Institutional Repository - Research Portal
Dépôt Institutionnel - Portail de la Recherche
researchportal.unamur.be
University of Namur

THESIS / THÈSE

MASTER EN SCIENCES BIOLOGIQUES DES ORGANISMES ET ÉCOLOGIE

Caractérisation structurale de fragments Fab et de peptides mimétiques de la chaîne O du LPS de Brucella sp

Albertani, Benjamin

Award date:
1999

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 02. May. 2026

FACULTES UNIVERSITAIRES NOTRE-DAME DE LA PAIX
NAMUR

Faculté des Sciences

Caractérisation structurale de fragments Fab et de peptides mimétiques de la chaîne O du LPS de *Brucella sp.*

Mémoire présenté pour l'obtention du grade de licencié en Sciences biologiques

Benjamin ALBERTANI

Juin 1999

Facultés Universitaires Notre-Dame de la Paix
FACULTE DES SCIENCES
Secrétariat du Département de Biologie
Rue de Bruxelles 61 - 5000 NAMUR
Téléphone: + 32(0)81.72.44.18 - Téléfax: + 32(0)81.72.44.20
E-mail: joelle.jonet@fundp.ac.be - http://www.fundp.ac.be/fundp.html

Caractérisation structurale de fragments Fab et de peptides mimétiques de la chaîne O du LPS de *Brucella sp.:*

Benjamin ALBERTANI

Résumé

Une série de peptides mimotopes du LPS de *Brucella* ont été isolés au sein du Laboratoire d'Immunologie et de Microbiologie. Afin de comprendre les bases structurales du mode de liaison de peptides mimotopes à un anticorps monoclonal, une étude structurale par modélisation et par cristallographie des peptides ainsi que des anticorps a été réalisée. Nous nous sommes focalisés sur 4 peptides, reconnus par l'anticorps 6B3 (spécifique d'un épitope du LPS M-dominant de *Brucella sp.*), dont les séquences permettent de dégager un consensus. Ces peptides ont été modélisés par dynamique et mécanique moléculaire : 20 conformères par peptide, potentiellement représentatifs d'une série de conformations possibles ont été engendrés. Afin de mettre en évidence une conformation qui pourrait être celle qui est reconnue par l'anticorps 6B3, nous avons tenté de trouver des conformations communes entre ces peptides. Pour ce faire, nous avons superposé l'ensemble des conformères de chaque peptide et nous les avons comparés sur base du squelette carboné. D'autre part, ces peptides fusionnés à une protéine "porteuse" ainsi qu'un fragment Fab d'un autre anticorps utilisé pour l'isolement de peptides mimotopes (4F9) ont été cristallisés. Bien que les cristaux soient inutilisables pour la diffraction par rayons X du fait de leur petite taille et de leur caractère polycristallin, il est possible d'affiner les conditions de cristallisation. Ces approches préliminaires nous ont permis de définir des stratégies prometteuses. Les résultats obtenus par modélisation et par cristallographie pourront dès lors être utilisés pour l'amélioration de mimotopes peptidiques.

Mémoire de licence en Sciences biologiques
Juin 1999
Promoteur: E. Depiereux

Remerciements

Je tiens tout d'abord à remercier le Professeur E.Depiereux pour m'avoir acceuilli au sein de son laboratoire.
Mes remerciements vont également au Professeur F.Durand ainsi qu'à tout le laboratoire de C.M.S. pour m'avoir ouvert les portes de leur laboratoire et de l'univers mystérieux de la cristallographie.

Je tiens à exprimer toute ma gratitude au Professeur JJ. Letesson pour ses encouragements et son aide précieuse.

Merci à toi Johan, pour tes conseils, ta motivation et ton efficacité dans les moments difficiles.

Je remercie Mr .VanCutsem, Damien Devos, Nathalie Dumont, et Johan Wouters pour avoir accepté de lire ce travail.

Un grand merci à Xavier et Poos pour leurs remarques pertinentes, leur disponibilité et également pour leurs anticorps et leurs peptides (!).

Toute ma gratitude à Bernard et Katalyn, disponibles à toutes heures, responsables moralement et physiquement de la présence de ce travail entre vos mains.... que dire de plus ?

Je remercie Thierry pour ses prises de « kwokosin-caille » (un sport de combat utilisé pour la chasse à la caille en Afrique du sud-ouest), ses conseil culinaires ainsi que pour la bonne humeur imperturbable qu'il arbore (même en manque de nourriture ! !...).

Merci aux autres mémorants BMS et URBM : Jean-Marc, Mitch, Géro, Emilie pour leur entraide et le visages aussi terrifié que le mien qu'ils montraient lors de ces cruelles mais ô combien bénéfiques « présentations URBM »...

Merci aux « autres », ceux qui n'étaient pas là pour la dernière ligne droite mais avec qui j'ai passé tant de bons moments pendant ces quatre années, il s se reconnaîtront...

Un grand merci à JYP et Steph' ainsi qu'à tous les membres du Laboratoire de BMS et d'Immuno pour leur acceuil et la sympathie qu'ils ont eu à mon égard.

Je remercie mes parents et toute ma famille, sans qui je n'aurais pas pu aller jusqu'au bout ... Merci , Merci , .encore Merci, maman, papa, Catherine, nonna, et tous les autres sans oublier zia Olga...

Enfin, je te remercie Sylviane, pour beaucoup de choses que tu sais déjà... mais surtout pour tes encouragements et ton sourire perpétuel..

J'allais oublier ! !
Merci Professeur Tournesol,... (figure 0)

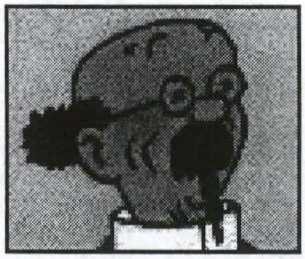


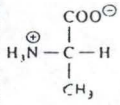
Figure 0 :Tryphon Tournesol ...

Table des abréviations

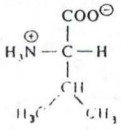
Ac	Anticorps
ADN	Acide désoxyribonucléique
ADNc	ADN complémentaire
AMBER	Homan's Forcefield Polysaccharide
BLAST	Basic Local Alignment Search Tool
BSA	Albumine de serum bovin
CDR	Région de complémentarité déterminante
CFF91	ClassII ForceField
CPU	Unité centrale du processeur
CSD	Cambridge Structure Database
CVFF	Consistent Valence ForceField
DEMAX	Dérivée maximale
DRX	Diffraction aux rayons X
DTT	Dithiothréitol
ELISA	Enzyme Like ImmunoSorbent Assay
ESFF	Extensive Systematic Forcefield
Fab	Fragment d'anticorps composé
Ff	Phage filamenteux
fs	femtoseconde
Hac/Ac^-	Acide acétique/acétate
HIV	Human immuno-defiency virus
K	Unité de degré Kelvin
LPS	Lipopolysaccharide
MAD	Méthode de diffusion anormale
MIR	Remplacement isomorphe multiple
MPD	2,4-méthyl-pentanediol
MSI	Molecular Simulated Inc.
PAGE	PolyAcrylamide Gel Electrophoresis
PBS	Phosphate Buffer Saline
PDB	Protein Data Bank
PEG	Polyéthylène glycol
pVIII et pIII	protéine de capside VIII et III
ps	picoseconde
RMN	Résonance magnétique nucléaire
RMS	Root Mean Square
RSFP	Réaction sérologique faussement positive
SDS	Sodium Dodécyl Sulfate
TBS	TRIS Buffer Saline
TRIS	Triéthyl-amino-méthane
VA09A	Fonction de minimisation à N variables

LISTE DES ACIDES AMINES NATURELS

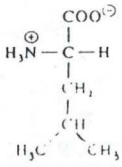
Groupes R hydrophobes (non-polaires)



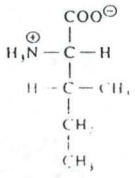
Alanine [A] (Ala)



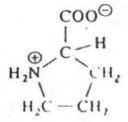
Valine [V] (Val)



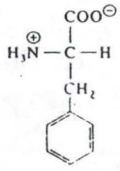
Leucine [L] (Leu)



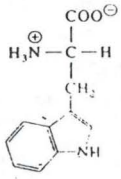
Isoleucine [I] (Ile)



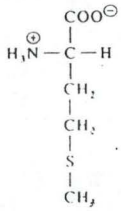
Proline [P] (Pro)



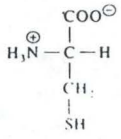
Phénylalanine [F] (Phe)



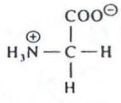
Tryptophane [W] (Trp)



Méthionine [M] (Met)



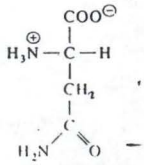
Cystéine [C] (Cys)



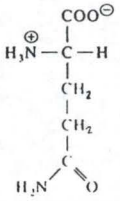
Glycine [G] (Gly)

Groupes R neutres

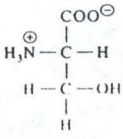
Groupes R hydrophiles (polaires)



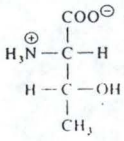
Asparagine [N] (Asn)



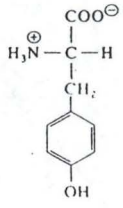
Glutamine [Q] (Gln)



Sérine [S] (Ser)

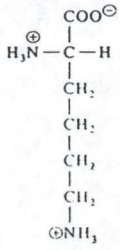


Thréonine [T] (Thr)

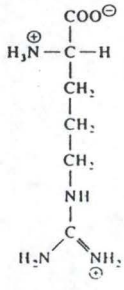


Tyrosine [Y] (Tyr)

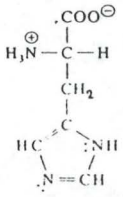
Groupes R basiques



Lysine [K] (Lys)

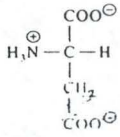


Arginine [R] (Arg)

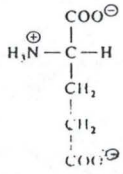


Histidine [H] (His)

Groupes R acides



Aspartate [D] (Asp)



Glutamate [E] (Glu)

name	code	side group	MW	HW	KD	EI
alanine	A ala	-CH3	71.0	-0.5	1.8	0.2
cysteine	C cys	-CH2-SH	103.0	-1.0	2.5	0.0
aspartic acid	D asp	-CH2-COOH	115.0	3.0	-3.5	-0.7
glutamic acid	E glu	-CH2-CH2-COOH	129.0	3.0	-3.5	-0.6
phenylalanine	F phe	-CH2-phi-	147.1	-2.5	2.8	0.6
glycine	G gly	-H	57.0	0.0	-0.4	0.2
histidine	H his	-CH2-imidazole	137.1	-0.5	-3.2	-0.4
isoleucine	I ile	-CH(CH3)-CH2-CH3	113.1	-1.8	4.5	0.7
lysine	K lys	-(CH2)4-NH2	128.1	3.0	-3.9	-1.1
leucine	L leu	-CH2-CH(CH3)2	113.1	-1.8	3.8	0.5
methionine	M met	-CH2-CH2-S-CH3	131.0	-1.3	1.9	0.3
asparagine	N asn	-CH2-CONH2	114.0	0.2	-3.5	-0.6
proline	P pro	[N]-(CH2)3-[CH]	97.1	0.0	-1.6	-0.1
glutamine	Q gln	-CH2-CH2-CONH2	128.1	0.2	-3,5	-0.7
arginine	R arg	-(CH2)3-NH-CNH-NH2	156.1	3.0	-4.5	-1.8
serine	S ser	-CH2-OH	87.0	0.3	-0.8	-0.3
threonine	T thr	-CH(CH3)-OH	101.0	-0.4	-0.7	-0.2
valine	V val	-CH-(CH3)2	99.1	-1.5	4.2	0.5
tryptophan	W trp	-CH2-indole	186.1	-3.4	-0.9	0.4
unknown	X ukw		0.0	0.0	0.0	0.0
tyrosine	Y tyr	-CH2-phi-OH	163.1	-2.3	-1.3	0.0

Table des matières

INTRODUCTION **5**

1) *BRUCELLA* ET LE LPS DE *BRUCELLA* 5

1.1) Généralités sur *Brucella* 5

1.2) La structure du LPS et de ses épitopes 5

1.3) Problèmes liés à l'utilisation du LPS comme outil diagnostic 6

1.4) Une solution: les mimotopes du LPS 7

2) BANQUES DE PEPTIDES ALÉATOIRES ET PRÉSENTATION EN PHAGE 7

2.1) La présentation du peptide en phage 8

2.2) Le phage Ff 8

2.3) Les banques de peptides présentés en phage 9

2.3.1) Contraintes conformationnelles 10

2.3.2) Présentation de peptides et sélection par un anticorps monoclonal 10

3) LIEN STRUCTURE-FONCTION DES ANTICORPS 11

a) Séquence 11

b) Structure 11

4) COMPLEXES ANTIGÈNES-ANTICORPS 12

5) CARACTÉRISATION DE COMPLEXES LIGAND-RÉCEPTEUR 13

6) MÉTHODES EXPÉRIMENTALES POUR LA DÉTERMINATION DE STRUCTURE DES PROTÉINES 14

6.1) Cristaux-cristallographie 15

6.2) Cristaux de petites molécules organiques et de protéines 16

6.3) La cristallogenèse des protéines 17

6.4) La diffraction de cristaux de protéines 18

6.5) L'apport de la biologie 19

6.6) Les progrès parallèles de la bioinformatique et de la diffraction 20

6.7) La cristallisation d'anticorps et de peptides 20

7) MODÉLISATION MOLÉCULAIRE: MÉTHODES THÉORIQUES POUR LA DÉTERMINATION DE STRUCTURE DES PROTÉINES 21

7.1) Champs de force 22

7.2) Minimisation de l'énergie potentielle 23

7.3) Dynamique moléculaire 25

7.4) Modélisation de peptides seuls ou en complexe avec un anticorps 26

OBJECTIF DU MÉMOIRE **27**

RÉSULTATS ET DISCUSSIONS **28**

1) RÉSULTATS PRÉLIMINAIRES 28

1.1) Tests de reconnaissance (ou d'antigénicité) 28

a) peptides fusionnés à la pVIII 28

b) peptides fusionnés à la thiorédoxine 29

1.2) Test «indirect» sur la présence du pont disulfure -"Test au DTT" 29

1.3) Test d'immunogénicité en souris 30

2) CONFORMATION DU RÉCEPTEUR: LES ANTICORPS 30

2.1) Séquences-Modélisation *30*

2.2) Détermination de structure par DRX *30*
2.2.1) Détermination de concentration 30
2.2.2) Cristallisation 30
2.2.3) Diffraction et analyse des cristaux de 4F9 32
2.2.4) Perspectives 32
3) CONFORMATION DU LIGAND: LES PEPTIDES 32
a) Peptides 32
b) Protéines-chimères («carrier»): 33
3.1) Détermination de structure par DRX *34*
3.1.1) Cristallisation 34
a) Test avec le lysozyme 34
b) Tests de cristallisation des protéines chimériques: 34
3.1.2) Diffraction 36
3.1.3) Perspectives 36
3.2) Analyse de séquence *36*
3.2.1) Comparaison globale 36
3.2.2) Comparaison de séquence des peptides 6B3-[X] 37
a) Observations et commentaires 37
b) Hypothèses et outils pour vérifier ces hypothèses 37
3.2.3) Recherche du motif dans les banques de données 37
a) Résultats 38
b) Conclusions 38
3.3) Analyse structurale: modélisation des peptides *38*
3.3.1) Mise au point d'un protocole de dynamique et de mécanique moléculaire pour des peptides contraints par un pont disulfure 38
3.3.2) Utilisation d'un cas-test 40
3.3.3) Modélisation des peptides 6B3-[X] et 4F9-[1] 40
a) Tests réalisés et résultats 40
b) Graphes des énergies 41
c) "Cluster graphs" 41
3.3.4) Analyse locale de la flexibilité du squelette carboné 42
a) Objectifs et stratégie 42
b) Remarques préliminaires 42
c) Résultats 43
d) Conclusions 44
3.3.5) Etablissement de familles pour chaque peptide sur base de la conformation du squelette carboné 44
a) Objectifs et stratégie 44
b) Résultats 45
c) Conclusions et hypothèses 45
3.3.6) Etablissement de familles entre tous les conformères de tous les peptides 6B3-[X]P 46
a) Objectifs et stratégie 46
b) Résultats 46
c) Hypothèses 47
d) Qualité du modèle 48

CONCLUSIONS GÉNÉRALES ET PERSPECTIVES **49**

ASPECT EXPÉRIMENTAL: CRISTALLISATION DU FRAGMENT FAB DE L'ANTICORPS 4F9 ET DE PROTÉINES DE FUSION THIORÉDOXINE-PEPTIDES MIMÉTIQUES DE LA CHAÎNE O DE BRUCELLA 49

ASPECT THÉORIQUE: MISE AU POINT D'UN PROTOCOLE DE MODÉLISATION PAR DYNAMIQUE MOLÉCULAIRE ET MINIMISATION D'ÉNERGIE DE PEPTIDES MIMÉTIQUES ET ÉTUDE D'UNE SÉRIE DE CONFORMÈRES GÉNÉRÉS PAR CETTE MODÉLISATION 50

MATÉRIEL ET MÉTHODES **53**

OUTILS EXPERIMENTAUX 53

1) Echantillons de protéines 53

a) Thiorédoxine-peptides 53

b) Fab 53

2) Concentration et dialyse de la solution de protéines 53

Remarques 54

3) Test de pureté et de concentration: TEST BCA (Bicinchoninic Acid) 54

a) Matériel 54

b) Méthode 54

4) Conservation et conditions d'utilisation 55

5) Elaboration de tampons (liqueur-mères) 55

a) Matériel 55

b) Méthode 55

Remarques 56

6) Application de la technique par gouttes suspendues 56

a) Matériel 56

b) Méthode 56

Remarques 57

7) Observation des cristaux 57

Remarques 57

8) Coloration, isolement des cristaux et tests de dureté 57

a) Matériel 57

b) Méthode 58

Remarques 58

9) Diffraction 58

OUTILS THÉORIQUES 59

1) Outil graphique et programmes de modélisation moléculaire *59*

a) Console graphique 59

b) MSI (Molecular Simulated Incorporating) 59

c) Modélisation de peptides 59

I) "Construction" du peptide 59

II) Mécanique et dynamique moléculaire 61

Remarques 61

d) Utilisation d'*insightII* pour l'étude de conformères issus d'une dynamique moléculaire 62

e) Superposition des peptides: modification des fichiers «.arc» 62

INTRODUCTION

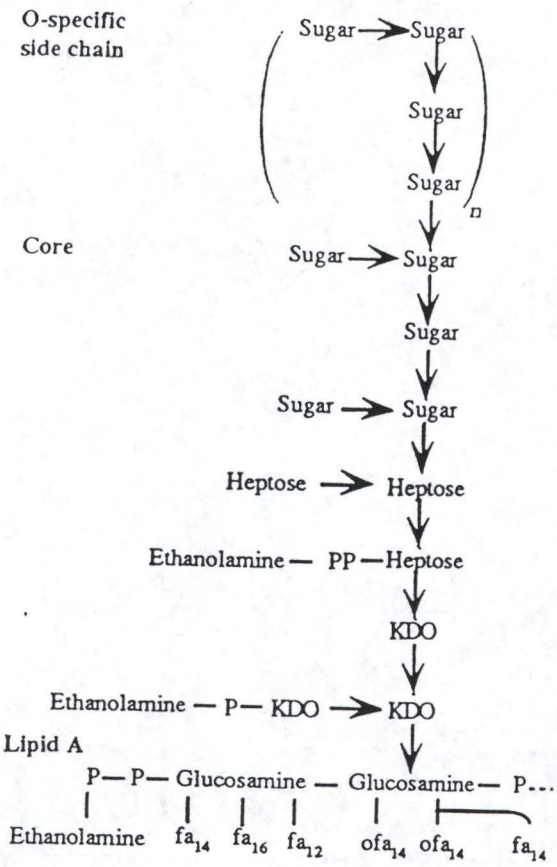


Figure 1: Diagramme schématique d'un monomère de lipopolysaccharide. (Letesson, 1997).

INTRODUCTION

1) *BRUCELLA* ET LE LPS DE *BRUCELLA*

1.1) Généralités sur *Brucella*

Brucella sp. est une bactérie Gram négative pathogène et intracellulaire facultative (Smith, 1985). C'est un coccobacille aérobie strict responsable d'une maladie sévissant mondialement: la brucellose. Elle touche principalement les animaux domestiques (tels que les bovins, les caprins et les ovins), mais l'homme peut être infecté par consommation de nourriture contaminée ou par contact avec un animal atteint (Corbel, 1997). Les bactéries du genre *Brucella* sont divisées en 6 espèces: *B.melintensis*, *B. abortus*, *B. suis*, *B. canis*, *B. neotomae* et *B. ovis*.

Le LPS (lipopolysaccharide) est la structure contre laquelle est dirigée la majorité de la réponse humorale de l'hôte infecté par *Brucella*.

1.2) La structure du LPS et de ses épitopes

Le lipopolysaccharide constitue en majorité le feuillet externe de la membrane externe des bactéries Gram négatives (figure 1).

Cette molécule comporte 3 parties: le lipide A, le "core" et la chaîne O. Le lipide A sert d'ancrage membranaire via plusieurs acides gras greffés sur un dimère de glucosamines phosphorylées. Le core est une courte série de sucres branchés sur lequel vient s'ajouter la chaîne O, un polymère de sucres, aussi appelé antigène O étant donné son caractère immunodominant lors d'une infection.

Il existe des souches de bactéries Gram négatives incapables de synthétiser une molécule complète de LPS: la chaîne O fait défaut. Ces souches sont nommées rough (R: rugueuses) car elles montrent une morphologie de colonies rugueuses. Les souches sauvages sont dites smooth (S: lisses) (Rick, 1987).

Chez *Brucella*, la chaîne O est un homopolymère de 4-formamido-perosamine.

Il existe plusieurs types d'épitopes sur la chaîne O: les mieux décrits sont les épitopes A et M. L'épitope A serait une succession de sucres unis par des liaisons α-1,2, tandis que l'épitope M serait un pentasaccharide comportant un lien α-1,3 et 3 liens α-1,2 (Bundle *et al.*, 1989). Sur base de la fréquence de ces 2 épitopes dans la chaîne, on classe les souches smooth de

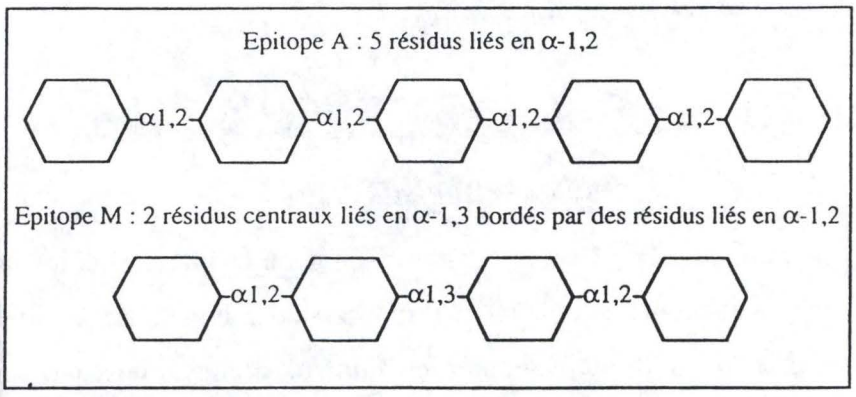


Figure 2: Structure hypothétique des épitopes A et M de *Brucella sp.* (Weynants, 1998), d'après Bundle *et al.* (1989).

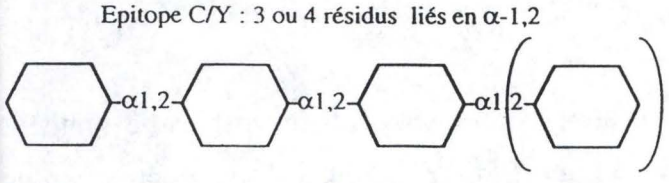


Figure 3: Structure hypothétique de l'épitope C/Y de *Brucella sp.* (Weynants, 1998), d'après Bundle *et al.* (1989).

Brucella en 3 sérotypes: A^+M^-, $A\text{-}M^+$, A^+M^+, par exemple les souches M-dominantes (A^-M^+) contiennent 25% de liaisons α-1,3 entre les résidus perosamines (figure 2).

Des évidences expérimentales ont permis de définir un épitope commun et spécifique à toutes les souches lisses de *Brucella*: l'épitope C. Il existe un autre épitope commun mais non-spécifique, car il est présent également dans la chaîne O de toutes les souches lisses de *Brucella* et de *Y. enterocolitica* O:9[1] (la chaîne O de cette bactérie est un homopolymère de perosamines liées en alpha-1,2): l'épitope C/Y (Palmer and Douglas, 1989) (figure 3).

1.3) Problèmes liés à l'utilisation du LPS comme outil diagnostic

Il existe 2 types de tests diagnostics pour la détection de la brucellose: les tests indirects et le diagnostic direct ou bactériologique. Les tests indirects analysent la réponse immunitaire: ils visent la réponse humorale (tests sérologiques: prélèvement de sérum) et parfois la réponse cellulaire (par exemple les tests intradermiques comme l'hypersensibilité retardée (Alton *et al.*, 1988). Le test direct est toujours nécessaire pour confirmer ou infirmer un test sérologique positif: cette méthode est la seule qui permette un diagnostic de certitude. Cependant, sa mise en œuvre est plus lourde qu'un simple test sérologique.

Les vaccins utilisés actuellement sont des souches de *Brucella* vivantes atténuées, utilisées sur certains animaux domestiques (bovins et petits ruminants) tandis qu'elles restent pathogènes pour l'homme (Blasco and Diaz, 1993).

Un problème majeur associé à la vaccination est que l'inoculation du vaccin induit la production d'anticorps anti-LPS lisse (par opposition au LPS rugueux qui ne contient pas de chaîne O), lesquels servent classiquement pour la détection de la maladie (détection indirecte, test sérologique). Dans ce cas, il est dès lors impossible de différencier un animal vacciné d'un animal atteint (Belzer *et al.*, 1991).

Un autre problème du diagnostic sérologique de la brucellose, lui aussi lié au LPS lisse, est que d'autres bactéries présentent un LPS structuralement très proche de celui de *Brucella* et donnent des réponses sérologiques faussement positives (RSFP; détection positive lors d'un diagnostic sérologique pour la brucellose alors que l'animal testé n'est pas atteint). Par exemple, un animal infecté par *Y. enterocolitica* O:9 va, entre autres, produire des anticorps dirigés contre le LPS lisse de cette bactérie. Comme ce LPS est très proche de celui des

[1] *Yersinia enterocolitica O:9 est une entérobactérie non-pathogène*

	Animal infecté par *Y. entero-colitica* >>> production d'anticorps anti-épitopes A et C/Y.	Animal infecté par *Brucella* >>>production d'anticorps anti-épitopes A, M, C et C/Y.
Détection via LPS de Brucella (épitopes A, M, C et C/Y)	+	+
Détection via un épitope C(commun à toutes les souches de *Brucella*)	-	+

Figure 4 : Tableau mettant en évidence le problème de réactions-croisées entre la chaîne O du LPS de *Y. enterocolitica* et la chaîne O du LPS de *Brucella.* L'isolement de l'épitope C permettrait de distinguer les animaux atteints de brucellose des animaux infectés par *Y.enterocolitica O:9* .

souches *Brucella*, ces anticorps reconnaissent le LPS lisse de *Brucella* lors des tests sérologiques (Weynants *et al.*, 1996). On parle de réactions croisées (figure 4).

Le LPS est donc au centre du problème diagnostique. Signalons également que le LPS pourrait faire partie d'une composition vaccinale. En effet, l'immunisation passive de souris avec des anticorps anti-LPS lisse permet une protection contre une infection ultérieure par *Brucella*. Cependant, la toxicité du LPS, les difficultés inhérentes à sa synthèse et à sa production ont empêché, jusqu'à présent, l'entrée du LPS dans la composition d'un vaccin anti-*Brucella*. Après la présentation de la structure du LPS, nous résumerons les approches originales mises en œuvre pour pallier au manque de spécificité du test sérologique.

L'épitope C isolé pourrait être utilisé pour améliorer le diagnostic. Afin d'isoler l'épitope C, on peut penser à cliver le LPS de manière spécifique, mais plusieurs problèmes sont liés à cette méthode. En effet, on ne connaît pas la structure du LPS. De plus, *Brucella* est une bactérie pathogène, il vaut mieux éviter une culture à grande échelle pour des problèmes de biosécurité. Enfin, la chimie et la synthèse des sucres est très peu développée à l'heure actuelle.

1.4) Une solution: les mimotopes du LPS

Une alternative théorique est l'identification d'un composé, reconnu uniquement par des anticorps spécifiques de l'épitope C, mais pas obligatoirement de même nature chimique Il serait possible d'isoler de tels composés en se servant de banques de peptides. La molécule sélectionnée sera appelée "mimotope" dans le cas d'une sélection par un anticorps.

2) BANQUES DE PEPTIDES ALEATOIRES ET PRESENTATION EN PHAGE

Ces banques sont utilisées pour cibler des épitopes linéaires[2], ligands de molécules d'intérêt.

Il existe des banques de type chimique et de type biologique (Gallop *et al.*, 1994).

En ce qui concerne les banques de type biologique, la diversité est générée par la synthèse chimique combinatoire de mixtures d'oligonucléotides, les séquences dégénérées sont ensuite placées dans un contexte génétique adéquat, c'est-à-dire qui permet la présentation du peptide encodé par l'oligonucléotide inséré. Des banques peuvent contenir jusqu'à 109 clones

---

[2] *Un épitope linéaire est un épitope dont les constituants sont agencés les uns à la suite des autres dans la séquence.*

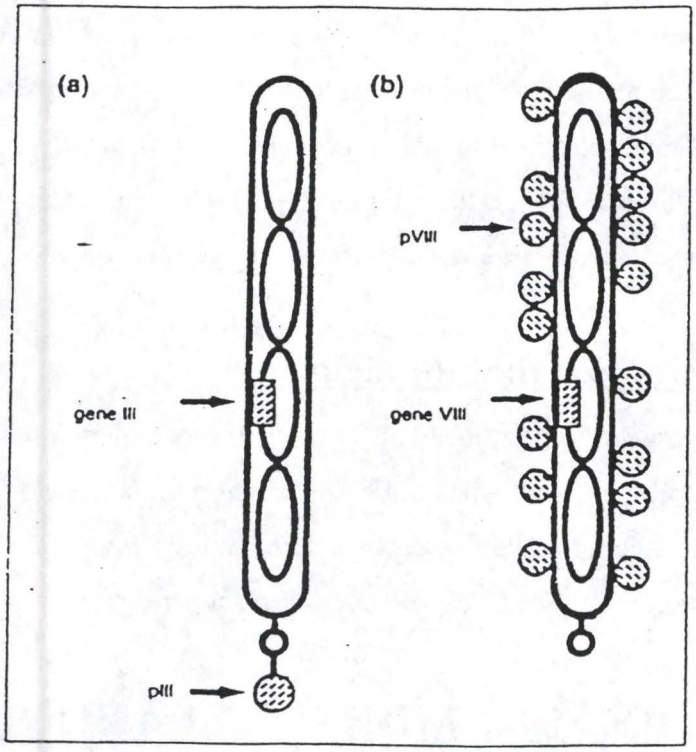


Figure 5: Représentation schématique d'un phage filamenteux présentant des épitopes peptidiques en fusion avec:
(a) la protéine pIII
(b) la protéine pVIII
L'ADN simple brin du phage est également représenté, doté de l'insert nucléotidique encodant le peptidique présenté. (Scott, 1992).

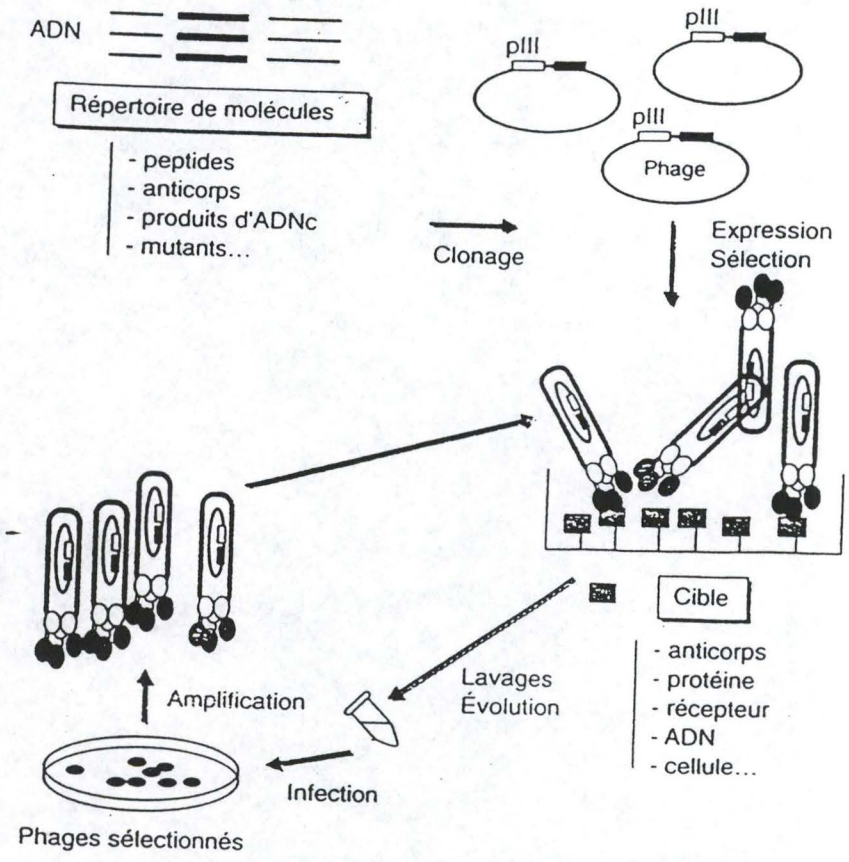


Figure 6: Le panning, sélection d'une banque de peptides exprimés à la surface de phages filamenteux par leur liaison à une cible. Le mélange de phages est incubé avec la cible. Après les lavages, les phages fixés sont élués puis amplifiés par infection de bactéries. Les phages amplifiés sont ensuite utilisés pour un nouveau cycle. (Souriau *et al.*, 1998).

indépendants (plus le nombre de peptides est élevé, plus les chances d'isoler les peptides d'intérêt sont élevées).

Plusieurs approches parmi les banques de type biologique, caractérisées par un mode de présentation distinct ont été mises en œuvre: banques sur plasmides, sur protéines membranaires et cytoplasmiques de bactéries, sur polysomes et sur phages.

Un des avantages des banques de type biologique est que ce n'est jamais le peptide lui-même qui est séquencé mais la séquence nucléotidique (ADN ou ADNc) qui l'encode. En effet, il existe toujours un lien physique (banques sur plasmide) ou topographique entre le peptide et la séquence codante. Le séquençage de l'ADN est nettement plus aisé et moins onéreux que le séquençage de peptides. Un autre avantage de ces banques est qu'elles ne s'épuisent pas.

2.1) La présentation du peptide en phage

La présentation des peptides se réalise à l'extrémité NH_2-terminale d'une protéine de capside du phage par insertion des séquences aléatoires au niveau du gène codant pour cette protéine. (Smith, 1985) (figure 5) Les peptides sont donc présentés en surface du phage et sont sélectionnés par leur capacité de liaison à une cible. Après les différents lavages, les phages sont élués et amplifiés par infection de bactéries. Les phages amplifiés sont sélectionnés de nouveau sur la même cible. Plusieurs cycles de sélection-amplification ("pannings") permettent de surreprésenter les phages d'intérêt. Après un certain nombre de tours de pannings (souvent de 1 à 3), un criblage des phages résultants permet de sélectionner des phages porteurs de peptides reconnus par la cible. Les phages sélectionnés sont ensuite séquencés dans la région codant pour le peptide (figure 6). Les phages qui présentent diverses séquences de peptides sont testés individuellement pour l'activité recherchée. Les trois phages filamenteux (aussi appelés Ff) utilisés sont M13, f1, et fd qui infectent *E.Coli* (Webster, 1996). Ils sont très proches au niveau de la séquence nucléotidique (98 % d'homologie). Une caractéristique importante du cycle de vie du phage est que l'infection ne mène pas à la lyse bactérienne ou à la mort cellulaire.

2.2) Le phage Ff

Le phage Ff sera décrit plus en détail, car il constitue l'outil principal qui a permis d'isoler les peptides mimétiques du LPS étudiés dans ce travail.

Le génome ADN simple brin du phage Ff est constitué de 6407 nucléotides répartis en 11 gènes. Ces gènes sont groupés selon leur fonction dans le cycle de vie du phage (Model and Russel, 1988). Dans la particule phagienne, l'ADN se situe dans un cylindre flexible comprenant environ 2700 unités monomériques de la protéine pVIII, la protéine de capside la

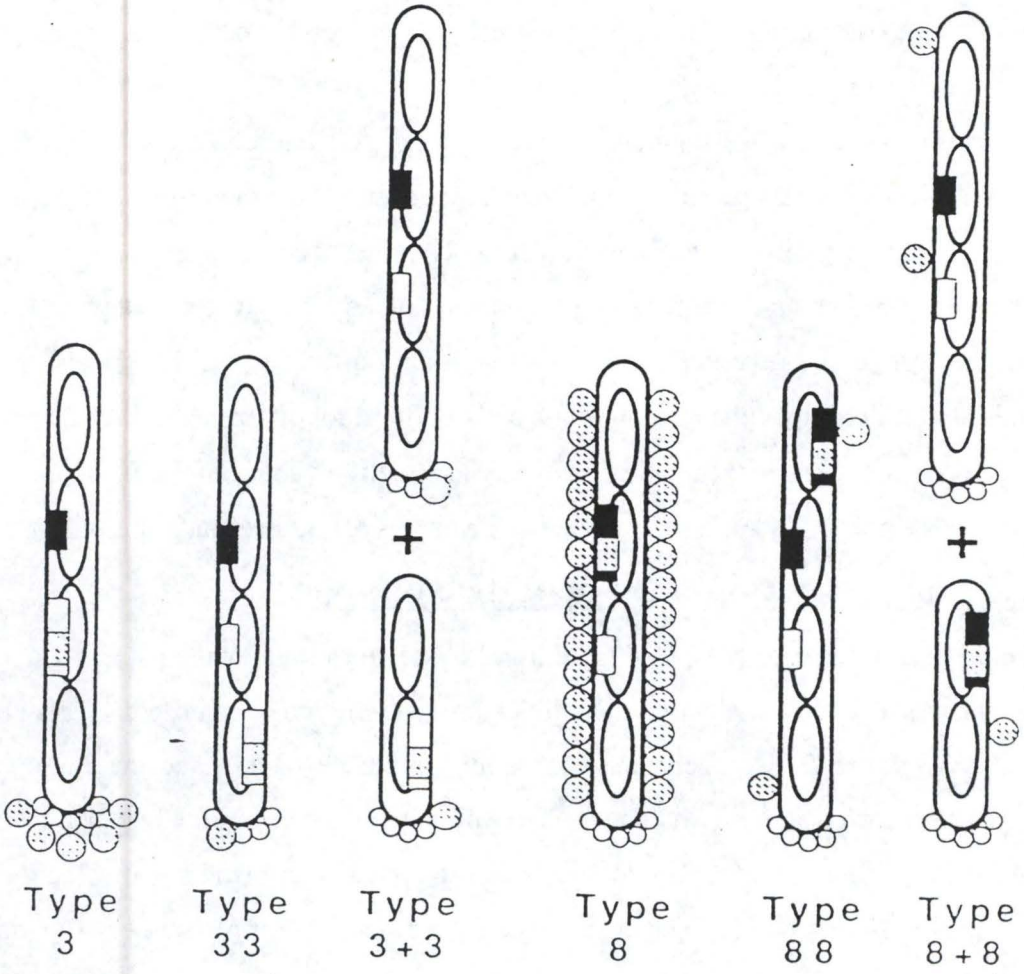


Figure 7: Types de présentation de peptides en phage (voir texte ci-contre pour l'explication) (Smith & Scott, 1993).

plus abondante. Aux extrémités de la capside, on retrouve d'un côté 5 protéines pVII et 5 protéines pIX et de l'autre 5 protéines pIII et 5 protéines pVI.

