMN

Un noyau atomique est observable en RMN s'il possède un spin nucléaire (I) non nul. Par exemple, l'atome d'hydrogène d'un seul proton, donc son spin nucléaire (qui correspond au moment magnétique du noyau), est égal à celui du proton : I = ± ½. Les autres atomes peuvent avoir un spin multiple de ½ (par exemple ^{13}C) ou nul (par exemple ^{12}C). En présence d'un champ magnétique ($\vec{B}_0$), on observe une levée de dégénérescence des niveaux d'énergie Zeeman du spin nucléaire. Par exemple, dans le cas du proton isolé, en l'absence de champ magnétique, le moment magnétique de spin du noyau est orienté de façon quelconque. Par contre, en présence d'un champ magnétique ($\vec{B}_0$) le moment magnétique de spin s'oriente, soit de façon parallèle au champ : m = + ½, c'est le niveau énergétique le plus stable ; soit de façon antiparallèle : m = – ½. Dans le cas d'une population d'atomes, les moments magnétiques vont se répartir dans les deux niveaux d'énergie, avec une préférence pour le plus stable : m = + ½ (distribution de Boltzman).

L'expérience RMN consiste alors dans un premier temps, à peupler le niveau d'énergie m = – ½ avec des moments magnétiques de spin du niveau m = + ½ en appliquant un champ magnétique oscillant ($\vec{B}_1$) perpendiculaire à $\vec{B}_0$. La fréquence de $\vec{B}_1$ est choisie en fonction de la nature du noyau que l'on veut étudier, et correspond à l'énergie permettant de passer de m = + ½ à m = – ½. Puis, on mesure le retour à l'équilibre de la répartition des moments magnétiques de spin. Un signal RMN sera détecté à la fréquence (déplacement chimique) correspondant à la différence d'énergie entre les deux niveaux.

Dans le cas d'un spin non isolé, les atomes voisins perturbent le champ magnétique local ressenti. La présence d'électrons dans le voisinage de l'atome crée une constante d'écrantage qui est à l'origine des variations de déplacement chimique. L'analyse du déplacement chimique des énergies d'absorption fournit des renseignements précieux sur l'environnement local des noyaux atomiques (symétrie, nature des voisins proches).

5.2 La préparation des échantillons, et la réalisation de l'expérience

Dans le cadre de ce travail, les spectres RMN enregistrés furent des types suivants :

- RMN 1D du proton : donnant les corrélations entre protons voisins ;
- RMN 1D du ^{13}C (en abondance naturelle) : donnant les corrélations entre carbones voisins ;
- RMN 2D ^{1}H - ^{13}C (HMQC) : donnant les corrélations entre protons et carbones liés ;
- RMN 2D TOCSY : donnant les corrélations entre protons voisins dans la séquence (séparés par 3 liaisons au maximum).

L'objectif de ces expériences était de caractériser la nature d'une molécule présente dans le site de fixation de NikA purifiée avec le protocole contenant du chloroforme.
Pour ce faire, deux échantillons protéiques furent utilisés. Le protocole de leur préparation est décrit dans le chapitre traitant de la spectroscopie de masse (chapitre 4.3 du matériel et méthodes). Les échantillons sont à une concentration maximale de 415 et 430 µM respectivement pour le traitement à l'EDTA et le traitement à la chaleur.

6 La Résonance Paramagnétique Electronique (RPE)

Ce travail a été réalisé en collaboration avec Jacques Gaillard, du Service de Chimie Inorganique et Biologique, du Département de Recherche Fondamentale sur la Matière Condensée (CEA, Grenoble).

6.1 Le principe de la RPE

La RPE est utilisée pour caractériser des ions ou des molécules ayant des électrons non appariés.

De la même façon que les noyaux avec un nombre de masse impair possèdent un spin nucléaire, l'électron possède un spin électronique : $S = ½$. Dans le cas où un atome est entouré de plusieurs électrons, ils vont s'apparier deux à deux (l'un adoptant $M_S = + ½$ ou $– ½$). Le spin total (S) de la molécule sera alors nul. Mais lorsque ce n'est pas le cas, le spin la molécule sera un multiple de ½, et la RPE peut s'appliquer ($S \neq 0$). Cette dernière situation se rencontre avec les radicaux libres (un électron non apparié) ou avec des molécules incorporant des atomes métalliques (un ou plusieurs électrons non appariés selon la nature du ou des atomes métalliques impliqués). De façon similaire à ce qui a été décrit pour la RMN, les différents états d'énergie M_S du spin, associés aux états de spin total S de la molécule en présence d'un champ magnétique ($\vec{B}_0$), se distinguent par des énergies différentes. Comme pour la RMN, on applique un champ magnétique $\vec{B}_1$ perpendiculaire à $\vec{B}_0$, et on observe les transitions entre les différents états M_S.

Pour l'étude des métalloprotéines, la spectroscopie RPE renseigne en particulier sur la nature du ou des atomes métalliques impliqués, sur leur degré d'oxydation, sur la répartition de la densité de spin et également sur la nature des atomes entourant les atomes métalliques, au travers de l'interaction hyperfine, si ceux-ci sont porteurs de spins nucléaires.

6.2 La préparation des échantillons, et la réalisation de l'expérience

Les spectres RPE ont été enregistrés sur un spectromètre Bruker EMX travaillant en bande X (fréquence ~ 9560 MHz), à la température de l'hélium liquide, atteinte grâce à l'utilisation d'un cryostat Oxford Instrument ESR 900 par Jacques Gaillard. Les tubes RPE remplis d'environ 200 µL d'un échantillon de protéine NikA à 200 µM dans une solution de Tris base pH 7,5 40 mM étaient congelés au préalable.

7 Le Dichroïsme Circulaire (CD)

Ce travail a été réalisé en collaboration avec Dominique Madern, du Laboratoire de Biophysique Moléculaire (LBM) de l'IBS.

7.1 Le principe du CD

Le dichroïsme circulaire est une technique spectroscopique reposant sur le fait que les structures secondaires des protéines n'absorbent pas de la même façon la lumière polarisée circulairement à droite et celle polarisé circulairement à gauche. La résultante de lumières polarisées à 90° l'une de l'autre et déphasées de ± π/2 est une lumière polarisée en spirale de pas droit ou gauche : c'est ce qu'on appelle la lumière polarisée circulairement à droite et à gauche (Figure 7.1-1a).

Dans l'expérience de dichroïsme circulaire, l'échantillon protéique est soumis simultanément à une lumière polarisée à gauche et à droite. Comme les structures secondaires de la protéine n'absorbent pas de façon égale ces deux rayonnements, la résultante présentera une déviation à droite ou à gauche. C'est elle que l'on appelle dichroïsme circulaire, que l'on enregistre en pratique sur une gamme de longueurs d'onde (Figure 7.1-1b).

C'est une technique de choix pour juger rapidement de l'état de repliement d'une protéine en fonction de plusieurs paramètres tels que la température ou la nature du solvant, sans se préoccuper de l'activité de l'enzyme. De plus, elle n'est pas destructrice pour l'échantillon.

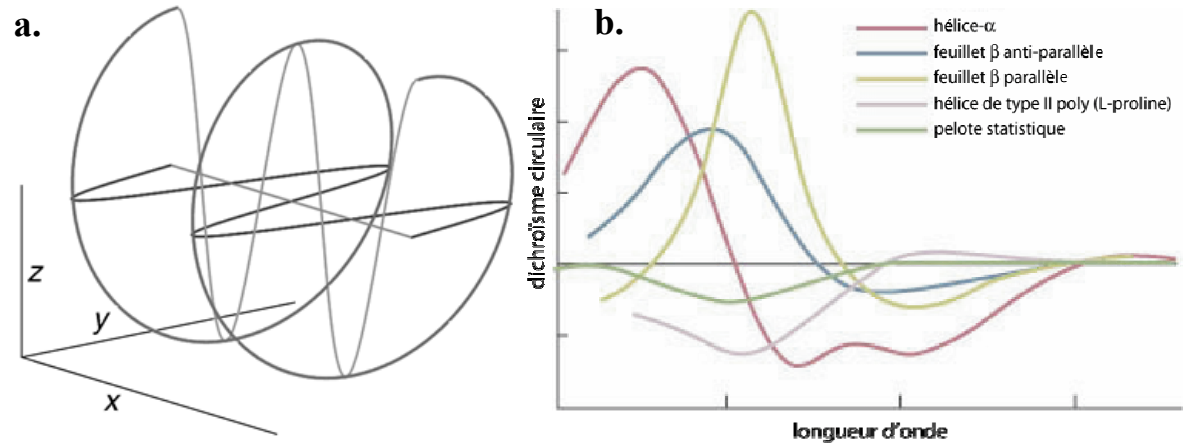


Figure 7.1-1 : Schéma du dichroïsme circulaire.
a. lumière polarisée circulairement à gauche (http://jchemed.chem.wisc.edu).
b. Spectres de dichroïsme circulaire théoriques (http://www.cryst.bbk.ac.uk).

7.2 La préparation des échantillons, et la réalisation de l'expérience

L'objectif de notre expérience de dichroïsme circulaire était de suivre le repliement de NikA en fonction de la température, afin de déterminer l'éventuelle existence d'un état intermédiaire dans le processus de dénaturation de la protéine, qui pourrait correspondre à une ouverture des deux domaines de la molécule.

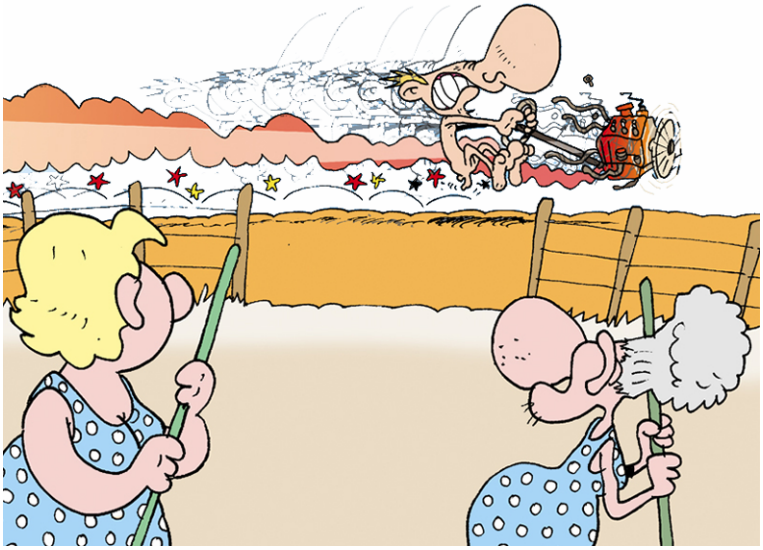
Pour ce faire, nous avons utilisé un échantillon de NikA concentré à 0,15 mg/mL dans un tampon phosphate pH 7,5 à 50 mM. Les spectres ont été enregistrés à 40, 45, 50, 55, 58, 61, 65 et 70°C (noté X°C dans la suite) de la façon suivante :

- La protéine est incubée à la température X°C pendant 15 min ;
- Chaque spectre est alors enregistré entre 190 et 260 nm à la température X°C.

ETUDE STRUCTURALE

DE LA PROTEINE

NIKR



L'étude structurale que nous avons effectuée sur la protéine NikR se fit sur une protéine fournie par Marie-Andrée Mandrand-Berthelot et Agnès Rodrigue, de l'Unité de Microbiologie et Génétique (UMG) de l'INSA de Lyon. Il s'agit en réalité d'un mutant Q2E de NikR capable de se fixer à l'ADN, mais en présence d'une concentration plus importante de nickel par rapport à la souche sauvage. En effet, cette mutation se trouve dans le domaine N terminal responsable de la fixation à l'ADN. La protéine est surexprimée dans *E. coli* et purifiée grâce à une colonne de nickel. Afin d'obtenir une population de protéine homogène ne contenant pas de nickel, 2 mM d'EDTA sont ajoutés.

1 La cristallogenèse de la protéine NikR

Lorsque nous avons reçu la protéine pour la première fois, nous avons tenté de caractériser l'homogénéité de la préparation en effectuant une expérience de diffusion de lumière. NikR s'est alors révélée être très hétérogène (données non présentées), puisque l'on a mis en évidence l'existence d'un grand nombre d'agrégats de taille importante. Nous avons tout de même réalisé les premiers tests de cristallisation avec une solution de protéine concentrée à 6,0 mg/mL, en utilisant la technique de cristallisation en gouttes suspendues à 20°C. Afin de déterminer le comportement de la protéine vis à vis du processus de cristallisation, nous avons utilisé les screens de cristallisation : *Jancarik* et *Screen II* (Hampton Research). De cette façon, nous avons pu évaluer le comportement de la protéine vis-à-vis d'un grand nombre de conditions de cristallisation différentes. Après une journée, deux conditions ont permis l'apparition de premiers cristaux, qui ont par la suite mis un temps variable pour atteindre leur taille définitive mais dont la qualité n'était pas suffisante pour être soumis à l'expérience de diffraction :

- 20% PEG 8.000 ; 0,1 M cacodylate de sodium pH 6,5 ; 0,2 M acétate de magnésium ;
- 20% isopropanol ; 0,1 M HEPES pH 7,5 ; 0,2 M $CaCl_2$.

Une fois ces premiers résultats obtenus, il nous a fallu améliorer la qualité des cristaux en affinant les conditions de cristallisation.

1.1 Les cristaux de NikR obtenus en utilisant l'isopropanol comme précipitant

Dans un premier temps, nous nous sommes focalisés uniquement sur l'amélioration des cristaux apparus dans la condition ayant l'isopropanol comme agent précipitant. En effet, ces derniers nous apparaissaient comme les plus gros et les mieux formés. Dans ces conditions, des agrégats de cristaux apparurent, mais certains semblèrent monocristallins et détachables (Figure 1.1-1), donc adaptés à une expérience de diffraction des rayons X.

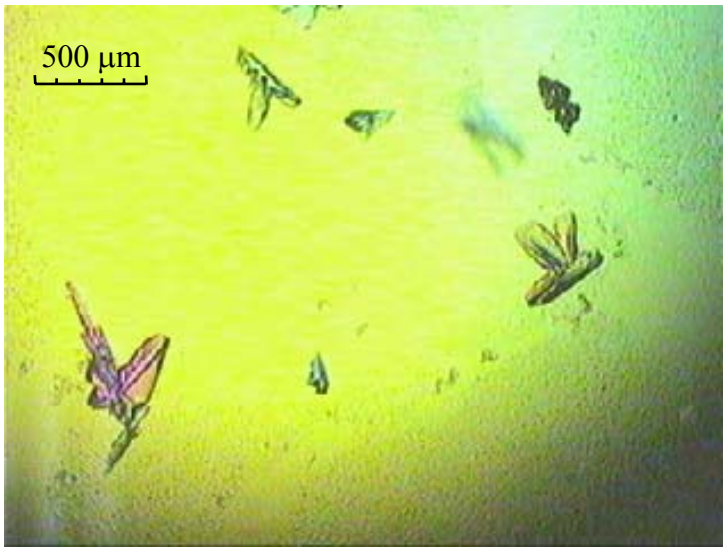


Figure 1.1-1 : Cristaux de NikR ayant poussé dans :
18% isopropanol ; 0,1 M HEPES pH 7,5 ; 200 mM $CaCl_2$; 4% PEG 400.

Afin d'améliorer la qualité et la taille de ces cristaux de NikR, nous avons fait varier différents paramètres :

- La concentration en isopropanol (l'agent précipitant) ;
- La nature et le pH du tampon utilisé ;
- La nature et la concentration du sel utilisé ;
- La température de cristallisation : tous les tests essayés préalablement sont réalisés à 20°C, on essaie donc de faire cristalliser la protéine à 4°C, souhaitant ainsi ralentir la croissance des cristaux et peut-être obtenir des cristaux plus gros et de meilleure qualité ;
- L'ajout d'additifs tels que par exemple 4% de PEG 400, dans l'espoir de ralentir, comme dans le cas de la température, la croissance cristalline.

Ces nombreux tests de conditions de cristallisation différentes, nous ont permis, en suivant une logique de recherche systématique, de faire les observations suivantes :

- Plus la concentration en isopropanol est importante, plus les cristaux sont petits et polycristallins. La concentration optimale de précipitant est comprise entre 14 et 16%. Lorsque la concentration en isopropanol est inférieure à 12% la protéine ne cristallise plus ;
- Le pH optimal est 7,5. En effet, à pH 8, des cristaux apparaissent, mais sont assez mal formés et d'assez petite taille ; à pH 7 les cristaux sont exploitables mais sont plus polycristallins et mosaïques qu'à pH 7,5. Ces résultats ont été obtenus avec une concentration de tampon d'HEPES 100 mM.
- Dans un premier temps, on a constaté que NikR cristallisait très bien en présence de $CaCl_2$. L'augmentation de la concentration de ce dernier permet d'augmenter la taille et la qualité des cristaux, avec une concentration optimale comprise entre 100 et 200 mM. Dans un second temps, afin d'évaluer l'importance de la présence de calcium dans les conditions de cristallisation, on a tenté de le remplacer par d'autres ions divalents, comme par exemple du $MgCl_2$. Des cristaux sont obtenus avec des concentrations de sel comprises entre 10 et 20 mM, mais présentent une faible qualité de diffraction : entre 7 et 9 Å de résolution maximum. De ce fait, le calcium est indispensable à une bonne croissance cristalline. Nous avons alors tenté d'évaluer l'importance du chlorure en le remplaçant par de l'acétate (en utilisant l'acétate de calcium). De la même façon qu'avec le $MgCl_2$, des cristaux apparaissent à une concentration comprise entre 100 et 200 mM, mais ils présentent une faible qualité de diffraction. Il est alors intéressant de remarquer que les deux ions (Ca^{2+} et Cl^-) sont indispensables à l'apparition de cristaux présentant une bonne qualité de diffraction ;
- La température a un effet radical sur la croissance cristalline ; en effet, les conditions qui donnent des cristaux à 20°C, font précipiter la protéine à 4°C ;
- Enfin, la présence de 4% de PEG 400 dans les gouttes de cristallisation, qui a été l'un des nombreux additifs testés, n'a pas modifié la qualité des cristaux. Mais d'un autre côté, n'empêchant pas la croissance cristalline, le PEG 400 pourrait faire un bon cryoprotectant.

Nous avons, par la suite, voulu enregistrer des données de diffraction de ces cristaux. Le premier problème qui s'est présenté à nous a été le caractère volatil de l'isopropanol. En effet, dès que l'on ouvre le système lamelle/puits, il s'évapore, entraînant des mouvements de convection dans la goutte de cristallisation. Il devient de ce fait très difficile de prélever les cristaux sans qu'ils ne s'abîment, car ils se mettent à tourner rapidement dans la goutte. Il ne nous était, de plus, pas possible d'attendre que l'équilibre ne se fasse avec l'évaporation de l'isopropanol, car son absence provoquait une détérioration très rapide du cristal. La seule solution restait de pêcher les cristaux très rapidement malgré leur rotation.

1.2 Les cristaux de NikR obtenus en utilisant les PEG comme précipitant

Du fait de la grande difficulté que présentait le prélèvement des cristaux ayant poussé en présence d'isopropanol, et de la faible qualité des données de diffraction (décrite dans le chapitre suivant), nous avons décidé de nous intéresser aux cristaux obtenus dans des conditions à base de PEG, qui atteignaient leur taille maximale moins vite : quinze jours au lieu de trois pour l'isopropanol. Afin d'améliorer la taille et la qualité des cristaux, nous avons testé, en suivant comme précédemment une logique de recherche systématique, les conditions suivantes :

- La variation de la nature et de la concentration des PEG : du PEG 1 000 jusqu'au PEG 20 000, plus le PEG 5 000 mme (monométhyléther) ;
- La variation de la nature et du pH des tampons ;
- La température de croissance des cristaux : 20 ou 4°C ;
- La concentration de NikR : au lieu de faire les gouttes a partir d'une protéine à 6,0 mg/mL on a utilisé une concentration de 5,0 mg/mL ;
- L'ajout d'additifs tels que : 200 mM $CaCl_2$, 5% glycérol, 100 mM NaCl, 3% isopropanol, 10 mM $CdCl_2$, et des détergents ;
- Ajout de nickel : $NiCl_2$ ou $NiSO_4$ à 200 ou 400 µM (soit respectivement 0,5 et 1 équivalent de nickel par rapport à NikR).

Deux faciès de cristaux ont alors été obtenus : des « prismes » épais (environ 0,15 x 0,15 x 0,050 mm^3), mais peu nombreux et des « plaquettes » peu épaisses (environ 0,40 x 0,05 x 0,01 mm^3) (Figure 1.2-1).

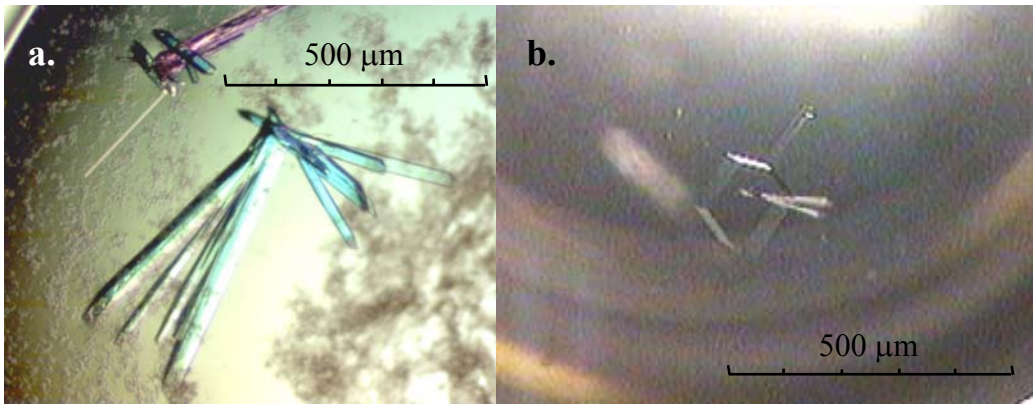


Figure 1.2-1 : Cristaux de NikR, ayant poussé dans des conditions à base de PEG.
a. forme « plaquettes ».
b. forme « prisme ».

Les tests que nous avons réalisés nous ont permis de déterminer les caractéristiques suivantes pour NikR :

- La forme « plaquette » est la plus fréquente, mais il paraît très difficile de conclure quant aux paramètres favorisant un faciès par rapport à un autre ;
- On constate que la nature du PEG est un paramètre primordial pour l'apparition et la qualité des cristaux, puisque plus la taille du PEG est importante, meilleure est la qualité des cristaux. Les meilleurs résultats sont obtenus avec les PEG 20 000, 10 000, 8 000, mais avec des concentrations optimales différentes : entre 8 et 16% pour le PEG 8 000 (en dessous on observe seulement l'apparition de microcristaux). Dans le cas des PEG 5 000 mme, 10 000, et 20 000, la concentration optimale semble être autour de 8% (la nucléation devient trop importante au-delà de 12 à 14%) ;
- Le pH optimal, quel que soit le PEG, est 6,5 (MES 0,1 M), ou 7 (Tris 0,1 M) ;
- Des cristaux sont obtenus à 20°C, par contre à 4°C la protéine précipite ;
- L'ajout de certains additifs améliore la qualité des cristaux : 100 mM de NaCl à faible concentration de PEG 5 000 mme ; et surtout 20 mM de βOG (n-octyl-β-D-glucopyranoside) avec 10 et 15% de PEG 5 000 mme, qui augmentent l'épaisseur des cristaux. Mais leur qualité de diffraction ne s'est pas révélée suffisante ;
- L'ajout de nickel ne semble pas modifier la croissance cristalline : on obtient des cristaux des deux faciès en présence et en absence des deux sels de nickel testés.

Il est à noter que tous ces résultats ont été obtenus avec la protéine NikR contenant 2 mM d'EDTA. Si l'on dialyse la protéine pour le supprimer, et que l'on essaye de la cristalliser, on ne constate pas de différence significative avec les résultats décrits plus haut. Par contre, en rajoutant du nickel à cette protéine dialysée, nous n'avons pas pu réobtenir de cristaux. Il se peut qu'un changement de conformation de la protéine perturbe l'empilement cristallin, et ainsi empêche l'apparition de cristaux. Mais ces résultats seraient à confirmer, puisque peu de tests ont été menés.

2 L'enregistrement et traitement des jeux de diffraction

2.1 Les cristaux de NikR obtenus en utilisant l'isopropanol comme précipitant

Le premier cristal dont nous avons obtenu les images de diffraction provient de la goutte à base d'isopropanol du screen *Jancarik*. Il a été congelé avec une solution cryoprotectrice, une concentration plus élevée en précipitant que la solution réservoir, et 30% de glycérol comme agent cryoprotectant, et a été enregistré sur la ligne ID13 (ligne microfocus) de l'ESRF. Nous avons enregistré 180° (en effectuant des pas de 1°), mais la congélation avait endommagé le cristal, comme le montrait son faciès dégradé et sa mosaïcité très importante. Le traitement des données n'a permis d'obtenir que les paramètres de la maille : le cristal de NikR appartenait donc au groupe d'espace monoclinique *P*2, avec **a** = 44,3 Å ; **b** = 122,0 Å ; **c** = 51,4 Å ; α = 90,0° ; β = 122,0° ; γ = 90,0°.

Le glycérol pouvant être à l'origine de la dégradation du cristal, nous avons par la suite testé un ensemble de cryoprotectants comme par exemple le sucrose, l'éthylène glycol, le PEG 400, les huiles… La plupart ne convenaient pas car :

- Soit ils ne cryoprotégaient pas les cristaux ;
- Soit il y avait une séparation de phase avec l'isopropanol ;
- Ou soit le cristal s'abîmait.

Les seuls qui ont semblé remplir toutes les conditions requises pour un cryoprotectant, étaient le PEG 400 à 30% et l'huile de silicone. Ils nous ont ainsi permis de congeler des cristaux et de les tester sur la ligne de l'ESRF ID14eh2. Malheureusement, aucun des trois jeux de données enregistrées (sur des cristaux d'environ 300 x 100 μm^2 qui diffractaient entre 4,5 et 3,2 Å de résolution) n'ont pu être indexés. En effet, il n'a pas été possible, à l'époque, de

déterminer une maille prédisant les taches de diffraction de ce cristal de NikR, du fait de la très mauvaise qualité des données de diffraction : ils se sont révélés présenter une grande mosaïcité et être polycristallins.

2.2 Les cristaux de NikR obtenus en utilisant les PEG comme précipitant

2.2.1 Les cristaux cryoprotégés dans l'azote liquide

Afin de pouvoir être testés en diffraction, les cristaux de NikR ont été cryoprotégés par remplacement progressif de la solution de cristallisation par une solution composée de : 3% de PEG de plus que la composition de la solution de cristallisation (afin de stabiliser les cristaux), 20% de PEG 400 (le cryoprotectant), le même tampon et les mêmes additifs que la solution mère.

Plusieurs cristaux ont été testés, ce qui nous a permis de constater que la forme « prisme » était souvent très fortement mosaïque et posait des problèmes lors de la cryoprotection (il leur arrivait de se casser), et que la forme « plaquette », présentait une meilleure qualité de diffraction, mais était très sensible aux rayons X. Dans les deux cas, on a constaté une chute de l'intensité de diffraction lorsque la résolution dépasse 3 Å. Mais différents jeux de données ont tout de même pu être enregistrés.

Un jeu de données de diffraction a pu être enregistré sur la ligne de lumière de l'ESRF ID14eh1 à partir d'une « plaquette » relativement épaisse (par rapport aux autres cristaux obtenus), de 0,50 x 0,04 x 0,01 mm^3. Elle est apparue dans les conditions de cristallisation suivantes :

[NikR] = 6 mg/mL;
14% de PEG 8 000 ;
0,1 M MES pH 6,5 ;
20°C en goutte suspendue.

Après traitement des données de diffraction (Tableau 2.2-1), ce cristal s'est révélé du groupe d'espace monoclinique *P*2, mais avec un axe pseudo 2_1, puisque l'on observe de faibles taches de diffraction là où des extinctions systématiques liées à l'axe hélicoïdal 2_1 correspondant à k = 2n + 1 devraient exister. Bien que le jeu de données soit relativement mosaïque (calculé avec XDS), il donne des statistiques de qualité raisonnable (Tableau 2.2-1).

Ligne de lumière de l'ESRF	ID14eh1
Longueur d'onde	0,934 Å
Groupe d'espace	*P*2
Paramètres de maille	a = 46,8 Å ; b = 117,2 Å ; c = 49,3 Å ; α = 90° ; β = 116,6° ; γ = 90°
Nombre de molécules par unité asymétrique	4
Mosaïcité	0,42°
Résolution	55,0–3,0 Å (3,2–3,0 Å)
Complétude	95,7% (90,3%)
I/σ(I)	8,6 (3,4)
Nombre total de réflexions mesurées	34 316 (5 737)
Redondance	3,6 (3,6)
R_{SYM}	7,7% (24,3%)

Tableau 2.2-1 : Statistiques des données de diffraction d'un jeu natif de la protéine NikR congelée dans l'azote liquide. Les valeurs entre parenthèses correspondent à la dernière tranche de résolution.

2.2.2 Les cristaux cryoprotégés dans le propane liquide

L'un des principaux problèmes que nous avions avec les cristaux de NikR, était leur grande mosaïcité et la relative mauvaise qualité de leurs données de diffraction. Puisque tous les essais d'amélioration de la qualité des cristaux n'ont pas permis de s'affranchir de ce problème, nous avons émis l'hypothèse qu'il devait provenir de l'étape de congélation des cristaux. Pourtant nous avions déjà testé plusieurs cryoprotectants, et aucun n'était apparu comme meilleur que le PEG 400. Cela ne nous a donc laissé plus qu'une seule possibilité d'amélioration de la qualité de diffraction : remplacer la congélation à l'azote liquide par une congélation au propane liquide (chapitre 2.3.1.1 du matériel et méthode).

C'est ce que nous avons réalisé en utilisant la solution suivante comme cryoprotectant : 3% de PEG supplémentaire, par rapport à la composition de la solution de cristallisation, 30% de PEG 400, le même tampon et les mêmes additifs que la solution mère. Les cristaux avaient la même limite de résolution que dans le cas d'une congélation à l'azote, mais par contre ils sont apparus peu mosaïques. Ceci est corrélé à une amélioration des paramètres statistiques reflétant la qualité de la diffraction des cristaux.

Dans ce cas, le meilleur jeu de données de diffraction natif que nous ayons obtenu provenait d'un cristal ayant poussé dans les conditions suivantes :

[NikR] = 6 mg/mL ;

PEG 8 000 14% ;

MES 0,1 M pH 6,5 ;

$NiCl_2$ 400 µM (1 équivalent Ni / NikR) ;

20°C en goutte suspendue.

Il s'agissait d'un cristal possédant un faciès de « plaquette ». Il est apparu très peu mosaïque et les données de diffraction d'une qualité meilleure (Tableau 2.2-2) que dans le cas des cristaux congelés dans l'azote.

Ligne de lumière de l'ESRF	ID14eh4
Longueur d'onde	0,9393
Groupe d'espace	$P2_1$
Paramètres de maille	a = 49,9 Å ; b = 116,5 Å ; c = 89,9 Å ; α = 90° ; β = 115,1° ; γ = 90°
Nombre de molécules par unité asymétrique	8
Mosaïcité	0,24°
Résolution	30,0–2,9 Å (3,0–2,9 Å)
Complétude	98,0% (92,1%)
I/σ(I)	13,5 (4,1)
Nombre total de réflexions mesurées	82 168 (7 419)
Redondance	3.8 (3.8)
R_{SYM}	6,0% (29,5%)

Tableau 2.2-2 : Statistiques des données de diffraction d'un jeu natif de la protéine NikR congelée dans le propane liquide. Les valeurs entre parenthèses correspondent à la dernière tranche de résolution.

2.3 Les Tentatives de résolution du problème de phases

Une étape importante du processus cristallographique est la résolution du problème de la phase. Pour ce faire, nous avons tenté d'utiliser plusieurs sources d'atomes lourds, afin d'employer les techniques de phasage MAD ou SAD.

2.3.1 Le jeu de données de diffraction SAD au seuil du nickel

Notre première tentative de phasage a consisté à utiliser le nickel, que l'on supposait présent dans la structure de NikR, car il est présent dans la goutte de cristallisation (mais pas dans la solution de cryoprotection, contrairement à la protéine native). Cette étude a été réalisée sur un cristal de protéine au faciès de « prisme » mesurant environ 0,230 x 0,080 x 0,070 mm^3. Il est apparu dans les conditions suivantes :

[NikR] = 6 mg/mL ;

PEG 20 000 9% ;

MES 0,1 M pH 6,5 ;

$NiCl_2$ 400 µM (1 équivalent Ni / NikR) ;

20°C en goutte suspendue.

Un jeu de données a été enregistré sur la ligne de lumière BM30A (FIP). La première étape a été d'enregistrer un spectre d'absorption au seuil du nickel. Le spectre obtenu était relativement bruité et décalé par rapport au seuil théorique. L'enregistrement des données de diffraction s'est alors fait au maximum de $^{\lambda}f''$, soit 8,349 KeV (1,48 Å).

Ligne de lumière de l'ESRF	BM30A (FIP)
Longueur d'onde	1,48 Å
Groupe d'espace	$P2_1$
Paramètres de maille	a = 50,6 Å ; b = 91,8 Å ; c = 49,4 Å ; α = 90°; β = 94,5° ; γ = 90°
Nombre de molécules par unité asymétrique	8
Mosaïcité	0,72°
Résolution	55,00–3,10 Å (3,14–3,10 Å)
Complétude	97,9% (90,4%)
I/σ(I)	12,4 (2,5)
Nombre total de réflexions mesurées	54 454 (1 007)
Redondance	3,5 (3,5)
R_{SYM}	12,2% (27,7%)

Tableau 2.3-1 : Statistiques des données de diffraction d'un jeu SAD au seul du nickel de la protéine NikR congelée dans l'azote liquide. Les valeurs entre parenthèses correspondent à la dernière tranche de résolution.

Après traitement des données de diffraction, le cristal s'est révélé du groupe d'espace monoclinique $P2_1$ avec des paramètres de maille (Tableau 2.3-1) différents des « plaquettes » décrites précédemment : les deux faciès correspondent donc à deux formes cristallines différentes. Bien que le cristal soit assez mosaïque puisque la congélation s'est faite dans l'azote liquide, ses autres statistiques sont correctes.

Afin de résoudre la structure de NikR par la technique du SAD, il nous faut dans un premier temps localiser la position du ou des nickels de la structure dans la maille cristalline. Pour ce faire, nous avons utilisé la fonction de Patterson. Mais après avoir calculé une carte de Patterson différence entre 55 et 3,5 Å de résolution, aucun pic de la section de Harker v = 0,5 (Figure 2.3-1) ne permit de déterminer de site de fixation du nickel, et ainsi de résoudre les équations de phase : les pics sont peu intenses, puisque aucun ne dépasse 4σ.

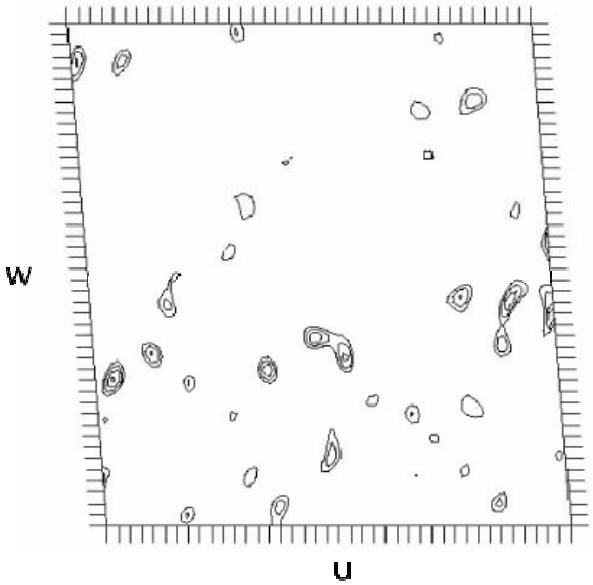


Figure 2.3-1 : Carte de Patterson différence dans sa section de Harker v = 0,5, calculée à partir du jeu SAD obtenu pour un cristal de NikR (sur la carte : base = 2σ ; contours = 0,5σ).

On peut émettre plusieurs hypothèses pour expliquer que le nickel n'apporte pas une diffusion anomale suffisante pour résoudre le problème des phases :

- Le nombre de nickels fixés à NikR peut ne pas être suffisant par rapport à la taille de la molécule, pour apporter une diffusion anomale suffisante (sachant que le nickel a un pouvoir anomal faible par rapport à d'autre atomes comme le sélénium) ;

- Le nickel peut également ne pas être fixé à la structure de manière très ordonnée : dans ce cas aucune diffusion anomale ne pourra pas être détectée, et la carte de Patterson différence ne fera pas apparaître de pics significatifs.

2.3.2 La recherche d'atome lourds se fixant à la structure de NikR permettant un phasage par la méthode MAD ou SAD

Nous avons donc tenté de réaliser des trempages et des co-cristallisations de la protéine avec différents types d'atomes lourds tels que par exemple le mercure ($HgCl_2$), l'or ($AuKCl_4$), ou même des sondes colorées ayant fixé un atome d'uranium (en collaboration avec l'équipe de Kahn R. de l'IBS).
Malheureusement, dans tous les cas, soit les cristaux ne supportaient pas l'ajout du métal et se détérioraient, soit l'atome lourd ne se fixait pas de façon ordonnée à la structure.

2.3.3 La production et cristallisation de la protéine NikR séléniée

La séquence de NikR ne présente que deux méthionines, dont une en N terminal, qui a toutes les chances d'être désordonnée dans la structure. Nous avons alors estimé qu'il serait difficile de mesurer le signal anomal de la protéine séléniée, au vu de la faible qualité de diffraction des cristaux, puisqu'il est communément admis qu'il faut au minimum un sélénium pour 14 kDa pour réussir à phaser une structure (ce qui correspond à un signal anomal de 3% ; (Sharff *et al.*, 2000)). Nous avons ainsi déterminé, grâce à un alignement de séquence et une comparaison structurale avec la protéine Arc, cinq leucines qu'il était possible de muter en méthionines, cette mutation n'induisant dans la plupart des cas que peu ou pas de changement quant à la stabilité et la fonction de la protéine (Bowie *et al.*, 1990; Lim *et al.*, 1992; Leahy *et al.*, 1994). De plus, deux de ces leucines (19 et 40) étaient remplacées par des méthionines chez des homologues de NikR d'*E. coli* (*B. suis* et *Pseudomonas putida*). Les mutations et la production de la protéine furent assurées par l'équipe de Marie-Andrée Mandrand-Berthelot. Plusieurs lots de protéine correspondant à différents mutants nous ont été fournis, mais aucun cristal exploitable ne put être obtenu.

3 Les structures de la protéine NikR

3.1 La structure de NikR d'*Escherichia coli*

Alors que nous n'avions pas encore résolu le problème de phasage de la structure de NikR, Schreiter et al. (Schreiter *et al.*, 2003) publièrent la structure de la forme apo de NikR

d'*E. coli* à 2,3 Å de résolution, ainsi que le domaine C terminal de la protéine fixant du nickel à 1,4 Å de résolution. Pour ce faire, ils produisirent la protéine séléniée et la cristallisèrent grâce aux conditions suivantes :

- Dans le cas de la protéine entière :

  [NikR] = 8 mg/mL ;

  PEG 6 000 6% ;

  Acide N-(2-(acétamido)imino)diacetique 0,1 M pH 6,5 ;

  $MgCl_2$ 50 mM ;

  Température ambiante, en goutte suspendue.

- Dans le cas du domaine C terminal de la protéine contenant du nickel :

  [NikR] = 10 mg/mL ;

  PEG 3 350 20% ;

  Titrate di-sodique 0,2 M ;

  Température ambiante, en goutte suspendue.

Dans tous les cas, les cristaux appartenaient au groupe d'espace orthorhombique $P2_12_12_1$ avec, pour la protéine entière, les paramètres de maille suivants :

$$a = 46{,}3 \text{ Å} ; b = 69{,}6 \text{ Å} ; c = 159{,}3 \text{ Å} ; \alpha = 90° ; \beta = 90° ; \gamma = 90°.$$

et pour le domaine C–terminal :

$$a = 45{,}9 \text{ Å} ; b = 78{,}3 \text{ Å} ; c = 81{,}7 \text{ Å} ; \alpha = 90° ; \beta = 90° ; \gamma = 90°.$$

3.1.1 La structure de la protéine apo entière

La structure de NikR est un homotétramère : les domaines C terminaux forment un tétramère auquel se rattachent les deux dimères du site de fixation à l'ADN (Figure 3.1-1).

Comme cela avait été prédit (Chivers and Sauer, 1999), le domaine N terminal de NikR adopte une conformation brin–hélice–hélice (βαα). En ce qui concerne le domaine C terminal, il apparaît avoir un repliement du type ferredoxine (βαββαβ).

La dimérisation de deux NikR se fait au niveau d'une interface hydrophobe constituée essentiellement des chaînes latérales des feuillets β des domaines C terminaux.

La tétramérisation est essentiellement assurée par des interactions hydrophobes et des liaisons hydrogènes entre les brins β.

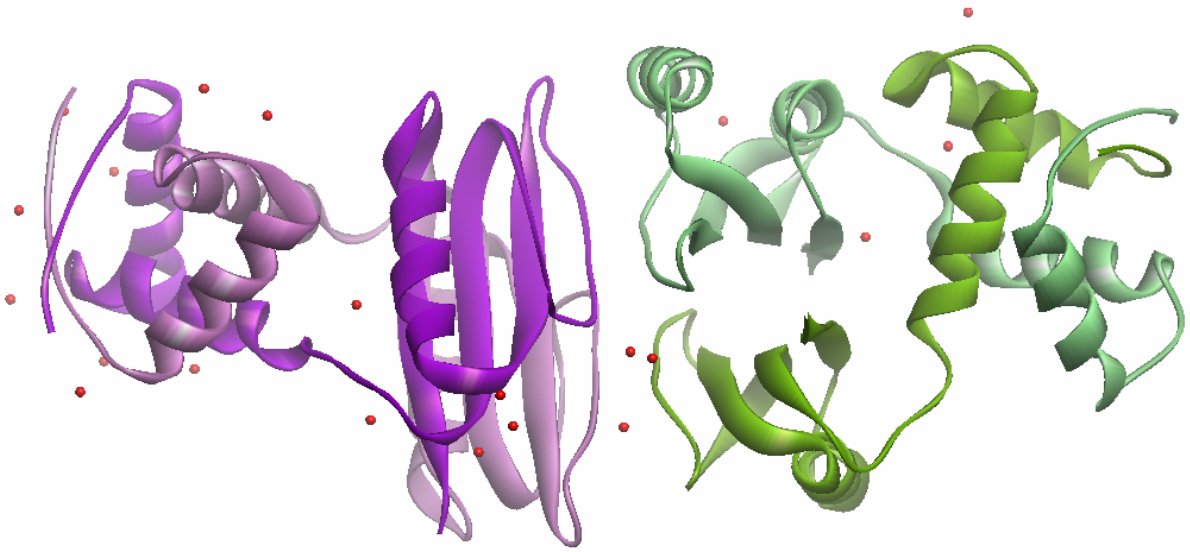


Figure 3.1-1 : Structure de l'homotétramère de NikR d'*E. coli* (code PDB : 1Q5V).
Le tétramère est constitué par la dimérisation de deux dimères ici représentés en vert et violet.
Au centre : le domaine fixant le nickel.
De chaque côté : les domaines de fixation à l'ADN.
Les molécules d'eau sont représentées par des sphères rouges

3.1.2 Le site de fixation du nickel

Les quatre nickels de la structure du domaine C terminal se trouvent à l'interface du tétramère. Dans chaque site de fixation, le métal adopte une conformation plan carré (elle correspond à une coordination d'un métal par quatre ligands présents dans le même plan, et à quatre-vingt dix degrés les uns des autres), ce qui est un élément qui doit assurer la spécificité de NikR pour le nickel vis-à-vis des autres métaux (le nickel (II) adoptant préférentiellement une géométrie plan carré (Cotton and Wilkinson)), assurée par les chaînes latérales des résidus His87, His89 et Cys95 d'un monomère de NikR et le résidu His76 du monomère adjacent dans le tétramère (Figure 3.1-2). Cette géométrie est en accord avec les études de SAX menées précédemment (Carrington *et al.*, 2003). Ces dernières montrent également que la coordination du nickel devient octaédrique lorsque NikR se fixe à l'ADN.

Ce site de fixation n'est pas équivalent aux autres sites de fixation du nickel des protéines décrites jusqu'à présent, et présentées dans l'introduction (Maroney, 1999).

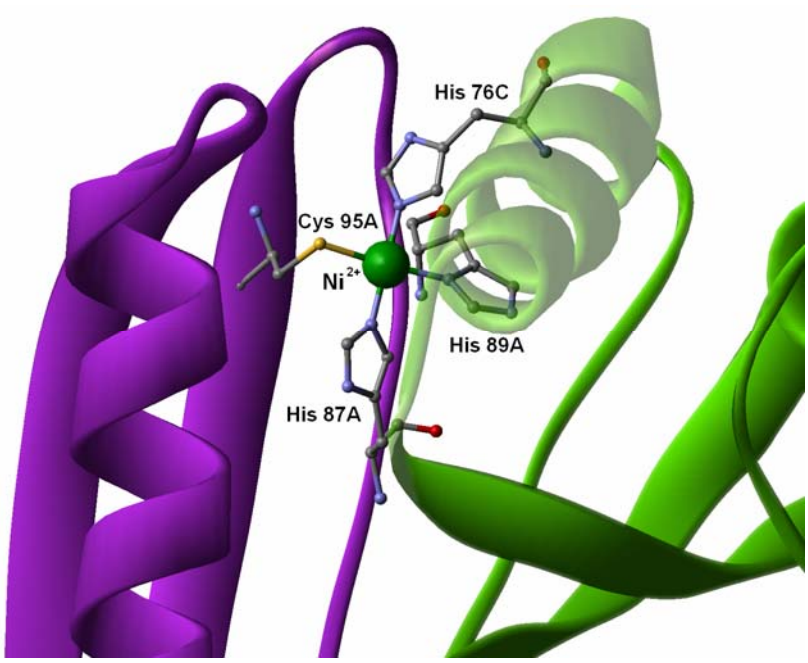


Figure 3.1-2 : Site de fixation du nickel dans le domaine C terminal de NikR.

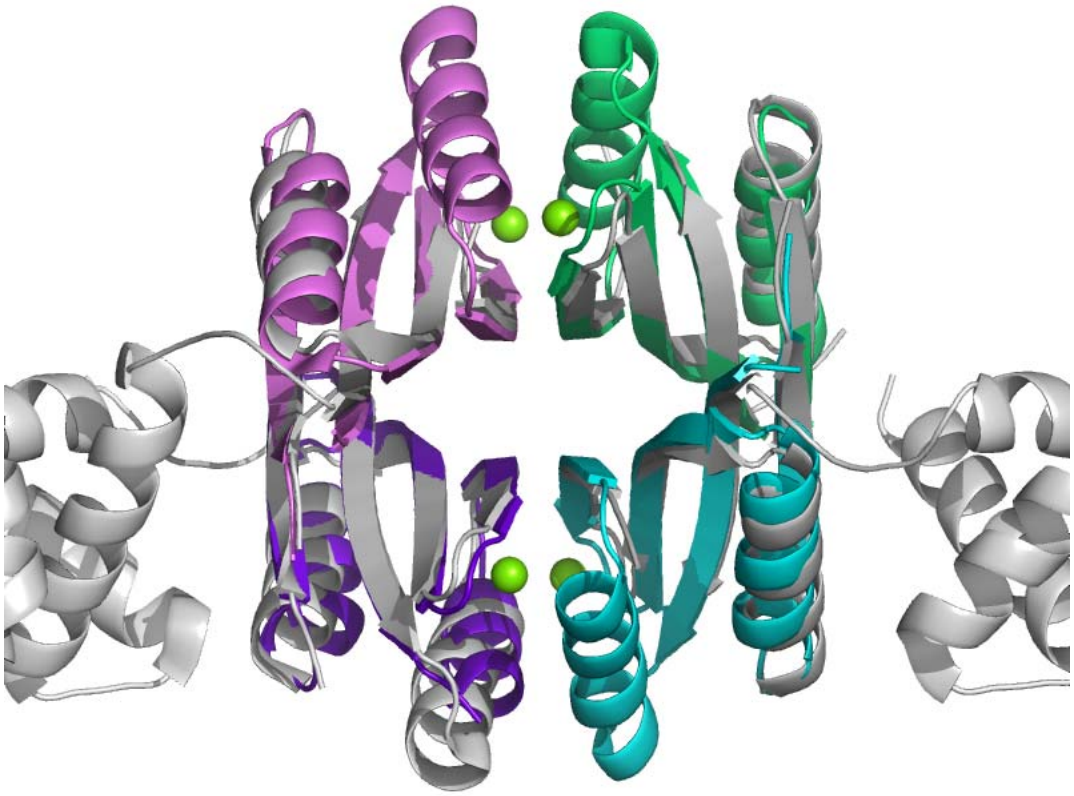


Figure 3.1-3 : Superposition de la structure de NikR entière apo (en gris ; code PDB : 1Q5V) et du domaine C terminal fixant le nickel (de violet à vert ; code PDB : 1Q5Y). En vert les atomes de nickel.

Les structures de la protéine entière apo et du domaine C terminal fixant le nickel ne présentent que peu de différences (rmsd entre les Cα = 1,6 Å ; Figure 3.1-3), néanmoins il est à en noter deux principales :

- Les brins contenant les résidus His87, His89 et Cys95, qui sont tous les trois constitutifs du site de fixation, présentent un mouvement de rapprochement en direction du nickel ;
- L'hélice α_3 contient le quatrième ligand du nickel : His76. Cette hélice est extrêmement désordonnée dans la forme apo, puisqu'elle n'est visible que dans le cas d'un des monomères avec un facteur d'agitation thermique B élevé.

Ces observations suggèrent que le site de fixation ne doit pas être pré-structuré.

3.1.3 La fixation à l'ADN

Le domaine N terminal, qui constitue le site de fixation du nickel, a un repliement très similaire aux répresseurs Mnt, Arc, MetJ, CopG et ω. Cette homologie, ainsi que des études de mutagénèse, suggèrent une interaction du feuillet β antiparallèle constitué par deux monomères, avec le grand sillon des deux séquences ADN répétées de l'opérateur de l'opéron *nikABCDE*. La fixation à l'ADN impliquerait également une des hélices du domaine, qui interagirait ainsi avec les phosphates de l'ADN.

Les répresseurs de la famille brin–hélice–hélice se fixent tous à l'ADN au niveau de séquences répétées séparées au maximum par un tour d'ADN. Seule NikR présente la particularité de posséder deux séquences répétées séparées par deux tours d'ADN (soit environ 80 Å). Cette distance est obtenue grâce à l'existence du domaine C terminal (absent chez les autres membres des répresseurs de la famille brin–hélice-hélice) : les deux feuillets β du domaine N terminal étant séparés d'environ 90 Å.

Par contre, les deux domaines de fixation à l'ADN sont, dans cette structure, dans une conformation rendant impossible leur interaction simultanée avec les deux séquences répétées d'ADN. De ce fait, Schreiter *et al.* proposent que la fixation du nickel puisse induire un changement de la protéine mettant les domaines de fixation dans une conformation plus favorable à l'interaction avec l'ADN. Il consisterait en une rotation du domaine N terminal par rapport au domaine C terminal grâce à la flexibilité de la région charnière (Figure 3.2-1 et Figure 3.2-2).

3.2 La structure de NikR de *Pyrococcus horikoshii*

Deux ans après la publication de la première structure de NikR, Chivers et al. publièrent la structure de NikR de *Pyrococcus horikoshii* (Chivers and Tahirov, 2005) qu'ils résolvaient par remplacement moléculaire. NikR de *P. horikoshii* présente 33,8% d'identité de séquence avec son homologue d'*E. coli*. Cinq structures furent alors publiées correspondant à trois conformation différentes du tétramère de NikR.

