HAL
open science

Etude structurale des aptamères peptidiques anti-Fur et de leur interaction avec leur cible

Cheickna Cisse

To cite this version:

Cheickna Cisse. Etude structurale des aptamères peptidiques anti-Fur et de leur interaction avec leur cible. Sciences agricoles. Université de Grenoble, 2012. Français. NNT : 2012GRENV012 . tel-00846781

HAL Id: tel-00846781

https://theses.hal.science/tel-00846781

Submitted on 20 Jul 2013

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.

UNIVERSITÉ DE GRENOBLE

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Biochimie**

Arrêté ministériel : 7 août 2006

Présentée par

Cheickna CISSE

Thèse dirigée par **Isabelle Michaud-Soret** et codirigée par **Serge CROUZY**

Préparée au sein du Laboratoire de biologie et chimie des métaux (**LCBM**)
dans l'École Doctorale Chimie et Science du Vivant **(EDCSV)**

Étude structurale des aptamères peptidiques anti-Fur et de leur interaction avec leur cible

Thèse soutenue publiquement le **« 19 janvier 2012 »**,
devant le jury composé de :

Dr Isabelle Schalk
Ecole supérieure de biotechnologie Strasbourg (ESBS), Rapporteur
Dr Norbert Garnier
Université d'Orléans, Rapporteur
Dr Agnes Rodrigue
Institut National des Sciences Appliquées (INSA) de LYON, Examinateur
Pr Franck Fieschi
Université Joseph Fourier (UJF), Examinateur
Dr Didier Boturyn
Université Joseph Fourier (UJF), Invité

Remerciements

J'adresse mes plus profonds et sincères remerciements à Isabelle Michaud-Soret, chef de l'équipe Biologie des métaux (BioMet) et Serge Crouzy, chef de l'équipe Modélisation et Chimie Théorique (MCT) mes co-directeurs depuis le stage de master 2.

A Isabelle, merci pour ton engagement, ton soutien et ton accompagnement sans lesquels cette thèse n'aura pas eu lieu. Merci pour tes conseils précieux, ta rigueur et ta pertinence dans le travail. Je n'oublierai jamais tout ce que tu as fait pour moi durant ces années passées sous ta responsabilité. Je te serai reconnaissant et redevable pour tout cela.

A Serge, merci d'avoir accepté de codiriger cette thèse. Ton implication dans le projet m'a permis de découvrir et d'apprécier la modélisation et ton encadrement a été très bénéfique pour moi. Merci pour ta sympathie, ta confiance et les discussions que nous avons souvent eues. Ton sens de l'humour restera le meilleur à mes yeux.

A vous deux, merci ! Vous serez toujours les bienvenus au Mali. Je vous inviterai quand j'aurais une bonne position là-bas.

Je n'aurais jamais obtenu ces résultats sans les travaux antérieurs de Nadia Abed et de Sylvia Vitale. Merci à Sylvia, Sandra Galop et Sophie Mathieu pour tous leurs apports dans les manipulations. Merci pour tous vos conseils, votre disponibilité, votre générosité et toute votre aide. Je vous souhaite bonne chance et bon courage pour la suite.

Je souhaiterais saluer et remercier les personnes du laboratoire de Chimie et Biologie des Métaux (LCBM), dirigé par Stéphane Ménage et principalement les membres des équipes BioMet et MCT que j'ai côtoyé durant ces années. J'ai passé de bons moments avec vous pendant les réunions, autour des repas. Je ne pourrais malheureusement pas tous vous citer mais je vous dis à tous, merci pour votre soutien spécialement ces derniers mois.

Merci à Didier Boturyn pour sa collaboration et à l'ensemble des membres du Département de Chimie Moléculaire de Grenoble (DCM) pour avoir facilité mes expériences au sein de leur laboratoire.

Merci à l'ensemble des membres du « Norvegian Structural Biology Centre (NorStruct) pour avoir rendu mon séjour à Tromso (Norvège) aussi agréable que possible que ce soit du point vue scientifique que relationnel. Je remercie particulièrement les Docteurs Rafi Hamad et Ingar Leiros pour leurs apprentissages au criblage virtuel et à la cristallogenèse respectivement. Merci pour vos discussions scientifiques et votre accueil au sein du laboratoire et pour les moments passés ensemble.

Je tiens à remercier les membres du jury d'avoir accepté de juger ce travail : les docteurs Isabelle Schalk et Norbert Garnier en tant que rapporteurs ; le docteur Agnes Rodrigue et le professeur Franck Fieschi en tant qu'examinateurs ; les docteurs Isabelle Michaud-Soret et Serge Crouzy en tant que directeur et co-directeur de thèse ; et enfin le docteur Didier Boturyn en tant qu'invité.

Je remercie ma famille et tous mes amis qui m'ont encouragé et soutenu tout au long de ce travail.

A tout le monde, « aw ni tché ! »

Liste des abréviations

Protéines

BsPerR	PerR de *Bacillus subtilis*
EcFur	Ferric uptake regulator d'*Escherichia coli*
HpFur	Ferric uptake regulator d'*Helicobacter pylori*
MtZur	Zinc uptake regulator de *Mycobacterium tuberculosis*
PaFur	Ferric uptake regulator de *Pseudomonas aeruginosa*
SaFur	Ferric uptake regulator de *Staphylococcus aureus*
VcFur	Ferric uptake regulator de *Vibrio cholerae*
VsFur	Ferric uptake regulator de *Vibrio salmonicida*
YpFur	Ferric uptake regulator de *Yersinia pestis*

Autres

Å	Angström
a.a	Acide aminé
ADN	Acide désoxyribonucléique
ARN	Acide ribonucléique
BET	Bromure d'éthidium
BTP	Bis-Tris Propane
CD	Circular dichroïsm ou dichroïsme circulaire
Cl-Trt	2-chlorotrityle®
C-ter	C-terminal
Da	Dalton
DCM	Dichlorométhane
DIEA	*N*,*N*-Diisopropylethylamine
DMSO	Diméthyl Sulfoxyde
DNase I	Désoxyribonucléase I
dNTP	Désoxyribonuléotide triphosphate
DO	Densité optique
DTT	Dithiothréitol

ε	Coefficient d'extinction molaire (en $M^{-1}.cm^{-1}$)
E. coli	*Escherichia coli*
EDTA	Acide éthylènediaminetétraacétique
eq	Equivalent
Fmoc	Fluorénylméthoxycarbonyle
HA	Hémagglutinine
HEPES	Acide (2-(4-(2-hydroxyéthyl)-1-pipérazinyl) ethanesulfonique
H. pylori	*Helicobacter pylori*
HSQC	Heteronuclear Single Quantum Coherence
IPTG	Isopropyl-ß-D-Thio-Galactopyranoside
Kd	Constante de dissociation
LB	Milieu Luria-Bertani
MALDI	Matrix-Assisted Laser Desorption/Ionisation
MM	Masse molaire
MOPS	Acide 3-(N-morpholino) propane sulfonique
N-ter	N-terminal
NLS	Nuclear localization sequence
SDS-PAGE	Sodium Dodécyl Sulfate- Protein Analysis Gel Electrophoresis
PAR	4-(2-pyridylazo) resorcinol
PCR	Polymerase Chain Reaction
pFx	Peptide linéaire de 13 acides aminés correspondant à la séquence variable de l'aptamère peptidique Fx (x=1 à 5)
pI	Point isoélectrique
PyBOP	(benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
RMN	Résonance magnétique nucléaire
SDS	Sodium DodécylSulfate
SPPS	Synthèse peptidique sur phase solide
TCA	Acide tricholoroacétique
TFA	Acide trifluoroacétique
TIS	Triisopropylsilane
Trx A	Thioredoxine A
UV-vis	Ultraviolet- Visible
X-Gal	5-bromo 4-chloro-3-indolyl-ß-D-galactopyranoside

Chapitre I :

Introduction bibliographique

CHAPITRE I : ..... **- 1 -**

INTRODUCTION BIBLIOGRAPHIQUE ..... **- 1 -**

AVANT-PROPOS ..... - 5 -

I| LES ANTIBIOTIQUES ET LA RESISTANCE ..... - 8 -

I. A| Les antibiotiques ..... *- 8 -*

I. B| La résistance bactérienne aux antibiotiques ..... *- 16 -*

I. C| Les nouvelles stratégies de lutte antibactérienne ..... *- 22 -*

II| LE FER DANS LE VIVANT ..... - 29 -

II. A| Le fer est indispensable à la vie ..... *- 29 -*

II. B| Le fer chez l'homme ..... *- 29 -*

II. C| L'homéostasie du fer chez les bactéries ..... *- 31 -*

II. D| Le fer et la relation hôte-pathogène ..... *- 39 -*

III| LA PROTÉINE FUR (FERRIC UPTAKE REGULATOR) ..... - 42 -

III. A| Historique ..... *- 42 -*

III. B| Fur, un régulateur global ..... *- 44 -*

III. C| La liaison de Fur à l'ADN ..... *- 46 -*

III. D| La famille des protéines Fur et « Fur-like » ..... *- 49 -*

III. E| Caractéristiques structurales des protéines Fur et « Fur-like » ..... *- 51 -*

III. F| Conclusion ..... *- 62 -*

IV| LA TECHNOLOGIE DES APTAMERES PEPTIDIQUES ..... - 63 -

IV. A| Description des aptamères peptidiques ..... *- 63 -*

IV. B| La technologie des aptamères peptidiques ..... *- 64 -*

IV. C| Intérêts de la technologie des aptamères peptidiques ..... *- 68 -*

IV. D| Conclusion ..... *- 70 -*

V| PRESENTATION DU TRAVAIL ..... - 71 -

BIBLIOGRAPHIE ..... - 73 -

Avant-propos

La vie sur terre est apparue avec les bactéries il y a de cela 3,5 milliards d'années. Ces microorganismes ont su évoluer et s'adapter à tous les biotopes auxquels elles ont été confrontées. Durant des millions d'années, ces minuscules êtres vivants ont formé un véritable empire avant l'arrivée des organismes supérieurs. On estime qu'il y aurait entre 4 à 6 quintillions (10^{30}) de bactéries dans le monde (Whitman *et al.*, 1998). Ce qui représente une grande partie de la biomasse (40%). Cette grande population est due à la division extrêmement rapide des bactéries, ce qui leur confère une énorme capacité de diversité génétique. La grande majorité de ces bactéries est inoffensive. Certaines sont même essentielles et vivent en symbiose avec les organismes supérieurs. C'est le cas des bactéries commensales qui vivent dans les intestins des mammifères et aident à la dégradation des aliments. Seulement une petite partie des bactéries communément appelées pathogènes est responsable de certaines maladies infectieuses comme la tuberculose, la diphtérie, le tétanos, le choléra, la maladie du charbon, les méningites et les pneumonies bactériennes pour n'en citer que quelques unes. Par ailleurs, toutes les maladies infectieuses ne sont pas d'origine bactérienne. Certaines maladies infectieuses sont causées par des virus comme le SIDA, la syphilis, la rougeole, la variole, l'Herpès ou par des parasites comme le paludisme.

Dès l'aube de son existence, l'Homme a toujours évolué avec les bactéries et a donc subi les ravages causées par les maladies infectieuses. Au XIXe siècle, les infections bactériennes étaient la première cause de mortalité et faisaient des millions de morts. L'homme restait impuissant, sans armes efficaces, face à ces maladies. On pensait même que ces maladies étaient héréditaires et incurables jusqu'à la fin du XIXe siècle.

En 1870, Louis Pasteur et d'autres microbiologistes découvrent que les microbes sont responsables des maladies infectieuses. Ces découvertes ouvrirent la voie à l'étude des microorganismes (microbiologie) et à la caractérisation des maladies. Ensuite, en 1880, Pasteur et Robert Koch découvrent respectivement les agents infectieux du choléra et de la tuberculose. Pasteur préconisait des mesures d'hygiène pour éviter la contamination et la propagation des microbes. Il mit au point un vaccin basé sur le principe de l'atténuation de la virulence contre le choléra des poules en 1880, puis le vaccin contre la rage en 1885. Par la suite, toute une génération de vaccins apparut avec les anatoxines tétaniques et diphtériques (Ebisawa 1987) ainsi que le B.C.G., contre la tuberculose (Feldberg 1995). Les anatoxines sont des exotoxines dénaturées par la chaleur et le formol et utilisé pour la vaccination. Ces

moyens de lutte préventive contre les microbes ont contribué à une amélioration significative de la santé, mais n'ont pas permis toutefois de résoudre tous les problèmes de santé causés par les bactéries.

En 1929, le bactériologiste anglais Sir Alexander Fleming découvre la pénicilline et ses propriétés antibactériennes (Fleming 1955), ouvrant la voie au traitement des maladies infectieuses par les antibiotiques. En 1941, la pénicilline fut utilisée avec succès pour traiter un patient atteint de septicémie à staphylocoques. Dès lors, la pénicilline fut considérée comme un véritable « produit miracle » et fut produite à grande échelle par les sociétés pharmaceutiques américaines, pour traiter dans un premier temps les nombreux blessés infectés de la seconde guerre mondiale (Latour 2001) puis mise sur le marché. Les décennies suivantes voient la découverte de plusieurs antibiotiques (Figure I.1). Ainsi, la période entre 1940 et 1960 considérée comme le « golden age » (l'âge d'or) voit la découverte de la grande majorité des antibiotiques.

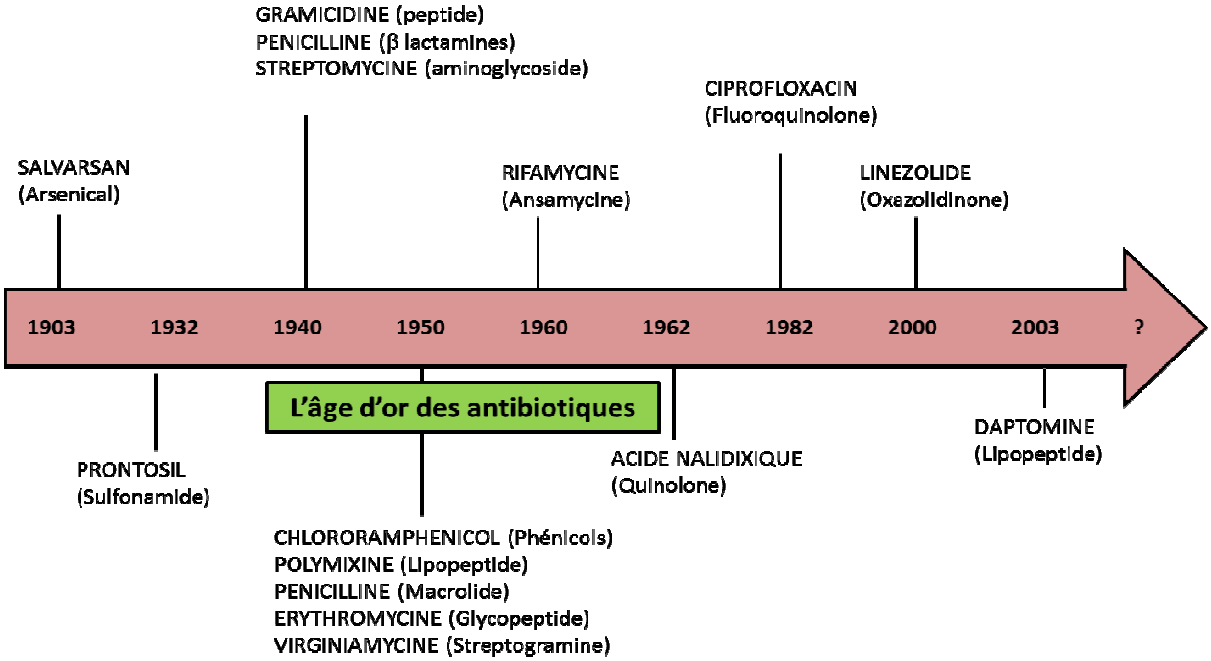


Figure I.1: les principales classes d'antibiotiques et leurs dates de découverte. Adapté à partir de la revue (Choffnes *et al.*, 2010).

Depuis leur découverte, l'utilisation massive et inappropriée des antibiotiques a provoqué une grande dissémination de molécules antibiotiques dans la nature, favorisant la sélection et le développement de bactéries (y compris les pathogènes) résistantes aux antibiotiques. Les maladies microbiennes qu'on pensait pouvoir vaincre et avoir vaincues sont toujours fréquentes pour la plupart, à cause du phénomène de résistance aux antibiotiques.

A l'heure actuelle, où il existe des résistances à l'ensemble des antibiotiques et connaissant la vitesse de propagation de ces résistances à travers le monde, il est impératif de trouver de nouvelles solutions. Cela passe par la découverte de nouvelles molécules antibiotiques.

La protéine Fur (Ferric Uptake Regulator) est un régulateur global transcriptionnel chez *Escherichia coli*, omniprésente chez toutes bactéries et absente chez les mammifères. Elle représente donc une cible antibactérienne très intéressante. Des études menées sur cette protéine par le passé au laboratoire ont montré qu'elle est bien une cible potentielle antibactérienne (Abed *et al.*, 2007). La technologie des aptamères peptidiques avait été utilisée afin de sélectionner des molécules capables d'inhiber spécifiquement la protéine Fur. Quatre inhibiteurs nommés F1 à F4 ont ainsi été sélectionnés puis ont été caractérisés (Abed 2005; Vitale 2009). Cette thèse s'inscrit comme la suite de ces travaux. Le but est de comprendre le mécanisme d'inhibition de Fur par ces aptamères peptidiques et plus précisément par les peptides nommés pF1 à pF4 issus des aptamères peptidiques anti-Fur. Un autre objectif consiste à identifier de petites molécules inhibitrices de Fur ayant des propriétés thérapeutiques, à partir des résultats obtenus. Une double approche, théorique et expérimentale a été utilisée pour ces travaux, en effectuant des allers-retours entre les deux parties.

Dans cette introduction bibliographique, nous reviendrons largement sur les thèmes abordés dans l'avant propos.

Nous ferons quelques rappels sur l'historique de la découverte des antibiotiques pour mieux les distinguer et nous présenterons les différentes classes d'antibiotiques et leurs modes d'action seront évoqués. Nous décrirons la résistance bactérienne aux antibiotiques, son origine et ses mécanismes. Nous aborderons le sujet des nouvelles stratégies de lutte antibactérienne.

Nous nous attarderons ensuite sur la protéine Fur et la famille de métallorégulateurs « Fur-like » après avoir bien établi le lien entre le Fer et la protéine Fur. Nous développerons le rôle crucial de ces protéines chez la bactérie. Nous ferons une synthèse des connaissances actuelles sur cette famille de protéines, principalement au niveau structural.

Nous décrirons enfin, de manière rapide la technologie des aptamères peptidiques et leur mode de sélection avant de décrire le projet de thèse.

I| Les antibiotiques et la résistance

I. A| Les antibiotiques

I. A. 1| Définition de l'antibiotique

L'antibiotique (du grec *anti* : « contre », et *bios* : « la vie ») désigne au sens large une substance ou molécule capable de détruire des bactéries ou de bloquer leur croissance. L'antibiotique désignera, dans ce cas-ci, toute « substance chimique organique d'origine naturelle ou synthétique inhibant ou tuant les bactéries pathogènes de façon ciblée » (Bentley *et al.*, 2003). Cette substance antibactérienne est employée à faible concentration et possède une toxicité sélective. Certains antibiotiques sont capables de détruire des microbes autres que les bactéries. On parle alors d'antibiotiques antifongiques ou antiviraux pour lutter respectivement contre les champignons et les virus. L'antibiotique se distingue d'un antiseptique, un désinfectant non ingéré qui détruit tous les germes.

I. A. 2| Histoire de la définition des antibiotiques

L'aventure antibiotique n'a pas commencé avec la découverte de la pénicilline comme on peut le penser. L'utilisation des antibiotiques a débuté bien avant l'arrivée de l'Homme sur la terre. En effet, les bactéries sont les inventeurs et les premiers utilisateurs des antibiotiques. En fait, les antibiotiques font partie de l'arsenal de molécules bioactives dont disposent les bactéries pour survivre ou communiquer dans un environnement. Les antibiotiques sont des molécules utilisées pour combattre d'autres micro-organismes pour la défense d'un nid écologique. Le mycologue Français Paul Vuillemin fut le premier à utiliser le mot « antibiotique » dans son ouvrage « Antibiose et symbiose » publié en 1890 pour désigner le phénomène d'antagonisme entre des espèces microbiennes. La pénicilline, découverte par Fleming en 1929 et considérée comme la première molécule antibiotique mise sur le marché fut qualifiée de « molécule miracle » (Papp 1954).

La définition telle que nous le concevons aujourd'hui provient du microbiologiste américain Selman A. Waksman qui, en 1943 qualifia la streptomycine de « médicament antibiotique ». La streptomycine a été largement utilisée pour traiter la tuberculose, maladie sur laquelle la pénicilline n'avait pas d'effet thérapeutique. Puis en 1947, Selman tenta de donner une définition de l'antibiotique pour le différencier des autres moyens de traitement antimicrobien. Il définissait alors l'antibiotique comme « une substance chimique d'origine

microbienne qui inhibe la croissance ou les activités métaboliques de bactéries ou d'autres micro-organismes » (Bentley *et al.*, 2003) (Waksman, 1947).

A partir des années 50, avec le développement des méthodes de synthèse artificielle, la définition des antibiotiques qui se limitait aux molécules d'origines microbiennes, fut actualisée et s'étend désormais aussi aux molécules antibactériennes artificielles ou semi-synthétiques (Bentley *et al.*, 2003).

I. A. 3| Les grandes classes d'antibiotiques

Aujourd'hui, il existe plus de dix mille molécules antibiotiques sur le marché, réparties en un peu plus d'une quinzaine de familles (Davies *et al.*, 2010). Une famille d'antibiotique a la particularité d'avoir une structure commune et un mode d'action similaire. Les principales classes d'antibiotiques actuelles, quelques exemples et leurs cibles sont représentés dans le Tableau I.1 ci-dessous.

Classes d'antibiotiques	**Exemples**	**Actions sur**
β-lactamines	Pénicilline (ampicilline), Cephalosporine (cephamycine), pénèmes (méropénème), monobactames (astréonam)	Synthèse de la paroi bactérienne
Glycopeptides	Vancomycine	Synthèse de la paroi bactérienne
Aminoglycosides	Streptomycine, Gentamycine	Traduction
Lincosamides	Clindamycine	Traduction
Tétracylines	Minocycline	Traduction
Macrolides	Erythromycine	Traduction
Streptogramines	Synercid	Traduction
Oxazolidinones	Linézolide	Traduction
Phénicols	Chloramphénicol	Traduction
Rifamycines	Rifampicine	Transcription
Quinolones	Ciprofloxacine	Réplication
Pyrimidines	Triméthoprime	Métabolisme
Sulfonamides	Triméthoprime	Métabolisme
Lipopeptides	Daphtomicine	Membrane plasmique
Peptides cationiques	Colistine	Membrane plasmique

Tableau I.1: Modes d'action des classes d'antibiotiques, avec quelques exemples et leurs cibles.

Tableau adapté selon les références (Davies *et al.*, 2010; Morar *et al.*, 2010; Wright *et al.*, 2010).

Bien qu'il existe des milliers de molécules antibiotiques disponibles, le nombre de cibles reste limité.

I. A. 4| Les cibles des antibiotiques et quelques modes d'action

Les antibiotiques se lient généralement à une enzyme en bloquant une étape essentielle du développement de la bactérie. Selon les cibles et le mode d'action on peut classer les antibiotiques actuels en cinq catégories (Figure I.2:) :

(1) ceux qui inhibent la synthèse de la paroi bactérienne

(2) ceux qui détruisent l'intégrité de la membrane cytoplasmique

(3) ceux qui bloquent la machinerie responsable de la synthèse protéique (traduction)

(4) ceux qui inhibent la transcription ou la réplication de l'ADN

(5) ceux qui agissent au niveau de la biosynthèse des métabolites

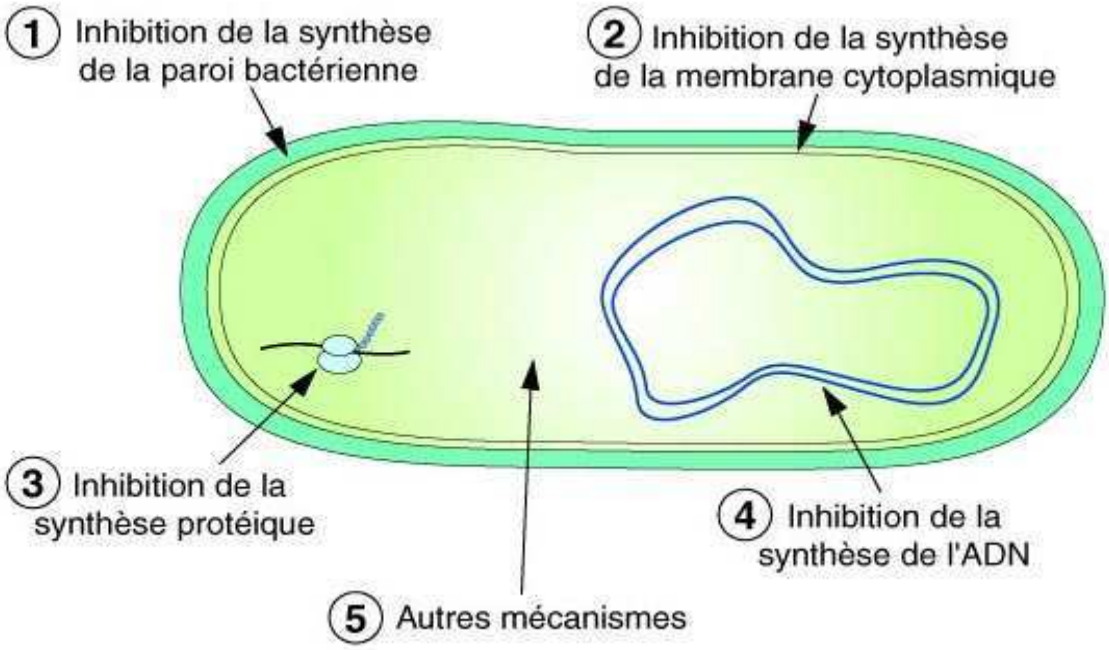


Figure I.2: Schéma des principales cibles d'antibiotiques dans une bactérie.

Des exemples caractéristiques non exhaustifs illustrant le type de mécanismes d'action de ces différentes catégories d'antibiotiques sont présentés ci-dessous.

I. A. 4. a| Les inhibiteurs de la paroi bactérienne

Comme inhibiteurs de la paroi bactérienne, la famille des β-lactamines est l'exemple le plus parlant. Cette famille représente à elle seule, un quart des antibiotiques utilisés. La particularité de cette grande famille est la possession d'un noyau β-lactame (Figure I.3).

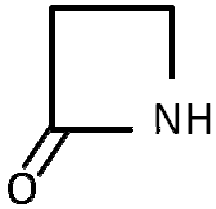


Figure I.3: un noyau β-lactame

La Figure I.4, ci-dessous montre la structure de la pénicilline (A) et son mécanisme d'action (B). Les pénicillines ont un cycle thiazolidine associé au noyau β-lactame, ressemblant au dipeptide D-Cys - D-Val, entrant dans la constitution de la paroi bactérienne. Cette ressemblance structurale permet aux pénicillines de se lier à deux enzymes appelées *penicillin-binding proteins* (PBP); la transglycosidase et la transpeptidase, impliquées dans la synthèse du peptidoglycane (Spratt *et al.*, 1988).

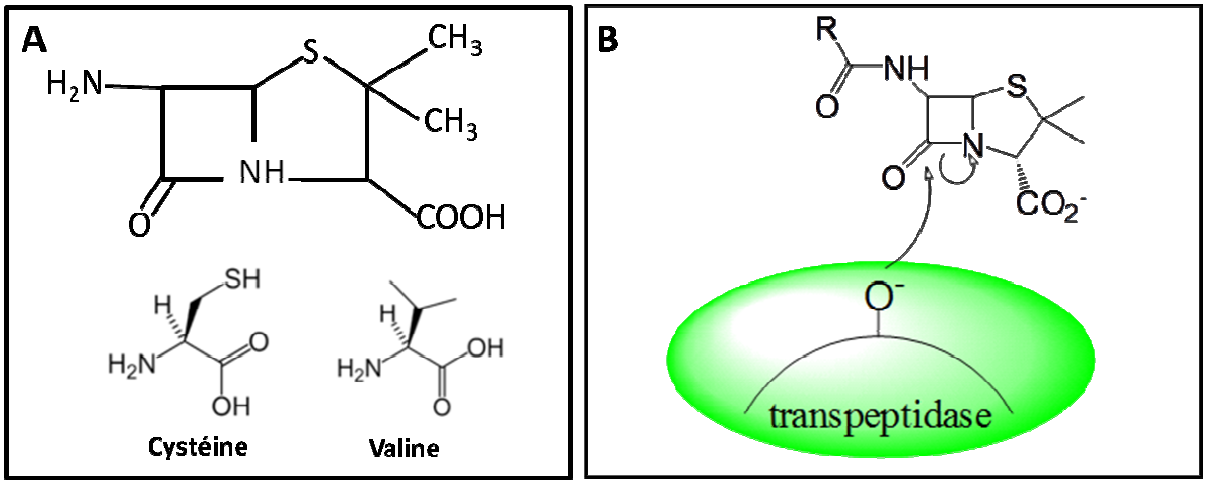


Figure I.4 : Structure des pénicillines. A : Analogie des pénicillines avec un dipeptide ; **B** : Mécanisme d'action de la transpeptidase.

I. A. 4. b| Les inhibiteurs de la synthèse protéique (au niveau de la traduction)

Le ribosome bactérien, différent de celui des eucaryotes, comprend deux sous-unités 30S et 50S, qui sont la cible de plusieurs antibiotiques (Tableau I.1) (Carter *et al.*, 2000; Hansen *et al.*, 2002).

La famille des oxazolidinones (Slee *et al.*, 1987), l'une des rares familles apparue au cours de ces 40 dernières années, sera évoquée ici. Les oxazolidinones sont des molécules développées artificiellement contre les germes Gram (+) et ils constituent la dernière ligne de défense contre certaines bactéries multi-résistantes comme *Staphylococcus aureus* résistant à la méticilline (SARM). Le linézolide (Figure I.5) est actuellement la seule molécule de la famille approuvée pour la commercialisation. Cette molécule agit en empêchant la formation du complexe d'initiation 70 S (formé de la sous-unité 50 S, de la sous-unité 30 S, de l'ARN messager, et de l'ARN de transfert).

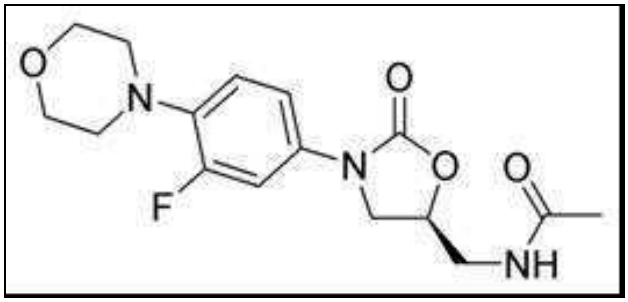


Figure I.5: Le linézolide

I. A. 4. c| Les inhibiteurs de la transcription ou de la réplication de l'ADN

La rifampicine (Figure I.6) se lie à l'ARN polymérase et empêche sa fixation à l'ADN, inhibant ainsi la transcription. Du fait de leur cible, certaines molécules de cette famille peuvent être utilisées comme antiviraux ou anti-tumoraux. La rifampicine est souvent associée à la streptomycine (Aminosides) dans le traitement de la tuberculose (provoquée par des mycobactéries) et est souvent efficace contre certains germes Gram (-).

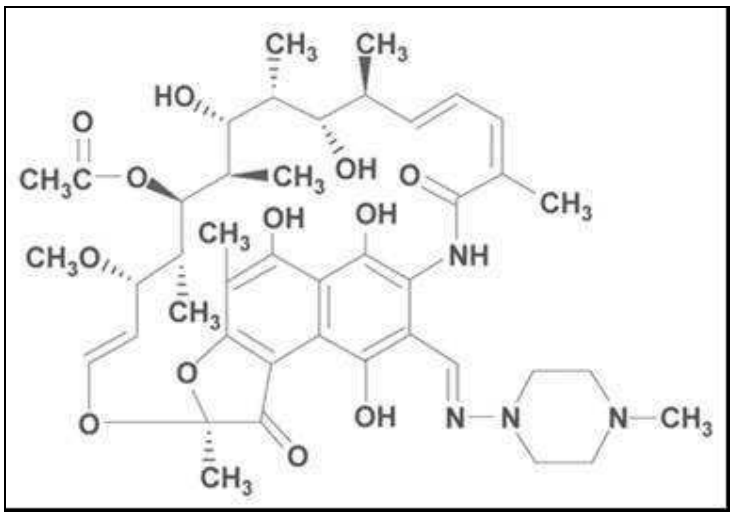


Figure I.6: La rifampicine ou N-amino-N'-méthylpipérazine-hydrazone

I. A. 4. d| Les inhibiteurs des voies de biosynthèses bactériennes

L'acide folique ou vitamine B_9, est le précurseur métabolique d'un coenzyme, le tétrahydrofolate (THFA), cofacteur de la synthèse ultérieure des bases puriques et pyrimidiques de l'ADN. Les enzymes impliquées dans cette voie de biosynthèse constituent des cibles de choix pour les molécules antibiotiques comme les sulfamides. Les sulfamides sont des dérivés de l'acide acide para-aminobenzoïque (PABA) (Figure I.7) (Tréfouël *et al.*, 1935).

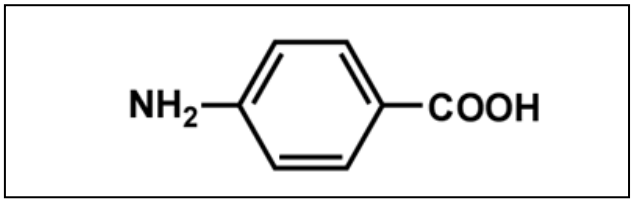


Figure I.7 : acide para-aminobenzoïque (PABA)

Les sulfamides sont généralement utilisés en co-traitement avec les diaminopyridines (triméthoprime) avec lesquels ils agissent en synergie en inhibant la synthèse de l'acide tétrahydrofolique (Figure I.8).

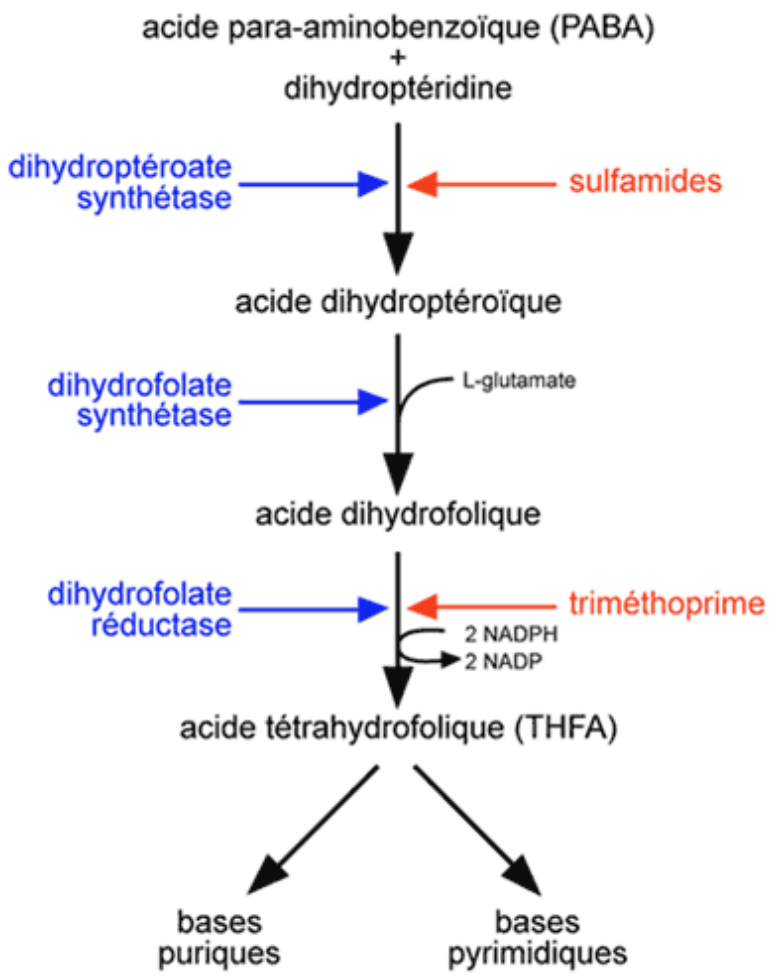


Figure I.8: Voie de synthèse de l'acide tétrahydrofolique et inhibition de l'action par des antibactériens (en rouge)

I. A. 4. e| Les inhibiteurs de la membrane cellulaire

Les polymyxines, actifs principalement sur les Gram (-), agissent comme des détergents cationiques. Grâce à leur caractère amphipathique, elles pénètrent dans la cellule bactérienne et s'insèrent parmi les phospholipides de la paroi, perturbant ainsi la perméabilité membranaire. La colistine (Figure I.9), l'une des rares polymyxines de la famille à être utilisée en thérapeutique est composé de la polymyxine E1, appelé colistine A et de la polymyxine E2, isolées à l'origine chez la bactérie *Bacillus colistinus* (Meyer *et al.*, 2004).

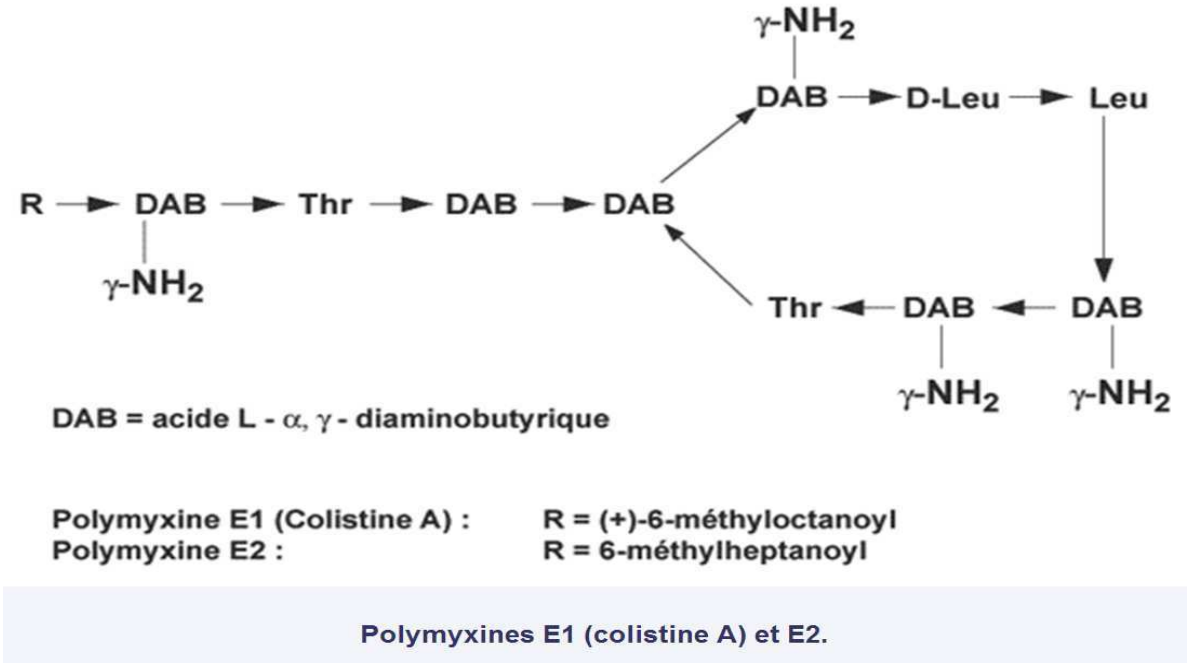


Figure I.9: Constituants de la colistine

I. A. 5| Conclusions sur les antibiotiques

Les antibiotiques ont largement contribué à l'amélioration de la santé. Le traitement antibiotique des infections bactériennes a constitué à partir de 1941 une véritable révolution, dans un premier temps sur la santé proprement dite, puis ensuite sur le développement de la recherche médicale en général. Les antibiotiques ont fait la puissance des grandes entreprises pharmaceutiques. La recherche sur les antibiotiques a ouvert la voie à des procédés de développement de médicaments. On estime que les antibiotiques ont augmenté la durée de vie de l'homme de près de 15 ans. Pour comparaison, les anticancéreux augmentent la durée de vie de quelques semaines à 5 ans maximum. L'utilisation massive et inadéquate de ces antibiotiques durant plusieurs décennies a cependant augmenté la vitesse d'apparition des bactéries résistantes aux antibiotiques. Depuis quelques années, ce nouveau fléau est devenu le problème majeur dans la lutte contre les maladies infectieuses. Nous allons maintenant décrire ces mécanismes de résistance avant de mentionner quelques stratégies envisagées pour remédier au problème de la résistance.

I. B| La résistance bactérienne aux antibiotiques

I. B 1| Définition

Le terme « résistance » est utilisée pour désigner tout agent pathogène qui est moins sensible que ses homologues à un composé antimicrobien selon le journal « American Academy of Microbiology ». La résistance aux antibiotiques n'est pas un phénomène nouveau. Elle n'est qu'une conséquence de la sélection naturelle, les antibiotiques ne faisant que sélectionner les bactéries aptes à survivre.

I. B 2| Quelques rappels des faits

En 1945, Alexander Flemming, le découvreur du premier antibiotique, anticipait les risques liés à une mauvaise utilisation de la pénicilline, en confiant au New York Times : *« Cela aboutirait à ce qu'au lieu d'éliminer l'infection, on apprenne aux microbes à résister à la pénicilline et à ce que ces microbes soient transmis d'un individu à l'autre jusqu'à ce qu'ils en atteignent un chez qui ils provoquent une pneumonie ou une septicémie que la pénicilline ne pourra pas guérir »*. Cette prédiction devint très vite une réalité. Effectivement, un an après cette alerte, plusieurs souches résistantes furent isolées chez des patients atteints de septicémies, selon un rapport de l'organisation mondiale de la santé. A l'époque, le problème a été sous-estimé et ignoré. Les médecins ont cru à un phénomène limité qui serait compensé par la découverte de nouvelles molécules. A la fin des années 60, après le « golden age » des antibiotiques (confère Figure I.1:), l'Homme pensait pouvoir éradiquer les maladies infectieuses, à l'image de ce chirurgien américain, le Dr William H. Stewart qui affirmait que les maladies infectieuses seraient vaincues par les antibiotiques et les vaccins disponibles et que par conséquent, il était temps de détourner l'attention du gouvernement américain et des ressources à la «guerre contre le cancer » (Choffnes *et al.*, 2010).

Dès lors, les grandes firmes pharmaceutiques abandonnèrent de façon « brutale » la production d'antibiotiques préférant se tourner vers d'autres molécules thérapeutiques plus rentables. D'autres raisons de cet abandon peuvent s'expliquer par le faible rendement des antibiotiques et la fréquence d'apparition de résistance. Ce qui fait que depuis quarante ans, les antibiotiques efficaces se raréfient. Effectivement, depuis 1970, seules deux nouvelles familles d'antibiotiques, les oxazolidinones (linézolide) (Slee *et al.*, 1987) et les lipopeptides

(daptomycine) (Tally *et al.*, 2000), ont été mis sur le marché. L'innovation qu'on pensait voir se poursuivre sans arrêt a pris fin.

Parallèlement à la raréfaction de nouveaux antibiotiques, le nombre de souches pathogènes résistantes aux antibiotiques disponibles s'est considérablement accru. Or, chaque nouvelle molécule antibiotique découverte a été suivie par l'apparition de souches résistantes à cette molécule. A l'heure actuelle le phénomène est plus que préoccupant au vu de la vitesse de propagation et de l'expansion de la résistance. Cette expansion globale et accélérée de la résistance n'est rien d'autre qu'une démonstration de l'évolution des bactéries et une réponse à l'utilisation massive des antibiotiques. En effet, l'amplification du phénomène de résistance prend effet après l'adoption des antibiotiques et l'utilisation abusive dans l'industrie agro-alimentaire des antibiotiques à partir des années 1950. Ceci provoqua une dissémination massive d'antibiotiques dans la nature touchant toute la flore microbienne. En réponse, les bactéries ont développés des stratégies pour survivre avec ces molécules antibiotiques ; celles-ci à l'origine du fléau de la résistance croissante.

I. B. 3| Les différents types de résistance

I. B. 3. a| La résistance naturelle

On parle de résistance naturelle lorsque le phénomène est observé chez toutes les souches d'une même espèce. Par exemple, l'espèce *Pseudomonas aeruginosa* possède une paroi cellulaire peu perméable, ce qui lui permet de bloquer l'entrée de certains antibiotiques. Globalement, toutes les bactéries productrices d'antibiotiques sont naturellement résistantes aux molécules qu'elles produisent. C'est le principe du poison et de l'antidote. Or une grande partie des antibiotiques utilisés par l'Homme est d'origine bactérienne (Davies *et al.*, 2010). De là, il n'est pas difficile de comprendre qu'il existe plusieurs bactéries dans la nature possédant des mécanismes de résistance contre les antibiotiques. Le problème concerne les bactéries pathogènes qui peuvent aussi développer ou acquérir ces mécanismes de résistance.

I. B. 3. b| La résistance acquise

La résistance acquise survient lorsque quelques souches d'une même espèce, normalement sensibles à un antibiotique deviennent résistantes. Cette faculté est une conséquence des modifications génétiques qui surviennent sur le génome bactérien. Ces

modifications génétiques peuvent être dues à une mutagenèse chromosomique ou acquises par le transfert d'éléments génétiques.

- Résistance acquise par mutagenèse chromosomique

C'est un phénomène spontané et rare apparaissant au cours de la réplication de l'ADN bactérien. Il se traduit à la suite d'une mutation intervenant à une fréquence moyenne de 10^{-6}. En conditions de stress, comme en présence d'antibiotique, cette fréquence de mutation est multipliée par un facteur 1000, générant ainsi plusieurs clones. Les clones capables de survivre se développent assurant la pérennité de l'espèce. Ce phénomène est connu sous le nom de « tâtonnement génétique » et peut conduire à la résistance à toute une famille d'antibiotiques (Choffnes *et al.*, 2010). On peut voir cela comme une conséquence de la pression de sélection.

- Résistance acquise par transfert

La bactérie récupère une information génétique exogène permettant, entre autres, l'acquisition de la résistance à un antibiotique. La grande majorité des gènes de résistance sont situés sur des éléments génétiques comme les intégrons, les transposons et les plasmides, facilement échangeables entre les bactéries (DeLeo *et al.*, 2009). Le transfert peut se faire par la conjugaison, la transduction ou par la transformation. Le transfert est dit « horizontal » lorsque le gène est transféré à une bactérie d'une autre espèce. Par exemple : une souche d'*Escherichia coli* peut acquérir une résistance aux sulfamidés par un transfert horizontal grâce à une souche de *Shigella dysenteriae* résistante aux sulfamidés.

La conjugaison est le transfert de l'ADN d'une bactérie donneuse à une bactérie réceptrice au cours d'un contact cellulaire étroit qui a lieu grâce aux pili.

Dans la transduction, le vecteur est un bactériophage pouvant incorporer une partie du génome bactérien infecté. En quittant la cellule, il peut emporter des gènes supplémentaires (bactériens) qui pourront être transfectés chez une autre bactérie.

Pour la transformation, la bactérie récupère et incorpore de l'ADN exogène présent dans son environnement.

I. B. 4| Les mécanismes moléculaires de défense contre les antibiotiques

Il existe plusieurs types de mécanismes décrits dans la littérature (Yoneyama *et al.*, 2006) (Figure I.10) :

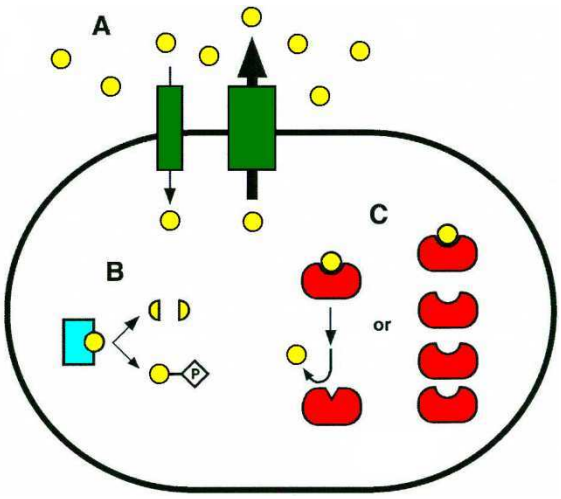


Figure I.10: Principaux mécanismes de résistance aux antibiotiques. A : La modification de l'entrée et de la sortie de l'antibiotique ; B : hydrolyse ou modification chimique de l'antibiotique; et C : La modification ou augmentation de la production de la cible de l'antibiotique. Les sphères jaunes représentent l'antibiotique. (Yoneyama *et al.*, 2006).

I. B. 4. a| La modification de la cible de l'antibiotique

Chaque antibiotique agit en se fixant sur une cible précise dans la cellule : paroi, ribosome, etc.… La présence d'une modification consécutive à une mutation modifie le site de fixation et empêche la liaison de l'antibiotique à sa cible. C'est un des mécanismes de résistance à la streptomycine, l'un des premiers antibiotiques utilisés pour traiter la tuberculose. Le pneumocoque *Streptococcus pneumoniae* a développé une résistance par modification d'une protéine membranaire spécifique où se fixent les pénicillines. La bactérie peut aussi diminuer l'affinité de l'antibiotique pour sa cible protéique (Spratt 1994). Des enzymes spécifiques peuvent effectuer une modification chimique post-traductionnelle covalente de la cible, par exemple une méthylation, ce qui inhibe la fixation de l'antibiotique, sans qu'il y ait altération du génome. Ce type de mécanisme est rencontré dans la résistance aux macrolides, où il existe une méthylase qui confère la résistance en modifiant l'ARN ribosomique au niveau du site de liaison de l'antibiotique.

I. B. 4. b| La modification de l'antibiotique

De nombreuses souches résistantes fabriquent des enzymes qui modifient chimiquement l'antibiotique, par exemple par phosphorylation (Walsh 2000), clivage ou hydrolyse, la rendant inactive. C'est le mécanisme principal de résistance aux β-lactamines (famille de la pénicilline et des céphalosporines) qui implique les enzymes de la famille des β-lactamases. Les staphylocoques méti-résistants (SARM) sont insensibles aux pénicillines grâce à la modification des enzymes liant les pénicillines (PBP) mais aussi, par la production d'une β-lactamase et d'une méticilinase (Walsh 2010) qui clivent l'antibiotique.

I. B. 4. c| La modification de la perméabilité membranaire

Par prévention de l'accumulation de l'antibiotique dans le cytoplasme, la bactérie peut inhiber les systèmes d'import en fermant des pores et augmenter la synthèse des systèmes d'efflux (Nikaido 1994). Les bactéries peuvent aussi expulser par pompage actif hors de la cellule les composés toxiques comme les antibiotiques (Gristwood *et al.,* 2008). C'est par exemple, l'un des principaux mécanismes de résistance de *Pseudomonas aeruginosa*, pathogène opportuniste (Davies 1996).

En plus de ces mécanismes, les bactéries peuvent bénéficier d'autres phénomènes pour résister contre les antibiotiques. Il s'agit de la défense altruiste pouvant se produire dans les biofilms où des bactéries résistantes synthétisent une substance en très grande quantité pour subvenir aux besoins des bactéries sensibles (Hancock *et al.,* 2008).

Les bactéries peuvent aussi perdre la résistance acquise lorsqu'elle n'est plus nécessaire, par exemple, en absence d'antibiotiques. Il faut noter que les mécanismes de résistance sont extrêmement coûteux en énergie pour la bactérie.

I. B. 5| La résistance bactérienne : un problème de santé majeur

L'organisation mondiale de la santé (OMS) a fait de la résistance microbienne sa priorité pour le 21ème siècle. En effet, en début du siècle, l'OMS qualifiait la résistance microbienne de « problème majeur pour la médecine du XXIème siècle ». Depuis, chaque année, l'OMS dédie une journée mondiale de la santé aux maladies infectieuses principalement axée sur le problème de la résistance. Les maladies infectieuses représentent

encore et depuis toujours la première cause de mortalité et de morbidité de l'humanité selon l'OMS. Le phénomène de résistance croissant et touchant les maladies pandémiques comme la tuberculose, est un véritable problème dans la lutte antimicrobienne. La tuberculose (TB) causée par le bacille *Mycobacterium tuberculosis,* qui fait entre 1,2 et 1,5 million de morts en 2008 (OMS 2011), représente l'un des pires fléaux de ces derniers siècles en matière de santé. Cette maladie qui infecte un tiers de la population mondiale est traitée avec une association de plusieurs antibiotiques antituberculeux, les plus efficaces étant l'isoniazide et la rifampicine. Selon l'OMS, il existe des preuves de l'existence de souches qui résistent à l'un, voire à l'ensemble de ces médicaments efficaces. Bien que de nouveaux médicaments soient en cours de développement pour cette maladie, comme le TMC207 (Lounis *et al.,* 2010a; Veziris *et al.,* 2006), près d'un demi-million de personnes ayant contracté des formes de tuberculose multi résistantes en sont morts en 2008 (OMS 2011). La résistance est aussi la principale cause des infections nosocomiales ou hospitalières. En France, ces infections sont passées de 296 059 cas à 358 353 cas entre 2001 et 2006 selon l'enquête de prévalence nationale réalisée en 2006 (RAISIN 2009). Parmi les infections nosocomiales, celles provoquées par les bactéries sont les plus nombreuses et les plus inquiétantes. Aux Etats-Unis, les infections nosocomiales dues à une résistance bactérienne feraient plus de 63 000 décès par an, selon un rapport de 2009 de l'organisation « Resources for the Future » (Choffnes *et al.,* 2010). En Europe, ce nombre concernant uniquement les bactéries multi-résistantes dépasserait les 25 000 cas par an (ECDC/EMEA 2009). Le surcoût engendré par le phénomène de la résistance serait estimé à 900 millions d'euros au niveau européen (ECDC/EMEA 2009).

La résistance, ayant d'abord été constatée à l'hôpital, touche aujourd'hui le milieu communautaire. Certaines pathogènes multi-résistants seraient insensibles à l'ensemble des antibiotiques utilisés pour leur traitement. Les souches bactériennes: *Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumanii, Pseudomonas aeruginosa* et les espèces *Enterobacter* regroupées sous l'acronyme ESKAPE sont responsables de la majorité des décès due aux infections nosocomiales (Choffnes *et al.,* 2010). Le cas le plus alarmant concernerait les souches pathogènes multi-résistantes de *Staphylococcus aureus* : celles résistantes à la méticilline (SARM) et son homologue communautaire celles résistantes à la vancomycine (VRSA). Aux USA, ces souches de Staphylocoques provoqueraient plus de décès que le virus du VIH (Choffnes *et al.,* 2010).

La bactérie *Escherichia coli* (*E. coli*) utilisée pour l'étude des inhibiteurs anti-Fur dans cette thèse est aussi concernée par le phénomène de résistance. Elle appartient aux Gram (-) qui possèdent une double membrane, jouant le rôle de bouclier contre les antibiotiques. Bactérie généralement commensale du tube digestif des mammifères, elle peut exister sous certaines formes pathogènes et virulentes parfois résistantes. D'ailleurs, une souche très rare et extrêmement virulente a été l'origine d'une épidémique meurtrière en Allemagne au printemps 2011. Cette souche qui avait contaminé 1500 personnes pour une vingtaine de décès, contenait des gènes lui conférant une résistance aux antibiotiques usuels.

Les souches résistantes ou « super bactéries », agents pathogènes meurtriers ne sont que l'expression extrême d'un phénomène beaucoup plus inquiétant. Les conséquences vont au delà de la résurgence de maladies mortelles et menacent de nombreuses autres interventions qui permettent de sauver des vies, comme les traitements anticancéreux, certaines opérations chirurgicales, ou les transplantations d'organes. Les hôpitaux devenant des foyers d'agents pathogènes hautement résistants, ces actes pourraient devenir dangereux.

Les souches pathogènes pourraient être aussi impliquées dans d'autres maladies pour lesquelles aucun remède n'est connu à ce jour, comme les maladies dégénératives. En effet, une récente étude américaine a montré qu'une infection liée à *Helicobacter pylori* joue un rôle important dans le développement de la maladie de Parkinson chez l'homme (McGee *et al.*, 2011). Cette étude a montré que le cholestérol était subtilisé par *H. pylori* pour en faire une molécule toxique pour le cerveau (McGee *et al.*, 2011).

Dans tous les cas, pour échapper au cauchemar d'un retour à l'ère pré-antibiotique, il est impératif d'inventer des nouveaux moyens permettant d'enrayer la diffusion des bactéries résistantes, de modifier radicalement la manière irresponsable de dilapider les antibiotiques et de disposer de nouveaux traitements efficaces contre les germes résistants de demain.

I. C| Les nouvelles stratégies de lutte antibactérienne

Il est admis par toute la communauté qu'il est impossible d'éradiquer les microbes de la terre. Ces micro-organismes sont capables de s'adapter et survivre dans les conditions les plus drastiques et sont aussi « vieux » que la vie. Comme le suggère Spellberg, on ne peut tenter qu'une utopique cohabitation paisible (Spellberg 2008). Il est toujours possible de développer des antibiotiques de plus en plus efficaces mais le phénomène de résistance persistera du fait de la protorésistance. De ce fait, Spellberg recommande la préservation de

l'effet des antibiotiques existants aussi longtemps que possible tout en encourageant le développement de nouveaux antibiotiques (Spellberg 2008). D'autres alternatives aux antibiotiques, non explorées et qui ne seront pas développées dans ce manuscrit, restent possible pour la lutte antibactérienne. On peut citer les vaccins antibactériens, l'immunothérapie et la phagothérapie (Cohen 2010). Cette dernière a été testée avec des bactériophages lysogéniques contre des infections dues à des souches résistantes d'*Acinetobacter* sur des soldats de l'armée américaine en Afghanistan et en Iraq (Lu *et al.*, 2007, 2009).

Ces autres moyens de lutte pourraient être intéressants si l'on ne parvenait pas à trouver une stratégie efficace contre le problème de résistance lié aux antibiotiques.

I. C. 1| Les campagnes de sensibilisation et les programmes de surveillance de la résistance

Il existe aujourd'hui dans plusieurs pays des programmes de surveillance et de contrôle sur l'apparition, la dissémination et la propagation des souches pathogènes résistantes. En France, ces programmes sur la résistance existent depuis 1978. Il existe aussi des campagnes de sensibilisation de la population pour une utilisation adéquate des antibiotiques. En France, ces campagnes sont organisées chaque année par les pouvoirs publics afin d'informer toute la population. En 2005, un « Programme national de lutte contre les infections nosocomiales » a été mis en place. Cette démarche, qui repose sur une utilisation plus restreinte des antibiotiques et une amélioration des conditions d'hygiène dans les hôpitaux, a commencé à porter ses fruits puisque le nombre de décès causé par les infections nosocomiales est passé de 10 000 à environ 4 000 entre 2001 et 2006 (RAISIN 2009). Pour mettre ces programmes en place au niveau mondial, l'OMS a lancé, pendant la journée mondiale de la santé 2011, un appel international pour une action concertée afin de stopper la propagation de la résistance aux antimicrobiens et recommande aussi aux gouvernements une politique qui se résume en ces points suivants:

- Un engagement des gouvernements dans la lutte contre la résistance aux médicaments
- Une mise en place mondiale de la surveillance des résistances et des mesures sur la propagation des infections avec des moyens appropriés
- Une offre de médicaments de qualité comme point d'action essentiel

- Une règlementation obligatoire de l'usage rationnel des médicaments dans le monde entier
- Un dialogue ouvert entre les gouvernements, les bailleurs de fonds et les industries pharmaceutiques pour identifier les nouvelles stratégies et de nouveaux outils qui permettront de faire face à la résistance de manière efficace.

Nous pouvons aussi parler des nouvelles mesures d'hygiène utilisées dans les hôpitaux pour stopper la propagation des microbes. Une de ces mesures, concerne l'utilisation du cuivre pour lutter contre les infections nosocomiales. En effet, une étude récente a montré que 90 à 100 % de staphylocoques dorés SARM, déposés sur une surface de cuivre sont tués en quelques minutes tandis que ces bactéries continuent de survivre sur une surface témoin en inox (Casey *et al.*, 2010). L'utilisation du cuivre dans les hôpitaux pourrait donc être une voie à exploiter pour limiter la propagation des infections.

I. C. 2| L'identification de nouvelles cibles et de nouveaux antibiotiques

Les antibiotiques actuellement disponibles sur le marché agissent contre un nombre limité de cibles comme nous l'avons évoqué précédemment. Il est donc possible d'identifier de nouvelles protéines bactériennes, cibles de nouvelles molécules potentiellement antibiotiques. La cible doit être indispensable à la croissance et/ou la viabilité cellulaire, absente chez l'hôte et exprimée par la bactérie lors de l'infection. Ce dernier point est critique puisque des études ont montré que les protéines exprimées par les bactéries durant l'infection peuvent être différentes de celles exprimées dans un milieu de culture (McDevitt *et al.*, 2001; Nathan 2004). L'identification de nouvelles cibles en laboratoire n'est donc pas simple. Notons au passage, que la protéine Fur, sur laquelle ce projet porte, respecte bien ce « cahier de charge ».

Les nouvelles molécules antibio-thérapeutiques, quant à elles, doivent être structurellement différentes des antibiotiques dont une résistance a été notée. C'est aussi le cas des aptamères peptidiques et des peptides dérivés pour lesquels, il n'existe aucune homologie de séquence dans la littérature. Les dernières générations d'antibiotiques sont en fait, pour la plupart, des molécules naturelles synthétisées par les bactéries pour leur survie, mais ayant subi des modifications artificielles afin d'acquérir d'autres propriétés comme un spectre d'action plus étendu et/ou une meilleure biodisponibilité dans l'organisme hôte. On

peut citer la famille des β-lactamines et ses différentes générations. Aujourd'hui, il existe plusieurs pathogènes (comme SARM) résistantes à l'ensemble de cette famille (Chambers *et al.,* 2009; DeLeo *et al.,* 2009). Il faut donc rechercher des molécules ayant un nouveau « scalfold » (Moy *et al.,* 2006). Il est aussi possible d'approfondir les recherches effectuées sur les produits naturels avec les nouveaux progrès faits et d'identifier de nouvelles molécules antibiotiques. En effet, de nouveaux produits naturels provenant de la vaste majorité des espèces bactériennes non-cultivables représentent une source pleine de molécules antibiotiques non utilisées (D'Onofrio *et al.,* 2010).

Les progrès récents en biologie, en chimie et en informatique ont vu naitre de nouveaux outils qui permettront de concevoir l'antibiotique idéal qui serait une pro-drogue : la molécule serait activée par une enzyme spécifique bactérienne à l'intérieur de la cellule et se lierait à sa cible de façon covalente pour éviter le mécanisme d'efflux et posséderait un effet antibiotique à large spectre (Choffnes *et al.,* 2010).

Dans la suite, nous évoquerons quelques cas d'études où ces nouveaux outils, qui laissent beaucoup d'espoirs pour le développement de futurs nouveaux antibiotiques efficaces, ont été utilisés.

I. C. 3| Quelques exemples d'antibiotiques potentiels en cours d'étude

Cette partie décrira les nouvelles techniques couramment utilisées pour la recherche de nouvelles molécules bioactives et de nouvelles cibles. Nous passerons également en revue quelques cibles bactériennes actuellement étudiées et les molécules inhibitrices de ces cibles où des méthodes, comme la génomique, les outils bioinformatiques, le criblage haut-débit et la chimie combinatoire, sont utilisées.

- Le séquençage de plus de 700 génomes bactériens ces dernières années, parmi lesquels de nombreux pathogènes, a permis l'identification et la validation de nouvelles cibles thérapeutiques (McDevitt *et al.,* 2001).

- Les librairies de produits synthétiques issues de la chimie combinatoire, appelés chimiothèques, sont aussi un moyen de produire des nouvelles molécules non exploitées. Récemment, la chimie combinatoire a permis la découverte d'une nouvelle classe de molécules antituberculeuses ; les diarylquinolines (Lounis *et al.,* 2010b). Un composé noté TMC207, sélectionné lors d'un criblage à haut débit parmi plus de 70 000 molécules, est capable d'inhiber la croissance de *Mycobacterium tuberculosis in vivo* chez des patients

atteints de tuberculose (Lounis *et al.,* 2010b). Cette molécule agit sur une nouvelle cible : l'ATP synthase, enzyme responsable de la synthèse d'ATP.

- Différents travaux reportent des candidats potentiels, comme la peptide deformylase, une enzyme indispensable à la maturation des protéines, ou la déacétylase, enzyme intervenant dans la biosynthèse du lipide A, un constituant de la membrane externe des bactéries à Gram (-) (Clements *et al.,* 2002). Ces auteurs ont en effet montré que des inhibiteurs spécifiques de ces deux enzymes entrainaient un arrêt de croissance puis la mort d'une souche d'*Escherichia coli* pathogène.

- En utilisant plusieurs des outils cités ci-dessus, le groupe de Levy est parvenu à la découverte d'une nouvelle classe de petites molécules inhibitrices de la virulence chez *Yersinia pseudotuberculosis*. Cette nouvelle famille, basée sur le « scalfold » du N-hydroxybenzimidazole agit sur une nouvelle cible : le facteur de transcription LcrF de MAR, en bloquant sa liaison à l'ADN (Alekshun *et al.,* 1999; Garrity-Ryan *et al.,* 2010; Kim *et al.,* 2009).

- Les peptides antimicrobiens cationiques et amphiphiles sont des molécules de 20 à 50 acides aminés, présentant une grande diversité structurale et un large spectre d'activité antimicrobien (Andres *et al.,* 2007). Considérés comme l'un des éléments clefs du système immunitaire inné des organismes multicellulaires, ces molécules tuent les microorganismes rapidement en perméabilisant leur membrane par un effet de type détergent ou par la formation de pores. Plusieurs de ces peptides ou analogues sont actuellement en cours de développement dans des infections systémiques et pourraient bientôt faire partie des médicaments anti-infectieux de demain (Andres 2011).

I. C. 4| La stratégie consistant à diminuer la virulence des pathogènes

Ces dernières découvertes semblent encourageantes, mais le problème reste le même. Il s'agit d'une course perpétuelle entre la mise au point des molécules antiobiotiques et l'apparition de la résistance bactérienne. C'est pourquoi une des nouvelles approche consiste non plus à tuer les bactéries ou bloquer leur croissance, mais plutôt de diminuer leur virulence, c'est-à-dire leur capacité à infecter l'hôte et causer une pathologie (Cegelski *et al.,* 2008) (Walsh 2000, 2003) (Garrity-Ryan *et al.,* 2010). C'est d'ailleurs l'approche envisagée

dans ce projet. Dans ces conditions, la bactérie ne serait plus soumise à une pression de sélection et serait libre de se diviser, mais ne serait plus pathogène pour l'homme.

Chacune des étapes représentées dans la Figure I.11 peut être une cible potentielle pour l'action d'un nouvel antibactérien agissant sur la virulence. Il est possible de cibler les facteurs de virulence comme les toxines, les adhésines ou d'autres déterminants de résistance aux antibiotiques (Clatworthy *et al.*, 2007).

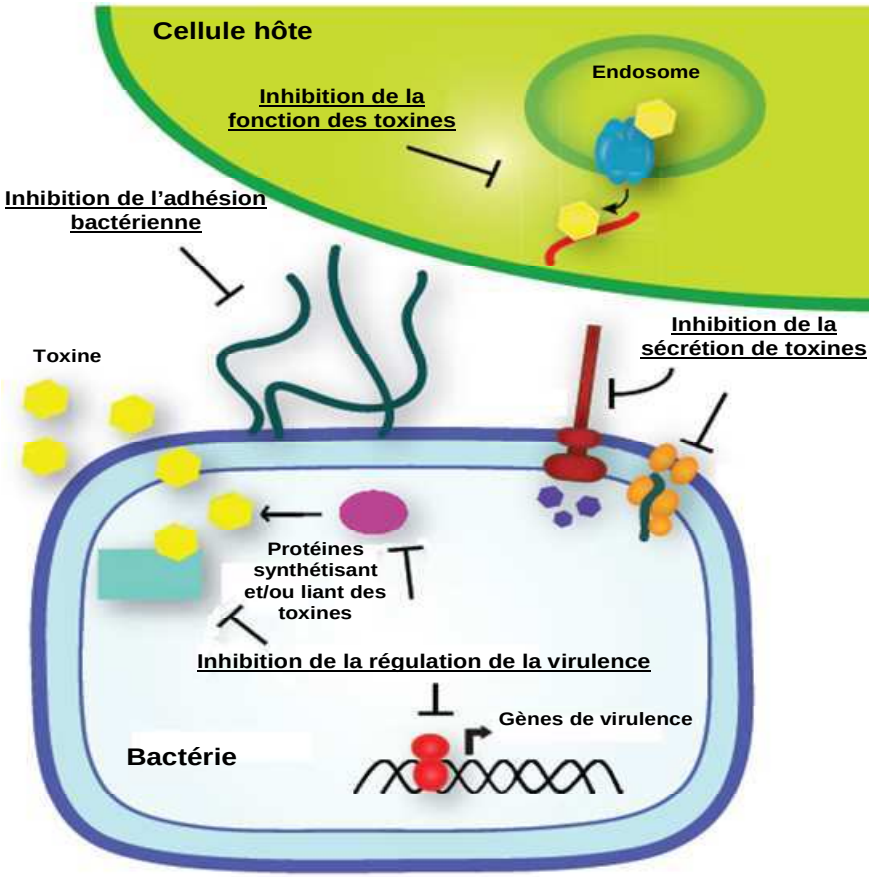


Figure I.11: Cibles potentielles des inhibiteurs de virulence. Inhibition de la transcription des gènes de virulence, de l'adhésion bactérienne, de la sécrétion des toxines ou de leur fonction. Modifié d'après (Clatworthy *et al.*, 2007).

Parmi les cibles concernées par cette nouvelle approche, on peut citer les régulateurs de transcription qui codent pour l'expression de gènes impliqués dans l'adhésion cellulaire, la production et la sécrétion de toxines. Ces dernires consituent des cibles potentielles très intéressantes pour la mise au point de nouveaux antibiotiques entrant dans la catégorie de la nouvelle approche (Garrity-Ryan *et al.*, 2010; Grier *et al.*, 2010; Kim *et al.*, 2009). Comme nous l'avons vu pour LcrF, le facteur de transcription ToxT de *Vibrio cholerae*, qui active l'expression du gène de la toxine cholérique, peut être une cible potentielle (Hung *et al.*,

2005). Outre ces facteurs, d'autres régulateurs transcriptionnels pourraient être intéressants comme cible potentielle.

Parmi eux, il y a la protéine Fur, qui module l'expression de plus de 190 gènes chez *Escherichia coli*, dont des facteurs de virulence, mais également de nombreux gènes impliqués dans l'assimilation du fer (Hantke 2001) (Abed *et al.,* 2007). Il a d'ailleurs été montré chez des pathogènes tels que *Pseudomonas aeruginosa* (Hassett *et al.,* 1996), *Erwinia chrysanthemi* (Franza *et al.,* 1999), *Staphylococcus aureus* (Horsburgh *et al.,* 2001), *Helicobacter pylori* (Bury-Mone *et al.,* 2004) et *Bacillus cereus* (Harvie *et al.,* 2005) qu'une inactivation du gène *fur* entraînait une diminution de la virulence de ces bactéries.

Différentes études par le passé ont montré que l'excès de fer augmente la virulence de nombreux genres bactériens dont *Escherichia et Neisseria* (Bullen *et al.,* 1968; Bullen *et al.,* 1971), *Listeria* et *Pasteurella* (Martinez *et al.,* 1990), *Salmonella* (Griffiths 1991), *Yersinia* et *Vibrio* (Chart *et al.,* 1985). Dans le même sens, les individus atteints de surcharge en fer chronique (hémochromatose) seraient plus sensibles aux infections par les pathogènes *Escherichia coli*, *Listeria monocytogenes* et *Vibrio vulnificus* (Khan *et al.,* 2007). La possession d'un système efficace d'assimilation du fer est donc indispensable à la survie d'un pathogène et à sa capacité à infecter l'hôte (Schaible *et al.,* 2004).

Afin de mieux comprendre le lien qu'il existe entre la virulence et l'homéostasie du fer, les différentes voies d'acquisitions, de transport et de gestion du fer chez l'hôte (l'Homme) et chez les bactéries seront décrites brièvement dans la section suivante.

II| Le fer dans le vivant

II. A| Le fer est indispensable à la vie

Hormis quelques bactéries, tous les organismes vivants ont besoin du fer pour vivre. Cet élément accompagne tous les processus indispensables à la vie tels que : la synthèse des acides aminés, la chaine de transfert d'électrons, la respiration, la fixation de l'azote, le cycle de l'acide citrique, le transport et le métabolisme de l'oxygène, la synthèse et la réplication de l'ADN (Sandy *et al.*, 2009). Cependant, sa bio-disponibilité pour des besoins physiologiques représente un challenge et un véritable enjeu de compétition entre les êtres vivants d'un même biotope (Guerinot 1994). En conditions d'aérobiose et à pH physiologique, le fer existe sous la forme d'hydroxyde ferrique $Fe(OH)_3$ ayant une constante de solubilité de seulement 10^{-39} M (Sandy *et al.*, 2009).

De plus, le fer réduit à l'état ferreux Fe(II) est extrêmement toxique pour la cellule en raison de sa capacité à réagir avec l'oxygène et à catalyser la production d'espèces réactifs à l'oxygène dans la réaction de Fenton (Fe^{2+} + H2O2→$Fe3^{+}$ + OH^{-} + $HO^{\cdot}$) (Touati 2000). La carence ou au contraire la surcharge en fer chez un individu conduisent à des pathologies. La carence en fer chez l'Homme peut provoquer des anémies qui touchent 1,62 milliards de personnes, selon une enquête de prévalence de l'OMS entre 1993-2005 (VMNIS 2005). Les microorganismes et les organismes supérieurs ont donc mis au point des moyens d'acquisition et de gestion du fer bio-disponible. De ces faits, l'homéostasie du fer est strictement régulée chez les êtres vivants, afin d'assurer un équilibre entre leur besoin d'acquérir le fer et le maintien d'une concentration non toxique pour les cellules. Dans les paragraphes qui suivent, nous verrons l'acquisition et l'homéostasie du fer chez l'Homme et chez les bactéries et l'enjeu que cela représente dans la lutte antibactérienne.

II. B| Le fer chez l'homme

Chez l'homme, le fer provenant de l'alimentation est absorbé au niveau intestinal puis transporté à travers les membranes apicales et basales des entérocytes. La Figure I.12 montre l'absorption intestinale du Fer et une partie des différents acteurs impliqués dans le

métabolisme du fer chez l'Homme, notamment des récepteurs membranaires, des protéines de stockage et un régulateur.

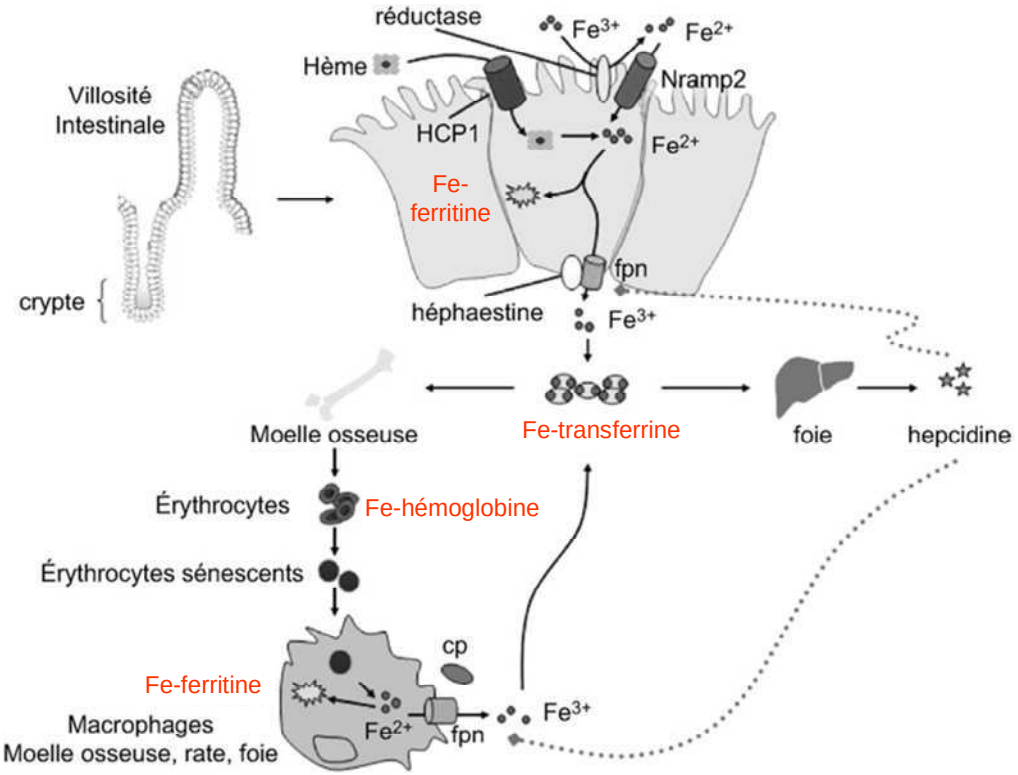


Figure I.12: Le maintien de l'homéostasie du fer chez l'homme. Le fer alimentaire est absorbé au pôle apical des entérocytes duodénaux, stocké dans les ferritines ou exporté au pôle basolatéral. Pris en charge par la transferrine plasmatique, le fer est distribué aux différents organes utilisateurs dont le foie (organe de réserve en fer) et la moelle osseuse qui assure l'érythropoïèse. Arrivés à sénescence, les érythrocytes sont reconnus et détruits par les macrophages tissulaires et le fer héminique est recyclé dans la circulation. L'hepcidine régule négativement l'absorption intestinale et le recyclage macrophagique du fer. Tf : transferrine ; HCP1 : Heme Carrier Protein 1 ; Nramp2 : Natural Resistance Associated Macrophage Protein ; fpn : ferroportine ; cp : céruloplasmine (Delaby *et al.*, 2007)

Le fer est généralement complexé à des protéines telles que la transferrine, la lactoferrine et la ferritine, par une liaison directe à des acides aminés, ou bien indirecte par l'intermédiaire d'un groupement héminique (hémoglobine, myoglobine).

Une fois arrivé dans le plasma sanguin, le fer est pris en charge par la transferrine (Delaby *et al.*, 2007) qui le transporte dans les tissus.

La ferritine, présente chez la plupart des organismes vivants, est la protéine majeure de stockage du fer. En cas d'excès de fer, la ferritine le stocke afin de protéger les cellules de sa toxicité. A l'inverse, elle le libère en cas de carence.

Chez l'Homme, le fer est donc majoritairement intracellulaire, principalement dans l'hémoglobine, qui en contient deux tiers. Ce qui limite sa disponibilité pour les organismes pathogènes (Braun *et al.*, 2011). L'hème contenu dans l'hémoglobine représente donc une source de fer intéressante pour la bactérie. Les microorganismes ont donc developpés plusieurs moyens d'acquisition du fer.

II. C| L'homéostasie du fer chez les bactéries

Le fer est un facteur de croissance pour presque toutes les bactéries à quelques exceptions près. Chez les lactobacilles comme *Lactobacillus plantarum*, le fer est remplacé par le manganèse (Weinberg 1997). Typiquement, une bactérie nécessite une concentration en fer de l'ordre du micromolaire pour une croissance optimale (Butler 1998). Dans la plupart des habitats naturels bactériens, le fer est présent sous forme d'oxydes ferriques peu solubles ou complexé à des petites molécules organiques comme le citrate, le malate ou l'oxalate (Expert 1999). Comme il a été décrit précédemment, chez un hôte tel que l'Homme, le fer est, soit lié à la transferrine et la lactoferrine dans le milieu extracellulaire, soit complexé à l'hème ou à la ferritine à l'intérieur des cellules. Ces différentes protéines prennent en charge la grande majorité du fer présent dans le corps humain, rendant la concentration en fer « libre » extrêmement faible : estimé à 10^{-24} M (Bridges KR 1995). Celles-ci qui constituent un moyen de défense pour diminuer une infection ou la virulence bactérienne. Compte tenu de l'importance du fer, les bactéries et particulièrement les pathogènes ont développé plusieurs mécanismes d'acquisition et de stockage du fer.

Nous porterons une attention particulière sur l'homéostasie du fer chez les bactéries, précisément sur les métallo-régulateurs et les transporteurs, qui sont en rapport direct avec la protéine Fur, objet de cette étude.

Les bactéries peuvent utiliser différentes sources de fer en fonction de leur environnement. Le fer peut être sous forme « libre » ou complexé à des protéines, au degré d'oxydation (II) ou (III).

II. C. 1| Acquisition *via* l'hème et les hémoprotéines

Chez l'homme, plus de 80 % du fer est lié à l'hème ou aux hémoprotéines (hémoglobine, myoglobine, cytochromes, etc…) (Wooldridge *et al.*, 1993).

Chez les bactéries Gram (-) la source de fer doit traverser les deux membranes (externe et cytoplasmique) constituant la paroi cellulaire en passant par le périplasme avant d'atteindre le cytoplasme. Le passage à travers la membrane externe peut se fait par l'intermédiaire des porines quand la taille du complexe ne dépasse pas 600 Da. De nombreuses espèces Gram (-) expriment des protéines spécifiques de l'hème et/ou des hémoprotéines sur la membrane externe pour transporter les complexes de fer (Figure I.13) (Sandy *et al.*, 2009).

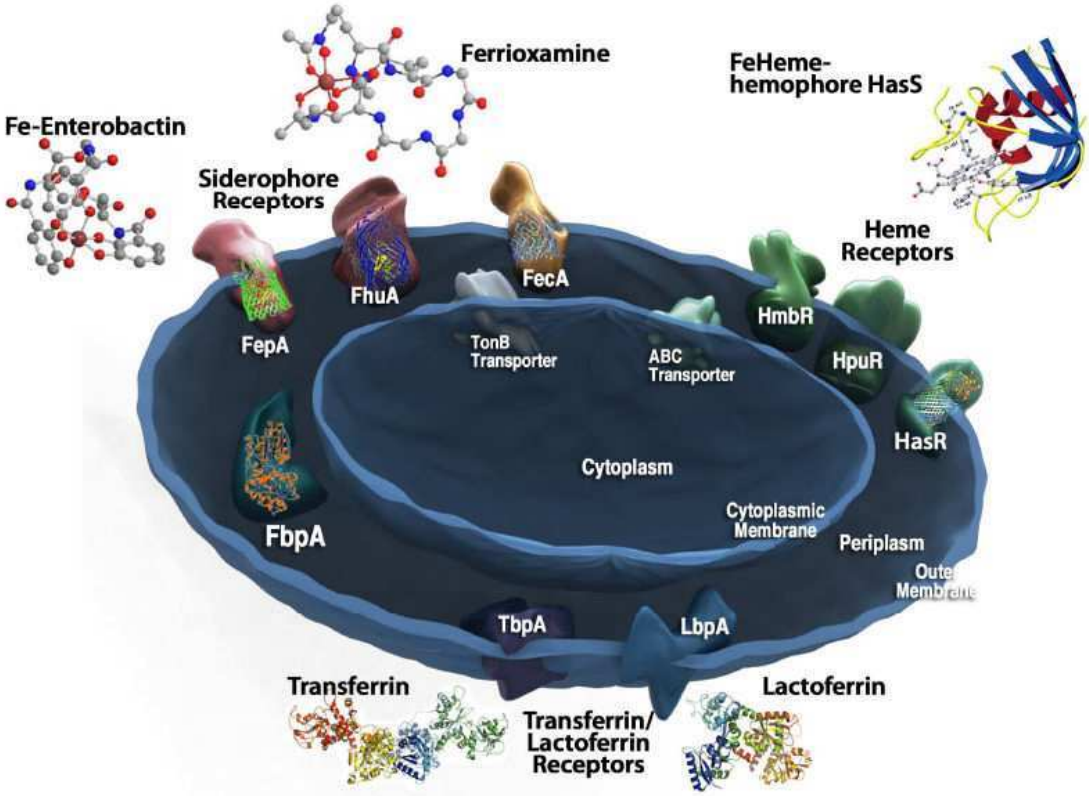


Figure I.13: Voies d'acquisition du fer chez les Gram (-) (Sandy *et al.*, 2009).

Le système de transport de l'hème le plus étudié est celui *Serratia marcescens* (Braun *et al.*, 2011) chez laquelle la structure du complexe de l'hémophore HasA et du récepteur HasR a été résolue (Figure I.14) (Braun *et al.*, 2011).

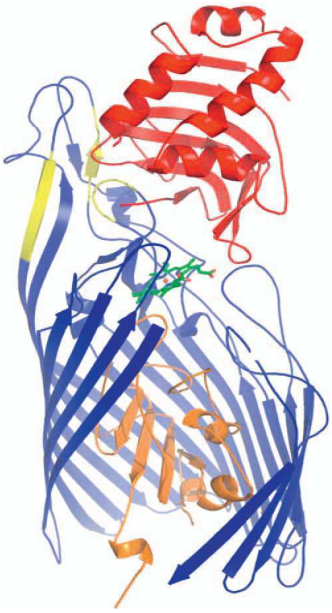


Figure I.14: Structure cristallographique de HasA (rouge) lié à HasR (bleu). Le plug (orange) joue le rôle de bouchon. L'hème est représenté en vert (Braun *et al.*, 2011).

HasA, qui a aussi été identifié aussi chez *Pseudomonas aeruginosa* et *Yersinia pestis* (Wandersman *et al.*, 2004) capte l'hème libre (K_d = 18 pM), puis le complexe HasA-hème est reconnu par le récepteur membranaire HasR. Ce dernier est également capable de reconnaitre l'hème seul (K_d = 0,2 µM). Seul l'hème est internalisé dans le périplasme (Ghigo *et al.*, 1997) grâce à la force proton motrice fournie par les protéines de la membrane interne TonB, ExbB et ExbD (Sandy *et al.*, 2009). Une fois dans le périplasme, l'hème est pris en charge par des protéines périplasmiques qui le transloquent à travers les transporteurs du type ABC (Braun *et al.*, 2011). Il serait ensuite dégradé dans le cytoplasme par oxydation par des oxygénases pour libérer le fer.

Chez les bactéries à Gram (+) le système d'acquisition des hémoprotéines Isd (Iron-regulated surface determinant) de *Staphylococcus aureus* est le mieux caractérisé. Ce système fait intervenir neuf protéines Isd qui agissent à la chaine pour transporter les hémoprotéines/hèmes à travers la paroi peptidoglycane (Maresso *et al.*, 2006) (Braun *et al.*, 2011).

II. C. 2| Acquisition du fer ferreux

Le mécanisme d'acquisition du Fer à l'état d'oxydation (II) chez la bactérie est relativement peu connu. Le Fe^{2+} du fait de sa forte solubilité à pH physiologique (0.1 M), pourrait être transporté facilement à travers la membrane externe par l'intermédiaire des porines, en condition réductrice ou anaérobie. Cependant, le fer ferreux ne prédomine sur le fer ferrique qu'en conditions réductrices, c'est-à-dire en absence d'oxygène ou à pH acide. On suppose qu'avant l'apparition de l'oxygène libre, il y a trois milliards d'années, le Fe^{2+} était la principale source de fer pour les microorganismes. Ces derniers ont donc développé des systèmes permettant son transport à travers la membrane plasmique. De façon non surprenante, de nos jours ce sont les bactéries anaérobies (strictes ou facultatives) ou aérotolérantes qui utilisent préférentiellement ce système de transport du Fe^{2+} (Cartron *et al.*, 2006). Parmi les bactéries aérobies, certaines possèdent à leur surface une réductase qui permet de réduire le Fe^{3+} avant de l'assimiler. De tels systèmes ont été identifiés chez différentes espèces, parmi lesquelles *Escherichia coli* et *Helicobacter pylori*, mais leurs mécanismes restent peu connus (Cowart 2002; Worst *et al.*, 1998).

Le système de transport du Fe^{2+} appelé Feo, identifié pour la première fois chez *Escherichia coli*, aide à maintenir les conditions anaérobies dans l'intestin de l'Homme (Hantke 1987). Ce système est composé de trois protéines majeures : FeoA, FeoB et FeoC, situées au niveau de la membrane cytoplasmique.

FeoA est une protéine cytoplasmique contenant un domaine « SH3-like » (Sandy *et al.*, 2009). Chez *Escherichia coli*, FeoB est une protéine transmembranaire de 773 acides aminés composée d'un domaine C-terminal, qui selon les prédictions, serait constitué de huit hélices α transmembranaires et d'un domaine N-terminal hydrophile contenant une activité GTPase, nécessaire au transport du Fe^{2+} à travers la membrane cytoplasmique vers la cellule. FeoC est une petite protéine fonctionnant comme un récepteur [Fe-S] dépendant (Cartron *et al.*, 2006).

II. C. 3| Acquisition du fer ferrique

Comme il a été décrit auparavant, la grande majorité du fer sur terre se trouve sous forme Fe^{3+} mais est très peu soluble à pH physiologique (10^{-18} M). Ce fer se retrouve généralement sous forme complexé soit à des petites molécules dans les océans ou sur terre ou

soit à des protéines chez des hôtes comme l'Homme. Ces complexes de fer souvent trop larges pour passer à travers les porines de la membrane externe (> 600 Da) sont prises en charge par un ensemble de protéines transmembranaires qui assurent le transport jusqu'au cytoplasme. Ces systèmes comportent un transporteur protéique, situé au niveau de la membrane externe, ayant une très forte affinité (de l'ordre de 10^9 M^{-1}) et une très forte sélectivité pour le complexe donné. L'acquisition du fer se fait alors soit par la transferrine/lactoferrine, soit par les sidérophores.

II. C. 4| Acquisition *via* la transferrine et la lactoferrine

Des récepteurs à la transferrine et à la lactoferrine ont été identifiés chez des espèces pathogènes telles que *Neissera* ou *Hemophilus* (Cornelissen *et al.*, 1994). Chez *Neissera meningitis*, la transferrine est reconnue par un récepteur composé de deux protéines, TbpA et TbpB (Transferrin binding protein) (Figure I.15). De même, la lactoferrine est reconnue par LbpA et LbpB (Lactoferrin binding protein). Dans les deux cas, les récepteurs sont capables de discriminer le transporteur sous sa forme Apo (sans le fer) ou Holo (avec le fer). Les mêmes types de récepteurs ont été identifiés chez *Hemophilus influenzae* (Gray-Owen *et al.*, 1996). Les sous-unités TbpA et LbpA interagissent avec le transporteur et permettent la translocation du Fe^{3+} à travers la membrane plasmique grâce à un mécanisme dépendant de la protéine TonB qui sera décrite plus loin. Les sous-unités TbpB et LbpB semblent jouer un rôle dans la discrimination entre les transporteurs de type Apo et Holo (Cornelissen *et al.*, 1994). Une fois dans le périplasme, le Fe^{3+} est pris en charge par le transporteur Fbp (Ferric binding protein), de type ABC (ATP binding cassette) : la sous-unité FbpA complexe le Fe^{3+} et le transporte jusqu'au canal FbpB qui va transloquer le fer dans le cytoplasme, grâce à l'énergie fournie par l'hydrolyse de l'ATP au niveau de FbpC.

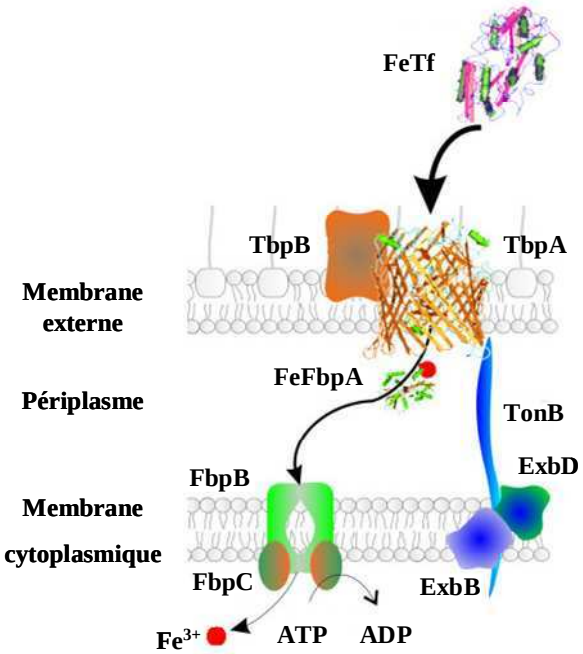


Figure I.15: Acquisition directe du fer lié à la transferrine à travers les systèmes Tbp et Fbp chez *Hemophilus influenzae*. FeTf: complexe Fer-Transferrine; FeFbpA: complexe Fer-FbpA; Tbp: transferrin binding protein; Lbp: lactoferrin binding protein; Fbp: ferric binding protein. L'entrée du fer par ces systèmes est dépendante de la machinerie énergétique TonB/ExbB/ExbD (Modifié d'après (Whitby *et al.*, 2009).

II. C. 5| Acquisition *via* les sidérophores

II. C. 5. a| Description des sidérophores

Les sidérophores sont des petites molécules de masse moléculaire inférieure à 1000 Da sécrétés par certains microorganismes pour lier le fer ferrique avec une très forte affinité (10^{30} M^{-1}) (Sandy *et al.*, 2009). Ces molécules permettent de solubiliser des complexes insolubles du métal comme les hydroxydes ferriques, ou d'entrer en compétition avec des protéines de transport telles que la transferrine (Byers *et al.*, 1998). Plus de 500 sidérophores différents ont été décrits, aussi bien chez les bactéries à Gram (+) qu'à Gram (-) (Sandy *et al.*, 2009). Les sidérophores sont classés en trois groupes en fonction du groupement chimique fonctionnel liant le fer ferrique : les hydroxamates (aerobactine et ferrichrome), les catécholates (entérobactine) et les hydroxyacides (pyochéline). Les plus étudiés, l'entérobactine, le ferrichrome et la pyochéline sont représentés sur la Figure I.16 avec leurs transporteurs respectifs.