Les processus de transcription-traduction, de réplication et d'assemblage sont bien tolérés par l'hôte: la bactérie infectée continue à croître et à se diviser avec un temps de génération un peu plus long que celui d'une bactérie non infectée. Environ 1000 particules sont produites durant la première génération suivant l'infection, après quoi l'hôte produit 100 à 200 phages par génération.

La protéine pVIII contient 50 acides aminés, elle peut être divisée en 3 domaines sur base de la composition en acides aminés (Hill and Stockey, 1996). L'extrémité amino-terminale est acide, la région centrale est hydrophobe et l'extrémité carboxy-terminale est basique.

La protéine pVIII présente une conformation en hélice α ininterrompue, excepté pour les 5 acides aminés de l'extrémité amino-terminale dirigée vers l'extérieure.

La protéine pIII comprend 406 acides aminés et possède 2 domaines de taille comparable: le domaine C-terminal est enfoui dans la particule et la partie N-terminale est protubérante vers l'extérieur.

Ces 2 protéines de capside (pVIII et pIII) sont intéressantes car leur arrangement et leur structure sont assez bien connus et des protéines fusionnées à des peptides ne semblent pas menacer l'intégrité du phage (Cesareni *et al.*, 1995).

2.3) Les banques de peptides présentés en phage

Les premiers essais ont été réalisés en utilisant la protéine pIII comme "carrier". Le site d'insertion a été localisé à l'extrémité N-terminale pour ne pas trop perturber la fonction de la protéine pIII. L'insertion de longues séquences a pu être réalisé par génération de phages possédant en même temps des protéines pIII chimériques et des protéines pIII sauvages (Smith, 1985). Cela peut être obtenu par l'ajout d'une copie supplémentaire du gène de la protéine pIII dans le génome du phage ou par l'utilisation de phagemides (banques de type 33 et 3+3 respectivement) (Smith and Scott, 1993).

L'inconvénient majeur de l'utilisation de la protéine pIII pour la présentation de séquences peptidiques est le nombre peu élevé d'exemplaires de protéines de fusion qui peuvent être insérées sur chaque particule (figure 7).

Pour cette raison des insertions dans la protéine pVIII ont été réalisées pour parvenir à présenter un plus grand nombre de peptides (intéressant pour certaines applications du "phage display").

Le site d'insertion de la séquence peptidique se situe au niveau des 4 ou 5 résidus N-terminaux de la protéine pVIII mature (sans la séquence signal).

L'insertion en protéine pVIII est plus contraignante qu'en protéine pIII: pour un phage "normal" les résidus ne peuvent pas être plus long que 6-8 résidus tandis que la présence simultanée de protéines pVIII sauvages et chimériques permet l'insertion de nombreux résidus. Il existe en effet le même type de banque pour la protéine pVIII que pour la protéine pIII (type 88, type 8+8) (Felici *et al.*, 1991).

2.3.1) Contraintes conformationnelles

Les peptides présentés sont dotés d'une grande flexibilité: ils ont une extrémité amino-terminale libre et sont souvent assez courts (moins de 20 résidus). Pour diminuer la perte d'entropie conformationnelle et donc augmenter leur affinité, lors de la liaison à la molécule-cible, plusieurs banques de peptides contraints ont été créées (Kay and Hoess, 1996). Les techniques utilisées se sont inspirées de ce que la nature utilise comme la formation de macrocycles à l'aide de ponts disulfures (McLafferty *et al.*, 1993). D'autres alternatives ont été envisagées: insertion du peptide dans un domaine protéique (boucles hypervariables d'anticorps, motifs en doigt de zinc) où le peptide y est "contraint" par son environnement protéique.

2.3.2) Présentation de peptides et sélection par un anticorps monoclonal

Une fois la banque de peptides en phage créée, la sélection des clones intéressants est réalisée grâce à l'interaction possible entre les peptides présentés à la surface du phage et de récepteurs cibles (protéines, ADN) (Souriau *et al.*, 1998). Dans certains cas, la sélection permet de mettre en évidence plusieurs peptides différents définissant un consensus au niveau de la séquence. Les outils de sélection peuvent être de l'ADN (reconnaissance par des molécules en doigt de zinc présentées sur phage), des protéines purifiées (Ac monoclonal, lectines, récepteur d'intérêt pharmacologique) ou encore un mélange complexe (cellules, animaux, sérum de patients).

L'outil le plus utilisé pour la sélection de peptides sur phages est l'Ac monoclonal. L'utilisation des procédures de panning avec des anticorps monoclonaux reconnaissant des épitopes linéaires a été particulièrement fructueuse. En effet, ces anticorps permettent souvent la sélection de peptides qui ont une séquence consensus en commun avec la séquence de la protéine reconnue par l'anticorps. Cette approche conduit donc à l'identification de la séquence reconnue par l'anticorps dans l'antigène protéique. Si l'épitope reconnu par l'anticorps est non-linéaire, alors l'exploitation de banques de peptides en phage va permettre

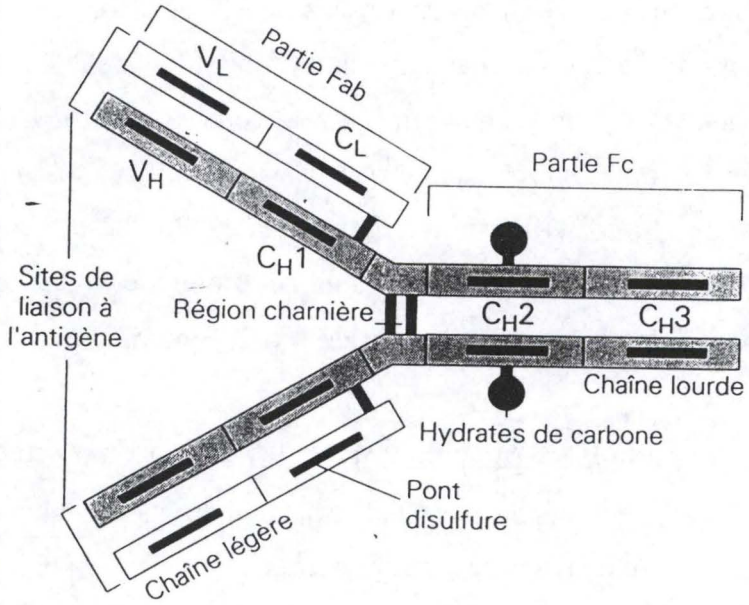


Figure 8: Structure de base de l'isotype IgG1 (Roitt & Bristoff, 1993)

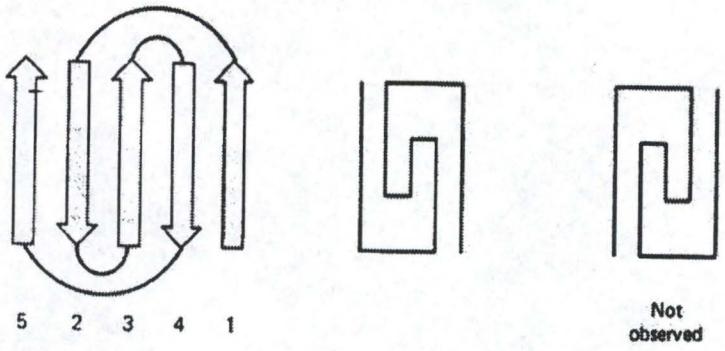


Figure 9: Diagramme topologique montrant un motif en clef grecque.

d'isoler des mimotopes, c'est-à-dire des peptides qui miment l'épitope naturel, mais qui ne présentent pas obligatoirement d'identités de séquence avec la macromolécule reconnue par l'anticorps.

3) LIEN STRUCTURE-FONCTION DES ANTICORPS

L'une des caractéristiques des anticorps en tant qu'effecteur indispensable du système immunitaire est leur spécificité de reconnaissance d'un épitope défini. Cette reconnaissance se fait via des interactions entre cet épitope et le paratope (partie de l'anticorps qui se lie à l'épitope) (Davies and Padlan, 1990).

La structure de base des anticorps comprend 2 chaînes légères identiques et 2 chaînes lourdes identiques. Il y a 5 isotypes (ou classes) de chaîne lourde: IgG, IgM, IgA, IgE et IgD, chacun caractérisé par une structure, une fonction et une localisation bien spécifiques. L'isotype IgG est l'immunoglobuline majeure du sérum des patients normaux. Il y a également 2 isotypes pour la chaîne légère (lambda et kappa) pour lesquels il ne semble pas y avoir de distinction au niveau fonctionnel (figure 8) (Branden and Tooze , 1991).

Les chaînes légères (L, "light") sont constituées de 2 domaines et les chaînes lourdes (H, "heavy") de 4 ou 5 domaines. Ces domaines sont composés d'environ 110 acides aminés chacun.

a) Séquence

En comparant les séquences d'anticorps, on peut mettre en évidence des domaines conservés ("constants": CH1, CH2, CH3, CL) et variables (VH et VL). Trois courtes régions situées à l'extrémité des domaines variables, les CDR (régions déterminantes de la complémentarité ou boucles hypervariables) sont responsables de la diversité et de la spécificité de reconnaissance des anticorps. Seule la composition et la longueur des CDR varient au sein des anticorps d'un même isotype; la localisation de ces CDR, quant à elle, ne varie pas. Ces boucles hypervariables sont toujours situées aux mêmes endroits: les acides aminés 24-34, 50-56 et 89-97 pour les chaînes légères et les acides aminés 31-35, 50-65 et 95-102 pour les chaînes lourdes.

Le reste de la chaîne variable et les domaines constants ont entre 30 et 40 % d'identité de séquence pour un même isotype.

b) Structure

Au niveau structural, les anticorps ont un repliement bien caractérisé ("immunoglobuline fold", en "clef grecque") (figure 9). Les différents domaines sont composés de brins β formant

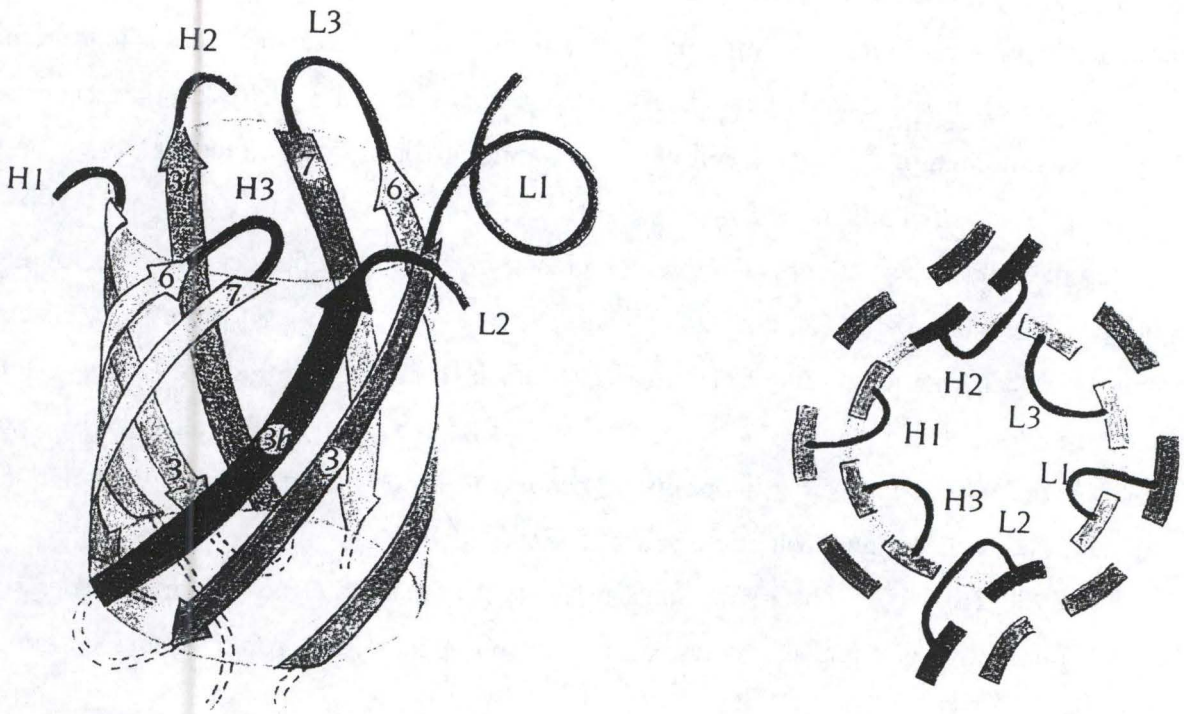


Figure 10: Arrangement schématique du tonneau β caractéristique du domaine Fab des Anticorps. Les six régions hypervariables sont situées à une extrémité de ce tonneau. (Branden & Tooze , 1991).

2 plans β antiparallèles. Ces structures, sont stabilisées par des ponts disulfures à l'intérieur des différents domaines mais aussi entre les domaines (entre les CH2 des chaînes lourdes et entre les CL des chaînes légères). Les domaines VH et VL sont également fortement liés via des interactions hydrophobes. Ils sont situés perpendiculairement l'un par rapport à l'autre, et permettent de cette manière la "protrusion" des 3 CDR des 2 domaines. Ces 6 boucles forment le site de reconnaissance de l'antigène, ou paratope (Branden and Tooze , 1991) (figure 10).
Les brins β, au nombre de 7 pour les domaines constants sont reliés par 6 boucles identiques (6) pour une même classe d'anticorps. Les domaines variables ont 2 brins β de plus (3b et 3c) et sont reliés par les 3 boucles hypervariables (entre les brins 6 et 7, 2 et 3, 3a et 3b) et 5 autres boucles, peu variables et qui n'ont pas de rôle dans la reconnaissance de l'antigène. L'extrême variabilité de séquence des CDR n'affecte pas la constance structurale des plans β.
Une étude de la structure des CDR révèle que les résidus importants pour le maintien de la conformation des boucles (en général glycine et proline) sont très conservés. On observe même que pour 5 des CDR, il n'existe qu'un petit répertoire de conformations de la chaîne principale. Les variations de séquences moduleraient la surface que ces structures "canoniques" présentent à l'antigène (Chothia, 1989).
Le CDR3 de la chaîne lourde varie en longueur et adapte une conformation "en tête d'épingle" ("hairpin"). Pour une telle structure, de nombreuses conformations sont envisageables à partir d'une même séquence, dépendant de la taille de la boucle et des interactions (par "packing") avec le reste de l'anticorps.
Ces données sont importantes pour la prédiction de structure du site de liaison à l'antigène. En effet, pour les 5 premières boucles, grâce à l'existence de structures canoniques, il sera aisé de prédire la conformation (cfr. "modélisation par homologie").

4) COMPLEXES ANTIGENES-ANTICORPS

Pour simplifier la reconnaissance entre l'antigène et l'anticorps, on la compare souvent à l'image "clef-serrure". Le gros problème de cette simplification est son aspect statique: la liaison de l'anticorps et de l'épitope induit un changement de conformation plus ou moins important du récepteur tandis que le ligand semble conserver une conformation proche de celle adoptée en solution. Par exemple, après examen de structure d'un complexe Fab-neuraminidase (petit peptide de VIH), on remarque que les boucles sont déplacées de 3 Å par rapport à leur position originale (Colman, 1987). L'interaction Fab-lysozyme est très

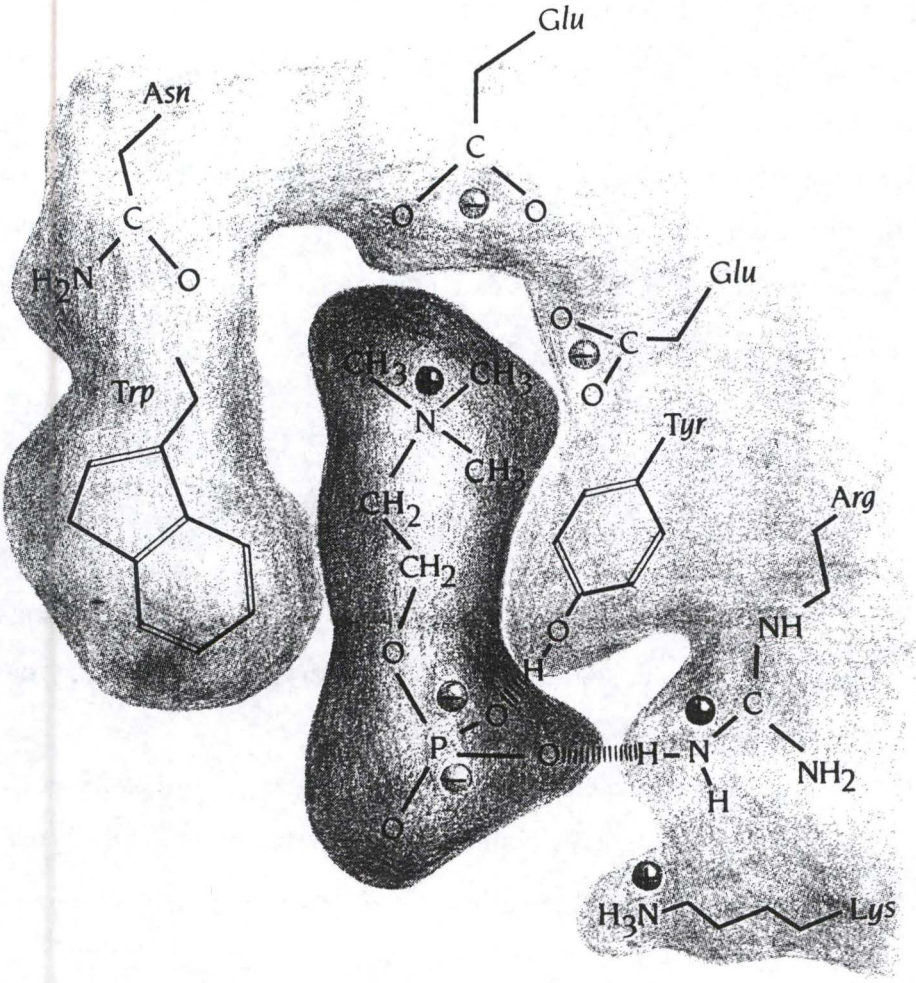


Figure 11: Représentation schématique d'une interaction spécifique entre une molécule organique (ici la phosphorylcholine) et certaines chaînes latérales des CDR. (Branden & Tooze , 1991).

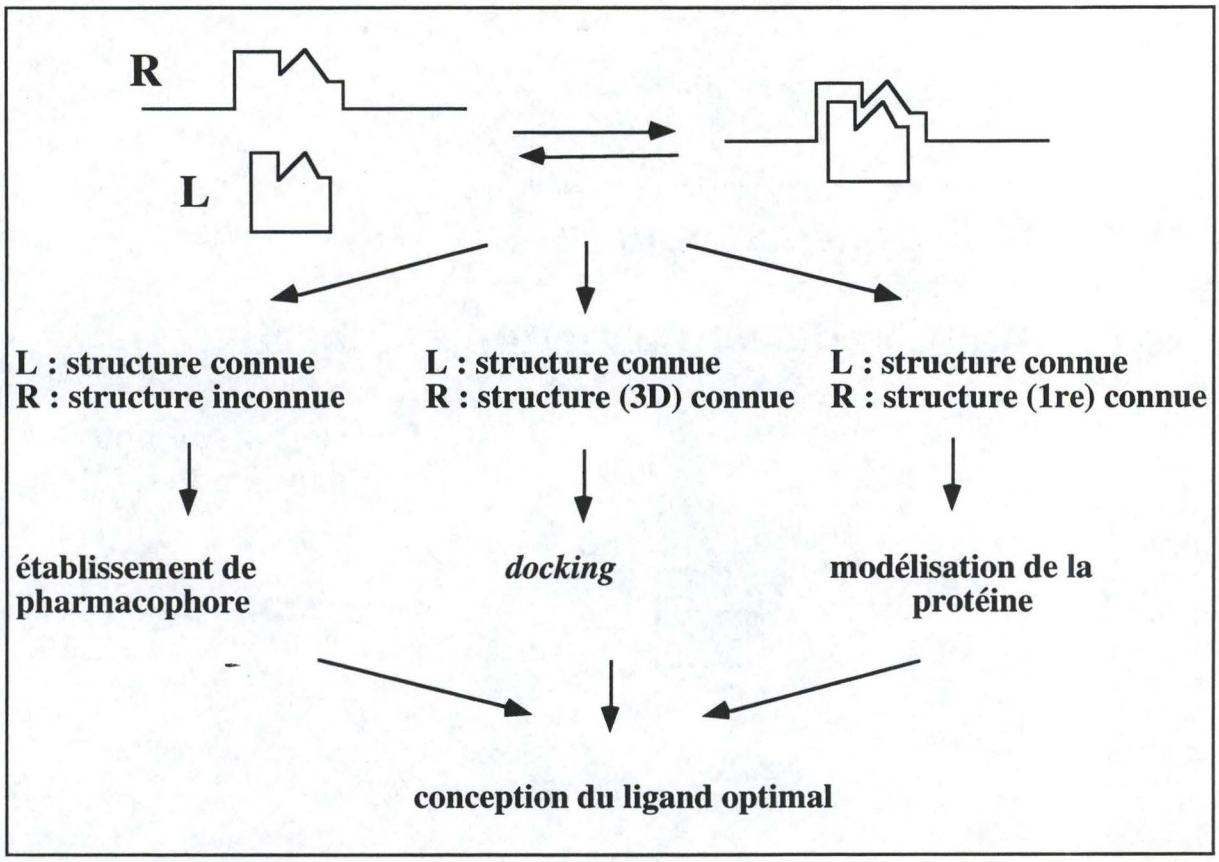


Figure 12: Schéma représentatif du type de stratégie envisageable selon le degré de connaissance de la structure du récepteur.

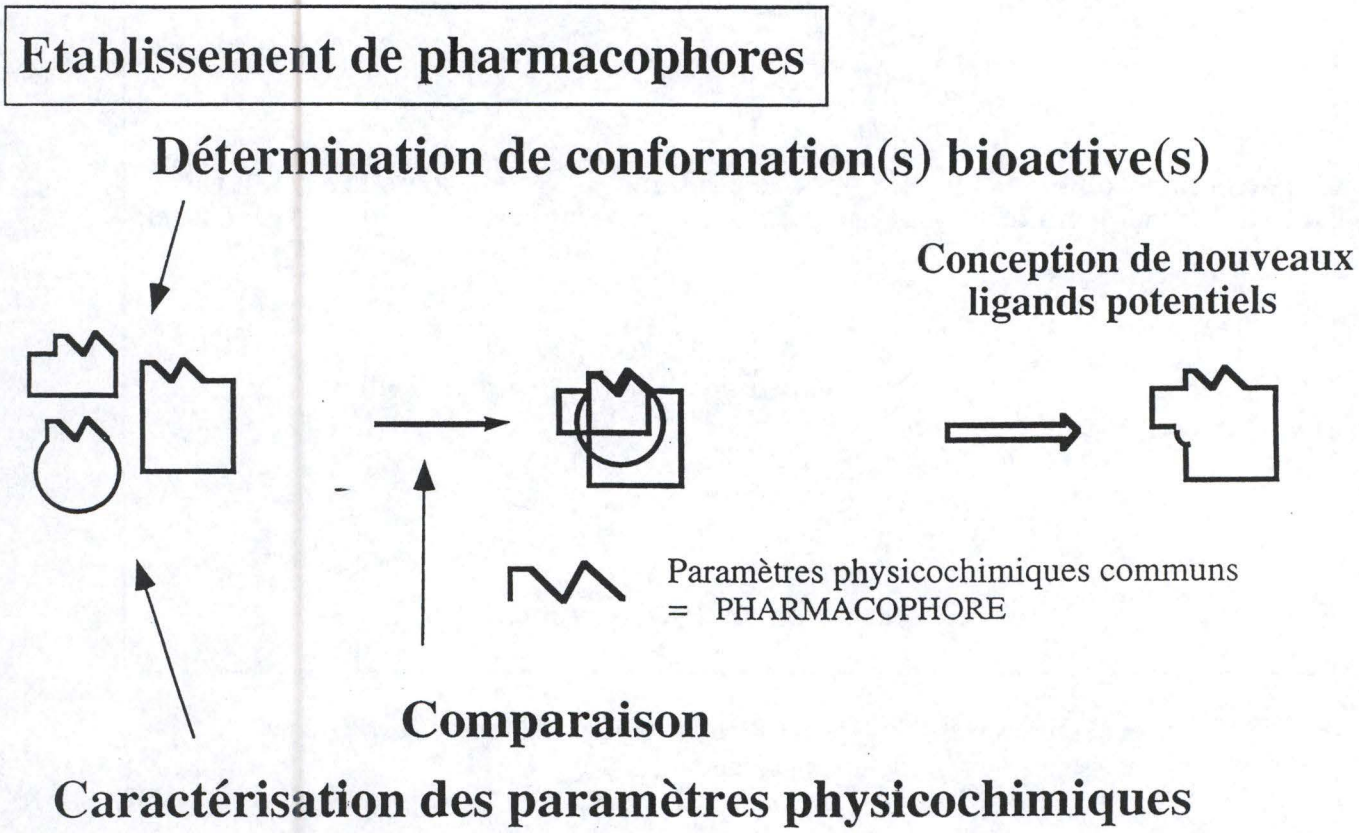


Figure 13 : Représentation de la démarche utilisée pour déterminer le pharmacophore (Wouters, 1998)

différente, on observe que la conformation du lysozyme et des boucles est quasiment identique, qu'ils soient seuls ou complexés.

En outre, cette image "clef-serrure" est intéressante car la reconnaissance ne se fait pas en surface des boucles mais par une imbrication de l'épitope entre les 6 CDR.

La cavité d'interaction entre les CDR et l'épitope est caractérisée par sa largeur en surface, sa profondeur et le nombre et le type de résidus impliqués dans la reconnaissance.

Dans cette cavité, on peut repérer plusieurs types de liaisons entre le paratope et l'épitope: liaisons ioniques (entre groupements de charge opposée), liaisons hydrogènes, liaisons hydrophobes (résultant de l'interaction de groupements par exclusion des molécules d'eau) et liaisons à courte portée (entre dipôles permanents ou induits) (figure 11).

Les résidus importants pour l'interaction doivent être placés correctement. Le paramètre stérique, en plus des caractères hydrophobes et hydrophiles est donc capital (Davies and Padlan, 1990).

5) CARACTERISATION DE COMPLEXES LIGAND-RECEPTEUR

Nous venons de décrire l'interaction entre un anticorps (récepteur) et un antigène (ligand). En général, pour déterminer les points d'ancrage structuraux lors de l'interaction ligand-recepteur, il existe plusieurs approches. On distingue la conception directe et la conception indirecte de nouveaux ligands en fonction des données et des outils disponibles (Wouters, 1998) (figure 12).

La conception directe est envisageable quand on dispose de la structure des ligands et de la structure ou de la séquence du récepteur. Dans le cas où la structure du récepteur serait connue, l'outil principal est le "docking": le ligand est imbriqué dans le récepteur à l'aide d'un programme informatique en optimalisant l'interaction (pour former un complexe stable). On peut introduire plusieurs biais en tenant compte de la structure du ligand: le ligand est considéré comme rigide ou flexible ainsi que le récepteur. Dans le cas où on ne connaîtrait pas la structure du récepteur mais que l'on dispose de la séquence, il est possible d'utiliser la modélisation (décrite au chapitre suivant) et de passer ensuite au "docking". En effet, si la structure de l'anticorps n'est pas connue, elle peut être modélisée à partir de la séquence d'autres anticorps tandis que les boucles hypervariables sont modélisées par dynamique moléculaire (cette méthode est explicitée dans la suite du travail).

La conception indirecte est envisagée si on ne possède que la structure ou la séquence des ligands. La seule alternative est la détermination du "pharmacophore" (figure 13). Le

pharmacophore est l'ensemble des caractéristiques physico-chimiques minimales du ligand (paramètre stérique, propriétés hydrophiles ou hydrophobes) qui permettent l'interaction avec le récepteur. Le pharmacophore peut être déterminé si une série de ligands sont connus. Les structures de ces ligands sont superposées et on peut de cette façon déterminer les groupements communs susceptibles d'être importants dans l'interaction avec le récepteur (Wouters, 1998). Des ligands de structure proche non-reconnus par le récepteur sont aussi utilisés pour éliminer du pharmocophore d'éventuels paramètres communs inutiles pour la reconnaissance. Ce type de stratégie est l'analogie structurale d'un consensus de séquences. Une alternative à la comparaison de séquences de peptides mimétiques peut être l'outil "pharmacophore" (ce qui est reconnu par le récepteur) que l'on peut relier à la notion de mimotope (Wouters, 1998).

6) METHODES EXPERIMENTALES POUR LA DETERMINATION DE STRUCTURE DES PROTEINES

Pour déterminer la conformation bioactive d'une molécule, il existe plusieurs méthodes expérimentales: la cristallographie, la RMN bidimensionnelle, et d'autres méthodes plus spécifiques comme le dichroïsme circulaire, et la cryo-microscopie à électrons (Duée and Dideberg, 1997).

La cristallographie sera décrite ci-dessous et expliquée plus en détail dans la chapitre "matériel et méthode".

La Résonance Magnétique Nucléaire (ou RMN) permet une analyse en solution des molécules organiques et de petites protéines grâce aux déplacements et aux couplages de certains noyaux (H, C_{13}) placés dans un champ magnétique: on trouve une conformation moyenne en solution (Ebel, 1999).

Le dichroïsme circulaire permet de détecter des structures secondaires dans les protéines, en particulier les hélices α et les plans β.

La cryo-microscopie à électrons est surtout utilisée pour étudier des récepteurs membranaires ainsi que certains centres photosynthétiques: ils forment des cristaux en deux dimension (planes) pouvant être visualisés par microscopie (Wouters, 1998).

Toutes ces méthodes ont leurs avantages et leurs inconvénients: la RMN permet une étude des molécules en solution (conformation "in vivo") mais est limitée par la taille des protéines. La cristallographie donne des informations sur l'empilement cristallin (forces de cohésion ou "packing"), l'aspect restrictif est ici la nécessité d'obtention de cristaux diffractables et la

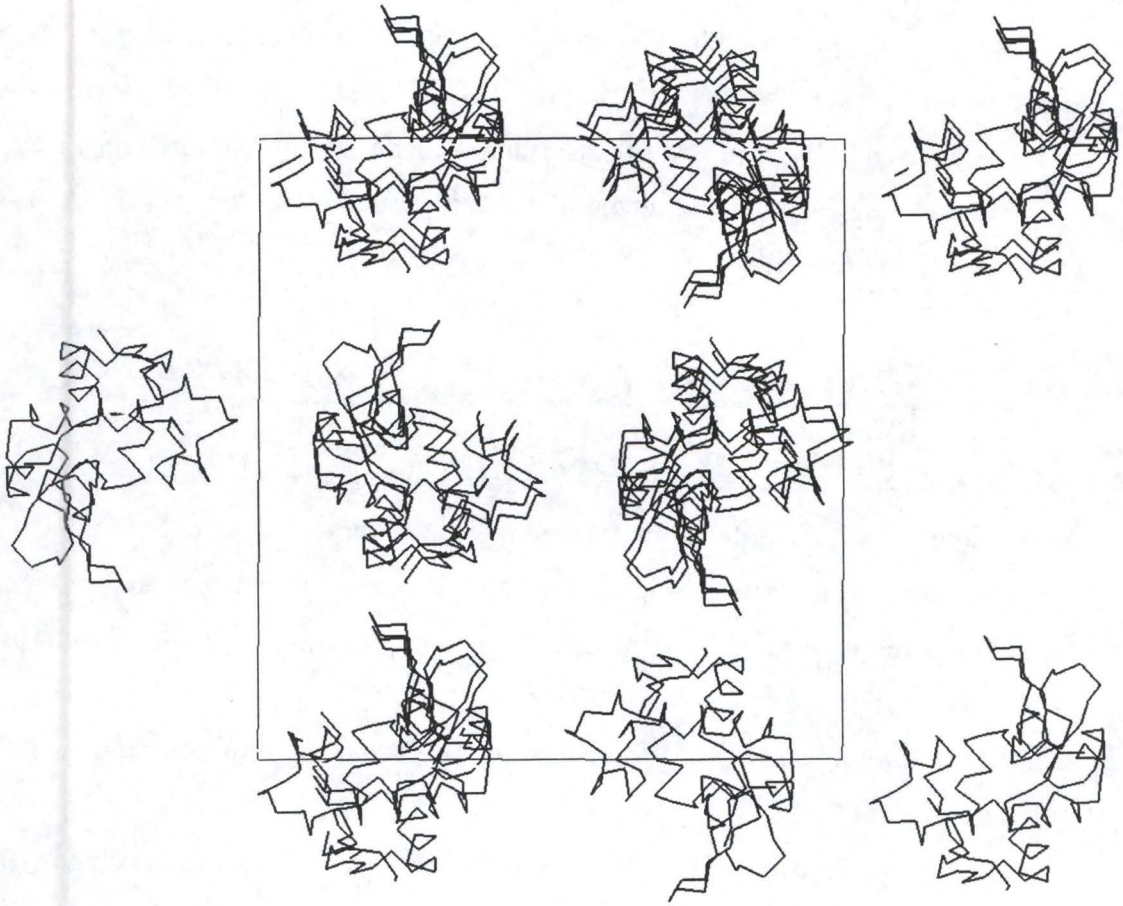


Figure 14: Représentation d'une maille d'un cristal de lysosyme.

conformation statique de la protéine cristallisée. La cryo-microscopie aux électrons souffre d'un manque de précision, tandis que le dichroïsme circulaire n'apporte rien dans le cas de petits peptides ou autres molécules sans structures secondaires.

Les progrès de l'informatique, de la biologie et de la conception du matériel (synchrotrons, RMN multidimensionnelle) ont permis de faire progresser toutes ces techniques, comme l'indique le nombre croissant de structures résolues.

6.1) Cristaux-cristallographie

Les premières études cristallographiques datent du XVIIIe siècle avec "Essai de cristallographie" (Romé de Lisle, 1772): l'existence de formes polyédriques était alors le critère essentiel pour la définition du cristal.

Ce sont les minéralogistes qui ont fondé la cristallographie en tant qu'étude des formes extérieures. R-J Hauy (1743-1822) a élargi cette définition en émettant l'hypothèse que l'existence des formes extérieures n'est qu'une conséquence d'un arrangement régulier périodique de la matière dans les solides.

Les observations macroscopiques sur les cristaux ont été mises en parallèle avec l'étude des propriétés géométriques des arrangements de la matière (fin du XIXe siècle) bien avant que ne soient utilisés les moyens d'étude à l'échelle atomique disponibles actuellement.

L. Bragg et M. von Laue ont découvert en 1912 la diffraction des rayons X et sa mise en œuvre (1914) pour des études de structure, apportant la preuve expérimentale directe de la nature périodique de l'arrangement cristallin.

Une série de techniques de diffraction: des rayons X, d'électrons et de neutrons permettent de déterminer non seulement la forme et les dimensions de la maille cristalline[3] mais également l'arrangement précis des atomes à l'intérieur de cette maille (figure 14).

Rares sont les solides qui ne soient pas ordonnés à l'échelle atomique, la plupart d'entre eux sont donc "anisotropes" au niveau atomique et moléculaire. Cette périodicité, entraîne souvent des propriétés de symétrie à l'échelle microscopique qui se manifeste à l'échelle macroscopique dans les propriétés physiques des cristaux.

La plupart des solides cristallins se présentent sous forme de polycristaux, c'est-à-dire d'agrégats de microcristaux plus ou moins gros orientés aléatoirement les uns par rapport aux autres. Dans cet état, étant donné la taille des polycristaux (autour du mm), on ne peut pas percevoir d'anisotropie.

---

[3] *La maille cristalline est le plus petit parallélépipède de matière, répété périodiquement dans le cristal*

Les critères principaux pour qu'un cristal soit potentiellement diffractable sont la durée de vie assez importante (il ne doit pas s'effriter pendant la diffraction) et la taille du monocristal[4] (un polycristal est inutilisable pour la cristallographie).

6.2) Cristaux de petites molécules organiques et de protéines

Depuis la découverte de la diffraction des cristaux par L.Bragg, la structure de nombreuses molécules a été résolue (il existe une banque de donnée, la C.S.D., Cambridge Structure Database où plus de 170000 coordonnées sont stockées) (Duée and Dideberg, 1997).

Pour de nombreuses petites molécules organiques, il est très facile d'obtenir des cristaux de dimension raisonnable (utilisables pour la diffraction), les conditions de cristallisation sont facilement reproductibles, et les cristaux ont une durée de vie allant de 1 mois à plusieurs années à l'air libre. La diffraction de tels cristaux est assez courante: le matériel utilisé n'est pas très coûteux (diffractomètre "classique") et les conditions de travail sont peu astringentes.

La détermination de la structure d'une série de molécules organiques va de pair avec la résolution de structures de protéines. Ces substances ont une série de rôles biologiques: inhiber ou activer une enzyme en se glissant dans le site actif, bloquer un site de phosphorylation. La détermination de la structure de ces molécules peut aider à trouver les sites importants pour la reconnaissance protéique intra- et extra-cellulaire, et inversement, la détermination de la structure des protéines peut aider à trouver de nouvelles substances "médicamenteuses"[5] (Wouters, 1998).

La cristallographie des protéines existe depuis de nombreuses années: en 1934 le premier cliché de diffraction de rayons X a été réalisé en Grande-Bretagne par l'équipe de J.D. Bernal sur un cristal de pepsine, mais ce n'est que depuis les années '60 que l'on peut voir apparaître une série de structures révélées par DRX (myoglobine, hémoglobine, lysozyme).

En 1976, on recensait quelque 75 protéines considérées alors "à haute résolution"[6] , alors que seulement 7 de ces structures avaient une résolution supérieure à 2 Å. A l'heure actuelle, environ 9500 structures de protéines sont disponibles dans la banque PDB (Protein Data Bank, conçues pour archiver les structures tridimensionnelles de macromolécules obtenues par DRX ou RMN-2D) (Bernstein, 1977). La plupart de ces coordonnées structurales ont une

---

[4] *On considère que, pour un cristal de protéine, la taille minimale diffractable est de 0,2 mm*

[5] *Médicament : toute substance organique présentée pour traiter, soigner ou prévenir la maladie chez l'animal*

[6] *La résolution est la capacité de distinguer 2 points, si la résolution n'est pas bonne, ces points sont trop proches et se confondent en 1 seul*

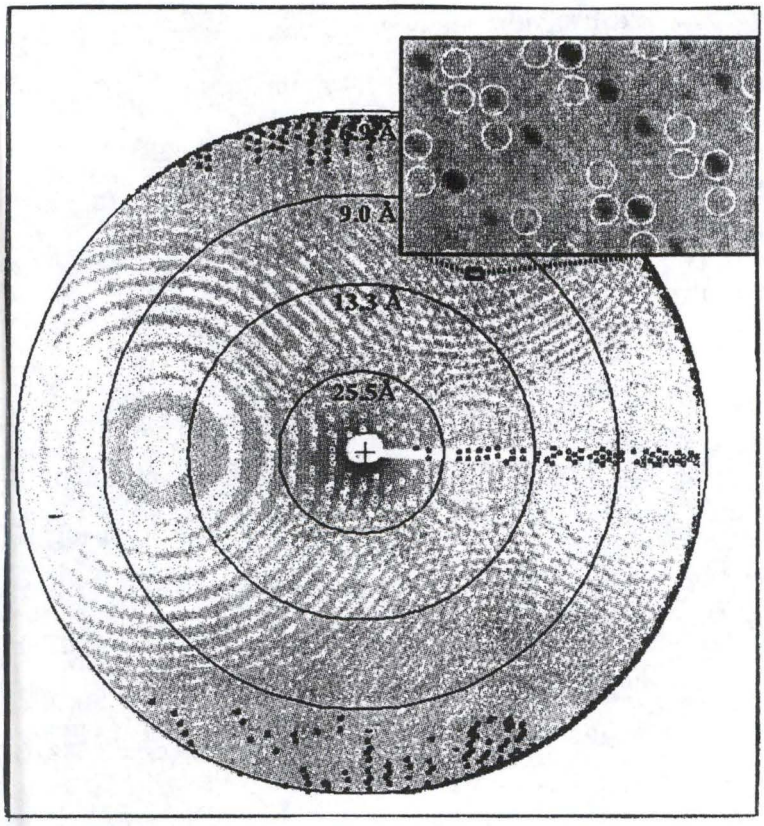


Figure 15: Diagramme de diffraction et son indexation obtenus à partir d'un cristal du "cœur" du "bluetongue virus". Cette image contient environ 20 000 taches de diffraction. (Gouet & al., 1996).

résolution meilleure que 2.5 Å. De plus, les structures de protéines de plus en plus complexes sont résolues: le "blue tongue virus" avec plus de 1000 chaînes peptidiques, est actuellement le plus gros assemblage macromoléculaire cristallisé (figure 15) (Grimes, 1997).