3.2.1 Les conformations des structures résolues

Les trois conformations de NikR publiées sont présentées dans la Figure 3.2-1 :

- La forme apo de la protéine correspond à une conformation dite ouverte. Sa structure est très similaire à celle de *E. coli* décrite précédemment (Figure 3.2-1a) ;
- Dans le cas de l'occupation des sites de haute affinité par le nickel, la protéine adopte la même conformation que dans le cas de la forme apo (Figure 3.2-1b) ;
- Par contre, dans le cas où les sites de haute et de faible affinité (chapitre 2.3.1.2 de l'introduction) sont occupés par le nickel, NikR adopte une conformation dite fermée–trans. Celle-ci correspond à une rotation du domaine N terminal (fixant l'ADN) de presque 90° par rapport à sa position dans la forme ouverte (Figure 3.2-1c).

3.2.2 Les sites de fixation du nickel de haute affinité

Le site de fixation de haute affinité, décrit précédemment, est parfaitement conservé chez *P. horikoshii*.

Par contre, un second site, qui peut aussi être considéré de haute affinité, existe également. Il se trouve à l'interface du tétramère (Figure 3.2-1b), et le nickel y adopte une géométrie pyramidale trigonale déformée, ressemblant à une géométrique octaédrique. Les ligands du nickel sont fournis par les chaînes latérales des résidus His64 et Asp65 (les deux oxygènes servant de ligand au nickel) d'un monomère, et le résidu Asp75 d'un second monomère. Seule l'histidine 64 est conservée dans la séquence d'*E. coli*, ce qui peut expliquer l'absence de ce site dans sa structure.

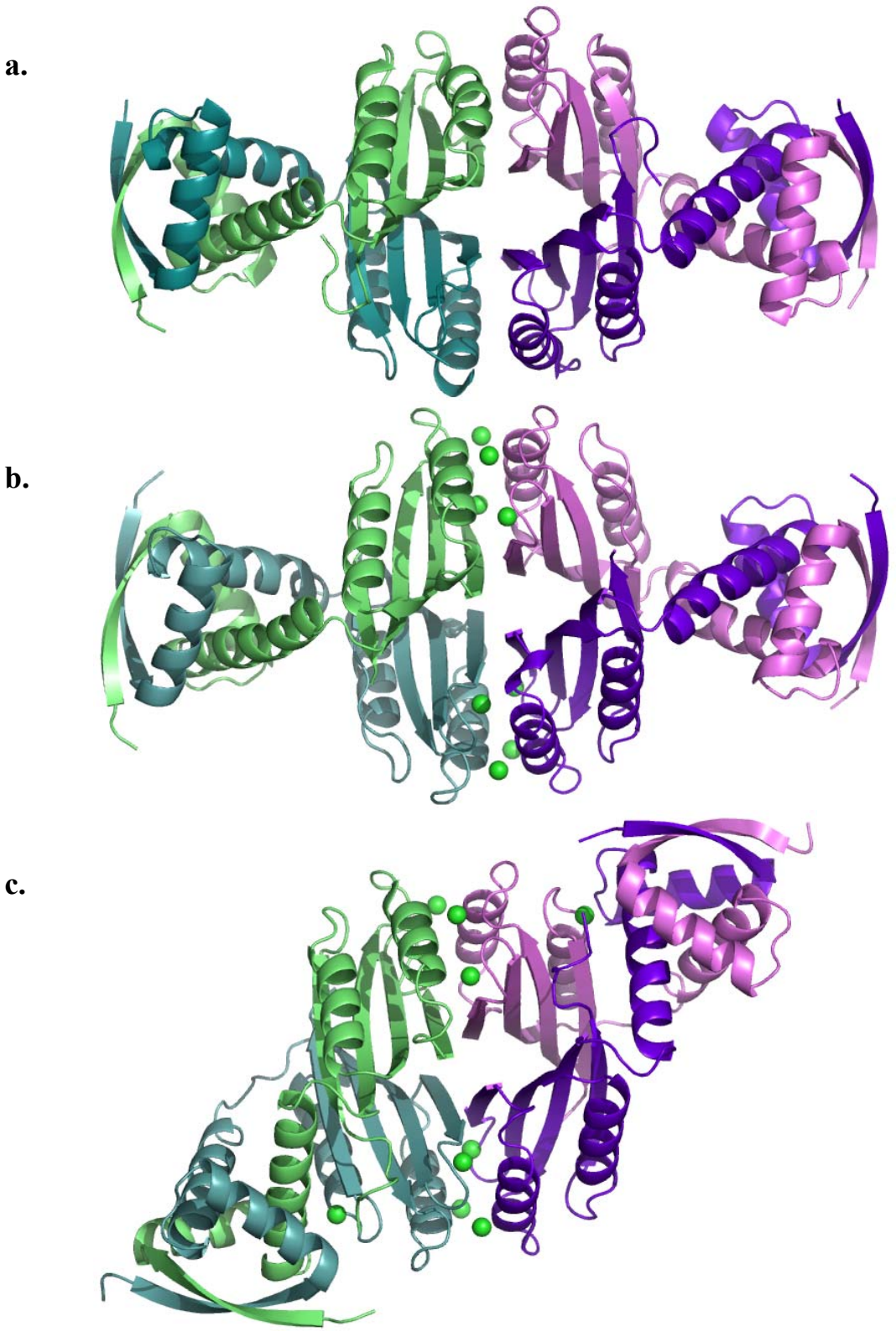


Figure 3.2-1 : Structures des tétramères (de violet à vert) de NikR de *P. horikoshii*. Le nickel est représenté par des sphères vertes.

a. forme ouverte apo (code PDB : 2BJ3) ;

b. forme ouverte contenant un nickel par monomère (code PDB : 2BJ1) ;

c. forme fermée-trans contenant deux nickels par monomère (code PDB : 2BJ9).

3.2.3 Les sites de fixation du nickel de faible affinité

Un site de faible affinité impliqué dans le processus de fixation à l'ADN avait depuis longtemps été mis en évidence (Chivers and Sauer, 2000; Chivers and Sauer, 2002; Bloom and Zamble, 2004). Dans le cas de *P. horikoshii*, il se situe à l'interface entre le domaine de fixation à l'ADN et celui fixant le nickel, provoquant l'adoption par NikR d'une conformation fermée (Figure 3.2-1c). Ce site correspond à une coordination incomplète du nickel, ce qui expliquerait sa faible affinité, assurée principalement par les chaînes latérales des résidus conservés Glu32 et Asp36.

Un second site de faible affinité a été identifié à la surface de NikR, proche (3,5 Å) du résidu Glu53. Ce dernier n'est pas conservé dans la séquence de NikR d'*E. coli*.

3.2.4 Le modèle de la régulation par le nickel de la fixation du tétramère de NikR à l'ADN

La possibilité d'envisager l'existence d'une conformation fermée–cis, a amené Chivers *et al.* à proposer un modèle quant à la régulation et la fixation du tétramère de NikR à l'ADN (Figure 3.2-2) qui implique que seule la forme fermée–cis peut se fixer aux deux séquences répétées de l'opérateur de l'opéron *nikABCDE*.

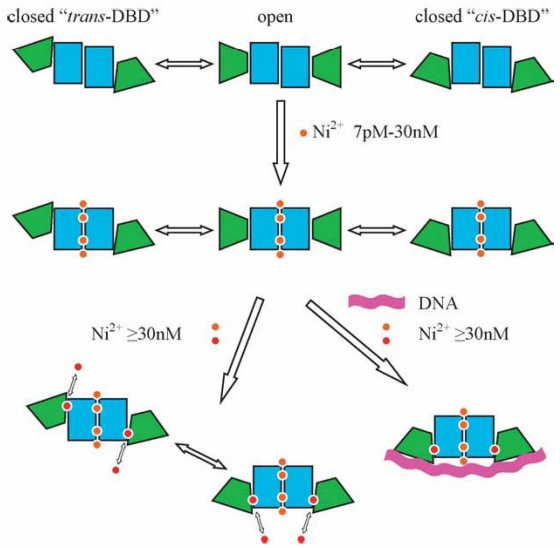


Figure 3.2-2 : Schéma illustrant les différentes conformations adoptées par le tétramère de NikR de *P. horikoshii*, en fonction de la concentration de nickel et de la présence ou non de l'ADN. La première étape consiste en l'occupation des sites de forte affinité. La seconde représentée consiste en l'occupation des sites de faible affinité en présence ou absence des deux séquences répétées de la région opératrice de l'opéron *nikABCDE (Chivers and Tahirov, 2005)*.

Dans un premier temps, l'occupation du site de haute affinité stabilise le tétramère de NikR, et permet la coexistence des formes ouvertes, fermées trans et cis. Dans le cas de la présence d'une quantité suffisante de nickel pour occuper également le site de faible affinité, seules les formes fermées cis et trans coexistent. L'addition d'ADN permet la fixation du tétramère de NikR au niveau des deux séquences répétées uniquement sous la forme fermée cis.

4 Le phasage de nos données de diffraction par remplacement moléculaire

Une fois ces résultats publiés, nous avons souhaité les utiliser pour résoudre la structure de NikR sur laquelle nous travaillions au laboratoire par remplacement moléculaire, et dont nous avions enregistré des données de diffraction. L'intérêt était de vérifier si le mutant sur lequel nous travaillions adoptait une conformation similaire à celles publiées dans le cas de la protéine sauvage. Pour ce faire nous avons utilisé les données de diffraction du cristal congelé dans le propane liquide décrit précédemment (chapitre 2.2.2).

4.1 Le remplacement moléculaire à partir de NikR d'*Escherichia coli*

Nous avons tout d'abord tenté le remplacement moléculaire avec la structure de la protéine NikR entière (apo) publiée par Schreiter *et al*. comme modèle (Schreiter *et al.*, 2003). Le calcul du coefficient de Matthews (chapitre 2.3.5.4 du matériel et méthodes) prédit que l'unité asymétrique des cristaux de NikR contient entre 6 et 8 molécules. Le fait que NikR cristallise sous la forme d'un tétramère, permet de supposer que notre cas présente 2 tétramères par unité asymétrique.

Les essais de remplacement moléculaire se sont faits à l'aide du programme Phaser. Dans un premier temps, nous avons essayé de placer dans notre maille entre un et huit monomères. De cette façon, quatre molécules ont pu être orientées, ce qui est insuffisant au vu des résultats donnés par le calcul du coefficient de Matthews.

Nous avons donc, dans un deuxième temps, réalisé des essais de remplacement moléculaire avec les domaines N et C terminaux séparément. De cette façon il fut facile de placer les huit domaines C terminaux dont la structure se révèle être très similaire à celle déjà publiée, mais malheureusement les domaines N terminaux sont restés invisibles.

Nous avons pu par la suite placer quelques hélices, mais la majeure partie du domaine n'est pas résolue (Figure 4.1-1). Un affinement par Refmac donne les statistiques suivantes : R = 29,5% ; R_{free} = 46,2%. Nous n'avons pas réussi à obtenir de meilleurs résultats quant à la résolution de cette structure.

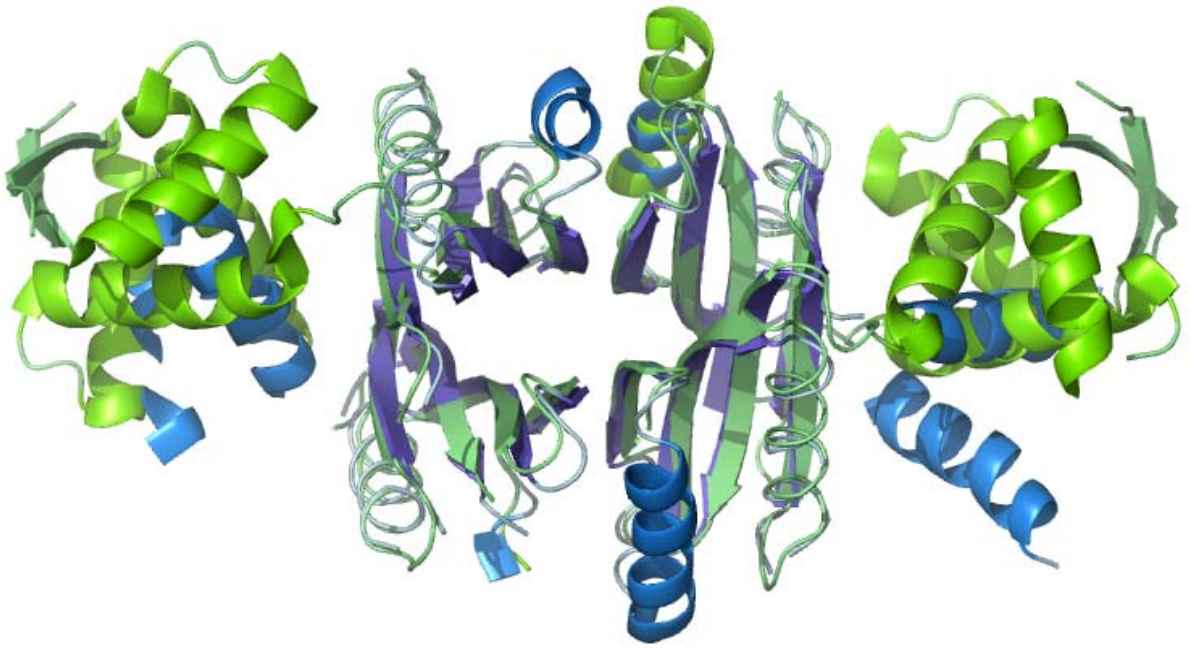


Figure 4.1-1 : Superposition de la structure de NikR apo publiée (vert ; code PDB : 1Q5V) et de la structure incomplète que nous avons résolue (bleu).

4.2 Le remplacement moléculaire à partir de NikR de *Pyrococcus horikoshii*

A la suite de la publication des structures de NikR de *P. horikoshii (Chivers and Tahirov, 2005)*, nous avons tenté d'améliorer notre modèle en utilisant un remplacement moléculaire avec nos données de diffraction. Nous l'avons tenté dans chaque cas (les cinq structures publiées) avec le tétramère, le dimère et le monomère de NikR, mais aucun n'a donné de résultats exploitables.

En définitive nous avons réussi à caractériser uniquement le domaine C terminal du tétramère de NikR. Le domaine de fixation à l'ADN (N terminal) doit être désordonné dans nos cristaux, ce qui explique les difficultés que nous avons à le localiser, ainsi que la faible résolution de nos données de diffraction. Il est possible que ceci vienne de la mutation que notre protéine contient (Q2E), puisqu'elle implique le deuxième acide aminé du brin β du site

de fixation à l'ADN. Le dimère peut alors être déstabilisé, et de cette façon la structure du domaine N terminal présente une agitation empêchant sa visualisation en cristallographie des protéines. On peut aussi imaginer que l'agitation provient de la possibilité que présente la structure du tétramère de NikR d'adopter différentes conformations : ouverte, fermée cis et trans.

ETUDE STRUCTURALE

DE LA PROTEINE

NIKA

Le travail principal de cette thèse a été l'étude de la structure de la protéine NikA par cristallographie des protéines, et notamment l'étude de son site de fixation du nickel.

1 L'étude de la protéine NikA, purifiée selon le protocole contenant de l'EDTA

Comme nous l'avons vu précédemment, des premières études visant à déterminer la structure de NikA avaient été menées au L.C.C.P. (Charon *et al.*, 1994). A l'époque, la protéine était fournie par l'équipe de L.F. Wu, et avait permis d'obtenir deux formes cristallines différentes : orthorhombique et hexagonale. La structure ne put être résolue ni par remplacement moléculaire, ni par addition d'atomes lourds.

Nous avons alors voulu reproduire ces cristaux et résoudre la structure par l'utilisation de la protéine séléniée (paragraphe 1.1.2.2 et1.2.2.2 du matériel et méthodes).

1.1 La surexpression de la protéine NikA

1.1.1 La protéine NikA surexprimée en milieu minimum

Bien que l'objectif était de cristalliser la protéine NikA séléniée, nous avons tout de même produit, dans un premier temps, la protéine native. En effet, tous les tests de surexpression en milieu minimum NYM, purification et cristallisation, ont été d'abord effectués sur cette dernière, du fait du coût élevé de la séléno-méthionine rendant rédhibitoire ce genre de mise au point directement sur la protéine séléniée.

1.1.1.1 La protéine NikA native

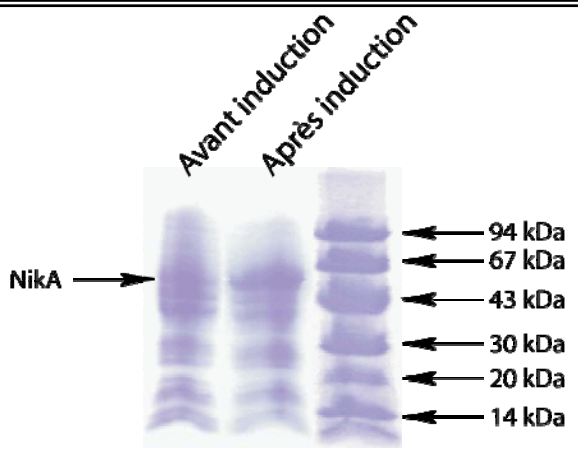


Figure 1.1-1 : Gel SDS - PAGE des protéines périplasmiques bactériennes avant et après la surexpression de la protéine NikA native.

Afin d'obtenir une quantité suffisante de bactéries, le protocole de surexpression a été répété de nombreuses fois. Les résultats de chaque surexpression furent comparables : on récupère environ 13 g de bactéries pour 2 L de culture en milieu minimum (Figure 1.1-1).

1.1.1.2 La protéine NikA séléniée

La surexpression de la protéine séléniée a été réalisée en milieu minimum NYM complémenté en séléno-méthionine. Mais par rapport à la protéine native, nous nous attendions à obtenir une quantité moindre de bactéries, et donc de protéine, à cause de la toxicité du sélénium. Nous avons donc été obligés de réaliser une culture plus importante, afin d'obtenir une quantité suffisante de NikA séléniée pour être purifiée. Les résultats sont à chaque fois comparables : on récupère environ 5 g de bactéries pour 2 L de culture. De la même façon que pour la protéine native, des témoins ont été réalisés à partir des protéines périplasmiques (Figure 1.1-2).

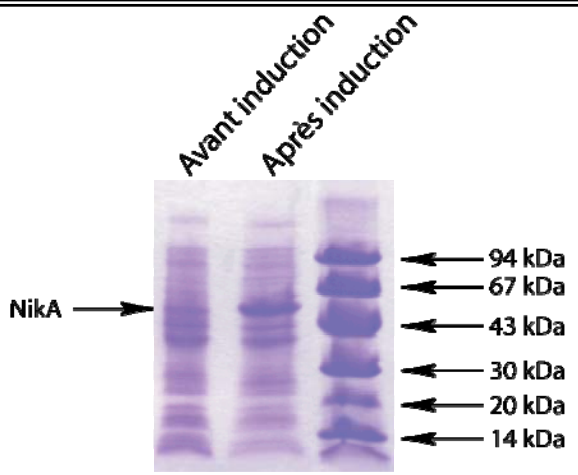


Figure 1.1-2 : SDS - PAGE des protéines périplasmiques bactériennes avant et après la surexpression de la protéine NikA séléniée.

1.1.2 La surexpression en milieu riche

La production de la protéine en milieu minimum s'est faite au début de notre étude de NikA. Mais par la suite, nous avons dû à nouveau produire la protéine. Nous obtenons alors de l'ordre de 6 g de bactéries pour 2 L de culture en milieu LB ou TB en moyenne : la croissance bactérienne est ralentie en milieu riche après induction de la surexpression.

Par contre on constate une amélioration de la surexpression en TB par rapport au LB. De ce fait, nous avons utilisé le TB comme milieu de culture lors de la surexpression de NikA pour les productions suivantes.

1.2 La purification de la protéine NikA

1.2.1 La purification de la protéine NikA surexprimée en milieu minimum

Le protocole de purification de NikA a été modifié par rapport à celui mis au point par L.F. Wu puisque ce dernier présentait un rendement trop faible (de l'ordre de 7 mg de protéine pour 100 g de bactéries) pour pouvoir envisager la purification de la protéine séléniée. Le même protocole de purification est employé dans le cas des protéines natives et séléniées. Dans les deux, cas la protéine présente le même comportement vis à vis des étapes de la purification.

1.2.1.1 La purification de la protéine NikA native

La mise au point du protocole de purification ayant été réalisée par Lydie Martin et Christine Cavazza du laboratoire, je ne décrirai en détail ici que la purification de la protéine séléniée dans laquelle j'ai été impliqué.

La purification de la protéine NikA native s'est faite à partir de 26,6 g de bactéries ayant surexprimé la protéine d'intérêt. Nous avons récupéré en fin de manipulation 650 µL d'une solution à 11,25 mg/mL (soit 7,3 mg) ce qui correspond à un rendement d'environ 27 mg de protéine pour 100 g de bactéries. Ce rendement peut être plus élevé, car nous avons des problèmes de pertes de protéine lors de la purification, qui ont été réglés par la suite.
Il y a donc bien eu amélioration du rendement de la purification, par rapport à l'ancien protocole, et on peut ainsi se lancer dans la purification de la protéine séléniée.

1.2.1.2 La purification de la protéine NikA séléniée

Dans le cas de la protéine séléniée, on utilise 20,6 g de bactéries ayant surexprimé la protéine NikA séléniée.

1.2.1.2.1 La colonne Q – Sépharose

Après extraction à l'EDTA des protéines périplasmiques et précipitation au sulfate d'ammonium, les protéines sont chargées sur une colonne Q – Sépharose (Figure 1.2-1). NikA est alors éluée entre 5 et 8% de NaCl 1 M, soit entre 5 et 8 mM de NaCl.
Ces premières étapes de purification ont ainsi permis d'obtenir un rendement de 251 mg de protéine pour 100 g de bactéries.

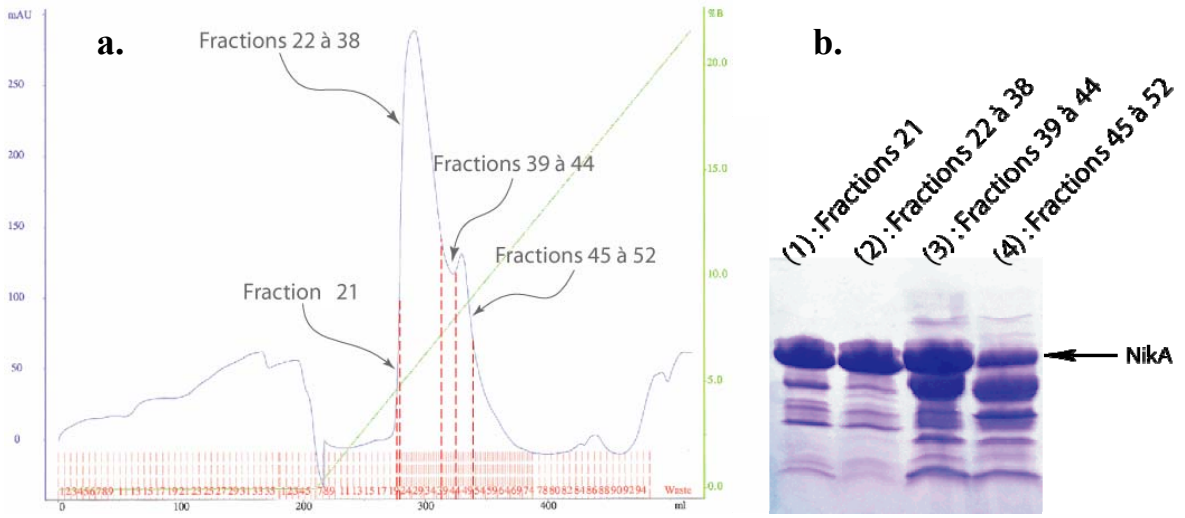


Figure 1.2-1 : Etape de purification sur colonne Q - Sépharose.

a. Profil d'élution de la colonne Q - Sépharose avec un gradient continu de NaCl 1 M en 10 volumes de colonne.

En bleu : la D.O. à 280 nm. En vert : le gradient de NaCl.

b. Gel SDS - PAGE des fractions purifiées sur colonne Q – Sépharose. Les fractions 22 à 38 sont gardées pour la suite de la purification.

1.2.1.2.2 La colonne Resource – Q

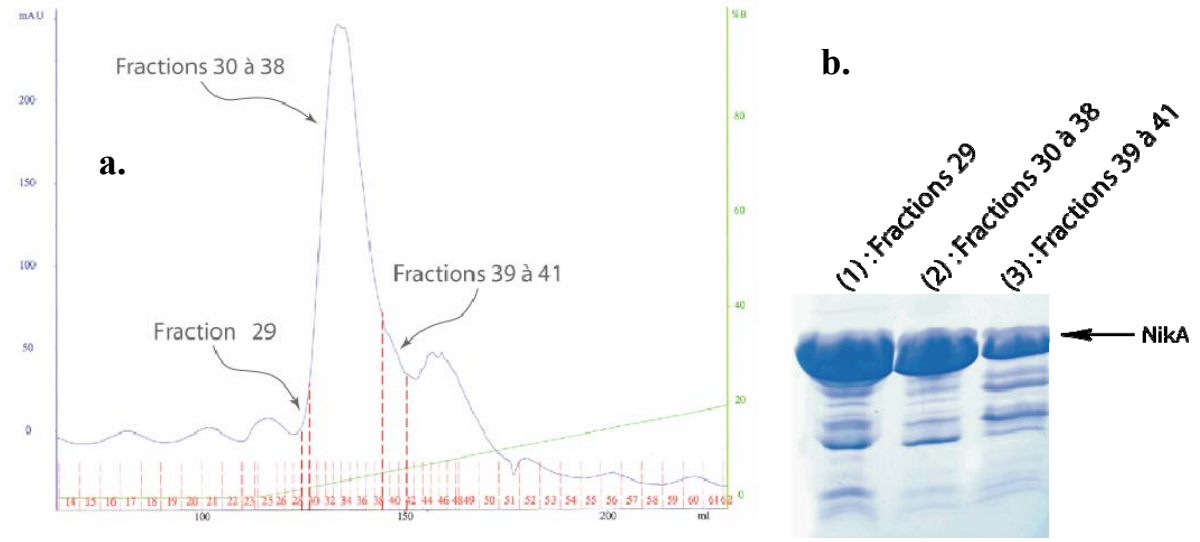


Figure 1.2-2 : Etape de purification sur colonne Resource - Q.

a. Profil d'élution de la colonne Resource - Q avec un gradient continu de NaCl 1 M en 25 volumes de colonne.

En bleu : la D.O. à 280 nm. En vert : le gradient de NaCl.

b. Gel SDS - PAGE des fractions purifiées sur colonne Resource – Q. Les fractions 30 à 38 sont gardées.

Après dialyse, les protéines sont chargées sur une colonne Resource – Q (Figure 1.2-2). NikA est alors éluée entre 3 et 6% de NaCl 1 M, soit entre 3 et 6 mM de NaCl.

1.2.1.2.3 Le rendement de purification

Après concentration de la protéine, on récupère 1,7 mL de solution de NikA séléniée à 11 mg/mL, soit 18,7 mg de solution de la protéine suffisamment pure pour l'expérience de cristallisation (Figure 1.2-2b). Le rendement de la purification est donc de 90,8 mg de protéine pour 100 g de bactéries.

Cette purification a relativement mieux fonctionné que celle de la protéine native : le sélénium n'a donc que peu, ou pas, affecté la quantité de protéine surexprimée par bactérie et la purification semble, cette fois-ci, s'être mieux passée, ce qui a permis de fortement limiter les pertes de protéine.

1.2.2 La protéine NikA surexprimée en milieu riche

Dans le cas de la protéine surexprimé en milieu riche, le protocole de purification est le même que celui décrit précédemment. De plus, la protéine se comporte de façon similaire à la protéine surexprimée en milieu minimum, vis-à-vis des différentes étapes de la purification. Le rendement de la purification est en moyenne de 100 mg de NikA pour 100 g de bactéries, ce qui est du même ordre de grandeur que celui de la purification de la protéine séléniée.

1.3 La cristallisation de la protéine NikA

1.3.1 Les tentatives de reproduction des cristaux de la protéine native

Une fois la protéine NikA produite, il fallait obtenir des cristaux exploitables pour l'analyse aux rayons X. Pour cela, nous avons testé sur la protéine native (surexprimée en milieu minimum) nouvellement purifiée les conditions de cristallisation de NikA obtenues par Charon *et al.*(Charon *et al.*, 1994) :

[NikA] = 11 mg / mL ;

Sulfate d'ammonium entre 1,2 et 1,7 M ;

Acétate de sodium 0,1 M pH 4,7 ou 4,9 ;

+/- DTT 1 mM ou +/- chlorure de nickel 500 µM (2,5 équivalents Ni / NikA) ;

20°C, en technique de goutte suspendue.

Malheureusement aucun cristal n'est apparu pour ces conditions, ce qui est sûrement à relier au changement du protocole de purification.

Nous avons donc tenté de cristalliser NikA en testant de nouveaux paramètres, ceci en suivant une logique de recherche systématique :

- Réalisation des screens de cristallisation type *Jancarik* et *Screen II* (Hampton Research) ;
- Ajout d'additifs tels que : 200 µM chlorure de cuivre ; 1 mM chlorure de nickel (5 équivalents Ni / NikA) ; EDTA 1 mM ; 5% glycérol ; 5 mM DTT…
- Variation du pH du tampon (acétate de sodium) entre 4,5 et 5,1 ;
- Changement de la nature du tampon et du pH : MES 0,1 M pH 6 ; HEPES 0,1 M pH 7,5 ; Tris base 0,1 M pH 8,5 ;
- Variation du rapport volume de protéine / volume de solution réservoir de la goutte, permettant de faire varier la concentration de NikA afin de ralentir ou d'accélérer la vitesse d'équilibration de la goutte de cristallisation. De façon classique, les gouttes sont réalisées à partir de 1 µL (ou 2 µL) de solution de protéine et 1 µL (ou 2 µL) de solution réservoir.

Mais aucun des nombreux essais de conditions de cristallisation n'ont permis d'obtenir de cristaux exploitables. Par contre, nous observons dans de nombreuses gouttes la présence d'un précipité cristallin indiquant que l'on se trouve dans la zone de précipitation et qu'il faut contrôler la nucléation pour qu'un cristal se développe.

1.3.2 La reproduction des cristaux par micro-ensemencement

1.3.2.1 La cristallisation de la protéine native

Pour tenter de s'affranchir du problème que nous pose la reproduction des cristaux de NikA, nous avons repris un dernier échantillon de protéine purifiée par L.F. Wu, et donc préparée par l'ancien protocole de purification, gardé à – 20°C depuis 1994. Nous nous en sommes alors servi afin de reproduire les cristaux obtenus à l'époque au laboratoire.

Contrairement à ce qui était observé lors de la première cristallisation, on n'obtient plus qu'une seule des deux formes cristallines, celle présentant le faciès des cristaux appartenant au système orthorhombique. Nous nous en sommes alors servi pour cristalliser la protéine nouvellement purifiée grâce à des micro-ensemencements successifs : un germe provenant d'un cristal de protéine de l'ancienne purification est placé dans une goutte pré-équilibrée

contenant de la protéine NikA issue de la nouvelle purification. Cette manipulation a permis de réobtenir des cristaux. La deuxième étape fut alors d'amener un germe de la nouvelle purification dans une goutte pré-équilibrée contenant de la protéine issue de la nouvelle purification. Là aussi, nous avons obtenu des cristaux assez bien formés (Figure 1.3-1), poussant en trois ou quatre jours, dans des conditions de cristallisation constituées de :

[NikA] = 11 mg / mL ;

Sulfate d'ammonium entre 1,3 et 1,5 M ;

Acétate de sodium 0,1 M pH 4,7 ;

+/- $NiCl_2$ 200 µM ou 500 µM (soit 1 et 2,5 équivalents Ni / NikA) ;

20°C en goutte suspendue.

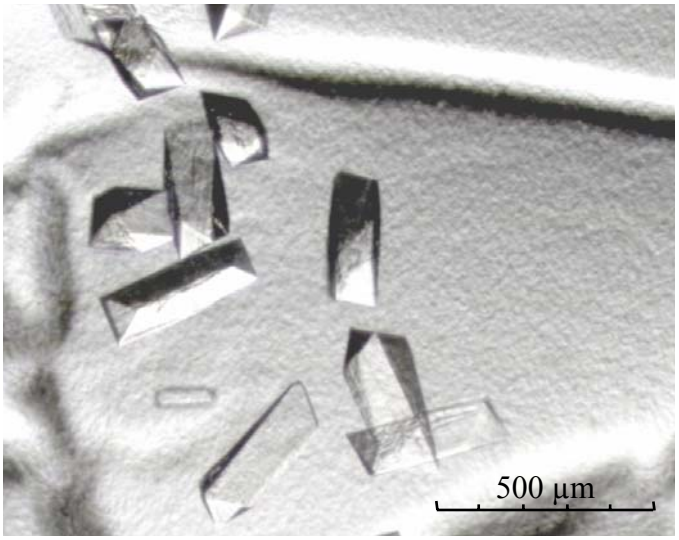


Figure 1.3-1 : Cristaux de la protéine NikA native réalisés par micro-ensemencement.

Afin d'améliorer la taille des cristaux, nous avons souhaité contrôler la nucléation. Pour cela, nous avons testé différentes solutions, et notamment, le lavage du poil de chien entre l'étape de prélèvement des germes cristallins, et l'ensemencement proprement dit. De cette façon on souhaite contrôler le nombre de germes que l'on apporte à la goutte de cristallisation.

La condition qui a alors donné les meilleurs résultats est le lavage du poil par une solution de sulfate d'ammonium 1,1 M et acétate de sodium 0,1 M pH 4,7.

1.3.2.2 La cristallisation de la protéine séléniée

Le même principe fut utilisé pour la réalisation des cristaux de NikA séléniée, le but n'étant pas, dans ce cas, de passer de la protéine purifiée par l'ancienne méthode à celle

obtenue grâce à la nouvelle, mais de passer de la protéine native à la protéine séléniée. Le protocole est identique, et l'on obtient des cristaux de bonne qualité (Figure 1.3-2) après environ une semaine, dans des conditions de cristallisation composées de :

[NikA] = 11 mg / mL ;

Sulfate d'ammonium 1,4 ou 1,5 M ;

Acétate de sodium 0,1 M pH 4,7 ;

DTT 5 mM ; +/- $NiCl_2$ 200 µM ou 500 µM (soit 1 et 2,5 équivalents Ni / NikA) ;

20°C en goutte suspendue.

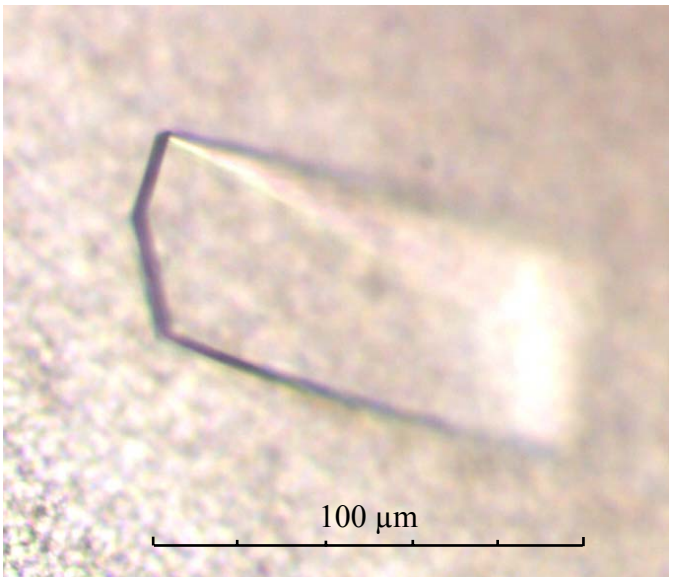


Figure 1.3-2 : Cristaux de la protéine NikA séléniée réalisés par micro-ensemencement.

Après deux ensemencements successifs, la qualité des cristaux semble diminuer, ce qui nous oblige alors à recommencer une nouvelle série d'ensemencements à partir de la protéine native. De plus, les cristaux semblent également s'abîmer lorsqu'ils restent plus de deux semaines dans la goutte de cristallisation. Ces deux limitations n'avaient cependant pas été observées dans le cas de la protéine native.

1.4 L'enregistrement et traitement des données de diffraction

La solution de cryoprotection utilisée pour NikA est ajoutée par un trempage rapide des cristaux, et se compose de sulfate d'ammonium 2 M, glycérol 25%, avec le même tampon et les mêmes additifs que la solution de cristallisation.

1.4.1 Le jeu de données de la protéine NikA native

Le meilleur jeu de données de diffraction de la protéine NikA native que nous ayons pu enregistrer est apparu dans la condition de cristallisation suivante :

NikA] = 11 mg / mL ;

Sulfate d'ammonium 1,4 M ;

Acétate de sodium 0,1 M pH 4,7 ;

20°C en goutte suspendue.

Le traitement des données de diffraction a permis d'obtenir des statistiques de bonne qualité (Tableau 1.4-1).

Ligne de lumière de l'ESRF	ID14eh4
Longueur d'onde	0,9998 Å
Groupe d'espace	$P2_12_12_1$
Paramètres de maille	a = 86,6 Å ; b = 94,0 Å ; c = 124,1 Å ; α = 90° ; β = 90° ; γ = 90°
Nombre de molécules par unité asymétrique	2
Résolution	30,0–2,2 Å (2,3–2,2 Å)
Complétude	97,4% (89,2%)
I/σ(I)	14,7 (5,8)
Nombre total de réflexions mesurées	239 144 (19 987)
Redondance	4,7 (3,5)
R_{SYM}	7,6% (28,4%)

Tableau 1.4-1 : Statistiques des données de diffraction d'un jeu de la protéine NikA native. Les valeurs entre parenthèses correspondent à la dernière tranche de résolution.

1.4.2 Le jeu de données MAD de la protéine NikA

1.4.2.1 L'enregistrement et indexation du jeu MAD

Lorsque les premiers cristaux de la protéine séléniée, de bonne qualité, ont été réalisés, nous avons enregistré, sur la ligne de lumière ID14eh4 de l'ESRF, un jeu MAD au seuil du sélénium sur un cristal d'environ 0,120 x 0,040 x 0,015 mm^3, ayant poussé dans :

[NikA] = 11 mg / mL ;

Sulfate d'ammonium 1,4 M ;

Acétate de sodium 0,1 M pH 4,7 ;

DTT 5 mM ;

20°C en goutte suspendue.

Pour cela, la première étape a été de vérifier que le sélénium était bien présent dans les cristaux, en enregistrant un spectre de fluorescence autour du seuil théorique d'absorption du sélénium. Le spectre alors obtenu (Figure 1.4-1) s'est révélé très proche d'un spectre théorique, ce qui nous a permis de déterminer les longueurs d'onde auxquelles le MAD a été réalisé :

- Au pic de fluorescence, qui correspond au maximum de $^{\lambda}f''$, soit 12,6569 KeV (0,97957 Å) ;
- A l'inflexion du pic du spectre de fluorescence, qui correspond au minimum de $^{\lambda}f'$, soit 12,6540 KeV (0,97997 Å) ;
- A une troisième longueur d'onde, éloignée des deux autres, qui correspond ici au maximum d'intensité de la ligne de lumière, soit 13,2 KeV (0,9392 Å).

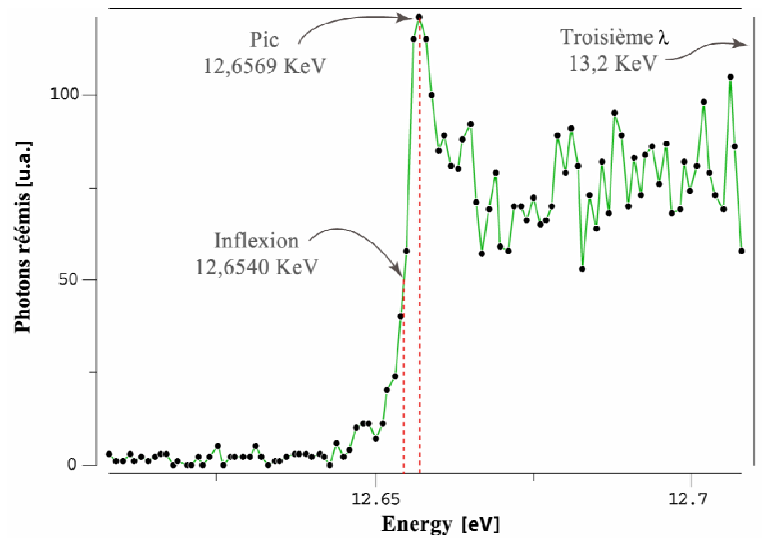


Figure 1.4-1 : Spectre de fluorescence d'un cristal de la protéine NikA séléniée.

Le cristal a été orienté afin d'obtenir, pour chaque longueur d'onde, une complétude proche de 100% en enregistrant le minimum d'images. Ceci avait pour objectif de conserver l'isomorphisme des trois jeux de données, en limitant au maximum la dégradation du cristal par dommages radiatifs.

Après traitement des données de diffraction, les statistiques se sont révélées correctes (Tableau 1.4-2), et l'on a constaté que, par rapport au jeu natif enregistré précédemment, un des axes 2_1 est devenu un axe 2, avec un changement de groupe d'espace de $P2_12_12_1$ à $P2_12_12$.

	maximum de $^{\lambda}f''$	minimum de $^{\lambda}f'$	troisième λ
Ligne de lumière de l'ESRF		ID14eh4	
Longueur d'onde	0,97957 Å	0,97997 Å	0,9392 Å
Groupe d'espace		$P2_12_12$	
Paramètres de maille	a = 70,0 Å ; b = 194,2 Å ; c = 75,9 Å ; α = 90° ; β = 90° ; γ = 90°		
Nombre de molécules par unité asymétrique		2	
Résolution	20,00–2,74 Å (2,75–2,74 Å)	20,00–2,81 Å (2,96–2,81 Å)	20,00–2,59 Å (2,75–2,59 Å)
Complétude	95,6% (89,4%)	96,2% (91,0%)	96,0% (90,6%)
I/σ(I)	7,9 (3,0)	8,5 (3,4)	9,2 (3,4)
Nombre total de réflexions mesurées	110 043 (1 083)	100 726 (15 376)	126 255 (19 300)
Redondance	2,2 (2,1)	2,1 (2,1)	2,1 (2,1)
R_{SYM}	9,4% (30,0%)	9,2% (29,6%)	7,4% (28,8%)

Tableau 1.4-2 : Statistiques des données de diffraction des trois longueurs d'onde d'un jeu MAD de la protéine NikA séléniée. Les valeurs entre parenthèses correspondent à la dernière tranche de résolution.

1.4.2.2 La localisation des atomes de sélénium

De la même façon que pour le jeu de données SAD enregistré sur la protéine NikR, dans le cas du MAD il nous faut également localiser les atomes de sélénium pour pouvoir résoudre la structure de NikA.

Après calcul de la carte de Patterson différence, les pics des sections de Harker se sont révélés trop faibles (en moyenne 5σ ; Figure 1.4-2) pour trouver les 10 séléno-méthionines

normalement présentes dans la séquence de NikA, ce qui a logiquement empêché la résolution de la structure de la protéine grâce à ce jeu de données.

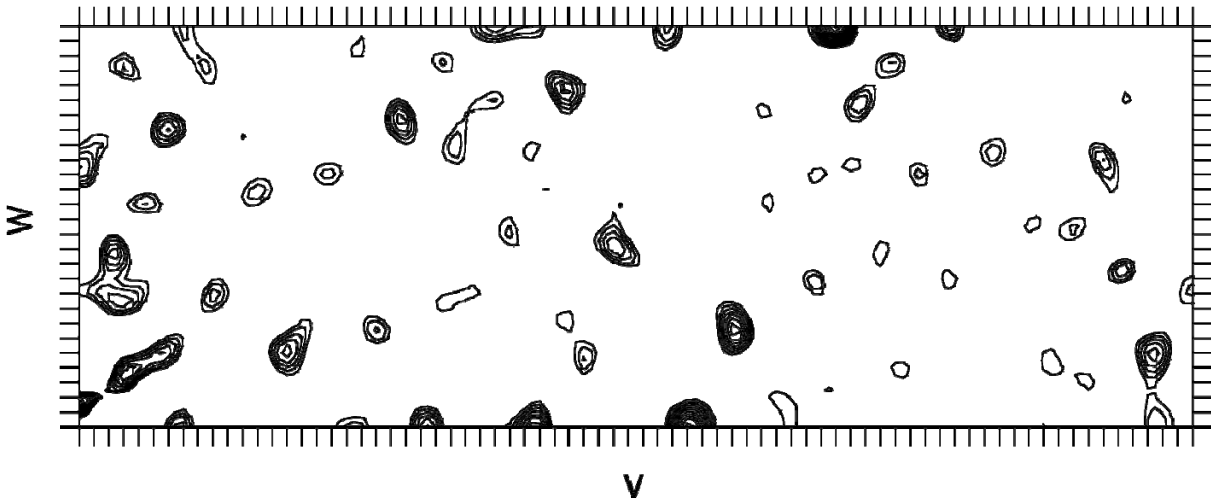


Figure 1.4-2 : Carte de Patterson différence dans sa section de Harker u = 0,5, calculée à partir du jeu MAD obtenu pour un cristal de NikA séléniée (sur la carte : base = 2σ ; contours = 0,5σ).

1.4.2.3 Des exemples de tentatives de phasage de la structure de NikA

Afin de résoudre la structure de NikA, nous avons essayé plusieurs techniques visant à introduire un diffuseur anomal dans les cristaux de NikA. Nous n'évoquerons ici que les deux méthodes les moins communes que nous ayons tentées.

Sharff *et al.* publient des résultats tendant à montrer que l'hétérogénéité de l'état d'oxydation des séléniums d'une protéine contenant des séléno-méthionines diminue de façon importante le signal anomal, jusqu'à le rendre insuffisant pour la localisation des atomes lourds, et donc le phasage de la structure de la protéine (Sharff *et al.*, 2000). Ils proposent alors d'oxyder les séléniums par un traitement au peroxyde d'hydrogène, et ainsi d'en augmenter le pouvoir de diffusion anomale.

Nous avons alors tenté d'oxyder nos cristaux de NikA séléniée en ajoutant progressivement du peroxyde d'hydrogène à la goutte de cristallisation. Dès la concentration finale atteinte (0,1 et 0,01% de H_2O_2), les cristaux sont congelés.

Le traitement a, dans la plupart des cas, réduit le pouvoir de diffraction des cristaux, mais il fut tout de même possible d'enregistrer des jeux de données SAD. Mais là encore, la diffusion anomale est insuffisante pour localiser les séléniums.

Nous avons également tenté d'introduire des atomes de xénon dans la structure de NikA. Pour ce faire, les cristaux ont été soumis à une pression de 10 bars de xénon pendant 10 secondes, grâce à la cellule haute pression du laboratoire. Cette technique a été mise au point

par Xavier Vernède (Vernede and Fontecilla-Camps, 1999), et elle permet de placer un cristal, préalablement pêché dans sa boucle, dans une enceinte rendue étanche par des joins toriques. Après le traitement, le cristal est immédiatement congelé.
Malheureusement le xénon n'a pas dû se fixer de façon ordonnée dans la structure de NikA, puisque il ne nous a pas été possible de localiser leur position grâce à sa diffusion anomale.

1.5 Les structures de NikA publiées par Heddle *et al.*

De la même façon que dans le cas de NikR, alors que nous cherchions un moyen de phaser la structure de NikA d'*E. coli*, Heddle *et al.* en publièrent la forme apo et la forme liée au nickel (Heddle *et al.*, 2003). Dans leur cas, la protéine est clonée dans un plasmide pET28b sans sa séquence signal d'exportation dans le périplasme. De cette façon, ils réalisent la purification de la protéine à partir des protéines cytoplasmiques. La structure fut résolue par la méthode MAD réalisée sur la protéine séléniée.
La structure de NikA apo a été obtenue grâce à un cristal appartenant au système trigonal ($P3_2$) ayant diffracté jusqu'à 1,85 Å de résolution, apparu dans les conditions suivantes :

[NikA] = 7,5 mg / mL ;
PEG 4 000 12,5% ;
Sulfate d'ammonium 0,3 M ;
Acétate de sodium 50 mM pH 5.0 ;
20°C en goutte suspendue.

Dans le cas de la protéine ayant fixé le nickel, la structure a été obtenue à partir d'un cristal orthorhombique ($P2_12_12$) ayant diffracté jusqu'à 1,95 Å de résolution, apparu dans les conditions suivantes :

[NikA] = 15 mg / mL ;
PEG 2 000 20% ;
Sulfate d'ammonium 0,3 M ;
Acétate de sodium 100 mM pH 5.5 ;
Chlorure de nickel 100 mM (375 équivalents Ni / NikA) ;
20°C en goutte suspendue.

Les paramètres de maille sont équivalents à ce que nous avons obtenus dans le cas de la protéine séléniée.
Dans les deux cas, leur protéine se trouve dans un tampon contenant un excès, d'un facteur 10, de nickel.

1.5.1 La description de la structure de NikA

La protéine NikA présente une structure composée de deux domaines à l'interface desquels se localise le site de fixation du nickel (Figure 1.5-1) :

- Le domaine I est composé des résidus 4 à 245 et 471 à 499 ;
- Le domaine II est composé des résidus 246 à 470.

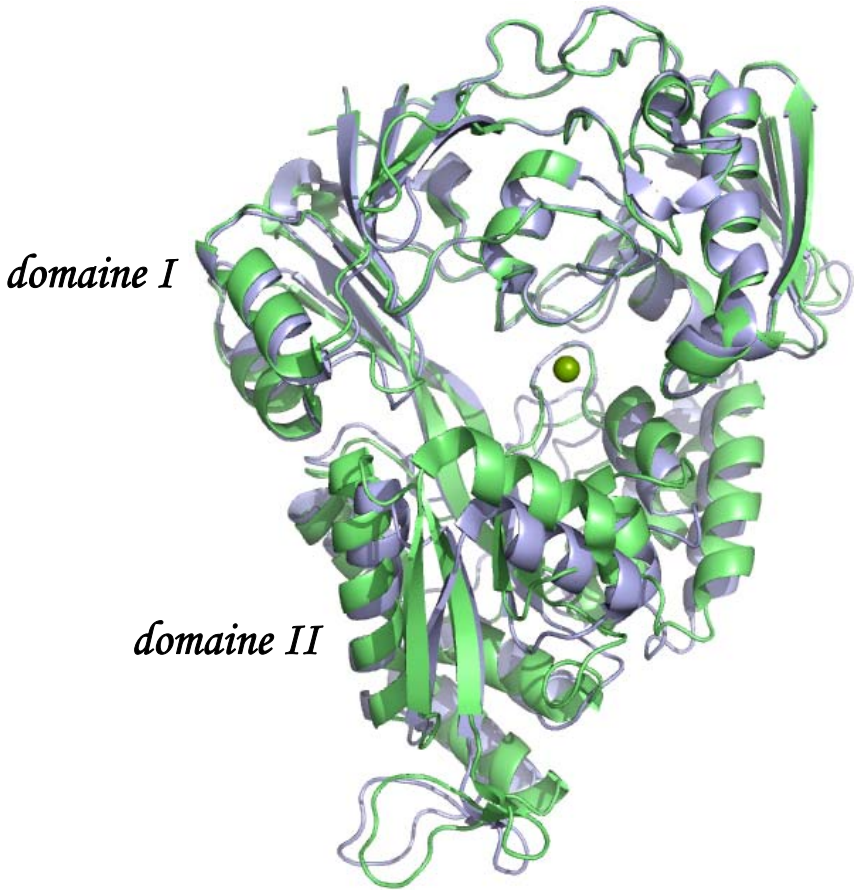


Figure 1.5-1 : Superposition des structures de NikA apo (en bleu ; code PDB : 1UIU) et NikA ayant fixé le nickel (en vert ; code PDB : 1UIV). Le nickel est représenté par une sphère verte.

