

50373
1987
323

50376
1987
323

THÈSE

présentée à

L'UNIVERSITÉ DES SCIENCES ET TECHNIQUES DE LILLE FLANDRÈS ARTOIS

pour l'obtention du titre de

DOCTEUR EN BIOCHIMIE

par

MARLENE IMBERT-CHOLLET

LA PROTEINE CHROMOSOMALE MC1 ISOLÉE D'ARCHAEBACTERIES DU GENRE

METHANOSARCINA :

ETUDES STRUCTURALES, LOCALISATION ULTRASTRUCTURALE ET

ÉTUDE DU COMPLEXE DÉSŒXYRIBONUCLÉOPROTÉIQUE.

présentée le 15 décembre 1987 devant la Commission d'Examen

Président : Professeur J. MONTREUIL

Rapporteurs : Docteur J. P. MORNON
Professeur J. KREMBEL

Examineurs : Docteur G. ALBAGNAC
Docteur P. SAUTIERE
Docteur B. LAINE

A Roxane,

A Thierry,

A mes beaux-parents

*"Qu'une chose soit difficile à faire
est une raison de plus pour l'entreprendre".....*

TABLE DES MATIÈRES

	Pages
INTRODUCTION	1
 <u>CHAPITRE I</u>	
GÉNÉRALITÉS	4
I - 1 LES ARCHAEBACTERIES	5
I - 1-1 Description phénotypique	6
I - 1-2 Principales caractéristiques	8
I - 1-3 Phylogénie	12
I - 1-4 Les Méthanogènes	18
I - 1-4-1 Principales caractéristiques	18
I - 1-4-2 Taxonomie	19
I - 2 ORGANISATION DE L'ADN	21
I - 2-1 Organisation de l'ADN dans la cellule eucaryote	21
I - 2-2 Organisation de l'ADN chez les eubactéries	23
I - 2-2-1 Propriétés de la protéine H1	24
I - 2-2-2 Propriétés de la protéine II	24
I - 2-3 Organisation de l'ADN chez les archaeobactéries	27
BIBLIOGRAPHIE	31
 <u>CHAPITRE II</u>	
PRÉPARATION, PURIFICATION ET PROPRIÉTÉS PHYSICO-CHIMIQUES DE LA PROTÉINE MC1	36
II - 1 CARACTERISTIQUES GÉNÉRALES DES BACTÉRIES MÉTHANOGENES : <i>M. BARKERI</i> ET <i>METHANOSARCINA</i> sp MST-A1	37

II - 2	PREPARATION DE LA PROTEINE MC1	40
II - 2-1	Principe de la méthode	40
II - 2-2	Résultats	40
II - 2-2-1	Préparation de la protéine MC1 de <i>Methanosarcina barkeri</i> souche MS	40
II - 2-2-2	Préparation de la protéine MC1 de <i>Methanosarcina sp</i> souche MST-A1	42
II - 3	PURIFICATION DE LA PROTEINE MC1	42
II - 3-1	Purification en milieu dénaturant	44
II - 3-2	Purification en milieu non dénaturant	46
II - 4	PROPRIETES PHYSICO-CHIMIQUES DE LA PROTEINE MC1	47
II - 4-1	Solubilité	47
II - 4-2	Comportement électrophorétique	50
II - 4-3	Détermination du point isoélectrique (pHi) de la protéine MC1 de <i>M.barkeri</i> MS	52
II - 4-4	Comportement à la lyophilisation	52
II - 4-5	Structure quaternaire de la protéine MC1	52

CHAPITRE III

DÉTERMINATION DE LA STRUCTURE PRIMAIRE DE LA PROTÉINE MC1 DE <i>METHANOSARCINA BARKERI</i> MS ET DE <i>METHANOSARCINA SP</i> MST-A1		57
PUBLICATION N°1		59
PUBLICATION N°2		66
III - 1	ETUDE DE LA PROTEINE MC1 DE <i>METHANOSARCINA BARKERI</i> MS	72
III - 1-1	Stratégie	72
III - 1-2	Résultats	74
III - 1-2-1	Contribution personnelle	74
III - 1-2-2	Séquence complète	75

III - 2	ETUDE DE LA PROTEINE MC1 DE <i>METHANOSARCINA</i> sp SOUCHE MST-A1	75
III - 2-1	Stratégie	75
III - 2-2	Matériel et méthodes	78
III - 2-3	Résultats	79
III - 2-4	Discussion	85

CHAPITRE IV

	CONFORMATION DE LA PROTÉINE MC1 DE <i>METHANOSARCINA BARKERI</i>	88
IV - 1	INTRODUCTION	89
IV - 2	METHODES DE PREDICTION DE STRUCTURE SECONDAIRE	90
IV - 2-1	Principe des méthodes de prédiction utilisées	91
IV - 2-1-1	Méthode de CHOU et FASMAN	91
IV - 2-1-2	Méthode de GARNIER ET coll.	91
IV - 2-1-3	Méthode de CID et coll.	92
IV - 2-1-4	Méthode de SETTE et coll.	92
IV - 2-1-5	Méthode de GABORIAUD et coll.	93
IV - 2-2	Application des méthodes de prédiction de structure secondaire à la protéine MC1 de <i>Methanosarcina barkeri</i>	95
IV - 3	METHODES SPECTROSCOPIQUES	100
IV - 3-1	Dichroïsme circulaire	100
IV - 3-2	Spectroscopie infrarouge	104
IV - 3-3	Fluorescence	107
IV - 4	COMPARAISON DE LA PROTEINE MC1 AVEC LA PROTEINE II DE DIFFERENTS MICRO-ORGANISMES ET AVEC L'HISTONE H4 DE THYMUS DE VEAU, A L'AIDE DE LA METHODE HCA	114
IV - 5	DISCUSSION	119

CHAPITRE V

ÉTUDE ULTRASTRUCTURALE DU COMPLEXE DÉSOXYRIBONUCLEOPROTEIQUE DE L'ARCHAEBACTERIE <i>Methanosarcina barkeri</i> .	122
INTRODUCTION	123
PUBLICATION N° 3	124
DISCUSSION	143
CONCLUSIONS ET PERSPECTIVES	144
APPENDICE TECHNIQUE	147
I - Souches bactériennes et culture des bactéries	148
II - Préparation de la protéine MCl	148
III - Concentration de la protéine MCl	151
IV - Dosage de l'ADN et des protéines dans le complexe désoxyribonucléoprotéique isolé de <i>M. barkeri</i> .	152
V - Isoélectrofocalisation	155
RÉFÉRENCES BIBLIOGRAPHIQUES	157

PUBLICATIONS

PUBLICATION N° 1

Primary structure of the chromosomal protein HMb from the archaeobacterium *Methanosarcina barkeri*.

LAINE, B., CHARTIER, F., IMBERT, M., LEWIS, R. and SAUTIERE, P. (1986).

Eur. J. Biochem. 161, 681-687.

PUBLICATION N° 2

Features of the primary structure of the *Methanosarcina barkeri* chromosomal protein HMb compared to those of eukaryotic and prokaryotic chromosomal proteins.

IMBERT, M., CHARTIER, F., LAINE, B. and SAUTIERE, P. (1986).

in Biology of Anaerobic Bacteria (DUBOURGUIER, H.C. *et al.*, eds) pp 236-241, Elsevier Science Publishers B.V. Amsterdam, The Netherlands.

PUBLICATION N° 3

The deoxyribonucleoprotein complex from the archaeobacterium *Methanosarcina barkeri* : characterization and ultrastructural localization of the chromosomal protein MC1.

IMBERT, M., LAINE, B., PRENSIER, G., TOUZEL, J.P. and SAUTIERE, P.

Soumise à Canadian Journal of Microbiology.

COMMUNICATIONS PAR AFFICHE

- Characterization of the chromosomal protein Hmb isolated from the archaeobacteria. *Methanosarcina barkeri*.

CHARTIER, F., IMBERT, M., LAINE, B. and SAUTIERE, P.
Proceeding of the EMBO workshop on Molecular Genetics of archaeobacteria and the international workshop on Biology and Biochemistry of archaeobacteria, Munich, Juin 1985 (KANDLER, O. and ZILLIG, W., eds) pp 403, Gustav Fisher Verlag. Stuttgart. New York.

- Structural characteristics of the *Methanosarcina barkeri* protein Hmb.

IMBERT, M., CHARTIER, F., LAINE, B. and SAUTIERE, P.
13th International Congress of Biochemistry 25-30 Août. Amsterdam The Netherlands.
Communication par affiche.

- CHARTIER, F., IMBERT, M., LAINE, B., SAUTIERE, P. and TOUZEL, J.P. (1986).

Characteristics of a chromosomal protein isolated from different species of methanosarcinaceae : implication in phylogeny.

Proceeding of the first international seminar on Biology of anaerobic bacteria, Villeneuve d'Ascq, 17-18 Juin 1986. (H.C. DUBOURGUIER *et al.*, eds) pp 242-244, Elsevier Science Publisher.

I N T R O D U C T I O N

Les micro-organismes, les seuls êtres vivants de notre planète pendant 2,5 milliards d'années peuvent être considérés comme une sorte de fossiles vivants qui portent inscrite dans leurs molécules, la longue histoire de l'origine et de l'évolution de la vie. Avec la découverte des archaebactéries une nouvelle perspective est apparue. En effet, quand on ne connaissait que deux lignées primaires (eucaryotes, procaryotes), il était difficile d'interpréter leurs différences. Avec trois lignées distinctes (eucaryotes, eubactéries, archaebactéries) les critères de comparaison étant plus nombreux, il devient plus facile de distinguer les propriétés ancestrales de celles qui sont apparues plus récemment. Plusieurs théories de l'évolution acceptent que les cellules eucaryotes doivent avoir évolué à partir d'un organisme de type bactérien et la condensation de l'ADN par les histones aurait une part importante dans cette transition de l'évolution. Aussi l'étude de protéines chromosomales chez les bactéries pourrait contribuer à une meilleure compréhension de l'origine des cellules eucaryotes.

Depuis quelques années, on sait que l'organisation de la chromatine des eubactéries n'est pas aussi bien définie que celle des eucaryotes, néanmoins on note la présence d'unités répétitives dans le nucléoïde (VARSHAVSKY et coll., 1977).

Jusqu'à présent l'organisation de l'ADN dans la cellule archaebactérienne a été peu étudiée. La présence de particules nucléoprotéiques a été signalée chez *Thermoplasma acidophilum* et *Sulfolobus acidocaldarius* mais l'organisation de la chromatine des halophiles et des méthanogènes reste inconnue.

Depuis 1983, une thématique nouvelle a été développée au laboratoire en collaboration avec la Station de Technologie Alimentaire de l'INRA (Villeneuve d'Ascq). Cette thématique

concerne l'organisation de la chromatine chez les bactéries méthanogènes et englobe des études structurales de la protéine chromosomale MCl isolée de souches de bactéries méthanogènes. Elle répond à deux objectifs :

- définir les caractéristiques structurales qui permettent à ces protéines d'organiser l'ADN. L'étude de la variabilité de séquence en acides aminés apporte de précieux renseignements sur le rôle de résidus d'acides aminés ou de domaines de la séquence dans la conformation et la fonction de la molécule.
- établir la phylogénie de souches méthanogènes qui ne se différencient pas sur la base de critères phénotypiques et génotypiques.

Mon activité dans ce domaine concerne les points suivants:

- l'étude de la structure primaire de la protéine MCl isolée de deux bactéries méthanogènes : *Methanosarcina barkeri* MS (souche type) et *Methanosarcina* sp MST-A1.

La protéine MCl de *Methanosarcina barkeri* MS a été isolée par F.CHARTIER dans le cadre de son D.E.A. Au commencement de mes travaux, la séquence N-terminale de la protéine était déterminée. J'ai contribué aux travaux qui ont permis d'établir la structure primaire complète de cette protéine.

- l'étude de la conformation de la protéine MCl isolée de *M.barkeri* à l'aide de méthodes spectroscopiques (dichroïsme circulaire, spectroscopie infrarouge, spectroscopie de fluorescence) et à l'aide de différentes méthodes de prédiction de structure secondaire.

- l'étude du complexe désoxyribonucléoprotéique isolé de *M. barkeri* MS par deux approches : la microscopie électronique et la biochimie, dans le but de déterminer le rôle de la protéine MCl dans la chromatine.

CHAPITRE I

G É N É R A L I T É S

I - 1 LES ARCHAEBACTÉRIES

L'esprit humain est satisfait lorsqu'il peut ordonner l'univers sur le mode dichotomique. Ainsi, jusqu'à une époque récente, les biologistes ont défini deux types de cellules dans le règne vivant : la cellule eucaryote qui possède un noyau et la cellule procaryote dépourvue de noyau. Cette distinction traditionnelle a été remise en cause en 1977 (WOESE et coll.) avec la découverte de micro-organismes dépourvus de noyau et possédant des macromolécules dont les caractéristiques sont telles que ces micro-organismes sont aussi différents des procaryotes que des eucaryotes.

Sur la base des comparaisons de séquences d'oligonucléotides obtenus après digestion par la ribonucléase T1 des ARN ribosomiaux (ARNr) 16 S bactériens et de quelques ARNr 18 S eucaryotes, WOESE et coll. (1977) ont montré que ces micro-organismes particuliers représentaient "une troisième forme de vie" sur terre. Ils ont proposé de leur donner le nom d'archaebactéries. Ce nom évoque la place que l'on a attribuée à ces micro-organismes dans l'évolution. En effet, certaines archaebactéries ont un métabolisme qui semble particulièrement adapté aux conditions ayant régné lors des débuts de la vie sur terre, il y a environ 3,5 milliards d'années. Il est donc possible que ce nouveau groupe soit le plus ancien.

La découverte de cette nouvelle lignée d'êtres vivants a, en outre, apporté une perspective sur l'évolution de la cellule en général. Elle permet en effet d'aborder deux questions fondamentales : celle de la nature de l'ancêtre commun à toutes les formes de vie et celle de l'évolution de la cellule eucaryote.

I - 1-1 DESCRIPTION PHENOTYPIQUE

Un facteur commun à toutes les archaebactéries est leur distribution écologique particulière : il s'agit d'habitats incompatibles pour la plupart des eubactéries. La classification des archaebactéries se fonde du reste sur cette distribution. Trois groupes ont été ainsi définis :

* *Les bactéries méthanogènes*

Ce sont des anaérobies strictes. En présence d'hydrogène, elles sont capables de réduire le gaz carbonique et des substrats organiques simples (formate, acétate, méthanol, méthylamines etc...) pour produire du méthane selon une voie métabolique particulière. Elles vivent donc dans des lieux sans oxygène où l'hydrogène et le dioxyde de carbone sont disponibles. Elles sont généralement associées à d'autres bactéries comme les *Clostridia* qui métabolisent les matières organiques en putréfaction et dégagent de l'hydrogène.

Les méthanogènes se trouvent dans les eaux stagnantes, dans les stations d'épuration, dans le rumen et les intestins des animaux mais également au fond des océans ainsi que dans des sources chaudes.

Ces bactéries présentent un intérêt industriel important puisqu'elles assurent la stabilité de transformation des matières organiques dans les digesteurs en anaérobiose et produisent du biogaz à partir des déchets agricoles.

* *Les bactéries halophiles aérobies*

Ces bactéries vivent dans des milieux très riches en sel (saumures, Grand Lac Salé, Mér Morte ...) et dans des conditions extrêmement alcalines. Bien qu'elles aient été étudiées depuis longtemps, leur intérêt n'est apparu que ces dernières années et ceci pour deux raisons :

- leur pouvoir de "pompes à ions" : elles se servent du gradient osmotique pour le transfert de molécules à travers la membrane cellulaire.

- leur pouvoir de photosynthèse relativement simple où n'intervient pas la chlorophylle : c'est un autre pigment de type rétinien (la rhodopsine) qui est mis en jeu.

Ces deux propriétés représentent des modèles d'intérêt biotechnologique.

* *Les bactéries thermoacidophiles*

On les trouve notamment dans des sources thermales très acides. Les différentes espèces vivent habituellement à des températures voisines de 80°C mais certaines variétés supportent des températures dépassant 90°C : croissance entre 100 et 110°C pour *Pyrodictium* (KANDLER, 1984). Bien que ces bactéries poussent à des pH acides (proches de 1,00), le pH du milieu interne de la cellule est voisin de la neutralité. Il existe donc un gradient de pH qui joue un rôle dans le pompage de molécules vers l'intérieur ou l'extérieur de la membrane cellulaire.

Les archaebactéries thermophiles sont capables de métaboliser le soufre. Les Sulfolobales oxydent le soufre en produisant de l'acide sulfurique, tandis que les Thermoprotéales le réduisent en H₂S. Cependant, *Thermoplasma acidophilum* ne métabolise pas le soufre. Cet organisme occupe d'ailleurs une position particulière dans les arbres phylogénétiques les plus récents (Figure I-7). *T.acidophilum* est un mycoplasme (organisme sans paroi), et jusqu'à présent on ne l'a trouvé que dans des terrils en combustion lente.

I - 1-2 PRINCIPALES CARACTERISTIQUES

Les archaebactéries sont classées parmi les procaryotes car elles en ont les caractéristiques : petite taille (1 micromètre de longueur), absence de membrane nucléaire et d'organites cellulaires, faible teneur en ADN. Cependant certains traits métaboliques et structuraux leur sont spécifiques :

* *Structure de la paroi*

La paroi des archaebactéries présente une grande diversité d'organisation structurale mais ne contient pas de muréine, composant rencontré dans la paroi des eubactéries. Cette dernière propriété rend les archaebactéries insensibles à la pénicilline. De nombreux autres antibiotiques actifs sur les eubactéries (chloramphénicol, kanamycine) sont également sans effet sur les archaebactéries.

Chez les Méthanobactériales, la muréine est remplacée par la pseudomuréine. Celle-ci diffère de son analogue eubactérien par le remplacement de l'acide muramique (D-Mur NAc) par l'acide L-N-acétyl talosaminuronique (L-NAc Tal NUA), par la nature des liaisons glycosidiques et par la séquence des éléments polypeptidiques (Figure I-1).

* *Lipides membranaires*

Alors que la majorité des lipides eubactériens et eucaryotes sont des esters de glycérol et d'acides gras, les lipides des archaebactéries sont des diéthers composés de glycérol et de deux molécules d'un alcool, le phytanol.

* *Synthèse protéique*

Par rapport aux deux autres règnes, la synthèse protéique des archaebactéries présente des particularités au niveau de la transcription et de la traduction.

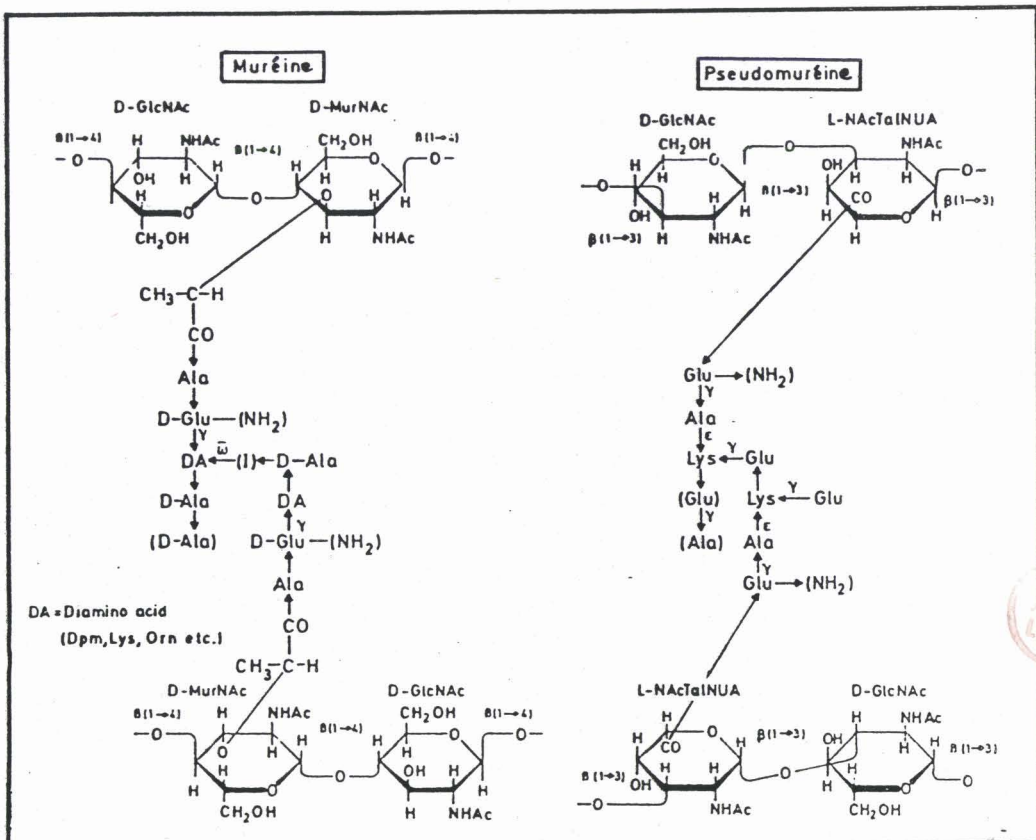


FIGURE I-1 - Comparaison de la structure primaire de la muréine d'eubactéries et de la pseudomuréine de Méthanobactériales (KANDLER, 1984).

ZILLIG et coll. (1979) et GROPP et coll. (1986) ont montré que les ARN polymérases ADN-dépendantes des archaebactéries diffèrent des ARN polymérases des eucaryotes et des eubactéries à la fois par leur composition en sous-unités et leur sensibilité aux antibiotiques.

Les protéines ribosomales laissent voir des propriétés singulières : leur nombre varie selon le groupe d'archaebactéries et diffère de celui des eubactéries (52 pour *E.coli*, 54 pour *Methanococcus* et 64 pour *Sulfolobus* (COCITO, 1986), tandis que leurs caractères structuraux sont partiellement eucaryotes ou partiellement eubactériens (ARNDT, 1986).

Chez les archaebactéries, le facteur d'élongation qui intervient dans la traduction, ressemble à celui des eucaryotes car il réagit avec la toxine diphtérique (KESSEL et coll., 1980).

* *Acides nucléiques et organisation du génome*

Les acides nucléiques des archaebactéries présentent non seulement des différences de séquence avec ceux des eucaryotes et eubactéries mais montrent également des propriétés particulières.

Les structures secondaires des ARN 5 S présentent une variabilité structurale importante chez les archaebactéries (WILLEKENS, 1986).

Parmi les espèces étudiées, on peut distinguer trois types de structure secondaire (Figure I-2). Cependant il existe un bras commun s'étendant des domaines B à H₁ qui s'apparente à celui des eubactéries, le reste de la structure est plus variable. Par ailleurs, chez les méthanogènes et les halophiles, la boucle I₂ a approximativement la même taille et la même forme que celle des eucaryotes.

Les séquences complètes des ARN 16 S ont révélé l'unicité des archaebactéries (WOESE et OLSEN, 1986) comme nous le verrons dans la partie I - 1-3.

L'absence de thymine et l'absence de dihydro-uracile dans l'ARN de transfert (ARNt) sont des particularités des archaebactéries. D'autre part l'ARNt initiateur porte la méthionine comme chez les eucaryotes.

La taille du génome archaebactérien n'a été déterminée que pour un petit nombre d'espèces, néanmoins elle varierait entre 2 millions de paires de bases (mbp) pour *Desulfurococcus mucosus* et 6,1 mbp pour *Methanobacterium thermoautotrophicum* (KLENK, 1986). Le code génétique est identique à celui des deux autres

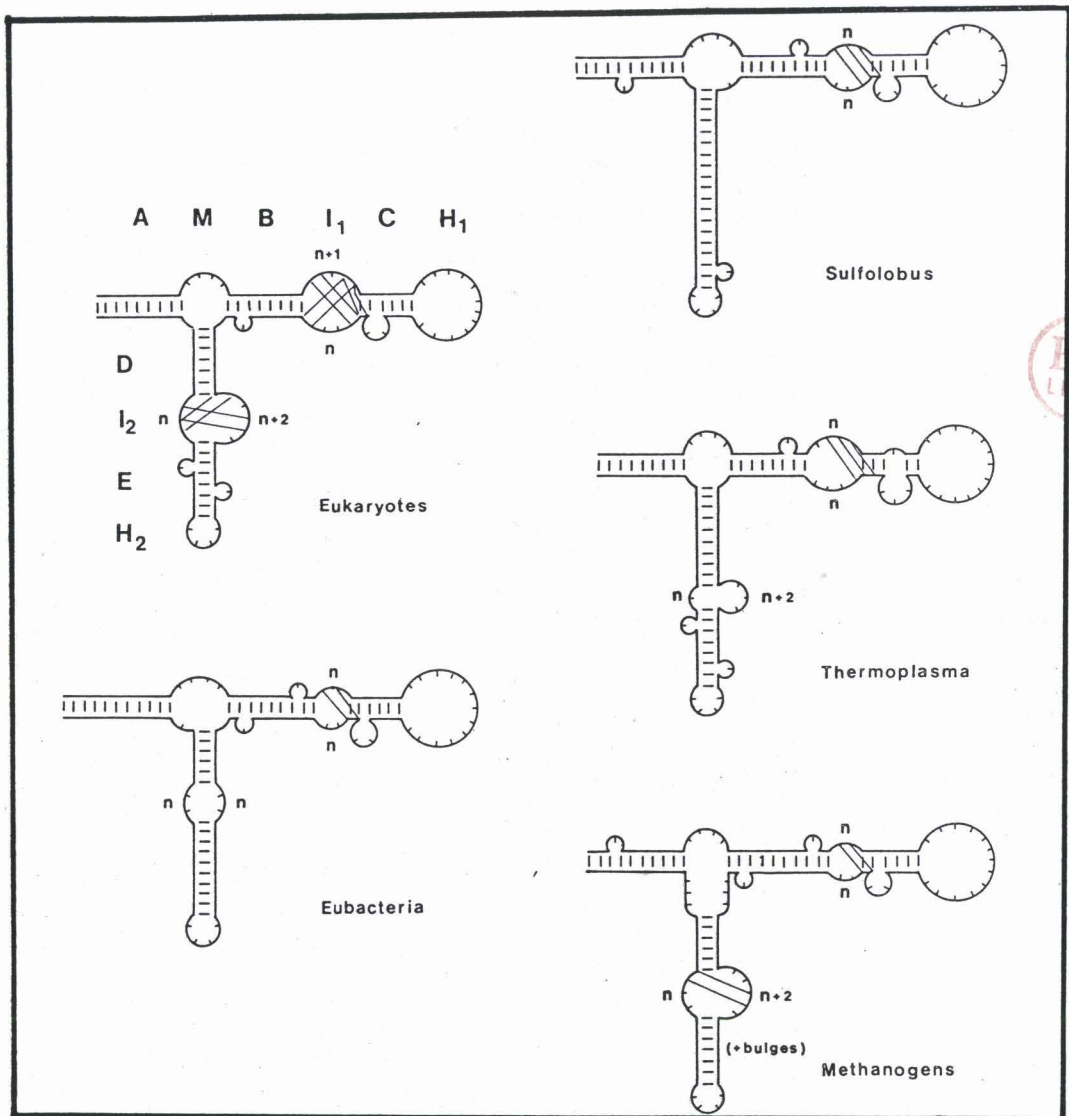


FIGURE I-2 - Caractères distinctifs des structures secondaires des ARN 5 S chez les eucaryotes, eubactéries et archaebactéries (WILLEKENS, 1986).

règles. Le pourcentage molaire en guanine et cytosine (G + C) est très variable au sein des archaebactéries. Ainsi dans l'ADN des méthanogènes on a signalé des taux de (G + C) allant de 25 à 55% (à l'intérieur d'un genre, généralement l'écart est inférieur à 10%). Ceci souligne la particularité et l'hétérogénéité des archaebactéries.

Comme chez les eucaryotes, des séquences de gènes dites silencieuses (introns) ont été trouvées chez une espèce halophile (DANIELS, 1985) et de Sulfolobales (KAINÉ, 1983).

Récemment, un enzyme qui introduit des supertours positifs dans l'ADN, la gyrase inverse a été découvert chez *S. acidocaldarius* (KIRUCHI et ASAI, 1984 ; FORTERRE et coll., 1986). Cet enzyme n'a jamais été identifié jusqu'à présent chez les eucaryotes ou les eubactéries. Ces dernières possèdent au contraire une gyrase qui introduit des supertours négatifs dans l'ADN.

I - 1-3 PHYLOGENIE

L'ensemble des données exposées ci-dessus montre clairement que les archaebactéries constituent un groupe d'organisation biochimiquement distinct aussi bien des eubactéries que des eucaryotes. Néanmoins, on peut se demander quelle place les archaebactéries occupent dans l'arbre généalogique des êtres vivants. Pour répondre à cette question, il fallait trouver de bons marqueurs phylogénétiques afin de pouvoir comparer les archaebactéries à la fois avec les eucaryotes et avec les eubactéries.

Ces marqueurs devaient répondre aux critères suivants :

- être distribués de façon universelle
- avoir joué un rôle constant pendant une longue durée à l'échelle de l'évolution
- avoir changé suffisamment lentement pour garder des caractéristiques ancestrales
- enfin être faciles à isoler.

Les molécules de choix qui répondent à ces critères sont les ARN ribosomiaux. Parmi ceux-ci deux molécules ont été étudiées, l'ARN 16 S et l'ARN 5 S. L'ARN 23 S a été écarté en raison de la difficulté à le caractériser. L'ARN 5 S a l'inconvénient de présenter parfois une trop forte variabilité structurale. L'ARN 16 S a été choisi par WOESE et ses collaborateurs (1977) à qui on doit la découverte des archaebactéries. L'approche choisie consiste en l'analyse comparative de séquences d'oligonucléotides obtenus à partir d'ARN 16 S. Des relations phylogénétiques pour un grand nombre de bactéries ont été quantifiées à l'aide du coefficient d'association S_{AB}^* .