6.3) La cristallogenèse des protéines

Le principe de la cristallisation est de faire évoluer la protéine vers un état "superprécipité" ou plus exactement dans une solution "sursaturée" en protéine qui va progresser vers la formation de cristaux.

Il existe de nombreuses techniques de cristallisation: par diffusion de vapeur, par interface de diffusion (par dialyse) ou par des méthodes en "batch". Les 2 méthodes les plus utilisées sont les méthodes de dialyse et les méthodes par diffusion de vapeur. La méthode par dialyse permet une variation aisée des différents paramètres qui influencent la cristallisation. La macromolécule biologique est séparée d'un grand volume de solvant par une membrane semi-perméable qui permet le passage de petites molécules mais pas celui de la macromolécule. La méthode de cristallisation pour de petites quantités de protéines la plus utilisée est la méthode par diffusion de vapeur: une goutte qui contient la macromolécule biologique (et d'autres constituants de la liqueur-mère) est "équilibrée" avec un réservoir qui contient la liqueur-mère (les agents précipitants y sont plus concentrés que dans la goutte). L'équilibrage se fait par diffusion des composés volatiles jusqu'à ce que la pression de vapeur dans la goutte soit la même que celle du réservoir. Le même principe est utilisé pour les différentes techniques que sont les "hanging drops" ("gouttes suspendues"), "sitting drops" et "sandwich drops" (figure 16).

On considère généralement que la cristallogenèse est assez empirique, et dépend de nombreux facteurs tels que la concentration en protéine et en composants du milieu (sels, tampons, précipitant, additifs divers), la température, le pH, la nature des composants. Il existe des "kits" échantillonnés sur des conditions de cristallisation "standard" qui permettent d'approcher le milieu "idéal" de cristallisation. Il faut par la suite affiner les résultats (améliorer la taille et le nombre de cristaux) en faisant varier chaque facteur. Pour arriver à définir de manière précise les conditions de cristallogenèse, il faut donc disposer d'une quantité suffisante de protéines de grande pureté (un contaminant en concentration très faible peut empêcher la cristallogenèse), la pureté de l'échantillon est contrôlée par microséquençage ou par spectrométrie de masse avant les essais de cristallisation. Les premières protéines, dont la structure a été résolue, sont donc celles que l'on pouvait trouver en grande quantité et dont les conditions de cristallisation étaient facilement reproductibles. Une série de protéines dont

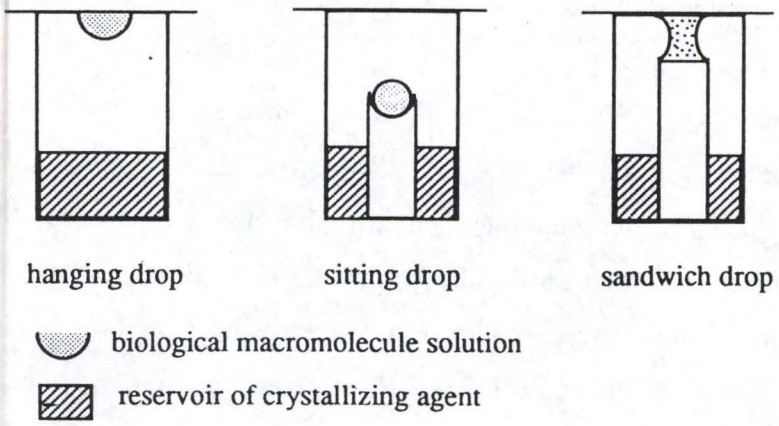


Figure 16: Représentation schématique des techniques de diffusion de vapeur : « hanging drop », « sitting drop » et « sandwich drop ». (Ducruix & Giegé , 1992).

les protéines membranaires, du fait de leur caractère peu soluble en milieu aqueux, sont difficiles à cristalliser: elles ne sont stables que dans un milieu dont les conditions physico-chimiques sont proches d'une membrane biologique, il faut les solubiliser avec un détergent doux choisi après de nombreux essais. Ces protéines sont depuis quelques années étudiées: la première protéine membranaire cristallisée a été le centre réactionnel photosynthétique de *Rhodopseudomonas viridis* en 1987 (Landau and Rosenbush, 1996).

6.4) La diffraction de cristaux de protéines

La diffraction d'un cristal par un faisceau monochromatique de rayons X donne un faisceau diffracté caractérisé par son amplitude et sa phase et par sa direction (décrite par les indices h, k, l) (Weiss, 1997).

Une opération mathématique, la synthèse de Fourier, permet de calculer à partir de l'amplitude et la phase de tous les rayons diffractés la carte de densité électronique de la protéine à l'intérieur du cristal.

L'information concernant la phase du rayon diffracté n'est pas directement accessible, tandis que l'intensité, grandeur proportionnelle au carré de l'amplitude est mesurable par les systèmes de détection des ondes.

Ce " problème de phase " peut être résolu par différentes méthodes (au nombre de quatre): le remplacement isomorphe multiple (MIR), la méthode de diffusion anormale multiple (MAD), les méthodes directes, et le remplacement moléculaire.

Le remplacement isomorphe multiple utilise des atomes lourds comme le Plomb, l'Uranium, le Platine, le Sélénium. Ils se fixent en des sites précis de chaque molécule de protéine du cristal et modifient l'intensité des taches de diffraction dans des longueurs d'ondes des rayons X caractéristiques de l'atome lourd. On doit disposer pour cette méthode de minimum 2 cristaux dérivés différents, ce qui apporte une source d'erreur en plus et de nombreux paramètres à ajuster.

La méthode de diffusion anormale, plus récente, n'utilise qu'un seul cristal. En faisant varier finement la longueur d'onde au voisinage du seuil d'absorption de l'atome lourd (via un rayonnement dit "synchrotron") à 3,4 longueurs d'ondes différentes, il est possible de déterminer la phase. La seule source d'erreur serait ici l'erreur expérimentale sur les mesures d'intensité.

Les méthodes directes, à l'heure actuelle seulement utilisées pour déterminer la structure de petites molécules, sont basées sur des relations statistiques entre les phases de différentes

réflexions. Ces méthodes n'utilisent elles aussi qu'un seul cristal, mais exigent que l'on dispose de données à haute résolution.

Enfin, le remplacement moléculaire est utilisable lorsque l'on dispose de la structure d'une protéine proche au niveau de la séquence. On essaie de placer la molécule connue dans le cristal de manière à ce que les intensités de diffraction soient les plus proches des intensités observées.

L'introduction de résidus particuliers, par des techniques de biologie, s'est révélée très importante pour l'étude des figures de diffraction: introduction de Sélénium comme diffuseur anormal (MAD) sous forme de séléno-méthionine en clonant le gène de la protéine dans un organisme incapable de synthétiser la méthionine dans un milieu riche en Se-Met, remplacement de thréonines par des cystéines pour pouvoir fixer du Zinc ou du Plomb (Pares *et al.*, 1996).

6.5) L'apport de la biologie

L'essor des techniques de génie génétique (comme la surexpression où un gène mis sous la dépendance d'un promoteur fort) et des techniques de purification (par exemple une protéine avec une queue poly-histidine sur une colonne de nickel), permet l'obtention d'une grande quantité de protéines d'intérêt industriel ou thérapeutique. Ceci étant surtout valable pour les molécules provenant d'organismes dont la culture est difficile.

Il est également possible de modifier la protéine (via l'ADN) afin de faciliter la purification (queue de poly-His), d'homogénéiser le produit, d'introduire des résidus nécessaires à la cristallographie ou la cristallogenèse.

Obtenir des cristaux de protéines ayant subit des phosphorylations et des glycosylations post-traductionnelles n'est pas facile, étant donné l'inhomogénéité d'un tel produit. Ce problème, ainsi que l'extraction de certaines protéines (solubilisation de protéines membranaires, purification d'un broyat pour obtenir une protéine en faible quantité) est résolu par le clonage en bactérie ou en levure.

D'autres techniques utilisées en parallèle, comme la mutagenèse dirigée, dans le cas d'interactions enzyme-substrat permettent de localiser les résidus importants pour la reconnaissance, la catalyse et la thermostabilité, et de faire une étude cristallographique en comparant *de visu* les différences au niveau atomique.

Il y a donc sans cesse des échanges d'informations entre biologistes et cristallographes: la production de protéines sur mesure et la production de mutants défectifs sont nécessaires aux

cristallographes, tandis que l'importance de la résolution de la structure pour affiner et comprendre la fonction d'une protéine n'est plus à démontrer (Wouters, 1998).

6.6) Les progrès parallèles de la bioinformatique et de la diffraction

Le développement de la cristallographie des protéines s'est fait en parallèle avec l'essor des outils informatiques. La capacité de calcul est primordiale à toutes les étapes: enregistrement et traitement des données, calculs cristallographiques, archivage des données, la communication entre chercheurs. L'apparition des consoles graphiques aide à la construction des modèles structuraux et à leur manipulation (voir "matériel et méthode": outils insightII).

Alors que les diffractomètres existant dans les années '80 mesuraient une seule intensité à la fois, ceci étant suffisant pour de petites molécules organiques, les diffractomètres utilisés aujourd'hui ("bidimensionnels") mesurent des centaines d'intensités simultanément. La durée de vie des cristaux biologiques soumis à une source de rayons X a été améliorée par la cryocristallographie (le cristal baigne dans un flux d'azote à 100 K). Enfin la source de rayonnements s'est sensiblement améliorée: les derniers grands instruments tels que les synchrotrons "de 3e génération"[7] créent des flux de rayons X plus intenses et des faisceaux plus fins couvrant une large gamme de longueurs d'onde (0.3-2 Å).

La forte intensité du faisceau permet des collectes rapides et l'exploitation de cristaux très petits.

D'autres améliorations sont en cours comme la microfocalisation du faisceau et l'étude d'intermédiaires d'une réaction enzymatique.

La mise au point de la structure de la bactériorhodopsine, de certains complexes de la catalase pour les protéines, de l'ARN de transfert, des ribozymes, du nucléosome pour les acides nucléiques (Luger, 1997) sont autant de résultats de l'amélioration des techniques citées ci-dessus.

6.7) La cristallisation d'anticorps et de peptides

Plusieurs fragments Fab (il est difficile de cristalliser un anticorps complet) et de nombreux complexes Fab-peptide ont été cristallisés. Leurs structures sont disponibles dans la banque PDB.

La cristallisation de petits peptides n'est pas très fréquente. Il est difficile d'obtenir des cristaux de telles molécules adoptent de nombreuses conformations (la chaîne carbonée d'un

[7] *un synchrotron est un anneau de stockage où circulent des particules à une vitesse proche de la lumière*

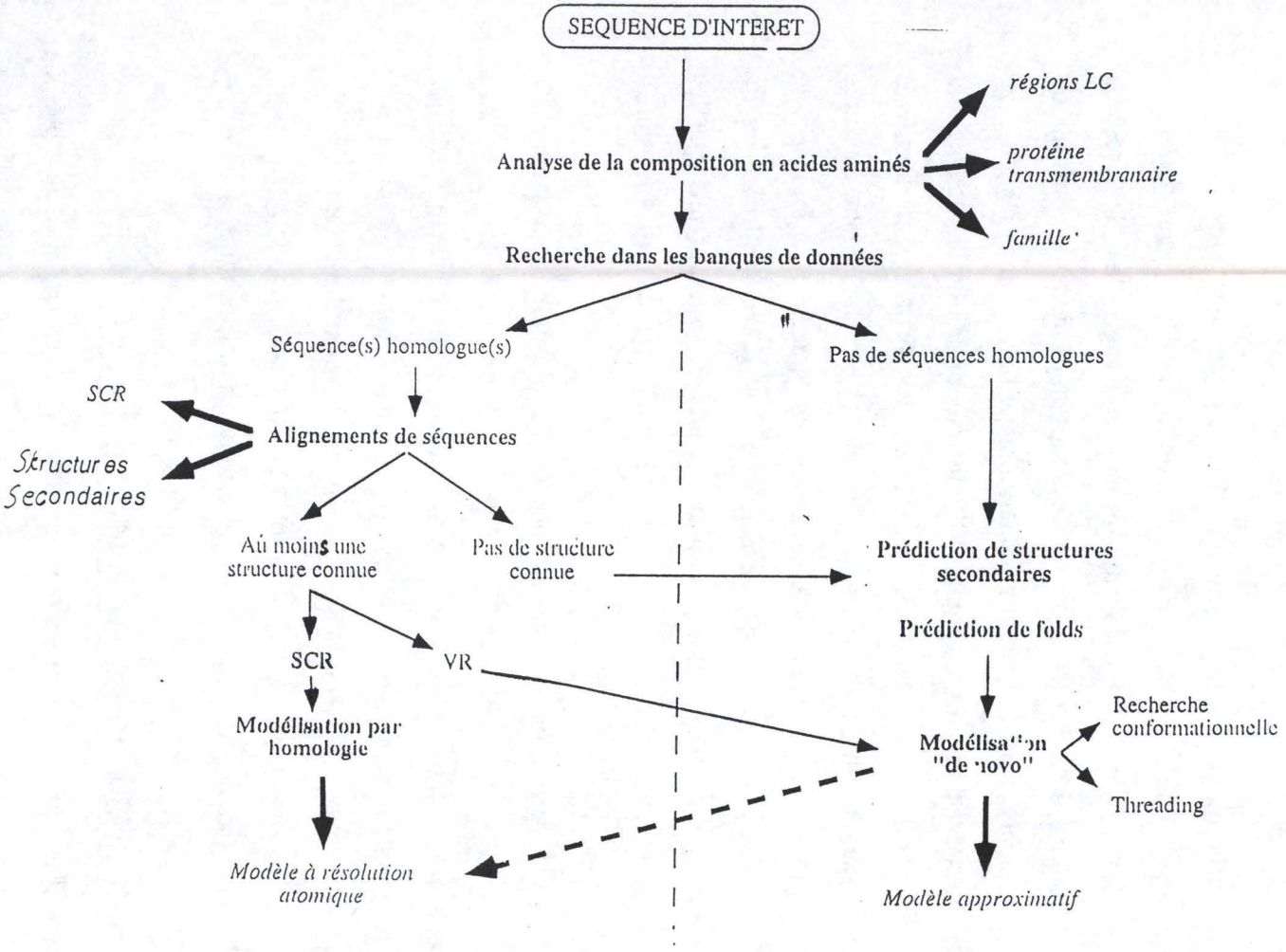


Figure 17: Schéma représentant la démarche utilisée pour modéliser une protéine à partir de sa séquence. (Vinals, 1997).

peptide en solution est assez flexible). On privilégie la RMN-2D: la conformation calculée sera une moyenne des conformations adoptées par le peptide.

Pour déterminer la structure des peptides par DRX, certains auteurs préconisent l'utilisation de protéines "carrier" ("porteuses"), dont on connaît déjà la structure, qui sont faciles à cristalliser et dont les résidus ne modifieraient pas la conformation que pourrait prendre le peptide seul (voir "matériel et méthode") (De Bolle, 1999).

En ce qui concerne la cristallisation de complexes Fab-peptides, il existe une méthode dite d' "ensemencement croisé" ("cross- seeding").En dispersant le complexe Fab-antigène dans le liquide où baigne le cristal de Fab, il est possible de le "glisser" dans le cristal à condition que le ligand ne soit pas trop grand et que la conformation du Fab en complexe soit proche de sa conformation en solution. On peut également utiliser un cristal de Fab que l'on plonge dans une goutte où se trouve le complexe. L'arrangement cristallin du complexe sera calqué sur celui du Fab seul.

Cette méthode est également utilisée pour agrandir des cristaux d'une protéine donnée: dans ce cas, la protéine en solution est identique à la protéine cristallisée. Ce type de méthode est appelé "ensemencement" ("seeding").

Il faut savoir que des études de structure de la protéine pVIII ont déjà été réalisées et ont donné des résultats satisfaisants (Opella *et al.*, 1987). Des phages ont également été cristallisés, mais la résolution était trop basse que pour pouvoir déterminer la structure de peptides situés en fin de chaîne sur les protéines de capside.

7) MODELISATION MOLECULAIRE: METHODES THEORIQUES POUR LA DETERMINATION DE STRUCTURE DES PROTEINES

La structure de nombreux biocomposés n'est pas encore connue via les techniques expérimentales comme la diffraction aux rayons X, la résonance magnétique nucléaire ou le dichroïsme circulaire. Il existe différents outils de prédiction pour approcher les structures secondaires et tertiaires de ces molécules (Vinals, 1997).

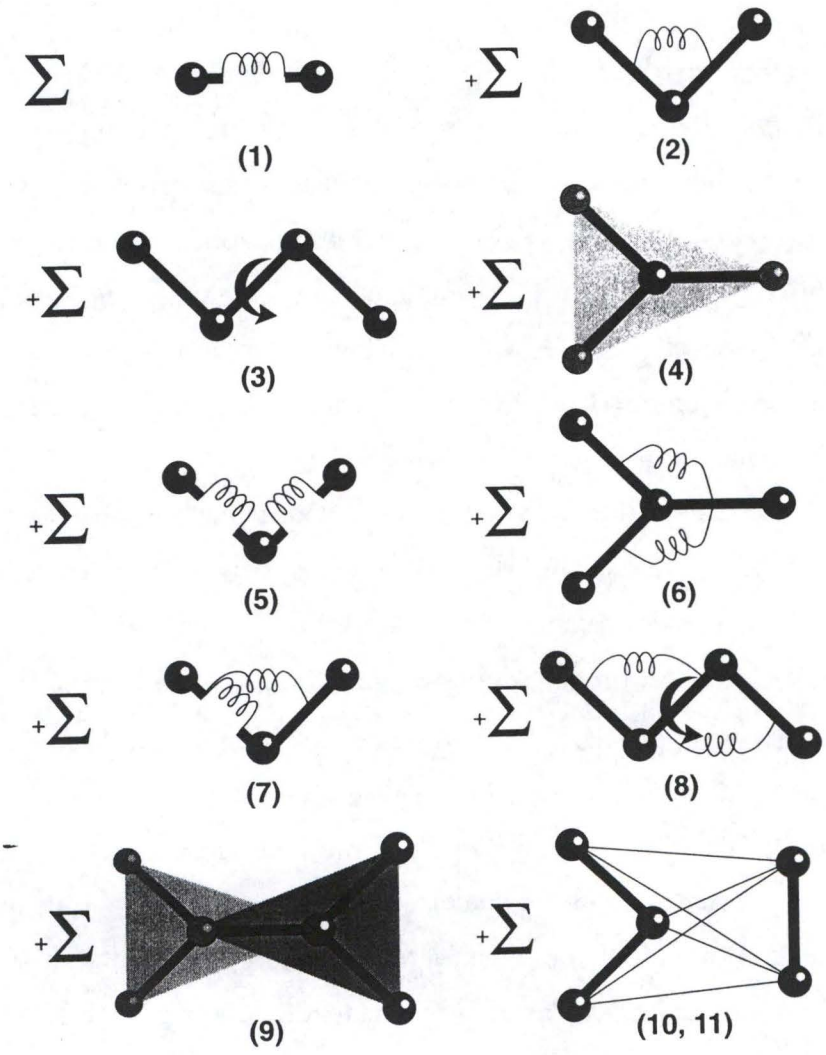
La première étape est l'analyse de la composition de la séquence d'intérêt en vue de trouver des domaines typiques de familles de protéines (figure 17): cette séquence est comparée à d'autres séquences connues reprises dans des banques de données. Le but consiste à détecter une similarité significative entre les séquences de la protéine cible et une séquence de protéine

pour laquelle une structure a été déterminée par cristallographie ou par RMN. Il est possible de modéliser par homologie sur base des séquences conservées (s'il y a plus de 20 % de similarité): la séquence de la protéine d'intérêt est "collée" de manière adéquate (afin de trouver un alignement qui correspond à un modèle énergétiquement favorable) sur la structure connue. Les séquences variables seront modélisées "de novo" (voir ci-dessous). Le modèle obtenu a une précision qui dépend du pourcentage d'identité entre la séquence de la protéine d'intérêt et la séquence de la protéine de structure connue.

S'il n'existe pas de séquences homologues ou si nous avons à notre disposition des séquences homologues sans structure connue, le modèle sera nettement plus approximatif et sera élaboré à partir de prédictions de structures secondaires (motifs caractéristiques de plans β, d'hélices α), de prédictions de structures super-secondaires (il existe aussi de nombreux motifs conservés, notamment chez les globines) ou de modélisation "de novo" (de Fays *et al.*, 1999). Ce type d'approche utilise le "threading" et la recherche conformationnelle. Le threading est une méthode qui "enfile" la séquence d'une protéine sur une série de structures tertiaires connues afin de déterminer un repliement énergétiquement compatible avec cette séquence. La recherche conformationnelle, quant à elle, "explore" l'espace conformationnel et trouve une série de conformations stables de la molécule. Elle est principalement utilisée pour affiner la structure de grosses molécules ou pour déterminer la structure globale de petits ligands (Vinals, 1997).

7.1) Champs de force

Lorsque l'on considère un système moléculaire comprenant un certain nombre de molécules, il est difficile de faire une description mathématique précise de l'énergie du système étant donné le nombre considérable de calculs nécessaires pour caractériser tous les mouvements électroniques. Par contre, la mécanique moléculaire, méthode de calcul par champs de force, ne tient compte que des positions atomiques du système pour en calculer l'énergie (MSI, 1996). Le principe est l'étude d'une gamme étendue de propriétés structurales en utilisant un ensemble de paramètres dans une équation décrivant la surface d'énergie potentielle, en fonction des coordonnées relatives des atomes. Ces équations et leurs paramètres constituent le champ de force et sont testés sur un certain nombre de situations connues. Le type atomique (carbone sp2-sp3, aromatique-aliphatique), les coordonnées (X, Y, Z) des différents atomes de la structure (les fichiers ".car") ainsi que le potentiel énergétique et la charge partielle pour chacun de ces atomes (les fichiers ".mdf") sont les principales données nécessaires à la



$$E_{pot} = \sum_{b} D_b\left[1 - e^{-\alpha(b-b_0)}\right] + \sum_{\theta} H_\theta (\theta - \theta_0)^2 + \sum_{\phi} H_\phi [1 + s\cos(n\phi)]$$

(1) (2) (3)

$$+ \sum_{\chi} H_\chi \chi^2 + \sum_{b}\sum_{b'} F_{bb'}(b-b_0)(b'-b'_0) + \sum_{\theta}\sum_{\theta'} F_{\theta\theta'}(\theta-\theta_0)(\theta'-\theta'_0)$$

(4) (5) (6)

$$+ \sum_{b}\sum_{\theta} F_{b\theta}(b-b_0)(\theta-\theta_0) + \sum_{\phi} F_{\phi\theta\theta'}\cos\phi\,(\theta-\theta_0)(\theta'-\theta'_0) + \sum_{\chi}\sum_{\chi'} F_{\chi\chi'}\chi\chi'$$

(7) (8) (9)

$$+ \sum \varepsilon\left[(r^*/r)^{12} - 2(r^*/r)^6\right] + \sum q_i q_i / \varepsilon r_{ij}$$

(10) (11)

Figure 18: Equation et illustration des paramètres du champ de force CVFF
(1): modification de la longueur de la liaison
(2): modification de l'angle de la liaison
(3): modification de l'angle de torsion
(4): modification de l'interaction "hors-plan"
(5) à (9): couplage des différentes modifications 1-2-3 et 4
(10) et (11): modification des interactions non-liantes de type Van der Waals et Coulombique. (MSI, 1996)

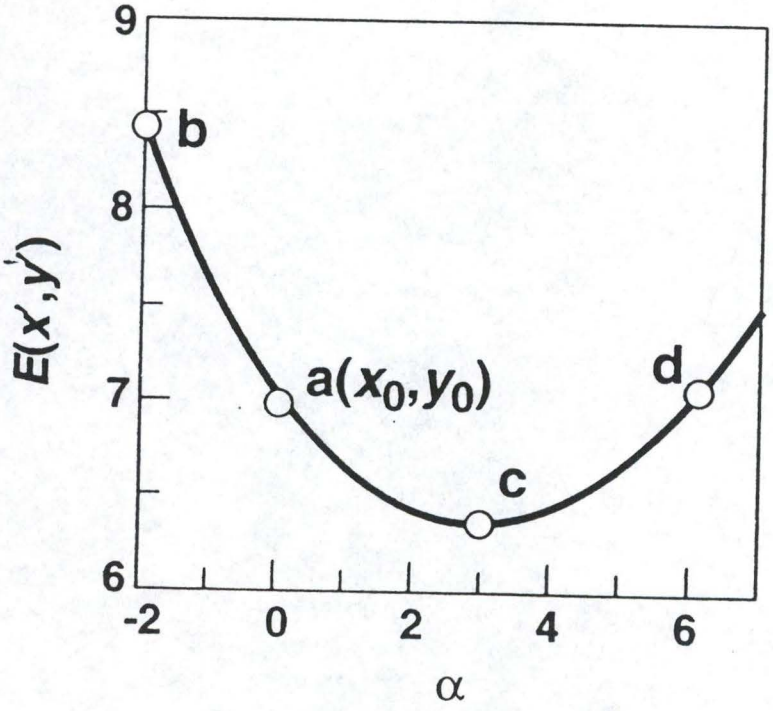


Figure 19: Intersection de la surface d'énergie définie par l'algorithme de recherche "Iine search" La variable indépendante α est un paramètre uni-dimensionnel qui est ajusté dans le but de minimiser la valeur de la fonction E (x', y'). Le point a correspond au point initial (quand α = O) et le point c est le minimum local uni-dimensionnel. Le vecteur passe d'abord par le point d. Comme il est plus énergétique, la recherche continue de l'autre côté, passe par le point c, et continue jusqu'au point d. Etant donné qu'il est également trop énergétique, le vecteur de recherche revient au point c , qui devient le départ d'une recherche itérative.(MSI, 1996).

caractérisation d'une molécule. Le potentiel de tous les atomes est décrit par une équation paramétrisée: le champ de force. (figure18)

Pour de nombreux systèmes biologiques, il existe des champs de force idéalement paramétrisés: pour les peptides et les acides nucléiques "AMBER"; pour les protéines et les peptides: "CVFF", pour les petites molécules organiques "CFF91" et pour tout le tableau périodique: "ESFF" (en développement, peu testé et peu caractérisé). Tous ces champs de force sont proposés par le programme Discover.

7.2) Minimisation de l'énergie potentielle

La manière dont l'énergie potentielle, fonction complexe et multidimensionnelle des coordonnées d'un système, varie avec ses coordonnées est appelée "surface d'énergie potentielle" (MSI, 1996).

Au cours d'une minimisation par mécanique moléculaire, on essaie d'atteindre les points les plus bas de cette surface d'énergie. Ceux-ci correspondent aux arrangements des atomes dans les états les plus stables de la molécule. N'importe quel mouvement qui s'écarte d'un de ces minima donne naissance à une conformation de plus haute énergie, donc moins stable (figure 19).

L'énergie potentielle d'une molécule est calculée en fonction des coordonnées de chaque atome, et du champ de force utilisé.

La minimisation d'énergie est un ajustement des coordonnées des atomes afin de réduire (minimiser) l'énergie totale du système considéré. C'est un processus itératif, où chaque itération consiste à calculer et à appliquer un mouvement à un atome en fonction de la surface d'énergie potentielle calculée. En effet, il n'y a pas de solution analytique à la dérivée seconde de l'énergie potentielle car la fonction est trop complexe, ce qui explique qu'on soit obligé de recourir à des méthodes itératives. La dérivée de l'énergie potentielle peut néanmoins être calculée pour chaque atome, et si la structure obtenue est proche d'un minimum d'énergie, la moyenne de cette dérivée pour tous les atomes du système sera proche de zéro. Au cours de la minimisation d'énergie, les itérations sont répétées jusqu'à l'obtention d'un minimum d'énergie. Ce minimum d'énergie peut être local, c'est-à-dire qu'il ne correspond pas au minimum global, qui est la valeur minimum que peut prendre l'énergie potentielle au système considéré.

Les valeurs des paramètres correspondant au minimum d'une fonction sont celles annulant la dérivée première de cette fonction. On doit estimer ces valeurs par des méthodes numériques

(il est impossible de trouver des solutions analytiques pour la dérivée première d'une fonction telle que celle décrivant l'énergie potentielle d'une protéine).

Les méthodes de minimisation d'énergie ont généralement deux composantes majeures: un algorithme de calcul de la dérivée de l'énergie, qui est un vecteur donnant la direction du minimum dans l'hypersurface du potentiel, et un algorithme de recherche du minimum dans la direction du vecteur: "line search".

Les coordonnées atomiques sont donc progressivement modifiées de manière à ce que l'énergie globale de la molécule tende à diminuer le plus rapidement possible.

L'efficacité d'un algorithme de minimisation dépend du temps nécessaire pour annuler la dérivée première et du nombre d'ajustements structuraux (itérations) nécessaires pour converger vers un minimum.

La méthode "steepest descent" permet un déplacement sur la surface d'énergie le long d'un gradient, jusqu'à ce que la fonction passe par un minimum (le gradient y est recalculé). Le nouveau gradient sera perpendiculaire au précédent, c'est-à-dire que les directions empruntées lors des minimisations successives sont orthogonales entre elles. Cette méthode va dès lors provoquer une oscillation autour du minimum très rapidement.

Ces caractéristiques en font une méthode puissante mais imprécise, principalement utilisée en début de minimisation. Ensuite, une autre méthode plus précise, les "conjugate gradients", est utilisée. Elle permet d'augmenter la vitesse de convergence en fin de minimisation. On préconise parfois l'utilisation d'une autre méthode "VA09A" (méthode de Newton-Raphson), aux caractéristiques analogues à l'algorithme " conjugate gradient ".

Quel que soit le protocole de minimisation, différentes données doivent être fixées:

1) Il est impossible d'atteindre la dérivée première nulle: pour définir une structure "minimisée", l'utilisateur doit fixer des critères de convergence (la dérivée maximale ou moyenne -ou les deux- de la molécule doit être inférieure à un seuil exprimé en kcal/mole Å).

2) Il faut définir un nombre d'itérations pour chaque minimisation dépendant du type d'algorithme, de la taille de la molécule, et du champs de force utilisé.

3) Il est possible de biaiser la minimisation, en ne tenant pas compte de certains paramètres à priori peu importants: la fonction de Morse (qui décrit les liaisons covalentes) peut être simplifiée, on peut ne pas tenir compte des termes mixtes et des charges dans des cas particuliers (lorsqu'ils ont peu d'influence sur la conformation de la molécule).

4) Des contraintes sont prises en compte pour "diriger" la minimisation, par exemple, si un substrat est lié à un enzyme via un pont disulfure, les atomes impliqués peuvent être fixés.

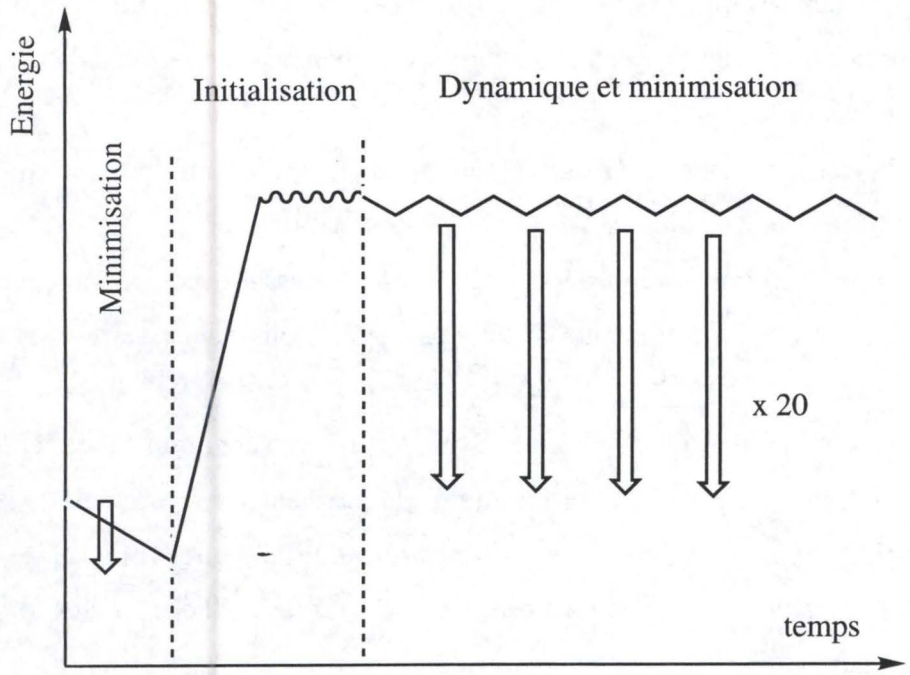


Figure 20: Graphique représentant le déroulement d'une dynamique moléculaire suivie d'une minimisation d'énergie: la molécule d'intérêt est d'abord minimisée. Une quantité importante d'énergie est donnée à cette conformation (par chauuffage): c'est l'étape d'initialisation. Vient ensuite la dynamique moléculaire proprement dite: 20 conformères de cette molécule sont générés, potentiellement représentatif de l'ensemble des conformation que cette molécule pourrait prendre dans des conditions prédéfinies. Ces structures sont ensuite minimisées dans le but d'obtenir 20 conformères stables.

Il faut garder à l'esprit la signification théorique d'une structure d'énergie minimum. Pour tous les champs de force, l'énergie zéro est arbitraire, on ne peut pas directement comparer les énergies potentielles de molécules différentes. La seule comparaison valable est entre molécules d'un même système, minimisées suivant le même protocole.

7.3) Dynamique moléculaire

Le but de la dynamique moléculaire est de franchir des barrières énergétiques afin de générer un grand nombre de conformations différentes qui pourront être analysées par la suite (figure 20).

Le passage de barrière se fait en augmentant l'énergie du système par chauffage (énergie cinétique additionnelle). La minimisation se déroule à 0 K (on ne tient compte que de l'énergie potentielle du système), et ne permet de trouver que des minima locaux. Il ne faut pas oublier que les objectifs de la minimisation sont différents de la dynamique mais sont complémentaires: la minimisation est surtout utilisée pour affiner la conformation de molécules dont on a préalablement déterminé la structure par des méthodes expérimentales (par DRX) ou pour "préparer" une dynamique (afin de ne pas avoir de structures aberrantes ou trop énergétiques).

La dynamique moléculaire permet de générer une série de conformères d'un système par intégration des équations classiques de Newton. Il en résulte une trajectoire qui précise la façon dont varient la position et la vitesse des particules en fonction du temps, la résolution de l'équation mouvement de Newton peut donner la trajectoire.

$F_i(t) = m_i\, a_i(t)$

F représente la force agissant sur l'atome i, de masse m et subissant une accélération a. Cette force peut être directement précisée via la dérivée de l'énergie potentielle en fonction des coordonnées.

Le "pas", ou intervalle de temps entre 2 changements de conformations est souvent arbitrairement fixé à un femtoseconde (1fs).

La première étape de la simulation de dynamique moléculaire est l'équilibration: on peut considérer que le système est équilibré quand l'énergie, la température, et la pression fluctuent autour d'une valeur moyenne. Le système a dès lors une conformation cohérente à la température à laquelle la simulation est effectuée.

La deuxième étape est la collecte des données, à des intervalles de temps prédéfinis, (par ex. 1picoseconde, soit 1 ps) une série de conformations (aussi appelées "conformères") sont "stockées" et ensuite minimisées ("mécanique moléculaire"). Ces conformations

correspondent à une série de minima locaux, certaines de ces conformations pouvant potentiellement exister dans les conditions simulées. Le nombre de conformères et l'intervalle de temps entre chaque collecte sont fonction du système et des informations que l'on veut extraire de la simulation.

Une série de valeurs expérimentales (température et volume) peuvent être gardées constantes pour mimer des conditions particulières.

7.4) Modélisation de peptides seuls ou en complexe avec un anticorps

La dynamique moléculaire peut être utilisée pour modéliser la structure de petites molécules comme des peptides. Lorsque l'on désire déterminer la conformation de peptides lors de la reconnaissance par un anticorps, on considère que la conformation adoptée par le peptide au sein du site de reconnaissance est proche d'une des conformations qu'il pourrait avoir en solution (cfr. "complexes anticorps-antigènes"). Le protocole utilisé lors de la modélisation de peptides en complexe sera donc assez proche d'un protocole déterminant la conformation de peptides seuls: par exemple une des différences sera la constante diélectrique utilisée car le peptide se trouve en grande partie dans une poche hydrophobe au sein du paratope.

Objectif du mémoire

Au sein de l'unité de recherche en Biologie Moléculaire (URBM), des peptides capables de se fixer au paratope d'anticorps monoclonaux anti-LPS de *Brucella* ont été obtenus (Walgraffe, 1998), (De Bolle, 1999). L'anticorps monoclonal A15-6B3, spécifique des LPS M-dominant, a ainsi permis l'isolement de nombreuses séquences peptidiques, qui se fixent au paratope de l'anticorps. Ces peptides sont des mimotopes antigéniques. Certains de ces peptides (appelés 6B3-[1], 6B3-[2], 6B3-[4] et 6B3-[6], et 6B3-[X] pour l'ensemble) ont été testés quant à leurs propriétés immunogéniques, et quelques-uns ont été capables de générer une réponse humorale anti-LPS après injections à des souris. Néanmoins la réponse immune est de très faible amplitude et seule une faible proportion des souris réagit contre le LPS. Il convient donc d'améliorer l'efficacité des mimotopes du LPS. Nous pensons que cette amélioration passe par une meilleure compréhension des interactions des anticorps avec les épitopes naturels (sur le LPS) et avec les peptides qui les miment.

Notre travail a consisté à proposer des conformations que peuvent prendre les peptides 6B3-[X] au sein du site de reconnaissance de l'anticorps. En effet, ces peptides, outre la présence d'un pont disulfure, sont très riches en résidus proline. Il est dès lors possible que ces caractéristiques rendent les peptides suffisamment rigides que pour qu'ils adoptent une série de conformations préférentielles. Les structures proposées ont été générées par modélisation moléculaire.

Une autre partie de notre travail a consisté à tenter de mettre au point les conditions qui permettent l'obtention de cristaux pour deux types de protéines. Le premier type de protéines consiste en une série de fusions entre la thiorédoxine de *E. coli* et les peptides 6B3-[X]. Le second type de protéines consiste en des fragments Fab d'anticorps anti-LPS.

Les peptides 6B3-[X], lorsqu'ils sont présentés en phages, sont reconnus par l'anticorps monoclonal A15-6B3. Par contre, lorsqu'ils sont fusionnés à la thiorédoxine, ils ne sont plus reconnus par l'anticorps. La résolution de la structure des protéines de fusion devrait permettre de comprendre pourquoi les peptides ne sont plus reconnus dans ce contexte.