La seule différence significative entre ces deux structures est la rotation d'environ 15° de l'un des deux domaines par rapport à l'autre, provoquant une fermeture de la protéine sur le nickel. De ce fait, la forme apo peut être qualifiée de forme ouverte, et celle ayant fixé le nickel de forme fermée (Figure 1.5-2).

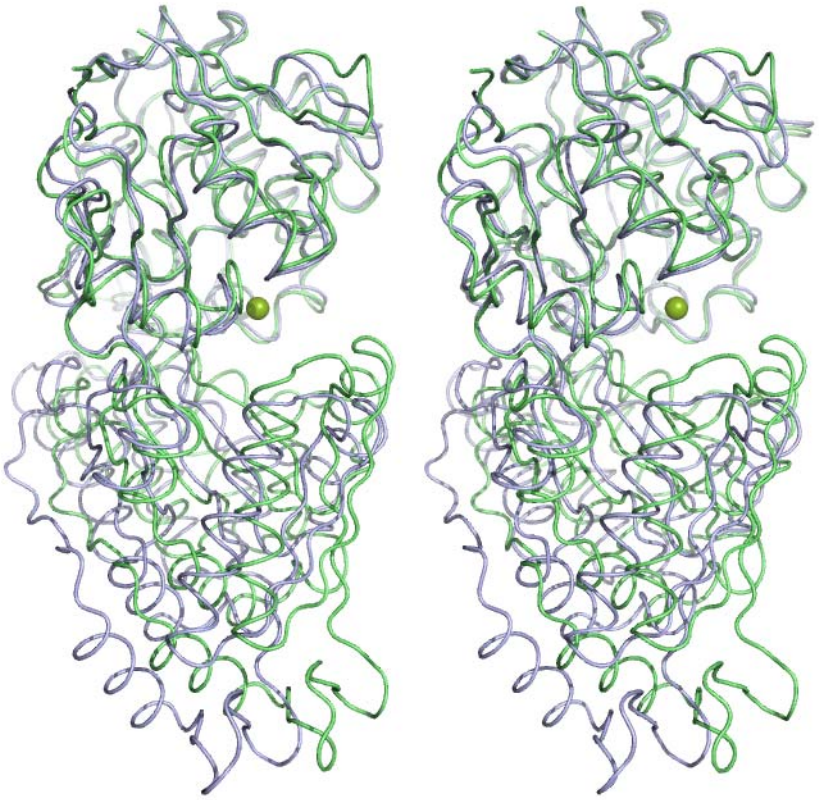


Figure 1.5-2 : Vue stéréoscopique de la superposition des formes ouverte (en bleu ; code PDB : 1UIU) et fermée de NikA (en vert ; code PDB : 1UIV). Le nickel est représenté par une sphère verte

1.5.2 Le site de fixation du nickel

Heddle *et al.* proposent que le nickel se fixe au sein de la protéine NikA sous une forme penta-hydratée dans une poche relativement hydrophobe. Il n'y a par ailleurs aucune liaison directe entre le nickel et la protéine. Il existe uniquement une liaison hydrogène entre l'arginine 97 et une des molécules d'eau (Figure 1.5-3).

La conformation des résidus du site de fixation reste, pour la plupart, inchangée entre les formes ouverte et fermée de NikA. Il n'y a que les résidus Glu221 et Arg97 qui présentent un réel changement de conformation. Dans le cas de ce dernier, son mouvement vers le nickel penta-hydraté chargé positivement est difficile à comprendre, puisqu'il est lui-même également chargé positivement.

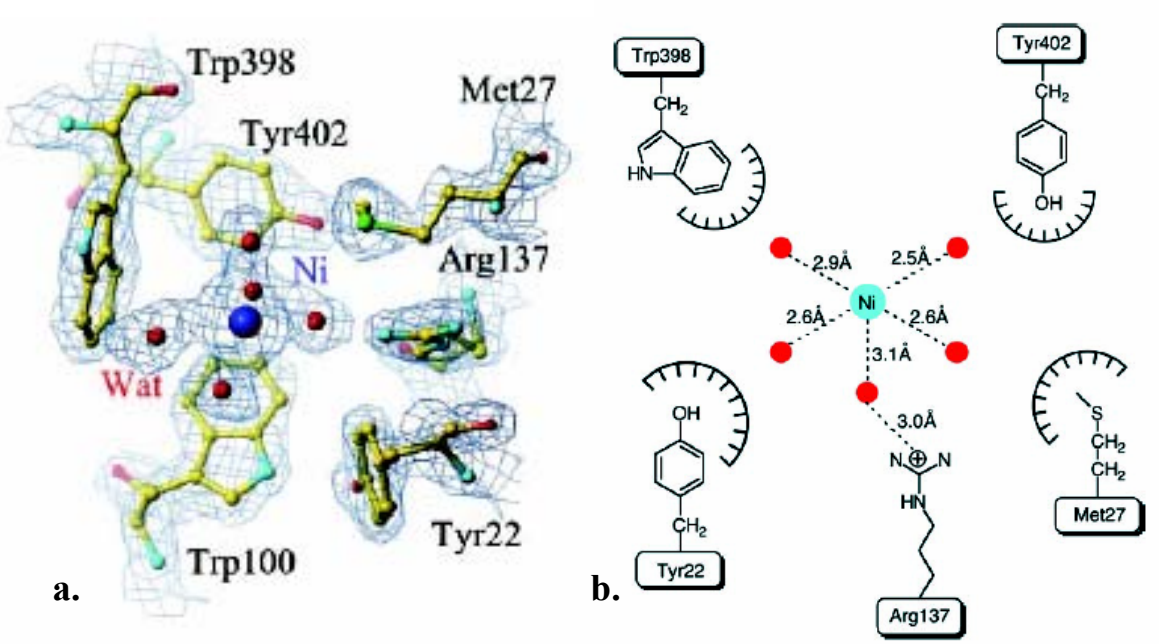


Figure 1.5-3 : Représentations du site de fixation du nickel dans NikA proposé par Heddle *et al.*
a. Densité électronique $2F_{obs}$-F_{calc} du nickel penta-hydraté dans son site de fixation.
b. Schéma des interactions entre le nickel penta-hydraté et la protéine.

1.5.3 La comparaison de la structure de NikA avec les protéines OppA et DppA

Les protéines OppA (Tame *et al.*, 1994) et DppA (Dunten and Mowbray, 1995) sont deux protéines de la famille des protéines fixatrices d'oligopeptides et de dipeptides du périplasme, qui ont une identité de séquence avec NikA (respectivement 24 et 26% d'identité de séquence). La comparaison de leur structure avec celle de NikA montre que, comme attendu, les trois structures présentent des repliements proches.

Pourtant, les différences sont suffisamment significatives pour avoir empêché la résolution de la structure de NikA par remplacement moléculaire en utilisant comme modèle ces deux protéines (Figure 1.5-4). En effet, les principales différences de la structure de NikA avec les structures d'OppA et DppA consistent en la position et la taille de certaines boucles et structures secondaires, et l'angle d'ouverture entre les deux domaines de la protéine.

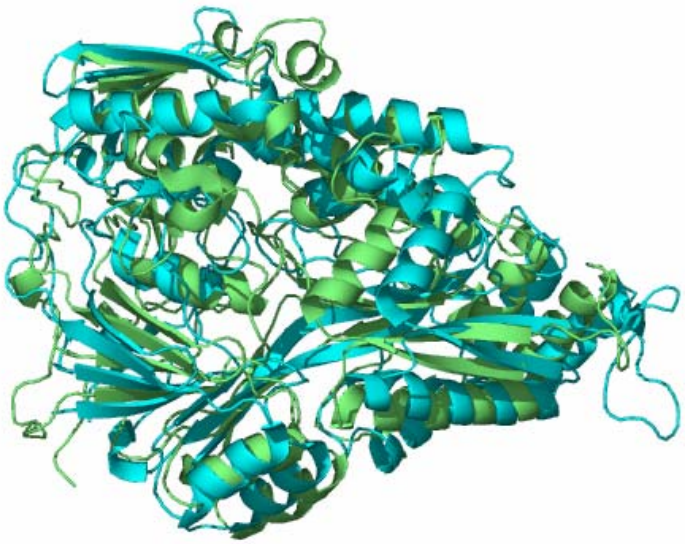


Figure 1.5-4 : Superposition de la structure de NikA apo (forme ouverte en bleu ; code PDB : 1UIU) et de OppA (forme fermée en vert ; code PDB : 1JEV). Le rmsd entre les C_α de ces deux protéines est de 5,18 Å.

1.6 La résolution de notre structure de NikA par remplacement moléculaire

Une fois les structures résolues par Heddle *et al.* disponibles, nous avons souhaité vérifier que la protéine NikA sur laquelle nous travaillions présentait la même structure.

Pour ce faire, nous avons réalisé un remplacement moléculaire à l'aide du programme AMoRe, à partir d'un jeu de données d'un cristal de protéine native ayant diffracté à 2,2 Å (chapitre 1.4.1), et de la structure publiée contenant du nickel.

Le remplacement moléculaire n'a présenté aucune difficulté (les paramètres de maille étant similaires dans les deux cas) et a permis de placer deux molécules de NikA par unité asymétrique, qui ont ensuite été affinées par le programme Refmac, sans appliquer de contraintes de symétrie non cristallographique (Tableau 1.6-1).

Notre structure s'est révélée très similaire à la structure de NikA contenant du nickel, déjà publiée (qui nous à servi de modèle pour le remplacement moléculaire). Le rmsd entre les carbones α des deux structures est de 0,58 Å. Cette grande similarité concerne également les résidus du site de fixation du nickel. Les seules différences concernent des boucles de surface en contact avec le solvant.

Les deux protéines présentes dans l'unité asymétrique sont équivalentes, mais la molécule B semble mieux définie, notamment au niveau du site de fixation. C'est la raison pour laquelle, dans la suite de ce manuscrit, c'est sur cette molécule que les analyses structurales ont été menées, puis vérifiées sur la molécule A.

Facteur R et R_{free} (%)	18,6 / 24,7
Nombre d'atomes de la protéine	7 987
Nombre de molécules d'eau	447
Facteur B moyen (Å^2)	26,1
Rmsd	
Liaison (Å)	0,011
Angles (°)	1,438
Ramamchandran	
Résidus dans la région la plus favorable (%)	92,0
Résidus dans la région favorable (%)	7,7
Résidus dans la région peu favorable (%)	0,3
Résidus dans la région interdite (%)	0,0

Tableau 1.6-1 : Statistiques d'affinement de la structure de NikA à 2,2 Å de résolution.

1.7 Pourquoi n'a-t-on pas pu résoudre la structure de NikA à partir de la protéine séléniée : un début de réponse

Nous avons voulu comprendre pourquoi les cristaux de protéine séléniée ne nous avaient pas fourni un signal anomal suffisant nous permettant la résolution de la structure de NikA, d'autant plus que Heddle *et al* ont utilisé cette méthode avec succès.

1.7.1 La vérification de la présence des atomes de sélénium par cristallographie

Dans un premier temps, nous avons résolu la structure de la protéine séléniée ayant permis d'enregistrer des données de diffraction MAD (chapitre 1.4.2.1). Pour ce faire, nous avons repris les données de diffraction enregistrées à la longueur d'onde du maximum de $^{\lambda}f''$ du sélénium (2,74 Å de résolution). Nous avons réalisé un remplacement moléculaire, à partir de la structure de NikA que nous venions de résoudre, qui n'a pas présenté de difficultés, les deux structures étant quasiment identiques. L'affinement de la structure par CNS et Refmac (R = 18,1% et R_{free} = 32,4%) a permis d'obtenir des cartes de densité électronique de bonne qualité.

Le modèle que nous avons affiné contenait des méthionines, par conséquent si la protéine était bien séléniée, l'on devrait observer un déficit de densité électronique (densité positive dans la carte $F_{obs} - F_{calc}$) au niveau des soufres des méthionines. Il fut alors surprenant de

constater que les 10 méthionines de NikA sont parfaitement définies et qu'aucun déficit en densité électronique n'était visible (Figure 1.7-1).

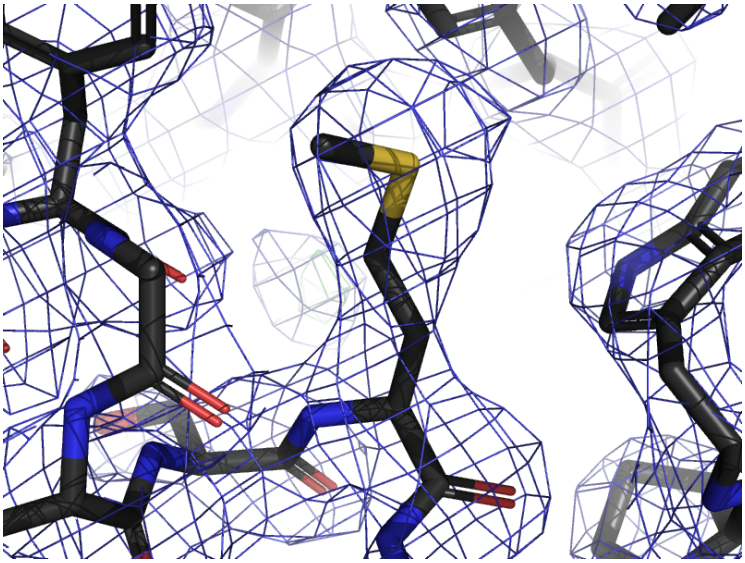


Figure 1.7-1 : Carte de densité électronique $2F_{obs}$-F_{calc} (en bleu), contourée à 1σ, et de la densité électronique F_{obs}-F_{calc} (en vert), contourée à 3σ, de la méthionine 32 de NikA.

Afin de confirmer ce résultat, nous avons calculé les cartes de différence anomale à partir de ce jeu de données.

Dans le cas de certaines méthionines une faible densité résiduelle F_{obs}–F_{calc} est observable, mais elle n'est pas suffisante pour être expliquée par la présence de sélénium à la place de soufre. Nous n'avons donc pas réussi à mettre en évidence la présence de séléno-méthionine à la place des méthionines.

Ces résultats tendent alors à montrer que la protéine n'a pas (ou peu) été séléniée. Pourtant, nous avons été capables d'enregistrer un spectre de fluorescence du sélénium à partir de ce cristal de protéine.

1.7.2 L'analyse du contenu en atomes lourds des cristaux de NikA séléniée par fluorescence des rayons X

Dans un second temps, nous avons alors décidé d'analyser le contenu en atomes lourds des cristaux de protéine séléniée par fluorescence des rayons X à l'aide du détecteur de fluorescence (Roentec®) de la ligne BM30A (FIP). Nous avons réalisé cette expérience sur plusieurs cristaux de NikA séléniée.

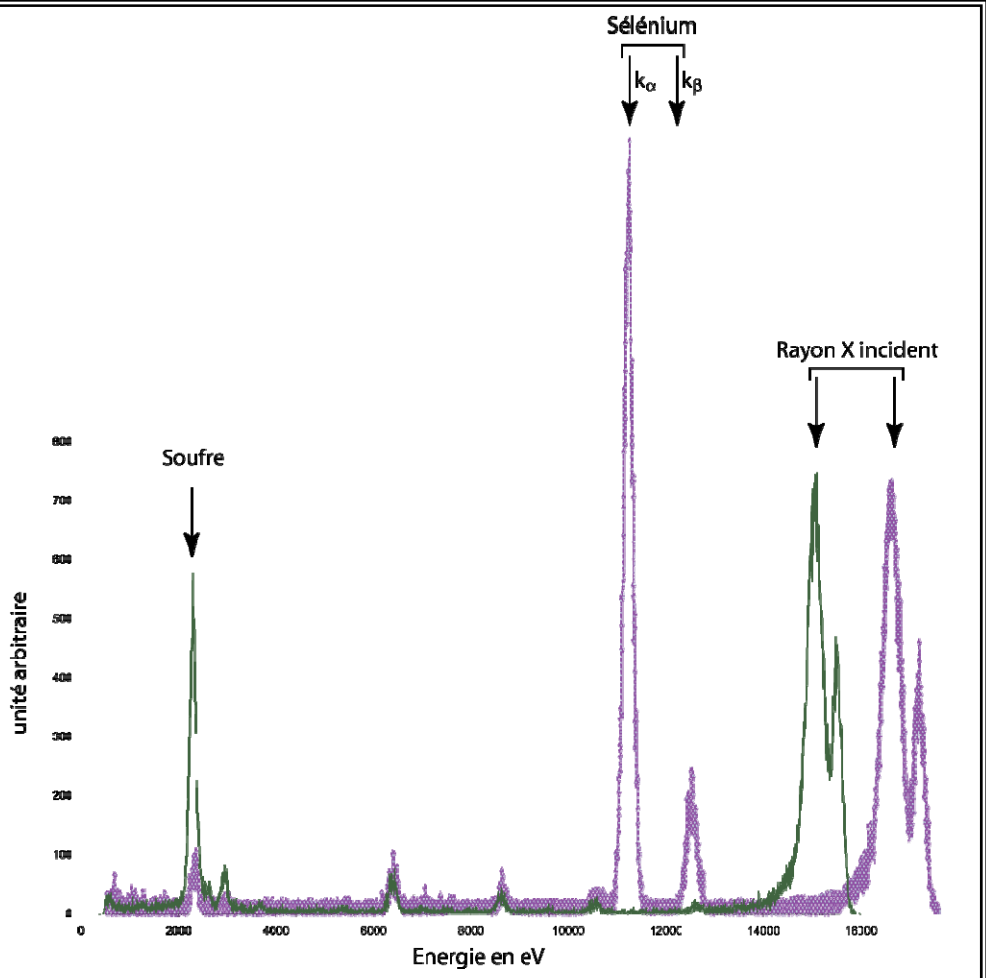


Figure 1.7-2 : Pics de fluorescence X mesurée sur des cristaux de NikA séléniée.

Deux cas de figure se sont alors présentés :

- Dans certains cas, le spectre de fluorescence n'a montré la présence que d'une faible quantité de sélénium (spectre vert de la Figure 1.7-2), ce à quoi nous nous attendions ;
- Mais nous avons été étonnés de constater que certains cristaux présentaient une quantité importante de sélénium corrélée à une diminution de la quantité de soufre par rapport aux cristaux précédents (spectre violet de la Figure 1.7-2). Ils présentent en outre une mauvaise qualité de diffraction.

Notre hypothèse est alors que le lot de protéine séléniée soit en réalité un mélange de protéine native et de protéine séléniée, et que la cristallisation les sélectionne en partie. En effet, il est possible de trouver dans la même goutte de cristallisation des cristaux :

- contenant du sélénium, mais diffractant très mal ;

- et d'autres n'ayant que peu de sélénium, mais quand même suffisamment pour obtenir une fluorescence acceptable, et présentant une diffraction correcte.

Lors de la recherche du meilleur cristal pour réaliser une expérience MAD, nous avons choisi les cristaux présentant la meilleure diffraction, et de ce fait nous avons sélectionné les cristaux ne possédant pas suffisamment de sélénium pour mener à bien la résolution de la structure de la protéine.
Nous ne sommes pas certains de l'origine de cette hétérogénéité, mais il est possible que certaines des cultures se soient faites avec une souche bactérienne n'ayant pas ou ayant perdu l'auxotrophie à la méthionine, et que ces dernières aient alors été mélangées avec des cultures de bactéries auxotrophes.

1.8 L'étude du site de fixation de NikA

1.8.1 La densité électronique du contenu du site de fixation

1.8.1.1 Les problèmes posés par la fixation du nickel à NikA sous une forme penta-hydratée

Comme nous l'avons déjà vu plus haut, il est surprenant de constater que Heddle *et al.* proposent un modèle de fixation du nickel sous une forme penta-hydratée ($Ni(H_2O)_5^{2+}$) dans une poche de NikA relativement hydrophobe, sans liaison directe entre le métal et la protéine. Mais ce modèle pose un autre problème : la géométrie et la coordination du nickel ne respectent pas ce qui a été déterminé pour le nickel hexa-hydraté ($Ni(H_2O)_6^{2+}$) (Figure 1.8-1) :

- L'angle H_2O-Ni-H_2O n'est pas de 90° comme ce qui est normalement constaté pour une géométrie octaédrique ;
- Les distances Ni-H_2O sont comprises entre 2,5 et 3,1 Å alors qu'elles devraient être comprises pour un nickel hexa-hydraté entre 2,05 et 2,10 Å.

De plus, le modèle de nickel penta-hydraté n'explique pas complètement la densité électronique, calculée à partir des facteurs de structure, de NikA contenant du nickel, déposée par Heddle *et al.*

Toutes ces raisons nous ont poussé à nous intéresser de plus près au contenu du site de fixation de NikA.

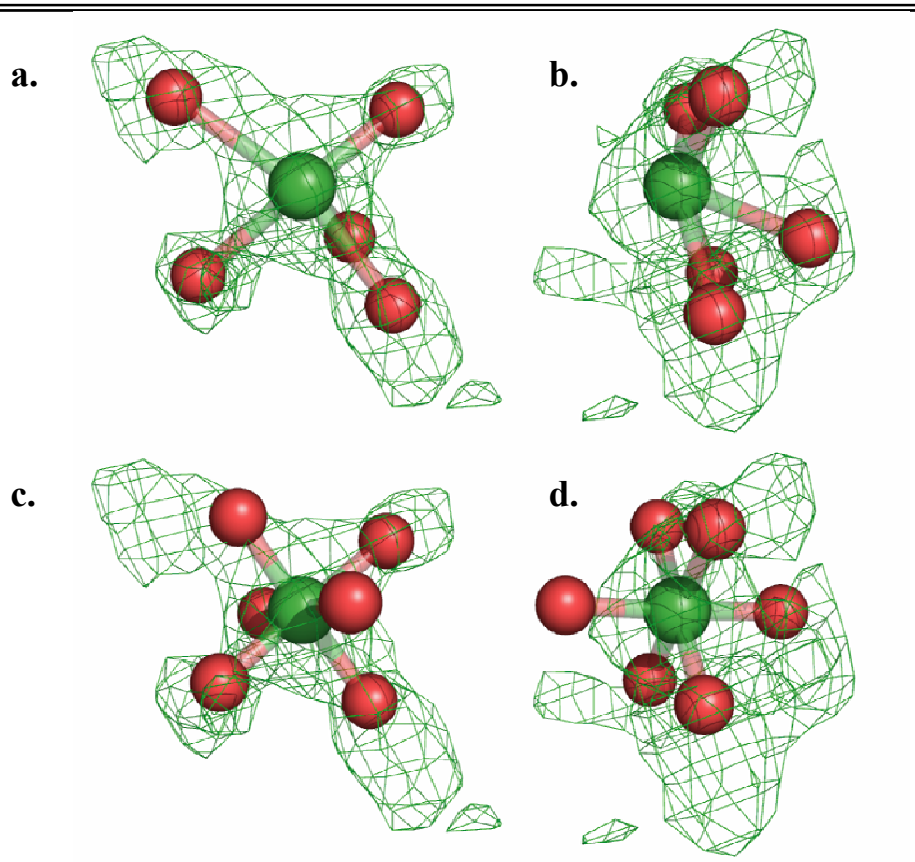


Figure 1.8-1 : Vues à 90° l'une de l'autre de la densité électronique F_{obs}-F_{calc} (carte « omit » ; contourée à 3σ) de la structure de NikA contenant un nickel résolue par *Heddle et al.*, avec un modèle de nickel penta-hydraté (a. et b.) et un nickel hexa-hydraté canonique (c. et d.).

Le nickel est représenté par une sphère verte et les molécules d'eau par des sphères rouges.

1.8.1.2 La densité électronique présente dans le site de fixation de la protéine NikA

Pour pouvoir étudier le site de fixation sans être biaisé par le modèle, nous avons calculé des cartes de densité « omit » en excluant le nickel et sa coordination. Nous avons fait ceci sur la structure que nous avons résolue à 2,2 Å de résolution (Figure 1.8-2a) ainsi que pour la structure contenant du nickel résolue par Heddle *et al.* à 1,95 Å de résolution (Figure 1.8-2b).

La densité électronique est très similaire dans les deux cas, et ne correspond pas à ce que l'on devrait s'attendre pour un nickel penta-hydraté.

Nous avons donc voulu caractériser la nature du métal et des ligands du métal présents dans le site de fixation de la protéine.

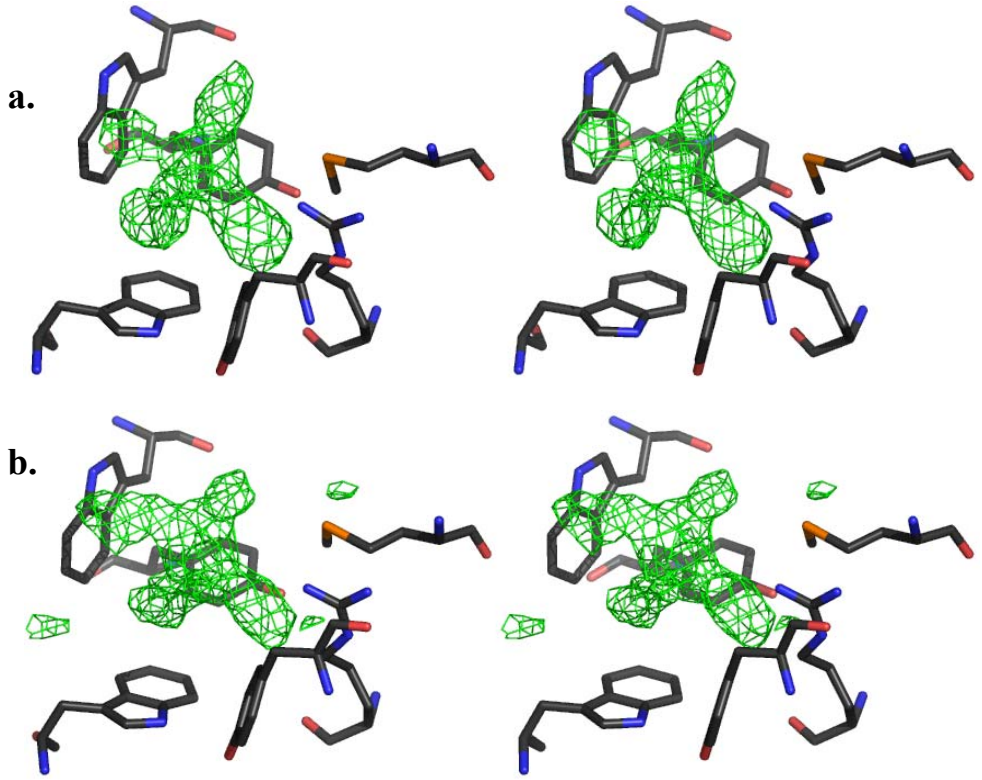


Figure 1.8-2 : Vue stéréoscopique de la densité électronique « omit » (contourée à 3σ) présente :
a. dans le site de fixation de la structure à 2,2 Å de résolution que nous avons résolue.
b. dans le site de fixation de la structure à 1,95 Å de résolution résolue par *Heddle et al.* (code PDB : 1UIV)

1.8.2 La caractérisation du métal présent dans le site de fixation de NikA

Le pic le plus intense de notre carte de densité « omit » est mesuré à 15σ, ce qui correspond bien à la densité électronique d'un métal. Sa localisation correspond parfaitement à celle du nickel dans la structure publiée. Nous avons donc bien le métal, dans le site actif de notre structure, dont nous avons voulu vérifier la nature.

1.8.2.1 La réalisation de cartes de double différence anomale

1.8.2.1.1 L'analyse du contenu en atomes lourds des cristaux de NikA par fluorescence des rayons X

Afin de caractériser la nature du métal présent dans NikA, nous avons dans un premier temps dû définir les atomes lourds présents dans les cristaux de protéine.
Pour ce faire, nous avons analysé par fluorescence des rayons X plusieurs cristaux, à l'aide du détecteur de fluorescence (Roentec®) de la ligne BM30A (FIP).

Dans tous les cas, les mêmes résultats furent obtenus, à savoir la mise en évidence de la présence de fer et de zinc, mais, ce qui était surprenant, une absence de nickel (Figure 1.8-3).

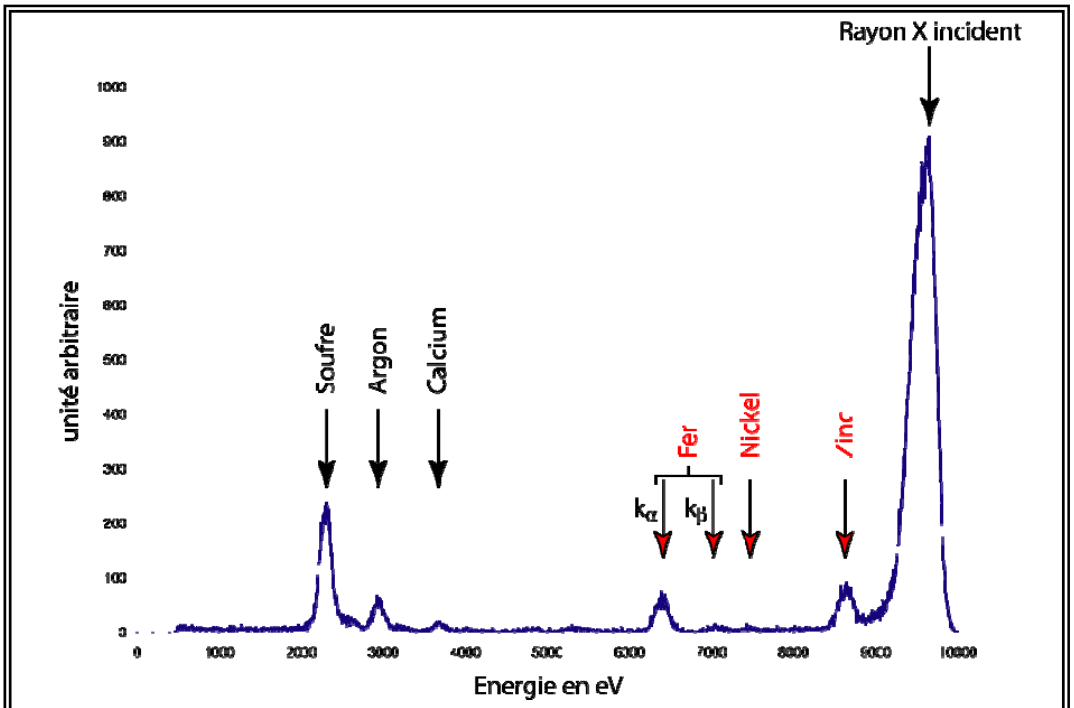


Figure 1.8-3 : Pics de fluorescence X mesurée sur des cristaux de NikA.

1.8.2.1.2 L'enregistrement des données de diffraction

Afin de localiser ces trois métaux grâce à des cartes de double différence anomale, nous avons enregistré des jeux de diffraction à quatre longueurs d'ondes permettant d'enregistrer la diffusion anomale (chapitre 2.5.2.3.3 du matériel et méthodes) :

- d'aucun des atomes suspectés d'être présents dans la structure, pour $\lambda_1 = 1.7438$ Å soit 7110,0 eV (légèrement inférieur en énergie au seuil d'absorption K du Fe) ;
- du Fe, pour $\lambda_2 = 1.7432$ Å soit 7125,0 eV (légèrement supérieur en énergie au seuil d'absorption K du Fe) ;
- du Fe et du Ni, pour $\lambda_3 = 1,4839$ Å soit 8355,0 eV (légèrement supérieur en énergie seuil d'absorption K du Ni) ;
- du Fe, du Ni et du Zn, pour $\lambda_4 = 1,2830$ Å soit 9663,5 eV (légèrement supérieur en énergie au seuil d'absorption K du Zn).

La structure a été résolue par remplacement moléculaire à partir de la structure de NikA que nous avons résolue dans le cas du jeu de données enregistré à la longueur d'onde λ_4. Cette

solution a été alors utilisée dans le cas des trois autres jeux de données, afin de calculer les cartes de différence anomale. Les statistiques d'indexation et d'affinement sont présentées dans le Tableau 1.8-1.

1.8.2.1.3 Le calcul des cartes de double différence anomale

Les jeux anomaux enregistrés ont permis de calculer des cartes Fourier de différence anomale. Pour réaliser ces dernières, les programmes de la suite CCP4 furent utilisés, afin de mettre à l'échelle les quatre jeux de données enregistrés (Scala ; (Kabsch, 1988)), et de générer les cartes différence anomale (FFT ; (CCP4, 1994)). Les cartes de double différence anomale ont été calculées en soustrayant les cartes de différence anomale deux à deux, à l'aide d'un programme développé par Anne Volbeda, du laboratoire, de façon à mettre en évidence en calculant :

- λ_2 - λ_1 : la présence éventuelle du Fe ;
- λ_3 - λ_2 : la présence éventuelle du Ni ;
- λ_4 - λ_3 : la présence éventuelle du Zn.

Nous avons ainsi pu montrer qu'il n'y aucune trace de zinc ni de nickel dans le site de fixation de NikA, zinc qui doit se fixer de façon non spécifique et avec une faible occupation à la protéine. Seul le fer y est présent (Figure 1.8-4).

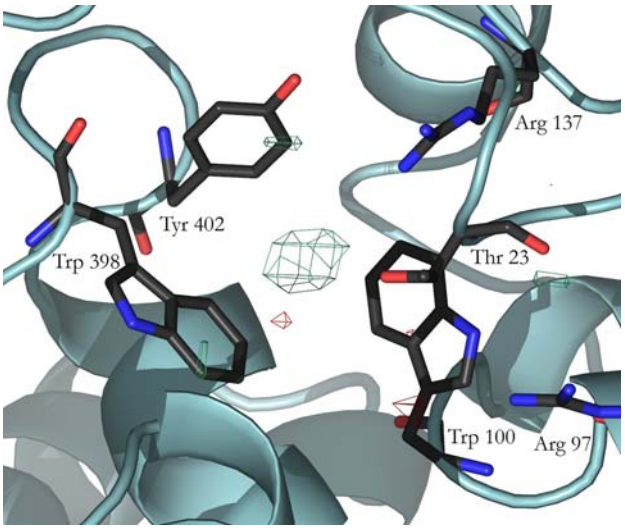


Figure 1.8-4 : Carte de densité électronique double différence anomale mettant en évidence la présence de fer (λ_2 - λ_1) dans le site de fixation de NikA.

En vert la densité positive contourée à 3σ. En rouge la densité négative contourée à -3σ.

	λ_1	λ_2	λ_3	λ_4
Ligne de lumière de l'ESRF	BM30A (FIP)			
Longueur d'onde	1.7438 Å	1.7432 Å	1,4839 Å	1,2830 Å
Groupe d'espace	$P2_12_12_1$			
Paramètres de maille	a = 85,78 ; b = 92,58 ; c = 122,41; α = 90° ; β = 90° ; γ = 90°			
Nombre de molécules par unité asymétrique	2			
Résolution	20,00–2,88	20,00–2,88	20,00–2,78	20,00–2,70
	(2,98–2,88)	(2,98–2,88)	(2,88–2,78)	(2,80–2,70)
Complétude	91,9 (85,3)	92,9 (87,1)	94,3 (91,0)	95,8 (94,8)
$I/\sigma(I)$	12,49	13,06	17,96	20,84
	(2,96)	(3,17)	(5,95)	(8,36)
Nombre total de réflexions mesurées	68 925	78 702	89 170	120 562
	(6 224)	(7 142)	(8 625)	(12 192)
Redondance	1,8 (1,8)	2,0 (2,0)	2,0 (2,0)	2,4 (2,4)
R_{SYM}	5,7 (27,0)	5,8 (27,3)	3,8 (15,0)	3,8 (14,1)
Facteur R et R_{free} (%)	19,5 / 28,8	20,6 / 29,3	20,3 / 29,3	20,7 / 29,3
Nombre d'atomes de la protéine	7 987			
Nombre de molécules d'eau	447			
Facteur B moyen (Å^2)	25,6			
Rmsd				
Liaison (Å)	0,013			
Angles (°)	1,539			
Ramachandran (%)				
Résidus dans la région la plus favorable	90,9			
Résidus dans la région favorable	9,6			
Résidus dans la région peu favorable	0,5			
Résidus dans la région interdite	0,0			

Tableau 1.8-1 : Statistiques des données de diffraction et d'affinement de la structure ayant permis le calcul des cartes de double différence anomale.

Une mesure de la quantité de fer par la technique colorimétrique de la bathophénanthroline, a permis de déterminer que cette dernière est proche de la stœchiométrie (chapitre 1.2.5 du matériel et méthodes).

1.8.2.2 La spectroscopie RPE

Nous avons par la suite souhaité déterminer l'état d'oxydation du fer : Fe^{2+} ou Fe^{3+}. Pour cela, nous avons enregistré un spectre RPE correspondant au fer (Figure 1.8-5).
NikA contient donc un Fe^{3+}, car le Fe^{2+}, qui ne possède pas d'électron célibataire (n'étant donc pas paramagnétique), n'aurait pas donné un signal en spectroscopie RPE.

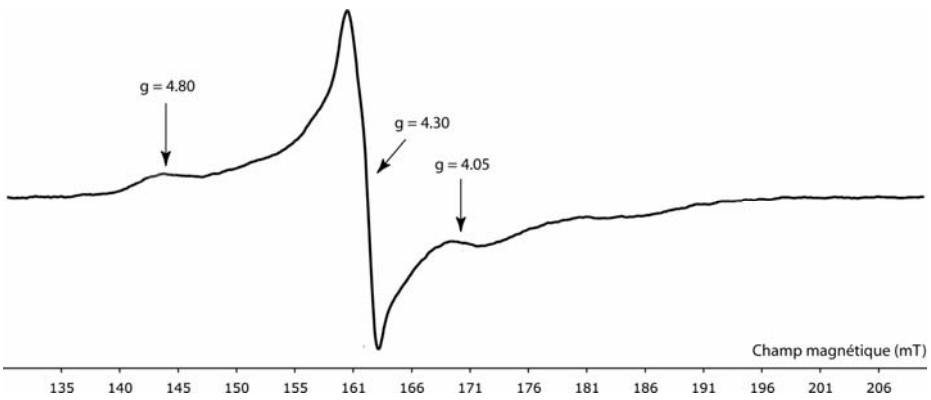


Figure 1.8-5 : Spectre RPE enregistré sur la protéine NikA.

1.8.3 La caractérisation du contenu du site de fixation de NikA

1.8.3.1 L'identification par cristallographie

1.8.3.1.1 Les tentatives d'identification du ligand de NikA

Afin d'identifier la nature du ligand du fer dans NikA, nous avons essayé de placer de nombreuses molécules dans le site de fixation de NikA.
Pour les choisir, nous nous sommes basés sur la forme de la densité électronique des cartes « omit », excluant le contenu du site de fixation de NikA, que nous avions calculé précédemment (Figure 1.8-2a). Une fois la ou les molécules placées autour de l'atome de fer dans la protéine, le modèle était affiné par dix cycles d'affinement par Refmac. On étudiait ensuite la densité résiduelle calculée à partir des coefficients F_{obs}-F_{calc} présente dans le site de fixation afin de déterminer la qualité du modèle. Il nous paraissait important que la coordination du fer respecte la géométrie octaédrique que semblait indiquer la forme de la densité électronique « omit ».

Nous avons dans un premier temps placé trois acétates et deux molécules d'eau (Figure 1.8-6a). Nous avons choisi les acétates car la densité électronique semblait présenter une forme aplatie caractéristiques des carboxylates, et que leur charge négative compense la charge positive du métal.

Ce modèle n'étant pas satisfaisant, nous avons par la suite remplacé deux des acétates par des glycinates (Figure 1.8-6b). Là encore, la densité résiduelle était trop importante pour accepter ce modèle.

De nombreuses coordinations furent testées, mais aucune ne fut satisfaisante. La qualité de la densité électronique était un facteur limitant empêchant l'identification du ou des ligands du fer.

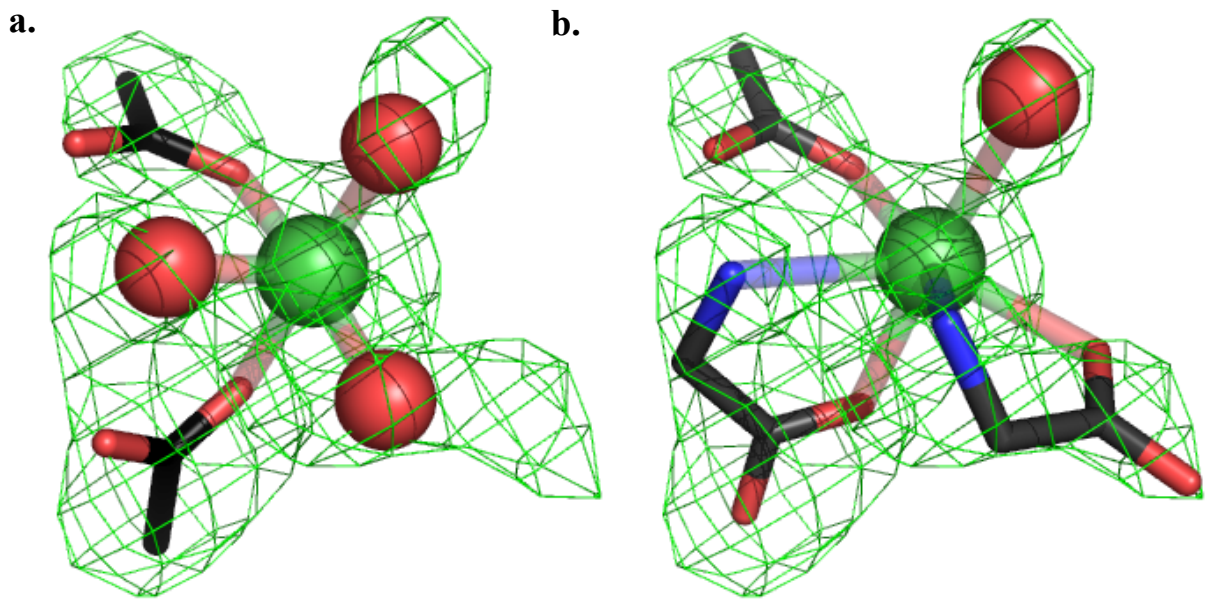


Figure 1.8-6 : Exemples de modèles de ligands du métal présent dans le site de NikA. Ici le métal est un nickel représenté en vert, car la présence du fer n'avait pas encore été mise en évidence lors de l'étude de ces modèles. En vert est représentée la densité électronique « omit » (F_{obs}-F_{calc} ; contourée à 3σ) calculée à partir de la structure à 2,2 Å de résolution que nous avons résolue.

1.8.3.1.2 L'obtention d'une structure diffractant à 1,8 Å de résolution

Nous avons essayé, de nouveau, d'améliorer la qualité de diffraction des cristaux de NikA, car la résolution de la structure sur laquelle nous travaillons (2,2 Å) n'est pas suffisante pour permettre l'identification du ou des ligands du fer.

Ceci fut réalisé en augmentant la concentration de la protéine dans les gouttes de cristallisation : au lieu d'utiliser une protéine concentrée à 11 mg/mL, comme ce que nous

avons précédemment utilisé, nous avons déterminé que la concentration optimale était de 21 mg/mL.

De cette façon, un jeu de donnée de diffraction allant jusqu'à une résolution de 1,8 Å fut enregistré sur la ligne de lumière ID29, grâce à un cristal étant apparu dans une goutte de cristallisation de la composition suivante :

[NikA] = 11 mg / mL ;

Sulfate d'ammonium 1,5 M ;

Acétate de sodium 100 mM pH 4.7 ;

20°C en goutte suspendue.

Ligne de lumière de l'ESRF	ID29
Longueur d'onde	0,9756 Å
Groupe d'espace	$P2_12_12_1$
Paramètres de maille	a = 86,83 Å ; b = 93,87 Å ; c = 124,49 Å α = 90° ; β = 90° ; γ=90°
Nombre de molécules par unité asymétrique	2
Résolution	40,0–1,8 Å (1,9–1,8 Å)
Complétude (%)	97,6 (94,9)
I/σ(I)	17,96 (5,19)
Nombre total de réflexions mesurées	302 914 (43 603)
Redondance	3,3 (3,3)
R_{SYM}	5,0 (30,4)
Facteur R et R_{free} (%)	16,5/21,8
Nombre d'atomes de la protéine	8172
Nombre de molécules d'eau	860
Facteur B moyen (Å^2)	20.6
Rmsd	
Liaison (Å)	0,013
Angles (°)	1,457
Ramachandran	
Résidus dans la région la plus favorable (%)	92,3
Résidus dans la région favorable (%)	7,4
Résidus dans la région peu favorable (%)	0,3
Résidus dans la région interdite (%)	0,0

Tableau 1.8-2 : Statistiques des données de diffraction et d'affinement de la structure de la protéine NikA native. Les valeurs entre parenthèses correspondent à la dernière tranche de résolution.

La structure fut résolue par remplacement moléculaire à l'aide du programme Phaser, et l'affinement a été réalisé à l'aide de Refmac (Tableau 1.8-2). Cette structure a été déposée dans la *Protein Data Bank* (PDB) sous le code 1ZLQ (Cherrier *et al.*, 2005).

1.8.3.1.3 L'identification de l'EDTA comme ligand du nickel

Nous avons calculé une nouvelle carte de densité électronique « omit », excluant le fer et ses coordinants, à partir de la nouvelle structure obtenue.

Cette dernière étant de meilleure qualité que celle sur laquelle nous travaillions jusqu'à présent, nous nous sommes aperçus que contrairement à ce que l'on avait supposé jusqu'à présent, le fer n'adopte pas une géométrie octaédrique (six ligands), mais bipyramidale pentagonale (sept ligands).

Nous nous sommes alors rendu compte que le Fe^{3+} est connu pour fixer l'EDTA (et ses dérivés) en formant une géométrie bipyramidale pentagonale à cause d'une encapsulation incomplète du fer (Cotton and Wilkinson).

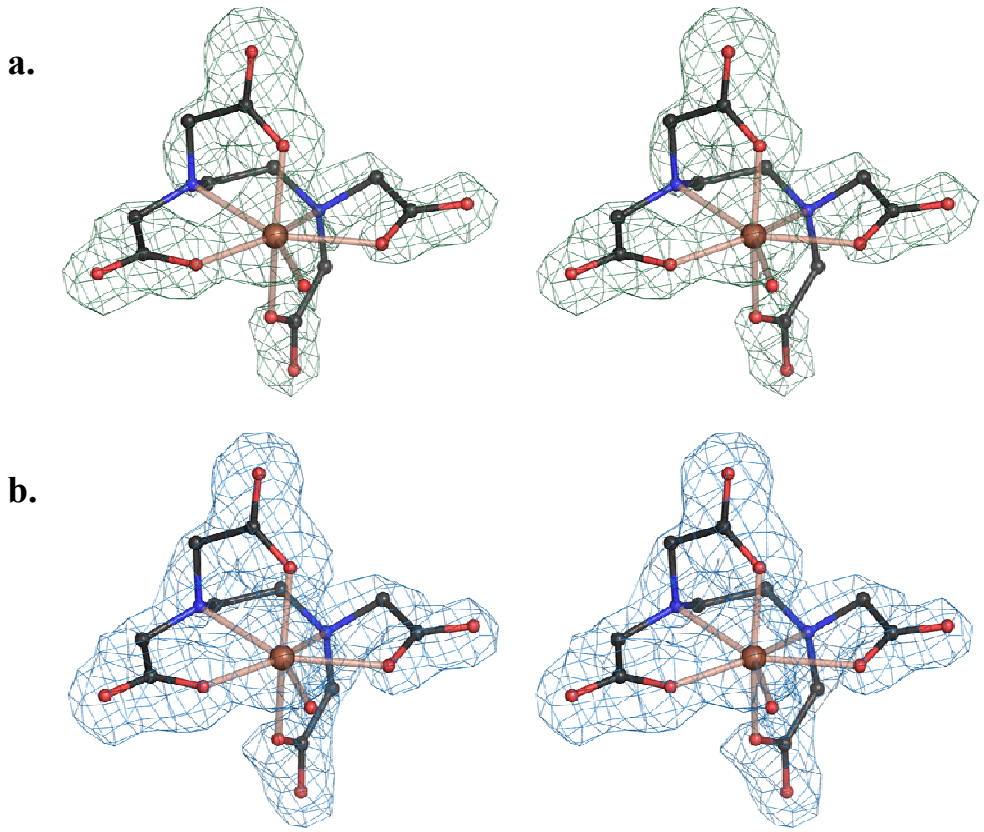


Figure 1.8-7 : Vue stéréoscopique de la densité électronique calculée autour de l'atome de fer.
a. Carte de densité électronique « omit » (F_{obs}-F_{calc}) contourée à 3σ calculée en omettant le FeEDTA(H_2O)$^-$ du modèle.
b. Carte de densité électronique finale (2F_{obs}-F_{calc}) contourée à 1σ.

Nous avons alors décidé de placer de l'EDTA dans la densité électronique nouvellement calculée, utilisant comme septième ligand une molécule d'eau labile (Schneppensieper *et al.*, 2000).

Ce nouveau modèle, contrairement aux précédents, s'est parfaitement adapté à la carte de densité électronique « omit » (Figure 1.8-7a) et la carte de densité $2F_{obs}$-F_{calc} finale (Figure 1.8-7b).

Afin de confirmer la validité de la géométrie du complexe de $FeEDTA(H_2O)^-$ du site de fixation de NikA, nous l'avons superposé à une structure cristallographique d'un complexe de (o-phénylènediamine-N,N,N',N'-tétraacéto) ferrate(III) : un dérivé de l'EDTA (Mizuno *et al.*, 1991), après inversion de ce dernier pour avoir la chiralité opposée.

Le calcul du rmsd entre les atomes équivalents donne une valeur de 0,27 Å (Figure 1.8-8), ce qui nous indique donc que le complexe $FeEDTA(H_2O)^-$ se fixe dans le site de fixation de NikA sans changement significatif par rapport à sa forme cristalline (Lind *et al.*, 1964).

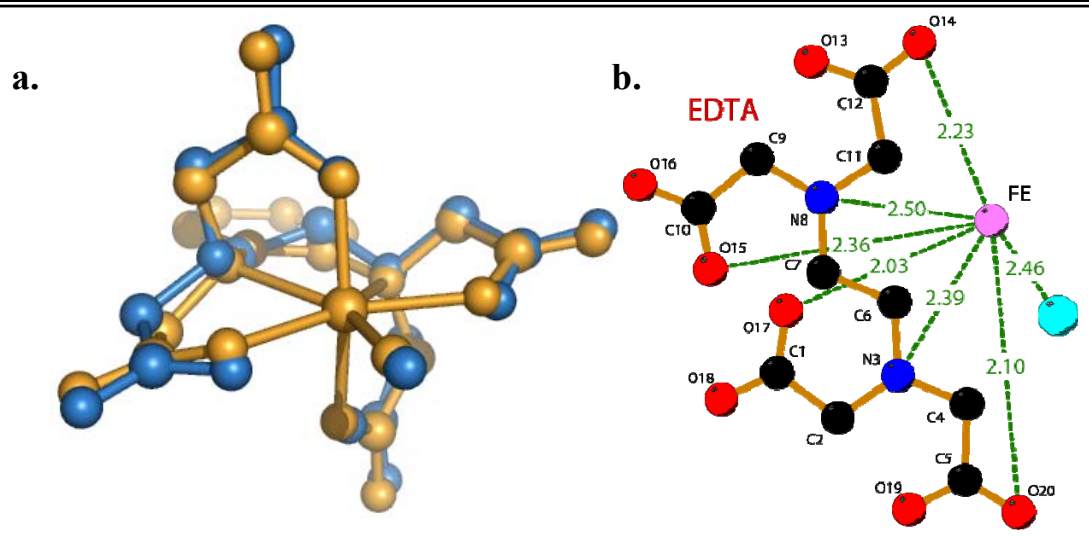


Figure 1.8-8 : Le complexe $FeEDTA(H_2O)^-$.

a. Superposition du modèle de $FeEDTA(H_2O)^-$ du site de fixation de NikA (en bleu), avec le (o-phénylènediamine-N,N,N',N'-tétraacéto) ferrate(III) (en orange). Le rmsd entre atomes équivalents est de 0,27 Å.

b. Figure générée par le programme LIGPLOT (Wallace *et al.*, 1995) représentant les liaisons (en vert ; les distances sont en Å) entre le fer (en rose), l'EDTA et la molécule d'eau labile (en bleu).