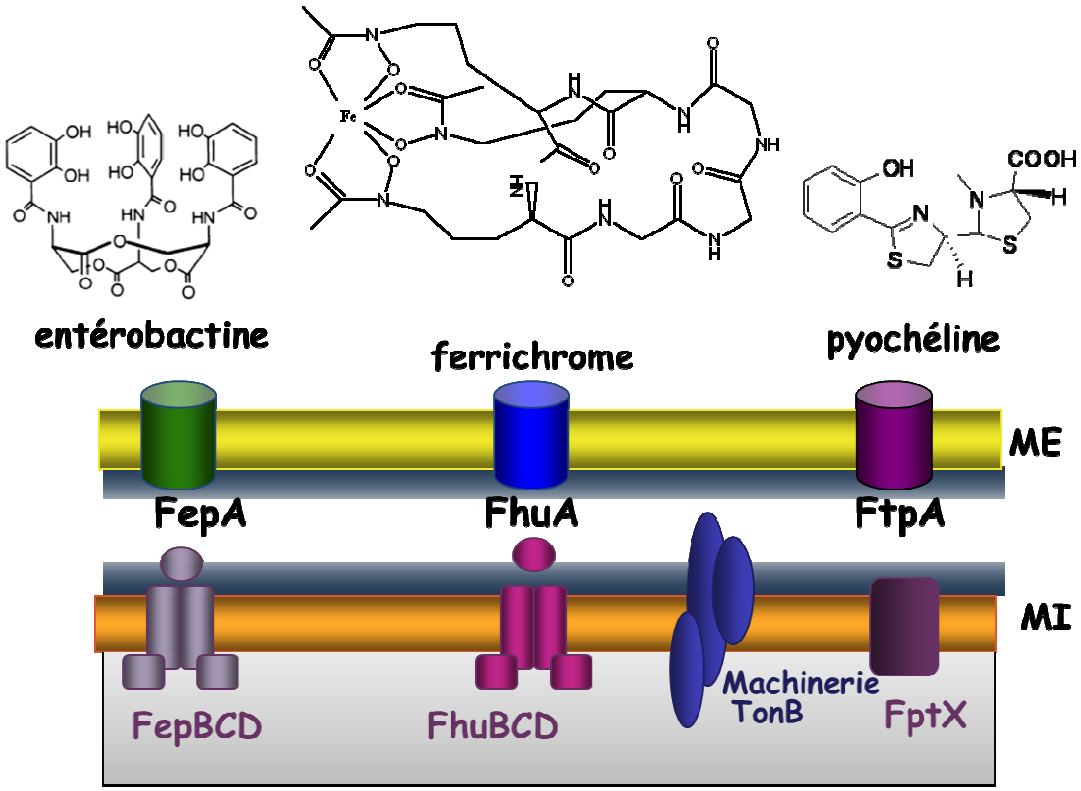


Figure I.16 : Structures de trois sidérophores : l'entérobactine, le ferrichrome et la pyocheline. Ces sidérophores sont reconnus spécifiquement (dans leur forme liée au fer) par des transporteurs de la membrane externe pour atteindre le périplasme de façon énergie dépendante (avec l'aide de la machinerie ExbBexbDTonB). Des perméases ou des transporteurs ABC permettent ensuite le transport à travers la membrane interne de manière à libérer le complexe fer-sidérophore dans le cytoplasme. Les transporteurs de fer spécifiques de l'entérobactine et du ferrichrome sont présents dans *E.coli*. Une bactérie telle que *Pseudomonas aeruginosa* peut aussi transporter le fer de ces deux sidérophores mais synthétise aussi d'autres sidérophores spécifiques tels que la pyochéline.

Certaines bactéries, en plus de leur propre système de synthèse, possèdent des récepteurs aux sidérophores capables de reconnaître des sidérophores synthétisés et sécrétés par d'autres espèces. Les bactéries pathogènes du genre *Neissera*, quant à elles, ne sécrètent pas de sidérophores, mais possèdent des protéines de surface leur permettant de capter et d'utiliser les sidérophores exogènes produits par les microorganismes environnants (Schryvers *et al.*, 1999).

II. C. 5. b| Systèmes de transport des sidérophores

Les bactéries Gram (-) possèdent des récepteurs qui reconnaissent avec une forte affinité et de façon spécifique (de l'ordre de 10^9 M^{-1}) des complexes Fe^{3+}- sidérophore à la

surface de la cellule. Ces complexes sont transportés activement dans la cellule selon un système faisant intervenir le récepteur de la membrane externe, les protéines périplasmiques et les transporteurs de la membrane interne de bactéries (Sandy *et al.,* 2009). Ces dernières années, la description des structures cristallographiques de FepA (ferric enterobactin permease), FhuA (ferric hydroxamate uptake), de FecA (ferric citrate) et de FpvA (ferric pyoverdine) ainsi que des modèles de FvtA (ferric vanchrobactine transport) et FatA (ferric anguibactin transport) montrent que ces transporteurs conservent des similarités structurales (Figure I.17) (Sandy *et al.,* 2009).

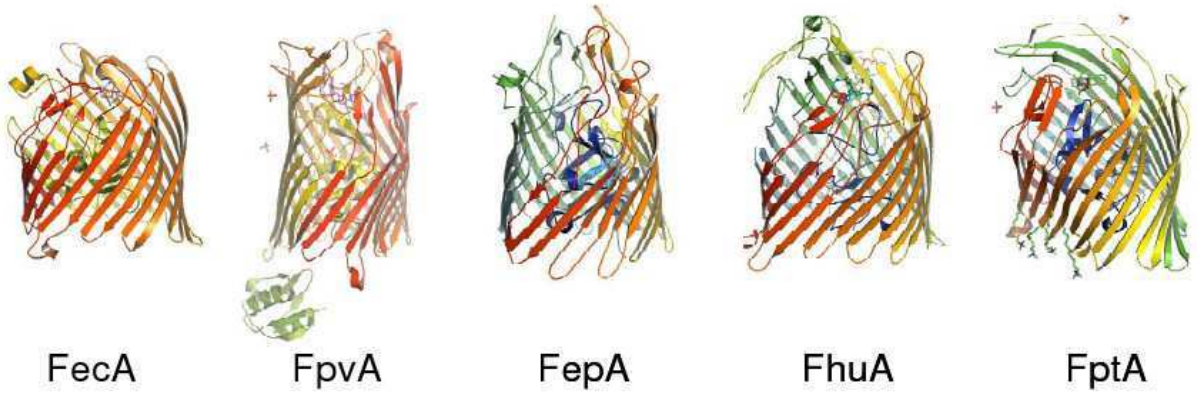


Figure I.17: Structures des transporteurs de la membrane externe. FecA, FepA et FhuA proviennent d.'*E. coli et* FpvA et FptA sont de *P. aeruginosa* (Sandy *et al.,* 2009).

Globalement, ces protéines sont constituées au niveau C-terminal d'un tonneau ß, composé de 22 feuillets ß antiparallèles transmembranaires permettant le passage du complexe de Fe^{3+}- sidérophore jusqu'au périplasme. Le domaine N-terminal est globulaire et forme une sorte de « bouchon » ou « plug » qui bloque l'accès au périplasme en absence de complexe Fe^{3+}- sidérophore. La grande différence entre ces protéines se trouve au niveau des sites de fixation, spécifiques des complexes Fe^{3+}- sidérophore.

La fixation du complexe Fe^{3+}- sidérophore sur le récepteur de la membrane externe entraîne un changement conformationnel au niveau du tonneau ß et du bouchon (Braun *et al.,* 2011; Ferguson *et al.,* 2002). Cependant, cette variation n'est pas suffisante pour laisser passer le complexe Fe^{3+}- sidérophore vers le périplasme. Le passage du complexe nécessite une deuxième transition allostérique énergie-dépendant. L'énergie va être fournie par le complexe protéique TonB-ExbB-ExbD (Braun 1995) (Higgs *et al.,* 1998) (Postle 1993) (Figure I.18).

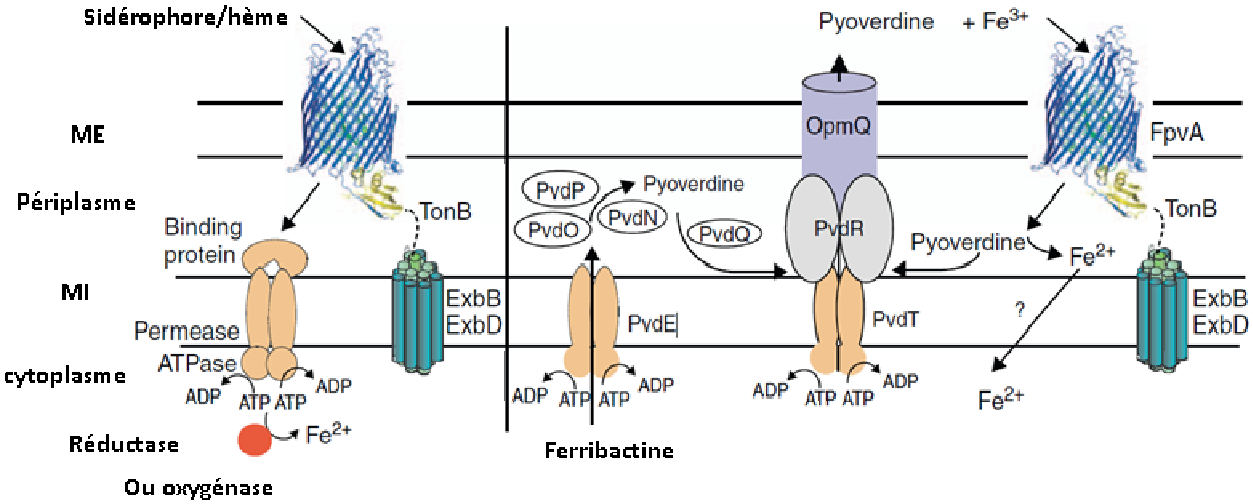


Figure I.18: Transport des sidérophores/hèmes. A gauche ; cas général, à droite, cas particulier de la pyoverdine chez *P. aeruginosa*. Schéma modifié à partie de (Braun *et al.*, 2011).

En effet, le premier changement conformationnel causé par l'arrivée du complexe va rendre accessible l'extrémité N-terminale du récepteur, où est située une séquence conservée de cinq acides aminés appelée « boîte TonB » (TXXV[S/T], où X est un résidu hydrophobe) (Sandy *et al.*, 2009). La protéine TonB va alors lier cette séquence et l'énergie qui lui est fournie par ExbB et ExbD va entraîner un second changement conformationnel, rendant possible le passage du complexe Fe^{3+}- sidérophore vers le périplasme (Braun 1995).

Une fois le complexe dans le périplasme, il est pris en charge par des protéines telles que FecB, FepB et FhuB qui vont le délivrer au niveau d'une perméase cytoplasmique de la même famille, respectivement FecCDE, FepCDEG et FhuCD (Andrews *et al.*, 2003). L'énergie fournie par l'hydrolyse de l'ATP permet le passage du complexe vers le cytoplasme, mais les mécanismes moléculaires impliqués restent mal connus.

II. D| Le fer et la relation hôte-pathogène

Nous venons de montrer l'importance du fer chez l'Homme et les bactéries. Le fer est un facteur de croissance et de survie chez la bactérie. Une concentration importante dans un milieu favorise également la capacité d'une bactérie pathogène à infecter l'hôte. Pour se protéger, l'organisme humain séquestre le fer pour éviter ou atténuer l'infection. De son côté, la bactérie utilise toutes les stratégies en sa procession pour subtiliser ce fer. On assiste donc à une véritable « bataille pour le fer » à laquelle l'hôte et le pathogène se livrent (Fischbach *et al.*, 2006a). Dans la suite, nous allons décrire les grandes lignes de cette bataille.

II. D. 1| La réponse immunitaire de l'hôte

Afin de limiter la biodisponibilité du fer pour prévenir les infections bactériennes, l'organisme humain synthétise et sécrète des protéines qui complexent le fer. Ces protéines qui vont complexer respectivement le fer et l'hémoglobine sont considérées comme des agents bactériostatiques naturels. En cas d'infection, leur production est augmentée dans les sites d'inflammation. La lactoferrine notamment est libérée par les neutrophiles afin de séquestrer le fer dans cette zone (Ellison 1994). L'haptoglobine, une protéine plasmatique, lie l'hémoglobine (Eaton *et al.,* 1982). Lors d'une infection, le complexe haptoglobine-hémoglobine est reconnu par un récepteur présent à la surface des monocytes et des macrophages avant d'être internalisé. Cela va induire une cascade de signalisation à l'origine de l'augmentation de la sécrétion de cytokines anti-inflammatoires (Kristiansen *et al.,* 2001). La lipocaline également appelée sidérocaline ou NGAL (Neutrophil gelatinase-associated lipocalin), une protéine sécrétée durant l'infection par les neutrophiles et les cellules épithéliales, a été décrite aussi comme un composé bactériostatique. En effet, elle est capable d'inhiber la croissance bactérienne en séquestrant des sidérophores chargés en fer, tels que l'entérobactine par exemple (Goetz *et al.,* 2002) ou d'autres types de sidérophores (Smith 2007). L'affinité du complexe fer-entérobactine pour la lipocaline est comparable à l'affinité pour le récepteur membranaire d'*Escherichia coli* FepA (Kd = 41 nM et 50 nM respectivement) (Buchanan *et al.,* 1999) (Goetz *et al.,* 2002). Il a été demontré que des souris déficientes en lipocaline sont beaucoup plus sensibles à l'infection par la souche d'*Escherichia coli* pathogène H9049 que des souris sauvages (Flo *et al.,* 2004).

II. D. 2| Le pathogène et l'infection

Face à une infection bactérienne, une partie de la réponse immunitaire de l'organisme humain consiste à séquestrer le fer et ainsi limiter le développement des pathogènes. Néanmoins, les bactéries ont su mettre en place différentes stratégies pour « contourner » ce problème.

La première réponse des bactéries face à un environnement appauvri en fer réside en la synthèse et la sécrétion de sidérophores, d'hémophores et de protéases comme l'hémolysine, qui va lyser les érythrocytes afin d'en récupérer l'hème complexé au fer.

La synthèse de ces molécules, considérées comme des facteurs de virulence, est sous le contrôle de la protéine Fur (Baumler *et al.,* 1996) (Russo *et al.,* 1999), objet de cette étude.. Chez certaines espèces telles que *Yersinia enterolitica*, les gènes codant pour ces facteurs de virulence sont localisés dans des régions de l'ADN appelées « îlots de pathogénicité » (Carniel *et al.,* 1996). Ces îlots contiennent essentiellement des gènes impliqués dans le métabolisme du fer qui sont régulés eux aussi par la protéine Fur.

Une autre stratégie mise en place par les bactéries a été découverte récemment. En réponse à la lipocaline sécrétée par l'hôte, certaines bactéries synthétisent des entérobactines portant un groupement glucose supplémentaire. Cette modification chimique entraîne un changement structural de l'entérobactine qui n'est plus reconnue par la lipocaline (Fischbach *et al.,* 2006b). Cette modification est sous le contrôle du cluster *iroA*, qui comporte cinq gènes : *iroB*, *iroC*, *iroD*, *iroE* et *iroN*, regroupés en deux opérons : *iroN* d'une part, et *iroEDCB* d'autre part. Ce cluster a été identifié pour la première fois chez *Salmonella* (Fischbach *et al.,* 2006a), mais a ensuite été retrouvé chez une souche d'*Escherichia coli* pathogène CFT073 (Welch *et al.,* 2002). De façon intéressante, il a été montré chez *Salmonella* que les deux opérons étaient sous le contrôle de Fur (Baumler *et al.,* 1996).

II. D. 3| La protéine Fur et la virulence

La pathogénicité d'une bactérie semble étroitement liée à sa capacité à contourner les stratégies mises en place par l'hôte pour limiter la biodisponibilité en fer. La protéine Fur, est un élément clé de l'homéostasie du fer chez les bactéries. L'inactivation du gène *fur* dans différents pathogènes tels que *Staphylococcus aureus* et *Erwinia chrysanthemi* entraine une diminution de la virulence des bactéries (Franza *et al.,* 1999). En effet, Fur régule l'expression de nombreux facteurs de virulence tels que les sidérophores, les hémophores, ainsi que différentes enzymes et toxines. Fur a donc un rôle central dans la pathogénicité des bactéries.

De plus, le rôle de régulation du métabolisme du fer assuré par Fur chez les bactéries à Gram (-) est fondamental. En effet, une délétion du gène *fur* chez *Escherichia coli* entraîne une assimilation constitutive du fer, ce qui provoque un stress oxydant et d'importantes mutations sur l'ADN (Touati *et al.,* 1995).

L'absence de Fur dans les cellules eucaryotes en fait une cible de choix dans la lutte antibactérienne. Dans la partie qui suivra, nous allons présenter cette protéine, son rôle en tant que régulateur global, son fonctionnement ainsi que ses caractéristiques moléculaires.

III| La protéine FUR (Ferric Uptake Regulator)

Après un rappel historique sur la découverte de Fur, son rôle dans la régulation de l'homéostasie du fer *via* l'interaction avec l'ADN, nous développerons les caractéristiques structurales des protéines Fur et Fur-like.

III. A| Historique

Les premières expériences ayant montré l'existence d'une régulation du métabolisme du fer chez les bactéries datent de 1956 (Garibaldi *et al.*, 1956). Garibaldi et ses collaborateurs ont montré que l'expression des systèmes d'acquisition du fer était régulée négativement par le fer lui-même. Par la suite, une souche mutante de *Salmonella typhimurium* produisant certains sidérophores de façon constitutive a été isolée (Ernst *et al.*, 1978). Ernst et ses collaborateurs ont suggéré que la mutation affectait un gène responsable de la régulation des systèmes d'acquisition du fer. Ce gène fut nommé *fur*, pour ferric uptake regulation. Chez *Escherichia coli*, les premiers mutants *fur* ont été identifiés en 1981 (Hantke 1981). Chez ces mutants, les gènes codant pour les récepteurs de sidérophores FecA et FhuA étaient exprimés constitutivement alors qu'ils étaient connus pour être régulés par le fer. Le gène *fur* d'*Escherichia coli* a ensuite été cloné (Hantke 1984) et séquencé (Schaffer *et al.*, 1985). Chez *Escherichia coli*, ce gène *fur* code pour une protéine homodimérique de 148 acides aminés avec une masse moléculaire de 16 795 Da. La séquence en acides aminés de la protéine Fur d'*Escherichia coli* est représentée sur la Figure I.19, ci-dessous.

```
MTDNNTALKKAGLKVTLPRLKILEVLQEPDNHHVSAEDLYKRLIEMGEEI
1         10        20        30        40        50

GLATVYRVLNQFDDAGIVTRHNFEGGKSVFELTQQHHHDHLICLDCGKVI
          60        70        80        90       100

EFSDDSIEARQREIAAKHGIRLTNHSLYLYGHCAEGDCREDEHAHEGK
          110       120       130       140     148
```

Figure I.19: Séquence d'acides aminés de Fur d'*Escherichia coli*.

Fur est une métalloprotéine possèdant un grand nombre d'acides aminés qui pouvant être des ligands pour des métaux. En particulier, elle contient 12 résidus histidines. Mais aussi des tyrosines, des aspartates et des cystéines.

La protéine Fur se comporte en fait comme un senseur du fer. Lorsque la concentration intracellulaire en fer est suffisante, elle s'active et réprime l'expression des protéines impliquées dans les processus comme la biosynthèse des sidérophores, les systèmes de transport du fer (Figure I.20).

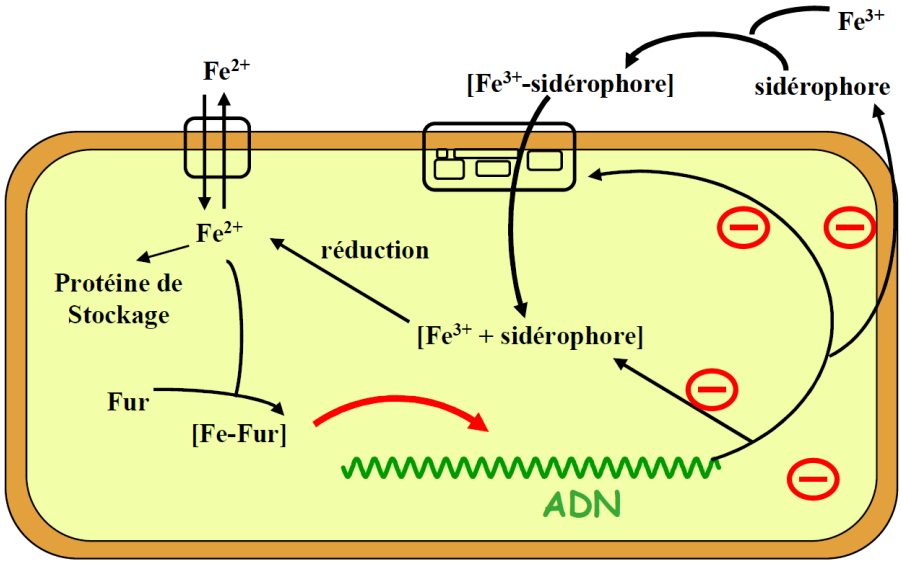


Figure I.20: Représentation schématique du mécanisme de contrôle de l'acquisition du fer par Fur chez la bactérie. En présence d'une concentration en fer suffisante, la protéine Fur est activée et réprime la synthèse de gènes impliqués, entre autres, dans l'assimilation du fer.

Dans ce mécanisme, le Fe^{2+} joue le rôle de cofacteur permettant l'activation de la protéine pour sa liaison à l'ADN. Lorsque la concentration en Fe^{2+} devient importante, celui-ci se fixe sur Fur pour son activation. La protéine activée peut alors se lier sur les séquences d'ADN spécifiques situées dans la région promotrice des gènes qu'elle régule (de Lorenzo *et al.*, 1987).

Le fer n'est pas le seul métal capable d'activer Fur. Des études *in vitro* ont montré que les cations divalents Mn^{2+}, Co^{2+}, Cd^{2+}, Cu^{2+}, Ni^{2+} et Zn^{2+} sont capables d'activer la protéine et de provoquer sa fixation sur la région promotrice du gène codant pour l'aérobactine. Par contre, *in vivo*, outre le Fe^{2+}, seule l'addition de Co^{2+}, Ni^{2+} ou Mn^{2+} dans le milieu de culture semble avoir un effet sur la répression des gènes régulés par Fur, alors qu'aucun effet n'a été observé pour le Zn^{2+} (Bagg *et al.*, 1987).

Plus tardivement, des homologues du gène *fur* ont été identifiés chez un grand nombre de bactéries Gram (-) dont plusieurs pathogènes humains tels que *Yersinia pestis* (Staggs *et al.*, 1991), *Helicobacter pylori* (Bereswill *et al.*, 1998) et *Pseudomonas aeruginosa*, mais

aussi chez certaines bactéries à Gram (+) comme *Staphylococcus aureus* (Kuroda *et al.,* 2001) et *Bacillus subtilis* (Bsat *et al.,* 1998). La plupart de ces homologues sont capables de complémenter une souche d'*Escherichia coli* déficiente en protéine Fur, suggérant un mécanisme d'action semblable de ces protéines dans ces microorganismes.

III. B| Fur, un régulateur global

Plusieurs travaux, dont les notres ont démontré que Fur régule l'expression d'environ 190 de gènes chez *E. coli* (Abed *et al.,* 2007) (Hantke 2001). La plupart de ces gènes sont impliqués dans l'acquisition et le stockage du fer. Les autres gènes codent pour des protéines impliquées dans des processus variés tels que le métabolisme (*cyoA*), la motilité (*flbB*) ou encore la réponse au stress oxydant (*sodB*) (Stojiljkovic *et al.,* 1994). De ce fait, Fur est considéré comme un régulateur global. A l'instar des opérons, l'ensemble des gènes régulés par Fur constitue un « modulon » (Andrews *et al.,* 2003).

L'inactivation du gène *fur* affecte les processus énoncés ci-dessous. Cela entraîne aussi une incapacité de croître sur milieu contenant uniquement des sources de carbone non fermentables comme le succinate, suggérant un défaut dans la respiration (Hantke 1987). De plus, ce même mutant est sensible à un stress oxydant induit par le fer. Ceci est probablement dû à une augmentation de la concentration en fer libre du cytosol (Touati *et al.,* 1995). Enfin, chez *Salmonella typhimurium* et *Helicobacter pylori*, l'inactivation du gène *fur* induit une sensibilité accrue au stress acide, suggérant un rôle de Fur dans la réponse à ce stress (Bijlsma *et al.,* 2002) (Foster *et al.,* 1992).

Dans la majorité des cas, Fur agit en tant que répresseur de l'expression des gènes. Mais depuis un certain temps, plusieurs études ont montré que certains gènes pouvent être régulés par Fur de manière positive, comme c'est le cas chez *Escherichia coli* pour *sodB* qui code pour la superoxyde dismutase à fer (Dubrac *et al.,* 2000, 2002; Fee 1991; Masse *et al.,* 2002).

III. B. 1| Régulation négative

La protéine activée par le métal va se fixer sur les séquences promotrices et bloquer l'accès de l'ARN polymérase. Mais, lorsque l'ARN polymérase est déjà fixée sur le promoteur, Fur, même activée, ne peut la déplacer (Escolar *et al.,* 1997).

Le mécanisme de répression de Fur en l'absence de cofacteur (ou « Apo-Fur ») a été proposé pour la première fois chez *Helicobacter pylori* en 2005 (Ernst *et al.,* 2005). D'après ces travaux, 16 gènes étaient concernés par ce type de régulation, comme le gène *pfr* codant pour la ferritine, impliquée dans le stockage du fer et le gène *sodB* qui code pour la superoxyde dismutase à fer. De façon intéressante, des études sur les protéines Fur de *Campylobacter jejuni* (Holmes *et al.,* 2005) et *Desulfovibrio vulgaris* (Bender *et al.,* 2007) ont également évoqué une répression possible de la transcription par la protéine Apo, mais ces hypothèses restent à confirmer. Un tel mécanisme n'a jamais été décrit chez *Escherchia coli*.

III. B. 2| Autorégulation de Fur

La régulation du gène *fur* par la protéine Fur elle-même est le mécanisme le plus conservé chez les bactéries. La protéine réprime sa propre expression en condition de carence en fer. Cette régulation a été mis en évidence très tôt chez *Escherichia coli* (de Lorenzo *et al.,* 1988a), et également par la suite chez *Helicobacter pylori* (Delany *et al.,* 2002; Delany *et al.,* 2003), *Neisseria gonorrhoeae* (Sebastian *et al.,* 2002) et *Edwardsiella tard* (Wang *et al.,* 2008). Certaines bactéries possèdent d'autres protéines qui régulent l'expression du gène *fur*, comme la protéine CAP (Catabolite activator protein) chez *Escherichia coli* (De Lorenzo *et al.,* 1988b), RpoS chez *Vibrio vulnificus* (Lee *et al.,* 2003) et NikR chez *Helicobacter pylori* (Delany *et al.,* 2005).

Dans la majorité des cas, l'autorégulation de Fur est un mécanisme de répression par la protéine métallée, tel que nous l'avons décrit précédemment. Néanmoins, chez *Helicobacter pylori*, la régulation proposée par Delany *et al*. semble être plus complexe. Elle ferait intervenir une activation de l'expression du gène *fur via* la protéine Apo (Delany *et al.,* 2005). De même, chez *Vibrio vulnificus*, Fur active la transcription de *fur* en l'absence de fer en se liant directement à l'ADN (Lee *et al.,* 2003).

Ces études sur l'autorégulation de Fur montrent qu'elle peut être un activateur de la transcription intervenant de façon directe ou indirecte.

III. B. 3| Régulation positive

Avant 1994, le mécanisme de régulation négative était le plus connu. C'est à partir des analyses protéomiques réalisées en 1994 sur des mutants *fur* de *Vibrio cholerae* qu'un

mécanisme d'activation par Fur a été mis en évidence (Litwin *et al.,* 1994). Les analyses de Litwin *et al.* ont montré que l'expression de nombreuses protéines nécessitait la présence de protéine Fur métallée. Depuis, ce mécanisme d'activation a été identifié chez différents microorganismes : *Helicobacter pylori* (gène *nifS*) (Alamuri *et al.,* 2006), *Neisseria meningitis* (gènes *norB*, *nuoA* et *panI*) (Delany *et al.,* 2004), *Yersinia pestis* (gènes *ftnA* et *napF*) (Gao *et al.,* 2008) et *Pseudomonas aeruginosa* (gène *bfr*) (Wilderman *et al.,* 2004).

Chez *Escherichia coli*, il avait été montré que certains gènes, notamment le gène *sodB* et le gène *ftnA* codant respectivement pour la superoxyde dismutase à fer et la ferritine étaient régulés positivement par Fur. Des études récentes ont montré que cette régulation positive par Fur était indirecte.

Dans le cas de *sodB* ceci est assuré *via* la répression de petits ARN régulateurs anti-sens (ARNs) qui eux mêmes répriment l'expression de *sodB* (Masse *et al.,* 2002). Ces ARNs agissent au niveau post-transcriptionnel en réprimant la traduction des gènes cibles. En effet, leur interaction avec les ARN messagers cibles entraîne une déstabilisation de ces derniers et leur traduction est inhibée.

Dans le cas de la ferritine, Fur agit en se liant en amont du site d'initiation de la transcription sur des sites de fixation de H-NS empêchant ainsi la repression par celle–ci (Nandal *et al.,* 2010).

III. C| La liaison de Fur à l'ADN

III. C. 1| La séquence consensus

En 1987, les expériences d'empreintes à la DNAase I réalisés par De Lorenzo et ses collaborateurs ont mis en évidence les sites de fixation de la protéine Fur sur le promoteur de l'aérobactine d'*Escherichia coli.* (de Lorenzo *et al.,* 1987). La comparaison de ces séquences avec celles de différents promoteurs reconnus par Fur (*fhuA*, *fepA* et *fur*) a identifié une séquence consensus palindromique de 19 paires de bases (Escolar *et al.,* 1998). Cette séquence, riche en adénines et thymines, a été nommée « boîte Fur » (Figure I.21).

```
A    5'- GATAATGAT w ATCATTATC -3'

B     5'- GATAATGATAATCATTATC -3'
      3'- CTATTACTATTAGTAATAG –5'

C     5'- GATAATGATAATCATTATC -3'
      3'- CTATTACTATTAGTAATAG –5'

D    5'- tGATAATGATAATCATTATCa -3'
     3'- aCTATTACTATTAGTAATAGt –5'

E     5'- GATAATGATAATCATTATC -3'
      3'- CTATTACTATTAGTAATAG –5'
```

Figure I.21 : Séquence consensus de la boîte Fur (A) et représentation des différents modèles d'interaction de Fur avec l'ADN (BCDE). Chaque flèche représente un monomère de Fur. A : w correspond à A/T ; B : liaison sur motif palindromique ; C : liaison sur motif hexamérique ; D : Liaison sur motif 7-1-7 ; E : liaison sur motif 6-1-6.

III. C. 2| Les modèles d'interaction Fur-ADN

Différentes études ont proposé des modèles d'interaction de Fur avec l'ADN (Baichoo *et al.,* 2002) (Escolar *et al.,* 1999) (Lavrrar *et al.,* 2002; Lavrrar *et al.,* 2003). La protéine Fur étant active sous forme dimérique, le premier modèle proposé pour son interaction avec l'ADN était que chaque dimère reconnaisse une séquence palindromique, c'est-à-dire deux dimères par « boîte Fur », chacun étant fixé sur un des brins d'ADN (Marmorstein *et al.,* 1992; Pabo *et al.,* 1992) (Figure I.21 B). Cette hypothèse n'était pas compatible avec les résultats expérimentaux d'empreintes à la DNaseI et de microscopie électronique qui montraient que Fur tendait à polymériser sur l'ADN et à en occuper plusieurs faces opposées ou adjacentes (de Lorenzo *et al.,* 1988a). Par la suite, Escolar *et al* ont proposé un modèle où le dimère de Fur reconnait une séquence élémentaire de 6 paires de bases : 5' NAT(A/T)AT 3' (Escolar *et al.,* 1999) (Figure I.21 C).

Toutefois, ce motif parait relativement court et les études biochimiques ont montré qu'un seul motif hexamérique, ou deux motifs adjacents, sont insuffisants pour la liaison de Fur. Baichoo et Helmann ont alors proposé un modèle où Fur reconnait une séquence de 15 paires de bases organisées en un motif 7-1-7 chez *Bacillus subtilis* (Baichoo *et al.,* 2002). Dans ce modèle, deux dimères de Fur peuvent se lier sur chaque boîte Fur, sur des faces opposées de la double hélice d'ADN (Figure I.21 D). Ce modèle est compatible avec l'enroulement de Fur sur l'ADN observé par De Lorenzo *et al.* (de Lorenzo *et al.,* 1988a).

Un modèle similaire a été proposé chez *Escherichia coli* (Lavrrar *et al.*, 2002; Lavrrar *et al.*, 2003). Toutefois, le motif reconnu serait de type 6-1-6 (soit 13 paires de bases) répété deux fois de façon inversée, et constituant le motif de 19 paires de bases lié par EcFur (Figure I.21 E). Ce modèle est compatible avec les données de positionnement d'EcFur sur l'ADN obtenues dans une étude de pontages covalents du complexe EcFur-ADN analysé par spectrométrie de masse (Tiss *et al.*, 2005b). Ces expériences ont montré que la tyrosine 55 (se trouvant dans l'hélice de reconnaissance de l'ADN) était très proche du doublet de thymine de la boîte Fur contenue dans les promoteurs régulés par Fur. Cette proximité doit être importante pour la liaison Fur/ADN de toutes les protéines de la famille Fur puisque la tyrosine et le doublet de thymine sont conservés. Tiss et ses collaborateurs ont proposé un modèle théorique d'interaction, représenté sur la Figure I.22: , ci-dessous où l'on voit un dimère de Fur de *Pseudomonas aeruginosa* dont les domaines d'interaction avec l'ADN se postionnent dans le grand sillon de l'ADN.

Bien évidemment, seule la résolution d'une structure du complexe Fur-ADN permettrait de valider un modèle et de résoudre cette question.

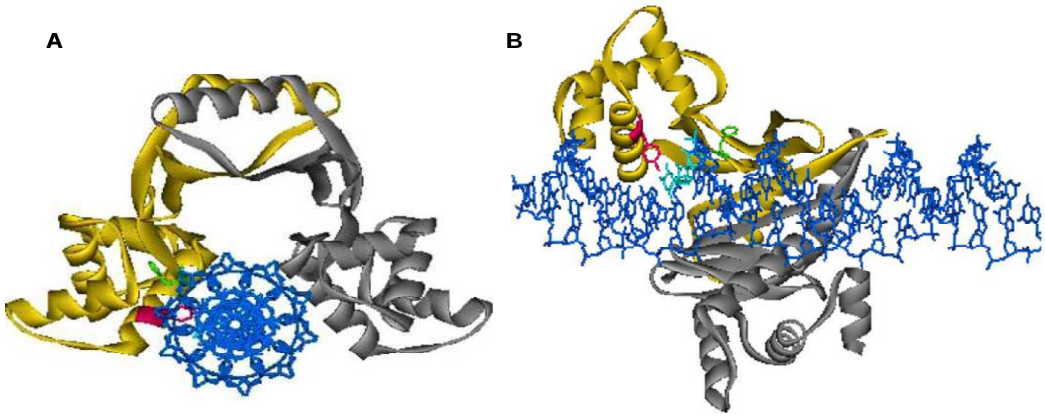


Figure I.22: Modèle du complexe Fur-ADN. A : vue axiale. B : vue longitudinale. La structure tridimensionnelle de Fur de *P.aeruginosa* a été utilisée (Pohl *et al.*, 2003). Le modèle a été réalisé à l'aide du logiciel Insight II (Accelrys). Jaune, gris : un monomère de Fur ; bleu : double hélice d'ADN ; rouge : tyrosine 55 ; vert : lysines 40 et 76 ; cyan : thymines 18 et 19. D'après (Tiss *et al.*, 2005b).

L'affinité de Fur pour ces boîtes Fur ne semble pas varier d'un promoteur à l'autre. Toutefois, le nombre de boîtes Fur présentes sur chaque promoteur serait à l'origine des différences observées dans les niveaux de répression. Plus un promoteur comprend de boîtes Fur, plus le niveau de la régulation par Fur est important (Abed *et al.*, 2007). Une diminution importante de la concentration en fer intracellulaire, cofacteur de la protéine est donc nécessaire pour atténuer la régulation par Fur. A l'inverse, un promoteur comportant

seulement une ou deux boîtes Fur, serait régulé par de petites variations de la concentration en fer dans le cytosol.

III. D| La famille des protéines Fur et « Fur-like »

Leur rôle de régulateur de la transcription suite à une activation par un ion métallique, ainsi que leurs homologies structurales avec Fur d'*Escherichia coli* et de *Pseudomonas aeruginosa* ont valu l'appellation de « Fur-like » à d'autres régulateurs transcriptionnels (Braun *et al.*, 1998). Depuis, Fur désigne non seulement une protéine, mais aussi une famille de métallorégulateurs avec des caractéristiques communes (Lee *et al.*, 2007a). Cette famille comprend des régulateurs impliqués dans l'homéostasie de métaux autres que le fer : zinc, nickel et manganèse (respectivement les protéines Zur, Nur et Mur), mais aussi PerR (senseur du stress oxydant) et Irr (senseur de l'hème). La Figure I.23 montre des alignements de séquences entre plusieurs protéines Fur d'intérêts pour notre étude ainsi que la description des structures secondaires de protéines dont des structures cristallographiques partielles ou totales ont été décrites.

```
HpFur    MKRLETLESILERLRMSIKKNGLKNSKQREEVVSVLYRSG-THLSPEEITHSIRQKD 56
EcFur    ----------MTDNNTALKKAGLKVTLPRLKILEVLQEPDNHHVSAEDLYKRLIDMG 47
VcFur    ----------MSDNNQALKDAGLKVTLPRLKILEVLQQPECQHISAEELYKKLIDLG 47
PaFur    ----------MVENS-ELRKAGLKVTLPRVKILQMLDSAEQRHMSAEDVYKALMEAG 46
MtZur    ----------------MSAAGVRSTRQRAAISTLLETLD-DFRSAQELHDELRRRG 39
ScNur    ----------MVSTDWKSDLRQRGYRLTPQRQLVLEAVDTLEHATPDDILGEVRKTA 47
BsPerR   -----MAAHELKEALETLKETGVRITPQRHAILEYL-VNSMAHPTADDIYKALEGKF 51

HpFur    KNTSISSVYRILNFLEKENFICVLETSKSGRRYEIAAKEHHDHIICLHCGKIIEFAD 113
EcFur    EEIGLATVYRVLNQFDDAGIVTRHNFEGGKSVFELTQQHHHDHLICLDCGKVIEFSD 104
VcFur    EEIGLATVYRVLNQFDDAGIVTRHHFEGGKSVFELSTQHHHDHLVCLDCGEVIEFSD 104
PaFur    EDVGLATVYRVLTQFEAAGLVVRHNFDGGHAVFELADSGHHDHMVCVDTGEVIEFMD 103
MtZur    ENIGLTTVYRTLQSMASSGLVDTLHTDTGESVYRRCSEHHHHHLVCRSCGSTIEVGD 96
ScNur    SGINISTVYRTLELLEELGLVSHAHLGHGAPTYHLADRHHHIHLVCRDCTNVIEADL 104
BsPerR   PNMSVATVYNNLRVFRESGLVKELTYGDASSRFDFVTS-DHYHAICENCGKIVDFHY 107

HpFur    PEIENRQNEVVKKYQAKLISHDMKMFVWCKECQESEC---------- 150
EcFur    DSIEARQREIAAKHGIRLTNHSLYLYGHC-AEGDCREDEHAHEGK-- 148
VcFur    DVIEQRQKEIAAKYNVQLTNHSLYLYGKCGSDGSCKDNPNAHKPKK- 150
PaFur    AEIEKRQKEIVRERGFELVDHNLVLY--------VRKKK-------- 134
MtZur    HEVEAWAAEVATKHGFSDVSHTIEIFGTCSDCRS------------- 130
ScNur    SVAADFTAKLREQFGFDTDMKHFAIFGRCESCSLKGSTTDS------ 145
BsPerR   PGLDEVEQLAAHVTGFKVSHHRLEIYGVCQECSKKENH--------- 145
```

Figure I.23: Alignement de séquences des protéines HpFur, EcFur, PaFur, VcFur, MtZur, BsPerR et ScNur. Les ligands des sites métalliques sont indiqués : en rouge : ligands du site structural ; en bleu : ligands du site de régulation ; en bleu souligné : ligands du site Zn2 d'HpFur ; en vert : ligands d'un site « annexe »; en vert souligné : ligands du site Zn3 d'HpFur, observé dans la structure. Les structures secondaires sont représentées : en vert les hélices α, en jaune les brins ß. Structures secondaires déterminées par cristallographie aux rayons X, sauf EcFur [83-148], déterminée par RMN.

Nous parlerons plus particulièrement des protéines HpFur, VcFur et PaFur ainsi que de BsPerR pour lesquelles des structures ont été résolues et qui seront l'objet de notre étude dans la partie théorique de ce travail.

- La protéine Fur d'*Helicobacter pylori* (HpFur)

Le gène *fur* d'*Helicobacter pylori* a été identifié et caractérisé en 1998 (Bereswill *et al.*, 1998). La protéine Fur d'*H. pylori* peut complémenter partiellement une souche d'*E. coli* dépourvue du gène *fur*. HpFur contrôle l'expression de gènes impliqués dans différentes fonctions, telles que le métabolisme du fer (gènes *pfr, fecA1, fecA2, nifS, exbB, fur*) (Ernst *et al.*, 2005), le stress oxydant (gènes *sodB, cat, nap*) (Ernst *et al.*, 2005), l'adaptation au stress acide (gènes *amiE, amiF* et *arsR*) ainsi que le métabolisme et l'homéostasie du nickel (gènes *nikR, hydABC, hpn-like*). Cette protéine a été caractérisée au laboratoire et sa structure a été publiée en 2009 (Dian *et al.*, 2011; Vitale *et al.*, 2009). Elle comporte une extremité N-ter unique et contient 3 types de sites métalliques dont l'importance relative a été étudiée par mutagénèse dirigée comme cela sera présenté plus bas.

- La protéine Fur de V*ibrio cholerae* (VcFur)

Vibrio cholerae est l'agent responsable du choléra qui a absolument besoin de systèmes d'acquisition du fer performants pour coloniser son hôte. Ces systèmes sont sous le contrôle de la protéine VcFur qui régule l'expression de plus d'une centaine de gènes chez *vibrio cholerae* (Sheikh *et al.*, 2009). Il a été montré que VcFur peut réguler de façon négative et positive les gènes qu'elle contrôle (Mey *et al.*, 2005). Le rôle de la protéine dans la pathogénicité a été mis en évidence et sa structure a été résolue (Sheikh *et al.*, 2009).

- La protéine Fur de *Pseudomonas aeruginosa* (PaFur)

Pseudomonas aeruginosa est une bactérie couramment responsable d'infections nosocomiales et pulmonaires. Elle est aussi connue pour sa capacité à former des biofilms et à résister à un antibiotique (voir acronyme ESKAPE, plus haut). La structure de PaFur publiée en 2003 a été la première structure d'un membre de la famille Fur (Pohl *et al.*, 2003).

➢ La protéine PerR de *Bacillus subtilis*

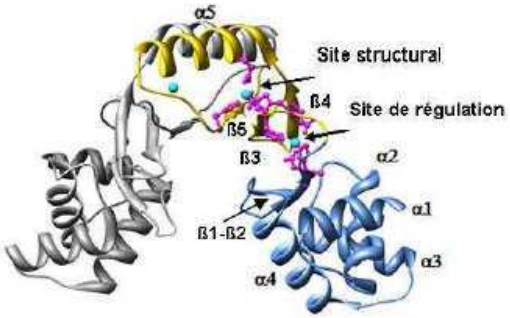
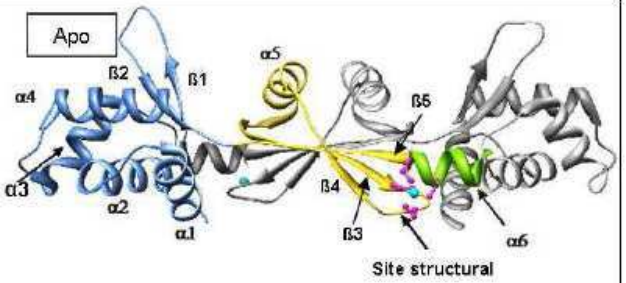
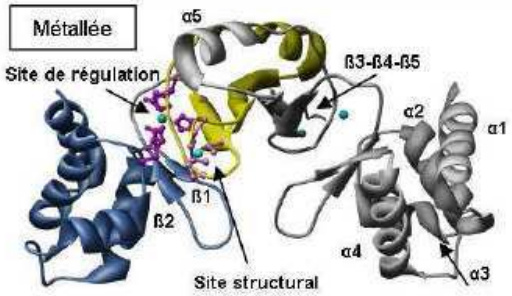
PerR (Peroxyde regulon repressor) a été identifiée comme le principal régulateur inductible par un stress au peroxyde d'hydrogène chez *Bacillus subtilis* (Bsat *et al.,* 1998). Depuis, PerR a été identifiée chez de nombreuses bactéries à Gram (+) telles que *Staphylococcus aureus* (Horsburgh *et al.,* 2001) ou *Listeria monocytogenes* (Rea *et al.,* 2004). Lorsque la protéine est oxydée par l'H_2O_2, on observe une levée de la répression des gènes régulés par elle. PerR régule principalement des gènes impliqués dans la réponse au stress oxydant, notamment le gène *katA*, codant pour la catalase, mais aussi des gènes impliqués dans le stockage du fer comme *ftn*, codant pour la ferritine, mais aussi le gène *fur* (Bsat *et al.,* 1998). Par ailleurs, des études ont montré que PerR était impliquée dans la virulence de différentes espèces, telles que *Staphylococcus aureus*, *Streptococcus pyogenes*, *Listeria monocytogenes* (Rea *et al.,* 2004) et *Enterococcus faecalis* (Verneuil *et al.,* 2005).

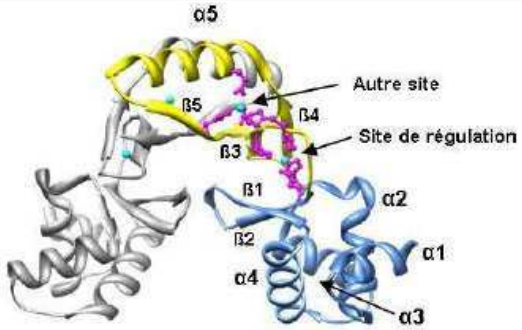
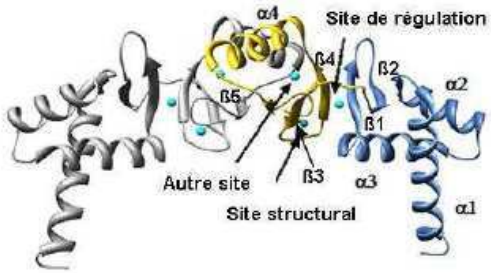
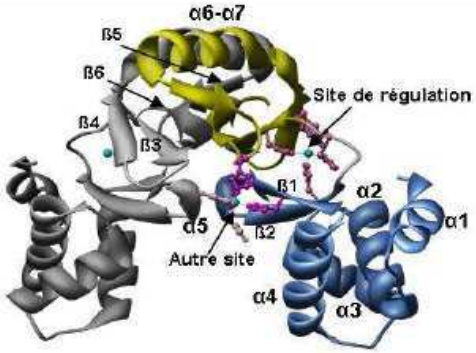
Pour l'ensemble de ces métallorégulateurs, la spécificité pour un métal donné provient de la structure et de la coordination des différents sites métalliques. La résolution récente de leurs structures tridimensionnelles de Fur et « Fur-like » nous a permis de proposer des modèles de la protéine Fur de *Escherichia coli* (EcFur) dont la structure entière n'est pas connue. Voici la description des propriétés structurales de ces protéines.

III. E| Caractéristiques structurales des protéines Fur et « Fur-like »

A ce jour, il existe moins d'une dizaine de structures cristallographies complètes des protéines de la famille de Fur. La première structure décrite en 2003 fût celle de *Pseudomonas aeruginosa* (PaFur) dans sa conformation active (Pohl *et al.,* 2003). Ensuite, la structure de PerR de *Bacillus subtilis* (BsPerR) en conformation proposée inactive a été résolue (Traore *et al.,* 2006). Au laboratoire, nous disposions de la structure cristallographique de domaine N-terminal et des données RMN de Fur d'*Escherichia coli* (EcFur) (Pecqueur *et al.,* 2006).
Ces dernières années de nouvelles structures de Fur et Fur-like ont été décrites : Zur de *Mycobacterium tuberculosis* (MtZur) (Lucarelli *et al.,* 2007), Fur de *Vibrio cholerae* (VcFur) (Sheikh *et al.,* 2009), PerR de *Bacillus subtilis* en conformation active (Jacquamet *et al.,* 2009), Nur de *Streptomyces coelicolor* (ScNur) (An *et al.,* 2009) et enfin Fur d'*Helicobacter pylori* (HpFur) par notre équipe en collaboration avec l'ESRF (Dian *et al.,* 2011). Toutes ces

protéines sont métallées et en conformation active, excepté peut-être MtZur. Les caractéristiques structurales des protéines dont les séquences ont été présentées ci-dessus (Figure I.23) sont répertoriées dans le Tableau I. 2. Ces différentes structures présentent plusieurs points communs : ces protéines sont des homodimères, chaque sous-unité monomérique (1 sous-unité grisée et l'autre en bleu et jaune) possède un domaine N-terminal de liaison à l'ADN (en bleu) et un domaine C-terminal de dimérisation (en jaune). Ces deux domaines sont séparés par une boucle appelée « région charnière ».

Protéine	Liaison à l'ADN	Dimérisation	Site 1 (structural)	Site 2 (charnière)	Site 3
 Fur de *Pseudomonas aeruginosa* (PaFur)	4 hélices α (α1- α4) 2 brins ß (ß 1- ß2) Distance entre les Tyr 55 : 28 Å	3 brins ß (ß 3- ß5) 1 hélice α (α5)	Site 3 proposé comme site structural 1 seule cystéine dans la protéine, n'est pas ligand d'un métal	Site à zinc His 32, Glu 80, His 89, Glu 100 + His 87 proposée Proposé comme site de régulation *In vivo* : site à fer	Site à zinc His 86, Asp 88 bidentate, Glu 107, His 124, H_2O
  PerR de *Bacillus subtilis* (BsPerR)	4 hélices α (α1- α4) 2 brins ß (ß 1- ß2) Distance entre les Tyr 59 : 35 Å (en conformation active)	3 brins ß (ß 3- ß5) 2 hélices α (α5- α6)	Site à zinc Motif CX_2C-CX_2C : Cys 96, 99, 136 et 139	Site à manganèse His 37, Asp 85, His 91, His 93, Asp 104 Proposé comme site de régulation *In vivo* : site à fer et manganèse	Non identifié Ligands non conservés

 Fur de *Vibrio cholerae* (VcFur)	4 hélices α (α1- α4) 2 brins ß (ß 1- ß2) Distance entre les Tyr 56 : 21 Å	3 brins ß (ß 3- ß5) 1 hélice α (α5)	Absence de métal Présence d'un motif CX_2C-CX_4C : Cys 93, 96, 133 et 138 Pont disulfure entre les cystéines 93 et 133	Site à zinc His 33, Glu 84, His 88, His 90 Proposé comme site de régulation *In vivo* : site à fer	Site à zinc His 87, Asp 89, Glu 108, His 125
 Zur de *Mycobacterium tuberculosis* (MtZur)	3 hélices α (α1- α3) 2 brins ß (ß 1- ß2) Distance entre les Tyr 49 : 61 Å	3 brins ß (ß 3- ß5) 1 hélice α (α4)	Site à zinc Motif CX_2C-CX_2C : Cys 86, 99, 126 et 129	Site à zinc Asp 62, Cys 76, His 81, His 83 Proposé comme site de régulation *In vivo* : site à zinc	Site à zinc His 80, His 82, Glu 101, His 118
 **Nur de de *Streptomyces coelicolor* (ScNur)**	4 hélices α (α1- α4) 2 brins ß (ß 1- ß2) Distance entre les Tyr 56 : 26 Å	4 brins ß (ß 3- ß6) 3 hélices α (α5- α7)	Absence de métal Présence d'un motif CX_2C-CX_2C : Cys 93, 96, 133 et 136	Site à zinc His 33, His 86, His 88, His 90 Proposé comme site de régulation *In vivo* : site à nickel	Site à nickel His 70, His 72, His 126, Malonate, Etylène glycol

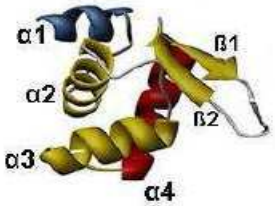
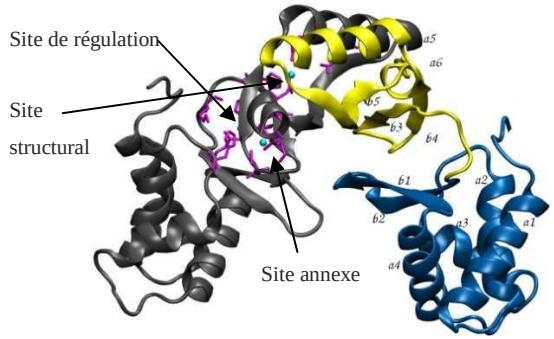
 Fur d'*Escherichia coli* [1-82] (EcFur)	4 hélices α (α1- α4) 2 brins ß (ß 1- ß2)	Non résolu	Site à zinc Présence d'un motif CX_2C-CX_4C : Cys 92, 95, 132 et 137. Ligands identifiés : Cys 92 et 95.	His 32 et His 89 proposées Proposé comme site de régulation *In vivo* : site à fer	Non identifié Ligands conservés
 **Fur d'*Helicobacter pylori* (HpFur)**	4 hélices α (α1- α4) 2 brins β (β1et β2) Distance entre les Tyr 56 : 25 Å	3 brins β (β3-β5) 3 hélices α (α5- α6)	Site à zinc motif CX_2C-CX_2C Cys 102, 105, 142 et 145	Site à zinc His 42, Glu 110, Glu 90, His 99, His 97, Lys 94 Proposé comme site de régulation : *In vivo* : site à fer	Site à zinc Gln 120, Glu 117, His 134, His 96, Asp 98

Tableau I. 2: Caractéristiques structurales des proteins PaFur, BsPerR, VcFur, MtZur, ScNur EcFur et HpFur. Les éléments de structure secondaire ainsi que les sites métalliques sont indiqués sur les structures (colonne 1). L'hélice α4 est l'hélice de reconnaissance de l'ADN, sauf chez MtZur où c'est l'hélice α3. La composition des domaines de liaison à l'ADN et de dimérisation sont indiqués (colonnes 2 et 3). Les ligands des sites métalliques sont indiqués dans les colonnes 4, 5 et 6. Références : PaFur (Ahmad *et al.*, 2009; Lee & Helmann, 2006; Pohl *et al.*, 2003), BsPerR apo (Traore *et al.*, 2006), BsPerR métallée (Jacquamet *et al.*, 2009), MtZur (Lucarelli *et al.*, 2007), VcFur (Sheikh & Taylor, 2009), ScNur (An *et al.*, 2009), EcFur (Gonzalez de Peredo *et al.*, 1999; Pecqueur *et al.*, 2006) et HpFur (Dian *et al.*, 2009).

III. E. 1| Domaine de liaison à l'ADN

Le domaine de liaison à l'ADN (en bleu) des différentes protéines comprend quatre hélices α et deux brins ß formant un feuillet ß antiparallèle (α1 à α4; ß1 et ß2). Le motif de ce domaine permettant la liaison à l'ADN est un motif hélice coude hélice ailé. L'hélice α4 est appelée hélice de reconnaissance et s'insère dans un grand sillon de l'ADN. Pour chaque dimère, la distance entre les deux hélices α4 peut sans doute être reliée à l'affinité de liaison entre la protéine et l'ADN. La mesure de la distance entre les Cα des tyrosines de l'hélice α4 de chaque monomère (conservées dans toutes les protéines et impliquées dans la reconnaissance de l'ADN chez EcFur (Tiss *et al.*, 2005a) indique cette différence d'orientation. Cette distance varie selon les protéines (Tableau I. 2).

BsPerR présente la plus grande distance : 35,3 Å (dans la conformation proposée active). Les autres protéines ont des distances inférieures : 21 Å pour VcFur, 25 Å pour HpFur, 26 Å pour ScFur et 28 Å pour PaFur. Pour MtZur, cette distance est de 61 Å et ne semble pas compatible avec une liaison sur l'ADN. L'hélice α1 est en revanche absente du domaine de liaison à l'ADN de MtZur. Cette dernière possède toutefois une première hélice très longue (notée tout de même α1, équivalente à α2 chez les autres protéines). L'hélice de reconnaissance de l'ADN de MtZur est donc α3. L'alignement des séquences suggère que le domaine de liaison à l'ADN d'HpFur est plus long que celui des autres protéines ce qui est confirmée dans la structure avec la présence d'une extension N-terminale structurée formant des liaisons hydrogènes avec des acides aminés de la région charnière : Arg 3 avec Tyr 37 et Ser 39 ; Leu 4 avec Ala 93 ; Glu 5 avec Tyr 37 (Vitale 2009). Ces liaisons hydrogène permettraient une conformation particulière du dimère, qui se fixerait à l'ADN en absence de métal, comme il a été observée (Delany *et al.*, 2002; Ernst *et al.*, 2005). A l'heure actuelle, HpFur est la seule des protéines « Fur » pour laquelle une fixation de la protéine « Apo » à l'ADN a été observée.

III. E. 2| Domaine de dimérisation

Tandis que les domaines de liaison à l'ADN des protéines Fur sont relativement proches, les domaines de dimérisation sont souvent parfois différents. Les domaines de dimérisation d'HpFur, PaFur, VcFur et MtZur sont constitués de trois brins ß antiparallèles (ß3, ß4 et d'un grand brin ß5) et d'une grande hélice α (α4 pour MtZur et α5 PaFur et VcFur).

Une petite hélice α (α6) supplémentaire est présente chez HpFur et BsPerR. Cette hélice permet le rapprochement dans l'espace de la Cystéine 138 (145) avec les Cystéines 96, 99 et 136 (102, 105 et 142) afin de former le site structural à zinc.

Le grand brin ß5 forme des liaisons hydrogène avec son homologue de l'autre sous-unité, afin de former un feuillet ß inter sous-unité. Ce feuillet ß aide à la structuration en dimère. Chez ScNur, pour laquelle le domaine de dimérisation est plus complexe, ce brin est le brin ß6. En effet, deux hélices α (α6 et α7) et un brin ß (ß6) supplémentaire sont présents chez ScNur. Le fait que les brins ß5 et ß6 ne soient pas rassemblés en un seul brin comme chez les autres protéines, est la conséquence de la présence d'un site à nickel.

III. E. 3| Les sites métalliques des protéines de la famille Fur

Les protéines de la famille Fur possèdent deux ou trois sites métalliques pour celles dont la structure a été résolue. Depuis la résolution de la structure HpFur, un troisième site a été mis en évidence. De plus, la plupart des ligands des trois sites métalliques d'HpFur est retrouvée chez les autres protéines. Même si à l'heure actuelle, parmi les structures résolues, aucune autre protéine Fur ne possède « exactement » les trois mêmes sites. Nous représenterons la protéine Fur avec la structure de HpFur bien caractérisée au laboratoire (Figure I.24: , ci-dessous).

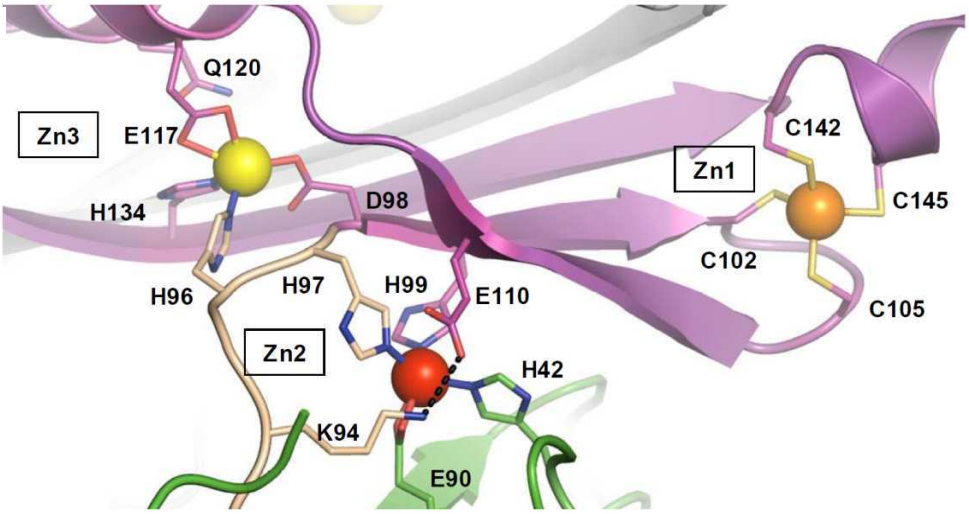


Figure I.24: Les sites métalliques de HpFur (Dian *et al*,. 2009). Zn1 correspond au site structural à zinc, Zn2 au site de régulation et Zn3 au site « annexe ».

III. E. 3. a| Le site structural à zinc « Zn1 »

Ce site pour la première fois identifié dans EcFur (D'Autreaux *et al.,* 2007; Gonzalez de Peredo *et al.,* 1999; Jacquamet *et al.,* 1998) où il a été montré comme étant essentiel à la dimérisation. Il a ensuite été trouvé dans de nombreuses protéines Fur et « Fur-like ». Il est constitué d'un ion Zn(II) dans une géométrie tétraèdrique avec deux à quatre cystéines comme ligands du zinc (voir Figure I.25). Ce site structural à zinc permet un rapprochement dans l'espace entre des brins ß3, ß4 et ß5, qui maintient la structuration d'une partie du domaine de dimérisation. Ce site a été identifié chez les protéines BsPerR, MtZur et HpFur (Lucarelli *et al.,* 2007; Traore *et al.,* 2009) comme étant celui où quatre cystéines sont rassemblées en un motif CX_2C-CX_2C liant un atome de zinc dans une géométrie tétraédrique.

Ces travaux ont démontré que pour ces trois protéines, la présence du zinc est indispensable à la dimérisation. Chez ScNur et VcFur, un tel motif est également présent, mais aucun atome de zinc n'est présent dans la structure cristallographique. VcFur, dispose en conséquence d'un pont disulfure le motif CX_2C et le motif CX_4C. Le zinc ne semble pas indispensable à la dimérisation chez ces protéines, contrairement à ce qui est observé chez HpFur, BsPerR et MtZur. La protéine PaFur quant à elle ne possède qu'une seule cystéine, qui n'est pas impliquée dans la coordination d'un métal.

La protéine EcFur, dont la structure entière n'a pas été résolue et qui est le cas d'étude principal de ce travail, possède aussi 4 cystéines rassemblées en un motif CX_2C-CX_4C. Une étude biochimique menée au laboratoire a montré que les cystéines de motif CX_2C lient un atome de zinc essentiel à la dimérisation (Gonzalez de Peredo *et al.,* 1999).

```
EcFur    ----------CxxC----------CxxxxC---
PaFur    ----------CxxT------------------
VhFur    -----C----CxxC----------CxxxxC---
ValFur   -----C----CxxC----------CxxxxC---
VanFur   -----C----CxxC----------CxxxxC---
EtFur    -----C----CxxC----------CxxxxC---
BsFur    ---C------CxxC----------CxxC-----
AbFur    ----------CxxC-------C--CxxC-----
HpFur    -----C----CxxC----------CxxC--C--
BsPerR   ----------CxxC----------CxxC-----
SrFurS   ----------CxxC----C--C--CxxC-----
MtFurB   --C-------CxxC----------CxxC--C--
EcZur    -C---C----CxxC----C-----CxxC--C-C
```

Figure I.25: Propriétés des proteins Fur et "Fur-like" pour lesquelles des informations structurales sur les sites à zinc sont connues. Les cystéines surlignées en gris ont été montré comme liées au zinc par cristallographie ou absortion des rayons X ou par spectrométrie de masse après modification chimique. C représente les Cystéines dans la séquence de protéine shématisée par ---. En gras sont représentées les cystéines essentielles à la liaison du métal ou de l'ADN. Les cystéines démontrées comme non essentielles par mutagenèse sont souslignées. EcFur, *E. coli* Fur; PaFur, *P. aeruginosa* Fur; VhFur, *V. harveyi* Fur; ValFur, *Vibrio alginolyticus* Fur; VanFur, *Vibrio anguillarium* Fur; EtFur, *E. tarda* Fur; BsFur, *B. subtilis* Fur; AbFur, *Anabaena PCC7119* Fur; HpFur, *H. pylori* Fur; BsPerR, *B. subtilis* PerR; SrFurS, *Streptomyces reticuli* FurS; MtFurB = Zur, *M. tuberculosis* Zur; EcZur, *E. coli* Zur.

III. E. 3. b| Le site de régulation

Ce site possède des ligands à la fois dans le domaine de liaison à l'ADN et dans le domaine de dimérisation. Il occupe une position stratégique, qui permet une orientation du domaine de liaison à l'ADN par rapport au domaine de dimérisation provoquant la liaison de la protéine à l'ADN. En effet, la présence d'un métal dans ce site entrainerait une diminution de l'angle entre le domaine C-terminal et le domaine N- terminal de chaque sous-unité, et par conséquent le rapprochement les deux domaines de liaison à l'ADN (Jacquamet *et al.*, 2009). Ce mécanisme n'a été prouvé que chez BsPerR grâce aux structures disponibles de la protéine dans les deux conformations (inactive et active). Cependant, puisque les ligands de ce site sont conservés chez les autres protéines, on peut s'attendre à ce que le mécanisme soit le même chez toutes les protéines Fur. L'Histidine 87 de PaFur, qui n'était pas impliquée dans la

coordination du zinc dans la structure cristallographique a été proposée par Ahmad *et al.* comme ligand supplémentaire du métal dans le site de régulation (Ahmad *et al.,* 2009).

Ce site de régulation n'est pas spécifique pour un métal donné *in vitro*: il est occupé par du zinc dans les structures d'HpFur, PaFur, VcFur et MtZur, par du manganèse dans la structure de BsPerR et par du nickel dans la structure de ScNur. Différentes travaux, dont les nôtres, ont montré que d'autres métaux pouvaient occuper ce site *in vitro* chez EcFur, BsPerR, VcFur et HpFur, et entrainer une activation de Fur (D'Autreaux *et al.,* 2007; Jacquamet *et al.,* 2009; Sheikh *et al.,* 2009). En fonction du métal et du nombre de ligands, ce site peut adopter une géométrie variable : plutôt tétraédrique en présence de zinc et octaédrique en présence de Fe, ou du Ni par exemple.

III. E. 3. c| Le site « annexe »

Comme nous l'avons évoqué précédemment, un troisième site a été identifié dans la structure d'HpFur (Dian *et al,*. 2009). Ce site est présent également chez MtZur et VcFur. De façon intéressante, les ligands de ce site sont conservés chez EcFur. Un tel site pourrait être présent chez cette protéine, mais non identifié jusqu'ici par les analyses spectroscopiques réalisées sur cette protéine. Ce site pourrait avoir un rôle structural ou encore un rôle de régulation secondaire comme proposée pour HpFur (Dian *et al.,* 2011). Une liaison supplémentaire du zinc dans ce site permettrait de « bloquer » la protéine, déjà activée et fixée sur l'ADN et donc favoriser une répression ou une activation plus forte des gènes cibles.
Chez PaFur, ce site, avait été initialement proposé comme le site de régulation (Pohl *et al.,* 2003). Mais depuis, des travaux recents ont montré que ce site pourrait être le site « structural » de PaFur (Lee *et al.,* 2007b) (Ahmad *et al.,* 2009). Chez HpFur il a été montré que contrairement au site de régulation qui est essentiel pour l'activité de HpFur, le site « annexe » n'est pas indispensable mais a un effet sur l'affinité de HpFur pour la séquence testée. Il semble moduler l'affinité pour l'ADN et pourrait servir à réguler plus finement le modulon Fur en fonction de la concentration en ions métalliques en excès. En effet ce site se métalle après le site de régulation dans les expériences d'addition successive d'ions métalliques sur HpFur (Dian *et al.,* 2011).

III. E. 4| Mécanisme d'activation des protéines de la famille Fur

A partir des données biochimiques et structurales sur les protéines Fur que nous venons de voir, nous pouvons proposer un mécanisme d'activation général pour ces protéines Figure I.26. Le métal présent dans le site structural de la protéine native permettrait de stabiliser sa structure quaternaire dimérique. Cependant, cela ne permet d'activer la protéine. La fixation d'un métal (Fe^{2+}, Mn^{2+}, Zn^{2+}, Ni^{2+} etc…) dans le site de régulation est indispensable à l'activation de la protéine par un changement conformationnel (Gonzalez de Peredo *et al.*, 2001), impliquant une exposition plus importante de l'hélice α4 au solvant. Ce changement permettrait ainsi une meilleure exposition des domaines de liaison à l'ADN (Pecqueur *et al.*, 2006).

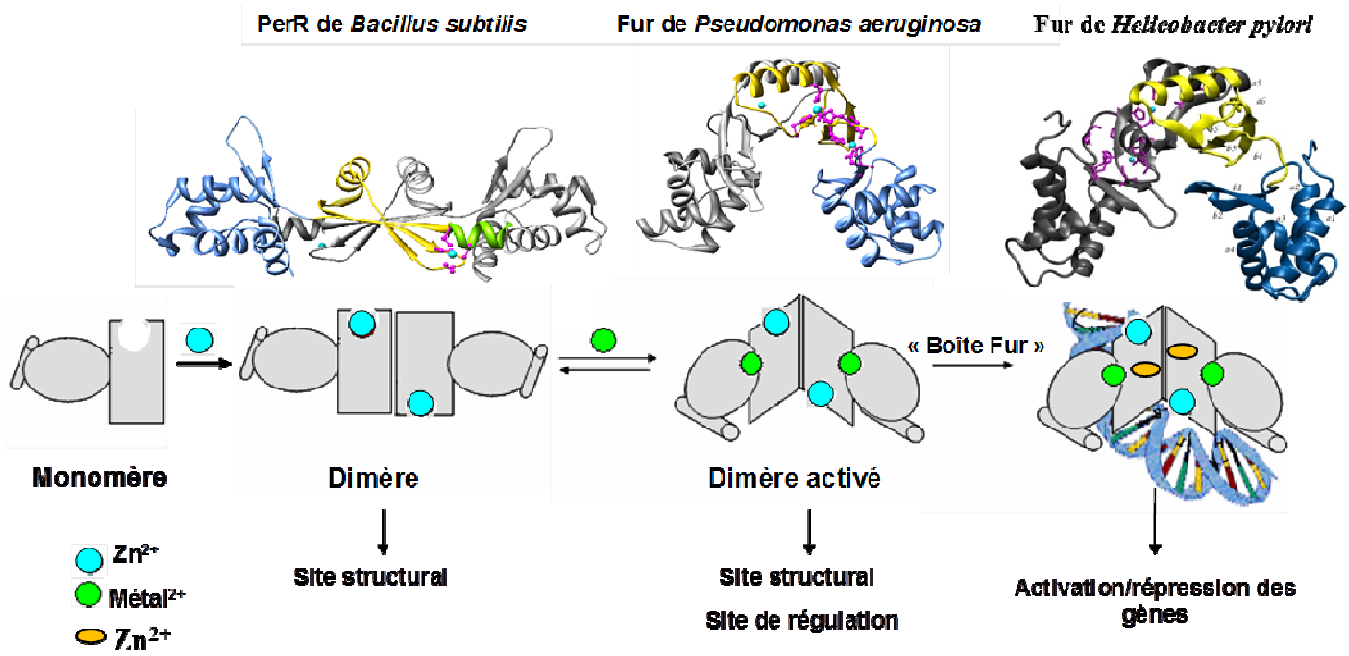


Figure I.26: Mécanisme d'activation proposé pour les protéines de la famille Fur. Trois sites liant des métaux : un site structural, un site de régulation et un site « annexe ». Le métal du site structural permet de maintenir la protéine à l'état de dimère inactif. La métallation du site de régulation entraine un changement conformationnel de la protéine et va permettre son activation et sa liaison à l'ADN. Un troisième site annexe est décrit dans plusieurs protéines. Il permettrait de stabiliser le site de régulation et d'agir légèrement sur la conformation pour moduler l'affinité pour certaines boîtes Fur.

III. F| Conclusion

Dans la première partie de cette introduction, nous nous sommes focalisés sur les antibiotiques et le phénomène de la résistance qui représente un véritable problème de santé publique. Par la suite, nous avons ensuite présenté le métabolisme du fer chez l'Homme et chez le pathogène et montré que cet élément est un facteur limitant de la virulence dont l'acquisition représente un vrai challenge pour la bactérie. Chez l'homme, plusieurs systèmes de régulation existent alors que chez la bactérie, cette régulation est essentiellement contrôlée par la protéine Fur.

Comme nous venons de le voir, la protéine Fur peut donc jouer un rôle important dans la virulence de certains pathogènes. L'inhibition ou l'absence de cette protéine chez la bactérie induit la synthèse constitutive des systèmes d'acquisition du fer et de facteurs de virulence, une sensibilité au stress oxydant et une diminution de la virulence. Si cela semble à première vue avantageux pour le pathogène, une augmentation de la concentration intracellulaire en fer va rapidement s'avérer toxique. Dans ce sens, cette protéine pourrait être une cible intéressante pour la mise au point de nouveaux antibactériens. C'est dans cette optique que la technologie des aptamères peptidiques chez la bactérie, décrite par Blum *et al.* (Blum *et al.*, 2000) a été utilisé pour sélectionner des inhibiteurs de la protéine Fur, en collaboration avec Pierre Colas, un des initiateurs de cette technologie (Abed 2005; Abed *et al.*, 2007). Dans la dernière partie, nous allons rapidement présenter cette technologie, ses avantages et ses applications.

IV| La technologie des aptamères peptidiques

La technologie des aptamères peptidiques consiste à exprimer directement dans la bactérie des molécules combinatoires et de sélectionner celles qui sont capables d'interagir et/ou d'inhiber l'activité des protéines cibles (Colas *et al.*, 1996). Grâce à leurs capacités à lier et à inactiver de manière spécifique une protéine cible d'intérêt, les aptamères fournissent une nouvelle manière très intéressante d'étudier les fonctions des protéines. Les aptamères peptidiques aident à la compréhension de mécanismes moléculaires complexes existant dans une cellule (Colas 2000). Dans les lignes qui suivent, nous décrirons la technologie des aptamères peptidiques pour mieux comprendre les principes, les avantages et les utilisations potentielles.

IV. A| Description des aptamères peptidiques

Etymologiquement, les aptamères peptidiques proviennent du mot latin « aptus » qui signifie « capable de » ou « apte à ». Ces molécules ont été baptisées aptamères peptidiques par analogie aux aptamères à acides nucléiques (Ellington *et al.*, 1990). Inspirés des immunoglobulines, les aptamères peptidiques sont des molécules combinatoires constituées d'une plate-forme protéique constante sur laquelle est greffée une boucle peptidique variable, de manière à être exposée à la surface protéique (Figure I.27).

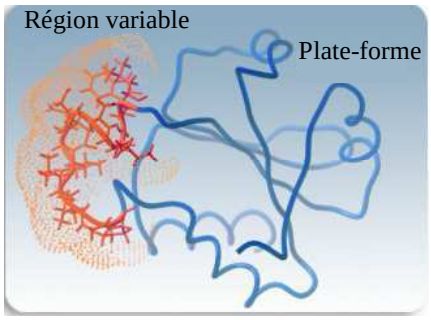


Figure I.27: Représentation schématique d'un aptamère peptidique. La région variable est représentée en orange et la plate-forme en bleu

La boucle peptidique est dite « doublement contrainte » puisqu'elle est attachée à chacune de ses extrémités à la plate-forme protéique.

IV. B| La technologie des aptamères peptidiques

La technologie des aptamères peptidiques consiste à exprimer une banque de molécules peptidiques avec des séquences variables et de sélectionner ces molécules soit pour leur habileté à interagir avec une cible protéique donnée, soit pour leur capacité à conférer un phénotype à une population cellulaire. Les aptamères peptidiques sont généralement sélectionnés pour leur capacité à interagir avec une protéine cible en utilisant la méthode du double hybride chez la levure, comme il a été le cas avec les aptamères peptidiques anti-Fur.

IV. B. 1| Sélection par double hybride

La découverte du système double hybride chez la levure à la fin des années 80 par Fields et Song (Fields *et al.,* 1989). Cette découverte a révolutionné la manière dont les interactions protéine-protéine pouvaient être détectées. Cette méthodologie repose sur des propriétés fondamentales de certains facteurs de transcription. En effet, tous les facteurs de transcription sont constitués d'au moins deux éléments : un domaine activateur et un domaine de liaison à l'ADN. Dans certains cas, ces éléments sont dits « modulables », c'est-à-dire que l'on peut reconstituer un facteur de transcription actif en fusionnant un domaine de liaison à l'ADN et un domaine d'activation provenant de deux facteurs de transcription différents (Brent *et al.,* 1985). D'autre part, ces deux domaines n'ont pas besoin d'être présents au sein d'un même polypeptide mais leur proximité dans l'espace est suffisante pour permettre l'activation de la transcription du gène cible (Ma *et al.,* 1988; Triezenberg *et al.,* 1988). C'est en se basant sur ces propriétés que Fields et Song ont montré que la transcription d'un gène rapporteur chez la levure pouvait être utilisée pour détecter l'interaction entre deux protéines : l'une fusionnée à un domaine activateur et l'autre fusionnée à un domaine de liaison à l'ADN. Ces auteurs ont également imaginé le fait que ce système double hybride permettrait de tester un grand nombre de protéines aléatoires (constituant une banque) pour leur interaction avec une protéine-cible donnée.

Les systèmes double-hybrides classiques utilisent trois vecteurs codant respectivement pour l'expression :

(1) d'un polypeptide constitué de la protéine d'intérêt fusionnée à un domaine de liaison à l'ADN (appelée « appât »), dans notre cas le domaine de liaison à l'ADN de LexA

(2) d'un polypeptide constitué d'une des protéines à tester, fusionnée à un domaine d'activation (appelée « proie »), dans notre cas l'aptamère lié au domaine d'activation B112

(3) d'un gène rapporteur, par exemple le gène *lacZ* d'*Escherichia coli* et le gène codant pour la luciférase.

Comme indiqué sur la Figure I.28: , si l'appât et la proie interagissent, le domaine d'activation et le domaine de liaison à l'ADN sont à proximité : un facteur de transcription complet et actif est alors reconstitué et le gène rapporteur est transcrit. L'expression de ce gène rapporteur permet donc de détecter si les deux protéines d'intérêt interagissent ou pas.

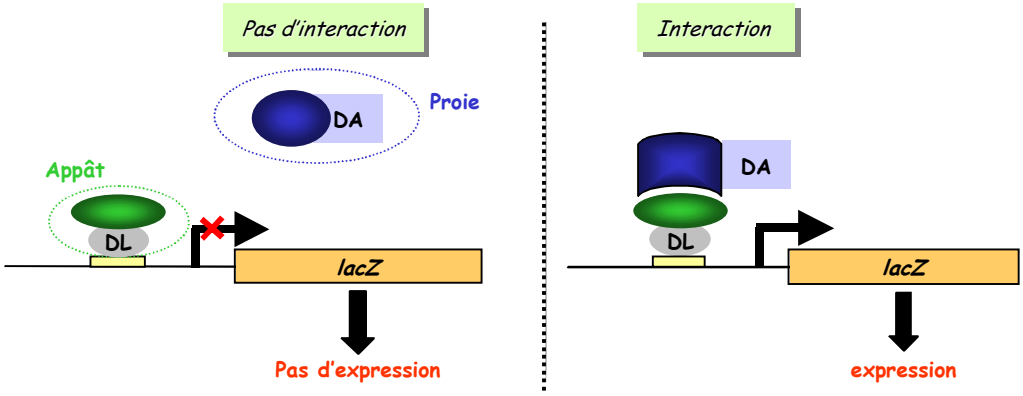


Figure I.28: Représentation schématique de la technique du double hybride. Les proies sont fusionnées à un domaine d'activation de la transcription et les appâts à un domaine de liaison à l'ADN. Une interaction entre la proie et l'appât reconstitue un facteur de transcription fonctionnel et l'expression du gène rapporteur, ici le gène *lacZ*, est détecté. DA : domaine d'activation ; DL : domaine de liaison à l'ADN (Abed 2005).

Lors d'une sélection d'aptamères peptidiques, l'appât est la protéine contre laquelle on souhaite obtenir des aptamères qui sont donc les proies. Concrètement, une banque d'aptamères est transformée dans une souche diploïde contenant le gène rapporteur et exprimant l'appât. Aujourd'hui, le domaine de liaison à l'ADN LexA et les domaines d'activation B42 et B112 sont les plus couramment utilisés (Gyuris *et al.*, 1993).

Colas *et al.* ont été les premiers à réaliser une sélection d'aptamères peptidiques par double hybride chez la levure (Colas *et al.*, 1996). Elle a été effectuée contre la protéine Cdk2 (Cyclin-dependent kinase 2) humaine. Les aptamères sélectionnés étaient spécifiques de Cdk2 avec laquelle ils interagissaient avec des constantes de dissociation à l'équilibre comprises entre 30 et 120 nM (Colas *et al.*, 1996).

IV. B. 2| La plate-forme Thioredoxine A

Les paramètres de choix d'une plate-forme sont très importants. Les principaux étant la taille, la stabilité et la solubilité. De plus, la plate-forme doit pouvoir être surexprimée dans un système procaryote sans être toxique. Elle doit permettre l'insertion de la séquence variable et l'exposer à sa surface.

Plusieurs protéines répondant à ces critères peuvent donc être utilisées comme plate-forme, comme le rapportent différentes travaux (Abedi *et al.,* 1998; McConnell *et al.,* 1995; Norman *et al.,* 1999). On peut ainsi citer la GFP (Green fluorescent protein) ; la STM (Stefin A triple mutant) ; la HA (head activator) ; la LiRP (Ligand regulated protein) ; la FKBP (FK506 binding protein) ; la GST (Glutathione S Transferase) et la Thioredoxine A.

Parce que la Thioredoxine A était la plus utilisée et concerne notre cas d'étude, nous nous focaliserons sur cette protéine ici.

La Thioredoxine A est une petite protéine ubiquitaire cytoplasmique qui compose, avec la Thioredoxine réductase et la NADPH, le « système Thioredoxine », conservée de l'archaebactérie à l'Homme. C'est l'enzyme principale de la réduction des ponts disulfures dans le cytoplasme. Ce qui permet de maintenir un environnement réducteur dans la cellule. Les Thioredoxines sont relativement bien conservées selon les organismes et montrent aussi des structures tridimensionnelles proches. Elles sont toutes constituées d'un noyau central composé de cinq feuillets ß entourés de quatre hélices α. Elles possèdent également toutes un site actif dont la séquence est Cys-X-X-Cys où X est un acide aminé quelconque, autre qu'une cystéine (Figure I.29) (Gleason *et al.,* 1988).

La Thioredoxine A d'*Escherichia coli* a été la première protéine utilisée comme plate-forme d'expression et d'exposition des régions variables. C'est une petite protéine de 109 acides aminés (environ 12 kDa), très rigide et très stable (80 % des résidus participent à sa structure secondaire). Sa structure tridimensionnelle est connue depuis 1975 (Holmgren *et al.,* 1975). Elle possède un site actif exposé, propice à l'insertion de régions variables (LaVallie *et al.,* 1993). Il est important de noter que l'insertion d'un peptide dans ce site fait perdre à la Thioredoxine son activité enzymatique.

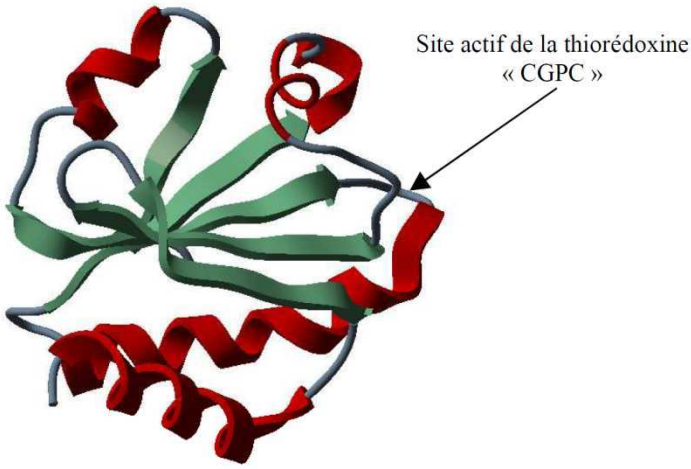


Figure I.29: Structure de la Thioredoxine A d'*Escherichia coli.* Le site actif composé d'un motif « CGPC » est le site d'insertion de la région variable pour la construction des aptamères peptidiques. Résolution : 2.8 Å. Code PDB: 1SRX

IV. B. 3| De l'aptamère peptidique au peptide dérivé

Différents travaux ont décrit l'utilisation des peptides issus des aptamères peptidiques, c'est-à-dire uniquement la partie variable (Armon-Omer *et al.,* 2008; Brien *et al.,* 2011; Buerger *et al.,* 2003; Fabbrizio *et al.,* 1999; Guida *et al.,* 2008; Martel *et al.,* 2006; Nouvion *et al.,* 2007; Schmidt *et al.,* 2002). Dans différents cas, l'interaction avec la cible est conservée, soit avec une affinité comparable à celle de l'aptamère complet (Buerger *et al.,* 2003; Martel *et al.,* 2006; Nouvion *et al.,* 2007; Schmidt *et al.,* 2002), soit avec une perte d'affinité d'un facteur 10 par rapport à l'aptamère complet (Fabbrizio *et al.,* 1999). Guida *et al* ont été les premiers à compléter le travail expérimental par la bioinformatique (Guida *et al.,* 2008). La même approche a été utilisée dans notre étude. En effet, après avoir identifié des aptamères peptidiques dirigés contre le régulateur du cycle cellulaire p53, les auteurs ont montré que les peptides (fusionnés à la biotine) interagissaient également avec la cible. Ils ont ensuite réalisé du « docking moléculaire » *in silico* pour prédire les zones de la protéine impliquées dans l'interaction avec les peptides. Pour cela, ils ont utilisé la structure cristallographique d'un mutant de la p53 (la seule disponible) ainsi qu'une structure « contrainte » minimisée et raffinée des peptides. Par des calculs de dynamique moléculaire, ils ont imposé aux peptides une contrainte similaire à celle exercée par la plate-forme protéique puis ont amarré le peptide sur la protéine. Parmi les différentes positions possibles du peptide sur la protéine, ils ont retenu celle possédant l'énergie libre de liaison la plus faible. Ces simulations ont permis de conclure que les aptamères peptidiques sélectionnés interagissaient probablement avec le

domaine Nter de la protéine (Guida *et al.,* 2008). Nous avons utilisé le même type d'approche mais de façon plus poussée comme nous le verrons au chapitre III.

A ce jour, seule l'étude réalisée au laboratoire par Nadia Abed, en collaboration avec Pierre Colas, l'ENS Lyon et l'entreprise Aptanomics SA, rapporte l'identification d'aptamères peptidiques par sélection double hybride contre une cible bactérienne, la protéine Fur d'*Escherichia coli* (Abed 2005; Abed *et al.,* 2007). Ce travail sera décrit en détail dans le chapitre II. Une étude précédente, publiée en 2000, décrit la sélection d'aptamères peptidiques chez la bactérie. Là, il ne s'agissait pas d'un criblage double hybride chez la levure, mais d'une sélection phénotypique (Blum *et al.,* 2000).

IV. C| Intérêts de la technologie des aptamères peptidiques

IV. C. 1| Avantages de la double contrainte

Différentes études ont montré que la contrainte imposée par la plate-forme protéique contribue à une grande spécificité et à une meilleure affinité des aptamères peptidiques pour leur cible par rapport à la séquence variable seule. Cette différence d'un facteur 100 dans la perte d'efficacité du peptide libre entre les deux études suggère que l'importance de la contrainte imposée par la plate-forme protéique est dépendante de la séquence de la région variable (Cohen *et al.,* 1998). En effet, certaines séquences variables peuvent avoir à elles seules une structure secondaire qui permet leur liaison à la cible avec une forte affinité, alors que d'autres régions variables, peu structurées, nécessitent d'être contraintes par la plate-forme pour interagir avec leur cible. Cet ancrage à une plate-forme présente également d'autres avantages. Tout d'abord, la plateforme favorise l'exposition de la région variable vers l'extérieur, comme sont exposées les régions déterminant la complémentarité chez les immunoglobulines. De plus, cette double contrainte rend possible, dans un contexte intracellulaire, l'exposition d'acides aminés hydrophobes à la surface de la protéine, alors qu'ils seraient le plus souvent enfouis dans le cas d'un peptide non contraint et une région variable ancrée est davantage protégée de la dégradation par les protéases intracellulaires (LaVallie *et al.,* 1993). Enfin, on peut aisément imaginer que la double contrainte devrait faciliter la résolution de la structure tridimensionnelle des régions variables ; surtout lorsque celle de la plate-forme est connue, comme c'est le cas pour la Thioredoxine A d'*Escherichia*

coli (Dyson *et al.,* 1989; Holmgren *et al.,* 1975; Jeng *et al.,* 1994). Cela n'est bien sûr envisageable que si la boucle variable présente une flexibilité limitée une fois ancrée à la plate-forme.

IV. C. 2| En recherche fondamentale

Les aptamères peptidiques ont été principalement utilisés comme outils de recherche pour étudier les fonctions des protéines, pour la compréhension des différents réseaux de régulation cellulaires et pour l'attribution de nouvelles fonctions à des protéines au sein de ces réseaux, notamment grâce à la sélection phénotypique (Abed 2005; Blum *et al.,* 2000; Geyer *et al.,* 1999). Ils représentent également une nouvelle alternative très prometteuse au traditionnel « Knock out » ou aux mutants dominants négatifs dans l'étude de la fonction des gènes (Abed *et al.,* 2007; Geyer *et al.,* 1999). Par exemple, en sélectionnant différents aptamères peptidiques dirigés contre une cible donnée, on disposerait dans la plupart des cas, de molécules agissant avec une « force » modulable comme cela a été le cas avec les aptamères peptidiques dirigés contre la protéine Fur (Abed 2005). L'utilisation *in vivo* de différents aptamères permettrait d'inhiber plus ou moins la fonction de la protéine cible.

IV. C. 3| En recherche appliquée

Les aptamères peptidiques ont trouvé de multiples applications dans la recherche appliquée. Plusieurs travaux dont la nôtre, ont montré que les aptamères peptidiques peuvent être utilisés pour la découverte de nouvelles molécules thérapeutiques, l'identification d'une cible thérapeutique et sa validation (Abed *et al.,* 2007; Bardou *et al.,* 2009; Brien *et al.,* 2011). En premier lieu, on peut imaginer que les aptamères peptidiques puissent être utilisés directement comme « médicaments » pour soigner certaines pathologies. L'utilisation possible de la Thioredoxine humaine comme plate-forme permet de limiter les problèmes d'immunogénicité (Borghouts *et al.,* 2008; de Chassey *et al.,* 2007). Récemment, plusieurs études ont montré que des aptamères peptidiques peuvent être utilisés pour inhiber des protéines impliquées dans le cancer. En effet, Brien *et al*, décrivent un peptide dérivé d'un aptamère peptidique qui interagit et inhibe avec Bf1-1, protéine anti-apoptotique de la famille de Bcl-2 et cible thérapeutique contre des cellules malignes (Brien *et al.,* 2011). D'autres travaux décrivent des aptamères peptidiques qui inhibent des protéines « heat shock proteins »

(Hsp), cibles émergentes dans la lutte contre le cancer (Gibert *et al.,* 2011; Rerole *et al.,* 2011). Les aptamères peptidiques sélectionnés pour leur interaction avec différents régulateurs du cycle cellulaire ou facteurs de transcription sont capables d'induire l'apoptose dans des cellules tumorales qui les surexpriment (Butz *et al.,* 2000; Guida *et al.,* 2008; Martel *et al.,* 2006; Nagel-Wolfrum *et al.,* 2004; Nauenburg *et al.,* 2001) ou de bloquer leur croissance (Buerger *et al.,* 2003; Fabbrizio *et al.,* 1999). D'autres travaux montrent l'utilisation de la technologie sur des protéines virales, cibles intéressantes puisqu'elles ne sont pas présentes dans les cellules (Armon-Omer *et al.,* 2008; Butz *et al.,* 2000; Kaiser *et al.,* 2009; Lopez-Ochoa *et al.,* 2006; Nauenburg *et al.,* 2001; Saccucci *et al.,* 2009).

Ces travaux semblent très encourageants, mais de longs essais cliniques devront être effectués avant de valider l'hypothèse d'une utilisation des aptamères peptidiques comme molécules thérapeutiques. La possibilité d'utiliser uniquement les séquences variables issues des aptamères peptidiques est aussi une opportunité très intéressante. Cela présenterait l'avantage d'utiliser une molécule plus petite, tout en vérifiant que l'interaction avec la cible n'est pas fortement diminuée, voire détruite, en supprimant la plate-forme. La capacité des peptides libres à interagir avec la cible avec des constantes d'affinité comparables à celles des aptamères complets a été montrée dans plusieurs études (Buerger *et al.,* 2003; Martel *et al.,* 2006; Nouvion *et al.,* 2007).

Lorsque les cibles ne sont pas exposées à la surface de la cellule, et cela est notre cas avec la protéine Fur, l'un des problèmes qu'il reste à résoudre est de trouver une stratégie afin de faire entrer l'aptamère peptidique ou le peptide dans les cellules.

IV. D| Conclusion

Les aptamères peptidiques sont capables de lier spécifiquement *in vitro* et *in vivo* une cible donnée et d'inhiber sa fonction, le plus souvent en inhibant les interactions protéiques de cette dernière. Différents travaux ont par ailleurs montré la grande spécificité des aptamères peptidiques pour leur cible. En effet, ils peuvent discriminer des membres très proches d'une même famille protéique présentant pourtant de fortes homologies de séquence (Colas *et al.,* 1996) ou parfois même plusieurs variants alléliques d'une même protéine (Colas *et al.,* 2000; Xu *et al.,* 2002). Par rapport aux immunoglobulines, les aptamères peptidiques sont plus faciles à obtenir, sans l'utilisation d'animaux. Leur manipulation est plus aisée et surtout, ils sont conçus pour fonctionner à l'intérieur de la cellule (Colas *et al.,* 2000).

V| Présentation du travail

Mon travail de thèse a été consacré à l'étude des interactions des aptamères peptidiques anti-Fur et de leur cible en utilisant une double-approche : théorique et expérimentale. La partie théorique consistait à proposer des modèles d'interactions entre la protéine Fur et les aptamères peptidiques. Ces modèles doivent être en accord avec les résultats antérieurs de double hybride et de caractérisations biochimiques des interactions en cours. La partie expérimentale était axée sur la caractérisation biochimique et structurale des interactions protéines/peptides.

L'objectif final du projet est l'obtention d'un nouvel antibactérien à partir d'aptamères peptidiques anti-Fur.

Lorsque j'ai débuté ma thèse, nous possédions quatre aptamères peptidiques F1 à F4 dirigés contre la protéine Fur d'*Escherichia coli*. Des tests de virulence chez la Drosophile avaient montré que la surexpression de ces aptamères chez une souche d'*E.coli* pathogène provoquait une diminution de sa virulence. La sélection de ces aptamères peptidiques ainsi que les différents tests effectués avant mon arrivée au laboratoire sont résumés au début du chapitre II.

Les aptamères peptidiques ont une taille relativement importante (environ 13 kDa) et il est difficile d'imaginer de les utiliser directement en tant qu'antibactériens. J'ai synthétisé et testé les peptides issus des aptamères peptidiques dans un premier temps et caractérisé les interactions entre la protéine et les peptides. Des peptides mutés aussi ont été synthétisés et testés afin de déterminer l'importance de chaque acide aminé dans l'interaction et afin d'en déduire des séquences minimales actives.

Ensuite, j'ai réalisé un criblage virtuel avec de petites molécules de la ZINC Data Base et sélectionné celles qui sont potentiellement inhibitrices de Fur. Ces molécules, si elles sont actives contre Fur, constitueront les inhibiteurs anti-Fur d'un autre génération.