Les résultats obtenus ont clairement montré que les archaebactéries forment un groupe cohérent, distinct à la fois des eucaryotes et des eubactéries. Ils ont fourni des renseignements précieux concernant les liens de parenté entre les différents groupes et fait ressortir des caractères communs ancestraux. Sur la base de ces données, WOESE et ses collaborateurs (1981) ont établi le schéma évolutif représenté figure I-3. Ce schéma indique que les archaebactéries ont dû se séparer des eubactéries très tôt dans l'évolution à partir d'un hypothétique ancêtre commun dit progénote.

Des travaux effectués par deux autres approches contestent ce schéma :

- HORI et ses collaborateurs (1982) discutent la validité de la "méthode S_{AB} " utilisée par WOESE et coll. Sur la base de séquences d'ARNr 5 S, ils ont établi un arbre phylogénétique différent où les archaebactéries seraient une déviation de la branche qui a conduit aux eucaryotes.
- A partir de l'analyse de la morphologie des ribosomes, LAKE et ses collaborateurs (1985) différencient quatre groupes principaux : les eubactéries, les eucaryotes, les archaebactéries constituées des halophiles et des méthanogènes, et les éocytes (Figure I-4). Ce nouveau groupe, les éocytes, rassemble les organismes dépendant du soufre.

* $S_{AB} = \frac{2 \times \text{nombre d'oligonucléotides communs à A et B}}{\text{nombre total d'oligonucléotides dans A et B}}$

(A et B sont des catalogues d'oligonucléotides de différents organismes)

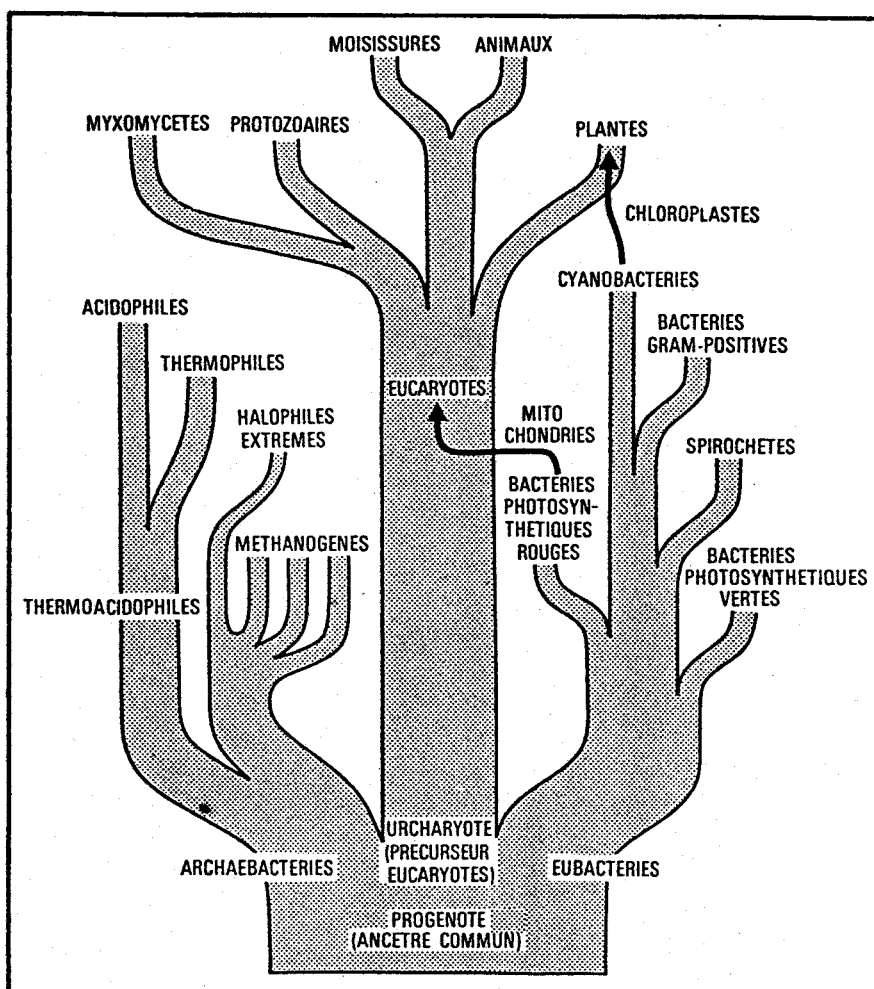


FIGURE I-3 - Schéma évolutif des êtres vivants d'après WOESE (1981) dans COCITO (1986).

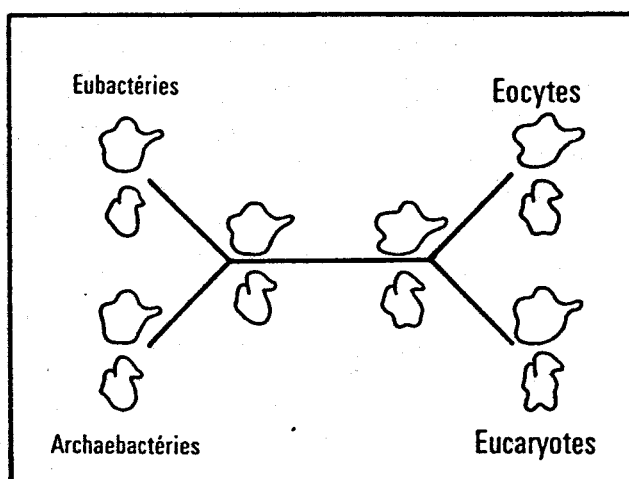


FIGURE I-4 - Représentation schématique des sous-unités ribosomales proposée par LAKE et coll. (1985)

Cependant les arguments de LAKE et coll. sont fortement contestés par plusieurs auteurs.

STÖFFLER et STÖFFLER-MEILICKE (1986), par des études en microscopie électronique des sous-unités ribosomales des archaebactéries, montrent que la classification de LAKE n'a pas de valeur phylogénétique. En effet ces auteurs ont trouvé des facteurs structuraux dans les sous-unités ribosomales de *Methanococcus vannieli* et *Halobacterium marismortui*, qui, selon LAKE et ses collaborateurs seraient spécifiques des éocytés.

WOESE et coll. (1986) reprochent à LAKE et coll. de donner un poids important à des arguments qui reposent sur l'absence supposée de certains caractères tels que l'absence d'introns à l'extérieur du groupe des archaebactéries dépendantes du soufre.

Depuis, WOESE et OLSEN (1986) ont souligné qu'il n'y avait pas d'information phylogénétique aussi fiable que celle qui est contenue dans des séquences. Actuellement les séquences complètes d'ARNr 16 S ont été déterminées chez dix espèces représentatives des archaebactéries. Elles révèlent que les archaebactéries ne sont sans doute pas plus proches des eucaryotes que des eubactéries. Elles occupent, semble-t-il, une position évolutive intermédiaire entre ces deux règnes (Figure I-5).

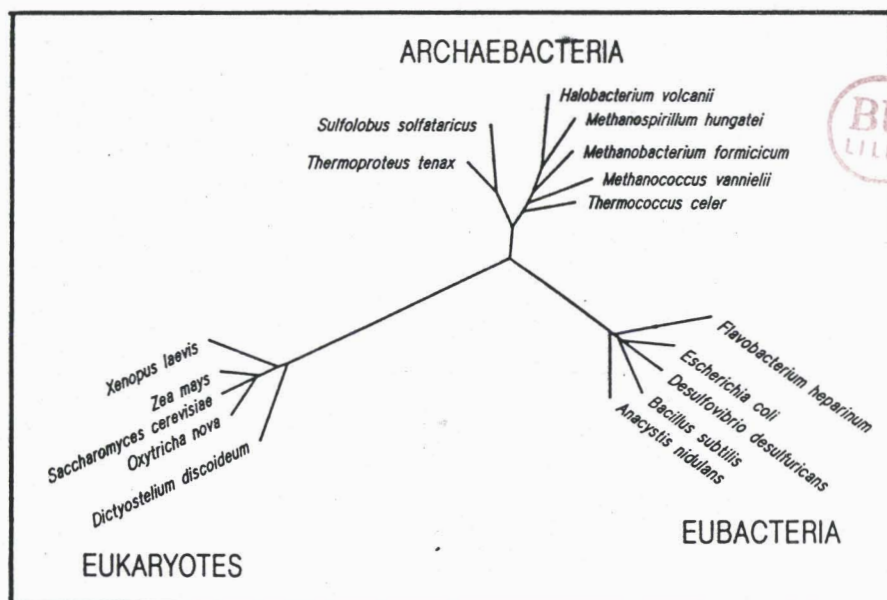


FIGURE I-5 - Arbre phylogénétique des trois règnes primaires selon WOESE et OLSEN (1986).

KLENK et coll. (1986) ont choisi la technique d'hybridation entre ADN et ARNr, méthode rapide et simple, pour analyser les relations phylogénétiques entre les différentes archaebactéries. L'arbre phylogénétique qu'ils ont établi est représenté figure I-7. On y distingue trois branches :

- Les méthanogènes, les halophiles et *Thermoplasma acidophilum*
- Les thermococcales
- Les sulfobactériales et thermoprotéales.

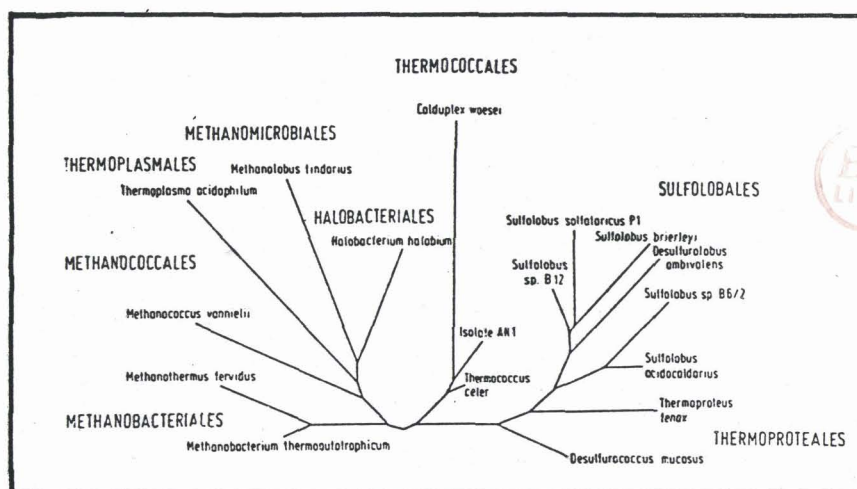


FIGURE I-7 - Arbre phylogénétique des archaebactéries d'après KLENK (1986)

Concernant l'ancêtre commun, il semblerait que les archaebactéries aient existé à un stade primaire de l'évolution (WOESE et OLSEN, 1986) ; les eubactéries et les eucaryotes se seraient formés ensuite à partir d'un ancêtre de nature archaebactérienne. Ceci justifierait alors le terme "urkingdom" pour définir ce groupe primaire d'organismes.

De même, les résultats acquis par GROPP et ses collaborateurs (1986) ont amené ces auteurs à suggérer que les trois ARN polymérases nucléaires des eucaryotes ont évolué à partir d'un enzyme de type ancestral rencontré dans les archaebactéries dépendantes du soufre.

Néanmoins, la nature de l'ancêtre commun à toutes les formes de vie reste encore incertaine. Par contre le fait que les archaebactéries constituent un groupe primaire spécifique est maintenant bien établi grâce aux études de la biologie moléculaire.

I - 1-4 LES METHANOGENES

En 1876, A.VOLTA a découvert que de "l'air combustible" se dégageait des marécages et des cours d'eau riches en végétaux en décomposition. Plus tard, on apprit que le gaz produit dans ces milieux était du méthane mais ce n'est qu'en 1956 qu'on a découvert que le méthane était synthétisé par les bactéries méthanogènes, ce qui a suscité un très grand intérêt pour ces bactéries.

I - 1-4-1 Principales caractéristiques

Les bactéries méthanogènes présentent, en plus des caractéristiques spécifiques aux archaebactéries, une hétérogénéité morphologique, structurale et trophique.

Toutes les formes morphologiques trouvées chez les eubactéries (cocci, bâtonnets, filaments, sarcines...) ont été mises en évidence chez les méthanogènes mais contrairement aux eubactéries, elles ne se développent que sur un nombre limité de substrats : l'hydrogène, le gaz carbonique, le formate, le méthanol, les méthylamines et l'acétate.

Les propriétés qui différencient les méthanogènes des autres archaebactéries sont les suivantes :

- a) La nécessité d'un environnement à faible potentiel d'oxydo-réduction : conséquence de l'anaérobiose stricte.
- b) La présence de coenzymes spécifiques jouant un rôle clé dans le mécanisme de biosynthèse du méthane : par exemple, le coenzyme M, les facteurs F420 et F430, les corrinoïdes, la méthanoptérine, le méthanofurane.

- c) L'assimilation du CO₂ ne se fait pas par le cycle de Calvin.
- d) L'absence de transporteurs d'électrons tels que les quinones et cytochromes. Cependant des cytochromes sont présents dans quelques cas (KUHN et GOTTSCHALK, 1983).

Récemment, BELAY et coll. (1984), MURRAY et ZINDER (1984) ont montré que certaines espèces de bactéries méthanogènes sont capables de fixer l'azote atmosphérique.

Les bactéries méthanogènes sont la principale source de méthane atmosphérique mais la formation de méthane nécessite la coopération d'au moins deux populations bactériennes différentes. En effet la dégradation de matières organiques fermentescibles s'effectue en plusieurs étapes. Les premières étapes sont assurées par une population bactérienne complexe et conduisent à la formation d'ammoniac, d'acétate et d'un mélange gazeux CO₂/H₂. Les bactéries méthanogènes qui interviennent dans la dernière étape ont un double rôle : elles éliminent du milieu ces composés dont l'accumulation est toxique pour les autres bactéries et synthétisent du méthane.

I - 1-4-2 Taxonomie

La classification des méthanogènes n'est pas une tâche facile compte tenu de leur grande diversité morphologique et physiologique.

La classification de BALCH et coll. (1979) reste à la base de la taxonomie des méthanogènes. Elle étudie plus précisément 18 souches connues en 1979. A partir de séquences partielles de l'ARN 16 S, BALCH et coll. différencient trois ordres principaux : les Méthanobactériales, les Méthanococcales et les Méthanomicrobiales. Cette classification a été complétée dernièrement par DUBACH et BACHOFEN (1985) avec des souches isolées récemment (Tableau I-1).

L'ordre des Méthanomicrobiales comprend trois familles dont la famille des Methanosarcinaceae. Selon la classification de BALCH et coll. (1979), cette famille comprend un seul

TABLEAU I-1

Classification des méthanoènes (DUBACH et BACHOFEN, 1985)



Order	Family	Genus	Species	Strain		
Methanobacterales	Methanobacteriaceae	<i>Methanobacterium</i>	<i>M. bacterium formicicum</i>	MF *		
			<i>M. bacterium bryantii</i>	M.o.H. *		
			<i>M. bacterium bryantii</i>	M.o.H.G. *		
			<i>M. bacterium thermoautotrophicum</i>	JH *		
				<i>Methanobrevibacter</i>	<i>M. bacterium thermoautotrophicum</i>	Marburg *
					<i>M. brevisbacter ruminantium</i>	M1 *
					<i>M. brevisbacter arboriphilus</i>	DH1 *
					<i>M. brevisbacter arboriphilus</i>	AZ *
					<i>M. brevisbacter arboriphilus</i>	DC *
					<i>M. brevisbacter smithii</i>	PS *
			Methanothermaceae	<i>Methanothermus</i>	<i>M. thermus fervidus</i>	V24S *
		Methanococcales	Methanococcaceae	<i>Methanococcus</i>	<i>M. coccus ramielli</i>	SB *
					<i>M. coccus voltae</i>	PS *
					<i>M. coccus maza</i>	MC3 *
					<i>M. coccus deltae</i>	JRC *
					<i>M. coccus maripaludis</i>	JJ *
					<i>M. coccus thermolithotrophicus</i>	SN1 *
					<i>M. coccus jamaeschii</i>	JAL1 *
Methanomicrobiales	Methanomicrobiaceae			<i>Methanomicrobium</i>	<i>M. microbium mobile</i>	BP *
					<i>M. microbium paynteri</i>	
		<i>Methanogenium</i>	<i>M. genium cariaci</i>	JR1 *		
			<i>M. genium marisnigri</i>	JR1 *		
			<i>M. genium olentangyi</i>	RC ER *		
			<i>M. genium tati</i>			
				<i>Methanospirillum</i>	<i>M. genium thermophilicum</i>	CR1 *
					<i>M. spirillum hungatei</i>	JF1 *
					<i>M. spirillum hungatei</i>	GP1 *
				<i>Methanothrix</i>	<i>M. thrix soehngenii</i>	Opfikon *
				<i>Methanococcoides</i>	<i>M. coccoides methylatus</i>	TMA10 *
		Methanosarcinaceae	<i>Methanosarcina</i>		<i>M. sarcina barkeri</i>	MS *
				<i>M. sarcina barkeri</i>	227 *	
				<i>M. sarcina barkeri</i>	W *	
				<i>M. sarcina barkeri</i>	UBS *	
				<i>M. sarcina barkeri</i>	DM *	
<i>M. sarcina barkeri</i>	FR1 *					
<i>M. sarcina acetivorans</i>	C2A *					
<i>M. sarcina vacuolata</i>	Z *					
<i>M. sarcina sp.</i>	TM1 *					
	Methanoplanaceae			<i>Methanoplanus</i>	<i>M. planus limicola</i>	M3 *

M = Methano; *based on classification scheme of Balch et al.; *classified by 16S rRNA oligonucleotide cataloguing.

genre : *Methanosarcina* . Il se caractérise par une morphologie particulière en forme de sarcines avec une paroi constituée d'hétéropolysaccharides. Le contenu en guanine et cytosine (G + C) varie au sein de ce genre entre 38,8 et 51%. Toutes les souches se caractérisent par une croissance sur acétate, méthanol, méthylamines. Une seule espèce *Methanosarcina barkeri* utilise également le mélange H₂/CO₂.

Notre travail a porté plus particulièrement sur deux souches de *Methanosarcina* :

- une bactérie mésophile : *Methanosarcina barkeri* souche MS, souche de référence qui a fait l'objet d'études structurales et métaboliques approfondies.
- une bactérie thermophile : *Methanosarcina* sp souche MST-A1.

Les caractéristiques de ces deux souches seront détaillées dans le chapitre III.

I - 2 ORGANISATION DE L'ADN

Avant d'aborder les travaux réalisés sur le nucléoïde des archaebactéries, il semble utile de rappeler quelques données générales concernant la condensation de l'ADN dans les cellules eucaryotes et eubactériennes.

I - 2-1 ORGANISATION DE L'ADN DANS LA CELLULE EUCARYOTE

Chez les eucaryotes, la plus grande partie de l'ADN est contenue dans le noyau et se trouve impliquée dans une structure complexe appelée chromatine. Celle-ci est constituée d'ADN, d'ARN et de protéines. L'ADN est associé à une grande variété de protéines parmi lesquelles on distingue les histones et les protéines chromosomiques non histones.

Les protéines non histones recouvrent de très nombreuses protéines de fonctions différentes. Au contraire les histones constituent une classe homogène de cinq protéines appelées H1, H2A, H2B, H3 et H4. Le rapport pondéral histones/ADN est voisin de 1. Les histones sont des protéines de faible masse moléculaire (M_r 11.000 à 22.000) riches en acides aminés chargés positivement (lysine et arginine). Elles sont pauvres en

acides aminés aromatiques ou soufrés et ne contiennent pas de tryptophane. La structure primaire des histones est caractérisée par une asymétrie de répartition des charges positives dans la chaîne polypeptidique. Les histones H2A, H2B, H3 et H4 possèdent une région N-terminale très basique qui représente environ le quart de la molécule. Cette région, de structure inorganisée, serait impliquée dans les interactions électrostatiques avec les groupes phosphate de l'ADN. Le reste de la chaîne protéique présente un caractère hydrophobe. Cette région hydrophobe, particulièrement bien conservée au cours de l'évolution est le siège d'interactions protéine-protéine.

L'histone H2A présente dans sa partie C-terminale une courte séquence basique et variable. L'histone H1 est composée d'un domaine globulaire de nature hydrophobe en position centrale et de deux domaines de structure inorganisée, de caractère très basique en positions amino- et carboxy-terminales.

La chromatine est organisée en structures élémentaires répétitives, les nucléosomes. Ces structures de forme globulaire donnent à la chromatine l'apparence d'un collier de perles en microscopie électronique. Le nucléosome est constitué d'une "core particle" et d'un fragment d'ADN appelé "DNA linker" auquel sont associées une molécule d'histone H1 et deux protéines non histones de type HMG ("High Mobility Group"). Le "DNA linker" a une longueur variable selon l'origine de la cellule. La "core particle" est un complexe octamérique formé de deux molécules de chacune des histones H2A, H2B, H3 et H4 autour duquel est enroulé un fragment de 146 paires de bases d'ADN. La "core particle" est une sous-unité constante quels que soient les tissus et les organismes. Dans la cellule vivante, la chromatine est maintenue sous une forme très compacte où les nucléosomes sont empilés en une structure solénoïdale qui apparaît en microscopie électronique comme une fibre de 30 nm de diamètre et correspond à une structure régulière d'ordre supérieur. L'histone H1 joue un rôle essentiel dans la formation de cette superstructure.

I - 2-2 ORGANISATION DE L'ADN CHEZ LES EUBACTERIES

La cellule procaryote est dépourvue de noyau délimité par une membrane nucléaire, cependant l'ADN chromosomique n'est pas dispersé dans toute la cellule. Il se trouve dans une région restreinte appelée corps nucléaire ou nucléoïde et est souvent associé à la membrane bactérienne par l'intermédiaire d'un mésosome.

A quelques exceptions près, les études concernant la structure du nucléoïde ont été effectuées chez *Escherichia coli*. Les recherches ont été limitées par la grande instabilité du nucléoïde et par la difficulté de préparer celui-ci exempt de membranes cellulaires. Cette difficulté a été levée par VARSHAVSKY et coll. (1977) en utilisant une digestion douce de la chromatine par les endonucléases endogènes ce qui leur a permis d'obtenir de grands fragments de complexe désoxyribonucléoprotéique débarrassé des protéines membranaires. Les composants protéiques majeurs de ce complexe sont l'ARN polymérase ainsi que deux protéines B1 et B2 dont la masse moléculaire respective est de 17.000 et 10.000. Les protéines B1 et B2 ont par la suite été identifiées respectivement aux protéines H1 (SPASSKY et BUC, 1977) et HU (ROUVIERE-YANIV et GROS, 1975). La protéine HU est maintenant appelée protéine II. Une molécule de chaque protéine est associée à des fragments d'ADN de 120 paires de bases qui sont résistants à la nucléase staphylococcale. Les protéines H1* et II sont distribuées uniformément le long de l'ADN. A côté de ces protéines, cinq autres protéines chromosomales ont été trouvées dans un extrait acido-soluble du nucléoïde d'*E.coli* (YAMAZAKI et coll., 1984), mais les protéines H1 et II ont reçu une plus grande attention. Ces protéines sont abondantes, thermostables et se comportent comme des dimères en solution.

* La protéine H1 ne doit pas être confondue avec l'histone H1 des eucaryotes.

I - 2-2-1 Propriétés de la protéine H1

Elle inhibe l'initiation de la transcription *in vitro*. Elle est constituée de trois variants de points isoélectriques différents. La synthèse du variant H1a est fortement augmentée en fin de phase stationnaire de croissance.

La protéine H1a peut replier et compacter ainsi fortement l'ADN bicaténaire.

I - 2-2-2 Propriétés de la protéine II

La fonction de la protéine II n'est pas encore clairement définie. La plupart des auteurs attribuent à cette protéine un rôle dans la condensation de l'ADN mais d'autres auteurs pensent que la protéine II aurait un rôle dans la régulation.

* Rôle de la protéine II dans la condensation de l'ADN

La protéine II stabilise l'ADN contre la dénaturation thermique (ROUVIERE-YANIV et coll., 1979 ; MIANO et coll., 1982). Cet effet est d'autant plus important que la quantité de protéine II associée à l'ADN augmente.

En présence de topoisomérase I, la protéine II d'*E.coli* introduit 18 supertours négatifs dans l'ADN circulaire relâché du virus SV40. La protéine II forme avec cet ADN des structures globulaires de 18 nm de diamètre et dans ce complexe le degré de condensation de l'ADN est similaire à celui observé dans les complexes reconstitués avec de l'ADN de SV40 et les histones (ROUVIERE-YANIV et coll., 1979).

La digestion à la nucléase staphylococcale des complexes reconstitués ADN-protéine II produit sur chaque brin d'ADN des coupures avec une périodicité de 8,5 paires de bases (BROYLES et PETTIJOHN, 1986) au lieu d'une périodicité de 10,5 paires de bases pour l'ADN seul. Ces résultats suggèrent

que lorsque l'ADN est associé à la protéine II, le pas de l'hélice est réduit ; l'ADN se trouve alors fortement compacté tout en gardant une configuration superhélicoïdale similaire à celle observée *in vivo* dans le chromosome d'*E.coli*. L'ADN chromosomal d'*E.coli* est en effet divisé en une centaine de "domaines" à l'intérieur desquels il est maintenu dans un état de torsion négative (SINDEN et PETTIJOHN, 1981). Les "domaines" sont indépendants c'est-à-dire que le relâchement de la superhélicité provoqué par une coupure sur un brin d'ADN à l'intérieur d'un "domaine" ne se propage pas dans les domaines adjacents.

* Rôle de la protéine II dans la régulation

Le rôle que pourrait jouer la protéine II dans l'organisation du nucléoïde a été contesté par plusieurs auteurs tandis que d'autres ont montré qu'elle serait impliquée dans la régulation de la transcription ou de la réplication de l'ADN.

En 1984, DIXON et KORNBERG ont étudié la réplication enzymatique de l'ADN double brin de phages recombinants ou de plasmides portant l'origine de réplication d'*E.coli* (*ori c*) et ont mis en évidence que la protéine II stimulait la réplication. Par ailleurs la protéine II inhibe la transcription de l'ADN de source procaryote ou eucaryote (LOSSO et coll., 1982). Toutefois ces auteurs notent qu'il paraît invraisemblable que cette propriété représente une fonction physiologique de la protéine.

A l'aide d'anticorps anti-protéine II marqués à la protéine A-or colloïdal, BJORNSTI et coll. (1986) ont localisé en microscopie électronique la protéine II dans les régions du nucléoïde contenant des ribosomes, ce qui laisse supposer que la protéine jouerait un rôle dans la transcription. Par ailleurs on notera que la protéine II a été associée à la sous-unité 30 S du ribosome d'*E.coli* (SURYANARAYANA et SUBRAMANIAN, 1978).

* Propriétés structurales de la protéine II

La protéine II a été isolée chez un grand nombre de bactéries. La protéine II est un polypeptide de 90 résidus qui ne contient ni cystéine, ni tryptophane et dans la plupart des cas est dépourvue de tyrosine. Chez plusieurs espèces eubactériennes cette protéine est un complexe hétérologue constitué de deux chaînes polypeptidiques de structure primaire étroitement apparentée.

La protéine II n'est que très légèrement basique (charge nette +4) et on n'observe pas dans cette molécule une asymétrie de répartition des résidus de lysine et d'arginine qui sont dispersés le long de la chaîne polypeptidique. Deux résidus d'arginine placés dans une région bien conservée dans toutes les protéines séquencées actuellement (résidus 53 à 61) jouent un rôle essentiel dans l'interaction de la protéine avec l'ADN : ils sont le siège d'interactions électrostatiques avec les groupes phosphate de l'ADN. La structure tridimensionnelle de la protéine II a été déterminée par l'étude de la diffraction des rayons X de cristaux de protéine II de *B.stearothermophilus* avec une résolution de 3 Å (Figure I-8).

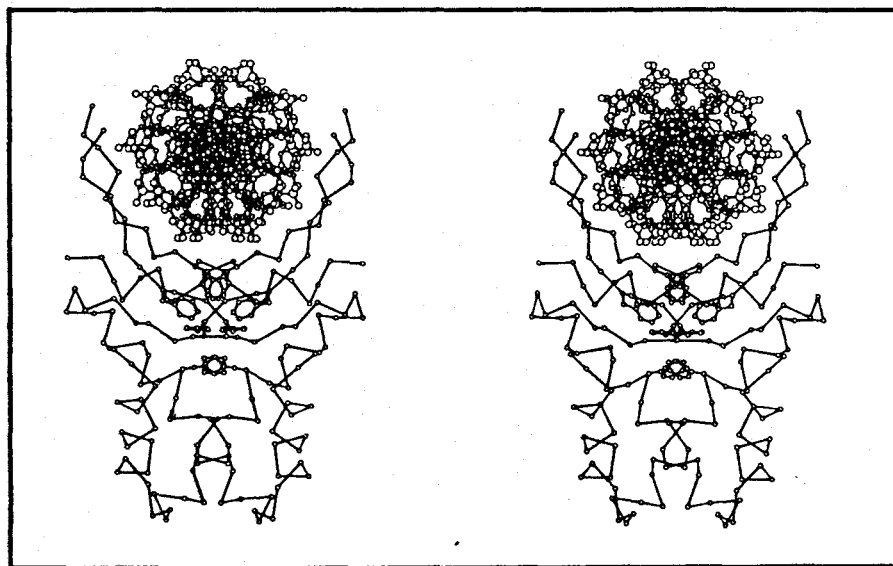


FIGURE I-8 - Structure du dimère de la protéine II de *B.stearothermophilus*. Vue stéréochimique montrant l'interaction avec l'ADN (TANAKA et coll., 1984).