RESULTATS ET DISCUSSIONS

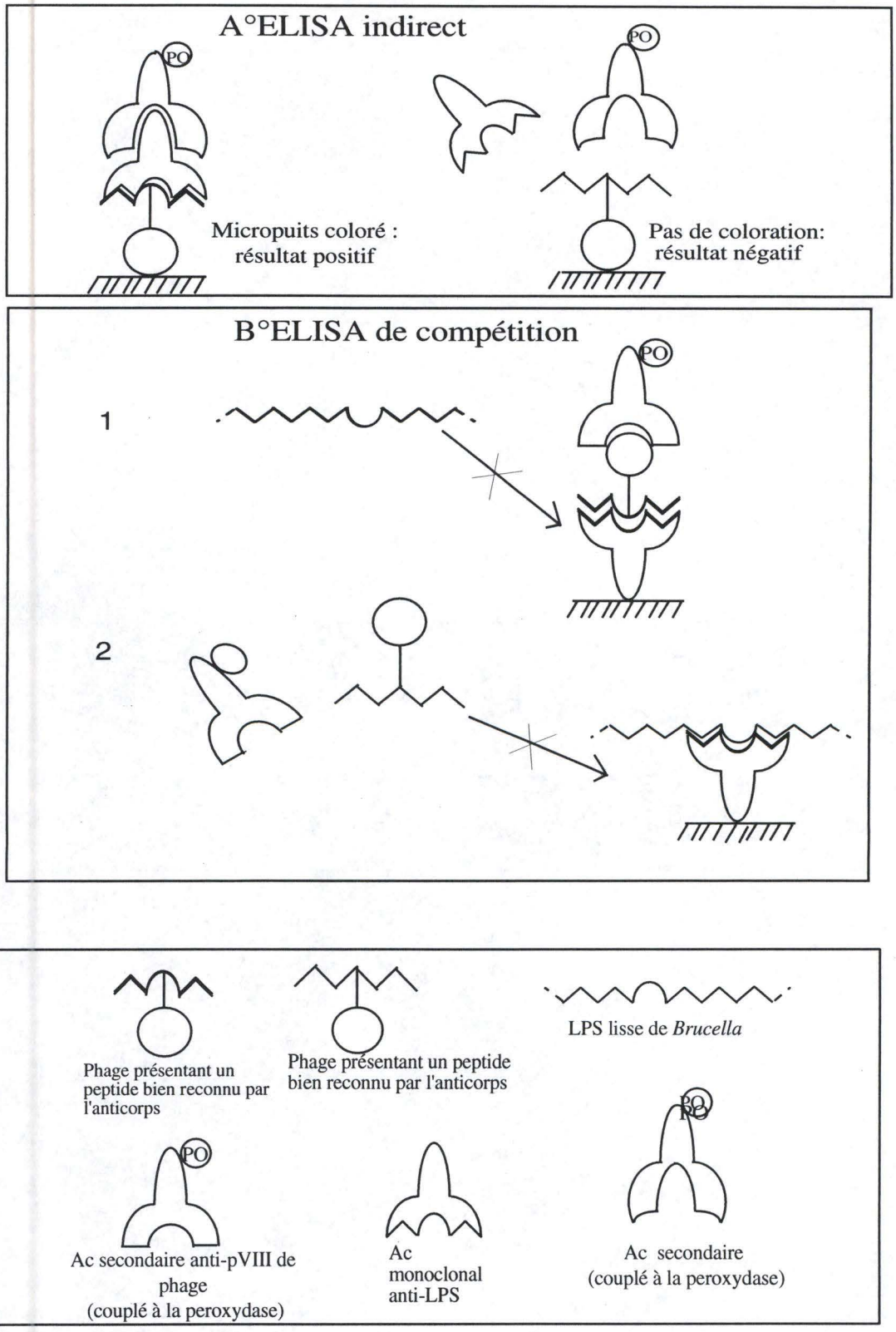


Figure 21:Représentation des tests ELISA indirects et de compétition réalisés au Laboratoire de Microbiologie et Immunologie sur les peptides mimétiques de la chaîne O de *Brucella* présentés à l'extrémité de la protéine pVIII du phage Rf.

A° ELISA indirecte: des phages présentants des peptides mimétiques sont coatés sur des plaques multipuits. Les anticorps anti-LPS vont reconnaître les peptides mimotopes. Un anticorps secondaire couplé à la péroxidase sert à la révélation.

B°ELISA de compétition: des anticorps spécifiques de la chaîne O de Brucella sont coatés sur des plaques multipuits. Les peptides suffisamment affins pourront déplacer l'anticorps de son épitope présent sur le LPS. La révélation est réalisée par des anticorps anti-pVIII.

Résultats et discussions

1) RESULTATS PRELIMINAIRES

1.1) Tests de reconnaissance (ou d'antigénicité)

Plusieurs tests de reconnaissance ont été réalisés sur les peptides mimétiques sélectionnés: les tests de type ELISA indirect et les tests ELISA de compétition (figure 21). Ces tests ont été réalisés sur l'ensemble des peptides sélectionnés, mais ne seront repris ici que les résultats pour 4 peptides sélectionnés par l'anticorps 6B3 (les peptides 6B3-[X]) et l'unique peptide sélectionné par l'anticorps 4F9 (4F9-[1]) (De Bolle and Mertens, 1997), (De Bolle, 1999). Les tests indirects consistent à adsorber des concentrations croissantes de phages présentateurs de peptides au fond de plaques multipuits, ces plaques sont mises en contact de l'anticorps monoclonal. Si la quantité d'anticorps fixés sur les plaques est proportionnelle à la quantité de phages adsorbés, le peptide est considéré comme spécifique à l'anticorps et son affinité relative peut être mesurée. Le test de compétition a pour but de valider le qualificatif « mimétique » au niveau antigénique. En effet, ces tests se basent sur la compétitivité entre le LPS et les peptides pour la reconnaissance avec un anticorps monoclonal. Si on détecte une compétition entre ceux-ci, cela signifie qu'ils se fixent au même endroit des CDR que le LPS.

a) peptides fusionnés à la pVIII

Test ELISA indirect

En ce qui concerne le peptide 4F9, on observe une augmentation de la concentration en anticorps fixés proportionnelle à l'augmentation de la concentration en phage. On peut également observer une gradation dans l'affinité des peptides 6B3-[X].

Test ELISA de compétition

L'anticorps est adsorbé sur des plaques multipuits, et est ensuite mis en présence d'un mélange de phages (avec un des mimotopes en pVIII) + LPS où le LPS est en concentration variable. La concentration en phage est fixe. On s'attend à ce que l'anticorps puisse fixer de moins en moins de phage au fur et à mesure que la concentration en LPS augmente. La fixation est ici révélée via la reconnaissance par une préparation d'anticorps polyclonaux de lapin dirigés contre la protéine pVIII du phage. Dans le cas du peptide 4F9, on observe un déplacement par l'ajout de quantité croissante de LPS. Dans le cas des peptides 6B3-[X], on n'observe un déplacement que pour le peptide 6B3-[6] (pour les autres peptides 6B3-[X], la compétition ne se distingue pas du bruit de fond) (figure 22).

Clone	Séquence du peptide	Reconnaissance en test ELISA	Compétitivité avec le LPS
6B3-1	PPSF**C**PNFIP**C**TDGL	++	+/-
6B3-2	APTQ**C**SQTWP**C**PSHL	+/-	+/-
6B3-4	PPPQ**C**SKAWP**C**PPTH	+	+/-
6B3-6	GPGQ**C**NTRNP**C**PRPM	++	+++

Figure 22: Tableau reprenant les résultats obtenus lors des tests de reconnaissance anticorps/peptide en test ELISA, ainsi que lors des tests de compétition entre peptide et LPS pour la liaison à l'anticorps. Les symboles « +/-» indiquent que la reconnaissance ne se distingue pas nettement du bruit de fond , « + » indiquent une reconnaissance détectable mais de faible intensité, « ++ » indiquent une reconnaissance moyenne, « +++ »indiquent une forte reconnaissance. L'intensité de la compétition est indiquée par les mêmes symboles.

Dans la suite du travail, nous parlerons d'affinité relative pour signaler une reconnaissance par l'anticorps, ce terme sera donc à mettre en parallèle avec les résultats du test ELISA indirect.

Le principal problème est que, par ce système de présentation, on ne tient compte que du nombre de phages reconnus. En effet, le nombre de peptides présentés en surface du phage est inconnu dans les banques de type 88 (il varie généralement de 5 à 15 %), lorsque l'on compare l'affinité des peptides reconnus par un même anticorps, on doit émettre l'hypothèse qu'en moyenne, le nombre de peptides présentés est équivalent. Il est intéressant d'utiliser un autre système de présentation des peptides, qui permettrait de mieux estimer la reconnaissance par peptide exposé. C'est entre autres pour cela, que l'on utilise des protéines-carrier. Dans ce cas, un seul peptide est fusionné par protéine, et on peut ainsi dissocier la réponse humorale dirigée contre un phage ou un peptide qu'il porte en ne mesurant que la réponse dirigée contre le peptide, car celui-ci est sorti de son contexte phagien (De Bolle and Mertens, 1997), (De Bolle, 1999).

b) peptides fusionnés à la thiorédoxine

Les différents tests indirects ou de compétition avec la thiorédoxine comme protéine présentatrice n'ont pas permis de détecter de reconnaissance avec l'anticorps pour l'ensemble des peptides. La résolution de la structure des peptides en thiorédoxine peut nous donner des informations quant aux types de conformation que les peptides pourraient adopter en solution. Des essais de cristallisation seront donc réalisés sur le stock de thiorédoxine-peptides mis à notre disposition.

1.2) Test «indirect» sur la présence du pont disulfure -"Test au DTT"

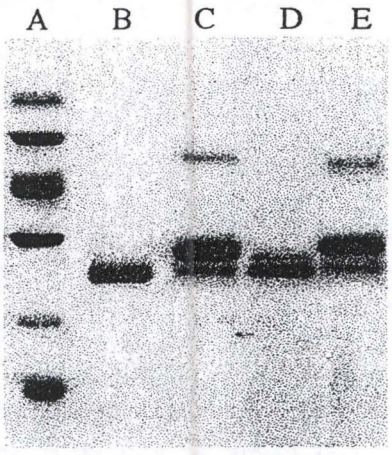
Les peptides présentés à la surface des phages peuvent être conformationnellement contraints par la présence d'un pont disulfure. En présence de DTT (dithiothréitol), les ponts disulfures sont rompus. Si la reconnaissance anticorps-peptide dépend de la présence du pont disulfure, une reconnaissance diminuée doit être observée.

C'est bien le cas du peptide 4F9-[1] et les 6B3-[X]: pour des concentrations très faibles en DTT (0,01 mM), la reconnaissance est fortement diminuée (De Bolle and Mertens, 1997), (De Bolle, 1999).

Préparation de fragments Fab à partir d'IgG de souris codées 4F9 et 2C8.

Contrôle de pureté

SDS-PAGE en présence de 2-mercaptoéthanol des fractions obtenues après digestion à la papaïne et purification sur protéine A sépharose.



A : Standards de poids moléculaires : 94 000, 67 000, 43 000, 30 000, 20 000 et 14 400.

B : 4F9 : fraction non retenue par la protéine A = fragments Fab.

C : 4F9 : fraction retenue par la protéine A = chaines lourdes et légères des IgG non digérées et fragments Fc.

D : 2C8 : fraction non retenue par la protéine A = fragments Fab.

E : 2C8 : fraction retenue par la protéine A = chaines lourdes et légères des IgG non digérées et fragments Fc.

Figure 23: SDS-PAGE

1.3) Test d'immunogénicité en souris

Les peptides 6B3-[X] fusionnés aux phages ont été injectés en souris dans le but de produire une réponse humorale dirigée contre le LPS de *Brucella* . Les peptides 6B3-[2], [4] et [6] ont permis une réponse humorale très faible tandis que l'injection du peptide 6B3-[1] n'a pas donné de réponse.

2) CONFORMATION DU RECEPTEUR: LES ANTICORPS

Les anticorps complets sont difficiles à cristalliser. Les échantillons de 4F9 ont donc été envoyés à une firme spécialisée dans l'obtention de Fab. La pureté et la concentration des échantillons qu'ils nous ont renvoyés ont été déterminées par leurs soins. Une image du gel SDS-PAGE utilisé est exposée ci-contre (figure 23).

2.1) Séquences-Modélisation

Ne disposant pas des séquences complètes des fragments Fab de l'anticorps 4F9, nous n'avons pas pu modéliser les CDR. Néanmoins, le séquençage est en cours: la modélisation reste une alternative théorique à l'obtention de la structure des CDR des fragments Fab.

2.2) Détermination de structure par DRX

2.2.1) Détermination de concentration

Le test utilisé pour la détermination de concentration utilisé est le BCA (Bicinchoninic acid) (décrit dans le chapitre "Matériel et méthode"). Le fragment Fab de l'anticorps 4F9 a été dialysé et concentré par centrifugation dans un tube muni d'une membrane de cellulose (avec un cut-off à 5000) pour avoir une solution à une concentration supérieure à 10mg/ml et pour que cet anticorps se trouve dans un tampon connu et utilisable pour la cristallisation.

En finale, nous avions une concentration évaluée à 10mg/ml pour un premier échantillon (appelé "échantillon n°1") et une concentration de 17,5 mg/ml pour un second échantillon (appelé "échantillon n°2"). La concentration de ces 2 échantillons est suffisante pour effectuer les essais de cristallisation, le premier ne sera utilisé que dans le cadre du screen "total" pour approcher les conditions de cristallisation du 4F9.

2.2.2) Cristallisation

Les conditions de cristallisation des Fab sont très différentes d'un anticorps à l'autre. Néanmoins, la présence d'un tampon-acétate semble être une constante dans une partie de ces

Fab 4F9

CONDITIONS	TEST 1 (ECHANTILLON N°1 AVEC LE "SCREEN TOTAL")	TEST 2 (ECHANTILLON N°2 AVEC LE "SCREEN TOTAL")	TEST 3 (ECHANTILLON N° 2 AVEC LES TAMPONS REFAIT A PARTIR DES CONDITIONS ORIGNALES)	TEST 4 (ECANTILLON N°2 AVEC LES TAMPONS REFAIT A PARTIR DES CONDITIONS ORIGNALES)
A 0,2 M $CaCl_2$ 0,1 M Na Hepes pH 7,5 28% v/v PEG 400	nombreux petits cristaux "en lamelle"	nombreux cristaux en lamelle **C++, S+**		quelques cristaux de petite taille en forme de lamelle (PEG 1000) **C-,S+**
B 0,2 M Acétate de Mg 0,1 M Na Cacodylate pH 6,5 20% w/v PEG 8000	2 gros cristaux polycristallins	un polycristal (0,2 mm) **S-**	aucun cristal	
C 0,2 M $MgCl_2$ 0,1 M Na Hepes pH 7,5 30% v/v PEG 400	1 polycristal	aucun cristal		quelques cristaux en forme de lamelle (PEG1000) **C++,S+**
D 0,2 M sulfate de NH_4 0,1 M Na Cacodylate Ph 6,5 30% w/v PEG 8000	peu de petits polycristaux"en lamelle"	un petit polycristal	polycristaux pour des dilution 1-1,5-2-2,5-3 et 5 fois de l'échantillon de 4F9	2 monocristaux de 0,2 mm **C-,S+**

Figure 24: Tableau reprenant les différents essais réalisés sur les fragments Fab de l'anticorps 4F9, ainsi que les faciès des cristaux observés après 1 semaine. Les cristaux de plus de 0,1 mm ont été caractérisés par leur dureté (S++: solide, S+: solidité moyenne, S-: fragile) et leur coloration par du bleu de Coomassie (C++: très coloré, C+: moyennement coloré, C-: faiblement coloré).

conditions (Gilliland and Bickham, 1990). Un screen "acétate" a donc été utilisé pour la cristallisation de fragments Fab de l'anticorps 4F9. Pour chacune des gouttes, le 4F9 précipite et on n'observe aucun cristal (la liste des composants des "screen" se trouve dans le chapitre "matériel et méthode").

D'autres tests ont été effectués à partir d'un screen "total" reprenant une série de conditions de cristallisation (au nombre de 50) très différentes les unes des autres (voir matériel et méthode: "Hampton research: Crystal screen I"). Après avoir réalisé ces essais, plusieurs observations ont été faites: après quelques heures, après 1 jour et après 1 semaine. Le tableau ci-contre reprend une description des cristaux après 1 semaine et les conditions dans lesquelles ils ont été obtenus (figure 24). Il est important de savoir si ces cristaux sont des cristaux de protéines ou des cristaux de sels, peut-être dus à une co-précipitation entre les constituants du tampon de la solution d'anticorps et le tampon de la liqueur-mère. Cette situation est malheureusement très fréquente. Nous avons utilisé 2 types de tests: le test basé sur la solidité (normalement plus importante des cristaux de sels) et le test de coloration (un colorant, ici le bleu de Coomassie, diffuse librement à l'intérieur du cristal de protéine et le colore, ce qui n'est pas le cas d'un cristal de sel). Ces tests ont été utilisés pour les cristaux de taille supérieure à 0,1 mm ("utilisable" pour l'ensemencement), ils ont également été utilisés pour les cristaux de thiorédoxine-peptides de taille comparable.

Après les premiers tests (TEST 1), des cristaux sont apparus dans 2 types de conditions très proches (conditions A-C et B-D). Malheureusement, les cristaux obtenus n'ont pas la taille requise pour pouvoir diffracter (ils avaient une taille maximale de 0,05 mm). Nous avons donc refait des tests (TEST 2) avec les tampons du "screen-total", mais cette fois en utilisant l'échantillon n°2 des Fab 4F9 qui est plus concentré. Des cristaux de même aspect morphologique mais en moins grand nombre sont apparus dans 3 des 4 conditions.

Dans le but d'obtenir des monocristaux de plus grande taille, nous avons refait des tests pour les conditions B et D (TEST 3) à l'aide de nouveaux tampons de même composition que le screen "total" mais avec des concentrations de l'anticorps dilué 1, 1.5, 2, 2.5, 3 et 5 fois. Dans ce cas, nous n'obtenions des cristaux que dans 1 des 2 conditions (condition "D") mais pour chacune des dilutions.[8].

De nouveaux essais ont été réalisés (TEST 4) dans les conditions D et dans les conditions A et C (en remplaçant le PEG 400 par du PEG 1000 dans ces dernières). Dans la liqueur-mère D,

[8] *Nous n'avons pas refait de tests pour les conditions A et C car ne disposions pas de PEG 400 (Polyéthylèneglycol 400).*

nous avons obtenu des polycristaux de grande taille, desquels il a été possible d'extraire deux monocristaux de 0,2 mm.

2.2.3) Diffraction et analyse des cristaux de 4F9

L'étape suivante est l'étude des cristaux par diffraction des rayons X. Pour ce faire, nous avons isolé un monocristal de taille suffisante (0,2 mm) qui semblait être positif (TEST 4, condition **D**) lors de la coloration par le bleu de Coomassie. La technique d'isolement du cristal avant diffraction est décrite dans le chapitre "matériel et méthode". L'image du spectre de diffraction du cristal nous a permis de constater que ce cristal n'était pas un cristal de protéine, mais bien un monocristal de sel. En effet, la maille d'un cristal de sel est beaucoup plus petite que celle d'un cristal de protéine, dès lors après un "scan" du spectre de ce cristal, les "taches" de diffraction étaient très éloignées.

2.2.4) Perspectives

Même si ce résultat nous a montré les limites du test de coloration (selon le test de coloration, ce cristal est un cristal de protéine), il faut savoir que ce cristal n'était pas très coloré et qu'il était assez solide en comparaison avec les autres polycristaux obtenus, par ailleurs très colorés.

Etant donné que ces monocristaux étaient isolés à partir d'un polycristal à l'aspect morphologique semblable aux autres polycristaux que nous avons eu dans les mêmes conditions (TEST1, 2 et 3), nous en concluons que tous les polycristaux élaborés dans ces conditions sont des cristaux de sel.

Ces essais préliminaires nous ont tout de même permis de trouver quelques conditions de cristallisations des Fab 4F9 (les conditions A, B et C), ces conditions devraient être affinées de différentes façons: par cristallisation et conservation à 4°C, par variation de la concentration de la protéine et de la liqueur-mère dans la goutte, par échantillonnage à des concentrations légèrement différentes en précipitant dans le puits, ainsi que par variation des autres paramètres (voir "matériel et méthode").

3) CONFORMATION DU LIGAND: LES PEPTIDES

a) Peptides

Les peptides mimétiques isolés au sein du laboratoire portent le nom de l'anticorps suivis d'un numéro caractéristique (ex: 4F9-[1]) (Walgraffe, 1998), (De Bolle and Mertens, 1997), (De

NOM	SEQUENCE
6B3-1	PPSFCPNFIPCTDGL
6B3-2	APTQCSQTWPCPSHL
6B3-4	PPPQCSKAWPCPPTH
6B3-6	GPGQCNTRNPCPRPM
4F9-[1]	RDPQCTEPELTCAREL

Figure 25: Tableau reprenant la séquence des 5 peptides qui seront étudiés lors de ce travail.

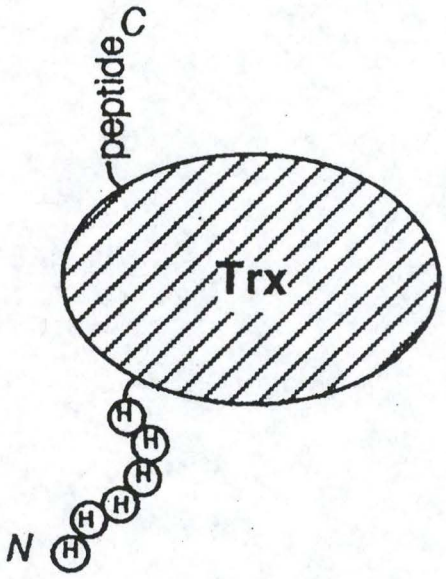


Figure 26: Représentation schématique de la fusion « thiorédoxine-peptide ». La structure de la thiorédoxine étant résolue, il a été vérifié que la queue poly-histidine (à l'extrémité N-terminale) est suffisamment éloignée du point d'insertion du peptide d'intérêt (à l'extrémité C-terminale) que pour postuler qu'il sera possible d'adsorber la protéine de fusion sur une surface tout en présentant le peptide sans encombrement stérique majeur (De Bolle, 1999).

Bolle, 1999). Ils proviennent tous des banques 88 ou des banques 8+8 (voir encadré ci-dessous)

Banque de type 88: $(X)_{4-5}C(X)_{3-6}C(X)_4$

Banque de type 8+8: $C-(X)_9-C$ **ou** $(X)_9$

X = n'importe quel acide aminé, C = cystéine

Dans ces banques, les peptides sont exprimés à l'extrémité N-terminale des protéines pVIII. Un pont disulfure entre les 2 cystéines (pour les banques où il y en a) est vraisemblablement présent (voir "test au DTT"). Nous nous sommes intéressés à 4 peptides sélectionnés par l'anticorps monoclonal A15-6B3[9] car ces peptides partagent un consensus au niveau de la séquence. Nous nous intéresserons également au peptide sélectionné par l'anticorps monoclonal B66-4F9[10] car nous avons à notre disposition une quantité suffisante d'anticorps 4F9 que pour pouvoir envisager sa cristallisation[11]. Ces 5 peptides appartiennent tous à une banque de type 88, les peptides 6B3-[X] sont tous de type $(X)_4\underline{C}(X)_5\underline{C}(X)_4$. et le peptide 4F9-[1] est de type $(X)_4\underline{C}(X)_6\underline{C}(X)_4$ (figure 25).

b) Protéines-chimères («carrier»):

Nous disposons de 4 échantillons de protéines-chimères réalisées par constructions génétiques: 4 peptides mimétiques (6B3-[X]) ont été fusionnés à l'extrémité C-terminale de molécules de thiorédoxine de *E.Coli* (figure 26). Le principal intérêt de ce mode de présentation de peptides réside dans le fait que le nombre de peptides présentés est connu précisément. Ce mode de présentation est également intéressant par le fait que les conditions de cristallisations de la protéine "chimère" sont hypothétiquement proches des conditions de cristallisation de la thiorédoxine seule (De Bolle, 1999). Les protéines chimériques se trouvent dans un tampon TBS, leur concentration est de 15 mg/ml.

---

[9] *appelé ici « 6B3 », un anticorps monoclonal spécifique de l'épitope M de la chaîne O du LPS de Brucella sp*

[10] *dit « 4F9 », un anticorps monoclonal spécifique de l' épitope C/Y , c'est-à-dire porté à la fois sur la chaîne O du LPS de Brucella sp. et celle de Yersinia Enterocolitica O:)*

[11] *les données de structure de CDR suite à la cristallisation des Fab porront dès lors être mise en commun avec les conformations modélisées des peptides*

HT6.1

MGSSHHHHHHSSGLVPRGSHMSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIA
PILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKE
FLDANLASSFAA**PPSFCPNFIPCTDGL**PAA

HT6.2

MGSSHHHHHHSSGLVPRGSHMSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIA
PILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKE
FLDANLASSFAA**APTQCSQTWPCPSHL**PAA

HT6.4

MGSSHHHHHHSSGLVPRGSHMSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIA
PILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKE
FLDANLASSFAA**PPPQCSKAWPCPPTH**PAA

HT6.6

MGSSHHHHHHSSGLVPRGSHMSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIA
PILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKE
FLDANLASSFAA**GPGQCNTRNPCPRPM**PAA

Figure 27 :Séquences des HT6.X (H pour Histidine Tail, T pour Thioredoxin et 6.1, 6.2, 6.4 et 6.6 pour le nom des peptides 6B3-[X]).
Les peptides sont représentés en gras.

3.1) Détermination de structure par DRX

3.1.1) Cristallisation

a) Test avec le lysozyme

Nous avons réalisé nos premiers essais sur du lysozyme pour plusieurs raisons: la première est que le lysozyme cristallise facilement, c'est donc une très bonne manière de s'entraîner. La seconde raison est que cette protéine est disponible en grande quantité, nous avons donc pu réaliser de nombreux essais. Enfin, comme les peptides vont être fusionnés à l'extrémité du lysozyme (en cours au laboratoire), il est intéressant de pouvoir connaître les conditions de cristallisation du lysozyme afin de les appliquer à la cristallisation du lysozyme fusionné aux peptides. Si nous obtenons des cristaux, ils devraient être conservés afin d'utiliser la technique du "cross-seeding" dès que nous les protéines-chimères en solution seront mises à notre disposition.

Nous disposions de 2 échantillons, l'un concentré à 50 mg/ml, l'autre à 25 mg/ml, dans un tampon acétate 0,1 M pH 4,6. Les conditions de cristallisations sont détaillées dans le chapitre "matériel et méthode", ce sont les conditions classiques d'obtention de la forme tétragonale du lysozyme.

Après quelques jours, nous avons obtenu des cristaux dans les gouttes à faible concentration en lysozyme et après quelques semaines, nous pouvions observer des cristaux dans chaque goutte. Ces monocristaux sont caractéristiques du lysozyme: ils sont assez volumineux (0,4 mm), ils apparaissent facilement dans une série de conditions assez différentes et ont un faciès propre aux cristaux de cette protéine. Ces cristaux n'ont pas été étudiés plus en détail car la structure du lysozyme est connue depuis longtemps, en revanche, ils nous serviront de contrôles positifs pour les tests de coloration au bleu de Coomassie. Ils seront conservés à température ambiante en vue de tentatives de "cross-seeding".

b) Tests de cristallisation des protéines chimériques:

Les premiers essais que nous avons réalisés sur la thiorédoxine fusionnée aux peptides 6B3-4 et 6 ont été guidés par les protocoles utilisés pour la cristallisation de la thiorédoxine seule. Les différences principales entre la thiorédoxine et la protéine-carrier d'intérêt sont: la présence d'une queue poly-histidine en N-terminal (6 résidus qui servent à la purification de la protéine) et le peptide fusionné en C-terminal (figure 27). Les conditions de cristallisation de la thiorédoxine sont les suivantes: 50 % v/v 2-4 méthyl-pentanediol (MPD) HAc/Ac 0,01M

Thiorédoxine-peptides

Conditions	TEST 1 "SCREEN TOTAL"	TEST 2 "AFFINAGE"
A 0,02M $CaCl_2$ 0,1 M acétate de Na pH 4,6 30 % w/v MPD	Petits monocristaux en forme de fines plaquettes	nombreux polycristaux ,présents dans toutes les gouttes **C++**
B 0,4 M KNa tartrate	1 seul polycristal (il a disparu après 1 semaine)	
C 0,2 M $CaCl_2$ 0,1 M Na Hepes pH 7,5 28% v/v PEG 400	nombreux polycristaux arborescents	
D 0,2 M $CaCl_2$ 0,1 M acétate de Na pH 4,6 20 % v/v propanol	petits polycristaux allongés	peu de polycristaux ,présents dans toutes les gouttes **C+**
E 0,2 M acétate de Ca 0,1 M Cacodylate de Na pH 6,5 18% w/v PEG 8000	nombreux polycristaux allongés (desquels nous avons pu isoler un monocristal) **C++,S+**	

Figure 28: Tableau reprenant les différents essais réalisés sur la thiorédoxine fusionnées aux pepitides 6B3-[X], ainsi que les faciès des cristaux observés après 1 semaine. Les cristaux de plus de 0,1 mm ont été caractérisé par leur dureté (S++: solide, S+: solidité moyenne, S-: fragile) et leur coloration par du bleu de Coomassie (C++: très coloré, C+: moyennement coloré, C-: faiblement coloré). Pour toutes ces conditions, les résultats sont identiques d'un peptide à l'autre.

pH 4,0 Cu^{++} 0,001 M (Holmgren and Sodeberg, 1970). Les différents essais que nous avons réalisés varient par la concentration en précipitant (voir "matériel et méthode"). Aucun cristal n'a été observé dans ces conditions. La présence de la queue poly-histidine du peptide, ainsi qu'une série de paramètres moins contrôlables (comme l'humidité et les variations de température) dont il faut tenir compte, modifieraient les conditions de cristallogenèse.

Une première solution serait de tenir compte de l'influence des 6 histidines en début de chaîne pour la cristallisation. Dans cette optique, Trakhanakov préconisait l'utilisation de cadmium à concentration élevée pour la cristallisation de protéines riches en histidine: le cadmium pourrait former un complexe avec l'histidine (pour éviter la précipitation précoce de ce type de protéine) (Trakhanakov *et al.*, 1998).

Nous avons donc réalisé une dizaine d'essais avec le Cadmium en concentration saturante (dans les conditions décrites dans "matériel et méthode"). Au cours des derniers essais, nous avons pu observer que la protéine précipitait systématiquement après le dépôt de la goutte de liqueur-mère. En général, la précipitation précoce de protéine aboutit rarement à la formation de cristaux. De fait, l'observation répétée à quelques jours d'intervalle de cette boîte ne montra aucun signe de cristallisation. Cette précipitation rapide serait due à la concentration excessive en cadmium, cette saturation causerait une augmentation importante de la force ionique de la goutte et une précipitation systématique de la protéine (salting out).

La même stratégie que pour les fragments Fab du 4F9 a ensuite été utilisée: un screen "total" a été effectué sur les solutions de thiorédoxine-peptides. Pour 5 conditions, nous avons pu observer des cristaux de morphologies très différentes (TEST 1) (figure 28).

En observant ces résultats, on remarque que la présence de chlorure de calcium aiderait à la cristallisation. Hormis la présence du chlorure de calcium dans la plupart des conditions, les autres composants de ces conditions sont très différents les uns des autres (pH, tampons, sels et précipitants). Nous avons choisi de nous attarder sur deux types de conditions très proches (A et E) dont le pH se rapprochait le plus du pH utilisé pour cristalliser la thiorédoxine seule et pour lesquels nous avions de nombreux cristaux. Dans une de ces conditions il y a du MPD (2-4 méthyl-pentanediol), précipitant également utilisé pour la cristallisation de la thiorédoxine seule.

Des essais ont été réalisés avec de nouvelles liqueurs-mères de même composition que celles du screen "total" dans le but de produire des cristaux en quantité (TEST 2). Nous espérons par là obtenir l'un ou l'autre cristal assez grand que pour pouvoir diffracter. En plus des 16 tests effectués sur les protéines-chimères, nous avons réalisé un test négatif pour chacune des conditions: une goutte où nous incorporons la liqueur-mère ainsi que le tampon d'origine de la

protéine (la protéine n'est pas présente). Cette stratégie est une alternative à l'identification de cristaux de sels: en effet, si nous obtenons des cristaux de même faciès avec ou sans protéine dans la même liqueur-mère, nous pouvons être sûr que ce sont bien des cristaux de sel et non de protéine. Nous n'avons pas obtenu de cristaux pour les 2 tests négatifs. Tous les cristaux obtenus dans les autres puits se colorent au bleu de Commassie. A ce stade, seule une image de diffraction pourrait nous permettre d'être sûr que ces cristaux sont effectivement des cristaux de protéine.

3.1.2) Diffraction

Tout comme pour le 4F9, nous avons isolé le plus grand monocristal disponible (isolé à partir du premier screen "total" dans la condition E). Aucune image de diffraction de ce cristal n'a été obtenue, le cristal est sans doute trop petit, il ne diffracte donc pas.

3.1.3) Perspectives

En ce qui concerne les cristaux obtenus dans les conditions proches de celles de la thiorédoxine (conditions A et E), de nouveaux essais (autour des conditions initiales) ainsi que d'autres contrôles négatifs devraient être réalisés afin d'obtenir des monocristaux fiables et assez gros. Si ces cristaux se révèlent inutilisables, il y a encore 2 des 5 conditions du screen à retester et à affiner (les conditions C et D). Des essais dans les mêmes conditions que pour la thiorédoxine seule, mais avec une concentration moindre en MPD (30 %) devraient être également réalisés.

3.2) Analyse de séquence

3.2.1) Comparaison globale

La composition des peptides 6B3-[X] a été analysée au Laboratoire de Microbiologie et d'Immunologie (Walgraffe, 1998), (De Bolle and Mertens, 1997), (De Bolle, 1999). Les peptides reconnus par un même anticorps semblent proches par leur composition globale: par exemple les peptides reconnus par l'anticorps 6B3 ont de nombreux résidus prolines (entre 3 et 6) et une série d'acides aminés aromatiques en commun. Des comparaisons de séquences ont également été réalisées suivant une série de règles prédéfinies. Un consensus pour une série de peptides a été mis en évidence. Les peptides 6B3-[X] semblent participer à un consensus (Walgraffe, 1998). Ils sont donc regroupés et étudiés plus en détail.

N°	SEQUENCE														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1:	P-	**P**-	S-	F-	**C**-	P-	N-	F-	I-	**P**-	**C**-	T-	D-	G-	L
2:	A-	**P**-	T-	Q-	**C**-	S-	Q-	T-	W-	**P**-	**C**-	P-	S-	H-	L
4:	P-	**P**-	P-	Q-	**C**-	S-	K-	A-	W-	**P**-	**C**-	P-	P-	T-	H
6:	G-	**P**-	G-	Q-	**C**-	N-	T-	R-	N-	**P**-	**C**-	P-	R-	P-	M
ID :	1	3	0	2	3	1	0	0	1	3	3	2	0	0	1
SIM:	2	3	3	2	3	3	2	1	2	3	3	3	1	3	2

Figure 29 :Alignement entre 4 peptides reconnus par 6B3.

L'abréviation ID. signifie score d'identité tandis que SIM. signifie score de similarité. Ces scores sont compris entre 0 (aucune identité/similarité) et 3 (identité/similarité pour les 4 séquences).

3.2.2) Comparaison de séquence des peptides 6B3-[X]

a) Observations et commentaires

Une comparaison de séquence a été réalisée en prenant en compte l'identité et la similarité entre résidus des 4 séquences des peptides 6B3-[X]. La comparaison de séquences par identité prend simplement en compte le nom des acides aminés, tandis que la comparaison de séquence par similarité est basée sur un caractère physico-chimique pour chaque résidu. Ces comparaisons ainsi qu'un score pour chaque position sont représentés à la figure 29. Par exemple, pour la position 4, pour 3 des séquences, une glutamine est présente. Le score d'identité est donc de 2. La phénylalanine étant hydrophobe, elle ne se situe pas dans la même catégorie que la glutamine: le score de similarité est de 2 également.

Deux prolines (résidus 2 et 10) sont alignées dans les 4 peptides, une autre proline et une glutamine (résidus 4 et 12) sont alignées pour les peptides 2-4-6 et une sérine ainsi qu'un tryptophane (résidus 6 et 9) sont alignés pour les peptides 2 et 4.

Cette comparaison permet donc de rapprocher ces peptides sur base de la séquence: le 6B3-[4] semble plus proche du 6B3-[2], un peu moins du 6B3-[6], le 6B3-[1] est plus éloigné. Au niveau de la similarité, on remarque qu'il n'y a que 2 positions où le score est de 1 (résidus 8 et 13), et qu'il y a plus de positions avec un score de 3 que de score 2. Cela montre une très grande similarité entre les séquences de chaque peptide.

b) Hypothèses et outils pour vérifier ces hypothèses

Si des résidus sont alignés, c'est sans doute parce qu'ils sont importants, soit pour le maintien de la conformation du peptide, soit pour l'interaction avec l'anticorps. En général, on considère que la présence de résidus prolines et de ponts disulfures (cystéines) dans un peptide le contraint fortement. Nous vérifierons ces hypothèses par l'analyse de la flexibilité locale de la chaîne carbonée (angles PHI-PSI) ainsi que par modélisation des peptides avec ou sans pont disulfure. Le rôle des autres résidus alignés n'a pas été analysé. Ils pourraient être importants pour l'interaction avec l'anticorps ou pour la formation de liaisons intramoléculaires. Cela est difficile à vérifier étant donné que l'on ne modélisera que le peptide seul et non le complexe peptide-anticorps.

3.2.3) Recherche du motif dans les banques de données

L'objectif de cette démarche est de trouver des parties de séquences apparentées aux peptides à l'aide d'outils de recherche dans les banques de données (où sont stockées les séquences de protéines de structure connue. Ces structures "sœurs" seraient une image (plus ou moins

contrainte selon leur situation dans la protéine) du type de conformation que pourrait adopter le peptide. Elles pourraient servir pour une modélisation par homologie des peptides.

a) Résultats

L'outil de recherche est BLAST 2.0.8 (Altshul, 1990) (brièvement décrit dans le chapitre "matériel et méthode"). Nous avons fait des recherches pour les peptides 6B3-[X] et le 4F9-[1]. Dans aucun cas, nous n'obtenions de séquences suffisamment proches: la valeur attendue était supérieure à 25[12]. Ces résultats ne sont donc pas très significatifs: ces peptides de 15 résidus ont maximum 6 résidus en commun avec les motifs trouvés. Nous avons tout de même visionné 3 motifs pour lesquels les 2 cystéines se superposaient. En effet, si le pont disulfure ainsi que 3 ou 4 autres résidus des peptides sont présents dans certains motifs, la conformation de ces motifs pourrait être utilisable pour y adapter la conformation des peptides. Malheureusement, il n'y a aucun pont disulfure formé au niveau des 3 structures sélectionnées.

b) Conclusions

Aucun des motifs trouvés n'est utilisable comme point de départ d'une modélisation par homologie. Ce résultat était prévisible, en effet ces peptides sont synthétiques et ne font partie d'aucune phylogénie "naturelle ". Les chances de trouver un homologue de ces peptides dans les banques de données étaient donc faibles.

3.3) Analyse structurale: modélisation des peptides

3.3.1) Mise au point d'un protocole de dynamique et de mécanique moléculaire pour des peptides contraints par un pont disulfure

Dans le cas présent, nous n'avions pas de protocole disponible pour la modélisation de peptides contraints d'une quinzaine d'acides aminés. Avant de passer à l'élaboration de ces protocoles, il faut savoir que même s'ils sont contraints, ces peptides restent très flexibles: ils ont une taille intermédiaire entre de petites molécules organiques, pour lesquelles il y a peu d'atomes et donc moins de conformations possibles et de plus longues protéines, qui adoptent des conformations plus stables par des liens intramoléculaires (structures secondaires: hélice α,plans β, turns). Le protocole devra donc être adapté à cette flexibilité.

---

[12] *la valeur attendue est la fréquence d'observation d'un score donné par hasard dans les bases de données, une valeur attendue de 0 signifie une identité parfaite entre les 2 séquences*

Figure 30: Séquence de l'agent actif de l'ANF.

Dans un premier temps, nous avons utilisé un protocole de type "annealing" pour modéliser les peptides 6B3-[X]T (c'est-à-dire le peptide plus les 4 résidus qui le bordent dans la thiorédoxine). Ce protocole s'est révélé inadapté[13], même en faisant varier la température (300-600-900 K). La dérivée maximale était chaque fois trop élevée, ce qui montre que la structure était incohérente à chaque fois, soit du fait de critères de convergence peu stringents lors de la minimisation préliminaire ou du fait de la température trop élevée (on peut plus facilement donner de l'énergie à une structure bien minimisée tandis que si on utilise une structure pas très bien minimisée, on obtient des conformations très instables). Dans le cas où la température serait trop basse, après examen des conformations à l'écran et du graphe des énergies, on peut constater que l'ensemble de l'espace conformationnel n'est pas exploré.