J'ai utilisé une double approche durant cette thèse :

- Une approche biochimique comprenant synthèse chimique peptidique, production et purification de la protéine Fur et des plasmides nécessaires aux tests d'activités et caractérisation des interactions. Le test de protection à la nucléase a été utilisé pour évaluer l'effet inhibiteur des peptides et des mutants peptidiques.

- Une approche théorique, qui a nécessité la mise au point d'un protocole adapté qui a permis de proposer des modèles d'interaction des complexes protéine-peptides cohérents avec

les tests expérimentaux. Un screening virtuel a été effectué dans le but d'identifier de petites molécules capables d'inhiber la protéine Fur sur la base du mode d'interaction des peptides anti-Fur. Enfin, la double approche a été utilisée pour prédire l'amarrage des peptides anti-Fur sur les protéines Fur d'autres microorganismes pathogènes.

Bibliographie

Abed N. (**2005**). "La protéine Fur nouvelle cible antibactérienne? Approches par utilisation des aptamères peptidiques", ***Université Joseph Fourier***.

Abed N., Bickle M., Mari B., Schapira M., Sanjuan-Espana R., Robbe Sermesant K., Moncorge O., Mouradian-Garcia S., Barbry P., Rudkin B. B., Fauvarque M. O., Michaud-Soret I. and Colas P. (**2007**). "A comparative analysis of perturbations caused by a gene knock-out, a dominant negative allele, and a set of peptide aptamers." *Mol Cell Proteomics* **6**(12): 2110-2121.

Abedi M. R., Caponigro G. and Kamb A. (**1998**). "Green fluorescent protein as a scaffold for intracellular presentation of peptides." *Nucleic Acids Res* **26**(2): 623-630.

Ahmad R., Brandsdal B. O., Michaud-Soret I. and Willassen N. P. (**2009**). "Ferric uptake regulator protein: binding free energy calculations and per-residue free energy decomposition." *Proteins* **75**(2): 373-386.

Alamuri P., Mehta N., Burk A. and Maier R. J. (**2006**). "Regulation of the *Helicobacter pylori* Fe-S cluster synthesis protein NifS by iron, oxidative stress conditions, and fur." *J Bacteriol* **188**(14): 5325-5330.

Alekshun M. N. and Levy S. B. (**1999**). "The mar regulon: multiple resistance to antibiotics and other toxic chemicals." *Trends in microbiology* **7**(10): 410-413.

An Y. J., Ahn B. E., Han A. R., Kim H. M., Chung K. M., Shin J. H., Cho Y. B., Roe J. H. and Cha S. S. (**2009**). "Structural basis for the specialization of Nur, a nickel-specific Fur homolog, in metal sensing and DNA recognition." *Nucleic acids research* **37**(10): 3442-3451.

Andres E. (**2011**). "Cationic antimicrobial peptides in clinical development, with special focus on thanatin and heliomicin." *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology*.

Andres E. and Dimarcq J. L. (**2007**). "[Cationic antimicrobial peptides: from innate immunity study to drug development. Up date]." *Médecine et Maladies Infectieuses* **37**(4): 194-199.

Andrews S. C., Robinson A. K. and Rodriguez-Quinones F. (**2003**). "Bacterial iron homeostasis." *FEMS Microbiol Rev* **27**(2-3): 215-237.

Armon-Omer A., Levin A., Hayouka Z., Butz K., Hoppe-Seyler F., Loya S., Hizi A., Friedler A. and Loyter A. (**2008**). "Correlation between shiftide activity and HIV-1 integrase inhibition by a peptide selected from a combinatorial library." *J Mol Biol* **376**(4): 971-982.

Bagg A. and Neilands J. B. (**1987**). "Ferric uptake regulation protein acts as a repressor, employing iron (II) as a cofactor to bind the operator of an iron transport operon in *Escherichia coli*." *Biochemistry* **26**(17): 5471-5477.

Baichoo N. and Helmann J. D. (**2002**). "Recognition of DNA by Fur: a reinterpretation of the Fur box consensus sequence." *J Bacteriol* **184**(21): 5826-5832.

Bardou C., Borie C., Bickle M., Rudkin B. B. and Colas P. (**2009**). "Peptide aptamers for small molecule drug discovery." *Methods in molecular biology* **535**: 373-388.

Baumler A. J., Tsolis R. M., van der Velden A. W., Stojiljkovic I., Anic S. and Heffron F. (**1996**). "Identification of a new iron regulated locus of *Salmonella typhi*." *Gene* **183**(1-2): 207-213.

Bender K. S., Yen H. C., Hemme C. L., Yang Z., He Z., He Q., Zhou J., Huang K. H., Alm E. J., Hazen T. C., Arkin A. P. and Wall J. D. (**2007**). "Analysis of a ferric uptake regulator (Fur) mutant of *Desulfovibrio vulgaris* Hildenborough." *Appl Environ Microbiol* **73**(17): 5389-5400.

Bentley R. and Bennett J. W. (**2003**). "What is an antibiotic? Revisited." *Advances in applied microbiology* **52**: 303-331.

Bereswill S., Lichte F., Vey T., Fassbinder F. and Kist M. (**1998**). "Cloning and characterization of the fur gene from *Helicobacter pylori*." *FEMS Microbiol Lett* **159**(2): 193-200.

Bijlsma J. J., Waidner B., Vliet A. H., Hughes N. J., Hag S., Bereswill S., Kelly D. J., Vandenbroucke-Grauls C. M., Kist M. and Kusters J. G. (**2002**). "The *Helicobacter pylori* homologue of the ferric uptake regulator is involved in acid resistance." *Infect Immun* **70**(2): 606-611.

Blum J. H., Dove S. L., Hochschild A. and Mekalanos J. J. (**2000**). "Isolation of peptide aptamers that inhibit intracellular processes." *Proc Natl Acad Sci U S A* **97**(5): 2241-2246.

Borghouts C., Kunz C., Delis N. and Groner B. (**2008**). "Monomeric recombinant peptide aptamers are required for efficient intracellular uptake and target inhibition." *Mol Cancer Res* **6**(2): 267-281.

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.

Braun V. and Hantke K. (**2011**). "Recent insights into iron import by bacteria." *Current Opinion in Chemical Biology* **15**(2): 328-334.

Brent R. and Ptashne M. (**1985**). "A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor." *Cell* **43**(3 Pt 2): 729-736.

Bridges KR S. P. (1995). Disorders of Iron metabolism. Blood, Principles and Practice of Hematology. L. S. Handlin RI, Stossel TP. London, JB Lippincott Company**:** 1433–1472.

Brien G., Debaud A. L., Bickle M., Trescol-Biemont M. C., Moncorge O., Colas P. and Bonnefoy-Berard N. (**2011**). "Characterization of peptide aptamers targeting Bfl-1 anti-apoptotic protein." *Biochemistry* **50**(23): 5120-5129.

Bsat N., Herbig A., Casillas-Martinez L., Setlow P. and Helmann J. D. (**1998**). "*Bacillus subtilis* contains multiple Fur homologues: identification of the iron uptake (Fur) and peroxide regulon (PerR) repressors." *Mol Microbiol* **29**(1): 189-198.

Buchanan S. K., Smith B. S., Venkatramani L., Xia D., Esser L., Palnitkar M., Chakraborty R., van der Helm D. and Deisenhofer J. (**1999**). "Crystal structure of the outer membrane active transporter FepA from *Escherichia coli*." *Nat Struct Biol* **6**(1): 56-63.

Buerger C., Nagel-Wolfrum K., Kunz C., Wittig I., Butz K., Hoppe-Seyler F. and Groner B. (**2003**). "Sequence-specific peptide aptamers, interacting with the intracellular domain of the epidermal growth factor receptor, interfere with Stat3 activation and inhibit the growth of tumor cells." *J Biol Chem* **278**(39): 37610-37621.

Bullen J. J., Leigh L. C. and Rogers H. J. (**1968**). "The effect of iron compounds on the virulence of *Escherichia coli* for guinea-pigs." *Immunology* **15**(4): 581-588.

Bullen J. J., Rogers H. J. and Lewin J. E. (**1971**). "The bacteriostatic effect of serum on *Pasteurella septica* and its abolition by iron compounds." *Immunology* **20**(3): 391-406.

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): 623-638.

Butler A. (**1998**). "Acquisition and utilization of transition metal ions by marine organisms." *Science* **281**(5374): 207-210.

Butz K., Denk C., Ullmann A., Scheffner M. and Hoppe-Seyler F. (**2000**). "Induction of apoptosis in human *papillomavirus* positive cancer cells by peptide aptamers targeting the viral E6 oncoprotein." *Proc Natl Acad Sci U S A* **97**(12): 6693-6697.

Byers B. R. and Arceneaux J. E. (**1998**). "Microbial iron transport: iron acquisition by pathogenic microorganisms." *Met Ions Biol Syst* **35**: 37-66.

Carniel E., Guilvout I. and Prentice M. (**1996**). "Characterization of a large chromosomal "high-pathogenicity island" in biotype 1B *Yersinia enterocolitica*." *J Bacteriol* **178**(23): 6743-6751.

Carter A. P., Clemons W. M., Brodersen D. E., Morgan-Warren R. J., Wimberly B. T. and Ramakrishnan V. (**2000**). "Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics." *Nature* **407**(6802): 340-348.

Cartron M. L., Maddocks S., Gillingham P., Craven C. J. and Andrews S. C. (**2006**). "Feo--transport of ferrous iron into bacteria." *Biometals* **19**(2): 143-157.

Casey A. L., Adams D., Karpanen T. J., Lambert P. A., Cookson B. D., Nightingale P., Miruszenko L., Shillam R., Christian P. and Elliott T. S. (**2010**). "Role of copper in reducing hospital environment contamination." *The Journal of hospital infection* **74**(1): 72-77.

Cegelski L., Marshall G. R., Eldridge G. R. and Hultgren S. J. (**2008**). "The biology and future prospects of antivirulence therapies." *Nat Rev Microbiol* **6**(1): 17-27.

Chambers H. F. and Deleo F. R. (**2009**). "Waves of resistance: Staphylococcus aureus in the antibiotic era." *Nature reviews. Microbiology* **7**(9): 629-641.

Chart H. and Griffiths E. (**1985**). "Antigenic and molecular homology of the ferric enterobactin receptor protein of *Escherichia coli*." *J Gen Microbiol* **131**(6): 1503-1509.

Choffnes E. R., Relman D. A. and Mack A. (**2010**). "Antibiotic Resistance: Implications for Global Health and Novel Intervention Strategies: Workshop Summary."

Clatworthy A. E., Pierson E. and Hung D. T. (**2007**). "Targeting virulence: a new paradigm for antimicrobial therapy." *Nat Chem Biol* **3**(9): 541-548.

Clements J. M., Coignard F., Johnson I., Chandler S., Palan S., Waller A., Wijkmans J. and Hunter M. G. (**2002**). "Antibacterial activities and characterization of novel inhibitors of LpxC." *Antimicrob Agents Chemother* **46**(6): 1793-1799.

Cohen B. A., Colas P. and Brent R. (**1998**). "An artificial cell-cycle inhibitor isolated from a combinatorial library." *Proc Natl Acad Sci U S A* **95**(24): 14272-14277.

Cohen R. (**2010**). "Résistance aux antibiotiques : état d'urgence." *Archives de Pédiatrie* **17, Supplement 4**(0): S119-S120.

Colas P. (**2000**). "Combinatorial protein reagents to manipulate protein function." *Curr Opin Chem Biol* **4**(1): 54-59.

Colas P., Cohen B., Jessen T., Grishina I., McCoy J. and Brent R. (**1996**). "Genetic selection of peptide aptamers that recognize and inhibit cyclin-dependent kinase 2." *Nature* **380**(6574): 548-550.

Colas P., Cohen B., Ko Ferrigno P., Silver P. A. and Brent R. (**2000**). "Targeted modification and transportation of cellular proteins." *Proc Natl Acad Sci U S A* **97**(25): 13720-13725.

Cornelissen C. N. and Sparling P. F. (**1994**). "Iron piracy: acquisition of transferrin-bound iron by bacterial pathogens." *Mol Microbiol* **14**(5): 843-850.

Cowart R. E. (**2002**). "Reduction of iron by extracellular iron reductases: implications for microbial iron acquisition." *Arch Biochem Biophys* **400**(2): 273-281.

D'Autreaux B., Pecqueur L., Gonzalez de Peredo A., Diederix R. E., Caux-Thang C., Tabet L., Bersch B., Forest E. and Michaud-Soret I. (**2007**). "Reversible redox- and zinc-dependent dimerization of the *Escherichia coli* fur protein." *Biochemistry* **46**(5): 1329-1342.

D'Onofrio A., Crawford J. M., Stewart E. J., Witt K., Gavrish E., Epstein S., Clardy J. and Lewis K. (**2010**). "Siderophores from neighboring organisms promote the growth of uncultured bacteria." *Chemistry & biology* **17**(3): 254-264.

Davies J. (**1996**). "Origins and evolution of antibiotic resistance." *Microbiologia* **12**(1): 9-16.

Davies J. and Davies D. (**2010**). "Origins and evolution of antibiotic resistance." *Microbiology and molecular biology reviews : MMBR* **74**(3): 417-433.

de Chassey B., Mikaelian I., Mathieu A. L., Bickle M., Olivier D., Negre D., Cosset F. L., Rudkin B. B. and Colas P. (**2007**). "An antiproliferative genetic screening identifies a peptide aptamer that targets calcineurin and up-regulates its activity." *Mol Cell Proteomics* **6**(3): 451-459.

de Lorenzo V., Giovannini F., Herrero M. and Neilands J. B. (**1988a**). "Metal ion regulation of gene expression. Fur repressor-operator interaction at the promoter region of the aerobactin system of pColV-K30." *J Mol Biol* **203**(4): 875-884.

De Lorenzo V., Herrero M., Giovannini F. and Neilands J. B. (**1988b**). "Fur (ferric uptake regulation) protein and CAP (catabolite-activator protein) modulate transcription of fur gene in *Escherichia coli*." *Eur J Biochem* **173**(3): 537-546.

de Lorenzo V., Wee S., Herrero M. and Neilands J. B. (**1987**). "Operator sequences of the aerobactin operon of plasmid ColV-K30 binding the ferric uptake regulation (fur) repressor." *J Bacteriol* **169**(6): 2624-2630.

Delaby C., Deybach J. C. and Beaumont C. (**2007**). "[Hepcidin and iron metabolism]." *Rev Med Interne* **28**(7): 510-512.

Delany I., Ieva R., Soragni A., Hilleringmann M., Rappuoli R. and Scarlato V. (**2005**). "In vitro analysis of protein-operator interactions of the NikR and fur metal-responsive regulators of coregulated genes in *Helicobacter pylori*." *J Bacteriol* **187**(22): 7703-7715.

Delany I., Rappuoli R. and Scarlato V. (**2004**). "Fur functions as an activator and as a repressor of putative virulence genes in *Neisseria meningitidis*." *Mol Microbiol* **52**(4): 1081-1090.

Delany I., Spohn G., Pacheco A. B., Ieva R., Alaimo C., Rappuoli R. and Scarlato V. (**2002**). "Autoregulation of *Helicobacter pylori* Fur revealed by functional analysis of the iron-binding site." *Mol Microbiol* **46**(4): 1107-1122.

Delany I., Spohn G., Rappuoli R. and Scarlato V. (**2003**). "An anti-repression Fur operator upstream of the promoter is required for iron-mediated transcriptional autoregulation in *Helicobacter pylori*." Mol Microbiol **50**(4): 1329-1338.

DeLeo F. R. and Chambers H. F. (**2009**). "Reemergence of antibiotic-resistant Staphylococcus aureus in the genomics era." *The Journal of clinical investigation* **119**(9): 2464-2474.

Dian C., Vitale S., Leonard G. A., Bahlawane C., Fauquant C., Leduc D., Muller C., de Reuse H., Michaud-Soret I. and Terradot L. (**2011**). "The structure of the Helicobacter pylori ferric uptake regulator Fur reveals three functional metal binding sites." *Molecular microbiology* **79**(5): 1260-1275.

Dubrac S. and Touati D. (**2000**). "Fur positive regulation of iron superoxide dismutase in *Escherichia coli*: functional analysis of the sodB promoter." *J Bacteriol* **182**(13): 3802-3808.

Dubrac S. and Touati D. (**2002**). "Fur-mediated transcriptional and post-transcriptional regulation of FeSOD expression in *Escherichia coli*." *Microbiology* **148**(Pt 1): 147-156.

Dyson H. J., Holmgren A. and Wright P. E. (**1989**). "Assignment of the proton NMR spectrum of reduced and oxidized thioredoxin: sequence-specific assignments, secondary structure, and global fold." *Biochemistry* **28**(17): 7074-7087.

Eaton J. W., Brandt P., Mahoney J. R. and Lee J. T., Jr. (**1982**). "Haptoglobin: a natural bacteriostat." *Science* **215**(4533): 691-693.

Ebisawa I. (**1987**). "The encounter of Gaston Ramon (1886-1963) with formalin: A biographical study of a great scientist." *Kitasato Archives of Experimental Medicine* **60**(3): 55-70.

ECDC/EMEA (**2009**). "The bacterial challenge: time to react." *European Centre for Disease Prevention and Control*.

Ellington J. E., Farrell P. B., Simkin M. E. and Foote R. H. (**1990**). "Method for obtaining bovine zygotes produced in vivo." *American journal of veterinary research* **51**(11): 1708-1710.

Ellison R. T., 3rd (**1994**). "The effects of lactoferrin on gram-negative bacteria." *Adv Exp Med Biol* **357**: 71-90.

Ernst F. D., Bereswill S., Waidner B., Stoof J., Mader U., Kusters J. G., Kuipers E. J., Kist M., van Vliet A. H. and Homuth G. (**2005**). "Transcriptional profiling of *Helicobacter pylori* Fur- and iron-regulated gene expression." *Microbiology* **151**(Pt 2): 533-546.

Ernst J. F., Bennett R. L. and Rothfield L. I. (**1978**). "Constitutive expression of the iron-enterochelin and ferrichrome uptake systems in a mutant strain of *Salmonella typhimurium*." *J Bacteriol* **135**(3): 928-934.

Escolar L., de Lorenzo V. and Perez-Martin J. (**1997**). "Metalloregulation in vitro of the aerobactin promoter of *Escherichia coli* by the Fur (ferric uptake regulation) protein." *Mol Microbiol* **26**(4): 799-808.

Escolar L., Perez-Martin J. and de Lorenzo V. (**1998**). "Binding of the fur (ferric uptake regulator) repressor of *Escherichia coli* to arrays of the GATAAT sequence." *J Mol Biol* **283**(3): 537-547.

Escolar L., Perez-Martin J. and de Lorenzo V. (**1999**). "Opening the iron box: transcriptional metalloregulation by the Fur protein." *J Bacteriol* **181**(20): 6223-6229.

Expert D. (**1999**). "WITHHOLDING AND EXCHANGING IRON: Interactions Between Erwinia spp. and Their Plant Hosts." *Annu Rev Phytopathol* **37**: 307-334.

Fabbrizio E., Le Cam L., Polanowska J., Kaczorek M., Lamb N., Brent R. and Sardet C. (**1999**). "Inhibition of mammalian cell proliferation by genetically selected peptide aptamers that functionally antagonize E2F activity." *Oncogene* **18**(30): 4357-4363.

Fee J. A. (**1991**). "Regulation of sod genes in *Escherichia coli:* relevance to superoxide dismutase function." *Mol Microbiol* **5**(11): 2599-2610.

Feldberg G. D. (**1995**). "Disease and class: tuberculosis and the shaping of modern North American society." *Rutgers University Press*: 1995-1274.

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): 1715-1719.

Fields S. and Song O. (**1989**). "A novel genetic system to detect protein-protein interactions." *Nature* **340**(6230): 245-246.

Fischbach M. A., Lin H., Liu D. R. and Walsh C. T. (**2006a**). "How pathogenic bacteria evade mammalian sabotage in the battle for iron." Nat Chem Biol **2**(3): 132-138.

Fischbach M. A., Lin H., Zhou L., Yu Y., Abergel R. J., Liu D. R., Raymond K. N., Wanner B. L., Strong R. K., Walsh C. T., Aderem A. and Smith K. D. (**2006b**). "The pathogen-associated iroA gene cluster mediates bacterial evasion of lipocalin 2." Proc Natl Acad Sci U S A **103**(44): 16502-16507.

Fleming A. (**1955**). "The story of penicillin." Bulletin. Georgetown University. Medical Center **8**(4): 128-132.

Flo T. H., Smith K. D., Sato S., Rodriguez D. J., Holmes M. A., Strong R. K., Akira S. and Aderem A. (**2004**). "Lipocalin 2 mediates an innate immune response to bacterial infection by sequestrating iron." Nature **432**(7019): 917-921.

Foster J. W. and Hall H. K. (**1992**). "Effect of *Salmonella typhimurium* ferric uptake regulator (fur) mutations on iron- and pH-regulated protein synthesis." J Bacteriol **174**(13): 4317-4323.

Franza T., Sauvage C. and Expert D. (**1999**). "Iron regulation and pathogenicity in *Erwinia chrysanthemi* 3937: role of the Fur repressor protein." Mol Plant Microbe Interact **12**(2): 119-128.

Gao H., Zhou D., Li Y., Guo Z., Han Y., Song Y., Zhai J., Du Z., Wang X., Lu J. and Yang R. (**2008**). "The iron-responsive Fur regulon in *Yersinia pestis*." J Bacteriol **190**(8): 3063-3075.

Garibaldi J. A. and Neilands J. B. (**1956**). "Formation of iron-binding compounds by micro-organisms." Nature **177**(4507): 526-527.

Garrity-Ryan L. K., Kim O. K., Balada-Llasat J. M., Bartlett V. J., Verma A. K., Fisher M. L., Castillo C., Songsungthong W., Tanaka S. K., Levy S. B., Mecsas J. and Alekshun M. N. (**2010**). "Small molecule inhibitors of LcrF, a Yersinia pseudotuberculosis transcription factor, attenuate virulence and limit infection in a murine pneumonia model." Infection and immunity **78**(11): 4683-4690.

Geyer C. R., Colman-Lerner A. and Brent R. (**1999**). ""Mutagenesis" by peptide aptamers identifies genetic network members and pathway connections." Proc Natl Acad Sci U S A **96**(15): 8567-8572.

Ghigo J. M., Letoffe S. and Wandersman C. (**1997**). "A new type of hemophore-dependent heme acquisition system of *Serratia marcescens* reconstituted in *Escherichia coli*." J Bacteriol **179**(11): 3572-3579.

Gibert B., Hadchity E., Czekalla A., Aloy M. T., Colas P., Rodriguez-Lafrasse C., Arrigo A. P. and Diaz-Latoud C. (**2011**). "Inhibition of heat shock protein 27 (HspB1) tumorigenic functions by peptide aptamers." Oncogene **30**(34): 3672-3681.

Gleason F. K. and Holmgren A. (**1988**). "Thioredoxin and related proteins in procaryotes." FEMS Microbiol Rev **4**(4): 271-297.

Goetz D. H., Holmes M. A., Borregaard N., Bluhm M. E., Raymond K. N. and Strong R. K. (**2002**). "The neutrophil lipocalin NGAL is a bacteriostatic agent that interferes with siderophore-mediated iron acquisition." Mol Cell **10**(5): 1033-1043.

Gonzalez de Peredo A., Saint-Pierre C., Adrait A., Jacquamet L., Latour J. M., Michaud-Soret I. and Forest E. (**1999**). "Identification of the two zinc-bound cysteines in the ferric uptake regulation protein from *Escherichia coli:* chemical modification and mass spectrometry analysis." Biochemistry **38**(26): 8582-8589.

Gonzalez de Peredo A., Saint-Pierre C., Latour J. M., Michaud-Soret I. and Forest E. (**2001**). "Conformational changes of the ferric uptake regulation protein upon metal activation and DNA binding; first evidence of structural homologies with the diphtheria toxin repressor." J Mol Biol **310**(1): 83-91.

Gray-Owen S. D. and Schryvers A. B. (**1996**). "Bacterial transferrin and lactoferrin receptors." Trends Microbiol **4**(5): 185-191.

Grier M. C., Garrity-Ryan L. K., Bartlett V. J., Klausner K. A., Donovan P. J., Dudley C., Alekshun M. N., Tanaka S. K., Draper M. P., Levy S. B. and Kim O. K. (**2010**). "N-Hydroxybenzimidazole inhibitors of ExsA MAR transcription factor in Pseudomonas aeruginosa: In vitro anti-virulence activity and metabolic stability." Bioorganic & medicinal chemistry letters **20**(11): 3380-3383.

Griffiths E. (**1991**). "Iron and bacterial virulence--a brief overview." Biol Met **4**(1): 7-13.

Gristwood T., Fineran P. C., Everson L. and Salmond G. P. (**2008**). "PigZ, a TetR/AcrR family repressor, modulates secondary metabolism via the expression of a putative four-component resistance-nodulation-cell-division efflux pump, ZrpADBC, in Serratia sp. ATCC 39006." Molecular microbiology **69**(2): 418-435.

Guerinot M. L. (**1994**). "Microbial iron transport." Annu Rev Microbiol **48**: 743-772.

Guida E., Bisso A., Fenollar-Ferrer C., Napoli M., Anselmi C., Girardini J. E., Carloni P. and Del Sal G. (**2008**). "Peptide aptamers targeting mutant p53 induce apoptosis in tumor cells." Cancer Res **68**(16): 6550-6558.

Gyuris J., Golemis E., Chertkov H. and Brent R. (**1993**). "Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2." Cell **75**(4): 791-803.

Hancock V., Ferrieres L. and Klemm P. (**2008**). "The ferric yersiniabactin uptake receptor FyuA is required for efficient biofilm formation by urinary tract infectious Escherichia coli in human urine." Microbiology **154**(Pt 1): 167-175.

Hansen J. L., Ippolito J. A., Ban N., Nissen P., Moore P. B. and Steitz T. A. (**2002**). "The structures of four macrolide antibiotics bound to the large ribosomal subunit." Mol Cell **10**(1): 117-128.

Hantke K. (**1981**). "Regulation of ferric iron transport in *Escherichia coli* K12: isolation of a constitutive mutant." Mol Gen Genet **182**(2): 288-292.

Hantke K. (**1984**). "Cloning of the repressor protein gene of iron-regulated systems in *Escherichia coli* K12." Mol Gen Genet **197**(2): 337-341.

Hantke K. (**1987**). "Selection procedure for deregulated iron transport mutants (fur) in *Escherichia coli* K 12: fur not only affects iron metabolism." Mol Gen Genet **210**(1): 135-139.

Hantke K. (**2001**). "Iron and metal regulation in bacteria." Curr Opin Microbiol **4**(2): 172-177.

Harvie D. R. and Ellar D. J. (**2005**). "A ferric dicitrate uptake system is required for the full virulence of Bacillus cereus." Curr Microbiol **50**(5): 246-250.

Hassett D. J., Sokol P. A., Howell M. L., Ma J. F., Schweizer H. T., Ochsner U. and Vasil M. L. (**1996**). "Ferric uptake regulator (Fur) mutants of Pseudomonas aeruginosa demonstrate defective siderophore-mediated iron uptake, altered aerobic growth, and decreased superoxide dismutase and catalase activities." J Bacteriol **178**(14): 3996-4003.

Higgs P. I., Myers P. S. and Postle K. (**1998**). "Interactions in the TonB-dependent energy transduction complex: ExbB and ExbD form homomultimers." J Bacteriol **180**(22): 6031-6038.

Holmes K., Mulholland F., Pearson B. M., Pin C., McNicholl-Kennedy J., Ketley J. M. and Wells J. M. (**2005**). "*Campylobacter jejuni* gene expression in response to iron limitation and the role of Fur." Microbiology **151**(Pt 1): 243-257.

Holmgren A., Soderberg B. O., Eklund H. and Branden C. I. (**1975**). "Three-dimensional structure of *Escherichia coli* thioredoxin-S2 to 2.8 A resolution." Proc Natl Acad Sci U S A **72**(6): 2305-2309.

Horsburgh M. J., Ingham E. and Foster S. J. (**2001**). "In *Staphylococcus aureus*, fur is an interactive regulator with PerR, contributes to virulence, and Is necessary for oxidative stress resistance through positive regulation of catalase and iron homeostasis." J Bacteriol **183**(2): 468-475.

Hung D. T., Shakhnovich E. A., Pierson E. and Mekalanos J. J. (**2005**). "Small-molecule inhibitor of *Vibrio cholerae* virulence and intestinal colonization." Science **310**(5748): 670-674.

Jacquamet L., Aberdam D., Adrait A., Hazemann J. L., Latour J. M. and Michaud-Soret I. (**1998**). "X-ray absorption spectroscopy of a new zinc site in the fur protein from *Escherichia coli*." Biochemistry **37**(8): 2564-2571.

Jacquamet L., Traore D. A., Ferrer J. L., Proux O., Testemale D., Hazemann J. L., Nazarenko E., El Ghazouani A., Caux-Thang C., Duarte V. and Latour J. M. (**2009**). "Structural characterization of the active form of PerR: insights into the metal-induced activation of PerR and Fur proteins for DNA binding." Molecular microbiology **73**(1): 20-31.

Jeng M. F., Campbell A. P., Begley T., Holmgren A., Case D. A., Wright P. E. and Dyson H. J. (**1994**). "High-resolution solution structures of oxidized and reduced *Escherichia coli* thioredoxin." Structure **2**(9): 853-868.

Kaiser N., Lischka P., Wagenknecht N. and Stamminger T. (**2009**). "Inhibition of human cytomegalovirus replication via peptide aptamers directed against the non-conventional nuclear localization signal of the essential viral replication factor pUL84." *J Virol*.

Khan F. A., Fisher M. A. and Khakoo R. A. (**2007**). "Association of hemochromatosis with infectious diseases: expanding spectrum." *Int J Infect Dis* **11**(6): 482-487.

Kim O. K., Garrity-Ryan L. K., Bartlett V. J., Grier M. C., Verma A. K., Medjanis G., Donatelli J. E., Macone A. B., Tanaka S. K., Levy S. B. and Alekshun M. N. (**2009**). "N-hydroxybenzimidazole inhibitors of the transcription factor LcrF in Yersinia: novel antivirulence agents." *Journal of medicinal chemistry* **52**(18): 5626-5634.

Kristiansen M., Graversen J. H., Jacobsen C., Sonne O., Hoffman H. J., Law S. K. and Moestrup S. K. (**2001**). "Identification of the haemoglobin scavenger receptor." *Nature* **409**(6817): 198-201.

Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I., Cui L., Oguchi A., Aoki K., Nagai Y., Lian J., Ito T., Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y., Takahashi N. K., Sawano T., Inoue R., Kaito C., Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M., Ogasawara N., Hayashi H. and Hiramatsu K. (**2001**). "Whole genome sequencing of meticillin-resistant *Staphylococcus aureus*." *Lancet* **357**(9264): 1225-1240.

Latour B. (**2001**). "Pasteur, guerre et paix des microbes."

LaVallie E. R., DiBlasio E. A., Kovacic S., Grant K. L., Schendel P. F. and McCoy J. M. (**1993**). "A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm." *Biotechnology (N Y)* **11**(2): 187-193.

Lavrrar J. L., Christoffersen C. A. and McIntosh M. A. (**2002**). "Fur-DNA interactions at the bidirectional fepDGC-entS promoter region in *Escherichia coli*." *J Mol Biol* **322**(5): 983-995.

Lavrrar J. L. and McIntosh M. A. (**2003**). "Architecture of a fur binding site: a comparative analysis." *J Bacteriol* **185**(7): 2194-2202.

Lee H. J., Park K. J., Lee A. Y., Park S. G., Park B. C., Lee K. H. and Park S. J. (**2003**). "Regulation of fur expression by RpoS and fur in *Vibrio vulnificus*." *J Bacteriol* **185**(19): 5891-5896.

Lee J. W. and Helmann J. D. (**2007a**). "Functional specialization within the Fur family of metalloregulators." *Biometals* **20**(3-4): 485-499.

Lee J. W. and Helmann J. D. (**2007b**). "Functional specialization within the Fur family of metalloregulators." *Biometals : an international journal on the role of metal ions in biology, biochemistry, and medicine* **20**(3-4): 485-499.

Litwin C. M. and Calderwood S. B. (**1994**). "Analysis of the complexity of gene regulation by fur in *Vibrio cholerae*." *J Bacteriol* **176**(1): 240-248.

Lopez-Ochoa L., Ramirez-Prado J. and Hanley-Bowdoin L. (**2006**). "Peptide aptamers that bind to a *geminivirus* replication protein interfere with viral replication in plant cells." *J Virol* **80**(12): 5841-5853.

Lounis N., Guillemont J., Veziris N., Koul A., Jarlier V. and Andries K. (**2010a**). "R207910 (TMC207) : un nouvel antibiotique pour le traitement de la tuberculose." *Médecine et Maladies Infectieuses* **40**(7): 383-390.

Lounis N., Guillemont J., Veziris N., Koul A., Jarlier V. and Andries K. (**2010b**). "[R207910 (TMC207): a new antibiotic for the treatment of tuberculosis]." *Medecine et maladies infectieuses* **40**(7): 383-390.

Lu T. K. and Collins J. J. (**2007**). "Dispersing biofilms with engineered enzymatic bacteriophage." *Proceedings of the National Academy of Sciences of the United States of America* **104**(27): 11197-11202.

Lu T. K. and Collins J. J. (**2009**). "Engineered bacteriophage targeting gene networks as adjuvants for antibiotic therapy." *Proceedings of the National Academy of Sciences of the United States of America* **106**(12): 4629-4634.

Lucarelli D., Russo S., Garman E., Milano A., Meyer-Klaucke W. and Pohl E. (**2007**). "Crystal structure and function of the zinc uptake regulator FurB from *Mycobacterium tuberculosis*." *J Biol Chem* **282**(13): 9914-9922.

Ma J. and Ptashne M. (**1988**). "Converting a eukaryotic transcriptional inhibitor into an activator." *Cell* **55**(3): 443-446.

Maresso A. W. and Schneewind O. (**2006**). "Iron acquisition and transport in *Staphylococcus aureus*." *Biometals* **19**(2): 193-203.

Marmorstein R., Carey M., Ptashne M. and Harrison S. C. (**1992**). "DNA recognition by GAL4: structure of a protein-DNA complex." *Nature* **356**(6368): 408-414.

Martel V., Filhol O., Colas P. and Cochet C. (**2006**). "p53-dependent inhibition of mammalian cell survival by a genetically selected peptide aptamer that targets the regulatory subunit of protein kinase CK2." *Oncogene* **25**(56): 7343-7353.

Martinez J. L., Delgado-Iribarren A. and Baquero F. (**1990**). "Mechanisms of iron acquisition and bacterial virulence." *FEMS Microbiol Rev* **6**(1): 45-56.

Masse E. and Gottesman S. (**2002**). "A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*." *Proc Natl Acad Sci U S A* **99**(7): 4620-4625.

McConnell S. J. and Hoess R. H. (**1995**). "Tendamistat as a scaffold for conformationally constrained phage peptide libraries." *J Mol Biol* **250**(4): 460-470.

McDevitt D. and Rosenberg M. (**2001**). "Exploiting genomics to discover new antibiotics." *Trends Microbiol* **9**(12): 611-617.

McGee D. J., George A. E., Trainor E. A., Horton K. E., Hildebrandt E. and Testerman T. L. (**2011**). "Cholesterol enhances Helicobacter pylori resistance to antibiotics and LL-37." *Antimicrob Agents Chemother* **55**(6): 2897-2904.

Mey A. R., Wyckoff E. E., Kanukurthy V., Fisher C. R. and Payne S. M. (**2005**). "Iron and fur regulation in Vibrio cholerae and the role of fur in virulence." *Infection and immunity* **73**(12): 8167-8178.

Meyer A., Deiana J. and Bernard A. (2004). Cours de microbiologie générale: avec problèmes et exercices corrigés, Doin.

Morar M. and Wright G. D. (**2010**). "The genomic enzymology of antibiotic resistance." *Annu Rev Genet* **44**: 25-51.

Moy T. I., Ball A. R., Anklesaria Z., Casadei G., Lewis K. and Ausubel F. M. (**2006**). "Identification of novel antimicrobials using a live-animal infection model." *Proceedings of the National Academy of Sciences of the United States of America* **103**(27): 10414-10419.

Nagel-Wolfrum K., Buerger C., Wittig I., Butz K., Hoppe-Seyler F. and Groner B. (**2004**). "The interaction of specific peptide aptamers with the DNA binding domain and the dimerization domain of the transcription factor Stat3 inhibits transactivation and induces apoptosis in tumor cells." *Mol Cancer Res* **2**(3): 170-182.

Nandal A., Huggins C. C., Woodhall M. R., McHugh J., Rodriguez-Quinones F., Quail M. A., Guest J. R. and Andrews S. C. (**2010**). "Induction of the ferritin gene (ftnA) of Escherichia coli by Fe(2+)-Fur is mediated by reversal of H-NS silencing and is RyhB independent." *Mol Microbiol* **75**(3): 637-657.

Nathan C. (**2004**). "Antibiotics at the crossroads." *Nature* **431**(7011): 899-902.

Nauenburg S., Zwerschke W. and Jansen-Durr P. (**2001**). "Induction of apoptosis in cervical carcinoma cells by peptide aptamers that bind to the HPV-16 E7 oncoprotein." *FASEB J* **15**(3): 592-594.

Nikaido H. (**1994**). "Prevention of drug access to bacterial targets: permeability barriers and active efflux." *Science* **264**(5157): 382-388.

Norman T. C., Smith D. L., Sorger P. K., Drees B. L., O'Rourke S. M., Hughes T. R., Roberts C. J., Friend S. H., Fields S. and Murray A. W. (**1999**). "Genetic selection of peptide inhibitors of biological pathways." *Science* **285**(5427): 591-595.

Nouvion A. L., Thibaut J., Lohez O. D., Venet S., Colas P., Gillet G. and Lalle P. (**2007**). "Modulation of Nr-13 antideath activity by peptide aptamers." *Oncogene* **26**(5): 701-710.

OMS (**2011**). "Global Tuberculosis control." *WHO report*.

Pabo C. O. and Sauer R. T. (**1992**). "Transcription factors: structural families and principles of DNA recognition." *Annu Rev Biochem* **61**: 1053-1095.

Papp D. (**1954**). "Histoire des antibiotiques." *Revue d'histoire des sciences et de leurs applications* **7**(2): 124-138.

Pecqueur L., D'Autreaux B., Dupuy J., Nicolet Y., Jacquamet L., Brutscher B., Michaud-Soret I. and Bersch B. (**2006**). "Structural changes of *Escherichia coli* ferric uptake regulator during metal-dependent dimerization and activation explored by NMR and X-ray crystallography." *J Biol Chem* **281**(30): 21286-21295.

Pohl E., Haller J. C., Mijovilovich A., Meyer-Klaucke W., Garman E. and Vasil M. L. (**2003**). "Architecture of a protein central to iron homeostasis: crystal structure and spectroscopic analysis of the ferric uptake regulator." *Mol Microbiol* **47**(4): 903-915.

Postle K. (**1993**). "TonB protein and energy transduction between membranes." *J Bioenerg Biomembr* **25**(6): 591-601.

RAISIN (**2009**). "Enquête nationale de prévalence des infections nosocomiales, France, juin 2006 – Volume 1 – Méthodes, résultats, perspectives." **1**: 1-81.

Rea R. B., Gahan C. G. and Hill C. (**2004**). "Disruption of putative regulatory loci in *Listeria monocytogenes* demonstrates a significant role for Fur and PerR in virulence." *Infect Immun* **72**(2): 717-727.

Rerole A. L., Gobbo J., De Thonel A., Schmitt E., Pais de Barros J. P., Hammann A., Lanneau D., Fourmaux E., Deminov O., Micheau O., Lagrost L., Colas P., Kroemer G. and Garrido C. (**2011**). "Peptides and aptamers targeting HSP70: a novel approach for anticancer chemotherapy." *Cancer research* **71**(2): 484-495.

Russo T. A., Carlino U. B., Mong A. and Jodush S. T. (**1999**). "Identification of genes in an extraintestinal isolate of *Escherichia coli* with increased expression after exposure to human urine." *Infect Immun* **67**(10): 5306-5314.

Saccucci L., Crance J. M., Colas P., Bickle M., Garin D. and Iseni F. (**2009**). "Inhibition of vaccinia virus replication by peptide aptamers." *Antiviral Res* **82**(3): 134-140.

Sandy M. and Butler A. (**2009**). "Microbial Iron Acquisition: Marine and Terrestrial Siderophores." *Chemical reviews* **109**(10): 4580-4595.

Schaffer S., Hantke K. and Braun V. (**1985**). "Nucleotide sequence of the iron regulatory gene fur." *Mol Gen Genet* **200**(1): 110-113.

Schaible U. E. and Kaufmann S. H. (**2004**). "Iron and microbial infection." *Nat Rev Microbiol* **2**(12): 946-953.

Schmidt S., Diriong S., Mery J., Fabbrizio E. and Debant A. (**2002**). "Identification of the first Rho-GEF inhibitor, TRIPalpha, which targets the RhoA-specific GEF domain of Trio." *FEBS Lett* **523**(1-3): 35-42.

Schryvers A. B. and Stojiljkovic I. (**1999**). "Iron acquisition systems in the pathogenic *Neisseria*." *Mol Microbiol* **32**(6): 1117-1123.

Sebastian S., Agarwal S., Murphy J. R. and Genco C. A. (**2002**). "The gonococcal fur regulon: identification of additional genes involved in major catabolic, recombination, and secretory pathways." *J Bacteriol* **184**(14): 3965-3974.

Sheikh M. A. and Taylor G. L. (**2009**). "Crystal structure of the *Vibrio cholerae* ferric uptake regulator (Fur) reveals insights into metal co-ordination." *Mol Microbiol*.

Slee A. M., Wuonola M. A., McRipley R. J., Zajac I., Zawada M. J., Bartholomew P. T., Gregory W. A. and Forbes M. (**1987**). "Oxazolidinones, a new class of synthetic antibacterial agents: in vitro and in vivo activities of DuP 105 and DuP 721." *Antimicrob Agents Chemother* **31**(11): 1791-1797.

Spellberg B. (**2008**). "Antibiotic resistance and antibiotic development." *The Lancet infectious diseases* **8**(4): 211-212; author reply 212-214.

Spratt B. G. (**1994**). "Resistance to antibiotics mediated by target alterations." *Science* **264**(5157): 388-393.

Spratt B. G. and Cromie K. D. (**1988**). "Penicillin-binding proteins of gram-negative bacteria." *Rev Infect Dis* **10**(4): 699-711.

Staggs T. M. and Perry R. D. (**1991**). "Identification and cloning of a fur regulatory gene in *Yersinia pestis*." *J Bacteriol* **173**(2): 417-425.

Stojiljkovic I., Baumler A. J. and Hantke K. (**1994**). "Fur regulon in gram-negative bacteria. Identification and characterization of new iron-regulated *Escherichia coli* genes by a fur titration assay." *J Mol Biol* **236**(2): 531-545.

Tally F. P. and DeBruin M. F. (**2000**). "Development of daptomycin for gram-positive infections." *The Journal of antimicrobial chemotherapy* **46**(4): 523-526.

Tiss A., Barre O., Michaud-Soret I. and Forest E. (**2005a**). "Characterization of the DNA-binding site in the ferric uptake regulator protein from Escherichia coli by UV crosslinking and mass spectrometry." *FEBS letters* **579**(25): 5454-5460.

Tiss A., Barre O., Michaud-Soret I. and Forest E. (**2005b**). "Characterization of the DNA-binding site in the ferric uptake regulator protein from *Escherichia coli* by UV crosslinking and mass spectrometry." *FEBS Lett* **579**(25): 5454-5460.

Touati D. (**2000**). "Iron and oxidative stress in bacteria." *Arch Biochem Biophys* **373**(1): 1-6.

Touati D., Jacques M., Tardat B., Bouchard L. and Despied S. (**1995**). "Lethal oxidative damage and mutagenesis are generated by iron in delta fur mutants of Escherichia coli: protective role of superoxide dismutase." *J Bacteriol* **177**(9): 2305-2314.

Traore D. A., El Ghazouani A., Ilango S., Dupuy J., Jacquamet L., Ferrer J. L., Caux-Thang C., Duarte V. and Latour J. M. (**2006**). "Crystal structure of the apo-PerR-Zn protein from *Bacillus subtilis*." *Mol Microbiol* **61**(5): 1211-1219.

Traore D. A., El Ghazouani A., Jacquamet L., Borel F., Ferrer J. L., Lascoux D., Ravanat J. L., Jaquinod M., Blondin G., Caux-Thang C., Duarte V. and Latour J. M. (**2009**). "Structural and functional characterization of 2-oxo-histidine in oxidized PerR protein." *Nat Chem Biol* **5**(1): 53-59.

Tréfouël J., Nitti F. and Bovet D. (**1935**). "Activité du p-aminophénylsulfamide sur l'infection streptococcique expérimentale de la souris et du lapin." *C. R. Soc. Biol.* **120**: 756.

Triezenberg S. J., Kingsbury R. C. and McKnight S. L. (**1988**). "Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression." *Genes Dev* **2**(6): 718-729.

Verneuil N., Rince A., Sanguinetti M., Posteraro B., Fadda G., Auffray Y., Hartke A. and Giard J. C. (**2005**). "Contribution of a PerR-like regulator to the oxidative-stress response and virulence of *Enterococcus faecalis*." *Microbiology* **151**(Pt 12): 3997-4004.

Veziris N., Truffot C., Lounis N., Chauffour A., Andries K. and Jarlier V. (**2006**). "R207910 : un nouvel antibiotique prometteur pour le traitement de la tuberculose." *Revue des Maladies Respiratoires* **23, Supplement 1**(0): 37.

Vitale S. (**2009**). "Inhibiteurs anti-Fur et caractérisation de Fur de Helicobacter pylori", ***Université J. Fourier***.

Vitale S., Fauquant C., Lascoux D., Schauer K., Saint-Pierre C. and Michaud-Soret I. (**2009**). "A ZnS(4) structural zinc site in the Helicobacter pylori ferric uptake regulator." *Biochemistry* **48**(24): 5582-5591.

VMNIS (**2005**). "Prévalence de l'anémie dans le monde 1993-2005." *OMS*.

Walsh C. (**2000**). "Molecular mechanisms that confer antibacterial drug resistance." *Nature* **406**(6797): 775-781.

Walsh C. (**2003**). "Where will new antibiotics come from?" *Nat Rev Microbiol* **1**(1): 65-70.

Walsh T. R. (**2010**). "Emerging carbapenemases: a global perspective." *International journal of antimicrobial agents* **36 Suppl 3**: S8-14.

Wandersman C. and Delepelaire P. (**2004**). "Bacterial iron sources: from siderophores to hemophores." *Annu Rev Microbiol* **58**: 611-647.

Wang F., Cheng S., Sun K. and Sun L. (**2008**). "Molecular analysis of the fur (ferric uptake regulator) gene of a pathogenic *Edwardsiella tarda* strain." *J Microbiol* **46**(3): 350-355.

Weinberg E. D. (**1997**). "The *Lactobacillus* anomaly: total iron abstinence." *Perspect Biol Med* **40**(4): 578-583.

Welch R. A., Burland V., Plunkett G., 3rd, Redford P., Roesch P., Rasko D., Buckles E. L., Liou S. R., Boutin A., Hackett J., Stroud D., Mayhew G. F., Rose D. J., Zhou S., Schwartz D. C., Perna N. T., Mobley H. L., Donnenberg M. S. and Blattner F. R. (**2002**). "Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*." *Proc Natl Acad Sci U S A* **99**(26): 17020-17024.

Whitby P. W., Seale T. W., VanWagoner T. M., Morton D. J. and Stull T. L. (**2009**). "The iron/heme regulated genes of *Haemophilus influenzae*: comparative transcriptional profiling as a tool to define the species core modulon." *BMC Genomics* **10**: 6.

Whitman W. B., Coleman D. C. and Wiebe W. J. (**1998**). "Prokaryotes: the unseen majority." *Proceedings of the National Academy of Sciences of the United States of America* **95**(12): 6578-6583.

Wilderman P. J., Sowa N. A., FitzGerald D. J., FitzGerald P. C., Gottesman S., Ochsner U. A. and Vasil M. L. (**2004**). "Identification of tandem duplicate regulatory small RNAs in *Pseudomonas aeruginosa* involved in iron homeostasis." *Proc Natl Acad Sci U S A* **101**(26): 9792-9797.

Wooldridge K. G. and Williams P. H. (**1993**). "Iron uptake mechanisms of pathogenic bacteria." *FEMS Microbiol Rev* **12**(4): 325-348.

Worst D. J., Gerrits M. M., Vandenbroucke-Grauls C. M. and Kusters J. G. (**1998**). "*Helicobacter pylori* ribBA-mediated riboflavin production is involved in iron acquisition." *J Bacteriol* **180**(6): 1473-1479.

Wright T. R., Shan G., Walsh T. A., Lira J. M., Cui C., Song P., Zhuang M., Arnold N. L., Lin G., Yau K., Russell S. M., Cicchillo R. M., Peterson M. A., Simpson D. M., Zhou N., Ponsamuel J. and Zhang Z. (**2010**). "Robust crop resistance to broadleaf and grass herbicides provided by aryloxyalkanoate dioxygenase transgenes." *Proceedings of the National Academy of Sciences of the United States of America* **107**(47): 20240-20245.

Xu C. W. and Luo Z. (**2002**). "Inactivation of Ras function by allele-specific peptide aptamers." *Oncogene* **21**(37): 5753-5757.

Yoneyama H. and Katsumata R. (**2006**). "Antibiotic resistance in bacteria and its future for novel antibiotic development." *Biosci Biotechnol Biochem* **70**(5): 1060-1075.

Chapitre II :

Approche expérimentale de l'étude de l'interaction des inhibiteurs anti-Fur et de leur cible

CHAPITRE II : ..... **- 85 -**

APPROCHE EXPERIMENTALE DE L'ETUDE DE L'INTERACTION DES INHIBITEURS ANTI-FUR ET DE LEUR CIBLE ..... **- 85 -**

II. A| RAPPEL DES DONNEES PREEXISTANTES ..... - 89 -

II. A. 1| La sélection des aptamères peptidiques F1 à F5 ..... *- 89 -*

II. A. 2| Les peptides linéaires dérivés des aptamères peptidiques anti-Fur ..... *- 91 -*

II. A. 3| Comparaison des interactions de EcFur-aptamères EcFur-peptides la « peptidisation ». - 92 -

II. B| CARACTERISATION DES PEPTIDES PF1 A PF5 ..... - 94 -

II. B. 1| Etude structurale des peptides par dichroïsme circulaire (CD) ..... *- 94 -*

II. B. 2| Oxydation, solubilité et interaction avec les métaux des peptides ..... *- 95 -*

II. B. 3| Interaction de pF1 à pF5 avec la protéine EcFur ..... *- 96 -*

II. C| SYNTHESE CHIMIQUE DES PEPTIDES ..... - 102 -

II. C. 1| Principe ..... *- 102 -*

II. C. 2| L'étape de protection de l'amine libre en N-terminal ..... *- 104 -*

II. C. 3| L'étape d'activation ..... *- 104 -*

II. C. 4| Décrochage du peptide de la résine ..... *- 105 -*

II. C. 5| Synthèse peptidique chimique en solution ..... *- 106 -*

II. C. 6| Résultats des synthèses peptidiques ..... *- 107 -*

II. D| ÉTUDE D'INHIBITION D'ECFUR PAR LES PEPTIDES SYNTHETISES ..... - 110 -

II. D. 1| Etude d'inhibition d'EcFur par des dérivés de pF1 ..... *- 110 -*

II. D. 2| Etude d'inhibition d'EcFur par des dérivés de pF2 ..... *- 117 -*

II. D. 3| Etude d'inhibition d'EcFur par des dérivés de pF3 et pF4 ..... *- 120 -*

II. D. 4| Etude d'inhibition d'EcFur par l'ester du benzyl-tryptophane : W_CO_Bz ..... *- 121 -*

II. D. 5| Essais de co-cristallisation des variants de pF1 avec EcFur ..... *- 123 -*

II. E| BILAN ET DISCUSSION ..... - 125 -

BIBLIOGRAPHIE ..... - 129 -

II. A| Rappel des données préexistantes

Dans cette partie, nous rappellerons principalement les résultats de thèse de Sylvia Vitale qui a beaucoup travaillé sur la caractérisation des aptamères peptidiques anti-Fur et de leurs dérivés peptidiques (Vitale 2009). Nous parlerons brièvement des travaux de thèse de Nadia Abed qui a été la première à travailler sur les aptamères peptidiques anti-Fur (Abed 2005).

Le projet d'obtention d'inhibiteurs dirigés contre Fur par la technologie des aptamères peptidique a été réalisé dans le cadre d'un projet soutenu par la région Rhône-Alpes en collaboration avec Pierre Colas (à l'époque à l'ENS Lyon et Aptanomics SA, maintenant à Roscoff) et notre laboratoire. C'est dans ce cadre que pendant sa thèse, Nadia a réalisé la sélection des aptamères peptidiques anti-Fur (Abed 2005). Elle a réalisé la sélection phénotypique d'aptamères peptidiques capables d'inactiver conditionnellement *in vivo* la protéine Fur d'*Escherichia coli.* Nous ne parlerons pas de ces aptamères dans cette thèse. Ce qui nous intéresse ici, ce sont les aptamères peptidiques interagissant spécifiquement avec Fur sélectionnés par criblage double-hybride. Les quatre aptamères peptidiques sélectionnés pour leurs interactions avec EcFur se sont avérés très intéressants par la suite. En effet, ces molécules sont capables d'inhiber EcFur et d'affecter la virulence d'une souche d'*Escherichia coli* pathogène dans un modèle d'infection de drosophile (Abed *et al.*, 2007). Sylvia Vitale, une ancienne doctorante du labo a consacré une partie de sa thèse sur la caractérisation de ces aptamères peptidiques ainsi que l'étude les peptides linéaires dérivés et de leurs effets sur la protéine Fur d'*Escherichia coli* (Vitale 2009). Une partie de mon travail de thèse a été réalisé conjointement avec cette thèse précédente. Avant de présenter les résultats, il est nécessaire de rappeler quelques résultats antérieurs permettant de mieux comprendre la démarche.

II. A. 1| La sélection des aptamères peptidiques F1 à F5

Le criblage double hybride a été réalisé chez la levure par le Dr Marc Bickle, (Aptanomics). La sélection des aptamères a été effectuée contre la protéine Fur d'*Escherichia coli* avec une banque estimée à 2.7 x 10^7 molécules (Abed *et al.*, 2007). Quatre clones, nommés F1 à F4 présentant un phénotype d'interaction double hybride avec la protéine Fur d'*Escherichia coli*, ont été isolés (Figure II.1).

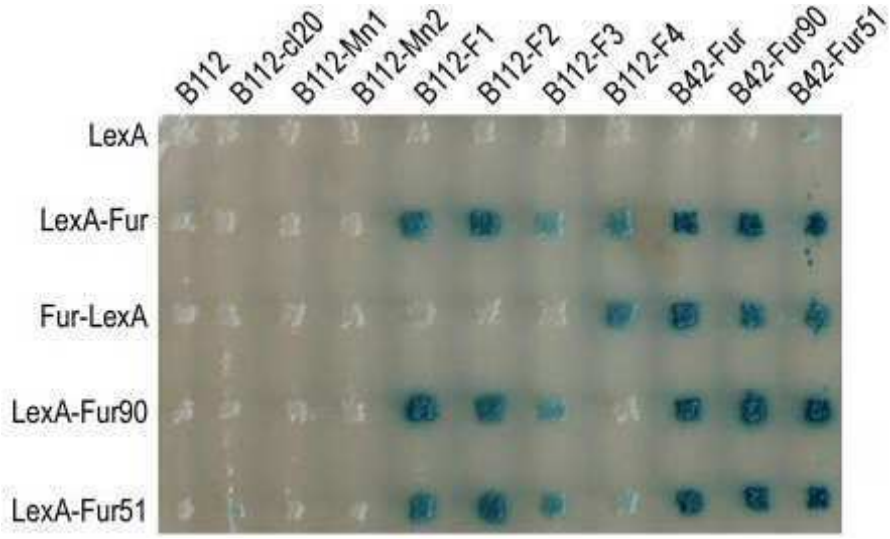


Figure II.1: Matrice d'interaction double hybride chez la levure. Les appâts sont les différentes protéines Fur (sauvage, Fur51 et Fur90) fusionnées au domaine LexA en N- ou C-terminal (respectivement LexA-Fur et Fur-LexA) et les proies sont les aptamères peptidiques et les différentes protéines Fur, fusionnés respectivement aux domaines B112 et B42. Les aptamères cl20 (renommé F5 pour la suite), Mn1 et Mn2, ainsi que le domaine LexA seul ont servi de contrôles négatifs (Abed 2005; Abed *et al.*, 2007).

Dans cette technique, la coloration bleue indique qu'il y a interaction et l'intensité de la couleur donne une indication quant aux affinités apparentes entre les molécules d'intérêt. D'après la Figure II.1, le domaine B112, utilisé ici comme contrôle de la fusion n'interagit avec aucune des protéines. On observe le même phénomène pour le contrôle cl20 (un aptamère choisi au hasard dans la banque et nommé par la suite F5) et aussi pour les aptamères Mn1 et Mn2 (issus de la sélection phénotypique). On voit bien que les aptamères F1 à F4 interagissent avec la fusion LexA-Fur. Par contre, l'interaction entre l'aptamère F3 et Fur est plus faible que pour les autres aptamères (colonne 7) (Figure II.1). De façon intéressante, F4 est le seul à interagir avec Fur lorsque celle-ci est fusionnée à LexA par son extrémité C-terminal, mais n'interagit pas avec les mutants Fur51 et Fur90 (mutants dominants négatifs utilisés comme contrôle positif dans les tests d'interaction car ils peuvent former des hétérodimères inactifs avec EcFur). Cette dernière information suggère que cet aptamère interagirait plutôt avec la partie C-terminal de Fur.

Afin de confirmer cela, différentes formes tronquées de la protéine Fur ont été utilisées en appât dans un nouveau test double hybride (Figure II.2).

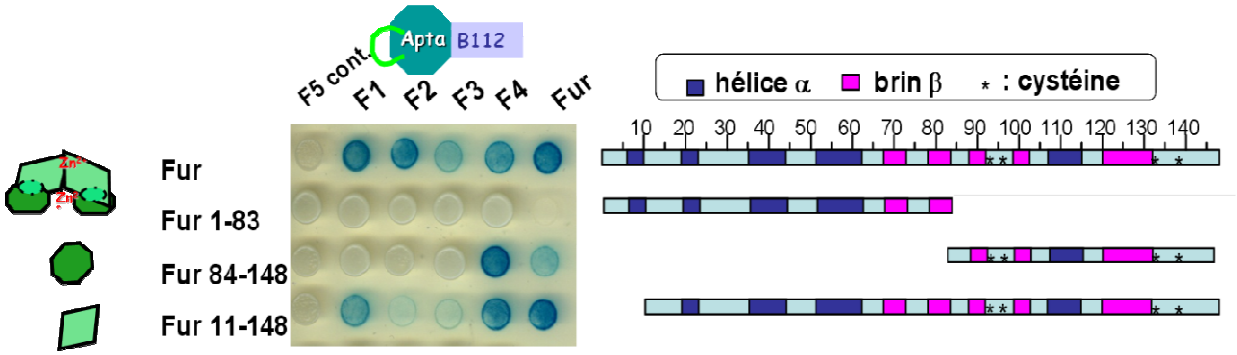


Figure II.2: Test d'interaction double hybride entre les différentes parties tronquées de Fur et les aptamères peptidiques.

Ce test montre qu'aucun aptamère peptidique n'est capable d'interagir avec le domaine N-terminal seul de la protéine (construction Fur 1-83) et confirme que F4 interagit bien avec le domaine C-terminal seul (construction Fur 84-148). De plus, F1 ne semble pas avoir besoin des 11 premiers acides aminés de Fur pour l'interaction contrairement à F2 et F3. Les résultats de ce test suggèrent que les aptamères F1 à F4 interagissaient différemment avec la protéine, probablement dans des zones différentes. L'interaction des aptamères F1 à F4 avec Fur a ensuite été confirmée chez *Escherichia coli* par des expériences de « GST pull-down » qui ont montré encore une fois de plus, comme la matrice double hybride, une interaction plus faible entre l'aptamère F3 et la protéine Fur. Cette tendance a été confirmée par un test double hybride plus quantitatif, utilisant le gène codant pour la luciférase comme gène rapporteur (Abed *et al.*, 2007).

La suite des caractérisations fonctionnelles par Nadia a montré que :

- L'aptamère F1 agit sur le régulon Fur qui contrôle l'expression de 196 gènes.
- Les aptamères peptidiques F1, F2, F3 et F4 induisent une diminution de la virulence de la souche pathogène d'*Escherichia coli* pathogène (souche 1106) dans un modèle d'infection de drosophiles (Abed *et al.*, 2007).

II. A. 2| Les peptides linéaires dérivés des aptamères peptidiques anti-Fur

Dans le but de mettre au point des molécules actives plus petites, dites «de deuxième génération», des peptides linéaires de 13 acides aminés pF1 à pF5 correspondant respectivement aux séquences variables des aptamères peptidiques F1 à F5 ont été étudiés.

Les séquences de la partie active des aptamères peptidiques F1 à F5, c'est-à-dire les peptides pF1 à pF5 et certaines de leurs caractéristiques sont reportées dans le Tableau II.1.

Nom du peptide	**Séquences**	**MM (Da) théorique sans contre-ions**	**pI**	**ε_{208nm} ($M^{-1}.cm^{-1}$)**
pF1	RLWCRYPHPPLTD	1654	8,2	6990
pF2	RQCNICGASLYSY	1478	8	2980
pF3	ETCKCGSQVWRHS	1521	8,2	5500
pF4	CARCGARVNVYKY	1502	9,4	2980
pF5 (contrôle)	RSLRGRCLSQHQD	1555	10,3	0

Tableau II.1 : Principales caractéristiques des peptides linéaires pF1 à pF5. MM : masse molaire ; Da : dalton ; pI : point isoélectrique ; ε280nm : coefficient d'extinction molaire à 280 nm. Les masses molaires correspondent aux masses théoriques. Le pI et le ε280nm sont des valeurs théoriques, prédites à l'aide du site web www.expasy.com

Les séquences des peptides ne contiennent aucune homologie dans la PDB. Elles possèdent toutes au moins une ou deux cystéines qui peuvent former des ponts disulfures et des résidus pouvant lier les métaux.

Les premiers peptides pF1 à pF5 ont été achetés chez Genscript, caractérisés puis testés pour leur capacité à inhiber Fur à l'aide du test de protection à la nucléase. Comme nous le verrons par la suite, nous avons synthétisé les autres peptides.

II. A. 3| Comparaison des interactions de EcFur-aptamères EcFur-peptides la « peptidisation »

Cette technique, également décrite par Bickle repose sur l'utilisation du test double hybride chez la levure (Abed *et al.*, 2007). Le principe est le même que pour les aptamères peptidiques complets, mais dans ce cas ce sont uniquement les séquences variables (ici pF1 à pF5) qui sont fusionnées au domaine d'activation B112. Deux plasmides de proie différents ont été utilisés, de façon à fusionner B112 à l'extrémité N- ou C-terminal du peptide. Ces différentes constructions sont illustrées ci-dessous.

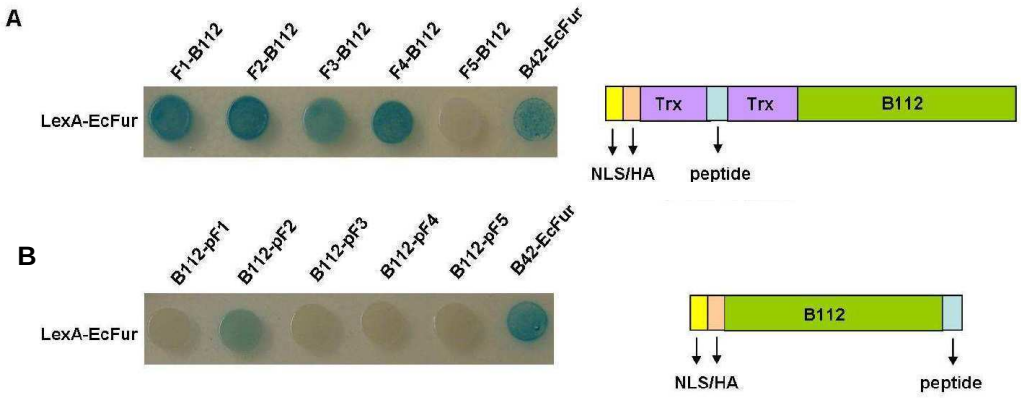


Figure II.3: Matrices d'interaction en double hybride chez la levure.
A : Interaction des aptamères F1 à F5 avec Fur d'*E.coli*.
B : Interaction des peptides pF1 à pF5 avec Fur d'*E.coli*.
Les appâts et les proies sont précisés respectivement à droite et au-dessus des matrices. Les différentes constructions de proies utilisées sont représentées à droite de la matrice d'interaction correspondante. Trx : Thioredoxine A d'*E.coli.* NLS : séquence de localisation nucléaire. HA : étiquette hémagglutinine. Tailles réelles : NLS : 12 aa ; HA : 13 aa ; peptide : 13 aa ; Trx : 108 aa (soit 2x54 ici) ; B112 : 360 aa.

La Figure II.3.A montre que les aptamères F1 à F4 interagissent avec la protéine Fur comme avec les résultats obtenus par Nadia Abed (Abed *et al.*, 2007). La Figure II.3.B montre une interaction de pF2 avec la protéine. Comme on estime qu'en utilisant le gène rapporteur *lacZ*, le double hybride permet de visualiser des interactions donc le Kd est de l'ordre du µM au nM, on peut dire que pF2 a une bonne affinité pour EcFur, même si cette affinité est plus faible que celle de l'aptamère F2, au vu des intensités de la coloration. Les autres peptides ne doivent pas avoir une affinité assez forte avec la protéine pour pouvoir observer une interaction dans ce test.

Cette technique de « peptidisation» permet de comparer l'activité des aptamères peptidiques complets F1 à F4 à leurs dérivés pF1 à pF. Ce qui n'a pas été possible avec le test de protection à la nucléase avec les aptamères purifiés (Vitale 2009).

Sylvia a aussi testé l'interaction des aptamères F1 à F4 avec les protéines Fur d'autres microorganismes pathogènes par double hybride. Elle a montré que l'aptamère F4 interagit avec d'autres protéines Fur. Nous reviendrons sur ces résultats dans le chapitre III, au moment de présenter les résultats théoriques équivalents.

II. B| Caractérisation des peptides pF1 à pF5

Dans cette partie, nous verrons des résultats où j'ai participé à l'obtention des données au cours de mon stage M2 en travaillant au côté de Sylvia Vitale.

II. B. 1| Etude structurale des peptides par dichroïsme circulaire (CD)

Des mesures de dichroïsme circulaire en UV-lointain ont été effectuées sur les peptides pF1 à pF5 synthétisés par Genscript afin d'étudier la structuration de ces peptides. L'allure des courbes enregistrées et la position des minima et des maxima permettent d'obtenir des informations structurales sur les protéines et les peptides. Les types d'informations structurales qu'on peut avoir sont donnés sur la Figure II.4.

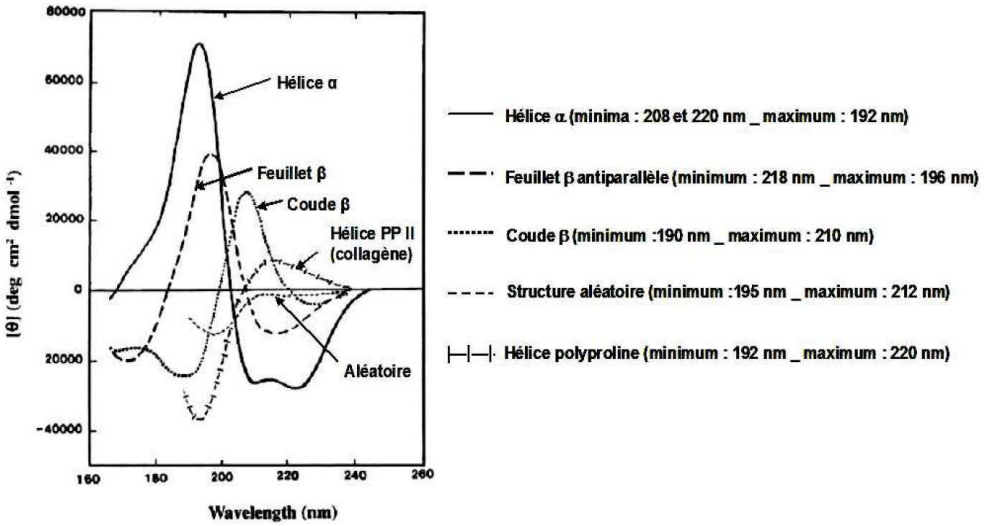


Figure II.4: Spectres CD dans l'UV-lointain de différentes structures secondaires (Kelly *et al.*, 2005)

Pour un peptide structuré, on doit observer des bandes positives autour de 190-200 nm et des bandes négatives autour de 210-230 nm. Par contre, une absence de structure secondaire conduira à des spectres ayant des composantes négatives autour de 200 nm.

Les spectres des peptides ont été enregistrés dans un tampon BTP 10 mM pH 7,5, KCl 10 mM, $MgSO_4$ 5 mM, dont la composition est proche de celle du tampon du test d'activité de protection à la nucléase, mais avec une concentration en sels plus faible. L'analyse des spectres (Figure II.5, ci-dessous) montre que les peptides pF1 à pF5 ne présentent aucune

structure secondaire particulière. Seul le peptide pF4 présente de faibles bandes correspondant peut-être à un coude β. Ces résultats ne sont pas surprenants pour des peptides de cette taille.

L'évolution des spectres des peptides au cours du temps a été suivie pendant 24h. Aucune modification notable n'a été observée ce qui était conforme aux résultats d'analyses Electrospray, qui ont montré que les peptides à 20 µM sont stables plusieurs heures à 25 °C.

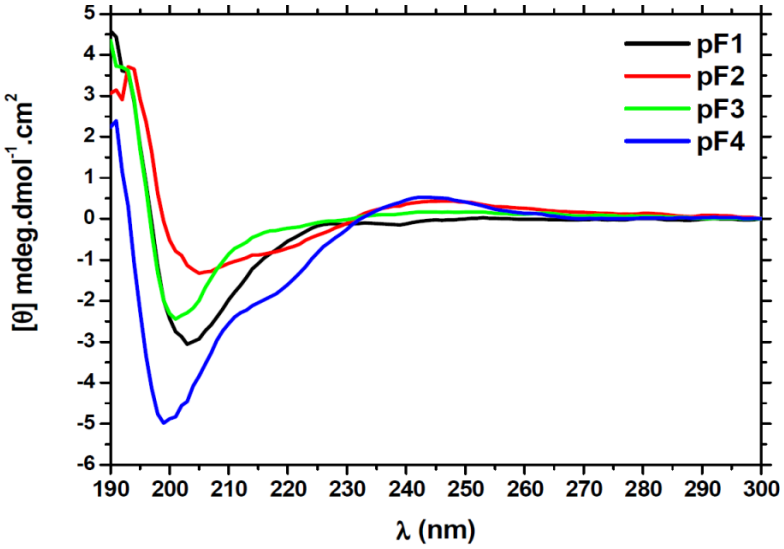


Figure II.5: Spectres CD enregistrés dans l'UV-lointain des peptides pF1 à pF4 à 20 µM, après soustraction du spectre du tampon. Les spectres ont été enregistrés dans le tampon BTP 10 mM pH 7,5, KCl 10 mM, MgSO4 5 mM à 25°C après 10 min d'équilibration.

II. B. 2| Oxydation, solubilité et interaction avec les métaux des peptides

Avant de réaliser les tests d'activité en présence des peptides, des études sur les capacités des peptides à lier les métaux ont été réalisée. Nous avons constaté également des problèmes de solubilité et d'oxydation avec certains peptides. Il est intéressant de les mentionner avant de présenter les résultats des tests.

Tout d'abord le peptide pF2, qu'on vient de voir comme interagissant très fortement avec la protéine Fur, n'est pas soluble dans le tampon BTP, utilisé pour le test de protection à la nucléase (que nous allons voir plus tard). Par contre, il est bien soluble dans le DMSO, qui peut avoir un effet inhibiteur sur l'enzyme de digestion Hinf1 (observé si la proportion de DMSO est importante dans le tampon), utilisé dans le test. De plus, Sylvia a montré que pF2 peut lier le Cobalt et former des ponts disulfures dans le DMSO (Vitale 2009). La formation

des ponts disulfures n'est pas spécifique de pF2 dans les conditions du test car ce phénomène a été observé également pour pF1 d'après des analyse de masse et de séquençage. Ces analyses n'ont pas été faites pour tous les peptides mais on peut s'attendre à ce que ce soit le cas. Ces ponts disulfures pourraient jouer un rôle important sur les conformations des peptides quant à leurs interactions avec la protéine dans les conditions du test. C'est donc avec certains de ces risques que les tests d'activité de la protéine en présence des peptides ont été réalisés. Nous avons très souvent utilisé des solutions de peptide fraîches mais aussi parfois des solutions qui ont été stockées à -20°C. Or, Genscript recommande maintenant le stockage de certains des peptides sous argon à -80°C surtout lorsqu'ils sont en solution. Désormais, tous les peptides sont stockés à -80°C dès qu'ils sont en solution.

II. B. 3| Interaction de pF1 à pF5 avec la protéine EcFur

Nous allons commencer par détailler la surproduction et la purification de la protéine. Ensuite, nous verrons le principe du test d'activité avant de présenter les résultats de ces tests.

II. B. 3. a| Surproduction et purification de la protéine EcFur

Le protocole pour surproduire et purifier la protéine EcFur, comme d'autres protéines Fur est bien maitrisé au laboratoire.

La protéine EcFur a été surproduite à partir de cellules BL21(DE3) commerciales transformées avec le plasmide pET-30c_*fur*. La surproduction est induite avec l'IPTG pendant 3h. L'extraction est réalisé dans du tampon MOPS contenant trois inhibiteurs de protéases : l'ICT (inhibiteur de la trypsine et de la chymotrypsine), le PMSF et la pepstatine. Les membranes bactériennes sont cassées par sonication et la protéine est précipitée par « salting-out » à 77,5% puis dissoute dans du tampon et dialysée. Une chromatographie sur une colonne d'affinité Chelating Superose chargée en Zn(II) permet de séparer EcFur des autres protéines. La protéine est alors éluée par une solution d'histidine et traitée à l'EDTA pour capter les ions Zn^{2+} élués avec la protéine.

La dernière étape de purification consiste en une chromatographie d'exclusion sur colonne « Superdex 75 prep grade » (HR16/60) GE Healthcare). Le chromatogramme de cette dernière est présenté à la Figure II.6, ci-dessous. Sur cette figure, la protéine Fur se retrouve dans les pics à 65 mL (dimère) et 73 mL (monomère). Les pics observés avant 60 mL

contiennent majoritairement de la protéine Fur sous forme d'oligomères (tétramère, hexamère et octamère) qui sont difficilement utilisables. Les fractions contenant du dimère ou du monomère purs sont regroupées et concentrées sur Ultrafree (« cutoff » 5000, Millipore).

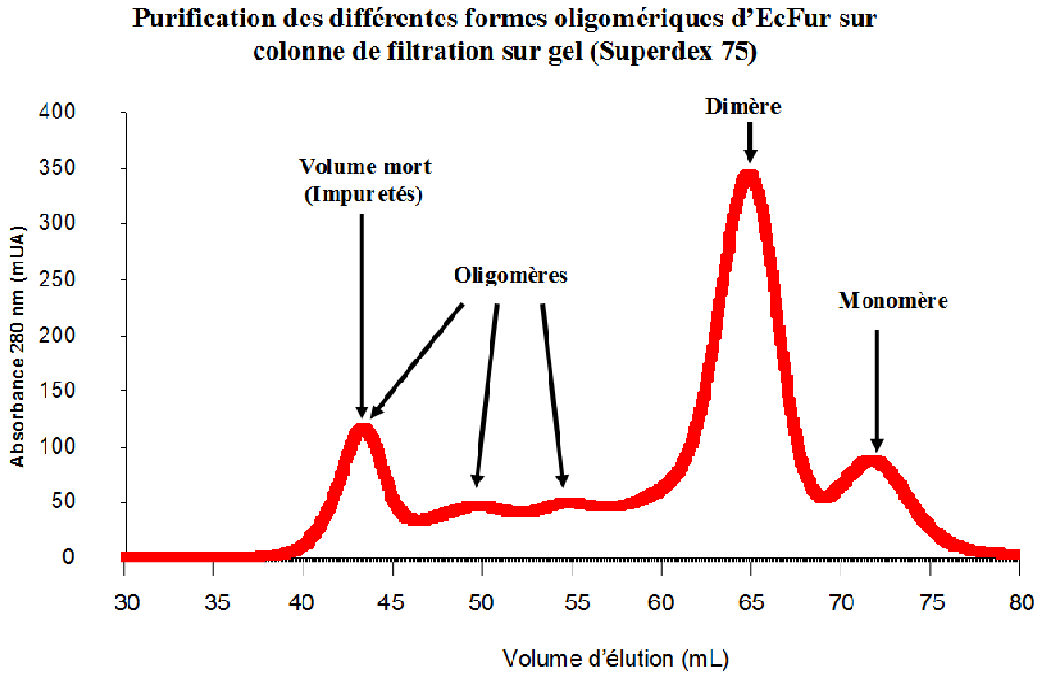


Figure II.6: Purification d'EcFur sur colonne de filtration Superdex 75

La pureté de la protéine est vérifiée par gel SDS-PAGE à 17 % (Figure II.7).

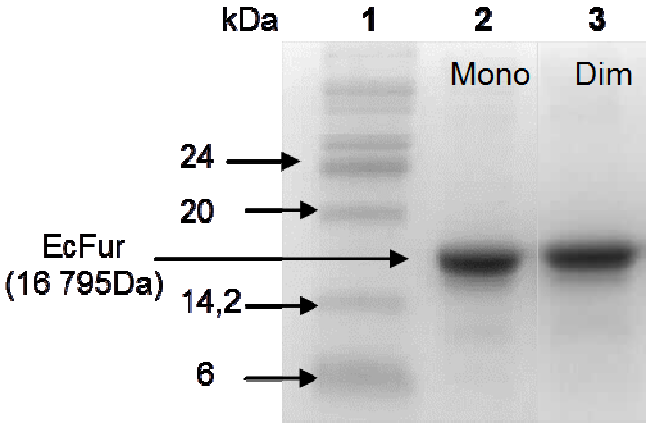


Figure II.7: Gels SDS-PAGE à 17% d'EcFur :

Ligne 1 : marqueurs de poids moléculaires : ligne 2 : fraction monomérique : ligne 3 fraction dimérique

La concentration en protéine a été déterminée par un dosage en spectroscopie UV-Visible à 280 nm, en utilisant un coefficient d'extinction massique de 0.4 mg/mL. Le

rendement de la purification a été d'environ 25 mg de dimère et 10 mg de monomère protéiques pour 3 L de culture bactérienne (Matériels et méthodes).

Nous allons maintenant présenter le test d'activité d'EcFur utilisé pour tester l'activité de la protéine en présence ou non des peptides dérivés des aptamères peptidiques.

II. B. 3. b| Présentation du test d'activité de protection à la nucléase

Le test d'activité utilisé pour évaluer l'effet des peptides sur la fixation de Fur à l'ADN a été décrit pour la première fois par Bagg *et al.*, (Bagg *et al.*, 1987) puis optimisé au laboratoire (D'Autreaux *et al.*, 2002).

Le plasmide pDT10 utilisé contient la séquence promotrice du gène codant pour l'aérobactine (Figure II.8A), qui contient elle-même deux boîtes Fur. Le plasmide pDT10 comporte également quatre sites de reconnaissance pour l'enzyme de restriction *Hinf*I, dont un contenu naturellement dans une des deux boîtes Fur (Figure II.8A). La présence de ce site de coupure naturel est utilisée pour tester la capacité d'EcFur à se lier à l'ADN. En effet, lorsque la protéine EcFur est activée par un dication métallique, elle se lie sur la boîte Fur et protège ainsi un site de restriction de la coupure par *Hinf*I. Après digestion et migration sur gel d'agarose, une bande de 1781 pb est visualisée, contenant la boîte Fur intacte (Figure II.8B). En revanche, lorsque la protéine est inactive, le site de coupure devient accessible et deux bandes, respectivement de 1530 pb et 251 pb, sont visualisées (Figure II.8B).

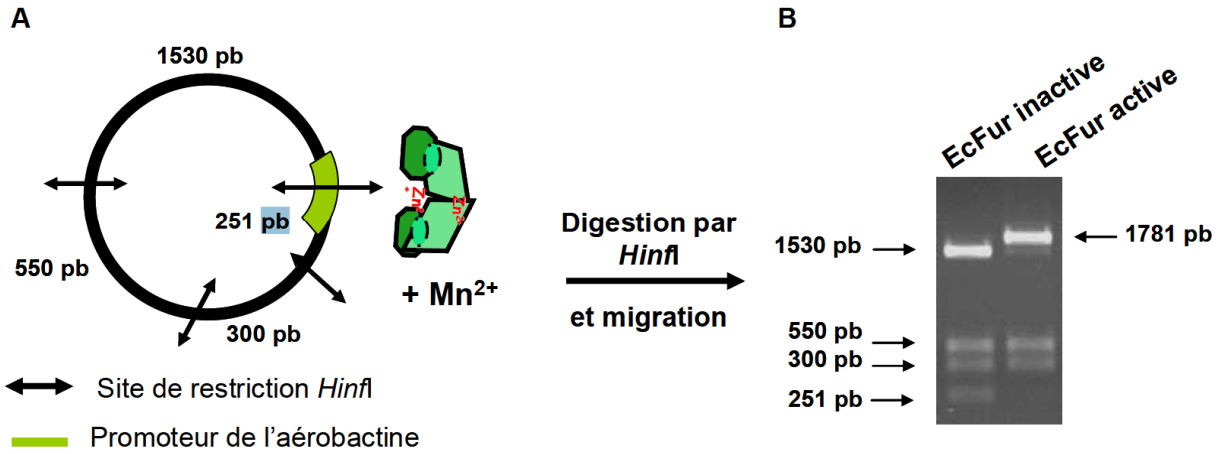


Figure II.8: Principe du test de protection à la coupure par la nucléase HinfI. A : carte des sites de coupure pour *Hinf*I dans le plasmide pDT10. Les tailles attendues des fragments sont indiquées. B : profil de migration sur gel d'agarose du plasmide pDT10 après digestion par *Hinf*I pendant 1 h à 37 °C en présence de protéine EcFur inactive ou activée par le Mn^{2+}.

Dans la pratique, le test de protection à la nucléase s'effectue en présence du dimère d'EcFur que l'on peut activer par ajout d'un métal (Mn^{2+} dans les conditions présentes, plus stable que le Fe^{2+} en conditions aérobies) ou inhiber par L'EDTA. Les puits correspondant à ces deux agents sur le profil servent respectivement de contrôles positifs et négatifs. L'incubation de la protéine avec un peptide linéaire ou une petite molécule avant l'ajout de l'ADN permet de déterminer si le peptide est capable d'inhiber l'interaction de Fur avec l'ADN *in vitro*.

Ce test repose sur une manière indirecte de visualiser la capacité des peptides à inhiber la liaison d'EcFur à l'ADN. En effet, on observe la conséquence de l'interaction des peptides avec EcFur (liaison ou non d'EcFur à l'ADN) mais pas une formation directe du complexe peptide-EcFur. Ce test ne permet donc pas de déterminer des constantes de dissociation pour ces différents complexes. Par contre, il permet la détermination d'IC_{50}, qui correspond à la concentration de molécules (peptides) nécessaire pour inhiber 50 % de l'activité de liaison d'EcFur à l'ADN. Cette valeur est relative et dépend des conditions expérimentales. Les IC_{50} des molécules testées peuvent être comparés uniquement si les tests sont réalisés dans les mêmes conditions.

II. B. 3.c| Effet des peptides linéaires pF1 à pF5 sur la liaison d'EcFur à l'ADN

Les premiers tests ont montré que les peptides pF1 à pF3 (commandés chez Genscript) inhibent l'activité de liaison du dimère d'EcFur contrairement à pF4 qui semble inactif. pF5, le peptide contrôle n'a aussi aucun effet (Figure II.9).

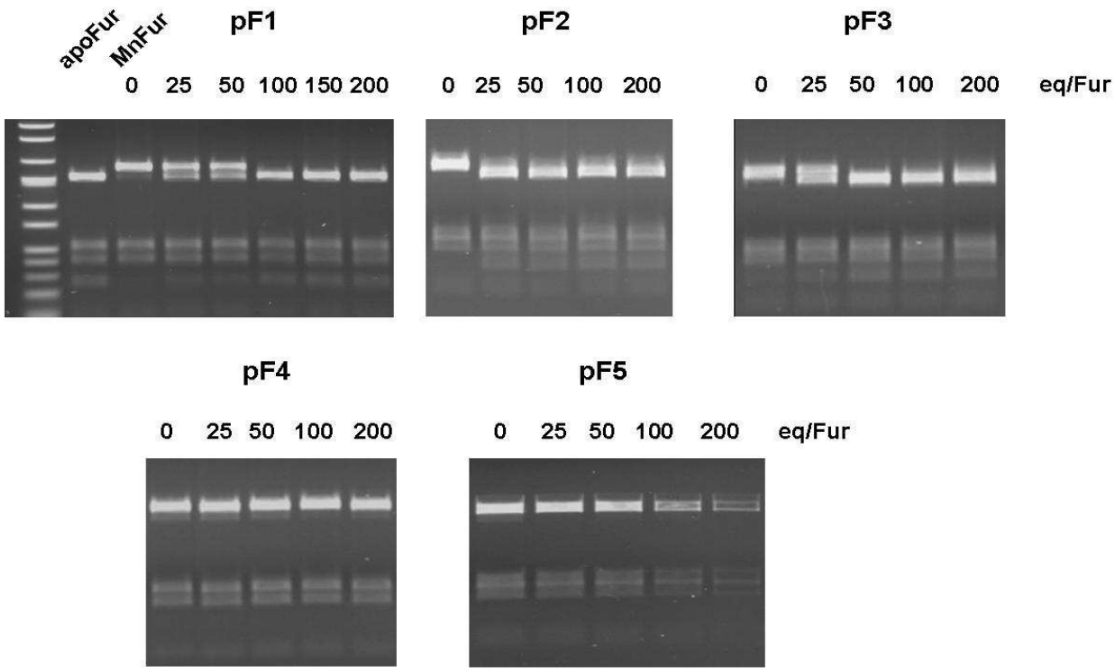


Figure II.9: Tests d'activité *in vitro* de Fur d'*E.coli* en présence des peptides linéaires pF1 à pF5. La protéine à 1 µM est métallée par 2 équivalents de Mn^{2+} par sous-unité (sel : $MnCl_2$). Différentes concentrations de peptide (0, 25, 50, 100 ou 200 équivalents par sous-unité) puis 10 nM de plasmide pDT10 sont ajoutés. Le mélange réactionnel est digéré 1 h à 37 °C avec 4 unités de HinfI puis déposé sur gel d'agarose 0,8 %. Tampon : Bis Tris Propane 100 mM pH 7,5, KCl 100 mM, $MgSO_4$ 5 mM. Les peptides pF1, pF3, pF4 et pF5 sont dissous dans le tampon, pF2 est dissous dans le DMSO. Pour chaque peptide, 3 tests indépendants ont été effectués. (Vitale 2009).

La Figure II.9 montre que les peptides pF1, pF2 et pF3 inhibent totalement la liaison d'EcFur à l'ADN respectivement à 100, 25 et 50 µM alors que pF4, inactif sur le dimère, est actif (IC_{50} = 25 µM) sur le monomère en empêchant la dimérisation (Figure II.10).

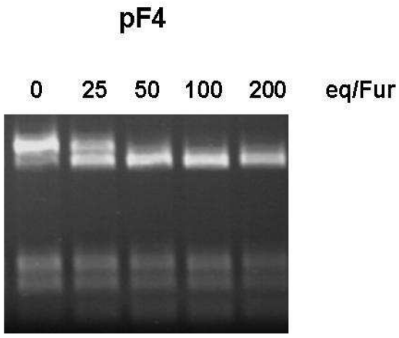


Figure II.10: Test d'activité *in vitro* du monomère de Fur d'*E.coli* en présence du peptide linéaire pF4. La protéine à 1 µM est incubée avec différentes concentrations de peptide (0, 25, 50, 100 ou 200 équivalents) puis 2 équivalents de Mn2+ (sel : MnCl2) et 10 nM de plasmide pDT10 sont ajoutés. Le mélange réactionnel est digéré 1 h à 37 °C avec 4 unités de HinfI puis déposé sur gel d'agarose 0,8 %. Tampon : Bis Tris Propane 100 mM pH 7,5, KCl 100 mM, MgSO4 5 mM. Le peptide est dissous dans le tampon. 3 tests indépendants ont été effectués.

Comme attendu, le peptide contrôle pF5 n'a aucun effet sur l'activité d'EcFur.

D'après ces travaux les peptides pF1, pF2 et pF3 ne semblent pas agir sur la métallation de la protéine EcFur, mais sur la protéine métallée et l'ordre d'ajout des réactifs n'a aucun effet sur le résultat du test (Vitale 2009).

La suite des tests d'activité a été réalisée avec des séquences tronquées ou mutées de ces peptides anti-Fur et d'une petite molécule sélectionnée par le criblage virtuel (Voir chapitre III théorique). Le but de ces tests sur ces séquences tronquées ou mutées est de déterminer les séquences minimales inhibitrices de chaque peptide et identifier les acides aminés essentiels à l'effet du peptide sur la protéine. La comparaison des modèles d'interactions théoriques avec les données expérimentales, s'il y a cohérence, donnera une idée de la zone d'interaction des peptides anti-Fur avec leur cible.

Pour réaliser tous ces tests, plusieurs synthèses peptidiques chimiques ont été réalisées en collaboration avec Didier Boturyn du Département de Chimie (DCM) moléculaire de Grenoble. Nous verrons les grandes lignes de la synthèse chimique avant de présenter les résultats de ces synthèses.

II. C| Synthèse chimique des peptides

II. C. 1| Principe

Contrairement à la biosynthèse peptidique, la synthèse peptidique chimique commence par l'acide aminé C-terminal de la séquence et se termine par le résidu N-terminal (Figure II.11).

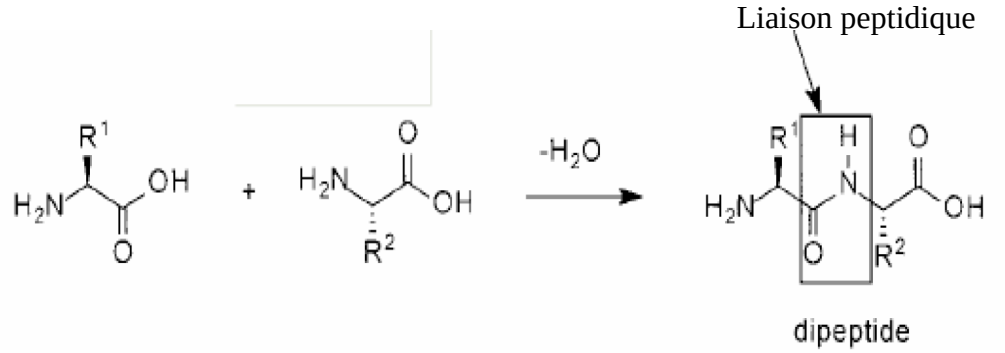


Figure II.11: Réaction de couplage peptidique

Cette synthèse peptidique chimique standard consiste en un cycle de trois étapes :

- Une étape de protection de l'amine N-terminal du premier résidu
- Une étape d'activation des acides aminés suivants
- Une étape de couplage. Cette étape peut être réalisée en solution ou sur support.

La synthèse sur support offre de nombreux avantages par rapport à la synthèse en solution. Le support étant séparé de la phase liquide par simple filtration, les étapes d'isolements et de purifications des produits sont simplifiées et les rendements peuvent être optimisés par l'utilisation d'excès de réactifs. L'excès du réactif étant facilement éliminable par filtration.

Nous avons utilisé la célèbre technique de synthèse peptidique sur phase solide (SPPS) pour obtenir les peptides linéaires utilisés dans cette étude. La SPPS, mise au point par Bruce Merrifield en 1963 (Lauréat du prix Nobel en 1984) consiste en l'addition successive d'acides aminés protégés sur un support insoluble, à partir de l'extrémité C-terminal. La procédure est expliquée sur la Figure II.12, ci-dessous.

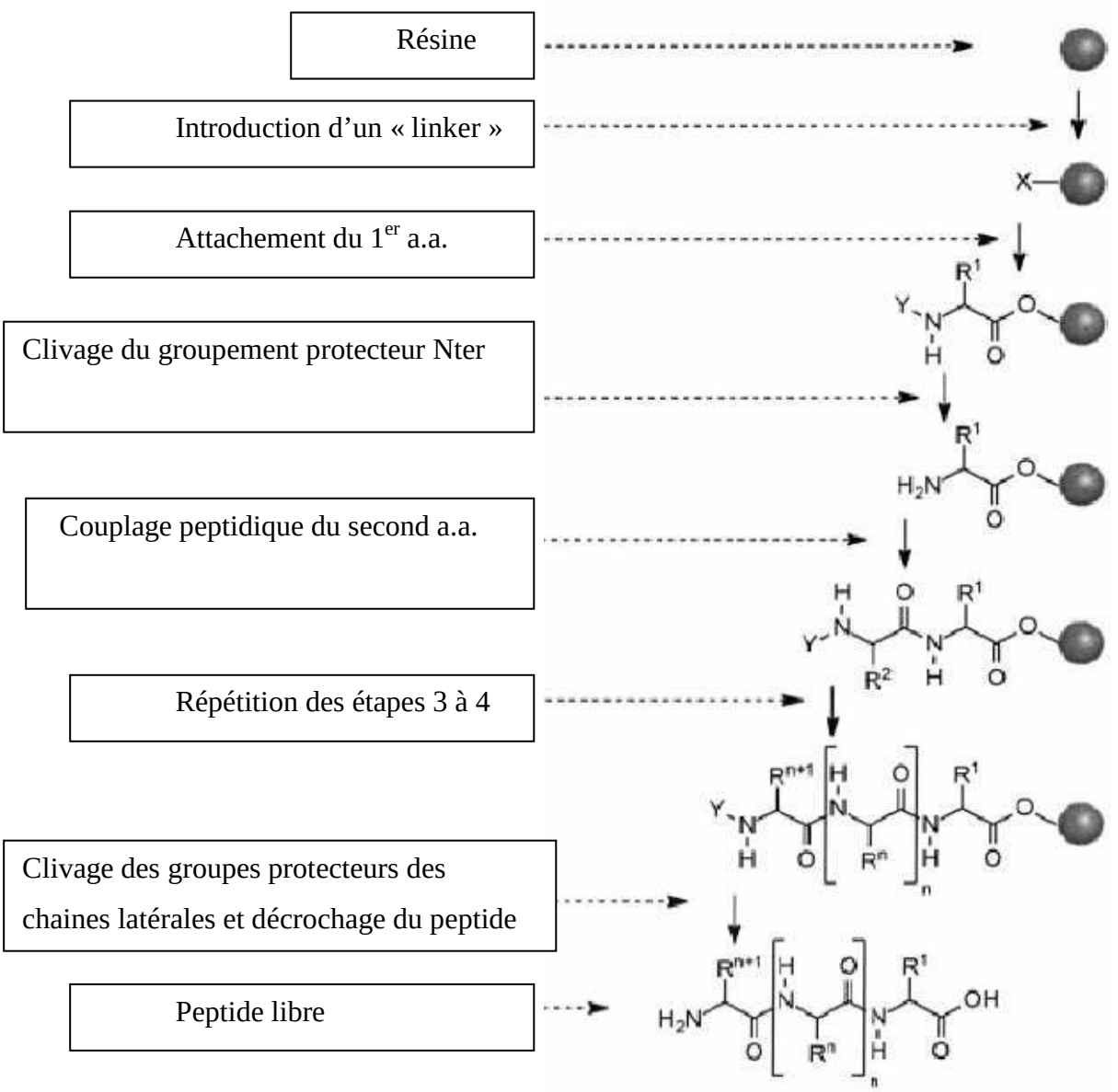


Figure II.12: Procédure générale de la SPPS. La lettre Y désigne le groupe protecteur temporaire ; R^1, R^2, R^n et R^{n+1} sont les chaines latérales des acides aminés, protégées temporairement si nécessaire par des groupements protecteurs.