La protéine II est un dimère qui correspond très certainement à son état fonctionnel dans sa liaison avec l'ADN. Dans le dimère, les segments de structure organisée (hélice α et feuillet β) forment une structure globulaire d'où émergent les deux bras externes qui délimitent un réceptacle où peut se loger la double hélice d'ADN en forme B. C'est sur ces bras que sont placés les deux résidus d'arginine cités précédemment.

I - 2-3 ORGANISATION DE L'ADN CHEZ LES ARCHAEBACTERIES

Seuls les travaux sur l'organisation du chromosome des archaebactéries *Thermoplasma acidophilum* et *Sulfolobus acidocaldarius* ont été publiés jusqu'à présent.

SEARCY et STEIN (1980) ont étudié la structure du complexe déoxyribonucléoprotéique chez *Thermoplasma acidophilum*. Ce complexe contient une protéine appelée HTa avec un rapport protéine/ADN égal à 0,6 (Nous verrons plus loin que la protéine HTa peut être assimilée à la protéine II des eubactéries). Le traitement de ce complexe par la nucléase staphylococcale libère des fragments de 40 paires de bases ; ces fragments résistants à l'action de la nucléase représentent 20% de l'ADN de cette bactérie. Quelle que soit la durée de la digestion, SEARCY et STEIN n'obtiennent pas de fragments d'ADN dont la longueur est un multiple de celle d'une sous-unité comme on peut en obtenir avec la chromatine des eucaryotes. La protéine II de *T. acidophilum* stabilise également l'ADN contre la dénaturation thermique (SEARCY, 1975 et STEIN, 1978). En microscopie électronique SEARCY et STEIN (1980) ont observé à la fois sur les complexes nucléoprotéiques natifs et reconstitués, des particules globulaires de 5 à 6 nm de diamètre qui consisteraient en une boucle d'ADN enroulée autour de 4 molécules de protéines II. Ces sous-unités répétitives rappellent

celles qui ont été mises en évidence dans la chromatine des eucaryotes. Par ailleurs la protéine II de *T.acidophilum* est riche en résidus basiques (22% du total des résidus, tableau I-2) souvent associés en doublets et triplets. L'ensemble de ces propriétés a conduit SEARCY à penser qu'elle ressemblait fortement aux histones. Cependant la protéine chromosomale de *T.acidophilum* doit être assimilée à la protéine II associée à l'ADN des eubactéries car elle présente avec cette dernière la même taille moléculaire et des homologues de structure primaire hautement significatives (DeLANGE et coll., 1981). En effet les acides aminés jouant un rôle important dans la fonction de la protéine sont strictement conservés.

Chez l'archaebactérie thermophile *Sulfolobus acidocaldarius*, GREEN et coll. (1983) ont découvert deux protéines chromosomales majeures : une protéine acido-soluble HSA (M_r 14.500 et rapport protéine/ADN = 0,21) et une protéine non acido-soluble NHSa (M_r 36.000 et rapport protéine/ADN = 0,22). Ces deux protéines protègent ensemble 5% de l'ADN contre la digestion à la nucléase staphylococcale et les fragments produits sont de différentes tailles. Le fragment moyen correspond à 25 paires de bases. Les protéines HSA et NHSa stabilisent également l'ADN contre la dénaturation thermique. Cependant, la protection de l'ADN contre la digestion par l'endonucléase et contre la dénaturation thermique est moins efficace que celle obtenue avec la protéine HTa de *Thermoplasma acidophilum*. La protéine HSA est moins basique que la protéine HTa et la protéine II d' *E.coli* (Tableau I-2).

La présence d'autres protéines chromosomales de faibles poids moléculaires a été mise en évidence chez *Sulfolobus brierleyi* (THOMM ET COLL., 1982).

Une protéine qui a une affinité pour l'ADN a été isolée de *Sulfolobus solfataricus* par KIMURA et coll. (1984). Sa séquence en acides aminés ne présente pas d'homologies avec celle des autres protéines chromosomales connues jusqu'alors.

TABLEAU I-2

Compositions en acides aminés des protéines chromosomales de *S.acidocaldarius* (HSa : GREEN et coll., 1983), de *T.acidophilum* (protéine II : DeLANGE et coll., 1981) et d'*E.coli* (protéine II : LAINE et coll., 1978)

Acides aminés	HSa <i>S.acidocaldarius</i>	Protéine II <i>T.acidophilum</i>	Protéine II <i>E.coli</i>
Acide aspartique	9,1*	5,6	8,9
Thréonine	3,7	5,6	6,8
Sérine	8,2	7,8	4,5
Acide glutamique	10,8	13,3	9,5
Proline	3	3,3	2,6
Glycine	9,5	6,7	7,2
Alanine	7,1	7,8	18,5
1/2 Cystine	1,8	0	0
Valine	4,5	8,9	7,4
Méthionine	1,1	1,1	1,2
Isoleucine	4,1	7,8	6,6
Leucine	14,7	3,3	6,9
Tyrosine	2,8	1,1	0
Phénylalanine	1,8	5,6	3,3
Histidine	0,2	0	0,9
<u>Lysine</u>	<u>10,2</u>	<u>15,6</u>	<u>11,1</u>
<u>Arginine</u>	<u>5,3</u>	<u>6,7</u>	<u>4,7</u>
Tryptophane	nd	0	0
Résidus basiques (%)	15,7	22,3	16,7

* Les résultats sont exprimés en moles/100 moles.

En 1986, GROTE et coll. ont isolé chez *Sulfolobus acidocaldarius*, trois classes de protéines chromosomales de masse moléculaire 7 kDa, 8 kDa et 10 kDa ; chaque classe correspond à un mélange d'au moins deux protéines de charge différente. L'étude en microscopie électronique des complexes reconstitués avec les protéines de 10 kDa et de l'ADN double brin ou simple brin ne montrent pas de diminution significative de la longueur de l'ADN (LURTZ et coll., 1986) ce qui suggère que la fonction de ces protéines ne serait pas de compacter l'ADN comme le font la protéine II et les histones. D'ailleurs il faut souligner que la présence de ces protéines dans le nucléoïde de *Sulfolobus acidocaldarius* n'a pas été démontrée puisque l'affinité pour l'ADN ne peut être un critère pour affirmer qu'une protéine est effectivement associée à l'ADN *in vivo* (GREEN et coll., 1983).

La protéine chromosomale de la bactérie méthanogène *Methanosarcina barkeri* a été isolée au laboratoire (CHARTIER et coll., 1985). Les propriétés de cette protéine appelée MC1, ainsi que les travaux effectués sur le chromosome de cette bactérie seront décrits dans les chapitres suivants.

BIBLIOGRAPHIE

- ARNDT, E., BREITHAUPT, G. and KIMURA, M. (1986)
The complete amino acid sequence of ribosomal protein H-S11 from the archaebacterium *Halobacterium marismortui*.
FEBS Lett. 194, 227-234.
- BALCH, W.E., FOX, G.E., MAGRUM, L.J., WOESE, C.R. and WOLFE, R.S. (1979)
Methanogens : reevaluation of a unique biological group.
Microbiol.Rev. 43, 260-296.
- BELAY, N., SPARLING, R. and DANIELS, L. (1984)
Dinitrogen fixation by a thermophilic methanogenic bacterium.
Nature 312, 286-288.
- BJORNSTI, M.A., HOBOT, J.A., KELUS, A.S., VILLIGER, W. and KELLENBERGER, E. (1986)
New electron microscopic data on the structure of the nucleoid and their functional consequences.
in Bacterial chromatin (GUALERZI, C.O. and PON, C.L. eds) pp 64-81, Springer-Verlag.
- BROYLES, S.S. and PETTIJOHN, D.E. (1986)
Interaction of the *Escherichia coli* HU protein with DNA. Evidence for formation of nucleosome-like structures with altered DNA helical pitch.
J.Mol.Biol. 187, 47-60.
- CHARTIER, F., LAINE, B., SAUTIERE, P., TOUZEL, J.P. and ALBAGNAC, G. (1985)
Characterization of the chromosomal protein Hmb isolated from *Methanosarcina barkeri*.
FEBS Lett. 183, 119-123.
- COCITO, C. (1986)
Archaeobactéries et évolution.
Médecine/Sciences 2, 436-444.
- DANIELS, C.J., GUPTA, R. and DOOLITTLE, W.F. (1985)
Transcription and excision of a large intron in the trNA (trp) gene of an archaebacterium *Halobacterium volcanii*.
J.Biol.Chem. 260, 3132-3134.

- DeLANGE, R.J., WILLIAMS, L.C. and SEARCY, D.G. (1981)
A histone-like protein (HTa) from *Thermoplasma acidophilum* :
II Complete amino acid sequence.
J. Biol. Chem. 256, 905-911.
- DIXON, N.E. and KORNBERG, A. (1984)
Protein HU in the enzymatic replication of the chromosomal
origin of *Escherichia coli*.
Proc. Natl. Acad. Sci. USA 81, 424-428.
- DUBACH, A. CH and BACHOFEN, R. (1985)
Methanogens : a short taxonomic overview.
Experientia 41, 441-446.
- FORTERRE, P. (1986)
Les archaebactéries : origine et évolution des lignées cel-
lulaires procaryotes et eucaryotes.
Ann. Biol. 25, 273-306.
- FOX, G.E., MAGRUM, L.J., BALCH, W.E., WOLFE, R.S. and WOESE, C.R.
(1977)
Classification of methanogenic bacteria by 16S ribosomal
RNA characterization.
Proc. Natl. Acad. Sci. USA 74, 4537-4541.
- GREEN, G.R., SEARCY, D.G. and DeLANGE, R.J. (1983)
Histone-like protein in the archaebacterium *Sulfolobus*
acidocaldarius.
Biochim. Biophys. Acta 741, 251-257.
- GROPP, F., REITER, W.D., SENTENAC, A., ZILLIG, W., SCHANBEL, R.,
THOMN, M. and STETTER, K.O. (1986)
Homologies of components of DNA-dependent RNA polymerases
of archaebacteria, eukaryotes and eubacteria.
System. Appl. Microbiol. 7, 95-101.
- GROTE, M., DIJK, J. and REINHARDT, R. (1986)
Ribosomal and DNA binding proteins of the thermoacidophilic
archaebacterium *Sulfolobus acidocaldarius*.
Biochim. Biophys. Acta 873, 405-413.
- HORI, H., ITOH, T. and OSAWA, S. (1982)
The phylogenetic structure of the metabacteria.
Zbl. Bakt. Hyg. I, Abt. Orig. C3, 18-30.
- KAINE, B.P., GUPTA, R. and WOESE, C.R. (1983)
Putative introns in tRNA genes of prokaryotes.
Proc. Natl. Acad. Sci. USA 80, 3309-3312.
- KANDLER, O. (1984)
Archaebacteria-biotechnological implications.
Preprint from : Proceedings third of the european congress
on biotechnology. Symposium : Futuristic aspects of bio-
technology. Munchen, 1984.

- KESSEL, M. and KLINK, F. (1980)
Archaeobacterial elongation factor is ADP-ribosylated by diphtheria toxin.
Nature 287, 250-251.
- KIMURA, M., KIMURA, J., DAVIE, P., REINHARDT, R. and DIJK, J. (1984)
The amino acid sequence of a small DNA binding protein from the archaeobacterium *Sulfolobus solfataricus*.
FEBS Lett. 176, 176-178.
- KIRUCHI, A. and ASAI, K. (1984)
Reverse gyrase, a topoisomerase which introduces positive superhelical turns into DNA.
Nature 309, 677-681.
- KLENK, H-P., HAAS, B., SCHWASS, V. and ZILLIG, W. (1986)
Hybridization homology : a new parameter for the analysis of phylogenetic relations, demonstrated with the urkingdom of the archaeobacteria.
J.Mol.Evol. 24, 167-173.
- KUHN, W. and GOTTSCHALK, G. (1983)
Characterization of the cytochromes occurring in *Methanosarcina* species.
Eur.J.Biochem. 135, 89-94.
- LAINE, B., SAUTIERE, P., BISERTE, G., COHEN-SOLAL, M., GROS, F. and ROUVIERE-YANIV, J. (1978)
The amino- and carboxy-terminal amino acid sequences of protein HU from *Escherichia coli*.
FEBS Lett. 89, 116-120.
- LAKE, J.A. (1985)
Evolving ribosome structure : domains in archaeobacteria, eubacteria, eocytes and eukaryotes.
Ann.Rev.Biochem. 54, 507-530.
- LAKE, J.A., CLARK, M.W., HENDERSON, E., FAY, S.P., OAKES, M., SCHEINMAN, A., THORNER, J.P. and MAH, R.A. (1985)
Eubacteria, halobacteria and the origin of photosynthesis : the photocytes.
Proc.Natl.Acad.Sci.USA 82, 3716-3720.
- LOSSO, M.A., MIANO, A., GIANFRANCESCHI, G.L. and GUALERZI, C.O. (1982)
Proteins from the prokaryotic nucleoid. II Inhibition of DNA transcription by NS1 and NS2 (HU proteins).
Biochem.Int. 5, 423-427.
- LURTZ, R., GROTE, M., DIJK, J., REINHARDT, R. and DOBRINSKI, B. (1986)
Electron microscopic study of DNA complexes with proteins from the archaeobacterium *Sulfolobus acidocaldarius*.
EMBO J. 5, 3715-3721.

- MIANO, A., LOSSO, M.A., GIANFRANCESCHI, G.L. and GUALERZI, C.O. (1982)
Proteins from the prokaryotic nucleoid. I - Effect of NS1 and NS2 (HU) proteins on the thermal stability of DNA.
Biochem.Int. 5, 415-422.
- MURRAY, P.A. and ZINDER, S.H. (1984)
Nitrogen fixation by a methanogenic archaeobacterium.
Nature 312, 284-286.
- ROUVIERE-YANIV, J. and GROS, F. (1975)
Characterization of a novel, low-molecular-weight DNA binding protein from *Escherichia coli*.
Proc.Natl.Acad.Sci.USA 72, 3428-3432.
- ROUVIERE-YANIV, J., YANIV, M. and GERMOND, J.E. (1979)
E.coli DNA binding protein HU forms nucleosome-like structure with circular double stranded DNA.
Cell 17, 265-274.
- SEARCY, D.G. (1975)
Histone-like protein in the prokaryote *Thermoplasma acidophilum*.
Biochim.Biophys.Acta 395, 535-547.
- SEARCY, D.G. and STEIN, D.B. (1980)
Nucleoprotein subunit structure in an unusual prokaryotic organism : *Thermoplasma acidophilum*.
Biochim.Biophys.Acta 609, 180-195.
- SINDEN, R.R. and PETTIJOHN, D.E. (1981)
Chromosomes in living *Escherichia coli* cells are segregated into domains of supercoiling.
Proc.Natl.Acad.Sci.USA 78, 224-228.
- SPASSKY, A. and BUC, H.C. (1977)
Physico-chemical properties of a DNA binding protein : *Escherichia coli* factor H1.
Eur.J.Biochem. 81, 79-80.
- STEIN, D.B. and SEARCY, D.G. (1978)
Physiologically important stabilization of DNA by a prokaryotic histone-like protein.
Science 202, 219-221.
- STÖFFLER, G. and STÖFFLER-MEILICKE, M. (1986)
Electron microscopy of archaeobacterial ribosomes.
System.Appl.Microbiol. 7, 123-130.
- SURYANARAYANA, T. and SUBRAMANIAN, A.R. (1978)
Specific association of two homologous DNA-binding proteins to the native 30 S ribosomal subunits of *Escherichia coli*.
Biochim.Biophys.Acta 520, 342-357.

- TANAKA, I., APPELT, K., DIJK, J., WHITE, S.W. and WILSON, K.S. (1984)
3-A resolution structure of a protein with histone-like properties in prokaryotes.
Nature 310, 376-381.
- THOMM, M., STETTER, K.O. and ZILLIG, W. (1982)
Histone-like proteins in eu- and archaebacteria.
Zbl. Bakt. Hyg. I, Abt. Orig. C3, 128-139.
- VARSHAVSKY, A.J., NEDOSPASOV, S.A., BAKAYEV, V.V., BAKAYEVA, T.G. and GEORGIEV, G.P. (1977)
Histone-like proteins in the purified *Escherichia coli* deoxyribonucleoprotein.
Nucleic Acids Research 4, 2725-2745.
- WILLEKENS, P., HUYSMANS, E., VANDENBERGHE, A. and DE WACHTER, R. (1986)
Archaeobacterial 5 S ribosomal RNA : nucleotide sequence in two methanogen species, secondary structure models and molecular evolution.
System. Appl. Microbiol. 7, 151-159.
- WOESE, C. (1981)
Les archaebactéries.
Pour la Science 46, 56-69.
- WOESE, C.R. and FOX, G.E. (1977)
Phylogenetic structure of the prokaryotic domain : the primary kingdoms.
Proc. Natl. Acad. Sci. USA 74, 5088-5090.
- WOESE, C.R. and OLSEN, G.J. (1986)
Archaeobacterial phylogeny : perspectives on the urkingdoms.
System. Appl. Microbiol. 7, 161-177.
- WOESE, C.R., PACE, N.R. and OLSEN, G.J. (1986)
Are arguments against archaebacteria valid ?
Nature 320, 401-402.
- YAMAZAKI, K., NAGATA, A., KANO, Y. and IMAMOTO, F. (1984)
Isolation and characterization of nucleoid proteins from *Escherichia coli*.
Mol. Gen. Genet. 196, 217-224.
- ZILLIG, W., STETTER, K.O. and JANEKOVIC, D. (1979)
DNA-dependent RNA-polymerase from the archaebacterium *Sulfolobus acidocaldarius*.
Eur. J. Biochem. 96, 597, 604.

CHAPITRE II**PRÉPARATION, PURIFICATION ET PROPRIÉTÉS PHYSICO-CHIMIQUES**
DE LA PROTÉINE MC1

Au laboratoire notre équipe s'est intéressée à l'étude de la protéine MCl isolée de quatre bactéries méthanogènes dont une souche mésophile *Methanosarcina barkeri* MS et trois souches thermophiles *Methanosarcina thermophila* TM1, *Methanosarcina* sp CHTI et *Methanosarcina* sp MST-A1. Notre travail a consisté en l'étude de la protéine MCl isolée des bactéries *M.barkeri* MS et *Methanosarcina* sp MST-A1 dont nous présentons les principales caractéristiques.

II-1 CARACTÉRISTIQUES GÉNÉRALES DES BACTÉRIES MÉTHANOGÈNES:

M.BARKERI ET METHANOSARCINA SP MST-A1

Ces bactéries sont des sarcines c'est-à-dire qu'elles ne vivent jamais isolées, elles sont collées entre elles par une sorte de ciment (Figure II-1), ce qui a pour conséquence de donner en culture des agrégats qui sédimentent. Les principaux caractères de ces bactéries sont présentés dans le tableau II-1. Elles se différencient principalement par la température à laquelle s'effectue leur croissance et par la nature du substrat de croissance : ainsi, seule *M.barkeri* est capable d'utiliser le mélange H₂/CO₂. Des hybridations ADN-ADN entre ces deux souches, récemment réalisées par J.P.TOUZEL (INRA) confirment que ces deux bactéries ne sont pas apparentées.

Par contre les deux bactéries *Methanosarcina thermophila* TM1 et *Methanosarcina* sp CHTI présentent les mêmes caractères que la bactérie *Methanosarcina* sp MST-A1 (même température de croissance, même substrat de croissance et même (G + C)%). Néanmoins la bactérie *Methanosarcina* sp MST-A1 se différencie des deux premières par l'aspect macroscopique.

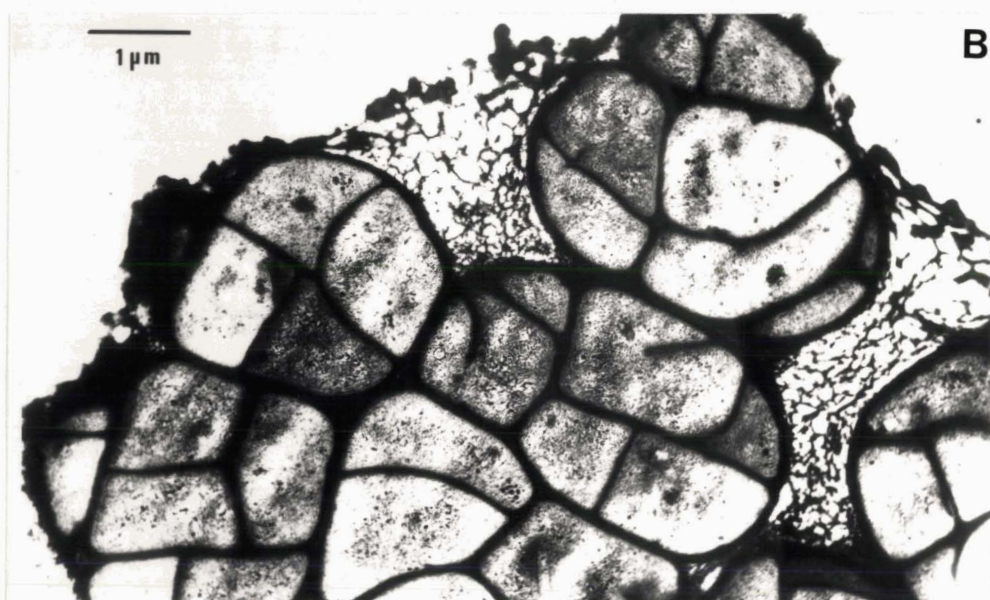
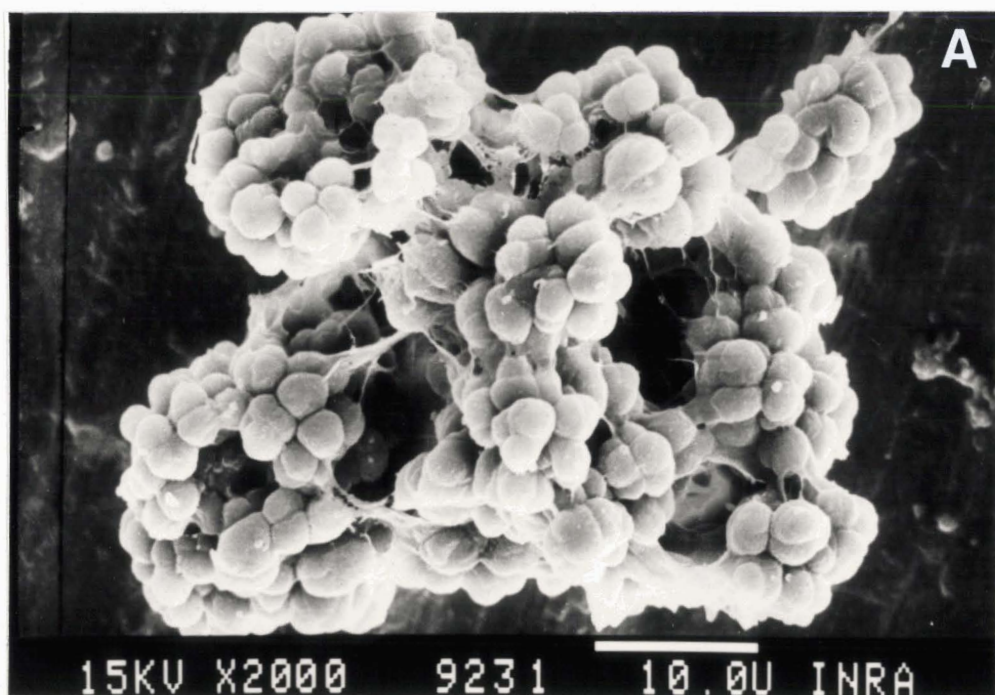


FIGURE II-1 - Sarcines de *Methanosarcina*.

A : Vue générale au microscope électronique à balayage

B : Vue en coupe au microscope électronique à transmission

TABLEAU II-1

Caractéristiques générales de deux bactéries méthanogènes

Nom	Morphologie	Aspect en culture	Température de croissance	Substrat de croissance	(G + C) %
<i>Methanosarcina barkeri</i>	Sarcines	Agrégats irréguliers	35°C	H ₂ /CO ₂ -acétate méthanol méthylamines	39
<i>Methanosarcina</i> sp MST-A1	Sarcines	Agrégats réguliers	55°C	acétate méthanol méthylamines	39

II - 2 PRÉPARATION DE LA PROTÉINE MCl

II - 2-1 PRINCIPE DE LA METHODE

La protéine MCl est préparée par chromatographie d'affinité sur une colonne d'ADN couplé à la cellulose selon la méthode décrite par ALBERTS et HERRICK (1971). Dans un premier temps, l'ADN de l'homogénat bactérien est éliminé par précipitation avec le polyéthylène glycol (PEG 6000) après dissociation du complexe nucléoprotéique à force ionique élevée (2 M NaCl). La force ionique est alors abaissée et les protéines qui ont une affinité pour l'ADN sont fixées sur la colonne d'où elles sont désorbées par un gradient de force ionique croissante.

La molarité en NaCl (exprimée en mmoles) des tampons utilisés lors de la chromatographie est indiquée par un chiffre qui suit la lettre P. Les conditions expérimentales sont présentées dans la partie technique (p. 151)

II - 2-2 RESULTATS

II - 2-2-1 Préparation de la protéine MCl de *Methanosarcina barkeri* souche MS

Les premières préparations de la protéine MCl ont été effectuées sur une colonne d'ADN d'*Escherichia coli*. Le diagramme d'éluion ainsi obtenu est présenté dans la figure II-2A. La protéine MCl est éluee dans la fraction 3 à une concentration en NaCl comprise entre 250 et 600 mM. La fraction 3 a été divisée en 3 sous-fractions notées a, b et c dont l'analyse électrophorétique sur gel de polyacrylamide en milieu SDS est présentée figure II-2B. La protéine MCl a également été préparée

FIGURE II-2A - Chromatographie sur colonne d'ADN d' *E.coli* d'un homogénat de *Methanosarcina barkeri*.

Les fractions 1, 2 et 4 sont respectivement éluées par les tampons P-150, P-250 et P-2000. La fraction 3 est éluée par un gradient linéaire de molarité en NaCl (250 mM - 600 mM) obtenu avec les tampons P-250 et P-600 (2 x 80 ml). Le débit de la colonne est de 40 ml/h et des fractions de 6,6 ml sont collectées.

FIGURE II-2B - Electrophorèse sur gel de polyacrylamide en milieu SDS de la fraction 3 de la figure II-2A.

M : témoins de masse molaire connue : phosphorylase b (94 000), albumine sérique de boeuf (67 000), ovalbumine (43 000), anhydrase carbonique (30 000), inhibiteur trypsique de soja (20 000) et α -lactalbumine (14 400).

Les couloirs 3a, 3b, 3c correspondent respectivement aux sous-fractions 3a, 3b, 3c de la figure II-2A.

L'électrophorèse est effectuée selon la technique de LAEMMLI (1970). Les échantillons (5 μ g) sont traités avec 5 p.100 de SDS, 2 p.100 de 2-mercaptoéthanol pendant 2 minutes à 100°C. L'électrophorèse est effectuée sous une intensité de 40 mA pendant 150 minutes sur une plaque de gel de polyacrylamide (120x 70x1 mm). La concentration d'acrylamide dans le gel varie linéairement de 5 à 30 p.100. Le tampon d'électrode (Tris 16 mM/glycine 128 mM ajusté à pH 8,3 avec HCl) et le tampon de gel (Tris 0,75 M/HCl, pH 8,9) contiennent 0,1 p.100 de SDS.

Le gel est coloré et décoloré selon la méthode décrite par STEPHANO et coll. (1986).

FIGURE II-3 - Electrophorèse sur gel de polyacrylamide en milieu SDS de la fraction 3 éluée d'une colonne d'ADN de *Rhizobium meliloti*.

M : témoins de masse molaire connue (voir figure II-2B). Les dépôts 3a, 3b et 3c correspondent respectivement aux sous-fractions 3a, 3b et 3c éluées de la colonne d'ADN de *R.meliloti*.