Les différentes étapes de la dynamique et de la mécanique ainsi que leur enchaînement ne paraissaient donc pas adaptés aux peptides que nous cherchons à modéliser (on ne quitte pas le minimum local de la conformation de départ).

Dans le même dessein, nous avons adapté un protocole utilisé pour la modélisation d'un peptide de séquence analogue aux peptides 6B3-[X], l'ANF (Atrial Natriurietic Factor), une hormone dont l'agent actif comprend 28 acides aminés et qui possède un pont disulfure entre les résidus 7 et 23 (figure 30). Cette modélisation est décrite dans le manuel de l'utilisateur de MSI (Molecular Simulated Incorporating) (MSI, 1996).

Les résultats ne se sont pas révélé plus fructueux que pour les premiers essais, certainement parce que l'ANF est presque 2 fois plus grand que nos peptides étudiés et qu'il y a beaucoup plus de résidus entre les cystéines: cette partie serait plus libre et formerait plus facilement des liaisons intramoléculaires.

Finalement, nous avons testé une série de protocoles-type [14]que MSI (Molecular Simulated Incorporating) met à la disposition des utilisateurs: le protocole "quenched" [15] est le plus intéressant pour la modélisation des peptides. Ce protocole a été modifié par nos soins, afin de le rendre optimal pour la modélisation des peptides 6B3-[X]: critères de convergence, types de minimisation et température de chaque étape ont été remaniés afin d'obtenir des conformations cohérentes (ces différents paramètres sont explicités dans le chapitre "matériel

---

[13] *Il faut savoir que les critères de succès d'une dynamique moléculaire sont l'exploration de l'espace conformationnel dans son entièreté et l'obtention de conformations valables dans les conditions demandées (dans ce cas-ci, en complexe dans le site de reconnaissance d'un Fab*

[14] *molecular dynamics, quenched dynamics, annealings dynamics et slow-quenched dynamics.*

[15] *il existe différents types de protocoles, caractérisés par un temps d'équilibration, de dynamique et de refroidissement , le protocole "quenched" est caractérisé par un refroidissement assez long*

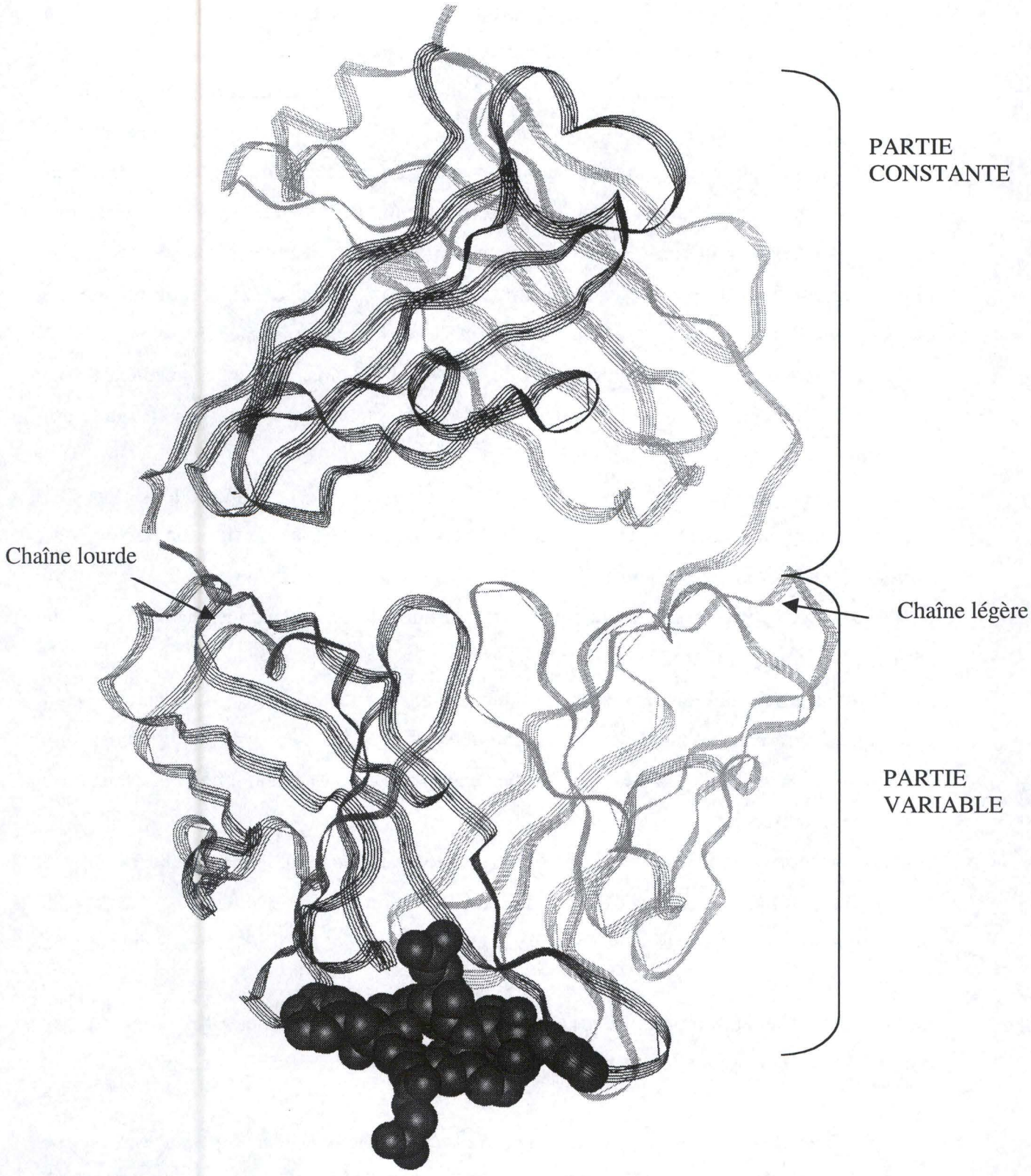
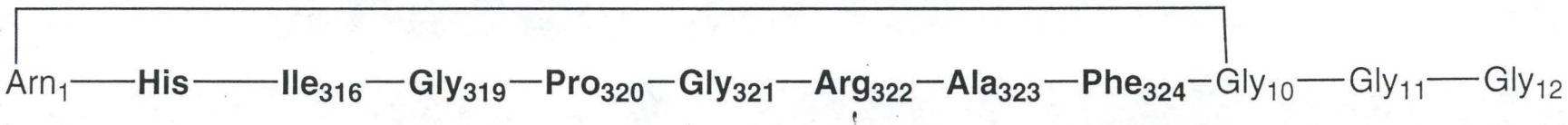


Figure 31: Représentation du complexe anticorps-épitope d'une protéine d'enveloppe de HIV.

Arn = 1-imino-5-pentanone

A)



B)

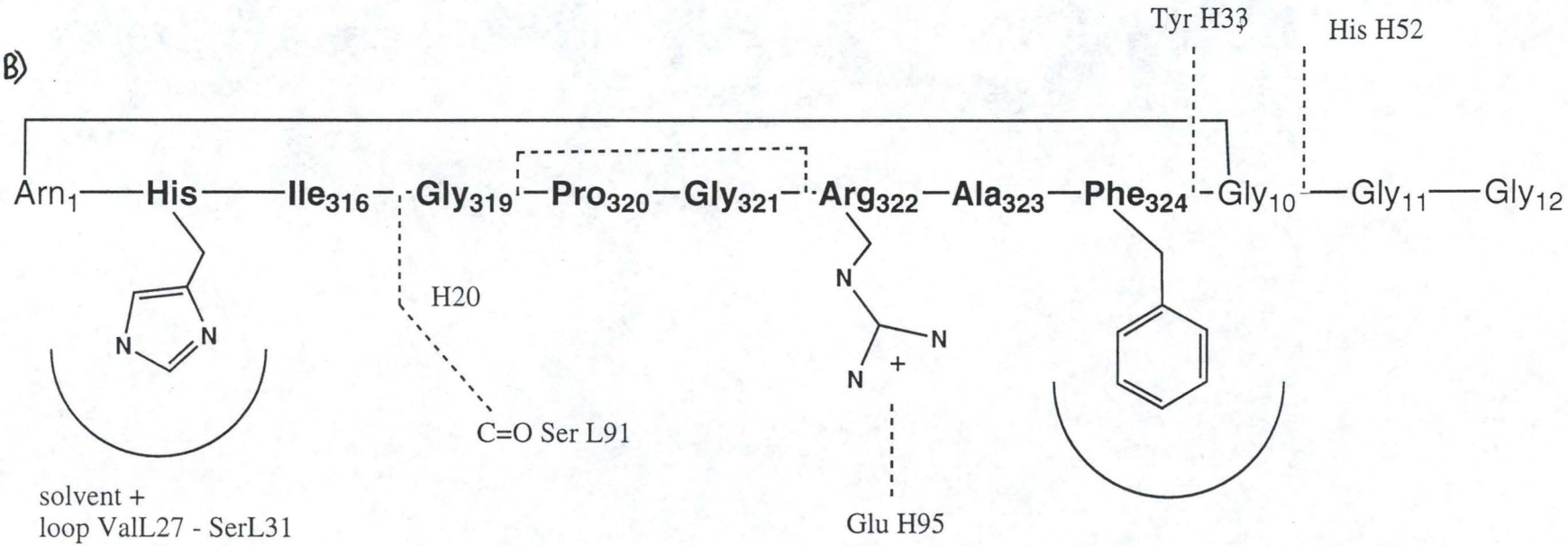


Figure 32 : a) Séquence du peptidoïde utilisé comme cas-test
b) Schéma reprenant les résidus du peptidoïde et de l'anticorps qui interagissent lors de la formation du complexe.

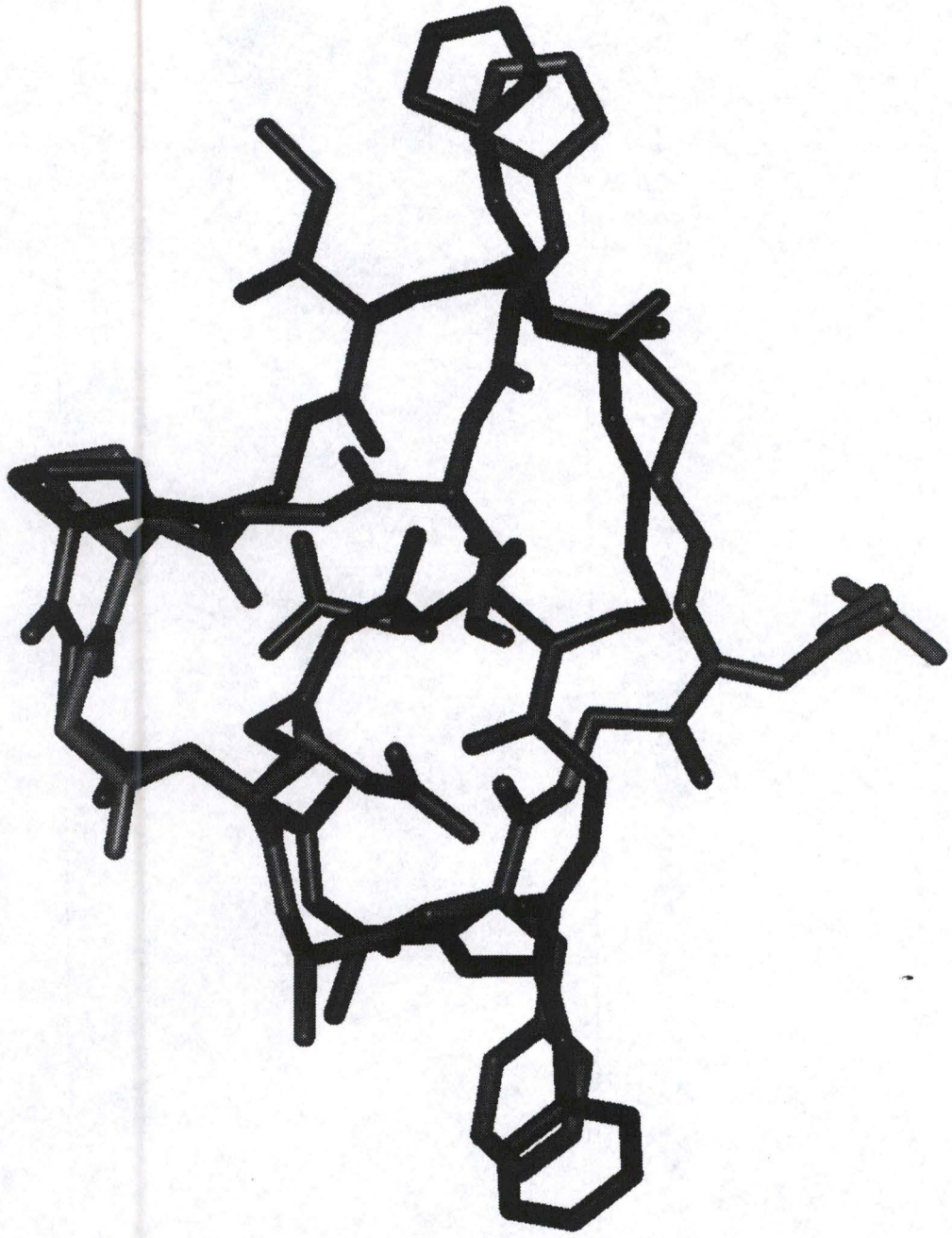


Figure 33: Représentation de la superposition entre la forme du peptide "cas-test" complexée à l'anticorps et un des conformères générés par dynamique moléculaire.

et méthode"). Nous avons également choisi de ne travailler que sur 20 conformères[16] finaux pour un même peptide: ce nombre est suffisant pour balayer l'espace conformationnel et l'analyse de plus de 20 conformères demanderait trop de temps. L'angle OMEGA a été bloqué à 180° pour éviter des conformations absurdes avec des angles OMEGA à 0°, que l'on risque d'obtenir lorsque l'on donne trop d'énergie au système. Le protocole que nous avons utilisé est décrit dans le chapitre «matériel et méthode».

3.3.2) Utilisation d'un cas-test

Afin de tester et valider le protocole que nous avons mis au point, il nous faut encore montrer que celui-ci nous permettra d'approcher la conformation du peptide en complexe. Pour ce faire, nous avons utilisé un peptidoïde synthétique de taille analogue aux peptides 6B3-[X] (12 résidus), contraint par une chaîne carbonée légèrement plus longue qu'un pont disulfure et dont on connaît la structure en complexe avec le Fab (figure 31). Ce peptidoïde est l'épitope d'une protéine d'enveloppe d'HIV auquel on a ajouté une série de glycines et un pont ARN afin de bloquer sa conformation (figure 32a). Ce peptidoïde a été modélisé en 20 conformères. En parallèle, nous avons repéré dans la structure complexe Fab-peptide, les résidus du peptide et du Fab qui sont en interaction. Ces résidus sont indiqués à la figure 32b. Les conformères ont été superposés un par un à la conformation en complexe. Deux des conformères se sont révélés être très proches du peptide complexé au niveau du RMS (*RootMeanSquare*[17]) global sur la chaîne carbonée. Le fait le plus remarquable est que l'ensemble des résidus importants pour la reconnaissance se sont révélés être superposables (figure 33).

Cette dernière donnée nous a permis de prouver que le protocole est utilisable pour la modélisation de petits peptides contraints en complexe: il est possible de trouver une conformation proche de la conformation en complexe d'un peptide en utilisant le protocole de dynamique et de mécanique moléculaire que nous avons établi. Nous pouvons donc modéliser tous les peptides d'intérêt (6B3-[X] et 4F9-[1]) sur base de ce protocole.

3.3.3) Modélisation des peptides 6B3-[X] et 4F9-[1]

a) Tests réalisés et résultats

Une série de modélisation par dynamique moléculaire ont été réalisées en vue de générer 20 conformères pour les peptides 4F9-[1] et 6B3-[X]. Ces peptides ont été modélisés soit isolés,

[16] *un conformère est une conformation potentiellement stable du peptide modélisé*

[17] *La mesureRMS correspond à la distance euclidienne moyenne minimale entre les atomes des deux squelettes protéiques impliqués dans la comparaison*

Figure 34 :

Séquence des peptides «en pVIII»: ALA-*peptide*--PRO-ALA-GLU

Séquence des peptides «en thiorédoxine»: ALA-ALA-*peptide*--PRO-ALA-ALA

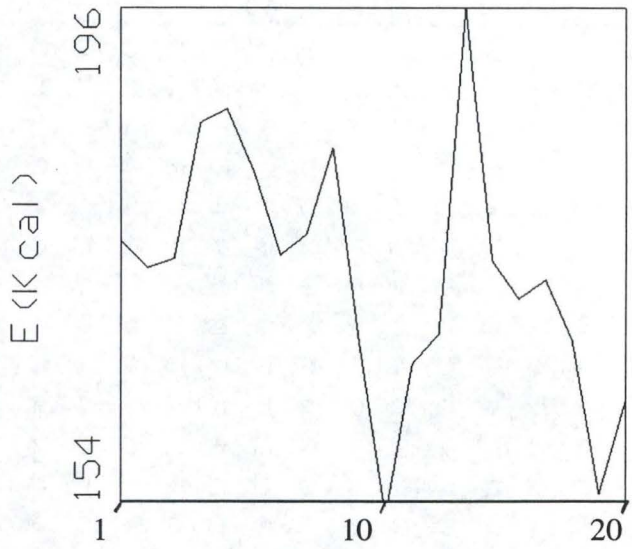


Figure 35: Graphe des énergies: les 20 conformères sont caractérisés par la somme des énergies de chaque atome (exprimée en kcal/mol).

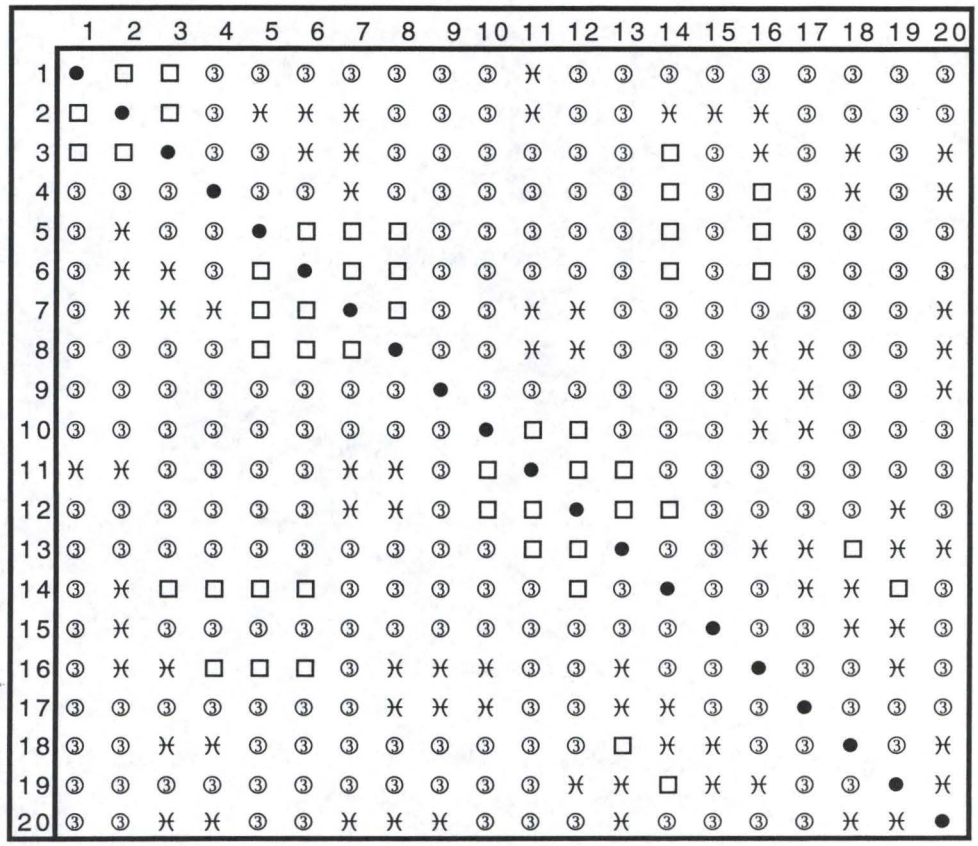


Figure 36: Exemple de "cluster graph" appliqué sur le peptide 6B3-6 "en thiorédoxine".
Les 20 conformères générés par dynamique moléculaire sont comparés un à un par un RMS sur l'ensemble des atomes de la chaîne carbonée.
les symboles ● représentent un RMS inférieur à 3.00 Å
les symboles □ représentent un RMS entre 3.00 et 4.60 Å
les symboles ③ représentent un RMS entre 4.60 et 6.40 Å
les symboles ⋓ représentent un RMS entre 6.40 et 9.25 Å
Par-exemple, un RMS global inférieur à 4.60 Å est observé entre le conformère 16 et les conformères 4, 18 et 13.

soit additionnés d'un certain nombre de résidus qui les bordent dans les protéines pVIII (6B3-[X]P) ou dans les thiorédoxines (6B3-[X]T)[18](figure 34). De cette façon, nous serons à même de comparer les différences de conformations qu'il pourrait y avoir entre les peptides 6B3-[X]P et les peptides 6B3-[X]T. En outre, un essai de modélisation a été réalisé sur le peptide 6B3-[6] sans pont disulfure afin de quantifier son influence.

b) Graphes des énergies

Dès que l'on possède l'ensemble des 20 conformères, une étape importante est l'étude du graphe des énergies. Chaque conformère est caractérisé par la somme de l'énergie de tous les atomes (calculée d'après les équations du champ de force). Le graphe des énergies est utile pour connaître la différence d'énergie entre les conformères pour plusieurs raisons:

- Si les premiers conformères générés sont de très haute énergie comparativement aux autres conformères, cela signifie que le protocole de dynamique n'est pas au point, le système n'est pas bien "refroidi".
- Si des conformères obtenus à des temps différents sont de très haute énergie, cela montre une mauvaise minimisation des conformères.
- Par contre, si l'ensemble des conformères se situe dans une fourchette de 30 kcal/mole, on peut considérer que les conformères ont bien été minimisés.

Pour l'ensemble des peptides que nous avons modélisés et utilisés par la suite, les différences d'énergie sont inférieures à 30 kcal/mole entre conformères. L'aspect des graphes est assez ressemblant d'un peptide à l'autre: il y a toujours de 1 à 4 conformères plus stables, 1 à 3 conformères de plus haute énergie, le reste des conformères se situant entre les deux (figure 35).

c) "Cluster graphs"

L'autre étape importante est l'élaboration de graphiques en "cluster" ("grappe"): une matrice (20*20) reprend tous les conformères en ordonnée et en abscisse. Chaque conformère est comparé aux autres conformères par un RMS sur les atomes que l'on a choisis pour cette comparaison (figure 36). Le but de cette opération est de trouver des conformères qui se ressemblent. Nous avons réalisé des graphiques en "cluster" pour chaque peptide en choisissant une comparaison de la chaîne carbonée (uniquement les carbones α). Les RMS furent délimités en 4 "zones" en fonction du RMS maximal (pris d'après les 2 conformères les

[18] *Lors des modélisations, nous avons repris les résidus voisins des peptides fusionnés en pVIII ou en thiorédoxine. En ce qui concerne les superpositions que nous détailleront dans la suite du mémoire, nous n'utiliseront que les 15 (6B3-[X]) ou 16 (4F9-[1]) résidus respectifs des peptides sans les résidus voisins*

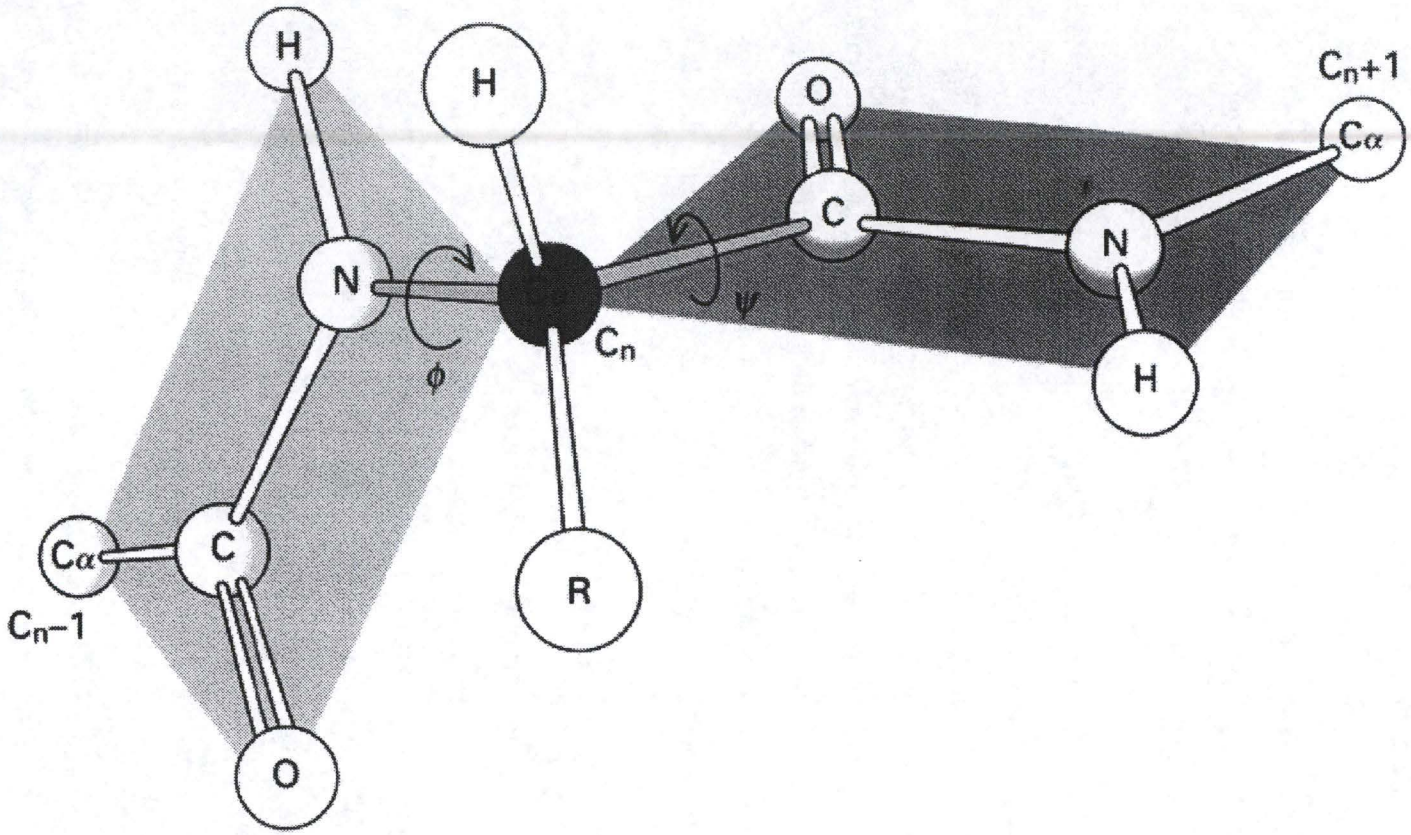


Figure 37: Représentation schématique de la localisation des angles PHI-PSI au sein d'un acide aminé.

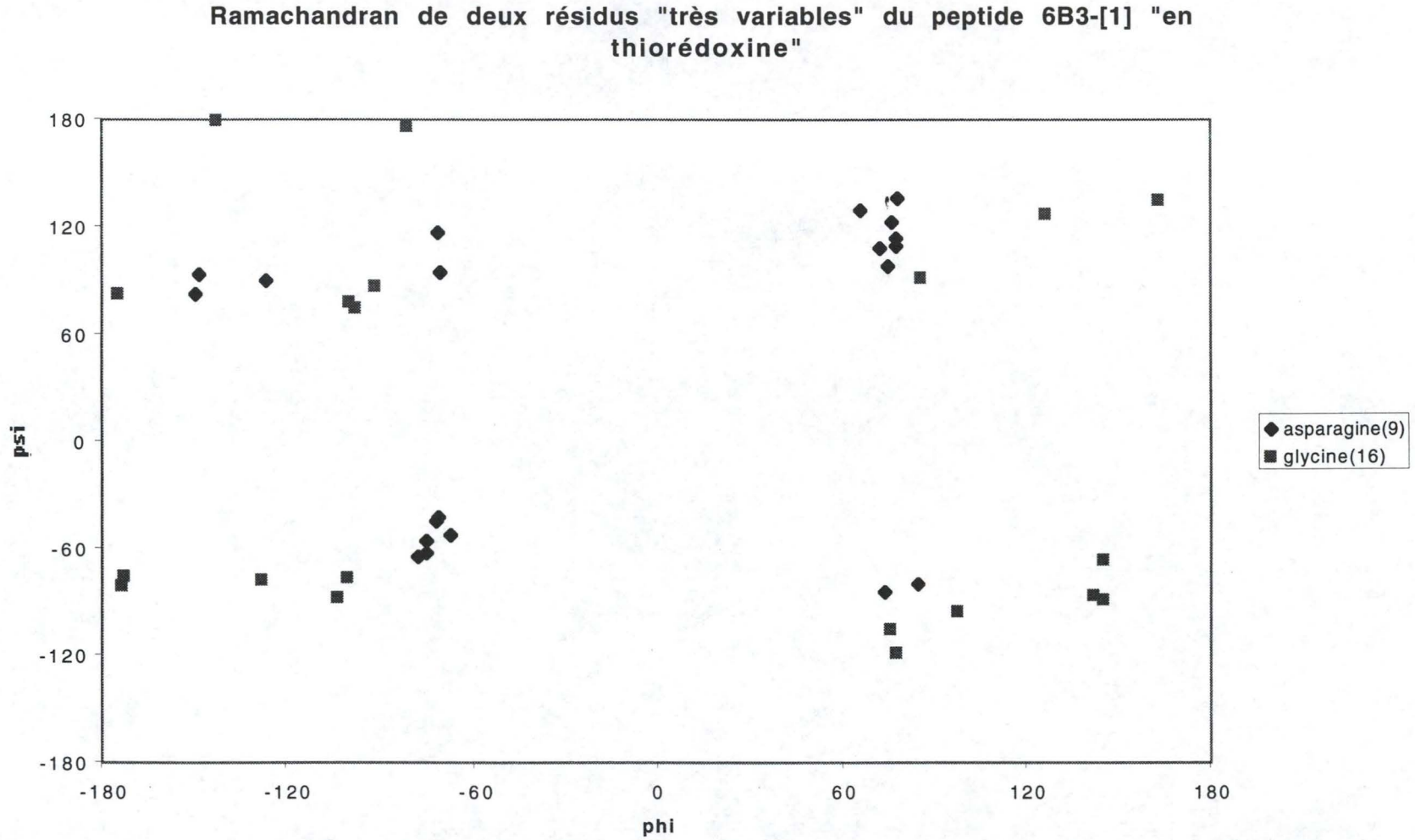


Figure 38 : Plot de Ramachandran pour les résidus du peptide 6B3-[1]T dont l'écart-type en PHI-PSI est très élevé. On peut remarquer que les diverses valeurs PHI-PSI sont dispersées aléatoirement entre les 4 cadrans.

plus éloignés). La différence de RMS la plus élevée entre les conformères des peptides à chaîne longue (en thiorédoxine et en pVIII) se situe autour de 11 Å, tandis que les peptides plus courts ont un RMS maximal proche de 9 Å. Les conformères plus longs semblent plus "éparpillés", ils forment moins de sous-groupes, certainement parce que les résidus additionnels induisent plus de variabilité, surtout en fin de chaîne, où ils sont plus libres.

3.3.4) Analyse locale de la flexibilité du squelette carboné

a) Objectifs et stratégie

Comme nous l'avons suggéré à partir des comparaisons de séquences des peptides 6B3-[X], il serait intéressant de quantifier l'influence de certains résidus sur la flexibilité locale des peptides (autour de ces résidus). En effet, on peut se demander dans quelles proportions les résidus prolines bloquent la chaîne carbonée, on peut se demander si les ponts disulfures ont une influence locale ou globale sur les différentes conformations, et si d'autres résidus contraignent également les peptides (par des ponts salins ou des ponts H intramoléculaires). Nous avons utilisé les informations tirées de la chaîne carbonée, plus exactement les angles PHI et PSI des résidus de tous les conformères (20 par peptide) (figure 37)[19]. La comparaison de ces angles entre les 20 conformères d'un même peptide par un plot de Ramachandran peut nous permettre de visualiser des différences de flexibilité. La variabilité de ces angles a également été évaluée par calcul de l'écart-type entre conformères pour chaque résidu.

b) Remarques préliminaires

Nous n'étudions dans ce cas que la flexibilité de la chaîne carbonée, il faut signaler qu'il existe d'autres angles indicatifs de la flexibilité locale (comme l'angle CHI entre le carbone α et le premier atome de la chaîne latérale). Néanmoins, comme la conformation de la chaîne carbonée ne permet que certaines positions des chaînes latérales, nous pensons que l'analyse des angles PHI-PSI nous donnera des informations suffisantes pour évaluer la contrainte exercée sur tous les résidus des peptides.

Les écart-types de ces angles ne sont qu'une indication de leur flexibilité. En effet, les écart-types sont fortement influencés par les valeurs extrêmes, or, étant donné que les conformères sont nécessairement très différents les uns des autres (car ils sont censés représenter l'ensemble des conformations que peut prendre le peptide), on se retrouve souvent avec l'un ou l'autre conformère présentant des angles très éloignés des autres conformères, ou alors avec 3 ou 4 groupes d'angles (figure 38).De plus, de par les valeurs que peuvent prendre ces

[19] *Nous n'avons pas tenu compte de l'angle OMEGA, étant donné qu'il reste bloqué autour de 180°*

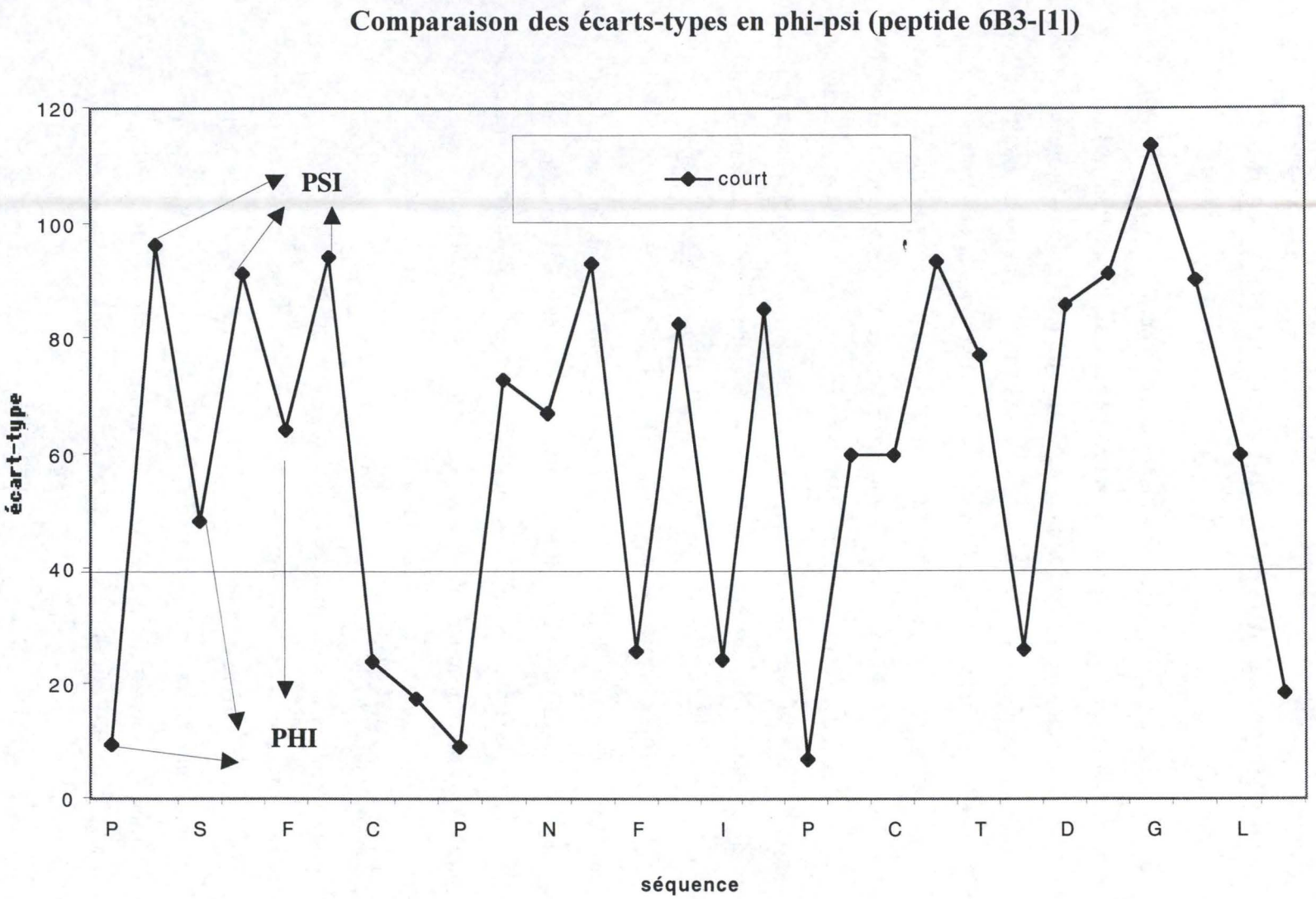


Figure 39a: Graphique reprenant les écarts-types des angles PHI-PSI pour tous les résidus du peptide 6B3-[1].
Deux valeurs sont reprises pour chaque résidu, la première est l'écart-type des angles PHI entre les 20 conformères, la seconde est l'écart-type de l'angle PSI. Les valeurs situées en-dessous du seuil (40°) sont considérées comme peu variables (par-exemple les prolines), les valeurs de plus de 80°sont considérées comme très variables (comme la glycine).

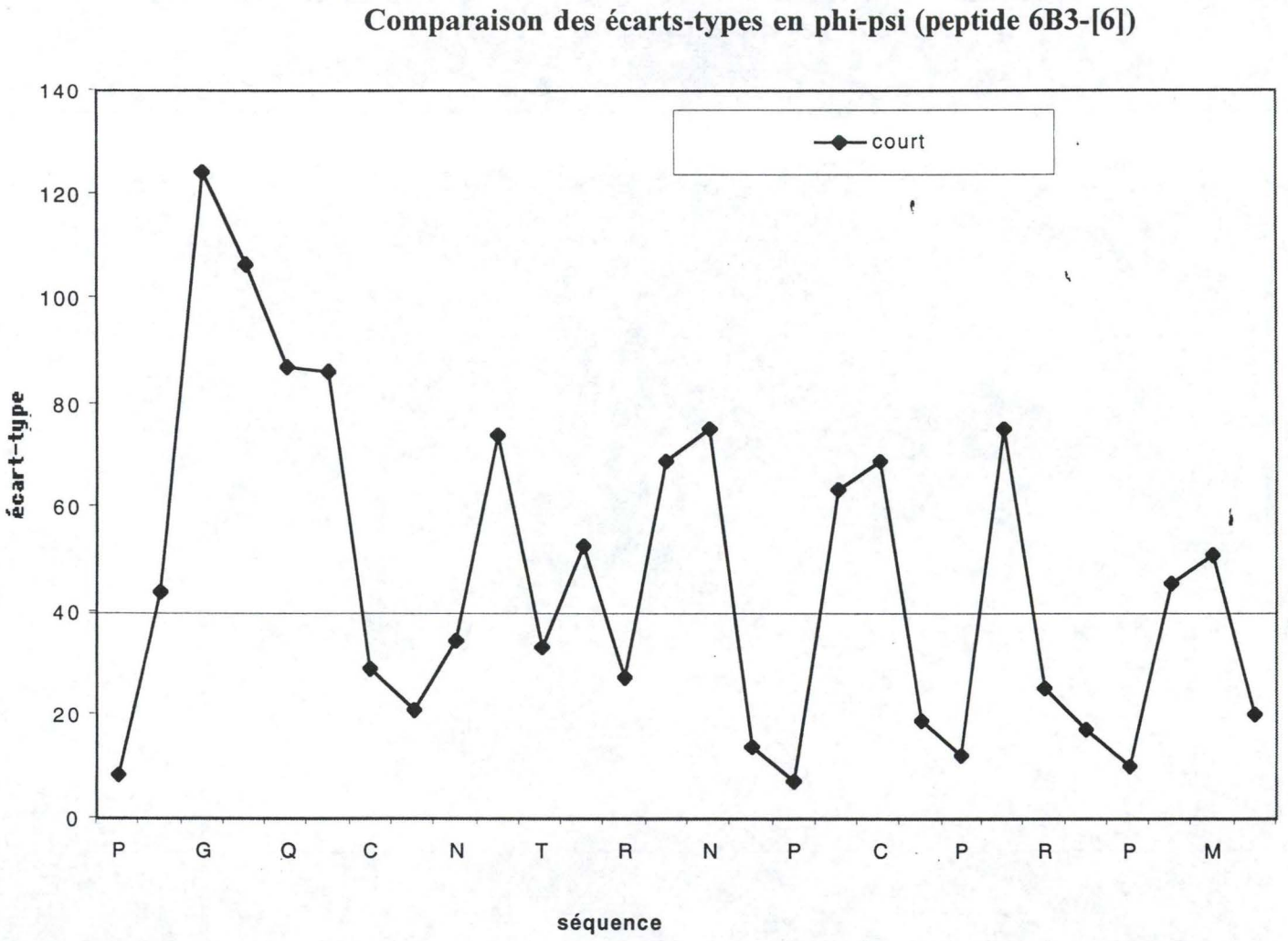


Figure 39b: Graphique reprenant les écarts-types des angles PHI-PSI pour tous les résidus du peptide 6B3-[6].
Deux valeurs sont reprises pour chaque résidu, la première est l'écart-type des angles PHI entre les 20 conformères, la seconde est l'écart-type de l'angle PSI. Les valeurs situés en-dessous du seuil (40°) sont considérées comme peu variable (par-exemple les prolines), les valeurs de plus de 80°sont considérées comme très variables (comme la glycine).