1.8.3.2 La spectroscopie de masse

La présence d'un complexe de $FeEDTA(H_2O)^-$ dans le site de fixation de NikA a été confirmée par la réalisation d'un spectre de masse de la protéine NikA en conditions natives (Figure 1.8-9).

Le pic A correspond à la masse de la protéine seule (56 301,4 ± 6,8 Da), ce qui est en accord avec les résultats obtenus en conditions dénaturantes (données non présentées) et le pic B correspond à une masse de 56 641,3 ± 5,6 Da. Le calcul de la différence de masse entre le pic B et le pic A donne une valeur de 340,9 ± 12,4 Da, ce qui est en accord, aux erreurs expérimentales près, avec la masse d'un FeEDTA déshydraté (la perte de la molécule d'eau labile est expliquée par les conditions expérimentales) : 346 Da.

La perte du complexe par une proportion importante des protéines peut être due à l'étape de dialyse contre l'acétate de d'ammonium 20 mM.

Par contre nous n'avons pas expliqué l'origine du pic C correspondant à une masse de 57 020,0 ± 8,8 Da.

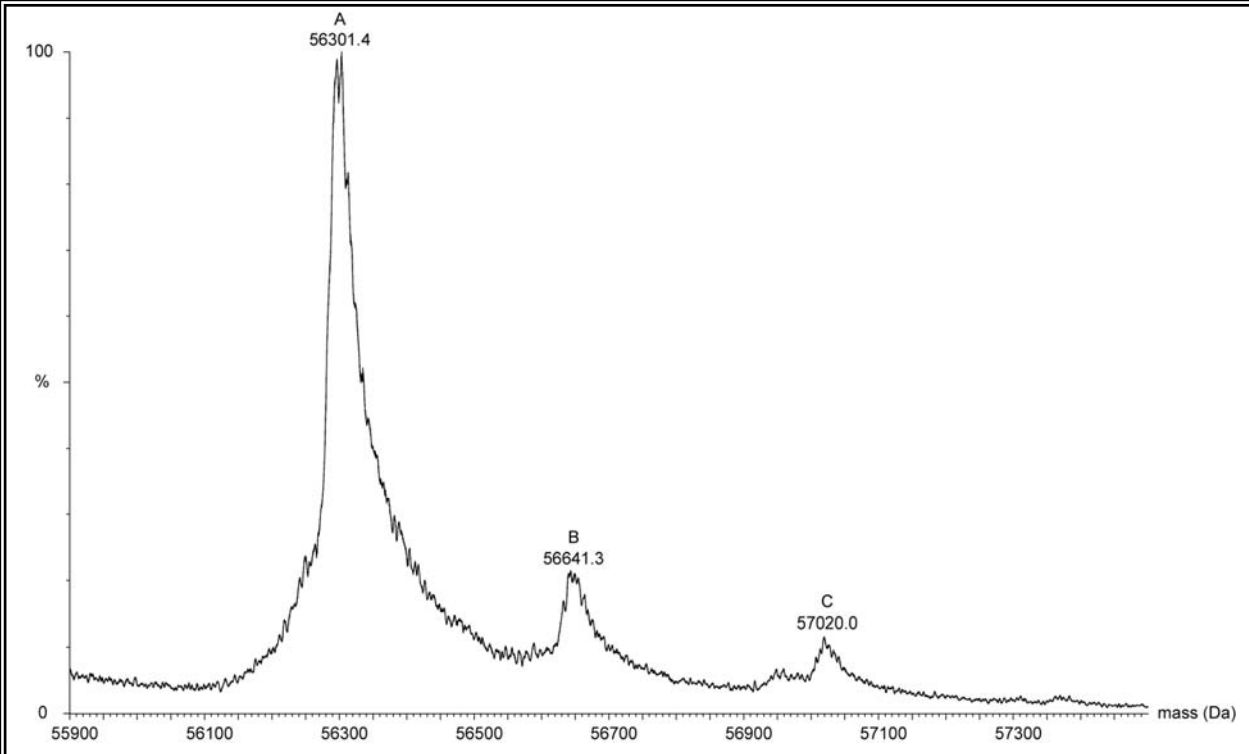


Figure 1.8-9 : Spectre de masse en conditions natives de la protéine NikA dialysée contre de l'acétate d'ammonium 20 mM.

1.8.3.3 Le contenu du site de fixation de NikA dans la structure de Heddle et al.

Nous avons montré que, dans notre cas, NikA contenait un complexe de FeEDTA mais qu'en est-il du contenu de la protéine publiée par Heddle *et al.* ?

Comme nous l'avons vu précédemment, le modèle de nickel penta-hydraté ne permet pas d'expliquer la carte de densité électronique « omit » calculée à partir des facteurs de structure déposés par ces auteurs en omettant le contenu du site de fixation.

Par contre, un complexe de $FeEDTA(H_2O)^-$ permet d'expliquer cette densité électronique « omit » (Figure 1.8-10a) et la carte de densité $2F_{obs}$-F_{calc} finale (Figure 1.8-10b), confirmant ainsi sa présence dans le site de fixation de NikA.

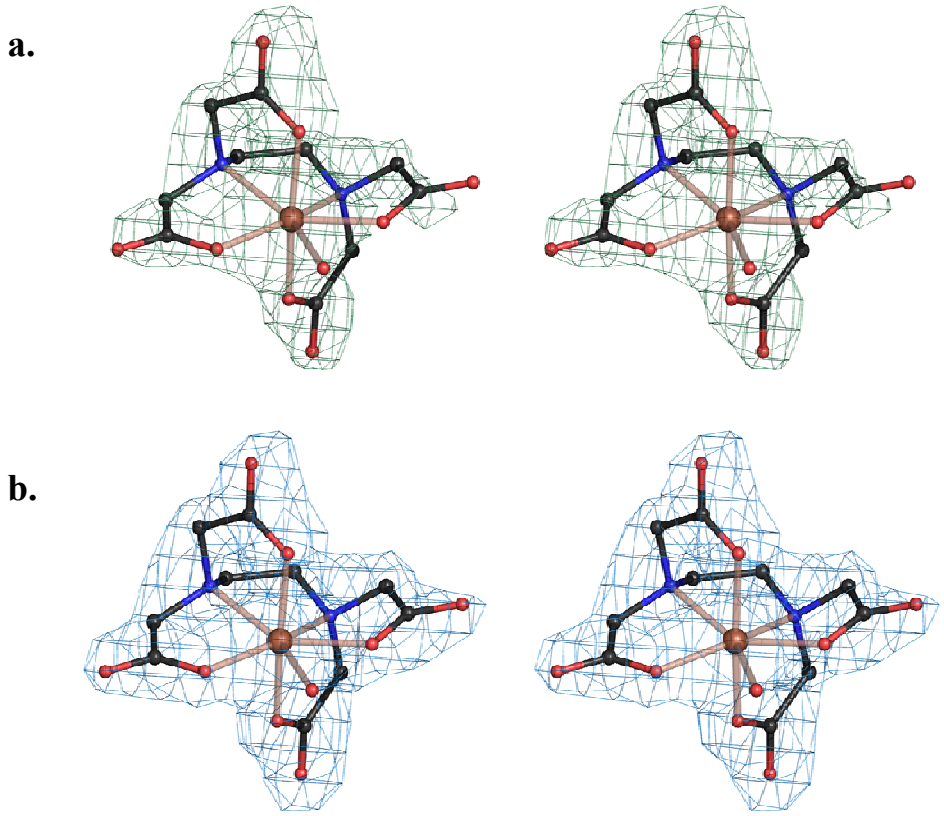


Figure 1.8-10 : Vue stéréoscopique de la densité électronique calculée autour du métal, dans le cas de la structure publiée par *Heddle et al.*

a. Carte de densité électronique « omit » (F_{obs}-F_{calc}) contourée à 3σ calculée en omettant le $FeEDTA(H_2O)^-$ du modèle.

b. Carte de densité électronique finale ($2F_{obs}$-F_{calc}) contourée à 1σ.

Par contre, il est impossible de caractériser la nature du métal à partir des facteurs de structure déposés : les auteurs n'ont pas déposé de jeux de données de diffraction aux longueurs d'ondes nécessaires au calcul de cartes de densité électronique de double différence anomale.

1.8.4 NikA contient un complexe FeEDTA(H_20)$^-$

1.8.4.1 La fixation du complexe FeEDTA(H_20)$^-$ à la protéine NikA

Le fer n'établit pas de liaisons directes avec la protéine, donc la fixation du métal se fait par l'intermédiaire de l'EDTA (Figure 1.8-11) :

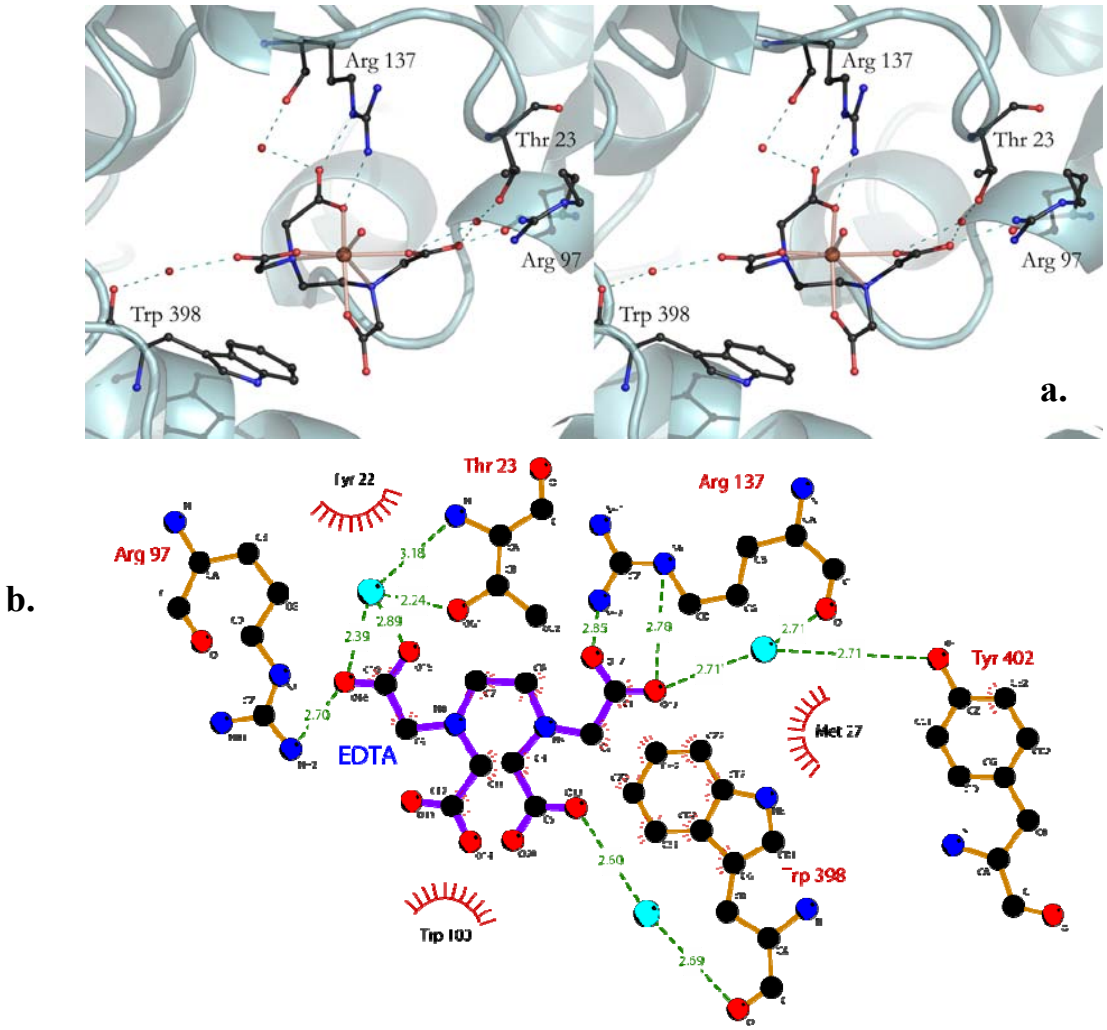


Figure 1.8-11 : Représentation de la fixation du complexe de FeEDTA(H2O)$^-$ dans le site de fixation de NikA. a. Vue stéréoscopique du site de fixation et des liaisons hydrogènes entre l'EDTA et la protéine. b. Figure générée par le programme LIGPLOT (Wallace *et al.*, 1995) représentant les liaisons (en vert ; les distances sont en Å) entre l'EDTA et les résidus de la protéine. Les molécules d'eau sont représentées en bleu.

- Deux des carboxylates de la molécule forment trois liaisons hydrogènes avec les arginines 97 et 137 ;
- S'y ajoutent trois molécules d'eau pontant l'EDTA aux résidus Thr23, Arg137, et Trp389 de la protéine.

La seule liaison entre le métal et la protéine est une interaction cation-π impliquant le Trp398 : la distance fer-Trp398 est 5,5 Å, et l'angle entre la perpendiculaire au plan des cycles du Trp398 et le fer est de 16° (Zaric *et al.*, 2000).
La présence dans le site de fixation d'un complexe de $FeEDTA(H_20)^-$ au lieu de $Ni(H_20)_5^{2+}$ explique le mouvement de l'arginine 97 entre la forme apo de la protéine et la forme ayant fixé le métal, qu'avait mis en évidence *Heddle et al.* mais qu'il n'avait pas interprété.

1.8.4.2 L'origine du FeEDTA

L'EDTA n'étant pas une molécule physiologique, il ne peut provenir que des solutions utilisées pour la production, la purification ou la cristallisation de la protéine.
Le seul moment où la protéine est en contact avec de l'EDTA est l'étape d'extraction périplasmique, c'est donc lors de cette étape de la purification que la protéine fixe le complexe FeEDTA.
Plusieurs éléments nous permettent d'estimer que l'affinité de la protéine pour le complexe est importante :

- Le $FeEDTA(H_2O)^-$ n'est pas dissocié par toutes les étapes de purification et de cristallisation qui ne contiennent pas d'EDTA ;
- Il n'est pas possible d'extraire complètement le fer lors d'un traitement des cristaux de NikA par une solution de ferrichrome (qui est un sidérophore, et présente donc de ce fait une affinité importante pour le fer, supérieure à celle de l'EDTA) ;
- Des cristaux de protéine trempés pendant plusieurs mois en présence de $NiCl_2$ 200 µM ne présentent aucune trace de nickel par mesure du spectre de fluorescence X.

Aucun métal ne fut ajouté aux étapes de surexpression, purification et cristallisation de la protéine. Lors de l'étape d'extraction périplasmique, L'EDTA s'est donc fixé au métal présent dans les tampons (sous la forme de contaminants) pour lequel il a la plus grande affinité. Il

s'agit d'un Fe(III) puisque la purification se fait en conditions aérobies, et que la constante de formation du complexe Fe(III)EDTA est de l'ordre de 10^{25}, en comparaison à celle de du Fe(II)EDTA et du Ni(II)EDTA qui sont respectivement de l'ordre 10^{14} et 10^{18}.
Ce résultat explique également la nature des deux types de cristaux obtenus lors des premières études de NikA menées au laboratoire, visant à en résoudre la structure. Ces deux formes cristallines sont :

- Une forme hexagonale obtenue uniquement avec de la protéine purifiée grâce à un protocole ne contenant pas d'EDTA ;
- Une forme orthorhombique, que nous avons reproduite, obtenue avec la même protéine que précédemment et en présence d'EDTA ajouté dans l'objectif d'obtenir la forme apo, ou de la protéine purifiée avec un protocole contenant de l'EDTA.

La forme orthorhombique ne correspond donc pas à une protéine apo, comme on avait pu le supposer à l'époque, mais à une protéine ayant fixé l'EDTA.
En ce qui concerne la forme hexagonale, étant données les connaissances de l'époque, nous avons supposé qu'il s'agissait de la forme apo de la protéine.

1.8.5 La mise en évidence de l'existence d'un métallophore spécifique du nickel

La présence d'un complexe de FeEDTA dans le site de fixation de NikA suggère fortement que le transport du nickel par NikA implique un métallophore spécifique du nickel, qui doit présenter de fortes similarités structurales avec l'EDTA. Le système de transport NikABCDE fonctionnerait alors de façon similaire à celui du fer par des sidérophores.
Cette hypothèse est confortée par le fait que, dans les deux cas, un système du type ABC est impliqué, et que comme le fer, le nickel se trouve dans l'environnement à faible concentration (de l'ordre de 10^{-9} M (Eitinger and Mandrand-Berthelot, 2000; Chen and Burne, 2003)). De plus, nous n'avons pas réussi à identifier un autre site dans la structure de NikA, susceptible de fixer le nickel seul.

L'existence d'un métallophore spécifique du nickel pourrait expliquer pourquoi la valeur du K_d du nickel à NikA est élevée par rapport à ce que l'on peut trouver dans le cas du fer, mais aussi pourquoi des valeurs très différentes sont trouvées dans la littérature : de 0,1 (de Pina *et al.*, 1995) à 10 µM (Salins *et al.*, 2002; Heddle *et al.*, 2003). Ces résultats sont très

difficiles à interpréter mais il est clair que si l'existence d'un métallophore n'a pas été prise en considération lors des expériences visant à mesurer la constante de dissociation du nickel pour NikA, les valeurs de K_d obtenues sont sûrement biaisées. De plus, la présence ou non d'EDTA lors de la purification de la protéine doit modifier le K_d apparent. Ceci pourrait d'ailleurs expliquer en partie le facteur 100 de différence entre le K_d trouvé par de Pina *et al.* (de Pina *et al.*, 1995) calculé sur une protéine n'ayant pas été purifiée en présence de nickel, et le K_d trouvé par Heddle *et al.* (Heddle *et al.*, 2003) calculé à partir d'une protéine purifiée en présence d'EDTA.

2 L'étude de la protéine NikA, purifiée selon le protocole contenant du chloroforme

L'étude de la protéine NikA purifiée selon le protocole de purification contenant de l'EDTA (Cherrier *et al.*, 2005) a permis de mettre en évidence que les cristaux orthorhombiques que nous avons reproduits, ainsi que ceux réalisés par Heddle *et al.*, contiennent une protéine possédant, dans son site de fixation du métal, une molécule d'EDTA, vraisemblablement située à la place d'un métallophore spécifique du nickel, jamais encore envisagé.

Nous avons donc décidé de produire NikA en absence totale d'EDTA, et d'étudier sa structure dans l'espoir de pouvoir découvrir la nature du nickelophore (métallophore spécifique du nickel).

2.1 La production de la protéine NikA, purifiée selon un protocole sans EDTA

2.1.1 La surexpression de la protéine en milieu riche

La surexpression de la protéine se fait de la même façon que précédemment, puisque les protocoles en milieu riche ne contiennent pas d'EDTA. Nous avons alors obtenu de l'ordre de 6 g de bactéries pour 2 L de culture, avec un bon niveau de surexpression.

2.1.2 La purification de la protéine selon le protocole contenant du chloroforme

2.1.2.1 L'extraction périplasmique au chloroforme et les précipitations au sulfate d'ammonium

La seule étape qui se réalisait en présence d'EDTA était l'extraction périplasmique. Nous avons donc remplacé cette étape par une extraction au chloroforme, comme c'était le cas pour le protocole utilisé lors des toutes premières études de NikA (de Pina *et al.*, 1995), qui avait permis d'obtenir des cristaux hexagonaux (Charon *et al.*, 1994).

Afin d'obtenir une quantité suffisante de protéine pour l'étape de cristallisation, nous avons réalisé ces purifications à partir de quantités importantes de bactéries (75 g de bactéries). Le problème qui s'est alors posé, par rapport aux cas de figures décrits dans la littérature (Ames *et al.*, 1984), a été la quantité de chloroforme à ajouter pour « solubiliser » les bactéries

(environ 2,5 mL de chloroforme par gramme de bactéries, soit ici 187,5 mL), mais surtout la quantité de Tris base 10 mM pH 7,5 (soit ici 1,875 L) à ajouter pour arrêter l'extraction (chapitre 1.2.2.3.1 du matériel et méthodes). Nous nous retrouvons donc, à la fin de cette étape d'extraction, avec un volume important de solution. On réalise alors après la première précipitation au sulfate d'ammonium à 40%, une seconde précipitation à 80% de sulfate d'ammonium permettant de concentrer la protéine.

Mais, l'efficacité de l'extraction périplasmique au chloroforme est moindre par rapport à celle effectuée à l'EDTA. Afin de tenter de pallier à ce problème, on laisse les bactéries sous agitation en présence du chloroforme jusqu'à 3 heures, pour extraire le maximum de protéine.

2.1.2.2 La colonne Q – Sépharose

Après extraction des protéines périplasmiques et précipitation au sulfate d'ammonium, les protéines sont chargées sur une colonne Q – Sépharose (Figure 2.1-1).

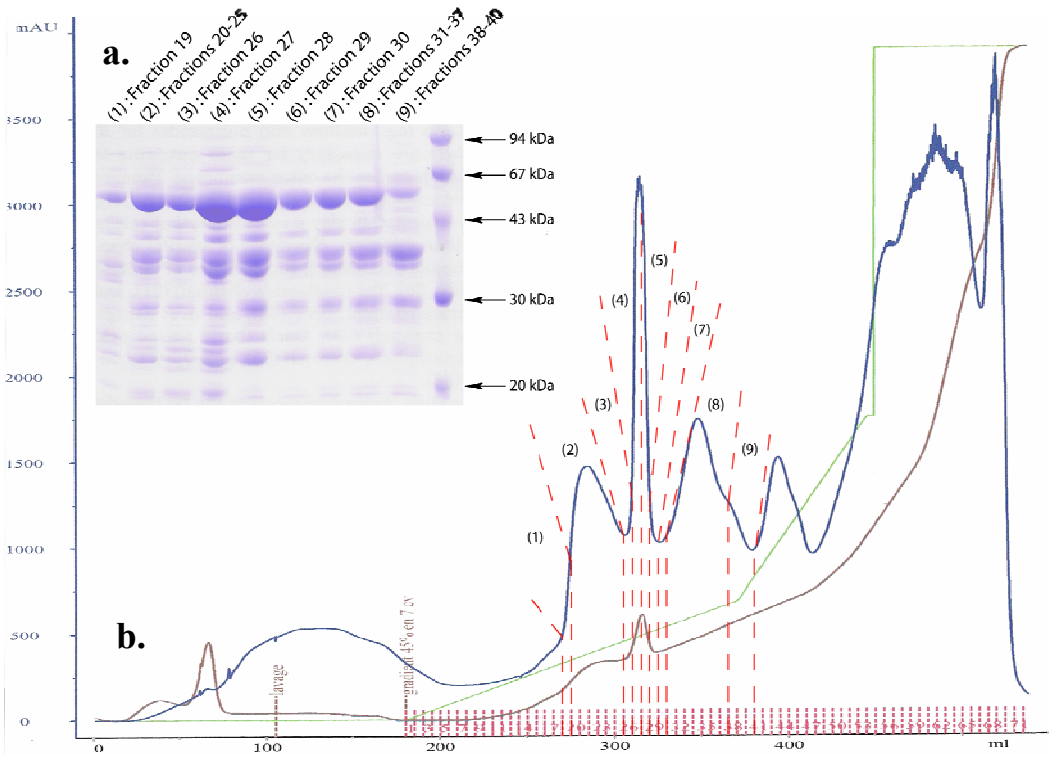


Figure 2.1-1 : Etape de purification sur colonne Q - Sépharose.

a. Gel SDS - PAGE des fractions purifiées sur colonne Q – Sépharose. Les fractions 19 à 37 sont gardées pour la suite de la purification.

b. Profil d'élution de la colonne Q - Sépharose avec un gradient continu de NaCl 1 M en 7 volumes de colonne.

En bleu : la D.O. à 280 nm. En vert : le gradient de NaCl. En marron : la conductivité.

NikA est alors éluée sur la colonne Q – Sépharose entre 13 et 16% de NaCl 1 M, soit 130 et 160 mM de NaCl.

2.1.2.3 La colonne Resource – Q

Après dialyse de la protéine contre un tampon Tris base 40 mM pH 7,5, celle-ci est chargée sur une colonne Resource – Q (Figure 2.1-2). NikA est alors éluée sur la colonne Resource – Q entre 6 et 9% de NaCl 1 M, soit 60 et 90 mM de NaCl.

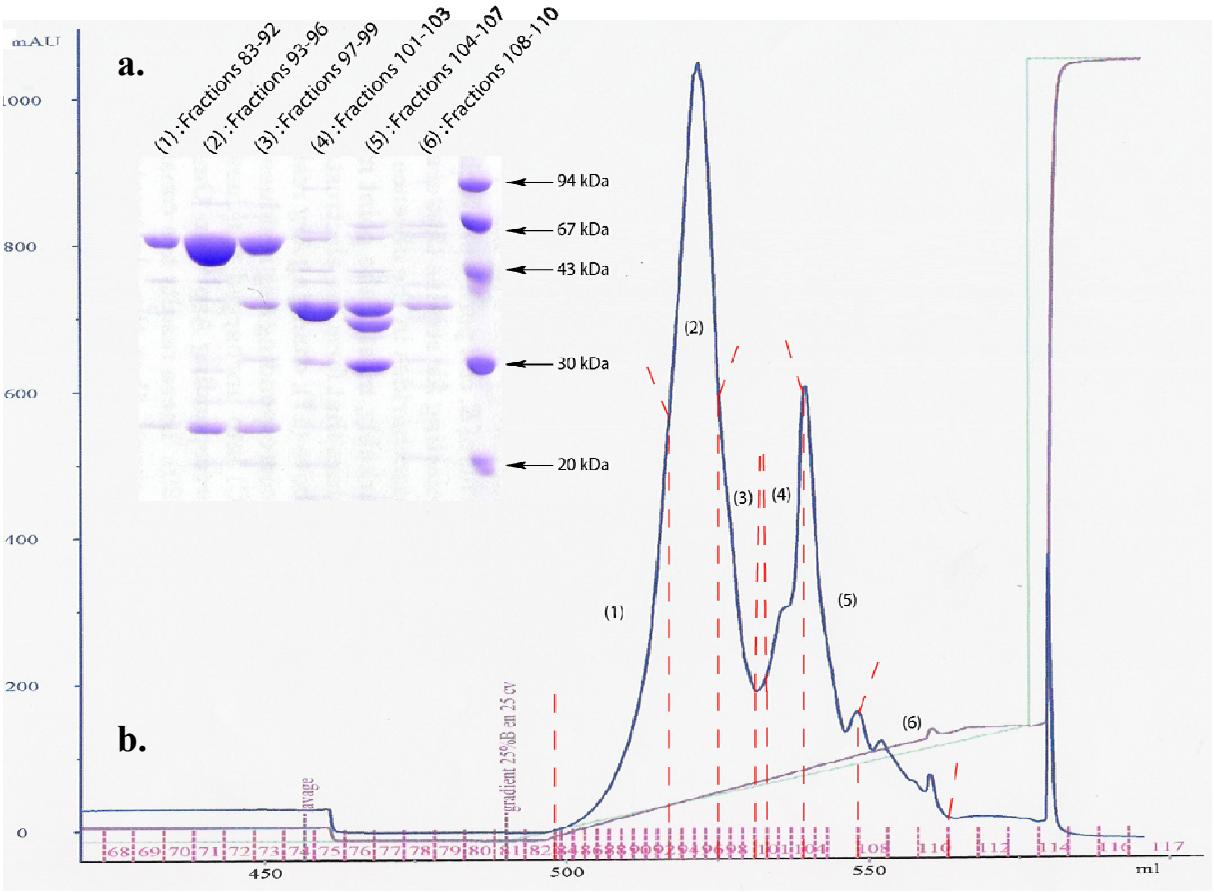


Figure 2.1-2 : Etape de purification sur colonne Resource – Q.

a. Gel SDS - PAGE des fractions purifiées sur colonne Resource – Q. Les fractions 83 à 99 sont gardées.

b. Profil d'élution de la colonne Resource – Q avec un gradient continu de NaCl 1 M en 25 volumes de colonne.

En bleu : la D.O. à 280 nm. En vert : le gradient de NaCl. En marron : la conductivité.

2.1.2.4 Le rendement de la purification

Après concentration de la protéine, on récupère 11,7 mL de solution de NikA à 3,6 mg/mL, soit 42,1 mg de protéine suffisamment pure pour être cristallisée. Le rendement

de la purification est donc de 56,1 mg de protéine pour 100 g de bactéries, ce qui est deux fois moins important que dans le cas de la purification selon le protocole contenant de l'EDTA. Cette différence de rendement provient en grande partie des pertes de protéine liées à l'utilisation de chloroforme, que nous avons décrite précédemment.

2.2 La cristallisation de la protéine

Nous espérions reproduire les cristaux hexagonaux de NikA (Charon *et al.*, 1994) grâce à cette protéine nouvellement purifiée. Pour cela, nous avons reproduit les conditions déterminées à l'époque :

[NikA] = 8 mg/mL ;
Sulfate d'ammonium entre 1,4 et 1,7 M ;
Acétate de sodium 0,1 M pH 4,7 ou 4,9 ;
20°C, en technique de goutte suspendue.

Des cristaux hexagonaux, de taille relativement petite, sont alors apparus dans les conditions testées, au bout de cinq à huit jours.

Afin d'améliorer la qualité et la taille des cristaux, nous avons décidé de changer le rapport protéine/solution réservoir de la goutte de cristallisation de la façon suivante : à 1 µL de solution réservoir, nous avons ajouté de 1 à 7 µL de protéine concentrée à 8 mg/mL. Les meilleurs résultats furent obtenus pour le rapport solution réservoir/protéine de 1 µL pour 4 µL (Figure 2.2-1).

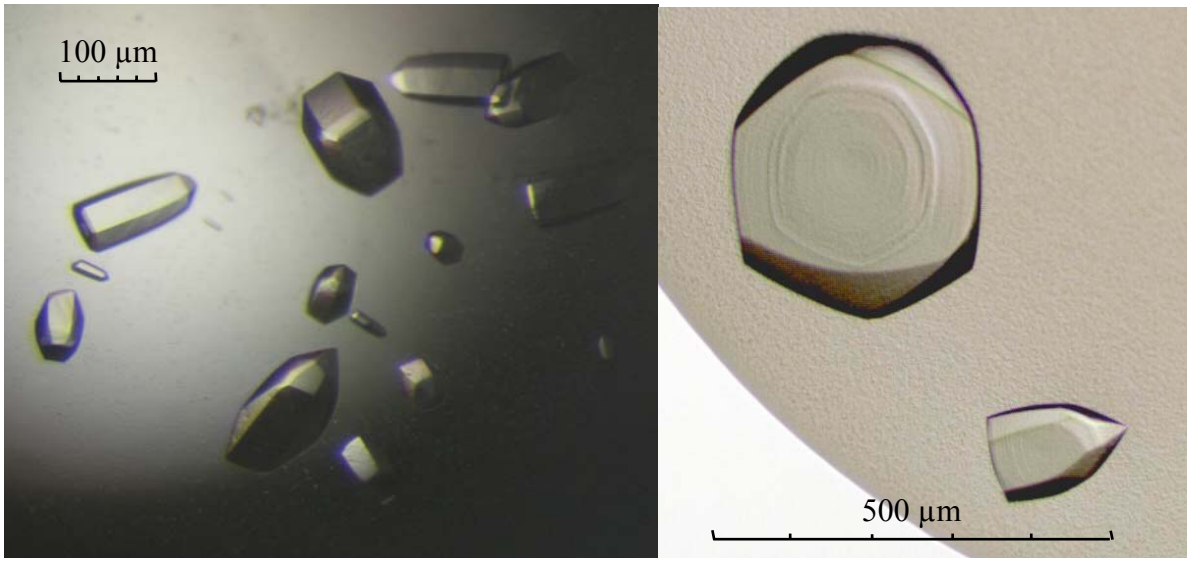


Figure 2.2-1 : Cristaux de la protéine NikA purifiée selon le protocole de purification contenant du chloroforme.

Les cristaux présentaient une grande diversité de taille, entre 10 et 500 µm, non corrélée à la concentration de l'agent précipitant (ici le sulfate d'ammonium).
Contrairement à la réalisation de cristaux orthorhombiques, aucun ensemencement n'a été nécessaire. Seule la forme cristalline hexagonale a pu être obtenue à partir de cette purification, même avec un micro-ensemencement à partir de cristaux appartenant au système orthorhombique.

2.3 L'enregistrement et le traitement des données de diffraction

Les cristaux de NikA sont cryoprotégés par un trempage rapide dans une solution composée de sulfate d'ammonium 2 M, glycérol 25%, avec le même tampon et les mêmes additifs que la solution de cristallisation.
Le meilleur cristal de la protéine NikA, purifiée selon le protocole de purification contenant du chloroforme, que nous ayons pu enregistrer sur la ligne de lumière ID14eh2 de l'ESRF, est apparu dans la condition de cristallisation suivante :

[NikA] = 8 mg/mL ;
Sulfate d'ammonium 1,6 M ;
Acétate de sodium 0,1 M pH 4,7 ;
Rapport solution réservoir/protéine : 1 µL pour 4 µL ;
20°C en goutte suspendue.

Ligne de lumière de l'ESRF	ID14eh2
Longueur d'onde	0,9330 Å
Groupe d'espace	$P6_2$
Paramètres de maille	a = b = 158,6 Å ; c = 134,9 Å ; α = 90° ; β = 90° ; γ = 120°
Nombre de molécules par unité asymétrique	3
Résolution	20,0–2,5 Å (2,6–2,5 Å)
Complétude	97,0% (92,7%)
I/σ(I)	18,6 (5,4)
Nombre total de réflexions mesurées	474 337 (50 477)
Redondance	3,7 (3,7)
R_{SYM}	4,6% (27,8%)

Tableau 2.3-1 : Statistiques des données de diffraction d'un jeu de la protéine NikA native, purifiée selon le protocole de purification contenant du chloroforme. Les valeurs entre parenthèses correspondent à la dernière tranche de résolution.

2.4 La résolution de la structure par remplacement moléculaire

Puisque les cristaux appartenant au système cristallin orthorhombique correspondaient à une protéine possédant de l'EDTA dans son site de fixation, nous nous attendions à ce que ceux appartenant au système cristallin hexagonal, qui n'avaient jamais été en contact avec de l'EDTA, correspondent à une protéine apo (sans métal) comme celle publiée par Heddle *et al.*

2.4.1 Les premières données de diffraction de NikA

Alors que nous reproduisions les cristaux hexagonaux de NikA, Frédérique Vellieux (du LBM de l'IBS) retrouva les données de diffraction qui avaient été collectées en 1994 à partir des premiers cristaux hexagonaux de NikA obtenus (Charon *et al.*, 1994).
Leur traitement montra qu'ils appartenaient bien au système cristallin $P6_2$ avec les paramètres de maille suivants : a = 160,3 Å ; b = 160,3 Å ; c = 138,4 Å ; α = 90° ; β = 90° ; γ = 120°. Malheureusement, les données se sont révélées très incomplètes puisque la complétude n'est que de 63,7%.

Elles nous ont tout de même permis de déterminer, par le calcul du coefficient de Matthews, qu'il y avait trois molécules par unité asymétrique.

Dans un premier temps, le remplacement moléculaire fut tenté en utilisant comme modèle la protéine apo publiée par *Heddle et al.* (code PDB : 1UIU), mais l'utilisation des programmes AMoRe et Phaser n'a pas permis d'obtenir une solution.

Nous avons donc décidé d'utiliser comme modèle la forme ayant fixé l'EDTA, et dans ce cas, le programme AMoRe donna une solution. On mit ainsi en évidence que les cristaux appartenant au système cristallin hexagonal ne correspondent pas à une protéine apo (ouverte), comme nous le suspections, mais à une forme fermée de NikA. Etant donnée la faible qualité de ces données de diffraction, nous n'avons pas continué leur étude.

2.4.2 Les nouvelles données de diffraction

Nous avons utilisé le jeu de données de diffraction allant jusqu'à une résolution de 2,5 Å décrit plus haut, afin de réaliser un remplacement moléculaire en utilisant comme modèle la structure résolue à partir des premières données de diffraction enregistrées (chapitre 2.4.1). Le programme Phaser permit de placer les trois molécules de l'unité asymétrique, et des cycles d'affinement par Refmac permirent l'obtention des statistiques du Tableau 2.4-1.

Facteur R et R_{free} (%)	21,7 / 28,4
Nombre d'atomes de la protéine	11 885
Nombre de molécules d'eau	235
Facteur B moyen (Å^2)	51,9
Rmsd	
Liaison (Å)	0,012
Angles (°)	1,481
Ramachandran	
Résidus dans la région la plus favorable (%)	88,5
Résidus dans la région favorable (%)	10,0
Résidus dans la région peu favorable (%)	0,8
Résidus dans la région interdite (%)	0,6

Tableau 2.4-1 : Statistiques d'affinement de la structure de NikA à 2,5 Å de résolution.

L'affinement de cette structure n'est pas totalement terminé, puisque certaines boucles de surface ne sont pas encore parfaitement placées. Les statistiques présentées dans le Tableau 2.4-1 ne sont donc pas des résultats définitifs.

L'empilement des cristaux hexagonaux a la particularité de favoriser la présence de canaux de solvants, ce qui réduit le nombre de liaisons entre les protéines qui assurent la cohésion du cristal.

Dans le cas de NikA, il y a 56,9% de solvant dans les cristaux (contre 37,6% de solvant dans les cristaux orthorhombiques), ce qui se traduit par un pouvoir de diffraction limité (on ne parvient pas à dépasser 2,5 Å de résolution) et par une agitation thermique globale plus importante (B_{moyen} = 51,9 $Å^2$ contre 26,1 $Å^2$ pour la structure obtenue à partir des cristaux appartenant au système cristallin orthorhombique ayant diffractés jusqu'à 2,2 Å de résolution).

Le facteur d'agitation thermique moyen est différent pour les trois molécules de l'unité asymétrique, et s'explique par leur position dans l'empilement cristallin (Figure 2.4-1) :

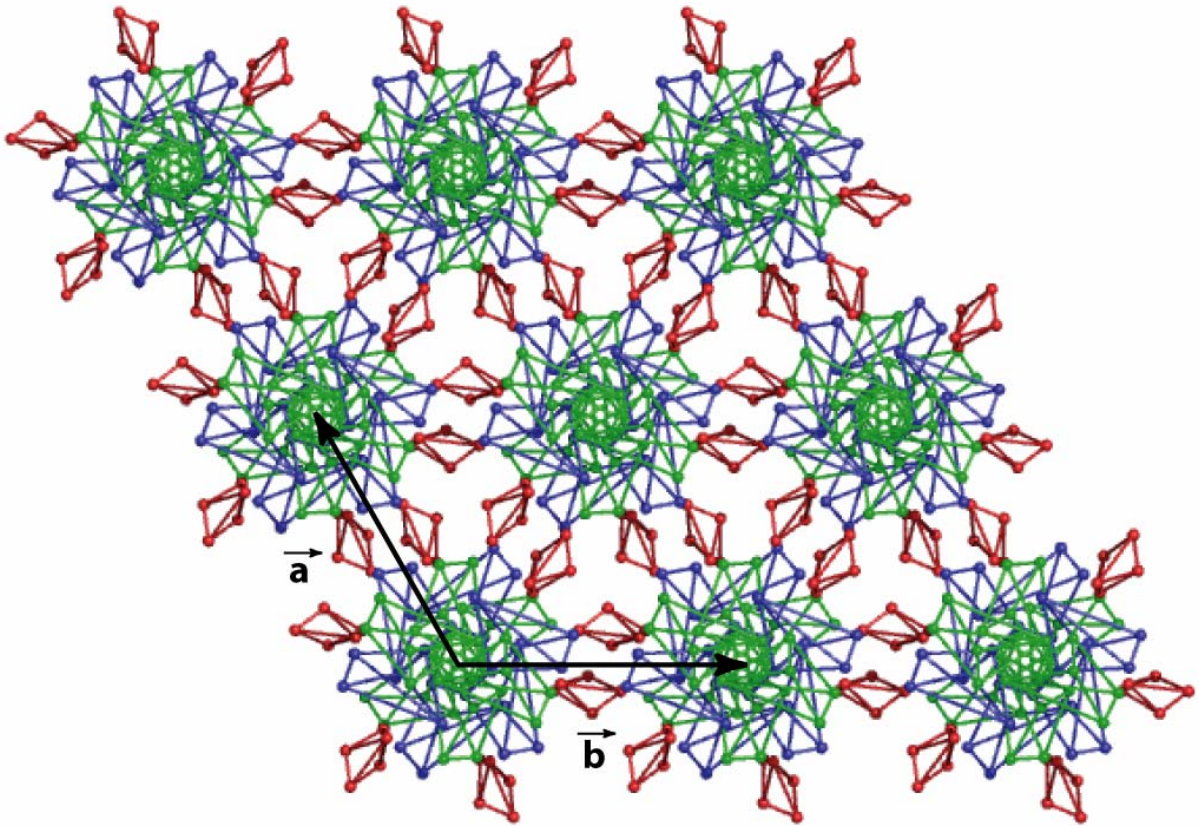


Figure 2.4-1 : Représentation de l'empilement cristallin dans la forme hexagonale des cristaux de NikA, purifiée selon le protocole contenant du chloroforme. Les protéines A B et C (respectivement en vert, bleu et rouge) sont représentées par cinq atomes définissant au mieux l'enveloppe de ces dernières.

- Les chaînes A et B ont un facteur B moyen respectivement de 48,5 et 46,8 $Å^2$. Ces molécules forment des « colonnes », séparées les unes des autres par des canaux de solvant, et établissant entre elles un nombre important de liaisons hydrogènes, ce qui les stabilise ;
- La chaîne C a un facteur B moyen de 60,5 $Å^2$, ce qui est bien supérieur aux deux autres molécules de l'unité asymétrique. Elle réalise des « ponts » entre les « colonnes » formées par les chaînes A et B. La chaîne C ne réalise donc que peu de liaisons hydrogènes avec les autres protéines, ce qui explique son facteur B moyen élevé, puisqu'elle est, de fait, peu stabilisée.

2.5 La comparaison de la structure hexagonale avec les structures orthorhombiques et apo

2.5.1 La comparaison des structures

Alors que l'on s'attendait à obtenir une structure ouverte de NikA (forme apo), nous avons été surpris de constater que la protéine purifiée en présence de chloroforme adopte une forme plus fermée que la protéine purifiée en présence d'EDTA (Figure 2.5-1).

En effet, nous pouvons estimer l'angle d'ouverture de la gorge, présente à l'interface des domaines I et II de NikA, à environ 26, 18 et 11° respectivement pour la forme apo, la forme de la protéine purifiée en présence d'EDTA et la forme purifiée en présence de chloroforme. Cette fermeture s'accompagne d'une augmentation du nombre de liaisons hydrogènes directes (sans implication d'une molécule d'eau) entre les deux domaines de la protéine : passant de 10 à 14 puis 20 liaisons hydrogènes respectivement pour la forme de la protéine apo (ouverte), la forme ayant fixé l'EDTA, et la forme purifiée en présence de chloroforme. Il s'agit de la seule différence significative entre les trois formes de NikA, puisque la comparaison de leurs domaines I et II pris séparément ne montre principalement que des mouvements de boucles de surface.

Les trois molécules de l'unité asymétrique n'étant pas équivalentes, il n'a pas été possible d'améliorer la qualité des cartes de densité électronique en utilisant la symétrie non cristallographique. La différence la plus importante concerne la molécule A qui, contrairement aux deux autres, adopte une conformation moins fermée, peut-être induite par l'empilement cristallin.

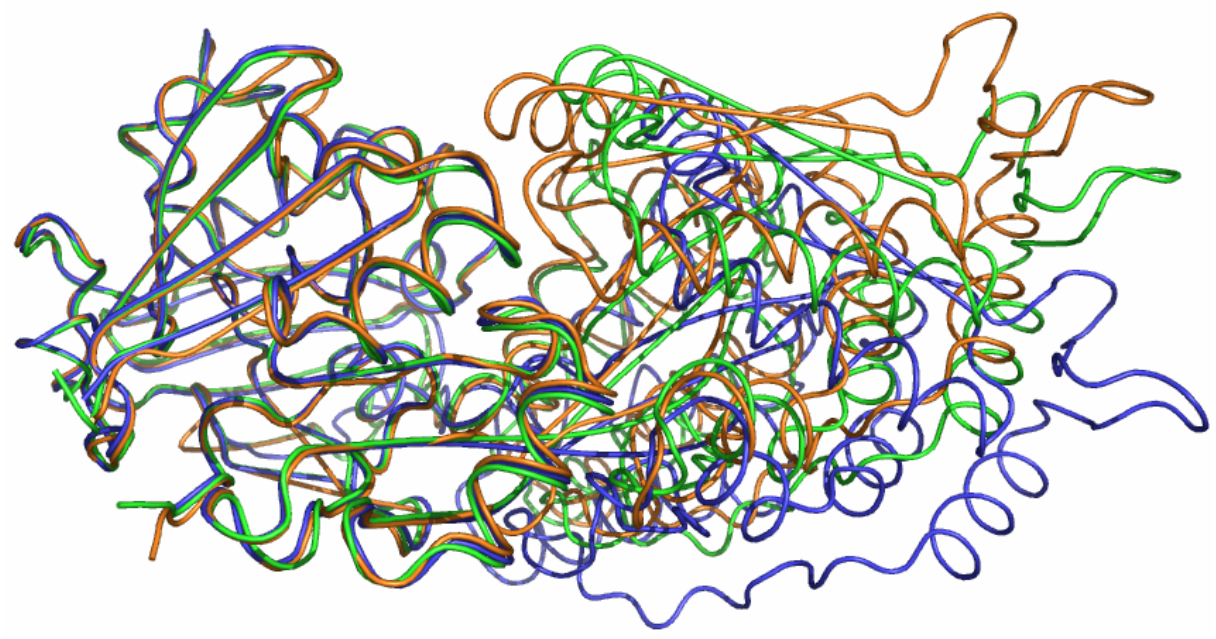


Figure 2.5-1 : Superposition (des domaines I) des structures de NikA apo (en bleu ; code PDB : 1UIU), NikA purifiée selon le protocole de purification contenant de l'EDTA (en vert : code PDB : 1ZLQ) et NikA purifiée selon le protocole de purification contenant du chloroforme (en orange).

2.5.2 Le contenu du site de fixation

Puisque les cristaux hexagonaux présentent une structure de NikA plus fermée que celles étudiées jusqu'à présent, nous avons décidé de comprendre l'origine de cette conformation en étudiant le contenu du site de fixation. Plusieurs éléments nous permettent d'affirmer que ce dernier contient un métallophore.

D'une part, nous avons calculé une carte de densité électronique « omit » à partir d'un modèle de la protéine possédant un site de fixation vide. Il nous montra que le site actif contient, comme dans le cas des cristaux orthorhombiques, un métal (21 σ dans la carte de densité électronique « omit ») entouré d'une densité électronique correspondant surement au métallophore dont nous avions supposé l'existence (Figure 2.5-2).

D'autre part, on constate un mouvement de l'histidine 416 et de l'hélice qui le contient, entre la forme de la protéine contenant de l'EDTA et cette nouvelle structure. En effet, pour pouvoir coordiner le métal du site de fixation, cet acide aminé se déplace respectivement de 4,3 et 8,6 Å par rapport la forme apo et la forme ayant fixé l'EDTA (Figure 2.5-3).

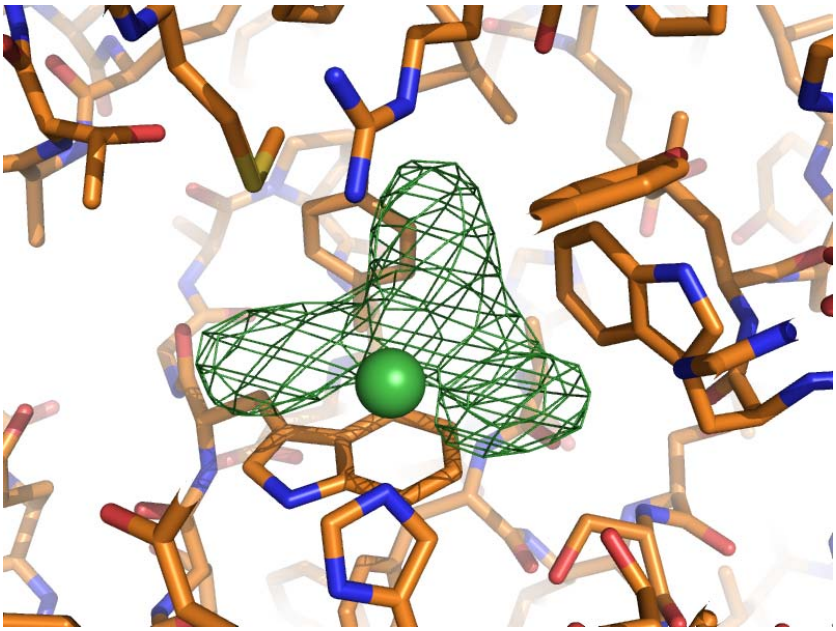


Figure 2.5-2 : Densité électronique « omit », dans le site de fixation de NikA purifiée selon le protocole contenant du chloroforme, calculée à partir d'un modèle de la protéine ne contenant que le métal (sphère verte) dans le site de fixation.

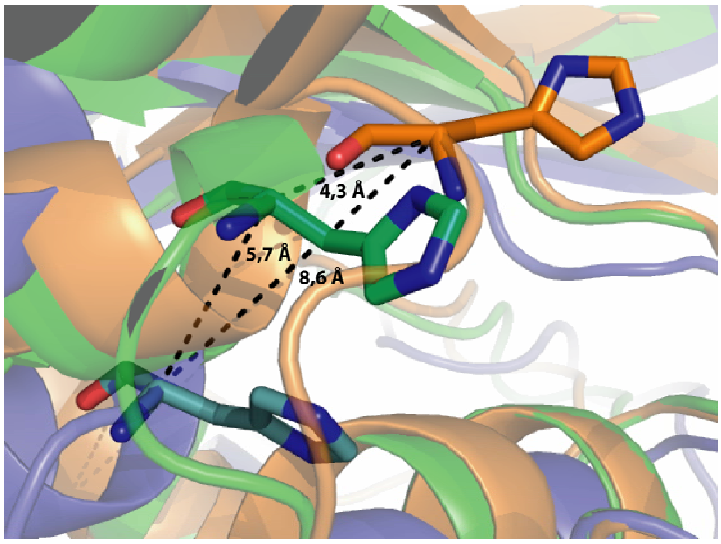


Figure 2.5-3 : Mouvement de l'histidine 416 entre les formes de NikA apo (en bleu), ayant fixé l'EDTA (en vert) et purifiée selon le protocole contenant du chloroforme (en orange).

L'histidine 416 est le seul résidu du site de fixation qui présente un changement de conformation et de position significatif par rapport à la forme ayant fixé l'EDTA (Figure 2.5-4). La position du métal est la même dans les deux cas.

Ces deux observations confortent notre hypothèse selon laquelle le site de NikA fixe un métallophore spécifique du nickel, possédant une structure similaire à celle de l'EDTA.