Les supports solides ou résines utilisés pour accrocher les peptides, sont généralement des billes de polymères réticulés de différentes natures (PEG, Polystyrène …) possédant des bras « espaceurs » fonctionnalisés pour la liaison du premier acide aminé. Sur le marché, il existe un grand nombre de résines commerciales possédant différentes propriétés de solvatation (solvants aqueux/organiques) et diverses fonctions chimiques. Le choix du bras « espaceur » permet le décrochage du peptide (protégé ou libre) en fonction de la modification voulue sur l'extrémité C-terminal (acide, amide, hydrazide, ester méthylique, aldéhyde).

II. C. 2| L'étape de protection de l'amine libre en N-terminal

En SPPS, l'addition de chaque acide aminé s'effectue selon un cycle de deux étapes : le couplage d'un acide aminé protégé et l'élimination du groupement protecteur N-terminal. Deux stratégies sont couramment utilisées pour réaliser ces étapes. Ces deux stratégies diffèrent par le choix du groupe protecteur de l'amine libre en N-terminal et les groupements protecteurs orthogonaux utilisés pour protéger les chaînes latérales des acides aminés. La stratégie la plus utilisée reste celle dite de Fmoc/*t*Bu. Cette stratégie consiste à utiliser un groupement alcalino-labile : le Fmoc (fluorénylméthoxycarbonyle), facilement éliminable en présence de pipéridine, comme groupement protecteur de l'amine libre en N-terminal. Le mécanisme de déprotection du Fmoc est présenté sur la Figure II.13, ci-dessous. Pour des raisons d'orthogonalité, les protections des chaînes latérales sont stables en milieu basique et s'enlèvent en conditions acides par un traitement à l'acide trifluoroacétique (TFA).

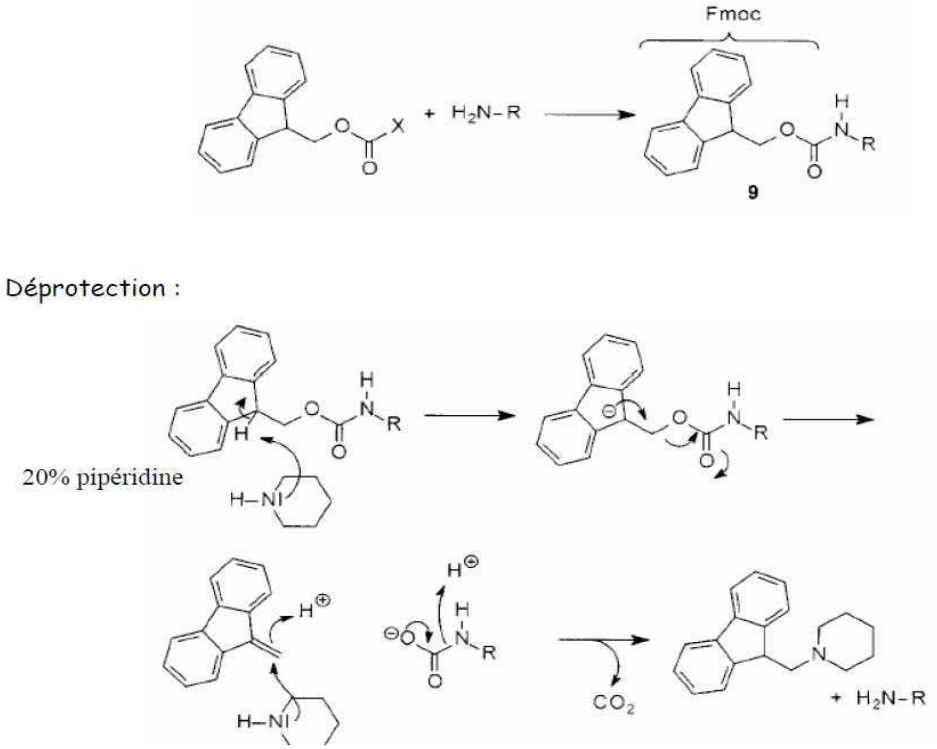


Figure II.13: Mécanismes de protection et déprotection du Fmoc

II. C. 3| L'étape d'activation

La réaction de couplage chimique entre deux acides aminés est réalisée en activant le groupe COOH du résidu à ajouter (Figure II.14).

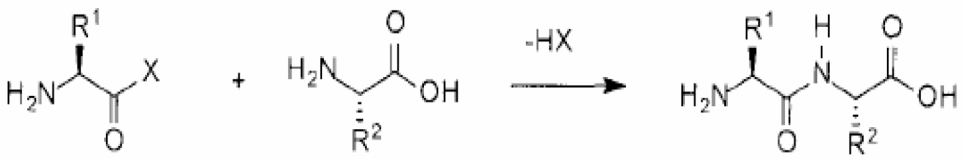


Figure II.14: Réaction de couplage avec un acide aminé activé

Pour l'étape d'activation, on utilise généralement des agents de couplage. Au cours des synthèses réalisées dans le cadre de ce travail, nous avons utilisé le benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) comme agent de couplage. Son mécanisme d'action est décrit sur la Figure II.15, ci-dessous.

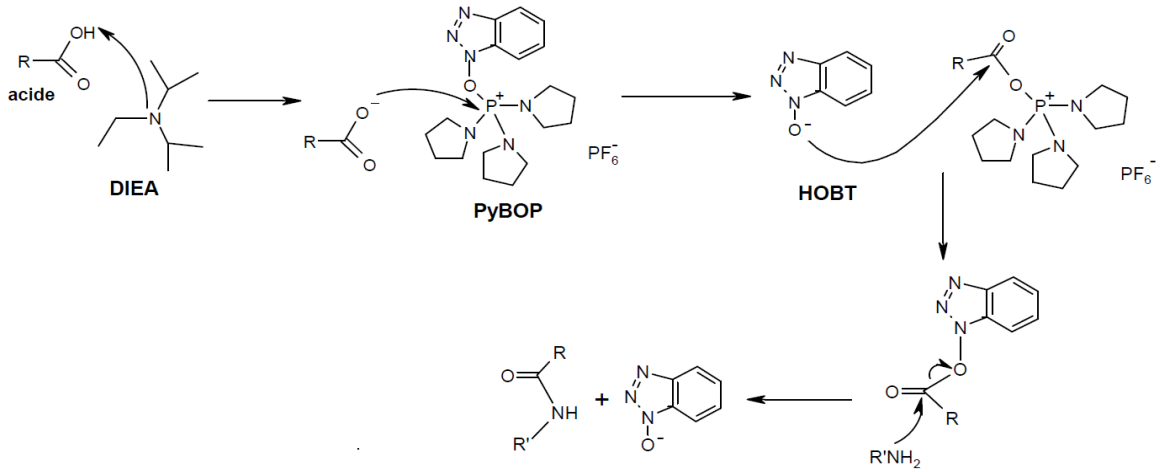


Figure II.15: Mécanisme de couplage peptidique en présence de l'agent PyBOP et de la base *N,N*-Diisopropylethylamine (DIEA).

II. C. 4| Décrochage du peptide de la résine

A la fin d'une synthèse peptidique chimique, le peptide doit être décroché du support (résine), lorsqu'on utilise la SPPS. Au cours de cette thèse, les résines Wang et 2-chlorotrityle (Cl-Trt) ont été utilisées en fonction des synthèses. La résine acido-labile Cl-Trt a été la plus utilisée et sa formule est présentée ci-dessous (Figure II.16).

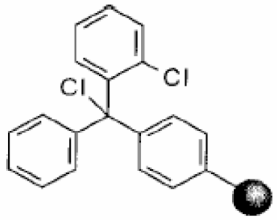


Figure II.16: Structure de la résine Chlorotrityle (Cl-Trt)

Cette résine présente l'avantage de pouvoir récupérer les peptides dans des conditions très douces (acide acétique/trifluoroéthanol/dichlorométhane) sans toucher les groupes protecteurs des chaines latérales. L'accrochage du premier résidu ne nécessite pas d'activation, ce qui limite les risques de racémisation. Le mécanisme de la réaction débute par l'attaque nucléophile du carboxylate du premier acide aminé sur la résine du carbone Sp^3 avec élimination du chlorure.

II. C. 5| Synthèse peptidique chimique en solution

Au cours de ce travail, nous avons eu à synthétiser un peptide cyclique en solution dont le protocole de synthèse est décrit ci-dessous.

La synthèse d'un peptide cyclique débute généralement par un résidu achiral comme la glycine permettant de limiter les phénomènes d'épimérisation qui peuvent apparaître lors de l'étape de cyclisation. La synthèse du peptide protégé a été réalisée au robot. Ensuite, les cycles de déprotection des amines libres en N-terminal sont effectués avec une solution de pipéridine à 20 %. Les couplages peptidiques sont réalisés au moyen d'agents de couplage comme le PyBOP en présence d'un excès de réactifs (2 équivalents d'acides aminés et d'agents de couplage). En fin de synthèse, le peptide protégé est libéré du support dans des conditions acides (TFA 1%) en gardant ces groupements protecteurs sur les chaines latérales. L'extrémité N-terminal est alors couplée à l'extrémité C-terminal dans une réaction de couplage normale. Après la cyclisation, le peptide est entièrement déprotégé au moyen d'une solution de TFA/TIS/H_2O (95/2,5/2,5). L'utilisation de l'eau est nécessaire à l'élimination du groupement éthoxyéthylidène tandis que le triisopropylsilane (TIS) est utilisé comme piégeur pour réduire les carbocations formés pendant la réaction. Le produit brut est vérifié par chromatographie liquide analytique puis purifié par chromatographie liquide préparative (confère matériels et méthodes). Les produits purifiés et lyophilisés ont ensuite été testés pour leur capacité à inhiber la protéine Fur pour sa liaison à l'ADN. Cette synthèse a été réalisée uniquement pour le peptide cyclique pF1(1-12) R1K : KLWCRYPHPPLT (confère Tableau II.3:, ci-dessous).

II. C. 6| Résultats des synthèses peptidiques

Au cours de cette thèse, nous avons eu à synthétiser plusieurs séquences peptidiques afin d'évaluer leurs effets sur l'activité de liaison de la protéine Fur à l'ADN par le test de protection à la coupure par une nucléase. Nous décrirons ce test plus tard dans ce chapitre.

La liste des peptides synthétisés et achetés est présentée dans le Tableau II.2.

Ces tableaux contiennent les notations des peptides ainsi que leurs séquences dans la colonne 1 et 2, respectivement. La notation des peptides est faite en fonction du numéro des résidus en partant du N-terminal. Exemple : le peptide pF1(1-13) désignera le peptide pF1 de 13 acides aminés et le pF2(3-10) sera la notation du peptide de huit acides aminés de pF2 allant du résidu 3 au résidu 10. Nous utiliserons ces notations tout au long du manuscrit pour désigner les peptides. Dans la colonne 2 (séquences), lorsqu'un acide aminé est muté, la position du résidu remplacé est désignée en rouge. Exemple : **RLWSRYPHP** est la séquence de pF1(1-9) où la Cystéine en position 4 est remplacée par une Serine. Cette séquence sera notée pF1(1-9) C4S. Les masses des produits après purification et lyophilisation sont données dans les tableaux. Les masses molaires (MM) théoriques ont été calculées sur le site http://web.expasy.org/protparam/ en ajoutant la masse de 114 du contre-ion TFA : CF_3COOH (MM=114) pour chaque charge positive sur les peptides (R, H, K et l'amine en N-terminal). Les masses des peptides purifiés ont été vérifiées expérimentalement par spectrométrie de masse à ionisation Electrospray (ESI-MS). Les masses des produits obtenus à la fin des synthèses et les MM sont indiqués dans le Tableau II.2. Le détail des produits achetés chez Genscript est donné au chapitre V.

Tableau II.2: Caractéristiques des peptides synthétisés au DCM et par Genscript

Tous les peptides synthétisés au DCM sont sur fond blanc, ceux achetés chez Genscript sont sur fond gris. En rouge la position d'une mutation comparée à la séquence initiale du peptide. Deux fractions ont été élués pour les séquences marquées par un (*).

Notation des Peptides	Séquences	Masse de produit pur obtenue (mg)	MM + masse TFA (114)
pF1(1-13)	RLWCRYPHPPLTD	10	2110 (=1654+4x114)
pF1(1-13)	RLWCRYPHPPLTD	10+10	2110
pF1(1-13) Y6A	RLWCRAPHPPLTD	5	2018
pF1(1-13) H8A	RLWCRYPAPPLTD	5	1930
pF1(1-13) R5A Y6A	RLWCAAPHPPLTD	5	1819
pF1(1-13) H8A T12A	RLWCRYPAPPLAD	5	1900
pF1(1-9)	RLWCRYPHP	10+5	1569
pF1(2-9)	LWCRYPHP	10	1413
pF1(3-9)	WCRYPHP	10	1300
pF1(1-9)	RLWCRYPHP	32	1569
pF1(1-9) C4S	RLWSRYPHP	31,2	1553
pF1(4-9)	CRYPHP	11,4	1000
pF1(4-9) C4S	SRYPHP	22,4	984
pF1(1-10)	RLWCRYPHPP	28	1780,5
pF1(2-9)	LWCRYPHP	57	1413
pF1(3-9) Δ8	WCRYP_P	4	1049
pF1(1-8)	RLWCRYPH	11	1586
pF1(3-9)	WCRYPHP	18,6	1300
pF1(5-13)	RYPHPPLTD	35	1437
pF1(8-13)	HPPLTD	11,5	907
pF1(1-12) R1K	KLWCRYPHPPLT	6,8	1966,8
pF1(1-12) R1K cyclique**	KLWCRYPHPPLT	7,8	1834,8
pF2(1-13)	RQCNICGASLYSY	10	1706
pF2(1-13) Y11A Y13A	RQCNICGASLYSY	5	1521

pF2(2-10)	**QCNICGASL**	**10**	**1022**
pF2(3-10)	**CNICGASL**	**10**	**894**
pF2(1-11)	**RQCNICGASLY**	**22**	**1455**
pF2(1-10)	**RQCNICGASL**	**26**	**1292**
pF2(2-10)*	**QCNICGASL**	**8,8+19,8**	**1004 et 1022**
pF2(3-10)	**CNICGASL**	**10,2**	**894**
pF2(1-10) C3S	**RQSNICGASL**	**1,2**	**1276**
pF2(1-10) C6S	**RQCNISGASL**	**en solution**	**1276**
pF2(1-10) C3S et C6S	**RQSNISGASL**	**en solution**	**1260**
pF2(1-11) Y11W	**RQCNICGASLW**	**3,8**	**1478**
pF2(3-13)	**CNICGASLYSY**	**10**	**1307**
pF3(1-13)	**ETCKCGSQVWRHS**	**10**	**1976**
pF3(1-11)	**ETCKCGSQVWR**	**2,4**	**1638**
pF3(3-13)	**CKCGSQVWRHS**	**0,5**	**1746**
pF4(1-13)	**CARCGARVNVYKY**	**10**	**1958**
pF4(1-11)	**CARCGARVNVY**	**4,2**	**1553**
pF4(3-13)	**RCGARVNVYKY**	**10**	**1784,5**
pF5(1-13)	**RSLRGRCLSQHQD**	**35**	**2124**

Plus de détails sur les caractéristiques sur ces peptides sont données au chapitre V.

II. D| Étude d'inhibition d'EcFur par les peptides synthétisés

II. D. 1| Etude d'inhibition d'EcFur par des dérivés de pF1

Le peptide pF1 est celui pour lequel le plus grand nombre de mutations ont été réalisées. Pas moins d'une vingtaine de séquences mutées ou tronquées ont a été synthétisées et testées. De plus, il faut trois expériences indépendantes réussies pour déterminer la valeur moyenne de l'IC_{50} et tous les gels ne seront pas présentés dans ce manuscrit. Au vu de ces critères, l'IC_{50} n'a pas été affiné pour toutes les séquences. De plus, les concentrations peptidiques (équivalents/Fur) sont comprises entre 0 et 150 µM. Au-delà de 150 équivalents / sous-unité de Fur, c'est à dire 150 µM en peptide dans le test nucléase quand on utilise Fur à 1 µM, on observe souvent une agrégation des échantillons dans les puits, probablement due aux problèmes de solubilité des peptides évoqués précédemment. Le Tableau II.3 présente le résumé des tests sur l'ensemble des dérivés de pF1.

Peptides	séquences	IC_{50} (µM)	Nombre de tests
pF1(1-13)	RLWCRYPHPPLTD	60±11	6
pF1(1-13) Y6A	RLWCRAPHPPLTD	inactif (*)	2
pF1(1-13) H8A	RLWCRYPAPPLTD	inactif (*)	2
pF1(1-13) R5A Y6A	RLWCAAPHPPLTD	inactif (*)	3
pF1(1-13) H8A T12A	RLWCRYPAPPLAD	inactif (*)	3
pF1(1-10)	RLWCRYPHPP	64±10	3
pF1(1-9)	RLWCRYPHP	75±5 (*)	3
pF1(1-9) C4S	RLWSRYPHP	inactif (*)	2
pF1(1-8)	RLWCRYPH	inactif	3
pF1(2-9)	LWCRYPHP	75	7
pF1(3-9)	WCRYPHP	50	6
pF1(3-9) Δ8	WCRYP_P	inactif	2
pF1(5-13)	RYPHPPLTD	inactif	2
pF1(8-13)	HPPLTD	inactif	2
pF1(1-12) R1K	KLWCRYPHPPLT	actif	3
pF1(1-12) R1K cyclique	KLWCRYPHPPLT	actif	3

Tableau II.3: Résumé des activités des dérivés mutés et/ou tronqués de pF1.

Les mesures d'IC_{50} marquées par une (*) ont été réalisées par Sylvia (Vitale 2009).

Parmi les différents dérivés de pF1 testés, seules les séquences pF1(1-13), pF1(1-10), pF1(1-9), pF1(2-9), pF1(3-9), pF1(1-12) R1K et son homologue cyclisé inhibent l'activité de liaison de la protéine à l'ADN dans les conditions des tests d'activité (équivalents/Fur <150).

Le but étant de déterminer les acides aminés clés de l'affinité du peptide pour la protéine, analysons finement les résultats de ce tableau. Les cinq premières lignes du tableau montrent que les acides aminés Y6, H8 sont importants pour inhiber la protéine. Le fait que le peptide pF1(1-9) inhibe la protéine montre que les quatre résidus C-terminal de pF1 ne sont pas essentiels à l'effet du peptide sur la protéine. Notons au passage que c'est le peptide qui a permis de faire avancer les simulations d'amarrage de pF1 sur la protéine. Nous reviendrons sur ce point dans la partie théorique. L'effet de ce peptide démontre que T12 n'est pas essentielle à l'activité du peptide pF1. De la même manière, l'effet de pF1(3-9) montre que les deux résidus N-terminal ne sont pas non plus indispensables pour inhiber la protéine. Aucune mutation dans ce tableau ne montre que le résidu R5 est important. En revanche, le tryptophane W3 est indispensable étant donné que pF1 (4-9) est inactif. La cystéine C4 est importante car pF1(1-9) C4S n'est plus actif. De façon surprenante, la proline P9 est importante car pF1(1-8) n'est pas actif contrairement à pF1(1-10) (Figure II.17).

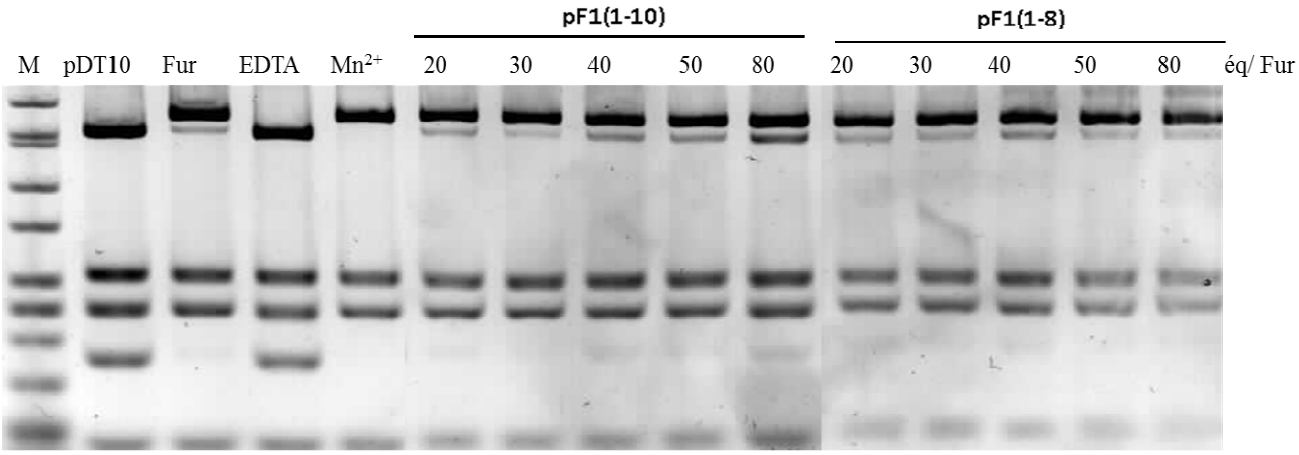


Figure II.17: Tests d'activité de Fur d'*E.coli* en présence de pF1(1-10) et de pF1(1-8). M est le marqueur de poids moléculaire. La protéine à 1 µM est inhibée par 1 mM d'EDTA (contrôle négatif) et métallée par 2 équivalents de Mn^{2+} par monomère d'EcFur. Différentes concentrations des peptides en équivalent de protéine (éq/Fur) puis 10 nM de plasmide pDT10 sont ajoutés. Le mélange réactionnel est digéré 1 h à 37 °C par 4 unités de HinfI puis déposé sur gel d'agarose 1 %. Tampon : Bis Tris Propane 100 mM pH 7,5, KCl 100 mM, $MgSO_4$ 5 mM. Les peptides sont dissous d'abord dans le DMSO puis dilués dans le tampon. Dans ce test, le peptide pF1(1-10) inhibe partiellement la protéine à 80 équivalents alors que pF1(1-8) n'inhibe pas.

De même, la troncation de la proline P10 diminue l'IC_{50}. De manière intéressante, pF1(3-9) est la séquence qui possède l'IC_{50} le plus faible, donc celle qui interagit la plus favorablement avec la protéine. Dans les conditions de ces tests où les peptides sont dilués dans le DMSO, pF1(3-9) de sept acides aminés est donc la séquence minimale active sur la protéine. La Figure II.18 présente deux profils de migration de ce peptide. La Figure II.18A montre un gel où le peptide inhibe 50% de l'activité de la protéine à 50 µM. Sur la Figure II.18B, on observe une inhibition totale à 100% de la protéine. On peut estimer l'IC_{50} de pF1(3-9) à 50 µM dans les conditions de ce test. Ce qui serait la plus active des séquences de pF1 (Figure II.18).

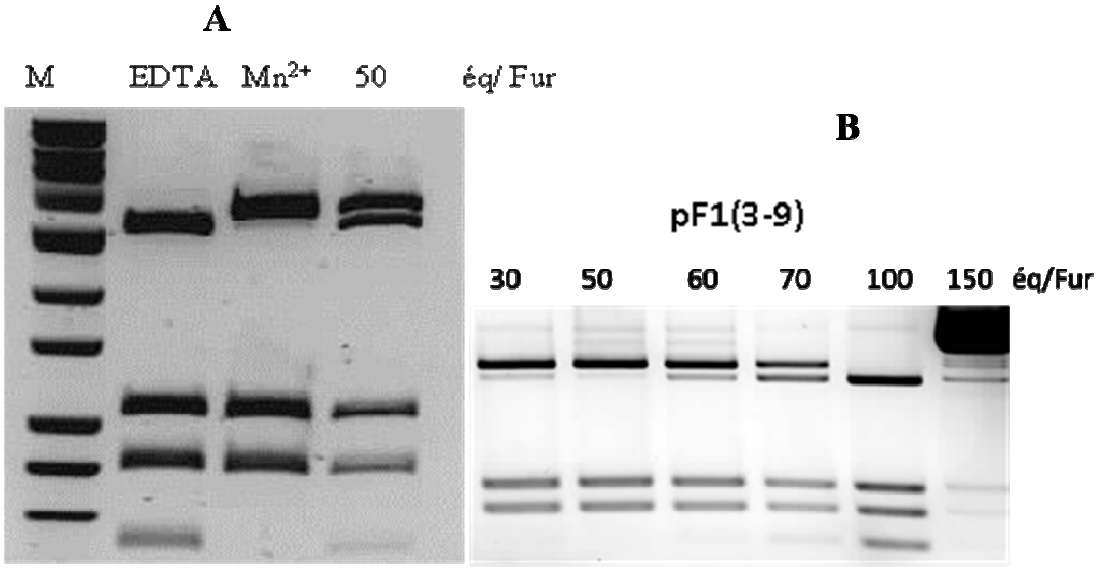


Figure II.18 : Tests d'activité de Fur d'*E.coli* en présence de pF1(3-9). M est le marqueur de poids moléculaire. La protéine à 1 µM est inhibée par 1 mM d'EDTA (contrôle négatif) et métallée par 2 équivalents de Mn^{2+} par monomère d'EcFur. Différentes concentrations des peptides en équivalent de protéine (éq/Fur) puis 10 nM de plasmide pDT10 sont ajoutés. Le mélange réactionnel est digéré 1 h à 37 °C par 4 unités de HinfI puis déposé sur gel d'agarose 1 %. Tampon : Bis Tris Propane 100 mM pH 7,5, KCl 100 mM, $MgSO_4$ 5 mM. Les peptides sont dissous d'abord dans le DMSO puis dilués dans le tampon. Gel présenté ici en contraste inversé.

Obtenir des séquences minimales actives plus actives que les peptides linéaires est très intéressant pour la suite du projet. Cela faciliterait les modèles de prédiction des amarrages des peptides sur la protéine, comme nous allons le voir dans le chapitre III. Afin de confirmer et d'affiner les valeurs d'IC_{50} obtenus sur les séquences de pF1, une série d'expériences a été réalisé récemment par Sophie Mathieu, actuellement en post-doc au laboratoire, en essayant d'autres conditions. La Figure II.19, montrée ci-dessous, présente les profils de migration de deux gels en présence de pF1(2-9) et pF1(3-9) achetés chez Genscript et dissous dans le tampon BTP en utilisant 2 µM de Fur et 8µM de Mn2+ et 2 unités de Hinf1.

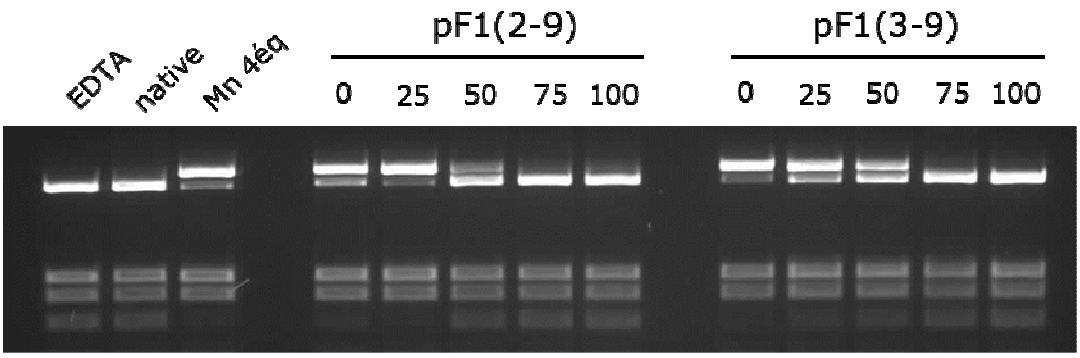


Figure II.19: Tests d'activité de Fur d'*E.coli* en présence de pF1(2-9)de pF1(3-9). La protéine à 2 µM est inhibée par 1 mM d'EDTA (contrôle négatif) et métallée par 4 équivalents de Mn^{2+} par monomère d'EcFur. Différentes concentrations des peptides en équivalent de protéine (éq/Fur) puis 10 nM de plasmide pDT10 sont ajoutés. Le mélange réactionnel est digéré 1 h à 37 °C par 2 unités de Hinf1 puis déposé sur gel d'agarose 1 %. Tampon : BTP 100 mM pH 7,5, KCl 100 mM, $MgSO_4$ 5 mM. Les peptides ont été dissous dans le BTP.

Dans ce test, le peptide pF1(2-9) est plus actif que celui de pF1(3-9). Ce gel montre que l'IC_{50} de pF1(2-9) peut être estimé à 35µM tandis que celui de pF1(3-9) est plus proche 45 µM. Ces résultats sont plus cohérents avec ceux des amarrages moléculaires (confère chapitre III) qui montrent que l'énergie libre de liaison de pF1(3-9) est supérieure que celle de pF1(2-9). Ces résultats démontrent et confirment aussi que pF1(3-9) est la séquence minimale active de pF1. Il sera intéressant de muter le résidu R5 afin de montrer son effet sur l'activité de cette séquence.

Ces valeurs d'IC_{50} du peptide pF1 linéaire sont relativement très faibles comparées à celles des affinités estimées pour l'aptamère F1, dont est issu le peptide. Afin de mimer la conformation du peptide pF1 telle qu'on l'imagine dans l'aptamère, la séquence pF1(1-12) R1K a été synthétisée dans le but de contraindre le peptide. Etant donné que pF1(3-9) est actif et que les quatre résidus C-terminal ainsi que les résidus R1 et L2 de pF1 ne sont pas indispensables pour inhiber l'activité de la protéine, la séquence pF1(1-12) R1K cyclisée entre l'amine de la chaine latérale de K1 et le COOH de T12 était une séquence raisonnable pour commencer les tests. La Figure II.20 illustre un modèle de ce peptide cyclisé. Ce modèle a été construit à partir de la conformation amarrée de pF1 sur la protéine Fur (confère chapitre III).

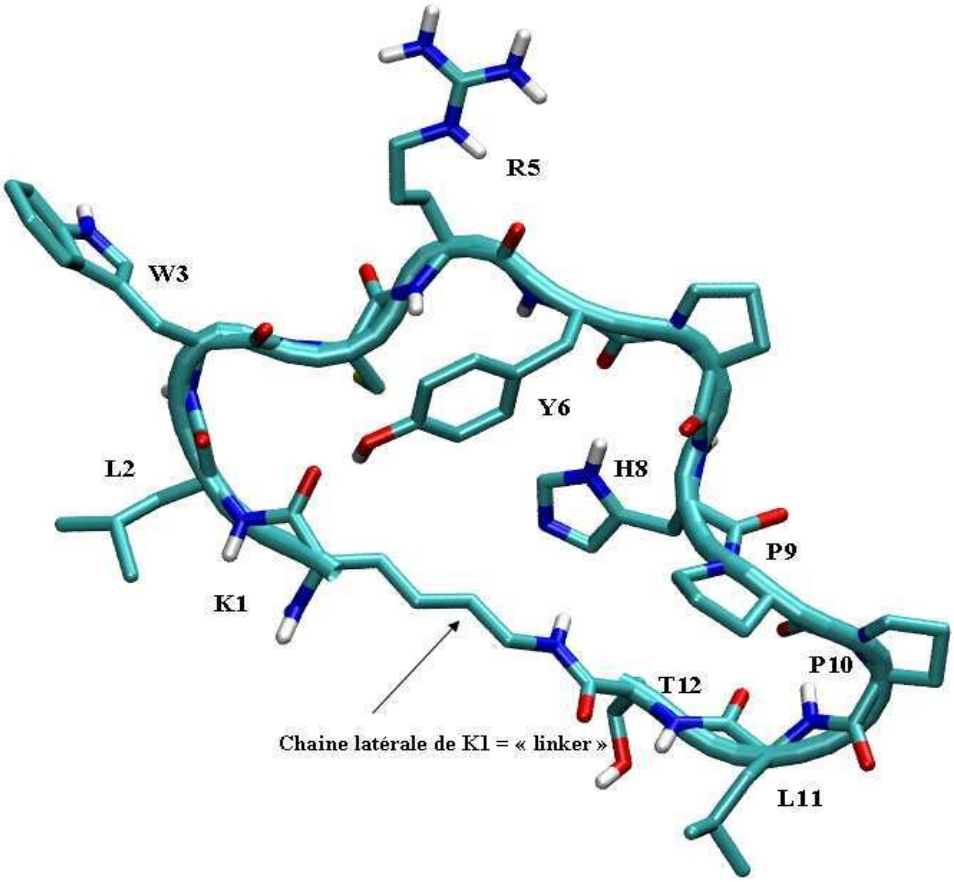


Figure II.20: Modèle du peptide pF1(1-12) R1K cyclisé

Les séquences pF1(1-12) R1K linéaires et pF1(1-12) R1K cyclisé, synthétisées au DCM, ont été testées. La Figure II.21 montre les profils des tests d'activité de pF1(1-13), pF1(1-12) R1K et pF1(1-12) R1K cyclisé. Ces tests réalisés par Sophie Mathieu avec 1,5µM de Fur et une unité de Hinf1 (nouvelle enzyme).

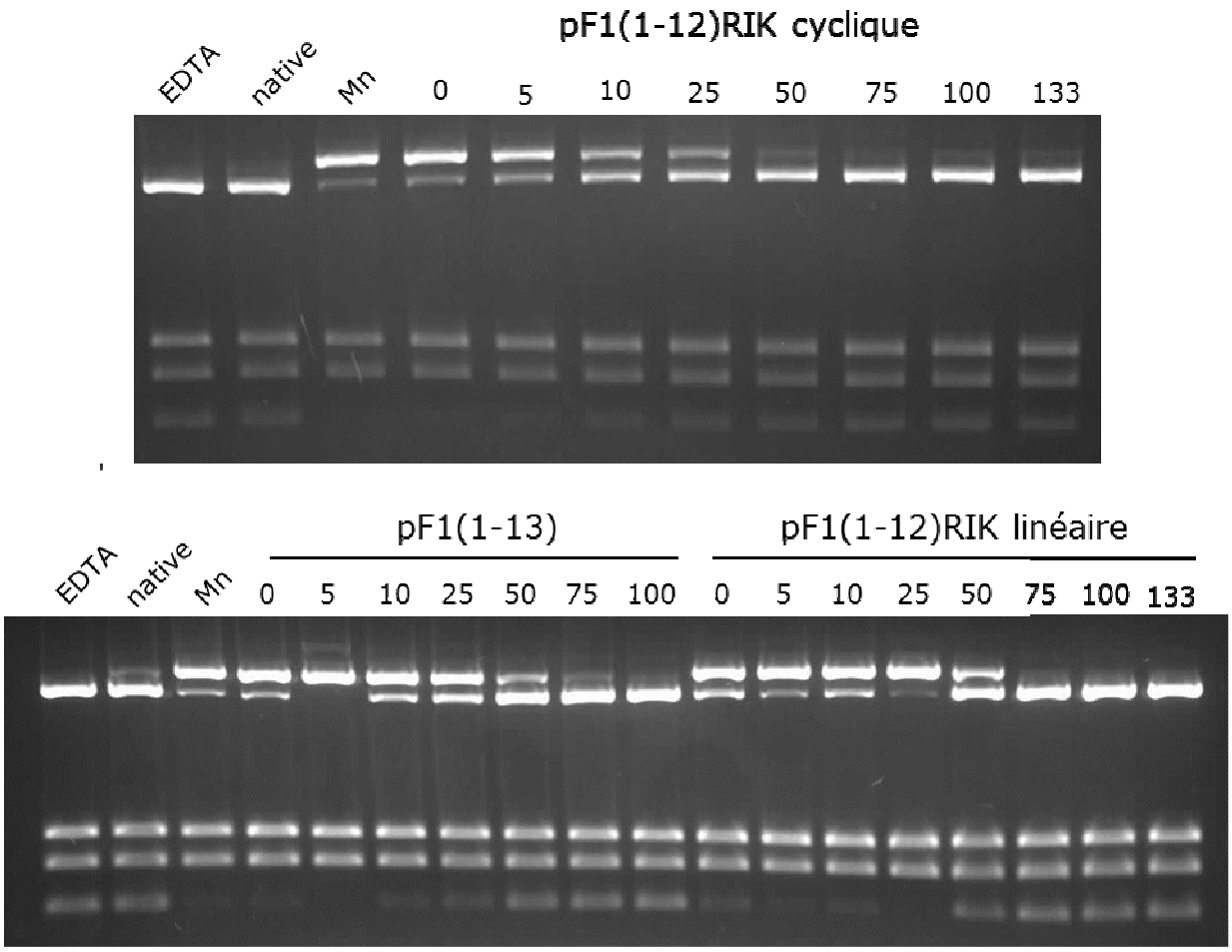


Figure II.21: Tests d'activité de Fur d'*E.coli* en présence de pF1(1-12) R1K cyclisé, pF1 (1-13) et pF1(1-12) R1K linéaire. La protéine à 1,5 µM est inhibée par 1 mM d'EDTA (contrôle négatif) et métallée par 4 équivalents de Mn^{2+} par monomère d'EcFur. Différentes concentrations des peptides en équivalent de protéine (éq/Fur) puis 10 nM de plasmide pDT10 sont ajoutés. Le mélange réactionnel est digéré 1 h à 37 °C par une unité de Hinf1 puis déposé sur gel d'agarose 1 %. Tampon : BTP 100 mM pH 7,5, KCl 100 mM, $MgSO_4$ 5 mM. Les peptides ont été dissous dans le BTP.

Dans ce test où les conditions sont les mêmes, on peut comparer l'effet des trois peptides sur l'activité de la protéine (Figure II.21). Le peptide pF1(1-12) R1K cyclique semble plus actif que le peptide pF1(1-13) linéaire qui est aussi plus actif que la séquence pF1(1-12) R1K linéaire. On peut estimer dans ces conditions des IC_{50} de l'ordre de 10 µM pour le peptide cyclique, de 40 µM pour le peptide pF1(1-12)R1K linéaire et de 30 µM pour le peptide pF1(1-13) Ces résultats montrent que le peptide pF1(1-12) R1K cyclique est 3 à 4 fois plus actif que les séquences linéaires. La cyclisation est donc bien une voie intéressante à suivre pour améliorer l'activité de ces peptides. La cyclisation devrait permettre aussi de stabiliser les peptides et de les protéger de l'action des exo-protéases.

De façon surprenante, les peptides pF1(1-9) et pF1(1-13) qui ont été synthétisés au DCM n'inhibent pas la protéine. Or, ces mêmes séquences achetées chez Genscript inhibent la liaison d'EcFur à l'ADN. Les peptides pF1(1-9) et pF1(1-13) (synthétisés et achetés) ont donc été analysés par HPLC analytique, par spectrométrie de masse à ionisation Electrospray (ESI-MS) et par séquençage. Les spectres ESI-MS ont montré que les peptides pF1(1-9) et pF1(1-13) synthétisés au DCM sont en présence de deux autres produits secondaires inconnus dans la solution analysée alors que celui de Genscript est seul. La pureté donnée par Genscript est de 96% (voir chapitre V matériel et méthodes) alors que pF1(1-9) et pF1(1-13) purifiés au DCM contiennent respectivement 27 et 26% d'abondance relative d'autres produits secondaires. En effet, quand on observe les spectres HPLC ci-dessous, on peut se rendre compte que le pic au temps de rétention RT = 6,53 minutes correspondant à pF1(1-13) possède un épaulement pour le peptide pF1(1-13) du DCM qu'il aurait été difficile de séparer contrairement à celui de Genscript dont le pic majoritaire sort à RT= 6,53 minutes avec un petit pic d'impureté à RT = 6,85 minutes (Figure II.22). Les peptides pF1(1-9) et pF1(1-13) purifiés au DCM contiendraient donc un ou deux produits secondaires difficilement séparables des produits espérés, probablement à l'origine de cette différence observée avec celui de Genscript. Ces résultats nous ont poussés à être plus vigilants sur l'étape de purification en préparatif des peptides en ne collectant que les sommets de pics pour obtenir des produits purs.

Le séquençage (réalisé par Jean-Pierre ANDRIEU, du laboratoire d'Enzymologie Moléculaire de l'Institut de Biologie Structurale IBS) a montré qu'une grande proportion de tous ces peptides contient des cystéines non libres et donc probablement impliquées dans des ponts disulfures. La séquence de pF1 ne contenant qu'une cystéine C4, on peut dire que ces peptides sont pour la plupart sous forme dimérique. L'activité des aptamères dont sont issus les peptides linéaires de 13 acides aminés nous indique que c'est une seule séquence de pF1(1-13) qui inhibe Fur comme nous l'avons vu dans le Tableau II.3:. D'après ces données, on peut penser que la proportion de dimères formés par ponts disulfures ainsi que la pureté jouent un rôle important sur l'activité inhibitrice de ces peptides. Dans l'avenir, il faudra être plus vigilant sur l'oxydation et la pureté des peptides synthétisés.

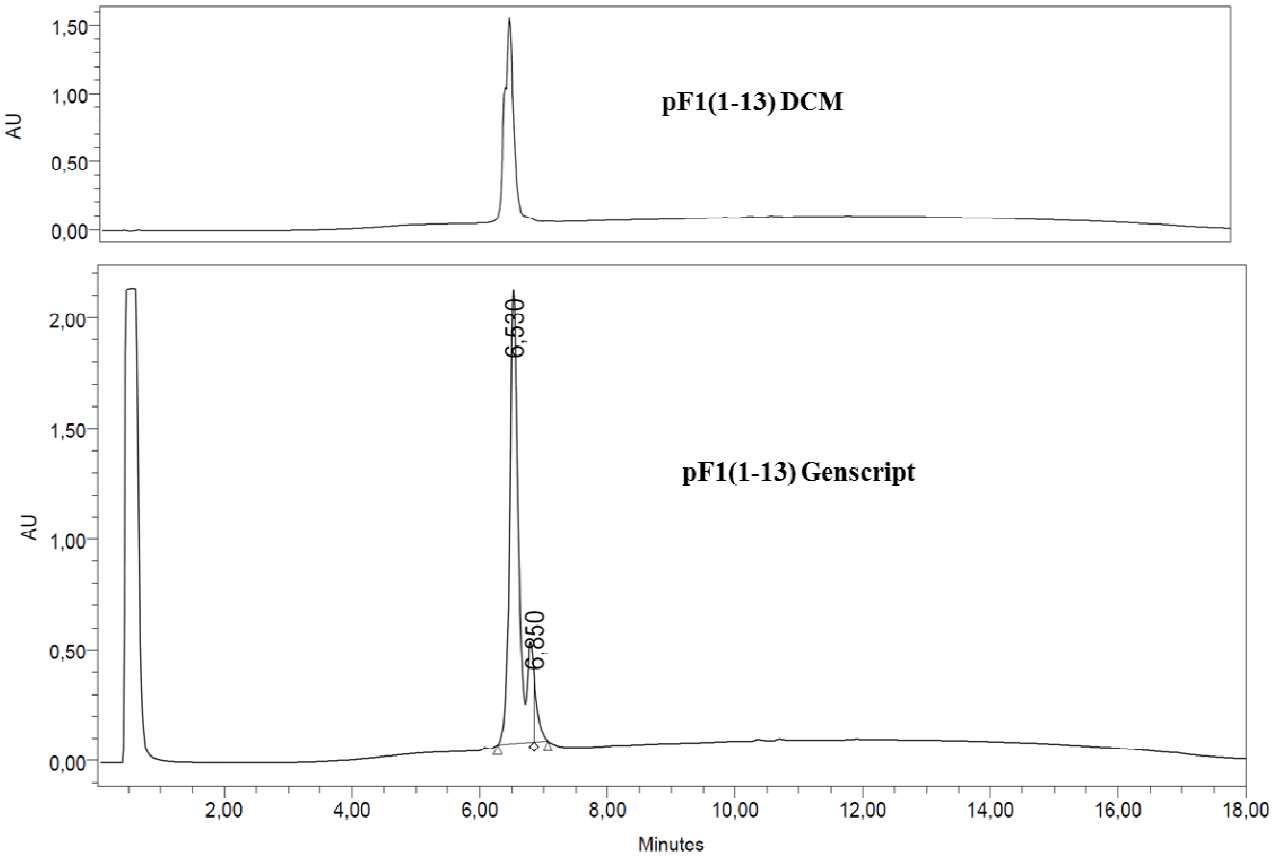


Figure II.22: Spectres HLPC de pF1(1-13) synthétisé au DCM et celui acheté chez Genscript. L'unité arbitraire (AU) montre l'évolution de l'absorption à 214 nm, spécifiques des liaisons peptidiques. Le pic majoritaire du peptide pF1(1-13) DCM possède un épaulement tandis que celui de Genscript possède 2 pics distincts. Colonne analytique *Nucleosil 120 Å 3 µm C18 particles, 30 x 4 mm*^{2} opérant à un flux de 1,3 mL.min^{-1} avec un gradient de 5 à 100% de solvant B (solution d'acétonitrille (CH_3CN) contenant 9,9% d'eau HPLC et 0,1% de TFA) pendant 15 min. Le solvant A est une solution d'eau HPLC contenant 0,1% de TFA.

II. D. 2| Etude d'inhibition d'EcFur par des dérivés de pF2

Nous savions que le peptide pF2 était le plus actif de tous les peptides dérivés des aptamères peptidiques. Cela s'est confirmé avec la technique du double hybride qui permet de visualiser des interactions *in vivo* avec les peptides avec une constante de dissociation (Kd) de l'ordre du µM au nM (Vitale 2009). Les séquences pF2(1-13) et pF2(1-13) Y11 A Y13A inhibent Fur avec des IC_{50} de 12,5 µM et de 15,5µM, respectivement. Les acides aminés Y11 à Y13 ne sont donc pas essentiels à l'activité du peptide. Le peptide pF2 est donc très intéressant puisqu'il présente la meilleure affinité pour EcFur parmi les peptides entiers, dérivés des aptamères peptidiques F1 à F4.

L'ensemble des dérivés mutés ou tronqués de pF2 testé dans cette étude est présenté dans le Tableau II.4, ci-dessous.

Peptides pF2	Séquences	IC_{50} (µM)	MW (g/mol)	Nombre de tests
pF2(1-13)	RQCNICGASLYSY	12,5±2 (*)	1706	5
pF2(1-13) Y11A Y13A	RQCNICGASLASA	15,5 (*)	1521,5	3
pF2(1-11)	RQCNICGASLY	7±2	1455	7
pF2(2-10)	QCNICGASL	75±5	1022	7
pF2(3-10)	CNICGASL	≥75	894	5
pF2(1-10)	RQCNICGASL	% d'Inhibition à 6 éq/Fur	1780,5	6
pF2(1-10) C3S	RQSNICGASL	inactif	1276	3
pF2(1-10) C6S	RQCNISGASL	Inactif	1049	1
pF2(1-10) C3S et C6S	RQSNISGASL	inactif	1049	1
pF2(1-11) Y11W	RQCNICGASLW	8±2	1586	3

Tableau II.4: Séquences et activité des variants mutés et/ou tronqués du peptide pF2. Les colonnes des IC_{50} marquées par une (*) ont été réalisées par Sylvia (Vitale 2009).

Les résultats montrent le peptide pF2(1-10) tronqué des résidus C-terminal Y11, S12 et Y13 est le plus actif. Ce peptide inhibe la protéine à 6 équivalents/Fur soit 6 µM (Figure II.23).

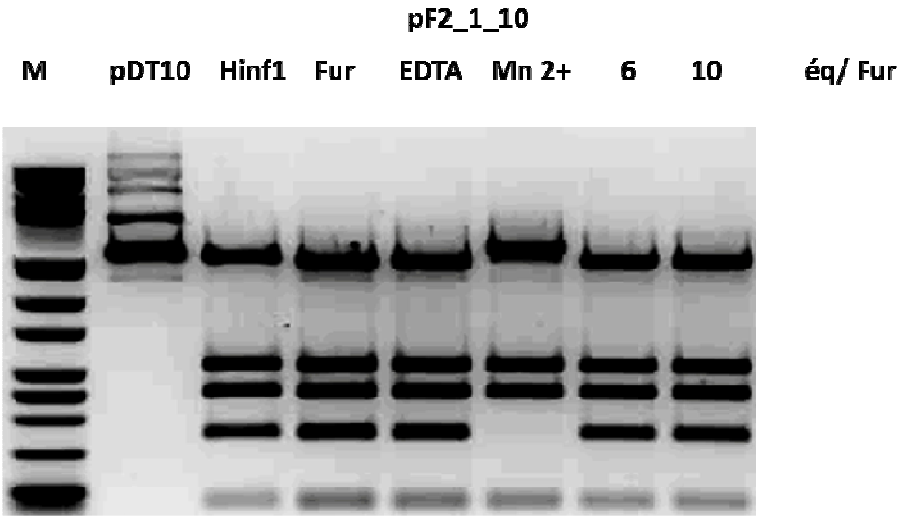


Figure II.23: Tests d'activité de Fur d'*E. coli* en présence de pF2(1-10). M est le marquer de poids moléculaire, pDT10 est le puits du plasmide seul et Hinf1 montre la digestion de pDT10. La protéine à 1 µM est inhibée par 1 mM d'EDTA (contrôle négatif) et métallée par 2 équivalents de Mn^{2+} par monomère d'EcFur. Différentes concentrations des peptides en équivalent de protéine (éq/Fur) puis 10 nM de plasmide pDT10 sont ajoutés. Le mélange réactionnel est digéré 1 h à 37 °C par 4 unités de Hinf1 puis déposé sur gel d'agarose 1 %. Tampon : BTP 100 mM pH 7,5, KCl 100 mM, $MgSO_4$ 5 mM. Les peptides ont été dissous dans le DMSO. Le peptide inhibe totalement la protéine à 6 et 10 éq/Fur. Photo de gel présenté en contraste inversé.

Bien évidemment, il faudra réaliser d'autres tests afin de déterminer l'IC_{50}. Ce résultat confirme que les trois derniers résidus C-terminal de pF2 ne sont pas indispensables à l'activité inhibitrice du peptide. Mieux encore, la troncation de ces résidus améliore l'affinité du peptide pour la protéine (Tableau II.4, ci-dessus). On passe d'un peptide de 13 acides aminés à une séquence de 10 résidus en améliorant l'activité du peptide. Ces résultats vont dans le même sens que les tests sur les séquences pF2(1-11) et pF2(1-11) Y11W qui inhibent Fur respectivement avec un IC_{50} de 7 et 8 µM (Tableau II.4).

La séquence pF2(1-11)Y11W a été synthétisée pour faire la spectroscopie de fluorescence du tryptophane (Reshetnyak *et al.,* 2001) qui permet de renseigner sur l'environnement du peptide. Cette méthode nous avait permis de savoir que le peptide pF1, qui possède un tryptophane, interagit avec la forme métallée de Fur. La tyrosine Y11 de pF2 n'étant pas indispensable à l'activité du peptide, elle a été remplacée par une tryptophane dans la séquence pF2(1-11)Y11W. Les résultats du test montre que ce peptide est aussi actif que la séquence pF2(1-11) (Tableau II.4). Il sera donc très intéressant d'étudier la fluorescence du tryptophane de ce peptide en présence de protéine métallée ou non afin de voir avec quelle forme d'EcFur pF2 interagit. Les résultats permettront d'améliorer les prédictions d'amarrage de ce peptide, qui ne semble pas aller dans le même sens que ceux de la partie expérimentale (confère chapitre III).

Tandis que les résidus C-terminal de pF2 améliorent l'affinité du peptide pF2, les résidus N-terminal semblent jouer un rôle non négligeable. En effet, les séquences tronquées des résidus R1, Q2 font augmenter les valeurs d'$IC_{50.}$ L'IC_{50} trouvé pour pF2(2-10) est encore de 75 µM, soit moins bon que celui de pF2(1-10). Ainsi, tandis que pF2(1-10) inhibe totalement la protéine à 6 µM les séquences pF2(2-10) et pF2(3-10) inhibent 50% de l'activité de la protéine à des concentrations supérieures ou égales à 75µM (Tableau II.4).

De plus, les mutations ponctuelles ou doubles des résidus C3 ou C6 de pF2(1-10) montrent que ces résidus jouent un rôle important dans l'inhibition de la protéine. Nous avions vu précédemment que pF2 pouvait lier le Co^{2+} et probablement les métaux de la protéine pour l'inhiber. Mais, sachant que les aptamères peptidiques ont été sélectionnés pour leurs interactions avec EcFur en double hybride ce n'est pas une simple compétition pour la liaison de l'ion métallique. Plusieurs hypothèses sont possibles pour expliquer l'effet de pF2 sur EcFur :

- Soit le peptide forme des ponts disulfures intramoléculaires par l'intermédiaire des résidus C3 et C6, conférant une conformation particulière au peptide pour inhiber la protéine. Ceci est peu probable si on en croit les modèles construits avec pF2 ponté par un pont disulfure

intramoléculaire C3-C6 qui ne sont pas favorables pour interagir avec la protéine (confère chapitre III).

- Soit le peptide interagit directement de façon plus forte avec la protéine quand on a des cystéines et pas avec des sérines. Cela voudrait dire que ce sont les caractères polarisables et plus facilement déprotonables des cystéines qui seraient essentiels pour l'interaction. Il sera difficile de lever cette ambiguïté en comparant ces résultats avec ceux des amarrages moléculaires car l'effet des métaux n'a pas été bien pris en compte dans les simulations pour les peptides. Il sera donc important de déterminer sur quelle forme de la protéine ce peptide est actif par le suivi de la fluorescence du tryptophane en utilisant la séquence pF2(1-11) Y11W synthétisée dans ce but. Il sera très intéressant de faire des essais de co-cristallisation de la protéine avec ce peptide qui semble interagir très fortement avec Fur. Cette interaction pourrait favoriser la stabilisation des cristaux et permettre la résolution de la structure de la protéine et mieux comprendre le mécanisme d'inhibition de ce peptide.

II. D. 3| Etude d'inhibition d'EcFur par des dérivés de pF3 et pF4

Pour ces deux peptides, seulement deux double-mutations ont été synthétisées pour le moment. Il s'agit des mutations des deux acides aminés N-terminal et C-terminal de la séquence qui ont été testées par Sylvia (Vitale 2009) (Tableau II.5).

Peptide	**séquences**	**IC_{50} (µM)**	**MW (g/mol)**	**Nombre de tests**
pF3(1-13)	**RQCNICGASLYSY**	**30±17**	**1706**	**5**
pF3(1-11)	**RQCNICGASLY**	**<50**	**1455**	**7**
pF3(1-13)	**RQCNICGASLYSY**	**30±17**	**1706**	**5**
pF3(3-13)	**CKCGSQVWRHS**	**Inactif**	**1307**	**2**
pF4(1-13)	**CARCGARVNVYKY**	**35,5±12**	**1022**	**7**
pF4(3-13)	**RCGARVNVYKY**	**Inactif**	**1307**	**2**
pF4(1-11)	**CARCGARVNVY**	**Inactif**	**1307**	**2**

Tableau II.5: Séquences et activité des peptides pF3 et pF4

Les peptides pF3(1-13) et pF4(1-13) sont actifs avec des IC_{50} respectifs de 30 et 35 ± 15 µM. Le peptide pF4 n'est pas actif sur le dimère, mais interagit avec le domaine de dimérisation

d'un monomère pour inhiber la dimérisation de la protéine (Vitale 2009). De plus, ce peptide interagit avec des protéines Fur, autres que celle d'EcFur comme nous le verrons au chapitre III (Vitale 2009). Malheureusement, les simulations d'amarrages n'ont pas pu été réalisées avec ce peptide, très intéressant. La structure du domaine de dimérisation du monomère d'EcFur n'étant pas stable d'après les analyse RMN (Pecqueur *et al.*, 2006), il ne sera pas possible d'effectuer des simulations d'amarrage avec pF4. Il faudra effectuer des mutations séquentielles à l'extrémité N-terminal et C-terminal, comme il a été fait avec pF1 et pF2.

Le peptide tronqué pF3(1-11) semble actif (Figure II.24).

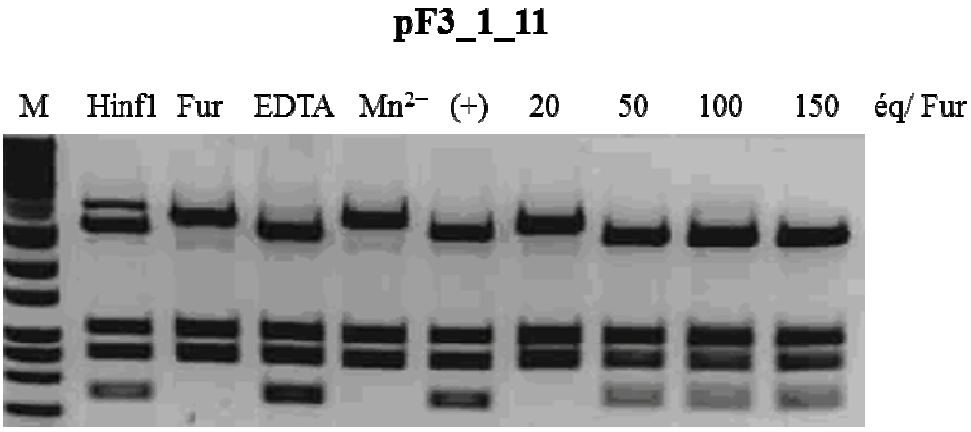


Figure II.24: Test d'activité de Fur avec pF3(1-11). pF2(1-10) (10 µM) est utilisé comme contrôle positif (+) de l'inhibition par un peptide**.** Les conditions du test sont identiques à celle de la Figure II.18.

Ce gel montre que pF3(1-11) inhibe totalement la protéine à 50 µM tandis qu'à 20µM, on observe aucune inhibition. Ce qui montre que les résidus S12 et Y13 ne sont pas indispensables à l'activité de pF3. Ce résultats va dans le même sens que les modèles de prédiction de pF3 (confère chapitre III) qui montrent que pF3(1-10) a une bonne affinité théorique pour EcFur. Plusieurs simulations d'amarrage moléculaires ont donc été réalisées avec des séquences tronquées de pF3 afin de prédire les mutations futures à réaliser sur ce peptide (confère chapitre III).

II. D. 4| Etude d'inhibition d'EcFur par l'ester du benzyl-tryptophane : W_CO_Bz

Les criblages virtuels que nous verrons dans la partie théorique ont permis d'identifier de petites molécules pouvant interagir et inhiber la protéine EcFur. Parmi ces molécules, nous avons commandé l'ester du benzyl-tryptophane noté ici W-CO-Bz (Figure II.25), qui avait la meilleure

affinité théorique pour EcFur lors du premier criblage que nous avons réalisé à Tromso (Norvège) avec le docteur Rafi Ahmad dans le cadre d'une collaboration (chapitre III).

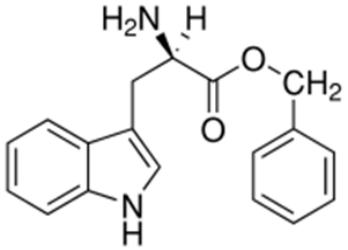


Figure II.25: Structure de l'ester du Benzyl-tryptophane

Le test d'activité de protection a été utilisé pour tester l'effet de la molécule sur la protéine EcFur. Le profil de deux gels est présenté sur la Figure II.26, ci-dessous.

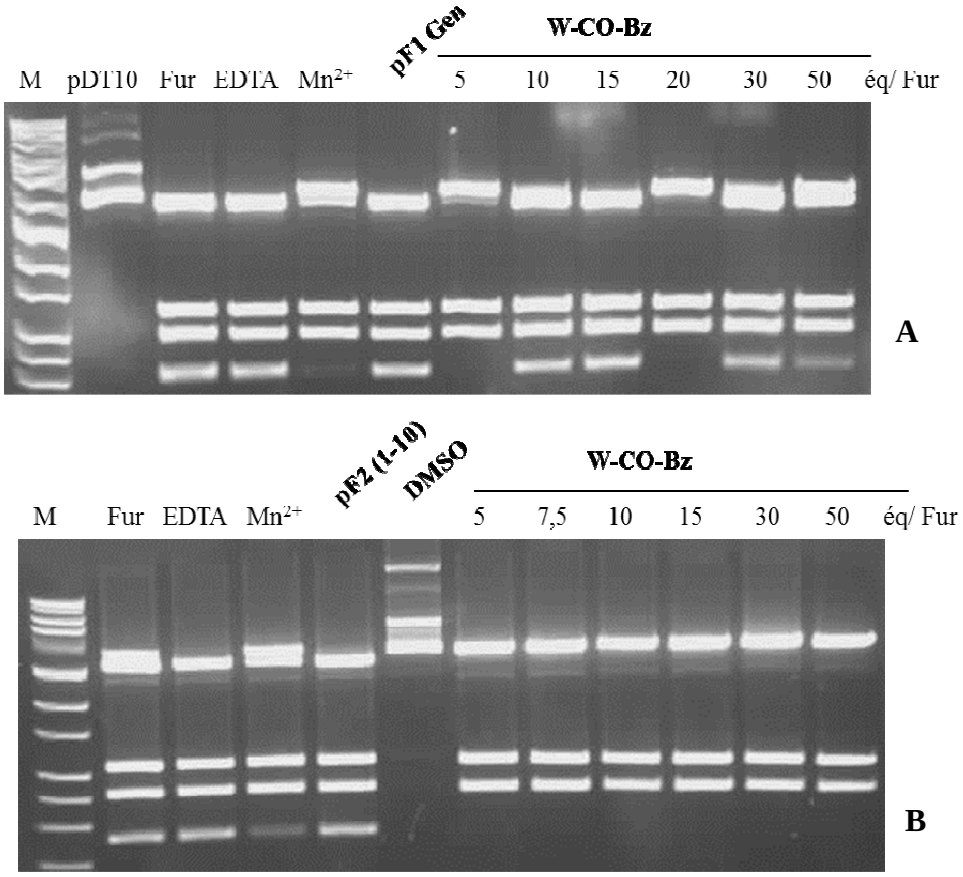


Figure II.26: Test d'activité de Fur avec le W-CO-Bz. Les peptides pF1(1-13) 100 µM et pF2(1-10) 50 µM sont utilisés comme contrôles positifs d'inhibition respectivement en A et B. M est le marqueur de poids moléculaire. La protéine à 1 µM est inhibée par 1 mM d'EDTA (contrôle négatif) et métallée par 2 équivalents de Mn^{2+} par monomère d'EcFur. Différentes concentrations des peptides en équivalent de protéine (éq/Fur) puis 10 nM de plasmide pDT10 sont ajoutés. Le mélange réactionnel est digéré 1 h à 37 °C par 4 unités de HinfI puis déposé sur gel d'agarose 1 %. Tampon : Bis Tris Propane 100 mM pH 7,5, KCl 100 mM, $MgSO_4$ 5 mM. Les peptides sont dissous d'abord dans le DMSO puis dilués dans le tampon. Le puits avec le DMSO contient 20à 200 fois plus de DMSO que les puits 5 à 50eq peptide/Fur.

Sur quatre tests réalisés, la molécule inhibe la protéine 2 fois (dans un test, le contrôle positif d'inhibition n'a pas inhibé la protéine, non montré). Dans le premier test (Figure II.26.A), la molécule W-CO-Bz inhibe Fur à partir de 10 équivalents/Fur mais ce résultat n'a pas été confirmé par la suite. Un exemple de cette non-inhibition par W-CO-Bz est montré sur la Figure II.26.B. Sur cette figure, on voit aussi l'effet du DMSO sur l'activité de HinfI.

Il reste d'autres petites molécules qui sont plus intéressant théoriquement qu'on pourra toujours commander et tester (confère chapitre III) pour passer à des molécules plus intéressantes précieuses pour des applications thérapeutiques.

II. D. 5| Essais de co-cristallisation des variants de pF1 avec EcFur

Les multiples tentatives de cristallisation de la protéine EcFur Apo ou métallée, dans le passé, n'ayant pas donné de résultats, nous avons réalisé des essais de co-cristallisation de la protéine avec des peptides actifs. L'interaction de la protéine avec des peptides actifs pourrait stabiliser les complexes EcFur-peptides qui auraient ainsi plus de chance de co-cristalliser.

Les conditions des essais de co-cristallisation sont les suivantes :

- Métallation de la protéine avec du Mn^{2+} à 1,1 équivalents
- Les peptides pF1(1-13), pF1(1-10), pF1(2-9) et pF1(3-9) ont été ajoutés à la protéine à une stœchiométrie peptide/dimère de 1:1.
- Les mélanges ont été concentrés à 10mg/mL

Les premiers essais ont été réalisés par Sandra Galop (actuellement en deuxième année de thèse, au laboratoire) à la fin de son stage de Master 2. Les échantillons ont été envoyés à l'EMBL à Grenoble. Ces essais ayant montré un début de cristallogenèse avec le peptide pF1(1-10) dans du dioxane à 35% (Qiagen/Nextal) (Figure II.27), nous avons effectué d'autres essais de co-cristallisation (avec 2 équivalents de peptide cette fois par rapport à la protéine métallée) et d'optimisation des conditions de EcFur-Mn^{2+}-pF1(1-10). Ces essais ont été réalisés à Tromso (Norvège) dans le cadre d'une collaboration (projet AURORA) avec «*The Norwegian Structural Biology Centre - Universitetet i Tromsø* » (Norstruct), où j'ai effectué un séjour de trois semaines. Nous avons testé six cent conditions avec des librairies locales et commerciales avec les peptides pF1(1-13), pF1(1-10), pF1(2-9) et pF1(3-9) à une stœchiométrie peptide/dimère EcFur de 1:2, au robot et dix conditions d'optimisation de pF1(1-10)/EcFur 1:1 (Dioxane 31, 33, 35, 37 et 39%) (Confère chapitre IV). A l'heure actuelle, et depuis plus d'un an malgré d'autres

tentatives de nos collègues norvégiens, les essais n'ont pas été concluants et les petits cristaux observés dans le Dioxane, n'ont malheureusement pas pu être reproduits.

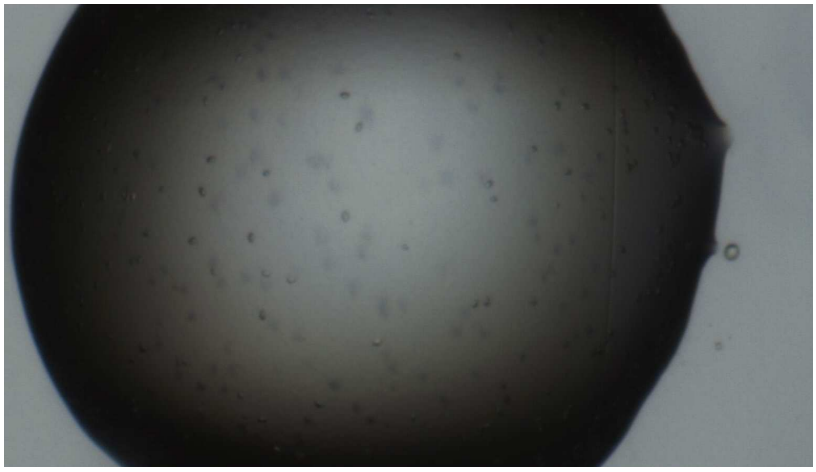


Figure II.27: Photo de cristallogenèse d'EcFur et pF1_1_10, réalisée à l'EMBL dans les conditions du Dioxane 35% (Qiagen/Nextal).

II. E| Bilan et discussion

Dans cette partie expérimentale, nous avons présenté les résultats des synthèses peptidiques effectuées et une caractérisation des interactions de la protéine EcFur avec ses inhibiteurs dérivés des aptamères peptidiques (Abed 2005). Nous avons commencé par présenter les résultats obtenus par le passé par Sylvia Vitale et Nadia Abed (Abed 2005) (Vitale 2009) sur la caractérisation de ces aptamères peptidique anti-Fur et leurs interactions avec leur cible : la protéine EcFur. Par la suite, nous avons présenté l'ensemble des études menés sur ces inhibiteurs anti-Fur dérivés des peptidiques.

Ces études sur les peptides linéaires pF1 à pF4 dérivés des aptamères peptidiques F1 à F4 et de leurs séquences tronquées ont permis de mieux comprendre les mécanismes d'inhibition de ces molécules. Dans ce travail, nous avons produit la protéine EcFur, synthétisé et purifié plusieurs peptides grâce à une collaboration avec Didier Boturyn du département de Chimie Moléculaire (UJF) qui m'a aidé à synthétiser et/ou purifier ces molécules. Nous avons vu les principes de la synthèse peptidique chimique et le mode opératoire pour la purification (confère Matériel et méthodes au chapitre V). Les peptides synthétisés ont ensuite été testées pour leur capacité à inhiber *in vitro* la liaison de Fur à l'ADN lors d'un test de protection à la coupure par une nucléase. Nous avons déterminé les IC_{50} de certains peptides contre l'activité de la protéine afin de pouvoir comparer l'effet inhibiteur de ces molécules et déterminer les séquences minimales inhibitrices. Il faut préciser que ce test étant dépendant des conditions, il faut pour pouvoir comparer les IC_{50} des inhibiteurs, que les tests soient réalisés dans les mêmes conditions, simultanément dans le meilleur des cas.

Globalement tous les peptides libres issus des aptamères inhibent la protéine EcFur, de manière différente, comme les aptamères peptidiques. Les différents peptides testés pF1, pF2, pF3 et pF4 inhibent EcFur avec des IC_{50} respectifs de 53 µM, 12.5 µM, 30 µM et 37 µM. Les diverses mutations réalisées sur les acides aminés prédits comme importants pour l'interaction des peptides pF1 et pF2 avec EcFur par la modélisation (confère chapitre III) ont permis d'avancer sur l'identification des acides aminés « clés » de ces peptides.

Nous avons montré que les séquences minimales inhibitrices de ces deux peptides sont pF1(3-9) et pF2(3-10) qui inhibent la protéine avec des IC_{50} de 50 µM et supérieur à 75 µM,

respectivement. Les acides aminés de pF1 identifiés comme essentiels pour observer un effet inhibiteur sont les résidus W3, C4, Y6, H8 et P9 pour pF1.

Concernant le peptide pF2, ce sont les résidus C3, C6 qui ont été mis en évidence. Le peptide pF2 pouvant lier les métaux, il se pourrait que ce peptide s'oxyde et forme des ponts disulfures intramoléculaires C3-C6, apportant une conformation particulière à la molécule pour son interaction forte avec la protéine. Le peptide pourrait aussi bien interférer avec la métallation de la protéine pour l'inhiber mais cela n'expliquerait pas son interaction forte visible en double hybride. Ce peptide pourrait lier l'ion métallique au sein de la protéine pour former un complexe inactif vis à vis de la liaison à l'ADN. Ceci pourrait expliquer la forte affinité de ce peptide pour EcFur et il sera intéressant d'étudier la perturbation des propriétés spectroscopiques des sites métalliques de la protéine par interaction avec pF2 et ces dérivés. Le peptide tronqué pF2(1-10) inhibe totalement la protéine à partir de 6 équivalents du peptide par rapport à la protéine. Il est encore possible d'effectuer des mutations pour déterminer l'ensemble des acides aminés responsable de cette forte inhibition d'EcFur.

De même, il faudrait réaliser d'autres mutations sur pF3 et pF4 afin de trouver les acides aminés essentiels à l'interaction avec EcFur. Pour le moment, les séquences minimales restent pF3(1-11) et le peptide entier pF4 pour ces deux peptides. Le peptide pF4 interagit sur des protéines Fur de la famille Fur autres que celle d'*E. coli* (Vitale 2009). Ce peptide sera très intéressant à étudier dans l'avenir dans la mesure où la compréhension de son mécanisme d'inhibition aiderait à comprendre l'inhibition des autres protéines Fur

On estime actuellement une perte d'affinité des peptides pF1 à pF4 par rapport aux aptamères peptidiques F1 à F4. Ce qui suggère que la contrainte imposée à la boucle variable par la plateforme thiorédoxine joue un rôle important dans l'interaction avec la cible. Dans le but d'obtenir des molécules ayant une affinité pour Fur similaire à celle des aptamères peptidiques, un peptide cyclique dérivé de pF1 a été synthétisé, conformément au modèle de l'amarrage de pF1 sur EcFur. Nous avons montré que ce peptide est plus actif que son homologue linéaire. Les peptides cycliques étant moins sensibles à la protéolyse que les peptides linéaires, les tests de leurs activités *in vivo* pourraient être envisagés. La cyclisation des peptides est donc une voie intéressante à suivre pour améliorer l'activité des peptides.

Nous avons souvent rencontré des problèmes de solubilité avec les peptides synthétisés, probablement à cause d'une instabilité des peptides en solution et à l'oxydation des groupements thiolates des cystéines contenues dans les séquences. Des analyses par spectrométrie de masse

« Electrospray» et des séquençages ont d'ailleurs confirmé qu'une partie des molécules présentes dans la solution mère des pF1 étaient oxydées, avec la formation de ponts disulfures intermoléculaires. Ces analyses n'ont pas été faites pour les autres peptides mais on peut facilement imaginer que c'est le cas car les autres peptides contiennent aussi des cystéines. Ils pourraient donc former des ponts disulfures intra et/ou intermoléculaires. Le dichroïsme circulaire nous a montré que les peptides ne possédaient pas de structure secondaire, mais la formation d'un pont disulfure peut apporter une structuration aux peptides, ou du moins « rigidifier » en partie leurs squelettes carbonés. Ce qui pourrait favoriser ou défavoriser leur interaction avec Fur. Il n'est pas impossible que ce soit la forme oxydée de ces peptides qui interagisse avec Fur. Cela pourrait expliquer la variabilité des résultats obtenus pour certains peptides. Le fait que les résultats obtenus avec pF2 soient reproductibles suggère que ce peptide est stable dans la solution de DMSO. L'utilisation de celui-ci pour la solubilisation des peptides avait été décidé afin d'être homogène avec les expériences sur pF2 (en raison de la solubilité limitée de pF2 en solution aqueuse). Or, le DMSO est un solvant connu pour ses propriétés oxydantes et il a pu accentuer les problèmes d'oxydations. Il faudra trouver une solution tampon, compatible avec le test (surtout compatibles avec l'activité de l'enzyme Hinf1 et de la protéine EcFur), qui solubilise bien tous les peptides.

Des précautions ont été prises pour stocker les peptides surtout en solution afin d'éviter au mieux ces problèmes d'instabilité. Les peptides qui étaient conservés à -20° sont désormais stockés sous argon à -80 C. Il y a donc encore un gros travail à faire pour maitriser l'oxydation et la solubilisation des peptides. Dans ce test, il faut pouvoir activer la protéine EcFur par métallation sans inhiber l'enzyme Hinf1. En effet, un excès de métal inhibe cette dernière et pourrait interférer avec les peptides dont certains peuvent lier le métal. Pour cela, il faudra utiliser un réducteur comme le DTT pour réduire les éventuels ponts disulfures qui pourraient se former et s'assurer qu'il n'a pas un effet sur la protéine et l'enzyme. Après avoir maîtrisé tous ces aspects, il faut être très vigilant aux mesures, aux pipetages de micros volumes et éviter toutes traces métalliques qui peuvent polluer les expériences sans oublier les contrôles pour chaque condition nouvelle. L'utilisation d'eau ultra pure, de pipettes et cônes spécifiques au métal a permis d'éviter les traces métalliques qui pourraient exister dans l'eau distillée conventionnelle et sur les instruments.

Certes, le test de protection à la nucléase renseigne sur la capacité des peptides à inhiber *in vitro* la liaison de Fur à l'ADN, mais il ne permet pas de déterminer directement des constantes de dissociation (Kd), ni de savoir si les peptides interagissent avec le dimère dans sa

conformation inactive (Apo) ou active (métallée ou Holo). Pour répondre à cette question, nous pouvons utiliser la spectroscopie de fluorescence pour suivre celle du tryptophane. C'est ce qui a déjà utilisée pour le peptide pF1 et qui a permis de montrer que ce peptide interagit avec la conformation active d'EcFur avec un Kd d'environ 30 µM, proche de l'IC_{50} obtenu lors des tests de protection à la nucléase. Ce résultat est présenté dans le chapitre III, car il a permis d'affiner les amarrages du peptide pF1 et les résultats sont maintenant cohérents avec les résultats expérimentaux. Le peptide pF2 ne possédant pas de tryptophane, nous avons synthétisé la séquence pF2(1-11) Y11W ou le tryptophane a été rajouté en position 11 sur pF2(1-10). Nous avons montré que cette séquence est le plus actif de tous les inhibiteurs dans les tests effectués. Il sera possible de faire la même modification pour pF4 afin de pouvoir utiliser ce test pour tous les peptides. Le peptide pF3 contient déjà un tryptophane.

A ce stade, on ne peut toutefois pas conclure qu'il n'existe pas du tout d'interaction entre les peptides trouvés inactifs et Fur. En effet, il se peut que certaines interactions, relevées dans l'étude théorique, soient trop faibles pour être visibles dans les conditions de nos tests. Le BIAcore sera précieux dans ce cas pour comparer les prédictions des simulations et les expériences pratiques. Des expériences sont actuellement réalisées au laboratoire pour essayer de caractériser mieux les interactions des peptides avec Fur par résonance plasmonique de surface (BIAcore) et la technique du double-hybride en utilisant le gène rapporteur *luc* de la luciférase. Nous espérons déterminer ainsi les constantes de dissociation (Kd) des complexes. Il sera alors intéressant de tester l'ensemble des peptides trouvés actifs en test nucléase pour confirmer les résultats de ces tests, et voir si les troncations entraînent une perte ou une augmentation de l'affinité des peptides pour Fur. Le test des petites molécules prédites par les simulations sera aussi intéressant pour la suite du projet. Cela ouvrira une voie vers des applications thérapeutiques.

Les structures cristallographiques ou RMN des complexes EcFur-peptides permettraient de mieux comprendre ces interactions et de valider les résultats des calculs. Nous avons réalisé des essais de co-cristallisation des complexes EcFur et des variants de pF1, trouvés actifs dans le test d'activité. Il serait intéressant d'effectuer d'autres essais, notamment avec les peptides pF4 et pF2 qui sont aussi intéressants comme nous l'avons décrits précédemment. Ceci sera fait avec EcFur mais aussi avec des Fur d'autres pathogènes.

Bibliographie

Abed N. (**2005**). "La protéine Fur nouvelle cible antibactérienne? Approches par utilisation des aptamères peptidiques", ***Université Joseph Fourier***.

Abed N., Bickle M., Mari B., Schapira M., Sanjuan-Espana R., Robbe Sermesant K., Moncorge O., Mouradian-Garcia S., Barbry P., Rudkin B. B., Fauvarque M. O., Michaud-Soret I. and Colas P. (**2007**). "A comparative analysis of perturbations caused by a gene knock-out, a dominant negative allele, and a set of peptide aptamers." *Mol Cell Proteomics* **6**(12): 2110-2121.

Bagg A. and Neilands J. B. (**1987**). "Ferric uptake regulation protein acts as a repressor, employing iron (II) as a cofactor to bind the operator of an iron transport operon in *Escherichia coli*." *Biochemistry* **26**(17): 5471-5477.

D'Autreaux B., Touati D., Bersch B., Latour J. M. and Michaud-Soret I. (**2002**). "Direct inhibition by nitric oxide of the transcriptional ferric uptake regulation protein via nitrosylation of the iron." *Proceedings of the National Academy of Sciences of the United States of America* **99**(26): 16619-16624.

Kelly S. M., Jess T. J. and Price N. C. (**2005**). "How to study proteins by circular dichroism." *Biochimica et biophysica acta* **1751**(2): 119-139.

Pecqueur L., D'Autreaux B., Dupuy J., Nicolet Y., Jacquamet L., Brutscher B., Michaud-Soret I. and Bersch B. (**2006**). "Structural changes of *Escherichia coli* ferric uptake regulator during metal-dependent dimerization and activation explored by NMR and X-ray crystallography." *J Biol Chem* **281**(30): 21286-21295.

Reshetnyak Y. K., Koshevnik Y. and Burstein E. A. (**2001**). "Decomposition of protein tryptophan fluorescence spectra into log-normal components. III. Correlation between fluorescence and microenvironment parameters of individual tryptophan residues." *Biophysical journal* **81**(3): 1735-1758.

Vitale S. (**2009**). "Inhibiteurs anti-Fur et caractérisation de Fur de Helicobacter pylori", ***Université J. Fourier***.

Chapitre III :

Approche théorique de l'étude de l'interaction des inhibiteurs anti-Fur et de leur cible

CHAPITRE III : ..... **- 131 -**

APPROCHE THEORIQUE DE L'ETUDE DE L'INTERACTION DES INHIBITEURS ANTI-FUR ET DE LEUR CIBLE ... **- 131 -**

III. A| OUTILS THEORIQUES ..... - 135 -
III. A. 1| Introduction ..... *- 135 -*
III. A. 2| La modélisation par homologie (MODELLER) ..... *- 136 -*
III. A. 3| La mécanique moléculaire ..... *- 137 -*
III. A. 4| La dynamique moléculaire ..... *- 146 -*
III. A. 5| L'amarrage moléculaire ..... *- 152 -*
III. B| MODE OPERATOIRE DE L'AMARRAGE ..... - 163 -
III. B. 1| Recherche et optimisation des structures initiales ..... *- 163 -*
III. B. 2| Recherche des zones d'amarrage des peptides sur la protéine ..... *- 173 -*
III. B. 3| Amarrage Assisté par Dynamique Moléculaire (AADM) ..... *- 177 -*
III. B. 4| Identifications des interactions clés à l'amarrage ..... *- 178 -*
III. B. 5| Exemple du processus d'optimisation de l'amarrage des peptides anti-Fur : Cas du peptide pF1 sur le modèle EcFur ←Pa. ..... *- 179 -*
III. B. 6| Conclusion sur la méthodologie ..... *- 186 -*
III. C| AMARRAGE DES PEPTIDES SUR LE MODELE ECFUR ← VC ..... - 187 -
III. C. 1| Le modèle EcFur ←Vc ..... *- 187 -*
III. C. 2| Amarrage des séquences de pF1 sur EcFur ←Vc ..... *- 188 -*
III. C. 3| Amarrage de pF2 sur EcFur ←Vc ..... *- 194 -*
III. C. 4| Amarrage de pF3 sur EcFur ←Vc ..... *- 200 -*
III. C. 5| Conclusions et discussions sur les amarrages des peptides pF1 à pF3 sur le modèle EcFur ←Vc- 206 -
III. D| AMARRAGE DU PEPTIDE PF1 SUR LES AUTRES PROTEINES FUR ..... - 208 -
III. D. 1| Amarrage sur la structure de PaFur ..... *- 209 -*
III. D. 2| Amarrage sur la structure de HpFur ..... *- 212 -*
III. D. 3| Amarrage sur la structure de VcFur ..... *- 214 -*
III. D. 4| Conclusion sur les amarrages de pF1 sur les protéines Fur de divers pathogènes ..... *- 215 -*
III. E| LE CRIBLAGE VIRTUEL DE PETITES MOLECULES SUR FUR ..... - 220 -
III. E. 1| Préparation de la protéine ..... *- 220 -*
III. E. 2| Préparation de la librairie ..... *- 222 -*
III. E. 3| Sélection des petites molécules ..... *- 223 -*
III. E. 4| Analyse des criblages ..... *- 227 -*
III. E. 5| Conclusion sur le criblage ..... *- 231 -*
BIBLIOGRAPHIE ..... - 230 -

III. A| Outils théoriques

III. A. 1| Introduction

Lorsqu'on s'intéresse à l'étude des mécanismes biologiques en utilisant des techniques de calcul, la connaissance de la structure des molécules d'intérêt est capitale. Ces structures sont généralement fournies par la cristallographie aux rayons X (RX) ou la résonance Magnétique nucléaire (RMN) et sont disponibles dans les banques de données. La banque dédiée aux structures des protéines est la *protein data bank* (PDB) (Berman *et al.* 2000). Aujourd'hui (2011), la PDB compte plus de 75 000 protéines (www.pdb.org)

La RMN et la cristallographie nous donnent certes des informations sur la structure des molécules biologiques, mais les coordonnées atomiques récupérées donnent une image statique. Des études sur les systèmes biologiques en perpétuel mouvement sont requises pour compléter les données obtenues de la PDB. Ces études sont rendues possibles grâce aux simulations de **mécanique** et de **dynamique moléculaire**. Ces techniques de simulation permettent d'étudier les propriétés dynamiques et structurales d'un système au niveau moléculaire. Elles permettent de "quantifier" des grandeurs inaccessibles par l'expérience comme le déplacement des atomes et de caractériser les interactions entre tous les atomes. Un autre avantage de ces techniques est de pouvoir faire varier très facilement des paramètres physiques difficilement modifiables dans l'expérience, tels que la densité, la température, la pression. Enfin, les simulations permettent de tester la validité d'une approche expérimentale physique, biologique, biochimique ou de cibler les paramètres à l'origine d'un phénomène particulier.

Lorsque la structure d'une protéine d'intérêt n'est pas connue et qu'il existe des homologues dans la PDB, on a recours à la **modélisation par homologie**. Cette méthode permet de construire des modèles à partir de la structure des protéines homologues.

La formation des complexes entre macromolécules est un phénomène quasi nécessaire pour tous les processus biologiques. La connaissance de la forme tridimensionnelle de ces complexes par la RMN ou la cristallographie est très utile pour une compréhension détaillée de la formation de ces complexes, mais souvent coûteuse à mettre en place. On a souvent

recours à des programmes d'**amarrage moléculaire automatique** « d***ocking*** » qui permettent de prédire la formation de complexe moléculaire et la sélection virtuelle des petites molécules prometteuses dans le développement de nouveaux médicaments

Dans le cadre de cette étude, il s'agit d'étudier l'amarrage de peptides inhibiteurs de la protéine Fur d'*Escherichia coli* (EcFur), d'identifier les interactions importantes et de prédire l'interaction de ces molécules avec d'autres homologues d'EcFur. Comme il a été présenté dans l'introduction, la protéine EcFur est active sous forme dimérique. La protéine dimérique peut adopter deux formes : la forme « ouverte », par analogie avec la structure de PerR, non métallée (Apo-PerR) (Traore *et al.* 2006), est non active, et la forme dite « fermée » ou (Holo) correspond à la protéine activée par le métal (Pohl *et al.* 2003). Chaque monomère contient un domaine de liaison à l'ADN (Nter) et un domaine de dimérisation (Cter) qui correspondent respectivement aux résidus 1 à 83 Nter (du résidu 1 au 83) et Cter (du résidu 84 à 148) chez *E. coli*. La structure cristallographique d'EcFur n'est pas connue dans sa totalité, seule la structure du domaine de liaison à l'ADN est décrite (Pecqueur *et al.* 2006). Au laboratoire, nous disposons également des données RMN sur la protéine Apo-EcFur qui donnent la structure secondaire de la protéine (Pecqueur *et al.* 2006). Ces données nous permettront de valider les modèles d'EcFur que nous allons construire. Les résultats expérimentaux antérieurs sur la caractérisation des aptamères peptidiques ainsi que des peptides dérivés des thèses de Nadia Abed et de Sylvia Vitale (Abed. 2005; Vitale. 2009) ont constitué une base solide pour ces simulations qui vont suivre. Des allers retours entre simulations et expériences biochimiques ont été nécessaires pour guider ces travaux.

Dans ce chapitre, nous allons décrire les principes des outils de simulation théorique, utilisés dans cette thèse. Nous développerons les méthodologies pratiquées avant la présentation des résultats d'amarrage et de criblages virtuels. La dernière partie sera consacrée à une discussion.

III. A. 2| La modélisation par homologie (MODELLER)

Comme il a été expliqué dans l'introduction, lorsque la structure d'une protéine n'est pas disponible, il est possible de construire des modèles à partir des structures de protéines de la même famille par la modélisation par homologie. Cette méthode s'appuie sur l'hypothèse

que les protéines homologues conservent en général une structure proche et une même fonction. C'est le cas de la protéine EcFur dont nous voulions construire des modèles par homologie à partir des structures de Fur publiées.

Pour établir le degré d'homologie entre deux protéines, on utilise l'alignement de séquence. On distingue deux types de ressemblance à ne pas confondre : les taux d'identité ou de similarité qui représentent respectivement les pourcentages d'acides aminés identiques ou similaires dans leurs séquences. On considère généralement qu'un pourcentage de similarité de 30% suffit pour construire des modèles corrects par homologie. Nous avons utilisé le programme MODELLER (Sali *et al.* 1993) pour construire les modèles utilisées dans cette thèse.

MODELLER est un programme libre d'accès très largement utilisé pour la construction de modèles par homologie et aussi pour la modélisation comparative de la structure des protéines. MODELLER est de surcroît très intéressant pour la modélisation de novo des boucles dans la structure des protéines ce qui a été très utile dans la construction des modèles des peptides issus des aptamères peptidiques. Le programme fonctionne sur un processus automatique de modélisation par homologie. Partant d'un alignement de séquences, le programme construit dix modèles (par défaut) de la protéine d'intérêt à partir de la (et/ou des) séquence (s) de structure(s) spécifiée(s). Ces modèles construits sont raffinés par minimisation d'énergie de manière à respecter des contraintes spatiales issues de la structure de référence, préalablement stockées dans le programme. Finalement, les modèles sont classés suivant une fonction de pénalité appelée « objective function value » permettant de choisir la meilleure conformation. A défaut d'autres informations structurales concernant la protéine à modéliser, la conformation ayant la fonction de pénalité la plus petite est conservée.

III. A. 3| La mécanique moléculaire

La mécanique moléculaire correspond à l'utilisation de la mécanique classique pour modéliser et raffiner les structures obtenues par diffraction des rayons X et RMN. L'idée directrice est de décrire les interactions intramoléculaires par des termes d'énergie empruntés à la mécanique classique. Les interactions sont modélisées empiriquement à l'aide d'une fonction énergétique reproduisant au mieux les données expérimentales telles que des spectres

vibrationnels, ou encore des grandeurs calculées *ab initio* obtenus par la mécanique quantique. Cette fonction énergétique doit être simple, rapide et suffisamment précise pour simuler de façon acceptable les propriétés structurales et thermodynamiques des macromolécules. Les coordonnées des atomes sont modifiées dans le but d'obtenir la conformation de la molécule d'énergie minimale, correspondant en théorie à la structure la plus stable.

NB : Contrairement à la mécanique quantique, les électrons ne sont pas traités explicitement. Contrairement à la mécanique quantique, on n'a pas de rupture des liaisons covalentes en mécanique moléculaire. De ce fait, elle ne permet pas l'étude de la formation ou de la rupture de liaisons covalentes entre les atomes, ni, bien sûr, d'obtenir des informations sur des transferts d'électrons entre des atomes ou des états électroniques excités des atomes.

III. A. 3. a| Principes

En mécanique moléculaire, on considère une structure moléculaire comme étant composée de sphères chargées (représentant les atomes) reliées par des ressorts (qui représentent les liaisons). L'énergie du système est décrite par une fonction analytique qui dépend des coordonnées atomiques et des interactions interatomiques. Cette fonction représente l'énergie potentielle (Ep) du système. Cette dernière est décomposée en une somme de termes internes (longueur de liaisons, angles de liaisons, angles dièdres et dièdres impropres) - qui permettent de maintenir la géométrie covalente de la molécule - et de termes externes (interactions de Van der Waals, électrostatiques et liaisons hydrogène), présentant les interactions entre atomes non liés.

Pour trouver la géométrie optimale d'un ensemble d'atomes, il faut minimiser la fonction énergétique. Par principe, l'énergie minimale correspond à la structure la plus stable. La forme analytique de l'énergie potentielle et le jeu de paramètres nécessaires aux calculs constituent le « champ de forces ».

III. A. 3. b| Champs de force

Le champ de forces est la base des techniques de simulations moléculaires. Il permet de quantifier les interactions interatomiques et intermoléculaires du système aussi précisément que possible. Il permet d'établir par le choix des fonctions énergétiques et des paramètres un

modèle mathématique, qui représente bien les variations de l'énergie potentielle avec la géométrie moléculaire. Les paramètres permettant de calculer les différents termes d'énergie sont déduits d'études détaillées sur des systèmes simples et transférés à des molécules plus grosses. Il est possible également d'introduire des termes de contraintes, de position, de géométrie, pour rendre compte des conditions expérimentales connues.

Il existe plusieurs types de champs de forces, en fonction du type de molécules. Les principaux champs de forces classiques utilisés pour la simulation des protéines et des acides nucléiques sont ceux de CHARMM (Brooks *et al.* 1983), AMBER (Weiner *et al.* 1981), etc…

Dans cette thèse, nous avons utilisé le programme CHARMM pour les simulations de mécanique et dynamique moléculaire. Une description des fonctions utilisées est donnée ci-dessous.

III. A. 3. c| CHARMM

CHARMM (Chemistry at HARvard Macromolecular Mechanics) est un programme de simulation moléculaire polyvalent, développé pour l'étude des molécules d'intérêt biologique (Brooks *et al.* 2009). Le programme permet de lire et construire des structures, de minimiser leur énergie, d'effectuer des simulations de dynamique moléculaire en solution, dans les cristaux, ou dans les environnements membranaires. Ce programme permet aussi d'analyser les structures générées et de déterminer les propriétés dynamiques et énergétiques des systèmes étudiés.

La fonction énergétique ($\boldsymbol{E_p}$) du champ de force CHARMM *c32b2*, comprend des termes internes et externes auxquels il est possible d'ajouter des contraintes.

i| Les termes internes

Les termes internes représentent les interactions entre atomes liés, c'est-à-dire séparés par au plus trois liaisons. Chacun de ces termes s'exprime sous la forme d'une fonction qui rend compte du coût énergétique lié à la déformation subie par l'édifice covalent du fait des divers modes de vibration et de rotation des atomes. Ce sont les termes d'élongation des liaisons covalentes entre deux atomes (stretching), de déformation des angles de valence entre

trois atomes (bending), de torsion des angles dièdres formés par quatre atomes liés consécutivement (appelés dièdres propres) et des angles formés par 4 atomes qui ne sont pas linéairement liés (dièdres impropre) (Figure III.1). Chacun de ces termes possède une ou plusieurs positions d'équilibre préférentiel (s) déterminé (s) par des calculs quantiques ou des mesures expérimentales.