Fig. II2A

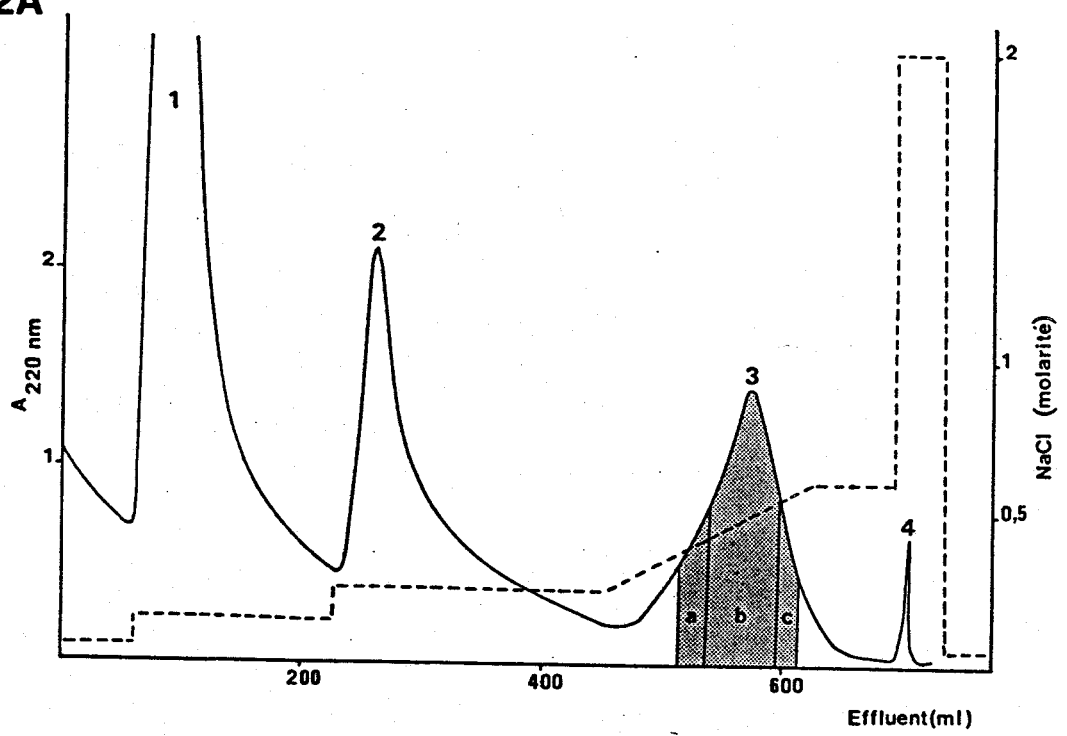
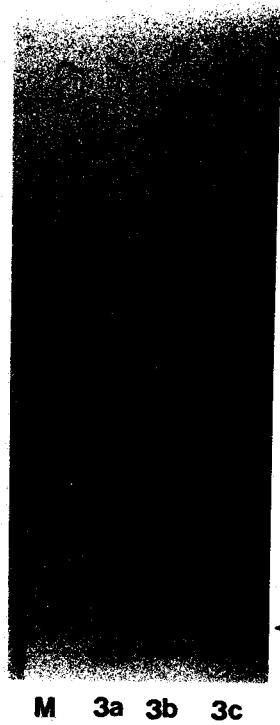


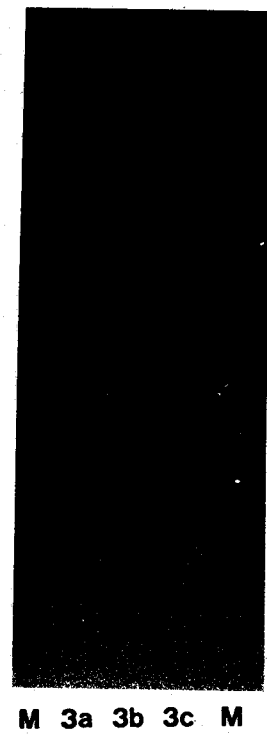
Fig. II2B



↓

+

Fig. II3



sur une colonne d'ADN de *Rhizobium meliloti* éluée par un gradient linéaire de 250 mM à 500 mM NaCl et nous avons obtenu un diagramme d'élution semblable à celui présenté dans la figure II-2A. Dans ce cas le nombre et la quantité de contaminants sont fortement réduits (Figure II-3). De plus, on ne retrouve pas dans les sous-fractions 3a et 3b le contaminant qui migre légèrement plus loin que la protéine MCl et qui est difficile à éliminer lors des étapes de purification en milieu non dénaturant. Cette colonne présente donc un très net avantage pour la suite de nos travaux.

II - 2-2-2 Préparation de la protéine MCl de *Methanosarcina* sp MST-A1

La protéine MCl de *Methanosarcina* sp MST-A1 a été préparée sur une colonne d'ADN d'*E.coli*, le diagramme d'élution obtenu est présenté figure II-4A. La figure II-4B montre le diagramme électrophorétique de la fraction 3 dans laquelle la protéine MCl est éluée avec une force ionique voisine de 470 mM NaCl.

II - 3 PURIFICATION DE LA PROTÉINE MCl

Les lots de protéine MCl utilisés pour la détermination de la structure primaire ont été purifiés par chromatographie de gel filtration en milieu dénaturant (HCl 0,01 M). Comme les premières études conformationnelles suggéraient que la protéine MCl avait une structure désorganisée, nous avons été amenée à mettre au point une méthode de purification dans des conditions non dénaturantes.

FIGURE II-4A - Chromatographie sur colonne d'ADN d' *E.coli* d'un homogénat de *Methanosarcina* sp souche MST-A1.

Les fractions 1, 2 et 4 sont respectivement éluées par les tampons P-150, P-300 et P-2000. La fraction 3 est éluée par un gradient linéaire de molarité en NaCl (300 mM - 600 mM) obtenu avec les tampons P-300 et P-600 (2 x 80 ml). Le débit de la colonne est de 40 ml/h et des fractions de 6,6 ml sont collectées.

FIGURE II-4B - Electrophorèse sur gel de polyacrylamide en milieu SDS de la fraction 3 de la figure II-4A.

M : témoins de masse molaire connue (voir figure II-2B). Les dépôts 3a, 3b et 3c correspondent respectivement aux sous-fractions 3a, 3b et 3c de la figure II-4A.

Fig. II4A

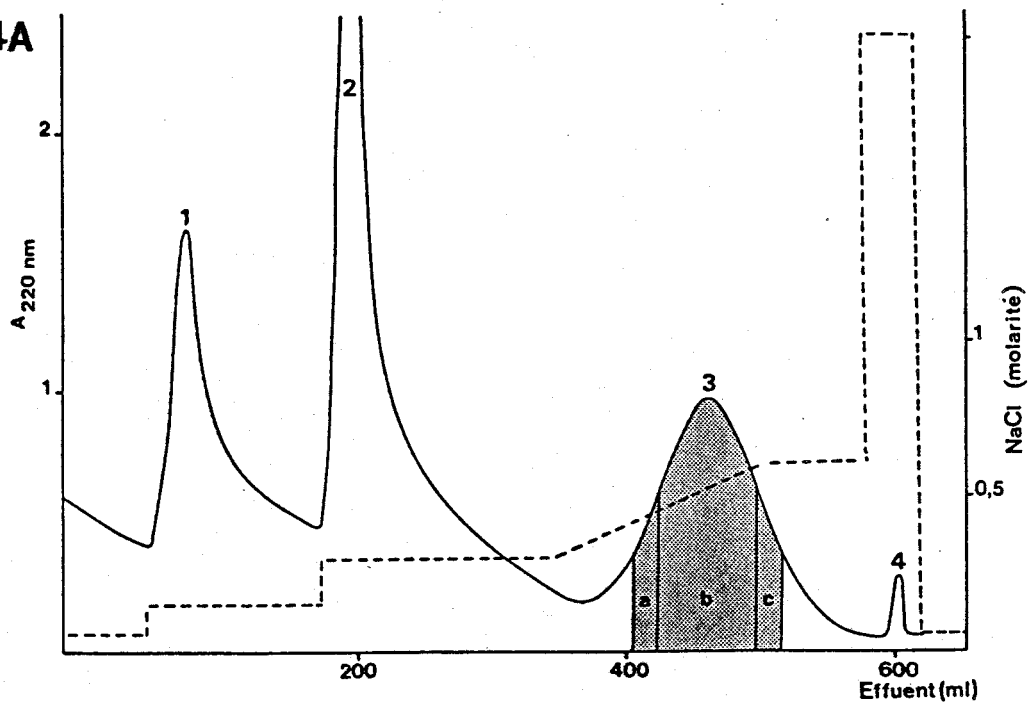
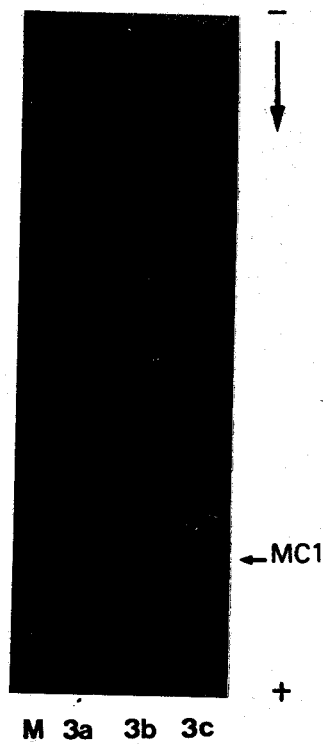


Fig. II4B



II - 3-1 PURIFICATION EN MILIEU DENATURANT

Les fractions qui contiennent de la protéine MCl sont concentrées par ultrafiltration sur une membrane AMICON YM5. La protéine est alors mise en présence de chlorhydrate de guanidine 6 M, 2-mercaptoéthanol 10 mM pendant 48 heures avant d'être déposée sur une colonne de Sephadex G-75 éluée par HCl 0,01 M. L'analyse électrophorétique en milieu SDS (Figure II-5) et en milieu acide acétique 0,9 M en présence d'urée 2,5 M (Figure II-6) montre que les protéines ainsi obtenues sont pures. Notons que les deux bandes de plus faible migration observées dans le diagramme électrophorétique de la protéine MCl de *Methanosarcina* sp MST-A1 (couloir 1 de la figure II-5) correspondent à des polymères obtenus par séchage de la protéine avant le dépôt sur le gel : ces deux bandes réagissent avec le sérum anti-protéine MCl de *M. barkeri* MS.

Des calculs de rendement, nous ont permis de mettre en évidence qu'une quantité importante de protéine s'adsorbait sur la membrane YM5 lorsque la concentration en protéine atteint 1 mg/ml. De plus, une turbidité importante était observée. Remarquons que la turbidité n'apparaît pas lors de la concentration d'une fraction de protéine MCl pure, ce qui suggère qu'elle est provoquée par les composants contaminants élués en même temps que la protéine MCl dans la fraction 3. Par la suite, nous avons concentré les fractions 3 par évaporation sur un évaporateur de type "Speed Vac" en présence de chlorhydrate de guanidine pour éviter l'agrégation. La quantité de cet agent dissociant est calculée de telle sorte que sa molarité soit égale à 6 M lorsque le volume est réduit à 2 ml. Ce protocole nous a permis d'augmenter les rendements notamment pour la protéine MCl de *Methanosarcina* sp MST-A1 qui s'agrégeait dès le début de la concentration.

FIGURE II-5 - Electrophorèse sur gel de polyacrylamide en milieu SDS, de la protéine MCl éluée d'une colonne de Sephadex G-75.

M : témoins de masse molaire connue (figure II-2B)

1 : protéine MCl de *Methanosarcina* sp souche MST-A1

2 : protéine MCl de *Methanosarcina barkeri* souche MS

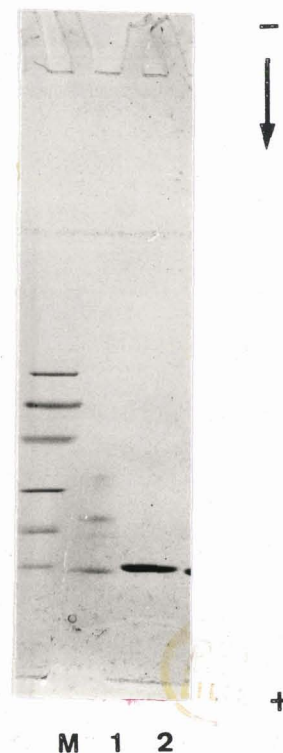


FIGURE II-6 - Electrophorèse sur gel de polyacrylamide en milieu acide acétique 0,9 M et en présence d'urée 2,5 M, de la protéine MCl éluée d'une colonne de Sephadex G-75

1 et 2 : légende de la figure II-5

L'électrophorèse est effectuée selon la technique de PANYIM et CHALKLEY (1969). Elle est réalisée sur une plaque de gel contenant 17 p.100 d'acrylamide.

Les échantillons sont dissous dans HCl 0,01 M, contenant de l'urée 8 M et du 2-mercaptoéthanol 0,5 M. L'électrophorèse est effectuée sous une intensité de 22 mA pendant 3 heures. Les gels sont colorés et décolorés suivant le protocole décrit dans la légende de la figure II-2B.



II - 3-2 PURIFICATION EN MILIEU NON DENATURANT

Les premiers spectres dichroïques obtenus sur la protéine MC1 de *M. barkeri* indiquaient l'absence de structure α -hélicoïdale, ce qui nous a conduit à penser que le traitement en milieu acide et en présence de chlorhydrate de guanidine pouvait dénaturer irréversiblement la protéine. La recherche d'une méthode de concentration et de purification excluant les traitements dénaturants s'imposait.

La concentration de la protéine en sortie de la colonne d'ADN a été l'étape la plus difficile à franchir compte tenu des faits suivants :

- 1) L'ultrafiltration provoque des phénomènes d'adsorption et d'agrégation de la protéine.
- 2) Lyophilisée, la protéine purifiée par gel filtration sur Sephadex G-75 éluée en HCl 0,01 M présente toujours plusieurs bandes correspondant à des polymères après électrophorèse en présence de SDS. Il y aurait donc agrégation.
- 3) La précipitation de la protéine MC1 par l'acétone (à pH 2,0 et à -20°C pendant 48 heures) n'a pas toujours été observée. Cette méthode est couramment utilisée pour la précipitation des histones et de la protéine II d'eubactéries : les protéines obtenues dans ces conditions ne sont pas dénaturées.
- 4) Lors d'essais de précipitation au sulfate d'ammonium directement sur la fraction 3, une grande partie de la protéine reste soluble à 95 p.100 de saturation ce qui s'explique en fait par une concentration en protéine trop faible.

Nous avons donc introduit une étape supplémentaire de concentration par chromatographie d'échange d'ions sur une colonne de carboxyméthyl-TSK (CM-TSK). Les conditions expérimentales utilisées pour concentrer la protéine MC1 sur cette colonne sont présentées dans la partie technique (p.46). La colonne est éluée par le tampon CM-1000 à pH 6,0. Cette valeur de pH a été choisie dans le but d'éviter l'agrégation.

En définitive, la méthode de concentration et de purification de la protéine MC1 est schématisée dans le tableau II-2. Cette méthode consiste en 4 étapes dont la première et la quatrième sont en fait des concentrations de la protéine sur la colonne de CM-TSK. La seconde étape consiste en une précipitation au sulfate d'ammonium à 80 p.100 de saturation. La totalité des contaminants est éliminée dans la fraction insoluble alors que peu de protéine MC1 précipite (20 % au maximum). La protéine est obtenue pure dans la fraction soluble (Figure II-7). L'élimination du sulfate d'ammonium est effectuée par chromatographie sur Sephadex G-25.

Bien que cette méthode donne un rendement en protéine deux fois moins important que celui obtenu avec la méthode de purification dans des conditions dénaturantes (2 mg de protéine MC1 sont préparés à partir de 100 g de bactéries humides), elle présente les avantages suivants :

- méthode douce, facteur indispensable pour les études de conformation : en effet le sulfate d'ammonium est couramment utilisé pour préparer des anticorps qui sont reconnus comme étant des molécules "fragiles".

- possibilité de traiter de grands volumes et obtention de solutions protéiques restant limpides à des concentrations atteignant 7 mg/ml.

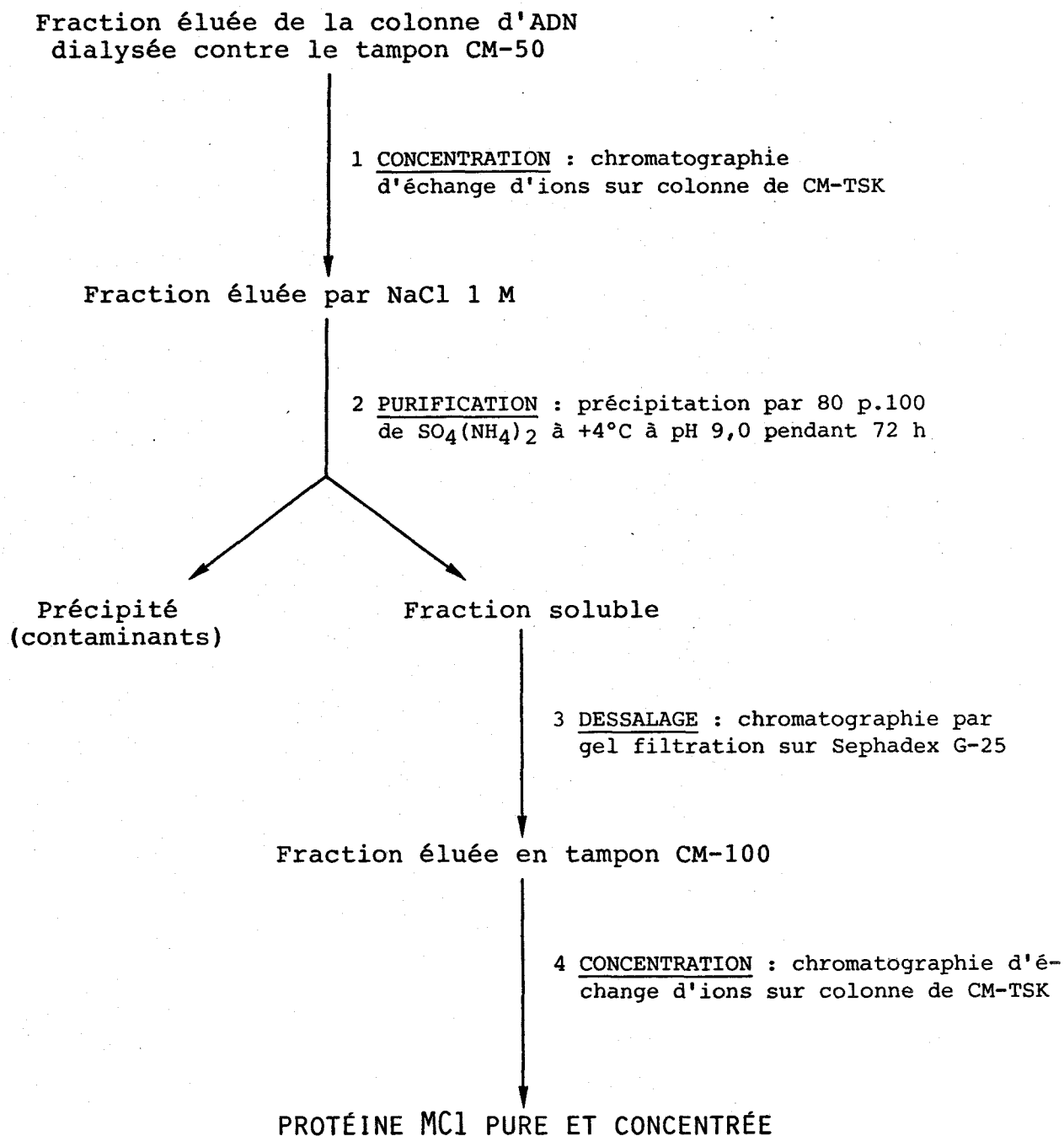
II - 4 PROPRIÉTÉS PHYSICO-CHIMIQUES DE LA PROTÉINE MC1

II - 4-1 SOLUBILITE

La protéine MC1 est acido-soluble (HCl 0,01 M, HCl 0,25 M). Elle est aussi soluble dans l'eau : une solution de protéine à 1 mg/ml a une absorbance nulle à 340 nm et donc ne présente pas de turbidité.

TABLEAU II-2

Protocole utilisé pour purifier et concentrer la protéine MC1 dans des conditions non dénaturantes



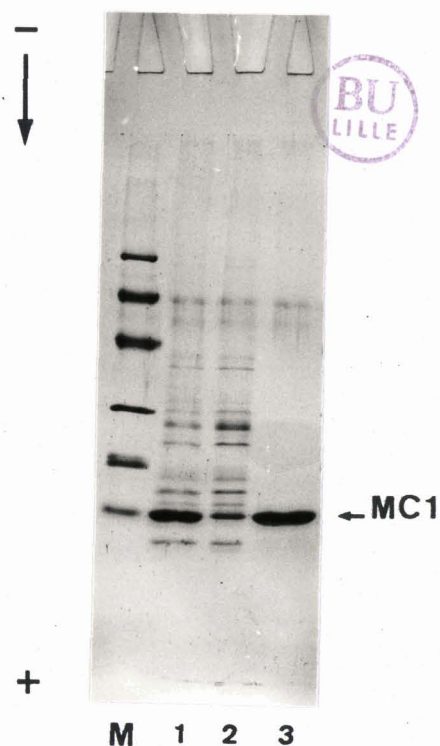


FIGURE II-7 - Electrophorèse sur gel de polyacrylamide en milieu SDS des fractions obtenues au cours de la purification au sulfate d'ammonium.

- M : témoins de masse molaire connue (voir figure II-2B)
 1 : fraction 3b éluée de la colonne d'ADN d'*E.coli*
 2 : fraction insoluble à 80 p.100 de saturation en sulfate d'ammonium
 3 : fraction soluble à 80 p.100 de saturation en sulfate d'ammonium

REMARQUE : Les 2 bandes de faible migration correspondent à des contaminants contenus dans le tampon échantillon puisque ces 2 bandes ont été trouvées dans les 10 couloirs que contenait ce gel.

A une concentration égale à 1 mg/ml, la précipitation de la protéine MC1 n'est pas totale à 95 p.100 de saturation de sulfate d'ammonium ce qui souligne le caractère hydrophile de la protéine.

II - 4-2 COMPORTEMENT ELECTROPHORETIQUE

Le comportement électrophorétique de la protéine MC1 en gel de polyacrylamide a été étudié dans deux systèmes différents. Les résultats ont été présentés dans le paragraphe II - 3-1 et sont rassemblés dans la figure II-8.

La protéine MC1 a, en gel de polyacrylamide en présence de SDS, une masse moléculaire relative apparente de 14.400 (Figure II-8A) or sa masse moléculaire relative calculée d'après la séquence en acides aminés est de 10.757 (publication n°1, chapitre III). Cette migration électrophorétique anormale est due à une teneur élevée en résidus basiques. Un comportement électrophorétique anormal est observé aussi pour l'histone H4 de veau. Cette protéine de M_r 11.000 a une masse moléculaire apparente de 14.400 (Figure II-9).

En milieu acide et en présence d'urée 2,5 M, les protéines MC1 de *M.barkeri* et de *Methanosarcina* sp MST-A1 n'ont pas la même migration électrophorétique (Figure II-8B). La protéine MC1 de *Methanosarcina* sp MST-A1 a une mobilité cathodique légèrement inférieure à celle de la protéine MC1 de *M.barkeri* ce qui suggère que la protéine MC1 de *Methanosarcina* sp MST-A1 est moins basique.

FIGURE II-8 - Comportement électrophorétique de la protéine MC1 de *M.barkeri* et *Methanosarcina* sp MST-A1.

II-8A : en présence de SDS

M : témoins de masse molaire connue (Figure II-2B)

1 : protéine MC1 de *Methanosarcina* sp souche MST-A1

2 : protéine MC1 de *Methanosarcina barkeri* souche MS

II-8B : en milieu acide (pH 2,7) et en présence d'urée 2,5 M

1 et 2 : légende de la figure 8A

FIGURE II-9 - Comportement électrophorétique en présence de SDS de la protéine MC1 de *Methanosarcina barkeri* et de l'histone H4

M : témoins de masse molaire connue (Figure II-2B)

1 : protéine MC1 de *Methanosarcina barkeri*

2 : histone H4 de thymus de veau

Fig. 8A

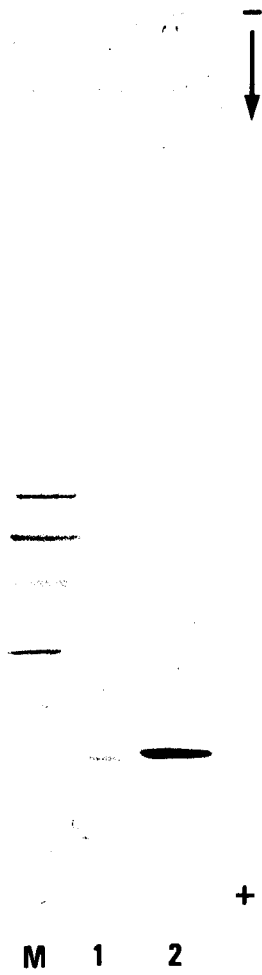


Fig. 8B

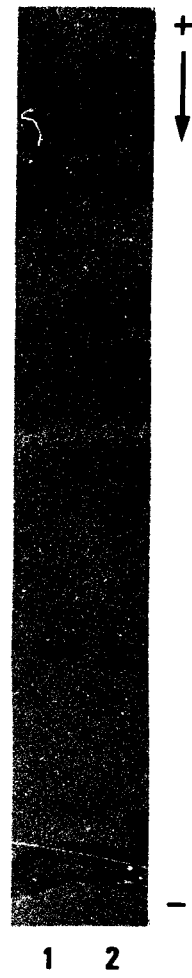
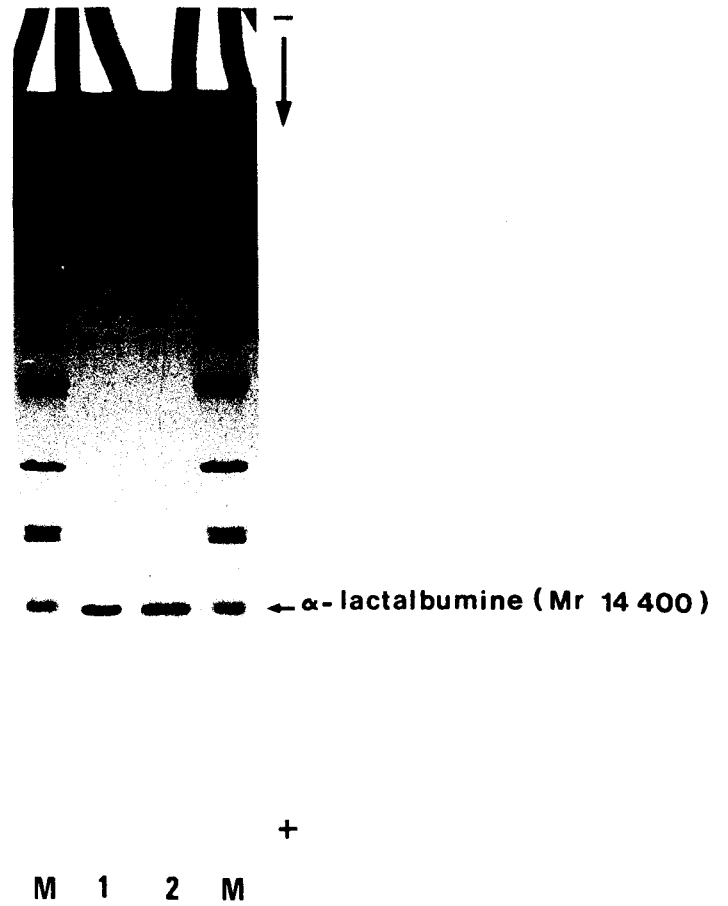


Fig. 9



II - 4-3 DETERMINATION DU POINT ISOELECTRIQUE (pHi) DE LA PROTEINE MCl DE *M.BARKERI* MS

La détermination du pHi a été effectuée par isoélectrofocalisation sur plaque de gel de polyacrylamide. Afin d'éviter une oxydation éventuelle des résidus de méthionine se traduisant par l'observation de trois bandes sur le gel d'isoélectrofocalisation, la protéine a été isolée et purifiée en présence de dithiothréitol (DTT) 0,5 mM. Avant d'être déposé sur le gel, l'échantillon de protéine est dessalé par dialyse contre l'eau distillée contenant du DTT 0,5 mM. Dans ces conditions, nous observons une bande correspondant à une valeur de pHi voisine de 9,7 (Figure II-10).

II - 4-4 COMPORTEMENT A LA LYOPHILISATION

Au cours de nos travaux nous avons remarqué que chaque fois que la protéine MCl était lyophilisée en milieu acide (HCl 0,01 M), elle apparaissait en gel de polyacrylamide en milieu SDS sous la forme de trois bandes correspondant à des polymères (Figure II-11). Ce phénomène ne s'observe pas lorsque la protéine est lyophilisée dans de l'eau distillée, de l'eau lourde (utilisée pour les études en infrarouge) ou du tampon phosphate de sodium 10 mM, NaCl 50 mM pH 6,0.

II - 4-5 STRUCTURE QUATERNAIRE DE LA PROTEINE MCl

La structure quaternaire de la protéine MCl a été déterminée par chromatographie de gel filtration sur Sephadex G-75 en tampon phosphate 10 mM pH 6,0, en présence de 20 mM NaCl et en

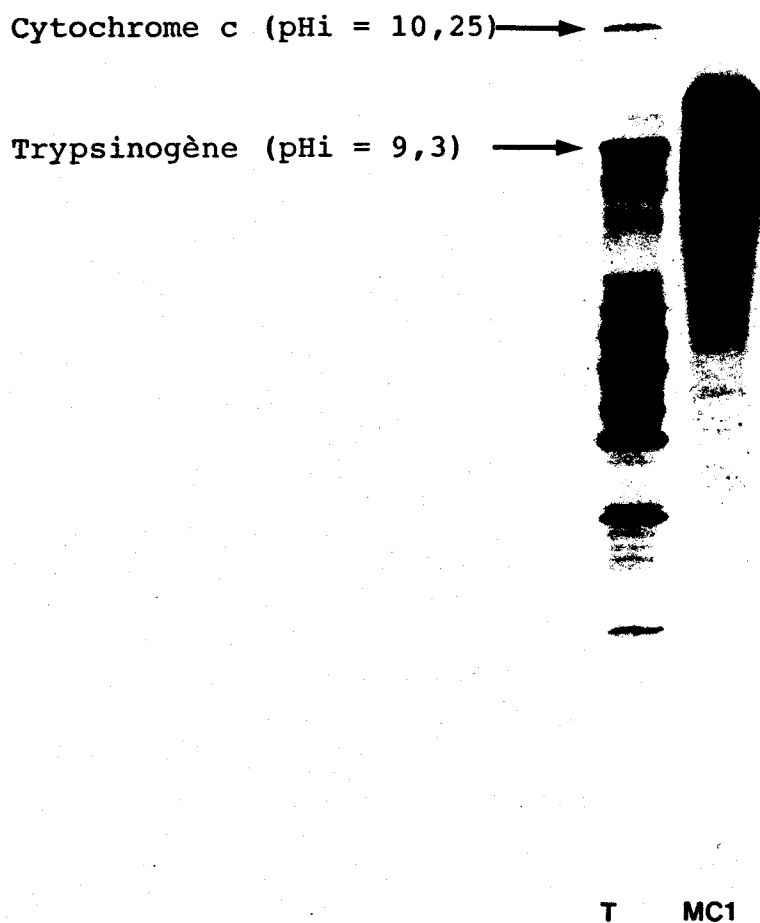


FIGURE II-10 - Etude en isoélectrofocalisation en plaque de gel de polyacrylamide de la protéine MC1 de *M.barkeri*.