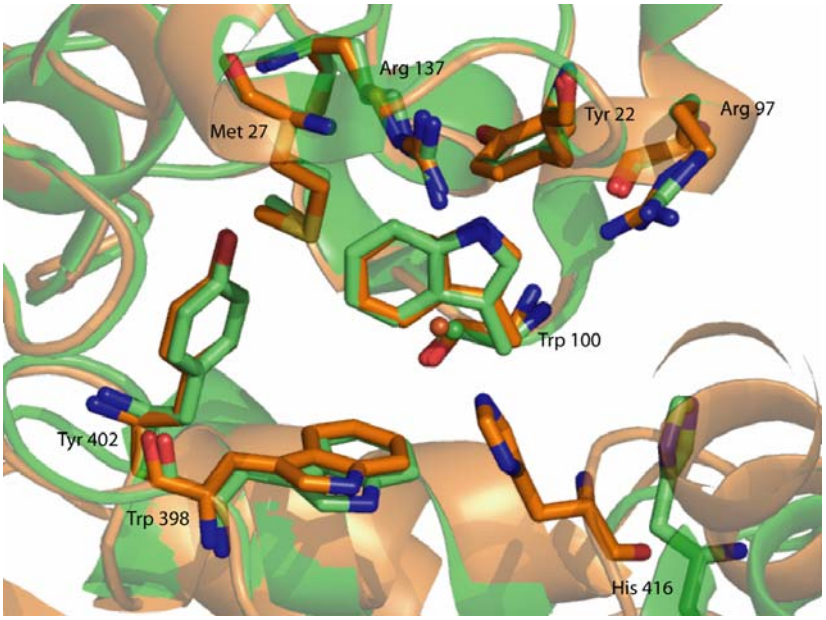


Figure 2.5-4 : Comparaison des résidus du site de fixation de la protéine ayant fixé l'EDTA (en vert) et la protéine nouvellement purifiée selon le protocole contenant du chloroforme (en orange). Le fer est représenté par une sphère rouge et le nickel par une sphère verte).

2.6 La caractérisation du métal présent dans le site de fixation de NikA

2.6.1 L'étude d'un cristal non mis en présence de nickel

2.6.1.1 L'analyse du contenu en atomes lourds des cristaux de NikA par fluorescence des rayons X

Nous avons voulu caractériser la nature du métal présent dans le site de fixation de NikA dans le cas de la protéine purifiée en absence d'EDTA.

Pour ce faire, nous avons analysé le contenu en atomes lourds des cristaux hexagonaux par spectroscopie d'absorption des rayons X à l'aide du détecteur de fluorescence (Roentec®) de la ligne BM30A (FIP) (Figure 2.6-1).

Nous avons ainsi remarqué qu'ils contiennent une grande proportion de zinc, mais aussi du fer et un pourcentage non négligeable de nickel.

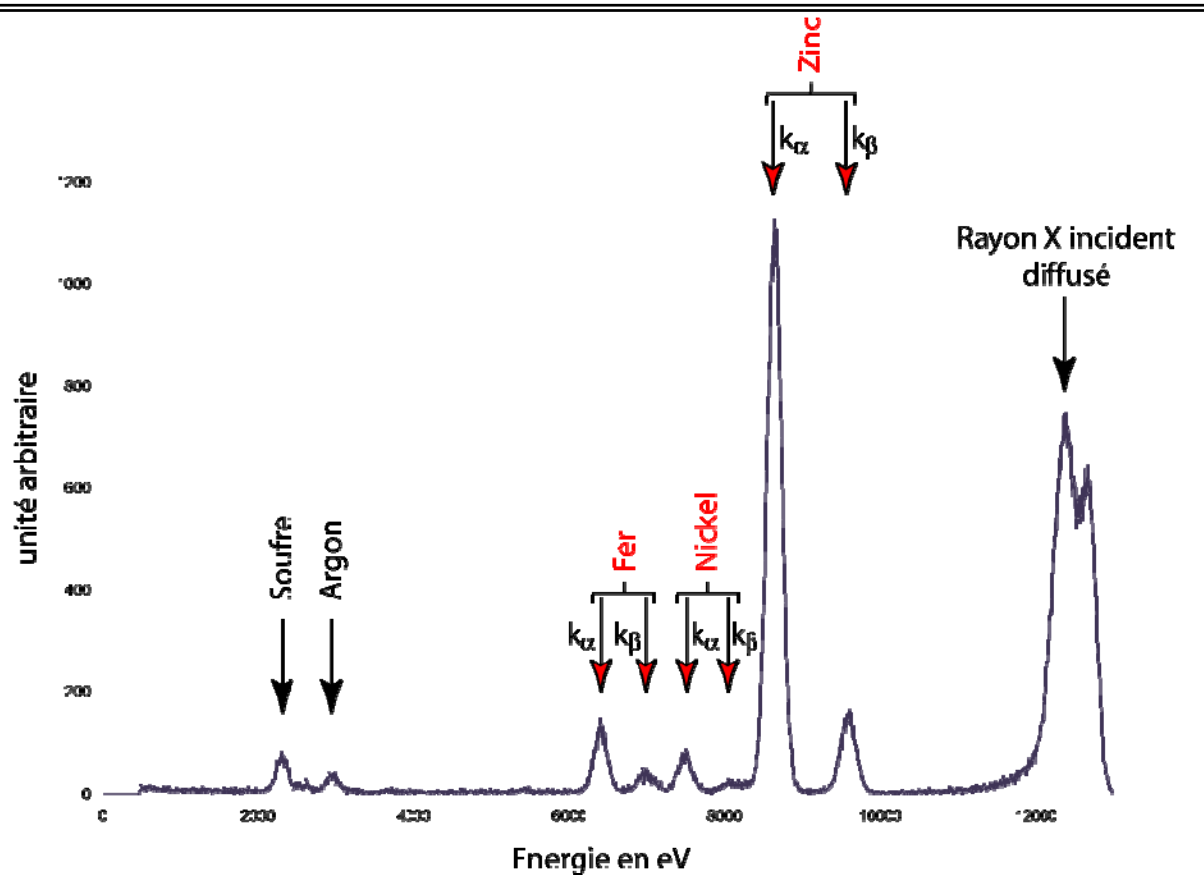


Figure 2.6-1 : Pics de fluorescence X mesurée sur des cristaux de NikA purifiée selon le protocole contenant du chloroforme.

2.6.1.2 L'enregistrement des données de diffraction

Afin de localiser ces trois métaux grâce à des cartes de double différence anomale, nous nous sommes servis du cristal sur lequel nous avions enregistré un jeu de données de diffraction allant jusqu'à 2,5 Å de résolution. Nous avons utilisé ces données de diffraction à longueur d'onde λ_4 = 0,9330 Å présentant la diffusion anomale du Fe, du Ni et du Zn. Pour compléter ces données, nous avons alors enregistré sur la ligne de lumière BM30A (FIP) de l'ESRF, les jeux de diffraction aux longueurs d'ondes permettant d'enregistrer la diffusion anomale :

- d'aucun des atomes suspectés d'être présents dans la structure, pour λ_1 = 1,7452 Å soit 7105,0 eV (légèrement inférieur en énergie au seuil d'absorption K du Fe) ;
- du Fe, pour λ_2 = 1,7394 Å soit 7128,0 eV (légèrement supérieur en énergie au seuil d'absorption K du Fe) ;
- du Fe et du Ni, pour λ_3 = 1,4852 Å soit 8348,0 eV (légèrement supérieur en énergie seuil d'absorption K du Ni).

L'indexation permit d'obtenir les statistiques présentées dans le Tableau 2.6-1.

	λ_1	λ_2	λ_3	λ_4
Ligne de lumière de l'ESRF	BM30A (FIP)			ID14eh2
Longueur d'onde	1,7450 Å	1,7394 Å	1,4852 Å	0,9330 Å
Groupe d'espace	$P6_2$			
Paramètres de maille	a = b = 158,6 Å ; c = 134,9 Å ; α = 90° ; β = 90° ; γ = 120°			
Nombre de molécules par unité asymétrique	3			
Résolution	20,0–3,4 Å (3,3–3,4 Å)	20,0–3,4 Å (3,5–3,4 Å)	20,0–3,0 Å (3,1–3,0 Å)	20,0–2,5 Å (2,6–2,5 Å)
Complétude	94,7% (89,5%)	95,4% (89,9%)	95,5% (89,2%)	97,0% (92,7%)
I/σ(I)	9,8 (3,7)	11,6 (3,9)	17,13 (3,9)	18,6 (5,4)
Nombre total de réflexions mesurées	121 779 (9 350)	146 036 (11 199)	212 559 (18 425)	474 337 (50 477)
Redondance	2,4 (2,4)	2,9 (2,8)	2,9 (2,9)	3,7 (3,7)
R_{SYM}	8,9% (29,1%)	9,2% (30,6%)	5,0% (31,1%)	4,6% (27,8%)

Tableau 2.6-1 : Statistiques des données de diffraction ayant permis le calcul des cartes de différence anomale à partir de la protéine NikA native purifiée selon le protocole contenant du chloroforme. Les valeurs entre parenthèses correspondent à la dernière tranche de résolution. Les statistiques de la longueur d'onde λ_4 sont données à titre de comparaison et correspondent à celles du Tableau 2.4-1.

2.6.1.3 Le calcul des cartes différence anomale

Les cartes de densité électronique différence anomale ont été calculées pour chaque longueur d'onde à partir du modèle affiné avec le jeu de données de diffraction à 2,5 Å de résolution (λ_4 ; affiné par Refmac : R = 21,7% ; R_{free} = 28,4%).

La situation est plus compliquée que dans le cas des cristaux orthorhombiques pour plusieurs raisons :

- La mauvaise qualité des données rend difficile la caractérisation des pics de diffusion anomale, puisque qu'ils ne dépassent que de quelques σ le bruit de fond de la carte ;
- Les trois métaux (fer, nickel et zinc) sont retrouvés dans les sites de fixation ;
- Les sites de fixation des trois molécules de l'unité asymétrique ne contiennent pas le même rapport zinc/nickel/fer.

Ces mêmes raisons ont empêché le calcul de cartes de densité électronique de double différence anomale facilement interprétables. L'étude de la composition métallique s'est donc faite à partir des cartes de densité électronique de différence anomale uniquement.

Cette expérience permet tout de même de montrer que la protéine purifiée selon le protocole contenant du chloroforme, possède dans son site de fixation un métal, qui peut être ici un zinc, un nickel ou un fer. Le fait de retrouver une proportion non négligeable de nickel dans la protéine suggère que le métallophore présente une haute affinité pour le nickel, puisque ce dernier ne se retrouve qu'à l'état de traces dans les tampons (inférieur à 0,0005% du tampon dans le cas du Tris Base) et que dans le cas de l'EDTA il n'était pas présent. De plus, NikA étant surexprimée, la quantité de nickel disponible doit être insuffisante par rapport à la quantité de protéine.

On peut tout de même essayer d'estimer la proportion de chaque métal présent dans chacune des molécules de l'unité asymétrique.

Pour cela, nous avons mesuré la hauteur des pics de diffusion anomale présents dans les sites de fixation des trois molécules de l'unité asymétrique, pour chacune des quatre longueurs d'ondes enregistrées. Puis, nous avons soustrait les valeurs obtenues afin d'estimer la contribution de chaque métal à la hauteur du pic (Tableau 2.6-2).

Molécule de l'unité asymétrique	Maximum du pic de diffusion anomale				Proportion du métal		
	λ_1	λ_2	λ_3	λ_4	Fe (λ_2-λ_1)	Ni (λ_3-λ_2)	Zn (λ_4-λ_3)
A	2,25σ	4,25σ	5,25σ	6,75σ	2,00σ	1,00σ	1,50σ
B	3,00σ	7,50σ	9,25σ	13,00σ	4,50σ	1,75σ	3,75σ
C	3,75σ	3,75σ	8,25σ	11,00σ	0,00σ	4,50σ	2,75σ

Tableau 2.6-2 : Estimation de la proportion de zinc, nickel et fer présents dans les trois molécules de l'unité asymétrique.

Mais les résultats ne doivent être considérés que de façon qualitative en raison des approximations qu'impliquent cette méthode (la diffusion anomale des métaux est considérée équivalente ; on raisonne uniquement sur la hauteur du pic de diffusion anomale au lieu d'en intégrer le volume …).

Contrairement à ce qui est observé pour les molécules B et C, la molécule A contient un site métallique ayant une faible occupation, ce qui est corrélé au fait que cette protéine adopte une conformation plus ouverte que les deux autres. De cette façon, l'histidine 416 n'est plus en position pour compléter la coordination du métal, et ce dernier s'en trouve donc plus labile.

2.6.2 L'étude d'un cristal mis en présence de nickel

2.6.2.1 L'analyse du contenu en atomes lourds des cristaux de NikA par fluorescence des rayons X

Après ces premiers résultats, nous avons réalisé le même type d'expérience sur un cristal apparu dans une goutte de cristallisation contenant un excès de nickel.

L'étude par fluorescence des rayons X montra qu'il y avait dans le cristal une grande proportion de nickel, un peu de fer mais pas de zinc.

2.6.2.2 L'enregistrement des données de diffraction

Bien que ce cristal soit relativement mosaïque et qu'il présente une mauvaise qualité de diffraction, nous avons enregistré les trois longueurs d'onde nécessaires à la caractérisation du métal présent dans le site de fixation (Fe et/ou Ni).

L'indexation permit d'obtenir les statistiques du Tableau 2.6-3. La qualité moyenne des données de diffraction est confirmée par les valeurs du R_{SYM}.

	λ_1	λ_2	λ_3
Ligne de lumière de l'ESRF	BM30A (FIP)		
Longueur d'onde	1,7450 Å	1,7389 Å	1,4848 Å
Groupe d'espace	$P6_2$		
Paramètres de maille	a = b = 157,5 Å ; c = 134,9 Å ; α = 90° ; β = 90° ; γ = 120°		
Nombre de molécules par unité asymétrique	3		
Résolution	50,0–3,5 Å (3,50–3,52 Å)	50,0–3,88 Å (3,90–3,88 Å)	50,0–4,38 Å (4,40–4,38 Å)
Complétude	94,6% (92,0%)	94,9% (93,8%)	95,5% (92,0%)
I/σ(I)	8,4 (3,5)	8,8 (4,5)	9,1 (4,7)
Nombre total de réflexions mesurées	130 983 (2 066)	96 930 (1 434)	68 692 (1 038)
Redondance	3,0 (2,9)	2,9 (2,8)	3,1 (3,0)
R_{SYM}	10,8% (31,4%)	11,2% (29,6%)	11,8% (30,3%)

Tableau 2.6-3 : Statistiques des données de diffraction ayant permis le calcul des cartes de différence anomale à partir de la protéine NikA native purifiée selon le protocole contenant du chloroforme, et cristallisée en présence d'un excès de nickel. Les valeurs entre parenthèses correspondent à la dernière tranche de résolution.

2.6.2.3 Le calcul des cartes différence anomale

L'analyse du contenu des sites de fixation se fit à partir des cartes de densité électronique de double différence anomale calculées grâce de la structure résolue à partir des données de diffraction enregistrées à la longueur d'onde λ_1 (affiné par Refmac : R = 14,5% ; R_{free} = 28,0%).

Malgré la faible qualité des données de diffraction, les cartes montrent que le nickel occupe les sites de fixation des molécules de l'unité asymétrique (Tableau 2.6-4). Il n'y a plus de zinc, et le fer est minoritaire. Par contre, comme c'est le cas dans tous les cristaux que nous

avons étudiés, la molécule A adopte toujours une forme plus ouverte que les deux autres protéines, et le nickel y est présent avec une faible occupation.
Ce résultat conforte l'hypothèse selon laquelle le métallophore est spécifique du nickel : lorsqu'il y a assez nickel par rapport à la quantité de protéine, ce dernier est capable de se substituer aux métaux déjà présents dans le site de fixation.

Molécule de l'unité asymétrique	Maximum du pic de diffusion anomale			Proportion du métal	
	λ_1	λ_2	λ_3	Fe (λ_2-λ_1)	Ni (λ_3-λ_2)
A	2,00σ	2,00σ	3,50σ	0,00σ	1,50σ
B	4,25σ	4,50σ	8,00σ	0,25σ	3,50σ
C	2,25σ	2,50σ	6,25σ	0,25σ	3,75σ

Tableau 2.6-4 : Estimation de la proportion de nickel et fer présents dans les trois molécules de l'unité asymétrique d'un cristal ayant poussé en présence de nickel.

2.7 La caractérisation du métallophore présent dans le site de fixation de NikA

Après avoir remarqué que, lorsque NikA est purifiée en absence d'EDTA, elle contient dans son site de fixation une molécule présentant une forte affinité pour le nickel (puisqu'elle est capable de fixer ce métal alors qu'il n'est présent dans le milieu qu'à faible concentration), nous avons souhaité mieux caractériser ce métallophore.

2.7.1 La cristallographie

2.7.1.1 La densité électronique

Dans un premier temps, nous avons tenté de déterminer la nature du métallophore à partir des données de diffractions que nous avions collectées jusqu'à une résolution de 2,5 Å.
Pour cela, nous avons calculé une carte de densité électronique « omit » à partir d'un modèle de la protéine ne contenant pas le métallophore (Figure 2.5-2). Cette étude s'est faite sur la molécule B de l'unité asymétrique car :

- La molécule A adopte une conformation plus ouverte que les deux autres molécules de l'unité asymétrique. De ce fait, le métal est labile et la densité électronique « omit » du métallophore y est mal définie ;

- La molécule C est plus agitée que les deux autres, ce qui ne permet pas d'obtenir une carte de densité électronique « omit » du métallophore de bonne qualité.

Nous avons souhaité employer la même technique que dans le cas de l'identification de l'EDTA, à savoir définir et placer une molécule dans la densité électronique « omit », et après dix cycles d'affinement par Refmac la valider ou non en tant que métallophore, sur la base de la géométrie du complexe formé avec le métal, et aussi de la densité résiduelle (positive et négative) de la carte de densité électronique F_{obs}-F_{calc}.

Mais, d'après la résolution et la qualité des données, inférieure à celle nous ayant permis d'identifier l'EDTA, l'étude de la densité électronique résiduelle F_{obs}-F_{calc} ne nous a apporté que peu d'informations. En effet, pour un grand nombre de molécules ayant des structures proches (comme par exemple différant par l'ajout d'un carbone) aucune densité résiduelle significative n'est visible dans la carte de densité électronique F_{obs}-F_{calc}. Afin de déterminer la structure du métallophore, nous n'avons donc pu nous baser que sur la validité de la géométrie du complexe métallophore – métal et sa corrélation avec la carte de densité électronique « omit ».

Pour toutes les molécules que nous avons étudiées, nous avons veillé à respecter la géométrie plan carré que semble adopter le métal dans le site de fixation, ce qui doit assurer sa spécificité pour le nickel. Trois des liaisons sont apportées par le métallophore, et la quatrième est assurée par l'histidine 416. Au vu de la forme de la densité électronique « omit » et de la comparaison avec le complexe de fer – EDTA, il nous paraissait très probable que la fixation du métal par le métallophore soit assurée par trois carboxylates.

Parmi toutes les molécules testées, c'est l'acide 3-(carboxy)hexanedioïque (aussi appelé acide butane-1,2,4-tricarboxylique) qui présenta la meilleure corrélation avec la carte de densité électronique « omit » et respectait la géométrie plan carré du complexe (Figure 2.7-1).

Il apparaît que les mêmes résidus sont impliqués dans la fixation du métallophore à la protéine et dans la fixation de l'EDTA à la protéine (Figure 2.7-1et Figure 2.7-2a) :

- Deux des carboxylates de la molécule forment trois liaisons hydrogènes avec les arginines 97 et 137 ;
- S'y ajoute une molécule d'eau pontant le métallophore à la protéine.

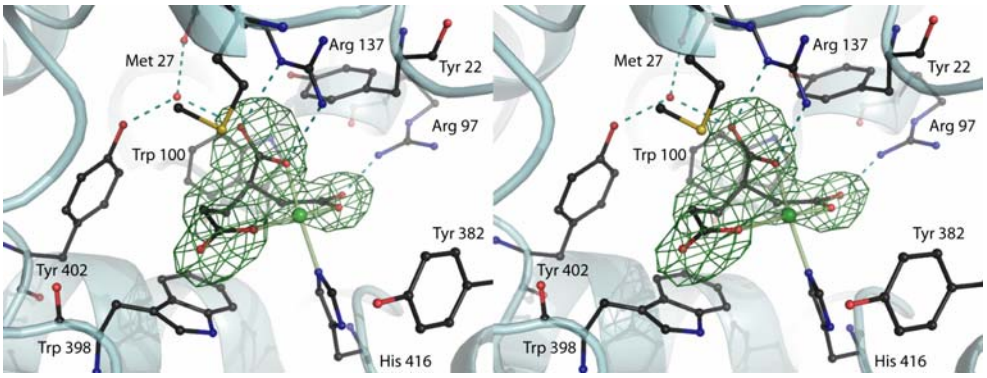


Figure 2.7-1 : Vue stéréoscopique du site de fixation et des liaisons hydrogènes entre l'acide 3-(carboxy)hexanedioïque et NikA. La densité électronique « omit » (contourée à 3σ) calculée à partir d'un modèle de protéine ne contenant pas de métallophore est représentée en vert. Le métal (ici le nickel) et les molécules d'eau sont représentés respectivement par des sphères vertes et rouges. Les liaisons hydrogènes sont représentées par les trais en pointillés.

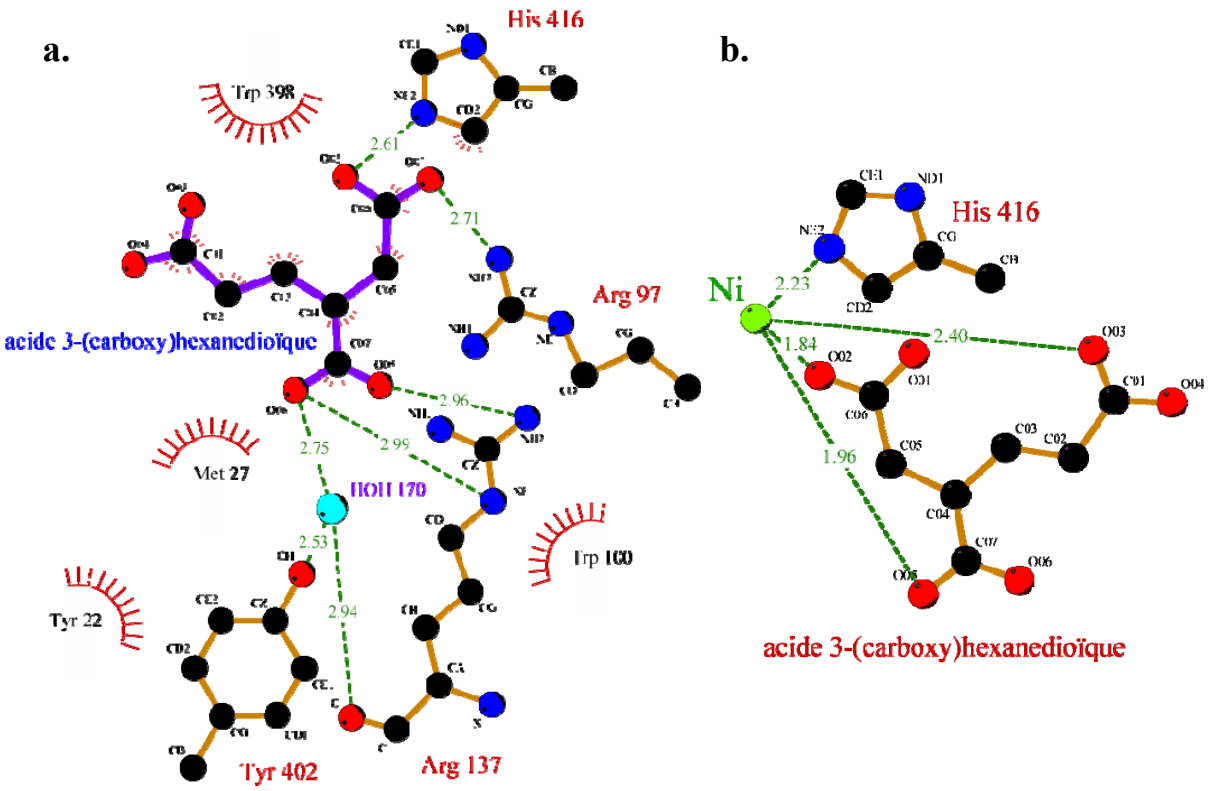


Figure 2.7-2 : Figures générées par le programme LIGPLOT (Wallace *et al.*, 1995) représentant :
a. Les liaison hydrogènes entre l'acide 3-(carboxy)hexanedioïque et la protéine.
b. La coordination du nickel par l'acide 3-(carboxy)hexanedioïque et l'histidine 416.
Les liaisons sont représentées par des pointillés verts, et les distances sont en Å.

Il semble s'établir la même interaction cation-π impliquant le Trp398 entre le métal et la protéine, que dans le cas du complexe FeEDTA (Figure 2.7-1). Mais il est également possible que la Tyr382 réalise aussi une liaison de ce type (la distance nickel-Tyr382 est 6,1 Å, et l'angle entre la perpendiculaire au plan des cycles du Trp398 et le nickel est de 18°).
Par contre les distances entre le nickel et ses coordinants ne sont pas tout à fait en accord avec une géométrie plan carré : elles devraient être en moyenne de l'ordre de 2,0 – 2,1 Å (Figure 2.7-2b).
L'acide 3-(carboxy)hexanedioïque est donc un bon candidat pour être le métallophore dont nous avions suspecté l'existence. Mais, pour pouvoir valider ce modèle il nous faut caractériser la structure du métallophore de façon certaine, ce qui est impossible ici par cristallographie à cause de la résolution limitée de nos données de diffraction.

2.7.1.2 La tentative d'amélioration de la résolution

Nous avons alors tenté d'améliorer la limite de diffraction des cristaux de NikA appartenant au système cristallin hexagonal en essayant différentes techniques :

- nous avons modifié les conditions de cristallisation en faisant varier la concentration de l'agent précipitant (le sulfate d'ammonium), le pH et la nature du tampon, la température ;
- nous avons testé les trois *screens* additifs de la société Hampton ;
- nous avons purifié et cristallisé la protéine en présence de nickel, avec l'espoir d'obtenir une population homogène de NikA dans les cristaux, et ainsi d'augmenter la limite de résolution des cristaux, ou à défaut de pouvoir utiliser la symétrie non cristallographique pour améliorer les cartes de densité électronique ;
- nous avons essayé les techniques d'*annealing* (qui consiste à décongeler et recongeler rapidement le cristal pour permettre à l'empilement cristallin de se réorganiser) et les techniques de déshydratation des cristaux (Heras *et al.*, 2003; Abergel, 2004; Heras and Martin, 2005).

Malheureusement aucune de ces techniques n'a permis d'obtenir des cristaux diffractant à une meilleure résolution que 2,5 Å.

Cette limite de résolution doit être liée à l'empilement cristallin qui présente d'importants canaux de solvants et un nombre limité de contacts cristallins entre les molécules (notamment en ce qui concerne la molécule C).

Pour obtenir une structure de NikA en complexe avec le métallophore à plus haute résolution, il conviendrait donc plutôt d'obtenir des cristaux appartenant à une nouvelle forme cristalline.

2.7.2 La spectroscopie d'absorption des rayons X (SAX)

Afin d'étudier de façon plus précise la coordination du nickel dans la protéine en solution (type de coordination et distances), nous avons décidé de réaliser une étude de spectroscopie d'absorption des rayons X de la protéine en complexe avec le métallophore.

2.7.2.1 Le XANES

Pour caractériser au mieux la géométrie que le nickel forme avec son métallophore, nous avons tenté d'enregistrer les spectres d'absorption des rayons X pour deux échantillons :

- Au seuil du fer :
  - ⇒ La protéine NikA purifiée selon le protocole de purification contenant du chloroforme, possédant donc, dans le site de fixation, le métallophore et du fer.
- Au seuil du nickel (Figure 2.7-3) :
  - ⇒ La protéine NikA purifiée selon le protocole de purification contenant du chloroforme, possédant donc, dans le site de fixation, le métallophore et du nickel ;
  - ⇒ Un complexe de NiEDTA adoptant une géométrie octaédrique.

L'enregistrement des spectres SAX de la protéine purifiée en présence de chloroforme n'a posé aucun problème au seuil du nickel, mais par contre n'a pas été possible à celui du fer, à cause d'une proportion de ce dernier dans la protéine trop faible.

On peut ainsi estimer qu'il y a environ 80% de nickel contre seulement 20% de fer en comparant l'intensité du signal mesurée dans les deux cas. Nous avons donc bien, en solution comme dans les cristaux, du nickel et du fer, même si les proportions mesurées par la SAX ne correspondent pas tout à fait à celles estimées par diffusion anomale (chapitre 2.6.1). De plus, la cristallisation peut, elle aussi, induire une différence de proportion entre le nickel et le fer, en imposant la conformation de la protéine par l'empilement cristallin.

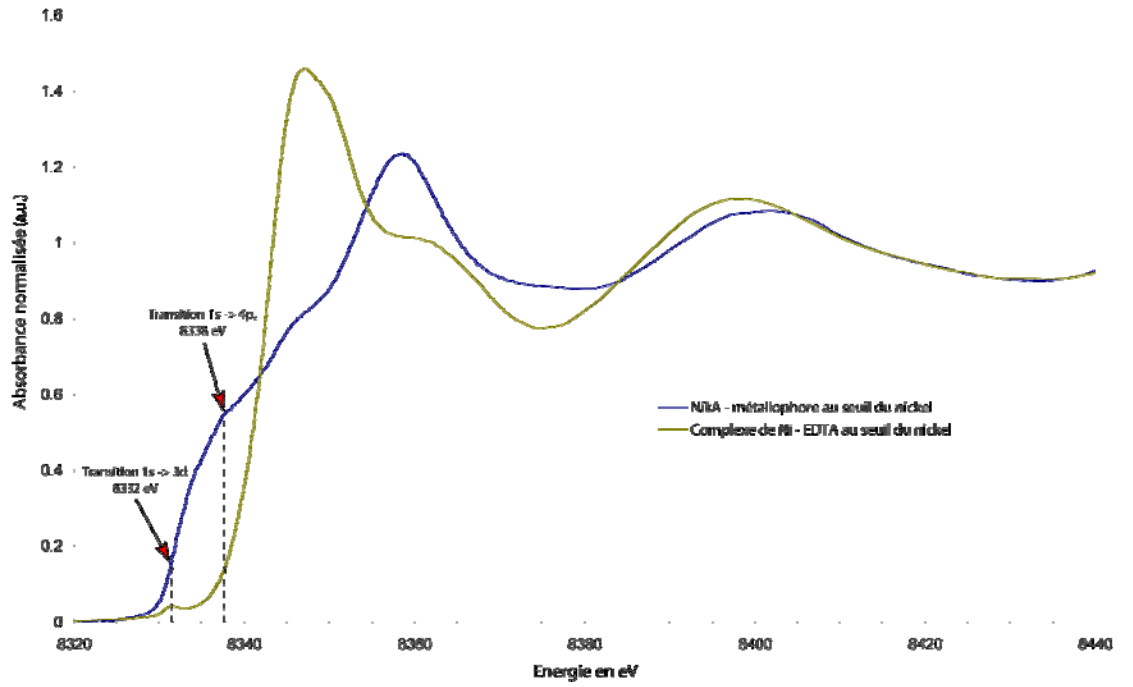


Figure 2.7-3 : XANES de la protéine NikA purifiée selon le protocole de purification contenant du chloroforme, du nickel (en bleu), et un complexe de NiEDTA en solution, enregistré au seuil du nickel (en vert).

Afin d'estimer avec une meilleure précision la quantité de nickel par rapport à celle du fer, il faudrait réaliser une analyse élémentaire (par spectroscopie de flamme) d'un échantillon protéique, qui permet d'avoir une estimation précise de la composition atomique d'un échantillon. Néanmoins, ce résultat conforte l'hypothèse selon laquelle le métallophore est spécifique du nickel.

Pour caractériser la géométrie du nickel en complexe avec l'EDTA ou dans la protéine, nous avons étudié les transitions 1s → 3d et 1s → $4p_z$ responsables des pics du XANES, présentes respectivement à environ 8332 et 8338 eV, puisque leur intensité est caractéristique de la coordination du métal (Colpas *et al.*, 1990) (Tableau 2.7-1).

Dans le cas du complexe de Ni-EDTA, la géométrie que l'on détermine grâce à l'étude du XANES est en accord avec celle observée dans la structure cristallographique, à savoir une géométrie octaédrique.

En ce qui concerne la coordination adoptée par le nickel dans la protéine, la présence de la transition 1s → $4p_z$ à 8338 eV nous indique que le nickel ne peut adopter que deux géométries possibles : pyramide à base carrée ou plan carré.

La transition 1s → 3d à 8332 eV est ici difficile à mettre en évidence du fait de la largeur du pic correspondant à la transition 1s → $4p_z$, mais doit tout de même être présente.
Dans le cas d'une géométrie plan carré, ce résultat est en accord avec ceux précédemment obtenus en cristallographie, mais aussi dans le cas de la géométrie pyramide à base carrée car il est tout à fait envisageable que le cinquième ligand soit agité et invisible en cristallographie, ou qu'il ne soit présent qu'en solution et non dans le cristal à cause des contraintes imposées par l'empilement cristallin.

Coordination	**Géométrie**	**Transition 1s → 3d**	**Transition 1s → $4p_z$**
6	octaédrique	✓ (faible)	∅
5	bipyramide trigonale	✓	∅
	pyramide à base carrée	✓	✓ (épaulement)
4	plan carré	✓ (faible)	✓
	tétraédrique	✓	∅

Tableau 2.7-1 : Transitions 1s → 3d et 1s → $4p_z$ visibles dans le spectre SAX en fonction de la géométrie du nickel (Colpas *et al.*, 1990).

2.7.2.2 L'EXAFS

2.7.2.2.1 Le complexe de nickel-EDTA

Nous nous sommes servis du complexe de NiEDTA comme référence puisque son spectre a été enregistré dans les mêmes conditions que le spectre de la protéine contenant le métallophore enregistré au seuil du nickel, qui constitue l'objet de l'étude EXAFS.
Connaissant sa structure cristallographique, et par conséquent la nature des ligands du métal et les distance nickel – EDTA, nous avons pu fixer le paramètre structural N_i, et ainsi calibrer les paramètres électroniques ΔE_0 et S_0^2 (Tableau 2.7-2) qui serons utilisés par la suite, et également déterminer les valeurs du paramètre structural R_i et du terme de Debye-Waller σ_i^2.

Liaison	Paramètres structuraux			Paramètres physiques		Résiduel
	R_i (Å)	N_i	σ_i^2 (Å^2)	ΔE_0 (eV)	S_0^2	
Ni - O/N	2,05	6	$7{,}07.10^{-3}$	-3,04	1,09	$4{,}62.10^{-2}$

Tableau 2.7-2 : Paramètres structuraux et physiques calculés à partir des données EXAFS enregistrées sur un complexe de NiEDTA au seuil du nickel.

Les distances cristallographiques (distance moyenne entre le nickel et un oxygène ou azote = 2,06 Å ; Figure 2.7-4) et les distances déterminées par l'EXAFS sont parfaitement en accord.

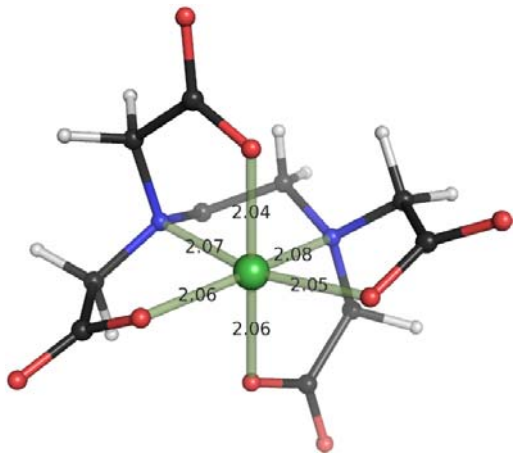


Figure 2.7-4 : Structure cristallographie d'un complexe de NiEDTA. Les distances sont en Å.

2.7.2.2.2 La protéine NikA au seuil du nickel

Nous avons alors réalisé l'analyse EXAFS de la protéine NikA, purifiée selon le protocole contenant du chloroforme, au seuil du nickel.

La transformée de Fourier des données expérimentales présente un pic autour de 1,5 Å, et un second pic autour de 2,0 Å majoritaire (sans correction de phase), correspondant à la première et à la seconde sphère de coordination, contre 1,6 et 2,2 Å dans le cas du complexe de NiEDTA (Figure 2.7-5).

Ceci nous permet alors de confirmer que ce n'est pas un complexe de NiEDTA qui est présent dans NikA. Nous avons alors décidé d'étudier la couche correspondant à ces deux pics en effectuant les filtres de Fourier sur deux couches (Figure 2.7-5).

Le meilleur affinement que l'on ait obtenu implique la considération de quatre couches de coordination autour du nickel (Tableau 2.7-3 et Figure 2.7-5). En effet, dans le cas où l'on n'en considère qu'une couche, composée d'oxygènes ou d'azotes, le résiduel obtenu est élevé, ce qui traduit un mauvais accord entre les données expérimentales et l'affinement du filtre de Fourier. D'un autre côté, dans le cas où l'on considère deux et trois couches de coordination, on obtient des valeurs de σ_i^2 négatives, ce qui n'est pas physique (σ_i n'étant pas un nombre complexe), et suggère un déficit électronique dans la composition de la coordination du nickel par rapport aux données expérimentales. Seule l'utilisation de quatre couches de coordination, impliquant quatre oxygènes ou azotes, la présence d'une chaîne carbonée (pouvant contenir des azotes) et

un chlore, permet d'affiner au mieux le filtre de Fourier des données expérimentales, et rendre compte de la hauteur du pic majoritaire de ces dernières.

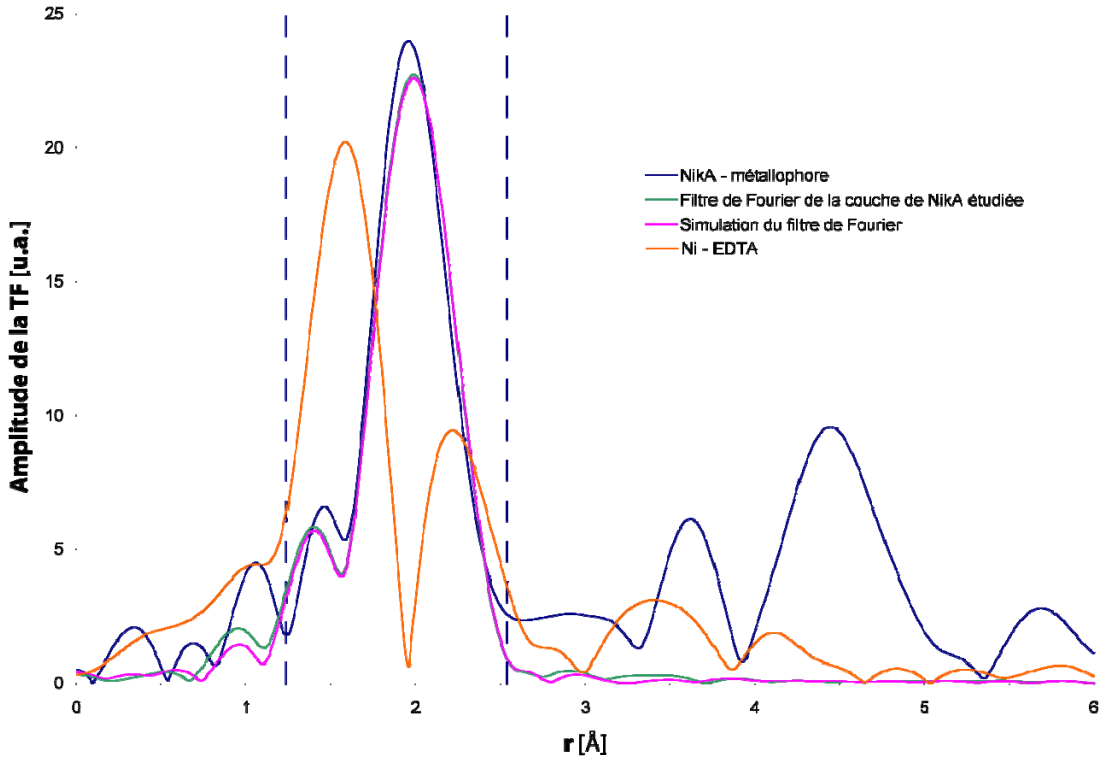


Figure 2.7-5 : Transformé de Fourier des données EXAFS.

En bleu : les données expérimentales enregistrées au seuil du nickel sur la protéine NikA purifiée en présence de chloroforme. Les traits en pointillés localisent la couche de coordination étudiée.

En vert : le filtre de Fourier de la couche étudiée.

En rose : la meilleure simulation du filtre de Fourier que nous ayons obtenue.

En orange : les données expérimentales enregistrées au seuil du nickel sur un complexe de NiEDTA.

Si l'on met en parallèle les résultats obtenus pour la première sphère de coordination (1 O/N à 2,00 Å et 3 O/N à 2,21 Å), avec les distances Ni – acide 3-(carboxy)hexanedioïque (1,84 Å ; 1,94 Å ; 2,40 Å) et Ni – histidine (2,23 Å) obtenues en cristallographie, on constate que l'on obtient un accord relativement bon.

La mise en évidence par l'EXAFS de la présence d'une chaîne carbonée, pouvant contenir des azotes, à 2,69 Å du nickel en moyenne, conforte l'hypothèse de la présence d'un métallophore. L'acide 3-(carboxy)hexanedioïque présente 7 carbones situés à, en moyenne, 2,89 Å du nickel, par conséquent, lors de l'étude des données EXAFS, nous avons fixé le nombre d'atomes de la chaîne carbonée à 7. Il est tout de même possible de modifier ce

nombre de quelques carbones ou azotes sans modifier fondamentalement tous les paramètres de l'ajustement au cas où l'acide 3-(carboxy)hexanedioïque ne soit pas le métallophore.

Une couche

Liaison	Paramètres structuraux			Résiduel
	R_i (Å)	N_i	σ_i^2 (Å^2)	
Ni - O/N	2,28	4	$8,85.10^{-4}$	$1,99.10^{-1}$

Deux couches

Liaison	Paramètres structuraux			Résiduel
	R_i (Å)	N_i	σ_i^2 (Å^2)	
Ni - O/N	2,21	1	$-7,03.10^{-3}$	$9,68.10^{-2}$
Ni - O/N	2,36	3	$-3,32.10^{-3}$	

Trois couches

Liaison	Paramètres structuraux			Résiduel
	R_i (Å)	N_i	σ_i^2 (Å^2)	
Ni - O/N	2,06	1	$-6,11.10^{-4}$	$9,68.10^{-3}$
Ni - O/N	2,26	3	$-3,56.10^{-3}$	
Ni - Cl	2,53	1	$-2,68.10^{-3}$	

Quatre couches

Liaison	Paramètres structuraux			Résiduel
	R_i (Å)	N_i	σ_i^2 (Å^2)	
Ni - O/N	2,00	1	$1,24.10^{-3}$	$2,11.10^{-3}$
Ni - O/N	2,21	3	$1,26.10^{-3}$	
Ni - Cl	2,44	1	$1,84.10^{-3}$	
Ni - C/N	2,70	7	$5,31.10^{-3}$	

Tableau 2.7-3 : Paramètres structuraux et physiques calculés à partir des données EXAFS enregistrées sur la protéine NikA purifiée en présence de chloroforme, au seuil du nickel.

Cette étude EXAFS suggère également que vient s'ajouter, en solution, aux quatre ligands du nickel, déjà identifiés en cristallographie, un chlore servant de ligand axial du métal par rapport du plan carré. En toute rigueur, on doit également considérer la possibilité qu'il s'agisse d'un soufre provenant d'une méthionine (la protéine ne contenant pas de cystéine). Le seul candidat possible est la méthionine 27 qui place, dans la structure cristallographique, son soufre à 5,2 Å du nickel. Il faudrait alors envisager des changements structuraux importants pour amener cette distance à 2,4 Å, qui de plus semblent impossibles car le métallophore masque l'accès du soufre à la méthionine.

Enfin, les pics supérieur à 2,5 Å de la transformée de Fourier des données expérimentales correspondent à un phénomène de diffusion multiple (qui pourrait entre autres provenir de l'histidine, mais pas uniquement) que nous n'avons pas pour l'instant interprété.

Ces résultats confirment la présence d'un métallophore spécifique du nickel, dans le site de fixation de NikA, en accord avec l'acide 3-(carboxy)hexanedioïque.

Mais il convient de confirmer ces résultats, car nous ne sommes qu'au début de cette étude EXAFS, et nous avons besoin de composés modèles supplémentaires afin de caractériser avec précision le complexe nickel – métallophore.

2.7.3 L'extraction et l'analyse de la molécule présente dans le site de fixation de NikA

2.7.3.1 L'extraction du nickelophore de la protéine

Afin d'extraire le nickelophore de la protéine nous avons employé deux approches :

- D'une part nous avons tenté de déplacer le métallophore du site de fixation en le remplaçant par de l'EDTA (de la même façon que ce qui se fit lors de l'extraction périplasmique à l'EDTA) ;
- D'autre part nous avons essayé de dénaturer la protéine par un changement de température.

2.7.3.1.1 La concentration d'EDTA permettant le déplacement du nickelophore

Cette concentration minimale d'EDTA permettant de déplacer le métallophore du site de fixation a été déterminée par des expériences de cristallisation : la protéine NikA purifiée selon le protocole contenant du chloroforme a été cristallisée en présence de différentes concentrations d'EDTA afin de déterminer à partir de laquelle il empêche l'apparition des cristaux hexagonaux (Figure 2.7-6).

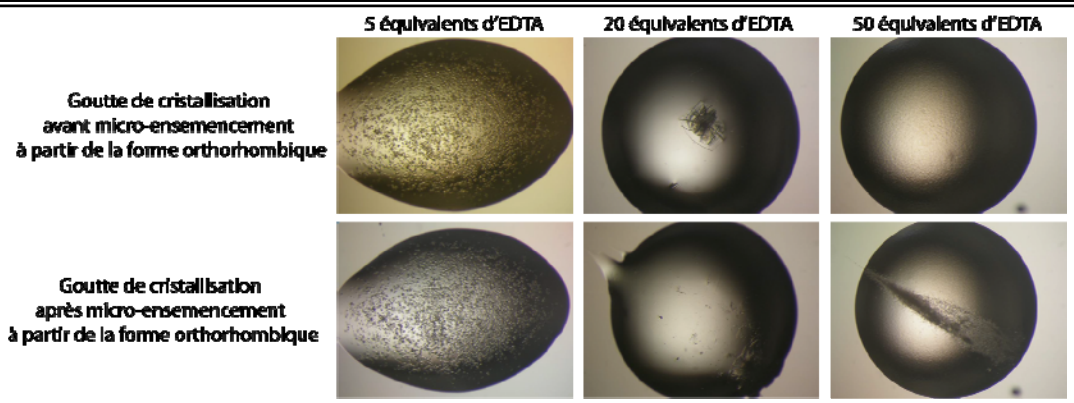


Figure 2.7-6 : Test de cristallisation de la protéine NikA purifiée en présence de chloroforme. Le micro-ensemencement provoque l'apparition de cristaux appartenant au système cristallin orthorhombique à partir d'un excès d'EDTA 50 fois supérieur à la concentration de protéine. En dessous de cette concentration, les cristaux appartiennent au système cristallin hexagonal, et le micro-ensemencement n'a pas d'effet.

Puis dans un second temps, un micro-ensemencement a été réalisé à partir de cristaux orthorhombiques, afin de vérifier si NikA avait fixé l'EDTA, car cette forme est caractéristique de la protéine ayant fixé ce dernier (Figure 2.7-6).
Ce traitement nous a permis de déterminer que l'EDTA déplace bien le métallophore du site de fixation de NikA à partir d'une concentration d'EDTA 50 fois supérieure à celle de la protéine.

2.7.3.1.2 L'étude de la protéine par dichroïsme circulaire

En ce qui concerne la dénaturation thermique de la protéine, nous l'avons réalisée dans un premier temps soit directement en plaçant la solution protéique à 70°C, soit en réalisant un gradient de température. Mais nous avons par la suite souhaité mieux la caractériser. Pour cela, nous avons soumis un échantillon protéique à un gradient de température de 40 à 70°C, et enregistré un spectre de dichroïsme circulaire à chaque palier de température (Figure 2.7-7a).
Jusqu' 50°C aucun changement significatif du spectre de dichroïsme circulaire n'est observé. Ce n'est qu'à 55°C que l'on constate une modification du deuxième minimum (222 nm) du spectre mais pas du premier (209 mn), et à partir de 58°C les deux minima se modifient.
La modification du spectre de dichroïsme circulaire de NikA traduit sa dénaturation et semble mettre en évidence l'existence d'un état intermédiaire vers 55°C, puisqu'à cette température le spectre présente un profil atypique par rapport aux autres températures (modification d'un seul des minima).
Une autre façon de mettre en évidence cet hypothétique état intermédiaire est de tracer l'évolution du minimum à 222 nm en fonction de la température (Figure 2.7-7b).

Il est donc possible que, lors du processus de dénaturation de NikA, la protéine passe par un état intermédiaire (vers 55°C) avant de se dénaturer complètement. On peut alors imaginer qu'à 55°C les deux domaines de la protéine (I et II) s'écartent l'un de l'autre, permettant ainsi une ouverture de la gorge présente à leur interface et contenant le site de fixation. Dans ce cas, en chauffant une solution protéique peu concentrée à 55°C, on pourrait augmenter l'efficacité d'extraction du métallophore.

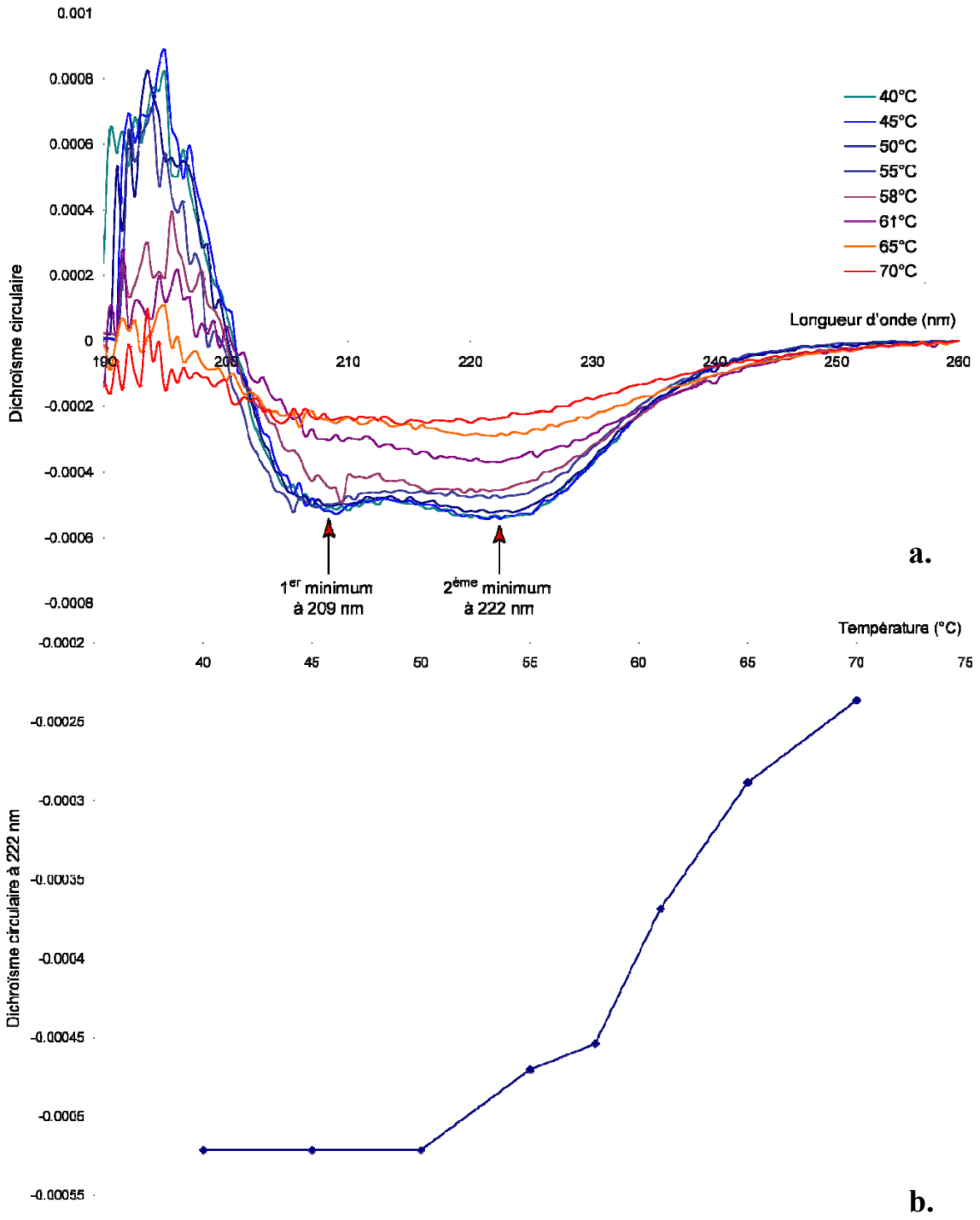


Figure 2.7-7 : Dichroïsme circulaire.

a. Spectres de dichroïsme circulaire mesurés aux différents paliers du gradient de température.

b. Dichroïsme circulaire à 222 nm en fonction de la température.