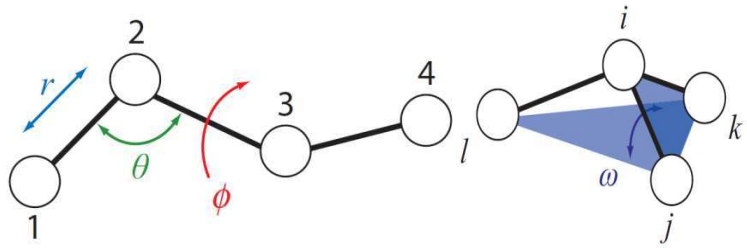
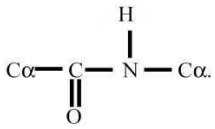


Figure III.1: Représentation des termes liés en mécanique moléculaire. A gauche : la liaison covalente ***r*** (bleu), l'angle de valence $\boldsymbol{\theta}$ (vert) et angle de torsion $\boldsymbol{\phi}$ (rouge). A droite : angle de dièdre impropre $\boldsymbol{\omega}$ (violet)

Les dièdres impropres maintiennent une certaine disposition spatiale d'un groupe de quatre atomes qui ne se suivent pas séquentiellement. C'est le cas par exemple de la liaison peptidique et des cycles aromatiques dont la planéité est assurée par un angle impropre.



Ces dièdres impropres empêchent également les inversions de configuration des centres énantiomériques tels que les carbones $\boldsymbol{\alpha}$ des acides aminés (à l'exception de la glycine) qui sont tous en configuration absolue L.

La fonction permettant de calculer les termes l'énergie des termes internes ($\boldsymbol{E_{internes}}$) est représentée ci-dessous:

$$E_{internes} = \sum_{liaisons} k_r (r - r_0)^2 + \sum_{angles} k_\theta (\theta - \theta_0)^2 + \sum_{dièdres} k_\phi \{1 + \cos(n\Phi - \delta)\} + \sum_{impropres} k_\omega (\omega - \omega_0)^2$$

L'ensemble des termes liés, à l'exception des termes de torsion, est traité de manière harmonique où $r - r_0$, $\theta - \theta_0$, et $\omega - \omega_0$ sont les déviations respectives des longueurs de liaison, des angles de valence et des dièdres impropres, par rapport à leur position d'équilibre respective, r_0, θ_0, ω_0 et ε_0. Les facteurs k_r, k_θ et k_ω sont les constantes de force associées aux interactions. La variation des angles dièdres propres est modélisée par les premiers termes

d'une série de Fourier où n est la périodicité et δ l'angle de phase du dièdre ϕ, k_ϕ est la constante de force du dièdre.

ii| Les termes externes

Les termes externes ou non liés concernent les atomes séparés par plus de 4 liaisons covalentes. Ces termes comprennent les liaisons hydrogènes ($\boldsymbol{E_{hb}}$), les interactions électrostatiques ($\boldsymbol{E_{el}}$) et de van der Waals ($\boldsymbol{E_{vdw}}$).

Les interactions de van der Waals sont représentées grâce à une fonction de Lennard-Jones (dit potentiel 6-12, à cause des puissances de la variable $\boldsymbol{r}$) :

$$E_{vdW} = \sum_{paires(ij)} \left\{ \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right\}$$

Les paramètres $\boldsymbol{A}_{ij}$ et $\boldsymbol{B}_{ij}$ sont des constantes de van der Waals dépendants de la nature des atomes i et j et sont évalués à partir de polarisabilité atomique. Ils correspondent à la profondeur du puits de la fonction de la courbe orange de la Figure III.2 ci-dessous.

L'énergie de van der Waals est composée d'un terme attractif et d'un terme répulsif, illustrés respectivement, par des pointillés (en vert) et des tirets (en bleu) sur la Figure III.2. Le terme attractif connu historiquement sous le nom d'énergie de dispersion de London (interaction dipôle induit dipôle induit) varie en $-\frac{1}{r_{ij}^{6}}$. Il rend compte de la dispersion de London: deux atomes sont soumis à des interactions attractives, indépendamment de l'électrostatique, dues aux mouvements des électrons autour des noyaux qui créent des dipôles instantanés. Le terme répulsif varie en $\frac{1}{r_{ij}^{12}}$ et répond au principe d'exclusion de Pauli : deux nuages électroniques ne peuvent pas s'interpénétrer car deux électrons ne peuvent pas occuper le même état quantique.

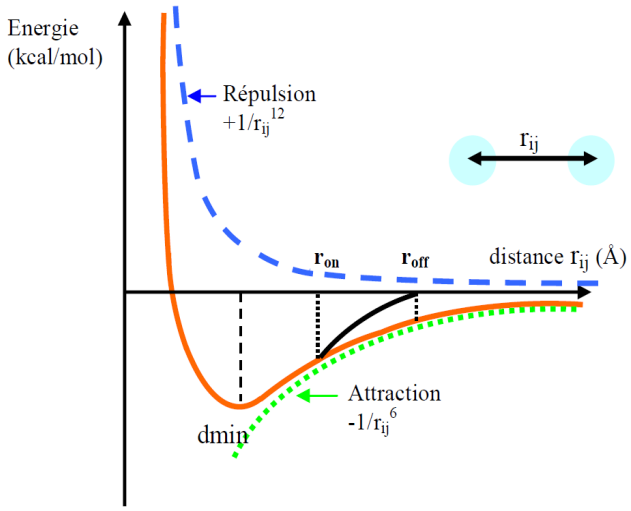


Figure III.2: Représentation de d'énergie de van der Waals en fonction de la distance r (en orange) : [somme des termes répulsif (en bleu) et attractif (en vert)].
Pour diminuer le temps de calcul, l'énergie est progressivement amenée à zéro (noir) par la fonction SWITCH dans l'intervalle $r_{on} < r < r_{off}$. Pour $r > r_{off}$, les interactions ne sont pas calculées.

Les interactions électrostatiques sont calculées suivant la loi de Coulomb:

$$E_{el} = \sum_{paires(i,j)} \frac{q_i q_j}{4\pi\varepsilon_0 \varepsilon_r r_{ij}}$$

Avec $\boldsymbol{q}_i$ et $\boldsymbol{q}_j$, les charges partielles des atomes i et j calculées par des méthodes de mécanique quantique. $\boldsymbol{r}_{ij}$, est la distance séparant les atomes i et j, $\boldsymbol{\varepsilon}_0$ la constante diélectrique du vide et $\boldsymbol{\varepsilon}_r$ la constante diélectrique relative. La valeur de la permittivité relative peut être utilisée pour simuler les effets du solvant autour de la molécule lorsque celui-ci n'est pas introduit explicitement.

Le calcul des termes externes est extrêmement couteux en temps de calcul lorsque l'on prend en compte toutes les paires d'atomes (i, j). Il est possible de définir une distance (rayon r_{off}) limite, appelé aussi « cut-off » au-delà de laquelle les termes externes ne sont pas calculés. Cette coupure brutale du traitement des interactions peut induire des artefacts de simulation et des instabilités énergétiques. CHARMM dispose des fonctions SHIFT et SWITCH qui permettent de remédier à ce phénomène en amenant progressivement le potentiel à 0. L'exemple de la fonction SWITCH est donné à la Figure III.2 (en noir).

Les liaisons hydrogènes ne sont pas explicitement prises en compte par des fonctions énergétiques dans les nouvelles versions de champ de forces de CHARMM. La

paramétrisation adaptée des potentiels de van der Waals et électrostatiques permet de prendre en compte la contribution des liaisons hydrogènes sans avoir à ajouter un terme explicite (***E_{hb}***). L'énergie des liaisons hydrogène est calculée simplement comme la somme des interactions de van der Waals et électrostatiques.

iii| Les termes de contraintes

Lors d'une simulation moléculaire, il peut être nécessaire de maintenir des informations structurales connues (cristallographie 3D ou RMN) ou supposées (modélisation par homologie). Il est possible de fixer ou d'imposer des contraintes sur l'ensemble ou une partie des coordonnées initiales des atomes (Ex : les atomes des chaines latérales dans le cas des protéines). CHARMM dispose de potentiels de contrainte harmoniques qui autorisent une certaine mobilité et donc une certaine flexibilité.

Les termes de contraintes harmoniques couramment utilisés dans CHARMM sont :

- Les contraintes harmoniques de positions atomiques

$$E_{cr} = \sum k_i \left(\vec{r}_i - \vec{r}_{io}\right)^2$$

Avec K_i constante dépendante de la masse de l'atome i, r_i la nouvelle position de l'atome i et r_{io}, sa position initiale.

Les contraintes harmoniques appliquées à la position des atomes permettent de maintenir une conformation particulière pour un ensemble donné d'atomes en évitant de grands déplacements mais en permettant tout de même la relaxation de la structure, par exemple, on peut maintenir le squelette de la protéine et relaxer les chaînes latérales.

- Les contraintes harmoniques de distances interatomiques (E_{cd})

$$E_{cd} = \sum_{ij} k_{ij} \left(r_{ij} - r_{ij0}\right)^2$$

Avec K_{ij}, constante de force, r_{ij} distance entre les atomes i et j et r_{ij0}, distance d'équilibre.

- Les contraintes d'angle de valence et d'angles dièdres ϕ ($E_{c\phi}$)

$$E_{c\phi} = \sum_{ij} k_i \left(\Phi_i - \Phi_{i0}\right)^2$$

Avec ϕ_i et ϕ_{i0}, angles dièdres calculés et initiaux respectivement. Les contraintes appliquées aux angles dièdres permettent le maintien d'une conformation locale.

Le processus d'optimisation de la géométrie d'une molécule passe par une minimisation de la fonction énergétique en utilisant les termes précédemment détaillés.

III. A. 3. d| La solvatation dans les simulations

Dans leur environnement, les molécules biologiques interagissent en permanence avec les molécules du solvant. Il est donc indispensable de prendre en compte ces interactions dans les modèles de simulation. Deux méthodes de solvatation sont utilisées dans les simulations pour prendre en compte l'effet du solvant : la **solvatation explicite** où les atomes du solvant sont représentés et la **solvatation implicite**. Dans cette dernière méthode, des termes énergétiques permettent de prendre en compte l'effet du solvant sans ajouter les atomes du solvant.

La **méthode explicite** consiste à définir une boîte virtuelle autour de la molécule dans laquelle on dispose des molécules d'eau. Il est également possible d'ajouter des ions positifs ou négatifs afin d'assurer la neutralité du système. Plusieurs modèles différents ont été proposés pour décrire les molécules d'eau, qui se différencient par le nombre de points utilisés pour définir la molécule, suivant la manière dont on veut traiter le solvant (si on considère la structure flexible ou rigide, ou encore si on veut inclure des effets de polarisation). Le modèle TIP3P (Jorgensen *et al.* 1983) utilisé dans CHARMM, modélise de façon simple et efficace les molécules d'eau.

La **méthode implicite** utilise une approximation de l'équation de Poisson-Boltzmann : le modèle de Born. L'équation de Poisson-Boltzmann décrit l'environnement électrostatique d'un soluté dans un solvant. Mais résoudre cette équation numériquement prend énormément de temps de calcul. L'approximation de Born permet, en rajoutant un terme supplémentaire à la fonction d'énergie potentielle, de traiter les effets du solvant de manière implicite.

La méthode « Effective Energy Function 1 » (EEF1) (Lazaridis *et al.* 1999) implémentée dans le champ de force CHARMM a été utilisée dans les processus de simulation de cette étude. Cette méthode de solvatation suppose que l'énergie libre de solvatation de chaque groupe dans la protéine est égale à l'énergie libre de solvatation de ce groupe dans un petit composé modèle moins l'énergie de solvatation perdue par exclusion de solvant des autres atomes de la macromolécule autour de lui, représentée ci-dessous.

$$\Delta G_i^{slv} = \Delta G_i^{ref} - \sum_j \int_{V_j} f_i(r) d_r$$

La méthode EEF1 prend 2 fois plus de temps que des simulations équivalentes dans le vide. Mais, elle est beaucoup plus rapide que les méthodes de solvatation explicite.

III. A. 3. e| Méthodes de minimisation d'énergie

La minimisation d'énergie est l'une des étapes clé d'un protocole de simulation moléculaire. Elle est couramment utilisée en alternance avec la dynamique moléculaire pour l'optimisation de la géométrie. Elle permet d'explorer l'espace conformationnel, d'éliminer les mauvais contacts interatomiques afin d'obtenir la conformation la plus stable d'une macromolécule (Figure III.3). Par hypothèse, cette conformation correspond généralement au *minimum* global. Généralement, les méthodes de minimisation permettent d'approcher le *minimum* local le plus proche de la structure de départ (Figure III.3). Comme une molécule peut avoir plusieurs *minima* locaux et il est très difficile de trouver le *minimum* global.

Nous allons décrire la méthode de la plus grande pente (*Steepest Descent* ou *SD*) et la méthode ABNR (*Adopted Basis Newton-Raphson*), utilisées au cours des simulations de cette thèse.

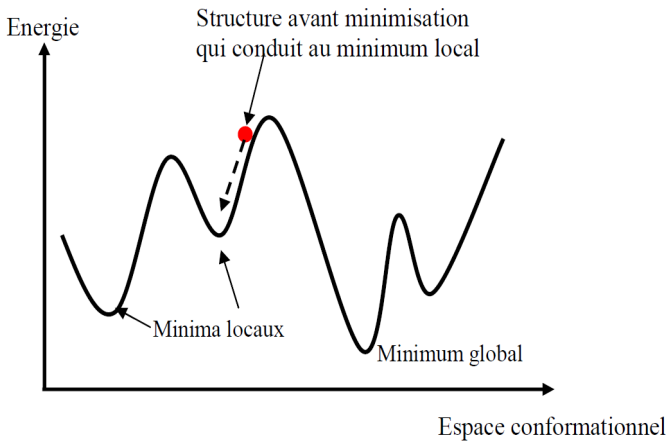


Figure III.3: Courbe d'énergie d'un espace conformationnel à une dimension

La méthode SD modifie la position des atomes en minimisant la fonction énergétique selon la dérivée première de l'énergie (ou gradient). C'est une méthode rapide et efficace lorsque la pente de la fonction est forte. Elle est très efficace lorsqu'on se trouve loin du minimum et permet de supprimer les mauvais contacts stériques et d'accéder rapidement au minimum local le plus proche de la conformation de départ. Toutefois, elle présente un

comportement oscillatoire autour de ce minimum local et ne converge que très lentement. Elle n'est donc pas efficace pour le raffinement autour du puits de potentiel, particulièrement pour les macromolécules. Pour contourner cet obstacle on fait appel aux méthodes dites de dérivée seconde, comme la méthode *ABNR*, particulièrement bien adaptée pour les gros systèmes tels que les protéines.

La méthode *ABNR* utilise l'algorithme de Newton-Raphson qui converge rapidement au tour du minimum. De plus, cette méthode nécessite beaucoup d'espace disque pour stocker les données et aussi un temps de calcul plus long. Cette méthode combine les avantages des méthodes du premier ordre (ici la SD) avec l'algorithme de Newton-Raphson, c'est-à-dire qu'elle s'appuie sur l'optimisation de la méthode SD pour affiner la géométrie.

Ces méthodes de minimisation décrites ci-dessus permettent de simples réajustements des coordonnées initiales des atomes afin d'approcher le *minimum* local le plus proche de la conformation initiale (optimisation locale). En revanche, elles ne permettent pas de franchir les barrières énergétiques (les sommets de la courbe) présentes dans un système moléculaire (Figure III.3).

Il existe d'autres méthodes permettant de franchir les barrières énergétiques et d'obtenir des mouvements plus importants au sein de la molécule : la méthode du recuit simulé « simulated annealing » (SA) (Goodsell *et al.* 1990) ; les méthodes Monté-Carlo (MC), les méthodes utilisant les algorithmes génétiques (GA) et la dynamique moléculaire (DM). Dans la suite, nous allons décrire la dynamique moléculaire utilisée pour obtenir des changements conformationnels importants dans les molécules.

III. A. 4| La dynamique moléculaire

La mécanique et la dynamique moléculaire décrites ici partagent le même champ de force CHARMM. La mécanique moléculaire permet d'obtenir des informations à partir d'une géométrie unique de la molécule. A l'inverse, la dynamique moléculaire produit une succession de conformations qui décrit l'évolution de la position des atomes au cours du temps (leur trajectoire). Les variations des positions atomiques résultent des diverses forces

auxquelles les atomes sont soumis au cours du temps. On a accès à une représentation dynamique des molécules où les atomes sont en perpétuel mouvement. Ces mouvements peuvent être locaux et rapides, ou collectifs et plus lents, impliquant ainsi des groupes d'atomes voire des domaines entiers de la molécule. Ces mouvements peuvent être concertés et essentiels à la fonction des molécules biologiques. Ils créent donc un ensemble de conformations possibles pour la molécule donnée à une température donnée.

III. A. 4. a| Principes

Le principe de la dynamique moléculaire est basé sur la résolution de l'équation fondamentale de Newton. Comme en mécanique moléculaire, la macromolécule est définie par la position dans l'espace de chacun de ses atomes. Chaque atome est considéré comme une masse ponctuelle de coordonnées r, au temps t. A l'instant t+δt, un atome i donné se trouvera à une nouvelle position $r_{t+\delta t}$ du fait des différentes forces qui s'exercent sur lui, liées à la présence des autres atomes environnants. Le but de la dynamique moléculaire est de calculer la position de l'atome à cet instant t+δt. Les équations employées permettent d'établir les relations entre ces deux instants t et t+δt. Ainsi, dans un intervalle de temps suffisamment court où les forces ne varient pas, on peut exprimer le mouvement des atomes grâce à l'équation fondamentale de la dynamique (Seconde loi de Newton) :

$$\sum_j F_{ij} = m_i a_i$$

Où $\boldsymbol{F_{ij}}$ sont les forces exercées par les atomes *j* sur l'atome *i* de masse m_i et d'accélération $\boldsymbol{a_i}$ au temps t. $\boldsymbol{F_{ij}}$ et $\boldsymbol{a_i}$ peuvent être facilement évaluées à partir des coordonnées atomiques :

$$a_i(t) = \frac{d^2 r_i}{dt^2}$$

$$F_{\perp}ij\ (t) = -(_r U_P\ (r_{\perp}i, t)$$

Avec $r_i(t)$ la position instantanée de l'atome i et $\boldsymbol{\nabla_r U_p}$ le gradient de l'énergie potentielle $\boldsymbol{E_p}$ **(U_p).** Connaissant l'accélération, on peut calculer la vitesse d'un atome à l'étape suivante, au temps t+δt. A partir des positions, vitesses et accélérations atomiques à tout moment au cours du temps, on détermine les positions, vitesses et accélérations à l'étape suivante. Ainsi l'intégration selon des pas infinitésimaux conduit à la variation de l'énergie potentielle au cours du temps, ce qui donne la trajectoire du système.

Dans les simulations de dynamique effectuées dans cette thèse, la dynamique de Langevin a été utilisée. Cette dynamique utilise l'équation de Langevin qui permet de coupler le mouvement des atomes à un bain thermostatique par l'inclusion de termes stochastiques dans l'équation de Newton de la dynamique classique (déterministe) :

$$m_i\, v'_i = -(rvp\,(r_i, t) - m_i\, \gamma_i\, v_i\,(t) + R_i\,(t)$$

Cette équation contient en plus de la dynamique newtonienne un terme de friction dissipatif $m_i \gamma_i v_i(t)$, et une force aléatoire $R_i(t)$. L'effet des molécules de solvant sur le mouvement des atomes de soluté est donc pris en compte par les forces de frictions $f_i = m_i \gamma_i v_i(t)$, qui s'opposent au mouvement de l'atome *i* et par la force stochastique $R_i(t)$, due aux collisions associées aux mouvements browniens des particules environnantes.

III. A. 4. b| Les algorithmes d'intégration

La résolution analytique de l'équation du mouvement est complexe. Les positions atomiques sont calculées à des intervalles de temps discrets par des algorithmes d'intégration, également nommés propagateurs, tels que celui de Verlet (Verlet. 1967) ou encore celui du saut du mouton « Leapfrog » (Hockney *et al.* 1981).

C. 2. b. i| L'algorithme de Verlet

Il repose sur le triplet [$r_i(t)$, $r_i(t - \Delta t)$, $a_i(t)$]. En faisant un développement de Taylor d'ordre 3 des positions $r_i(t)$, on a

$$\mathbf{r}_i(t+\Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t)\Delta t + \frac{1}{2}\frac{\mathbf{F}_i(t)}{m_i}(\Delta t)^2 + \frac{1}{6}\frac{d^3\mathbf{r}}{dt^3}(\Delta t)^3 + \mathcal{O}(\Delta t)^4$$

De manière similaire,

$$\mathbf{r}_i(t-\Delta t) \;=\; \mathbf{r}_i(t) - \mathbf{v}_i(t)\Delta t + \frac{\mathbf{F}_i(t)}{2m_i}(\Delta t)^2 - \frac{1}{6}\frac{d^3\mathbf{r}}{dt^3}(\Delta t)^3 + \mathcal{O}(\Delta t)^4$$

En sommant ces 2 équations, on obtient l'expression de propagation des positions :

$$\mathbf{r}_i(t+\Delta t) + \mathbf{r}_i(t-\Delta t) = 2\mathbf{r}_i(t) + \frac{\mathbf{F}_i(t)}{m_i}(\Delta t)^2 + \mathcal{O}(\Delta t)^4$$

Le calcul des nouvelles positions est donc effectué avec une précision de l'ordre de $(\Delta t)^4$. Les vitesses sont quant à elles déterminées ainsi :

$$\mathbf{v}_i(t) = \frac{\mathbf{r}_i(t+\Delta t) - \mathbf{r}_i(t-\Delta t)}{2\Delta t} + \mathcal{O}(\Delta t)^2$$

Cette méthode de calcul des vitesses introduit des approximations dommageables à l'évaluation de l'énergie cinétique. Il est important de noter que l'algorithme de Verlet est réversible dans le temps. Il est dit symétrique. La conséquence de cette propriété est que, si à un instant *t* de la simulation on inverse le temps, la trajectoire revient sur ses pas, mais les erreurs d'arrondis de la simulation font que le processus n'est pas totalement réversible.

C. 2. b. ii| L'algorithme du saut du mouton

Les vitesses sont calculées pour des intervalles de temps semi-entiers et les positions sont obtenues dans les intervalles de temps entiers. Cet algorithme utilise le triplet [$r_i(t)$, $v_i(t - \Delta t/2)$, $a_i(t)$]. Si on définit les vitesses pour les temps $t + \frac{\Delta t}{2}$ et $t - \frac{\Delta t}{2}$:

$$\mathbf{v}_i\left(t + \frac{\Delta t}{2}\right) = \frac{\mathbf{r}_i(t+\Delta t) - \mathbf{r}_i(t)}{\Delta t}$$
$$\mathbf{v}_i\left(t - \frac{\Delta t}{2}\right) = \frac{\mathbf{r}_i(t) - \mathbf{r}_i(t-\Delta t)}{\Delta t}$$

On obtient alors immédiatement

$$\mathbf{r}_i(t+\Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i\left(t + \frac{\Delta t}{2}\right)\Delta t$$
$$\mathbf{r}_i(t-\Delta t) = \mathbf{r}_i(t) - \mathbf{v}_i\left(t + \frac{\Delta t}{2}\right)\Delta t$$

L'algorithme du leapfrog est identique à celui de Verlet quant au calcul des trajectoires. Les vitesses aux temps demi-entiers n'apparaissent que comme des intermédiaires de calcul. En revanche, les deux algorithmes diffèrent pour évaluer des grandeurs thermodynamiques. Ainsi, tandis que la moyenne de l'énergie potentielle peut être calculée aux temps entiers (grâce aux $\boldsymbol{r}_i(t)$), celle de l'énergie cinétique l'est aux temps demi-entiers (en faisant intervenir les vitesses) (Figure III.4).

En pratique, les valeurs calculées et gardées en mémoire sont, pour tous les atomes du système, les positions $r_i(t)$ et accélérations $a_i(t)$ instantanées et les vitesses au demi-temps d'intégration précédent v_i $(t - \delta_t/2)$. L'équation des vitesses est d'abord incrémentée de Δt pour obtenir vi $(t + \delta_t/2)$. Ensuite, on évalue les vitesses instantanées des atomes :

$$\mathbf{v}_i(t) = \frac{1}{2}\left[\mathbf{v}_i\left(t - \frac{\Delta t}{2}\right) + \mathbf{v}_i\left(t + \frac{\Delta t}{2}\right)\right]$$

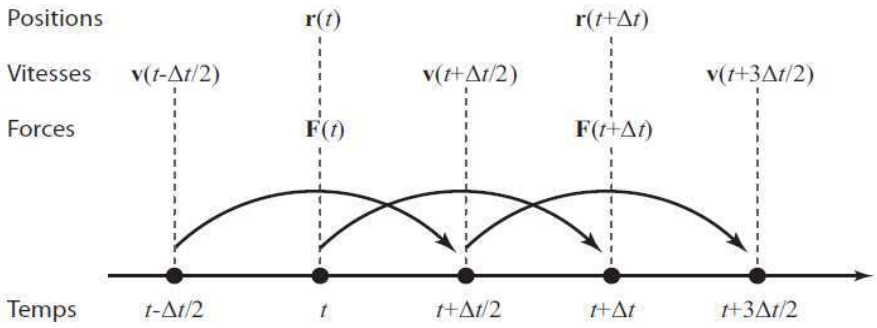


Figure III.4: Schéma de principe de l'algorithme du leapfrog.

III. A. 4. c| Durée du pas de temps

Les équations du mouvement sont résolues numériquement et propagées grâce à des algorithmes relativement rapides à mettre en œuvre. Néanmoins, ils imposent une restriction sévère sur la longueur du pas d'intégration. Un pas de temps Δt trop grand introduit des erreurs significatives dans l'évaluation de l'énergie et peut faire diverger la simulation. Il faut tenir compte des fréquences de vibrations les plus rapides. Dans les molécules organiques, les liaisons qui vibrent le plus rapidement sont les liaisons C−H (avec une période de vibration de l'ordre de 10 fs). La valeur maximale du pas d'intégration Δt doit être approximativement le dixième de cette période, soit 1 fs. L'algorithme **Shake** (Ryckaert *et al.* 1977) utilisé dans CHARMM permet de réduire les variations des longueurs de liaison dans la limite d'un seuil de tolérance choisi. Il permet d'augmenter l'intervalle de temps (δt = 2fs) en éliminant les oscillations les plus rapides associées aux liaisons chimiques du type X—H/X (X étant les atomes N, O, et S dans les protéines). On peut donc générer des simulations plus longues sans avoir à accroître le temps de calcul.

III. A. 4. d| Conditions initiales

Pour initier la simulation d'un système, l'algorithme d'intégration a besoin des coordonnées (x, y, z) et les vitesses (v_x, v_y, v_z) de tous les atomes au temps t_0. La position des atomes est fournie par l'expérience (structure RMN ou cristallographique) ou déduite des expériences *in silico* (construction de modèles par homologie). Une vitesse initiale est

nécessaire pour débuter la simulation et correspond à un apport énergétique. Cette énergie dépend elle-même de la température de simulation désirée qui est la seule information concernant les vitesses initiales. Plus la température est élevée, plus il y aura de changements conformationnels du système. Une vitesse initiale est attribuée aléatoirement à tous les atomes du système de masse *m*, selon une distribution de Maxwell-Boltzmann centrée sur la température T et donnée par la densité de probabilité ***f*** (v_i) :

$$f(\mathbf{v}_i) = \sqrt{\frac{m_i}{2\pi k_\mathrm{B} T}} \exp\left(-\frac{m_i v_i^2}{2k_\mathrm{B} T}\right)$$

Cette distribution des vitesses initiales n'est pas équilibrée. Une période d'équilibration est nécessaire afin de stabiliser la simulation et d'éviter des zones « chaudes » ou « froides » dans le système. La stabilisation est habituellement obtenue en élevant lentement la température, placée initialement au zéro absolu et en amenant progressivement le système vers l'équilibre.

III. A. 5| L'amarrage moléculaire

III. A. 5. a| Introduction

L'amarrage moléculaire automatique ou « docking » peut être défini comme la prédiction de la formation d'un complexe récepteur-ligand, où le récepteur est une protéine ou oligomère et le ligand considéré comme une petite molécule (peut-être aussi une autre protéine). Le principe repose sur l'hypothèse que les ligands formant des interactions favorables avec le récepteur sont ceux présentant une affinité élevée. Aujourd'hui le « docking » est un outil indispensable dans la recherche moderne de molécules thérapeutiques où il peut aider à la découverte et l'optimisation de molécules prototypes « lead » à travers le criblage virtuel « Virtual Screening » (Kitchen *et al.* 2004; Simmons *et al.* 2010).

Il existe différents algorithmes de « docking ». Les critères de choix sont le plus souvent le nombre de molécules à tester et la précision souhaitée, critères qui ont un impact sur le temps de calcul et la disposition des moyens nécessaires. Les méthodes de « docking » les plus rapides sont les méthodes dites de « docking » rigide, où le récepteur et le ligand sont figées, historiquement les premières conçues (Kuntz *et al.* 1982) (Levinthal *et al.* 1975) (Salemme. 1976) (Wodak *et al.* 1978). Aujourd'hui avec l'augmentation des puissances de calculs, les méthodes de « docking » dites «semi-flexibles» et même « flexibles », où les partenaires d'interaction sont partiellement déformables, se sont imposées, car elles permettent d'obtenir des résultats plus précis, même dans des approches coûteuses de « criblage virtuel haut débit ».

Les processus de « docking » comprennent généralement deux étapes essentielles : une étape, dite d'*exploration*, et une autre dite de « *scoring* ».

III. A 5. a. i| L'étape d'*exploration*

Cette étape permet aux ligands d'explorer toutes les configurations accessibles dans une zone ou un site du récepteur afin de trouver le meilleur candidat, sous-entendu le complexe ayant l'énergie libre de liaison la plus basse. Le ligand adopte plusieurs conformations possibles au cours de cette exploration, et les programmes génèrent plusieurs positions et conformations possibles du ligand dans le récepteur. Plusieurs conformations du ligand dans le site d'interaction résultent du processus de « docking » (poses). Ces poses

correspondent à des conformations favorables que le ligand peut adopter dans un site d'action. Trois approches pour l'exploration du ligand sont couramment utilisées.

- La recherche systématique

Ces méthodes exploitent tous les degrés de liberté des angles de torsion du ligand sous certains critères de contraintes (permettant de limiter l'exploration si la conformation du ligand ne remplit pas certaines conditions). Dans un premier temps, la molécule est découpée en parties rigides et en liaisons flexibles. Les parties rigides sont positionnées en premier et le ligand est reconstruit progressivement, explorant diverses conformations lors de la mise en place des liaisons flexibles. Parmi les logiciels fonctionnant sur ce principe, figure Glide qui a été utilisé aussi dans cette thèse.

- La recherche stochastique

Le ligand est considéré dans son entier. Les mouvements de rotations, de translations et les variations des angles de torsion se font au hasard. Cette recherche aléatoire peut se faire par des approches de type Monte Carlo ou algorithmes génétiques. L'algorithme de Monte Carlo permet de contrôler cette exploration aléatoire grâce au critère de Métropolis. AutoDock fut le premier à utiliser cet algorithme, même si ce programme offre d'autres approches stochastiques jugées plus performantes (algorithme génétique, algorithme hybride Lamarckien). Les algorithmes génétiques génèrent une population de conformations et de positions du ligand au hasard, qui évolue selon des événements génétiques (mutations, crossing-over) régis par un taux de probabilité et une pression de sélection fixés. Un taux de survie définit le nombre d'individus passant d'une génération à l'autre. Le nombre de générations maximum est fixé, mais un critère de convergence peut arrêter l'algorithme avant. Nous détaillerons très largement ces algorithmes par la suite.

- La recherche déterministe

La dynamique moléculaire et la minimisation d'énergie sont utilisées pour trouver des conformations d'énergie la plus basse possible. Les conformations futures des molécules sont guidées en fonction des conformations antérieures.

III. A 5. a. ii| L'étape de « *scoring* »

Cette étape permet d'évaluer les conformations du complexe rencontrées lors de l'étape d'exploration pour ne retenir que les meilleures. Cette évaluation repose généralement sur une fonction énergétique qui doit être la plus réaliste possible en tenant compte de toutes les interactions et reproduire les résultats observés expérimentalement. La fonction de « scoring », qui permet de classer les modèles de complexes, est basée généralement sur l'évaluation des contributions entropiques et enthalpiques à la stabilité du complexe et dépend du logiciel utilisé. Les méthodes de « scoring » permettent d'évaluer l'énergie de liaison du complexe formé et de donner une valeur énergétique aux poses obtenues lors de la phase de « docking ». Ces valeurs permettront de retenir la meilleure pose parmi toutes celles proposées.

Plusieurs programmes ont été testés au début de cette thèse afin de trouver le meilleur logiciel adapté au cas particulier de cette étude, c'est-à-dire prédire l'amarrage d'inhibiteurs « *peptidiques* » sur la protéine EcFur dont le site d'action n'est pas connu. Il s'agit du « docking » en aveugle ou « blind docking » (prédire l'amarrage de molécules sur une protéine sans une connaissance préalable du site d'action du ligand) (Hetenyi *et al.* 2006; Iorga *et al.* 2006). Ces programmes sont :

- 3D-dock (Moont *et al.*, 2001) version 2 et FTDOCK.
- MULTIDOCK v1.0 (Jackson *et al.* 1998).
- DOCK v4.0 (Ewing *et al.* 2001).
- Le module FlexX de SYBYL (Keseru. 2001).
- GOLD (Verdonk *et al.* 2003).
- GLIDE (Friesner *et al.* 2004; Friesner *et al.* 2006).
- AutoDock4 (Huey *et al.* 2007; Morris *et al.* 2009).

Les trois premiers logiciels cités utilisent une approche d'exploration du partenaire moléculaire basée sur la complémentarité stérique. Cette approche est utile lorsqu'il s'agit de prédire l'interaction entre deux protéines. Les quatre derniers logiciels utilisent une approche plus précise basée sur les interactions chimiques et prennent en compte l'effet du solvant, la flexibilité du ligand et souvent de la protéine.

Dans cette étude, nous avons finalement choisi AutoDock4 pour prédire l'amarrage des peptides anti-Fur et partant des résultats obtenus, des études de criblages virtuels ont été entreprises avec GLIDE et GOLD pour cribler virtuellement des petites molécules, sélectionnées dans la base de données « ZINC » (Irwin *et al.* 2005).

III. A. 5. b| Amarrage avec AutoDock

AutoDock est un programme d'amarrage automatique de ligands flexibles sur des récepteurs. Il combine une méthode d'évaluation de l'énergie à partir de grilles de potentiels avec différentes méthodes d'exploration allant de la dynamique moléculaire (méthode Métropolis : recuit simulé Monte Carlo - RSMC) aux algorithmes génétiques. AutoDock est sous licence gratuite, disponible sur le site http://autodock.scripps.edu/.

AutoDock fait partie, avec GOLD et Glide, des logiciels couramment utilisés pour les études de « docking ». Plusieurs études ont démontré l'efficacité d'AutoDock dans l'amarrage en aveugle (Morris *et al.* 2009) (Huey *et al.* 2007) (Hetenyi *et al.* 2006). Nous avons choisi d'utiliser AutoDock pour la prédiction de l'amarrage des peptides anti-Fur sur la protéine Fur, car il permet une plus grande flexibilité du ligand pour explorer la protéine, comparé à GOLD et Glide. Ce qui s'est avéré très intéressant car ces peptides contenant 13 acides aminés ont un nombre important d'angles de torsion à échantillonner.

III. A. 5. b. i| Modélisation du potentiel énergétique

Pour calculer l'énergie libre du complexe ligand-récepteur, AutoDock4.2 utilise un champ de force basé sur l'énergie libre calculée de façon semi-empirique. Ce champ de force a été paramétré en utilisant un grand nombre de complexes protéine-inhibiteur pour lesquels les structures et les constantes d'inhibition sont connues (Morris *et al.* 2009). Il évalue la liaison protéine-ligand en deux étapes (Figure III.5) : le champ de force permet d'évaluer les énergies intramoléculaires protéine/ligand lors de la transition de l'état libre à l'état lié, puis estime les énergies intermoléculaires combinant le ligand et la protéine dans l'état lié.

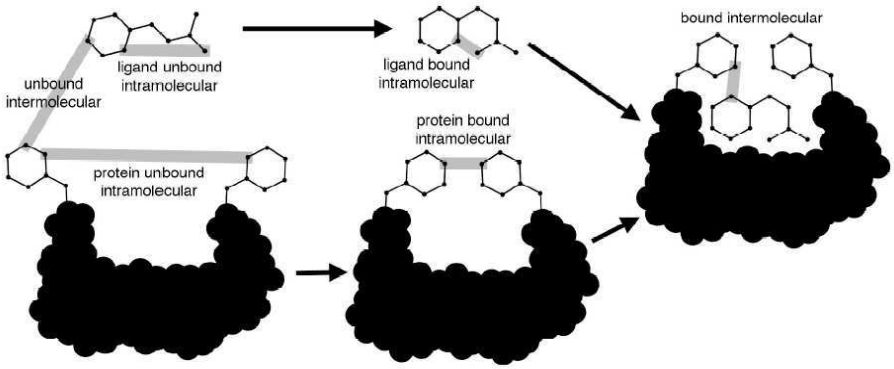


Figure III.5: Les étapes d'évaluation de la liaison protéine-ligand par AutoDock4.2 (Morris *et al.* 2009)

L'équation décrivant l'énergie libre de liaison du ligand à la protéine, inclut six évaluations d'énergies par paires (V) auxquelles on ajoute une estimation de la perte d'entropie lors de la liaison du ligand au récepteur. L'énergie libre est décrite de la manière suivante :

$$\Delta G = (V_{\downarrow}lié^{\uparrow}(L-L) - [\![V]\!]_{\downarrow}(non\ lié)^{\uparrow}(L-L)) + (V_{\downarrow}lié^{\uparrow}(P-P) - [\![V]\!]_{\downarrow}(non\ lié)^{\uparrow}(P-P)) + (V_{\downarrow}lié^{\uparrow}(P-L) - V_{\downarrow}(non\ lié)^{\uparrow}(P-L) + [\![\Delta S]\!]_{\downarrow}conf)$$

Les lettres L et P réfèrent respectivement au ligand et au récepteur (protéine).

Le terme ΔS_{conf} est un terme qui traduit l'augmentation d'énergie du système due à la restriction des angles dièdres libres « rotors » du ligand et à la restriction des rotations et translations du ligand lors de la complexation au récepteur. Cette perte de degrés de liberté est une perte d'entropie du ligand lorsqu'il se lie au récepteur, prise en compte dans la fonction d'énergie. Elle est proportionnelle au nombre de liaisons sp^3 dans le ligand.

Chaque fonction V peut être décomposée comme suit :

$$V = W_{vdw}\sum_{i,j}\left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right) + W_{hbond}\sum_{i,j}E(t)\left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}}\right) + W_{elec}\sum_{i,j}\frac{q_i q_j}{\varepsilon(r_{ij})r_{ij}} + W_{sol}\sum_{i,j}\left(S_i V_j + S_j V_i\right)e^{(-r_{ij}^2/2\sigma^2)}$$

Les trois premiers termes sont des termes classiques de mécanique moléculaire déjà détaillés; il s'agit respectivement des énergies dispersion / répulsion des atomes, des liaisons hydrogène et des interactions électrostatiques, décrits précédemment auxquels on rajoute un autre terme lié à l'entropie de solvatation qui décrit les variations d'énergie du système lors de la désolvatation du ligand au moment de la complexation au récepteur. C'est une modélisation partielle de ce que l'on appelle « l'effet hydrophobe » dû aux variations de l'entropie du solvant aux interfaces solvant – soluté.

Les liaisons hydrogènes sont modélisées par un potentiel directionnel de Lennard-Jones (12-10). Ce potentiel est pondéré en fonction de l'angle t formé entre la direction du doublet libre de l'atome accepteur (A), l'atome d'hydrogène (H) et l'atome qui le porte (D) (Figure III.6).

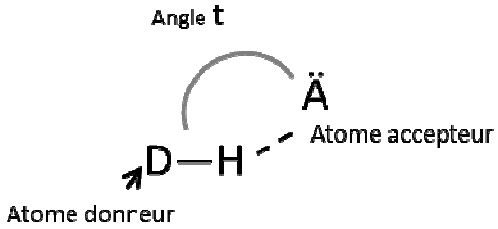


Figure III.6: Formation de la liaison hydrogène

L'angle optimum est de 180° alors que pour un angle inférieur à 90° la liaison devient impossible et la fonction ***E(t)*** devient nulle. Cette fonction pénalise aussi les atomes polaires du ligand qui ne forment pas de liaisons hydrogène avec le récepteur, favorisant ainsi les solutions exploitant au mieux le potentiel de liaisons hydrogène du site actif.

L'énergie nécessaire à la désolvatation du ligand lors de l'association au récepteur est rapportée par le dernier terme de l'équation ci-dessus. Elle est évaluée par une méthode basée sur les volumes, permettant de calculer la contribution de chaque atome à l'énergie totale de désolvatation du ligand. Pour chaque atome considéré, on calcule la proportion du volume ***V*** des atomes qui l'entourent qu'on pondère par le paramètre de solvatation atomique S et un terme exponentiel avec un coefficient de pondération de la distance $\boldsymbol{r_{ij}}$: σ = 3,5Å.

Les termes ***W,*** qui pondèrent chaque terme de l'équation ci-dessus, sont les coefficients empiriques déterminés par l'analyse en régression linéaire d'un jeu de complexes protéine–ligand pour lesquels la constante d'inhibition Ki est connue (Huey *et al.* 2007). Les sommes sont faites sur l'ensemble des couples formés d'un atome du ligand i et d'un atome de la protéine j et sur toutes les paires d'atomes du ligand séparés par au moins 3 liaisons covalentes.

La relation permettant le lien entre la constante d'inhibition et l'énergie libre de liaison est donnée par l'Équation suivante :

$$\Delta G = RT \ln Ki$$

Ou R est la constante des gaz parfaits (8,31 J. K^{-1}.mol^{-1}) et T la température absolue.

III. A. 5. b. ii| Le processus AutoDock

Le processus d'AutoDock comprend 4 étapes qui sont :

- La préparation des coordonnées initiales
- La préparation de la grille par AutoGrid
- L'amarrage automatique par AutoDock
- L'évaluation des résultats de « docking »

➢ La préparation des coordonnées initiales

La première étape consiste à préparer le ligand et le récepteur en définissant les charges partielles et les types des atomes et en ajoutant les hydrogènes polaires avec le module *AutoTors*, implémenté dans *AutoDockTools* (ADT). Cette opération transforme le fichier PDB au format PDBQT, fichier qui contient les informations sur les angles à bouger pendant l'exploration du récepteur par le ligand. Les charges attribuées aux atomes sont de type "Kollman" et sont calculées avec le champ de forces empirique Amber (Weiner et Kolmann, 1981).

➢ La préparation de la grille par AutoGrid

Un très grand nombre d'évaluations d'énergie est nécessaire tout au long de la simulation pour optimiser l'amarrage des ligands et ce calcul doit donc être extrêmement rapide. Le récepteur étant rigide, son champ d'interaction intermoléculaire est constant. AutoDock utilise une méthode d'évaluation de l'énergie par l'intermédiaire de grilles de potentiel d'affinité atomique pré-calculées pour chaque type d'atome. Une grille est une matrice tridimensionnelle qui englobe l'intégralité ou une région intéressante du récepteur étudié (Figure III.7). Chaque point de la grille enregistre le potentiel d'interaction entre une partie du ligand et l'ensemble des atomes du récepteur. Le calcul de ces grilles est effectué par AUTOGRID, un programme d'AutoDock, qui calcule automatiquement les grilles nécessaires parmi les types C, N, O, H, C, X et M (où X représente les halogènes et M les ions métalliques communs). Une grille supplémentaire est calculée pour le potentiel d'interaction électrostatique avec pour sonde une charge ponctuelle de +1 eV.

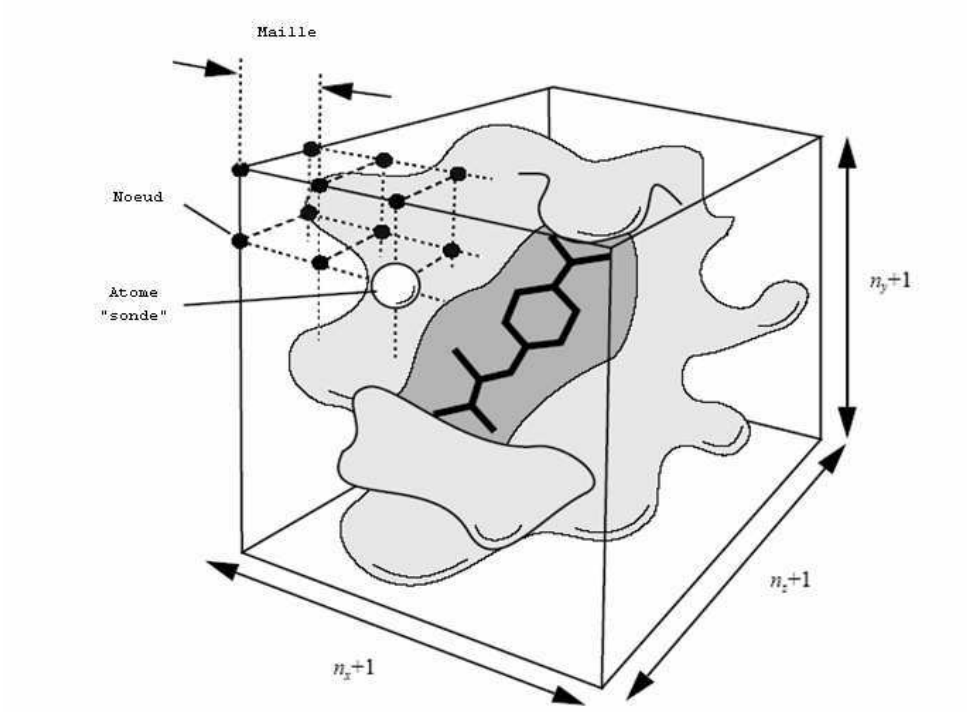


Figure III.7: Schéma d'une grille de potentiel englobant le site actif d'une macromolécule

Les dimensions de la matrice définissent une boîte qui sera la zone à explorer par le ligand. La boîte doit englober le site d'intérêt mais le temps de calcul pour l'explorer sera fonction de ses dimensions. Le volume explorable est donc limité par le temps et la puissance de calcul dont on dispose. L'énergie d'interaction du ligand dans une configuration particulière est évaluée en faisant une interpolation tri-linéaire des valeurs d'affinité des huit points de grille qui entourent chaque atome du ligand. Le temps de calcul de l'énergie dépend uniquement du nombre d'atomes dans le ligand et est indépendant de la taille de la macromolécule.

- L'amarrage automatique par AutoDock

AutoDock permet d'explorer l'espace conformationnel avec, au choix, un algorithme génétique darwinien (AG), un algorithme génétique lamarckien (AGL) ou par recuit simulé par la méthode de Monté Carlo (RSMC). Des trois méthodes, l'algorithme génétique lamarckien qui donne les résultats les plus fiables et les temps de calcul les plus courts (Brooijmans *et al.* 2003) a été utilisé pour tous les amarrages réalisés dans cette étude. Avant de décrire cette méthode, nous allons expliquer le principe des algorithmes génétiques.

- Les Algorithmes Génétiques (AG).

Les algorithmes génétiques reprennent des mécanismes et la terminologie de la génétique naturelle et de l'évolution biologique. L'organisation d'un ligand en complexe avec un récepteur peut être définie par un jeu de paramètres décrivant la position, l'orientation et la conformation du ligand par rapport au récepteur. Ces paramètres sont les « variables d'état » et dans un AG, chaque variable d'état correspond à un gène. La valeur de ces variables correspond au génotype et les coordonnées atomiques associées correspondent au phénotype. Chaque état défini du ligand correspond à un individu. Pour faire évoluer le ligand dans le champ d'interaction du récepteur, on emploie un certain nombre d'opérateurs qui vont agir sur le génotype. Le phénotype qui en découle sera amélioré par une méthode de sélection. Des paires aléatoires d'individus sont combinées selon le principe du croisement (*crossover*) pour donner des individus fils qui héritent de gènes provenant de leurs deux parents. D'autre part, certains enfants peuvent être le résultat de mutations où un gène est modifié de façon aléatoire. La sélection des individus fils constituant la nouvelle génération est basée sur la qualité de leur interaction avec le récepteur : les solutions qui s'ajustent mieux au récepteur que leurs parents persistent alors que les autres disparaissent. Le critère d'évaluation de la qualité d'une solution d'amarrage est l'énergie libre totale du système ligand– récepteur.

- L'Algorithme Génétique Lamarckien (AGL).

La plupart des algorithmes génétiques reproduisent le comportement de l'évolution darwinienne en appliquant le principe de la génétique de Mendel c'est-à-dire le transfert à sens unique des informations du génotype vers le phénotype. Ce comportement est illustré par la partie droite de la Figure III.8. Par contre, dans le cas où il existe un mécanisme de transcription inverse, un génotype peut être induit par un phénotype. Il est alors possible, pour un individu, d'acquérir de nouveaux caractères génétiques en fonction de son environnement. Les enfants pourront hériter, à leur tour, de ces caractères acquis durant la vie de leur parent. Dans le cadre de l'interaction ligand-récepteur, on peut ainsi effectuer une optimisation locale du ligand par rapport au récepteur et remonter les informations du phénotype optimisé vers le génotype de l'individu. Ce comportement est illustré par la partie gauche de la Figure III.8. C'est ce qu'on appelle un algorithme génétique lamarckien par analogie avec la théorie, aujourd'hui discréditée, de Jean Baptiste de Lamarck selon laquelle les caractéristiques acquises par un individu durant sa vie pouvaient devenir héréditaires.

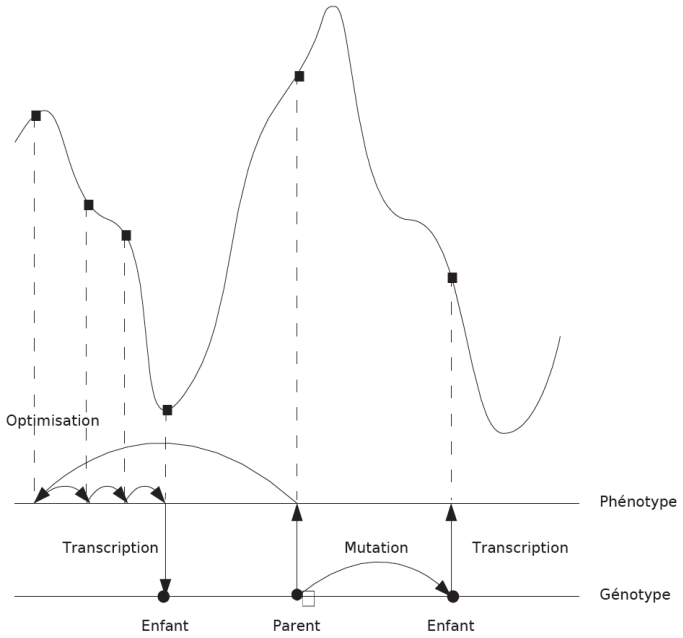


Figure III.8: Comportement des algorithmes génétiques. Principes de Darwin (à droite) et de Lamarck (à gauche)

AutoDocK dispose d'un AGL dont la phase d'optimisation locale est particulière. Elle utilise une variante de la méthode de Solis - Wets (Huey *et al.* 2007) où l'opérateur travaille sur l'espace génotypique du ligand pour minimiser son énergie alors que la plupart des algorithmes classiques travaillent sur l'espace phénotypique. L'intégration d'une fonction de traduction inverse n'est donc pas nécessaire mais cette combinaison de recherche globale et de recherche locale reste de type lamarckien. En effet, toutes les adaptations environnementales du ligand acquises pendant l'optimisation locale seront transmises à ses enfants, s'il en a. La méthode de Solis - Wets procède avec un opérateur semblable à l'opérateur de mutation de l'AG mais le gène affecté subit une modification qui n'est pas aléatoire. Au contraire, l'opérateur modifie les gènes affectés, un par un, par pas réguliers et identiques pour chaque gène. De plus la méthode est adaptative. Elle ajuste la taille du pas en fonction de l'historique énergétique des optimisations : après un nombre déterminé de hausses consécutives de l'énergie, la taille du pas est doublée. A l'inverse, après un nombre déterminé de baisses consécutives de l'énergie, la taille du pas est divisée par deux. La méthode d'optimisation locale implémentée dans AutoDock est une variante de Solis – Wets dans laquelle la taille du pas est différente pour chaque type de gène : une variation de 1 Å dans un gène de position aura beaucoup plus d'impact qu'une variation de 1 degré dans un gène d'orientation ou de torsion. Aussi la taille du pas pour un gène de position est paramétrée par

défaut à 1,0 Å alors que la taille du pas pour un gène d'orientation ou de torsion est par défaut de 50°.

A chaque génération, il est possible de faire une optimisation locale sur une fraction de la population. L'efficacité de l'amarrage est sensiblement améliorée avec une fréquence d'optimisation locale de seulement 6 % alors que le gain supplémentaire pour une fréquence de 100 % est très faible.

- L'évaluation des résultats de « docking »

A la fin d'un processus de « docking » par AutoDock4, les coordonnées de chaque conformation du ligand sont stockées dans un fichier avec les informations d'énergie et de classement en familles de conformations « cluster ». Ce fichier est lu par l'interface ADT afin de visionner les différentes configurations du complexe, de sélectionner et enregistrer les meilleures conformations d'amarrage.

III. A. 5. c| Criblage virtuel avec Glide

Nous avons également utilisé Glide et Gold qui sont des logiciels commerciaux d'amarrage molécule et criblage virtuel à haut débit. Dans cette étude, nous avons utilisés le logiciel d'amarrage automatique de Schrödinger (http://www.schrodinger.com/): Glide. Ce logiciel largement utilisé pour le criblage des petites molécules provenant des bases de données virtuelles. La Direction des Sciences du Vivant (DSV) du CEA dispose d'une licence pour un ensemble des programmes de Schrödinger à travers son « cluster ». Nous avons utilisé Glide pour le criblage virtuel de petites molécules provenant de la banque de données ZINC. Nous n'allons pas développer le mode opératoire de ces deux programmes. Le prince et la méthodologie de Glide seront donnés en annexe.

III. B| Mode opératoire de l'amarrage

La méthodologie employée afin de prédire l'amarrage des peptides issus des aptamères anti-Fur sur la cible Fur peut être décomposée en quatre parties:

1. Recherche et optimisation des structures initiales de la protéine et des peptides.
2. Recherche de la zone d'amarrage des peptides sur la protéine.
3. Amarrage moléculaire Assisté par la Dynamique Moléculaire (AADM).
4. Identification des interactions clés.

III. B. 1| Recherche et optimisation des structures initiales

Comme il a été abordé dans l'introduction de cette partie théorique, les structures initiales sont pré-requises pour pouvoir effectuer des simulations. Pour réaliser l'amarrage des peptides sur la protéine Fur dans les meilleures conditions possibles, il aurait fallu avoir les structures de ces molécules. A défaut de ces structures, il faut des modèles. Nous avons construit et optimisé des modèles avec les logiciels MODELLER et CHARMM. Les modèles sélectionnés ont constitué les coordonnées initiales pour l'amarrage moléculaire par AutoDock et les autres logiciels de « docking ». Le processus d'obtention de ces coordonnées initiales est développé ci-dessous.

III. B. 1. a| Structures initiales des aptamères peptidiques et des peptides

Les peptides de 13 acides aminés issus des aptamères peptidiques (confère tableau Tableau III.1, ci-dessous) n'ont aucune homologie avec des structures connues, d'après une recherche BLAST (http://blast.ncbi.nlm.nih.gov) (Altschul *et al.* 1990). Des études de dichroïsme circulaire ont également montré que les peptides ne disposent d'aucune structure particulière. Les peptides ne sont pas assez longs pour pouvoir effectuer une recherche de structure secondaire sur les serveurs de prédiction. Des modèles ont donc été construits avec CHARMM et MODELLER en utilisant la méthode décrite ci-dessous.

La structure de la thiorédoxine A d'*E. coli*, résolue en 1990 à 1,68 Å (Code PDB : 2TRX) (Katti *et al.* 1990) (Figure III.9) a été utilisée comme plateforme pour construire les aptamères peptidiques (confère Introduction) de manière similaire aux travaux récents de

Brown *et al* (Brown *et al.* 2010). Ces derniers ont montré que l'insertion des peptides déstabilisait un peu la thiorédoxine de sorte que les peptides se trouvaient dans une conformation leur permettant d'interagir avec des protéines cibles.

Tableau III.1: Séquences des peptides issus des aptamères anti-EcFur

peptides	Séquences
pF1	RLWCRYPHPPLTD
pF2	RQCNICGASLYSY
pF3	ETCKCGSQVWRHS
pF4	CARCGARVNVYKY
pF5 (contrôle inactif)	RSLRGRCLSQHQD

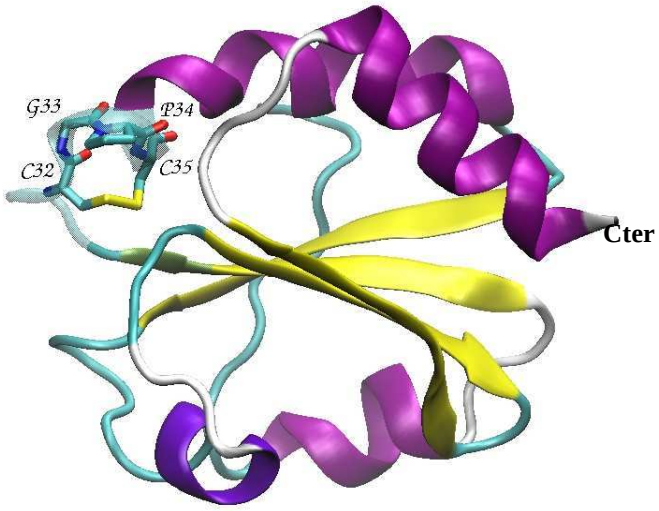


Figure III.9: Structure de la thiorédoxine A d'*E. coli* (2TRX) (Katti *et al.* 1990). Représentation en ruban où seules les chaînes latérales des résidus 32 à 34 sont montrées.

Cette structure comprend 4 feuillets β en jaune, 3 hélices α en rose une hélice D10 en violet, reliés entre eux par des boucles. Les résidus C32, G33, P34 et C35 du motif CXXC constituant du site actif de la thiorédoxine sont indiqués.

Pour mimer les aptamères peptidiques, les peptides de 13 acides aminés sont insérés dans le site actif entre la Pro34 et la Cys35 de la thiorédoxine (Brown *et al.* les ont insérés entre G33 et P34) (Brown *et al.* 2010). Deux résidus (Gly et Pro) sont ajoutés à l'extrémité Cter du peptide pour servir de « linker ». Les Cystéines 32 et 35 de la thiorédoxine sont

remplacées par des Sérines pour rester conforme à la séquence de l'aptamère utilisée dans les tests double-hybride antérieurs (Abed *et al.* 2007) (confère Figure III.10).

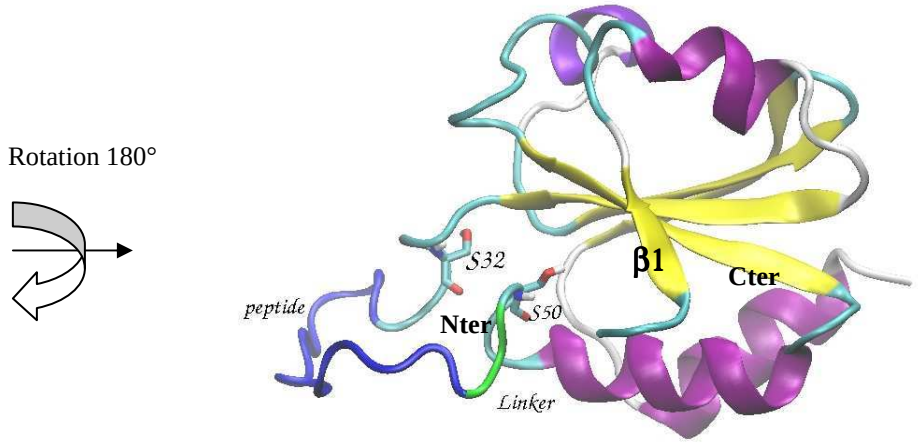


Figure III.10 : Modèle de construction d'un aptamère peptidique. Représentation en ruban où seules les chaînes latérales des résidus sérines remplaçant les cystéines sont montrées.

Pour construire la boucle constituée par le peptide, on utilise une fonctionnalité de MODELLER (Loop Modelling) qui permet de générer des modèles de boucles compatibles avec le reste de la protéine supposé fixe. Ainsi, dix modèles pour chaque aptamère sont construits dans MODELLER et classés suivant une fonction de pénalité (« objective function »). Afin d'obtenir une structure optimisée suivant le champ de force de MM habituellement utilisé dans ce travail, le modèle ayant la pénalité la plus petite est choisi puis optimisé dans CHARMM en figeant la chaine carbonée peptidique - appelée squelette ou « backbone » - de la thiorédoxine. Ainsi les chaines latérales de la thiorédoxine et le peptide+le *linker* sont flexibles pendant la simulation d'optimisation. Le peptide peut explorer l'espace conformationnel tout en étant fixé au niveau de ses extrémités à la thiorédoxine. Ces contraintes forcent le peptide à rester conforme à l'aptamère expérimental. L'ensemble est soumis à une dynamique moléculaire de 5 ns à 298 K dans un solvant implicite (méthode EEF1) qui génère 5000 « poses » ou « frames ». Les « frames » sont les enregistrements des coordonnées atomiques au cours du temps dans un processus de dynamique moléculaire qui produit une trajectoire. L'analyse de la trajectoire permet de choisir la conformation ayant l'énergie minimale après minimisation avec 100 de SD et 1000 pas d'ABNR.

Nous présenterons ci-dessous un modèle de l'aptamère F1 et l'extraction du peptide pF1 qui sera utilisé pour l'amarrage sur la protéine Fur (Figure III.11).

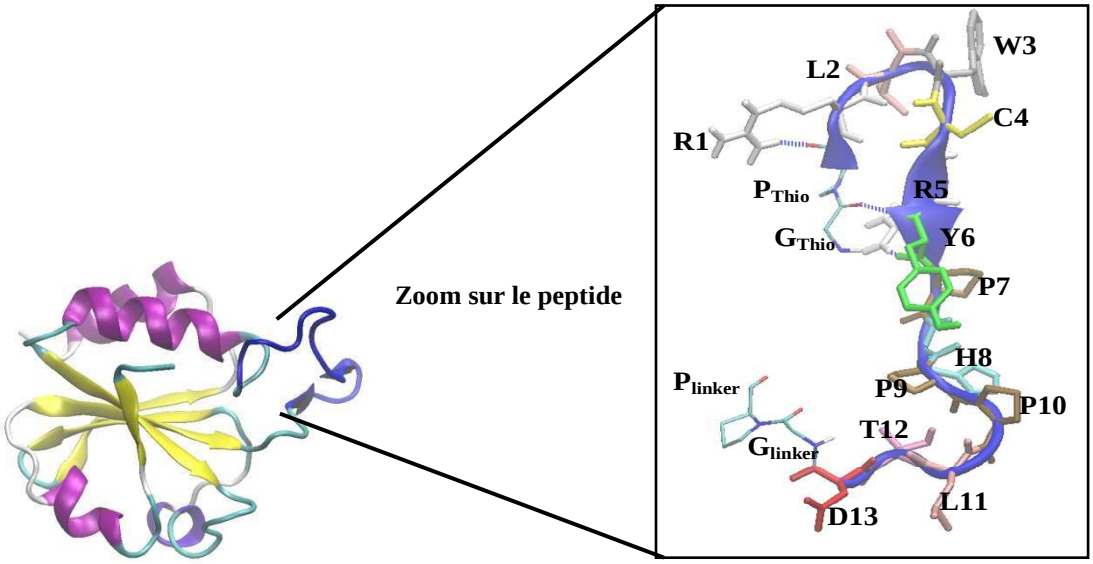


Figure III.11: Modèle de l'aptamère F1 et extraction du peptide pF1. Représentation en ruban et zoom sur la partie du peptide inséré où les chaînes latérales des résidus sont montrées ainsi que les liaisons hydrogènes en hachurés et le nom des résidus.

Le peptide de 13 acides aminés est extrait de l'aptamère pour le « docking ». On constate que certains résidus du peptide peuvent former des liaisons hydrogènes avec la thiorédoxine ce qui participe à la formation d'une structure secondaire. En effet sur la Figure III.11, les résidus Gly et Pro forment des liaisons hydrogènes avec l'Arg5 et la Tyr6 du peptide pF1, formant un feuillet β. L'aptamère pourrait donc aider à la formation d'une conformation particulière du peptide permettant une meilleure interaction avec la protéine. Toutefois l'inactivation de la protéine Fur est due à la spécificité du peptide de 13 acides aminés, la partie active. Notons également que le peptide extrait est rendu flexible pendant le processus d'amarrage moléculaire sur la protéine. Nous verrons aussi que le peptide libre pF1 peut perdre et la retrouver sa structure secondaire dans l'approche de « docking » que nous utilisons.

Des simulations d'amarrage ont été réalisées avec les peptides pF1 à pF3 extraits des meilleures conformations des aptamères F1 à F3 obtenues avec MODELLER et CHARMM pour rechercher les meilleures zones d'amarrage sur la protéine EcFur.

Le squelette carboné de ces peptides ne bouge pas au cours des processus d'amarrage qui vont suivre. De plus, les tests expérimentaux ayant montré que les peptides libres de 13 acides aminés inhibent la protéine Fur, des conformations peptidiques libres (sans contraintes de l'aptamère) sont aussi construites pour l'amarrage sur la protéine. Ces derniers sont flexibles au niveau de la chaine carbonée pendant le processus d'arrimage. Nous disposons

ainsi de deux modèles (libres et contraints) de chaque peptide qui vont être amarrés sur la protéine afin de choisir le meilleur mode d'amarrage.

Ne disposant pas d'informations structurales concernant le domaine de dimérisation du monomère de Fur et sachant que le peptide pF4 interagit à ce niveau, nous avons décidé de ne pas poursuivre les simulations d'amarrage avec ce peptide.

III. B. 1. b| structures initiales de la protéine

La structure de la protéine EcFur qui nous intéresse n'étant pas entièrement connue, une modélisation par homologie avec MODELLER a été effectuée. Dans cette approche, le logiciel génère plusieurs modèles à l'image de la structure de protéines homologues. Pour choisir les homologues les plus proches d'EcFur, un alignement de séquences des protéines Fur a été réalisé (confère Figure III.12).

```
                                   10        20        30        40
EcFur   ..........MTDNNTALKKAGLKVTLPRLKILEVLQEPDNHHVSAEDLYKRLI    44
YpFur   ..........MTDNNKALKNAGLKVTLPRLKILEVLQNPACHHVSAEDLYKILI    44
VcFur   ..........MSDNNQALKDAGLKVTLPRLKILEVLQQPECQHISAEELYKKLI    44
VsFur   ..........MSDNNQALKKAGLKVTLPRLKILEILQQPESQHISAEDLYKKLI    44
PaFur   ..........MVENS.ELRKAGLKVTLPRVKILQMLDSAEQRHMSAEDVYKALM    43
HpFur   MKRLETLESILERLRMSIKKNGLKNSKQREEVVSVLYR.SGTHLSPEEITHSIR    53
SaFur   ........MNTNDAIKILKENGLKYTDKRKDMLDIFVE.EDKYINAKYIQQVMD    45
MtZur   ........MASAAGVRSTRQR.....AAISTLLETLD..DFR..SAQELHDELR    37
ScNur   ........MVSTDWKSDLRQRGYRLTPQRQLVLEAVD..TLEHATPDDILGEVR    44
BsPrr   .....MAAHELKEALETLKETGVRITPQRHAILEYLVN.SMAHPTADDIYKALE    48

            50        60        70        80        90
EcFur   DMGEEIGLATVYRVLNQFDDAGIVTRHNFEGGKSVFELTQQHHHDHLICLDCGK    98
YpFur   DIGEEIGLATVYRVLNQFDDAGIVTRHNFEGGKSVFELTQQHHHDHLICLDCGK    98
VcFur   DLGEEIGLATVYRVLNQFDDAGIVTRHHFEGGKSVFELSTQHHHDHLVCLDCGE    98
VsFur   DIGEEIGLATVYRVLNQFDDAGIVTRHHFEGGKSVFELATQHHHDHLVCLDCGK    98
PaFur   EAGEDVGLATVYRVLTQFEAAGLVVRHNFDGGHAVFELADSGHHDHMVCVDTGE    97
HpFur   QKDKNTSISSVYRILNFLEKENFICVLETSKSGRRYEIAAKEHHDHIICLHCGK   107
SaFur   ENYPGISFDTIYRNLHLFKDLGIIENTELDGEMKFRIACTNHHHHHFICEKCGD    99
MtZur   RRGENIGLTTVYRTLQSMASSGLVDTLHTDTGESVYRRCSEHHHHHLVCRSCGS    91
ScNur   KTASGINISTVYRTLELLEELGLVSHAHLGHGAPTYHLADRHHHIHLVCRDCTN    98
BsPrr   GKFPNMSVATVYNNLRVFRESGLVKELTYGDASSRFDFVTS.DHYHAICENCGK   101

        100       110       120       130       140
EcFur   VIEFSDDSIEARQREIAAKHGIRLTNHSLYLYGHCAE.GDCREDEHAHEGK.   148
YpFur   VIEFSNESIESLQREIAKQHGIKLTNHSLYLYGHCET.GNCREDESAHSKR.   148
VcFur   VIEFSDDVIEQRQKEIAAKYNVQLTNHSLYLYGKCGSDGSCKDNPNAHKPKK   150
VsFur   VIEFSDDLIEERQKDIANKYNVRLTNHSLYLYGHCIT.GDCIENTTAHNG..   147
PaFur   VIEFMDAEIEKRQKEIVRERGFELVDHNLVLY..............VRKKK.   134
HpFur   IIEFADPEIENRQNEVVKKYQAKLISHDMKMFVWCKE...CQESEC......   150
SaFur   TKVIDYCPIDQIKLSLP...GVNIHKHKLEVYGVCES...CQD.........   136
MtZur   TIEVGDHEVEAWAAEVATKHGFSDVSHTIEIFGTCSD...CRS.........   131
ScNur   VIEADLSVAADFTAKLREQFGFDTDMKHFAIFGRCES...CSLKGSTTDS..   145
BsPrr   IVDFHYPGLDEVEQLAAHVTGFKVSHHRLEIYGVCQE...CSKKENH.....   145
```

Figure III.12: Alignement des homologues de Fur de chez *Escherichia coli* (EcFur) : *Yersinia pestis* (YpFur), *Vibrio cholerae* (VcFur), *Vibrio salmonicida* (VsFur), *Pseudomonas aeruginosa* (PaFur), *Helicobacter pylori* (HpFur), *Staphylococcus aureus* (SaFur) et les « Fur-like »: Zur de chez *Mycobacterium tuberculosis* (MtZur), Nur de *Streptomyces coelicolor* (ScFur) et PerR de *Bacillus subtilis (*BsFur). Ces alignements ont été réalisés avec le programme CLUSTAL (Thompson *et al.* 1994).

Parmi ces séquences, sept structures cristallographiques de protéines Fur, y compris celle du domaine de liaison d'EcFur, sont actuellement connues :

- PaFur (Pohl *et al.* 2003)dans une conformation active (code PDB 1MZB)
- VcFur (Sheikh *et al.* 2009) dans une conformation active (code PDB 2W57)
- BsPerR dans la conformation inactive (forme Apo) (code PDB 2FE3) (Traore *et al.* 2006) et active (code PDB 3F8N) (Jacquamet *et al.* 2009)
- Le domaine de liaison d'EcFur (Pecqueur *et al.* 2006)(code PDB 2FU4)
- MtZur (Lucarelli *et al.* 2007)(code PDB 2O03)
- ScNur (An *et al.* 2009)(PDB code 3EYY)
- HpFur (Dian *et al.* 2011) (code PDB 2XIG)

Le tableau Tableau III.2 suivant résume les pourcentages de similarité entre les protéines EcFur, VsFur et FtFur et les homologues de structures connues : PaFur, VcFur, HpFur et BsPerR. Ce pourcentage a été calculé avec BLAST en divisant le nombre de « positifs » par la longueur de la séquence de la protéine cible à prédire.

Tableau III.2: Pourcentage de similarité des différentes protéines Fur

Modèle \ Cible	**EcFur**	**VsFur**	**FtFur**
PaFur	77 (66)	78 (67)	64 (59)
VcFur	91 (82)	93 (86)	64 (60)
HpFur	51 (51)	53 (53)	60 (57)
BsPerR	49 (49)	56 (49)	50 (48)

Les chiffres entre parenthèse représentent ce même pourcentage de similarité calculé à partir du nombre d'acides aminés effectivement présents dans la structure servant de modèle. Ce pourcentage est pris en compte pour choisir la structure à utiliser pour construire les modèles.

Parmi les structures connues actuellement, EcFur partage avec VcFur et PaFur la meilleure similarité, respectivement 91% et 77%. Au début de mes travaux, les structures de PaFur (1MZB) et de BsPerR (2FE3) ont été utilisées pour construire des modèles d'EcFur dans sa conformation inactive (Apo) et active (Holo). Par la suite, la structure de VcFur (2W57), qui partage la meilleure similarité avec EcFur, a été utilisée à la place de PaFur.

Les structures cristallographiques sont raffinées dans un premier temps, par minimisation de l'énergie, sous fortes contraintes avec CHARMM. Dans un second temps, ces structures optimisées sont utilisées comme « moule » pour construire les modèles d'EcFur avec MODELLER. La meilleure conformation parmi les dix conformations générées par le logiciel, est choisie sur la base de la fonction de score et raffinée avec CHARMM par une simple minimisation. La comparaison avec la structure cristallographique du domaine Nter d'EcFur disponible permet de valider le modèle choisi. Nous reviendrons sur cette partie dans la présentation du modèle d'EcFur construit à partir de la structure de VcFur.

Nous présentons ci-dessous les modèles d'EcFur construits à partir de PaFur (1MZB) et de BsPerR (2FE3), représentant respectivement la protéine dans sa forme inactive (Apo) et active (Holo).

La Figure III.13, ci-dessous montre la structure optimisée de la protéine BsPerR dans sa forme Apo. Cette structure a été raffinée dans CHARMM, par une minimisation (SD de 100 pas et ABNR 2000 pas) suivie d'une dynamique de 5 ns à 298 K dans un solvant implicite EEF1, en maintenant une contrainte harmonique sur le « backbone ». Ainsi les atomes de la protéine oscillent autour d'une position d'équilibre définie. Cela permet d'optimiser la structure de la protéine sans perdre la structuration initiale. Cette structure contient six hélices α et cinq feuillets β par monomère. Les brins $\beta3-\beta5$ et les hélices $\alpha5-\alpha6$ s'unissent pour former le domaine de dimérisation, encadré en vert. Le domaine de liaison à l'ADN est formé des hélices $\alpha1-\alpha4$ et des brins $\beta1$ et $\beta2$ Cette structure a servi à construire le modèle d'EcFur dans sa configuration inactive.

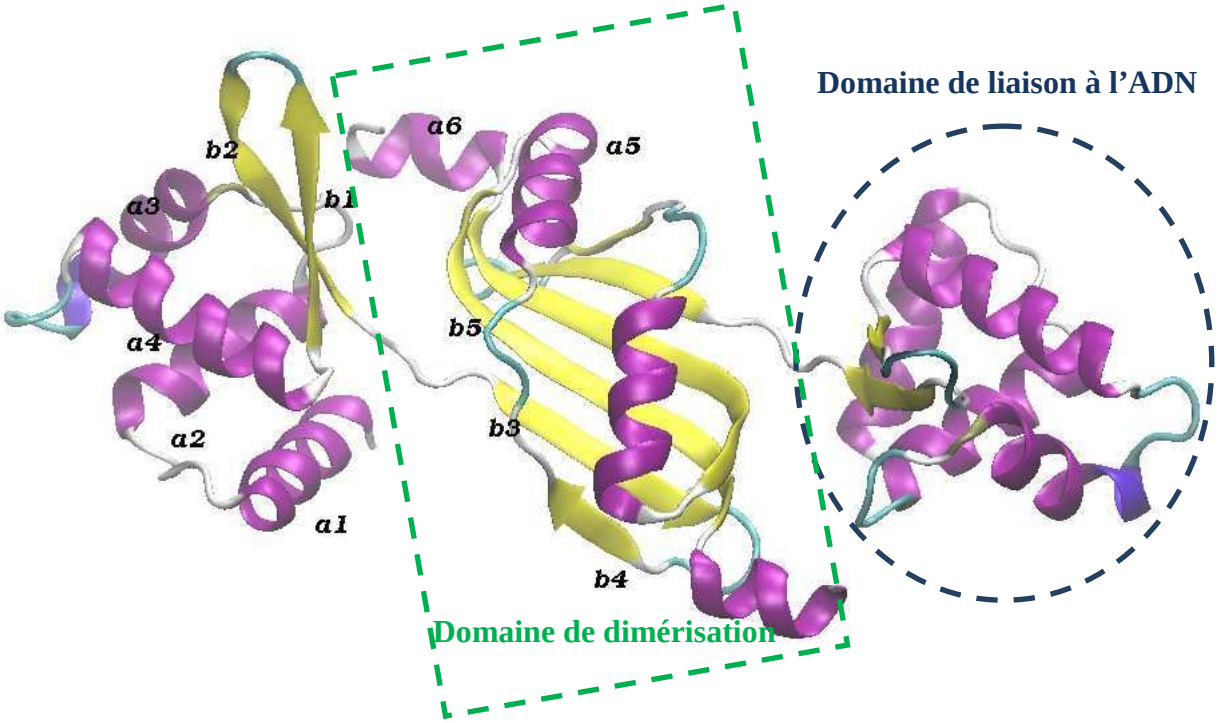


Figure III.13: Structure optimisée de BsPerR dans sa configuration inactive (code PDB 2FE3)

La Figure III.14, ci-dessous présente le modèle d'EcFur construit par MODELLER (fonction de score = 1382) à partir de la structure optimisée de BsPerR ci-dessus (Figure III.13). Le modèle a été optimisé dans CHARMM en utilisant les mêmes paramètres de minimisation et de dynamique que la structure précédente de BsPerR. Comme on peut le voir sur la Figure III.14, ce modèle contient également six hélices α et cinq feuillets β par monomère. On peut noter la présence d'un petit brin β supplémentaire, visible au niveau du domaine de liaison à l'ADN et les boucles à l'extrémité Cter sur ce modèle.

Cette conformation représentera la forme inactive de la protéine EcFur utilisé pour les amarrages des peptides.

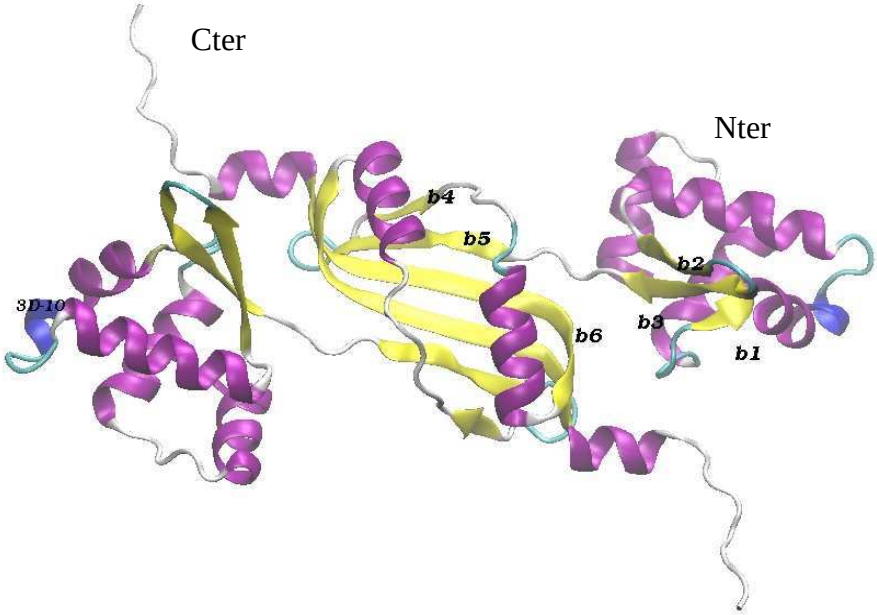


Figure III.14: Modèle d'EcFur dans la forme inactive

Pour construire le modèle d'EcFur dans sa forme active, PaFur (1MZB) dont la structure optimisée est présenté ci-dessous (Figure III.15) a été utilisée dans un premier temps. Ce modèle sera noté EcFur←Pa par la suite. Plus tard la structure de VcFur a été utilisée pour construire le modèle d'EcFur (EcFur←Vc).

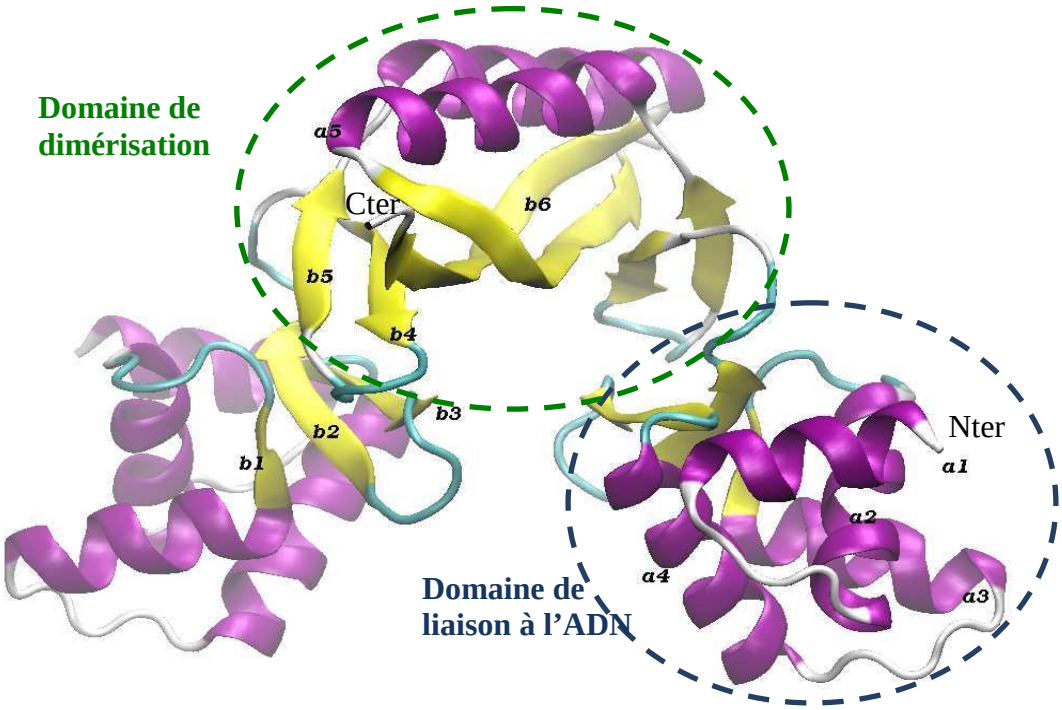


Figure III.15: Structure optimisée du dimère de PaFur (1MZB)

La Figure III.16 ci-dessous montre le modèle d'EcFur dans sa forme active construit par MODELLER (fonction de score = 1341) à partir de la structure optimisée de PaFur ci-dessus. Ce modèle a été optimisé par la suite avec CHARMM avec les paramètres précédents.

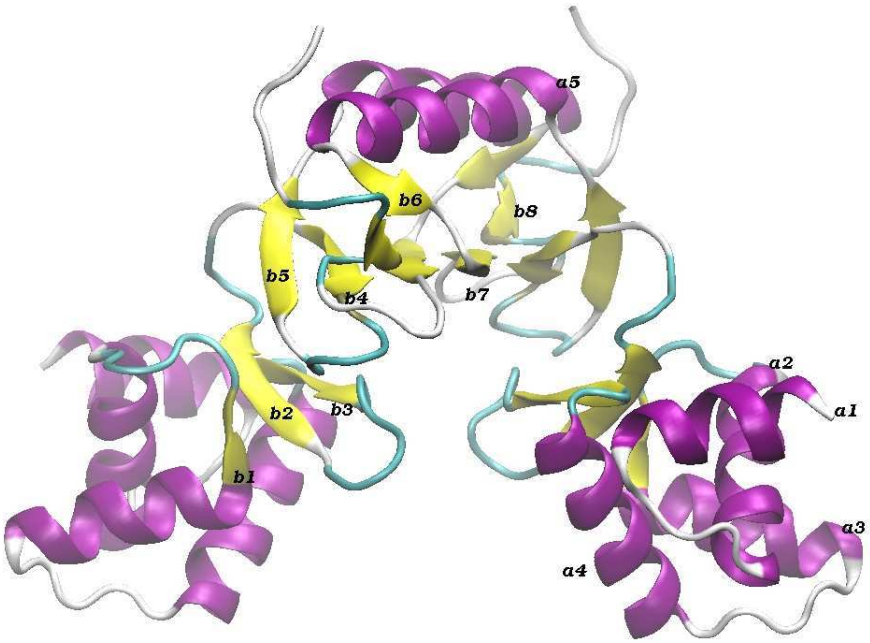


Figure III.16: Modèle du dimère d'EcFur dans sa configuration active (réalisé par homologie avec PaFur)

Cette conformation est partiellement différente de celle de la structure de PaFur, pourtant utilisé pour sa construction. Ce modèle compte huit feuillets β contrairement à la structure de PaFur qui en contient six. En effet, le long feuillet β6 de PaFur ci-dessus a généré ici 3 brins β de taille plus courte (Figure III.16). Cette différence peut être liée à la grande différence des acides aminés entre EcFur et PaFur (voir Alignement de séquence Figure III.12).

Les deux modèles d'EcFur construits à partir de la structure de PaFur (1MZB) et de BsPerR (2FE3), respectivement sur la Figure III.16 et Figure III.14ont été utilisés pour l'amarrage des peptides anti-Fur. Nous verrons plus tard que c'est la structure de VcFur (2W57) qui a été utilisée ensuite pour construire le modèle EcFur dans sa forme active et sur lequel les simulations ont continué.

III. B. 2| Recherche des zones d'amarrage des peptides sur la protéine

Avant de commencer le processus d'amarrage moléculaire, il est intéressant de rappeler quelques résultats antérieurs donnant des indications sur les interactions des aptamères peptidiques avec la protéine Fur et les difficultés de prédire le mode d'arrimage de ces molécules.

Les aptamères F1 à F3 nécessitent à la fois le domaine de liaison à l'ADN (Résidus 1 à 83) et le domaine de dimérisation de Fur (résidus 84 à 148) pour interagir avec la protéine EcFur et l'inhiber. L'aptamère F4 peut interagir uniquement avec le domaine de dimérisation de la protéine pour l'inhiber (Abed *et al.* 2007). Comme il a été évoqué dans le chapitre II, le peptide pF4 interagit avec une sous-unité monomérique et empêche l'association des deux monomères de la protéine.

Partant de ces travaux, il a fallu trouver des conformations peptidiques amarrées favorablement sur la protéine et éventuellement compatibles avec un amarrage de l'aptamère peptidique entier. Sachant que le nombre maximal d'angles de torsion flexibles dans un processus d'amarrage avec AutoDock4 est de 35 et qu'il faut en utiliser au maximum une douzaine pour avoir des résultats reproductibles (Huey *et al.* 2007), l'amarrage des peptides de 13 résidus avec tous les angles de torsion était impossible. Notons qu'une séquence peptidique composée de 13 Glycines (sans chaines latérales) contient déjà 24 angles de torsion **φ** et **ψ.**

La solution logique a été d'utiliser des séquences contraintes du peptide sur le « backbone », la protéine étant fixe. Ainsi, nous avons commencé par tenter d'amarrer les peptides de 13 résidus issus des modèles des aptamères peptidiques avec des angles de torsions flexibles uniquement sur les chaines latérales, mais les résultats d'arrimage n'étaient pas favorables. Par la suite, lorsque des séquences peptidiques plus courtes inhibant l'activité de liaison de Fur à l'ADN ont été découvertes expérimentalement, l'amarrage de ces séquences réduites a été plus aisé. Partant de ces derniers amarrages des peptides de séquence raccourcie, nous sommes revenus, dans un second temps, à l'amarrage des peptides de 13 acides aminés.

Les zones de liaison des peptides anti-Fur ne sont pas connues. La méthode utilisée a été l'amarrage en aveugle ou « blind docking » dans laquelle le logiciel AutoDock4.2 avait montré quelques résultats intéressants (Hetenyi *et al.* 2006; Morris *et al.* 2009). En analysant la structure des protéines Fur et vu la taille des ligands (pF1 à pF3), les zones d'amarrage

choisies sont représentées par des grilles avec AutoDock4.2 sur la Figure III.18, ci-dessous. Les paramètres de ces grilles sont donnés dans le Tableau III.3. Ces grilles sont définies par un point central choisi. Le nombre de points à partir du centre désigne la taille de la grille. Le nombre maximal de points est 126. L'espacement entre les points est de 0,375 Å par défaut. Pour explorer la protéine entière il a fallu placer 126 points sur chaque axe et augmenter l'espacement pour couvrir toute la protéine (confère Figure III.18 et Tableau III.3).

Nous savions aussi grâce aux travaux de Sylvia que le peptide pF1 interagit avec la forme métallée de Fur. Cela a été démontré par la spectroscopie d'émission de fluorescence du tryptophane de pF1 (Figure III.17).

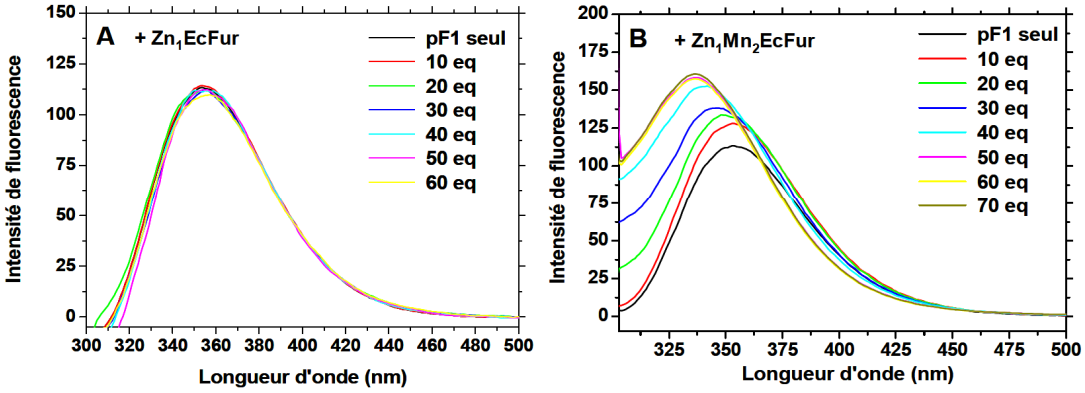
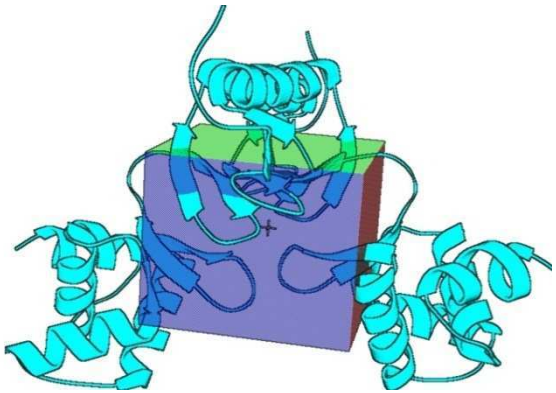


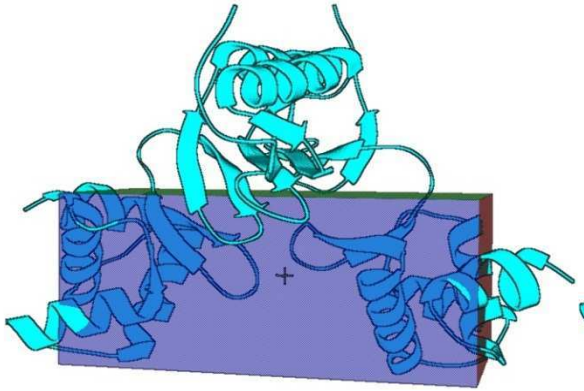
Figure III.17: Spectres d'émission de fluorescence du tryptophane de pF1 wt à 1 µM après excitation à 295 nm en présence de différents équivalents (A) EcFur (Apo) ou (B) de EcFur (métallée). La protéine a été métallée au préalable par 1 équivalent de Mn^{2+} par sous-unité (sel : MnCl2). Spectres enregistrés dans le tampon BTP 100 mM pH 7,5, KCl 100 mM, MgSO4 5 mM, TCEP 10 µM à 25 °C après 10 min d'équilibration. Bande passante d'excitation : 3 nm ; bande passante d'émission : 5 nm ; 3 spectres d'émission sont enregistrés entre 300 nm et 500 nm puis sont moyennés.

La Figure III.17**A** montre que le spectre de pF1 n'est pas modifié par les différents ajouts d'EcFur Apo (Zn1-EcFur). Cette absence de variation suggère que le peptide n'a pas interagit avec la protéine. En revanche, les ajouts successifs d'EcFur métallée (Zn1Mn2-EcFur) entraînent un décalage progressif du ëmax d'émission de fluorescence, qui sature à 337 nm à partir de 50 équivalents de protéine (Figure III.17**B**). Ce décalage rend compte du changement d'environnement du tryptophane qui tend à être enfoui dans la protéine. La constante de dissociation Kd entre la protéine et le peptide serait de 32,6 ± 0,6 µM, compatible avec l'IC_{50} de pF1 déterminé à partir des tests de protection à la nucléase.

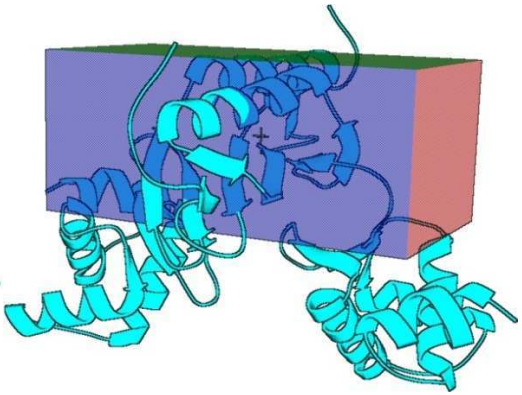
Partant de ces résultats, nous nous sommes focalisés sur la forme active de la protéine. De façon intéressante, elle donne aussi de meilleurs résultats comparés au modèle Apo de Fur.