T : témoins de pHi (PHARMACIA)



FIGURE II-11 - Etude par électrophorèse sur gel de polyacrylamide en milieu SDS du comportement de la protéine MCl de *M.barkeri* MS après lyophilisation.

M : témoins de masse molaire connue (Figure II-2B)

1 : protéine MCl de *Methanosarcina barkeri* après lyophilisation dans l'eau

2 : protéine MCl de *Methanosarcina barkeri* MS après lyophilisation (ou séchage) en milieu acide (HCl 0,01 M).

présence de 400 mM NaCl. Cette étude est fondée sur les modifications de conformation de la protéine en fonction de la force ionique (Chapitre IV).

Les conditions de chromatographie sont décrites dans l'appendice technique p.156 . Les protéines de référence ont été choisies parmi les protéines chromosomales : la protéine II d'*Escherichia coli* dont la forme fonctionnelle est le dimère, l'histone H5 et le complexe formé par les histones H2A et H2B.

A force ionique faible ou élevée (20 mM NaCl et 400 NaCl), la masse moléculaire apparente de la protéine MCl est voisine de 11.000 (Figure II-12) ce qui correspond à la masse moléculaire relative déterminée à partir de la séquence en acides aminés (M_r 10.757). La protéine se comporte donc comme un monomère.

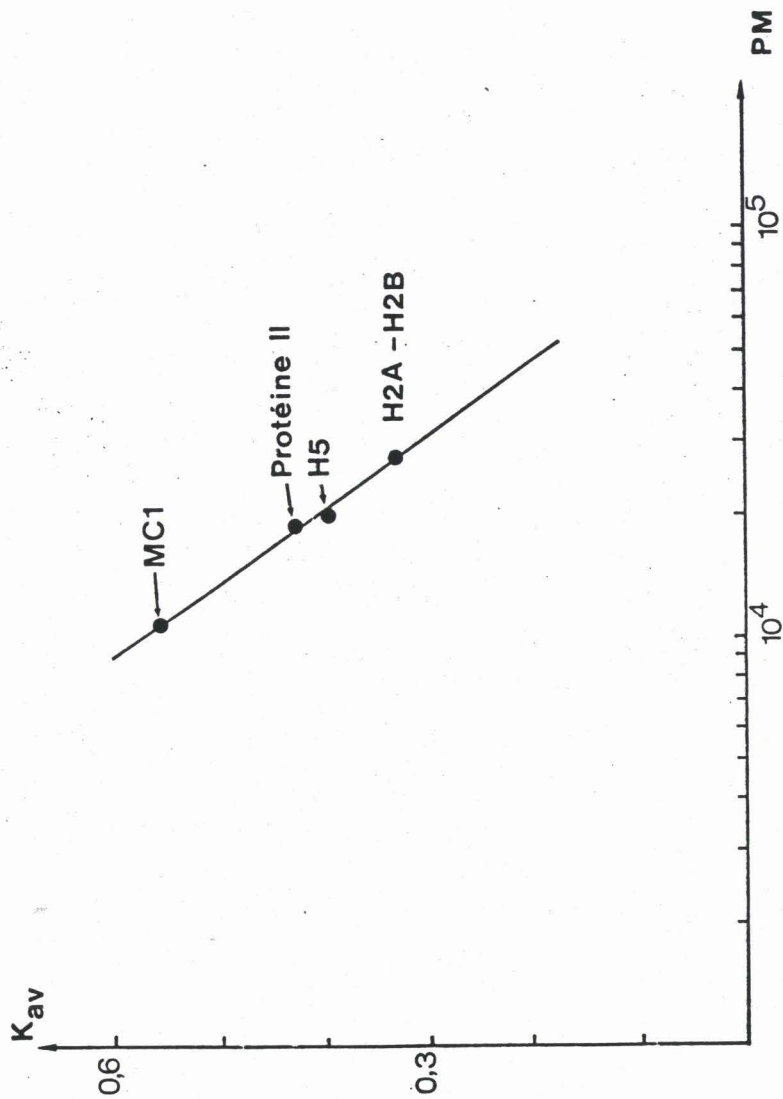


FIGURE II-12 - Détermination de la structure quaternaire de la protéine MC1 par chromatographie de gel filtration sur Sephadex G-75 en tampon phosphate 10 mM pH 6,0, 20 mM NaCl.

PM : poids moléculaire ; K_{av} (voir appendice technique).

CHAPITRE III

DÉTERMINATION DE LA STRUCTURE PRIMAIRE DE LA PROTÉINE MCI DE *METHANOSARCINA BARKERI* MS ET DE *METHANOSARCINA* SP MST-A1

Les travaux concernant la détermination de la séquence en acides aminés de la protéine de *Methanosarcina barkeri* MS, ont fait l'objet des publications suivantes :

Publication n°1

Primary structure of the chromosomal protein Hmb from the archaeobacterium *Methanosarcina barkeri*.
LAINE, B., CHARTIER, F., IMBERT, M., LEWIS, R. and SAUTIERE, P. (1986)
Eur. J. Biochem. 161, 681-687.

Publication n°2

Features of the primary structure of the *Methanosarcina barkeri* chromosomal protein Hmb compared to those of eukaryotic and prokaryotic chromosomal proteins.
IMBERT, M., CHARTIER, F., LAINE, B. and SAUTIERE, P. (1986)
in Biology of Anaerobic Bacteria (DUBOURGUIER, H.C. *et al*, eds)
pp. 236-241, Elsevier Science Publishers B.V., Amsterdam, The Netherlands.

Eur. J. Biochem. 161, 681–687 (1986)
© FEBS 1986

Primary structure of the chromosomal protein HMB from the archaeobacteria *Methanosarcina barkeri*

Bernard LAINE¹, François CHARTIER¹, Marlène IMBERT¹, Randy LEWIS² and Pierre SAUTIERE¹

¹ Unité Associée 409 du Centre National de la Recherche Scientifique, Lille

² Department of Biochemistry, University of Wyoming, Laramie

(Received July 2/September 5, 1986) – EJB 86 0703

The amino acid sequence of the protein HMB, a protein of 93 residues (M_r 10757) which represents the major acid-soluble component of the *Methanosarcina barkeri* nucleoprotein complex, has been established from automated sequence analysis of the protein and from structural data provided by peptides derived from cleavage of the protein at aspartic acid, arginine and methionine residues.

The protein HMB is mainly characterized by a high amount of charged residues (15% of acidic residues and 26.8% of basic residues) which are distributed all along the polypeptide chain. The amino acid sequence of the protein HMB is not homologous to any eubacterial, archaeobacterial or eukaryotic chromosomal proteins known up to now.

Methanogens constitute with thermoacidophiles and halophiles a group of prokaryotic organisms called archaeobacteria [1]. Although some macromolecules in archaeobacteria exhibit physico-chemical properties similar to those found in eubacteria or in eukaryotes, others such as membrane lipids, cell wall peptidoglycans or enzymes have unique structural features which place archaeobacteria in a distinct line of descent [1, 2]. The fundamental differences observed in the packaging of the DNA between eubacteria and eukaryotes result from the nature of their chromosomal proteins. Packaging of the DNA inside the nucleus of eukaryotic cells is mainly ensured by histones whereas in eubacteria, HU-type proteins (M_r 9500) are involved in DNA condensation [3–5]. It is therefore of great interest to investigate the features of chromosomal proteins from archaeobacteria. Chromosomal proteins have been isolated from various archaeobacteria: three strains of thermoacidophiles, *Thermoplasma acidophilum* [6], *Sulfolobus acidocaldarius* [7] and *Sulfolobus solfataricus* [8] and one strain of methanogens *Methanosarcina barkeri* [9]. Only two of these proteins have been sequenced: on the criteria of their primary structure, the protein HTa from *T. acidophilum* [10] is homologous to the HU-type proteins [11] whereas the protein (M_r 7000) from *S. solfataricus* [8] is related neither to histones nor to HU-type proteins.

The protein HMB is a low-molecular-mass polypeptide (M_r 10757) which has been shown to be the major acid-soluble component of the *Methanosarcina barkeri* nucleoprotein complex [9]. We report here the complete amino acid sequence of this protein. The primary structure of protein HMB is not homologous to any another chromosomal protein from prokaryotes or eukaryotes.

Correspondence to B. Laine, Unité Associée CNRS 409, Place de Verdun, F-59045 Lille Cedex, France

Abbreviations. SDS, sodium dodecyl sulfate; HPLC, high-pressure liquid chromatography.

Enzymes. Carboxypeptidase A (EC 3.4.17.1); carboxypeptidase B (EC 3.4.17.2); carboxypeptidase Y (EC 3.4.16.1); endoproteinase Arg-C (EC 3.4. —. —).

MATERIALS AND METHODS

Materials

Methanosarcina barkeri (strain DSM 800) was grown as indicated in [9]. The protein HMB was isolated as in [9]. Endoproteinase Arg-C from mouse submaxillary gland and carboxypeptidase Y were purchased from Boehringer. Carboxypeptidase A treated with phenylmethylsulfonyl fluoride and carboxypeptidase B treated with diisopropyl-fluorophosphate were from Sigma. Cyanogen bromide was from Eastman Kodak (Rochester, USA). Acetonitrile for reverse-phase high-pressure liquid chromatography was from Carlo Erba. Reagents and solvents for the liquid-phase sequencer were as indicated in [12]. Reagents and solvents for the gas-phase sequencer were from Applied Biosystems.

Cleavage at aspartyl residues

Protein HMB (200 nmol) was hydrolysed with 1 ml of 0.25 M acetic acid in an evacuated sealed tube for 15 h at 105°C. The hydrolysate was freeze-dried prior to fractionation.

Cleavage with the endoproteinase Arg-C

Protein HMB (300 nmol) was dissolved in 2.2 ml of 0.1 M sodium bicarbonate pH 7.9 and digested with endoproteinase Arg-C at 37°C for 4 h. Enzyme was added at the beginning of the reaction and after 2 h; the final enzyme/substrate ratio was 1/50 (w/w). The digestion was stopped by lowering the pH with formic acid. The hydrolysate was freeze-dried prior to fractionation.

Cleavage at methionyl residues

Protein HMB (1 μ mol) was dissolved in 4 ml of 70% (v/v) formic acid. Cyanogen bromide was added in a 100-fold molar excess over one methionine residue. The reaction mixture was

flushed with N₂ and incubated for 24 h under gentle stirring at 37°C in the dark. The reaction mixture was then evaporated in a vacuum concentrator (Savant Instruments) to 0.7 ml; the concentration was repeated twice after washing with 1 ml water. The peptide solution was made 6 M in guanidine hydrochloride and submitted to fractionation.

Fractionation of peptides

Peptides generated by cleavage with 0.25 M acetic acid or with the endoproteinase Arg-C were fractionated by reverse-phase high-pressure liquid chromatography as in [11], on a C₁₈ μ Bondapak column (Waters Associates) and using gradient systems indicated in Fig. 1.

Peptides generated by cleavage with cyanogen bromide were fractionated by chromatography on a Sephadex G-50 superfine column (95 \times 1.6 cm) equilibrated and eluted with 0.01 M HCl saturated with chloroform.

Amino acid analyses

Peptide samples were hydrolysed in evacuated sealed tubes at 110°C for 24 h in 6 M HCl.

The hydrolysates of peptides generated by cleavage with cyanogen bromide were treated with a pyridine acetate buffer (pyridine/glacial acetic acid/water, 25:1:224; pH 6.5) at 105°C for 1 h in order to convert homoserine lactone to homoserine.

Amino acid analyses were performed on a Beckman 119 CL amino acid analyser, using a single-column system. The column was packed with a Durrum DC 6A resin.

Carboxy-terminal analyses

The carboxy-terminal sequence of protein HMB and of the carboxy-terminal peptide obtained by cleavage at methionine residues (peptide CN-3) were determined simultaneously. Each substrate (28 nmol) was successively digested with carboxypeptidases Y, A and B. After every step, an 8-nmol aliquot of each digest was acidified with formic acid and immediately freeze-dried. The released amino acids were analysed on the amino acid analyser. Digestion by carboxypeptidase Y (56 μ g) was performed for 7 h at 37°C in 0.1 M acetate pyridine buffer (0.35 ml) adjusted at pH 5.5. The carboxypeptidase Y was then denatured in a boiling water bath for 2 min and the buffer was removed by evaporation. The dry residue was then dissolved in 0.2 M *N*-methylmorpholine acetate pH 8.0 (0.25 ml) and the digestion by carboxypeptidase A (11 μ g; 0.55 unit) was allowed to proceed at 37°C for 0.5 h. Carboxypeptidase B (1 μ g; 0.14 unit) was then added and the digestion was carried on for 1 h at 37°C.

Peptides derived by cleavage of the protein with endoproteinase Arg-C were digested successively by carboxypeptidase B (0.6 μ g; 0.08 unit) for 15 min and by carboxypeptidase A (3.6 μ g; 0.18 unit) for 15 min. Peptides (12 nmol) were dissolved in 0.2 M *N*-methylmorpholine acetate pH 8.0 (0.24 ml) and the digestion was performed at 37°C. Aliquots (5 nmol) were treated as above for amino acid analyses.

Sequence determination

Automated Edman degradation of the protein HMB and of peptides was carried out on a gas-phase sequencer Applied Biosystems using the 02 N VAC Program.

Table 1 Edman degradation of protein HMB

The Edman degradation of protein HMB was performed on the gas-phase sequencer using the 02 N VAC program. A single run was carried out with 12.9 nmol of protein. Phenylthiohydantoin derivatives of amino acids were identified and quantified by HPLC. The repetitive yield calculated from valine residues at positions 7 and 18 was 93 %.

Cycle	Identified residue	Yield (nmol)	Cycle	Identified residue	Yield (nmol)
1	Ser	3.50	25	Arg	0.17
2	Asn	7.95	26	Gln	0.21
3	Thr	4.31	27	Ala	0.28
4	Arg	1.81	28	Ala	0.41
5	Asn	5.69	29	Leu	0.35
6	Phe	10.99	30	Lys	0.31
7	Val	6.69	31	Ala	0.37
8	Leu	8.97	32	Ala	0.41
9	Arg	2.73	33	Asn	0.20
10	Asp	4.10	34	Arg	0.10
11	Glu	4.41	35	Gly	0.16
12	Glu	4.28	36	Asp	0.09
13	Gly	3.19	37	Gly	0.12
14	Asn	5.08	38	Thr	0.03
15	Glu	4.13	39	Lys	-
16	His	1.41	40	Ser	-
17	Gly	2.28	41	Asn	0.05
18	Val	2.96	42	Pro	0.14
19	Phe	4.07	43	Asp	0.04
20	Thr	1.13	44	Val	0.07
21	Gly	1.66	45	nd	-
22	Lys	0.58	46	nd	-
23	Gln	0.30	47	Leu	0.23
24	Pro	0.31			

nd : not determined.

Alternatively automated Edman degradation of large peptides was performed on a liquid-phase Beckman 890 C sequencer using the 0.33 M quadrol program (11 1978) in the presence of polybrene [13]. Manual Edman degradation of peptides was carried out as in [14]. The phenylthiohydantoin derivatives of amino acids were identified by high-pressure liquid chromatography as indicated in [12].

Nomenclature of peptides

Peptides derived from cleavage at aspartyl and methionyl residues are designated by the symbols Ac- and CN-, respectively, whereas peptides generated by digestion with the endoproteinase Arg-C were designated by PR-. Peptides were numbered according to their position in the amino acid sequence of the protein.

RESULTS

The sequence of the amino-terminal half of the protein HMB (residues 1–44) was determined by automated Edman degradation of the protein (Table 1). The remainder of the sequence was elucidated from data provided by large adjacent or overlapping peptides generated by highly specific cleavages of the protein.

Cleavage at aspartic acid residues

Peptides derived from hydrolysis of protein HMB with 0.25 M acetic acid were fractionated by reverse-phase high-pressure liquid chromatography (Fig. 1A). Five peptides were obtained directly in pure form. Their amino acid compositions are given in Table 2. Due to partial deamidation of glutamine residues, peptide Ac-2 was eluted in fractions 2 and 3.

Only the large peptides Ac-4 and Ac-5 which overlap the carboxy-terminal half of the protein (residues 44–93) were completely sequenced by automated Edman degradation performed with the gas-phase sequencer (Table 3).

Cleavage by the endoproteinase Arg-C

Peptides generated by cleavage of the protein HMB with endoproteinase Arg-C were fractionated by reverse-phase high-pressure liquid chromatography (Fig. 1B). Each of the nine fractions obtained corresponds to a pure peptide. Only

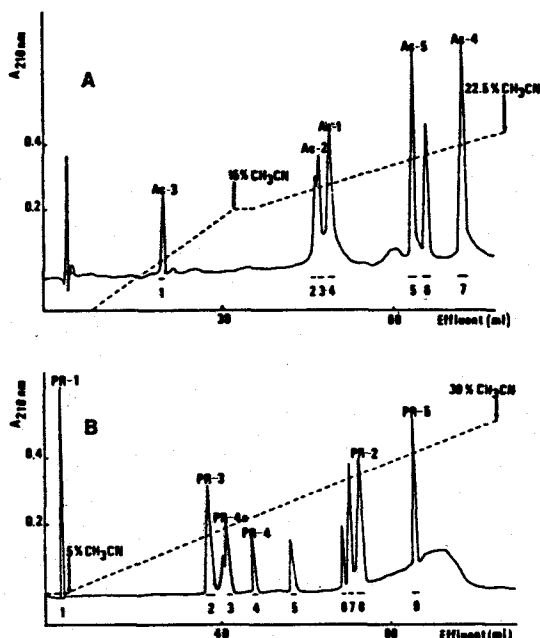


Fig. 1. Fractionations of peptides by reverse-phase HPLC. Chromatography was performed on a C_{18} μ Bondapak column (30×0.39 cm): (---) gradient of acetonitrile in 0.05% trifluoroacetic acid; (—) absorbance at 210 nm. Fractions were numbered by order of elution. Peptides used for sequence determination were numbered according to their position in the sequence of the protein. (A) Peptides derived by cleavage of protein HMB at aspartyl residues. 40 nmol of hydrolysate were loaded; flow rate was 1.5 ml/min; fractions of 0.75 ml were collected. (B) Peptides generated by hydrolysis of protein HMB with the endoproteinase Arg-C. 65 nmol of digest were loaded; flow rate was 2.0 ml/min; fractions of 0.8 ml were collected.

the amino acid compositions of peptides PR-1, PR-2, PR-3, PR-4, PR-4a and PR-5, which account for the entire sequence of the protein HMB, are given in Table 2.

Data obtained by automated Edman degradation performed on the gas-phase sequencer (for peptide PR-5) and on the liquid-phase sequencer (for peptides PR-2 and PR-4a) are given in Tables 3 and 4 respectively.

Besides the usual lack of cleavage of the Arg-Pro bond at position 71-72, the endoproteinase Arg-C did not cleave either the bond Arg-Asp at position 9-10 or the two Arg-Gly bonds at positions 34-35 and 50-51. The cluster Asp¹⁰-Glu-Glu¹² makes the environment of the susceptible bond Arg-Asp too acidic and the activity of the endoproteinase Arg-C may be hampered by the conformation of the polypeptide chain which has a high potential for β -turn structure at these glycine and proline residues. On the other hand, unexpected cleavage was observed at two Lys-Ala bonds at positions 30-31 and 59-60: the hydrolysis of both bonds was complete.

Half of the protein molecules were cleaved at the Arg⁴-Asn⁵ bond but only 16% of the molecules were cleaved at the Arg²⁵-Gln²⁶ bond; therefore, in addition to peptides PR-1 (residues 1-4) and PR-2 (residues 5-30), two peptides covering the sequences 1-30 and 1-25 were obtained in fractions 7 and 6 respectively. The bonds Arg-Leu at position 46-47 and Arg-Glu at position 48-49 were partially cleaved. Consequently, in addition to peptides PR-3 (residues 31-46) and PR-4 (residues 47-59), two peptides covering the sequences 31-48 and 49-59 (peptide PR-4a) were obtained with a yield of 8% and 10% respectively; the former was eluted in fraction 5 and the latter was eluted in fraction 3.

Cleavage at methionyl residues

Fragments derived from cleavage of the protein HMB at methionyl residues with cyanogen bromide, were separated by gel-filtration chromatography on a Sephadex G-50 superfine

Table 2 Amino acid compositions of peptides derived from protein HMB

Values are expressed as number of residues per mole of peptide. The numbers of residues in parentheses are from the established sequence. No correction was made for degradative losses of threonine and serine during hydrolysis of peptides. a, 72 h hydrolysis values. b, determined spectrophotometrically. c, determined as homoserine.

	Ac-1 Residues 1-9	Ac-2 Residues 11-35	Ac-3 Residues 37-42	Ac-4 Residues 44-72	Ac-5 Residues 74-93	PR-1 Residues 1-4	PR-2 Residues 5-30	PR-3 Residues 31-46	PR-4 Residues 47-59	PR-4a Residues 49-59	PR-5 Residues 60-93	CN-1 Residues 1-64	CN-2 Residues 65-75	CN-3 Residues 76-93	Protein HMB
Aspartic acid	2.0 (2)	2.1 (2)	1.0 (1)	1.5 (1)		1.0 (1)	3.2 (3)	3.9 (4)			1.8 (2)	7.6 (8)	1.8 (2)		9.6 (10)
Threonine	1.0 (1)	1.1 (1)	1.0 (1)	0.8 (1)		1.1 (1)	1.1 (1)	1.0 (1)	1.0 (1)	0.8 (1)		3.6 (4)			3.8 (4)
Serine	0.8 (1)		1.1 (1)		1.3 (1)	0.9 (1)		1.0 (1)			1.0 (1)	1.8 (2)		0.9 (1)	2.9 (3)
Glutamic acid		4.7 (5)		3.3 (3)	4.2 (4)		5.0 (5)		1.1 (1)	1.0 (1)	6.0 (6)	7.3 (7)	1.3 (1)	4.2 (4)	12.1 (12)
Proline		1.1 (1)	1.0 (1)	1.8 (2)	2.1 (2)		1.0 (1)	1.0 (1)			4.0 (4)	2.4 (2)	1.6 (2)	1.8 (2)	6.0 (6)
Glycine		3.4 (3)	1.1 (1)	1.8 (1)			3.0 (3)	2.2 (2)	1.1 (1)	1.2 (1)		5.7 (6)			6.3 (6)
Alanine		3.6 (4)		2.1 (2)			2.0 (2)	2.1 (2)			2.0 (2)	5.0 (5)	1.0 (1)		6.0 (6)
Valine	1.1 (1)	1.2 (1)		4.1 ^a (4)	1.4 (1)		1.9 (2)	1.0 (1)	1.7 (2)	1.9 (2)	2.0 (2)	4.7 (5)	1.0 (1)	1.3 (1)	6.5 (7)
Methionine				0.8 (1)	1.2 (1)						2.0 (2)	1.1 ^c (1)	1.0 ^c (1)	0.0 ^c (0)	1.9 (2)
Isoleucine				1.2 ^a (1)	2.6 ^a (3)			0.8 (1)			2.9 (3)	0.9 (1)		2.8 (3)	3.9 (4)
Leucine	1.0 (1)	1.1 (1)		1.3 (1)			1.9 (2)		1.0 (1)			3.2 (3)			3.2 (3)
Phenylalanine	1.1 (1)	1.0 (1)		1.0 (1)	1.2 (1)		1.8 (2)		0.9 (1)	1.0 (1)	1.1 (1)	2.9 (3)		1.1 (1)	3.9 (4)
Histidine		0.9 (1)		1.0 (1)			0.9 (1)		0.8 (1)	1.0 (1)		1.6 (2)			2.0 (2)
Lysine		1.9 (2)	0.9 (1)	5.0 (5)	5.7 (6)		1.8 (2)	0.7 (1)	3.0 (3)	3.2 (3)	8.2 (8)	7.2 (7)	1.3 (1)	5.8 (6)	14.4 (14)
Arginine	2.1 (2)	1.8 (2)		3.6 (4)		0.9 (1)	2.4 (2)	2.0 (2)	2.0 (2)	0.9 (1)	1.1 (1)	7.1 (7)	1.0 (1)		8.4 (8)
Tryptophan				1 ^b (1)	1 ^b (1)						2 ^b (2)	1 ^b (1)	1 ^b (1)		1.9 (2)
TOTAL	9	26	6	29	20	6	26	16	13	11	36	64	11	18	93
Fraction Number	6	2+3	1	7	5	1	8	2	6	3	9	1	3	2	
% yield of recovery	69	23	36.7	19.8	38.5	47	32	51.3	16	10	22.5	36	52	61	

Table 3 Edman degradation of protein HMB peptides performed on the gas-phase sequencer.

The runs were performed once for each peptide using the 02 N VAC program.

Phenylthiohydantoin derivatives of amino acids were identified and quantified by HPLC.

Peptides	Ac-4	Ac-3	PR-3	CN-3				
Quantity of material subjected to degradation	6 nmol	3 nmol	6 nmol	6 nmol				
Repetitive yield (%)	89.7	89.1	91.7	89.9				
Cycle	Identified residue	Yield (nmol)	Identified residue	Yield (nmol)	Identified residue	Yield (nmol)	Identified residue	Yield (nmol)
1	Val	5.00	Trp	0.81	Ala	5.00	Pro	2.99
2	Ile	4.90	Met	2.00	Trp	3.17	Glu	4.50
3	Arg	3.91	Pro	1.40	Lys	5.00	Lys	3.09
4	Leu	7.21	Glu	1.17	Glu	6.00	Ile	2.93
5	Arg	3.87	Lys	2.00	Met	5.00	Ser	0.43
6	Glu	3.80	Ile	0.76	Val	5.00	Lys	2.40
7	Arg	3.60	Ser	0.53	Glu	4.00	Pro	2.56
8	Gly	2.61	Lys	1.00	Ala	4.00	Phe	2.18
9	Thr	3.27	Pro	0.69	Pro	4.00	Val	1.36
10	Lys	3.07	Phe	0.76	Lys	4.00	Lys	1.43
11	Lys	4.19	Val	1.00	Asn	4.00	Lys	1.88
12	Val	2.55	Lys	1.00	Arg	2.44	Glu	1.74
13	His	1.65	Lys	1.00	Pro	3.00	Lys	1.88
14	Val	1.67	Glu	0.37	Asp	3.00	Ile	1.01
15	Phe	2.19	Lys	0.11	Trp	1.02	Glu	2.73
16	Lys	2.45	Ile	0.24	Met	3.00	Lys	1.15
17	Ala	2.55	Glu	0.36	Pro	3.00	Ile	0.78
18	Trp	0.17	Lys	0.16	Glu	2.06		
19	Lys	2.01	Ile	0.21	Lys	3.00		
20	Glu	1.07			Ile	1.64		
21	Met	1.00			Ser	0.23		
22	Val	0.86			Lys	1.00		
23	Glu	0.93			Pro	2.00		
24	Ala	0.88			Phe	1.01		
25	Pro	0.73			Val	0.72		
26	Lys	0.44			Lys	1.00		
27	Asn	0.32			Lys	1.00		
28	Arg	0.21			Glu	0.99		
29					Lys	1.00		
30					Ile	0.69		
31					Glu	0.93		
32					Lys	1.00		
33					Ile	0.66		

column (Fig. 2). Their amino acid compositions are presented in Table 2.

The lack of homoserine in peptide CN-3 and the data showing that peptide CN-3 and protein HMB have the same carboxy-terminal sequence determined as Ile-Glu-OH (Table 5) led us to assign the peptide CN-3 to the carboxy-terminal position in the protein. The complete sequence of peptide CN-3 was determined by automated Edman degradation performed on the gas-phase sequencer (Table 3).

Complete amino acid sequence of the protein HMB

The complete amino acid sequence of the protein HMB is presented in Fig. 3. The sequence of the amino-terminal half of the protein HMB (residues 1–44) was determined by automated Edman degradation of the protein. This sequence was confirmed by the data provided by peptides Ac-1, Ac-3,

PR-1, PR-2 and PR-3. The sequence of the carboxy-terminal half of the protein HMB (residues 44–93) was unambiguously established from sequence analyses of peptides Ac-4, Ac-5, PR-4, PR-5, CN-2 and CN-3. Alignment of fragments Ac-3 (residues 37–42) and Ac-4 (residues 44–72) was provided by the overlapping peptide PR-3 (residues 31–46). Alignment of fragments Ac-4 and Ac-5 (residues 74–93) was deduced from sequence analysis of peptide PR-5 (residues 60–93). The assignment of peptides Ac-5 and PR-5 to the carboxy-terminal position in protein HMB was supported by the sequence analyses of fragment CN-3.