2.7.3.2 La spectroscopie de masse

2.7.3.2.1 L'étude de la protéine NikA en conditions natives

Afin d'estimer la masse du métallophore en complexe avec le nickel dans le site de fixation de NikA, nous avons réalisé une expérience de spectroscopie de masse en conditions natives (Figure 2.7-8).

Le pic A correspond à la masse de la protéine seule (56 306,9 Da), équivalent aux résultats obtenus en conditions dénaturantes (données non présentées) et le pic B correspond à une masse de 56 588,0 Da. Le calcul de la différence de masse entre le pic B et le pic A donne une valeur de 281,1 Da. Cette expérience a été réalisée plusieurs fois, ce qui permet d'estimer que la masse du complexe nickel – métallophore est comprise entre environ 275 et 285 Da. On peut alors calculer la masse du métallophore en retirant un nickel (58 Da pour l'isotope le plus abondant) et un chlore (35 Da pour l'isotope le plus abondant), ce qui suppose que le pic B correspond à NikA complexant uniquement ces derniers. Dans ce cas, la molécule doit avoir une masse comprise entre 182 et 192 Da.

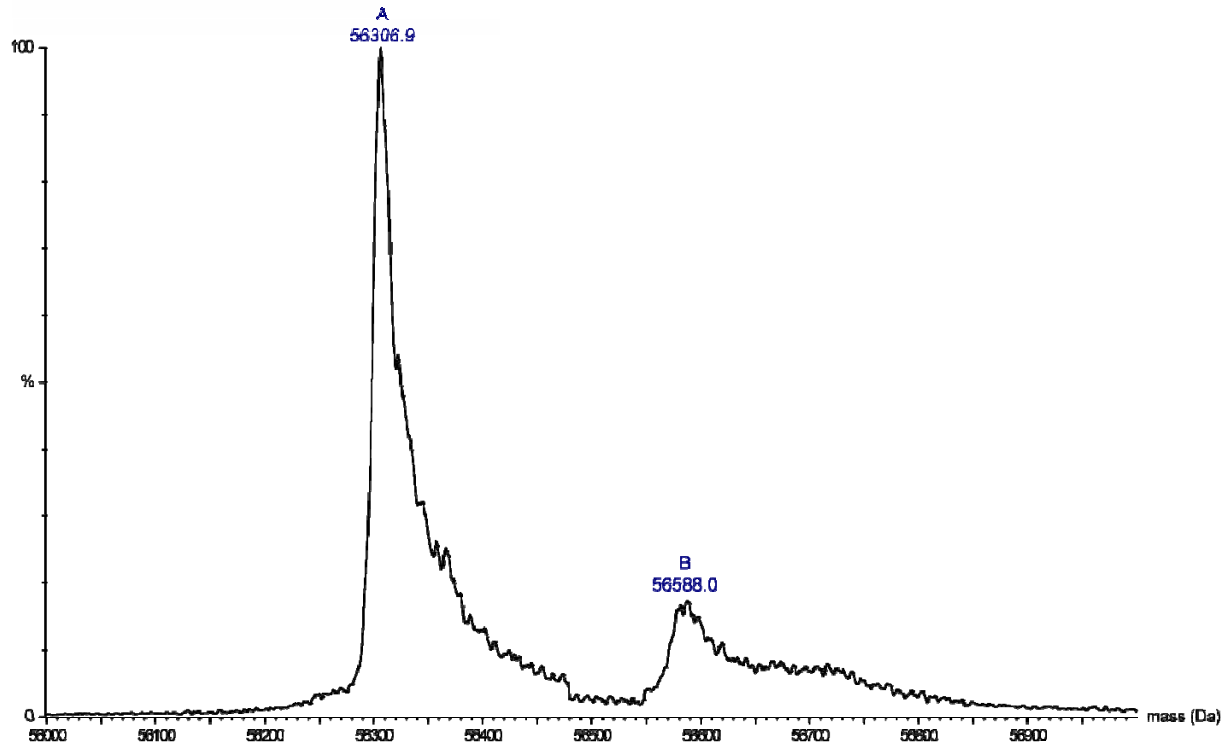


Figure 2.7-8 : Spectre de masse en conditions natives de la protéine NikA purifiée selon le protocole contenant du chloroforme, et dialysée contre de l'acétate d'ammonium 20 mM.

L'acide 3-(carboxy)hexanedioïque a une masse de 190 Da, et, dans le cas où il est en complexe avec le nickel, les carboxylates perdent leur hydrogène, ce qui diminue la masse à 187 Da.

Cette expérience conforte donc l'idée que l'acide 3-(carboxy)hexanedioïque est candidat possible pour être le nickelophore.

2.7.3.2.2 La caractérisation de la masse du métallophore

Afin de mieux caractériser la molécule, nous avons essayé de l'extraire de la protéine afin de pouvoir analyser plus facilement la masse du métallophore.

Dans un premier temps, nous avons analysé le surnagent d'une solution protéique concentrée chauffée à 70°C.

Ceci nous permit d'identifier une molécule ayant une masse de 242,7 Da (Figure 2.7-9). Cette dernière peut perdre 58 Da, soit la masse d'un nickel, pour donner une molécule de 184,7 Da. Le spectre ayant été enregistré en mode positif (enregistrement des cations), on doit retirer un proton à la masse de 184,7 Da pour obtenir la molécule électriquement neutre, soit 183,7 Da.

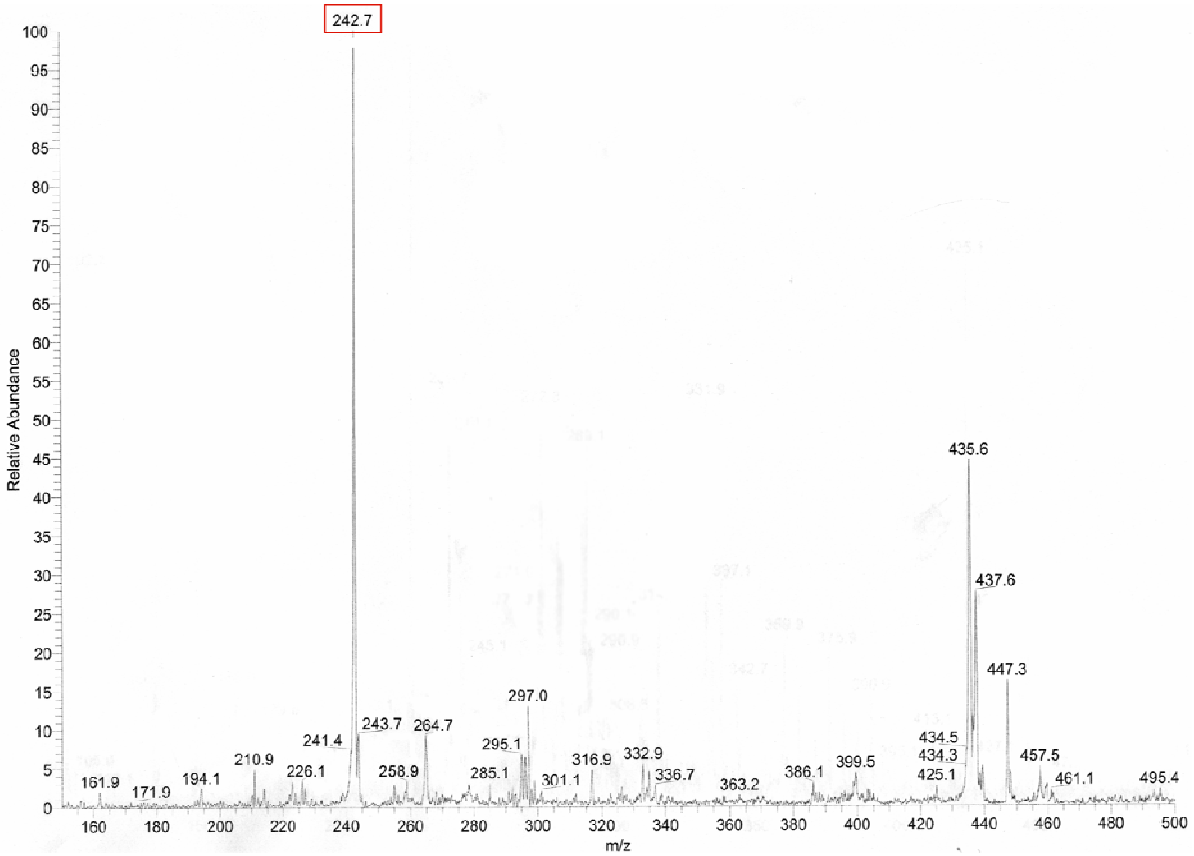


Figure 2.7-9 : Spectre de masse du surnageant d'une solution de NikA purifiée selon le protocole contenant du chloroforme, chauffée à 70°C.

Cette expérience est un indice supplémentaire indiquant que NikA contient un métallophore spécifique du nickel dans son site de fixation. De plus, elle nous permet d'avoir une estimation plus précise de sa masse, par rapport à l'expérience réalisée sur la protéine dialysée contre de l'acétate d'ammonium 20 mM. Par contre la masse de l'acide 3-(carboxy)hexanedioïque (190 Da) n'est pas en accord avec les 184 Da trouvés ici, si l'on considère que l'expérience donne une masse juste au dalton près.

Dans un second temps, nous avons essayé d'extraire la molécule de la protéine :

- En réalisant un gradient de température visant à dénaturer lentement la protéine peu concentrée en espérant favoriser l'extraction de la molécule ;
- En incubant la protéine avec une concentration en EDTA 50 fois supérieure, visant à remplacer le métallophore.

Dans les deux cas, les spectres de masse ont montré la présence d'une molécule à 187,9 Da, qui peut être cassée et perdre l'équivalent d'un CO_2 (44 Da), ce qui signifie que la molécule contient au moins un groupement carboxylate (Figure 2.7-10). Ceci est en accord avec ce que nous avions supposé grâce à la forme de la densité électronique « omit », présente dans le site de fixation de NikA, et par comparaison avec la structure. Le spectre a été enregistré en mode négatif (enregistrement des anions), on doit donc ajouter un proton à la molécule pour avoir une estimation de la masse du métallophore électriquement neutre, soit 188,9 Da.
Le spectre de masse a également montré que si le métallophore contient de l'azote, il y a un nombre impair de ce dernier (1 ou 3).

En définitive, ces expériences de spectroscopie de masse ont permis de confirmer la présence dans NikA d'une molécule complexant un nickel et possédant au moins une fonction carboxylate.
Ceci est parfaitement en accord avec notre hypothèse de l'existence d'un métallophore spécifique du nickel ayant une structure similaire à celle de l'EDTA.

Le problème qui se pose maintenant est que nous obtenons deux masses pour le métallophore : 184 et 189 Da. L'acide 3-(carboxy)hexanedioïque possède une masse de 190 Da, de ce fait, si ce n'est pas le métallophore, ce dernier doit avoir une structure très proche.

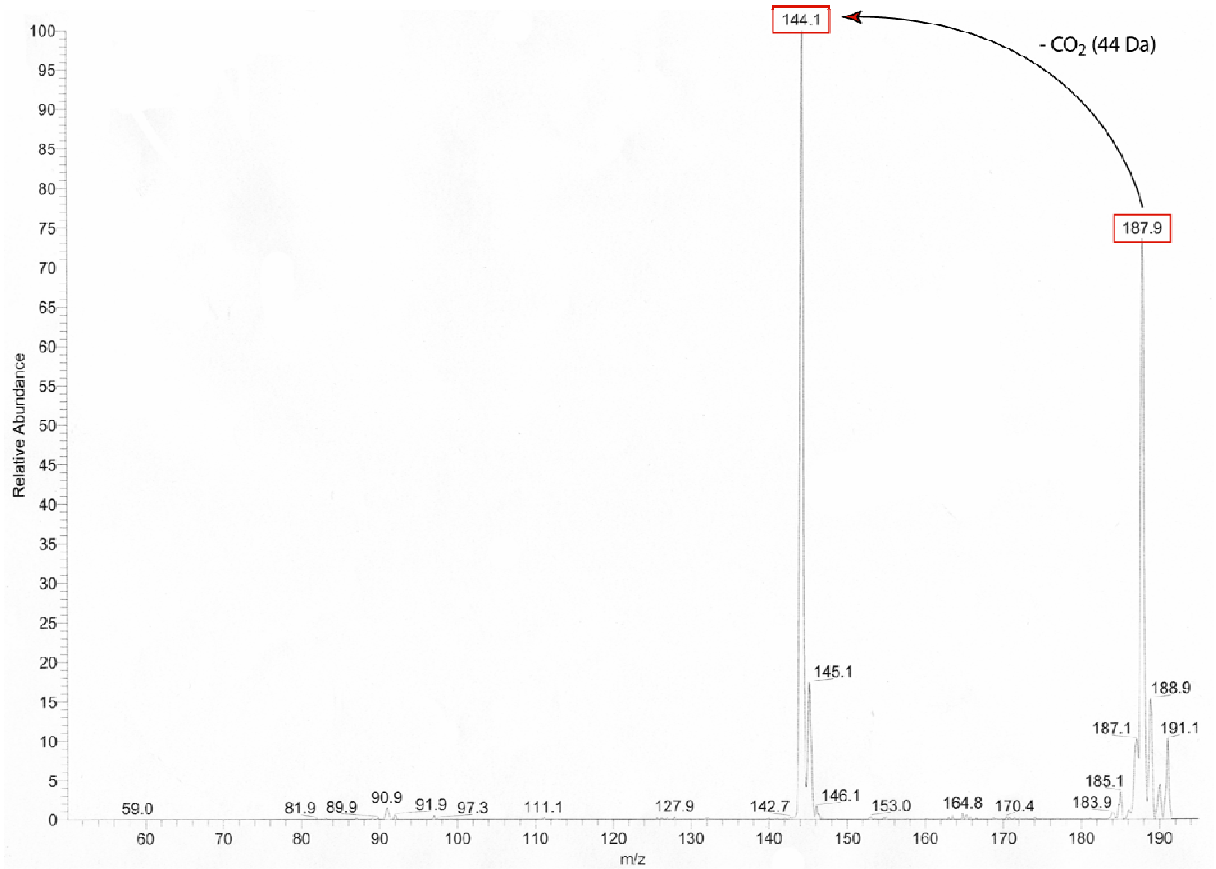


Figure 2.7-10 : Spectre de masse de la molécule extraite de NikA purifiée selon le protocole contenant du chloroforme par traitement à l'EDTA.

Ces résultats permettent d'envisager un autre nickelophore qui serait l'acide N-carboxy N-éthanoïque prop-2-énoïque car sa masse moléculaire est la même que celle déterminée plus haut : soit 189 Da (Figure 2.7-11). Même si la structure de cette molécule est très proche de celle de l'acide 3-(carboxy)hexanedioïque, elle présente plusieurs formes mésomères qui impliquent la planarité de l'azote et impose que deux des carboxylates soient dans le même plan. Ceci empêche un bon accord avec les résultats cristallographiques.

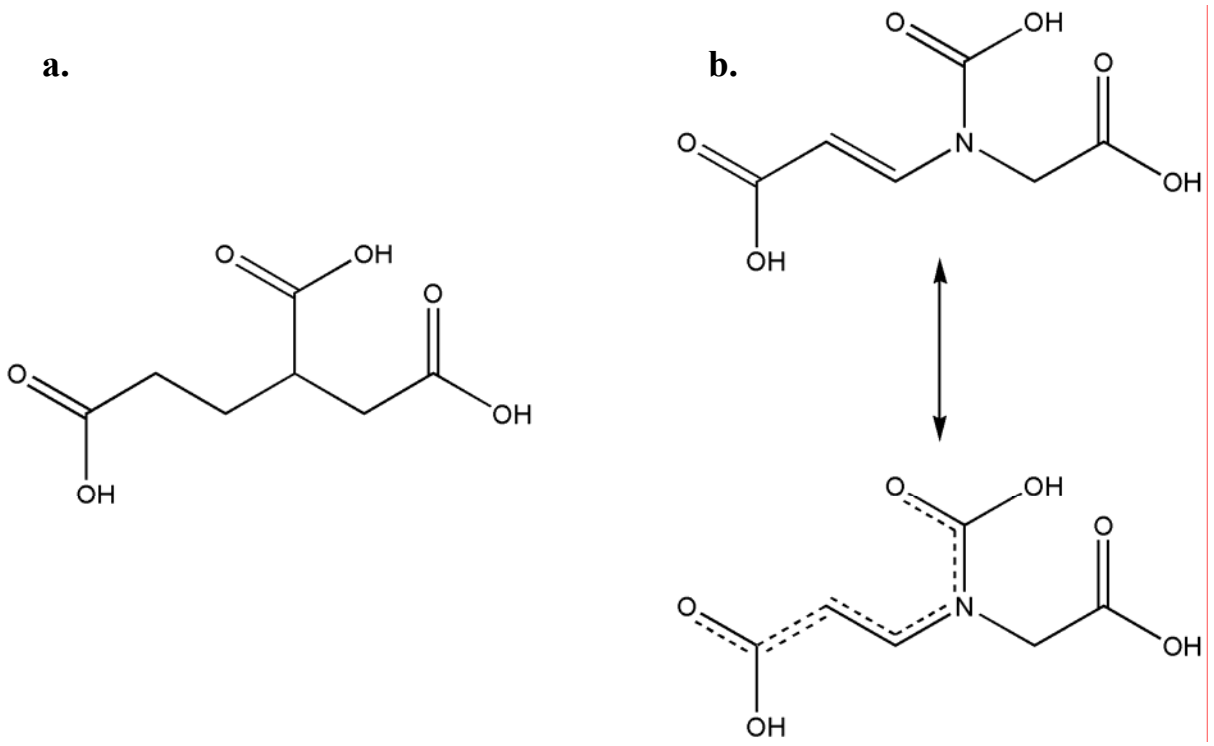


Figure 2.7-11 : Comparaison des structures de (a.) l'acide 3-(carboxy)hexanedioïque et de (b.) l'acide N-carboxy N-éthanoïque prop-2-énoïque.

2.7.3.3 La spectroscopie RMN

Afin d'identifier avec certitude la nature du métallophore, nous avons décidé de réaliser une étude RMN des échantillons analysés en spectroscopie de masse et décrits précédemment, c'est-à-dire la molécule extraite par le gradient de température et celle extraite par le traitement à l'EDTA.

Dans les deux cas, le spectre 1D du proton a présenté le même profil (Figure 2.7-12).

Afin de caractériser la structure du métallophore, on enregistra sur l'échantillon traité à la température un spectre 1D ^{13}C (en utilisant l'abondance naturelle du ^{13}C), ainsi qu'un spectre 2D HMQC permettant de visualiser les corrélations proton – carbone.

Malheureusement, le spectre 1D ^{13}C ne montra que deux carbones : le premier correspondant à un $-CH_2$, et le second pouvant être attribué en tant que -CH ou $-CH_2$.

Le spectre 2D HMQC présenta la même situation.

Ce nombre de carbones ne correspond pas aux résultats obtenus en cristallographie, en EXAFS et en spectroscopie de masse.

De plus, aucun carboxylate ne put être mis en évidence, ce qui, là aussi, est en désaccord avec les résultats précédents.

Mais, il est connu que la présence d'un atome paramagnétique peut rendre invisible à l'expérience RMN les atomes voisins, par exemple dans le cas de la résolution de la structure de l'acireductone dioxygenase, dans un rayon de 8 Å autour de l'atome de nickel aucun proton ne put être visualisé (Pochapsky *et al.*, 2002). Dans notre cas, le métallophore pourrait être en complexe avec le Ni^{2+}, et ce dernier présentant un caractère paramagnétique il nous empêcherait alors de voir la totalité des protons et des carbones de la molécule, et notamment les carboxylates.

Il conviendra donc de dissocier au préalable le complexe Ni-métallophore pour pouvoir caractériser la structure du métallophore par spectroscopie RMN.

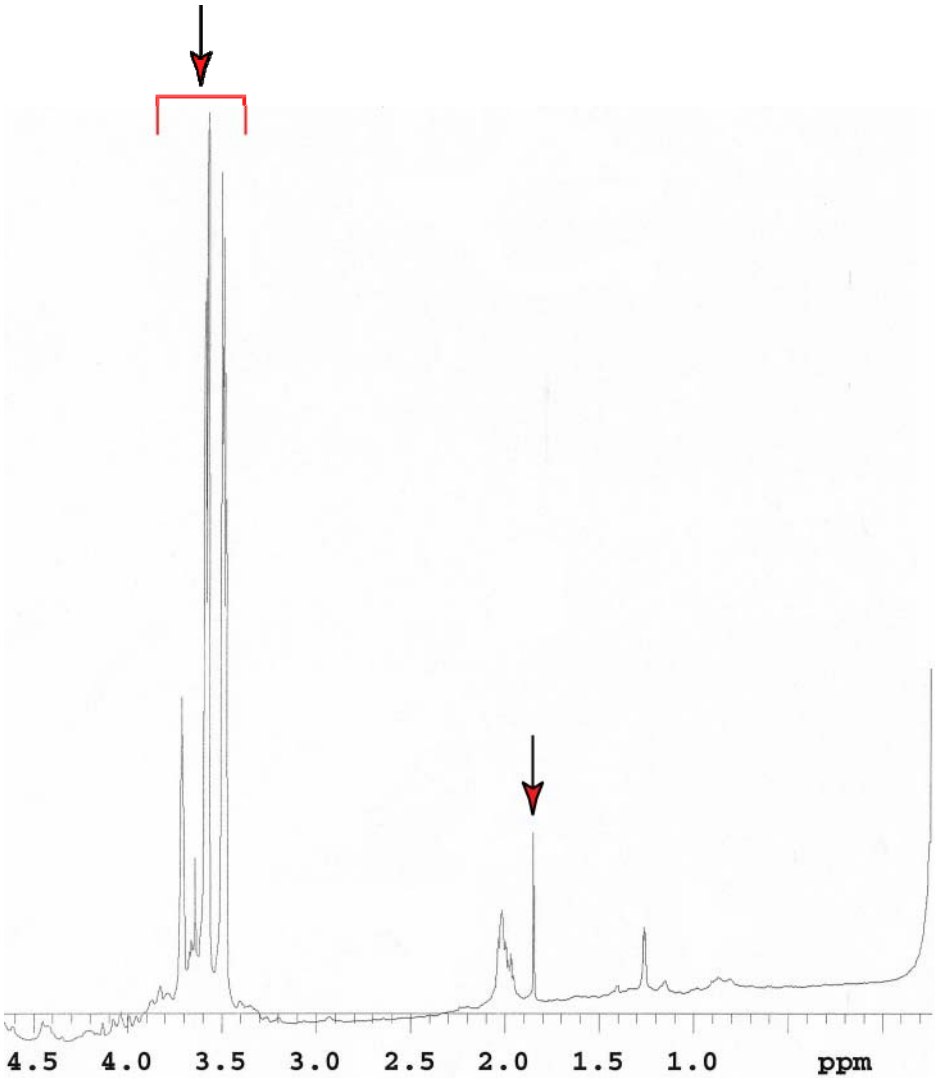


Figure 2.7-12 : Spectre RMN 1D du proton de la molécule (flèches rouges) extraite de NikA purifiée selon le protocole contenant du chloroforme, par un gradient de température.

2.8 La comparaison des systèmes de transport du nickel et fer

Comme nous venons de le montrer, le transport du nickel par le système NikABCDE implique l'existence d'un métallophore spécifique du nickel, que l'on peut désigner sous le terme de nickelophore. Ce système d'import présente une similarité avec celui du fer par les sidérophores, puisque dans les deux cas le il ne reconnaît pas directement le métal, mais le complexe métal-métallophore.

Dans notre cas, le nickelophore semble présenter une structure moins complexe, et être plus léger (entre environ 185 et 190 Da) que la majorité des sidérophores décrits (par exemple 324 Da pour la pyocheline qui est l'un des sidéropores les plus simples). Il se peut que cet écart vienne du fait de la différence de coordination que nécessitent ces deux métaux : le nickel adoptant généralement une géométrie plan carré à quatre ligands, contre une géométrie octaédrique à six ligands pour le fer.

Contrairement au fer, aucun système de transport à travers la membrane externe n'a été mis en évidence pour le nickel. Dans ce cas, on peut se demander si le métallophore reste dans le périplasme et de ce fait le nickel traverse seul la membrane interne grâce, par exemple, à des porines, ou s'il existe un récepteur fonctionnant de façon similaire à ceux impliqués dans le transport des complexes fer-sidérophores, permettant au complexe nickel-nickelophore de traverser la membrane externe. Dans le cas où ce dernier fonctionne comme pour le fer, l'import dans le périplasme pourrait alors être couplé à un transport de protons vers le cytoplasme (chapitre 3.3.2 de l'introduction).

C'est un système du type ABC (sous famille 8) qui est impliqué dans l'import du fer, de la même façon que celui du nickel par NikABCDE (sous famille 2). Mais la protéine NikA ne présente pas d'identité de séquence particulière (ainsi que les autres protéines du système de transport : NikB, NikC, NikD, NikE), ni un repliement similaire aux protéines périplasmique impliquées dans le transport des complexes fer-sidérophore (Figure 2.8-1).

Le fait que les deux systèmes soient du type ABC, nous amène à nous demander si leur fonctionnement est similaire : si, comme c'est le plus souvent le cas pour le fer, le nickelophore est importé avec le nickel, ou si le complexe est dissocié avant le passage dans le cytoplasme.

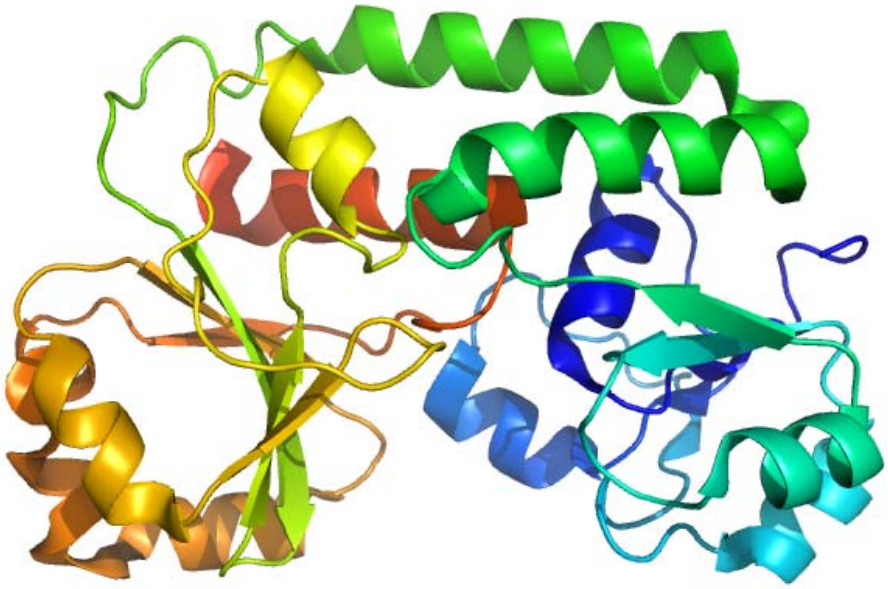


Figure 2.8-1 : Structure de la protéine périplasmique Fhud d'*E. coli* (code PDB : 1EFD) impliquée dans le transport du fer par l'intermédiaire des sidérophores de la famille des hydroxamates.

2.9 La comparaison de la protéine NikA d'*E. coli* avec ses homologues

Nous avons montré que plusieurs résidus du site de fixation de NikA avaient un rôle important dans la fixation du nickel et du métallophore.

Nous avons alors décidé de vérifier la conservation des résidus du site de fixation chez les autres protéines décrites comme étant des homologues de NikA d'*E. coli*.

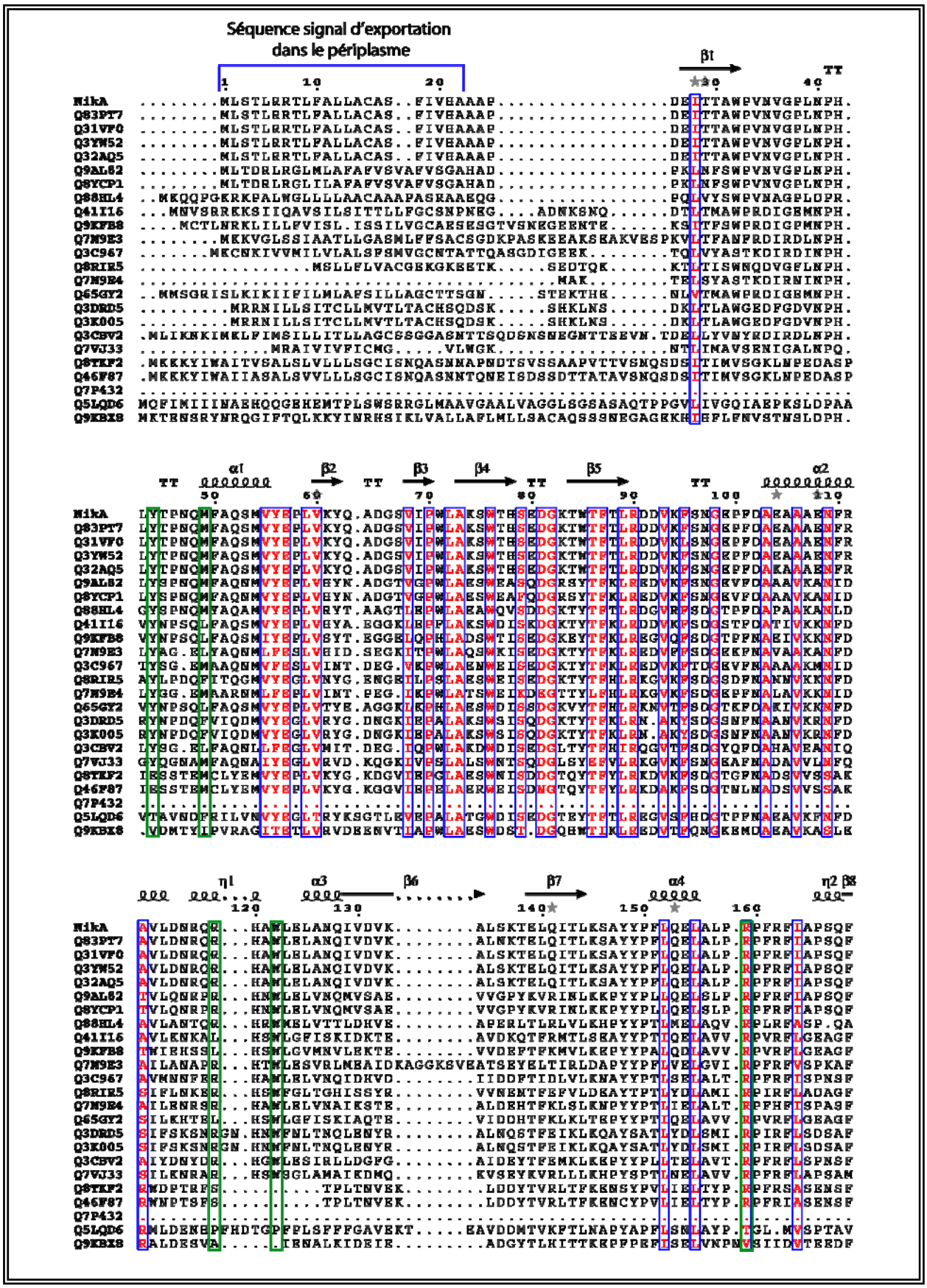
Pour cela, nous avons sélectionné les protéines possédant une identité de séquence significative avec NikA (calculée avec le programme Blastp (Madden *et al.*, 1996)) et annotées comme étant ses homologues ou comme étant des protéines capables de fixer le nickel. Une fois triées, l'alignement multiple des séquences protéiques est assuré par le programme ClustalW (Chenna *et al.*, 2003), et sa présentation est réalisée par le programme ESPript (Gouet *et al.*, 1999) (Figure 1.1-1et Tableau 2.9-1).

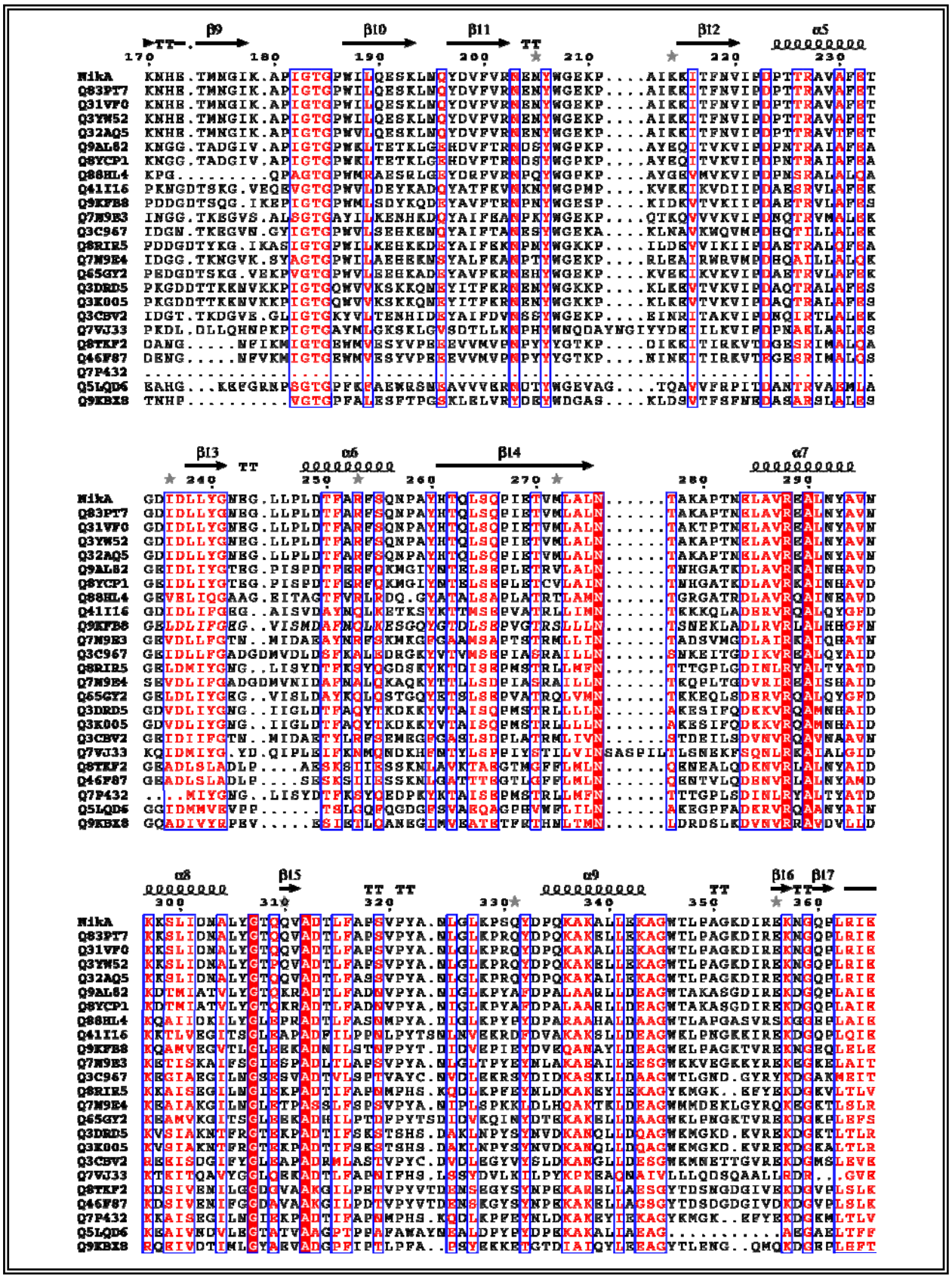
En se basant sur la conservation des acides aminés du site actif, et en particulier ceux impliqués dans la fixation du complexe nickel – métallophore (Arg 97, Arg 137 et His 416), on peut reconsidérer l'annotation en tant qu'homologue de NikA faite pour certaines protéines (Tableau 2.9-1). Par exemple la protéine Q9KFB8 de *Bacillus halodurans* n'a que peu de chance d'être, comme elle est décrite, un homologue, puisqu'on ne retrouve dans sa séquence que peu de résidus du site actif conservés, et l'Arg 97 et l'His 416 n'en font pas partie. D'un autre côté, cette analyse permet de conforter le fait que des organismes tels que *B. suis* ou *Helicobacter hepaticus*, qui sont deux pathogènes, possèdent un système de transport du

nickel du type ABC. Chez *Wolinella succinogenes*, deux protéines avaient été annotées comme des homologues potentiels de NikA. Sur la base de la conservation des résidus du site de fixation, il semble que ce soit la protéine Q7M9E4 qui soit l'homologue de NikA, notamment car elle conserve l'His 416.

Code de la protéine	Organisme	Identité de séquence	Nombre de résidus comparés à NikA	Possibilité d'être un homologue de Nika
Q83PT7	*Shigella flexneri*	99%	499/502	+++
Q31VF0	*Shigella boydii*	99%	499/502	+++
Q3YW52	*Shigella sonnei*	99%	499/502	+++
Q32AQ5	*Shigella dysenteriae*	98%	494/502	+++
Q9AL82	*Brucella suis*	64%	323/498	+++
Q8YCP1	*Brucella melitensis*	64%	321/498	+++
Q88HL4	*Pseudomonas putida*	50%	250/495	+++
Q41I16	*Exiguobacterium sp.*	43%	214/496	--
Q9KFB8	*Bacillus halodurans*	42%	208/493	--
Q7M9E3	*Wolinella succinogenes*	43%	218/504	+
Q3C967	*Alkaliphilus metalliredigenes*	43%	210/487	+++
Q8RIR5	*Fusobacterium nucleatum subsp. nucleatum*	40%	199/496	-
Q7M9E4	*Wolinella succinogenes*	41%	205/497	+++
Q65GY2	*Bacillus licheniformis*	39%	197/495	--
Q3DRD5	*Streptococcus agalactiae*	41%	207/504	++
Q3K005	*Streptococcus agalactiae serotype Ia*	40%	206/504	++
Q3CBV2	*Alkaliphilus metalliredigenes*	39%	190/486	+
Q7VJ33	*Helicobacter hepaticus*	35%	180/503	+++
Q8TKF2	*Methanosarcina acetivorans*	34%	166/484	--
Q46F87	*Methanosarcina barkeri*	32%	159/488	--
Q7P432	*Fusobacterium nucleatum subsp.vincentii*	37%	92/248	---
Q5LQD6	*Silicibacter pomeroyi*	28%	135/471	---
Q9KBX8	*Bacillus halodurans*	28%	105/367	---

Tableau 2.9-1 : Protéines annotées comme homologues à la protéine NikA d'*E. coli* et présentant une identité de séquence significative. La possibilité que la protéine soit un homologue de NikA va de (+++) très probable à (---) très peu probable.


Séquence signal d'exportation
dans le périplasme
NikA
Q83PT7
Q31VF0
Q3YW52
Q32AQ5
Q9AL82
Q8YCP1
Q88HL4
Q41II6
Q9KFB8
Q7M9E3
Q3C967
Q8RIR5
Q7M9B4
Q65GY2
Q3DRD5
Q3K005
Q3CBV2
Q7VJ33
Q8TKF2
Q46F87
Q7P432
Q5LQD6
Q9KBX8


β9
β10
β11
β12
α5
β13
α6
β14
α7
α8
β15
α9
β16
β17
NikA
Q83PT7
Q31VF0
Q3YW52
Q32AQ5
Q9AL82
Q8YCP1
Q88HL4
Q41I16
Q9KFB8
Q7N9E3
Q3C967
Q8RIR5
Q7N9E4
Q65GY2
Q3DRD5
Q3K005
Q3CBV2
Q7VJ33
Q8TKF2
Q46F87
Q7P432
Q5LQD6
Q9KBX8

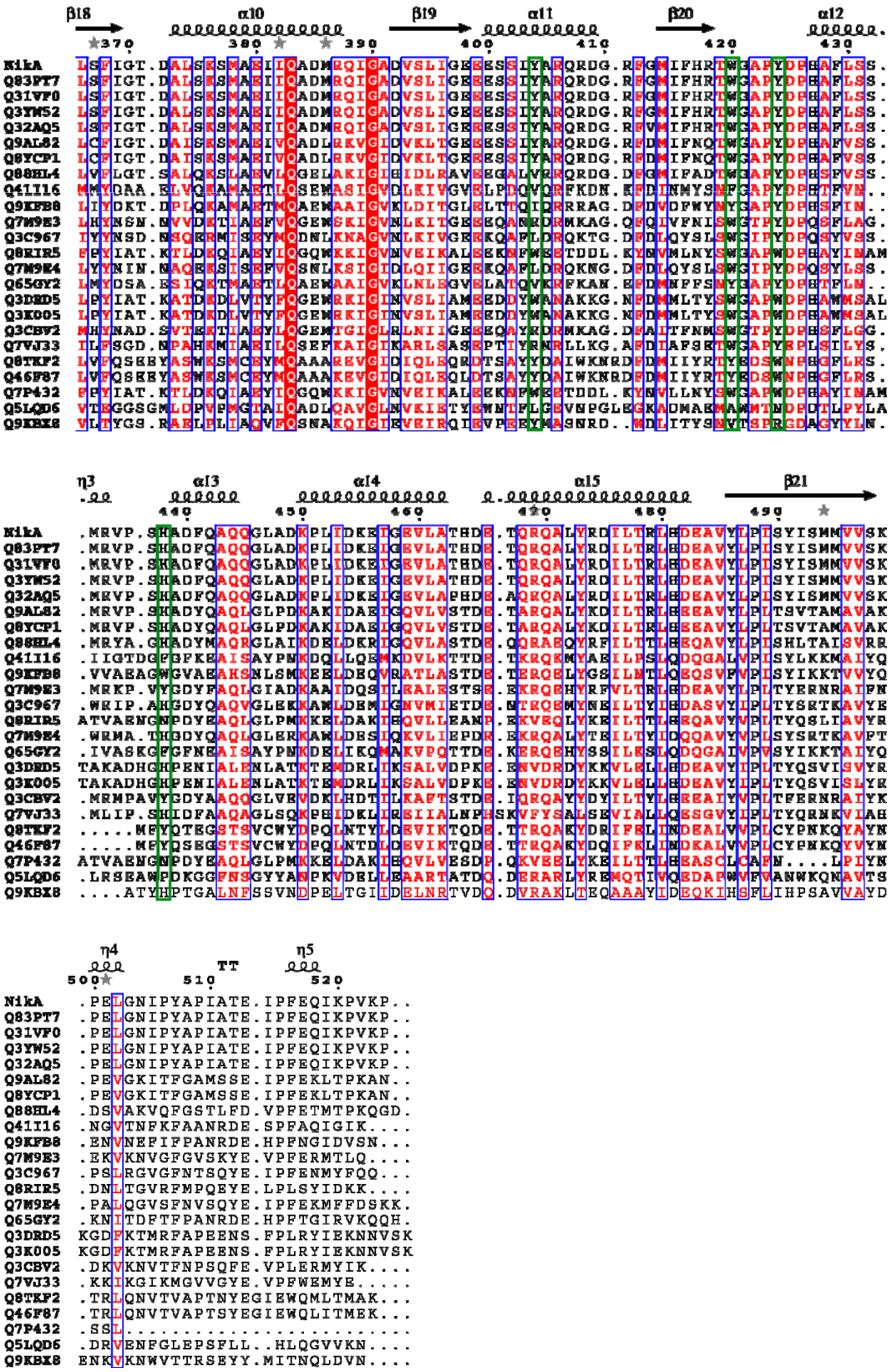
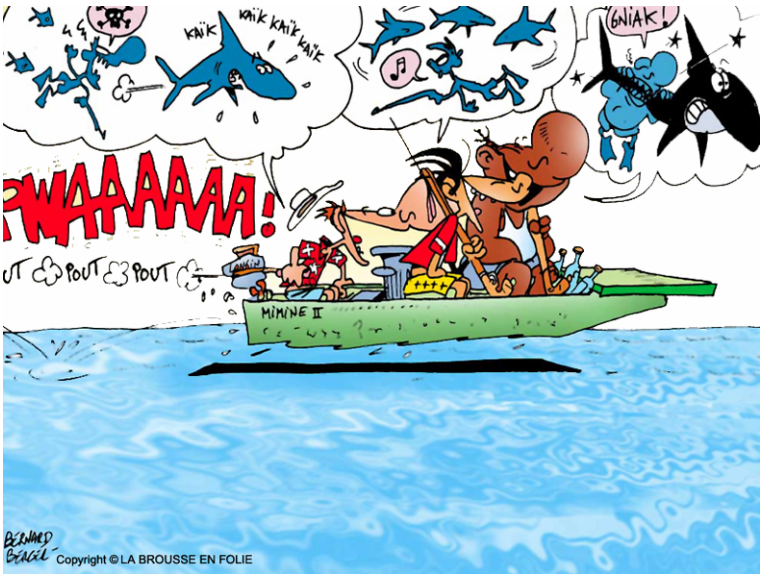


Figure 2.9-1 : Alignement des séquences de NikA d'*E. coli* et de ses homologues. La séquence de NikA ici présentée contient la séquence signal d'exportation dans le périplasme. Les résidus du site de fixation du nickel sont encadrés en vert.

CONCLUSION

ET

PERSPECTIVES



1 La protéine NikR

Une partie de mon travail de thèse consistait en la résolution de la structure par cristallographie aux rayons X de la protéine NikR d'*E. coli*, qui est le répresseur de l'opéron *nikABCDE* impliqué dans l'import du nickel.

Bien que nos travaux nous aient permis d'obtenir des cristaux de protéine de bonne qualité, grâce en partie à l'utilisation de propane liquide pour congeler ces derniers, nous n'avons pas réussi à résoudre la structure de NikR avant que *Schreiter et al.* ne publient la structure apo de cette protéine, ainsi que celle du domaine C terminal ayant fixé le nickel d'*E. coli* (Schreiter *et al.*, 2003). Ils ont montré que, comme cela était suspecté, la protéine est constituée de deux domaines :

- Un domaine N terminal se fixant à l'ADN ;
- Un domaine C terminal contenant le site de fixation du nickel de haute affinité et assurant la tétramérisation de la protéine.

Par la suite, plusieurs structures de NikR de *P. horikoshii* ont été résolues, ce qui a permis de localiser les sites de faible affinité du nickel, et de proposer un modèle de la fixation de NikR à l'ADN.

Ces structures étant disponibles, nous avons alors décidé de les utiliser pour effectuer un remplacement moléculaire sur les données de diffraction que nous avions collectées. La protéine que nous étudions étant un mutant, il était en effet intéressant d'en résoudre la structure, et ainsi d'essayer de comprendre comment la mutation Q2E modifie le domaine N terminal pour le rendre incapable de se fixer à l'ADN. Malheureusement, seul le domaine C terminal a pu être placé dans la densité électronique, le domaine N terminal resta en grande partie invisible. La mutation Q2E de la protéine que nous étudions doit induire un désordre structural dans le site de fixation à l'ADN de NikR, ce qui est en corrélation avec le fait que ce mutant ne se fixe au niveau du promoteur de l'opéron *nikABCDE* qu'en présence d'une concentration de nickel supérieure à celle permettant la fixation de la protéine sauvage.

En l'état actuel des choses, les seuls travaux encore envisageables sur NikR consistent en l'étude de la fixation de l'EDTA à la protéine. En effet, des résultats récents tendraient à monter que la protéine a une affinité pour l'EDTA (communication personnelle d'Isabelle Michaud-Soret du Laboratoire de Physicochimie des Métaux en Biologie du CEA).

On peut alors envisager que, comme dans le cas de NikA, NikR soit capable de fixer l'EDTA dans un site adapté au métallophore. Il serait alors intéressant d'obtenir une structure cristallographique de NikR avec l'EDTA et ainsi de vérifier l'existence ou non d'un site capable de fixer spécifiquement l'EDTA. D'autre part, si nous nous parvenons à déterminer avec certitude la nature du nickelophore naturel lié à NikA, il serait ensuite intéressant de tenter de le lier avec NikR dans le site de liaison du nickel.

2 La protéine NikA

Nous voulions également, dans le cadre de cette thèse, résoudre la structure de NikA, qui est la protéine périplasmique du système de transport du nickel.
Dans ce cas, la première difficulté a été la reproduction des cristaux obtenus lors de la première étude de cristallisation menée au laboratoire (Charon *et al.*, 1994).
La solution a été alors de réaliser un micro-ensemencement à partir de cristaux obtenus avec une protéine utilisée lors des premières études. Mais dans ce cas, comme dans celui de NikR, nous n'avons pas réussi à résoudre la structure de la protéine avant que Heddle *et al.* ne publient les structures de NikA d'*E. coli* apo et fixant le nickel (Heddle *et al.*, 2003).

La protéine présente deux domaines à l'interface desquels Heddle *et al.* ont proposé la fixation d'un nickel pentahydraté dans une poche relativement hydrophobe.

Nous avons alors montré que NikA était en réalité capable de fixer un complexe de $FeEDTA(H_2O)^-$ avec une haute affinité (Cherrier *et al.*, 2005).

Ce résultat nous a alors permis d'envisager l'existence d'un métallophore spécifique du nickel possédant une structure proche de celle de l'EDTA, puisqu'aucun autre site de la protéine ne semble adapté à la fixation du nickel. Le système ABC codé par l'opéron *nikABCDE* transporterait alors le nickel à l'image de celui du fer par les sidérophores. L'existence d'un métallophore spécifique du nickel n'avait jamais encore été envisagée, et peut permettre de mieux comprendre pourquoi toutes les valeurs de la constante de dissociation du nickel vis-à-vis de la protéine, données dans la littérature, sont élevées et très hétérogènes. Même s'il est difficile de réinterpréter ces études, il est clair qu'elles ont été réalisées en ne considérant ni la présence du métallophore, ni le fait que l'EDTA occupe le site de fixation de NikA avec une haute affinité. Il ne faut donc pas mesurer l'affinité du nickel pour la protéine, mais l'affinité du complexe nickel – métallophore pour NikA.