Ramachandran de deux a.a. "peu variables" et d'un a.a. "très variable" pour les 20 conformères du peptide 6B3-[1]

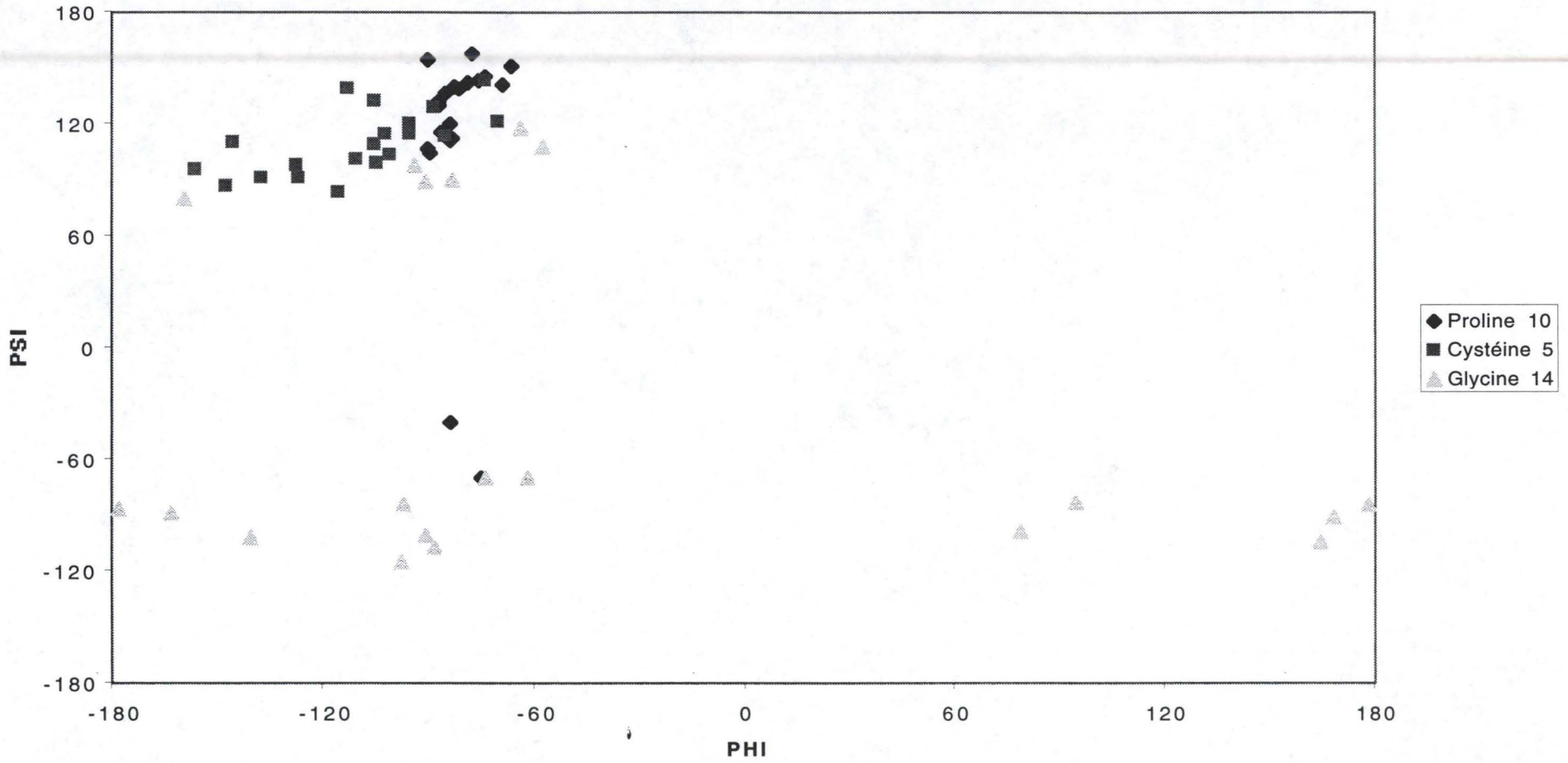


Figure 40: Plot de Ramachandran reprenant les angles PHI-PSI considérés comme peu variables et très variables après étude des écarts-types. On peut en effet observer que dans le cas de la glycine, les valeurs sont beaucoup plus dispersées que pour la proline ou la cystéine.

		6B3-1	6B3-1T	6B3-1TS	6B3-1P			6B3-2	6B3-2T	6B3-2TS	6B3-2P			6B3-4	6B3-4T	6B3-4TS	6B3-4P			6B3-6	6B3-6T	6B3-6TS	6B3-6P
A	PHI	0,00	70,62	57,66	0,00	A	PHI	0,00	50,83	43,82	0,00	A	PHI	0,00	22,81	35,97	0,00	A	PHI	0,00	22,26	59,70	0,00
	PSI	0,00	27,13	65,16	0,00		PSI	0,00	84,54	90,97	0,00		PSI	0,00	37,62	11,54	0,00		PSI	0,00	67,70	99,41	0,00
P	PHI	0,00	6,47	9,54	9,60	A	PHI	0,00	83,98	25,27	53,88	P	PHI	0,00	9,05	7,83	7,56	G	PHI	0,00	122,67	113,95	148,23
	PSI	0,00	17,48	18,55	17,52		PSI	0,00	26,60	43,67	21,63		PSI	0,00	16,38	10,93	17,64		PSI	0,00	89,43	108,93	107,86
P	PHI	9,25	10,33	11,18	10,28	P	PHI	8,45	7,71	10,86	8,27	P	PHI	8,07	9,78	7,16	9,64	P	PHI	8,46	9,67	9,71	9,67
	PSI	96,29	80,74	92,12	85,05		PSI	90,93	74,17	77,18	93,66		PSI	15,58	19,47	18,48	17,59		PSI	43,59	79,20	66,33	71,83
S	PHI	48,26	46,62	33,68	62,76	T	PHI	29,15	25,80	60,59	71,22	P	PHI	9,83	10,09	8,67	9,87	G	PHI	124,15	108,99	84,47	136,84
	PSI	91,24	102,45	85,82	91,31		PSI	84,18	97,37	75,10	92,24		PSI	88,88	89,30	78,73	95,77		PSI	106,50	71,14	89,34	111,34
F	PHI	64,21	61,40	84,92	85,09	Q	PHI	61,04	81,87	29,90	31,75	Q	PHI	26,91	94,03	43,92	75,62	Q	PHI	86,89	25,25	64,77	32,77
	PSI	93,89	96,94	88,46	97,00		PSI	88,89	70,30	71,02	95,98		PSI	16,98	82,32	14,03	14,39		PSI	86,12	56,91	87,48	76,72
C	PHI	23,76	62,39	25,84	25,83	C	PHI	25,43	51,48	68,36	31,27	C	PHI	28,80	78,09	25,99	29,21	C	PHI	29,21	65,57	24,43	30,37
	PSI	17,24	24,90	18,01	25,80		PSI	58,36	88,99	86,69	58,88		PSI	43,43	103,80	40,45	83,43		PSI	20,84	60,22	92,70	106,25
P	PHI	8,86	11,69	12,43	8,64	S	PHI	94,13	86,71	60,39	52,82	S	PHI	29,66	72,23	22,90	65,33	N	PHI	34,30	27,41	50,86	51,15
	PSI	72,65	109,29	86,65	56,76		PSI	96,51	105,20	88,49	77,95		PSI	86,11	59,09	14,85	95,25		PSI	73,95	74,59	91,77	96,43
N	PHI	67,09	88,36	69,88	31,05	Q	PHI	72,63	58,09	20,51	28,83	K	PHI	26,62	59,94	54,10	81,63	T	PHI	33,41	30,80	80,26	90,37
	PSI	92,88	86,03	71,46	76,58		PSI	103,78	47,29	66,75	32,64		PSI	20,64	82,07	64,17	100,81		PSI	52,74	104,95	99,88	81,93
F	PHI	25,55	94,39	52,55	59,11	T	PHI	62,08	32,20	31,39	67,86	A	PHI	26,22	103,25	72,03	54,67	R	PHI	27,38	54,19	28,77	50,79
	PSI	82,48	52,59	98,50	64,25		PSI	75,50	77,31	89,29	63,93		PSI	116,97	105,29	71,22	52,30		PSI	68,94	83,53	86,25	46,09
I	PHI	24,19	9,66	30,08	17,63	W	PHI	38,56	30,56	47,76	103,48	W	PHI	50,29	33,80	49,23	36,48	N	PHI	75,28	68,06	87,27	72,09
	PSI	84,99	9,32	17,34	11,65		PSI	19,51	14,22	22,17	17,68		PSI	13,85	6,25	15,88	59,52		PSI	14,07	8,85	21,18	12,99
P	PHI	6,87	11,28	10,13	10,46	P	PHI	8,89	8,26	11,19	12,83	P	PHI	4,88	7,87	10,97	10,05	P	PHI	7,16	7,10	8,80	9,81
	PSI	59,91	64,66	91,39	70,80		PSI	67,47	21,55	84,35	86,80		PSI	7,80	22,15	90,54	82,69		PSI	63,66	77,28	82,07	26,81
C	PHI	59,95	34,08	90,84	74,39	C	PHI	30,72	29,62	40,15	82,79	C	PHI	44,49	26,98	27,75	59,78	C	PHI	68,84	93,17	91,41	21,84
	PSI	93,24	26,33	101,07	97,28		PSI	85,36	20,65	69,47	24,65		PSI	18,33	66,83	20,32	34,97		PSI	18,93	21,46	25,62	17,57
T	PHI	76,95	64,45	27,53	58,54	P	PHI	9,05	5,59	11,58	10,24	P	PHI	8,15	9,05	8,19	12,04	P	PHI	12,01	10,08	9,03	11,95
	PSI	25,77	94,15	60,57	17,53		PSI	10,99	4,01	18,91	115,73		PSI	18,91	18,14	46,04	18,02		PSI	75,22	79,85	16,48	80,62
D	PHI	85,70	31,84	74,59	73,67	S	PHI	28,37	5,98	12,00	40,14	P	PHI	11,82	10,62	9,86	11,34	R	PHI	25,08	84,70	29,90	78,69
	PSI	91,23	91,90	86,15	86,01		PSI	24,33	99,17	92,40	86,03		PSI	81,68	79,04	98,75	87,42		PSI	17,05	19,44	19,41	16,99
G	PHI	113,58	128,10	101,04	121,04	H	PHI	30,44	92,73	52,13	51,56	T	PHI	59,89	34,92	62,08	83,29	P	PHI	9,98	11,46	10,75	10,27
	PSI	89,92	106,98	108,14	98,03		PSI	97,03	88,25	82,51	23,37		PSI	89,79	79,94	83,76	72,30		PSI	45,54	81,57	67,35	82,50
L	PHI	59,67	27,94	24,26	66,62	L	PHI	92,00	60,08	82,29	14,19	H	PHI	29,91	47,84	40,44	50,36	M	PHI	50,98	71,79	25,00	85,49
	PSI	18,32	22,88	21,28	20,33		PSI	11,07	23,15	20,44	9,50		PSI	23,04	16,07	88,97	17,60		PSI	20,17	22,41	17,46	24,17
P	PHI	0,00	8,23	11,39	8,48	P	PHI	0,00	9,65	9,30	8,33	P	PHI	0,00	8,87	6,81	10,70	P	PHI	0,00	10,96	9,89	10,01
	PSI	0,00	68,57	54,72	45,81		PSI	0,00	49,89	72,57	113,86		PSI	0,00	94,69	51,35	74,89		PSI	0,00	79,08	63,42	54,90
A	PHI	0,00	35,27	56,73	85,72	A	PHI	0,00	33,55	59,07	27,44	A	PHI	0,00	51,03	70,82	29,51	A	PHI	0,00	35,47	59,42	104,83
	PSI	0,00	56,22	72,31	60,91		PSI	0,00	81,58	75,49	65,58		PSI	0,00	15,17	84,82	85,14		PSI	0,00	92,35	91,37	74,96

Figure 41 : Valeurs des écart-types des angles PHI-PSI de tous les résidus entre les conformères des peptides 6B3-[X], 6B3-[X]T, 6B3-[X]P et 6B3-[X]TS ("en thiorédoxine" sans pont disulfure).

angles (-180° à 180°), si, par exemple, la moyenne pour l'angle PHI du résidu X en position 4 du peptide A est de –100° et qu'un des conformères à une valeur de 170° pour l'angle PHI, la différence à la moyenne sera de 270° au lieu de 90°. Ce type d'erreur est inhérent au système utilisé. Bien que pour la plupart des résidus que nous ayons étudiés, il n'y a jamais de telles différences, seul le plot de Ramachandran est donc utilisable quantitativement[20].

c) Résultats

Les peptides 6B3-[X] ont été analysés en premier, d'abord suivant la variabilité à chaque position (par les écarts-types) (figure 39a et 39b). Nous pouvons remarquer une très faible variabilité de l'angle PHI des résidus prolines ainsi que des voisins directs de la proline étant donné la cyclisation entre l'azote de la chaîne carbonée et la chaîne latérale. La chaîne carbonée semble également contrainte autour d'autres acides aminés: par exemple la thréonine 7 (en position 7 dans la séquence) et l'arginine 8 pour le peptide 6B3-[6], la phénylalanine 8 et l'isoleucine 9 du peptide 6B3-[1],.... En ce qui concerne les cystéines, on n'observe pas de contraintes locales au niveau de la chaîne carbonée, de plus, leur variabilité est différente d'un peptide à l'autre. Les acides aminés dont les angles varient le plus significativement sont évidemment les glycines (14 du peptide 6B3-[1], 3 du peptide 6B3-[6]).

Les résidus dont l'écart-type se situait en–dessous d'un certain seuil (fixé à 40) et les résidus très variables ont ensuite été étudiés via la position des 20 conformères sur un plot de Ramachandran comme indiqué à la figure 40 pour le peptide 6B3-[1] (les autres peptides ont les mêmes caractéristiques). On peut remarquer que les angles PHI-PSI pour les résidus prolines des 20 conformères se situent toujours entre –100° et –60° pour l'angle PHI et 120° et 180° pour l'angle PSI. Les plots de Ramachandran des autres résidus peu variables ont été réalisés et dans chaque cas, on remarque que les angles PHI-PSI se situent toujours dans la même portion du plot. Les résidus très variables sont souvent présents dans les 4 cadrans du plot.

Les mêmes analyses ont été réalisées sur les 6B3-[X] sans ponts disulfures, sur les 6B3-[X]P et sur les 6B3-[X]T (figure 41). Nous observons les mêmes profils que pour les peptides courts en ce qui concerne les résidus prolines et glycines. Par contre, les résidus contraints

---

[20] *Il faut également noter que les acides aminés situés en fin de chaîne (le premier et le dernier) ne sont pas repris étant donné que leur position induit automatiquement une plus grande flexibilité (ils sont plus « libres » et de plus , on ne peut pas détrminer les angles PHI en N-terminal et les angles PSI en C-terminal)*

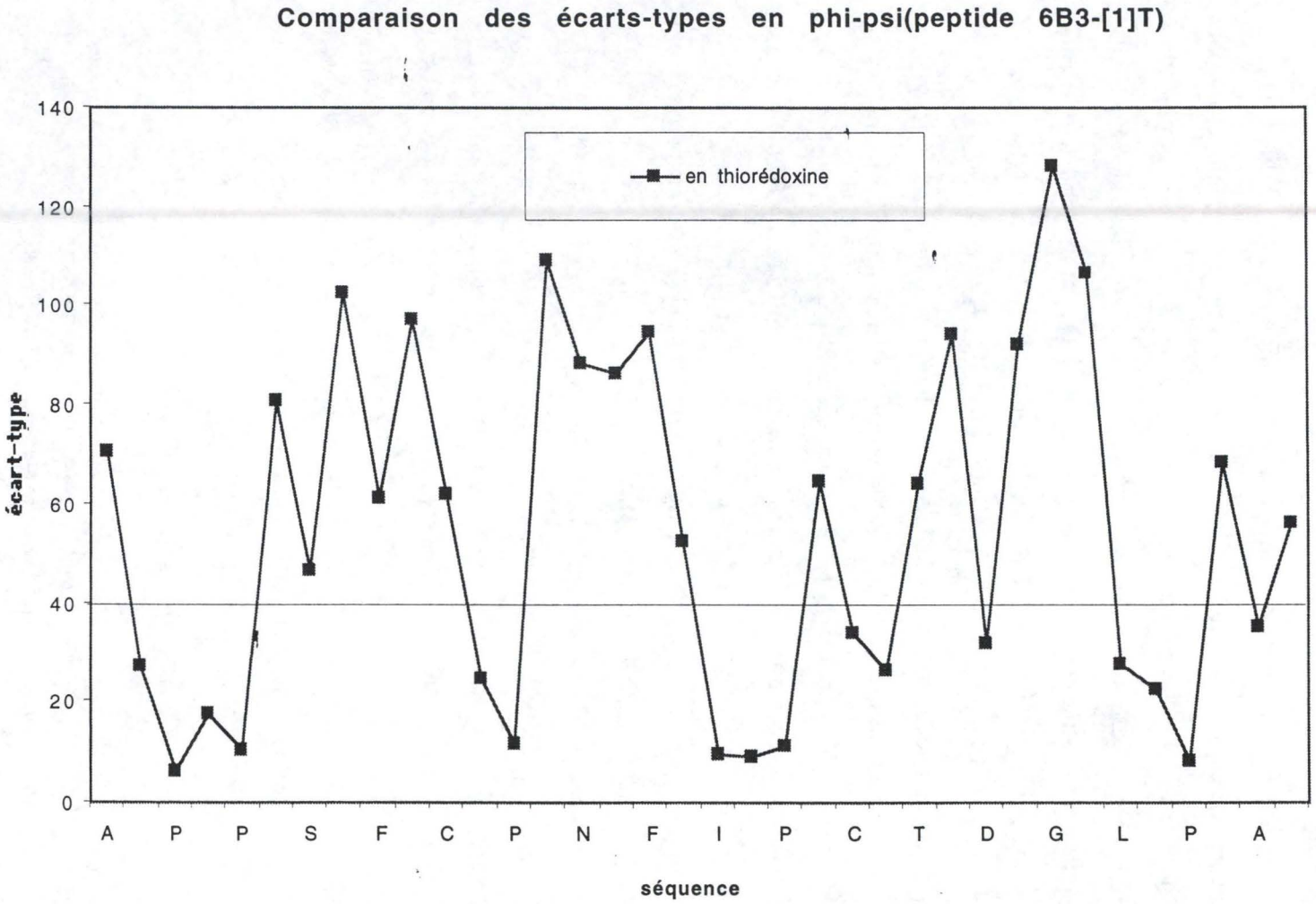


Figure 42 : Graphique reprenant les écarts-types des angles PHI-PSI pour tous les résidus du peptide 6B3-[1]T.
Les résidus peu variables hormis les prolines ne sont pas les mêmes que pour le peptide 6B3-[1] (figure 39a).
Par exemple, la cystéine est dans ce cas beaucoup plus contrainte, tandis que l'aspartate est "plus libre".

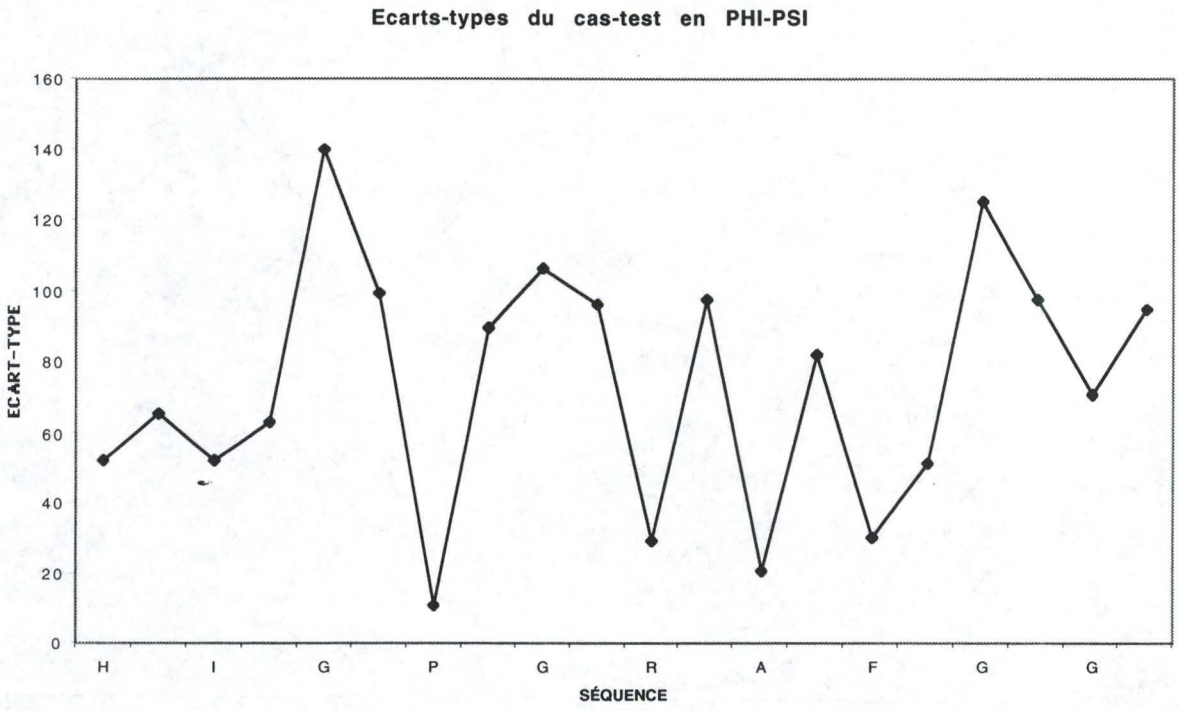
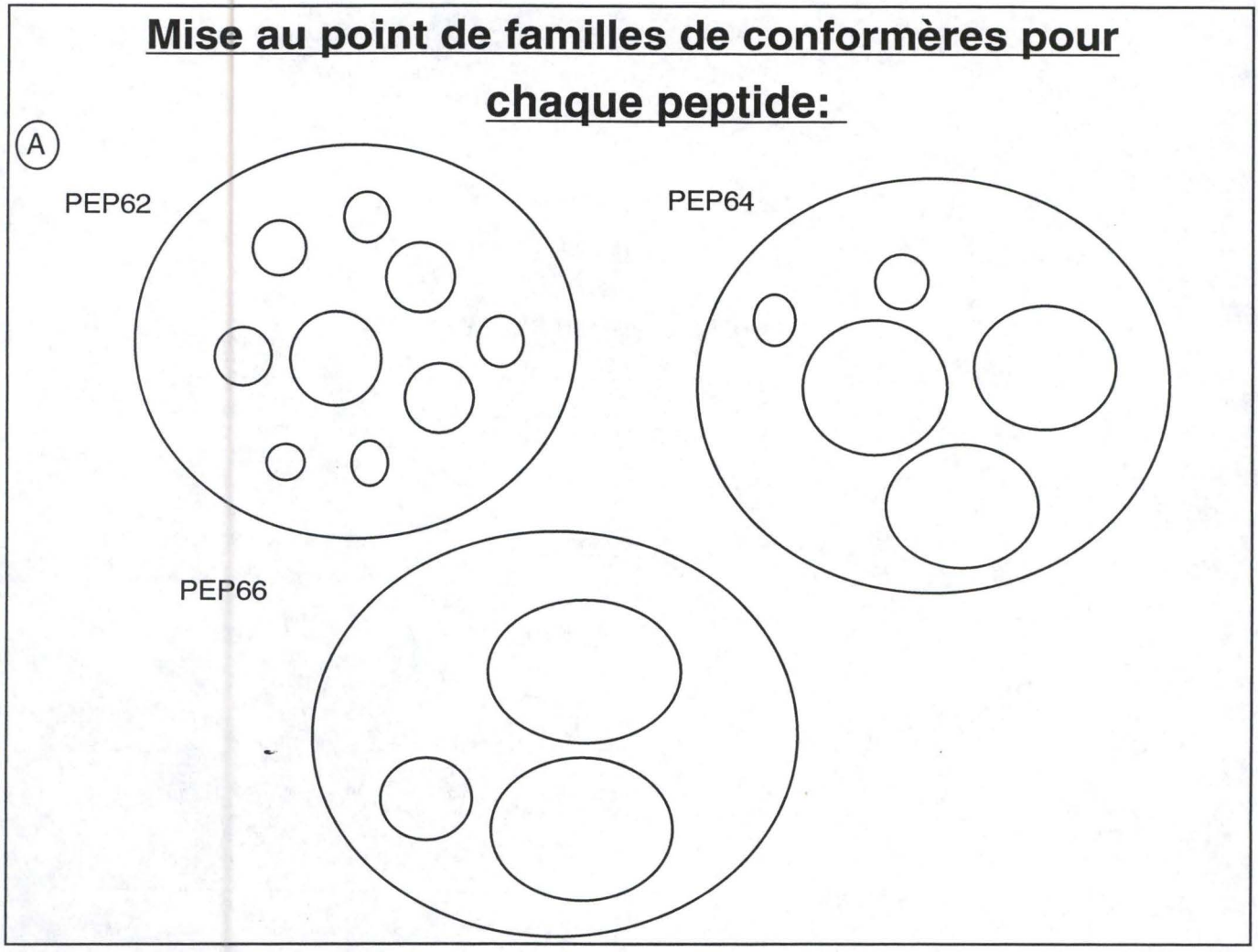


Figure 43: Graphique reprenant les écarts-types des angles PHI-PSI pour le "cas-test". Nous pouvons observer les mêmes caractéristiques que pour les peptides 6B3-[X]: faible variabilité des prolines et d'autres résidus (arginine, alanine et phénylalanine) et forte variabilité des glycines.



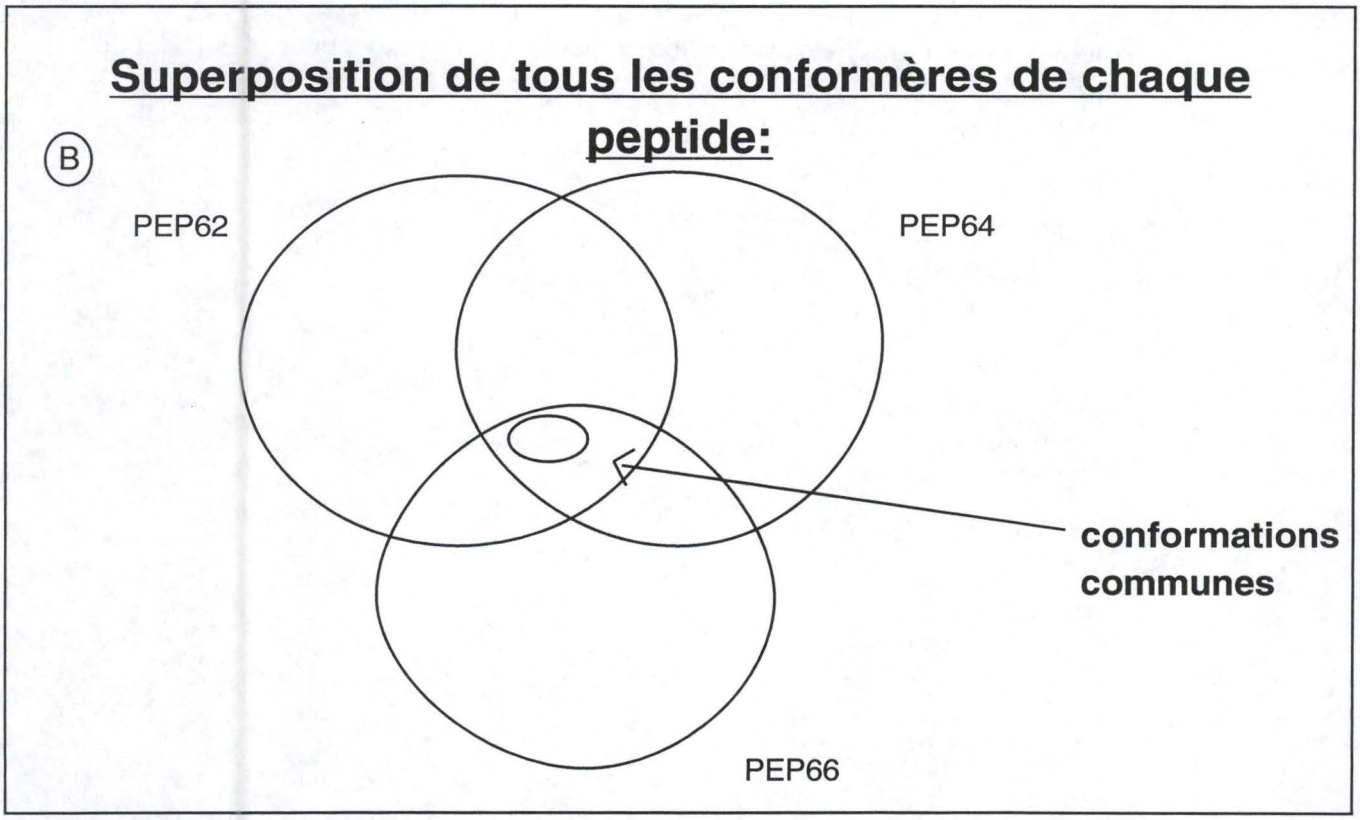


Figure 44: Schéma expliquant la méthodolgie utilisée pour la superposition de conformères de 6B3-[2],[4] et [6].
A) Un certain nombre de familles sont élaborées pour chacun des peptides (par exemple 3 pour 6B3-[6]).
B) Des familles sont mises au point pour l'ensemble des peptides en regroupant l'ensemble de leurs conformères dans le but de trouver des conformations communes.

dans le peptide court paraissent plus lâches, et d'autres résidus semblent plus contraints (glutamine 4 dans le peptide 6B3-[6] en thiorédoxine à voir dans la figure 42).

Nous disposons également des valeurs des angles PHI-PSI du peptidoïde qui nous a servi de cas-test pour valider le protocole de modélisation (figure 43) et du peptide 4F9-[1] avec ou sans pont disulfure. La proline du peptidoïde est extrêmement contrainte (écart-type de 11°) et la glycine est très lâche. Dans le cas du 4F9-[1], les résidus prolines sont tout aussi contraints.

d) Conclusions

Il ne fait aucun doute que la proline contraint fortement la chaîne carbonée et de cette manière les résidus voisins. Le pont disulfure, visiblement, ne contraint pas localement la chaîne carbonée. Il faudrait certainement reporter son effet sur l'ensemble de la conformation des peptides. La glycine induit une flexibilité locale très importante. En ce qui concerne les autres acides aminés, on ne peut, à l'heure actuelle tirer aucune conclusion. Pour étudier plus en profondeur ces flexibilités, il serait intéressant d'étudier un plus grand nombre de conformères générés par plusieurs dynamiques sur base d'une séquence particulière.

3.3.5) Etablissement de familles pour chaque peptide sur base de la conformation du squelette carboné

a) Objectifs et stratégie

Nous avons tenté d'élaborer des familles de conformères [21]pour chaque peptide sur base de la chaîne carbonée et nous avons ensuite comparé ces résultats avec les données concernant l'affinité relative de l'anticorps spécifique de chaque peptide (voir "test de reconnaissance") (figure 44a). Le but était de savoir si un peptide qui adopte des conformations plus diversifiées à moins de chance d'être reconnu par un anticorps. En effet, si le peptide modélisé peut être caractérisé par un nombre peu important de familles de conformères, et en admettant que la conformation bioactive se situe dans une de ces familles, cette conformation a beaucoup de chance d'être reconnue par l'anticorps, car le peptide est en partie préformé: il y a moins de conformères parasites ("bruit de fond") que s'il y avait de nombreuses conformations possibles. Pour ce faire, nous avons superposé les quinze résidus du peptide "seul[22] des conformères des peptides 6B3-[X]P et des peptides 6B3-[X]T ainsi que les conformères du

---

[21] *les familles sont des sous-groupes de conformères d'un même peptide entre lesquels le RMS est inférieur à une valeur-seuil prédéfinie*

[22] *cette opération permet de comparer les résultats obtenus pour les superpositions des 6B3-[X]P et des 6B3-[X]T*

	Numéro de la famille	Conformère chef de file	Nombre de conformères	Composition de la famille
6B3-[1]	1	20	5	4 13 18 19 20
	2	17	7	3 6 7 14 15 16 17
	3	12	4	11 12 13 18
	4	10	9	1 2 4 5 6 8 10 11 1
	5	9	2	8 9
6B3-[2]	1	20	3	18 19 20
	2	17	6	2 12 14 15 17 18
	3	16	4	12 13 15 16
	4	11	9	2 3 6 9 10 11 12 14 15
	5	8	3	2 7 8
	6	5	6	1 3 5 6 7 9
	7	4	1	4
6B3-[4]	1	20	14	1 3 4 5 6 7 8 9 14 15
	2	16	3	11 16 17
	3	13	8	6 7 8 9 10 11 12 13
	4	2	2	2 3
6B3-[6]	1	20	11	4 7 11 12 13 14 17 18 19 1
	2	16	8	2 3 5 7 15 16 19 20
	3	10	10	2 3 5 8 9 10 11 15 19 1
	4	6	7	2 3 5 6 7 9 1

Figure 45: Tableau reprenant les familles de conformères réalisées pour chacun des peptides 6B3-[X] avec une valeur de RMS de 4 Å.

peptide 4F9-[1] "seul" (16 résidus). Nous avons également réalisé un contrôle: la superposition des conformères d'un des peptides sans pont disulfure. La méthodologie de superposition est explicitée dans le chapitre "matériel et méthode": elle se base sur le graphe d'énergie et sur le "cluster graph".

b) Résultats

La première étape a été la superposition des peptides 6B3-[X]P. Le seuil de RMS pour l'élaboration des familles était fixé à 4 Å ou 3 Å (deux conformères dont la somme des RMS pour chaque atome est inférieure à 4 Å ou 3 Å appartiennent à la même famille). Les résultats pour chaque peptide sont repris à la figure 45 avec un RMS de 4 Å.

Dans les 2 cas, on peut observer que le peptide 6B3-[2] présente des conformations plus diversifiées. Les conformations des peptides 6B3-[1], 6B3-[4] et 6B3-[6] semblent plus contraintes. Les conformères du peptide 4F9 sont regroupés en un nombre de familles significativement plus élevé alors qu'ils ne comportent qu'un seul résidu en plus que les peptides 6B3-[X][23].

Nous avons également formé des familles de peptides 6B3-[X]T avec ou sans pont disulfure. Bizarrement, nous obtenons moins de familles pour le peptide 6B3-[2]T comparativement au 6B3-[2]P et aux autres peptides 6B3-[2]T.

En ce qui concerne le peptide sans pont disulfure, nous pouvons observer un nombre significativement plus élevé de familles que lorsqu'il est présent.

c) Conclusions et hypothèses

Le nombre de familles des peptides 6B3-[X]P est inversement proportionnel au nombre de prolines présentes dans chacun des peptides. Cette observation montre que les prolines ne contraignent pas seulement le squelette carboné des résidus voisins, mais également l'ensemble de la chaîne carbonée des peptides.

Les tests d'affinité de chaque peptide 6B3-[X]P montrent que le peptide 6B3-[6]P est beaucoup moins bien reconnu par l'anticorps que les autres peptides, or c'est pour ce peptide que l'on obtient le plus de familles. Nous pouvons dès lors émettre l'hypothèse que ce peptide est moins bien reconnu car la conformation qu'il adopte au sein de l'anticorps est "noyée" au milieu d'autres conformations. Malgré tout, il faut savoir que l'affinité des autres peptides pour cet anticorps n'est pas proportionnelle au nombre de familles que nous avons obtenu

[23] *un résidu supplémentaire induit un peu plus de variabilité sur l'ensemble de la chaîne carbonée*

pour chacun d'entre eux. Un nombre minimum de prolines dans la séquence des peptides 6B3-[X]P serait donc important pour la reconnaissance.

Le nombre de conformations est également influencé par l'absence ou la présence du pont disulfure. Il a été prouvé que le pont disulfure est important dans la reconnaissance par l'anticorps (voir "test au DTT"). Nous pouvons relier cette reconnaissance au fait qu'en l'absence du pont disulfure, le peptide adopte au moins deux fois plus de conformations différentes que lorsque cette liaison est présente.

3.3.6) Etablissement de familles entre tous les conformères de tous les peptides 6B3-[X]P

a) Objectifs et stratégie

L'objectif suivant est l'application du principe du "pharmacophore" (figure 44b). Après modélisation des peptides, nous pouvons définir une série de paramètres stériques et hydropathiques. Si on admet que chaque peptide 6B3-[X] est reconnu de la même façon par l'anticorps 6B3 (avec les mêmes conformations[24]), un consensus au niveau structural peut nous permettre d'établir le pharmacophore.

La première étape est la superposition de la chaîne carbonée de tous les conformères des peptides 6B3-[2], [4]et [6]. Cette superposition est une donnée importante puisque l'orientation des chaînes latérales dépend fortement de la conformation de la chaîne carbonée (ce sont les atomes des chaînes latérales qui confèrent la spécificité et l'affinité à l'interaction avec les CDR de l'anticorps).

b) Résultats

Nous avons superposé les 3 peptides (6B3-[2], [4] et [6]) et un peptide sélectionné par un autre anticorps (2C8): le peptide 2C8-[2] [25] dans le but de voir si les peptides sélectionnés par 6B3 sont significativement plus proches entre eux au niveau de la chaîne carbonée que d'un peptide de même longueur sélectionné par un autre anticorps.

Nous avons donc comparé les 80 conformères entre eux (20 conformères pour chaque peptide) en fixant le seuil de RMS à 4 Å et ensuite à 3 Å. Nous avons observé plusieurs types de familles pour les deux valeurs de RMS:

1) Certaines sont composées exclusivement de conformères d'un seul peptide.