The sequence data led us to discover that the apparent molecular size of protein HMB determined by SDS/polyacrylamide gel electrophoresis was overestimated: the protein HMB migrates as a protein of M_r 14500 which corresponds to a polypeptide chain of 123 amino acid residues [9]. But in all the sets of peptides generated by chemical or enzymatic

cleavages of the protein HMB (including a cleavage at tryptophan residues, data not shown) we never found peptides which might overlap a sequence beyond the residue at position 93. This was confirmed by determining the carboxy-terminal sequence of the protein HMB. The relative molecular mass of the protein HMB calculated from the sequence data is 10757. The abnormal electrophoretic migration of protein HMB in SDS/polyacrylamide gel is due to its high content of basic amino acids. Such an abnormal electrophoretic behaviour was also observed for calf thymus histone H4: protein HMB and histone H4 (calculated M_r 11 280) exhibit the same migration which is slightly slower than that of α -lactalbumin (M_r 14400) (Fig. 4).

DISCUSSION

The protein HMB is mainly characterized by a high amount of charged residues (15% of acidic residues and 26.8% of basic residues) which are distributed all along the polypeptide chain (Fig. 5). However, three sequences enriched in charged residues are obvious in the molecule: an acidic sequence (residues 10–15), a basic sequence (residues 46–

56) and the carboxy-terminal sequence (residues 85–93) where the four positive charges are balanced by the negative charges given by three glutamic acid residues and the free α -carboxyl group.

Most of the hydrophobic residues are associated in doublets or triplets. They are probably involved in strong hydrophobic interactions which might account for the high tendency of the protein HMB to aggregate during its purification process. The four residues of proline at positions 68, 72, 76 and 82 give to this region of the protein a highly folded

Table 4. Edman degradation of protein HMB peptides performed on the liquid-phase sequencer.

The runs were performed once for each peptide using the 0.33 M quadrol program (11 19 78) in the presence of polybrene. Phenylthiohydantoin derivatives of amino acids were identified and quantified by HPLC and when necessary by gas chromatography indicated by GC.

Peptide	PR-2	PR-4a
Quantity of material subjected to degradation	65 nmol	50 nmol
Repetitive yield (%)	77.7	82.8

Cycle	Residue	Yield (nmol)	Residue	Yield (nmol)
1	Asn	17.9	Glu	50
2	Phe	37.1 GC	Arg	13.9
3	Val	29.8 GC	Gly	17.2
4	Leu	25.3	Thr	2.7
5	Arg	14.9	Lys	8.3
6	Asp	19.5	Lys	8.3
7	Glu	1.5	Val	7.3 GC
8	Glu	1.7	His	3.8
9	Gly	6.0	Val	3.0 GC
10	Asn	0.8	Phe	3.4 GC
11	Glu	2.0	Lys	2.4
12	His	-	-	-
13	Gly	2.8	-	-
14	Val	0.6 GC	-	-
15	Phe	1.4 GC	-	-
16	Thr	-	-	-
17	Gly	0.6	-	-
18	Lys	-	-	-
19	Gln	-	-	-
20	Pro	-	-	-
21	Arg	-	-	-
22	nd	-	-	-
23	Ala	-	-	-
24	Ala	-	-	-

nd: not determined.

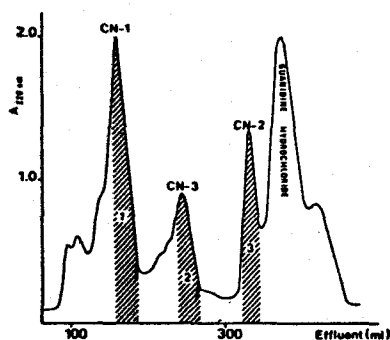


Fig. 2. Fractionation by gel-filtration chromatography of peptides derived from cleavage at methionyl residues. Peptides (0.5 μ mol in 0.7 ml of 6 M guanidine hydrochloride, 0.01 M HCl) were loaded on a Sephadex G-50 superfine column (90 \times 1.6 cm). The column was equilibrated and eluted with 0.01 M HCl saturated with chloroform. 1.3-ml fractions were collected at a flow rate of 4.0 ml/h. (—) Absorbance at 220 nm

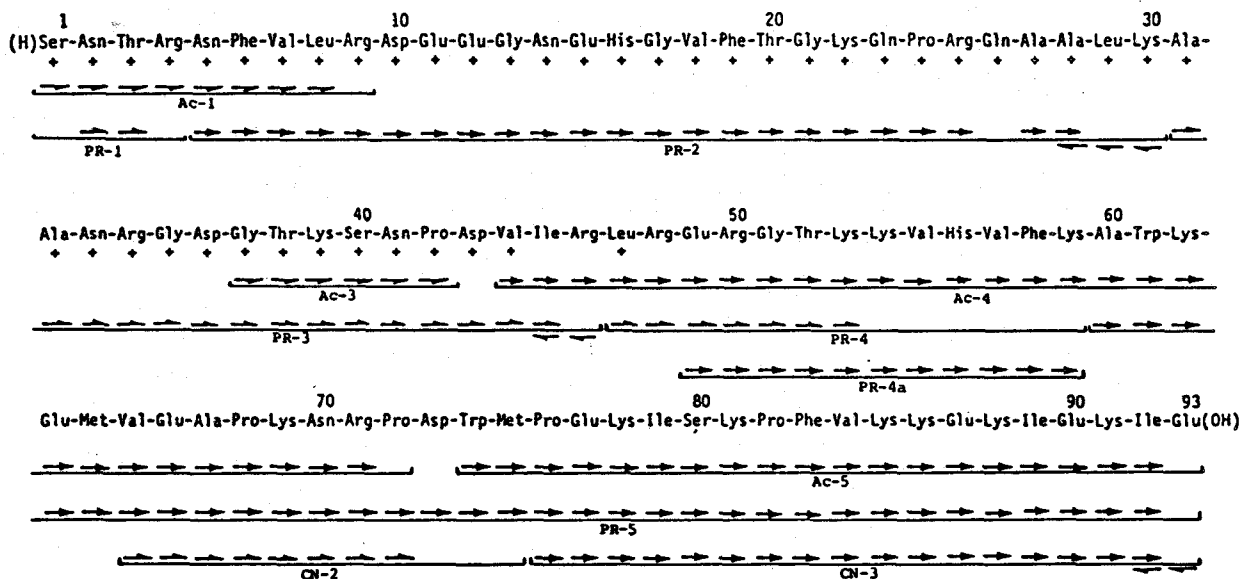


Fig. 3. Complete amino acid sequence of the *Methanosarcina barkeri* protein HMB. Methods used for determination of the sequence are indicated as follows: +++ automated Edman degradation of the protein; \rightarrow automated Edman degradation of peptides; \leftarrow manual Edman degradation of peptides; \leftarrow carboxypeptidase hydrolysis

conformation where β turns are placed side by side as predicted by the statistical method of Fasman et al. [15].

The protein HMB does not show any amino acid sequence homology either to eubacterial or to archaeobacterial chromosomal proteins. With an amount of basic residues equal to 27% and a net charge of +10, the protein HMB is much more basic than eubacterial HU-type proteins: the *Escherichia coli* HU protein contains 16.7% basic residues and has a net charge of +4. Moreover the distribution of basic residues in the two types of proteins is quite different; in HU-type proteins which lack highly basic sequences, the basic residues are scattered along the polypeptide chain (Fig. 5). The protein HTa, isolated from the archaeobacteria *T. acidophilum* [10], has an amino acid sequence clearly homologous to that of eubacterial HU-type proteins [11]. However a higher basicity

(net charge equal to +13) and the presence of clusters of two or three basic residues, differentiate the protein HTa from eubacterial HU-type proteins. The DNA-binding protein isolated from *Sulfolobus solfataricus* [8] differs from protein HMB in size (M_r 7000) and in primary structure. Proteins of $M_r \approx 14500$ and exhibiting physico-chemical properties different from those of protein HMB have been characterized in *Methanosarcina thermoautotrophicum* [16] and in *Sulfolobus acidocaldarius* (protein HSA) [7]. Unlike protein HMB, the former is not acid-soluble and the protein HSA is distinct from the protein HMB particularly with respect to the amounts of lysine, arginine and leucine.

On the other hand, the protein HMB presents some analogies with the smallest core histone, the histone H4 in: (a) its molecular size; (b) its tendency to form aggregates and (c) the large number of basic residues, several of which are associated in doublets or in sequences where they are separated from each other by one or two non-basic residues. However the following differences between these two proteins have to be emphasized: (a) due to its high content in dicarboxylic amino acids, the protein HMB is less basic than histone H4 which has a net charge equal to +19; (b) the protein HMB contains two residues of tryptophan, an amino acid residue which is not encountered in any histones; (c) in histone H4, the highly basic character is located in the amino-terminal half of the polypeptide chain whereas an opposite distribution is encountered in protein HMB: two thirds of its basic residues are located in the carboxy-terminal half of the polypeptide (Fig. 5); (d) no significant amino acid sequence homology has been detected between the protein HMB and calf histone H4 by a computer-based survey.

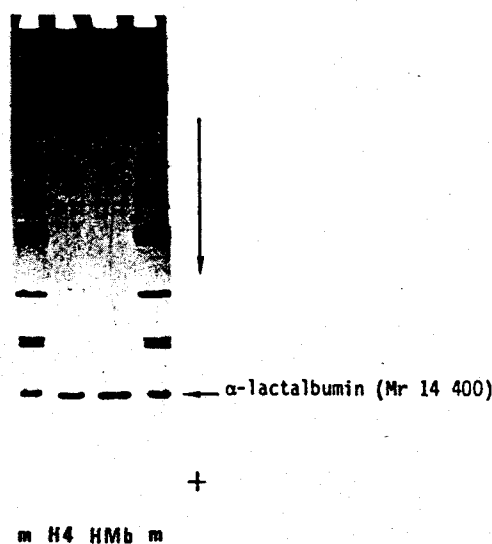


Fig. 4. SDS/polyacrylamide gel electrophoresis of *Methanosarcina barkeri* protein HMB and calf thymus histone H4. The M_r markers in lane m are: phosphorylase b (94000), bovine serum albumin (67000), ovalbumin (43000), carbonic anhydrase (30000), soybean trypsin inhibitor (20100), α -lactalbumin (14400). Samples (7 μ g) were treated with 5% SDS, 2% 2-mercaptoethanol for 2 min at 100°C and run at 40 mA for 150 min on a 5–30% polyacrylamide gradient gel [17]. Electrode buffer (Tris/glycine, pH 8.3) and gel buffer (Tris/HCl, pH 8.9) were made 0.1% in SDS. The gels were stained with 0.5% Coomassie blue R-250 in acetic acid/ethanol/water (1:2:7, v/v/v) for 15 h. The gel was destained by diffusion in acetic acid/ethanol/water (4:5:31, v/v/v)

Table 3 Determination of the carboxy-terminal sequences of the protein HMB and of peptide CN-3.

Results are expressed in number of residues released in the digests (see Methods).

Carboxypeptidase Y for 7 h	protein HMB peptide CN-3	Glu 1.0, Ile 0.6 Glu 0.9, Ile 0.0
Carboxypeptidase A for 0.5 h	protein HMB peptide CN-3	Glu 1.0, Ile 0.9 Glu 0.9, Ile 0.8

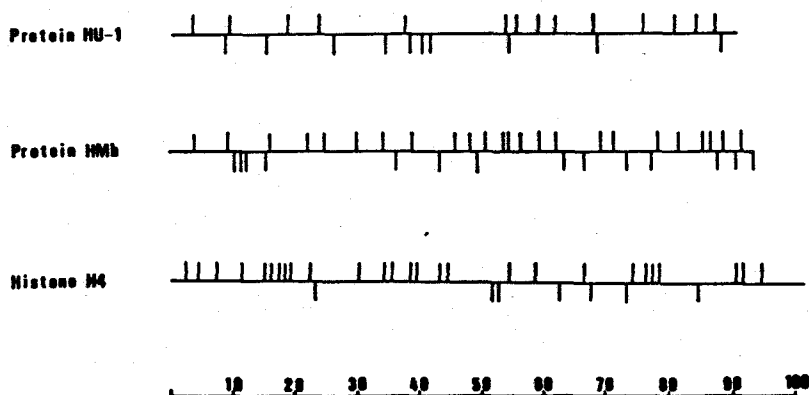


Fig. 5. Distribution of basic (\uparrow) and acidic (\downarrow) residues in protein HU-1 (*Escherichia coli*) [18], protein HMB (*Methanosarcina barkeri*) and histone H4 (calf thymus) [19]

Circular dichroic spectroscopy indicates that the protein Hmb is almost devoid of α -helix structure. We are now looking for the conformational state of the protein and the way it can organize the DNA in the archaeobacterial nucleoid. We also have to determine whether the protein Hmb is able by itself to play this function or if it must be associated with another chromosomal protein, such as the 7-kDa protein found in the *Methanosarcina barkeri* deoxyribonucleoprotein complex [9].

The authors are grateful to Dr J. P. Touzel (Station de Technologie Alimentaire, INRA, Villeneuve d'Ascq) who provided the cultures of *Methanosarcina barkeri* and for valuable discussions. They are indebted to A. Lemaire, M. J. Dupire, and W. Kruggel for skilful technical assistance and to M. Lestiennes for editorial assistance. They also thank the *service commun de séquence Unités 409 CNRS/16 INSERM* for sequencer runs. This work was supported by grants from the *Centre National de la Recherche Scientifique*, from the *Pôle Régional Anaérobies* and from the *Fondation à la Recherche Médicale*.

REFERENCES

1. Fox, G. E., Stackebrandt, E., Hespell, R. B., Gibson, J., Maniloff, J., Dyer, T. A., Wolfe, R. S., Balch, W. E., Tanner, R. S., Magrum, L. J., Zablen, L. B., Blakemore, R., Gupta, R., Bonem, L., Lewis, B. J., Stahl, D. A., Luchsen, K. R., Chen, K. N. & Woese, C. R. (1980) *Science (Wash. DC)* **209**, 457–463.
2. Fox, G. E. (1985) in *The bacteria, a treatise on structure and function* (Woese, C. R. & Wolfe, R. S., eds) vol. VIII, pp. 257–301, Academic Press, New York.
3. Pettijohn, D. E. (1982) *Cell* **30**, 667–669.
4. Rouvière-Yaniv, J., Yaniv, M. & Germond, J. E. (1979) *Cell* **17**, 265–274.
5. Broyles, S. S. & Pettijohn, D. E. (1986) *J. Mol. Biol.* **187**, 47–60.
6. Searcy, D. G. (1975) *Biochim. Biophys. Acta* **395**, 535–547.
7. Green, G. R., Searcy, D. G. & De Lange, R. J. (1983) *Biochim. Biophys. Acta* **741**, 251–257.
8. Kimura, M., Kimura, J., Davie, P., Reinhardt, R. & Dijk, J. (1984) *FEBS Lett.* **176**, 176–178.
9. Chartier, F., Laine, B., Sautière, P., Touzel, J. P. & Albagnac, G. (1985) *FEBS Lett.* **183**, 119–123.
10. De Lange, R. J., Williams, L. C. & Searcy, D. G. (1981) *J. Biol. Chem.* **256**, 905–911.
11. Laine, B., Bélaïche, D., Khanaka, H. & Sautière, P. (1983) *Eur. J. Biochem.* **131**, 325–331.
12. Laine, B., Bélaïche, D., Sautière, P. & Biserte, G. (1982) *Biochem. Biophys. Res. Commun.* **106**, 101–107.
13. Tarr, G. E., Beccher, J. F., Bell, M. & McKeaw, D. (1978) *Anal. Biochem.* **84**, 622–627.
14. Edman, P. & Henschen, A. (1975) in *Protein sequence determination* (Needleman, S. B., ed.) 2nd edn, pp. 232–279, Springer-Verlag, Berlin, Heidelberg, New York.
15. Fasman, G. D., Chou, P. Y. & Adler, A. J. (1976) *Biophys. J.* **16**, 1201–1238.
16. Thomm, M., Stetter, K. O. & Zillig, W. (1982) *Zentralbl. Bakteriol. Hyg. I. Abt. Orig. C3*, 128–139.
17. Laemmli, U. K. (1970) *Nature (Lond.)* **227**, 680–685.
18. Laine, B., Kmićik, D., Sautière, P., Biserte, G. & Cohen-Solal, M. (1980) *Eur. J. Biochem.* **103**, 447–461.
19. De Lange, R. J., Fambrough, D. M., Smith, E. L. & Bonner, J. (1969) *J. Biol. Chem.* **244**, 319–334.

FEATURES OF THE PRIMARY STRUCTURE OF THE *METHANOSARCINA BARKERI* CHROMOSOMAL PROTEIN HMB COMPARED TO THOSE OF EUKARYOTIC AND PROKARYOTIC CHROMOSOMAL PROTEINS

M. IMBERT, F. CHARTIER, B. LAINE and P. SAUTIERE

Unité Associée 409 CNRS, Institut de Recherches sur le Cancer, Place de Verdun
59045 Lille Cédex, France.

SUMMARY

The basic protein HMB is the major component of the *Methanosarcina barkeri* nucleoprotein complex. The main features of protein HMB primary structure are described and compared to those of eukaryotic, eubacterial and archaeobacterial chromosomal proteins. Protein HMB is not homologous to any of these chromosomal proteins.

INTRODUCTION

The composition and the organization of the chromosome in eukaryotes and eubacteria are quite different (ref.1). Inside the nucleus of eukaryotic cells packaging of DNA is mainly ensured by basic proteins called histones, whereas in eubacteria, HU-type proteins are involved in DNA condensation. The discovery of archaeobacteria, a line of descent distinct from eubacteria and eukaryotes (ref. 2), led several authors to investigate the protein composition of their nucleoid. Chromosomal proteins have been isolated from three strains of thermoacidophiles : *Thermoplasma acidophilum* (ref.3), *Sulfolobus acidocaldarius* (ref.4), and *Sulfolobus solfataricus* (ref.5). Up to now, only one chromosomal protein (protein HMB) has been characterized from a methanogenic bacteria : *Methanosarcina barkeri* strain MS (ref.6). The main features of protein HMB primary structure are described and compared to those of eukaryotic, eubacterial and archaeobacterial chromosomal proteins.

RESULTS AND DISCUSSION

Protein HMB is the major component of the acid extract of the *Methanosarcina barkeri* deoxyribonucleoprotein complex. It is a polypeptide of 93 amino acid residues (Mr 10,757) which tightly binds to the double stranded DNA : a sodium chloride molarity of 0.47 M is necessary to desorb the protein HMB from a column of *E. coli* DNA bound to cellulose. The protein HMB has a high content of basic residues (Lys 14 ; Arg 8 and His 2) and of acidic residues (Asp 4 and Glu 10). Half of the basic residues are clustered in sequences such as B-B-X-B (residues 53-56 and 85-88) or B-X-B (residues 46-48, 48-50, 70-72) where B and X respec-

238

TABLE 1

Chemical characteristics of chromosomal proteins

	EUKARYOTES				EUBACTERIA	ARCHAEBACTERIA			
	H2B	H2A	H3	H4	HU	HTa	HMb	HSa	7kD
Molecular size	13770	14000	15300	11280	9500	9930	10760	14500	7000
Basic residues %	25	23	24	26	17	22	27	16	27
Net charge	+18	+16	+20	+16	+4	+13	+10	nd	+5

The proteins HU, HTa, HMb, HSa and 7kD are respectively isolated from *E.coli*, *T.acidophilum*, *M.barkeri*, *S.acidocaldarius* and *S.solfataricus*. The values given for histones H2B, H2A, H3 and H4 are those of calf thymus histones.

stitutions by comparison with *E.coli* protein HU-1 but the amino acid residues which play a crucial role in the function of the protein are strictly conserved (ref.10). With 102 residues (Mr 11,280) histone H4 is the smallest of the five species of histones, it is the most conserved during evolution and constitutes with histone H3 a tetramer (H3-H4)₂ which is essential in the formation and the stability of the nucleosome protein core.

The comparison of the hydrophobicity diagrams indicates that the protein HMb is radically different from the other proteins (Fig.2). The protein HMb exhibits a marked hydrophilic character due to its high content of charged residues (42% of the total amino acid residues) distributed all along the polypeptide chain. On the contrary large hydrophobic domains are encountered in the three other proteins ; they play an important role in the conformation of these proteins and in the formation of oligomers. In HU-type proteins, including the archaeobacterial protein HTa, these hydrophobic domains are well conserved and are involved in hydrophobic interactions between the structural segments of the protein (α helices or β stranded sheets). These hydrophobic interactions give a globular shape to the dimer HU which corresponds to the functional state of this chromosomal protein (ref.10). On the other hand, the protein HMb behaves as a monomer and preliminary data indicate that, when dissociated from the DNA, it does not exhibit a conformation as well organized than that of HU-type proteins.

The protein HMb also differs from the HU-type proteins in its amount of basic residues and in its net charge (Table 1). Moreover in protein HMb, half of the basic residues are clustered and constitute highly basic sequences (Fig.3). Such a distribution is not encountered in eubacterial HU-type proteins where the basic residues are scattered all along the polypeptide chain. Only the *T.acidophilum* protein HTa displays a strong basicity with a net charge equal to +13 and with the presence of clusters of two or three basic residues.

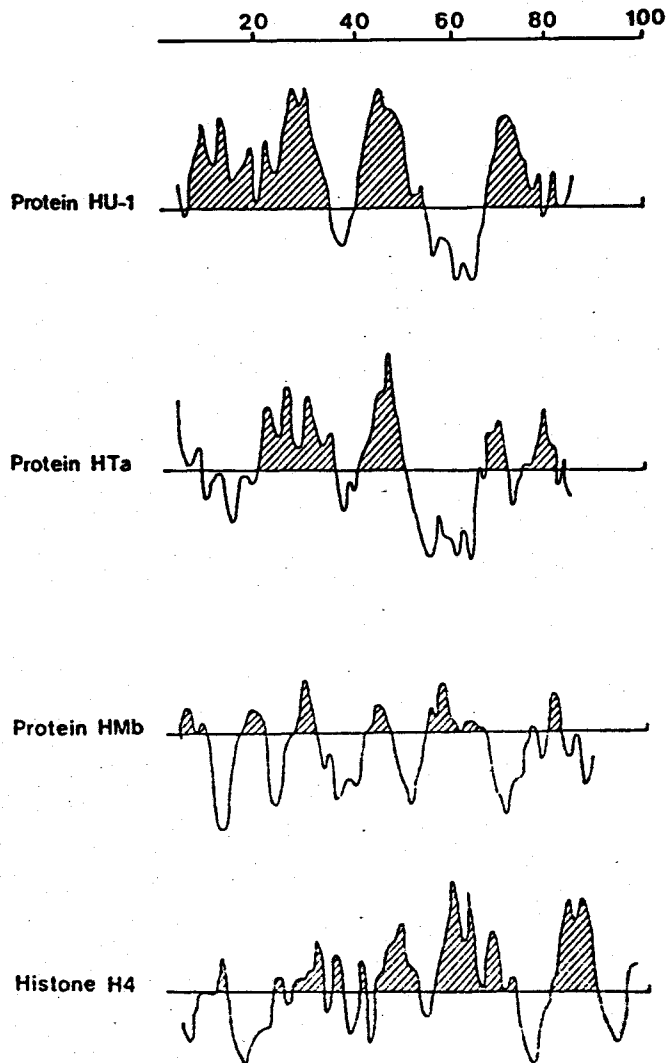


Fig.2. Hydrophobicity diagrams of proteins HU-1 (*E. coli*), HTa (*T. acidophilus*), HMB (*M. barkeri*) and calf thymus histone H4. The diagrams were drawn according to (ref.11) and using a computer program. Hydrophobic sequences appear in shaded areas and hydrophilic sequences in non-shaded areas.

The amount of basic residues in protein HMB is similar to that found in histones. However a higher content of dicarboxylic amino acid residues makes the net charge of protein HMB lower than that of histones. Moreover, unlike in protein HMB, acidic residues are not found in the sequences of histones where basic residues are clustered. Consequently the basic domains of histones have a very high net charge : the amino-terminal half of histone H4 (residues 1 through 45)

240

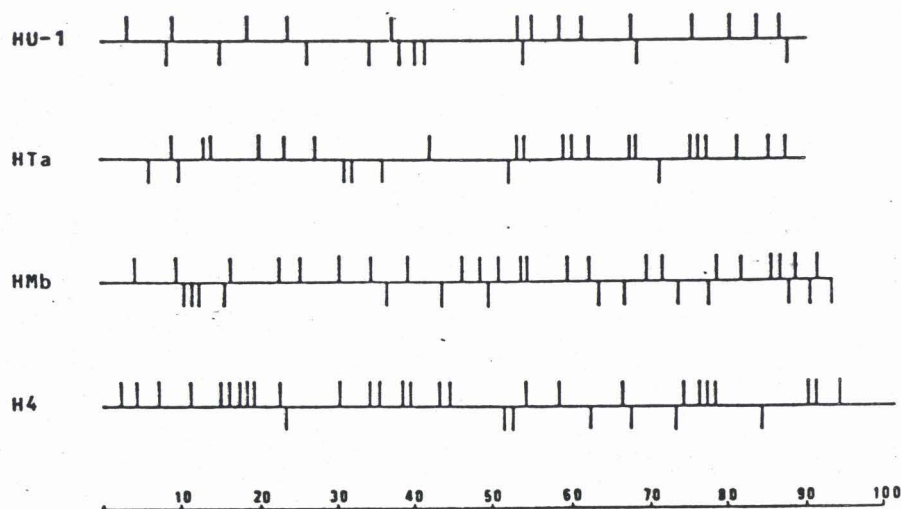


Fig.3. Distribution of basic (—) and acidic (—) residues in proteins HU-1 (*E. coli*)(ref.7), HTa (*T.acidophilum*)(ref.8), HMB (*M.barkeri*) and histone H4 (calf thymus)(ref.9).

has a net charge of +13. A last character differentiates protein HMB from histones : it is the presence in protein HMB of two residues of tryptophan, an amino acid residue which is not encountered in any histone.

In conclusion, by its high basicity and the presence of basic clusters, the protein HMB appears to be closer to eukaryotic histones and to *T.acidophilum* protein HTa than to eubacterial HU-type proteins. However any sequence homology can be observed between protein HMB and histones or protein HTa.

Work is in progress to localize the protein HMB in the nucleoid of *Methanosarcina barkeri* using an immunochemical method. We also are looking for the conformational state of the protein and the way it can condense the DNA of this bacteria.

The authors are grateful to Dr J.P.Touzel from the Station de Technologie Alimentaire, INRA, Villeneuve d'Ascq, for valuable discussions and for growing the bacteria. They are indebted to A.Lemaire-Poitau, T.Ernout and A.Hemez for skiful technical assistance. This work was supported by grants from the Centre National de la Recherche Scientifique and the Pôle des Anaérobies de la Région Nord-Pas de Calais.

REFERENCES

- 1 D.E. Pettijohn, Structure and properties of the bacterial nucleoid, *Cell*, 30 (1982) 667-669.
- 2 C.R. Woese and G.E. Fox, Phylogenetic structure of the prokaryotic domain : the primary kingdoms, *Proc.Natl.Acad.Sci.USA*, 74 (11)(1977) 5088-5090.

- 3 D.G. Searcy, Histone-like protein in the prokaryote *Thermoplasma acidophilum*, *Biochim.Biophys.Acta*, 395 (1975) 535-547.
- 4 G.R. Green, D.G. Searcy and R.J. De Lange, Histone-like protein in the archaeobacterium *Sulfolobus acidocaldarius*, *Biochim.Biophys.Acta*, 741 (1983) 251-257.
- 5 M. Kimura, J. Kimura, P. Davie, R. Reinhardt and J. Dijk, The amino acid sequence of a small DNA-binding protein from the archaeobacterium *Sulfolobus solfataricus*, *FEBS Letters*, 176(1)(1984) 176-178.
- 6 F. Chartier, B. Laine, P. Sautière, J.P. Touzel and G. Albagnac, Characterization of the chromosomal protein HMB isolated from *Methanosarcina barkeri*, *FEBS Letters*, 183 (1)(1985) 119-123.
- 7 B. Laine, D. Kmiećik, P. Sautière, G. Biserte and M. Cohen-Solal, Complete amino acid sequences of DNA-binding proteins HU-1 and HU-2 from *Escherichia coli*, *Eur.J.Biochem.*, 103 (1980) 447-461.
- 8 R.J. De Lange, L.C. Williams and D.G. Searcy, A histone-like protein (HTa) from *Thermoplasma acidophilum* II. Complete amino acid sequence, *J.Biol.Chem.* 256 (2)(1981) 905-911.
- 9 R.J. De Lange, D.M. Fambrough, E.L. Smith and J. Bonner, Calf and Pea histone IV. II. The complete amino acid sequence of calf thymus histone IV ; presence of ϵ -N-acetyllysine, *J.Biol.Chem.*, 244 (1969) 319-334.
- 10 I. Tanaka, K. Applet, J. Dijk, S.W. White and K.S. Wilson, 3-Å resolution structure of a protein with histone-like properties in prokaryotes, *Nature*, 310 (1984) 376-381.
- 11 J. Kyte and R. Doolittle, A simple method for displaying the hydrophobic character of a protein, *J.Mol.Biol.*, 157 (1982) 105-132.

III - 1 ETUDE DE LA PROTÉINE MC1 DE *METHANOSARCINA BARKERI* MS

III - 1-1 STRATEGIE

La figure III-1 schématise la stratégie que nous avons adoptée pour déterminer la séquence en acides aminés de la protéine MC1 de *M.barkeri*. Dans un premier temps, la séquence N-terminale de la protéine (44 résidus) a été établie par dégradation d'Edman automatisée. Nous avons ensuite cherché à obtenir un nombre restreint de fragments peptidiques faciles à isoler. Pour cela la protéine a été coupée spécifiquement par des méthodes chimiques ou enzymatiques, le choix des coupures découlant des caractéristiques de la composition en acides aminés de la protéine MC1.