La capacité de NikA à fixer un complexe de $FeEDTA(H_2O)^-$ présente d'autre part un intérêt chimique. En effet, la formation de complexes artificiels protéines/molécules de synthèse, ayant des propriétés catalytiques, est devenue un sujet de premier ordre en chimie. Le Complexe $FeEDTA(H_2O)^-$ est connu comme capable d'activer les peroxydes (Brausam and van Eldik, 2004), qui sont alors disponibles pour réaliser des oxydations. Il serait intéressant d'étudier cette réaction (mais aussi des réactions analogues, où l'on utiliserait des

complexes analogues d'EDTA), dans l'environnement protéique qu'est NikA, et de comparer sa réactivité par rapport à l'espèce en solution.

Le Complexe $FeEDTA(H_2O)^-$ a également la particularité de présenter deux énantiomères en fonction du mode de fixation de l'EDTA au fer. La fixation de ce dernier à NikA correspond à un dédoublement racémique du complexe, puisque uniquement l'un des deux énantiomères est reconnu par la protéine. Ainsi, en utilisant des peroxydes activés par le complexe $FeEDTA(H_2O)^-$ fixé à la protéine, il serait possible de réaliser des oxydations présentant un caractère chiral, et donc de n'obtenir, comme produit de réaction, qu'un seul des énantiomères. Pour ces raisons, nous avons entamé une collaboration avec Stéphane Ménage (du Laboratoire de Chimie Biologique (CB) du CEA) visant à mener ces études.

Afin d'identifier et de caractériser le métallophore, nous avons par la suite décidé de recommencer l'étude cristallographique de la protéine, mais cette fois en veillant à ce qu'elle ne soit jamais en contact avec de l'EDTA. Le seul moment où NikA est en présence d'EDTA est l'étape d'extraction périplasmique, de ce fait nous l'avons donc remplacée par une extraction périplasmique au chloroforme. Ceci nous a alors permis d'obtenir des cristaux appartenant au système cristallin hexagonal, alors que les cristaux étudiés précédemment appartenaient au système cristallin orthorhombique. Les différentes études que nous avons menées nous ont alors permis de conforter notre hypothèse selon laquelle NikA accepte dans son site de fixation un métallophore présentant une haute affinité pour le nickel, et de structure proche de celle de l'EDTA. Ce dernier adopte, dans la protéine, une géométrie plan carré impliquant trois carboxylates du métallophore et une histidine de la protéine qui n'était pas impliquée dans la fixation de l'EDTA. Cette histidine joue donc un rôle central quant à l'affinité de NikA vis-à-vis du nickel. L'étude EXAFS suggère qu'un ion chlorure complète la coordination du nickel, en plus des quatre ligands. La sélectivité de la protéine vis-à-vis du nickel par rapport aux autres métaux, doit en grande partie être induite par cette géométrie de coordination.

Enfin, les études que nous avons menées concernant la caractérisation de la structure du nickelophore nous ont permis à l'heure actuelle, d'envisager deux candidats possibles : l'acide 3-(carboxy)hexanedioïque et l'acide N-carboxy N-éthanoïque prop-2-énoïque. En tout état de cause, même si aucune de ces deux molécules ne s'avère être le métallophore

physiologique, nos résultats suggèrent que ce dernier doit avoir une structure proche des molécules proposées.

Afin de déterminer avec certitude la structure du métallophore, deux solutions sont envisageables :

- La première consiste à obtenir une structure cristalline de la protéine à plus haute résolution que celle obtenue (2,5 Å). Jusqu'à présent, nous avons tenté sans succès d'améliorer les cristaux appartenant au système cristallin hexagonal. Mais l'étude de l'empilement cristallin montre la présence de canaux de solvant importants et de l'existence d'un nombre limité d'interactions entres les molécules de l'unité asymétrique (notamment entre la molécule C et les molécules A et B), qui doivent être la cause de la diffraction limitée que nous avons observée pour ces cristaux. De ce fait, si l'on espère déterminer la structure du nickelophore grâce à la cristallographie, il nous faut obtenir une nouvelle forme cristalline ;
- La seconde solution consiste à obtenir la structure de la molécule par RMN. Pour cela il nous faut dissocier le complexe métallophore – nickel, car le caractère paramagnétique de ce dernier nous a, jusqu'à présent, empêché de voir lors de l'expérience RMN toute la molécule, et notamment les carboxylates dont l'existence est suggérée par la cristallographie, et confirmée par la spectroscopie de masse. Une façon de dissocier le complexe serait d'acidifier la solution afin de protoner le métallophore, puis de l'extraire à l'aide d'un solvant organique tel que le chloroforme ou le dichlorométhane.

Une fois la structure du métallophore caractérisée, ce qui est notre priorité, il sera possible de recalculer la constante de dissociation de NikA vis-à-vis du complexe nickel – métallophore. Mais pour cela, il nous faut une protéine NikA apo, ce que nous n'avons pas réussi à obtenir jusqu'à présent. Pour y parvenir, nous avons demandé à l'équipe de J.R. Tames de nous fournir le plasmide leur ayant permis de cristalliser la protéine apo. Il contient le gène de NikA sans sa séquence d'exportation vers le périplasme. On peut alors envisager qu'en restant dans le cytoplasme, la protéine ne peut être en contact avec le métallophore si l'on fait l'hypothèse que ce dernier n'est pas transporté avec le nickel. Ceci expliquerait pourquoi ils ont réussi à obtenir une protéine ayant son site de fixation vide.

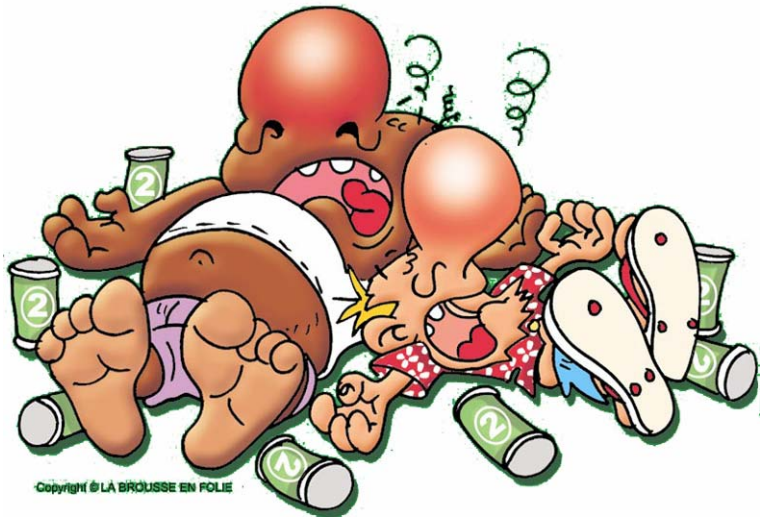
Il serait également intéressant de réaliser une étude de mutagenèse des résidus du site de fixation et notamment l'histidine 416, puisqu'elle joue un rôle important dans la coordination du nickel. La mutation des arginines 97 et 137 permettrait également d'obtenir des informations importantes sur l'affinité du métallophore pour la protéine.

Lorsque l'on réalise une recherche des homologues de NikA en fonction de leur identité de séquence, nous avons vu que certaines protéines semblent avoir été mal annotées. Il conviendrait donc de poursuivre cette étude afin de vérifier si de vrais homologues n'ont pas, à l'inverse, été désignés comme une autre protéine, comme par exemple un homologue de OppA ou de DppA. Pour cela, il faudrait dans un premier temps vérifier si les résidus du site du site actif ont une importance dans ces protéines.

A plus long terme, la meilleure connaissance du mécanisme de transport du nickel par le système ABC pourrait avoir des débouchés médicaux puisque le nickel est un élément rare, mais indispensable à de nombreux pathogènes humains. Dans un premier temps, elle pourrait permettre des avancées pour les maladies impliquant des pathogènes possédant un transporteur de nickel du type NikABCDE tel que *B. suis* ou *H. hepaticus*. Mais on peut se demander si les transporteurs basés sur des perméases aux métaux de transition ne fonctionnent pas également avec un « métallophore » que l'on n'aurait pas encore identifié. Dans ce cas, ces études s'élargiraient à un nombre important de pathogènes. Au vu du nombre croissant de résistances aux antibiotiques qui apparaissent chez les microorganismes pathogènes, il devient indispensable de rechercher de nouvelles cibles thérapeutiques, et ainsi développer de nouvelles molécules actives.

La biorémédiation est un autre domaine où la meilleure connaissance du transport du nickel par les bactéries pourrait avoir son importance, puisque l'on pourrait alors l'utiliser pour permettre la décontamination par des bactéries de milieux pollués par le nickel.

BIBLIOGRAPHIE



A

Aberdam, A. (1998). "SEDEM, a software package for EXAFS data extraction and modeling." J. Synchrotron Rad. **5**: 1287-1297.

Abergel, C. (2004). "Spectacular improvement of X-ray diffraction through fast desiccation of protein crystals." Acta Crystallogr D Biol Crystallogr **60**(Pt 8): 1413-6.

Adams, M. W., Mortenson, L. E. and Chen, J. S. (1981). "Hydrogenase." Biochim. Biophys. Acta. **594**: 105-176.

Ahn, B. E., Cha, J., Lee, E. J., Han, A. R., Thompson, C. J. and Roe, J. H. (2006). "Nur, a nickel-responsive regulator of the Fur family, regulates superoxide dismutases and nickel transport in Streptomyces coelicolor." Mol Microbiol **59**(6): 1848-58.

Al-Mjeni, F., Ju, T., Pochapsky, T. C. and Maroney, M. J. (2002). "XAS investigation of the structure and function of Ni in acireductone dioxygenase." Biochemistry **41**(21): 6761-9.

Allan, C. B., Wu, L. F., GU, Z., Choudhury, S. B., Al-Mjeni, F., Sharma, M. L., Mandrand-Berthelot, M. A. and Maroney, M. J. (1998). "An X-ray absorption spectroscopie structural investigation of the nickel site in Escherichia coli NikA protein." Inorg Chem **37**: 5952-5955.

Ames, G. F., Prody, C. and Kustu, S. (1984). "Simple, rapid, and quantitative release of periplasmic proteins by chloroform." J Bacteriol **160**(3): 1181-3.

ASPC http://www.phac-aspc.gc.ca.

B

Baader, E. W. (1937). "Berufkrebs." Neu. Ergeb. Geb. Krebskranhk: 104-128.

Bartha, R. and Ordal, E. J. (1965). "Nickel-Dependent Chemolithotrophic Growth Of Two Hydrogenomonas Strains." J Bacteriol **89**: 1015-9.

Bauerfeind, P., Garner, R. M. and Mobley, L. T. (1996). "Allelic exchange mutagenesis of nixA in Helicobacter pylori results in reduced nickel transport and urease activity." Infect Immun **64**(7): 2877-80.

Blokesch, M., Magalon, A. and Bock, A. (2001). "Interplay between the specific chaperone-like proteins HybG and HypC in maturation of hydrogenases 1, 2, and 3 from Escherichia coli." J Bacteriol **183**(9): 2817-22.

Bloom, S. L. and Zamble, D. B. (2004). "Metal-selective DNA-binding response of Escherichia coli NikR." Biochemistry **43**(31): 10029-38.

Blundell, T. L. and Johnson, L. N. (1976). Protein Crystallography. New-York, Academic Press Inc.

Bonhivers, M., Desmadril, M., Moeck, G. S., Boulanger, P., Colomer-Pallas, A. and Letellier, L. (2001). "Stability studies of FhuA, a two-domain outer membrane protein from Escherichia coli." Biochemistry **40**(8): 2606-13.

Bowie, J. U., Reidhaar-Olson, J. F., Lim, W. A. and Sauer, R. T. (1990). "Deciphering the message in protein sequences: tolerance to amino acid substitutions." Science **247**(4948): 1306-10.

Boyer, E., Bergevin, I., Malo, D., Gros, P. and Cellier, M. F. (2002). "Acquisition of Mn(II) in addition to Fe(II) is required for full virulence of Salmonella enterica serovar Typhimurium." Infect Immun **70**(11): 6032-42.

Braun, V. (1995). "Energy-coupled transport and signal transduction through the gram-negative outer membrane via TonB-ExbB-ExbD-dependent receptor proteins." FEMS Microbiol Rev **16**(4): 295-307.

Brausam, A. and van Eldik, R. (2004). "Further mechanistic information on the reaction between FeIII(edta) and hydrogen peroxide: observation of a second reaction step and importance of pH." Inorg Chem **43**(17): 5351-9.

Brunger, A. T. (1992a). "Free R value: a novel statistical quantity for assesseing the accuracy of crystal structure." Nature **355**: 475.

Brunger, A. T. (1992b). X-PLOR version 3.1, a system for X-ray Crystallography and NMR. New Haven and London, Yale University press.

Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. and Warren, G. L. (1998). "Crystallography & NMR system: A new software suite for macromolecular structure determination." Acta Crystallogr D Biol Crystallogr **54 (Pt 5)**: 905-21.

Budisa, N., Steipe, B., Demange, P., Eckerskorn, C., Kellermann, J. and Huber, R. (1995). "High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in Escherichia coli." Eur J Biochem **230**(2): 788-96.

Bury-Mone, S., Thiberge, J. M., Contreras, M., Maitournam, A., Labigne, A. and De Reuse, H. (2004). "Responsiveness to acidity via metal ion regulators mediates virulence in the gastric pathogen Helicobacter pylori." Mol Microbiol **53**(2): 623-38.

C

Carrington, P. E., Chivers, P. T., Al-Mjeni, F., Sauer, R. T. and Maroney, M. J. (2003). "Nickel coordination is regulated by the DNA-bound state of NikR." Nat Struct Biol **10**(2): 126-30.

Cavani, A. (2005). "Breaking tolerance to nickel." Toxicology **209**(2): 119-21.

CCHST http://www.cchst.ca/reponsessst/diseases/allergic_derm.html.

CCP4, C. C. P. N. (1994). "The CCP4 suite: Programs for Protein Crystallography." Acta Cryst D **50**: 760-763.

Charon, M. H., Wu, L. F., Piras, C., de Pina, K., Mandrand-Berthelot, M. A. and Fontecilla-Camps, J. C. (1994). "Crystallization and preliminary X-ray diffraction study of the nickel-binding protein NikA of Escherichia coli." J Mol Biol **243**(2): 353-5.

Chen, Y. Y. and Burne, R. A. (2003). "Identification and characterization of the nickel uptake system for urease biogenesis in Streptococcus salivarius 57.I." J Bacteriol **185**(23): 6773-9.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G. and Thompson, J. D. (2003). "Multiple sequence alignment with the Clustal series of programs." Nucleic Acids Res **31**(13): 3497-500.

Cherrier, M. V., Martin, L., Cavazza, C., Jacquamet, L., Lemaire, D., Gaillard, J. and Fontecilla-Camps, J. C. (2005). "Crystallographic and spectroscopic evidence for high affinity binding of FeEDTA(H2O)- to the periplasmic nickel transporter NikA." J Am Chem Soc **127**(28): 10075-82.

chimie, S. f. d. http://www.sfc.fr/Donnees/metaux/ni/texni.htm.

Chivers, P. T. and Sauer, R. T. (1999). "NikR is a ribbon-helix-helix DNA-binding protein." Protein Sci **8**(11): 2494-500.

Chivers, P. T. and Sauer, R. T. (2000). "Regulation of high affinity nickel uptake in bacteria. Ni2+-Dependent interaction of NikR with wild-type and mutant operator sites." J Biol Chem **275**(26): 19735-41.

Chivers, P. T. and Sauer, R. T. (2002). "NikR repressor: high-affinity nickel binding to the C-terminal domain regulates binding to operator DNA." Chem Biol **9**(10): 1141-8.

Chivers, P. T. and Tahirov, T. H. (2005). "Structure of Pyrococcus horikoshii NikR: nickel sensing and implications for the regulation of DNA recognition." J Mol Biol **348**(3): 597-607.

Ciurli, S., Safarov, N., Miletti, S., Dikiy, A., Christensen, S. K., Kornetzky, K., Bryant, D. A., Vandenberghe, I., Devreese, B., Samyn, B., Remaut, H. and van Beeumen, J. (2002). "Molecular characterization of Bacillus pasteurii UreE, a metal-binding chaperone for the assembly of the urease active site." J Biol Inorg Chem **7**(6): 623-31.

Clugston, S. L., Barnard, J. F., Kinach, R., Miedema, D., Ruman, R., Daub, E. and Honek, J. F. (1998). "Overproduction and characterization of a dimeric non-zinc glyoxalase I from Escherichia coli: evidence for optimal activation by nickel ions." Biochemistry **37**(24): 8754-63.

Clugston, S. L. and Honek, J. F. (2000). "Identification of sequences encoding the detoxification metalloisomerase glyoxalase I in microbial genomes from several pathogenic organisms." J Mol Evol **50**(5): 491-5.

Colbeau, A., Elsen, S., Tomiyama, M., Zorin, N. A., Dimon, B. and Vignais, P. M. (1998). "Rhodobacter capsulatus HypF is involved in regulation of hydrogenase synthesis through the HupUV proteins." Eur J Biochem **251**(1-2): 65-71.

Colpas, G. J., Maroney, M. J., Bagyinka, C., Kumar, M., Willis, W. S., Suib, S. L., Baidya, N. and Mascharak, P. K. (1990). "X-ray Spectroscopic Studies of Nickel Complexes, with Application to the Structure of Nickel Sites in Hydrogenases." Inorg Chem **30**: 920-928.

Contreras, M., Thiberge, J. M., Mandrand-Berthelot, M. A. and Labigne, A. (2003). "Characterization of the roles of NikR, a nickel-responsive pleiotropic autoregulator of Helicobacter pylori." Mol Microbiol **49**(4): 947-63.

Cotton, F. A. and Wilkinson, G. Iron. Advanced Inorganic chemistry, Wiley Interscience**:** 711.

Cotton, F. A. and Wilkinson, G. nickel. Advanced Inorganic chemistry, Wiley Interscience**:** 743.

Cromer, D. T. and Liberman, D. (1970). "Relativistic calculation of anomalous scattering factors for x-rays." J. Chem. Phys. **53**: 1891-1898.

Crowther, R. A. and Blow, D. M. (1967). "A method for positioning a know molecule in a unknown crystal structure." Acta Cryst **23**: 544-548.

D

Dai, Y., Wensink, P. C. and Abeles, R. H. (1999). "One protein, two enzymes." J Biol Chem **274**(3): 1193-5.

Darnault, C., Volbeda, A., Kim, E. J., Legrand, P., Vernede, X., Lindahl, P. A. and Fontecilla-Camps, J. C. (2003). "Ni-Zn-[Fe4-S4] and Ni-Ni-[Fe4-S4] clusters in closed and open subunits of acetyl-CoA synthase/carbon monoxide dehydrogenase." Nat Struct Biol **10**(4): 271-9.

de Pina, K., Desjardin, V., Mandrand-Berthelot, M. A., Giordano, G. and Wu, L. F. (1999). "Isolation and characterization of the nikR gene encoding a nickel-responsive regulator in Escherichia coli." J Bacteriol **181**(2): 670-4.

de Pina, K., Navarro, C., McWalter, L., Boxer, D. H., Price, N. C., Kelly, S. M., Mandrand-Berthelot, M. A. and Wu, L. F. (1995). "Purification and characterization of the periplasmic nickel-binding protein NikA of Escherichia coli K12." Eur J Biochem **227**(3): 857-65.

Degen, O., Kobayashi, M., Shimizu, S. and Eitinger, T. (1999). "Selective transport of divalent cations by transition metal permeases: the Alcaligenes eutrophus HoxN and the Rhodococcus rhodochrous NhlF." Arch Microbiol **171**(3): 139-45.

DeLano, W. L. (2002). http://www.pymol.org.

Denkhaus, E. and Salnikow, K. (2002). "Nickel essentiality, toxicity, and carcinogenicity." Crit Rev Oncol Hematol **42**(1): 35-56.

Dixon, N. E., Gazzola, T. C., blakeley, R. L. and Zermer, B. (1975). "Letter: Jack bean urease (EC 3.5.1.5). A metalloenzyme. A simple biological role for nickel?" J Am Chem Soc **97**(14): 4131-3.

Dobbek, H., Svetlitchnyi, V., Gremer, L., Huber, R. and Meyer, O. (2001). "Crystal structure of a carbon monoxide dehydrogenase reveals a [Ni-4Fe-5S] cluster." Science **293**(5533): 1281-5.

Dosanjh, N. S. and Michel, S. L. (2006). "Microbial nickel metalloregulation: NikRs for nickel ions." Curr Opin Chem Biol.

Drapal, N. and Bock, A. (1998). "Interaction of the hydrogenase accessory protein HypC with HycE, the large subunit of Escherichia coli hydrogenase 3 during enzyme maturation." Biochemistry **37**(9): 2941-8.

Drennan, C. L., Heo, J., Sintchak, M. D., Schreiter, E. and Ludden, P. W. (2001). "Life on carbon monoxide: X-ray structure of Rhodospirillum rubrum Ni-Fe-S carbon monoxide dehydrogenase." Proc Natl Acad Sci U S A **98**(21): 11973-8.

Drenth, J. Principles of protein X-Ray Crystallography., Springer.

Dunten, P. and Mowbray, S. L. (1995). "Crystal structure of the dipeptide binding protein from Escherichia coli involved in active transport and chemotaxis." Protein Sci **4**(11): 2327-34.

E

Eberz, G., Eitinger, T. and Friedrich, B. (1989). "Genetic determinants of a nickel-specific transport system are part of the plasmid-encoded hydrogenase gene cluster in Alcaligenes eutrophus." J Bacteriol **171**(3): 1340-5.

Eidsness, M. K., Scott, R. A., Prickril, B. C., DerVartanian, D. V., Legall, J., Moura, I., Moura, J. J. and Peck, H. D., Jr. (1989). "Evidence for selenocysteine coordination to the active site nickel in the [NiFeSe]hydrogenases from Desulfovibrio baculatus." Proc Natl Acad Sci U S A **86**(1): 147-51.

Eitinger, T. and Friedrich, B. (1991). "Cloning, nucleotide sequence, and heterologous expression of a high-affinity nickel transport gene from Alcaligenes eutrophus." J Biol Chem **266**(5): 3222-7.

Eitinger, T. and Friedrich, B. (1994). "A topological model for the high-affinity nickel transporter of Alcaligenes eutrophus." Mol Microbiol **12**(6): 1025-32.

Eitinger, T. and Friedrich, B. (1997). Microbial nickel transport and incorporation into hydrogenases. Transition metal in microbial metabolism. G. Winkelmann and C. J. Carrano. Harwood, Amsterdam**:** 235-256.

Eitinger, T. and Mandrand-Berthelot, M. A. (2000). "Nickel transport systems in microorganisms." Arch Microbiol **173**(1): 1-9.

Eitinger, T., Suhr, J., Moore, L. and Smith, J. A. (2005). "Secondary transporters for nickel and cobalt ions: theme and variations." Biometals **18**(4): 399-405.

Eitinger, T., Wolfram, L., Degen, O. and Anthon, C. (1997). "A Ni2+ binding motif is the basis of high affinity transport of the Alcaligenes eutrophus nickel permease." J Biol Chem **272**(27): 17139-44.

Emsley, P. and Cowtan, K. (2004). "Coot: model-building tools for molecular graphics." Acta Crystallogr D Biol Crystallogr **60**(Pt 12 Pt 1): 2126-32.

Engh, R. A. and Huber, R. (1991). "Accurate bound and angle parameters for X-ray structure refinement." Acta Cryst A **47**: 392-400.

Ermler, U., Grabarse, W., Shima, S., Goubeaud, M. and Thauer, R. K. (1997). "Crystal structure of methyl-coenzyme M reductase: the key enzyme of biological methane formation." Science **278**(5342): 1457-62.

F

Faraldo-Gomez, J. D. and Sansom, M. S. (2003). "Acquisition of siderophores in gram-negative bacteria." Nat Rev Mol Cell Biol **4**(2): 105-16.

Ferguson, A. D., Chakraborty, R., Smith, B. S., Esser, L., van der Helm, D. and Deisenhofer, J. (2002). "Structural basis of gating by the outer membrane transporter FecA." Science **295**(5560): 1715-9.

Ferry, J. G. (1995). "CO dehydrogenase." Annu Rev Microbiol **49**: 305-33.

Fish, W. W. (1988). "Rapid colorimetric micromethod for the quantitation of complexed iron in biological samples." Methods Enzymol **158**: 357-64.

Fricke, H. (1920). Phys. Rev. **16**: 202.

Fritsche, E., Paschos, A., Beisel, H. G., Bock, A. and Huber, R. (1999). "Crystal structure of the hydrogenase maturating endopeptidase HYBD from Escherichia coli." J Mol Biol **288**(5): 989-98.

Fu, C., Javedan, S., Moshiri, F. and Maier, R. J. (1994). "Bacterial genes involved in incorporation of nickel into a hydrogenase enzyme." Proc Natl Acad Sci U S A **91**(11): 5099-103.

Fu, C., Olson, J. W. and Maier, R. J. (1995). "HypB protein of Bradyrhizobium japonicum is a metal-binding GTPase capable of binding 18 divalent nickel ions per dimer." Proc Natl Acad Sci U S A **92**(6): 2333-7.

Fulkerson, J. F., Jr. and Mobley, H. L. (2000). "Membrane topology of the NixA nickel transporter of Helicobacter pylori: two nickel transport-specific motifs within transmembrane helices II and III." J Bacteriol **182**(6): 1722-30.

G

Gardina, P. J., Bormans, A. F., Hawkins, M. A., Meeker, J. W. and Manson, M. D. (1997). "Maltose-binding protein interacts simultaneously and asymmetrically with both subunits of the Tar chemoreceptor." Mol Microbiol **23**(6): 1181-91.

Garner, C. D. (1991). "X-ray absorption spectroscopy and the structures of transition metal centers en proteins." Adv. Inorg. Chem. **36**: 303-339.

Ge, R., Watt, R. M., Sun, X., Tanner, J. A., He, Q. Y., Huang, J. D. and Sun, H. (2006). "Expression and characterization of a histidine-rich protein, Hpn: potential for Ni2+ storage in Helicobacter pylori." Biochem J **393**(Pt 1): 285-93.

Gouet, P., Courcelle, E., Stuart, D. I. and Metoz, F. (1999). "ESPript: analysis of multiple sequence alignments in PostScript." Bioinformatics **15**(4): 305-8.

Grabarse, W., Mahlert, F., Duin, E. C., Goubeaud, M., Shima, S., Thauer, R. K., Lamzin, V. and Ermler, U. (2001). "On the mechanism of biological methane formation: structural evidence for conformational changes in methyl-coenzyme M reductase upon substrate binding." J Mol Biol **309**(1): 315-30.

Grabarse, W., Mahlert, F., Shima, S., Thauer, R. K. and Ermler, U. (2000). "Comparison of three methyl-coenzyme M reductases from phylogenetically distant organisms: unusual amino acid modification, conservation and adaptation." J Mol Biol **303**(2): 329-44.

Grimsrud, T. K., Berge, S. R., Martinsen, J. I. and Andersen, A. (2003). "Lung cancer incidence among Norwegian nickel-refinery workers 1953-2000." J Environ Monit **5**(2): 190-7.

H

Ha, N. C., Oh, S. T., Sung, J. Y., Cha, K. A., Lee, M. H. and Oh, B. H. (2001). "Supramolecular assembly and acid resistance of Helicobacter pylori urease." Nat Struct Biol **8**(6): 505-9.

Hausinger, R. P. (1993). Biochemistry of nickel. New York, Plenum.

Hausinger, R. P. and Karplus, P. A. (2001). Urease in Handbook of Metalloproteins. K. Wieghardt, R. Huber, T. L. Poulos, A. Messerschmidt and Eds. Wiley, West Sussex**:** 867-879.

He, M. M., Clugston, S. L., Honek, J. F. and Matthews, B. W. (2000). "Determination of the structure of Escherichia coli glyoxalase I suggests a structural basis for differential metal activation." Biochemistry **39**(30): 8719-27.

Healey, F. P. (1970). "Hydrogen evolution by several algae." Plant (Berl.) **91**: 220-226.

Heddle, J., Scott, D. J., Unzai, S., Park, S. Y. and Tame, J. R. (2003). "Crystal structures of the liganded and unliganded nickel-binding protein NikA from Escherichia coli." J Biol Chem **278**(50): 50322-9.

Hendricks, J. K. and Mobley, H. L. (1997). "Helicobacter pylori ABC transporter: effect of allelic exchange mutagenesis on urease activity." J Bacteriol **179**(18): 5892-902.

Hendrickson, W. A., Smith, J. L. and Sheriff, S. (1985). "Direct phase determination based on anomalous scattering." Methods Enzymol **115**: 41-55.

Heras, B., Edeling, M. A., Byriel, K. A., Jones, A., Raina, S. and Martin, J. L. (2003). "Dehydration converts DsbG crystal diffraction from low to high resolution." Structure **11**(2): 139-45.

Heras, B. and Martin, J. L. (2005). "Post-crystallization treatments for improving diffraction quality of protein crystals." Acta Crystallogr D Biol Crystallogr **61**(Pt 9): 1173-80.

HERCULES Neutron And Synchrotron Radiation For Condensed Matter Studies. Neutron And Synchrotron Radiation For Condensed Matter Studies. J. Baruchel, J. L. Hodeau, M. S. Lehmann, R. J.R. and C. Schlenker, Springer-Verlag. **Volume 1: Theory, Instruments and Methods:** 179.

Hertz, G. (1920). Zeit. F. Physik. **3**: 19.

Hery, M., Herrera, A., Vogel, T. M., Normand, P. and Navarro, E. (2005). "Effect of carbon and nitrogen input on the bacterial community structure of Neocaledonian nickel mine spoils." FEMS Microbiol Ecol **51**(3): 333-40.

Hery, M., Nazaret, S., Jaffre, T., Normand, P. and Navarro, E. (2003). "Adaptation to nickel spiking of bacterial communities in neocaledonian soils." Environ Microbiol **5**(1): 3-12.

Hostynek, J. J. and Maibach, H. I. (2002). Nickel and the Skin. C. Press, Boca Raton**:** 1-249.

I

IARC, I. A. f. R. o. C. (1990). Monographs on the Evaluation of Carcinogenic Risks to Humans. Lyon, IARC Scientific Publication. **49, Chromium Nickel and Welding:** 257-445.

Ifeffit http://cars9.uchicago.edu/ifeffit/.

J

Jabri, E. and Karplus, P. A. (1996). "Structures of the Klebsiella aerogenes urease apoenzyme and two active-site mutants." Biochemistry **35**(33): 10616-26.

Jacquamet, L. (1999). Etude structurale de la protéine FUR (Ferric Uptake Regulation) d'escherichia coli par spectroscopie d'absorption des rayons X: XANES et EXAFS. chimie. Grenoble, UJF.

Jubier-Maurin, V., Rodrigue, A., Ouahrani-Bettache, S., Layssac, M., Mandrand-Berthelot, M. A., Kohler, S. and Liautard, J. P. (2001). "Identification of the nik gene cluster of Brucella suis: regulation and contribution to urease activity." J Bacteriol **183**(2): 426-34.

K

Kabsch, W. (1988). J. Appl. Cryst. **21**: 916-924.

Kabsch, W. (1993). "Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants." J. Appl. Cryst. **26**: 795-800.

Kammler, M., Schon, C. and Hantke, K. (1993). "Characterization of the ferrous iron uptake system of Escherichia coli." J Bacteriol **175**(19): 6212-9.

Kantardjieff, K. A. and Rupp, B. (2003). "Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals." Protein Sci **12**(9): 1865-71.

Kasprzak, K. S., Sunderman, F. W., Jr. and Salnikow, K. (2003). "Nickel carcinogenesis." Mutat Res **533**(1-2): 67-97.

Kim, H. and Maier, R. J. (1990). "Transcriptional regulation of hydrogenase synthesis by nickel in Bradyrhizobium japonicum." J Biol Chem **265**(31): 18729-32.

Kim, H., Yu, C. and Maier, R. J. (1991). "Common cis-acting region responsible for transcriptional regulation of Bradyrhizobium japonicum hydrogenase by nickel, oxygen, and hydrogen." J Bacteriol **173**(13): 3993-9.

Kim, J. S., Kang, S. O. and Lee, J. K. (2003). "The protein complex composed of nickel-binding SrnQ and DNA binding motif-bearing SrnR of Streptomyces griseus represses sodF transcription in the presence of nickel." J Biol Chem **278**(20): 18455-63.

Klebba, P. E., Rutz, J. M., Liu, J. and Murphy, C. K. (1993). "Mechanisms of TonB-catalyzed iron transport through the enteric bacterial cell envelope." J Bioenerg Biomembr **25**(6): 603-11.

Kleywegt, G. J. and Brunger, A. T. (1996). "Checking your imagination: applications of the free R value." Structure **4**(8): 897-904.

Knight, K. L., Bowie, J. U., Vershon, A. K., Kelley, R. D. and Sauer, R. T. (1989). "The Arc and Mnt repressors. A new class of sequence-specific DNA-binding protein." J Biol Chem **264**(7): 3639-42.

Koster, W. (2001). "ABC transporter mediate update of iron, siderophores, heme and vitamin B12." Res. Microbiol. **152**: 291-301.

Kronig, R. d. L. (1932). Zeit. F. Physik. **75**: 468.

L

Leach, C. A., Jr. and Sunderman, F. W., Jr. (1987). "Hypernickelemia following coronary arteriography, caused by nickel in the radiographic contrast medium." Ann Clin Lab Sci **17**(3): 137-44.

Leach, C. N., Jr. and Sunderman, F. W., Jr. (1985). "Nickel contamination of human serum albumin solutions." N Engl J Med **313**(19): 1232.

Leahy, D. J., Erickson, H. P., Aukhil, I., Joshi, P. and Hendrickson, W. A. (1994). "Crystallization of a fragment of human fibronectin: introduction of methionine by site-directed mutagenesis to allow phasing via selenomethionine." Proteins **19**(1): 48-54.

Leclere, V., Boiron, P. and Blondeau, R. (1999). "News & notes: diversity of superoxide-dismutases among clinical and soil isolates of Streptomyces species." Curr Microbiol **39**(6): 365-0368.

Lee, M. H., Pankratz, H. S., Wang, S., Scott, R. A., Finnegan, M. G., Johnson, M. K., Ippolito, J. A., Christianson, D. W. and Hausinger, R. P. (1993). "Purification and characterization of Klebsiella aerogenes UreE protein: a nickel-binding protein that functions in urease metallocenter assembly." Protein Sci **2**(6): 1042-52.

Lim, W. A., Farruggio, D. C. and Sauer, R. T. (1992). "Structural and energetic consequences of disruptive mutations in a protein core." Biochemistry **31**(17): 4324-33.

Lind, M. D., Hamor, M. J., Hamor, T. A. and Hoard, J. L. (1964). Inorg Chem **3**: 34-43.

Lindahl, P. A. (2002). "The Ni-containing carbon monoxide dehydrogenase family: light at the end of the tunnel?" Biochemistry **41**(7): 2097-105.

Linton, K. J. and Higgins, C. F. (1998). "The Escherichia coli ATP-binding cassette (ABC) proteins." Mol Microbiol **28**(1): 5-13.

Locher, K. P. (2004). "Structure and mechanism of ABC transporters." Curr Opin Struct Biol **14**: 426-431.

Lyon, E. J., Shima, S., Boecher, R., Thauer, R. K., Grevels, F. W., Bill, E., Roseboom, W. and Albracht, S. P. (2004). "Carbon monoxide as an intrinsic ligand to iron in the active site of the iron-sulfur-cluster-free hydrogenase H2-forming methylenetetrahydromethanopterin dehydrogenase as revealed by infrared spectroscopy." J Am Chem Soc **126**(43): 14239-48.

M

Madden, T. L., Tatusov, R. L. and Zhang, J. (1996). "Applications of network BLAST server." Methods Enzymol **266**: 131-41.

Magalan, A. and Bock, A. (2000). "Dissection of the maturation reaction of the [NiFe] hydrogenase 3 from E. coli taking place after nickel incorporation." FEBS Lett **473**: 254-258.

Maier, R. J. (2005). "Use of molecular hydrogen as an energy substrate by human pathogenic bacteria." Biochem Soc Trans **33**(Pt 1): 83-5.

Maier, T., Jacobi, A., Sauter, M. and Bock, A. (1993). "The product of the hypB gene, which is required for nickel incorporation into hydrogenases, is a novel guanine nucleotide-binding protein." J Bacteriol **175**(3): 630-5.

Maier, T., Lottspeich, F. and Bock, A. (1995). "GTP hydrolysis by HypB is essential for nickel insertion into hydrogenases of Escherichia coli." Eur J Biochem **230**(1): 133-8.

Maris, A. E., Sawaya, M. R., Kaczor-Grzeskowiak, M., Jarvis, M. R., Bearson, S. M., Kopka, M. L., Schroder, I., Gunsalus, R. P. and Dickerson, R. E. (2002). "Dimerization allows DNA target site recognition by the NarL response regulator." Nat Struct Biol **9**(10): 771-8.

Maroney, M. J. (1999). "Structure/function relationships in nickel metallobiochemistry." Curr Opin Chem Biol **3**(2): 188-99.

Marquardt, D. W. (1963). "An algorithm for leasr-squares estimation of nonlinear parametres." J. Soc. Indus. Appl. Math. **11**: 431-441.

Matthews, B. W. (1968). "Solvent content of protein crystals." J Mol Biol **33**(2): 491-7.

McGregor, D. B., Baan, R. A., Partensky, C., Rice, J. M. and Wilbourn, J. D. (2000). "Evaluation of the carcinogenic risks to humans associated with surgical implants and other foreign bodies - a report of an IARC Monographs Programme Meeting. International Agency for Research on Cancer." Eur J Cancer **36**(3): 307-13.

McPherson, A., Malkin, A. J. and Kuznetsov, Y. G. (1995). "The science of macromolecular crystallization." Structure **3**(8): 759-68.

Mizuno, M., Funahashi, S., Nakasuka, N. and Tanaka, M. (1991). Inorg Chem **30**: 1550-1553.

Mobley, H. L. (1996). "The role of Helicobacter pylori urease in the pathogenesis of gastritis and peptic ulceration." Aliment Pharmacol Ther **10 Suppl 1**: 57-64.

Mobley, H. L., Garner, R. M. and Bauerfeind, P. (1995). "Helicobacter pylori nickel-transport gene nixA: synthesis of catalytically active urease in Escherichia coli independent of growth conditions." Mol Microbiol **16**(1): 97-109.

Mobley, H. L. and Hausinger, R. P. (1989). "Microbial ureases: significance, regulation, and molecular characterization." Microbiol Rev **53**(1): 85-108.

Mobley, H. L., Island, M. D. and Hausinger, R. P. (1995). "Molecular biology of microbial ureases." Microbiol Rev **59**(3): 451-80.

Moeck, G. S. and Coulton, J. W. (1998). "TonB-dependent iron acquisition: mechanisms of siderophore-mediated active transport." Mol Microbiol **28**(4): 675-81.

Moncrief, M. B. and Hausinger, R. P. (1997). "Characterization of UreG, identification of a UreD-UreF-UreG complex, and evidence suggesting that a nucleotide-binding site in UreG is required for in vivo metallocenter assembly of Klebsiella aerogenes urease." J Bacteriol **179**(13): 4081-6.

Muller, K., Matzanke, B. F., Schunemann, V., Trautwein, A. X. and Hantke, K. (1998). "FhuF, an iron-regulated protein of Escherichia coli with a new type of [2Fe-2S] center." Eur J Biochem **258**(3): 1001-8.

Mulrooney, S. B. and Hausinger, R. P. (2003). "Nickel uptake and utilization by microorganisms." FEMS Microbiol Rev **27**(2-3): 239-61.

Murshudov, G. N., Vagin, A. A. and Dodson, E. J. (1997). "Refinement of macromolecular structures by the maximum-likelihood method." Acta Crystallogr D Biol Crystallogr **53**(Pt 3): 240-55.

$\mathcal{N}$

Navarro, C., Wu, L. F. and Mandrand-Berthelot, M. A. (1993). "The nik operon of Escherichia coli encodes a periplasmic binding-protein-dependent transport system for nickel." Mol Microbiol **9**(6): 1181-91.

Navaza, J. (1994). "AMoRe: an automated pakage for molecular replacement." Acta. Cryst. A **50**: 157-163.

Norseth, T. and Piscator, M. (1979). Handbook on the Toxicology of Metals. E. N.-H. B. Press**:** 541-553.

NSA, N. A. o. S. (1975). Nickel, Medical and Biologic Effects of Environemental Pollutants. Washington, DC, NSA press**:** 1-277.

P

Park, I. S. and Hausinger, R. P. (1995). "Evidence for the presence of urease apoprotein complexes containing UreD, UreF, and UreG in cells that are competent for in vivo enzyme activation." J Bacteriol **177**(8): 1947-51.

Park, K. S., Iida, T., Yamaichi, Y., Oyagi, T., Yamamoto, K. and Honda, T. (2000). "Genetic characterization of DNA region containing the trh and ure genes of Vibrio parahaemolyticus." Infect Immun **68**(10): 5742-8.

Pautsch, A. and Schulz, G. E. (1998). "Structure of the outer membrane protein A transmembrane domain." Nat Struct Biol **5**(11): 1013-7.

Payne, M. J., Chapman, A. and Cammack, R. (1993). "Evidence for an [Fe]-type hydrogenase in the parasitic protozoan Trichomonas vaginalis." FEBS Lett **317**(1-2): 101-4.

Pearson, M. A., Schaller, R. A., Michel, L. O., Karplus, P. A. and Hausinger, R. P. (1998). "Chemical rescue of Klebsiella aerogenes urease variants lacking the carbamylated-lysine nickel ligand." Biochemistry **37**(17): 6214-20.

Pelmont, J. (1993). Bactéries et environement, adaptations physiologiques, PUG.

Pochapsky, T. C., Pochapsky, S. S., Ju, T., Mo, H., Al-Mjeni, F. and Maroney, M. J. (2002). "Modeling and experiment yields the structure of acireductone dioxygenase from Klebsiella pneumoniae." Nat Struct Biol **9**(12): 966-72.

Postle, K. (1993). "TonB protein and energy transduction between membranes." J Bioenerg Biomembr **25**(6): 591-601.

R

Ragsdale, S. W. and Kumar, M. (1996). "Nickel-Containing Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase." Chem Rev **96**(7): 2515-2540.

Ramakrishnan, C. and Ramachandran, G. N. (1965). "Stereochemical criteria for polypeptide and protein chain conformations. II. Allowed conformations for a pair of peptide units." Biophys J **5**(6): 909-33.

Read, R. J. (2001). "Pushing the boundaries of molecular replacement with maximum likelihood." Acta Crystallogr D Biol Crystallogr **57**(Pt 10): 1373-82.

Rodionov, D. A., Hebbeln, P., Gelfand, M. S. and Eitinger, T. (2006). "Comparative and Functional Genomic Analysis of Prokaryotic Nickel and Cobalt Uptake Transporters: Evidence for a Novel Group of ATP-Binding Cassette Transporters." J Bacteriol **188**(1): 317-27.

Rodrigue, A., Batia, N., Muller, M., Fayet, O., Bohm, R., Mandrand-Berthelot, M. A. and Wu, L. F. (1996). "Involvement of the GroE chaperonins in the nickel-dependent anaerobic biosynthesis of NiFe-hydrogenases of Escherichia coli." J Bacteriol **178**(15): 4453-60.

Rodrigue, A., Effantin, G. and Mandrand-Berthelot, M. A. (2005). "Identification of rcnA (yohM), a nickel and cobalt resistance gene in Escherichia coli." J Bacteriol **187**(8): 2912-6.

Rossman, M. G. and Blow, D. M. (1962). "The detection of sub-unit within the crystallographic asymetric unit." Acta. Cryst. **15**: 24-31.

Roussel, A. and Cambillaud, C. (1989). "Turbo-Frodo." Silicon Graphics **77**.

Rowe, J. L., Starnes, G. L. and Chivers, P. T. (2005). "Complex transcriptional control links NikABCDE-dependent nickel transport with hydrogenase expression in Escherichia coli." J Bacteriol **187**(18): 6317-23.

S

Salins, L. L., Goldsmith, E. S., Ensor, C. M. and Daunert, S. (2002). "A fluorescence-based sensing system for the environmental monitoring of nickel using the nickel binding protein from Escherichia coli." Anal Bioanal Chem **372**(1): 174-80.

Sayers, D. E., Lytle, F. W. and Stern, E. A. (1970). "Point scattering therory of x-ray absorption fine structure." Adv. X-ray Anal. **13**: 248-271.

Sayers, D. E., Stern, E. A. and Lytle, F. W. (1971). "New technique for investigation noncrystalline structures: fourier analysis of the extended x-ray absorption fine structure." Phys. Rev. Lett. **27**: 1204-1207.

Schalk, I. J., Yue, W. W. and Buchanan, S. K. (2004). "Recognition of iron-free siderophores by TonB-dependent iron transporters." Mol Microbiol **54**(1): 14-22.

Schneppensieper, T., Liehr, G., van Eldik, R., Ensling, J. and Gutlich, P. (2000). "A new cylindrical, six-membered iron(III) inclusion cluster consisting of three oxo-bridged diiron subunits linked by carboxylate bridges." Inorg Chem **39**(24): 5565-8.

Schreiter, E. R., Sintchak, M. D., Guo, Y., Chivers, P. T., Sauer, R. T. and Drennan, C. L. (2003). "Crystal structure of the nickel-responsive transcription factor NikR." Nat Struct Biol **10**(10): 794-9.

Scott, R. A. (1988). "Measurement of metal-ligand distances by EXAFS." Methods Enzymol **117**: 414-459.

Sebbane, F., Mandrand-Berthelot, M. A. and Simonet, M. (2002). "Genes encoding specific nickel transport systems flank the chromosomal urease locus of pathogenic yersiniae." J Bacteriol **184**(20): 5706-13.

Seravalli, J. and Ragsdale, S. W. (2000). "Channeling of carbon monoxide during anaerobic carbon dioxide fixation." Biochemistry **39**(6): 1274-7.

Sharff, A. J., Koronakis, E., Luisi, B. and Koronakis, V. (2000). "Oxidation of selenomethionine: some MADness in the method!" Acta Crystallogr D Biol Crystallogr **56 (Pt 6)**: 785-8.

Shulman, R. G., Eisenberger, P., Blumberg, W. E. and Stombaugh, N. A. (1975). "Determination of the iron-sulfur distances in rubredoxin by x-ray absorption spectroscopy." Proc Natl Acad Sci U S A **72**(10): 4003-7.

Silver, S. and Phung, L. T. (1996). "Bacterial heavy metal resistance: new surprises." Annu Rev Microbiol **50**: 753-89.

Skibinski, D. A., Golby, P., Chang, Y. S., Sargent, F., Hoffman, R., Harper, R., Guest, J. R., Attwood, M. M., Berks, B. C. and Andrews, S. C. (2002). "Regulation of the hydrogenase-4 operon of Escherichia coli by the sigma(54)-dependent transcriptional activators FhlA and HyfR." J Bacteriol **184**(23): 6642-53.

Snavely, M. D., Gravina, S. A., Cheung, T. T., Miller, C. G. and Maguire, M. E. (1991). "Magnesium transport in Salmonella typhimurium. Regulation of mgtA and mgtB expression." J Biol Chem **266**(2): 824-9.

Stephenson, M. a. S. L. H. (1931). "Hydrogenase, a Bacterial Enzyme Activating Molecular Hydrogen. I. The Properties of the Enzyme." Biochem. J. (London) **25**: 205-214.

Sunderman F.W. J.r (2004). Elements and their Compounds in the Environment - Occurrence, Analysis and Biological Relevance (in 3 Volumes), Wiley-VCH.

$\mathcal{T}$

Takishima, K., Suga, T. and Mamiya, G. (1988). "The structure of jack bean urease. The complete amino acid sequence, limited proteolysis and reactive cysteine residues." Eur J Biochem **175**(1): 151-65.

Tame, J. R., Murshudov, G. N., Dodson, E. J., Neil, T. K., Dodson, G. G., Higgins, C. F. and Wilkinson, A. J. (1994). "The structural basis of sequence-independent peptide binding by OppA protein." Science **264**(5165): 1578-81.

Tickle, I. J., Laskowski, R. A. and Moss, D. S. (1998). "Rfree and the rfree ratio. I. Derivation of expected values of cross-validation residuals used in macromolecular least-squares refinement." Acta Crystallogr D Biol Crystallogr **54 (Pt 4)**: 547-57.

U

Unden, G. and Bongaerts, J. (1997). "Alternative respiratory pathways of Escherichia coli: energetics and transcriptional regulation in response to electron acceptors." Biochim Biophys Acta **1320**(3): 217-34.

V

van der Westen, H. M., Mayhew, S. G. and Veeger, C. (1978). "Separation of hydrogenase from intact cells of Desulfovibrio vulgaris. Purification and properties." FEBS Lett **86**(1): 122-6.

van Vliet, A. H., Ernst, F. D. and Kusters, J. G. (2004). "NikR-mediated regulation of Helicobacter pylori acid adaptation." Trends Microbiol **12**(11): 489-94.

van Vliet, A. H., Kuipers, E. J., Stoof, J., Poppelaars, S. W. and Kusters, J. G. (2004). "Acid-responsive gene induction of ammonia-producing enzymes in Helicobacter pylori is mediated via a metal-responsive repressor cascade." Infect Immun **72**(2): 766-73.

van Vliet, A. H., Poppelaars, S. W., Davies, B. J., Stoof, J., Bereswill, S., Kist, M., Penn, C. W., Kuipers, E. J. and Kusters, J. G. (2002). "NikR mediates nickel-responsive transcriptional induction of urease expression in Helicobacter pylori." Infect Immun **70**(6): 2846-52.

Velayudhan, J., Hughes, N. J., McColm, A. A., Bagshaw, J., Clayton, C. L., Andrews, S. C. and Kelly, D. J. (2000). "Iron acquisition and virulence in Helicobacter pylori: a major role for FeoB, a high-affinity ferrous iron transporter." Mol Microbiol **37**(2): 274-86.

Vernede, X. and Fontecilla-Camps, J. C. (1999). "A method to stabilize reduced and/or gas treated protein crystals by flash-cooling under a controlled atmosphere." J. Appl. Cryst. **32**: 505-509.

Vignais, P. M., Dimon, B., Zorin, N. A., Tomiyama, M. and Colbeau, A. (2000). "Characterization of the hydrogen-deuterium exchange activities of the energy-transducing HupSL hydrogenase and H(2)-signaling HupUV hydrogenase in Rhodobacter capsulatus." J Bacteriol **182**(21): 5997-6004.

Volbeda, A., Charon, M. H., Piras, C., Hatchikian, E. C., Frey, M. and Fontecilla-Camps, J. C. (1995). "Crystal structure of the nickel-iron hydrogenase from Desulfovibrio gigas." Nature **373**(6515): 580-7.

Volbeda, A., Fontecilla-Camps, J. C. and Frey, M. (1996). "Novel metal sites in protein structures." Curr Opin Struct Biol **6**(6): 804-12.

W

Wages, H. E., Taft, K. L. and Lippard, S. J. (1993). "[Ni2(OAc)3(urea)(tmen)2](OTf) and [Ni2(OAc)(urea)(tmen)](OTf), model complexes for thr enzyme urease." Inorg Chem **32**: 4985-4987.

Wallace, A. C., Laskowski, R. A. and Thornton, J. M. (1995). "LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions." Protein Eng **8**(2): 127-34.

Wandersman, C. and Delepelaire, P. (2004). "Bacterial iron sources: from siderophores to hemophores." Annu Rev Microbiol **58**: 611-47.

Wang, B. C. (1985). "Resolution of phase ambiguity in macromolecular crystallography." Methods Enzymol **115**: 90-112.

Wikipédia http://fr.wikipedia.org/wiki/Accueil.

Winkelmann, G. (2001). Microbial Transport Systems, Weinheim: Wiley - VCH.

Wlodawer, A. and Hodgson, K. O. (1975). "Crystallization and drystal data for monellin." Proc. Nat. Acad. Sc. **72**: 398-399.

Wolfram, L., Friedrich, B. and Eitinger, T. (1995). "The Alcaligenes eutrophus protein HoxN mediates nickel transport in Escherichia coli." J Bacteriol **177**(7): 1840-3.