---

[24] *l'épitope doit comprendre plus ou moins les mêmes fonctionnalités chimiques et doit se trouver au même endroit au-niveau de la séquence, sans quoi les peptides ne seraient pas superposables*

[25] *ce peptide provient de la même banque de départ* $(X)_4\underline{C}(X)_5\underline{C}(X)_4$

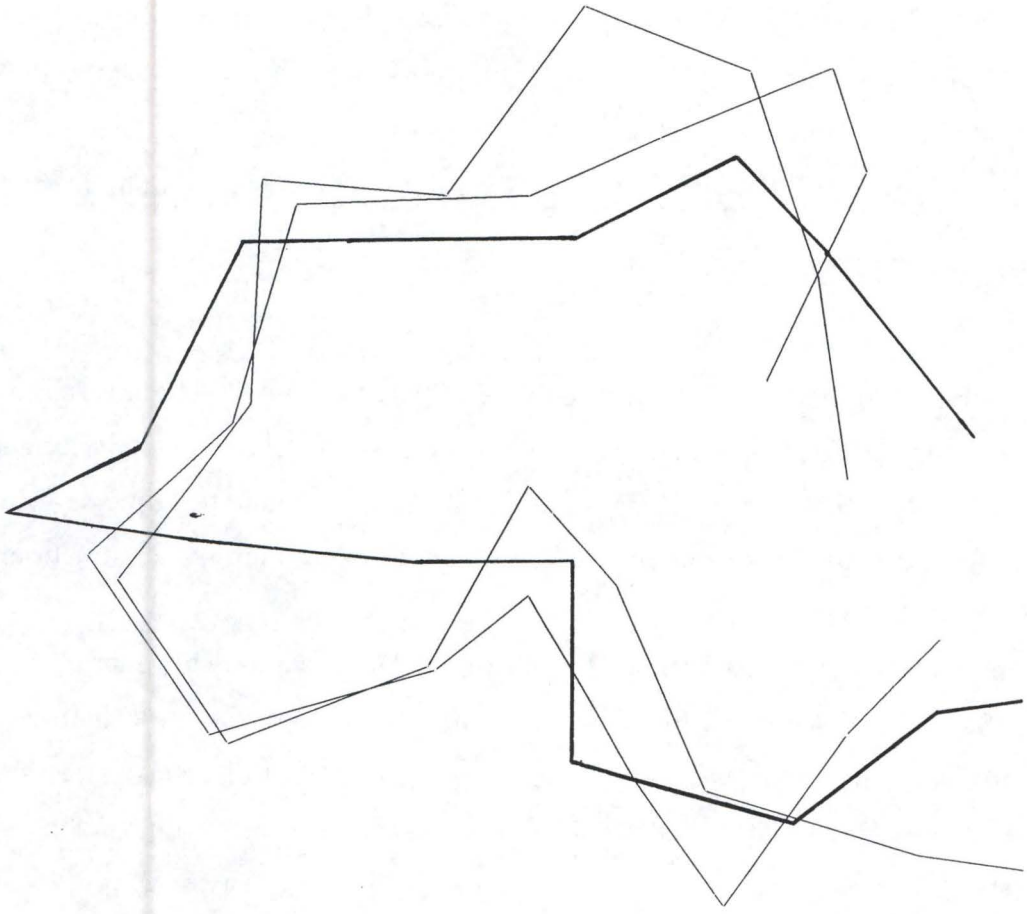


Figure 46: Image d'une des familles (14) après superposition de tous les conformères de tous les peptides , cette famille est composée de conformères provenant du peptide 6B3-[2], 6B3-[4] et 2C8-[2]. Les conformères des peptides 6B3-[2] et 6B3-[4] sont très proches pour les résidus situés entre les cystéines tandis que les résidus terminaux semblent moins proches. Du fait de sa situation entre les deux autres conformères pour les résidus terminaux, le peptide 2C8-[2] appartient à la même famille que les deux autres conformères , et ce, malgré une différence marquante pour les résidus centraux. La chaîne carbonée du conformère du peptide 2C8-[2] est représenté en gras, et les 2 conformères des peptides 6B3-[X] sont en trait plus fin.

RMS 3 Å

Numéro de la famille	Conformère "chef de file"	Nombre de conformères	Composition de la famille
1	60	33	5 6 7 10 11 12 13 15 16 21
			40 41 42 44 45 47 48 50 51 52
			53 54 55 56 57 58 59 60 63 66
			68 69 74
2	49	55	3 5 10 12 13 14 15 16 17 20
			21 22 23 24 25 26 27 28 29 30
			31 32 33 34 35 37 38 39 40 41
			42 43 44 45 46 47 48 49 50 51
			52 53 61 63 65 66 70 71 72 74
			75 76 78 79 80
3	36	19	4 19 23 24 25 26 27 28 29 30
			31 32 33 34 35 36 37 38 75
4	18	51	1 2 3 4 5 7 8 9 10 12
			13 14 15 16 17 18 19 20 23 24
			26 29 30 35 37 39 40 41 42 48
			50 51 52 53 61 62 63 64 65 67
			70 71 72 73 74 75 76 77 78 79
5	80	1	80

RMS 2 Å

Numéro de la famille	Conformère "chef de file"	Nombre de conformères	Composition de la famille
1	60	7	45 54 55 56 57 59 60
2	58	6	54 57 58 59 68 69
3	53	14	2 40 41 42 43 44 47 48 49 50 51 52 53 57
4	46	11	37 41 43 44 45 46 47 48 49 51 52
5	39	17	3 22 23 26 29 30 37 38 39 40
			41 43 48 49 50 51 52
6	36	2	24 36
7	35	14	22 23 24 25 28 34 35 37 38 72 75 77 79 80
8	21	5	21 45 55 56 61
9	20	5	4 18 19 20 62
10	17	20	3 7 12 14 15 17 18 40 61 62
			63 65 70 71 72 74 76 78 79 80
11	16	3	15 16 61
12	13	11	5 12 13 14 15 44 47 61 63 65 66
13	11	7	6 11 57 66 67 68 69
14	10	17	1 3 5 6 7 8 9 10 12 63

RMS 1 Å

Numéro de la famille	Conformère "chef de file"	Nombre de conformères	Composition de la famille
1	60	1	60
2	59	1	59
3	58	1	58
4	57	1	57
5	56	2	55 56
6	54	1	54
7	53	2	42 53
8	52	6	41 44 47 48 51 52
9	50	2	40 50
10	49	3	43 48 49
11	46	1	46
12	45	1	45
13	39	3	29 39 40
14	38	1	38
15	37	1	37

RMS 1 Å
(suite)

Numéro de la famille	Conformère "chef de file"	Nombre de conformères	Composition de la famille
16	36	1	36
17	35	3	24 28 35
18	34	3	25 28 34
19	33	4	27 30 32 33
20	31	2	27 31
21	26	2	26 29
22	23	2	22 23
23	21	1	21
24	20	2	19 20
25	18	1	18
26	17	2	14 17
27	16	1	16
28	15	1	15
29	13	1	13
30	12	1	12
31	11	1	11
32	10	1	10
33	9	1	9
34	8	1	8
35	7	2	7 65
36	6	1	6
37	5	1	5
38	4	1	4
39	3	1	3
40	2	2	1 2
41	80	7	71 72 76 77 78 79 80
42	75	1	75
43	74	2	72
44	73	1	73
45	70	1	70
46	69	3	67 68 69
47	66	1	66
48	64	1	64
49	63	1	63
50	62	1	62
51	61	1	61

Figure 47: Tableau reprenant les familles élaborées entre l'ensemble des conformères de chaque peptide avec les résidus centraux selon un RMS de 1, 2 ou 3 Å.

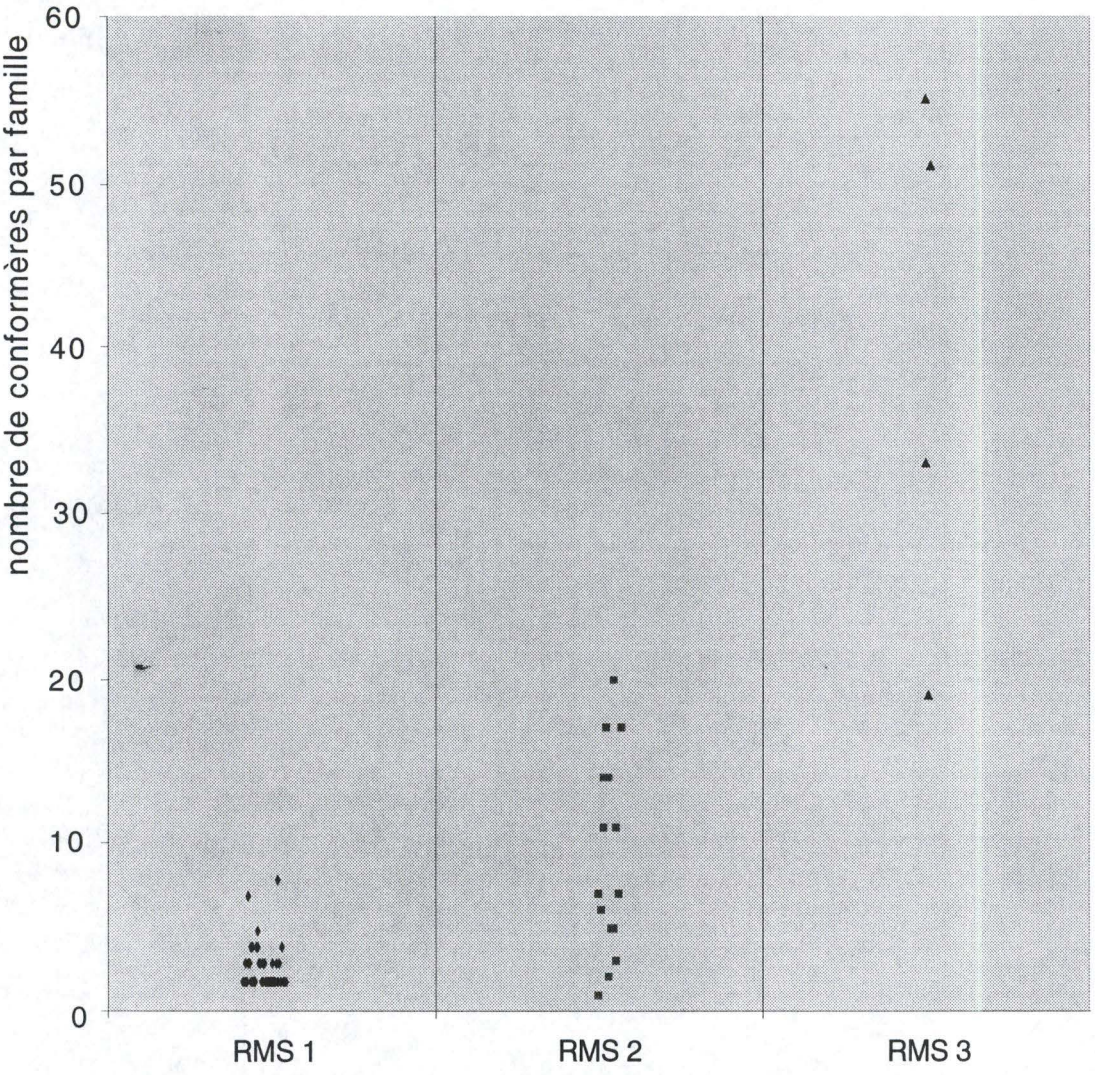


Figure 48: Graphique exposant le nombre de conformères par famille pour la superposition des peptides 6B3-[2], [4] et [6] selon 3 valeurs de RMS.
Une valeur seuil de 2 Å semble bien adaptée à l'étude en familles: un seuil de 1Å est trop bas: les familles que l'on obtient ne sont composées en moyenne que de 2 ou 3 conformères, et un seuil de 3 Å est trop élevé: les familles sont trop fournies, elles ne sont donc pas significatives.

2) D'autres sont composées de conformères provenant de 2 des peptides 6B3-[X].

3) Enfin, les dernières familles sont composées de conformères des peptides 6B3-[X] et des conformères du peptide 2C8-[2].

Aucune famille ne rassemble à la fois des conformères des peptides 6B3-[2], [4] et [6], sans conformères du 2C8-[2]. Par cette méthode, les peptides 6B3-[2], [4] et [6] ne sont donc pas significativement superposables. Nous pouvons interpréter ces résultats de 2 manières:

1) Les peptides ne sont pas assez proches au niveau de la chaîne carbonée parce que les peptides qui comprennent 5 résidus entre leurs cystéines adoptent le même type de conformation (les peptides 6B3-[2], [4] et [6] ressemblent donc autant au peptide 2C8-[2] qu'ils ne se ressemblent entre eux).
2) La méthode de regroupement en familles que nous utilisons n'est pas au point: les peptides 6B3-[2], [4] et [6] se ressemblent significativement plus au niveau de la chaîne carbonée qu'ils ne ressemblent au peptide 2C8-[2].

Après visualisation des différentes familles à l'écran, nous avons observé que la chaîne carbonée des peptides 6B3-[X] n'adoptait pas le même type de conformation que la chaîne carbonée du peptide 2C8-[2], surtout en ce qui concerne les résidus situés entre les cystéines (figure 46). La superposition étant globale, il est possible que la flexibilité des extrémités des peptides compense la différence de conformation des résidus situés entre les cystéines. C'est sans doute également pour cette raison que la valeur-seuil pour la formation de familles était si élevée (on ne trouve plus de familles en utilisant un RMS de 2 Å). Nous avons donc décidé à ce moment de ne superposer que 9 acides aminés centraux: les 5 résidus entre les cystéines, les cystéines ainsi que les résidus voisins des cystéines ($X\underline{C}X_5\underline{C}X$).

Dans ce cas, nous avons fixé les valeurs-seuil de RMS pour la formation de familles à 3, 2 et 1 Å. Le nombre de conformères de chaque famille est repris ci-contre (figure 47). Un seuil de 1 Å est trop strict: on ne trouve pas de familles significatives (elles sont trop peu fournies). Le seuil de 3 Å est trop large: on se retrouve avec des familles beaucoup trop fournies. Le seuil de 2 Å semble être un bon compromis pour pouvoir différencier les conformères des peptides 6B3-[X] des conformères 2C8-[2] (figure 48). En effet, 2 familles sont composées exclusivement de conformères des peptides 6B3-[2], [4] et [6].

c) Hypothèses

En résumé, les peptides 6B3-[2], [4] et [6] sont tous reconnus par l'anticorps 6B3 et sont proches au niveau de la séquence. On peut donc penser qu''ils sont reconnus de la même

manière par l'anticorps 6B3. Ils sont significativement superposables pour une partie de leurs résidus (en comparaison avec un peptide non-reconnu par 6B3).

Nous émettons donc l'hypothèse que la conformation reconnue par l'anticorps se situe dans une des deux familles citées ci-dessus. Ces deux familles de conformères doivent donc être étudiées plus en détail pour la recherche du pharmacophore.

d) Qualité du modèle

Il y a malgré tout quelques remarques à faire quant à l'application de ce type de démarche:

1) Les conformations que nous avons superposées ont été modélisées: ces conformations ne tiennent donc pas compte de contraintes imposées par l'interaction avec un paratope. Même si nous avons généré un nombre important de conformères, il est donc possible que nous soyons passé "à côté" de la conformation des peptides en complexe.

En général, on considère tout de même que la conformation en complexe est proche des conformations stables dans un environnement homogène.

2) Nous devons remarquer que nous n'avons utilisé qu'un seul contrôle (2C8-[2]). Nous pensons qu'il serait intéressant d'utiliser d'autres peptides non-reconnus par 6B3 dans la même optique pour pouvoir regrouper significativement les 6B3-[X].

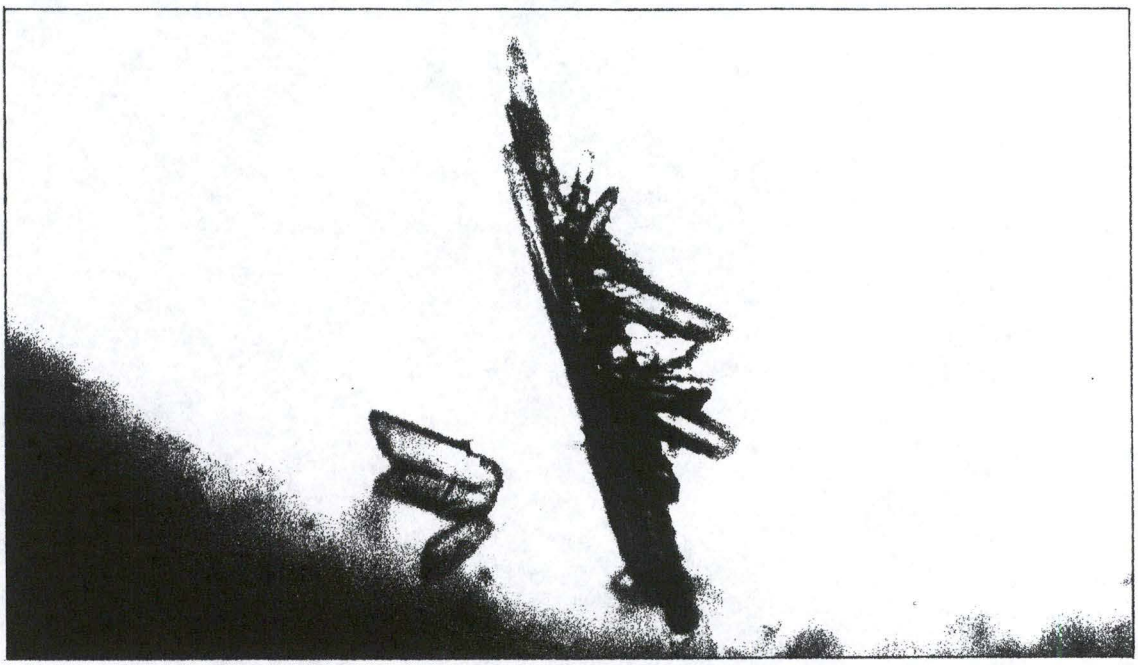
3) Dans le cas où l'hypothèse qui consiste à dire que les peptides 6B3-[X] sont significativement plus proches entre eux que d'un autre peptide de même composition est fausse[26], il est tout de même intéressant d'utiliser ce regroupement par familles. En effet, la formation de familles limitera le nombre de conformations à analyser pour la recherche de la conformation du phamacophore: comme les conformères d'une famille sont considérés comme très proches (un RMS de 2 Å sur 9 atomes est un seuil assez strict), il suffira d'utiliser un conformère de chaque famille où sont présents des conformères des peptides 6B3-2,4 et 6 (par exemple le "chef de file") pour la recherche du pharmacophore.

---

[26] *il y a autant de familles composées exclusivement de conformères des peptides 6B3-[X] que de familles composées de conformères provenant des peptides 6B3-[X] et 2C8-[2]!*

CONCLUSIONS ET PERSPECTIVES

A



B

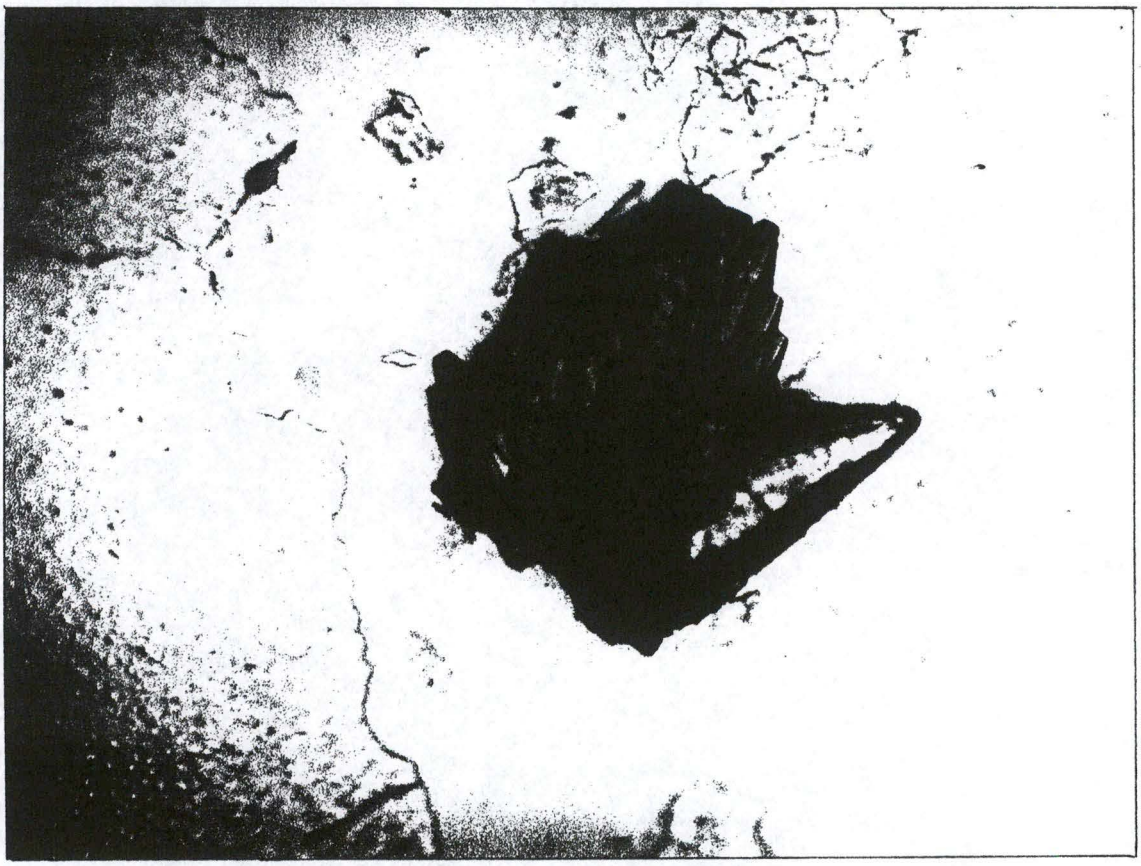


Figure 49: Photographie de deux des cristaux qui ont été obtenus lors des essais de cristallisation. Ces 2 cristaux ont été colorés par du bleu de Coomassie.
a) polycristal de 4F9
b) polycristal d'une thiorédoxine fusionnée à un des peptide 6B3-[X]

Conclusions générales et perspectives

Au terme de ce mémoire, les objectifs que nous nous étions fixés ont été partiellement atteints:

Une série de conditions de cristallisation ont été déterminées pour l'anticorps 4F9, spécifique de la chaîne O de *Brucella* et d'une série de protéines de fusion (thiorédoxines) portant des peptides mimétiques de la chaîne O.

Nous avons mis au point un protocole de modélisation des peptides 6B3-[X] et du peptide 4F9-1 par dynamique et mécanique moléculaire; ce protocole nous a permis de générer 20 conformères représentatifs de l'ensemble des conformations que pourraient prendre ces peptides en complexe avec l'anticorps. De plus, leur flexibilité a été étudiée via la variabilité des angles PHI-PSI.

Enfin, nous avons regroupé les conformères de chaque peptide en familles, au sein desquelles il est théoriquement probable de trouver une conformation proche de la conformation reconnue par l'anticorps.

ASPECT EXPERIMENTAL: CRISTALLISATION DU FRAGMENT FAB DE L'ANTICORPS 4F9 ET DE PROTEINES DE FUSION THIOREDOXINE-PEPTIDES MIMETIQUES DE LA CHAINE O DE BRUCELLA

Dès le début des essais d'obtention de cristaux de Fab (4F9) et de peptides en fusion, nous avons été confronté au caractère empirique des tests de cristallisation: les premiers tests réalisés sur les Fab et sur les protéines-chimères inspirés de conditions de cristallisation reprises dans la littérature n'ont pas donné de résultats satisfaisants. Dans les 2 cas, des "screen", reprenant diverses compositions en agents précipitants, tampons et sels ont dû être réalisés et ont permis de délimiter des conditions intéressantes pour la cristallisation des protéines d'intérêt. A la suite de screen successifs qui ont permis d'affiner les conditions de cristallisation, des polycristaux des protéines d'intérêt ont été formés en grande quantité (figure 49).

Bien que les cristaux obtenus ne soient pas directement utilisables pour la diffraction aux rayons X, leur croissance pourra être améliorée par isolement de monocristaux et réensemencement d'une nouvelle liqueur-mère. D'autres tests seront réalisés grâce aux quantités importantes de protéines purifiées dont nous disposons, soit en modifiant les

conditions établies dans ce travail soit en testant de nouveaux kits (par exemple un "screen" spécial pour les fragments Fab).

Il semble important de réaliser de nombreux tests de coloration différents afin d'augmenter les chances d'identification des vrais cristaux de protéine.

Une fois les conditions affinées, les tests de cristallisation seront étendus à de nouveaux fragments Fab, provenant d'autres anticorps monoclonaux spécifiques de la chaîne O de *Brucella* (2C8 et 6B3).Signalons également que d'autres protéines "carrier" comme le lysozyme, vont être utilisées pour la présentation des peptides mimétiques de la chaîne O de *Brucella*. Il serait intéressant de cristalliser ces protéines chimères (par la technique habituelle de "hanging drops" ou par "cross-seeding" avec les cristaux de lysozyme disponibles au laboratoire) afin de trouver d'autres conformations des peptides mimétiques (reconnues ou non par les anticorps spécifiques de la chaîne O).

Une alternative expérimentale à la détermination de structure des peptides pourrait être la RMN 2D. Cet outil permet une analyse en solution des peptides seuls. Les données sur la flexibilité et la stabilité des peptides en solution fourniraient des informations complémentaires aux données disponibles par diffraction (qui permet d'étudier les interactions intra- et intermoléculaires au sein de l'empilement cristallin).

ASPECT THEORIQUE: MISE AU POINT D'UN PROTOCOLE DE MODELISATION PAR DYNAMIQUE MOLECULAIRE ET MINIMISATION D'ENERGIE DE PEPTIDES MIMETIQUES ET ETUDE D'UNE SERIE DE CONFORMERES GENERES PAR CETTE MODELISATION

Un protocole de modélisation par dynamique moléculaire et minimisation d'énergie de peptides composés de 10 à 20 résidus, contraints par un pont disulfure a été adapté d'un protocole général conseillé par les concepteurs du logiciel de modélisation.

Ce protocole a été validé par un "cas-test" de taille analogue aux peptides étudiés (la structure expérimentale de ce "cas-test" en complexe est très proche de deux des conformations modélisées). Une validation plus poussée du protocole devrait être réalisée avec d'autres peptides contraints de taille analogue dont on connaît la structure en complexe (à l'heure actuelle, il n'y en a aucun autre disponible).

Une approche originale, permettant l'étude de la flexibilité de la chaîne carbonée via des plots de Ramachandran et la variabilité des angles PHI-PSI pour chaque résidu de tous les conformères de chaque peptide a été réalisée. Les résultats apportés par cette étude pour les

peptides 6B3-[X] a permis de quantifier l'importance des prolines et des glycines dans la flexibilité locale des peptides. Nous pouvons également supposer que le pont disulfure ne contraint pas directement la chaîne carbonée. Son effet serait plutôt marqué sur la conformation globale du peptide. La variabilité des autres résidus communs entre les peptides 6B3-[X] n'est pas la même d'un peptide à l'autre. Il faut donc analyser les résidus de chaque peptide indépendamment, afin de mettre en évidence les acides aminés importants pour le maintien de la conformation en solution via des interactions intramoléculaires (ces liaisons sont certainement spécifiques d'un peptide à l'autre). Nous avons également mis en évidence l'importance de la présence ou de l'absence de résidus bordant sur la variabilité des angles PHI-PSI.

Il serait intéressant d'affiner ces résultats en utilisant plus de conformères générés par différents types de dynamiques, et en les étudiant plus en profondeur par visualisation à l'écran. Dans cette optique, l'étude de la variabilité des angles "CHI" est aussi envisageable. Nous pensons également que la flexibilité telle qu'étudiée par ce type d'analyse devrait être comparée aux résultats que l'on pourrait obtenir par RMN. Cette comparaison permettrait de valider et de mieux comprendre les données "PHI-PSI" qui sont en notre possession.

Finalement, des superpositions entre conformères ont été réalisées selon un RMS sur la chaîne carbonée, d'abord pour chacun des peptides 6B3-[X] séparément, et ensuite pour l'ensemble des conformères des peptides 6B3-2,4 et 6 dans le but d'appliquer le principe du pharmacophore.

Nous avons émis l'hypothèse que le peptide 6B3-[2] est moins bien reconnu car il génère plus de familles et donc plus de conformations différentes que les autres peptides 6B3-[X]. Il est visiblement important de contraindre les peptides 6B3-[X] par un nombre minimal de prolines placées à intervalle régulier dans le peptide.

Nous avons regroupé les conformères des peptides 6B3-[2], [4] et [6] en familles sur base du même principe (selon le RMS des atomes de la chaîne carbonée). Des conformères de chacun de ces peptides sont superposables. Ces conformères sont nos candidats privilégiés pour la recherche d'une conformation très proche de la conformation du peptide en complexe.

De nouveaux peptides issus de la même banque et reconnus par 6B3 (comme le peptide 6B3-[1]) pourraient augmenter la pertinence des résultats obtenus lors de cette première superposition.

Cette démarche est également adaptable aux autres peptides mimétiques de la chaîne O du LPS de *Brucella*, s'ils possèdent une séquence assez proche et s'ils ont été isolés par le même

anticorps (par exemple pour une série de peptides reconnus par l'anticorps 2C8, spécifique de l'épitope A).

La méthode suivie au cours de ce travail est novatrice: la détermination du pharmacophore est courante en chimie structurale (pour de petits ligands qui interagissent avec le même récepteur). En revanche, à notre connaissance, cette approche n'a jamais été réalisée à l'aide de peptides reconnus par un même anticorps.

Si par ce type d'approche, il semble possible de trouver une ou plusieurs conformations communes, il serait intéressant de la (les) tester via une approche par "docking" avec les CDR dont on aurait déterminé la structure par modélisation ou par cristallographie.

Cette approche par regroupement en famille est intéressante pour la modification rationnelle des acides aminés des peptides mimétiques de la chaîne O de *Brucella*. C'est une méthode prometteuse pour la détermination de structure de peptides en complexes avec des anticorps.

MATERIEL ET METHODES

Matériel et méthodes

OUTILS EXPERIMENTAUX

1) Echantillons de protéines

a) Thiorédoxine-peptides

Le matériel ainsi que la méthode utilisée pour la fusion de la thiorédoxine avec les peptides sont détaillés dans les «rapport d'activités» de P. Mertens et X. De Bolle. La thiorédoxine a été dialysée et concentrée au Laboratoire d'Immunologie.

b) Fab

Tous les anticorps utilisés pour la sélection des peptides mimétiques étaient disponibles au laboratoire d'Immunologie. L'obtention des Fab a été réalisée via digestion à la papaïne par le Centre de bioexpérimentation Valbex (figure 50): le pont disulfure situé entre 2 cystéines des chaînes lourdes est coupé par cette enzyme.

2) Concentration et dialyse de la solution de protéines

Pour concentrer et dialyser l'échantillon d'anticorps, nous avons utilisé des tubes qui ont une membrane de dialyse en biseau au fond. Ils sont déposés dans un autre tube plus large de 15 ml qui sert à récolter le liquide qui s'écoule par la membrane. Cette membrane à un cut-off de 5000, ce qui est approprié pour la concentration de Fab (50000 kDa).

Nous avons choisi d'utiliser la centrifugation pour concentrer et dialyser nos solutions de protéines: les 2 opérations sont réalisées en même temps. De cette manière, nous avons concentré et dialysé 2 échantillons de l'anticorps 4F9, le premier de 8 ml et le second de 4 ml que nous avons dialysé contre un tampon. Le protocole est détaillé ci-dessous:

- Rinçage préalable de la membrane (pour enlever la couche de glycérol de protection): 4 ml de tampon sont centrifugés pendant 20 minutes à 4000 tour/minute.
- Prélèvement de la solution de protéines vers le tube de dialyse (lui-même déposé dans le tube de 15 ml)
- 20 minutes de centrifugation à 4000 tour /minute à 4°C
- Ajout de 3 ml de tampon PBS pH 6 $Na_2 HPO_4$ 1/15 M, $KH_2 PO_4$ 1/15 M
- Ces 2 dernières opérations sont répétées 3 fois en prélevant le "float through" après chaque centrifugation.
- Stockage de la solution concentrée et dialysée dans des eppendorfs (1,5 ml) à 4°C.

En finale, nous disposons de 1,5 ml pour le premier échantillon et de 1 ml pour le second.

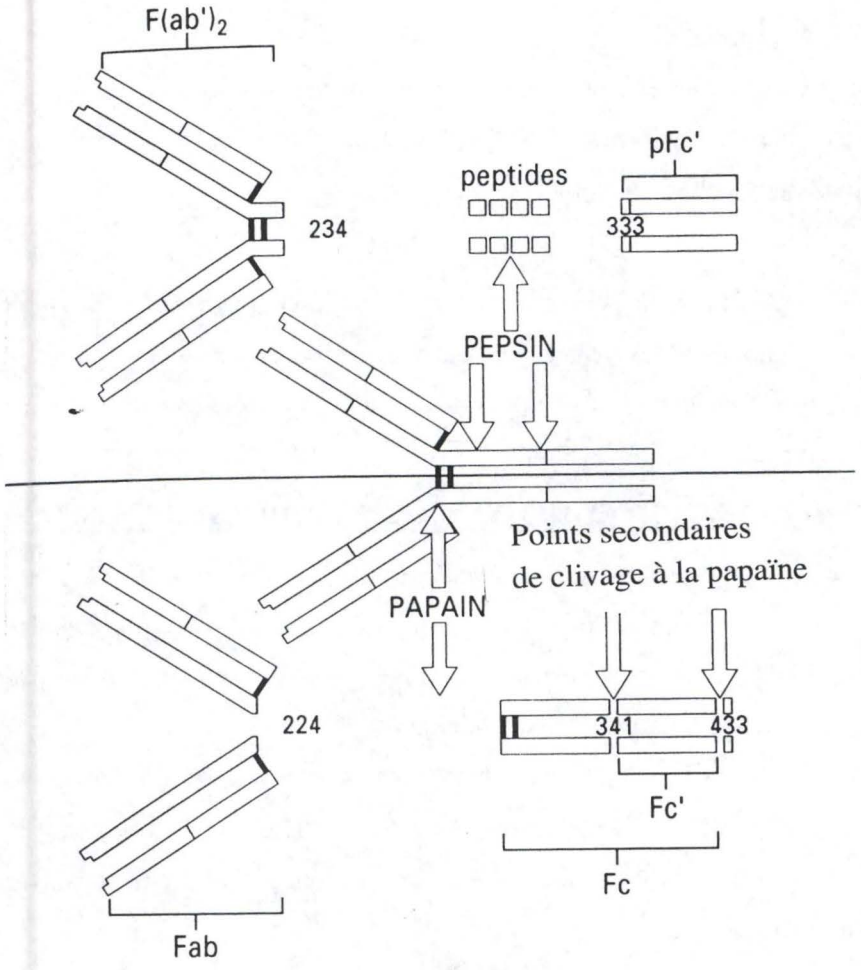


Figure 50: Schéma de la digestion d'un anticorps par la papaïne et la pepsine : la pepsine clive la chaîne lourde d'un IgG1 humain aux positions 234 et 333 pour obtenir le fragment $(Fab)_2$ et le fragment Fc. La papaïne coupe à la position 224, on obtient 2 fragments 2 fragments Fab et un fragment Fc.

Remarques

Le float through doit être conservé et testé, car si la membrane est défectueuse ou abîmée en cours de centrifugation, la protéine risque de s'y trouver.

Il est possible d'ajuster la durée de la dernière centrifugation afin d'obtenir un échantillon plus ou moins concentré. Il faut tout de même veiller à ne pas centrifuger trop longtemps, pour ne pas avoir la membrane "à sec" avec la protéine qui y est accolée. Dans ce cas, il est difficile de récupérer la protéine en bon état.

3) Test de pureté et de concentration: TEST BCA (Bicinchoninic Acid)

Ce test consiste à mesurer la concentration en protéines d'une solution. L'absorbance à 550 nm est proportionnelle à la concentration en protéines.

a) Matériel

- Boîtes 96 puits à fond plat
- BSA (Bovine Serum Albumin)1mg/ml
- Solutions A et B du test ("BCA Protein Assay Reagent A" et "BCA Protein Assay Reagent B")
- Echantillon de protéine dilué 5 fois et «Float through» de cet échantillon.
- Spectrophotomètre relié à un ordinateur (ici Macintosh)
- Programme d'analyse des boîtes 96 puits (Asoft)

b) Méthode

Chaque puit doit contenir 10 µl d'échantillon.

1) Dans les 16 premiers puits, diluer l'échantillon de BSA 1mg/ml 1, 2, 4, 6, 8, 10 et 20 fois plus un contrôle négatif (un puit sans BSA) pour arriver à ce volume. Cette opération est réalisée en duplicat. Dans les 32 puits suivants, faire la même opération avec l'échantillon de protéine diluée et avec le "float through".
2) Préparer le réactif BCA dans les proportions suivantes: 2 % de solution B et 98% de solution A.
3) Verser 200 µl de cette solution dans chaque puit (attention ce mélange ne peut être conservé plus de 30 minutes après préparation).
4) Laisser agir 30 minutes à 37°C.
5) Calcul de la concentration en protéines dans les différents puits par mesure de l'absorbance.
6) Elaboration du tableau des concentrations présentes dans chaque puits: l'ordinateur, à l'aide du programme Asoft, va mettre au point une droite d'étalonnage en fonction de l'absorbance de la BSA et va ensuite calculer la concentration de l'échantillon en fonction de cette droite.

CRYSTAL SCREEN REAGENT FORMULATION

TUBE NUMBER	*Salt*	TUBE NUMBER	*Buffer* †	TUBE NUMBER	*Precipitant*
1.	0.02 M Calcium chloride dihydrate	1.	0.1 M Na Acetate trihydrate pH 4.6	1.	30% v/v 2-methyl-2,4-pentanediol
2.	None	2.	None	2.	0.4 M K, Na Tartrate tetrahydrate
3.	None	3.	None	3.	0.4 M Ammonium dihydrogen phosphate
4.	None	4.	0.1 M Tris HCl pH 8.5	4.	2.0 M Ammonium sulfate
5.	0.2 M tri-sodium citrate dihydrate	5.	0.1 M Na Hepes pH 7.5	5.	30% v/v 2-methyl-2,4-pentanediol
6.	0.2 M Magnesium chloride hexahydrate	6.	0.1 M Tris HCl pH 8.5	6.	30% w/v PEG 4000
7.	None	7.	0.1 M Na Cacodylate pH 6.5	7.	1.4 M Sodium acetate trihydrate
8.	0.2 M tri-sodium citrate dihydrate	8.	0.1 M Na Cacodylate pH 6.5	8.	30% v/v 2-propanol
9.	0.2 M Ammonium acetate	9.	0.1 M tri-sodium Citrate dihydrate pH 5.6	9.	30% w/v PEG 4000
10.	0.2 M Ammonium acetate	10.	0.1 M Na Acetate trihydrate pH 4.6	10.	30% w/v PEG 4000
11.	None	11.	0.1 M Na Citrate pH 5.6	11.	1.0 Ammonium dihydrogen phosphate
12.	0.2 M Magnesium chloride hexahydrate	12.	0.1 M Na Hepes pH 7.5	12.	30% v/v 2-propanol
13.	0.2 M tri-sodium citrate dihydrate	13.	0.1 M Tris HCl pH 8.5	13.	30% v/v PEG 400
14.	0.2 M Calcium chloride dihydrate	14.	0.1 M Na Hepes pH 7.5	14.	28% v/v PEG 400
15.	0.2 M Ammonium sulfate	15.	0.1 M Na Cacodylate pH 6.5	15.	30% w/v PEG 8000
16.	None	16.	0.1 M Na Hepes pH 7.5	16.	1.5 M Lithium sulfate monohydrate
17.	0.2 M Lithium sulfate monohydrate	17.	0.1 M Tris HCl pH 8.5	17.	30% PEG 4000
18.	0.2 M Magnesium acetate tetrahydrate	18.	0.1 M Na Cacodylate pH 6.5	18.	20% PEG 8000
19.	0.2 M Ammonium acetate	19.	0.1 M Tris HCl pH 8.5	19.	30% v/v 2-propanol
20.	0.2 M Ammonium sulfate	20.	0.1 M Na Acetate trihydrate pH 4.6	20.	25% w/v PEG 4000
21.	0.2 M Magnesium acetate tetrahydrate	21.	0.1 M Na Cacodylate pH 6.5	21.	30% v/v 2-methyl-2,4-pentanediol
22.	0.2 M Sodium acetate trihydrate	22.	0.1 M Tris HCl pH 8.5	22.	30% w/v PEG 4000
23.	0.2 M Magnesium chloride hexahydrate	23.	0.1 M Na Hepes pH 7.5	23.	30% v/v PEG 400
24.	0.2 M Calcium chloride dihydrate	24.	0.1 M Na Acetate trihydrate pH 4.6	24.	20% v/v 2-propanol
25.	None	25.	0.1 M Imidazole pH 6.5	25.	1.0 M Sodium acetate trihydrate
26.	0.2 M Ammonium acetate	26.	0.1 M tri-sodium Citrate dihydrate pH 5.6	26.	30 % v/v 2-methyl-2,4-pentanediol
27.	0.2 M tri-sodium citrate dihydrate	27.	0.1 M Na Hepes pH 7.5	27.	20% v/v 2-propanol
28.	0.2 M Sodium acetate trihydrate	28.	0.1 M Na Cacodylate pH 6.5	28.	30% w/v PEG 8000
29.	None	29.	0.1 M Na Hepes pH 7.5	29.	0.8 M K, Na Tartrate tetrahydrate
30.	0.2 M Ammonium sulfate	30.	None	30.	30% w/v PEG 8000
31.	0.2 M Ammonium sulfate	31.	None	31.	30% w/v PEG 4000
32.	None	32.	None	32.	2.0 M Ammonium sulfate
33.	None	33.	None	33.	4.0 M Sodium formate
34.	None	34.	0.1 M Na Acetate trihydrate pH 4.6	34.	2.0 M Sodium formate
35.	None	35.	0.1 M Na Hepes pH 7.5	35.	0.8 M Na dihydrogen phosphate, 0.8 M K dihydrogen phosphate monohydrate
36.	None	36.	0.1 M Tris HCl pH 8.5	36.	8% w/v PEG 8000
37.	None	37.	0.1 M Na Acetate trihydrate pH 4.6	37.	8% w/v PEG 4000
38.	None	38.	0.1 M Na Hepes pH 7.5	38.	1.4 M tri-Sodium citrate dihydrate
39.	None	39.	0.1 M Na Hepes pH 7.5	39.	2% v/v PEG 400, 2.0 M Ammonium sulfate
40.	None	40.	0.1 M tri-sodium Citrate dihydrate pH 5.6	40.	20% v/v 2-propanol, 20% w/v PEG 4000
41.	None	41.	0.1 M Na Hepes pH 7.5	41.	10% v/v 2-propanol, 20% w/v PEG 4000
42.	0.05 M Potassium dihydrogen phosphate	42.	None	42.	20% w/v PEG 8000
43.	None	43.	None	43.	30% w/v PEG 1500
44.	None	44.	None	44.	0.2 M Magnesium formate
45.	0.2 M Zinc acetate dihydrate	45.	0.1 M Na Cacodylate pH 6.5	45.	18% w/v PEG 8000
46.	0.2 M Calcium acetate hydrate	46.	0.1 M Na Cacodylate pH 6.5	46.	18% w/v PEG 8000
47.	None	47.	0.1 M Na Acetate trihydrate pH 4.6	47.	2.0 M Ammonium sulfate
48.	None	48.	0.1 M Tris HCl pH 8.5	48.	2.0 M Ammonium dihydrogen phosphate
49.	1.0 M Lithium sulfate monohydrate	49.	None	49.	2% w/v PEG 8000
50.	0.5 M Lithium sulfate monohydrate	50.	None	50.	15% w/v PEG 8000

† Buffer pH is that of a 1.0 M stock prior to dilution with other reagent components. pH with HCl or NaOH.