La protéine MC1 de *M.barkeri* MS renferme 10 résidus d'acide aspartique ou d'asparagine (Tableau III-1, p.76). Le caractère basique de la protéine MC1, déterminé par sa migration électrophorétique, suggérant que les résidus d'acide aspartique étaient peu nombreux, nous avons choisi le clivage des liaisons aspartyles par l'acide acétique 0,25 M.

La coupure de la protéine MC1 par le bromure de cyanogène a été réalisée en fonction de la présence de deux résidus de méthionine dans la protéine.

L'endoprotéinase Arg-C a été choisie du fait de son étroite spécificité et de la présence de 8 résidus d'arginine dans la protéine.

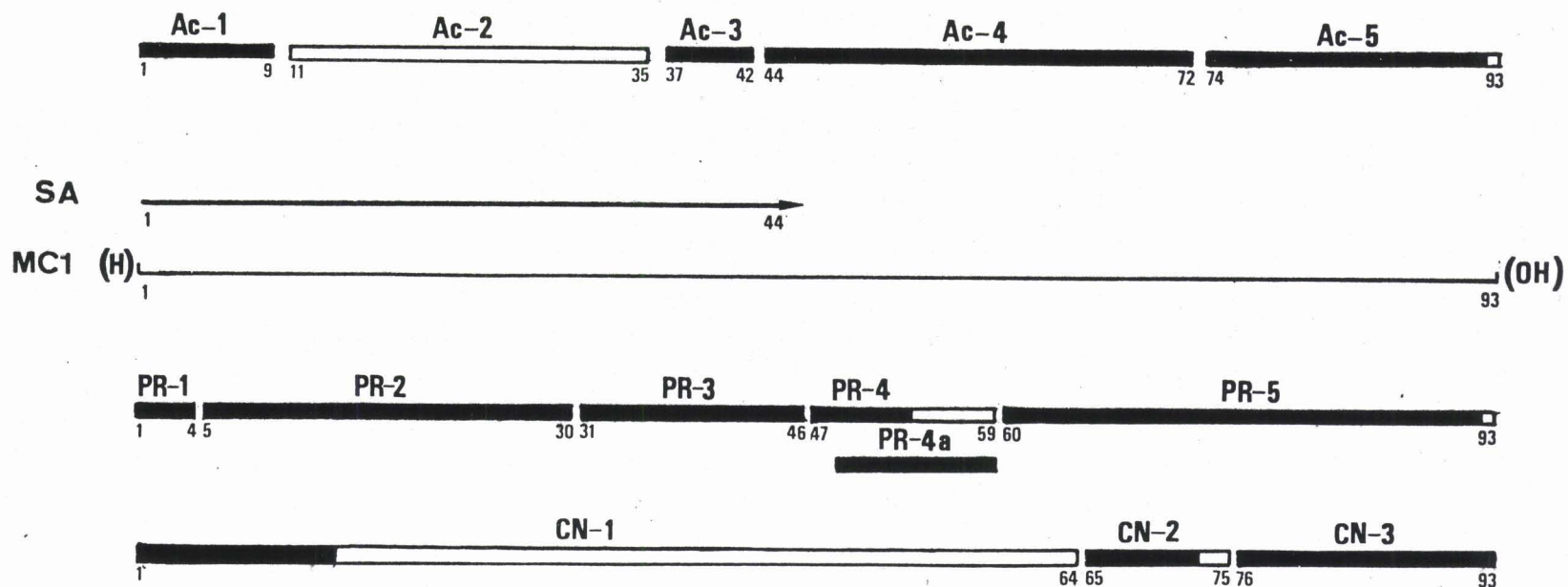


FIGURE III-1 - Stratégie adoptée pour la détermination de la séquence en acides aminés de la protéine MCl de *M. barkeri*.

S.A : séquence N-terminale de la protéine MCl établie par séquençage automatique (dégradation d'Edman).

Peptides obtenus par hydrolyse de la protéine MCl par l'acide acétique (Ac-), avec l'endoprotéinase Arg-C (PR-), par le bromure de cyanogène (CN-).

Les rectangles noirs représentent les séquences déterminées dans les fragments peptidiques.



III - 1-2 RESULTATS

III - 1-2-1 Contribution personnelle

L'étude de la structure primaire de la protéine MC1 de *Methanosarcina barkeri* MS constitue un ensemble de travaux réalisés par Bernard LAINE, François CHARTIER, Pierre SAUTIERE et nous-même et qui a fait l'objet de la publication n°1.

Pour notre part, nous avons étudié les peptides obtenus par action de l'endoprotéinase Arg-C et nous avons déterminé la séquence de l'extrémité C-terminale de la protéine par hydrolyse avec les carboxypeptidases Y et A.

L'hydrolyse enzymatique de la protéine MC1 par l'endoprotéinase Arg-C a permis d'obtenir des peptides de longueur moyenne et qui recouvrent ensemble la protéine entière. Le peptide PR-5 (résidus 60-93) est particulièrement intéressant car il a permis l'alignement des peptides CN-2 et CN-3 ainsi que des peptides Ac-4 et Ac-5. Certaines liaisons arginyles ont été coupées partiellement, ces résultats sont discutés dans la publication n°1 : il faut souligner en particulier la coupure imprévue des liaisons Lys-Ala en positions 30-31 et 59-60. L'hydrolyse de ces liaisons a été complète.

Avec une masse moléculaire apparente de 14.400, en électrophorèse sur gel de polyacrylamide SDS, la protéine MC1 semblait correspondre à un polypeptide de 125 résidus (CHARTIER et coll., 1985) mais les travaux de séquence ont montré sans équivoque que la protéine MC1 ne contenait que 93 résidus.

Par ailleurs, la séquence C-terminale de la protéine a été confirmée en hydrolysant la protéine et le peptide CN-3, successivement avec les carboxypeptidases Y et A. Le peptide CN-3 obtenu par coupure de la protéine avec le bromure de cyanogène est dépourvu d'homosérine. Il correspond donc à la séquence C-terminale de la protéine. L'hydrolyse avec la carboxy-

peptidase Y à pH 5,5 a donné d'excellents résultats puisqu'un résidu d'acide glutamique a été libéré par molécule de protéine et de peptide CN-3. L'action de la carboxypeptidase A a ensuite libéré le résidu d'isoleucine en position 92.

III - 1-2-2 Séquence complète

Les travaux concernant la détermination de la séquence complète de la protéine MC1 de *M.barkeri* MS ont fait l'objet de la publication N°1.

III - 2 ETUDE DE LA PROTÉINE MC1 DE *METHANOSARCINA* SP SOUCHE MST-A1

III - 2-1 STRATEGIE

La figure III-2 schématise la stratégie que nous avons adoptée pour déterminer la séquence en acides aminés de la protéine MC1 de *Methanosarcina* sp MST-A1. La stratégie est basée sur le fait que les protéines MC1 de *M.barkeri* et de *Methanosarcina* sp sont des protéines homologues : elles présentent la même migration électrophorétique en gel de polyacrylamide en milieu SDS (Figure II-5, p.45) et leurs compositions en acides aminés sont très proches (Tableau III-1). Par conséquent, nous avons utilisé les mêmes méthodes de coupures chimiques et enzymatiques (acide acétique 0,25 M, bromure de cyanogène et endoprotéinase Arg-C) pour élucider la structure primaire de la protéine MC1 de *Methanosarcina* sp MST-A1. Néanmoins, le peptide 36-65 obtenu après hydrolyse à l'acide acétique 0,25 M n'a pas pu être séquencé. Un autre type de coupure a dû être envisagé afin de couvrir la région médiane de la protéine. Le clivage aux liaisons glutamyles par la protéase

TABLEAU III-1

Compositions en acides aminés de la protéine MCl de *Methanosarcina barkeri* souche MS et de *Methanosarcina* sp souche MST-A1

	<i>Methanosarcina barkeri</i> souche MS	<i>Methanosarcina</i> sp souche MST-A1
Asp	10*	9,5 (10)
Thr	4	4,2 (4)
Ser	3	3,7 (2)
Glu	12	11,4 (11)
Pro	6	6,1 (6)
Gly	6	6,3 (6)
Ala	6	8,3 (8)
Cys	0	0 (0)
Val	7	6,2 (6)
Met	2	0,9 (1)
Ile	4	4,7 (5)
Leu	3	4,1 (4)
Tyr	0	0 (0)
Phe	4	4,0 (4)
His	2	2,0 (2)
Lys	14	13,7 (14)
Arg	8	7,9 (8)
Trp	2	2,0 (2)
Total	93	93

* Les résultats sont exprimés en nombre de résidus d'acides aminés par molécule de protéine et les chiffres entre parenthèses correspondent aux résidus effectivement présents dans la séquence.

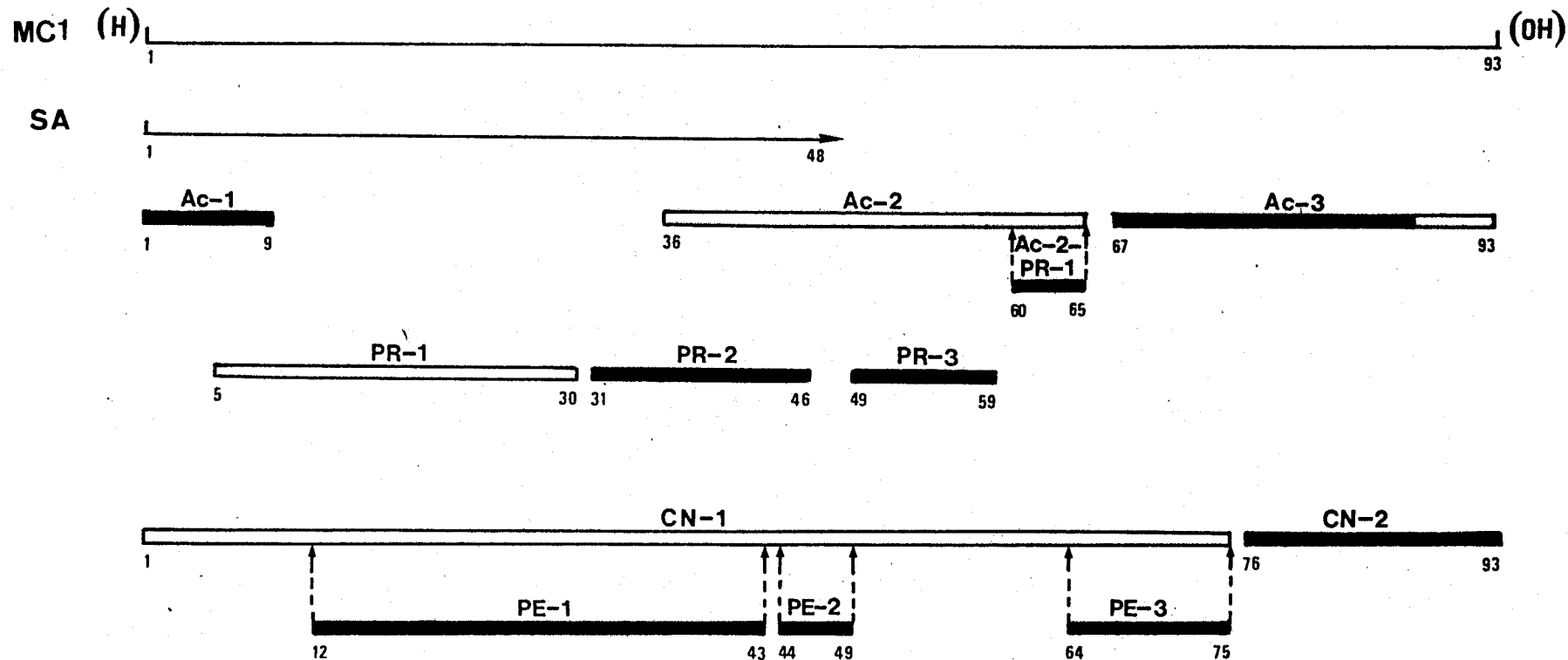


FIGURE III-2 - Stratégie adoptée pour la détermination de la séquence en acides aminés de la protéine MCl de *Methanosarcina* sp souche MST-A1.

S.A : séquence N-terminale de la protéine MCl obtenue par séquençage automatique (dégradation d'Edman)

Peptides obtenus par hydrolyse de la protéine MCl par l'acide acétique (Ac-), avec l'endoprotéinase Arg-C (PR-), par le bromure de cyanogène (CN-), avec la protéase staphylococcale (PE-).

Les rectangles noirs représentent les séquences déterminées dans les fragments peptidiques.

staphylococcale a donc été réalisé, non pas sur la protéine intacte mais sur le fragment 1-75 obtenu après hydrolyse de la protéine MC1 par le bromure de cyanogène. Nous avons ainsi évité l'obtention d'un hydrolysats complexe dû à la libération de nombreux petits peptides provenant de la partie C-terminale de la protéine très riche en résidus d'acide glutamique.

III - 2-2 MATERIEL ET METHODES

Les conditions d'hydrolyse au niveau des résidus d'acide aspartique, de méthionine, et par l'endoprotéinase Arg-C sont décrites dans la publication n°1.

** Hydrolyse avec la protéase staphylococcale V8*

La protéine MC1 (350 nM) a été hydrolysée avec la protéase staphylococcale V8 (Miles) dans 2 ml de tampon acétate d'ammonium 0,05 M à pH 4,0 pendant 14 h à 37°C en utilisant un rapport enzyme/substrat de 1/40 en poids. La digestion est arrêtée par chauffage à 100°C pendant 2 minutes. L'hydrolysats est ensuite évaporé à sec.

** Fractionnement des peptides*

Les peptides obtenus après hydrolyse de la protéine MC1 par l'acide acétique ou avec l'endoprotéinase Arg-C ou avec la protéase staphylococcale ont été séparés par chromatographie liquide haute performance (CLHP) en phase inverse sur une colonne C₁₈ µBondapak (Waters Associates) et en utilisant les systèmes de gradients indiqués dans les figures III-3, III-4 et III-6.

Les peptides produits par la coupure au bromure de cyanogène sont séparés par chromatographie de tamisage moléculaire sur une colonne de Sephadex G-50 superfine (95x1,6 cm) équilibrée et éluée avec HCl 0,01 M.

III - 2-3 RESULTATS

* Coupure aux résidus d'acide aspartique

Après hydrolyse par l'acide acétique 0,25 M de la protéine MCl de *Methanosarcina* sp MST-A1, trois peptides (Ac-1, Ac-2, Ac-3) ont été obtenus purs (Figure III-3). Seuls les peptides Ac-1 et Ac-3 ont pu être séquencés par dégradation d'Edman automatisée. Le peptide Ac-2 qui est élué dans les fractions 5 et 6 n'a pas pu être dégradé : ce point sera discuté ultérieurement. Néanmoins, les caractéristiques de sa composition en acides aminés (présence d'un résidu de phénylalanine et d'un résidu d'histi-dine, richesse en lysine) permettent de placer sans ambiguïté ce peptide en position 36-65 (Figure III-7). L'hydrolyse du peptide Ac-2 avec l'endoprotéinase Arg-C a permis d'obtenir le peptide Ac-2-PR-1. Compte tenu de la spécificité de l'endopro-téinase Arg-C, ce peptide qui ne possède pas d'arginine à son extrémité C-terminale peut être placé en position C-terminale du fragment Ac-2 (Fig. III-7).

* Coupure avec l'endoprotéinase Arg-C

12 fractions souvent hétérogènes ont été obtenues après hy-drolyse de la protéine MCl de *Methanosarcina* sp MST-A1 avec l'en-doprotéinase Arg-C. Cependant trois peptides ont pu être obte-nus directement à l'état pur : les peptides PR-1 (résidus 5-30), PR-2 (résidus 31-46) et PR-3 (résidus 49-59) (Figure III-4 et III-7).

Les liaisons Arg²⁵-Gln²⁶, Arg⁴⁶-Leu⁴⁷ et Arg⁵⁰-Gly⁵¹ ont été coupées partiellement. Pour cette raison, trois peptides couvrant les séquences 26-30, 40-48 et 51-59 ont été obtenus respectivement dans les fractions 1, 8 et 4. De plus, une coupure imprévue avec l'endoprotéinase Arg-C a eu lieu au niveau des liaisons Lys-Ala aux positions 30-31, 39-40 et 59-60. Des coupu-res totales au niveau de la liaison Lys-Ala avaient déjà été

FIGURE III-3 - Diagrammes d'élution des peptides obtenus par coupe et
FIGURE III-4 l'endoprotéinase Arg-C sur la protéine MCl de *Methanosarcina* sp MST-A1, fractionnés par CLPH en phase inverse.

La chromatographie est effectuée sur une colonne C₁₈ µBondapak (300 x 3,9 mm) ; (---) gradient d'acétonitrile dans de l'acide trifluoroacétique à 0,05 % ; (—) absorbance à 210 nm. Les fractions sont numérotées par ordre d'élution. Les peptides utilisés pour la détermination de la séquence ont été numérotés selon leur position dans la séquence de la protéine.

Figure III-3 : Peptides produits par clivage de la protéine MCl de *Methanosarcina* sp MST-A1 des liaisons aspartyles.

120 nM de l'hydrolysats sont déposées sur la colonne. Le débit est de 1,5 ml/min. et des fractions de 0,75 ml sont collectées.

Figure III-4 : Peptides produits par hydrolyse de la protéine MCl de *Methanosarcina* sp MST-A1 avec l'endoprotéinase Arg-C.

300 nM de l'hydrolysats sont déposées sur la colonne. Le débit est de 2 ml/min. et des fractions de 0,8 ml sont collectées.

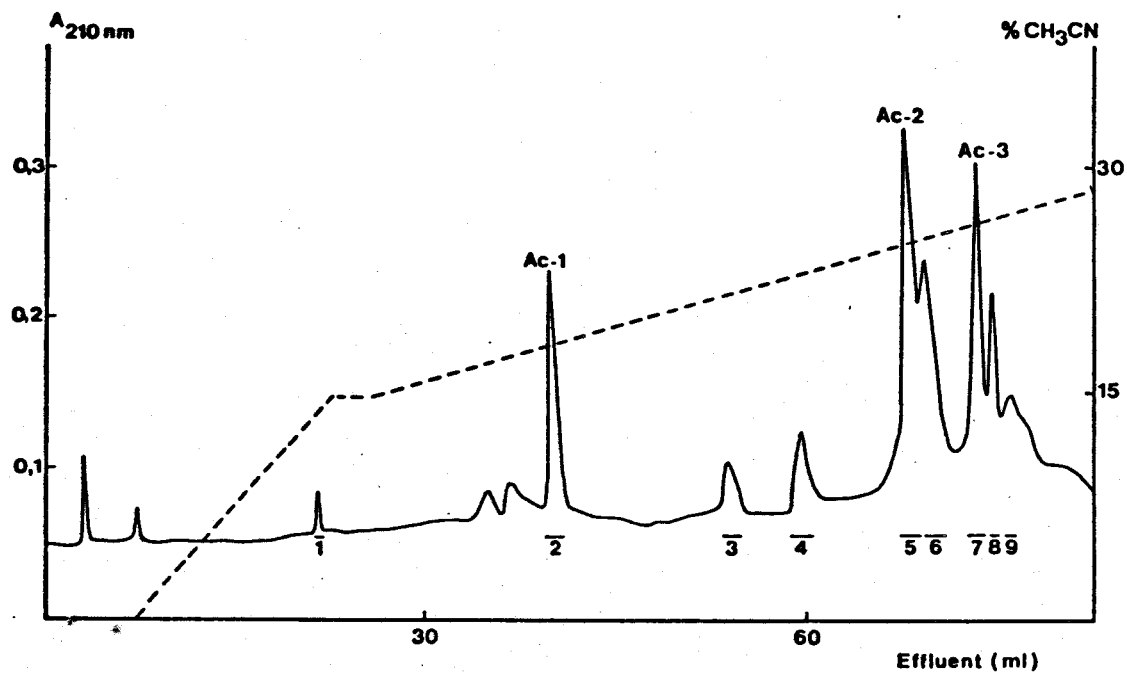


Figure III-3

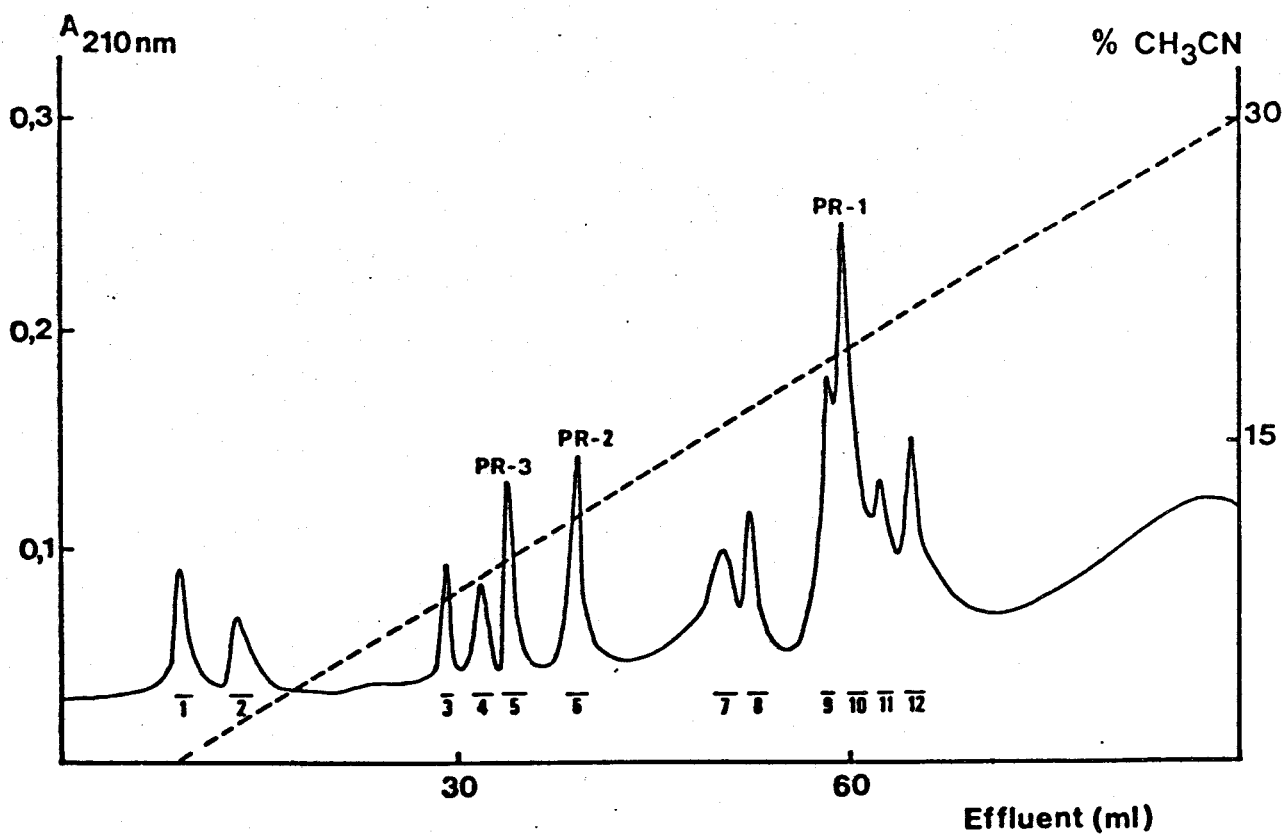


Figure III-4

observées lors de la détermination de la séquence de la protéine MCl de *Methanosarcina barkeri* (publication n°1). Par contre la coupure de la liaison Lys³⁹-Ala⁴⁰ dans la protéine MCl de *Methanosarcina* sp MST-A1 est incomplète.

* Coupure au résidu de méthionine

Le clivage de la liaison méthionyle de la protéine MCl de *Methanosarcina* sp MST-A1 par le bromure de cyanogène a permis d'obtenir deux peptides CN-1 et CN-2 facilement isolés sur une colonne de Sephadex G-50 (Figure III-5). Le peptide CN-2 a été séquencé par dégradation d'Edman automatisée mais le résidu en position 88 n'a pas pu être identifié. Nous en reparlerons ultérieurement. Le peptide CN-1 a été hydrolysé avec la protéase staphylococcale. Au cours de cette hydrolyse, de nombreuses coupures partielles se sont produites et nous avons obtenus de nombreux peptides difficiles à séparer (Figure III-6). Toutefois, le peptide PE-1 (résidus 12-43) élué dans la fraction 5, le peptide PE-2 (résidus 44-49) élué dans la fraction 4 et le peptide PE-3 (résidus 64-75) élué dans la fraction 9 ont pu être séquencés par dégradation d'Edman automatisée.

Séquence complète de la protéine MCl de *Methanosarcina* sp MST-A1

La séquence complète en acides aminés de la protéine MCl de *Methanosarcina* sp MST-A1 est présentée dans la figure III-7.

La dégradation d'Edman automatisée effectuée directement sur la protéine a donné des informations de séquence jusqu'au résidu 48. Au cours de cette dégradation, les résidus en positions 16, 28, 36, 38, 40, 41, 42, 43, 44 n'ont pu être identifiés. Cependant la détermination de la séquence des fragments PE-1 et PR-2 obtenus par hydrolyse enzymatique de la protéine a permis d'identifier ces résidus (à l'exception toutefois du résidu 36) et d'établir ainsi la séquence des 48 premiers résidus de la protéine. A partir de ces éléments de structure, nous avons pu aligner les peptides Ac-1, PE-1, PE-2 et PR-2.

FIGURE III-5 - Fractionnement par chromatographie de tamisage moléculaire des peptides produits par le clivage au résidu de méthionine.

L'hydrolysate (500 nM, dissous dans 1,4 ml de chlorhydrate de guanidine 6 M, 0,01 M HCl) est déposé sur une colonne de Sephadex G-50 superfin (90 x 1,6 cm). La colonne est équilibrée et éluee avec de l'acide chlorhydrique 0,01 M saturé avec du chloroforme. Des fractions de 1,2 ml sont collectées à un débit de 4 ml/h. (—), absorbance à 220 nm.

FIGURE III-6 - Diagramme d'élution des peptides obtenus après hydrolyse du peptide CN-1 (350 nM) avec la protéase staphylococcale.

La chromatographie liquide haute performance en phase inverse est effectuée sur une colonne C₁₈ μ Bondapak dans les mêmes conditions que celles indiquées dans les figures III-3 et III-4. Le débit est de 2 ml/min. et des fractions de 0,8 ml sont collectées.