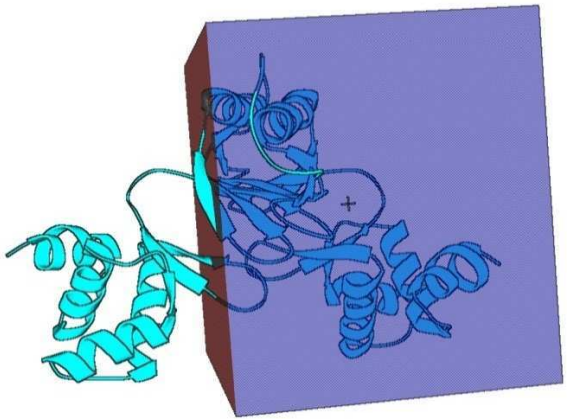
FurEcG1



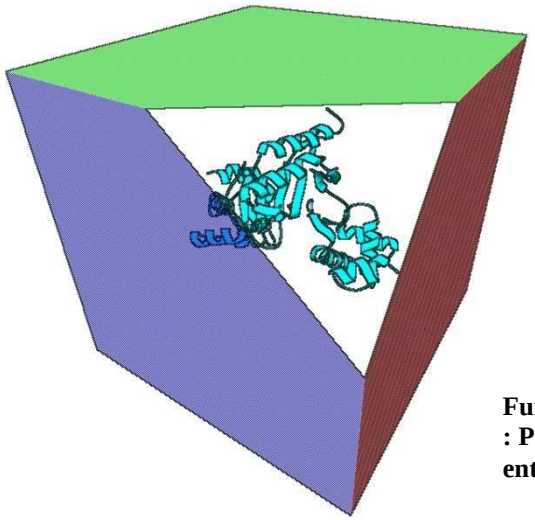
FurEcG2



FurEcG4



FurEcGsy



FurEcG_2 : Protéine entière

Figure III.18: Grilles explorées par les peptides

Tableau III.3: Paramètres des grilles utilisées

Grille	Nombre de points			Distance entre les points	Coordonnées du centre de la grille		
	X	Y	Z	en Angström	X	Y	Z
FurEcG1	66	54	64	0,375	-0,14	0,81	1,73
FurEcG2	126	54	64	0,375	0,72	-7,07	1,73
FurEcG3	110	74	86	0,375	0,32	-2.70	7.76
FurEcG4	126	54	64	0,375	0,72	11,52	1,73
FurEcGsy	108	126	118	0,375	13,29	3,99	1,73
FurEc_2	126	126	126	0,702	-0,14	0,81	1,73

Chacune de ces grilles avait pour but de pouvoir explorer une zone spécifique de la protéine. Les grilles FurEcG1, FurEcG2, FurEcG4, FurEcGsy devaient permettre d'explorer respectivement le centre de la protéine, le domaine de liaison à l'ADN, le domaine de dimérisation et un monomère de la protéine. Ces grilles ont été criblées virtuellement (voir ci-dessous) avec les paramètres par défaut d'AutoDock4.2 (Morris *et al.* 2009) par les peptides pF1 à pF4 (conformations contraintes sur le « backbone » issues de l'optimisation décrite précédemment, avec douze angles de torsion flexibles sur les chaines latérales. Avec ces paramètres le peptide peut bouger en translation et en rotation en gardant son squelette peptidique fixe et les chaines latérales flexibles. Cette opération a été répétée jusqu'à l'obtention d'un amarrage favorable ($\Delta G < 0$).

Les résultats d'amarrages d'AutoDock sont classés sous la forme de familles ou « cluster » mais aussi en fonction de l'énergie libre de liaison ΔG. Le « cluster » regroupe les conformations les plus proches « root-mean-square deviation ou RMSD » < à 2 Å. Plus le nombre de conformations dans un cluster est élevé à l'issu d'un criblage, plus cette conformation est significative statistiquement. L'énergie libre de liaison est la fonction d'évaluation de la formation d'un complexe (plus cette valeur est négative, plus la formation du complexe est favorable).

Le choix d'une zone d'interaction entre les peptides et la protéine s'est effectué sur les meilleures conformations du complexe (Tableau III.4).

Tous les autres paramètres d'amarrage étant identiques, les diverses grilles d'amarrage ont été choisies en fonction des énergies libres de liaison obtenues pour les divers peptides. Un amarrage a également été réalisé avec la grille correspondant à la protéine entière afin de s'assurer que cet amarrage, moins précis, donnait des énergies de liaisons plus élevées. La taille des clusters n'a pas été indiquée dans le tableau mais tous les amarrages correspondent à de structures statistiquement acceptables.

Tableau III.4 : Énergies libres de liaison des peptides dans les grilles (kcal/mol)

	pF1	pF2	pF3	pF4
FurEcG1	-1.7	+5.8	+3.3	+2.8
FurEcG2	-3.5	+7.6	+0.8	+4.1
FurEcG4	+1.5	-2.1	+5.2	-0.2
FurEcGsy	+0.6	+0.8	-1.9	-1.3
FurEc_2	-0.9	+2.5	-0.7	+3.2

Ces résultats montrent que les énergies libres de liaison entre peptides pF1, pF2, pF3 et pF4 et EcFur pourront être évaluées avec les grilles FurEcG2, FurEcG4, FurEcGsy et FurEcGsy, respectivement (Tableau III.4). Il n'y a bien sûr aucune obligation dans le choix de ces grilles différentes pour chaque peptide : il y a fort à parier que la même grille FurEc_2 englobant toute la protéine aurait pu convenir pour tous, mais le temps de calcul ne sera pas raisonnable. Les valeurs d'énergie libre de liaison reportées dans le Tableau III.4 ne sont toutefois pas favorables comparées à ce qu'on peut trouver dans la littérature, c'est-à-dire proches de -10 kcal/mol. Pour tenir compte de la flexibilité de la protéine et optimiser les amarrages, une première solution a été d'effectuer des simulations de dynamique moléculaire de la protéine seule, de sélectionner les meilleurs « frames » de la protéine et d'effectuer des explorations sur les conformations sélectionnées.

Cette méthode d'exploration n'a pas permis d'affiner l'énergie d'amarrage et ne tient pas compte du changement conformationnel engendrée lors de l'interaction entre la protéine et les peptides. Nous avons donc tenté une autre approche combinant l'amarrage par AutoDock et la Dynamique Moléculaire que j'ai nommé « amarrage moléculaire assisté par la dynamique moléculaire ». Cette approche est expliquée ci-dessous.

III. B. 3| Amarrage Assisté par Dynamique Moléculaire (AADM)

Le principe de ce protocole consiste à prendre en compte la flexibilité de la protéine et des peptides engendrée simultanément lors de l'interaction. Le protocole peut être décomposé en trois parties :

- Réalisation d'un amarrage initial sur les différentes grilles de la protéine, décrites précédemment et définition de la zone d'interaction du ligand (site d'action).
- Réalisation d'une DM de Langevin avec les meilleures conformations du complexe protéine Fur-peptide (meilleure énergie libre de liaison lors de l'amarrage initial précédent), pour prendre en compte la flexibilité des deux

partenaires simultanément. La dynamique moléculaire de 5 ns à 289 K, générant 5000 conformations est précédée et suivie par une minimisation (100 pas de SD et 1000 pas d'ABNR) en solvatation implicite (EEF1) est appliquée. Pendant cette dynamique, le peptide au complet ainsi que les chaines latérales des résidus de la protéine proches du peptide (distance ≥ 5 Å) sont flexibles. Les cinq meilleures conformations du complexe, choisies sur la base de l'énergie d'interaction (***E*$_{inter}$**) entre peptide et protéine sont utilisées pour d'autres amarrages d'optimisation.

- Réalisation d'amarrage d'optimisation avec des paramètres permettant une évaluation plus poussée des possibilités d'interaction du peptide sur la protéine. A l'issue de ces amarrages d'optimisation, les conformations des complexes sont analysées avec CHARMM (paragraphe suivant).

Cette approche AADM peut être répétée plusieurs fois en veillant toutefois à ce que des contraintes sur les atomes de la protéine éloignés du peptide soient toujours imposées afin de ne pas s'éloigner de la structure initiale de celle-ci.

L'apport de ce processus ADDM combinant la dynamique et l'amarrage moléculaires sera démontré à travers le processus d'optimisation de l'amarrage de pF1 sur le modèle EcFur←Pa donné en exemple dans la suite.

III. B. 4| Identifications des interactions clés à l'amarrage

Après avoir choisi les coordonnées des meilleures conformations du complexe Fur-peptides selon AutoDock, nous avons utilisé CHARMM pour calculer l'énergie d'interaction et les liaisons hydrogènes. La description de l'énergie d'interaction (E_{inter}) est donnée par la fonction suivante :

$$E_{Inter} = E_p - (E_{Intra_{Lig}} + E_{Intra_{Pro}})$$

E_p est l'énergie potentielle totale du système moléculaire, $E_{Intra_{Lig}}$ est l'énergie interne du ligand et $E_{Intra_{Pro}}$ celle de la protéine (et des ions métalliques si ils sont disponibles). L'énergie d'interaction a été calculée pour les résidus des peptides (ligands). Elle représente la contribution énergétique du résidu à la liaison à la protéine. L'analyse par CHARMM a également permis d'identifier des liaisons hydrogènes et de quantifier ces liaisons par l'énergie d'interaction des atomes impliqués.

III. B. 5| Exemple du processus d'optimisation de l'amarrage des peptides anti-Fur : Cas du peptide pF1 sur le modèle EcFur←Pa.

III. B. 5. a| Amarrage sur EcFur←Pa Apo

Comme nous l'avons évoqué précédemment, les « dockings » ont été réalisés dans un premier temps sur des peptides tronqués (moins de 13 acides aminés), généralement actifs *in vitro* dans le test d'activité vu dans la partie expérimentale. Dans un second temps, la séquence totale des peptides est construite à partir de la conformation du peptide tronqué issu d'un amarrage optimisé par le processus AADM.

Dans cette section, nous allons voir le processus d'optimisation de l'amarrage du peptide pF1 sur le modèle EcFur←Pa. Nous avons vu sur les Figure III.11 et Figure III.16, respectivement, les modèles optimisés du peptide pF1 et d'EcFur←Pa qui constituent les coordonnées initiales de ce complexe. La séquence des neuf acides aminés Nter de pF1 : pF1(1-9) est extraite de la conformation optimisée du peptide entier pour l'amarrage. Les grilles FurEcG1 et FurEcG2 (Tableau III.3) ont été utilisées avec les paramètres par défaut d'AutoDock dans le protocole AADM. La Figure III.19, ci-dessous montre l'amarrage de ce peptide sur la protéine. Cet amarrage a été obtenu avec une énergie libre de liaison ΔG de -13.6 kcal/mol et une constante d'inhibition estimée, Ki de 105 pM (pico Molaire) indiquant une affinité très favorable. Rappelons que les valeurs d'énergie libre de liaison ΔG ne sont pas directement comparables aux IC_{50} des tests d'activité car elles reposent sur une calibration empirique à partir d'une base de complexes de structure connue. Par contre, elles permettront de comparer les amarrages des divers peptides sur les diverses protéines Fur présentés par la suite dans ce rapport.

.

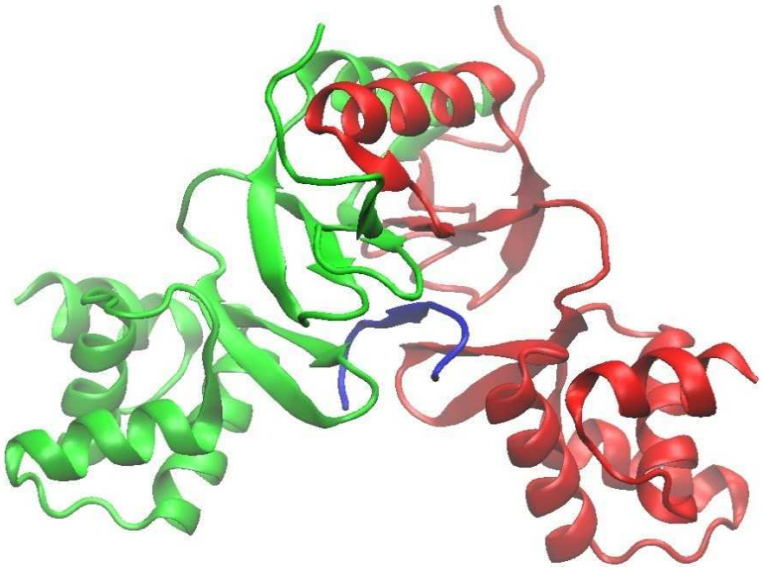


Figure III.19: Amarrage du peptide pF1(1-9) sur EcFur obtenu par AutoDock4.
En vert la sous-unité A, en rouge la sous-unité B de la protéine et en bleu le peptide.

Le peptide interagit à la fois avec les domaines Nter et Cter de la protéine. Cette information est cohérente avec les résultats des tests double-hybride antérieurs (Abed *et al.*, 2007), rappelés dans la chapitre II.

Il est à noter que le peptide forme un feuillet β avec le brin β_2 de la sous-unité B de la protéine. Nous détaillerons ci-dessous ces interactions une fois l'étape d'optimisation de l'amarrage réalisé. En utilisant une nouvelle version du programme de « docking » : AutoDock4.2, apparue en cours de thèse, et avec les mêmes paramètres par défaut, nous obtenons un ΔG de -15.8 kcal/mol et un Ki de 2.7 pM (Figure III.20), plus favorables que ceux d'AutoDock4 (Figure III.19).

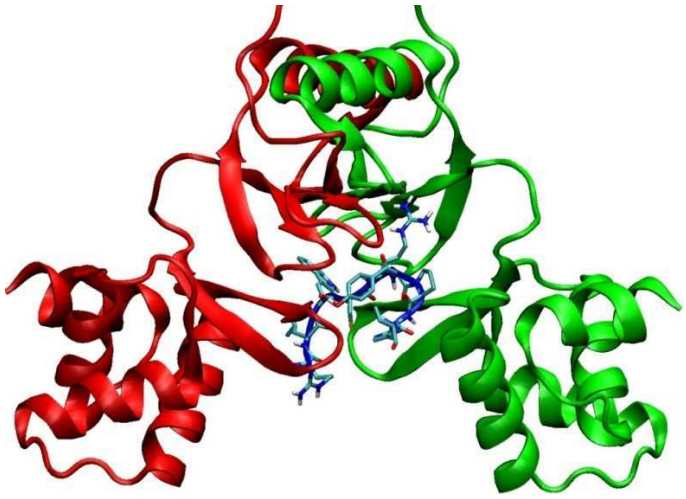


Figure III.20: Amarrage du peptide pF1 (1-9) sur EcFur obtenu par AutoDock4.2. Ici les chaines latérales du peptide sont montrées

Sur les deux figures le peptide est amarré dans la même zone et de la même façon. L'affinité a été bien améliorée par la nouvelle version d'AutoDock. En effet, la nouvelle version d'AutoDock (version 4.2) contient de nouveaux paramètres optimisés pour les calculs des termes de désolvatation basés sur les charges, des liaisons hydrogènes et l'état non lié du complexe, probablement à l'origine de cette amélioration de l'énergie libre de liaison.

Cette version AutoDock4.2 a été utilisée pour la suite.

Partant de la conformation de pF1(1-9) obtenu ici sur la Figure III.20, le peptide de 13 acides aminés est construit avec CHARMM en ajoutant les quatre résidus Cter grâce aux coordonnées internes correspondant à leur conformation la plus stable. Le peptide entier est ensuite raffiné par minimisation d'énergie (100 pas de SD et 1000 pas ABNR) en fixant les neuf acides aminés Nter, issus de l'amarrage précèdent. Une nouvelle grille combinant les grilles FurEcG1 et FurEcG2 : FurEcG3 (Tableau III.3), plus adaptée à la taille du peptide de 13 acides aminés a été utilisée. Le protocole AADM est appliqué pour obtenir l'amarrage du peptide pF1(1-13). Les paramètres d'amarrage d'optimisation utilisés sont les suivants : 50 pour la taille de la population initiale, 60000000 pour le nombre maximal d'évaluations d'énergie avec 50 conformations finales. La meilleure conformation du complexe obtenu (ΔG = - 17,5 kcal/mol et Ki = 160 fM : femto Molaire) est présentée sur la Figure III.21.

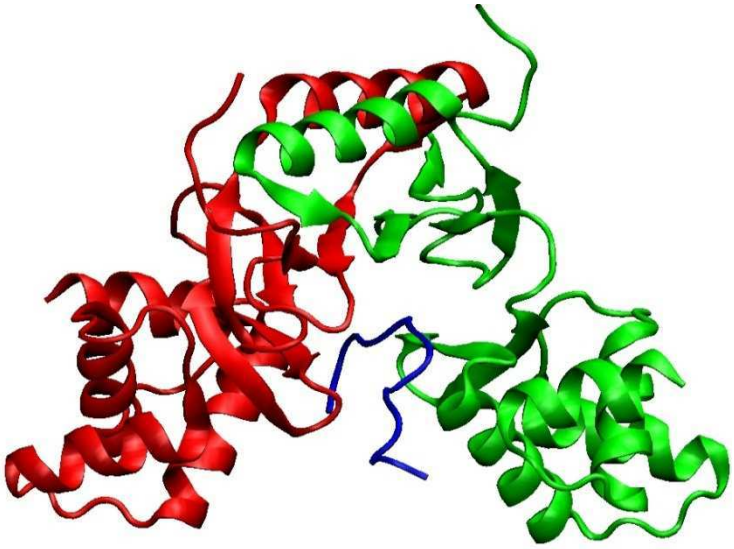


Figure III.21 : Amarrage du peptide pF1 sur EcFur←Pa

Dans cette configuration, le peptide a une meilleure affinité pour la protéine comparé au peptide pF1(1-9). Ce qui va dans le sens des tests d'activités où on a vu que pF1(1-13) a un

IC_{50} plus faible que pF1(1-9). L'ensemble des interactions du peptide a été évalué par CHARMM *via* le calcul de l'énergie d'interaction, décrite auparavant. Le Tableau III.5, ci-dessous présente l'énergie d'interaction des résidus du peptide avec la protéine pour des valeurs d'$E_{inter} \leq$ -8 kcal/mol.

Résidus pF1	E_{inter} (kcal/mol)
R1	-15,3
W3	-19,1
R5	-18,2
Y6	-18,8
D13	-8

Tableau III.5: Énergies d'interaction des acides aminés de pF1 interagissant fortement avec EcFur←Pa.

Les résidus R1, W3, R5, Y6 et D13 du peptide sont donc ceux qui interagissent le plus favorablement avec la protéine. Ces résidus et ceux de la protéine, sont représentés avec une vue plus détaillée (Figure III.22).

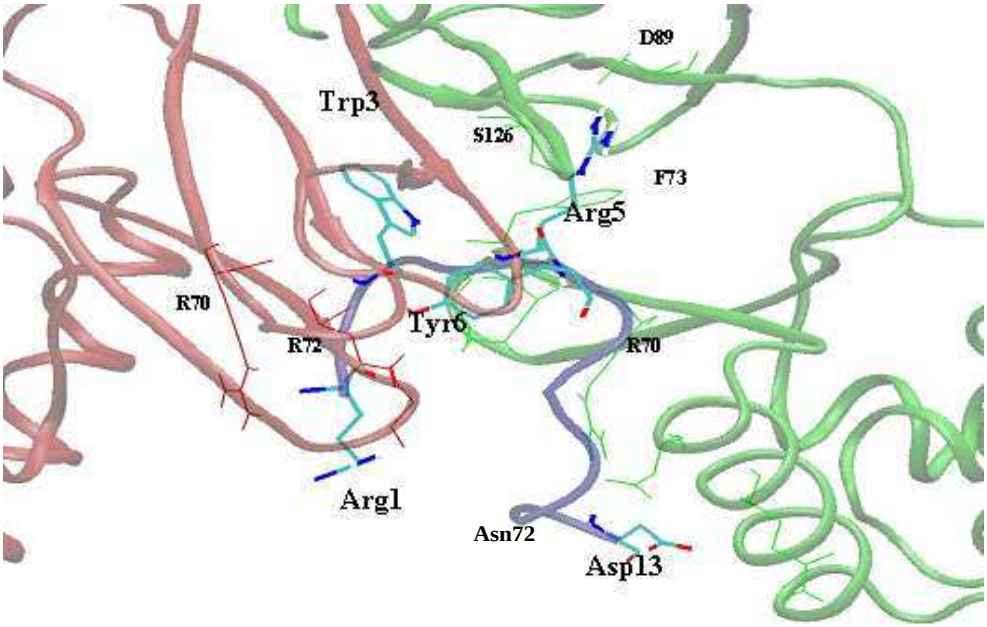


Figure III.22: Zoom sur les interactions de pF1 et de la protéine Fur. Les résidus du peptide en bleu sont marqués en gros caractères, ceux de la protéine en rouge et vert, sont notés en petits caractères.

III. B. 5. b| Amarrage sur EcFur←Pa « métallée »

Une analyse approfondie de cet amarrage montre que certains acides aminés du peptide pF1 interagissent avec des résidus importants, probables ligands des métaux du site structural et/ou de régulation de la protéine. C'est le cas des résidus His125 et Asp88 de la protéine. Il est alors apparu nécessaire de prendre en compte les sites métalliques de la

protéine, jusqu'ici ignorés. Comme nous l'avons expliqué dans l'introduction, nous allons construire avec CHARMM trois sites métalliques par monomère au lieu de deux, comme décrit dans la littérature sur EcFur, notamment en tenant compte entre autres des données récentes sur HpFur (Dian *et al.* 2011). Les ions métalliques sont dans ce cas des atomes de fer, considérés comme des boules avec une charge +2. La nouvelle conformation de la protéine (« métallée ») ; contenant les sites métalliques, a ensuite été optimisée par une minimisation (SD 100 pas, ABNR 1000 pas ou gradient ≤ 0,1 kcal/mol/Å) dans un solvant implicite (méthode EEF1) avec des contraintes harmoniques sur l'ensemble de la protéine, sauf les atomes se trouvant à 8 Å de rayon autour de fer. La nouvelle conformation de la protéine « métallée » et ces sites métalliques sont présentés sur la Figure III.23. Notons que ces sites métalliques ont été construits dans l'unique but d'orienter les chaines latérales des ligands du site métallique vers l'atome de fer et qu'il faudra améliorer ces constructions pour avoir de meilleurs modèles prenant en compte les propriétés électronique des ions Fe(II). L'amarrage de peptide pF1 sera repris afin de vérifier dans la suite s'il peut interagir avec des résidus ligands du site métallique.

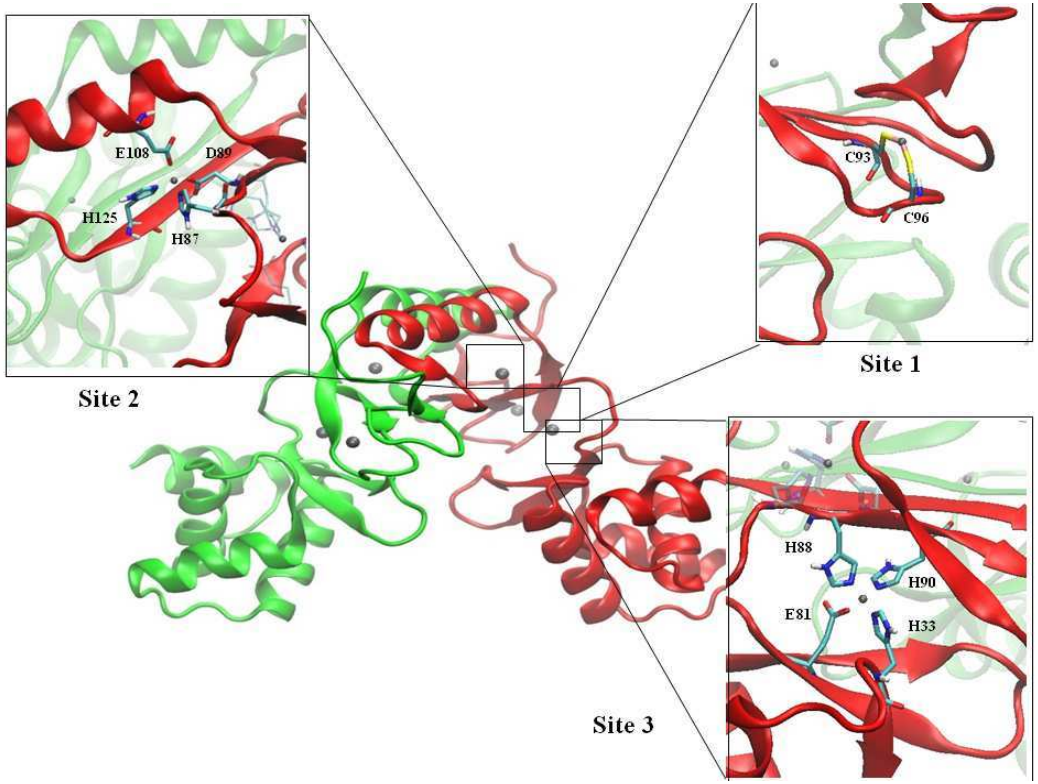


Figure III.23: Vue d'ensembles des sites métalliques construits dans le modèle EcFur←Pa et zoom sur ces sites

Site 1 ; le M^{2+} est chelaté par les cystéines C93 et C96 (annotation EcFur)
Site 2 ; les ligands sont D89, H125, H87 et E108.
Site 3 ; le M^{2+} est coordonné ici par les histidines H33, 88 et 90 et la glutamine E81.

Le protocole d'AADM est appliqué comme précédemment avec la conformation du peptide pF1 et la nouvelle conformation optimisée de la protéine « métallée ». Les paramètres d'amarrage d'optimisation utilisés sont les suivants : 40 pour la taille de la population initiale, 15000000 pour le nombre maximal d'évaluations d'énergie avec 30 conformations finales sur la grille EcFurG3 (Tableau III.3). La meilleure conformation, obtenue avec un ΔG de - 20,2 kcal/mol et une constante Ki = 2 fM (femto molaire) est montrée sur la Figure III.24.

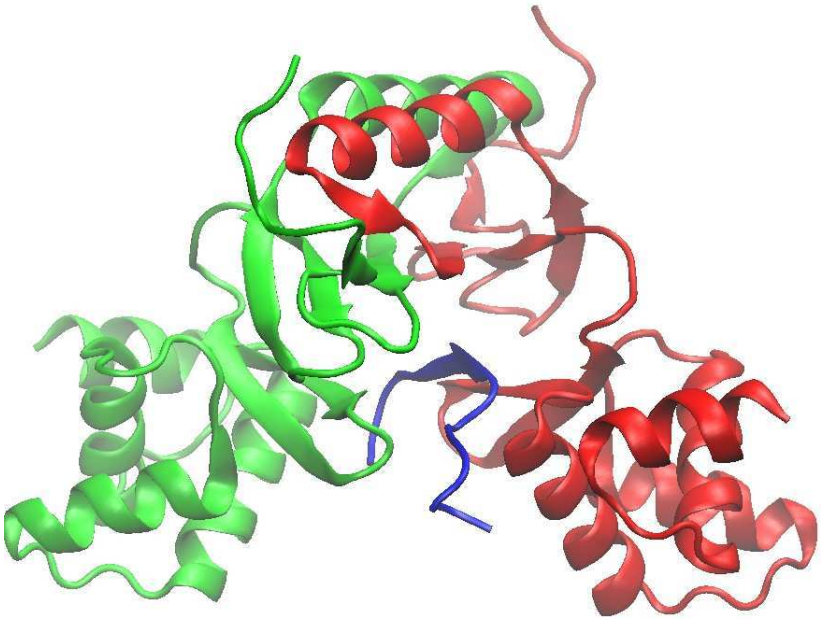


Figure III.24: Amarrage du peptide pF1 sur EcFur←Pa « métallée ». Les ions métalliques ne sont pas montrés sur cette figure

De façon intéressante, le peptide s'amarre sur la protéine lorsque celle-ci est « métallée » avec une meilleure affinité qu'il ne le faisait sur la protéine Apo.

Pour comprendre pourquoi cette affinité est meilleure que celle trouvée avec la conformation précédente, non « métallée », il faut observer les énergies d'interaction des deux amarrages (Tableau III.5et Tableau III.6 ci-dessous.

Résidus pF1	E_{inter} (kcal/mo
R1	-15
W3	-21,5
R5	-19,2
Y6	-18,2
D13	-17

Tableau III.6: Énergies d'interaction ≥ -10 kcal/mol entre les acides aminés de pF1 et la protéine EcFur←Pa « métallée ».

Les interactions de l'Asp13 du peptide ont été fortement améliorées dans le cas de la protéine métallée. La grande différence viendrait du repositionnement de pF1 qui se rapproche du domaine Nter. Les chaines latérales des résidus-ligands du site métallique de la protéine étant réorientées vers le site à fer, le peptide a trouvé d'autres résidus de la protéine vers le domaine Nter pour interagir. La construction des sites métalliques a donc permis de réaliser un modèle plus réaliste de la protéine qui donne des amarrages encore plus favorables.

Nous venons de mettre en évidence l'apport de notre approche AADM avec l'amarrage de pF1 sur le modèle EcFur←Pa.

III. B. 6| Conclusion sur la méthodologie

Dans ce premier chapitre de la partie théorique, nous avons décrit les principes des méthodes de simulation et d'amarrage moléculaire en notant au passage les programmes utilisés au cours de cette étude. Ainsi, nous avons vu les programmes CHARMM et AutoDock utilisés respectivement pour les minimisations d'énergie et les amarrages moléculaires. Nous avons décrit l'approche combinant les deux programmes permettant d'optimiser l'amarrage des peptides sur la protéine Fur et montré le processus d'optimisation à travers l'exemple de l'amarrage du peptide pF1 sur le modèle EcFur construit à partir de la structure de Pa : EcFur←Pa. Les résultats obtenus avec l'approche décrite sont cohérents avec les résultats expérimentaux antérieurs sur lesquels ces travaux théoriques s'appuient. Cette approche peut donc être valide. Précisons que la mise en place de l'approche ADDM a été nécessaire car il n'y avait pas de logiciel de « docking » adapté pour des peptides de cette taille. Elle a été appliquée par la suite pour obtenir l'amarrage de tous les peptides sur les différents modèles d'EcFur construits. VcFur étant la protéine la plus proche d'EcFur et dont la structure est disponible, nous allons présenter les résultats d'amarrage des peptides pF1 à 3 sur le modèle EcFur←Vc, obtenus avec le protocole AADM. Plus tard, nous verrons l'amarrage du peptide pF1 sur des protéines Fur d'autres organismes, ce peptide étant celui pour lequel nous disposons de la plus grande quantité d'informations expérimentales.

III. C| AMARRAGE des peptides sur le modèle EcFur←Vc

III. C. 1| Le modèle EcFur←Vc

A la parution de la structure de VcFur (Sheikh *et al.* 2009) qui présente une meilleure similarité avec EcFur, un nouveau modèle par homologie a été construit avec MODELLER : EcFur←Vc. Rappelons que nous disposons également de la structure 3D du domaine Nter d'EcFur seul (code pdb 2FU4) (Pecqueur *et al.* 2006) que nous avions négligée jusqu'ici. Afin de voir l'influence de l'utilisation de telle ou telle structure sur la qualité du modèle final, nous avons construit 2 modèles d'EcFur←Vc, l'un à partir de la structure de VcFur seul (Code pdb 2W57) (Sheikh *et al.* 2009), l'autre à partir de l'alignement de structures composées de VcFur et des 2 domaines Nter de EcFur. La Figure III.25 présente une superposition des deux modèles obtenus ainsi que des structures des 2 domaines Nter (2FU4).

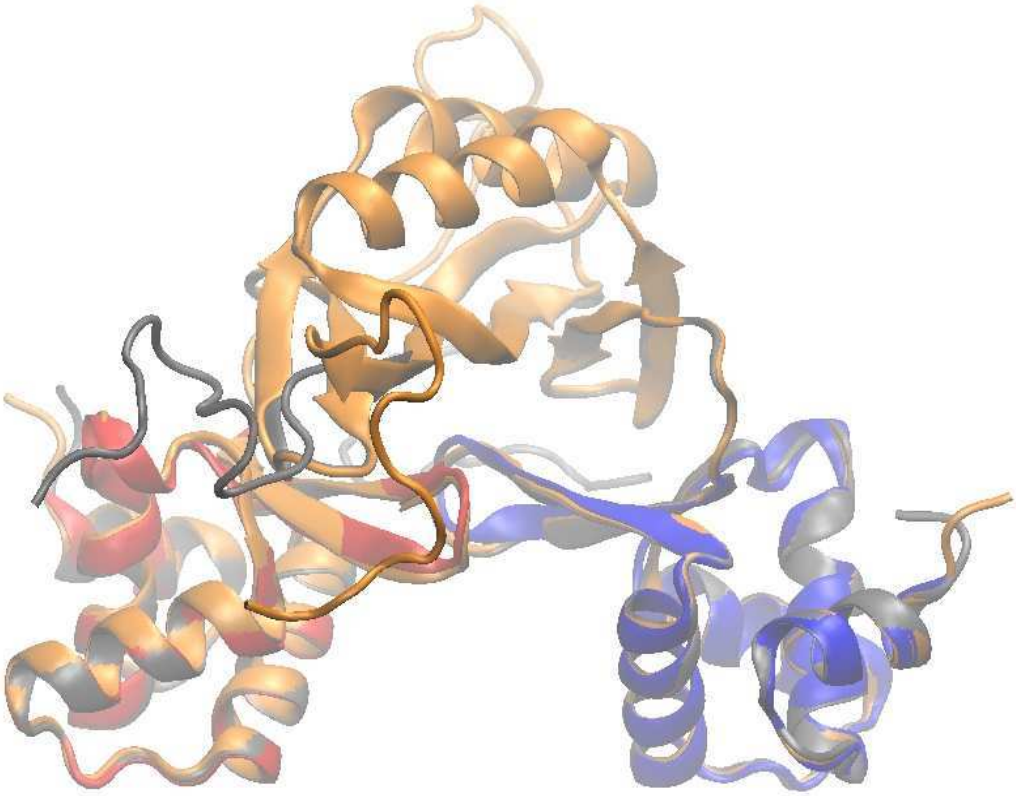


Figure III.25: Superposition de deux modèles d'EcFur←Vc avec la structure du domaine Nter d'EcFur. En bleu et rouge ; les structures du domaine Nter de la sous-unité A et B d'EcFur respectivement**.** En Gris, le modèle EcFur←Vc à la fois à partir des deux domaines Nter de la sous-unité A et B (code pdb et de la structure de VcFur (Code pdb 2W57) (Sheikh *et al.* 2009). En orange, le modèle construit à partir de la structure de VcFur.

Cette figure montre que les deux modèles par homologie d'EcFur se superposent parfaitement aux structures du domaine Nter (monomères A et B) (Pecqueur *et al.* 2006). La superposition du « backbone » du domaine Nter du modèle construit et de la structure d'EcFur donne un « RMSD » de moins de 1 Å. Cela montre aussi que les domaines Nter de VcFur et EcFur sont très proches ce qui valide notre modèle. On voit aussi que les résidus manquant dans la structure de VcFur (résidus 1 à 2 et 133 à 148) (Sheikh *et al.* 2009), modélisés sans homologie, ne sont pas structurés.

Le modèle construit à partir de la seule structure de VcFur (fonction de pénalité de MODELLER = 1576,8) a donc été utilisé pour les « dockings » des peptides qui vont suivre. Les sites métalliques ont été construits comme pour le modèle EcFur←Pa. La conformation a été raffinée par minimisation avec CHARMM avec des contraintes permettant de fixer la protéine et d'optimiser les ligands autour du site métallique. Comme dans la structure 3D de VcFur, le modèle contient cinq hélices α et six feuillets β (Sheikh *et al.* 2009). Ce modèle a été utilisé pour le « docking » des peptides pF1 à pF3, comme nous allons le voir. La boucle Cter (134-148) du modèle d'EcFur a été coupée avant l'amarrage des peptides, car non modélisée par homologie. Les résidus de cette boucle sont absents dans la structure de VcFur. Dans le « docking » de pF2, il a fallu garder cette boucle pour obtenir des amarrages favorables.

III. C. 2| Amarrage des séquences de pF1 sur EcFur←Vc

III. C. 2. a| Amarrage de pF1 entier : RLWCRYPHPPLTD

L'amarrage optimisé de pF1 sur le modèle EcFur←Pa nous donne les coordonnées initiales du peptide. Le complexe pF1-Fur est soumis au processus AADM, décrit précédemment avec une dynamique de 5 ns. Pour chaque « frame », l'énergie d'interaction entre le peptide et la protéine est évaluée par CHARMM. Les cinq conformations d'énergie minimale du complexe sont sélectionnées pour le « docking ». Plusieurs « dockings » en faisant varier les paramètres : nombre d'évaluations et taille de la population initiale, ont été effectués pour chaque « frame » sélectionnée afin d'obtenir une énergie libre favorable. La conformation ayant la meilleure énergie libre est présentée ci-dessous dans la Figure III.26

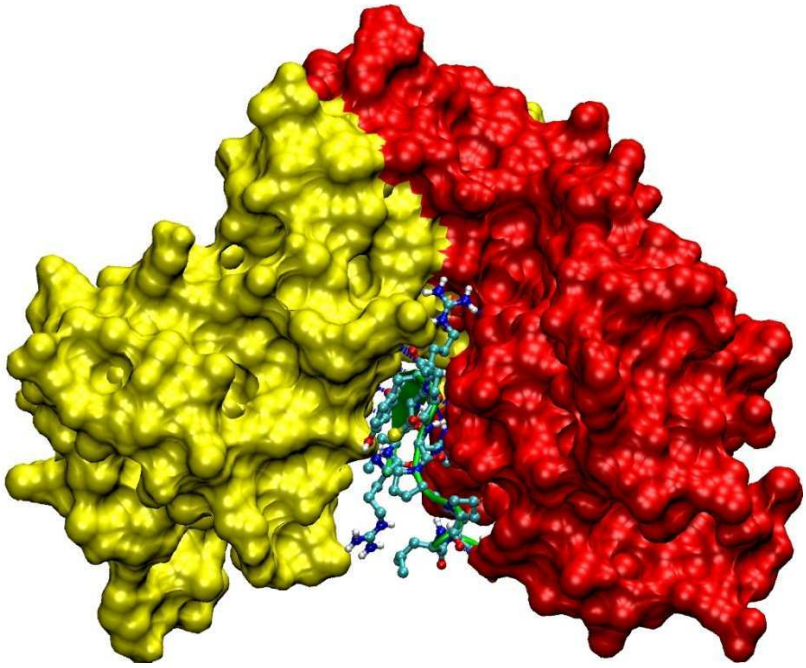


Figure III.26: Amarrage de pF1 sur EcFur←Vc. Représentation en surface de la protéine, en jaune le monomère A et en rouge le monomère B. Le peptide est représenté avec ses chaines latérales.

L'énergie libre de la liaison de ce complexe est de -18 kcal/mol avec une constante d'inhibition de 70 fM. Le peptide pF1 interagit ici aussi avec la conformation active de la protéine, comme l'a montré la spectroscopie du tryptophane (Vitale 2009). Le peptide est amarré entre les deux domaines Nter du dimère de Fur. Il interagit avec les domaines Nter et Cter du monomère ce qui est parfaitement en accord avec les résultats de double-hybride (Abed *et al.* 2007) (Vitale 2009). La valeur de l'énergie libre de liaison est légèrement plus faible que celle obtenue sur le modèle EcFur←Pa qui est de -20 kcal/mol. Cette petite différence peut être due à la disposition des chaines latérales et/ou aux contraintes stériques exercées sur le peptide. Les principales interactions après l'amarrage sont montrées sur la Figure III.27.

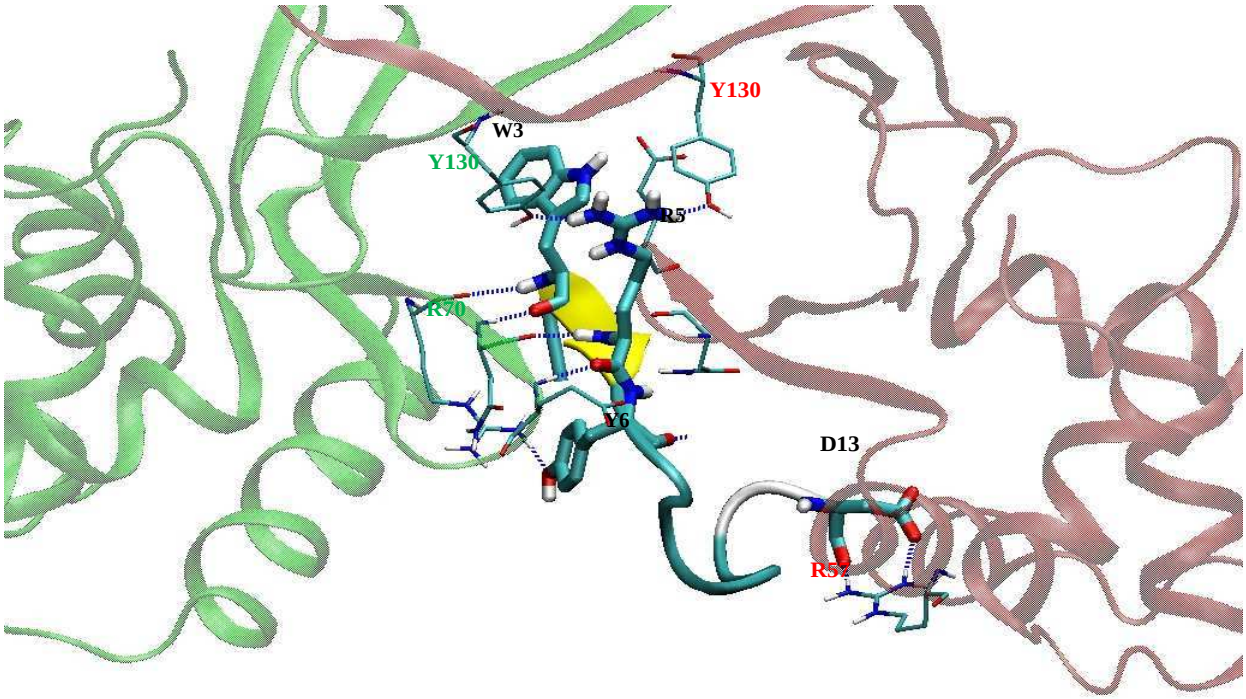


Figure III.27: Vue d'ensemble des principales interactions entre pF1 et EcFur←Vc. Le peptide est représenté en jaune avec ses chaines latérales en licorice (taille 3) et la protéine Chaine A et B, respectivement en vert et rouge, chaines latérales en licorice (taille 1) VMD (Humphrey *et al.* 1996).

Comme nous venons de la voir sur le modèle EcFur←Pa, le peptide pF1 interagit avec la protéine principalement *via* les résidus R1, W3, R5, Y6 et D13. Dans cet amarrage avec le modèle EcFur←Vc, pF1 interagit favorablement avec les résidus 1 à 6, 8 et 13 (énergie d'interaction < -10 kcal/mol) (Tableau III.7).

Tableau III.7 : Énergies d'interaction ≥ -10 kcal/mol entre les acides aminés de pF1 et la protéine EcFur←Vc.

Résidus pF1	E_{inter} (kcal/mol)
R1	-11,8
L2	-10,5
W3	-24,3
C4	-11,6
R5	-34,8
Y6	-12
H8	-13,3
D13	-13,8

En comparaison avec les valeurs énergétiques du Tableau III.6, on constate que les interactions du peptide ont changé. Certaines se sont renforcées tandis que d'autres se sont affaiblies. Ces variations sont dues à un léger changement conformationnel du peptide lors de l'amarrage. En effet, la boucle peptidique située entre les résidus 7 et 13, est différente dans

les deux amarrages de pF1 sur EcFur←Pa et EcFur←Vc et cela a peut-être joué sur l'énergie libre ΔG observée.

L'analyse des liaisons hydrogènes entre le peptide et la protéine après un raffinement par minimisation d'énergie (100 pas de SD suivi de 1000 pas d'ABNR) avec CHARMM avec un « cutoff » de distance de 2,4 Å et d'angle de 120° est présentée sur le Tableau III.8.

Peptide		E_{inter} (kcal/mol)	Protéine	
Résidus	Atomes		Résidus	Atomes
R1	HH12	-2,1	D64 (A)	OD2
	HH22	-2,8	D64 (A)	OD2
	O	-3,3	G75 (B)	H
L2	O	-3,1	E74 (B)	H
W3	H	-2,5	N72 (A)	O
	O	-3,4	N72 (A)	H
	HE1	-3,6	Y130 (B)	OH
C4	H	-3,2	N72 (B)	O
	O	-3,4	N72 (B)	H
R5	H	-3,6	N72 (A)	O
	HH21	-6,4	Y130 (A)	OH
	O	-3,5	E74 (A)	H
	HH11	-3,1	S126 (B)	OG
	HH12	-6,4	Y128 (B)	OH
	HH22	-3,9	S126 (B)	OG
Y6	OH	-6,7	R70 (A)	HH21
H8	O	-4,1	R70 (B)	HH11
T12	HG1	-4,7	Y56 (A)	OH
D13	OD1	-4,5	R57 (B)	HH11
	O	-6,7	R57 (B)	HH21

Tableau III.8: Énergies d'interaction entre atomes de pF1 et VcFur←Vc formant des liaisons hydrogènes. Les résidus et atomes de la protéine sont colorés en vert et en rouge respectivement pour le monomère **A** et **B**. Les atomes HH, HG et OG sont des notations des types d'atomes des chaines latérales dans CHARMM (voir **Figure III.26**).

Nous venons de voir dans le Tableau III.7, les énergies d'interaction des résidus de pF1 interagissant fortement avec la protéine. Les résidus de la protéine intervenant dans la liaison avec le peptide sont illustrés dans le Tableau III.8. Ces résidus protéiques contribuant à l'interaction avec pF1 par des liaisons hydrogènes sont les **Y56, R57, D64, R70, N72, D64, E74, G75, S126, Y128** et **Y130**. Les énergies d'interaction des résidus protéiques interagissant essentiellement *via* le « backbone » ; **E74, N72, G75**, sont faibles comparées aux interactions des chaines latérales.

Les résidus protéiques qui interagissent donc fortement avec pF1 par leurs chaines latérales sont :

- D64 qui interagit avec R1 ou D13 de pF1
- R57 interagissant avec D13 de pF1
- R70 qui interagit avec Y6 de pF1
- Y56 interagissant avec T12 de pF1. Cette tyrosine avait été identifiée comme interagissant avec les bases de l'ADN (Tiss *et al.* 2005).
- Y128 et Y130 qui interagissent avec pF1.
- Y130 interagit avec W3 de pF1
- S126 avec R5 de pF1.

Nous allons maintenant voir les amarrages des autres séquences tronquées et/ou mutées de pF1 avec EcFur←Vc et pouvoir comparer les résultats avec ceux des tests expérimentaux.

III. C. 2. b| Amarrage de séquences tronquées de pF1

Afin de vérifier l'importance des interactions de pF1 proposées et aussi pour prédire les mutations à effectuer dans les tests d'activité biochimique, nous avons réalisé des amarrages des peptides tronqués/mutés de pF1 sur le modèle EcFur←Vc. Le Tableau III.9 résume les résultats des séquences de pF1 qui inhibent EcFur *in vitro* dans les tests d'activité biochimiques (confère chapitre II).

Notations	séquences pF1	ΔG (kcal/mol)	Ki (pM)
pF1 (1-13)	RLWCRYPHPPLTD	-19	0,007
pF1 (1-10)	RLWCRYPHPP	-18	0,07
pF1 (1-9)	RLWCRYPHP	-17	1
pF1 (1-8)	RLWCRYPH	-17	1
pF1 (2-9)	LWCRYPHP	-14,5	24
pF1 (3-9)	WCRYPHP	-13	395

Tableau III.9: Récapitulatif des énergies libres et constantes d'inhibition calculées de séquences tronquées de pF1 interagissant avec le modèle EcFur→Vc

Tous ces amarrages montrent que le peptide interagit favorablement dans le site actif d'EcFur que nous venons de voir (Figure III.26 et Figure III.27). L'énergie libre de liaison est inférieure à -12 kcal/mol pour toutes ces séquences. Les résidus **W3, C4, R5, Y6** et **H8** contenus dans la séquence pF1(3-9), interagissent fortement avec la protéine (Tableau III.7).

On avait montré expérimentalement que cette séquence était la plus petite pour inhiber EcFur. Lorsqu'une mutation touche à l'un de ces résidus, le peptide n'inhibe plus la protéine (confère chapitre II). De plus, les expériences biochimiques ont montré que les résidus **R1-L2** et **P10-D13** ne sont pas indispensables pour inhiber la protéine. De la même façon, lorsqu'un des résidus **W3, C4, R5, Y6** et **H8** est absent dans la séquence de pF1, l'énergie libre de liaison avec la protéine diminue Ces résidus sont donc très importants pour l'interaction avec la protéine comme nous venons de le voir dans ces résultats de « docking ». Ils sont aussi indispensables pour inhiber l'activité de la protéine comme on l'a vu dans les résultats des tests biochimiques. Ces résultats d'amarrage de pF1 sur la protéine sont donc en accord avec ceux des expériences biochimiques réalisées. Toutefois, la séquence pF1(1-8) interagit favorablement en théorie avec Fur avec une énergie libre de liaison identique à celle obtenue pour le peptide pF1(1-9). Or pF1(1-8) n'est pas actif expérimentalement. Le « docking » prévoit donc une affinité forte de pF1(1-8) pour Fur qui n'a pas été confirmée par les expériences *in vitro*. Les tests biochimiques indiquent que la proline **P9** de pF1 est essentielle à l'inhibition de la protéine. La Figure III.28 présente l'amarrage de pF1(3-9) qui est donc la séquence minimale active de pF1 (sept acides aminés).

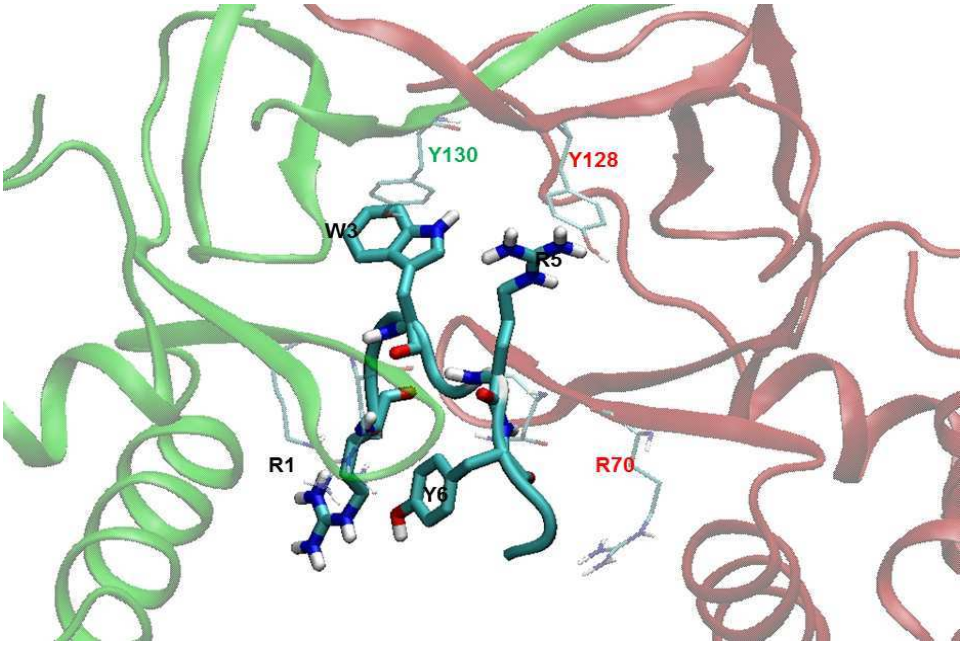


Figure III.28: Amarrage de pF1(3-9) sur EcFur←Vc.

Cet amarrage a une énergie libre de liaison de -13 kcal/mol, très favorable, compatible avec les résultats biochimiques comme nous venons de le voir.

III. C. 3| Amarrage de pF2 sur EcFur←Vc

III. C. 3. a| Conformation de pF2 entier : RQCNICGASLYSY

Le modèle du peptide pF2 utilisé, provient de la conformation optimisée de l'aptamère F2 construit par CHARMM, comme il a été expliqué précédemment à la page - 165 -. Le modèle de l'aptamère F2 et le peptide pF2 qui en est extrait sont montrés sur la Figure III.29, ci-dessous.

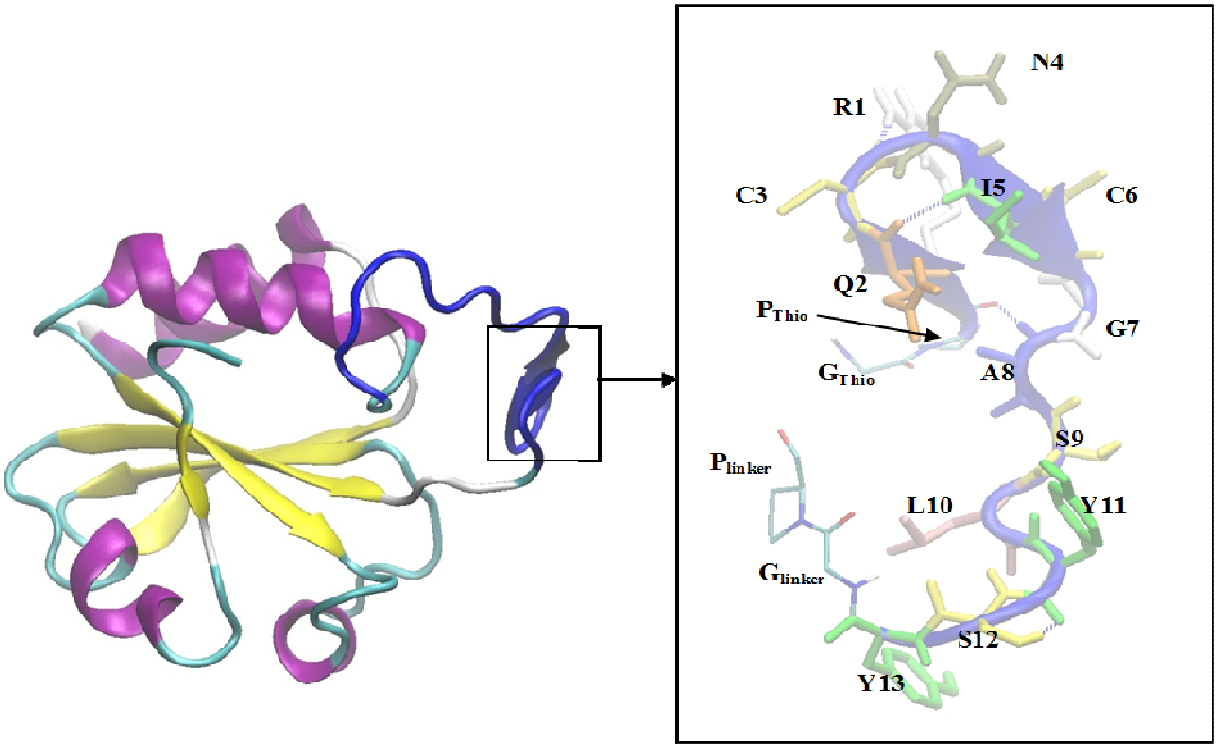


Figure III.29 : Modèle de l'aptamère F2 et extraction du peptide pF2 (rotation 180°).
A gauche, le modèle de l'aptamère F2 construit et optimisé par CHARMM. A droite, le peptide pF2, en bleu est extrait avec le « linker ». Les résidus de la séquence et les liaisons hydrogènes sont représentés.

La conformation du peptide montre un feuillet β entre les résidus R1 et C6 et une boucle non structurée des résidus G7 à Y13. Cette conformation du peptide pF2 a été utilisée pour l'amarrage sur EcFur←Vc. Les premiers « dockings » ont montré statistiquement une interaction favorable avec les grilles FurEcGsy et FurEc_2 (Figure III.18 et Tableau III.3), c'est-à-dire sur le domaine de dimérisation et entre les deux monomères. Les optimisations de l'amarrage se sont focalisées sur cette zone.

III. C. 3. b| Amarrage de pF2 sur EcFur←Vc

Il n'a pas été évident de trouver un amarrage favorable pour ce peptide, qui pourtant est plus actif expérimentalement. Et ce, pour les raisons suivantes :

- Nous n'avions aucune information sur la configuration de la protéine, active ou inactive sur laquelle les activités inhibitrices de pF2 à pF4 s'exercent. Les « dockings » ont donc été réalisés sur les deux formes de la protéine. Ce qui fait un nombre important de calculs à effectuer.
- Les expériences de métallation ont montré une grande affinité de pF2 pour le cobalt (Vitale 2009). Or, le logiciel utilisé pour l'amarrage ne reconnait pas les sites métalliques que nous avons construits. Rappelons que nous construisons le site métallique uniquement pour orienter correctement les ligands du métal.
- Les amarrages ne sont pas effectués sur une structure complète de la protéine. Effectivement, il manque la boucle Cter (134-148) non modélisée par homologie.

Nous expliquions précédemment que l'extrémité Cter (résidus 134-148) a été coupée dans le modèle d'EcFur généré par homologie à partir de VcFur car ces résidus sont absents dans cette structure. Or, les premiers résultats d'amarrage de pF2 sur Fur montraient que le peptide pourrait interagir avec la boucle Cter. Pour avoir des chances d'obtenir un amarrage favorable de pF2, un modèle construit par homologie avec MODELLER où la boucle (résidus 134-148) est présente, a été utilisé. Cette boucle a été raffinée par 100 pas de minimisation SD, suivis de 1000 pas d'une minimisation ABNR.

Nous avons commencé l'amarrage de pF2 en partant de la séquence pF2(1-6) (pas testée expérimentalement) structurée en feuillet β (Figure III.29). Ensuite, les séquences correspondant aux 10 et 11 premiers résidus de pF2, qui étaient actifs *in vitro,* ont été construites à partir de la conformation de pF2(1-6) amarré sur EcFur←Vc. Les conformations de pF2(1-10) et pF2(1-11), amarrées dans les grilles FurEcGsy et FurEc_2, ont été utilisées pour construire pF2 dans sa totalité. Le processus d'AADM a été appliqué avec une DM de 5 ns sur le complexe Fur-pF2. Plusieurs « dockings » ont été lancés en changeant les paramètres du nombre d'évaluations, taille de population initiale et nombre de « runs » pour optimiser les amarrages. La Figure III.30 montre l'amarrage de la meilleure conformation du peptide pF2 obtenue sur le modèle EcFur←Vc.

Le peptide interagit avec le domaine de dimérisation de la protéine, entre les deux sous-unités A (E_{inter} de -60 kcal/mol) et B (E_{inter} de -70 kcal/mol). L'affinité théorique de l'interaction est estimée à -16 kcal/mol (Ki = 1,9 pM).

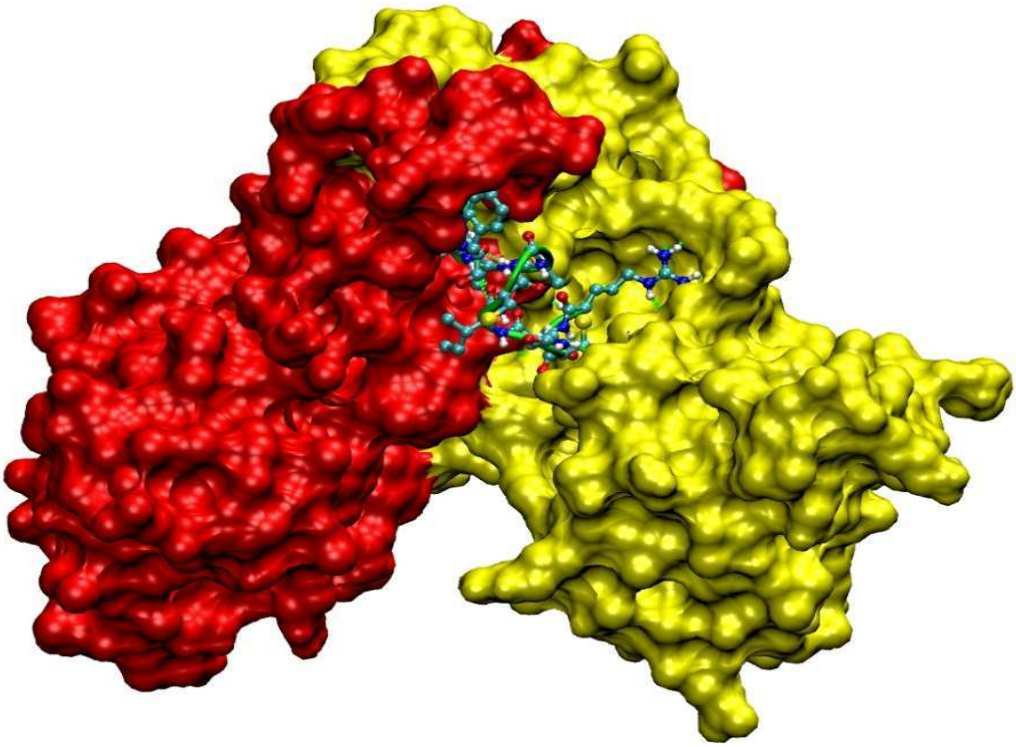


Figure III.30 : Amarrage de la meilleure conformation de pF2 sur EcFur←Vc. Les sous-unités A et B de la protéine sont respectivement en rouge et jaune et le peptide est en vert avec ses chaines latérales.

L'amarrage de pF2 sur Fur est très favorable (**Δ**G = -16 kcal/mol), proche de celui de pF1. Expérimentalement, ce peptide est le plus actif parmi les peptides dérivés des aptamères peptidiques (confère Chapitre II). Le modèle d'interaction proposé, ici, bien que très favorable, pourrait être différent de celui qu'il peut y avoir en réalité. En effet, ce peptide, ayant une grande affinité pour les métaux, pourrait interférer avec la métallation de Fur et empêcher son activation. Ce qui expliquerait son mécanisme d'inhibition très efficace, observé expérimentalement. Ce mode d'inhibition n'a pas été pris en compte dans les protocoles d'amarrage des peptides. Par contre, un peptide pF2 ponté par ses C3 et C6 (C3-C6) et amarré sur Fur, n'a pas l'affinité théorique espérée pour cette conformation peptidique. Les aptamères peptidiques ayant été sélectionnés pour leurs interactions avec Fur, le peptide pF2 pourrait tout à fait interagir avec Fur pour l'inhiber autrement que l'interférence avec l'activation de la protéine. Les interactions de l'amarrage de pF2 avec Fur ont été analysées plus finement par les calculs d'énergies d'interaction avec CHARMM. L'analyse de ces interactions a lieu après une minimisation (100 pas de SD et 1000 pas ABNR) qui a permis

une légère structuration de la conformation du peptide en hélice π (Figure III.31). Les acides aminés présentant une énergie d'interaction avec la protéine (E_{inter}) inférieure à - 8 kcal/mol sont listés dans le Tableau III.10 ci-dessous.

Résidus pF2	E $_{inter}$ (Kcal/mol)
R1	-30,3
Q2	-31,3
S9	-8,7
Y11	-13,1
Y13	-23,4

Tableau III.10: Énergies d'interaction ≥ -8 kcal/mol entre les acides aminés de pF2 et la protéine EcFur←Vc

Ce tableau montre que les résidus **R1, Q2, Y11** et **Y13** de pF2 interagissent fortement avec la protéine (E_{inter} inférieure à - 10 kcal/mol) ainsi que le résidu **S9**. La Figure III.31 ci-dessous montre une vue générale des interactions de pF2 avec la protéine EcFur.
Les atomes de pF2 et EcFur←Vc formant des liaisons hydrogènes sont résumés dans le Tableau III.11.

Peptide		E_{inter} (kcal/mol)	Protéine	
Résidus	Atomes		Résidus	Atomes
R1	H	-2,8	S126 (A)	OG
	HE	-1,7	E81 (A)	OE1
	HH12	-5,7	H86 (A)	O
	HH21	-3,5	E81 (A)	OE1
	HH22	-5,4	Q84 (A)	O
	O	-4,7	Y130 (B)	HH
Q2	H	-3	R70 (A)	O
	OE1	-3,2	N72 (A)	H
	OE1	-5,3	Y128 (A)	HH
	O	-3,3	R70 (A)	H
	HE21	-4,7	F73 (B)	O
	HE22	-3,1	N72 (B)	O
S9	O	-3,5	T123 (A)	HG1
Y11	HH	-5,2	C138 (B)	O
Y13	H	-2,5	E140 (B)	OE
	OH	-1,6	D141 (B)	H
	OH	-2,9	E142 (B)	H
	HH	-5,3	R139 (B)	O

Tableau III.11: Énergies d'interaction entre atomes de pF2 et EcFur←Vc formant des liaisons hydrogènes. Les résidus et atomes de Fur sont colorés en vert et rouge respectivement pour le monomère B et A. Les atomes HH, HG et OG sont des notations des types d'atomes des chaines latérales dans CHARMM.

Les deux premiers résidus de pF2 : **R1** et **Q2** interagissent très fortement avec Fur dans ce modèle d'amarrage. Expérimentalement, nous avons vu également l'effet de ces résidus sur les valeurs d'IC_{50} des séquences pF2(2-10) et pF2(3-10), ne contenant pas ces

deux résidus. Ces deux résidus semblent interagir fortement sans être essentiels pour inhiber Fur, d'après cette comparaison. De plus, ces deux acides aminés agissent dans la même zone que le résidu **R5** de pF1. Par contre, les troncations des résidus **Y11** et **Y13** sont bénéfiques pour l'activité *in vitro* de pF2. Or, ces deux derniers résidus interagissent aussi fortement avec la protéine, d'après les calculs. Signalons que ces résidus de pF2 : **Y11** et **Y13** interagissent avec la boucle (134-148), absente de la structure utilisée pour construire ce modèle de Fur. Ces interactions stabilisatrices avec la boucle Cter, par le calcul, pourraient donc être artefactuelles. Le modèle prédit ici peut ne pas être la conformation que le peptide et la protéine adoptent *in vitro*. Ces interactions des résidus Y11 et Y13 de pF2 avec Fur pourraient également être à l'origine du changement conformationnel observé sur le peptide pF2. La conformation du peptide amarré a beaucoup changé par rapport à celle d'avant l'amarrage (Figure III.29). Le peptide passe ainsi d'une conformation où pF2 comporte un feuillet β entre les résidus R1 et G7 à une configuration où pF2 contient une hélice **π** entre les résidus (I5 et L10 Figure III.29). Ces changements de conformations sont intervenus après les divers étapes de dynamique du complexe en vue d'améliorer l'énergie libre de liaison du peptide.

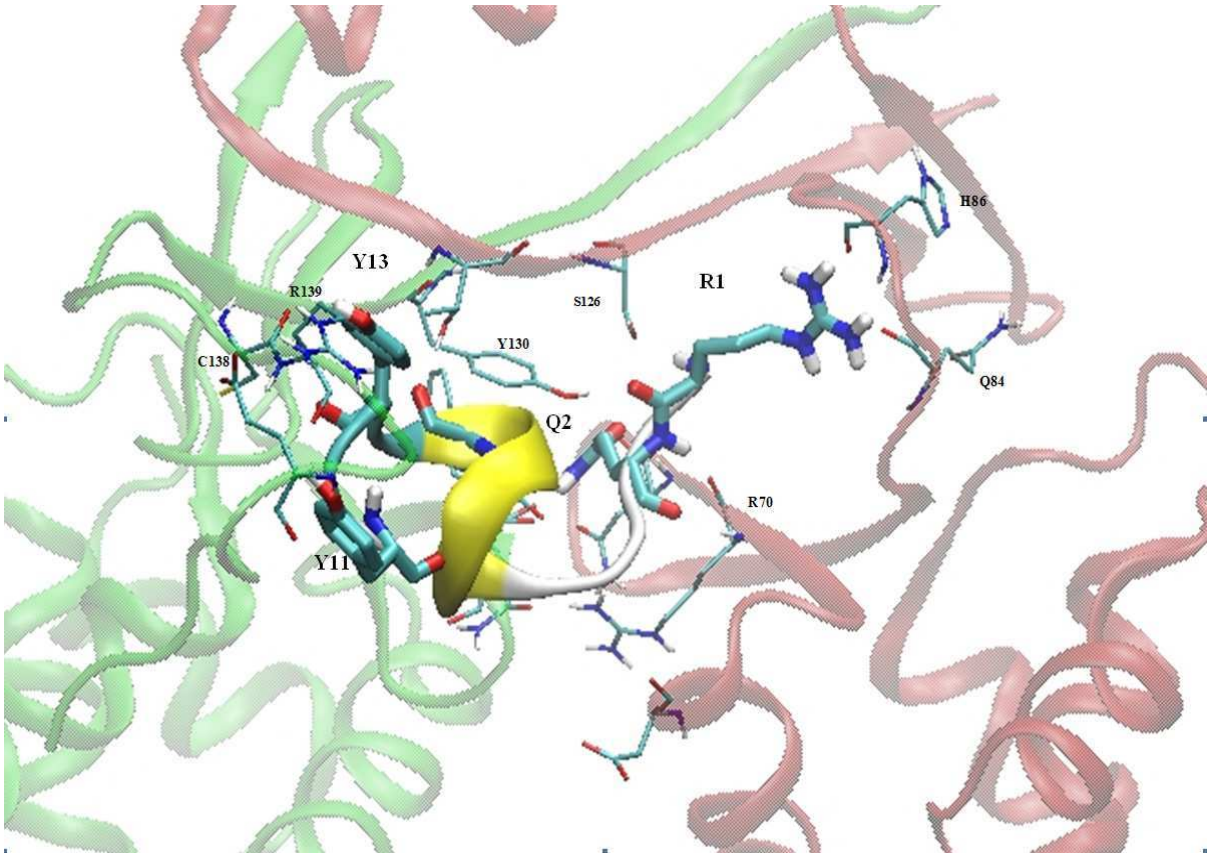


Figure III.31: Vue d'ensemble du « docking » de pF2 sur la protéine EcFur←Vc.
Les sous-unités A et B de la protéine sont respectivement en rouge et vert, le peptide est en jaune. Les chaines latérales de pF2 sont en licorice (taille 3), celles de la protéine en licorice (taille 1).

III. C. 3. c| Amarrage des séquences tronquées de pF2

Comme pour le peptide pF1, plusieurs mutations dans la séquence de pF2 et des peptides tronqués ont été testées. Nous avons réalisé des amarrages des peptides pF2 plus courts sur le modèle EcFur←Vc en utilisant le procédé AADM. Le Tableau III.12 résume les résultats d'amarrage pour les séquences de pF2, dont on a montré qu'elles inhibent la protéine Fur *in vitro*.

Notations	**séquences pF2**	**ΔG kcal/mol)**	**Ki (nM)**
pF2(1-13)	**RQCNICGASLYSY**	**-16**	**1,9. 10^{-3}**
pF2(1-11)	**RQCNICGASLY**	**-11,85**	**2,07**
pF2(1-10)	**RQCNICGASL**	**-11,64**	**2,94**
pF2(2-10)	**QCNICGASL**	**-7,13**	**5,94. 10^{3}**
pF2(3-10)	**CNICGASL**	**-2,82**	**8,6. 10^{6}**

Tableau III.12: Récapitulatif des énergies libres et constantes d'inhibition calculées de séquences tronquées de pF2 interagissant avec le modèle EcFur→Vc

Ces résultats semblent montrer une fois de plus que les résidus **R1, Q2, Y11 et Q13** sont importants pour l'amarrage des séquences tronquées de pF2. Ces peptides tel que la séquence minimale de pF2 active *in vitro* sur la protéine EcFur : pF2(3-10) CNICGASL s'amarrent dans la même région que pF2 entier (Figure III.32). L'énergie libre de liaison de pF2(3-10) est seulement de -2,8 kcal/mol, ce qui est très faible quand on sait que pF2(1-13) a une énergie libre de liaison de -16 kcal/mol. Cette faible affinité pour la protéine est probablement due à l'absence des résidus **R1, Q2, Y11** et **Q13** qui interagissent fortement avec la protéine. Expérimentalement, on a trouvé que l'absence de ces résidus affecte fortement la valeur de l'IC_{50} qui passe de 6 µM pour pF2(1-13) à 75 µM pF2(3-10). Ces résidus pourraient donc être impliqués dans l'interaction et l'inhibition de la protéine mais peut être qu'ils ne réagissent pas de la façon proposée.

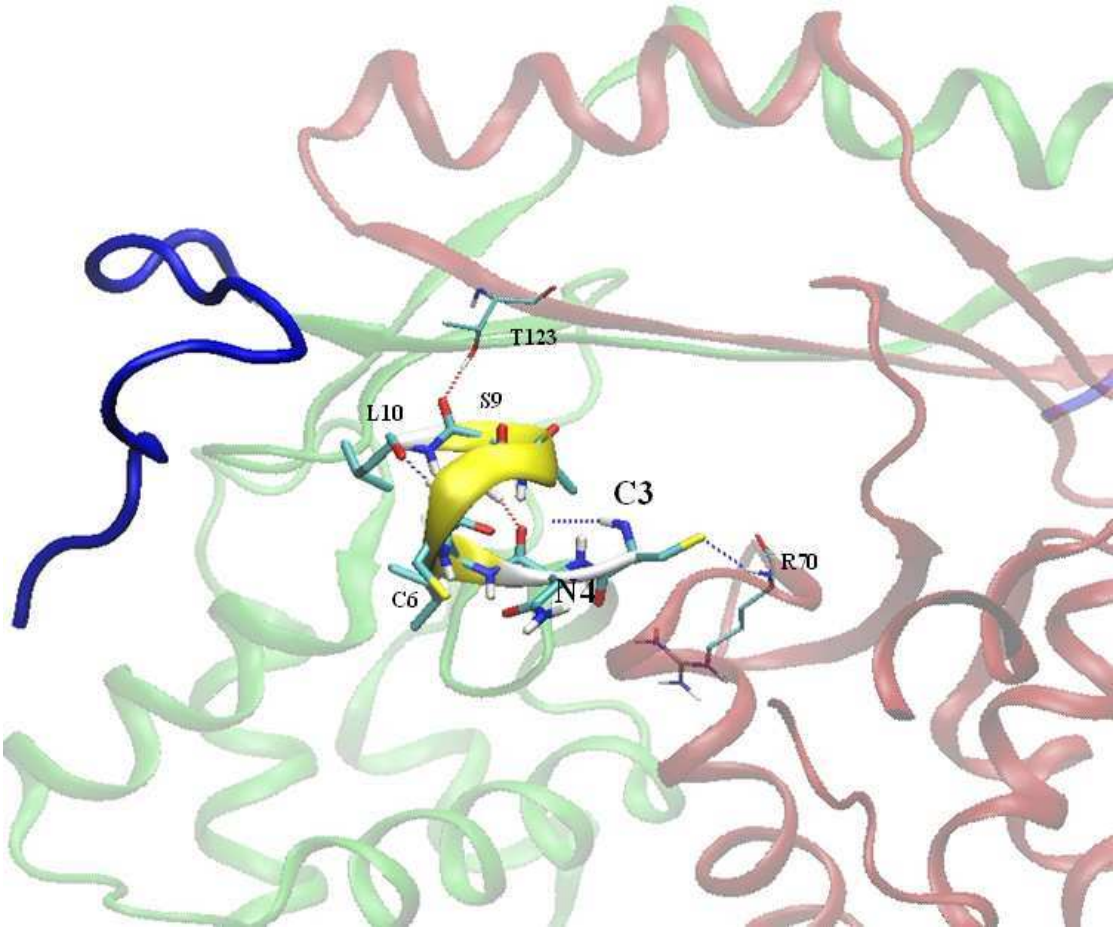


Figure III.32 : Amarrage de pF2(3-10) sur EcFur←Vc. Le monomère A en rouge et B en vert, la boucle Cter (134-148) de la protéine en bleu et le peptide en jaune avec ses chaines latérales.

On remarque également dans les Figure III.32 et Figure III.31 que pF2 interagit avec des résidus de la protéine qui interagissaient aussi avec pF1. C'est le cas des résidus **R70, N72, S126, Y128** et **Y130**, des résidus d'ailleurs très conservés chez la famille Fur.

III. C. 4| Amarrage de pF3 sur EcFur←Vc

III. C. 4. a| Conformation du peptide pF3 : ETCKCGSQVWRHS

Le modèle de l'aptamère F3 et du peptide pF3 qui en est extrait sont montrés sur la Figure III.33.

La conformation initiale du peptide pF3 utilisée provient de l'aptamère F3 construit et optimisé par CHARMM, comme il a déjà été expliqué précédemment à la page - 165 -. Il n'y a pas de structuration particulière comme dans les conformations de pF1 et pF2. Les premiers

amarrages montrent que ce peptide peut être convenablement amarré en utilisant les grilles FurEcGsy et FurEc_2 (Tableau III.4 et Tableau III.3), mais différemment du peptide pF2 que nous venons de voir.

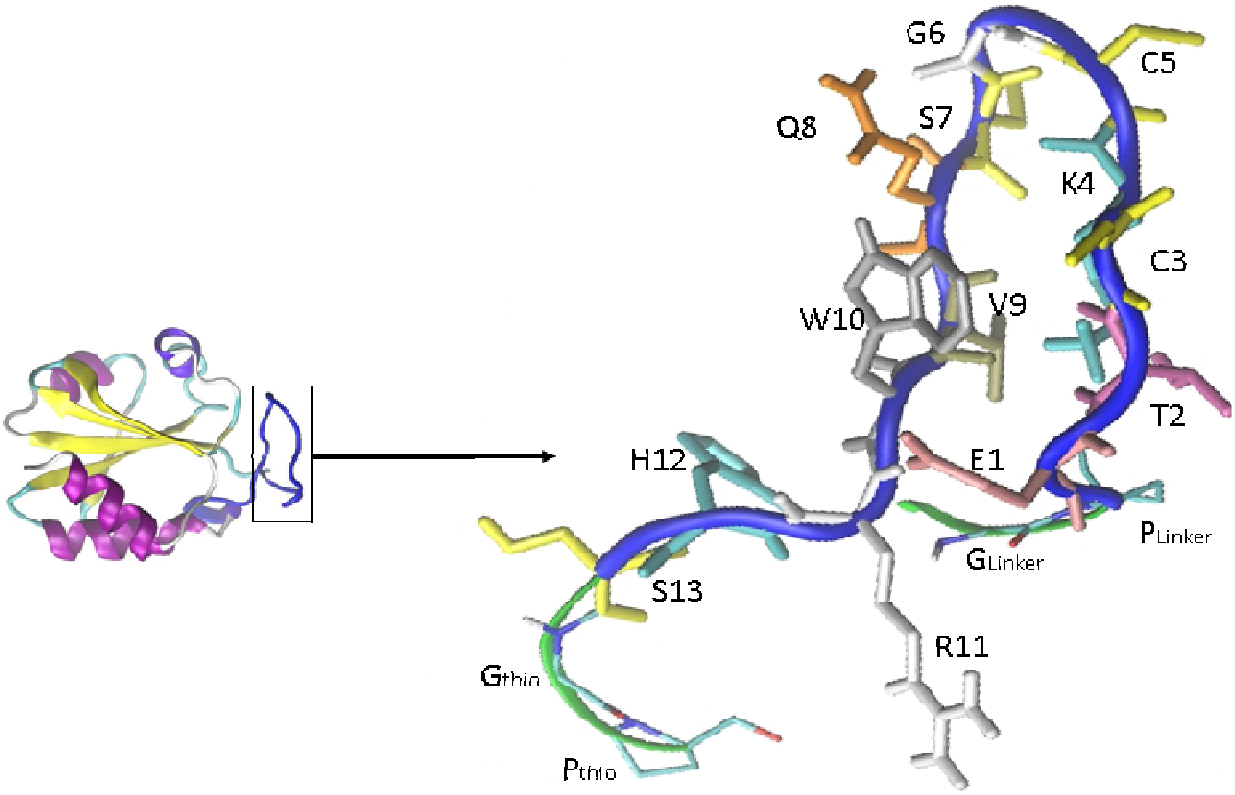


Figure III.33: Modèle de l'aptamère F3 et extraction du peptide pF3.
A gauche, le modèle de l'aptamère construit et optimisé par CHARMM. A droite, le peptide pF2, en bleu est extrait avec le « linker ». Les résidus de la séquence et les liaisons hydrogènes sont représentés.

III. C. 4. b| Amarrage de pF3

Pour obtenir l'amarrage du peptide pF3, nous partons de l'amarrage optimisé de pF3(1-10) sur le modèle EcFur←Pa, comme coordonnée initiale du peptide. Le peptide s'amarrait sur un seul monomère de la protéine. A partir de la conformation de pF3(1-10), le peptide entier pF3(1-13) a été construit en rajoutant les 3 acides aminés Cter avec CHARMM. Le complexe pF3-EcFur←Vc a été soumis au processus AADM. Les cinq meilleures conformations où pF3(1-13) interagissait favorablement avec la protéine ont été sélectionnées. Plusieurs « dockings » en faisant varier les paramètres du nombre d'évaluations et de taille de population initiale, ont été effectués pour chaque « frame » sélectionnée afin d'obtenir une énergie libre favorable. La conformation ayant la meilleure énergie libre est présentée sur la Figure III.34.

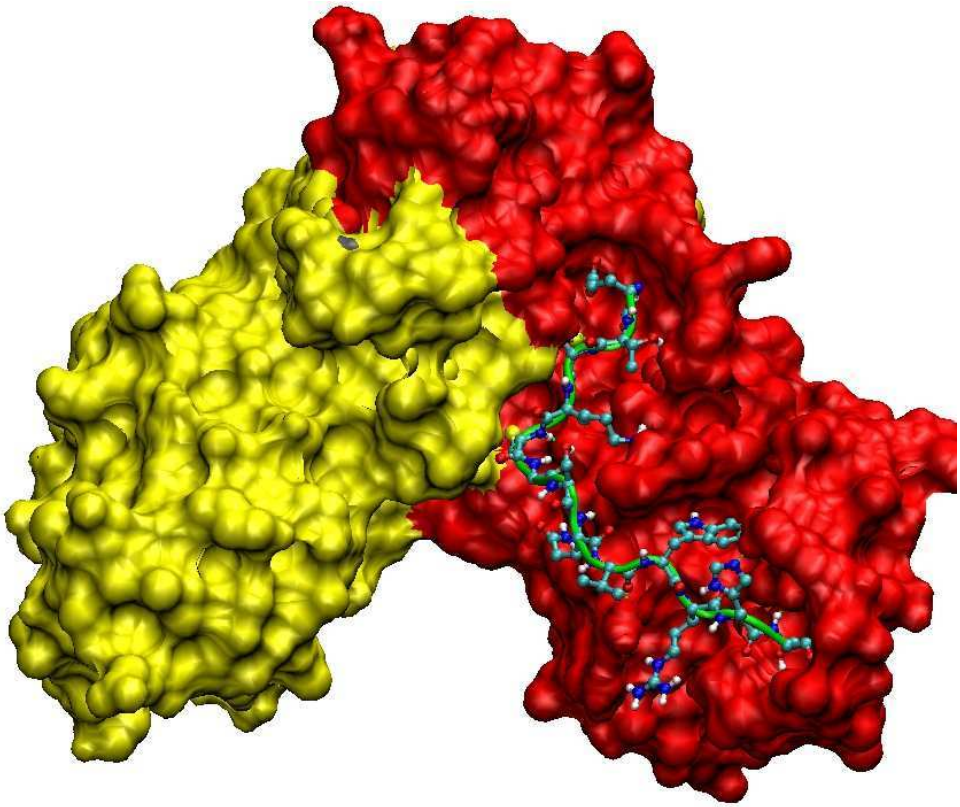


Figure III.34 : Amarrage de pF3 sur EcFur←Vc. Représentation en surface de la protéine : en jaune le monomère A et en rouge le monomère B. Le peptide est représenté avec ses chaines latérales.

L'énergie libre de liaison de ce complexe est de -16 kcal/mol avec une constante d'inhibition de 994 fM. Le peptide pF3 interagit presque exclusivement avec un seul monomère : sur la Figure III.34, la sous-unité B (en rouge). En effet, l'énergie d'interaction calculée est de -138 kcal/mol pour le monomère B et seulement de -12 kcal/mol pour son homologue A. On peut avoir une vue plus détaillée des résidus de pF3 et de la protéine qui interagissent sur la Figure III.35 ci-dessous.

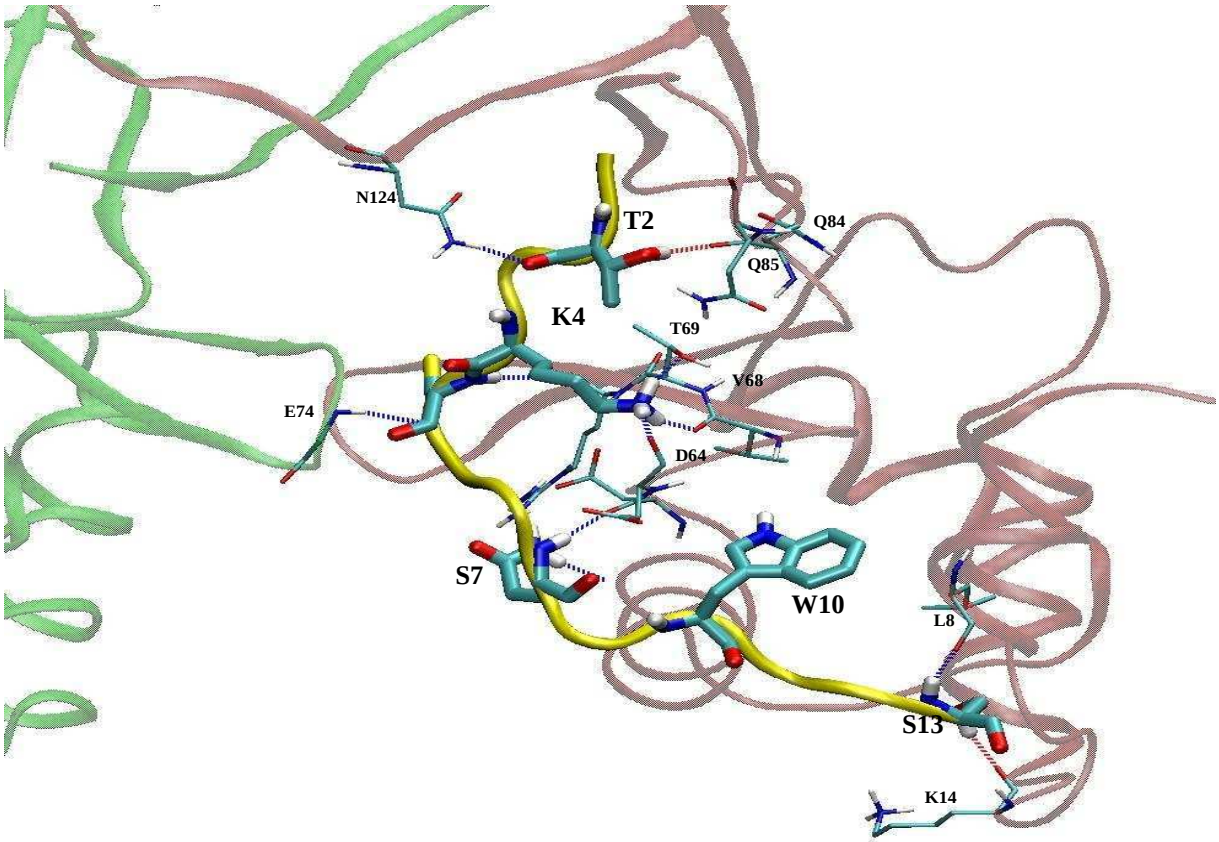


Figure III.35 : Vue d'ensemble des principales interactions entre pF3 et EcFur←Vc. Le peptide est représenté en jaune avec ses chaines latérales en « licorice » (taille 3) et la protéine monomères A et B, respectivement en vert et rouge, chaines latérales en « licorice » (taille 1).

Le peptide pF3 interagit avec une forte affinité (ΔG = -16 kcal/mol) à la fois avec le domaine de liaison à l'ADN et le domaine de dimérisation de la protéine, ce qui est en parfait accord avec les résultats des tests double-hybride sur les domaines protéiques de EcFur réalisés par Nadia Abed (Abed et al. 2007). Les interactions de pF3 avec EcFur←Vc ont été évaluées avec CHARMM par le calcul de l'énergie d'interaction et nous présentons dans le Tableau **III.13** les acides aminés de pF3 liant fortement les résidus de la protéine (énergie d'interaction < -10 kcal/mol).

Résidus pF3	E_{inter} (kcal/mol)
T2	-13,2
K4	-19,1
C5	-11,9
Q8	-15,3
W10	-13,5
H12	-13,5
SER13	-28,0

Tableau III.13: Énergies d'interaction ≥ -10 kcal/mol entre les acides aminés de pF3 et la protéine EcFur←Vc.

Ce tableau montre que les résidus **T2, K4, C5, Q8, W10, H12** et **S13** sont très importants dans l'amarrage de pF3 présenté sur les Figure III.34 et Figure III.35. Ces résidus sont tout au long de la séquence du peptide. Peut-être sera-t-il difficile de trouver une séquence minimale peptidique active sur la protéine Fur ? Seulement deux séquences tronquées de pF3 ont été testées expérimentalement : pF3(3-13) et pF3(1-11). La séquence pF3(1-11) étant active, les résidus **S13** et **T12** ne doivent pas être indispensables à l'activité du peptide. Or, ces résidus interagissent très fortement avec la protéine (E_{inter} = -28 et -13,5 kcal/mol respectivment). Comme pour pF1 et pF2, il se pourrait que certaines peptidiques aient une forte interaction avant la protéine mais ne soient pas vraiment indispensables pour inhiber la protéine.

Les atomes de pF3 qui établissent des liaisons hydrogènes avec EcFur ont été analysés par CHARMM, afin de connaitre les résidus de la protéine qui sont impliqués dans l'interaction avec le peptide. L'analyse a été effectué après un raffinement (SD 100 pas et ABNR 1000 pas) du modèle du complexe EcFur-pF3 précédent avec un « cutoff » de distance de 2,4 Å et d'angle de 120°. Les résultats sont présentés dans le Tableau III.14.

Peptide	pF3	E_{inter} (kcal/mol)	Protéine	
Résidus	Atomes		Résidus	Atomes
T2	OG1	-2,2	Q85 (B)	HE21
	HG1	-5,2	Q84 (B)	O
	O	-3,4	N124 (B)	HH22
K4	H	-2,1	E74 (A)	OE2
	HZ1	-5,3	D64 (B)	O
	HZ2	-6,5	V68 (B)	O
	HZ3	-5,0	T69 (B)	OG1
C5	O	-2,0	G75 (B)	H
	O	-2,4	G76 (B)	H
	H	-3,6	R70 (B)	O
Q8	H	-3,4	Q63 (B)	O
	HE21	-3,3	N60 (B)	O
	HE22	-2,2	D63 (B)	OD2
W10	HE1	-2,9	F62 (B)	OG
H12	O	-5,3	R19 (B)	HH11
S13	H	-3,3	L8 (B)	O
	OG	-5,0	R19 (B)	HH12
	HG	-5,3	K14 (B)	O
	O	-2,7	L13(B)	H
	O	-2,8	K14 (B)	H

Tableau III.14: Énergies d'interaction entre atomes de pF3 et EcFur←Vc formant des liaisons hydrogènes. Les résidus et atomes de la protéine sont colorés en vert et en rouge respectivement pour le monomère A et B. Les atomes HH, HG et OG sont des notations des types d'atomes des chaines latérales dans CHARMM.

Nous retrouvons les résidus **T2, K4, C5, Q8, W10, H12** et **S13** du peptide pF3 qui interagissent fortement avec la protéine. Ces acides aminés qui interagissent avec les résidus d'EcFur en formant des liaisons hydrogènes (confère le tableau ci-dessus). Cet amarrage montre que pF3 interagit sur les deux domaines Nter et Cter d'une sous-unité d'EcFur, en formant des liaisons hydrogènes le long du monomère. Les principaux résidus de la protéine qui interviennent dans l'interaction pF3 dans ce modèle sont **K14, R19, R70, T69, V68, D64, Q84**, sans exclure les acides aminés **N124** et **S126** du domaine de dimérisation de Fur (Tableau III.14) ($E_{inter} \geq$ -3,6 kcal/mol). On retrouve une fois de plus des résidus de Fur qui interagissaient avec les peptides pF1 et pF2 ; comme **R70, D64, S126**. Ces résidus de la protéine semblent délimiter une zone dans la protéine qui serait sensible à la perte d'activité de Fur lorsque les peptides interagissent avec la protéine.

Nous allons maintenant voir les amarrages des autres séquences tronquées et/ou mutées de pF3 avec EcFur←Vc.

III. C. 4. c| Amarrage des séquences tronquées de pF3

Ces séquences proposées ici n'ont pas encore été testées expérimentalement. Ainsi, les amarrages et les valeurs de ΔG permettront de prédire de prochaines mutations. Les résultats sont listés dans le **Tableau III.15** ci-dessous.

Notations	séquences pF3	ΔG (kcal/mol)	Ki
pF3 (1-13)	ETCKCGSQVWRHS	-16	994 fM
pF3 (1-11)	ETCKCGSQVWR	-11,7	2,5 nM
pF3 (1-10)	ETCKCGSQVW	-12,2	1,5 nM
pF3 (1-9)	ETCKCGSQV	-9,5	105,8 nM
pF3 (1-8)	ETCKCGSQ	-8,7	434,7nM
pF3 (2-13)	TCKCGSQVWRHS	-18,2	42 fM
pF3 (3-13)	CKCGSQVWRHS	-16,8	474 fM
pF3 (4-13)	KCGSQVWRHS	-14,9	12,4 pM
pF3 (5-13)	CGSQVWRHS	-12,9	356 pM

Tableau III.15: **Récapitulatif des énergies libres et constantes d'inhibition calculées pour des séquences tronquées de pF3 interagissant avec le modèle EcFur→Vc**

Dans ce tableau résumant les amarrages des séquences de pF3, tous les résidus ont été enlevés sauf les acides aminés **C5, G6, S7** et **Q8**. Nous avons vu dans le

Tableau **III.13** que les résidus **T2, K4, C5, Q8, W10, H12** et **S13** de pF3 interagissaient fortement avec Fur d'après le calcul des énergies d'interaction du modèle de « docking » de pF3. Dans le Tableau III.15, on constate que lorsque la troncation touche les résidus Cter de pF3 ; **W10, H12** et **S13**, l'énergie libre de liaison diminue fortement, passant de -16 kcal/mol pour pF3(1-13) à -8,7 kcal/mol pour pF3(1-8), qui n'est d'ailleurs pas très favorable ($\Delta G > -10$ kcal/mol). Cela indique une très forte contribution de ces résidus Cter de pF3 dans l'interaction avec la protéine dans le modèle d'amarrage proposé. Contrairement aux résidus Cter, les troncations touchant les résidus Nter ne semblent pas être importantes dans l'affinité du peptide pour la protéine. La troncation du résidu **E1** de pF3 améliore l'énergie libre de liaison du peptide, comme si ce résidu affectait négativement l'interaction avec la protéine. Comme nous l'avons vu dans le chapitre II, expérimentalement, les séquences tronquées de pF3 : pF3(1-11) et pF3(3-13) ont été testées. La séquence pF3(3-13), testée une seule fois ne semble pas inhiber la protéine. Par contre, la séquence pF3(1-11) a inhibé Fur dans 2 cas sur 5 avec un IC_{50} proche de 150 µM. Quand on sait que le peptide pF3(1-13) inhibe Fur avec un IC50 de 30±17 µM, cela montre tout de même l'effet des résidus **H12** et **S13** dans l'inhibition de la protéine. Il faudra plus de tests biochimiques sur ces séquences de pF3 pour affiner nos conclusions.