PEG = polyethylene glycol.

Crystal Screen contains fifty unique reagents. To determine the formulation of each reagent, simply read across the page.

HAMPTON RESEARCH

25431 Cabot Road, Suite 205 Laguna Hills, CA 92653-5527 U.S.A.
Tel: (714) 699-1040 ◊ Fax: (714) 586-1453
E-mail: xtalrox@aol.com
WEB: http://www.hamptonresearch.com

Figure 51: « Crystal screen I » (Hampton screen)

4) Conservation et conditions d'utilisation

Tous les échantillons biologiques doivent être conservés à 4°C. Lors de l'utilisation, ils doivent être exposés le moins possible.

5) Elaboration de tampons (liqueur-mères)

a) Matériel

1. la liqueur-mère utilisée pour la cristallisation du lysozyme

Solution-stock de NaCl 3M

Solution-stock de lysozyme 50mg/ml dans un tampon acétate pH 4,5

Solution de tampon 50 mM acétate de sodium pH 4,5

2. Le criblage (screen)

Les screens sont des kits reprenant une série de conditions de cristallisation prédéfinies qui permettent de délimiter des conditions optimum pour les protéines étudiées. 2 types de screen commerciaux (Hampton) ont été utilisés:

Le screen "total" ("crystal screen I" *Hampton Research*) pour la cristallisation de l'anticorps 4F9 et de la thiorédoxine couplées à différents peptides

Le screen acétate pour la cristallisation du 4F9, c'est un screen où la présence d'acétate est une constante et où les autres constituants de la liqueur sont très différents (différents pH et variation de la concentration en $(NH_4)_2SO_4$)

Pour les screen "totaux", étant donné que nous disposions de 2 "linbro box" de 48 puits, 2 des 50 conditions n'ont pas été testées. Pour la thiorédoxine, il s'agit des conditions 49 et 50, et pour le 4F9, les conditions 1 et 2

3. les tampons

Nous avons dû refaire une série de liqueurs-mères identiques à celles des « screen. » Pour ce faire, nous bénéficions des réactifs du laboratoire d'Immunologie et de CMS (Chimie Moléculaire Structurale). Ces liqueurs-mères ont été réalisées sur base de protocoles suggérés dans la littérature (pour la thiorédoxine), sur base des données du screen "total" (voir figure 51) ou sur base de nouveaux protocoles proches des données de ce screen (voir «résultats», nous avons remplacé le PEG 400 utilisé dans une des conditions du screen par du PEG 1000).

b) Méthode

Pour la cristallisation du lysozyme, nous avons reproduit un protocole conseillé par A. Ducruix (Ducruix and Giegé , 1992), décrit dans la partie "application de la technique par gouttes suspendues".

Pour la cristallisation de la thiorédoxine, nous avons reproduit les conditions de cristallisations de la littérature décrites pour la thiorédoxine seule: 50 % v/v MPD Hac/Ac 0,01M pH 4,0 Cu^{++} 0,001 M.

Nous avons testé d'autres conditions en utilisant diverses concentrations en précipitant (MPD, 2-4 méthyl-pentanediol): 10-20-30-40% en MPD avec le même tampon et la même concentration en

cuivre. Par la suite, nous avons refait ces liqueurs-mères en remplaçant le Cu^{++} 0,001 M par du Cd^{++} en concentration saturante.

- préparation du tampon seul
- ajout du sel
- ajustement de la solution par ajout d'eau bidistillée
- ajout du précipitant
- filtration avec un filtre de 0,22 µ

Remarques

- Tous les réactifs doivent être purs!
- Il ne faut pas oublier de vérifier que le précipitant ajouté ne modifiera pas le pH de la solution.
- Pour la préparation d'autres types de tampons, il faut veiller à limiter la production de sels additionnels, ce qui modifierait les conditions de cristallisation (par exemple: lorsqu'il faut refaire le tampon 0,1 M Na cacodylate pH 6,5, pour mettre à pH, on préfèrera l'utilisation d'acide cacodylique plutôt que d'HCl pour éviter la production de NaCl)
- Une autre manière de garder la solution intacte est de l'autoclaver avant l'ajout de MPD (ou de n'importe quel précipitant)

6) Application de la technique par gouttes suspendues

a) Matériel

Les "linbro box" sont des boîtes en plastique normalement utilisées pour les cultures de cellules. Ces boîtes contiennent 24 puits agencés en 4 lignes et 6 colonnes.

Chaque puit a un volume approximatif de 2 ml et un diamètre interne de 16 mm. Un petit rebord sert à sceller le système. De petites lamelles de 18 à 22 µ sont appliquées au-dessus du puit avec la goutte en suspension à l'intérieur du puit. Ces lamelles ont une surface hydrophobe afin que la goutte ne s'étale pas et reste bien suspension. De l'huile de paraffine, une couche de graisse ou de silicone sont utilisées pour sceller le système. Le remplissage des puits et le dépôt des gouttes se fait à l'aide de micropipettes.

b) Méthode

- Dépôt d'une couche de silicone ou d'huile de paraffine sur le bord du puit
- Remplissage du réservoir (puit) avec 500 µl de liqueur-mère (ou plus selon la disponibilité en tampon)
- Mélange d'une goutte de 4 µl de la solution de protéine avec la même quantité de liqueur-mére (tirée du puit)
- Dépôt d'1 ou 2 goutte (s) de ce mélange sur une fine lamelle à surface hydrophobe (siliconée) de 18 ou 22 µ (selon le type de Linbro Box)
- Retournement de la lamelle sur le puit à l'aide d'une pince.

- Appui sur les bords de la lamelle et ajustement pour bien sceller le puit (si le puit n'est pas bien scellé la protéine se concentrera tout comme le réservoir, et il n'y aura pas d'équilibre atteint).
- Application d'une serviette de type Kleenex sur la boîte de façon à empêcher le contact entre le silicone ou l'huile et le couvercle lors de la fermeture).

Remarques

- Pour la cristallisation du lysozyme, nous avons utilisé l'entièreté de la boîte pour une série de tests différant par la concentration en NaCl (de 0,6 à 2,0 M) et en lysozyme (de 50 mg/ml à 12,5 mg/ml).
- L'avantage de l'utilisation de l'huile est importante pour les cristallisations à 4°C, l'huile ne se solidifie pas: on peut aisément enlever la lamelle et recueillir le cristal, tandis qu'en utilisant le silicone, il faut laisser la boîte à température ambiante (ce qui très mauvais, pour les protéines qui cristallisent à basse température).
- On préfèrera les gouttes de 3/3 µl (au lieu de 4/4) lorsque l'on ne dispose pas de grosses quantités de protéines.
- Dès que la lamelle est retirée du sommet du puits, l'équilibre est rompu, la cristallisation s'arrête ou se dirige vers un autre état d'équilibre si on redépose la lamelle sur le puit d'origine.

7) Observation des cristaux

Pour l'observation des cristaux, nous avons utilisé un binoculaire muni d'un lampe "chaude" avec un grossissement de 50 fois. La hauteur de l'objectif permet l'observation de boîtes épaisses (les "linbro box" ont 3 cm d'épaisseur).

Remarques

Il ne faut pas laisser la boîte trop longtemps exposée sous la lampe, la température (et de ce fait, la cristallogenèse) s'en trouve à chaque fois modifiée. Pour les mêmes raisons, il est important de ne pas faire trop d'observations rapprochées (il est souhaitable de faire des observations après quelques heures, 1 jour, 1semaine, et après 1 ou 2 mois).

8) Coloration, isolement des cristaux et tests de dureté

Le but de l'opération est de laisser diffuser un colorant à travers la maille du cristal de protéine. Si c'est un cristal de sel, il ne se colore pas.

Au préalable, il est conseillé de faire un test de dureté: le cristal est estimé comme étant très solide, solide, moyennement solide et fragile, d'après le type de choc qu'on lui fait subir. Les cristaux de sels sont en général très durs, et les cristaux de protéines sont plus fragiles.

a) Matériel

- Pince
- Fine tigette
- Bleu de Coomassie
- Cristaux de NaCl et de Lysozyme (test négatif et positif)

- Cristaux de la protéine d'intérêt (thiorédoxine couplée aux peptides 6B3-[X] et l'anticorps 4F9)
- Lame de verre
- Microscope optique

b) Méthode

- Oter la lamelle où se trouve le cristal, la retourner et la déposer sur une lame
- Observation au microscope optique
- Test de dureté:
    - A l'aide d'une tigette, donner quelques légers coups. Si le cristal cède, il est considéré comme «fragile».
    - Application de coups de plus en plus longs et de plus en plus appuyés pour estimer la dureté.
- Test de coloration:
    - Dépôt de quelques mg de bleu de Coomassie dans le réservoir et homogénéisation.
    - Application d'une goutte de ce mélange sur la lamelle à côté de la goutte d'intérêt.
    - Glissement du cristal dans la goutte de colorant
    - Laisser agir 24 heures
    - Remettre le cristal dans sa goutte d'origine et observer au microscope

Remarques

Les cristaux de protéines sont en général moins solides (sauf exception type «lysozyme») et bien colorés. On peut utiliser n'importe quel colorant à la place du Bleu de Coomassie.

9) Diffraction

Comme le diffractomètre disponible aux Facultés n'est utilisable que pour la diffraction de petites molécules, nous avons employé celui de l'Université de Liège qui est du type bidimensionnel. Ce type de détecteur permet d'analyser rapidement les données provenant de la diffraction d'un cristal de protéines. Il permet la diffraction par des cristaux de protéines à travers un capillaire contenant une goutte de liqueur-mère.

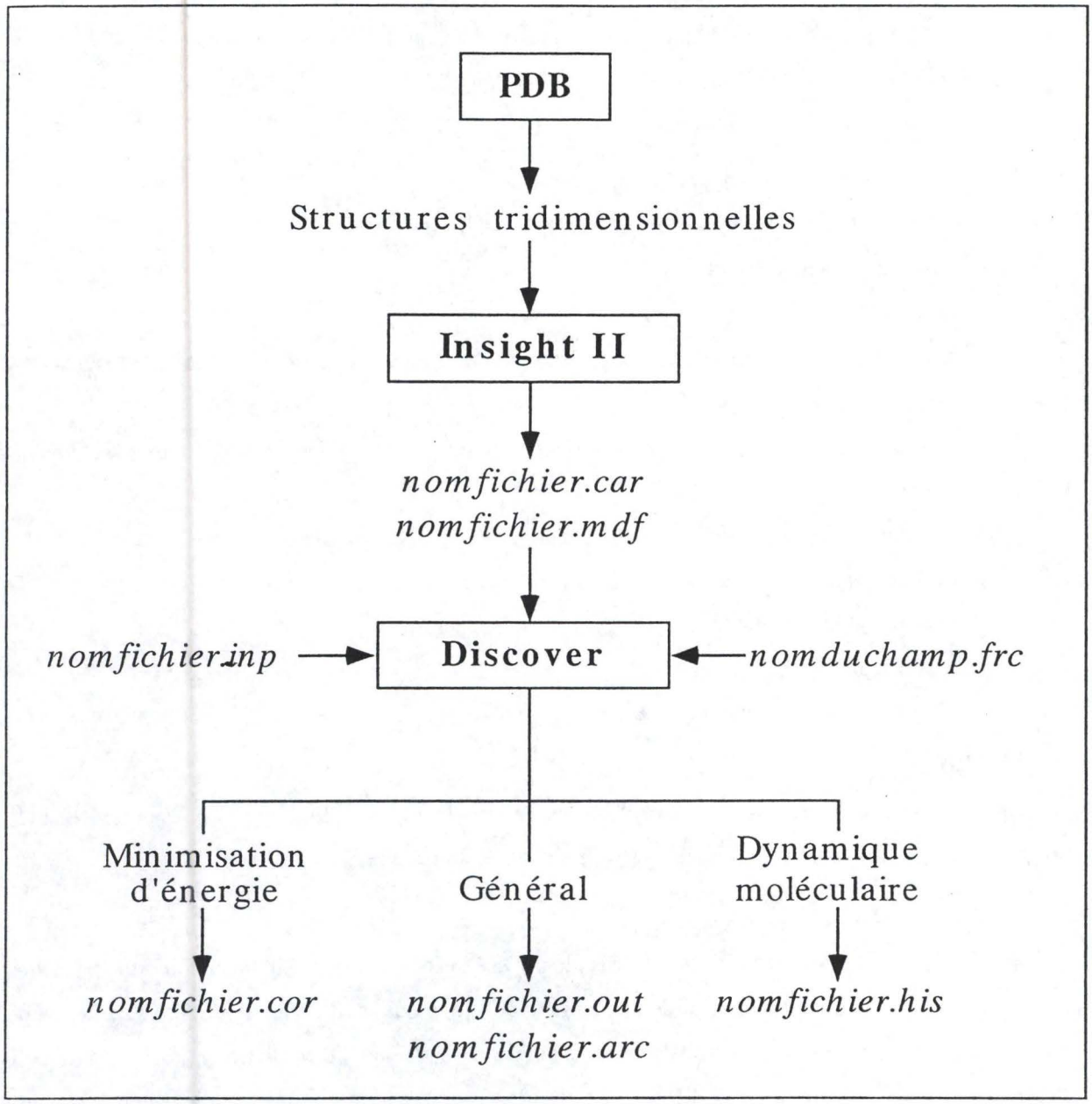


Figure 52: Schéma représentant la stratégie utilisée lors d'une modélisation par dynamique moléculaire.

OUTILS THEORIQUES

1) Outil graphique et programmes de modélisation moléculaire

a) Console graphique

Le support informatique que nous avons utilisé est une station de travail *Silicon Graphics Octane duo* fonctionnant sur le système d'exploitation IRIX 6,5. Cette station nous permet d'appliquer nos protocoles de mécanique et dynamique en des temps relativement court (entre 30 minutes et 1 heure en temps CPU). Cette console nous permet également de visualiser les peptides et de modifier rapidement les gros fichiers (.arc) nécessaires à la modélisation.

b) MSI (Molecular Simulated Incorporating)

Nous disposons au laboratoire d'une série de programmes mis au point par MSI. L'avantage de ce type d'outil est qu'il ne faut pas utiliser de programmes sélectionnés dans plusieurs filières pour mener à bien une étude: toutes les étapes sont réalisables à partir des programmes de la filière "MSI".

Au cours de ce travail, nous avons utilisé le programme "discover" pour la modélisation des peptides et le logiciel *insightII 97.5* pour la visualisation, l'analyse et la modification des peptides modélisés. Le logiciel *insightII* comporte une série de sous-programmes pour des applications bien précises: par exemple la partie "homology" permet une analyse de séquence et de la structure des molécules étudiées.

c) Modélisation de peptides

I) "Construction" du peptide

La première étape est la construction du peptide par simple addition des acides aminés du peptide sur base de la séquence grâce au programme "biopolymer" (*insightII*). Il suffit de nommer la molécule et de sélectionner un à un les résidus dans l'ordre dans lequel il se trouve en commençant par l'extrémité N-terminale(figure 52).

Pour contraindre les peptides à l'aide d'un pont disulfure, il faut préalablement imposer la liaison entre les soufres ("modify bond").

Avant la modélisation proprement dite, il est indispensable de minimiser grossièrement le peptide ("optimize") afin de l'optimaliser avant la dynamique en éliminant les contacts défavorables ou les géométries absurdes (par exemple au niveau du lien covalent S-S).

Après ces différentes étapes, 2 types de fichiers sont écrits. Il s'agit des fichiers ".car" et ".mdf". Le fichier ".car" reprend les coordonnées X, Y, Z de tous les atomes ainsi que l'énergie de la molécule. Le fichier ".mdf" reprend les énergies attribuées aux différents atomes d'après les paramètres du champ de

force utilisé. Ces énergies sont fonction du type d'atome. Le champ de force dont nous nous sommes servi est CVFF, étant donné qu'il a été testé à de nombreuses reprises et que les paramètres utilisés semblent adaptés à la description des peptides étudiés:

CVFF (Consistent Valence Force Field)

On peut retrouver au sein de l'équation de champ de force plusieurs paramètres; les termes 1 à 4 représentent les interactions covalentes et caractérisent les déformations des liaisons, les angles de valence, les angles de torsion, et les interactions "hors-plan". Les termes mixtes (de 5 à 9) rendent compte du couplage des déformations des coordonnées internes (angles de valence, ...). Les termes 10 et 11 représentent les interactions non-covalentes (van der Waals et interactions électrostatiques) et sont appelés dans leur ensemble "énergie d'interaction".

Simplifications:

Diverses approximations se révèlent indispensables pour diminuer le nombre de calculs (ils augmentent exponentiellement avec le nombre d'atomes):

a) Fonction de Morse

Cette fonction est normalement utilisée pour représenter les liaisons covalentes, comme elle est très fidèle à la réalité. Cependant, elle peut engendrer des résultats irréalistes si la structure est mal définie (par exemple en début de minimisation). Elle a donc été remplacée par une fonction de type harmonique (plus rapide).

b) Cut-off:

En général, à partir d'une certaine distance entre atomes, on considère que leurs interactions sont négligeables. *Discover* utilise une fonction qui annule progressivement les interactions à longues distances, pour éviter les discontinuités au niveau de l'énergie via les modérateurs "cutdis" et "swtdis" ou "cutoff".

c) Atomes "wild card":

Tous les paramètres ne sont pas importants pour chaque atome: par exemple, les paramètres des termes de torsions ne sont pas affectés par les atomes terminaux, il est possible de ne pas en tenir compte.

d) Constante diélectrique:

Le terme d'énergie électrostatique est un des plus délicats à évaluer: il dépend de la constante diélectrique et de l'exactitude des charges partielles attribuées aux différents atomes. Le problème principal est que la constante diélectrique change fortement d'une partie de la protéine à l'autre (site actif: 4, dans l'eau: 80), or pratiquement aucun logiciel ne propose l'utilisation de gradient. Une constante diélectrique dépendante de la distance est utilisée, sans tenir compte de la présence d'eau (qui augmenterait substantiellement le temps de calcul).

II) Mécanique et dynamique moléculaire

Pour minimiser la géométrie d'une molécule, on dispose de plusieurs types d'algorithmes dépendant de la stabilité de la molécule de départ et de la stabilité que l'on désire obtenir. Les méthodes "steepest descent" et "conjugate gradient" sont les 2 principaux algorithmes que nous avons utilisés. Ces 2 algorithmes se suivent dans le protocole que nous avons établi: la conformation utilisée pour la dynamique sera minimisée en fonction de la valeur de dérivée choisie ainsi que des paramètres "cutdis" et "swtdis".

Avant de trouver le protocole adapté aux peptides contraints, nous avons testé une série de protocoles disponibles. Le principal problème rencontré était une valeur trop élevée de la dérivée du système (conformation trop énergétique). Pour pallier ce problème nous avons successivement modifié la valeur de dérivée maximale ("DEMAX" de 10^4 à 10^8) ainsi que les paramètres de minimisation (cd 1, 0.1 ou 0.01; sd 0.5, 0.05 ou 0.005; ff 0.1, 0.01 ou 0.001) afin de réaliser la dynamique sur une conformation plus stable. Aucune de ces modifications n'a permis d'adapter correctement le protocole de départ. Nous avons ensuite remanié ce protocole sur base d'informations sur l'ANF (voir "résultats et discussion"): nous avons modifié le nombre d'itérations des étapes de la dynamique (1 itération = 1 femtoseconde). Finalement, nous avons opté pour un protocole "MSI" que nous avons également modifié. Ce protocole est décrit dans la figure 53.

Remarques

1° les phrases reprises après les "!" ne sont pas prises en compte pour la modélisation, ce sont des commentaires pour les utilisateurs (la plupart sont très importants).

2° une série de paramètres sont des simplifications des champ de force, déjà décrites: "overlap", "cutoff", constante diélectrique ("set dielectric")

3° les paramètres de convergence "cutdis" et "swtdis" dépendent du "cutoff"

4° ce protocole est utilisé sous forme de fichier ".inp" (il est avec les fichiers ".car", ".mdf" le troisième fichier indispensable à l'optimisation de géométrie)

5° on peut délimiter le protocole en plusieurs parties:

1° paramètrage

2° minimisations préalables (au nombre de 2)

3° initialisation de la dynamique (équilibration)

4° dynamique proprement dite (génération des 20 conformères)

5° minimisations terminales (au nombre de 2)

6 ° In fine, on se retrouve avec une série de fichiers:

Les fichiers ".cor" reprennent les nouvelles coordonnées reprises lors des dynamiques ;

Les fichiers ".out" reprennent l'ensemble des données de la dynamique ;

Figure 53: Protocole de dynamique moléculaire et minimisation d'énergie que nous avons utilisé pour la modélisation des peptides mimétiques de la chaîne O du LPS de *Brucella*

```
!   INPUT FILE FOR DISCOVER GENERATED BY INSIGHT
! 1
! Quenched Molecular Dynamics Protocol
!
     overlap = 0.01
! Change the next line to the cutoff value that you want to use:
     cutoff = 20.000000
! cutdis = cutoff - 1
     cutdis = 19.000000
! swtdis is always 1.5
     swtdis = 1.5
! timtmp is the exponential rate constant during
! the "Initialize dynamics" and "resume dynamics" steps.
! Default timtmp = 0.1
! Normally, this is NOT specified for straight Molecular Dynamics,
! since we only change the temperature at the beginning, and we
! usually want to do this fairly quickly.
! Uncomment and edit the next line if you want to change this.
! timtmp = 0.1
     begin simulation
     *    add-automatic bond torsion valence out-of-plane
     reduce
!
! Change the dielectric constant to the value you want to use.
! distance-dependent dielectrics:  set dielectric = 1.0*r
     set dielectric  = 1.000000 * r
!
!              Bias the peptide bonds to trans
! !!! define the residues !!!!!!!!
     force omega to 180 in molecule 1 residue ALAn 1 to ALAN 20
     *     using 5.0 kcal/rad**2
!
! 2
! First minimization to reduce very high energies.
! If you want charges, delete the "and no charges" line.
     Minimize
     *    no cross terms
     *    no morse
     *    and no charges
     *    for 100 iterations
     *    using steep descents
     *    until the maximum derivative is less than 10.000000001 kcal/A
```

```
!
!
! Second minimization to drive to a low minimum.
 We don't need to minimize REALLY low, since we are about to
! heat up the molecule.
! If you want charges, delete the "and no charges" line.
! If you have > 200 non-fixed atoms, change "VA09A" to
! "conjugate gradient".
     Minimize
    *    no cross terms
    *    no morse
    *    and no charges
    *    for 5000 iterations
    *    using  conjugate gradient
    *    until the maximum derivative is less than 1.000000001 kcal/A
!
!3
! The "for ### iterations" line specifies the number of
! fsec for heating and equilibration.
! Change the temperature to suit your molecule.  If
! your molecule is restrained (e.g., rings,
! double bonds, etc), then 1000-1500K is usually good.
! If your molecule is unconstrained, 600-1000K is good.
! Steps of 1 fsec are required.
! If you want charges, delete the "and no charges" line.
! The "write history" line reduces the history file output, which is
! highly recommended unless you want to restart the calculation if it fails.

     initialize dynamics
    *    write history file every 9000000 steps
    *    for 5000 iterations
    *    at 900.000 K
    *    steps of    1.000
    *    no cross terms
    *    no morse
    *    and no charges
!
!4
!    Dynamics Loop
!
     ifile = 1
!
!
! The next line specifies how many fsec to run the dynamics
! before saving the conformation.
! If you want charges, delete the "and no charges" line.
100   resume dynamics
    *    for 5000 iterations
    *    no cross terms
    *    no morse
    *    and no charges
!
!
!
!
!    Record Results
!
     archive as file number ifile
```

```
    ifile = ifile + 1
! Change the number below (currently 1000) to the number
! of structures you want to calculate
    if ifile .le. 20 then 100
!
!
! The next loop retrieves each structure, minimizes it, and stores
! the results
    ifile = 1
200  retrieve file number ifile
!5
! First minimization to reduce very high energies.
! If you want charges, delete the "and no charges" line.
    Minimize
    * no cross terms
    * no morse
    * and no charges
    * for 500 iterations
    * using steep descents
    * until the maximum derivative is less than 10.0 kcal/A
!
!
! Second minimization to drive to a low minimum.
! Adjust the "maximum derivative is less than ###"
! number to suit the level of minimization you require.
! If you want charges, delete the "and no charges" line.
    Minimize
    * no cross terms
    * no morse
    * and no charges
    * for 5000 iterations
    * using conjugate gradient
    * until the maximum derivative is less than 0.001 kcal/A
!
!
! Record Results
!
    print energy summary
    print torsions
    archive as file number ifile
    ifile = ifile + 1
! Change the number below (currently 1000) to the number
! of structures you want to minimize (this should match the number
! of structures that were initially archived).
    if ifile .le. 20 then 200
    end
```

Les fichiers ".his" et les fichiers ".arc" reprennent les coordonnées ainsi que l'énergie des conformères (toujours calculée d'après le champ de force "CVFF"). Ce sont ces derniers que nous utilisons et modifions lors de l'étude des peptides sur *insight*.

d) Utilisation d'*insightII* pour l'étude de conformères issus d'une dynamique moléculaire

La première étude à réaliser est la comparaison des énergies des conformères grâce au programme "analysis". Pour ce faire, il ne faut disposer que du fichier ".arc". Il faut renseigner *insightII* sur les fichiers à utiliser ("get trajectory") et ensuite lui demander de construire un graphe d'après ces données ("construct graph").Nous avons utilisé des graphiques en 2 dimensions, basés sur le numéro du conformère et sur l'énergie totale (X = "frame", Y = "energy total").

Le "cluster graph" est réalisable par un processus semblable ("construct graph"): il est possible de comparer les conformères sur l'ensemble des atomes, uniquement sur la chaîne carbonée ou sur des atomes choisis préalablement.

Pour grouper les conformères en familles, nous avons utilisé la macro "family" qui regroupe les conformères sur base du graphe des énergies et des "clusters graphs" en fonction d'un RMS choisi par l'utilisateur. Le principe est le suivant: le programme repère le conformère le plus stable en partant de la fin du fichier et trouve les conformères proches (pour lesquels le RMS est inférieur à une valeur fixée). Le programme fait la même chose avec le deuxième conformère en terme de stabilité et ainsi de suite. De cette manière, il n'y a pas de redondances: un conformère se trouve dans une famille uniquement! Afin de montrer que ce n'est pas l'ordre dans lequel on forme les familles qui est important pour établir leur nombre, nous avons modifié les fichiers ".arc" (voir point suivant). Dans la plupart des cas, nous avons donné à chaque conformère soit la même énergie, soit des énergies fixées *à priori* (0, 1, 2, 3, 4, 5 kcal/mole).

e) Superposition des peptides: modification des fichiers «.arc»

Pour pouvoir superposer un ensemble de conformères ou partie de conformères (par-exemple leurs chaînes carbonées), ils doivent faire partie du même fichier ".arc". Or, dans un fichier ".arc", les différentes molécules (ici les conformères) doivent avoir nécessairement le même nombre d'atomes. Il est donc nécessaire de modifier les fichiers ".arc" des différents peptides (par-exemple les 6B3-1,2,4,6) pour les grouper (puisqu'ils ne comportent pas le même nombre d'atomes) et de cette manière pouvoir les comparer. En effet, il suffit de ne sélectionner que la chaîne carbonée de tous les peptides et de les mettre dans le même fichier. Pour ce faire, il faut suivre les étapes suivantes:

par la commande "dirtoseq", disposer du fichier ".arcs" (puisque le fichier ".arc" n'est pas directement lisible).

modification par le programme (script) "jot": sélection des carbones par la commande "grep" sous *Unix* et copier-coller vers un fichier "jot", en n'oubliant pas de coller les valeurs d'énergies et les données pour la lecture des fichiers dans *insightII*. Dans le cas où on désire supprimer certains atomes (peptides courts en pVIII ou en thiorédoxine), ils suffit de le faire directement dans le fichier "jot".
retransformation du fichier "arcs" en "arc" par la commande "seqtodir" (on obtient un fichier ".arcsd" qu'il suffit de renommer en ".arc").

2) Analyse des angles phi-psi

Les angles PHI-PSI que nous utilisons pour comparer les flexibilités d'un ensemble de conformères sont des angles de torsion: ils sont disponibles au départ de nos peptides grâce à l'outil *insightII* (programme "homology" et "prostat"). Dès que l'on dispose des valeurs des angles PHI-PSI, il suffit de les enregistrer et de les transférer sous format *Excel*. Les données de tous les angles sont utilisées pour établir des plots de Ramachandran ou pour un calcul des écarts-types.

3) Recherche de motif dans les banques de données: outil "BLAST"

BLAST (Basic Local Alignement Search Tool) est l'algorithme heuristique utilisé par les programmes blastn, blastp, blastx, tblastn et tblastx. Ces programmes ont été conçu pour les recherches de similarité de séquence. Ces 5 programmes ainsi qu'une série de données concernant les paramètres, la stratégie et le type de matrices utilisées sont disponibles dans le manuel en ligne d'aide à l'utilisateur («BLAST HELP MANUAL»).

BIBLIOGRAPHIE

Alton, G.G., Jones, L.M., Angus, R.D. and Verger, J.M. (1988) *Techniques for the brucellosis laboratory*. Institut National de la Recherche Agronomique, Paris.

Altshul, S.F. (1990) Basic Alignement Search Tool. *J Mol Biol*, **215**, 403-410.

Belzer, C.A., Tabatabai, L.B. and Deyoe, B.L. (1991) Differentiation by western blotting of immune responses of cattle vaccinated with *Brucella abortus* strain 19 or infected experimentally or naturally with virulent *Brucella abortus*. *Vet.Microbiol.*, **27**, 79-90.

Bernstein, F.C. (1977) The protein Data Bank. A computer-based archival file for macromolecular structures. *J. Mol. Biol.*, **112 (3)**, 535-542.

Blasco, J.M. and Diaz, R. (1993) *Brucella melitensis* Rev1 vaccine as a cause of human brucellosis. *Lancet*, **342**, 805-800.

Branden, C. and Tooze , J. (1991) *Introduction to protein structure*.

Bundle, D.R., Cherwonogrodzky, J.W., Gidney, M.A.J., Meikle, P.J., Perry, M.B. and Peters, T. (1989) Definition of Brucella A et M epitopes by monoclonal typing reagents and synthetic oligosaccharides. *Infect.Immun.*, **57**, 2829-2836.

Cesareni, G., Minenkova, O., Dente, L., Iannolo, G., Zucconi, A., Helmer Citterich, M., Lanfrancotti, A., Castagnoli, L. and Vetriani, C. (1995) Structural and functional constraints in the display of peptides on filamentous phage capsids. In Cortese, R. (ed.) *Combinatorial Libraries. Synthesis, screening and application potential*. de Gruyter, Berlin; New York, pp. 113-126.

Chothia, C. (1989) Conformations of immunoglobin hypervariable regions. *Nature*, **343**, 877-883.

Colman, P.M. (1987) Three-dimensionnal structure of a complex of antibody with influenzae virus neuraminidase. *Nature*, **326**, 358-363.

Corbel, M.J. (1997) Brucellosis an overview. *Emerging Infectious Diseases*, **3**, 213-221.

Davies, D.R. and Padlan, E.A. (1990) Antigen-antibody complex. *Annual review of biochemistry*, **59**, 439-473.

De Bolle, X. (1999) Thorédoxine-Carrier. .

De Bolle, X. and Mertens, P. (1997-99) "Rapport d'activités: Peptidomimétiques et glycomimétiques: applications diagnostiques et vaccinales". .

de Fays, K., Tibor, A., Lambert, C., Vinals, C., Denoel, P., Debolle, X. and Wouters, J. (1999) Structure and function prediction of the Brucella abortus P39 protein by comparative modeling with marginal sequence similarities. *Protein Engineering*, **12(3)**, 217-223.

Ducruix, A. and Giegé , R. (1992) *Crystallisation of nucleic acids and proteins.*

Duée, E. and Dideberg, O. (1997) La cristallographie macromoléculaire en mouvement. *Biofutur: le technoscope*, Vol. cahier n°98.

Ebel, J.-P. (1999) La RMN des macromolécules. *Biofutur "le technoscope"*, **189**, 1-15.

Felici, F., Castagnoli, L., Musacchio, A., Jappeli, R. and Cesareni, G. (1991) Selection of antibody ligands from a large library of oligopeptides expressed on a multivalent exposition vector. *J.Mol.Biol.*, **222**, 301-310.

Gallop, M.A., Barrett, R.W., Dower, W.J., Fodor, S.P.A. and Gordon, E.M. (1994) Applications of combinatorial technologies to drug discovery. 1. Background and peptide combinatorial libraries. *J.Med.Chem.*, **37**, 1233-1251.

Gilliland and Bickham. (1990) Biological macromolecule crystallization database. *Methods in Enzymology*, **1**, 6-11.

Grimes, J.M. (1997) An atomic model of the outer layer of the blue tongue virus core derived from X-ray cristallography and electron microscopy. *Structure*, **5(7)**, 885-893.

Hill, H.R. and Stockey, P.G. (1996) Phage presentation. *Molecular Microbiology*, **20**, 685-692.

Holmgren, A. and Sodeberg, B.-O. (1970) Cristallization and preliminary crystallographic data for thioredoxin from *Escherichia coli* B. *Journal of molecular biology*, **54**, 387-390.

Kay, B.K. and Hoess, R.H. (1996) Principles and applications of phage display. In Kay, B.K., Winter, J. and MacCafferty, J. (eds.), *Phage display of peptides and proteins.* Academic Press, San Diego, pp. 21-34.

Landau, E.M. and Rosenbush, J.P. (1996) Lipidic cubic phase : a novel concept for the crystallisation of the membrane proteins. *Proc Natl Acad Sciences USA*, **93(25)**, 14532-14535.

Luger, K. (1997) Charactrerization of nucleosomes core particles containing histone proteins made in bacteria. *J Mol Biol*, **272(3)**, 301-311.

McLafferty, M.A., Kent, R.B., Ladner, R.C. and Markland, W. (1993) M13 bacteriophage displayind disulfide-constrained microproteins. *Gene*, **128**, 29-36.

Model, P. and Russel, M. (1988) Filamentous bacteriophage. In Calendar, R. (ed.) *The Bacteriophages*. Plenum, NeW York, pp. 375-456.

MSI. (1996) *Discover user guide.*

Opella, S.J., Stewart, P.L. and Valentine, K.G. (1987) Protein structure by solid-state NMR spectroscopy. *Q.Rev.Biophys.*, **19**, 7-49.

Palmer, D.A. and Douglas, J.T. (1989) Analysis of *Brucella* lipopolysaccharide with specific and cross-reacting monoclonal antibodies. *J.Clin.Microbiol.*, **27**, 2331-2337.

Pares, S., Mouz, N., Petillot, Y., Hackenbeck, R. and O, D. (1996) X-ray structure of Streptococcus pneumoniae PBP2X, a primary penicillin target enzyme. *Nat. Struct. Biology*, **3(3)**, 284-289.

Rick, P.D. (1987) Lipopolysaccharide biosynthesis. In Neidhart, F.C., Ingraham, J.L., Low, K.B., Magasanik, B., Schaechter, M. and Umbarger, H.D. (eds.), *Escherichia coli and Salmonella typhimurium. Cellular and molecular biology*. American Society for Microbiology, Washington, pp. 648-662.

Smith, G.P. (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science*, **228**, 1315-1316.

Smith, G.P. and Scott, J.K. (1993) Libraries of peptides and proteins displayed on filamentous phage. In s (ed.) *Methods in Enzymology Vol 217*. Academic Press, New York, pp. 228-257.

Souriau, C., Hua, T., Lefranc, M.-P. and Weill, M. (1998) Présentation à la surface des phages filamenteux : les multiples applications du phage display. *m/s*, **14**, 300-309.

Trakhanakov, S., Kreimer, D.I., Parkin, S., Ferro-Luzzi Ames, G. and Rupp, B. (1998) Cadmium-induced crystallisation of proteins:

II.Crystallisation of the salmonella typhimurium histidine-binding protein in complex with l-histidine, l-arginine, or l-lysine. *Protein Science*, **7**, 600-604.

Vinals, C. (1997) *Structure et conformation des macromolécules*.

Walgraffe, D. (1998) Sélection de mimotopes peptidiques de la chaîne O du lipopolysaccharide de *Brucella sp. Immunologie*. F.U.N.D.P., Namur, p. 78.

Webster, R.E. (1996) Biology of the filamentous bacteriophage. In Kay, B.K., Winter, J. and MacCafferty, J. (eds.), *Phage display of peptides and proteins*. Academic Press, San Diego, pp. 1-20.

Weiss, T.J. (1997) REVIEW: Biocristallography, structure determination and beyond. *Biotechnol. Appl. Biochem.*, **26**, 127-142.

Weynants, V., Gilson, D., Cloeckaert, A., Denoel, P.A., Tibor, A., Thiange, P., Limet, J.N. and Letesson, J.-J. (1996) Characterization of a monoclonal antibody specific for *Brucella* smooth lipopolysaccharide and development of a competitive enzyme-linked

immunosorbent assay to improve the serological diagnosis of Brucellosis. *Clin.Diagn.Lab.Immunol.*, **3**, 309-314.

Wouters, J. (1998) *Conception de médicaments*.