Wu, L. F. and Mandrand-Berthelot, M. A. (1986). "Genetic and physiological characterization of new Escherichia coli mutants impaired in hydrogenase activity." Biochimie **68**(1): 167-79.

Wu, L. F. and Mandrand-Berthelot, M. A. (1995). "A family of homologous substrate-binding proteins with a broad range of substrate specificity and dissimilar biological functions." Biochimie **77**(9): 744-50.

Wu, L. F., Mandrand-Berthelot, M. A., Waugh, R., Edmonds, C. J., Holt, S. E. and Boxer, D. H. (1989). "Nickel deficiency gives rise to the defective hydrogenase phenotype of hydC and fnr mutants in Escherichia coli." Mol Microbiol **3**(12): 1709-18.

Wu, L. F., Navarro, C. and Mandrand-Berthelot, M. A. (1991). "The hydC region contains a multi-cistronic operon (nik) involved in nickel transport in Escherichia coli." Gene **107**(1): 37-42.

Wuerges, J., Lee, J. W., Yim, Y. I., Yim, H. S., Kang, S. O. and Carugo, K. D. (2004). "Crystal structure of nickel-containing superoxide dismutase reveals another type of active site." Proc Natl Acad Sci U S A **101**(23): 8569-74.

Y

Yeates, T. O. (1997). "Detecting and overcoming crystal twinning." Methods Enzymol **276**: 344-58.

Youn, H. D., Kim, E. J., Roe, J. H., Hah, Y. C. and Kang, S. O. (1996). "A novel nickel-containing superoxide dismutase from Streptomyces spp." Biochem J **318 (Pt 3)**: 889-96.

Youn, H. D., Youn, H., Lee, J. W., Yim, Y. I., Lee, J. K., Hah, Y. C. and Kang, S. O. (1996). "Unique isozymes of superoxide dismutase in Streptomyces griseus." Arch Biochem Biophys **334**(2): 341-8.

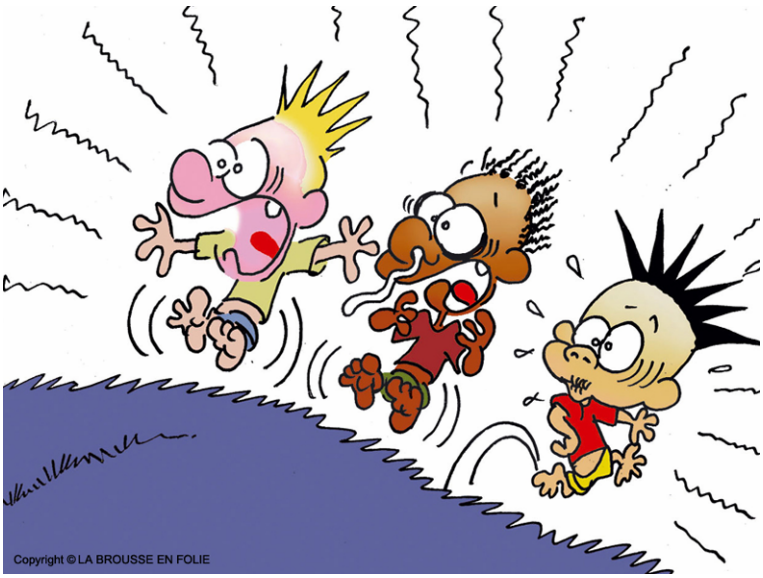
Z

Zabinsky, S. I., Rehr, J. J., Ankudinov, A., Albers, R. C. and Eller, M. J. (1995). "Multiple-scattering calculations of x-ray-absorption spectra." Physical Review. B. Condensed Matter. **52**(4): 2995-3009.

Zaric, S. D., Popovic, D. M. and Knapp, E. W. (2000). "Metal ligand aromatic cation-pi interactions in metalloproteins: ligands coordinated to metal interact with aromatic residues." Chemistry **6**(21): 3935-42.

Zhang, K. Y. J. and Main, P. (1990). "The use of Sayre's equation with solvent flattening and histogram matching for phase extension and refinement of protein structures." Acta Cryst **A43**: 377-381.

ANNEXES



J|A|C|S
ARTICLES
Published on Web 06/15/2005

Crystallographic and Spectroscopic Evidence for High Affinity Binding of FeEDTA(H_2O)$^-$ to the Periplasmic Nickel Transporter NikA

Mickaël V. Cherrier,[§] Lydie Martin,[§] Christine Cavazza,[§] Lilian Jacquamet,[§] David Lemaire,[#] Jacques Gaillard,[+] and Juan C. Fontecilla-Camps*,[§]

Contribution from the Laboratoire de Cristallographie et de Cristallogenèse des Protéines and Laboratoire de Spectroscopie de Masse des Protéines, Institut de Biologie Structurale J.P. Ebel (CEA-CNRS-UJF), 41 rue Jules Horowitz, 38027 Grenoble Cedex 1, France, and Service de Chimie Inorganique et Biologique, Departement de Recherche Fondamentale sur la Matire Condense, CEA Grenoble - 17, Rue des Martyrs 38054 Grenoble, Cedex 1, France

Received March 23, 2005; E-mail: juan.fontecilla@ibs.fr

Abstract: Because nickel is both essential and toxic to a great variety of organisms, its detection and transport is highly regulated. In *Escherichia coli* and other related Gram-negative bacteria, high affinity nickel transport depends on proteins expressed by the *nik* operon. A central actor of this process is the periplasmic NikA transport protein. A previous structural report has proposed that nickel binds to NikA as a pentahydrate species. However, both stereochemical considerations and X-ray absorption spectroscopic results are incompatible with that interpretation. Here, we report the 1.8 Å resolution structure of NikA and show that it binds FeEDTA(H_2O)$^-$ with very high affinity. In addition, we provide crystallographic evidence that a metal−EDTA complex was also bound to the previously reported NikA structure. Our observations strongly suggest that nickel transport in *E. coli* requires the binding of this metal ion to a metallophore that bears significant resemblance to EDTA. They also provide a basis for the potential use of NikA in the bioremediation of toxic transition metals and the design of artificial metalloenzymes.

Introduction

Although nickel is a rare element, generally present at only nanomolar concentrations in the environment,[1] it is essential for a variety of microbial enzymes, such as NiFe hydrogenases, ureases, acetyl coenzyme A synthases, and one class of carbon monooxide dehydrogenases.[2] In the well-studied Gram-negative bacterium *Escherichia coli* and other related organisms, high affinity Ni transport is controlled by the five-gene *nikABCDE* operon. The encoded Nik proteins, which are only expressed under anaerobic conditions,[3,4] are similar to those of other ABC-type transporters involved in nutrient and signaling peptides import processes.[5,6] They correspond to the two pore-forming integral inner membrane proteins NikB and NikC, the two inner membrane-associated proteins with ATPase activity NikD and NikE, and the periplasmic Ni-binding protein NikA. The expression of the Nik proteins is under the direct control of NikR, a DNA-binding nickel sensor that is indirectly controlled by the fumarate−nitrate regulator FNR.[7]

In addition to being central to the synthesis of NiFe hydrogenases,[8] the NikA transporter is involved in Tar-dependent negative chemotaxis.[9] NikA is a 56 kDa protein of known structure[10] that binds a single Ni(II) ion with a K_d value that has been reported to be either 0.1 or 10 μM.[3,10,11] Heddle et al. have described the metal-binding site of NikA as being a pocket rich in aromatic and arginine residues that lodges a Ni-(H_2O)$_5^{2+}$ species.[10] However, the average Ni−O_w bond distance in the proposed nickel pentahydrate is about 2.7 Å, far from the value of about 2.0−2.2 Å normally expected for transition-metal hydrates, and the coordination is not the expected octahedral one. In addition, there is only one proposed polar interaction between the protein and the hydrated nickel, which seems incompatible with the expected selectivity of NikA for this ion. This model also disagrees with an X-ray absorption spectroscopic study that indicated extensive O, N coordination to the Ni ion at 2.06 Å.[12] These observations prompted us to resume the crystallographic analysis of NikA.[9] Here, we report

[§] Laboratoire de Cristallographie et de Cristallogenèse des Protéines.
[#] Laboratoire de Spectroscopie de Masse des Protéines.
[+] Departement de Recherche Fondamentale sur la Matire Condense.

(1) Chen, Y. Y.; Burne, R. A. *J. Bacteriol.* **2003**, *185*, 6773−6779.
(2) *The Bioinorganic Chemistry of Nickel*. Lancaster, J. R., Ed.; Wiley: New York, 1988; p 337.
(3) De Pina, K.; Navarro, C.; McWalter, L.; Boxer, D. H.; Price, N. C.; Kelly, S. M.; Mandrand-Berthelot, M. A.; Wu, L. F. *Eur. J. Biochem.* **1995**, *227*, 857−865.
(4) Yohannes, E.; Barnhart, D. M.; Slonczewski, J. L. *J. Bacteriol.* **2004**, *186*, 192−199.
(5) Wu, L. F.; Mandrand-Berthelot, M. A. *Biochimie* **1995**, *77*, 744−750.
(6) Navarro, C.; Wu, L. F.; Mandrand-Berthelot, M. A. *Mol. Microbiol.* **1993**, *9*, 1181−1191.
(7) De Pina, K.; Desjardin, V.; Mandrand-Berthelot, M. A.; Giordano, G.; Wu, L. F. *J. Bacteriol.* **1999**, *18*, 670−674.
(8) Eitinger, T.; Mandrand-Berthelot, M. A. *Arch. Microbiol.* **2000**, *173*, 1−9.
(9) Charon, M. H.; Wu, L. F.; Piras, C.; De Pina, K.; Mandrand-Berthelot, M. A.; Fontecilla-Camps, J. C. *J. Mol. Biol.* **1994**, *243*, 353−355.
(10) Heddle, J.; Scott, D. J.; Unzai, S.; Park, S. Y.; Tame, J. R. *J. Biol. Chem.* **2003**, *278*, 50322−50329.
(11) Salins, L. L.; Goldsmith, E. S.; Ensor, C. M.; Daunert, S. *Anal. Bioanal. Chem.* **2002**, *372*, 174−180.

10.1021/ja0518530 CCC: $30.25 © 2005 American Chemical Society

Table 1. Data Collection and Refinement Statistics

	2.2 Å resolution set	1.8 Å resolution set
wavelength (Å)	0.9998	0.9756
space group	$P2_12_12_1$	$P2_12_12_1$
unit cell (Å)	$a = 86.59$; $b = 94.05$; $c = 124.06$	$a = 86.83$; $b = 93.87$; $c = 124.49$
resolution range (Å)	30.0−2.2 (2.3−2.2)	40.0−1.8 (1.9−1.8)
R_{merge}[a] (%)	7.6 (28.4)[c]	5.0 (30.4)
$R_{measured}$[b] (%)	8.5 (33.5)	5.9 (36.3)
I/σ	14.69 (5.84)	17.96 (5.19)
completeness (%)	97.4 (89.2)	97.6 (94.9)
redundancy	4.7 (3.5)	3.3 (3.3)
$N_{measured}$	239144 (19987)	302914 (43603)
N_{unique}	50791 (5729)	92472 (13305)
R factor/R_{free} factor (%)	17.1/24.1	16.3/21.8
No. protein atoms	7987	8172
No. water molecules	447	860
average B-factor (Å^2)	27.6	20.6
	rmsd	
bonds (Å)	0.022	0.013
angles (°)	1.986	1.457
	Ramachandran plot	
residues in most favorable region (%)	91.0	92.3
residues in additional allowed region (%)	8.7	7.4
residues in generously allowed region (%)	0.3	0.3
residues in disallowed region (%)	0.0	0.0

[a] $R_{sym} = R_{merge} = \sum_h |\hat{I}_h - I_{h,i}| / \sum_h \sum_i^{n_h} I_{h,i}$. [b] $R_{meas} = \sum_h \sqrt{\frac{n_h}{n_h - 1}} \sum_i^{n_h} |\hat{I}_h - I_{h,i}| / \sum_h \sum_i^{n_h} I_{h,i}$ with $\hat{I}_h = \frac{1}{n_h} \sum_i^{n_h} I_{h,i}$.

[c] Data in parentheses correspond to the highest resolution shell.

the 1.8 Å resolution structure of NikA and provide crystallographic, EPR spectroscopic, and mass spectrometric evidence that, unexpectedly, aerobically overexpressed, recombinant NikA binds Fe(III)EDTA(H_2O)$^-$ with high affinity when purified in the presence of the metal chelator. Furthermore, electron density maps calculated with X-ray data from Heddle et al. show that their metal-binding site is almost identical to ours, strongly suggesting that their crystals also contain a NikA−metal−EDTA(H_2O)$^-$ complex. The surprisingly high affinity binding of the aminocarboxylate chelator EDTA to the putative Ni-binding site of NikA suggests that a chemically similar natural metallophore is involved in periplasmic Ni transport. In addition, the very tight binding of NikA to the metal−EDTA complex may find applications in environmental bioremediation and in the design of artificial metalloenzymes.

Experimental Methods

NikA Expression and Purification. NikA was purified from 5l of BL21(DE3) *E. coli* culture grown in Luria broth medium at 37 °C. The strain carried pET22B (NOVAGEN) with the *nikA* gene sequence inserted into the polylinker cloning site (a kind gift from Long-Fei Wu). Production of NikA was induced by adding 1 mM IPTG when the culture reached 0.8 OD_{600} units. After 3 h incubation at 37 °C, cells were harvested by centrifugation and resuspended in 50 mM EDTA and 50 mM Tris-HCl pH 9.0 containing a protease inhibitor cocktail (ROCHE).[13] The periplasmic NikA was extracted by stirring the cells for 30 min at 37 °C. The cell debris was spun down at 15 000 rpm for 45 min at 4 °C, and the supernatant solution was adjusted to 40% saturation ammonium sulfate. The resulting soluble fraction was dialyzed against 40 mM Tris-HCl pH 7.5 (buffer A) and applied to a Q sepharose fast flow column (PHARMACIA) equilibrated with the same buffer. A gradient of 0−1 M NaCl was applied to the column. The NikA fraction that eluted at 150 mM NaCl was dialyzed against buffer A and subsequently applied to a resource Q column (PHARMACIA) equilibrated with the same buffer. Again, a 0−1 M NaCl gradient was applied, and the NikA protein was eluted at 100 mM NaCl. A total of 78 mg of pure NikA per 100 g of cells was obtained by this procedure. The protein was concentrated to 12 mg/mL in buffer A. No nickel was added at any of the stages of the purification procedure.

Crystallization. De novo crystallization of NikA using the conditions described in Charon et al.[9] was unsuccessful with the protein preparation described in the previous section. Consequently, we resorted to a small amount of NikA left from an old sample purified by the chloroform-shock procedure[14] and subsequently supplied with EDTA (L. F. Wu, personal communication) that we had crystallized in the past. Hanging drops prepared as previously described[9] yielded orthorhombic plates with dimensions of 0.2 × 0.1 × 0.05 mm^3 after a week. Identical conditions were then used with the new NikA sample prepared as described in the preceding section. One day after being set up, the drops were seeded with a dog whisker that had touched one of the crystals obtained with the old NikA sample. Reproducible crystals appeared after a week. The best crystals were obtained when the drops were prepared by mixing 2 μL of a 21 mg/mL protein solution in 40 mM Tris pH 7.5, with 2 μL of the 1.5 M ammonium sulfate, 100 mM sodium acetate pH 4.7 reservoir solution. Identical crystals were obtained upon initial addition of 200 μM $NiCl_2$ to the crystallization solution.

Data Collection and Structure Solution. Native X-ray diffraction data sets were collected to 2.2 and 1.8 Å resolution in the usual way, at the ID14-EH4 and ID29 beam lines of the European Synchrotron

(12) Allan, C. B.; Wu, L. F.; Gu, Z.; Choudhury, S. B.; Al-Mjeni, F.; Sharma, M. L.; Mandrand-Berthelot, M. A.; Maroney, M. J. *Inorg. Chem.* **1998**, *37*, 5952−5955.
(13) Van der Westen, H. M.; Mayhew, S. G.; Veeger, C. *FEBS Lett.* **1978**, *86*, 122−126.
(14) Ferro-Luzzi Ames, G.; Prody, C.; Kustu, S. *J. Bacteriol.* **1984**, *160*, 1181−1183.

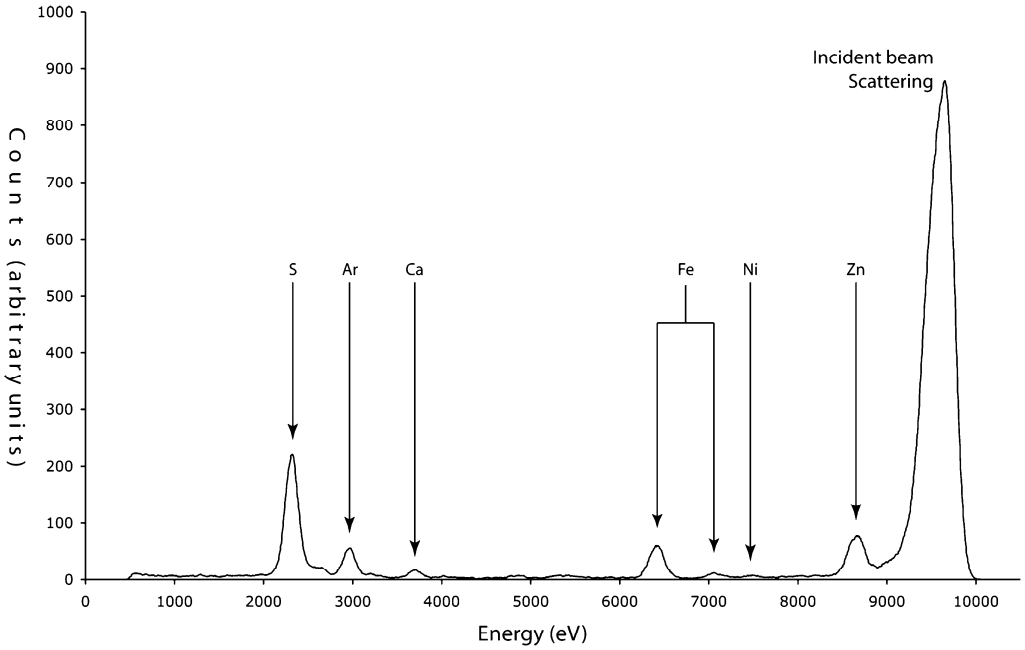


Figure 1. X-ray fluorescence spectrum measured on a single NikA crystal at beam line BM30A of the ESRF. Note the absence of nickel and the presence of iron and zinc. See also Figure 4.

Radiation Facility (ESRF), respectively. To identify the nature of the metal bound to NikA (see Results and Discussion), four additional sets were subsequently collected at the ESRF French BM30A beam line from a single crystal at the following wavelengths: 1.2830 Å (maximum f'' for zinc), 1.4839 Å (maximum f'' for nickel), 1.7432 Å (maximum f'' for iron), and 1.7438 Å (Supporting Information table). The structure was initially solved by molecular replacement with AMoRe[15] using the 2.2 Å resolution data set and the atomic coordinates deposited by Heddle et al.[10] (PDB code 1uiv). A well-contrasted solution was observed that included two NikA molecules in the asymmetric unit (not shown).

Refinement. The two NikA molecules in the asymmetric units were refined with REFMAC[16] and programs from the CCP4 suit[17] without applying noncrystallographic symmetry restrains. The model was periodically examined and corrected using computer graphics and the program TURBO−FRODO.[18] Solvent was progressively added using ARP/WARP.[19] The final stages of refinement were carried out with the 1.8 Å resolution data set using anisotropic individual temperature factors (Table 1). For the data collected at BM30A, the NikA model was refined with the set corresponding to $\lambda = 1.2830$ Å. This model was subsequently used to phase the remaining three data sets without further refinement (Supporting Information table).

X-ray Fluorescence Measurements. X-ray fluorescence was monitored with a solid-state Röntec XFlash detector operating at the BM30A beam line (Figure 1). The detector was placed 90° from the incident X-ray beam to minimize scattering. The device server to run the detector on ESRF beam lines and the Graphical User Interface to monitor the X-ray fluorescence have been developed by A. Beteva (ESRF) and J. Joly (BM30A), respectively.

(15) Navaza, J. *Acta Crystallogr.* **1994**, *A50*, 157−163.
(16) Murshudov, G. N.; Vagin, A. A.; Lebedev, A.; Wilson, K. S.; Dodson, E. J. *Acta Crystallogr.* **1999**, *D55*, 247−255.
(17) The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr.* **1994**, *D50*, 760−763.
(18) Roussel, A.; Cambillaud, C. *Silicon Graphics*; Mountain View, CA, 1989, 77.
(19) Lamzin, V. S.; Wilson, K. S. *Acta Crystallogr.* **1993**, *D49*, 129−147.

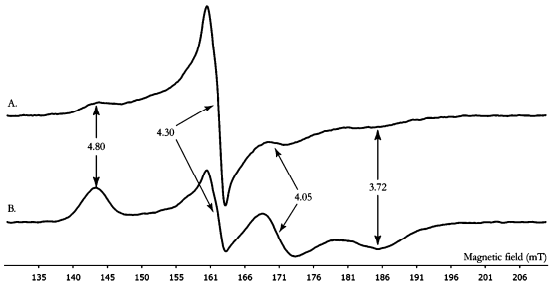


Figure 2. 4K X-band EPR spectra of NikA before (A) and after (B) treatement by EDTA (see text). In spectrum (B), the relative intensity of the signal at $g \sim 4.3$ decreases when compared to signals at g values of 4.8, 4.05, and 3.72. Experimental conditions were: frequency modulation, 100 kHz; amplitude modulation, 1 mT; microwave power, 0.25 mW. The amplifier gain was multiplied by 8.0 to record spectrum B in order to take into account the difference in protein concentrations.

Protein and Iron Concentration Determination. NikA concentration was determined by O.D. measurements at 280 nm with an $\epsilon = 72.83$ mM^{-1} cm^{-1}.[3] The iron concentration in NikA was measured colorimetrically by the Fish method using bathophenanthroline disulfonate after having denatured the protein with perchloric acid.[21]

Electron Paramagnetic Resonance (EPR) Spectroscopy. EPR spectra were recorded on a Bruker EMX spectrometer equipped with a Hewlett-Packard 5350B microwave frequency counter. A constant temperature was kept inside a helium cryostat with an ITC 503 controller (Oxford Instruments). Spectra were obtained for both an untreated 200 μM NikA solution and a 36 μM NikA solution previously incubated for 1 h in 10 mM EDTA, 40 mM Tris pH 7.5 and then dialyzed against the same buffer without EDTA (Figure 2).

Mass Spectrometry. Noncovalent mass spectrometry measurements were performed using a Q-TOF Micromass spectrometer (Micromass,

(20) Diederichs, K.; Karplus, P. A. *Nat. Struct. Biol.* **1997**, *4*, 269−275.
(21) Fish, W. W. *Methods Enzymol.* **1988**, *158*, 357−364.

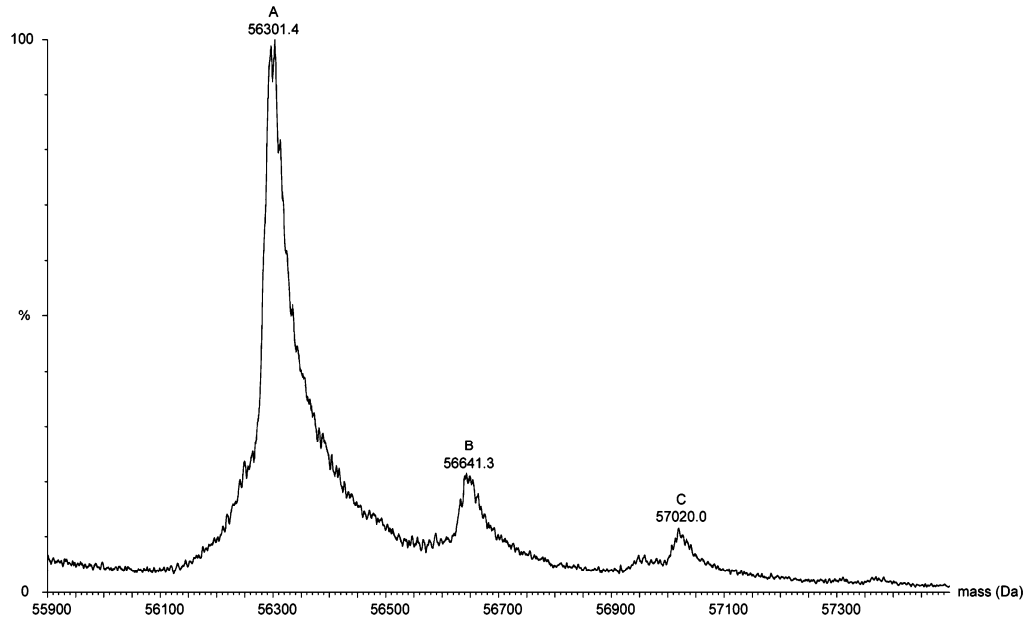


Figure 3. Noncovalent mass spectrum of as-prepared NikA. Peak A corresponds to the metal-free protein, whereas peak B matches the mass of a dehydrated FeEDTA−NikA complex within experimental error. Peak C corresponds to NikA plus 719 g/mol (see text and mass spectrometry section).

Manchester, U.K.) equipped with an electrospray ion source. It operated with a needle voltage of 2.7 kV and sample cone and extraction cone voltages of 90 and 5 V, respectively. Backing Pirani pressure was set at 6.48 mbar. The mass spectra were recorded in the 2700−4700 mass-to-charge range. The sample at a concentration of 25 μM in 20 mM ammonium acetate was continuously infused at a flow rate of 7 μL/min. Data were acquired in the positive mode, and calibration was performed using a solution of 0.5 mg/mL of CsI in water/isopropyl alcohol (1/1 v/v). Mass spectra were acquired and data were processed with MassLynx 4.0 from Micromass (Figure 3).

Results and Discussion

Crystal Structure of NikA. Our refined 1.8 Å resolution structure is very similar to the metal-containing NikA structure reported by Heddle et al.[10] with an average r.m.s. deviation of 0.58 Å for all Cα's between the two superimposed models. Small differences are confined to surface loops, where the packing influences the polypeptide conformation. As explained below, the close similarities between the two structures include the metal-binding site.

NikA Binds Iron. The highest feature in the Fo−Fc electron density map of the refined NikA, at 15σ, corresponds to a metal ion placed at the site where Heddle et al. modeled a $Ni(H_2O)_5^{2+}$. Surprisingly, X-ray fluorescence spectra indicated the absence of nickel and the presence of iron and zinc in our NikA crystals or in the surrounding solution (Figure 1). Nearly stoichiometric amounts of iron were found by a colorimetric assay using bathophenanthroline.[21] Binding of iron at the reported NikA nickel-binding site[10] was unambiguously determined using X-ray anomalous scattering effects (Figure 4). No zinc sites were detectable using this technique, probably because this ion binds many sites with low occupancy and affinity.

The electron density surrounding the iron ion in the omit map is complex and cannot correspond to a simple hydration sphere (Figure 5a). Consequently, we attempted the interpretation of

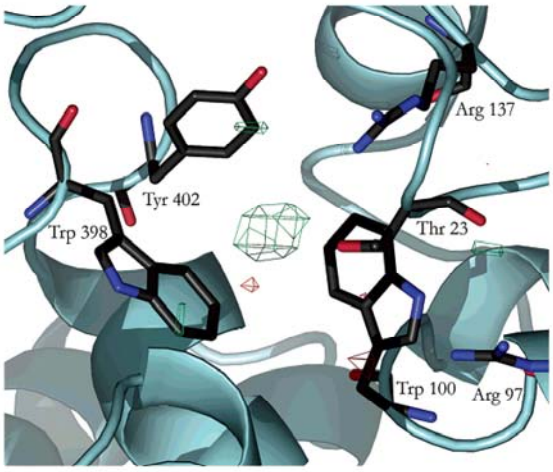


Figure 4. Double anomalous difference (ΔΔ) electron density map calculated using $\Delta_{anom(\lambda=1.7432Å)} - \Delta_{anom(\lambda=1.7438Å)}$ coefficients. The λ values correspond to the high and low energy sides of the iron absorption edge, respectively (Supporting Information table). The positive peak (green) corresponding to iron is contoured at the 3.2σ level; the very minor negative peaks (red) are contoured at −3σ. Equivalent ΔΔ maps calculated with data collected at either side of the nickel and zinc absorption edges were featureless (not shown). This figure, along with Figures 5 and 7, was prepared with Pymol.[22]

the electron density by a gradual approach. Our modeling of the iron coordination sphere started by introducing three acetates and two water molecules. Subsequently, and based on this partial model, we replaced two of the acetates by glycinates. Fo−Fc maps calculated including this model with both water molecules removed indicated that the iron ion was hepta-coordinated by a single complex ligand and one water molecule (not shown).

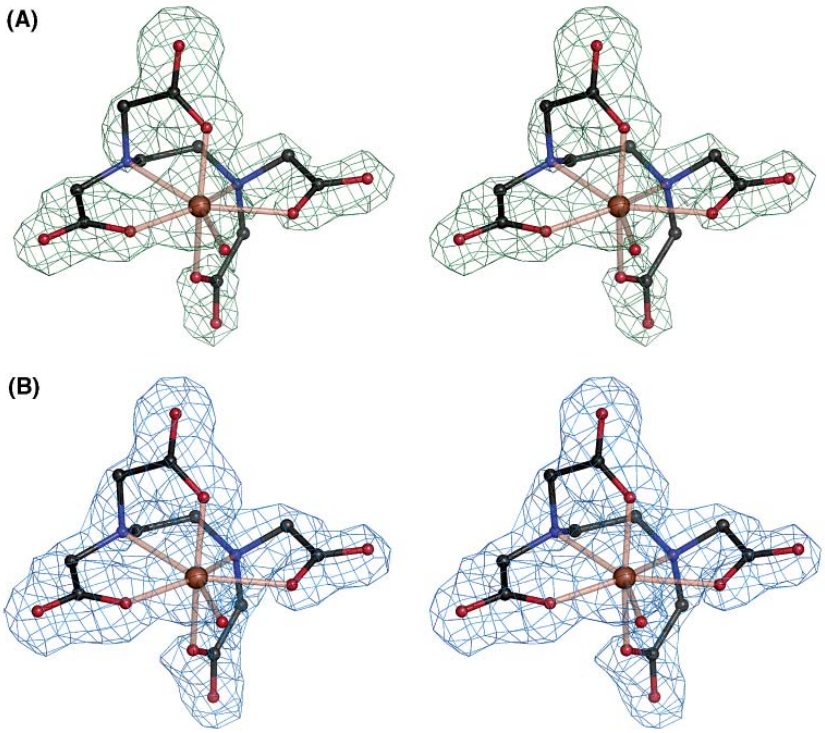


Figure 5. Stereopairs corresponding to electron density maps calculated around the iron ion. (A) Fo−Fc ("omit") electron density map contoured at 3σ, calculated with refined phases and structure factors with FeEDTA(H_2O)$^-$ omitted from the model. Ten cycles of crystallographic refinement using REFMAC[17] were run prior to the calculations in order to remove bias that could have been introduced by previous modeling of portions of the FeEDTA(H_2O)$^-$; (B) final 2Fo−Fc electron density map contoured at the 1σ level. Atoms are colored as usual; the central feature is the iron ion.

High Affinity Binding of FeEDTA(H_2O)$^-$ to NikA. The partial glycinate-containing model suggested to us that the metal coordination sphere corresponded to a truncated pentagonal bipyramid. By inspecting the relevant literature, we found out that Fe(III) forms hepta-coordinated pentagonal bipyramidal complexes with EDTA and its derivatives due to incomplete encapsulation of the iron center. The latter prevents optimal octahedral geometry and allows the binding of a water molecule as a seventh, more labile, ligand.[23] Both the omit map calculated with refined parameters and a 2Fo−Fc map of NikA including FeEDTA(H_2O)$^-$ showed that this model was correct (Figure 5). We confirmed this by superimposing our FeEDTA(H_2O)$^-$ structure to the relevant portions of the (*o*-phenylenediamine-*N,N,N′,N′*-tetraacetato)ferrate(III) complex (Cambridge Structural Database entry SOCFAR) reported by Mizuno et al.[24] After inversion of their structure, due to opposite chirality, and superposition of the two models, the calculated r.m.s. deviation is 0.27 Å, indicating that FeEDTA(H_2O)$^-$ binds NikA without any significant conformational changes relative to its unbound counterpart. We have obtained further evidence for the binding of FeEDTA to NikA from mass spectrometry. Three significant peaks are present in the mass spectrum (Figure 3). Peak A corresponds to metal-free NikA with a mass of 56 301.4 ± 6.8 g/mol; peak B corresponds to a mass of 56 641.3 ± 5.6 g/mol, and peak C represents a mass of 57 020.0 ± 8.8 g/mol. The difference between peaks A and B is 340.9 g/mol, a value that within experimental error corresponds to dehydrated FeEDTA, with a calculated mass of approximately 346 g/mol. Loss of the labile water molecule[23] was to be expected under these experimental conditions. The possible nature of peak C will be discussed below.

The EDTA−protein interactions are summarized in Figure 6. There are three bonds between Arg137 and Arg97 and two of the carboxylates of the chelator. In addition, there is an unshielded cation−π interaction between the iron ion and Trp398, with an indol-to-metal distance of 5.5 Å and a θ_0 angle of 16°.[26] Three water-mediated interactions complete the binding of the chelator to NikA (Figures 6 and 7).

Origin of the FeEDTA(H_2O)$^-$. The initial hexagonal crystals we reported several years ago[9] were obtained from a NikA sample where the periplasmic extraction was performed using chloroform[14] in the absence of EDTA.[3] When we added this chelator to that NikA solution, we obtained orthorhombic crystals (unpublished results) identical to those reported here that were obtained using a sample extracted from the periplasm with EDTA.[13] These observations strongly suggest that the orthorhombic crystals are only obtained when a metal−EDTA-(H_2O)$^-$ complex is bound to NikA. The Fe−ligand complex

(22) DeLano, W. L. http://www.pymol.org, 2002.
(23) Schneppensieper, T.; Seibig, S.; Zahl, A.; Tregloan, P.; van Eldick, R. *Inorg. Chem.* **2000**, *40*, 3670−3676.
(24) Mizuno, M.; Funahashi, S.; Nakasuka, N.; Tanaka, M. *Inorg. Chem.* **1991**, *30*, 1550−1553.
(25) Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. *Protein Eng.* **1995**, *8*, 127−134.
(26) Zaric, S. D.; Popovic, D. M.; Knapp, E. W. *Chem.—Eur. J.* **2000**, *21*, 3935−3942.

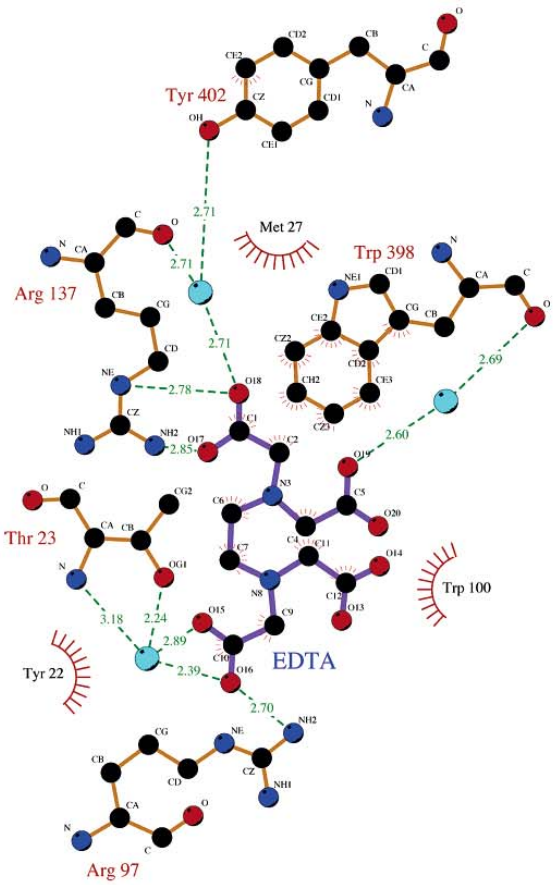


Figure 6. LIGPLOT[25] depicting the EDTA–NikA interactions (Å). The iron ion, which does not directly interact with NikA, has been omitted for clarity. See also Figure 7 and Supporting Figure 1.

reported here must be very tightly bound, as it did not dissociate from NikA during the purification and crystallization procedures or upon treatment of the crystals with the natural siderophore ferrichrome (not shown). In addition, since iron was not added at any stage, the $FeEDTA(H_2O)^-$ complex, that is, the predominant form in solution,[27] must have formed from iron that was naturally present in the medium. The EPR results indicate that most of the iron is found as specifically bound Fe(III) that gives rise to the signal with g values of 4.8, 4.05, and 2.72 (Figure 2). This signal is specific of the NikA-bound Fe–EDTA complex as it is absent from the free metal–ligand complex (not shown). The presence of Fe(III) was to be expected because the purification was performed aerobically, and the formation constant for Fe(III)EDTA is about 10^{25} whereas those for Fe(II)EDTA and Ni(II)EDTA are approximately 10^{14} and 10^{18}, respectively.[28] Treatment of the NikA sample with EDTA removed most of the $g = 4.3$ Fe(III) EPR signal, whereas the specific signal showed a relative increase (Figure 2).

It is worth mentioning here that the EXAFS results of Allan et al. suggested a pentagonal bipyramidal coordination of nickel by NikA[12] and that the sample was purified in the presence of EDTA (M. A. Mandrand-Berthelot, personal communication).

$Ni(H_2O)_5^{2+}$ versus $FeEDTA(H_2O)^-$. The nickel pentahydrate model proposed by Heddle et al.[10] is compatible neither with stereochemical considerations nor with our crystallographic results. Consequently, in order to compare their structure to ours, we have calculated refined omit and 2Fo–Fc maps using atomic coordinates and structure factors deposited by these authors (PDB codes 1uiv and r1uivsf). We find that the omit map corresponding to their metal-binding site is almost identical to ours, and that $FeEDTA(H_2O)^-$ refines with their X-ray data as well as with ours (Supporting Figure 2). It is not possible from the deposited X-ray data to determine the identity of their metal ion. We believe that it should be a ferric iron because, like us, they used EDTA during the purification of NikA[10] and this ligand has a much higher affinity for Fe(III) than for Ni(II).[28] We have also checked crystals that were grown in the presence of 200 μM $NiCl_2$ and found that only iron was bound to NikA (not shown). This means that once the $FeEDTA(H_2O)^-$ complex attaches to the protein, it will not be displaced by free nickel. Unfortunately, Heddle et al. did not check the nature of their metal ion by collecting anomalous scattering X-ray data and calculating electron density maps, such as the one shown in Figure 4, so at this point, its identity remains uncertain.

Besides the crystallographic and stereochemical arguments supporting a metal–EDTA complex in the structure of Heddle et al., there is also the issue of the metal-binding mode by NikA. In the nickel–pentahydrate model, there is only one nickel-bound water–Arg137 interaction at $d = 3.1$ Å,[10] instead of the three direct arginine–carboxylate and three water-mediated bonds observed in the FeEDTA model (Figures 6 and 7). The latter model also explains an observation that puzzled Heddle et al. (see legend to Figure 6 in ref 10): one of the major differences in going from the metal-free to the metal-bound NikA is the change in orientation of Arg97. The side chain of this residue rotates from a solvent-exposed position in the former to approach the metal ion in the latter. In the nickel–pentahydrate model, this is difficult to explain because a positively charged arginine side chain would point at a positively charged metal ion. In the metal–EDTA model, there is a nice interaction between Arg97 and one of the carboxylates from the ligand (Figures 6 and 7).

Physiological Binding of Nickel to NikA. The implausibility of NikA binding a nickel hydrate and the very tight binding of $FeEDTA(H_2O)^-$ to the protein strongly suggest that nickel transport by NikA involves a metallophore that displays significant structural similarities to EDTA. Metallophores have been extensively studied in the case of iron transport.[29] Under low iron conditions, microorganisms and fungi secrete low-molecular weight, ferric-specific chelating agents, called siderophores.[30] They are needed because, at physiological pH values, the free Fe(III) concentration is 10^{-18} M, much below the minimal effective iron concentration required by microorganisms that is about 10^{-8} M.[31] Siderophores can be divided into several groups depending on their chemical nature; the most

(27) Lind, M. D.; Hamor, M. J.; Hamor, T. A.; Hoard, J. L. *Inorg. Chem.* **1964**, *3*, 34–43.
(28) http://www.chm.bris.ac.uk/motm/edta/edtah.htm.
(29) Raymond, K. N.; Dertz, E. A.; Kim, S. S. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3584–3588.
(30) Neilands, J. B. *J. Biol. Chem.* **1995**, *270*, 26723–26726.
(31) Braun, V.; Hantke, K. In *Transition Metals in Microbial Metabolism*; Winkelmann, G., Carrano, C. J., Eds.; Harwood: Amsterdam, 1997; pp 81–101.

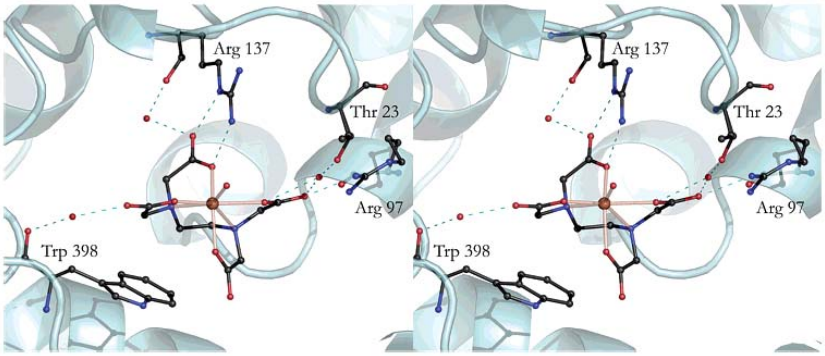


Figure 7. Stereopair drawing depicting the FeEDTA(H_2O)$^-$—NikA complex. See also Figure 6. Fe—EDTA distances are shown in Supporting Figure 1.

common ones belong to the catecholate, hydroxypyridonate, and hydroxamate groups, but aminocarboxylates have also been described.[32] The best-studied siderophore is enterobactin.[29] In *E. coli* and other Gram-negative bacteria, ferric—siderophore uptake is both receptor- and energy-dependent. After being actively translocated across the outer membrane, the ferric complex binds to its cognate periplasmic transporter protein, with a K_d lower than 0.1 nM in the case of ferric—enterobactin,[33] and it is actively moved across the cytoplasmic membrane by an ATP-dependent transporter system. The low K_d value reflects the fact that periplasmic iron levels are in the submicromolar range, and a high affinity carrier is needed for its effective transport and to prevent release of the metal ion.[34] The analogy of this system to the *nik*-dependent nickel transport machinery is significant, including a periplasmic carrier and cytoplasmic membrane-bound ATP transporters.[5] Like iron, nickel is present at low concentrations in the environment (10^{-9} M) [1]. In vitro K_d measurements for the Ni—NikA complex have given values of either 0.1 [3] or 10 μM.[10,11] These experiments are likely to be very difficult to reinterpret because of the results reported here. The 100-fold difference between the K_d values mentioned above may be related to the periplasmic extraction method. De Pina et al. used EDTA-free, nickel-free protein, whereas Heddle et al. probably had a significant amount of the NikA—metal—EDTA complex. The latter authors reported a small enthalpy of binding and undetectable intrinsic fluorescence related to nickel binding.[10] Alternatively, De Pina et al. observed weak fluorescence effects that were interpreted as resulting from a small nickel-induced conformational change that alters the environment of one or more tryptophan side chains.[3] However, the conformational change is rather significant as it involves a 15° rotation of one of the lobes of NikA relative to the other.[10] A calculation using NACCESS[35] shows that the change from open to FeEDTA-containing NikA reduces the solvent-exposed surface of Trp100 and Trp398 from about 40 to 2.0%, and 36 to 8%, respectively. (Even when an empty binding site is used for the calculation for the closed form, Trp100 is about 30% more buried than in the metal-free NikA structure.) These changes in the environment of Trp100 and Trp398 should have generated a significant shift in the emission maximum. The absence of such shift probably reflects the nonphysiological conditions of these experiments due to either the presence of a NikA—metal—EDTA complex or the absence of an endogenous ligand. Perhaps for the same reason, the reported NikA—nickel K_d values are significantly higher than those found for many ferric—siderophore complexes and have indicated weak binding of nickel to the transporter. However, physiological binding of nickel to NikA could be much tighter if the metal ion is incorporated as a nickel—metallophore complex.

Aerobically purified and overexpressed NikA samples are unlikely to contain any significant amounts of a hypothetical endogenous nickelophore. Nevertheless, it cannot be ruled out that small amounts of this ligand, co-purified with NikA, could have caused some of the effects observed in previous nickel-binding studies. For example, our mass spectrometric spectrum contains a minor peak that could correspond to NikA bound to a species with a mass of approximately 719 g/mol (Figure 3).

Conclusions and Perspectives

Maybe the main conclusion from this work is that NikA has not evolved to bind directly to metal ions. On the contrary, our results show that NikA binds FeEDTA(H_2O)$^-$ very tightly as it cannot be dissociated by either competition with nickel ions or ferrichrome, or by dilution. By extension, NikA should be able to bind nickel, provided it forms a complex with a metallophore. (One of the referees has suggested that a small metallophore could serve as an amplifier of the subtle differences among metal ions and in our case favor nickel binding to NikA.) As discussed above, previous nickel-binding studies using EDTA or metal-free protein are likely to be flawed. To avoid this problem, we intend to grow *E. coli* anaerobically, without overexpressing NikA, and without using EDTA. We expect that the nickel transporter sample obtained from these bacteria will be in the biologically relevant form. If there is, indeed, a metallophore specific for nickel, we will try to determine its structure. We will also check the binding of nickel to apo NikA in the absence of chelator.

The surprisingly tight binding of FeEDTA to NikA opens the possibility of using this protein for the bioremediation of toxic transition metals and the synthesis of complexes with enzymatic activity, such as peroxidase.[36] To better characterize the binding of metal ions to NikA, we are in the process of

(32) Stintzi, A.; Barnes, C.; Xu, J.; Raymond, K. N. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 10691—10696.
(33) Newton, S. M.; Igo, J. D.; Scott, D. C.; Klebba, P. E. *Mol. Microbiol.* **1999**, *32*, 1153—1165.
(34) Sprencel, C.; Cao, Z.; QI, Z.; Scott, D. C.; Montague, M. A.; Ivanoff, N.; Xu, J.; Raymond, K. M.; Newton, S. M.; Klebba, P. E. *J. Bacteriol.* **2000**, *182*, 5359—5364.
(35) http://wolf.bms.umist.ac.uk/naccess/.
(36) Brausam, A.; Van Eldick, R. *Inorg. Chem.* **2004**, *43*, 5351—5359.

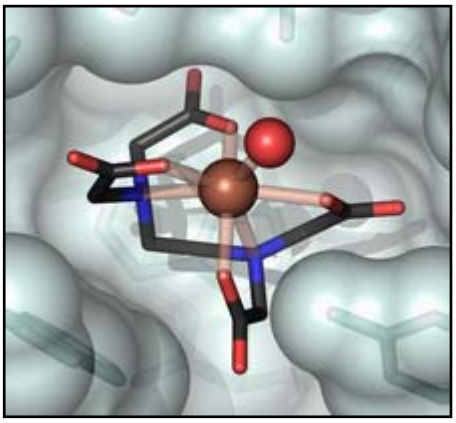
preparing the metal-free protein using the chloroform-shock procedure[14] and will determine the K_d values for the binding of different metal−EDTA complexes.

Coordinates and structure factors have been deposited with PDB codes 1zlq and r1zlqsf.

Acknowledgment. We thank the local contact staff of the ID14-EH4, ID-29, and BM-30A beam lines of the ESRF for help with X-ray data collection, and Dr. LongFei Wu for providing the genetic material used in this work. We also thank Prof. Marc Fontcave and Dr. Stephan Menage for pointing out that NikA−EDTA−metal complexes could be used as artificial metalloenzymes.

Supporting Information Available: One table with additional crystallographic statistics, a LIGPLOT depicting EDTA−Fe bond distances, a figure showing omit, and 2Fo−Fc electron density maps using data from Heddle et al.[10] This material is available free of charge via the Internet at http://pubs.acs.org.

JA0518530



Résumé

Le nickel est un cofacteur essentiel à l'activité de certaines protéines bactériennes telles que les uréases ou les hydrogénases, mais il est également toxique à trop forte concentration. De ce fait, il est indispensable aux bactéries de posséder un système d'import du Ni^{2+} spécifique et hautement régulé. Dans le cas d'*Escherichia coli*, il s'agit d'un système ABC (*ATP Binding Casette*) composé des cinq protéines de l'opéron *nikABCDE* : NikA, une protéine périplasmique, NikB et NikC, deux protéines transmembranaires formant un pore à travers la membrane interne, et les protéines NikD et NikE, deux protéines cytoplasmiques hydrolysant l'ATP et apportant ainsi l'énergie nécessaire au système. La régulation de l'expression de cet opéron est assurée par FNR (*Fumarate Nitrate Regulator*), l'activateur, et de NikR, le répresseur. Dans un premier temps, nous avons travaillé sur une structure de la protéine NikA, résolue par remplacement moléculaire, à partir de cristaux appartenant au système orthorhombique. Elle nous a permis de montrer que NikA ne fixait pas le nickel sous une forme pentahydratée comme cela était décrit dans la littérature, mais sous la forme d'un complexe $FeEDTA(H_2O)^-$. L'EDTA n'étant pas une molécule naturelle, dans notre cas il provient de l'étape d'extraction périplasmique, ce résultat suggère que le nickel ne se fixe pas directement à la protéine, mais par l'intermédiaire d'un métallophore de structure proche de celle de l'EDTA. En remplaçant l'étape d'extraction périplasmique à l'EDTA par une étape utilisant le chloroforme, nous avons résolu, par remplacement moléculaire, une seconde structure de NikA à partir de cristaux appartenant au système cristallin hexagonal. Dans ce cas, on observe, dans le site de fixation de la protéine, un nickel en complexe avec une molécule similaire à l'EDTA, laquelle a de fortes chances d'être le métallophore physiologique, que nous avons tenté de caractériser. Nous abordons aussi des études cristallographiques préliminaires concernant NikR.

Mots-clés : cristallographie, structure, nickel, transport, métallophore, EDTA.

Abstract

Nickel is an essential cofactor for several proteins found in bacteria, examples are ureases and hydrogenases, but is also toxic at high concentrations. Consequently, microorganisms need a highly regulated and specific import system of Ni^{2+}. In *Escherichia coli*, this system is composed by five proteins encoded by the *nikABCDE* operon (which is an *ATP Binding Cassette)* : NikA, a periplamic protein, NikB and NikC, two transmembrane proteins, which create a pore in the inner membrane, and the cytoplasmic proteins NikD and NikE, which hydrolyse ATP and provide the energy for the transport process. FNR (*Fumarate Nitrate Regulator*) is the activator, and NikR the repressor of the transcription of the *nik* operon. We solved the NikA protein structure by molecular replacement and found that, contrary to what was reported in a previous crystallographic analysis, the protein doesn't bind a pentahydrate nickel, but an $FeEDTA(H_2O)^-$ complex. The EDTA, which is not a natural molecule, was introduced during the periplasmic extraction. These results suggest that NikA doesn't fix nickel directly, but complexed with a natural metallophore. Indeed, when we changed the periplasmic extraction with EDTA to a protocol using chloroform, we obtained a new structure where the nickel is complexed to a molecule similar to EDTA. Given that it seems to be the physiological metallophore, we have tried to characterise it using a variety of techniques. We also report a preliminary crystallographic analysis of NikR.

Key-words: crystallography, structure, nickel, transport, metallophore, EDTA.