III. C. 5| Conclusions et discussions sur les amarrages des peptides pF1 à pF3 sur le modèle EcFur←Vc

Pour prédire l'amarrage des peptides pF1 à pF3 sur la protéine EcFur, nous avons construit des modèles par homologie avec les structures de Fur disponibles avec le logiciel MODELLER. Les modèles générés et choisis ont été validés par une superposition des domaines Nter des modèles et de la structure 3D du domaines Nter d'EcFur disponible (Pecqueur et al., 2006). Les amarrages ont d'abord débuté avec le modèle construit à partir de la première structure de Fur décrite : PaFur (Pohl *et al.* 2003) qu'on a nommé/noté EcFur←Pa : pourcentage de similarité avec EcFur 77% dans la séquence, mais seulement 66% d'homologie dans la structure. Ce modèle EcFur←Pa a permis de mettre en place une méthodologie permettant de prédire l'amarrage des peptides de 13 acides aminés issus des aptamères peptidiques. Compte tenu de la difficulté pour obtenir des énergies libres de liaison favorables dans les « dockings », à cause surtout de la grande flexibilité des peptides, nous avons utilisé une approche que j'ai nommée AADM, pour amarrage moléculaire assisté par

dynamique moléculaire, qui consiste à prendre un modèle d'amarrage Fur-peptide obtenu avec amarrage moléculaire automatique (AutoDock4) avec le récepteur fixe et de soumettre ce complexe à une dynamique en solvatation implicite (méthode EEF1) puis de sélectionner les « frames » où le peptide interagit favorablement avec la protéine ($E_{inter} < 0$). Ces « frames » avec de nouvelles orientations des chaines latérales de la protéine et de nouvelles conformations du peptide ont servi de point de départ à de nouveaux amarrages avec AutoDock4.2. Cette étape peut être répétée plusieurs fois jusqu'à convergence (conformation du complexe Fur-peptide où le ΔG n'évolue plus).

Nous avons montré dans un premier temps un exemple avec pF1 où nous optimisons l'amarrage du peptide sur la protéine EcFur←Pa avec les différentes stratégies utilisées.

A la parution de la structure 3D de VcFur (Sheikh *et al.* 2009) qui présente un pourcentage de similarité avec EcFur de 91% dans la séquence dont 82% ont été utilisés pour l'homologie, donc beaucoup plus proche d'EcFur, un nouveau modèle, noté EcFur←Vc, a été construit. Les amarrages des peptides pF1 à pF3 ont été repris sur ce modèle, représentant mieux la protéine EcFur. Les deux modèles d'EcFur étant très semblables, les conformations des peptides déjà amarrées sur le premier modèle EcFur←Pa, ont été réutilisées pour le deuxième modèle EcFur←Vc. Les amarrages des peptides sur ce dernier modèle étant plus compatibles avec les résultats expérimentaux, ont été présentés comme les modèles d'interaction des peptides pF1 à pF3 sur EcFur.

Parmi les modèles d'amarrages présentés, ceux de pF1 semblent être les plus cohérents avec les tests expérimentaux. Cela est probablement dû à un nombre assez important de données expérimentales. En effet, la double approche étant validée par les résultats expérimentaux, cela a permis d'affiner les amarrages de pF1. Le fait de savoir que pF1 interagit sur la forme métallée de Fur a été un grand pas dans la prédiction du modèle d'interaction de ce peptide avec la protéine. Nous trouvons aussi les meilleurs amarrages des autres peptides pF2 et pF3 sur la forme métallée de Fur mais il faudra des expériences pratiques pour confirmer ces prédictions théoriques.

Voyons à présent dans la quatrième partie de ce chapitre, l'amarrage de pF1 sur les autres protéines Fur.

III. D| Amarrage du peptide pF1 sur les autres protéines Fur

Dans cette section, nous présenterons l'amarrage du peptide pF1 sur d'autres protéines Fur pour lesquelles nous connaissons la structure et/ou nous avons des données expérimentales. Ces amarrages nous donnent d'importantes informations sur la spécificité de l'interaction du peptide pF1 pour EcFur. Nous allons rappeler brièvement les résultats expérimentaux à notre disposition concernant l'interaction des aptamères peptidiques sur les protéines Fur d'autres organismes. Par le test double hybride, nous savons que les aptamères F1, F2 et F3 sont spécifiques d'EcFur. L'aptamère F4 peut interagir également avec les protéines Fur de *Yersinia pestis* (YpFur), *Vibrio salmonicida* (VsFur) et *Pseudomonas aeruginosa* (PaFur) (Figure III.36).

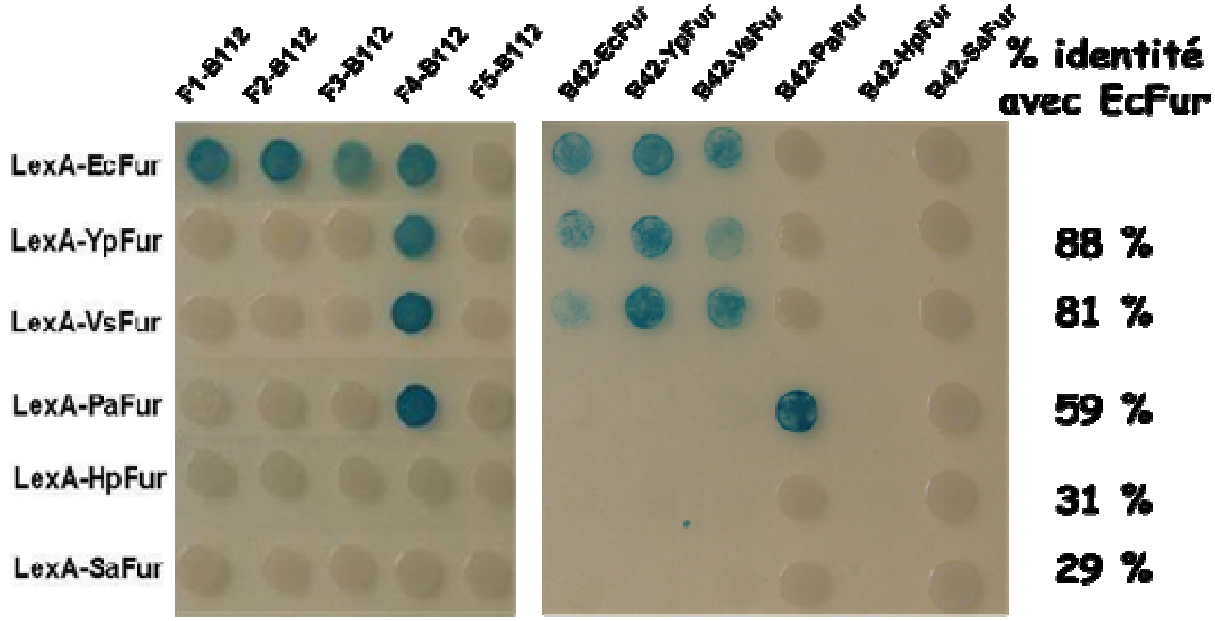


Figure III.36: Matrice d'interaction entre les protéines Fur de différents microorganismes pathogènes et les aptamères peptidiques F1 à F5. La présence d'une coloration bleue montre qu'il y a interaction. Les appâts et les proies sont précisés, respectivement à gauche et au dessus de la matrice. Ec : *Escherichia coli* ; Yp : *Yersinia pestis* ; Vs : *Vibrio salmonicida* ; Pa : *Pseudomonas aeruginosa* ; Hp : *Helicobacter pylori* ; Sa : *Staphylococcus aureus* (Vitale. 2009).

L'absence de contrôle positif (couleur bleue rendant compte de l'homodimérisation de la protéine) dans ce test n'a pas permis de conclure concernant l'interaction protéines Fur d'*Helicobacter pylori* (HpFur) et Fur de *Staphylococcus aureus* (SaFur).

Nous allons voir les résultats théoriques d'amarrage de pF1 sur les protéines PaFur, VcFur et HpFur dont les structures sont connues.

III. D. 1| Amarrage sur la structure de PaFur

Pour obtenir l'amarrage de pF1, nous partons de l'amarrage optimisé de pF1(1-9) sur le modèle EcFur←Vc, comme il a été fait pour les modèles EcFur←Pa et EcFur←Vc. A partir de cette conformation de la séquence pF1(1-9), le peptide entier pF1(1-13), est construit en rajoutant les quatre acides aminés Cter avec CHARMM. Le complexe est alors utilisé dans un protocole AADM. Plusieurs amarrages d'optimisation ont été réalisés en faisant varier les paramètres du nombre d'évaluations et de taille de population initiale, afin d'obtenir une énergie libre favorable. La conformation ayant obtenue la meilleure énergie libre de liaison est présentée sur la Figure III.37 ci-dessous.

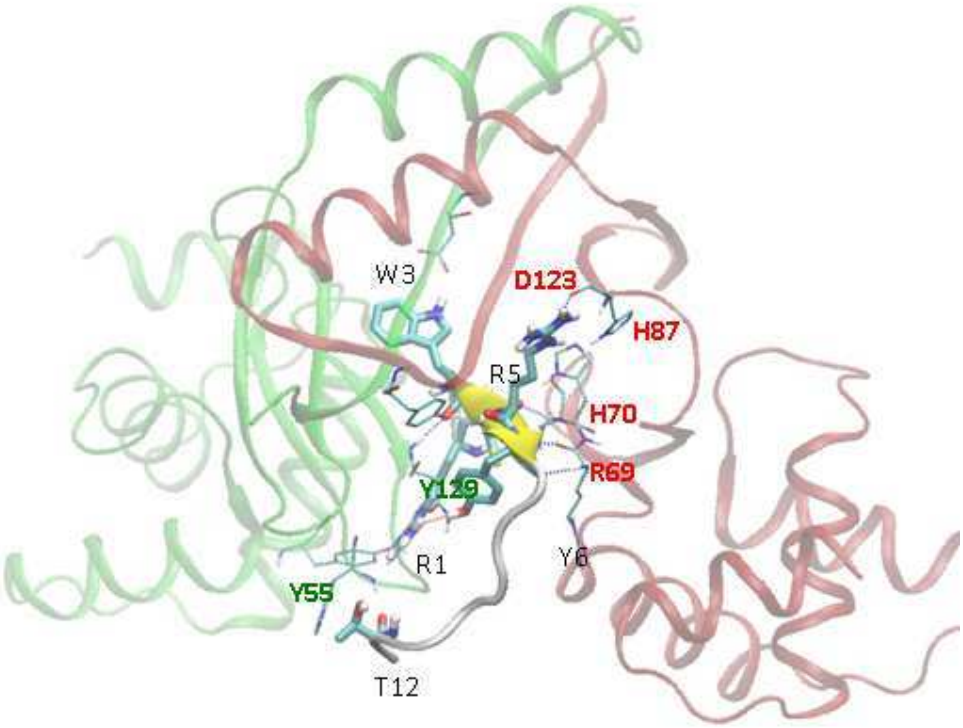


Figure III.37: Amarrage de pF1 sur la structure de PaFur. Représentation de la protéine, en jaune le monomère A et en rouge le monomère B. Le peptide est représenté avec ses chaines latérales. .

L'énergie libre de liaison de ce complexe est de -21,3 kcal/mol avec une constante d'inhibition de 258 aM (atto molaire), ce qui représente la meilleure affinité théorique trouvée jusqu'ici. Le peptide pF1 interagit comme dans les autres amarrages que l'on a vus à l'exception qu'ici, la boucle Cter (résidus 10-13) de pF1 interagit avec le monomère A (Figure III.37) au lieu du monomère B comme dans les amarrages précédents. L'évaluation de l'énergie d'interaction entre peptide et protéine avec CHARMM permet d'apporter plus d'informations : E inter = -78 kcal/mol avec le monomère A et seulement -48 kcal/mol avec le monomère B. Les valeurs résidu par résidu sont présentées dans le Tableau III.16, qui montre

les acides aminés de pF1 liant fortement la protéine (énergie d'interaction < -9 kcal/mol dans ce tableau).

Résidus pF1	E_{inter} (kcal/mol)
R1	-16,3
W3	-20,5
C4	-9,1
R5	-27
Y6	-20,3
T12	-12,6

Tableau III.16 : Énergies d'interaction ≥ -9 kcal/mol entre les acides aminés de pF1 et la protéine PaFur.

Ce tableau montre que les résidus R1, W3, C4, R5, Y6 **et** T12 **sont très importants dans l'amarrage de pF1 sur la protéine. Le changement majeur dans l'interaction du peptide provient du résidu** T12 **qui lie fortement la protéine dans cette configuration. Dans les amarrages précédents de pF1, c'était le résidu** D13 **qui liait fortement la protéine. Dans cet amarrage, le peptide a adopté une conformation qui permet au résidu** T12 **d'interagir avec les résidus** H76, G75 **et** D73 **du monomère A de PaFur. Les atomes de pF1 qui établissent des liaisons hydrogènes avec ceux de PaFur sont listés dans le**

Tableau III.17 ci-dessous.

Peptide		E_{inter} (kcal/mol)	Protéine	
Résidus	Atomes		Résidus	Atomes
R1	HH11	-7	Y55 (A)	OH
	HH12	-2,1	G75 (A)	O
	HH21	-4,2	G75 (A)	O
	HH22	-6,8	N71 (A)	OD1
L2	H	-2,4	E62 (A)	OE1
W3	H	-2,8	R69 (A)	O
	HE1	-3,6	D123 (A)	OD2
	O	-3,6	N71 (A)	H
R5	O	-5,2	Y129 (A)	HH
Y6	OH	-2,7	N71 (A)	HD21
	HH	-5,8	F72 (A)	O
L11	H	-2,9	D73 (A)	O
T12	H	-3,1	D73 (A)	O
	HG1	-3,4	D73 (A)	OD1
	O	-3,1	H76 (A)	H
D13	O	-1,6	G75 (A)	H
C4	O	-4,3	H70 (A)	HD1
	O	-3,3	N71 (B)	H
R5	HH12	-5,2	H87 (B)	O
Y6	H	-3,5	R69 (B)	O
	O	-3,1	R69 (B)	H

Tableau III.17: Énergies d'interaction entre atomes de pF1 et PaFur formant des liaisons hydrogènes. Les résidus et atomes de la protéine sont colorés en vert et en rouge respectivement pour le monomère A et B. Les atomes HH, HG et OG sont des notations des types d'atomes des chaines latérales dans CHARMM

Les principaux résidus de la protéine qui interviennent donc dans l'interaction avec le peptide dans ce complexe sont les résidus **Y55 (A), N71 (A), G75 (A), Y129 (A) F72 (A) H70 (A) H87 (A)** et **R69** (numérotation PaFur) qui établissent des liaisons hydrogènes fortes

($E_{inter} \leq -4$ kcal/mol). Ces résidus avaient déjà été identifiés dans l'interaction de pF1 avec les modèles d'EcFur←Vc et EcFur←Pa. Cela confirme l'interaction de pF1 entre les deux monomères et à la fois avec les domaines Nter et Cter de la protéine Fur. D'après les résultats expérimentaux précédemment présentés, l'aptamère F1 n'interagit pas avec cette protéine. Mais le modèle d'amarrage prédit une bonne affinité entre le peptide pF1 issu de F1 et PaFur. Il sera intéressant de tester le peptide libre avec cette protéine et valider ou non ces amarrages théoriques quand la protéine PaFur, actuellement en cours de purification au laboratoire sera testable. Il se peut que l'aptamère F1 se comporte différemment du peptide pF1 qui a plus de liberté pour pouvoir adopter une conformation présentant une meilleure affinité pour PaFur. Une vue plus détaillée des interactions du complexe pF1-PaFur est montrée sur la Figure III.38, ci-dessous.

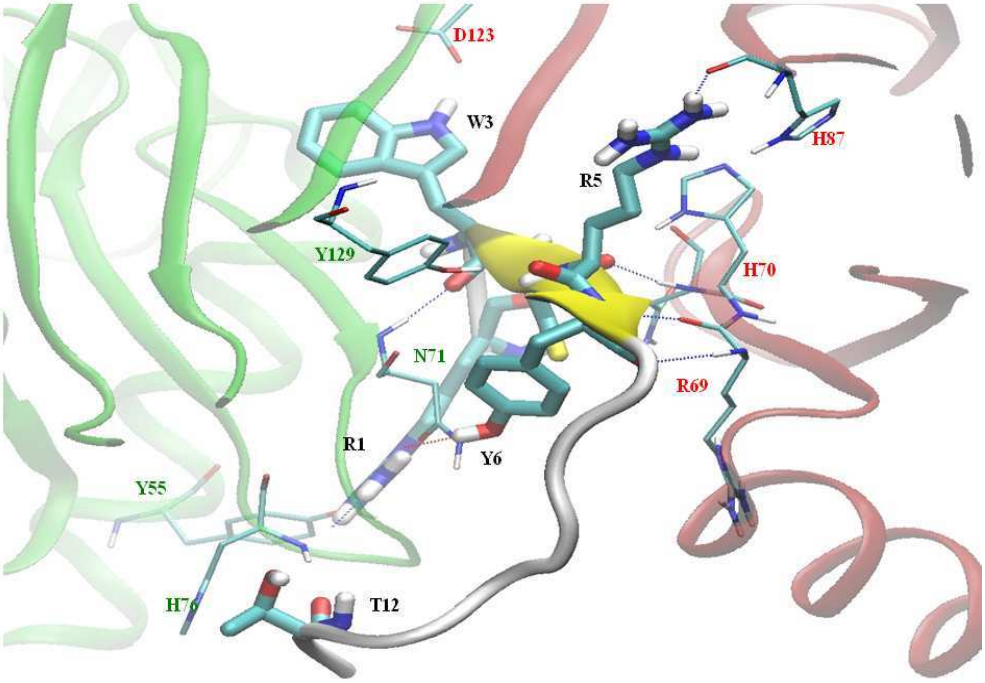


Figure III.38 : Vue d'ensemble des interactions de pF1 avec la structure de PaFur

III. D. 2| Amarrage sur la structure de HpFur

La structure de la protéine HpFur a été publiée en 2009 par notre laboratoire en collaboration avec l'EMBL (Dian *et al.*, 2009). Nous disposons de résultats expérimentaux concernant le test d'activité des peptides linéaires. Nous savons que le peptide linéaire pF1 n'inhibe pas la protéine *in vitro*. Nous avons commencé l'amarrage avec la séquence pF1(1-9).. Comme dans le cas précédent, plusieurs optimisations ont été réalisées en changeant les paramètres de taille de la population et nombre d'évaluation pour optimiser l'amarrage. Malgré toutes ces tentatives, on n'obtient pas une affinité très forte entre le peptide et la protéine. La meilleure conformation obtenue a une énergie libre de liaison de -6,5 kcal/mol et une constante d'inhibition Ki de 18 µM. Ce résultat est en accord avec les résultats du test d'activité du peptide pF1 sur la protéine HpFur. Les valeurs de l'énergie libre de liaison et du Ki suggèrent qu'il y a pas d'interactions fortes entre pF1 et HpFur pouvant inhiber l'activité de la protéine.

En partant de la conformation du peptide pF1(1-9), celle du peptide pF1(1-13) est construite et l'amarrage est réalisé en utilisant le protocole AADM. La Figure III.39 ci-dessous montre une vue d'ensemble des interactions du peptide pF1 avec HpFur.

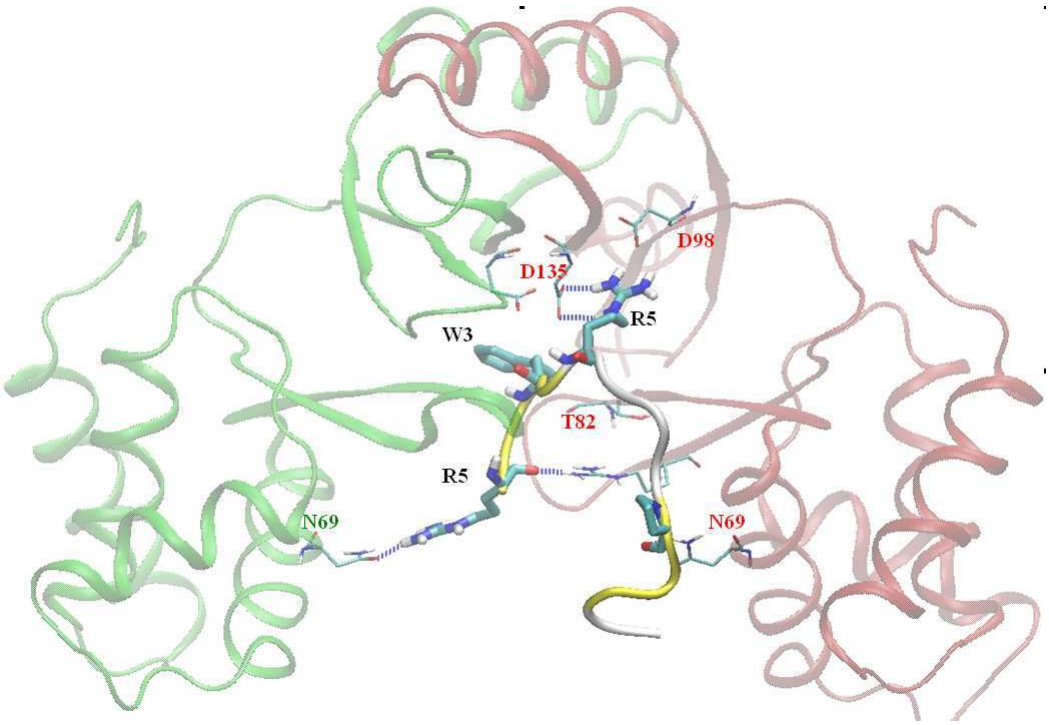


Figure III.39: Amarrage de pF1 sur la structure de HpFur. Représentation de la protéine, en vert le monomère A et en rouge le monomère B. Le peptide est représenté en jaune avec ses chaines latérales.

L'énergie libre de liaison du complexe est de -8 kcal/mol avec une constante d'inhibition de 2 µM (micro molaire). Le peptide pF1 interagit toujours avec les deux monomères de la protéine, mais ici l'écart entre les deux brins β semble trop grand pour stabilisant le peptide. L'écart entre les carbones α des Valines 79 de HpFur (équivalent de la R70 conservée chez les autres protéines) est de 19,8 Å, ici. En comparaison, cette distance R70-R70 entre monomères chez PaFur (meilleure affinité théorique de pF1) est de 17,8 Å. Ainsi les résidus protéiques qui formaient des liaisons hydrogènes stabilisant le feuillet β du peptide pF1 sont éloignés du peptide dans HpFur. Cela pourrait expliquer la baisse d'affinité de pF1 pour HpFur et le manque d'effet du peptide sur la protéine. Les interactions du peptide avec HpFur sont résumées dans le
Tableau III.18, qui montre les acides aminés de pF1 liant fortement les résidus de la protéine (énergie d'interaction < -8 kcal/mol).

Résidus pF1	E_{inter} (kcal/mol)
R1	-9
W3	-8,2
R5	-12
D13	-9

Tableau III.18: Énergies d'interaction ≥ -8 kcal/mol entre les acides aminés de pF1 et la protéine HpFur.

Ici, seulement quatre résidus : R1, W3, R5 et D13 interagissent fortement avec la protéine HpFur. Seule, le résidu R5 du peptide lie la protéine avec une énergie d'interaction ≥ à -10 kcal/mol. L'analyse des liaisons hydrogènes par CHARMM après une minimisation du complexe montre une fois de plus un faible nombre d'interactions avec pF1 et HpFur. Les atomes de pF1 qui établissent des liaisons hydrogènes avec HpFur sont présentés dans le
Tableau III.19. .

Peptide		E_{inter} (kcal/mol)	Protéine	
Résidus	Atomes		Résidus	Atomes
R1	HH12	-7,5	N69 (A)	OD1
W3	O	-7,5	K137 (A)	HZ1
R1	O	-6,8	R87 (B)	HH21
W3	HE1	-2,1	T82 (B)	O
R5	HE	-2,0	D135 (B)	OD2
	HH11	-8,3	D135 (B)	OD1
	HH12	-8,2	D98 (B)	OD1
	HH22	-7,8	D98 (B)	OD1
P9	O	-3,3	N69 (B)	HD22

Tableau III.19: Énergies d'interaction entre atomes de pF et PaFur formant des liaisons hydrogènes. Les résidus et atomes de la protéine sont colorés en vert et en rouge respectivement pour le monomère A et B. Les atomes HH, HZ, OD et OG sont des notations des types d'atomes des chaines latérales dans CHARMM

Dans cet amarrage, le peptide pF1 interagit avec HpFur principalement en ne formant que quelques liaisons hydrogènes avec la protéine. Les résidus impliqués sont R1, W3 et R5 principalement pour pF1 et les résidus N69, K137, D98 et 135 pour HpFur. Ces résidus de la protéine HpFur sont les équivalents des résidus N60, Y128, D89 et S126 chez EcFur, respectivement. Théoriquement, nous avions vu que les résidus Y128, S126 et 130 interagissaient fortement avec le résidu R5. Ces résidus de la protéine sont remplacés dans

HpFur par les acides aminés K137, D135 et F139 (numérotation HpFur). L'absence d'interaction de pF1 avec ces résidus pourrait aussi jouer dans la perte d'affinité du peptide pour HpFur. Pour valider ces hypothèses, il faudrait effectuer des mutations sur la protéine et tester l'activité du peptide ou bien suggérer des modifications pour augmenter l'interaction.

III. D. 3| Amarrage sur la structure de VcFur

Nous ne disposons pas d'informations expérimentales sur cette protéine concernant une éventuelle interaction avec pF1 pour le moment. Toutefois, nous savons d'après les données de double hybride que l'aptamère F1 n'interagit pas avec VsFur qui a un pourcentage de similarité de 98% avec VcFur. On peut s'attendre à ce que ces deux protéines se comportent de la même manière. Partant de cette analyse, l'amarrage de pF1 sur la structure de VcFur a été réalisé en utilisant le mode opératoire identique aux deux derniers cas. La conformation du complexe après la dynamique est présentée ci-dessous dans la Figure III.40.

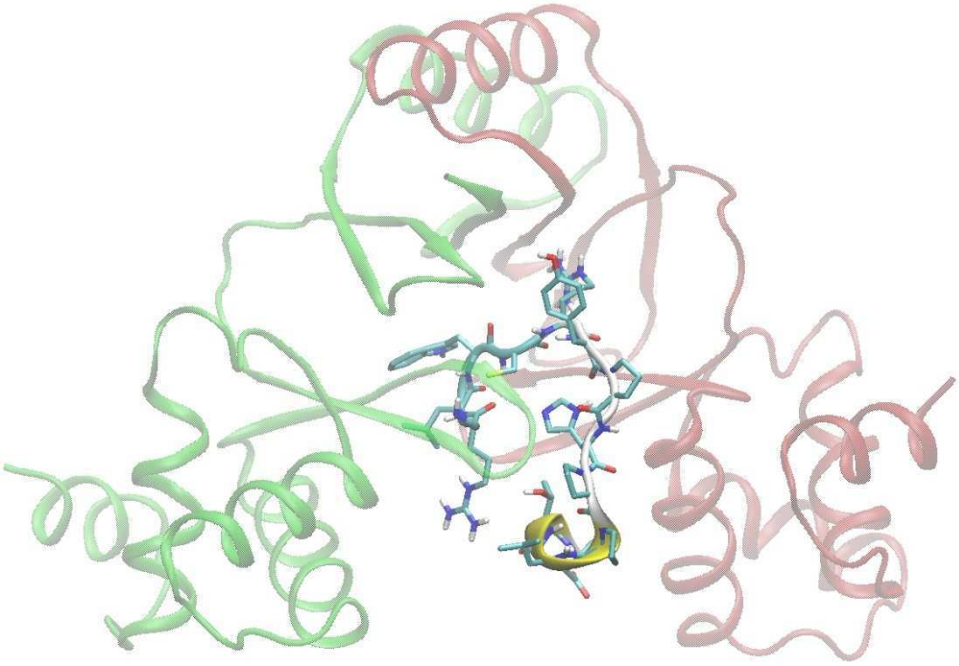


Figure III.40: Conformations du peptide pF1 et de VcFur après la dynamique moléculaire

Le peptide pF1 est initialement entre les deux sous-unités de la protéine comme dans les autres Fur, mais, après le protocole AADM, le peptide s'est déplacé. L'énergie libre de liaison de la meilleure conformation du complexe pF1-VcFur est seulement de -2,6 kcal/mole (Figure III.41), ce qui est très faible pour des complexes de cette taille et comparativement aux amarrages précédents.

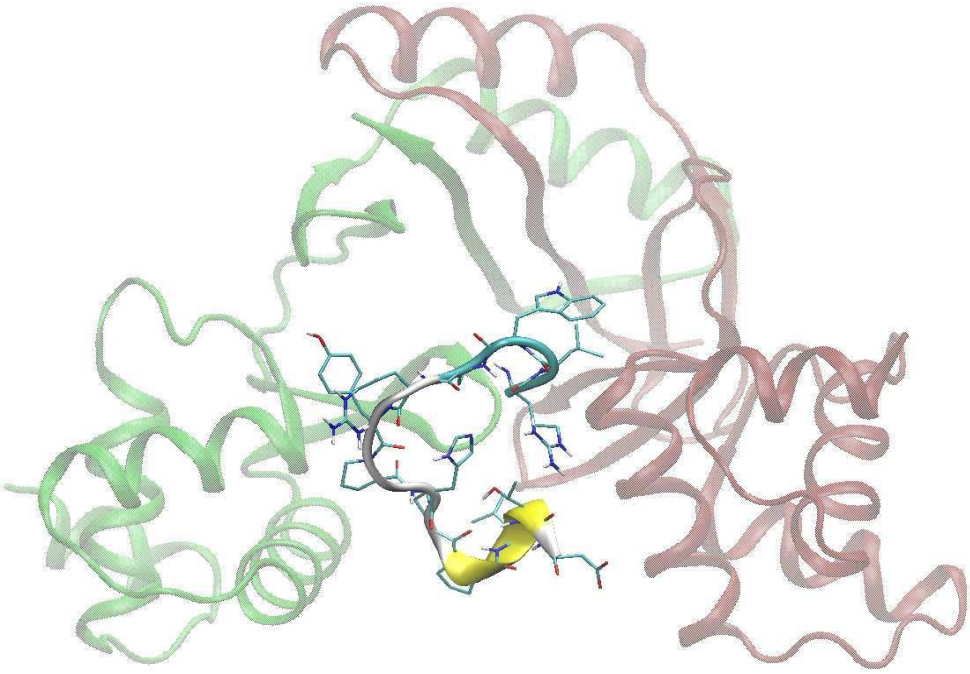


Figure III.41: Amarrage de pF1 sur la structure de VcFur

Même si nous n'avons pas de résultats expérimentaux sur VcFur, par analogie avec VsFur, nous pouvons prédire que ces deux protéines se comportent de la même façon. Nous avons testé cette protéine car sa structure était disponible et nous avons aussi construit le modèle d'EcFur avec cette structure. Le fait que pF1 interagisse favorablement en théorie avec le modèle EcFur←Vc et non avec la structure de VcFur est assez surprenant, a priori, et en même temps, cela prouve qu'un modèle par homologie ne reproduit pas forcément toutes les propriétés de la structure dont il est issu. On pense que cela peut venir, bien sûr, de petites différences dans la séquence des deux protéines ou encore de l'optimisation de l'amarrage mais aussi de la distance entre les deux domaines Nter de la protéine. Nous en discuterons dans le paragraphe suivant. Quoi qu'il en soit, on attend avec impatience des résultats expérimentaux sur cette protéine. Les amarrages prédisent une très faible affinité théorique.

III. D. 4| Conclusion sur les amarrages de pF1 sur les protéines Fur de divers pathogènes

Dans les amarrages des peptides pF1, pF2 et pF3 sur la protéine EcFur, nous avons prédit des conformations du complexe pF1-Fur relativement cohérentes avec les résultats expérimentaux, particulièrement dans le cas des amarrages de pF1. Nous avons donc réalisé l'amarrage de pF1 sur les protéines PaFur, HpFur et VcFur dont les structures sont

disponibles. Ces amarrages permettront d'étudier la spécificité de pF1 sur les protéines de la famille Fur.

Partant de la conformation du peptide amarré sur les modèles d'étude EcFur←Pa et EcFur←Vc, nous avons réalisé plusieurs processus d'optimisation d'amarrage de pF1 sur les protéines PaFur, HpFur et VcFur, en nous focalisant sur la grille adaptée FurEcG3, décrite plus haut. Ces amarrages montrent que le peptide pF1 interagit très favorablement avec la structure de PaFur, un peu moins avec la structure HpFur et très peu avec VcFur tandis que les résultats expérimentaux montrent une interaction exclusive de ce peptide avec EcFur. Excepté HpFur, ce peptide n'a pas été testé sur les autres protéines expérimentalement, mais nous savons que l'aptamère F1 interagit exclusivement avec EcFur et pas avec VsFur PaFur et HpFur. Nous savons que le peptide pF1 n'inhibe pas l'activité de HpFur *in vitro vitro* dans les conditions de notre test. Dans les prédictions d'amarrage de pF1 sur cette protéine, l'énergie libre de liaison est -6,5 kcal/mol ce qui reste relativement faible par rapport aux affinités de ce même peptide sur PaFur et les modèles d'EcFur, respectivement de -21,3 kcal/mol et - 20,5 kcal/mol. Cette faible valeur d'énergie libre de liaison pourrait expliquer l'absence d'effet du peptide sur la protéine HpFur. De même, par extrapolation, on peut s'attendre à ce que le peptide pF1 ne soit pas actif sur les protéines VcFur et VsFur d'après ces prédictions.

Les amarrages prédisent aussi une forte affinité de pF1 pour PaFur et le modèle EcFur←Pa, alors que le double hybride ne visualise pas d'interaction de PaFur avec F1. Il se peut que l'aptamère F1 se comporte différemment du peptide pF1 qui a plus de liberté pour pouvoir adopter une conformation compatible avec l'amarrage proposé.

Globalement, les acides aminés de pF1 pourraient se décomposer en trois groupes :

- Les résidus Nter, R1 à W3, qui constituent une petite boucle et interagissent soit avec le monomère A, soit avec le monomère B.
- Les résidus de la partie centrale du peptide ; résidus C4 à H8 dont une partie forme un feuillet β avec les brins β2 de la protéine (Figure III.42).
- Les résidus de la partie Cter du peptide, pouvant être une hélice ou une boucle, interagissant aussi selon les protéines avec les domaines Nter du monomère A ou du monomère B.

La séquence pF1(3-9) étant active sur EcFur, focalisons notre attention sur les résidus **C4** à **H8** de pF1(Figure III.42).

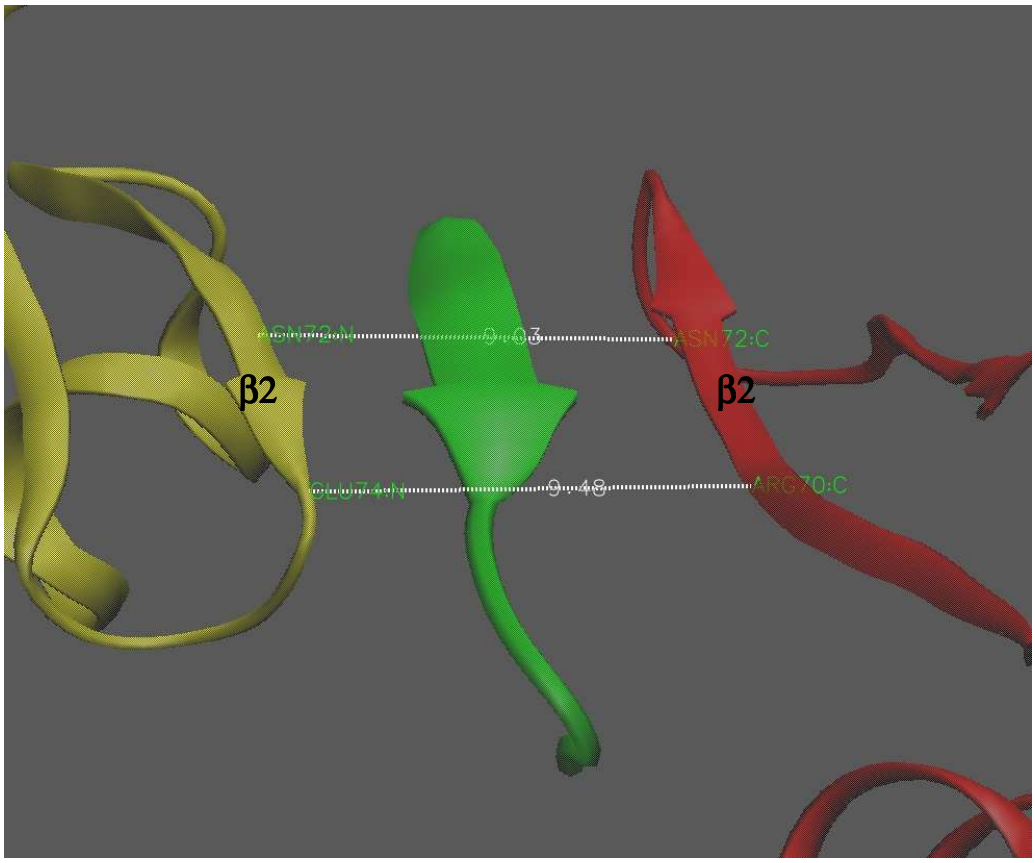


Figure III.42: Écart entre les brins β2 des monomères A et B et la partie centrale de pF1 (C4 à H8) dans l'amarrage pF1-EcFur←Vc

La capacité du peptide pF1 à s'insérer entre les brins β2 de Fur (Figure III.42) contribue pour beaucoup à l'interaction de pF1 avec la protéine et est probablement à l'origine de la forte affinité théorique observée. Il faut pour cela une distance entre les deux brins β de Fur, favorable à l'insertion du peptide. D'après le résumé du Tableau III.20, ci-dessous, il se pourrait que l'écart entre les brins β de la protéine justifie l'interaction défavorable de VcFur et de HpFur. Le constat est que pF1 interagit mieux avec la protéine lorsque la distance entre les deux brins β2 est comprise entre 9 et 10,5 Å. Lorsque cette distance est inférieure à 9Å ou supérieure à 12,5Å, l'interaction du peptide n'est plus favorable.

	ΔG (kcal/mol)	E_{inter} (kcal/mol)	Aβ2-Bβ2 (Å)
PaFur	-21,3	-126,3	9,5 à 10,5
EcFur←Pa	-20,5	-130,5	9,5 à 10
EcFur←Vc	-18	-151,8	9 à 9,5

HpFur	**-6,5**	**-98**	**12,5 à 14,3**
VcFur	**-2,6**	**-85,2**	**8,3 à 8,8**

Tableau III.20: Résumé des affinités de pF1 pour les protéines Fur. Le ΔG est l'énergie libre de liaison du complexe, calculée avec AutoDock4.2, l'E_{inter} est l'énergie d'interaction calculée par CHARMM. Aβ2-Bβ2 est l'écart entre les brins β2 (résidus 69 à 73) des monomères A et B de Fur mesuré avec VMD.

Une analyse plus poussée sur les interactions a été effectuée avec CHARMM pour ne pas valider les résultats qu'avec les énergies libres de liaison d'AutoDock4.2. Pour cela, protéines et peptides ont été préparés à partir des fichiers de sortie d'AutoDock de type *.pdbqt* (voir annexe) pour être transformés en fichiers de type *.pdb* avec des noms d'atomes compatibles avec le champ de force « carbones étendus » : *param19* de CHARMM. Des contraintes harmoniques de position ont été appliquées sur les atomes du squelette peptidique du dimère de Fur (constante de force 2 kcal/mol/Å^2) ; le peptide a été laissé libre. Des dynamiques de Langevin de 4 ns (coefficient de friction de 5 ps^{-1} sur les atomes lourds) ont ensuite été lancées en présence de solvant implicite (potentiel EEF1) sur le peptide seul, la protéine seule et le complexe peptide+protéine. La température était fixée à 300 K.

Les résultats des énergies d'interaction moyennes sur la dynamique sont reportés dans le

Tableau **III.21**, suivant.

	E(AB)d dimère	E(Pep)p peptide	E(Pep-A+Pep-B)c complexe	E(Pep)c	E(AB)c
EcFur	-134(8)	16(10)	-114(6)	52 (9)	-141 (7)
PaFur	-147(7)	9(10)	-80(6)	32 (13)	-148 (7)
VcFur	-155(7)	10(10)	-64(7)	29 (9)	-157 (7)
HpFur	-139(7)	9 (10)	-102(7)	42 (9)	-138 (7)

Tableau III.21: Énergies d'interactions moyennes (écarts types entre parenthèses) entre peptide et dimères de Fur calculées à partir de la dynamique de 4 ns du dimère seul, du peptide seul et du complexe. E(AB) est l'énergie d'interaction entre les 2 monomères de Fur. Les autres énergies d'interaction sont : entre peptide et peptide : E(Pep) ; entre peptide et monomère A : E(Pep-A) ; et entre peptide et monomère B : E(Pep-B).

Les résultats obtenus lors de cette analyse confirment l'interaction très favorable entre pF1 et EcFur et montrent également une bonne interaction entre pF1 et HpFur. Par contre les interactions entre pF1 et VcFur sont beaucoup moins stabilisantes, ce qui était déjà le cas après le calcul d'énergie libre de liaison avec AutoDock. Les énergies d'interaction entre les deux monomères de Fur dans les simulations sans peptide (d) varient exactement dans le sens opposé des interactions pF1-dimère dans le complexe avec une meilleure interaction entre les

deux monomères de VcFur. Ceci rejoint notre analyse précédente de la distance entre les brins β2 des deux monomères (voir tableau 2): les deux monomères de VcFur, proches l'un de l'autre, ont une bonne énergie d'interaction mais ne laissent pas suffisamment de place à l'insertion du peptide. Le peptide inséré dans EcFur semble être dans une conformation plus contrainte : en effet, l'énergie de « self-interaction » E(pep) est sensiblement supérieure à celle obtenue avec les autres protéines Fur. Même simulé seul, E(pep)p, pF1 semble incapable de retrouver une meilleure conformation dans le temps de la dynamique. Cette observation est validée par les tests d'activité sur le peptide pF1 cyclique, donc très contraint, qui est un bon inhibiteur d'EcFur.

Il serait intéressant de confronter les prédictions d'amarrage aux résultats de « peptidisation » en double hybride, les tests nucléases avec les peptides linéaires et le BIAcore. Nous verrons grâce à l'étude de la « peptidisation » en test double hybride qui est en cours avec Sophie Mathieu. Ce test avec le gène *luc* de la Luciférase comme « rapporteur » est en effet beaucoup plus sensible et quantitatif. Mais il faudrait estimer si l'amarrage de pF1 sur PaFur (Figure III.37 et Figure III.38) est compatible avec F1. Cette étude mérite d'être réalisée pour connaitre la spécificité de pF1 pour EcFur.

III. E| Le criblage virtuel de petites molécules sur Fur

Dans le but d'obtenir de petites molécules ayant des propriétés thérapeutiques, nous avons réalisé un criblage virtuel avec des composés provenant de la base de données ZINC (Irwin *et al.* 2005). ZINC est une base de données virtuelle de petites molécules commerciales, gratuitement disponibles pour le criblage virtuel. Cette base qui comprenait un peu plus de 700 000 composés en 2004 contient aujourd'hui plus de 13 millions de molécules, prêtes pour le « docking » (http://zinc.docking.org/). Le serveur Web permet de rechercher et de présélectionner les molécules en fonction de plusieurs critères de choix. Un des avantages de cette base est qu'elle dispose de petites molécules sous un format directement utilisable sans une préparation ultérieure par les programmes de « docking » comme Glide et Gold que nous avons utilisés.

Dans ce travail, nous avons réalisé deux criblages virtuels, afin de sélectionner quelques molécules interagissant favorablement en théorie avec la protéine Fur. Ces molécules constitueront la troisième génération de molécules anti-Fur, potentielles « leads » dans la voie de développement de nouveaux antibiotiques.
Nous avons utilisé le logiciel d'amarrage Glide de Schrödinger (http://www.schrodinger.com/), un programme sous license payante pour la sélection des petites molécules qui interagiront favorablement avec la protéine Fur. Le programme Glide et son interface de visualisation MAESTRO fournissent un ensemble de fonctionnalités permettant la préparation rapide et efficace des molécules avant criblage. Le protocole de criblage virtuel comprend trois étapes :
1) la préparation de la protéine, 2) la constitution de la librairie à cribler et 3) la sélection de petites molécules.

III. E. 1| Préparation de la protéine

Nous avons utilisé le modèle d'EcFur construit à partir de la structure de PaFur pour les criblages qui vont suivre. La protéine doit être soigneusement préparée avant le criblage virtuel à haut débit. La conformation du meilleur amarrage de pF1(3-9), séquence minimale active de pF1 (Figure III.43) a été choisie comme ligand par défaut pour définir l'espace à explorer. Cela permet de limiter la taille de la zone à explorer, donc le temps de calculs.

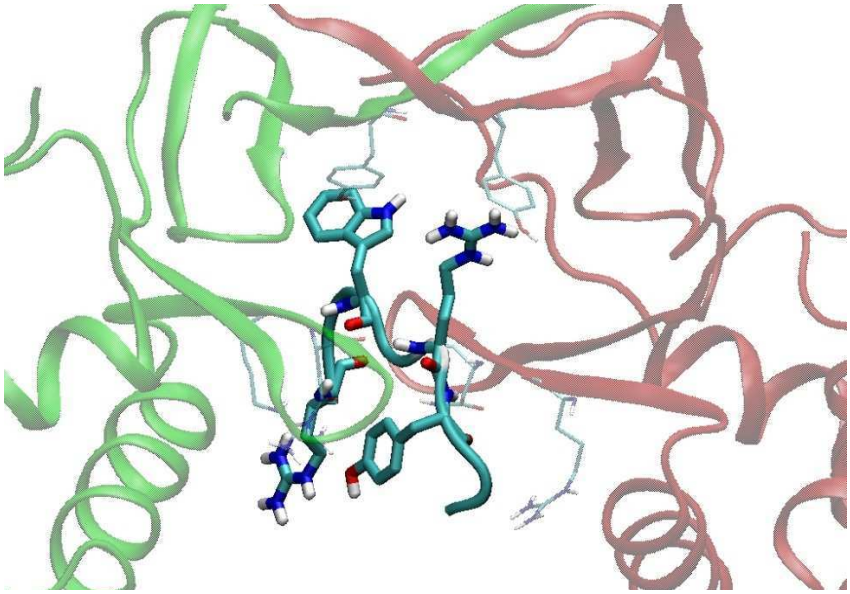


Figure III.43: Amarrage de pF1(3-9) sur EcFur←Vc

La zone à explorer a été définie pour ressembler à la grille FurecG1 (Figure III.44), une boite de 40 Å de côté centrée sur l'atome C_α du résidu C4 de pF1 amarré sur Fur.

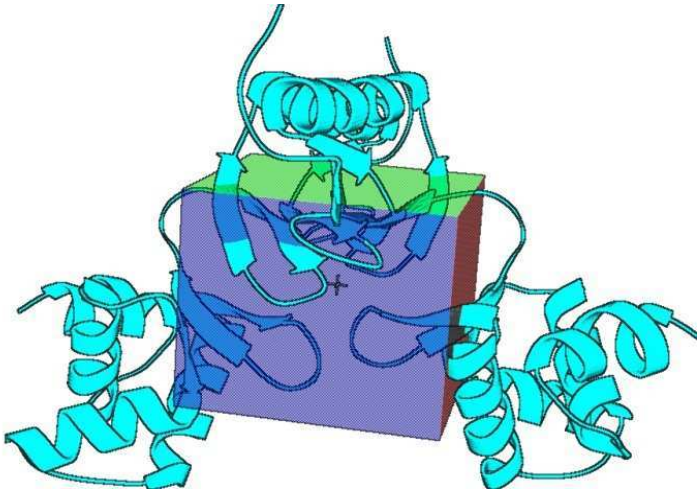


Figure III.44: Zone à explorer pour le criblage virtuel. La croix centrale situe l'atome C_α du résidu C4 de pF1 amarré.

Cette zone à explorer couvre tous les sites d'interaction de Fur précédemment identifiés lors de l'amarrage du peptide pF1. Nous recherchons donc des petites molécules pouvant mimer l'interaction de pF1 avec la protéine Fur. L'état de protonation des résidus ; arginine, glutamates, histidines, lysines et aspartates de la protéine, inclus dans l'espace à explorer, est ajusté avec le programme (confère annexe). Les groupements hydroxyles et thiols de tous les résidus protéiques concernés sont orientés vers le ligand pour optimiser la

formation de liaisons hydrogènes. La protéine est finalement optimisée par minimisation par le programme Glide.

III. E. 2| Préparation de la librairie

Nous avons sélectionné les petites molécules dans le but de les tester expérimentalement. Dans ce contexte, nous avons choisi de limiter notre recherche de petites molécules aux produits disponibles dans un délai de deux semaines chez les commerciaux Sigma Aldrich et ChemBrigde mais aussi des produits de l'institut National du Cancer (NCI). Notre but étant d'obtenir des molécules déjà approuvées par la FDA, nous avons choisi comme critère la célèbre règle des cinq de Lipinski (Lipinski *et al,* 1997), qui est la suivante :

- Poids moléculaire ≤ 500 Dalton
- Nombre de donneurs d'hydrogènes ≤ 5
- Nombre d'accepteurs d'hydrogènes ≤ 10
- Coefficient de partition Octanol-Eau log P ≤ 5
- Nombre d'angles de torsions ≤ 5

Les molécules respectant cette règle sont des composés pouvant être absorbés, distribués, métabolisés et excrétés (ADME) dans le corps humain.

Dans un premier temps, lors d'une mission de trois semaines à Tromso (Norvège), nous avons réalisé en collaboration avec le docteur Rafi Ahmed (NorStruct) de l'université de Tromso (UIT) un premier criblage avec les logiciels Glide et Gold en utilisant 2435 molécules. Ces molécules ont été choisies dans la base ZINC pour leur ressemblance structurale, soit aux chaines latérales d'une arginine (16 molécules), d'une tyrosine (42 molécules) ou d'un tryptophane (408 molécules), soit aux squelettes carbonés (sans les chaines latérales) de deux (769 molécules) ou trois résidus (1200 molécules). Nous espérions trouver de petites molécules mimant les interactions des chaines latérales de la tyrosine 6, de l'arginine 5, du tryptophane 3 et du squelette carboné des résidus W3 à R5 de pF1, formant généralement un feuillet β avec la protéine d'après les modèles d'interaction que nous venons de voir.

Dans un second temps, en utilisant comme moyen de calcul, le « cluster » Gabriel du CEA de Grenoble, un criblage a été réalisé avec 8000 petites molécules de la base ZINC

choisies avec comme critère, cette fois, les composés disponibles uniquement chez Sigma-Aldrich et la règle de Lipinski. Avec les différents conformères/rotamères de certaines molécules que génère Glide après la sélection de la librairie, nous obtenons plus de 13000 molécules pour le criblage.

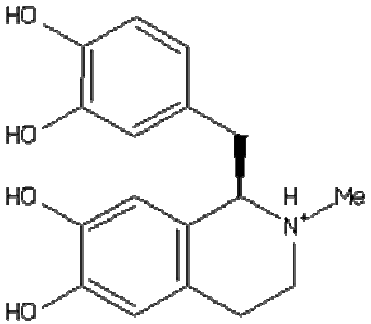
III. E. 3| Sélection des petites molécules

Pour simplifier, les résultats de sélection des petites molécules obtenus lors des deux criblages ont été regroupés dans le Tableau III.22 ci-dessous. Le tableau présente la liste des composés sélectionnés en fonction du score donné par l'évaluation XP de Glide. Les valeurs d'évaluation des programmes AutoDock4.2 et Gold (Fitness pour le criblage réalisé à TROMSO) sont également données.

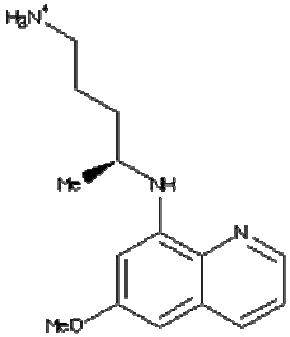
Molécules	code ZINC	MW	xLogP	NDH	NAH	Charge	AT	XP Glide	énergie Glide	Fitness	ΔG	N° criblage
1	ZINC00056408	302.35	1.76	5	5	1	2	-11,08	-42,6		-12,5	2
2	ZINC01530862	260.36	2.10	4	4	1	6	-10,68	-51,64		-10,1	2
3	ZINC00056410	302.35	1.76	5	5	1	2	-10,62	-42,27		-11,4	2
4	ZINC09211296	226.21	-1.96	4	7	1	1	-10,7	-39,65		-10,3	2
5	ZINC03860908	184.2	-0.17	5	6	2	4	-10,37	-51,71		-7,8	2
6	ZINC02324120	366.42	4.04	1	6	0	7	-8,7	-40,5	38	-6,8	1
7	ZINC02293058	336.39	4.00	1	5	0	6	-8,5	-39,62	28	-7,5	1
8	ZINC02805535	475.60	1.98	1	8	-1	6	-9,1	-45,84	46	-8,2	1
9	ZINC02805587	461.57	1.56	1	8	-1	6	-9,5	-48,41	56	-8,3	1
10	ZINC02805589	461.57	1.56	1	8	-1	6	-9,2	-50,54	51	-7,7	1
11	ZINC01736286	294,45	1,9	2	4	0	4	-9,6	-52,4	54	-12	1
12	ZINC01563670	292.33	-1.51	3	6	0	6	-9	-38,5	24	-7,9	1
13	ZINC01708201	261.28	-1.54	5	6	0	5	-9,2	-37,81	49	-12,1	1
14	ZINC01569529	222.24	-1.70	4	5	0	5	-8,6	-38,3	43	-7,1	1

Tableau III.22: Listes des petites molécules sélectionnées et leurs propriétés. Le code ZINC est la référence de la molécule dans la banque ZINC. MW est le poids moléculaire en g/mol. Les colonnes NDH, NAH et AT représentent respectivement le nombre de donneurs/accepteurs d'hydrogènes et d'angles de rotation. La colonne XP est la fonction de « *scoring* » Glide utilisée pour sélectionner les petites molécules (en kcal/mol). Le Fitness est la fonction de « *scoring* » du programme Gold (n'a pas été utilisé lors du second criblage. Le ΔG est l'énergie libre de liaison AutoDock4.2 (en kcal/mol).

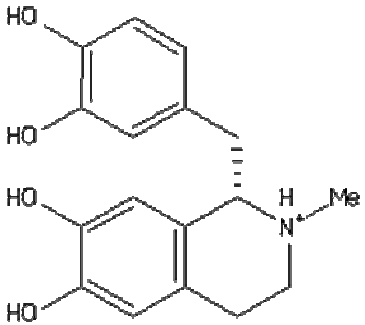
Le Tableau III.23 présente les structures des molécules sélectionnées avec leurs codes d'identification ZINC et le numéro des molécules du Tableau III.22.



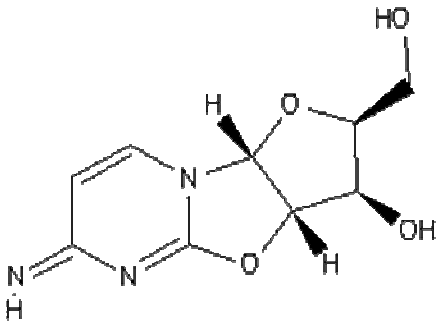
ZINC00056408: molécule 1



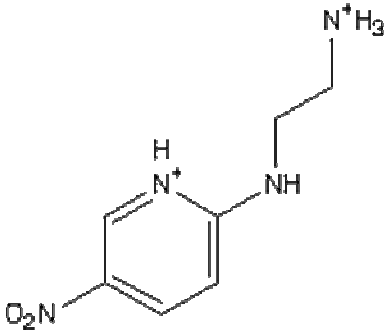
ZINC01530862: molécule 2



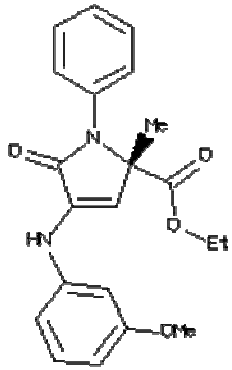
ZINC00056410: molécule 3



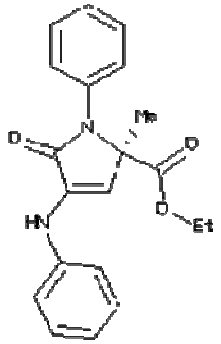
ZINC09211296: molécule 4



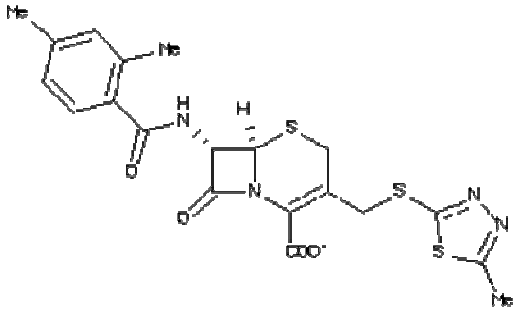
ZINC03860908: molécule 5



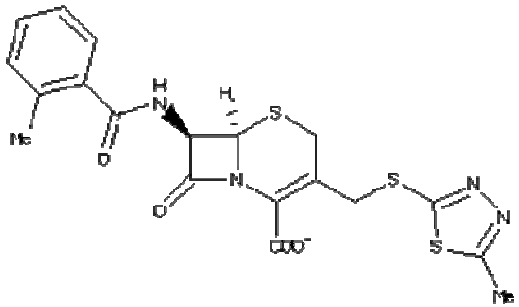
ZINC02324120: molécule 6



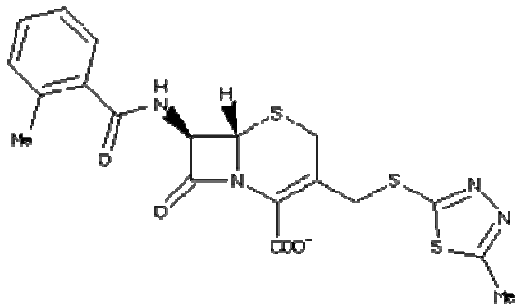
ZINC02293058: molécule 7



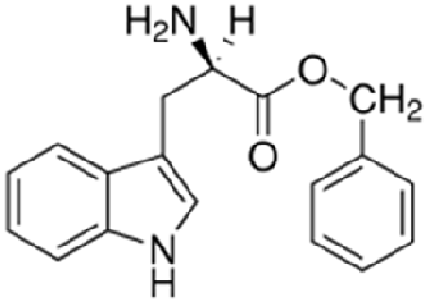
ZINC02805535: molécule 8



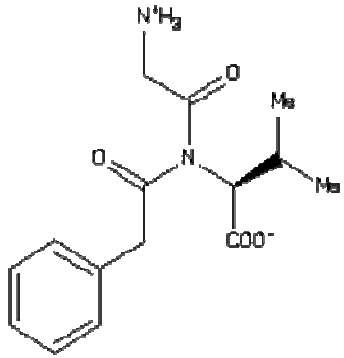
ZINC02805587: molécule 9



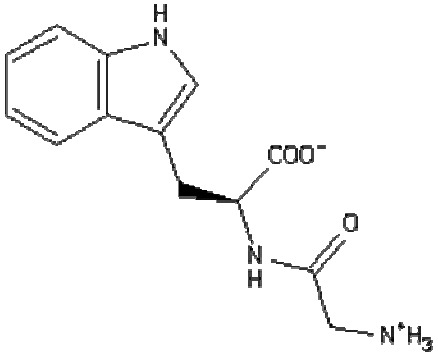
ZINC02805589: molécule 10



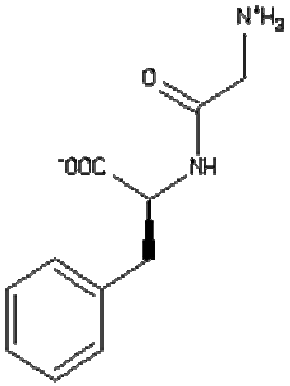
ZINC01736286: molécule 11



ZINC01563670: Molécule 12



ZINC01708201: Molécule 13



ZINC01569529: molécules 14

Tableau III.23: Les petites molécules sélectionnées par les criblages virtuels et leurs structures (ZINC) ou Sigma Aldrich (molécule 11).

III. E. 4| Analyse des criblages

Les molécules sélectionnées sont réparties en deux groupes (Tableau III.22):

- Le premier groupe présente les molécules 6 à 14 Tableau III.23) obtenues après criblage réalisé au CEA. Toutes les molécules ont une énergie de « *scoring* » Glide ≤ - 10 kcal/mole, amarrages très favorables.
- Le second groupe concerne les molécules sélectionnées lors du criblage réalisé en Norvège. Ces molécules ont été sélectionnées avec une énergie de « *scoring* » Glide ≤ - 8,5 kcal/mol.

III. E. 4. a| Criblage réalisé à TROMSO (N°1, Tableau III.22)

Les petites molécules ayant une structure proche du « squelette » de 2-3 résidus sont les plus représentées dans le second groupe. Les valeurs de leurs énergies d'interactions avec la protéine sont toutes favorables (énergie Tableau III.22 et Tableau III.23). On s'attendait à ce que ces groupements interagissent avec Fur comme pF1: c'est-à-dire forment un feuillet avec les deux brins β2 de Fur, mais, ces molécules sont toutes amarrées dans la zone située entre les domaines Nter et Cter inter-sous unité.

Dans cette première sélection, la molécule ayant obtenu la valeur énergétique la plus favorable est un composé proche d'un résidu tryptophane : la molécule 11 (Tableau III.22 et Tableau III.23) (code ZINC : ZINC01736286), l'ester du benzyl-tryptophane qui s'amarre sur la protéine avec une énergie d'interaction de -9,6 kcal/mol (Glide) et une fonction de fitness de 54 (Gold), favorables tous les deux. Cette molécule a été notée **W-CO-Bz** et testée expérimentalement (confère chapitre II). Un « docking » a été réalisé avec cette molécule en utilisant le programme AutoDock4.2 afin de comparer son affinité pour la protéine avec celle des peptides anti-Fur. La meilleure conformation du complexe de Fur avec cette molécule, donne une énergie libre de liaison **ΔG** = - 12 kcal/mol. Cette valeur d'affinité est très favorable, surtout pour une petite molécule. Les énergies d'interaction entre W-CO-Bz et Fur n'ont pas été calculées avec CHARMM faute de temps pour terminer sa paramétrisation (Voir chapitre V pour la méthode de paramétrisation de la molécule).

La Figure III.45 montre l'amarrage de cette molécule sur la protéine : EcFur←Pa. Le mode d'amarrage prédit, permet d'envisager l'interaction simultanée de 2 molécules de **W-CO-Bz** par dimère de Fur.

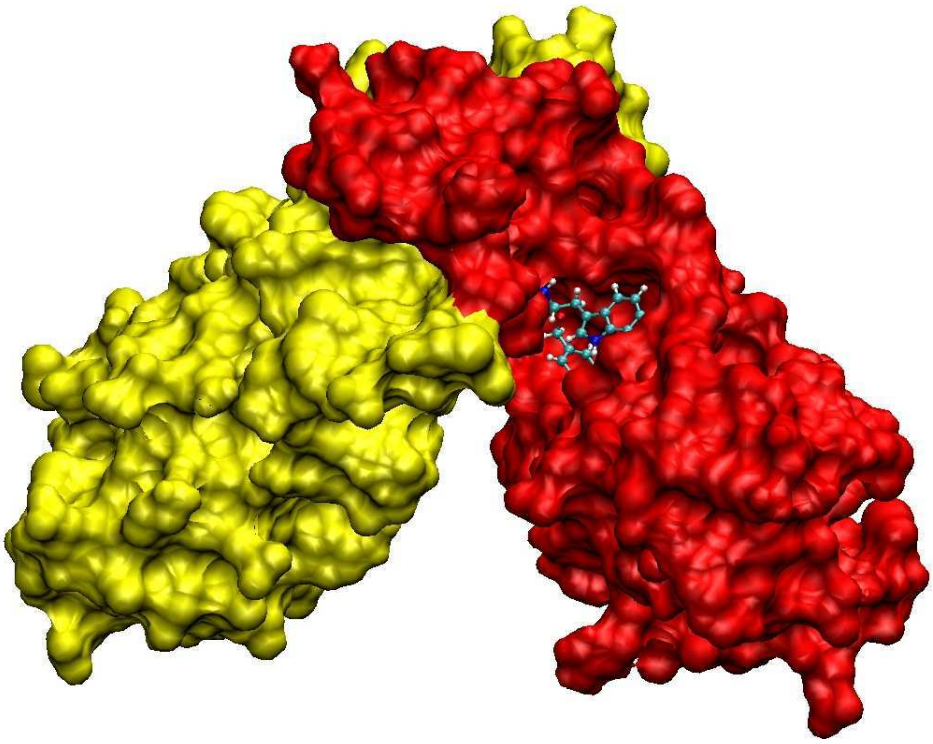


Figure III.45: Amarrage du W-CO-Bz sur la protéine EcFur←Pa

Une analyse détaillée des interactions montre que la molécule interagit à l'interface des domaines Nter et Cter d'une sous unité en établissant des liaisons hydrogènes avec les résidus T69, H125 et L127 d'EcFur (**Figure III.46**).

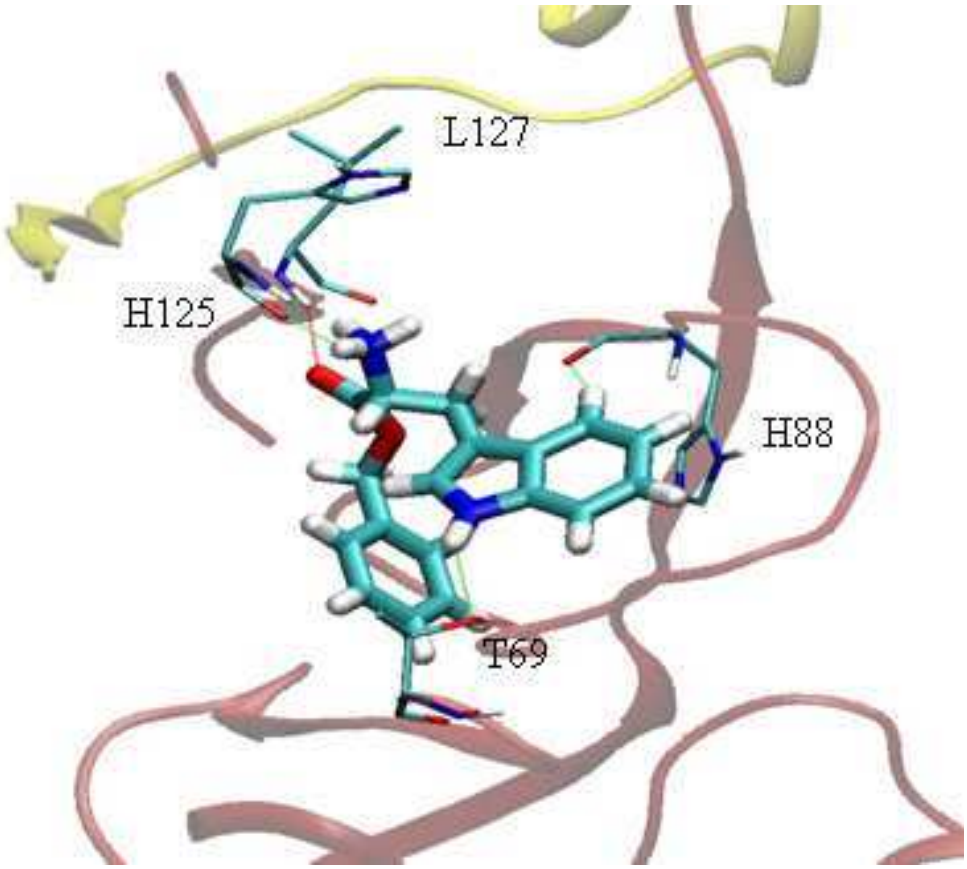


Figure III.46: Interactions du W-CO-Bz avec EcFur←Pa

Les interactions observées sur la **Figure III.46**, ont lieu dans la zone de Fur où l'on a observé les interactions de la chaine latérale du résidu R5 de pF1 avec EcFur, montrant une fois de plus que cette poche protéique, est accessible pour une interaction forte. Ces petites molécules peuvent bien mimer certaines interactions des peptides sans avoir la taille et lier la protéine très fortement. Mais le test expérimental sur cette molécule n'a pas été concluant. En effet, La molécule a inhibé la liaison de la protéine lors d'un test mais ces résultats n'ont pas pu être reproduits après trois autres tentatives. On remarque aussi que la petite molécule est relativement proche du site métallique « annexe » de Fur (confère chapitre I). Le fait que, dans les tests d'activité, la protéine soit métallée avant l'ajout de la molécule pourrait avoir un effet sur l'interaction de la molécule. Rappelons que dans cette étude théorique, nous avons pris soin d'orienter les chaines latérales des résidus ligands de ce site afin de mimer la protéine métallée. Cela ne devrait donc pas être la cause de l'effet non inhibiteur de la molécule. Pour en être sûr, il faut effectuer le test en mettant la molécule avant le métal et voir l'effet produit. De toute façon, il faudra optimiser le test d'activité ou utiliser d'autres tests plus sensibles pour valider l'effet de ces petites molécules.

III. E. 4. b| Criblage réalisé au CEA (N°2, Tableau III.22)

Dans le second criblage, la librairie de molécules présélectionnées n'était pas basée sur des similarités de structure avec le peptide pF1. Les molécules sélectionnées (molécule 1 à 5,Tableau III.23) ont une affinité très forte pour la protéine Fur énergie de « *scoring* » XP-Glide ≤ - 10 kcal/mol). Le meilleur amarrage de ce criblage est celui du ZINC00056408 (molécule 1, Tableau III.23ci-dessus), obtenu avec une énergie Glide de -11,1 kcal/mol et représenté sur la Figure III.47 ci-dessous.

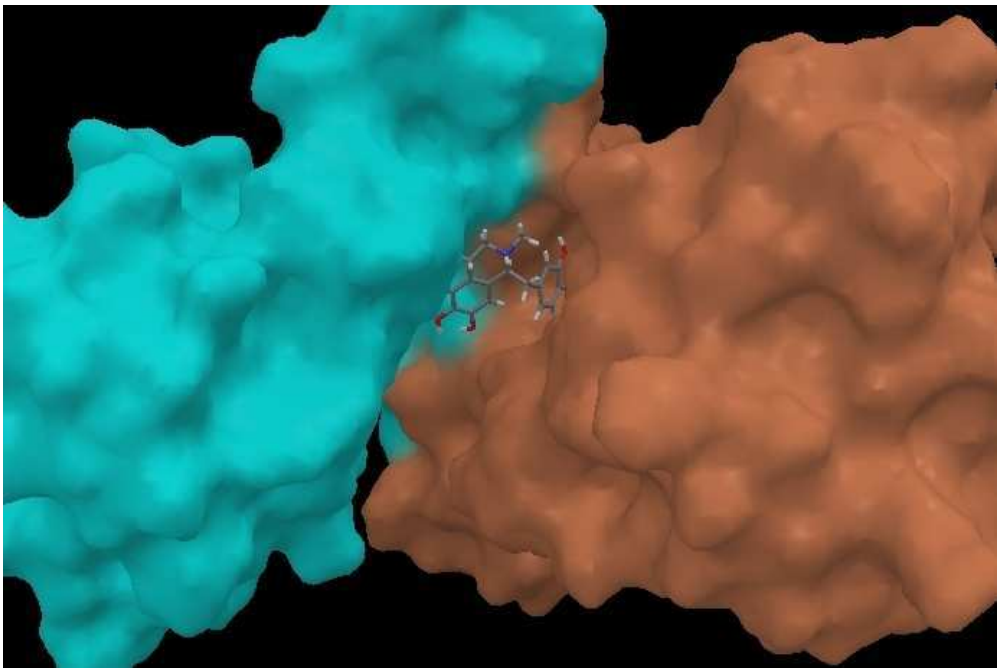


Figure III.47: Amarrage de la meilleure interaction de ZINC00056408 sur EcFur obtenue lors du second criblage avec Glide.

La molécule est amarrée dans une poche hydrophobe, sous le domaine de dimérisation d'EcFur. L'amarrage de cette molécule avec AutoDock4.2 donne une énergie libre de liaison ΔG de -12,5 kcal/mol, un peu plus favorable que le meilleur amarrage du premier criblage. Ici encore, il aurait fallu terminer la paramétrisation de cette molécule pour le champ de force de CHARMM (voir chapitre V) pour effectuer l'analyse des liaisons hydrogènes et des énergies d'interaction. Malheureusement, nous n'avions pas pu faire cela avant l'écriture de ce manuscrit. Ces travaux sont toutefois en cours de réalisation. Il serait intéressant de commander ces petites molécules et de les tester expérimentalement, car les affinités théoriques obtenues sont encore plus favorables que

celles du premier criblage. Avec un peu de chance, une ou plusieurs de ces molécules inhiberont la protéine, ouvrant ainsi la voie aux molécules anti-Fur de troisième génération.

III. E. 5| Conclusion sur le criblage

Après avoir identifié une zone sur la protéine EcFur qui pourrait être la zone d'inhibition du peptide pF1, nous avons recherché de petites molécules dans la base de données ZINC, en espérant que ces molécules miment les interactions de pF1 ou trouvent d'autres interactions plus favorables pour inhiber la protéine. Il faut savoir que l'obtention d'une petite molécule inhibitrice de Fur sera une grande avancée dans ce travail. Nous avons donc choisi des critères qui permettent d'obtenir de petites molécules ayant des propriétés thérapeutiques potentielles. Nous avons sélectionné 14 molécules interagissant favorablement, en théorie, avec la protéine Fur grâce au programme Glide. La comparaison avec les **ΔG** d'AutoDock montre aussi que ces molécules ont des affinités apparentes favorables. Une molécule a été testée, pour le moment, mais les résultats obtenus ne permettent pas de conclure sur un éventuel effet inhibiteur de la molécule sur la protéine.

Il sera intéressant de se procurer toutes ces molécules afin de les tester comme cela a été fait pour le W-CO-Bz. Mais une optimisation du test sera nécessaire. Il est également possible d'utiliser d'autres tests d'interaction comme le Biacore qui est déjà en cours concernant l'interaction des peptides anti-Fur. Ce test est sensé être plus sensible et on pourra espérer obtenir la constante de dissociation des molécules anti-Fur avec leur cible.

L'obtention d'une petite molécule inhibitrice de Fur ouvrira la voie aux molécules anti-Fur de troisième génération.

Bibliographie

Abed, N. (2005). La protéine Fur nouvelle cible antibactérienne? Approches par utilisation des aptamères peptidiques, Université Joseph Fourier.

Abed, N., M. Bickle, B. Mari, M. Schapira, R. Sanjuan-Espana, K. Robbe Sermesant, O. Moncorge, S. Mouradian-Garcia, P. Barbry, B. B. Rudkin, M. O. Fauvarque, I. Michaud-Soret and P. Colas (2007). "A comparative analysis of perturbations caused by a gene knock-out, a dominant negative allele, and a set of peptide aptamers." *Mol Cell Proteomics* **6**(12): 2110-2121.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." *Journal of molecular biology* **215**(3): 403-410.

An, Y. J., B. E. Ahn, A. R. Han, H. M. Kim, K. M. Chung, J. H. Shin, Y. B. Cho, J. H. Roe and S. S. Cha (2009). "Structural basis for the specialization of Nur, a nickel-specific Fur homolog, in metal sensing and DNA recognition." *Nucleic acids research* **37**(10): 3442-3451.

Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne (2000). "The Protein Data Bank." *Nucleic acids research* **28**(1): 235-242.

Brooijmans, N. and I. D. Kuntz (2003). "Molecular recognition and docking algorithms." *Annu Rev Biophys Biomol Struct* **32**: 335-373.

Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus (1983). "CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations." *J. Comp. Chem.* **4**: 187-217.

Brooks, B. R., C. L. B. III, A. D. M. Jr, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York and M. Karplus (2009). "CHARMM: The biomolecular simulation program." *Journal of Computational Chemistry* **30**(10): 1545-1614.

Brown, C. J., S. G. Dastidar, H. Y. See, D. W. Coomber, M. Ortiz-Lombardia, C. Verma and D. P. Lane (2010). "Rational design and biophysical characterization of thioredoxin-based aptamers: insights into peptide grafting." *J Mol Biol* **395**(4): 871-883.

Dian, C., S. Vitale, G. A. Leonard, C. Bahlawane, C. Fauquant, D. Leduc, C. Muller, H. de Reuse, I. Michaud-Soret and L. Terradot (2011). "The structure of the Helicobacter pylori ferric uptake regulator Fur reveals three functional metal binding sites." *Molecular microbiology* **79**(5): 1260-1275.

Ewing, T. J., S. Makino, A. G. Skillman and I. D. Kuntz (2001). "DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases." *Journal of computer-aided molecular design* **15**(5): 411-428.

Friesner, R. A., J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis and P. S. Shenkin (2004). "Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy." *Journal of medicinal chemistry* **47**(7): 1739-1749.

Friesner, R. A., R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin and D. T. Mainz (2006). "Extra precision glide: docking and scoring

incorporating a model of hydrophobic enclosure for protein-ligand complexes." *Journal of medicinal chemistry* **49**(21): 6177-6196.

Goodsell, D. S. and A. J. Olson (1990). "Automated docking of substrates to proteins by simulated annealing." *Proteins* **8**(3): 195-202.

Hetenyi, C. and D. van der Spoel (2006). "Blind docking of drug-sized compounds to proteins with up to a thousand residues." *FEBS letters* **580**(5): 1447-1450.

Hockney, R. W. and J. W. Eastwood (1981). "Computer simulation using particles." *McGraw-Hill Book Company*.

Huey, R., G. M. Morris, A. J. Olson and D. S. Goodsell (2007). "A semiempirical free energy force field with charge-based desolvation." *Journal of Computational Chemistry* **28**(6): 1145-1152.

Humphrey, W., A. Dalke and K. Schulten (1996). "VMD: visual molecular dynamics." *Journal of molecular graphics* **14**(1): 33-38, 27-38.

Iorga, B., D. Herlem, E. Barre and C. Guillou (2006). "Acetylcholine nicotinic receptors: finding the putative binding site of allosteric modulators using the "blind docking" approach." *Journal of molecular modeling* **12**(3): 366-372.

Irwin, J. J. and B. K. Shoichet (2005). "ZINC--a free database of commercially available compounds for virtual screening." *Journal of chemical information and modeling* **45**(1): 177-182.

Jackson, R. M., H. A. Gabb and M. J. Sternberg (1998). "Rapid refinement of protein interfaces incorporating solvation: application to the docking problem." *Journal of molecular biology* **276**(1): 265-285.

Jacquamet, L., D. A. Traore, J. L. Ferrer, O. Proux, D. Testemale, J. L. Hazemann, E. Nazarenko, A. El Ghazouani, C. Caux-Thang, V. Duarte and J. M. Latour (2009). "Structural characterization of the active form of PerR: insights into the metal-induced activation of PerR and Fur proteins for DNA binding." *Molecular microbiology* **73**(1): 20-31.

Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein (1983). "Comparison of simple potential functions for simulating liquid water." *J. Chem. Phys.* **79**(926).

Keseru, G. M. (2001). "A virtual high throughput screen for high affinity cytochrome P450cam substrates. Implications for in silico prediction of drug metabolism." *Journal of computer-aided molecular design* **15**(7): 649-657.

Kitchen, D. B., H. Decornez, J. R. Furr and J. Bajorath (2004). "Docking and scoring in virtual screening for drug discovery: methods and applications." *Nat Rev Drug Discov* **3**(11): 935-949.

Kuntz, I. D., J. M. Blaney, S. J. Oatley, R. Langridge and T. E. Ferrin (1982). "A geometric approach to macromolecule-ligand interactions." *Journal of molecular biology* **161**(2): 269-288.

Lazaridis, T. and M. Karplus (1999). "Effective energy function for proteins in solution." *Proteins* **35**(2): 133-152.

Levinthal, C., S. J. Wodak, P. Kahn and A. K. Dadivanian (1975). "Hemoglobin interaction in sickle cell fibers. I: Theoretical approaches to the molecular contacts." *Proceedings of the National Academy of Sciences of the United States of America* **72**(4): 1330-1334.

Lucarelli, D., S. Russo, E. Garman, A. Milano, W. Meyer-Klaucke and E. Pohl (2007). "Crystal structure and function of the zinc uptake regulator FurB from *Mycobacterium tuberculosis*." *J Biol Chem* **282**(13): 9914-9922.

Morris, G. M., R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson (2009). "AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility." *Journal of Computational Chemistry* **30**(16): 2785-2791.

Pecqueur, L., B. D'Autreaux, J. Dupuy, Y. Nicolet, L. Jacquamet, B. Brutscher, I. Michaud-Soret and B. Bersch (2006). "Structural changes of *Escherichia coli* ferric uptake regulator during metal-dependent dimerization and activation explored by NMR and X-ray crystallography." *J Biol Chem* **281**(30): 21286-21295.

Pecqueur, L., B. D'Autreaux, J. Dupuy, Y. Nicolet, L. Jacquamet, B. Brutscher, I. Michaud-Soret and B. Bersch (2006). "Structural changes of Escherichia coli ferric uptake regulator during metal-dependent dimerization and activation explored by NMR and X-ray crystallography." *The Journal of biological chemistry* **281**(30): 21286-21295.

Pohl, E., J. C. Haller, A. Mijovilovich, W. Meyer-Klaucke, E. Garman and M. L. Vasil (2003). "Architecture of a protein central to iron homeostasis: crystal structure and spectroscopic analysis of the ferric uptake regulator." *Mol Microbiol* **47**(4): 903-915.

Ryckaert, J. P., G. Ciccotti and B. H. C. (1977). "Numerical integration of cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes." *J. Comput Phys* **23**(327).

Salemme, F. R. (1976). "An hypothetical structure for an intermolecular electron transfer complex of cytochromes c and b5." *Journal of molecular biology* **102**(3): 563-568.

Sali, A. and T. Blundell (1993). "Comparative protein modelling by satisfaction of spatial restraints." *J. Mol. Biol.* **234**: 779-815.

Sheikh, M. A. and G. L. Taylor (2009). "Crystal structure of the Vibrio cholerae ferric uptake regulator (Fur) reveals insights into metal co-ordination." *Molecular microbiology* **72**(5): 1208-1220.

Simmons, K. J., I. Chopra and C. W. Fishwick (2010). "Structure-based discovery of antibacterial drugs." *Nat Rev Microbiol* **8**(7): 501-510.

Thompson, J. D., D. G. Higgins and T. J. Gibson (1994). "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice." *Nucleic acids research* **22**(22): 4673-4680.

Tiss, A., O. Barre, I. Michaud-Soret and E. Forest (2005). "Characterization of the DNA-binding site in the ferric uptake regulator protein from *Escherichia coli* by UV crosslinking and mass spectrometry." *FEBS Lett* **579**(25): 5454-5460.

Traore, D. A., A. El Ghazouani, S. Ilango, J. Dupuy, L. Jacquamet, J. L. Ferrer, C. Caux-Thang, V. Duarte and J. M. Latour (2006). "Crystal structure of the apo-PerR-Zn protein from *Bacillus subtilis*." *Mol Microbiol* **61**(5): 1211-1219.

Traore, D. A., A. El Ghazouani, S. Ilango, J. Dupuy, L. Jacquamet, J. L. Ferrer, C. Caux-Thang, V. Duarte and J. M. Latour (2006). "Crystal structure of the apo-PerR-Zn protein from Bacillus subtilis." *Molecular microbiology* **61**(5): 1211-1219.

Verdonk, M. L., J. C. Cole, M. J. Hartshorn, C. W. Murray and R. D. Taylor (2003). "Improved protein-ligand docking using GOLD." *Proteins* **52**(4): 609-623.

Verlet, L. (1967). "Computer "experiments" on classical fluids: I. Thermodynamical properties of Lennard-Jones molecules." *Phys Rev* **159**: 89-103.

Vitale, S. (2009). Inhibiteurs anti-Fur et caractérisation de Fur de Helicobacter pylori, Université J. Fourier.

Weiner, P. K. and P. A. Kollman (1981). "AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions." *Journal of Computational Chemistry*: 281-304.

Wodak, S. J. and J. Janin (1978). "Computer analysis of protein-protein interaction." *Journal of molecular biology* **124**(2): 323-342.

Chapitre IV :

Bilan et discussions

La résistance bactérienne est depuis longtemps un fléau majeur pour lequel il est nécessaire de trouver de nouveaux antibiotiques efficaces et dans le meilleur des cas, des molécules structuralement bien différentes de celles déjà disponibles sur le marché. La métalloprotéine Fur est un régulateur global chez *Escherichia coli* qui est absente chez les eucaryotes et omniprésente chez les autres bactéries Gram négatives. Des travaux antérieurs au laboratoire ont consisté à sélectionner des aptamères peptidiques anti-Fur afin d'étudier cette protéine en tant que cible antibactérienne agissant sur la virulence. Ce projet rentre donc dans la nouvelle approche d'antibiothérapie qui consiste à atténuer la virulence des bactéries plutôt que de les tuer. En effet, plusieurs études ont montré que la protéine Fur intervenait dans l'expression de facteurs de virulence, comme les transporteurs membranaires et la sécrétion de toxines.

La technologie des aptamères peptidiques a permis de sélectionner quatre molécules interagissant spécifiquement avec cette protéine. Ces aptamères nommés F1 à F4 inhibent l'activité de la protéine *in vivo* et interfèrent avec la virulence bactérienne. Ils constituent donc la première génération très intéressante d'inhibiteurs anti-Fur.

Les peptides de 13 acides aminés pF1 à pF4 correspondant à la partie active des aptamères F1 à F4 et considérés comme la seconde génération d'inhibiteurs anti-Fur, ont fait l'objet de cette thèse dans laquelle une double-approche théorique et expérimentale a été utilisée.

La combinaison des études théorique et expérimentale a permis de mieux caractériser les peptides et surtout leurs interactions avec leur cible et de déboucher vers l'obtention de petites molécules ayant des propriétés thérapeutiques potentielles.

De façon non surprenante pour des peptides de cette taille, l'analyse de ceux-ci par dichroïsme circulaire a montré que ces molécules ne possèdent pas de structuration en solution. Cela va dans le sens des analyses BLAST qui ont montré que ces molécules ne disposent d'aucune homologie dans la PDB. La modélisation moléculaire a confirmé que les peptides linéaires ne sont pas structurés. Par contre, lorsqu'on impose des contraintes aux peptides de façon à mimer la conformation du peptide dans l'aptamère on observe une structuration pour pF1, pF2 et le contrôle pF5 inactif en modélisation. Cette contrainte correspond peut-être au rôle de la plate-forme (la Thiorédoxine) dans la différence d'affinité observée entre les aptamères peptidiques F1 à F4 et les peptides pF1 à pF4. En effet, cela réduit les possibilités de flexibilité des peptides qui adoptent plus facilement une

conformation favorable pour interagir avec la protéine. Il a donc été intéressant de mimer la conformation du peptide dans l'aptamère pour les études d'amarrage.

Ces peptides possédant une à deux cystéines, leur état redox et leur capacité à former des ponts disulfures intra et/ou intermoléculaires ou à lier des métaux ont été un peu étudiés.

Des tests en spectrométrie de masse Electrospray ont montré que dans les peptides en solution, les cystéines peuvent s'oxyder et former des ponts disulfures. Cela a été observé pour les peptides pF1(1-13) et pF1(1-9) également en séquençage. Ces analyses présagent que toutes les séquences contenant au moins une cystéine peuvent s'oxyder et former des ponts disulfures en solution. De plus, l'analyse en spectroscopie UV-Visible a montré que pF1, pF2 et le peptide contrôle pF5 sont capables de lier le cobalt. Ces peptides pourraient interférer avec l'activation de la protéine, par exemple en formant un complexe protéine métal peptide avec le peptide partiellement lié au métal. Ce type de complexe n'a pas pu été considéré dans les prédictions théoriques. Cela peut biaiser la comparaison que nous faisons entre les résultats expérimentaux et théoriques. C'est pourquoi la prise en compte des propriétés électroniques des ions métalliques dans les études de « docking » sur des métalloprotéines fait partie de nos perspectives à court terme.

Les tests de protection à la coupure par une nucléase des différents peptides ont montré que pF1, pF2 et pF3 sont actifs sur la protéine dimérique et pF4 sur le monomère en empêchant sa dimérisation. Bien qu'indicatives et dépendantes des conditions du test d'activité (1 à 2 μM de Fur), les valeurs d'IC_{50} déterminées (confère chapitre II) nous permettent de penser que les constantes de dissociation entre les peptides et la protéine Fur sont comprises entre 1 μM et 100 μM, et donc supérieures à celles des aptamères peptidiques, estimées à l'ordre du nM.

Comme on l'a souligné précédemment, la contrainte imposée à la séquence variable par la plateforme jouerait un rôle important dans l'interaction du peptide avec la protéine. Pour étudier cette hypothèse, la distance entre les extrémités Nter et Cter de pF1 amarré sur la protéine EcFur←Pa a été mesurée afin de construire un peptide cyclique de pF1. Cette distance était de 15,4 (et 15,7) Å entre les carbones α de R1 et P9 (et D13). Etant donné que pF1(1-9) est actif et que les quatre résidus Cter et le résidu R1 de pF1 ne sont pas indispensables pour inhiber l'activité de la protéine, la séquence pF1(1-12) R1K a été synthétisée dans le but de cycliser le peptide (entre l'amine de la chaine latérale de K1 et le

COOH de T12). La Figure IV.1 montre un modèle optimisé de ce peptide. Les séquences linéaire et cyclique de ce peptide ont été synthétisées et testées expérimentalement. Nous avons montré que le peptide cyclique pourrait être plus actif que son homologue linéaire.

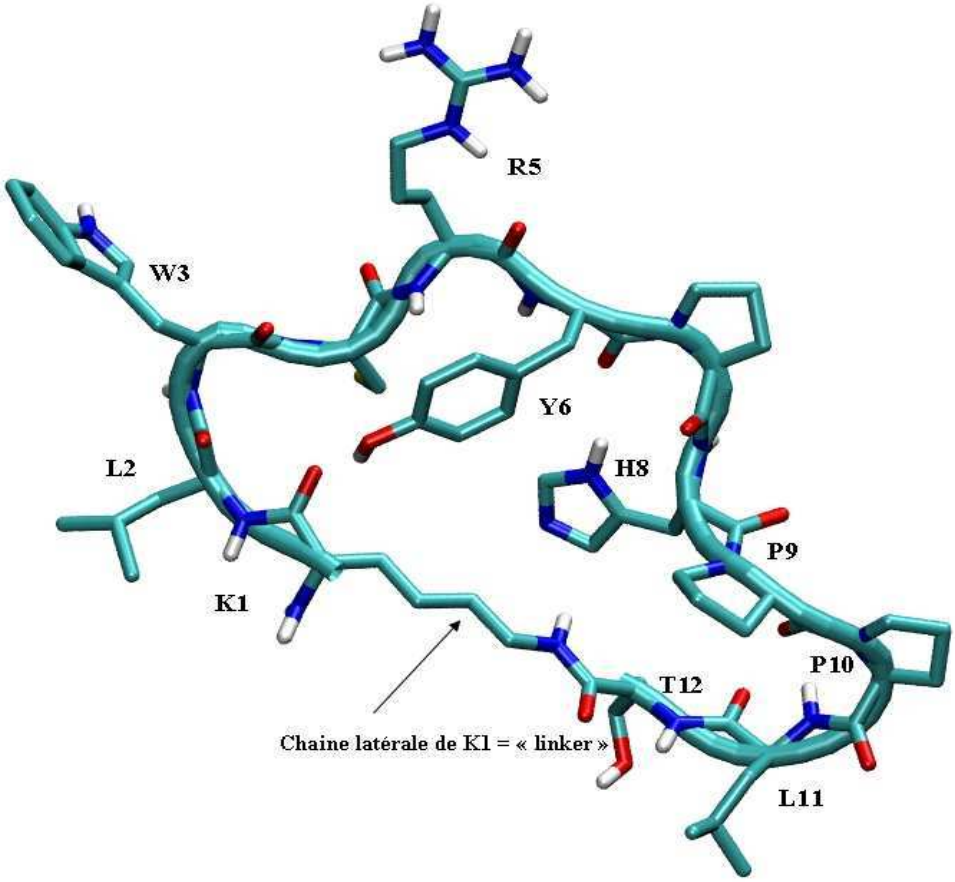


Figure IV.1: Modèle du peptide pF1(1-12) R1K

Il sera donc intéressant de synthétiser les autres peptides cycliques contenant la boucle variable des aptamères peptidiques puis de les tester *in vitro* et *in vivo* pour leur capacité à interagir avec Fur. De plus, les peptides cycliques présentent l'avantage d'être moins sensibles à la protéolyse que les peptides linéaires. Le test de leur activité *in vivo* devrait en être facilité.

Pour revenir sur la comparaison des différentes valeurs d'IC_{50} des peptides, rappelons que le test utilisé est condition-dépendante. Il faudrait dans le meilleur des cas réaliser les expériences des peptides simultanément pour pouvoir comparer. Comme ce procédé n'est pas encore automatisé, il sera difficile de réaliser ces conditions. On peut quand même essayer de comparer l'activité des peptides sur la protéine. Le peptide pF2 est le plus actif avec des séquences qui inhibent la protéine avec seulement quelques équivalents et des IC_{50} de l'ordre

du micromolaire. Ce résultat est cohérent avec celui de la « peptidisation » qui a montré que ce peptide interagit avec la protéine *in vivo* alors que les interactions avec les autres peptides ne sont pas visualisables par double hybride par manque de sensibilité.

Le peptide pF1 semble moins actif que pF2 dans les tests d'activité, or les aptamères peptidiques F1 et F2 ont montré une intensité de coloration bleue identique en test double hybride. Il peut y avoir plusieurs raisons à cela, l'une d'entre elle pourrait être la présence d'interactions entre Fur et la plateforme thiorédoxine de F1 mais pas de F2 en plus de la région variable. Si c'était le cas, ces interactions seraient perdues dans le cas du peptide pF1. On peut bien sûr avoir d'autres hypothèses comme le fait que la contrainte imposée par la plateforme est plus importante pour l'activité dans le cas de F1 que dans le cas de F2.

Quelle que soit la raison, je pense qu'il faut en premier lieu quantifier directement les interactions cible-inhibiteurs pour toutes ces molécules par une technique comme le BIAcore ou d'ITC (« isothermal calorimetry ») par exemple, avant de pouvoir comparer et aussi mieux étudier l'interaction de pF2 avec Fur métallée ou pas. Ceci pourra être étudié par fluorescence avec le peptide muté pF2(1-11)Y11W actif où un tryptophane a été introduit en vue de ces études. Hormis le fait que pF2 pourrait être favorisé par la formation de ponts disulfures intramoléculaires pour interagir avec Fur, pF1 forme des ponts disulfures intermoléculaires qui peuvent défavoriser son interaction avec Fur dans les conditions du test nucléase où aucun réducteur n'a été ajouté. En effet, d'après les modèles théoriques d'interaction proposés de pF1 avec Fur, la formation d'un pont disulfure intermoléculaire C4-C4 déstabiliserait fortement son interaction avec Fur. De plus, nous avons vu que le mutant C4S de pF1 n'inhibe plus Fur dans les tests d'activité. Il sera donc intéressant de tester ce peptide en présence d'un réducteur comme le DTT ou le Tris(2-carboxyéthyl)phosphine (TCEP) et de comparer les effets. Même si aucun changement structural n'a été observé au CD, on a vu que ces peptides peuvent s'oxyder et former des ponts disulfure en spectrométrie de masse MS-ESI. Ces phénomènes pourraient influencer l'activité des peptides sur la liaison de Fur à l'ADN.

Les peptides pF3 et pF4 ont été moins étudiés pendant cette thèse. En effet, Sylvia Vitale avait observé que pF3 donnait des résultats très variables dans le test d'activité. L'aptamère F3 ayant montré la plus faible intensité au niveau de la coloration des tests double-hybride de Nadia Abed, on s'attend à une interaction moins forte avec Fur. Par contre les valeurs d'IC_{50} de pF3 sont proches de celles obtenues pour pF1 même si celles de pF3 sont

plus variables. Il sera intéressant de tester les prédictions des amarrages afin de valider ce modèle qui est le seul à interagir avec uniquement un seul monomère.

Le peptide pF4, quant à lui, n'est actif que sur le monomère de Fur. Ce résultat n'est pas surprenant, puisque lors des tests double-hybride chez la levure réalisés avec des formes tronquées en N-terminal ou C-terminal de Fur, l'aptamère F4 était le seul capable d'interagir avec le domaine C-terminal seul de Fur, qui correspond au domaine de dimérisation de la protéine. De plus, lors d'un test double hybride inversé, F4 était le seul aptamère capable d'inhiber l'homodimérisation de Fur. Il semblerait donc que pF4 interagisse avec le domaine C-terminal du monomère de Fur et que cette interaction empêche la dimérisation de la protéine. Au vu de ce résultat, on ne peut toutefois pas savoir si c'est la liaison du zinc sur les cystéines du site structural de Fur (C92 et C95) qui est inhibée par la présence du peptide, ou si c'est le rapprochement dans l'espace entre les deux monomères métallés.

Si le test de protection à la nucléase renseigne sur la capacité des peptides à inhiber *in vitro* la liaison de Fur à l'ADN, il ne permet pas de déterminer si les peptides pF1, pF2 et pF3 interagissent avec le dimère dans sa conformation inactive (Apo) ou active (métallée). Dans le but d'étudier ce fait, la spectroscopie de fluorescence peut être suivie pour les peptides pF1, pF3 et pF4 qui possèdent un tryptophane. Comme on l'a vu avec pF1, il a été possible de savoir que ce peptide interagissait avec la protéine métallée (qui ne possède pas de tryptophane). Les mêmes expériences avec les peptides pF3, pF4 et pF2(1-11)Y11W (synthétisé dans ce but) permettraient d'élucider la conformation (métallée ou Apo) sur laquelle ces peptides agissent. La connaissance de la forme protéique qui interagit avec chaque peptide enlèvera une difficulté supplémentaire quant à la prédiction des modèles d'amarrage des peptides sur Fur.

Partant de ces résultats, des modèles d'interaction des peptides pF1 à pF3 ont été proposés. Rappelons que les simulations du peptide pF4 n'ont pas continué car on sait que celui-ci interagit sur le domaine de dimérisation d'un seul monomère de Fur et les données structurales en solution obtenues par RMN montrent que ce domaine n'est pas structuré dans le monomère et donc très difficile à modéliser.

Nous avons montré que le peptide pF1 interagit entre les deux monomères de Fur, au niveau des domaines de liaison à l'ADN. De plus, le peptide interagit à la fois avec les domaines de liaison à l'ADN et de dimérisation de Fur métallée. L'ensemble de ces

interactions est compatible avec les résultats de double-hydride et de la spectroscopie de fluorescence du tryptophane. Nous avons montré aussi que la T12 de pF1 interagit avec Y56 de Fur (présente chez toutes les protéines Fur) qui a été identifiée comme interagissant avec les bases de l'ADN. Le peptide est amarré dans une conformation compatible avec l'aptamère F1, c'est-à-dire qu'on peut imaginer l'accrochage de la plate-forme Thiorédoxine sans gène stérique avec la protéine. Ce qui n'est pas le cas pour les autres peptides pF2 et pF3 amarrés. L'amarrage prédit de pF1 est celui qui correspond le plus aux résultats expérimentaux. Ce fait est dû aussi aux données expérimentales nombreuses concernant ce peptide. Il en faudrait d'avantage pour les autres peptides afin d'affiner les modèles prédits. On peut donc facilement imaginer le mécanisme par lequel pF1 pourrait interférer avec l'activité de la protéine Fur d'après le modèle prédit (confère Figure IV.2 et chapitre III).

Le modèle de pF2 n'est pas très cohérent avec les résultats comme nous l'avons vu dans la partie théorique. Plusieurs raisons peuvent expliquer cela. Nous ne savons pas sur quelle forme (Apo ou métallée) de Fur le peptide interagit. De plus, nous avons émis l'hypothèse que le peptide pourrait former des ponts disulfures qui peuvent influencer l'interaction du peptide avec la protéine. Des modèles théoriques construits afin de mimer la conformation du peptide pontée par ces deux cystéines n'ont pas résolu le problème. Rappelons que l'amarrage de ce peptide a été très difficile. Le modèle obtenu avec une énergie libre de liaison favorable n'interfère pas avec les sites métalliques construits dans la protéine. Ces sites métalliques ont été construits, ici, dans l'unique but d'orienter les chaines latérales des ligands du site métallique vers l'atome de fer. Il aurait fallu améliorer ces constructions pour avoir de meilleurs modèles prenant en compte les propriétés électroniques des ions Fe (II).

Ne disposant pas de suffisamment d'informations expérimentales concernant pF3, des prédictions ont été faites concernant les prochaines mutations à effectuer pour mieux caractériser les résidus indispensables et les interactions avec Fur.

Les mutations et les troncations effectuées sur les peptides afin d'identifier les acides aminés importants pour inhiber l'activité de la protéine ont permis de sélectionner les séquences minimales actives :

pF1(3-9), (7 a.a)

pF2(3-10), (8 a.a)

probablement pF3(1-11), (11 a.a).

Aucune mutation, à l'heure actuelle sur pF4 n'a permis de déterminer une séquence minimale active pour ce peptide. Il sera intéressant dans l'avenir de réaliser des mutations et des troncations sur ces peptides pF3 et pF4

L'ensemble de ces résultats (théoriques et expérimentaux) permet de proposer des zones d'interaction des peptides pF1 à pF4 avec les différentes conformations d'EcFur (monomère, dimère inactif et dimère actif) (Figure IV.2).

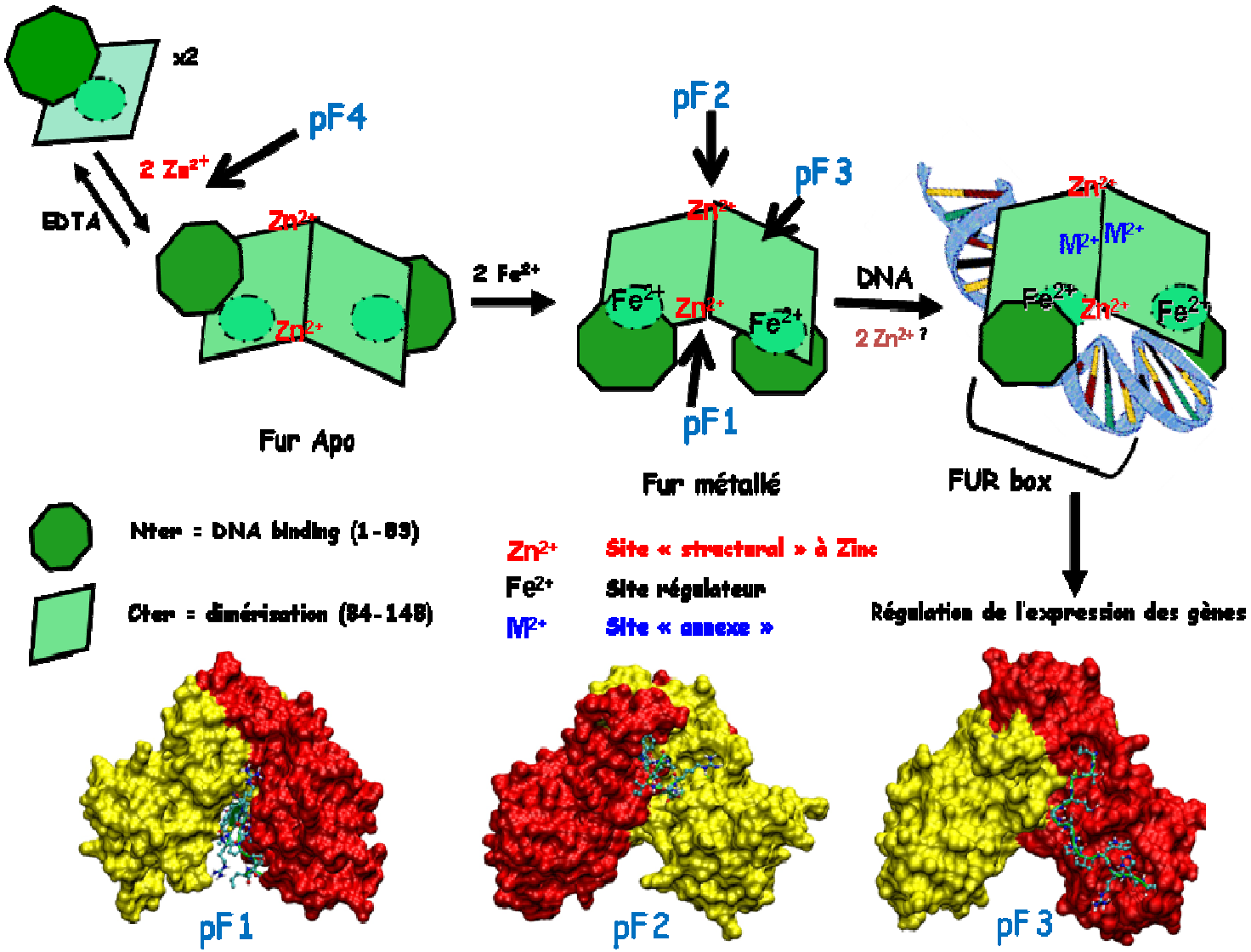


Figure IV.2: Schéma du mécanisme d'action de Fur, des sites d'interaction de pF1 à pF4 et les modèles de pF1 à pF3 sur Fur. Mécanisme général d'activation de la protéine Fur avec les différentes conformations et les sites métalliques.

Pour résumer la Figure IV.2, le peptide pF4 interfère avec la dimérisation de la protéine Fur. Le peptide pF1 interagit entre les deux monomères au niveau des domaines de liaison à l'ADN. Le peptide pF2 interagit au niveau des domaines de dimérisation et pF3 interagit avec une seule sous-unité.

Le tableau suivant résume les propriétés des séquences des peptides actifs sur Fur à travers le test d'activité. Les valeurs des énergies libres de liaison théoriques y sont regroupées également.

Notations	séquences pF1	ΔG (kcal/mol)	IC_{50} (µM)
pF1 (1-13)	RLWCRYPHPPLTD	-18	60±
pF1 (1-10)	RLWCRYPHPP	-18	64±10
pF1 (1-9)	RLWCRYPHP	-17	75±5
pF1 (1-8)	WCRYPH	-17	inactif
pF1 (2-9)	LWCRYPHP	-14,5	75
pF1 (3-9)	WCRYPHP	-13	50
pF2(1-13)	RQCNICGASLYSY	-16	12,5±2
pF2(1-11)	RQCNICGASLY	-11,9	7±2
pF2(1-10)	RQCNICGASL	-11,6	3±1
pF2(2-10)	QCNICGASL	-7,1	75±5
pF2(3-10)	CNICGASL	-2,8	≥75
pF3 (1-13)	ETCKCGSQVWRHS	-16	30±17
pF3 (1-11)	ETCKCGSQVWR	-11,7	<50

Tableau IV.1: Récapitulatif de l'interaction des séquences inhibitrices d'EcFur

Ce Tableau IV.1 montre que les énergies libres de liaison (**ΔG**) et les IC_{50} de pF1 évoluent dans le même sens. En effet, lorsque le **ΔG** augmente (défavorable pour l'interaction), l'IC_{50} diminue et vice versa.

Concernant les amarrages de pF2, nous avons déjà parlé du fait de l'incohérence avec les tests d'activité. Un point essentiel concerne les tyrosines Cter Y11 et Y13 qui paraissent importantes dans les modèles d'amarrage alors qu'elles sont inutiles, voire déstabilisantes dans les tests d'activité. En effet, la troncation de pF2 à l'extrémité Cter montre une activité meilleure avec la protéine que la troncation à l'extrémité Nter.

Nous ne disposons pas de suffisamment d'informations expérimentales concernant pF3 pour pouvoir faire une comparaison des données théoriques et expérimentales. Néanmoins, on sait que ce peptide aussi se lie à la protéine, lorsque les domaines Nter et Cter

sont présents. Le modèle d'amarrage montre également que le peptide interagit avec une sous-unité en reliant le domaine de liaison à l'ADN et le domaine de dimérisation (Figure IV.2).

Un criblage double hybride chez la levure a été réalisé dans le passé afin d'évaluer la spécificité des aptamères peptidiques comme inhibiteur de la protéine EcFur. Les aptamères peptidiques F1 à F4 ont été testés pour leurs interactions avec des protéines Fur de différents microorganismes pathogènes. Cette spécificité s'avéra très forte entre les aptamères F1, F2 et F3 avec EcFur, puisque ces aptamères n'interagissent ni avec Fur de *Yersinia pestis* (YpFur) ni avec Fur de *Vibrio salmonicida* (VsFur) alors que ces protéines ont respectivement 85 % et 80 % d'identité de séquence avec EcFur. L'aptamère F4, quant à lui, est capable d'interagir avec ces deux protéines, mais aussi avec Fur de *Pseudomonas aeruginosa* (PaFur) qui ne présente « que » 55 % d'identité de séquence avec EcFur. De même, nous avons étudié *in silico* l'interaction des peptides avec les protéines PaFur, VcFur et HpFur dont les structures sont disponibles. Les résultats d'amarrage de pF1 ont été présentés. Comme nous l'avons mentionné, c'est le peptide pour lequel simulations et expériences théoriques sont en accord. Les amarrages du peptide ont montré que pF1 interagissait favorablement avec PaFur, un peu moins avec HpFur et très peu avec VcFur. Le fait que pF1 n'inhibe pas HpFur dans le test de protection à la nucléase est cohérent avec l'amarrage du peptide présenté. La faible valeur d'énergie libre de liaison obtenue pourrait expliquer l'absence d'effet de pF1 sur HpFur. Par contre l'amarrage de pF1 sur PaFur est très favorable. L'explication qu'on peut donner à ce résultat est le fait que le peptide pF1 peut avoir un comportement différent de l'aptamère F1 dont il est issu. En effet, le peptide pF1 peut avoir un degré de liberté important lorsqu'il est libre, favorisant l'adoption d'une conformation interagissant mieux avec la protéine. L'amarrage du peptide sur VsFur n'a pas été présenté mais cette protéine présente un pourcentage de 98% avec VcFur dont la structure est disponible. Par analogie, on peut comparer l'amarrage de pF1 sur VcFur qui a montré une affinité très faible. Ce fait est assez surprenant puisque le modèle EcFur←Vc, construit à partir de la structure de VcFur interagit favorablement avec pF1. Pour élucider ce fait, une analyse poussée de l'interaction de pF1 par CHARMM sur toutes les protéines a été menée. Les résultats d'énergies d'interactions au cours d'une dynamique sont en accord avec les énergies libres de liaison obtenues par AutoDock4.2.

Pour la suite de ce travail, il sera important d'optimiser les conditions de stockage (-80°C au lieu de -20°C) et le suivi de l'état d'oxydation et de solubilité des peptides Le test de

protection à la nucléase est très délicat en soi et ne permet qu'une mesure indirecte de l'effet du peptide sur la protection de l'ADN par la protéine. Il fait intervenir plusieurs produits qui peuvent chacun perturber l'expérience. Pour une expérience réussie, il faut une « bonne » digestion du plasmide pDT10 qui est l'indicateur de l'activité de la protéine. Il faut que la protéine soit activée par le métal Mn^{2+} et inhibée par l'EDTA. Idéalement, un contrôle doit être fait pour vérifier tous ces faits avant de tester l'effet des peptides sur l'activité de la protéine. Pour déterminer l'IC_{50} d'un peptide, il faut une gamme de concentration importante permettant d'observer un effet « zéro » (pas d'inhibition de la protéine activée), un effet inhibiteur à 50% et à 100% de l'activité de la protéine. Tout cela pour observer une migration des bandes d'ADN sur un gel, dont le profil montre le résultat d'interaction du peptide avec la protéine. Toutefois, certaines séquences ont montré des résultats encourageants. Il sera important à l'avenir d'utiliser des méthodes d'études des interactions moléculaires directes plus quantitatives.

Des méthodes comme la résonance plasmonique de surface (BIAcore), l'ITC et la technique du double hybride sont très intéressantes. L'utilisation de ces méthodes a commencé avec Sophie Mathieu, actuellement en post-doc au laboratoire. Le BIAcore devrait permettre de mieux quantifier les interactions et de déterminer des constantes de dissociations (Kd). On pourra alors mieux comparer l'effet des peptides sur les protéines contrairement au test de protection à la nucléase, qui dépend fortement des conditions et qui visualise une interaction indirecte *via* la coupure des bandes d'ADN. La technique du double-hybride a été utilisée pour sélectionner les aptamères peptidiques F1 à F4 pour leur interaction avec la protéine Fur. Elle a été utilisée pour évaluer l'activité des peptides pF1 à pF4 contre la protéine *in vivo*. Cela a d'ailleurs montré que pF2 est actif *in vivo* et confirmé que ce peptide a une affinité très forte pour Fur. Par contre les autres peptides pF1, pF3 et pF4 n'étaient pas actifs alors qu'ils le sont dans le test *in vitro*. De plus, il n'a pas été possible de bien quantifier l'intensité de la coloration bleue observée dans ces expériences, avec le gène rapporteur de la β-galactosidase *lacZ*. L'utilisation du gène rapporteur *luc* de la Luciférase, devrait remédier à cela et permettre de mieux quantifier l'interaction *in vivo*.

Avec ces méthodes, il sera possible de mieux caractériser *in vivo* et *in vitro* les interactions des peptides et de savoir si les peptides agissent sur la protéine Apo ou métallée. Il sera intéressant de tester les peptides tronqués actifs en test nucléase, afin de voir si la troncation entraîne une perte ou une augmentation de l'affinité du peptide pour Fur et évoluer

vers des petites molécules de plus en intéressantes pour des applications thérapeutiques. Il sera intéressant de tester expérimentalement les petites molécules sélectionnées par le criblage virtuel. Certaines molécules ont montré une interaction très favorable *in silico* avec la protéine Fur. Il sera toujours possible d'élargir la recherche de petites molécules à partir des bases de données et de tester virtuellement leurs interactions avec Fur. Il faudra pour cela prendre en compte l'effet des métaux dans les simulations en faisant intervenir la mécanique quantique.

Signalons que l'approche théorique que nous avons utilisée présente également quelques limitations :

- La structure de la protéine EcFur est inconnue et nous nous sommes basés sur un modèle par homologie à partir de VcFur, présentant certes une très bonne similarité de séquence mais qui pourrait tout de même présenter des erreurs au niveau du positionnement des chaines latérales des acides aminés.
- Les programmes d'amarrage moléculaire comme AutoDock ne sont pas capables de prendre en compte de manière fiable le très grand nombre de degrés de liberté présents dans des peptides de 13 acides aminés. Les résultats obtenus peuvent alors ne plus être statistiquement significatifs.
- L'introduction de la dynamique moléculaire permet de pallier le problème de la rigidité de la protéine lors des amarrages mais la nécessité de l'utilisation d'une méthode de solvatation implicite pour avoir des temps de calcul raisonnables, enlève toute possibilité d'interaction entre peptides et molécules d'eau ou de prise en compte de la protonation des cystéines ou des histidines.
- Les métaux n'ont pas été introduits de manière explicite dans les simulations de dynamique des complexes ou dans les amarrages.

Dans la suite de ce projet, nous nous intéresserons plus particulièrement à l'interaction de petites molécules avec Fur pour lesquelles les programmes d'amarrage commerciaux sont très efficaces. Nous effectuerons des paramétrisations complètes de ces molécules pour le champ de force de CHARMM à partir de calculs de chimie quantique et nous représenterons explicitement tous les atomes y compris les molécules d'eau proches du site de liaison sur Fur. Cette approche quantique nous permettra également de représenter les deux états possibles Apo ou métallé des protéines Fur.

L'entrée des molécules inhibitrices de Fur sera aussi un enjeu crucial pour étudier leurs effets *in vivo*. Les premiers tests préliminaires réalisés avec les peptides libres sur des cultures n'avaient montré aucun effet sur l'activité de la protéine. Ceux-ci soit ont été dégradés soit n'ont pas passé les membranes bactériennes ce qui n'est pas étonnant en soi. C'est pourquoi le laboratoire développe une stratégie efficace pour les faire entrer dans la bactérie. Cette stratégie consiste à conjuguer des molécules antibactériennes à des sidérophores et à les internaliser par le système naturel de transport des sidérophores. Cette stratégie portant le nom de « Cheval de Troie » fait l'objet d'une collaboration avec Gaëtan Mislin de Strasbourg dans le cadre du projet PepSiFur, financé par l'ANR en 2011. Un schéma de la stratégie envisagée est illustré sur la Figure IV.3, ci-dessous.

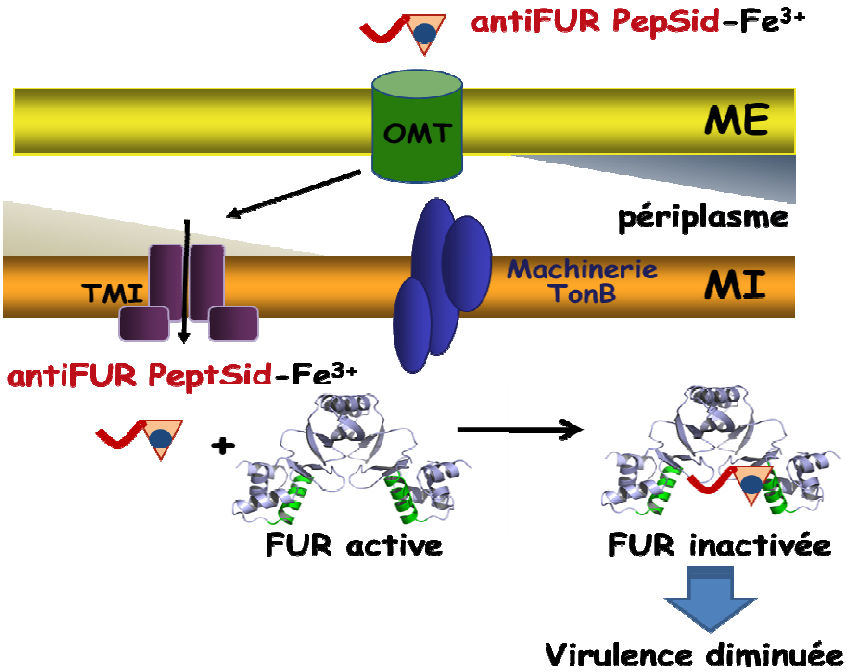


Figure IV.3: Schéma de stratégie du « Cheval de Troie » avec inhibiteurs anti-Fur

Les peptides seront accrochés par une seule ou les deux extrémités à un sidérophore afin d'être internalisés par le système naturel de transport des sidérophores (Figure IV.3).

Les peptides pouvant être dégradés, la cyclisation des peptides pourrait résoudre ce problème. De plus, les peptides seront plus proches des aptamères peptidiques au niveau de la conformation du peptide, constituant la partie active.

D'ailleurs, nous avons montré dans le cas de pF1 que le peptide cyclique peut être plus actif que son homologue linéaire.

Même si les modèles permettent de mieux comprendre les interactions peptides / protéines, seule, la résolution des structures des protéines en complexation avec un inhibiteur peptidique (ou non), permettra à coup sûr d'élucider le mécanisme d'inhibition de la protéine. Pour cela, plusieurs essais de co-cristallisation ont été effectués, mais pour le moment, nous n'avons observé aucune cristallogenèse. Il faudra insister sur ces essais de co-cristallisation des complexes Fur-inhibiteurs, tout en essayant différents produits chimiques mais aussi en optimisant les conditions prometteuses.

D'autre part, plusieurs autres protéines Fur de pathogènes (*Y.pestis, P. aeruginosa, F. tularensis et L. pneumophila*) sont en cours de purification et de caractérisation au laboratoire en collaboration avec l'IBS. Nous espérons grâce à la double approche expérimentée ici, pouvoir découvrir des inhibiteurs dirigés contre les protéines Fur de ces pathogènes à partir de ceux étudiés en détail dans cette thèse.

Chapitre V :

Matériels et méthodes

CHAPITRE V : ..... **- 249 -**

MATERIELS ET METHODES ..... **- 249 -**

I| SYNTHESE CHIMIQUE PEPTIDIQUE ..... - 253 -

I. A| Principes et mode opératoire ..... *- 253 -*

I. B| Récapitulatif des propriétés des peptides synthétisés au DCM ..... *- 257 -*

I. C| Récapitulatif des caractéristiques des peptides synthétisés par Genscript ..... *- 260 -*

I. D| Étude des peptides par dichroïsme circulaire ..... *- 262 -*

II| TEST D'ACTIVITE DE LIAISON A L'ADN DE LA PROTEINE FUR D'*E. COLI* ..... - 263 -

II. A| Principe ..... *- 263 -*

II. B| Mode opératoire du test ..... *- 264 -*

II. C| Surproduction et purification des protéines EcFur ..... *- 264 -*

II. D| Essais de co-cristallisation ..... *- 266 -*

III| AMARRAGE MOLECULAIRE ..... - 268 -

III. A| L'amarrage automatique avec Glide ..... *- 268 -*

III. B| Méthode de paramétrisation des petites molécules pour CHARMM ..... *- 270 -*

III. C| Matériels informatiques ..... *- 271 -*

I| Synthèse chimique peptidique

Les synthèses ont été réalisées grâce à la collaboration avec Didier Boturyn du département de chimie moléculaire (DCM). L'ingénieur de recherche Remy Garcia et le technicien Miguel Calin, responsables des synthétiseurs automatiques ont réalisé les synthèses au robot. Certains peptides ont été achetés chez Genscript.

I. A| Principes et mode opératoire

La synthèse chimique des peptides se fait de l'extrémité Cter vers le Nter de la séquence peptidique. Il peut être réalisé en solutions ou généralement en phase solide (SPPS). La SPPS peut se faire manuellement ou sur un robot synthétiseur. Deux stratégies peuvent être utilisées : la stratégie Fmoc/tBu et la stratégie Boc/Bz. Dans notre cas, nous avons utilisé la SPPS manuellement dans un premier temps et automatiquement lorsque le nombre d'acides aminés à synthétiser devenait important en stratégie (Fmoc / tBu).

I. A. 1| La SPPS manuelle

En SPPS manuelle, le dispositif est composé d'un ballon en verre à deux cols, équipé d'un fritté au fond du ballon. Ce dispositif permet d'éliminer les solvants et les excès de réactifs sous air comprimé lors des lavages. Avant l'utilisation, le ballon est traité pendant 12 h (typiquement une nuit) avec du Dimethyldichlorosilane $(CH3)_2SiCl_2$ comme lubrifiant pour empêcher les billes de résine de coller à la paroi intérieure lors de la synthèse. Le $(CH3)_2SiCl_2$ est acide. Pour éliminer toute trace d'acidité le ballon est soigneusement lavé avec du CH_2CI_2 jusqu'au dégagement complet des traces d'acides. La résine généralement couplée à un premier acide aminé (protégé sur sa chaine latérale par un groupement protecteur, différent selon l'acide aminé et l'amine Nter par un groupement Fmoc) est préchargée pendant 30 min dans du CH_2Cl_2 (DCM) anhydre à 1 mmol / g. On rajoute 6 mL de méthanol / *N*,*N*-Diisopropylethylamine (DIPEA) / DCM (2/1/17) pendant 5 min qu'on lave et on répète l'opération une deuxième fois pendant 10 min. Au début de la synthèse et après chaque lavage à l'éther, la résine est lavée et gonflée deux fois avec du CH_2CI_2 (20 mLmL / g de résine) pendant 15 min et une fois avec du Diméthylformamide (DMF) (20 mL / g de résine) pendant 15 min. Les réactions de couplage sont effectuées dans le DMF en utilisant, 1,5 à 2

équivalents (par rapport à la charge de résine ; loading) de Na-Fmoc-acide aminé protégé, *in situ.* L'acide aminé est activé par 1,5 à 2 éq de PyBOP et 3-4 éq de DIPEA, (10 mL / g de résine) pendant 30 min. La résine est ensuite lavée deux fois avec du DMF (20 mL / g de résine) pendant 1 min et deux fois avec du CH_2CI_2 (20 mL / g de résine) pendant 1 min. L'état de couplage de la réaction est vérifié par le test de l'acide trinitrobenzènesulfonique (TNBS).

I. A. 1.a| Test TNBS :

Après le dernier lavage au DCM, on prélève quelques billes qu'on met dans un tube à hémolyse. On rajoute 4 gouttes d'une solution de DIEA/DMF (1 : 9) et 4 gouttes d'une solution 1% de TNBS (couleur jaune). On agite pendant une minute et on observe si on voit une coloration orange. Le TNBS réagit spécifiquement avec les amines libres pour donner une coloration rouge (test positif) indiquant un couplage non total. La réaction doit être reprise jusqu'à l'obtention des billes de résines incolores (test négatif).

I. A. 1.b| Déprotection du groupement Fmoc

Lorsque la réaction de couplage est totale, il faut enlever le groupement Fmoc de l'acide aminé couplé pour pouvoir en rajouter d'autres. Cette déprotection (de-Fmocage) est réalisée par un traitement avec de la pipéridine / DMF (1:4) (10 mL / g de résine) pendant 10 min. Le processus est répété trois fois et la résine est lavée cinq fois avec du DMF (10 mL / g de résine) pendant 1 min. La déprotection du Fmoc est vérifiée par mesure UV. En effet, le groupement Fmoc libère du dibenzofulvène qui réagit avec de la pipéridine et absorbe la lumière dans la gamme des UV (λ = 299 nm, ε = 7800 $M^{-1}.cm^{-1}$). Les solutions de clivage du Fmoc et des lavages sont collectées ensemble dans une fiole jaugée de 250 mL (le volume de la solution est ajusté avec du méthanol à une valeur connue V= 250 mL). L'absorbance de ce mélange à 299 nm permet de calculer le taux de chargement de la résine grâce à la relation suivante :

$$\mathbf{n_{fmoc}\ (mmol) = (A_{299nm} \times d/7800) \times V}$$

d étant le facteur de dilution.

Le taux de chargement permet de calculer la quantité de réactifs nécessaire au couplage suivant.

I. A. 2| La SPPS automatique

La synthèse SPPS des peptides par la voie automatique a été effectuée par le robot ABI 433 A1 Peptide Synthesizer (Applied Biosystems) qui fonctionne avec la méthode standard de synthèse en phase solide ou par le robot 348 Synthétiseur Ω (Advance ChemTech) qui utilise les mêmes conditions que la procédure manuelle.

I. A. 2.a| Décrochage et déprotection des peptides

La libération du peptide de la résine et des groupements protecteurs de chaines latérales peut se faire en une seule étape selon la nature de ces composés.

La résine couplée au peptide est mise à « gonfler » dans du dichloromethane (DCM). Pendant ce temps, la solution à utiliser pour la déprotection est préparée avec 92,5% de Trifloroacétonitrille (TFA), 2,5% de d'eau distillée, 2,5% de composés captant les carbocations (C^+) libérés pendant la réaction et 2,5% de Triisopropylsilane (TIS) dans cet ordre d'ajout. Le TIS est ajouté à la solution qui est bien mélangée juste avant l'ajout sur la résine. On doit observer une coloration jaune due à la capture des carbocations par les scavengers. A l'absence de Scavengers, les C^+ donnent une coloration vive rouge. La réaction se déroule sous forte agitation pendant 3h (ce temps depend de la nature des groupements protecteurs et de la résine). A la fin de la réaction, la solution de déprotection (le jus) est récupérée dans un ballon (préalablement taré) et évaporée dans un rotavapor. Après l'évaporation du TFA, le produit est repris dans de l'éther qui permet de cristalliser le peptide. Cette opération est accélerée par le froid. Après la décantation, le surnageant est enlevé et le produit est récupéré après une évaporation rapide du solvant restant. Le produit est pésé et analysé avant une purification.

I. A. 2.b| Analyse et purification de peptide par HPLC :

L'analyse des peptides synthétisés et déprotégés a été effectuée sur un dispositif composé :

- D'un équipement Waters comprenant un analyseur *Waters 600 controller*, un détecteur *Waters 2487 Dual Absorbance Detector* et un *Waters In-Line Degasser*.
- Une colonne analytique *Nucleosil 120 Å 3 µm C18 particles, 30 x 4 mm*^{2} opérant à un flux de 1,3 mL.min^{-1} avec un gradient de 5 à 100% de solvant B (solution d'acétonitrile (CH_3CN) contenant 9,9% d'eau HPLC et 0,1% de TFA) pendant 15 min. Le solvant A est une solution d'eau HPLC contenant 0,1% de TFA.

L'eau HPLC est obtenue après filtration de l'eau distillée par le système *Milli-Q® cartridge*. L'acétonitrile et le TFA sont aussi de qualité nécessaire pour une utilisation HPLC.

L'analyse s'effectue avec une injection 10 µL de la solution filtrée du peptide dissout par ultra-sons dans l'acétonitrile (environ 1 mg/mL). L'enregistrement est suivi par l'aborbance à 214 et 250 nm, correspondant respectivement à l'absorption des liaisons peptidiques et des groupements aromatiques de certaines chaines latérales et des groupements protecteurs. La Figure V.1 montre le profil de pF1(1-10) après déprotection des chaines latérales et décrochage de la résine.

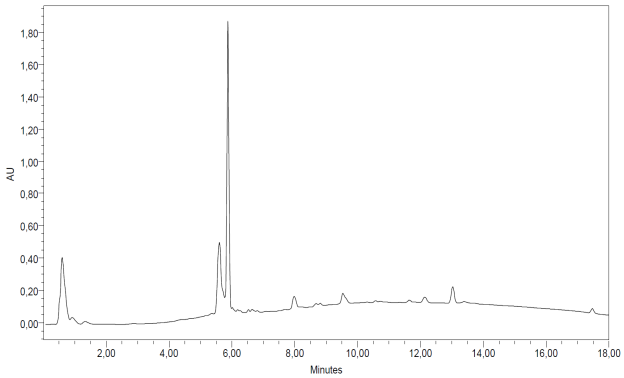


Figure V.1: Spectres HLPC de pF1(1-10) synthétisé au DCM. L'unité AU (arbitrary unit) montre l'évolution de l'absorption à 214 nm, spécifiques des liaisons peptidiques, en fonction du temps. 3 µL de la solution du peptide dissout dans d'acétonitrile/eau 50 :50 + 0,1 TFA à 1mg/ml. Le pic sortant à 5,9 minutes contient un produit de masse attendue pour le peptidesynthétisé (vérifié par spectométrie de masse) . Ce pic est collecté dans une étape de purification.

La purification s'effectue sur le même type de dispositif avec cette fois-ci un dégazage des solvants avec de l'azote à 50 mL.min^{-1} en utilisant une colonne préparative *Delta-Pak™ 300 Å 15 µm C18 particles, 200 x 25 mm*^{2} opérant à un flux de 22 mL.min^{-1} sur un gradient allant de 5 à 100% de solvant B en 30 min. Elle se fait par injections successives de 2 mL d'une solution concentrée constituée du peptide libre dissout dans un mélange d'eau HPLC/acétonitrile (50/50). L'enregistrement est suivi par l'aborbance à 214 et 250 nm. Les fractions correspondant aux pics ayant une intensité importante sont collectées et analysées par spectrométrie de masse.

I. A. 2.c| Analyse par spectrométrie de masse

Il a été réalisé sur la machine *Esquire 3000* (Bruker) en utilisant la méthode d'ionisation par pulvérisation d'électrons (ESI-MS). Le spectromètre de masse convertit les données de fragmentation en poids moléculaire. Les fractions collectées contenant le produit

dont la masse moléculaire correspond à la masse théorique calculée, sont assemblées et lyophilisées.

Dans les conditions d'analyse par masse (exemple de pF1(1-10) sur la Figure V.2), on voit qu'il n'était pas possible de savoir si les cystéines de certains peptides en contenant 2 étaient oxydées en pont disulfure ou réduites. Il aurait fallu enregistrer avec une fenêtre couvrant 2 fois plus large, c'est-à-dire un « scan » m/z autour de 3400 pour voir la masse du dimère.

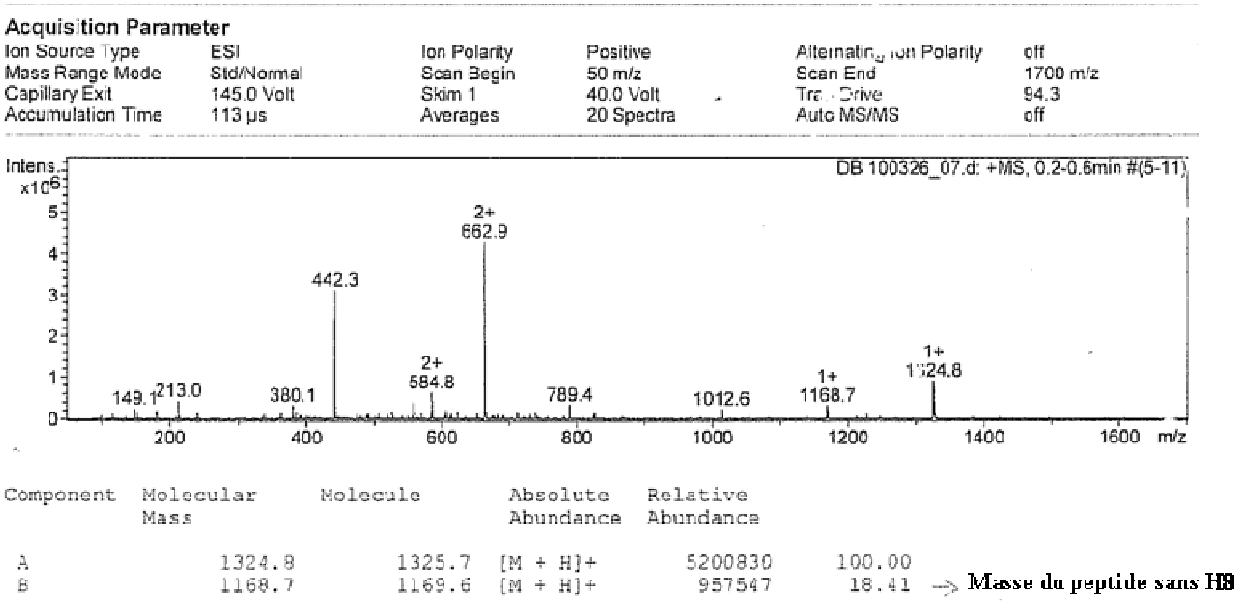


Figure V.2: Spectre MS-ESI de pF1(1-10) purifié

I. A. 3| Procédure générale pour la cyclisation des peptides

Le peptide linéaire est dissous dans le DMF et le pH est ajusté à 8-9 par addition de DIPEA. On ajoute le PyBOP (1,2 éq) et la solution est agitée à température ambiante pendant 1 h. Le solvant est éliminé et le produit est dissout dans un volume minimum de CH_2Cl_2. L'éther diéthylique est ajouté pour précipiter le peptide. Le mélange est broyé et lavé trois fois à l'éther éthylique pour obtenir du produit brut qui peut être utilisé sans étape de purification supplémentaire.

I. B| Récapitulatif des propriétés des peptides synthétisés au DCM

Voir tableau ci-dessous.

En rouge la position d'une mutation comparée à la séquence initiale du peptide. Deux fractions ont été élués pour les séquences marquées par un (*). Les temps de rétention ne sont pas donnés pour tous les peptides. La pureté de certains peptides ont été pris en compte à partir de l'analyse MS.

Notation des Peptides	**Séquences**	**Masse de produit pur obtenue (mg)**	**MM + masse TFA (114)**	**Temps de rétention HPLC (min)**
pF1(1-13)	**RLWCRYPHPPLTD**	**10**	**2110 (=1654+4x114)**	**6,53**
pF1(1-9)	**RLWCRYPHP**	**32**	**1569**	**6,28**
pF1(1-9) C4S	**RLWSRYPHP**	**31,2**	**1553**	**6,5**
pF1(4-9)	**CRYPHP**	**11,4**	**1000**	**-**
pF1(4-9) C4S	**SRYPHP**	**22,4**	**984**	**-**
pF1(1-10)	**RLWCRYPHPP**	**28**	**1780,5**	**6,4**
pF1(2-9)	**LWCRYPHP**	**57**	**1413**	**6,3**
pF1(3-9) Δ8	**WCRYP_P**	**4**	**1049**	**5,7**
pF1(1-8)	**RLWCRYPH**	**11**	**1586**	**6,1**
pF1(3-9)	**WCRYPHP**	**18,6**	**1300**	**5,8**
pF1(5-13)	**RYPHPPLTD**	**35**	**1437**	**7,1**
pF1(8-13)	**HPPLTD**	**11,5**	**907**	**7,4**

pF1(1-12) R1K	KLWCRYPHPPLT	6,8	1966,8	6,7
pF1(1-12) R1K cyclique**	KLWCRYPHPPLT	7,8	1834,8	10
pF2(1-11)	RQCNICGASLY	22	1455	-
pF2(1-10)	RQCNICGASL	26	1292	6,3
pF2(2-10)*	QCNICGASL	8,8+19,8	1004 et 1022	6,4
pF2(3-10)	CNICGASL	10,2	894	6,5
pF2(1-10) C3S	RQSNICGASL	1,2	1276	-
pF2(1-10) C6S	RQCNISGASL	en solution	1276	-
pF2(1-10) C3S et C6S	RQSNISGASL	en solution	1260	-
pF2(1-11) Y11W	RQCNICGASLW	3,8	1478	-
pF2(3-13)	CNICGASLYSY	10	1307	-
pF3(1-11)	ETCKCGSQVWR	2,4	1638	-
pF3(3-13)	CKCGSQVWRHS	0,5	1746	-
pF4(1-11)	CARCGARVNVY	4,2	1553	5,8
pF4(3-13)	RCGARVNVYKY	10	1784,5	7,7
pF5(1-13)	RSLRGRCLSQHQD	35	2124	5,2

I. C| Récapitulatif des caractéristiques des peptides synthétisés par Genscript

Notation des Peptides	Séquences	Masse de produit pur obtenue (mg)	MM sans les masses TFA	Pureté %	Recommandations de Genscript	Volume de Rétention (mL)
pF1(1-13)	**RLWCRYPHPPLTD**	**10+10**	**1653,4**	**96 et 94,8**		14,7*
pF1(1-13) Y6A	**RLWCRAPHPPLTD**	**5**	**1563,6**	**95,3**		12**
pF1(1-13) H8A	**RLWCRYPAPPLTD**	**5**	**1587,9**	**97,8**		12,66**
pF1(1-13) R5A Y6A	**RLWCAAPHPPLTD**	**5**	**1476,7**	**96**		12,66**
pF1(1-13) H8A T12A	**RLWCRYPAPPLAD**	**5**	**1557,8**	**95,8**		7,8**
pF1(1-9)	**RLWCRYPHP**	**10+5**	**1227,8**	**92,8 et 96,7**	Stockage à -20 °C, soluble dans l'eau à pH 7	11,43**** et 12,5**
pF1(2-9)	**LWCRYPHP**	**10**	**1071,9**	**95,1**	Stockage à -20 °C, soluble dans l'eau à pH 7	17,01****
pF1(3-9)	**WCRYPHP**	**10**	**958,3**	**93,4**	Stockage à -20 °C, soluble dans l'eau à pH 7	14,541****
pF2(1-13)	**RQCNICGASLYSY**	**10**	**1477,5**	**95,7**		13,99*
pF2(1-13) Y11A Y13A	**RQCNICGASLYSY**	**5**	**1291,9**	**98,3**		12,3*
pF2(2-10)	**QCNICGASL**	**10**	**949,4**	**91,1**	Stockage à -80 °C, soluble	16,11*

					dans l'eau à pH 4, facilement oxydable en solution neutre ou basique	
pF2(3-10)	**CNICGASL**	**10**	**779,4**	**97**	Stockage à -80 °C, soluble dans l'eau à pH 10, très facilement oxydable en solution neutre ou basique	16,58****
pF3(1-13)	**ETCKCGSQVWRHS**	**10**	**1520,8**	**97,9**		18,04**
pF4(1-13)	**CARCGARVNVYKY**	**10**	**1502,8**	**96,5**		8,55***

***** Colonne VP-ODS 4,6x150mm, Débit à 1mL/min, détection à 214 nm. Pompe A : 0,065 % acide trifluoroacétique (TFA) dans l'eau, Pompe B : 0,05% acide trifluoroacétique dans l'acétonitrile (ACN), T0 : 5% de B et T25min de B

**Colonne Venusil MP C18-5, 4,6x150mm, Débit à 1mL/min, détection à 220 nm. Pompe A : 0,1 % acide trifluoroacétique (TFA) dans l'eau, Pompe B : 0,1% acide trifluoroacétique dans l'acétonitrile (ACN), T0 : 5% de B et T25min de B

******* Colonne SinoChrom ODS-BP-5 4,6x250mm, Débit à 1mL/min, détection à 214 nm. Pompe A : 0,065 % acide trifluoroacétique (TFA) dans l'eau, Pompe B : 0,05% acide trifluoroacétique dans l'acétonitrile (ACN), T0 : 5% de B et T25min de B

**** Colonne Alltima^{TM} C18 5 4,6x250mm, Débit à 1mL/min, détection à 220 nm. Pompe A : 0,065 % acide trifluoroacétique (TFA) dans l'eau, Pompe B : 0,05% acide trifluoroacétique dans l'acétonitrile (ACN), T0 : 15% de B et T25min 65 ou 75% de B

I. D| Étude des peptides par dichroïsme circulaire

Les expériences ont été réalisées sur un spectrophotomètre Chirascan d'Applied photophysics à 25°C. Dans l'UV-lointain, les spectres ont été enregistrés entre 195-250 nm dans des cuves de 300 µL et de 1 mm de trajet optique. Les conditions de mesures sont les suivantes : Bande passante : 1 nm ; Pas : 1 nm ; Vitesse : 50 nm/min ; Réponse du détecteur : 1 sec. 4 scans par expériences sont enregistrés et moyennés. Les spectres, obtenus, ont tous été normalisés afin de faire abstraction de la concentration en peptide utilisée, et du nombre de résidus de ces derniers (au préalable la contribution du tampon est soustraite : $\theta_{peptide}$-θ_{tampon}). Les spectres équivalents du tampon contenant les mêmes quantités de métal ont été enregistrés afin de soustraire la contribution des sels à l'ellipticité des échantillons peptidiques.

L'équation de normalisation est la suivante :

$$[\theta] = \frac{\theta \times 0.1}{([C] \times T \times Naa)}$$

Où $[\theta]$ correspond à l'ellipticité molaire par résidu en deg.dmol-1.cm², θ correspond à l'ellipticité mesurée en mdeg, [C] correspond à la concentration peptidique par sous unité monomérique en M, T précise le trajet optique en cm et *Naa* correspond au nombre de résidus.

Les solutions mères de peptides ont été diluées soit dans du tampon Phosphate 20 mM pH 7.5, soit dans du BisTris Propane 10 mM pH 7.5 KCl 10 mM $MgSO_4$ 5 mM pour obtenir des solutions à 20 µM. Les spectres CD en UV-lointain des peptides pF1 à pF4 à 20 µM ont été enregistrés après 15 minutes d'équilibration.

II| Test d'activité de liaison à l'ADN de la protéine Fur d'*E. coli*

Pour évaluer l'effet des molécules sur l'activité de liaison de la protéine EcFur, le test de la protection de la coupure par une nucléase a été utilisé.

II. A| Principe

Le test à la nucléase permet de tester la capacité des protéines purifiées Fur d'*Escherichia coli* métallées ou non à se lier ou non à l'ADN.

Ce test nécessite un plasmide (pDT10) dans lequel a été insérée la séquence d'ADN qu'EcFur est supposée lier. Cette séquence correspond au promoteur de l'aérobactine qui contient naturellement un site (GANTC) pour l'enzyme de restriction *HinfI*. Le plasmide pDT10 correspond à un plasmide dérivé du pUC19 et qui a été construit par Danièle Touati (D'autréaux *et al*., 2002).

Selon les conditions testées, si la protéine ne lie pas à sa séquence, l'enzyme choisie va couper le plasmide aux différents sites de coupure de HinfI sauf si l'un d'entre eux est protégé par la liaison de Fur. La digestion de ce plasmide par *HinfI*, enzyme qui présente 6 sites de coupure, a été utilisée pour estimer la liaison d'EcFur à l'ADN (Figure V.3).

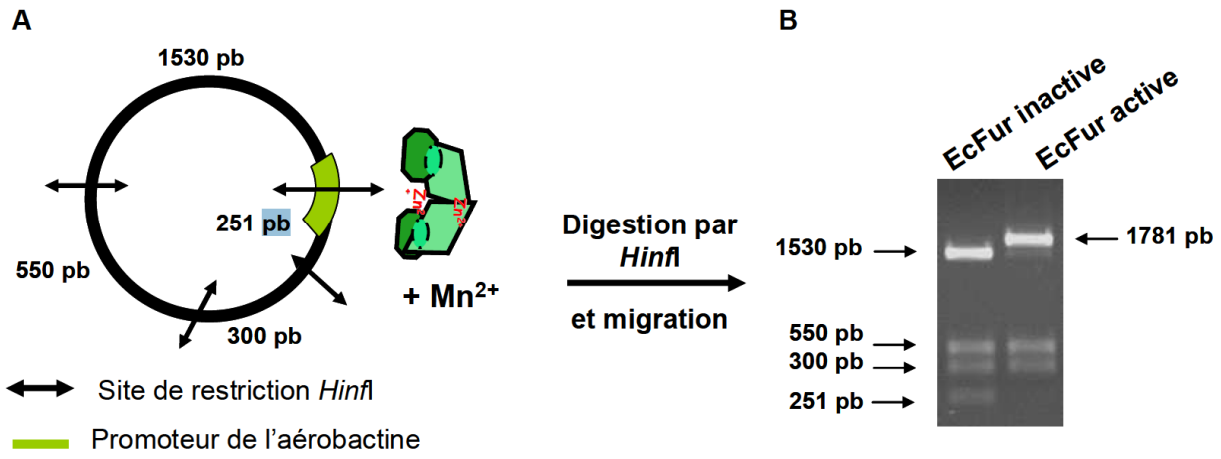


Figure V.3: Principe du test de protection à la coupure par la nucléase HinfI. A : carte des sites de coupure pour *Hinf*I dans le plasmide pDT10. Les tailles attendues des fragments sont indiquées. B : profil de migration sur gel d'agarose du plasmide pDT10 après digestion par *Hinf*I pendant 1 h à 37 °C en présence de protéine EcFur inactive ou activée par le Mn^{2+}..

II. B| Mode opératoire du test

Le mélange réactionnel de 10 µL comprend 1 µM de EcFur dans du tampon 100 mM BisTrisPropane pH7.5, 100 mM KCl, 5mM $MgCl_2$ auxquels sont ajoutés 2 µM de $MnCl_2$. Après 10 min d'incubation à température ambiante, 10 nM de plasmide sont ajoutés. Une incubation de 30 minutes à température ambiante est requise. 1 à 4 unités de *HinfI*, selon l'état d'activité de l'enzyme, sont alors ajoutées et le mélange réactionnel est incubé une heure à 37 °C. 1 µL d'EDTA à 500 mM est ajouté afin de stopper la réaction. Le produit de la digestion est déposé sur un gel d'agarose de0.8% dans du tampon Tris, Acétate, EDTA (TAE) 0,5X contenant du BET ou du Gel red (15 µL de solution concentrée 10 000X dans 50 mL).

Pour étudier la capacité des peptides pF1 à pF5 ainsi que des formes mutées et/ou tronquées à inhiber la liaison d'EcFur à l'ADN, des quantités croissantes de peptides (de 0 à 500 µM) sont ajoutées à la protéine EcFur avant ou après métallation par le $MnCl_2$. Une incubation de 20 min à température ambiante est effectuée. Les étapes successives (ajout ou non de métal, ajout de pDT10 puis digestion enzymatique) sont effectuées comme décrit ci-dessus.

Il est à noter que ce test est assez délicat à cause de sa sensibilité aux traces métalliques en particulier et que des dépôts dans des puits contrôles sont nécessaires dans chaque gel.

II. C| Surproduction et purification des protéines EcFur

La protéine EcFur a été surproduite et purifiée comme précédemment décrit (ref)à partir de cellules BL21(DE3) fraîchement transformées avec le plasmide pET-30c_*fur* (D'autreaux *et al*, 2002). Une pré-culture d'un transformant est faite sur la nuit à 37 °C avant de diluer 100 fois la suspension bactérienne dans 3 L de LB contenant de la Kanamycine à 50 µg/mL.

Après culture des cellules à 37 °C sous agitation pour atteindre une absorbance à600nm de 0.6, la surproduction est induite avec de l'IPTG à 0.5 mM final pendant 3h. Les cellules sont centrifugées à 6000 g pendant 10 minutes à 6°C. Les culots correspondant à 1 L sont resuspendus dans 30 mL de MOPS 100 mM pH 8, EDTA 20 mM, 10 % m/v sucrose, 10 % glycérol v/v contenant trois inhibiteurs de protéases : l'ICT (inhibiteur de la trypsine et de la chymotrypsine) 10 mg/mL, le PMSF à 10 µg/mL et la pepstatine A à 4 µg/mL. La

suspension bactérienne est soniquée, sur un bain de glace, pendant 5 cycles de 30 secondes (pulser 3.5 ; amplitude 75-90 ; Tuner < 25%). Les débris cellulaires sont éliminés par centrifugation à 20000 g pendant 20 minutes à 6 °C. Les protéines du surnageant sont précipitées par « salting-out » en amenant progressivement la concentration en sulfate d'ammonium à 77.5 % La précipitation se déroule sous agitation sur la nuit à 6 °C. Le précipité est récupéré par une centrifugation à 35000 g pendant 30 min à 6 °C et redissous dans 30 mL de Tris/HCl 20 mM pH8 contenant 10% v/v de glycérol. Afin d'éliminer les sels, les extraits protéiques sont dialysés deux fois 2 h dans 2 litres de tampon Tris/HCl 50mM pH 8 contenant 10 % v/v de glycérol. Le matériel insoluble est éliminé par une centrifugation à 35000 g pendant 30 min à 6 °C.

Les extraits protéiques sont chargés à 0.5 mL/min sur une colonne d'affinité Chelating Superose, GE healthcare), préalablement saturée par une solution de ZnSO4 à 20 mM puis lavée avec du Tris/HCl 20 mM pH 8 jusqu'à ce que le pH remonte à 8. La colonne est alors rincée avec différents tampons afin d'éliminer les contaminants :

· 320 mL deTris/HCl 20 mM pH 8

· 180 mL de Tris/HCl 20 mM pH 8, (NH4)2SO4 500 mM

· 120 mL de BisTris 100 mM pH 6

· 200 mL Tris/HCl 20 mM pH 8, glycine 25 mM

La protéine est alors éluée par une solution de α D,L histidine à 50 mM diluée dans du Tris/HCl 20 mM pH 8, glycine 25 mM.

Les fractions contenant la protéine sont regroupées, du glycérol est ajouté de manière à avoir 10 % v/v final. Le tout est traité par l'EDTA pH 8 à une concentration finale de 20 mM pendant 30 minutes sous agitation à 6 °C afin de se débarrasser du Zn2+ élué par l'histidine.

La protéine est alors précipitée par l'ajout de sulfate d'ammonium à 77,5 % sur la nuit à 6 °C sous agitation. Le précipité est récupéré par une centrifugation de 35000 g pendant 30 minutes à 6 °C puis redissous dans 3 à 4 mL de tampon Tris/HCl 100 mM pH 8, KCl 100 mM.

La dernière étape de purification consiste en une colonne de filtration sur gel (Superdex 75 prep grade (HR16/60) GE Healthcare). La colonne est équilibrée avec Tris/HCl pH8 100 mM, KCl 500 mM. La protéine Fur se retrouve dans les pics à 58 mL (dimère) et 63 mL (monomère). Les pics observés avant 50 mL contiennent majoritairement de la protéine Fur sous forme d'oligomères (tétramère, hexamère et octamère) qui sont difficilement utilisables.

Les fractions contenant du dimère ou du monomère purs sont regroupées et concentrées sur Ultrafree (cut-off 5000, Millipore).

Passage du monomère au dimère :

La protéine EcFur monomérique est incubée à 70 µM dans le tampon 20 mM MOPS pH 7.4, 500 mM NaCl avec 1 équivalent de DTT et 1 équivalent de CdSO4 ou ZnSO4 pendant 15 min à 25 °C puis est injectée sur colonne de filtration sur gel (Superdex 75 10/300 GL). Les fractions contenant le dimère sont récupérées et concentrées sur un « Ultrafree » de 15 mL avec un « cutoff » de 10 kDa.

La pureté de la protéine est vérifiée par gel SDS-PAGE à 17 %. La concentration en protéine est déterminée par un dosage en spectroscopie UV-Visible à 277 nm, en utilisant un coefficient d'extinction massique de 0.4 mg/mL. La protéine est stockée à -80°C avec 10% de glycérol. Le rendement de purification est de 15 à 20 mg de protéine pure par litre de culture bactérienne.

II. D| Essais de co-cristallisation

II. D. 1| Dépôt automatique

Les essais de co-cristallisation automatiques réalisés à Tromso (Norvège) ont été effectués par le robot de cristallisation de protéine Phoenix DT (RIGAKU). Cinq librairies de produits chimiques (96 puits) ont été testées :

- UCSG+
- Hampton Cristal Screen (1 et 2)
- Tromso Kanj Tromso-stock ten (locale)
- Tromso Stockpegs eleven
- Wizard 1 et 2

Cela fait 576 conditions pour chaque concentration du complexe EcFur-Mn^{2+}-peptide. Les conditions des essais de co-cristallisation étaient les suivantes :

- Métallation de la protéine 10 mg/mL avec du Mn^{2+} à 1,1 équivalents
- Les peptides pF1(1-10), pF1(2-9) et pF1(3-9) ont été ajoutés à la protéine à une stœchiométrie peptide/dimère de 2:1.
- Les mélanges ont été concentrés à 10mg/mL.

- Un autre échantillon a été réalisé avec le peptide pF1(3-9) peptide/dimère (1:1) équivalent dans lequel a été rajouté 1µL de dithiothréitol (DTT) 500 mM pour réduire d'éventuels ponts disulfures.
-

Après avoir placé de nouvelles plaques de 96 puits, 60 µL de produits chimiques (conditions) sont déposés. Ensuite 0,2 µL d'échantillons des complexes Fur-peptides sont ajoutés aux produits chimiques. Les plaques sont recouvertes de façon hermétique d'un film plastique de manière à ne laisser aucun espace entre les puits. Les plaques sont centrifugées et observées au microscope pour vérifier si les gouttes sont bien déposées et ordonnées. Les plaques sont ensuite rangées dans le robot incubateur Alchemist II relié à un ordinateur équipé du logiciel de visualisation CrystalTrak version V2_1.2.

II. D. 2| Optimisation manuelle de EcFur-Mn^{2+}-pF1(1-10)

Comme il a avait été observé un début de cristallisation de ce complexe dans du Dioxane à 35% à l'EMBL, nous avons entrepris d'optimiser manuellement les conditions de ces essais.

Le Dioxane (Fluka Mw : 88,11 g/mol) a été dilué pour obtenir 30, 32, 34, 36 et 38% dans l'eau ultra pure (Volume finale = 500 µL) et déposé en duplicata sur une plaque 6X4 puits. On rajoute dans chaque puits 0,8 µL du complexe EcFur-Mn^{2+}-pF1(1-10) (1 : 1,1 : 1) concentré à 10 mg/mL. La plaque est soigneusement recouverte et déposée dans l'incubateur.

III| Amarrage moléculaire

III. A|L'amarrage automatique avec Glide

Glide est un programme d'amarrage automatique de la société Schrödinger disponible sous licence. Il utilise une méthodologie d'amarrage basée sur une grille comme AutoDock avec une évaluation des interactions par une fonction énergétique. Glide a été conçu pour effectuer une recherche exhaustive de la position, de l'orientation et de la conformation du ligand avec une grande précision dans un temps très raisonnable. La durée de prédiction de l'amarrage d'une molécule ayant moins de 10 angles flexibles est de moins d'une minute sur un processeur Linux AMD Athlon MP 1800+. Les affinités de liaison prédites sont comparables à celles des données cristallographiques de complexes avec un « RMSD » moyen de 2,3 kcal/mol {Friesner, 2004 #54}.

.

III. A. 1| Méthodologie

Glide utilise une série de filtres hiérarchiques pour la recherche des positions des ligands dans le site actif de la protéine Figure V.3.

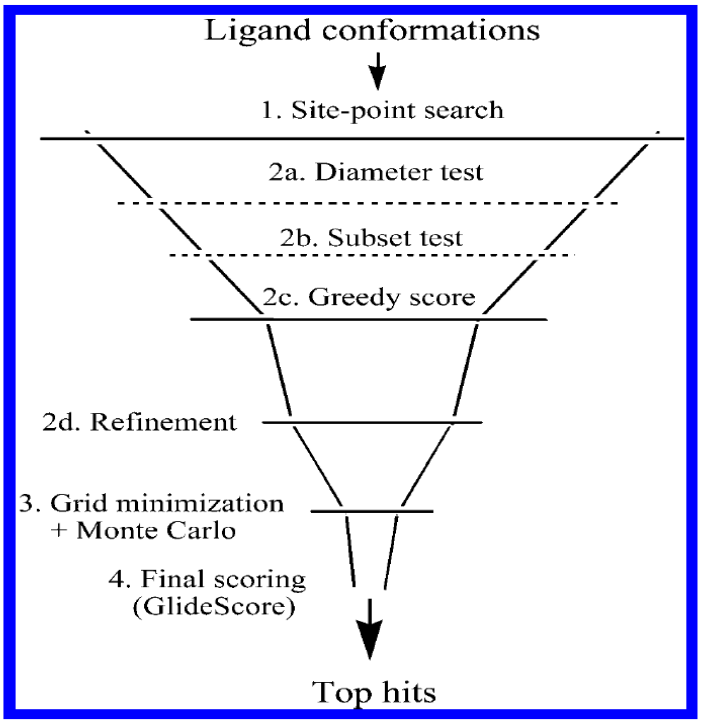


Figure V.3: Processus d'amarrage de Glide {Friesner, 2004 #54}

La forme et les propriétés du récepteur sont représentées par une grille avec différents champs de force. Ces champs sont générés comme une étape de prétraitement du récepteur qui est exécutée une seule fois pour chaque récepteur. La prochaine étape produit un ensemble de conformations du ligand initial. Ces conformations sont sélectionnées à partir d'une évaluation des *minima* du terme d'énergie des angles de torsion du ligand. Un premier criblage est réalisé sur ces conformations pour sélectionner les « poses » prometteuses du ligand. A partir de la position du ligand issu de cette présélection, Glide réduit l'espace de recherche du ligand pour minimiser le temps de calcul extrêmement long. A partir de la « pose » choisie par le criblage initial, l'énergie du ligand est minimisée dans le champ du récepteur en utilisant le champ de force de mécanique classique OPLS-AA en conjonction avec un modèle diélectrique dépendant de la distance. Enfin, trois à six des « poses » de plus basses énergies sont soumises à une procédure de Monte Carlo utilisée pour minimiser l'énergie des angles de torsion. Cette procédure permet aussi d'orienter correctement certains groupes périphériques.

III. A. 2| La fonction de score

Glide utilise deux fonctions de score pour évaluer l'interaction des ligands avec le récepteur : la précision standard: « SD » et l'extra précision: « XP » à choisir selon le temps de calcul et la précision désirés. La fonction SD est présentée ci-dessous {Friesner, 2004 #54} :

$$\begin{aligned}\Delta G_{\text{bind}} = {} & C_{\text{lipo-lipo}} \sum f(r_{\text{lr}}) + \\ & C_{\text{hbond-neut-neut}} \sum g(\Delta r)\ h(\Delta\alpha) + \\ & C_{\text{hbond-neut-charged}} \sum g(\Delta r)\ h(\Delta\alpha) + \\ & C_{\text{hbond-charged-charged}} \sum g(\Delta r)\ h(\Delta\alpha) + \\ & C_{\text{max-metal-ion}} \sum f(r_{\text{lm}}) + C_{\text{rotb}} H_{\text{rotb}} + \\ & C_{\text{polar-phob}} V_{\text{polar-phob}} + C_{\text{coul}} E_{\text{coul}} + \\ & C_{\text{vdW}} E_{\text{vdW}} + \text{solvation terms}\end{aligned}$$

La fonction de score de Glide est basée sur ChemScore {Eldridge, 1997 #158} tout en incluant de nombreuses modifications. Le terme lipophile-lipophile $C_{\text{lipo–lipo}}$, par exemple, est tiré de ChemScore. Le terme de liaison hydrogène C_{hbond} est également du type Chem-Score, mais est séparé en différentes composants pondérées qui varient selon que l'on traite un

donneur ou un accepteur d'hydrogène. Le terme d'interaction métal-ligand $C_{max\text{-}metal\text{–}ion}$ utilise la même fonctionnelle que dans ChemScore mais avec quelques variantes. Par exemple, ce terme ne considère que les interactions avec les atomes accepteurs anioniques (tels que, par exemple, les deux oxygènes d'un groupe carboxylate). Cette modification dans Glide permettrait de reconnaître la préférence des ligands anioniques pour le centre métallique dans les métalloprotéases. Le septième terme, issu de la cartographie du site actif favorise un atome polaire, mais non impliqué dans une liaison hydrogène (classification de ChemScore) qui se retrouve dans un environnement hydrophobe.

La deuxième composante majeure de cette fonction $C_{coul}E_{coul} + C_{vdW}E_{vdW}$ représente les énergies d'interactions de Coulomb et de vdW entre le ligand et le récepteur.

Les termes de solvatation utilisent une composante explicite où les molécules d'eau sont en compétition avec les « poses » du ligand et une composante d'énergie empirique qui mesure l'exposition des différents groupes au solvant.

Après l'amarrage, Glide peut classer les « poses » du ligand en fonction de la valeur de l'énergie SD ou en fonction du RMSD. La méthode de préparation du ligand et de la protéine a déjà été décrite dans le chapitre III.

III. B| Méthode de paramétrisation des petites molécules pour CHARMM

La paramétrisation d'une nouvelle molécule dans un champ de force de mécanique moléculaire tel que CHARMM est un processus complexe et coûteux en temps de calcul. Nous avons utilisé une méthode simplifiée dans ce travail car la molécule paramétrée était proche d'un résidu tryptophane et donc le transfert de la plupart des paramètres était possible. Nous avons utilisé le programme Gaussian (G03) pour optimiser la géométrie de la molécule au niveau de théorie MP2 avec des bases 6-31G* pour tous les atomes. Une charge totale de 0 et une multiplicité de spin de 1 ont été utilisées. Les charges partielles de tous les atomes ont été calculées sur la molécule optimisée par la méthode de Merz-Kollman basée sur l'ajustement du potentiel électrostatique. Nous avons ensuite confié à VMD et son module « Molefacture » le soin de créer un fichier de structure (*.psf*) contenant la connectivité de la molécule. Les paramètres liés et les paramètres de van der Waals ont été extraits d'atomes équivalents dans la topologie CHARMM (principalement du résidu tryptophane).

III. C| Matériels informatiques

Nous disposons dans le groupe de Modélisation et Chimie Théorique dans lequel j'ai travaillé de 5 stations de travail de type mono- ou biprocesseur quadri-cœur Xeon muni d'une dizaine de Go de mémoire vive, 200 Go de disque dur pour le stockage des données et de beaux écrans 22 à 24 pouces pour la visualisation moléculaire. A ces stations, s'ajoute un « cluster » de 5 stations biprocesseur quadri-cœur reliées par réseau Gigabit-Ethernet permettant d'effectuer des calculs parallèles.

J'avais également accès au « cluster » de 128 processeurs de la Direction des Sciences du Vivant du CEA sur lequel est installée la suite des logiciels Schrödinger tels que Glide.

La majorité de mes calculs (AutoDock, CHARMM) ont été effectués sur ma station de travail dédiée en environnement Linux ou sur les autres stations du groupe. Toutes les stations et les 2 clusters sont accessibles sur le réseau interne du CEA via le protocole « ssh ».
Enfin, j'ai utilisé pendant quelques jours les ressources de calcul de l'Université de Norvège à Tromso où était implémenté le logiciel Gold.

Chapitre VI :

Annexes

CHAPITRE VI : ..... **- 273 -**

ANNEXES ..... **- 273 -**

I| LE PROGRAMME AUTODOCK4.2 ..... - 277 -

I. A| Préparation des fichiers ..... - 277 -

I. B| Calcul des grilles de potentiels ..... - 279 -

I. C| Recherche des solutions d'amarrage ..... - 279 -

I. D| Analyse des résultats ..... - 281 -

II| SYNTHESE PEPTIDIQUE ..... - 283 -

II. A| Liste des acides aminés ..... - 283 -

II. B| Les structures peptidiques ..... - 285 -

II. C| Protection des groupes fonctionnels des peptides ..... - 286 -

I| Le programme AutoDock4.2

AutoDock est constitué d'un ensemble de programmes et d'utilitaires exécutés en ligne de commandes ou sur une interface graphique sous environnement Unix. Cette section permet une prise en main rapide du logiciel. Elle présente la séquence de commandes à exécuter pour effectuer un calcul d'amarrage. Ce guide à l'usage des utilisateurs débutant avec AUTODOCK, permet de lancer rapidement les premiers calculs à partir des fichiers bruts du ligand et du récepteur. Des indications sur la nature et l'influence des différents outils et paramètres jalonnent la procédure ainsi que quelques recommandations.

I. A| Préparation des fichiers

La macromolécule doit être représentée avec ses hydrogènes polaires ainsi que les charges partielles de tous ses atomes. Cette opération se fait avec le module *AUTOTORS* du programme. A partir d'une structure PDB ou d'une construction de modèles, il faut ajouter les atomes d'hydrogènes dits « essentiels » c'est à dire polaires. AutoDockTools (ADT) crée un fichier *.pdbqt* qui contient les informations sur les charges partielles, les atomes d'hydrogènes polaires, les pivots et le nombre d'angles de torsions défini. Un exemple de fichier *.pdbqt* est donné ci-dessous en exemple.

Sample PDBQT file

```
REMARK  4 active torsions:
REMARK  status: ('A' for Active; 'I' for Inactive)
REMARK    1  A    between atoms: N_1  and  CA_5
REMARK    2  A    between atoms: CA_5  and  CB_6
REMARK    3  A    between atoms: CA_5  and  C_13
REMARK    4  A    between atoms: CB_6  and  CG_7
ROOT
ATOM      1  CA  PHE A   1      25.412  19.595  12.578  1.00 12.96     0.287 C
ENDROOT
BRANCH   1   2
ATOM      2  N   PHE A   1      25.225  18.394  13.381  1.00 13.04    -0.065 N
ATOM      3  HN3 PHE A   1      25.856  17.643  13.100  1.00  0.00     0.275 HD
ATOM      4  HN2 PHE A   1      25.558  18.517  14.337  1.00  0.00     0.275 HD
ATOM      5  HN1 PHE A   1      24.247  18.105  13.350  1.00  0.00     0.275 HD
ENDBRANCH   1   2
BRANCH   1   6
ATOM      6  CB  PHE A   1      26.873  20.027  12.625  1.00 12.45     0.082 C
BRANCH   6   7
ATOM      7  CG  PHE A   1      27.286  20.629  13.923  1.00 12.96    -0.056 A
ATOM      8  CD2 PHE A   1      27.470  22.001  14.050  1.00 12.47     0.007 A
ATOM      9  CE2 PHE A   1      27.877  22.571  15.265  1.00 13.98     0.001 A
ATOM     10  CZ  PHE A   1      28.108  21.754  16.360  1.00 13.84     0.000 A
ATOM     11  CE1 PHE A   1      27.919  20.380  16.242  1.00 13.77     0.001 A
ATOM     12  CD1 PHE A   1      27.525  19.821  15.027  1.00 11.32     0.007 A
ENDBRANCH   6   7
ENDBRANCH   1   6
BRANCH   1  13
ATOM     13  C   PHE A   1      25.015  19.417  11.141  1.00 13.31     0.204 C
ATOM     14  O2  PHE A   1      24.659  20.534  10.507  1.00 12.12    -0.646 OA
ATOM     15  O1  PHE A   1      25.024  18.283  10.608  1.00 13.49    -0.646 OA
ENDBRANCH   1  13
TORSDOF 4
```

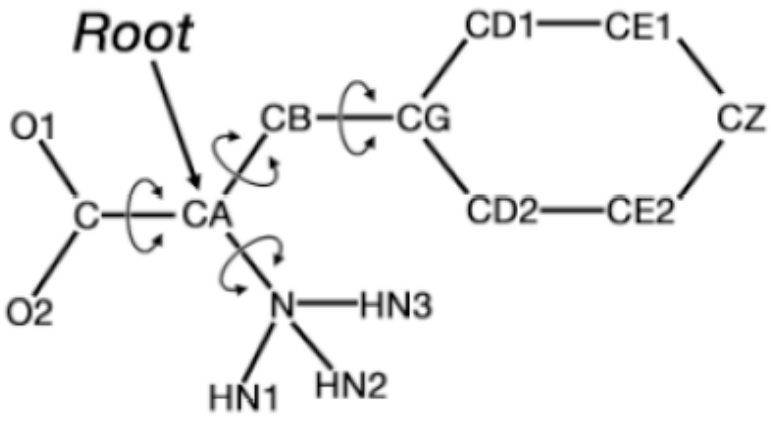


Figure VI.1: Exemple d'un fichier *.pdbqt* d'AutoDock

I. B| Calcul des grilles de potentiels

Les grilles de potentiels sont calculées par le programme AUTOGRID selon les instructions du fichier de paramétrage qu'on lui fournit. Le fichier de paramétrage est un fichier texte dont l'extension est *.gpf* (*Grid Parameters File*).

Les paramètres, codés par des mots clé, indiquent :

· Les noms de fichiers du ligand et du recepteur, aux formats *.pdbqt*

· Le centre et les dimensions de la boîte ainsi que l'espacement entre les points de la grille respectivement (choix d'une coordonnée d'un atome, 40, 40 et 40 points en x, y et z et 0,375 Å d'espacement) par défaut;

· Les types d'atomes présents dans le ligand et pour lesquels il faut calculer une grille ;

· Les termes des constantes à calculer.

Chaque grille calculée est enregistrée dans un fichier dont l'extension est *map*. Le fichier de paramètres doit être généré simultanément à partir des fichiers du ligand et du récepteur. La boîte est alors centrée sur le ligand et ses dimensions sont proportionnelles à la taille de ce ligand. Par conséquent, il est nécessaire d'éditer ce fichier généré pour modifier les paramètres de position et de dimension de la boîte afin de l'ajuster au mieux à la région du récepteur que l'on étudie. Les grilles ainsi calculées peuvent être réutilisées pour tout ligand ne comportant pas de nouveau type d'atome et devant explorer la même région de la macromolécule.

I. C| Recherche des solutions d'amarrage

C'est le programme AUTODOCK qui va faire cette recherche en fonction des paramètres qu'on lui transmet par l'intermédiaire d'un fichier *.dpf* (*Dock Parameters File*).

Les paramètres d'un calcul d'AUTODOCK sont :

· Les noms des fichiers contenant le ligand et les grilles de potentiels à utiliser ;

· L'état initial du ligand (position, orientation et conformation aléatoire ou précisée) ;

· La méthode de recherche à utiliser (RSMC, AG ou AGL) avec les paramètres associés et les paramètres de l'algorithme d'optimisation locale pour la méthode AGL.

Les paramètres de l'algorithme génétique sont :

· La taille de la population. La valeur par défaut est de 150 mais peut être modifiée en fonction du nombre de degrés de liberté et de l'étendue de la boîte à explorer.

· Le nombre maximum d'évaluations d'énergie, 2 500 000 par défaut. C'est un paramètre d'arrêt de l'amarrage utile lorsque le processus met trop de temps, ou ne parvient pas, à converger.

· Le nombre maximum de générations qui est un autre paramètre d'arrêt, 2700 par défaut. Ces deux derniers paramètres sont à adapter conjointement en fonction de la taille de la population, du nombre de degrés de liberté, de l'étendue de la boîte à explorer et du degré de convergence à atteindre. Le degré de convergence est corrélé au nombre de solutions distinctes, ou classes, que l'on veut obtenir.

· Le niveau de sélection élitiste : le nombre des meilleurs individus qui survivent automatiquement à la génération suivante. La valeur standard est 1.

· Le taux de croisements. Le croisement est l'opérateur d'exploration globale, sa probabilité doit être élevée pour que l'AG ait ce caractère global. La valeur standard est 0,8.

· Le taux de mutations. Cet opérateur joue un rôle d'optimisateur local pour les AGs purs (darwiniens) en opérant des variations de faible amplitude sur les gènes, chose que le croisement fait très difficilement. Avec l'opérateur d'optimisation locale explicite des AGL, ce rôle devient inutile et la mutation sert seulement à réintroduire des allèles éliminés par la sélection. Sa probabilité doit donc être faible dans un AGL. La distribution de Cauchy donne un compromis entre des variations radicales et une exploration trop détaillée de la topographie conformationnelle. La valeur standard est 0,02.

· Le nombre de générations précédentes N dans lesquelles on recherche l'individu le plus mal adapté pour calculer le facteur de sélection proportionnelle. Si cette valeur est trop grande, les plus anciennes générations qui sont peu optimisées, vont perturber la sélection en gommant les écarts entre les nouvelles solutions. Les meilleurs individus de la génération en cours n'auront pas beaucoup plus d'enfants que les moins bons individus et la population convergera très lentement. Si N est trop petit, la population convergera rapidement vers le minimum local le plus proche. On peut faire l'analogie entre ce paramètre et à un facteur d'échelle : si on est trop près du relief énergétique on n'en voit qu'une partie et on trouve une solution locale, si on est trop loin on ne distingue plus le relief. La valeur optimale est de 10 générations.

· Les valeurs de la moyenne et la variance de la distribution de Cauchy pour la mutation des gènes. Les valeurs standard sont 0 et 1 respectivement.

· Le nombre de cycles « Run » à exécuter c'est-à-dire le nombre de solutions souhaitées.

Dans le cadre d'une recherche par AGL, les paramètres de l'optimisateur local (Solis-Wets) sont :

· Le nombre maximal d'itérations d'optimisation, 300 pour une petite molécule organique ;

· Le nombre de succès consécutifs de l'optimisation avant de modifier la taille du pas de l'opérateur. La valeur standard est 4.

· Le nombre d'échecs consécutifs de l'optimisation avant de modifier la taille du pas de l'opérateur. La valeur standard est 4.

· La taille du pas de l'opérateur sur les gènes de position : entre 0,2 Å et 1 Å.

· La taille du pas de l'opérateur sur les gènes d'orientation : entre 5° et 50°. Les multiples de 360 sont à écarter si l'on veut éviter de re-tester périodiquement les mêmes valeurs d'angle. Il est également souhaitable que ce ne pas soit un réel ou un entier grand, supérieur à 36, pour assurer l'exploration fine de ces gènes. Les pas de translation et de rotation sont des pas de recherche locale.

· La taille initiale de « l'espace de recherche locale ». C'est le facteur qui est doublé ou divisé par deux après quatre optimisations consécutives de même tendance et qui pondère les pas de recherche locale. Sa valeur initiale est de 1 et ne devrait pas être modifiée par l'utilisateur.

· La taille limite inférieure de l'espace de recherche locale. C'est un paramètre d'arrêt de l'optimisation locale. Il indique la « finesse » d'optimisation en dessous de laquelle on ne descend pas. La valeur par défaut est de 0,01. A titre d'exemple, il faut au minimum 4x7 itérations d'optimisation pour que la taille de l'espace de recherche passe de 1 à 0,01.

· Le nombre maximum de cycles d'optimisation.

Un cycle d'amarrage s'arrête soit :

· Lorsque le nombre maximal de générations est atteint ;

· Lorsque le nombre maximal d'évaluation d'énergie est atteint ;

· Lorsque la population a convergé vers une unique solution.

I. D| Analyse des résultats

A la fin de chaque cycle d'exploration, AutoDock enregistre la meilleure solution d'amarrage du cycle c'est à dire le phénotype de l'individu de la dernière génération ayant la

meilleure interaction avec le récepteur. Le fichier de résultats contiendra donc autant de solutions que de cycles exécutés. Ce fichier porte l'extension *.dlg* et peut être visualisé à l'aide du programme d'interface ATD. AutoDock peut faire une première analyse des résultats en regroupant les solutions en classes (clusters) en fonction de la racine de la moyenne des carrés des écarts « *Root Mean Square Deviation* » (RMSD) des coordonnées atomiques. Si le RMSD entre deux molécules est inférieure à 2 Å (par défaut), ces deux solutions sont dans la même classe.

La meilleure solution est celle de plus basse énergie et une convergence de toutes les solutions vers la même classe paraît souhaitable. Malheureusement, un ligand peut adopter différents modes de liaison avec le récepteur et le modèle de l'énergie n'est pas parfait. Il est donc intéressant d'obtenir plusieurs classes sur une gamme d'une dizaine de kcal/mol. Elles peuvent révéler les différents modes de liaison d'un ligand à un récepteur. Le nombre de classes obtenu dépend du degré de convergence de la recherche et c'est en jouant sur le nombre d'évaluations de l'énergie et le nombre de générations qu'on module le nombre de classes obtenues.

II| Synthèse peptidique

Voici quelques informations de base sur les propriétés des acides aminés qui sont tirées du cours de Master 2 de Chimie Bio-Organique et Bio-Inorganique (CBOBI) réalisé en 2008 à l'université J. Fourier.

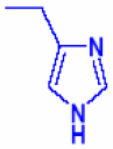
II. A| Liste des acides aminés

- Hydrophobes :

R	Nom	Codes		Propriétés
H	Glycine	Gly	G	*flexible*
CH_3	Alanine	Ala	A	*hydrophobe*
	Valine	Val	V	*hydrophobe*
	Isoleucine	Ile	I	*hydrophobe*
	Leucine	Leu	L	*hydrophobe*
N H COOH	Proline	Pro	P	*cyclique, hydrophobe* casse les *helices-α*
$-CH_2-$	Phénylalanine	Phe	F	*hydrophobe*
$-CH_2-$ OH	Tyrosine	Tyr	Y	*hydrophobe* *Liaisons H*
$-CH_2-$ NH	Tryptophane	Trp	W	*hydrophobe* *Liaisons H*

Hydrophobe → Intérieur de la protéine

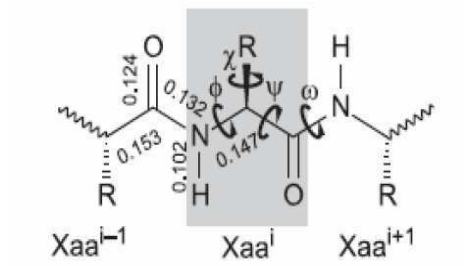
- Polaires

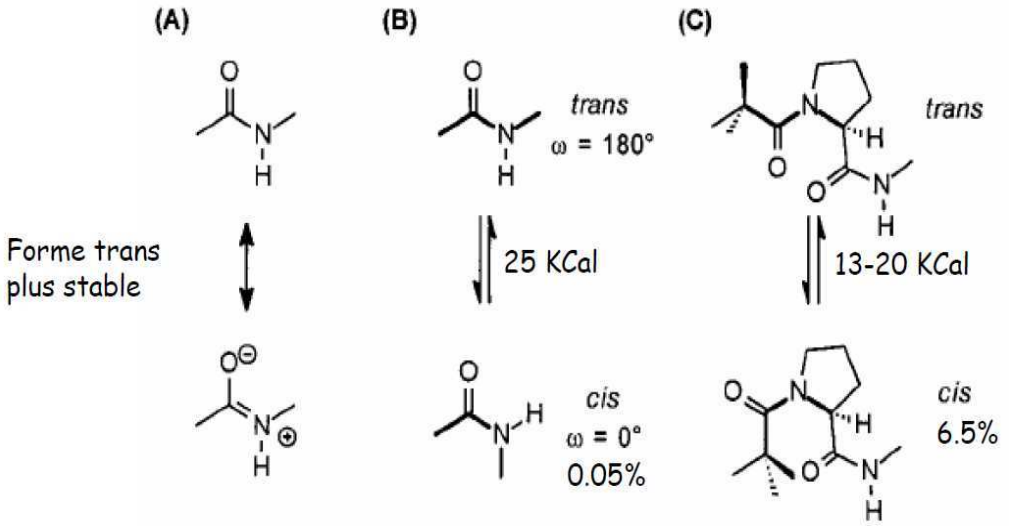
R	Name	Codes		Properties
CH_2OH	Sérine	Ser	S	*Liaisons H*
$CH(CH_3)-CH_3$	Thréonine	Thr	T	*Liaisons H*
CH_2SH	Cystéine	Cys	C	*ponts disulfure*
$CH_2-CH_2-CH_2-S-CH_3$	Methionine	Met	M	*hydrophobe*
CH_2-CONH_2	Asparagine	Asn	N	*Liaisons H*
$CH_2-CH_2-CONH_2$	Glutamine	Gln	Q	*Liaisons H*
CH_2-COOH	Acide aspartique	Asp	D	*Liaisons ioniques, Liaisons H*
CH_2-CH_2-COOH	Acide glutamique	Glu	E	*Liaisons ioniques, Liaisons H*
$CH_2-CH_2-CH_2-CH_2-NH_3^+$	Lysine	Lys	K	*Liaisons ioniques, Liaisons H*
CH2-CH2-CH2-NH(NH2)=NH	Arginine	**Arg**	**R**	*Liaisons ioniques, Liaisons H*
	Histidine	His	H	*Liaisons H*

Polaire ➡ Extérieur de la protéine

II. B| Les structures peptidiques

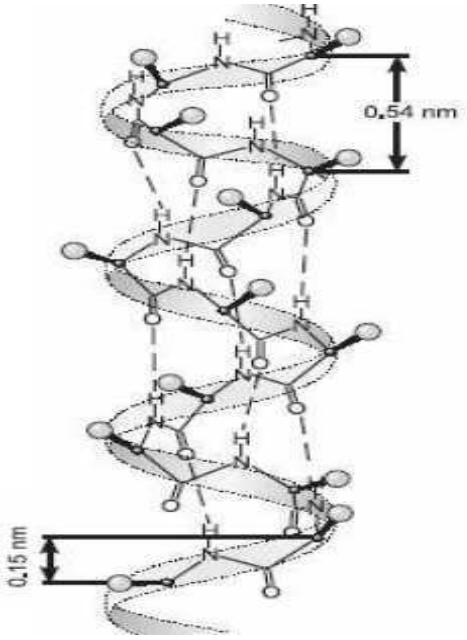
- Les caractéristiques de la liaison peptidique.





- Les structures secondaires

- Les hélices



Hélice α : 3.6 a.a./tours
φ=57°, ψ=47°

- Les feuilles β

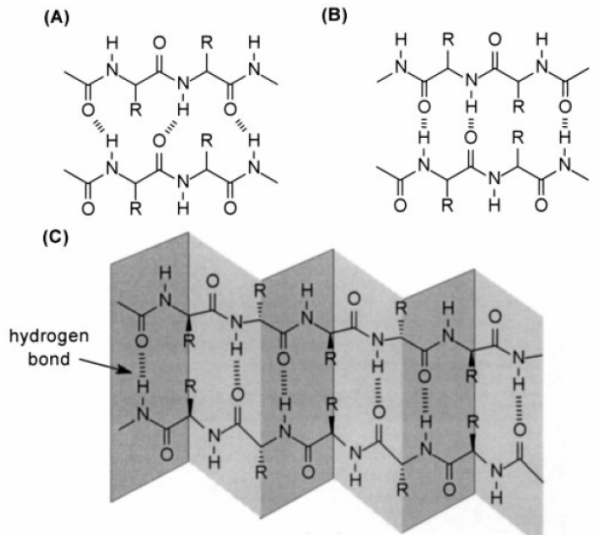


Figure VI.2: (A) Feuillets β parallèles, et (B) (C) feuillets β anti-parallèles

- Les coudes

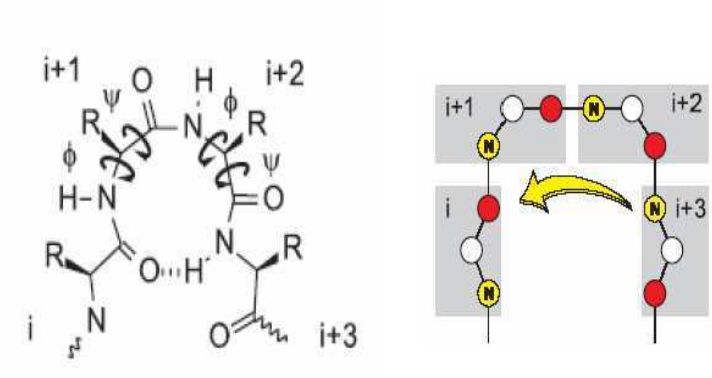


Figure VI.3: Coude β (4 acides aminés impliqués)

II. C| Protection des groupes fonctionnels des peptides

La protection des chaines latérales des peptides permet d'éviter des réactions secondaires.

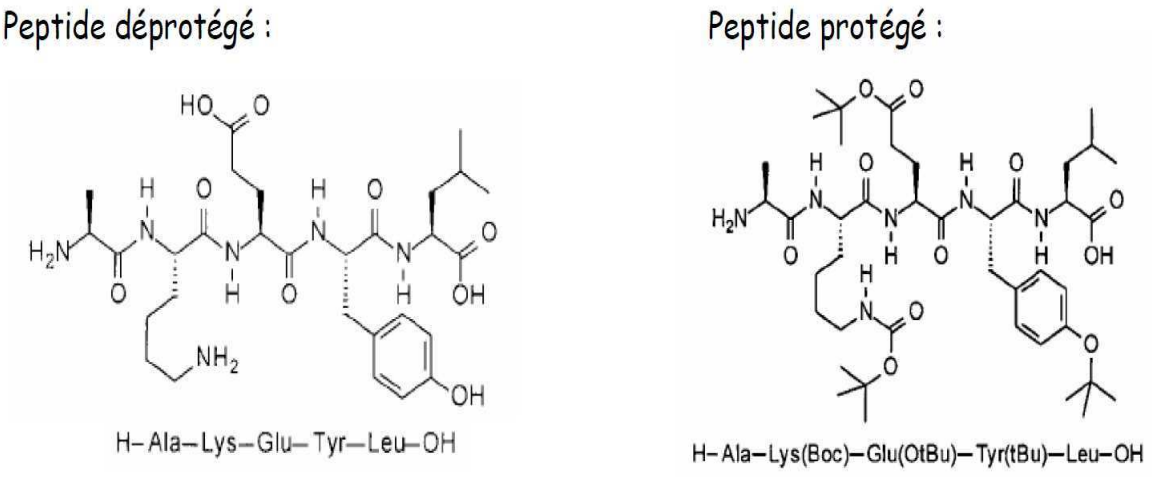
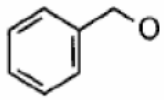
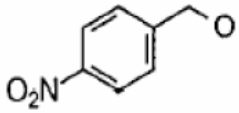
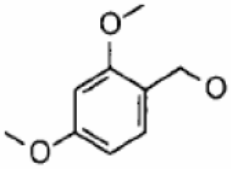
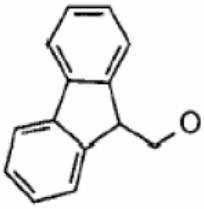


Figure VI.4: Exemple de peptide protégé (à droite) et déprotégé (à gauche)

- Protection de la fonction amine en α: Lien du type carbamate : Y-O-CO-NH-R

Group	*Symbol*	*Y-O*	*Cleavage conditions*
Benzyloxycarbonyl	Z		H_2/Pd, HBr/AcOH, Na/liq. NH_3
4-Methoxybenzyloxy-carbonyl	Z(OMe)		TFA; H_2/Pd; Na/liq. NH_3
2-Nitrobenzyloxy-carbonyl	Z(2-NO_2)		H_2/Pd (more labile than Z) HBr/AcOH (more stable than Z)
tert-Butoxycarbonyl	Boc		TFA, TFA/CH_2Cl_2, HCl/organic solvent
Fluorenyl-9-methoxy-carbonyl	Fmoc		Piperidine, DBU, 2-aminoethanol, Morpholine, Liq. NH_3
Allyloxycarbonyl	Aloc		Pd^0/nucleophile

- Protection de la fonction carboxy en α: Lien du type ester

Ethyl	Et		Alkaline or enzymatic hydrolysis
Benzyl	Bzl		H_2/Pd, HBr/AcOH, Liq. HF, Alkaline hydrolysis
4-Nitrobenzyl	Nbz		H_2/Pd, alkaline hydrolysis, stable to HBr/AcOH
2,4-Dimethoxybenzyl	2,4-Dmb		1% TFA/CH_2Cl_2
tert-Butyl	tBu		TFA, HCl/AcOH, HBr/AcOH, $BF_3 \cdot OEt_2$/AcOH
9-Fluorenylmethyl	Fm		Piperidine
Allyl	Al		Pd(0), nucleophile

Résumé:

Fur (Ferric Uptake Regulator) est un régulateur transcriptionnel spécifique des bactéries qui intervient dans le contrôle de l'homéostasie du fer, ce qui en fait une cible antibactérienne intéressante.

Avant mon arrivée au laboratoire, quatre inhibiteurs interagissant spécifiquement avec Fur avaient été isolés. La partie active de ces inhibiteurs consiste en des peptides de 13 a.a.

Au cours de cette thèse, j'ai utilisé une double-approche : théorique et expérimentale pour étudier l'interaction de ces peptides avec Fur afin de comprendre le mécanisme d'inhibition. J'ai synthétisé plusieurs séquences peptidiques, montré par des tests biochimiques que certaines inhibaient Fur et déterminé les interactions importantes à l'activité inhibitrice. J'ai obtenu des modèles théoriques des complexes Fur/peptides par amarrage moléculaire, cohérents avec les résultats expérimentaux, qui ont mis en évidence une zone d'inhibition de Fur. Des criblages *in silico* dans cette zone ont permis de sélectionner de petites molécules, inhibitrices potentielles de Fur et donc intéressantes pour des applications thérapeutiques.

Abstract:

Fur (Ferric Uptake Regulator) is a transcriptional regulator involved in the control of iron homeostasis. Specific to bacteria, Fur is an attractive antibacterial target.

Before my arrival in the laboratory, four inhibitors interacting specifically with Fur had been isolated. The active part of these inhibitors consists of peptides of 13 aa.

In this thesis I have used both theoretical and experimental approaches to study interactions of these peptides with Fur in order to understand the inhibition mechanism. I have synthesized several peptide sequences, shown through biochemical assays that some of them could inhibit Fur and I have identified residues important to the inhibitory activity. I've obtained theoretical models of Fur/peptide complexes consistent with experimental results, which reveal an inhibition pocket in Fur. Small molecules have then been selected though *In silico* screening of this pocket, that could potentially inhibit Fur, and thus be interesting for therapeutic applications.