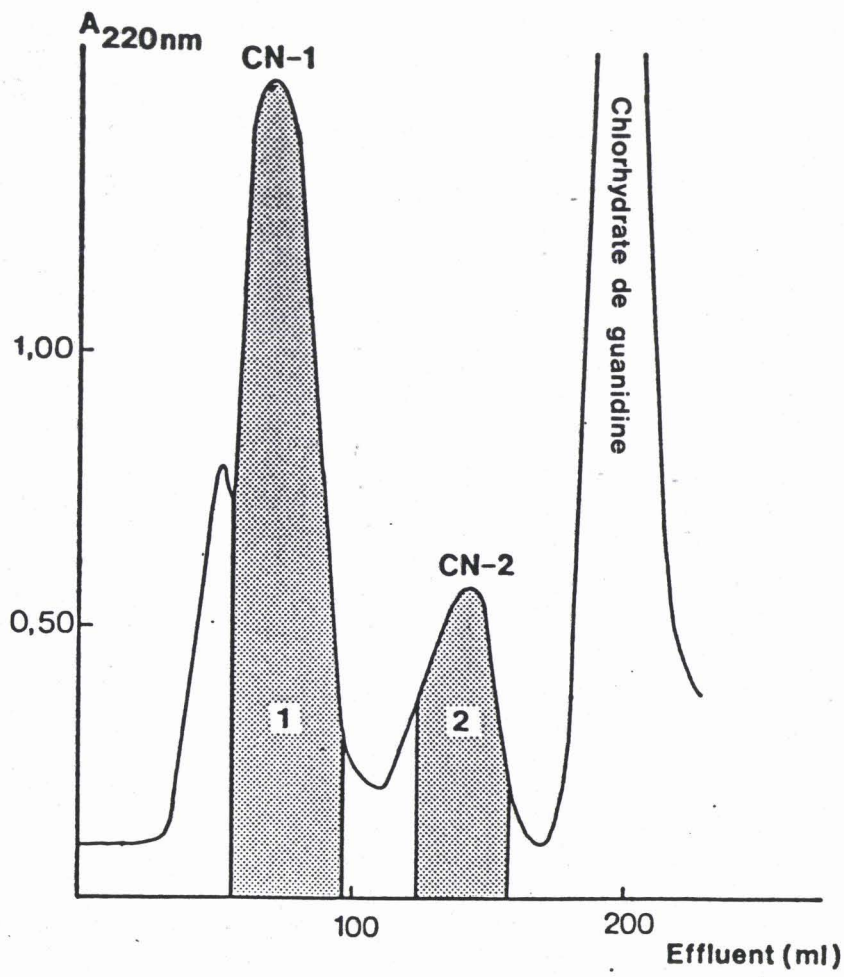


Figure III-5

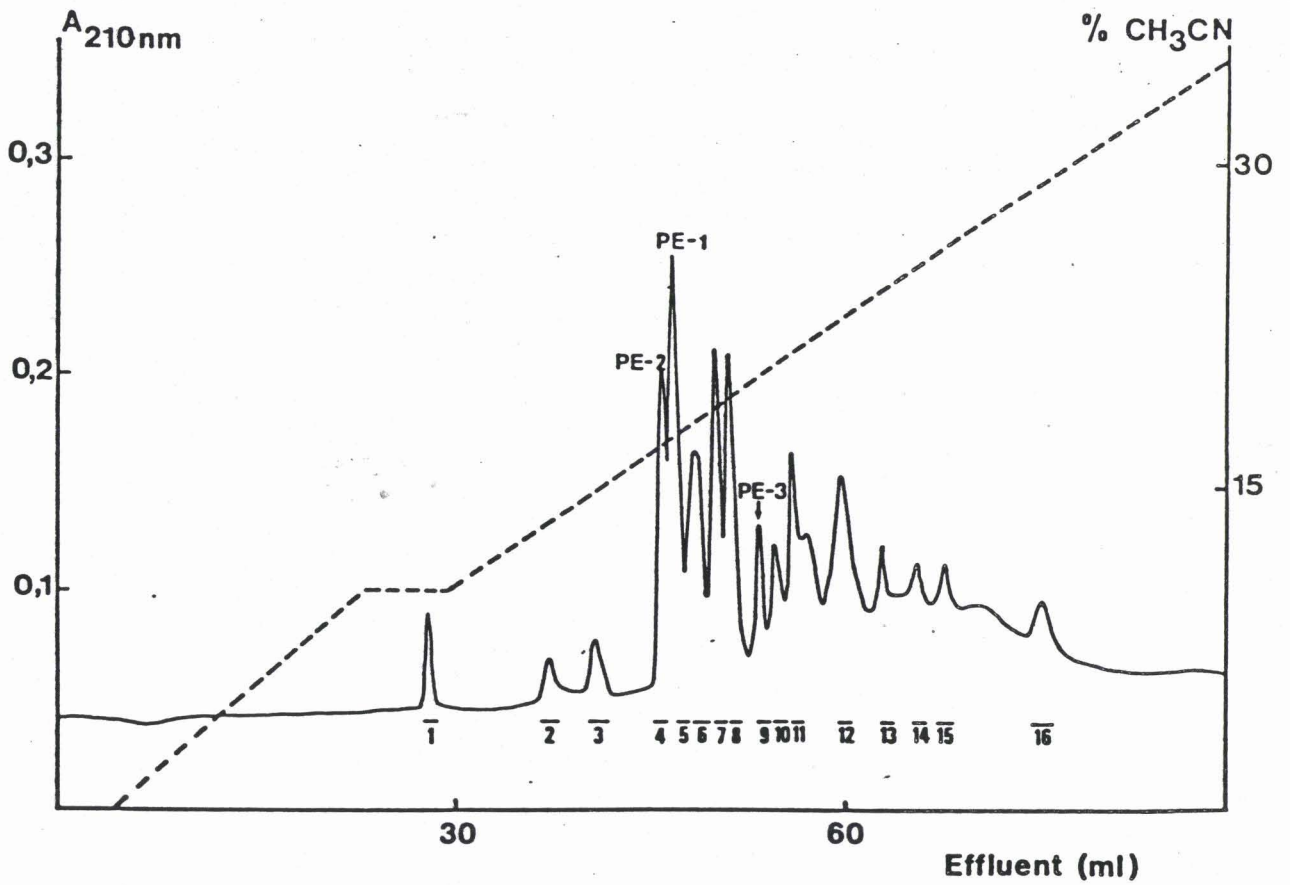


Figure III-6

Le peptide PE-2 recouvre le résidu d'acide glutamique en position amino-terminale du peptide PR-3. Bien que très court, ce recouvrement est suffisant pour placer le peptide PR-3 et établir la séquence jusqu'au résidu 59 car la liaison Arg-Glu ne se rencontre qu'une seule fois dans la protéine. De plus, la composition du peptide PR-3 caractérisée par la présence d'histidine, de phénylalanine et de lysine, qui constituent autant d'acides aminés marqueurs, présente des analogies évidentes avec la composition du fragment Ac-2 qui recouvre la séquence 36 à 65. Pour des raisons non encore établies la séquence du fragment Ac-2 n'a pas pu être déterminée directement. Ce fragment a donc été soumis à une hydrolyse par l'endoprotéinase Arg-C qui a libéré le peptide C-terminal Ac-2-PR-1 (Figure III-7).

Par ailleurs, par son recouvrement, le peptide PE-3 permet d'aligner les peptides Ac-2-PR-1 et Ac-3. Le peptide PE-3 ne recouvre le peptide Ac-2-PR-1 que par la séquence Ile-Val mais celle-ci n'est retrouvée qu'une seule fois dans la protéine. Etant dépourvu d'homosérine, le peptide CN-2 correspond à la séquence carboxy-terminale de la protéine et confirme la position C-terminale du peptide Ac-3.

Ainsi sur les 93 résidus déterminés dans la composition en acides aminés de la protéine (Tableau III-1), 91 ont été identifiés. Les deux acides aminés manquants correspondent aux positions 36 et 88 dans la séquence de la protéine. Concernant le résidu en position 36, lors de la dégradation d'Edman automatisée de la protéine et des peptides PR-2 et PE-1, le dérivé phénylthiohydantoïne de cet acide aminé est élué en CLHP en phase inverse, entre le résidu d'histidine et de tyrosine. Il semblerait que ce résidu soit un acide aspartique modifié puisque l'hydrolyse de la protéine par l'acide acétique a produit le peptide Ac-2 (résidus 36 à 65). Ce peptide contient deux résidus d'acide aspartique (ou asparagine). Un résidu d'asparagine ayant été placé en position 41 grâce aux peptides PE-1

et PR-2, un résidu d'acide aspartique n'a pas été énucléé lors de la coupure à l'acide acétique. Ce résidu se place en position amino-terminale car le peptide Ac-2-PR-1 est dépourvu d'acide aspartique. La présence de ce résidu d'acide aspartique modifié pourrait expliquer le blocage de la dégradation d'Edman.

Par ailleurs, il est vraisemblable que le résidu en position 88 soit un résidu de lysine modifié compte tenu de la composition en acides aminés de la protéine dont 13 des 14 résidus de lysine ont été identifiés. Néanmoins, ce résidu n'a pas pu être identifié avec certitude dans les peptides CN-2 et le peptide 84-93 (CN-2-ch1) obtenu après hydrolyse du peptide CN-2 par la chymotrypsine.

Afin d'obtenir des informations sur la nature des résidus en positions 36 et 88, nous envisageons de préparer de courts fragments peptidiques renfermant ces résidus pour les analyser ultérieurement en spectroscopie de masse.

III - 2-4 DISCUSSION

La protéine MC1 des deux bactéries *Methanosarcina* sp MST-A1 et *M. barkeri* MS contient 93 résidus d'acides aminés. Chaque résidu d'une protéine peut être aligné avec le résidu correspondant de l'autre protéine (Figure III-8) et les séquences en acides aminés de ces protéines diffèrent seulement par 10 résidus (soit 10,7 %). Excepté pour les résidus inconnus en positions 36 et 88, toutes ces différences peuvent être expliquées par un changement d'une seule base dans les codons correspondants et 6 d'entre elles sont conservatives. Le remplacement d'un résidu acide (Asp⁷³) par un résidu neutre (Ala⁷³) est la seule substitution non conservative.

Par rapport à la protéine de *Methanosarcina* sp souche CHTI (isolée et séquencée par F.CHARTIER), la protéine MC1 de *Methanosarcina* sp MST-A1 diffère uniquement par deux résidus d'acides aminés : le résidu d'acide glutamique en position 43 et le résidu non encore identifié (lysine modifiée ?) en position 88. Ces résidus sont remplacés respectivement par un résidu d'acide aspartique et un résidu d'arginine dans la protéine MC1 de la souche CHTI. Ces deux changements conservatifs qui distinguent la protéine MC1 de *Methanosarcina* sp MST-A1 de la protéine MC1 de *Methanosarcina* sp CHTI établissent de manière irréfutable que les deux souches sont distinctes. On peut ici apprécier l'apport de la séquence peptidique à la taxonomie car dans cet exemple précis ces deux bactéries ne peuvent être différenciées sur des critères génotypiques et phénotypiques.

La protéine MC1 de *Methanosarcina barkeri* MS et de *Methanosarcina* sp MST-A1 ont les mêmes caractéristiques générales de structure primaire. La protéine MC1 est caractérisée par une quantité importante en résidus chargés (15 % de résidus acides et 26,8 % de résidus basiques) qui sont distribués tout le long de la chaîne polypeptidique ; néanmoins certaines régions (10-15) sont acides tandis que d'autres sont basiques (46-56). La protéine MC1 se caractérise également par l'absence de domaines hydrophobes mais la plupart des résidus hydrophobes sont associés en doublets ou triplets.

La comparaison de la séquence en acides aminés de la protéine MC1 avec celles d'autres protéines chromosomales eucaryotes et procaryotes a fait l'objet de la publication n°2. La protéine MC1 de *M.barkeri* se différencie des protéines chromosomales d'eubactéries et d'eucaryotes par l'absence de domaines hydrophobes. De plus elle se différencie des protéines II d'eubactéries par sa quantité importante en résidus chargés et par sa charge nette relativement élevée. Bien que son caractère basique la rapproche des histones et de la protéine HTa de *T.acidophilum*, elle ne montre aucune homologie de séquence avec ces protéines.

CHAPITRE IV

CONFORMATION DE LA PROTÉINE MC1 DE *METHANOSARCINA BARKERI*

IV - 1 INTRODUCTION

La séquence en acides aminés d'une protéine contient l'information nécessaire pour lui imposer sa conformation native. Cette conformation désigne, en fait, l'ensemble des structures secondaire, tertiaire et quaternaire. Trois types de structure secondaire ont pu être définis : l'hélice α , les feuillets plissés β et le coude β . Les deux premières structures sont parfaitement caractérisées tandis que plusieurs types de coudes β ont été décrits et classés dans une famille de structures appelées boucles (MILNER-WHITE et POET, 1987). La chaîne peptidique peut aussi se replier sur elle-même pour produire différentes conformations, telle la structure globulaire, constituant la structure tertiaire. Enfin, un autre niveau d'ordre structural dans les protéines est dû à l'interaction des chaînes protéiques entre elles conduisant à la structure quaternaire.

Ainsi la conformation d'une protéine va déterminer son activité biologique. L'étude de la structure tridimensionnelle de la protéine MCl pourrait donc apporter des éléments d'information importants concernant son rôle dans le complexe nucléoprotéique.

La structure tridimensionnelle des protéines se détermine avec précision par l'analyse de la diffraction des rayons X par des cristaux de ces protéines. La cristallisation des protéines est un travail long, nécessitant des concentrations importantes de protéines et de caractère aléatoire : les protéines de structure globulaire cristallisent facilement, tandis que les protéines ayant des domaines flexibles, comme les histones et la protéine MCl, ne peuvent être cristallisées si elles ne sont pas stabilisées. Il est donc évident que pour la protéine MCl, il fallait faire appel aux différentes méthodes de prédiction de structure (CHOU et FASMAN, 1978 ; GARNIER ET COLL., 1978 ; CID et coll., 1982 ; SETTE et coll., 1986 ; GABORIAUD et coll., 1987) susceptibles d'apporter quelques éléments d'information sur la structure probable de cette protéine.

Nous avons également tenté une approche de la conformation par des méthodes spectroscopiques : dichroïsme circulaire, spectroscopie infrarouge et spectroscopie de fluorescence, chacune de ces méthodes apportant des informations complémentaires.

Par ailleurs, nous avons vu au chapitre III que la séquence en acides aminés de la protéine MCl ne présentait pas d'homologies avec la protéine II isolée d'eubactéries et de l'archaebactérie *Thermoplasma acidophilum* et avec l'histone H4 de thymus de veau. La méthode comparative de GABORIAUD et coll. (1987) permet de retrouver des parentés entre des protéines qui, malgré une fonction analogue, présentent de faibles homologies de structure primaire. Nous avons donc utilisé cette méthode pour comparer la protéine MCl avec d'autres protéines chromosomales (protéine II de différents micro-organismes et histone H4 de thymus de veau).

L'ensemble des travaux qui font l'objet de ce chapitre ont été menés en collaboration avec :

- le Docteur J.P.MORNON de l'Université Pierre et Marie Curie (PARIS VI) pour le développement des méthodes de prédiction de structure (GARNIER et coll., 1978 ; GABORIAUD et coll., 1987).
- les Docteurs N.HELBECQUE (Département de Biochimie, Faculté de Médecine de LILLE II) et J.P.HENNICHART (Unité 16 INSERM) pour l'application des méthodes spectroscopiques.

IV - 2 MÉTHODES DE PRÉDICTION DE STRUCTURE SECONDAIRE

Les méthodes de prédiction de structure secondaire à partir de la séquence en acides aminés, considèrent que l'environnement de chaque acide aminé détermine la conformation du segment polypeptidique où est situé cet acide aminé. Bien que ces méthodes ne soient fiables qu'à 50 ou 60%, elles apportent des renseignements utiles pour une première approche de la conformation.

IV - 2-1 PRINCIPE DES METHODES DE PREDICTION UTILISEES

IV - 2-1-1 Méthode de CHOU et FASMAN (1978)

C'est une des méthodes les plus employées car la micro-informatique a rendu son utilisation aisée ; de plus, elle permet la prédiction des trois structures organisées : l'hélice α , le feuillet plissé β et le coude β . La probabilité (P_α , P_β , $P_{\beta t}$) que présente chacun des vingt acides aminés de se trouver dans une structure en hélice α , en feuillet β ou en coude β est déterminée selon leur fréquence d'apparition dans des conformations de protéines dont la structure tertiaire est connue par cristallographie.

La chaîne polypeptidique à étudier est divisée en blocs de quatre résidus d'acides aminés. Les probabilités moyennes P_α^m , P_β^m , $P_{\beta t}^m$, pour chaque résidu i dans ce bloc, sont déterminées en tenant compte des résidus voisins. Ces probabilités moyennes sont calculées à partir des valeurs respectives P_α , P_β ou $P_{\beta t}$ de chaque résidu (CHOU et FASMAN, 1978). La séquence de la protéine est ainsi "balayée" pour rechercher les groupes d'acides aminés susceptibles d'être en hélice α ou en feuillet β ou en coude β . La confrontation des résultats obtenus en cristallographie et par cette méthode montre que le pourcentage de structure prédite correctement est de 50% (STERNBERG, 1986).

IV - 2-1-2 Méthode de GARNIER et coll. (1978)

La méthode de GARNIER et coll. est basée sur le même principe que celui de la méthode de CHOU et FASMAN mais, dans ce cas, l'environnement de chaque acide aminé à la position i est pris en compte en considérant l'effet des résidus de la position $i - 8$ à $i + 8$.

Le pourcentage de structure prédite correctement est de 56%, il est donc supérieur à celui obtenu par la méthode de CHOU et FASMAN (STERNBERG, 1986).

IV - 2-1-3 Méthode de CID et coll. (1982)

CID et ses collaborateurs tiennent compte de la corrélation qui existe entre le caractère hydrophobe de résidus d'acides aminés et leur accessibilité au sein de la structure, pour prédire la structure secondaire d'une protéine. La prédiction est basée sur la forme de quatre types de profils d'hydrophobicité. Les considérations suivantes sont également prises en compte :

- les segments en hélice α sont généralement localisés en surface de la protéine et présentent des surfaces hydrophiles et hydrophobes.
- les segments en feuillet β ont tendance à être enfouis à l'intérieur de la protéine. Ceux qui sont en surface montrent une alternance de résidus hydrophobes et hydrophiles.
- une séquence de deux résidus hydrophiles interromperait une structure en feuillet β .

IV - 2-1-4 Méthode de SETTE et coll. (1986)

La méthode de SETTE recherche le caractère amphipathique de séquences protéiques : les régions amphipathiques sont des régions montrant une alternance de surfaces hydrophobes et de surfaces hydrophiles. Ainsi des périodicités dans le caractère polaire/apolaire d'une séquence en acides aminés sont recherchées en attribuant à chaque résidu une valeur d'hydrophobicité. La périodicité est exprimée par un angle δ . Cet angle est d'environ 100° pour une structure en hélice α et compris entre 160° et 180° pour une structure en feuillet β . SETTE et ses collaborateurs ont mis au point un programme sur micro-ordinateur permettant de déterminer le caractère amphipathique d'une séquence en acides aminés. Ce caractère peut être calculé selon trois échelles différentes d'hydrophobicité. Dans cette étude nous avons appliqué l'échelle d'EISENBERG et coll. (1984) qui est la plus utilisée. L'hydrophobicité de chaque résidu d'acide aminé est déterminée en tenant compte des résidus adjacents : en général on examine la séquence de la protéine par segments de sept acides aminés.

IV - 2-1-5 Méthode de GABORIAUD et coll. (1987)

Cette méthode d'analyse des amas hydrophobes appelée méthode HCA (Hydrophobic Cluster Analysis) a été mise au point dans le Laboratoire de Minéralogie-Cristallographie (J.P. MORNON).

Alors que la plupart des méthodes de prédiction de structure citées précédemment analysent la chaîne polypeptidique sous forme de séquence alignée, la méthode HCA tient compte des interactions des acides aminés dans l'espace. Le but principal est de différencier les zones hydrophobes en amas (sur une hélice α ou un feuillet β) et les zones hydrophiles ou indifférentes les entourant. Pour cela, la méthode HCA utilise une représentation des séquences de protéines en hélice α . Cette représentation avait été proposée par LIM en 1974, l'hypothèse étant que le polypeptide synthétisé au niveau des ribosomes était sous forme d'une hélice α avant que la protéine n'acquiert son organisation "définitive".

Pratiquement, la séquence en acides aminés est représentée en hélice α sur un cylindre en plaçant 3,6 résidus par tour. Pour rendre la représentation plus facile, le cylindre est coupé parallèlement à son axe et déroulé dans le plan du papier. Pour recréer l'environnement des acides aminés éloignés par le dépliement du cylindre et pour suivre plus facilement la séquence, la représentation est dupliquée (Figure IV-1a,b,c).

Cette représentation permet de mettre en évidence des amas de résidus hydrophobes séparés les uns des autres par des zones de transition. Dans ces zones de transition, on met en évidence les résidus de proline et de glycine qui favorisent les boucles; on peut aussi y faire figurer des amas de résidus hydrophiles.

La protéine peut alors être découpée en blocs matérialisant les segments structurés en hélice α ou en feuillet β . Ces deux structures sont déterminées d'après la morphologie et l'orientation des amas hydrophobes à partir de statistiques effectuées sur des protéines dont la structure tridimensionnelle est connue (Figure IV-1d).

FIGURE IV-1 - Principe de l'analyse de séquence de protéine par la méthode HCA.

L'exemple représente une partie de la séquence de l' α_1 antitrypsine humaine.

²⁴⁶GNATAIFFFLPDEGK²⁶⁴LQHLE d'après GABORIAUD et coll., 1987.

IV - 2-2 APPLICATION DES METHODES DE PREDICTION DE STRUCTURE
SECONDAIRE A LA PROTEINE MC1 DE *METHANOSARCINA BARKERI*

Quatre méthodes décrites ci-dessus (CHOU et FASMAN, 1978 ; GARNIER et coll., 1978 ; SETTE et coll., 1986 ; GABORIAUD et coll., 1987) ont été choisies en fonction de leur fiabilité et de leur complémentarité. Elles ont été appliquées individuellement et la confrontation de l'ensemble a permis de déduire la structure la plus probable selon les règles suivantes:

- un segment est considéré en hélice α ou en feuillet β si l'une de ces structures est prédite par au moins deux méthodes
- une hélice α ou un feuillet β sont pris en considération quand ils contiennent respectivement au moins quatre résidus en conformation α ou au moins trois résidus en conformation β
- un coude β est pris en considération quand il est déterminé, à la fois, par les méthodes de CHOU et FASMAN et de GARNIER et coll.
- lorsque pour un même segment, deux méthodes prédisent de la structure en hélice α et deux autres de la structure en feuillet β , la méthode de CID et coll. (1982) est utilisée pour lever l'ambiguïté (MANAVALAN et coll., 1986).

Parmi les méthodes de prédiction qui ont été présentées, la méthode HCA, récemment proposée, se distingue par son principe et la perception immédiate des zones structurées qu'elle permet d'avoir. Nous avons donc jugé utile de développer plus particulièrement les informations structurales obtenues sur la protéine MC1 par cette méthode (Figure IV-2).

Si l'on se réfère à la position des résidus de glycine et de proline, la protéine MC1 peut être divisée en 11 segments structuraux (S1-S11). Parmi ceux-ci, 7 amas hydrophobes (C1-C7) ont pu être mis en évidence. La morphologie et l'orientation des amas C1, C2, C3, C5 et C6 indiquent une forte probabilité de structure en feuillet β . Par contre il n'est pas possible d'assigner avec certitude une structure pour les amas C4 et C7.

Les résultats de chacune des quatre méthodes de prédiction ainsi que la synthèse que l'on peut en faire sont présentés dans la figure IV-3.

La prédiction "de synthèse" indique que la protéine comprendrait deux segments de structure α -hélicoïdale (séquences 26-32 et 59-66) et quatre segments de structure en feuillets β (séquences 6-8, 17-19, 44-47 et 55-58). Les structures en hélice α et en feuillet β représentent 16% et 15% des résidus de la protéine.

Il faut souligner par ailleurs, le nombre relativement important de coudes β dans lesquels 30% des résidus de la molécule seraient impliqués.

Toutes les méthodes utilisées sont concordantes excepté pour la région 55-66. D'abord, on observe une divergence au niveau du dipeptide Val-Phe. La méthode de CID et coll. (1982) prédit que ces deux résidus sont en conformation β , car le profil d'hydrophobicité de la séquence 55-58 est caractéristique d'un feuillet β enfoui. D'autre part, les résultats des méthodes 1 et 3 qui prédisent une structure β pour le segment 52-54 n'ont pas été pris en considération dans la prédiction "de synthèse" car selon la méthode de CID, les deux résidus de lysine adjacents en positions 53 et 54 interrompraient une structure en feuillet β .

En conclusion, on trouverait dans la région N-terminale de la protéine (résidus 1-42) deux feuillets β (résidus 6-8 et 17-19) et une hélice α (résidus 26-32) située entre deux coudes β induits par les résidus de glycine en positions 21 et 35. La partie médiane (résidus 43-66) s'avère très bien structurée : un feuillet β (résidus 44-47) et une plus longue séquence structurée constituée d'un feuillet β et d'une hélice α (résidus 55-66), les deux feuillets β étant séparés par un coude β . Au contraire, la partie C-terminale de la protéine serait peu organisée et perturbée par des résidus de proline initiateurs de coudes β .

Cependant il faut considérer les hélices α avec une certaine prudence car ces structures sont attribuées à des segments comprenant des résidus d'alanine qui sont souvent responsables de surestimation de structure α -hélicoïdale (J.P.MORNON, comm.pers.).

Pour tenter de lever cette incertitude, nous avons appliqué aux régions 26-32 et 59-66 la méthode de WU et KABAT (1971). Si un segment est en structure α , on doit trouver sur sa représentation en forme de roue une surface hydrophile et une surface hydrophobe.

La séquence 26-32, ne présentant pas de surface hydrophobe ne serait pas en hélice α (Figure IV-4a). Par contre, la séquence 59-66 aurait une plus forte probabilité d'être en hélice α , car on observe une surface hydrophobe (Trp⁶¹, Met⁶⁴ et Val⁶⁵) et une surface hydrophile (Lys⁵⁹, Lys⁶², Glu⁶³ et Glu⁶⁶) (Figure IV-4b).

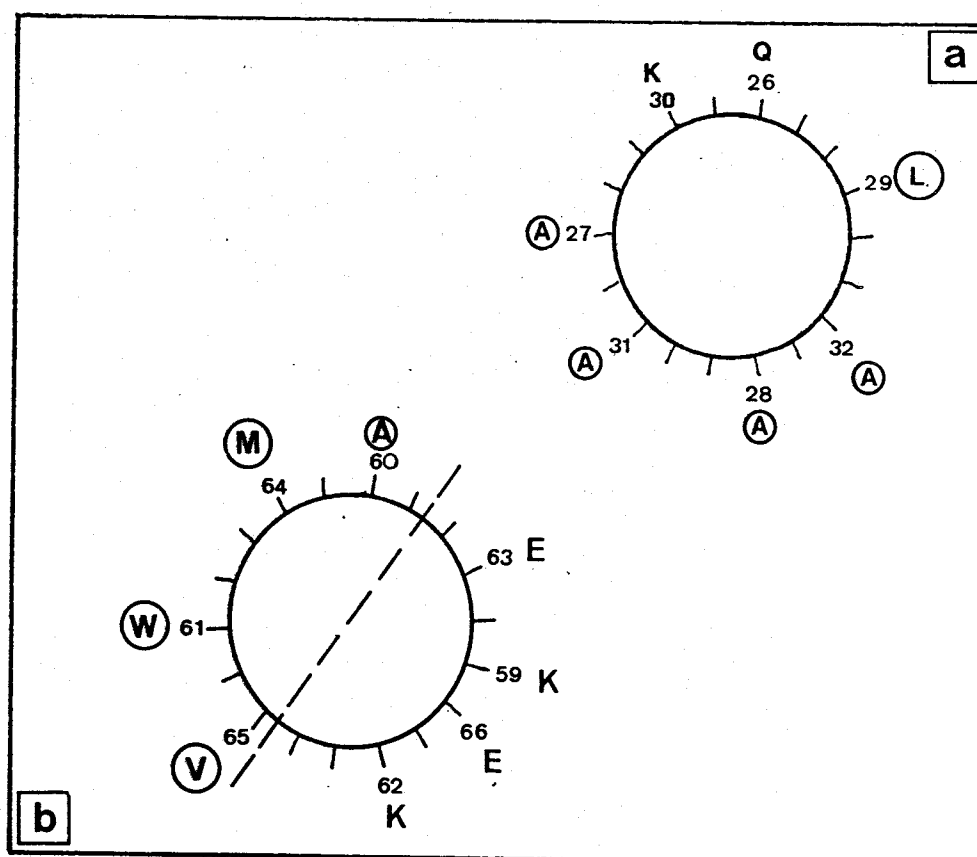


FIGURE IV-4 - Représentation en roue hélicoïdale des séquences 26-32 (a) et 59-66 (b) de la protéine MC1 de *M. barkeri* selon la méthode de WU et KABAT (1971).

IV - 3 MÉTHODES SPECTROSCOPIQUES

IV - 3-1 DICHROÏSME CIRCULAIRE

Le dichroïsme circulaire permet en théorie d'apprécier la structure secondaire d'une protéine, mais en fait le signal dichroïque obtenu correspond à la somme des signaux des différentes structures contenues dans la protéine (hélice α , feuillet β , coude β et structure inorganisée).

En pratique, on constate que plus il y a d'hélices α dans une protéine, plus il est facile de déterminer sa structure secondaire de façon précise (AUBERT, 1981). Au contraire, il est plus difficile de déterminer la structure secondaire d'une protéine riche en feuillets β car le signal dichroïque de cette structure est très faible (Figure IV-5). La détermination de la structure β est beaucoup plus précise si l'on travaille à très basse longueur d'onde c'est-à-dire entre 200 et 150 nm (BRAHMS et BRAHMS, 1980) ce qui nécessite un appareillage particulier dont nous n'avons pas pu disposer pour nos expérimentations.

** Matériel et méthodes*

La protéine MCl a été préparée selon la méthode décrite p. 46 et le fragment 1-64 a été obtenu par coupure de la protéine au bromure de cyanogène. Les méthodes de coupure de la protéine et de purification de ce peptide ont été décrites dans la publication n°1. La concentration des échantillons a été déterminée par analyse acide totale.

Les spectres dichroïques de la protéine et du fragment 1-64 ont été enregistrés à différentes concentrations en fluorure de sodium (NaF) dans du tampon phosphate 10 mM pH 6,0 et l'étude a été menée dans le domaine spectral 250-190 nm.

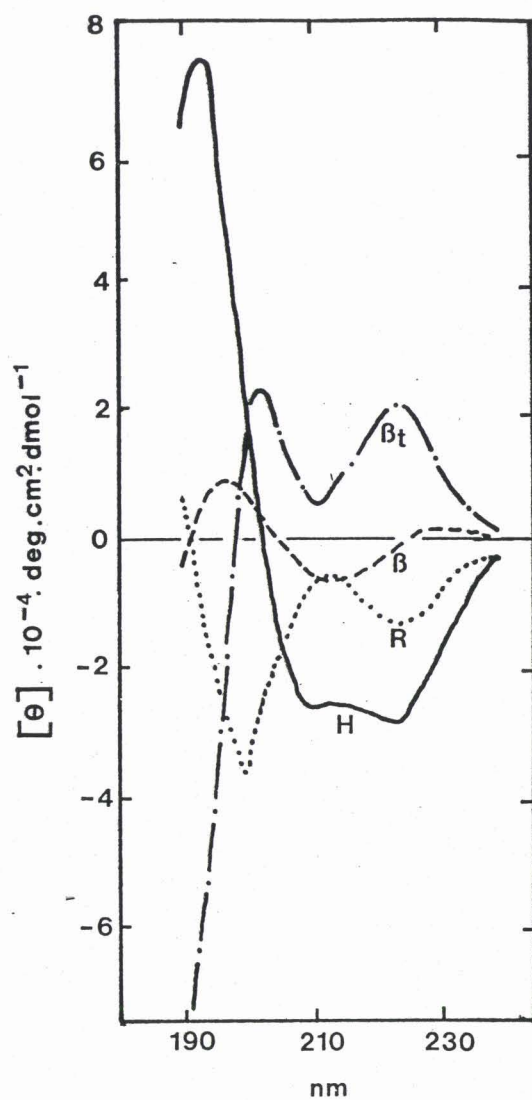


FIGURE IV-5 - Spectres de références de la structure en hélice α (H), de la structure en feuillet β (β), des coudes β (β_t) et de la structure inorganisée (R) (CHANG et coll., 1978)

Les spectres dichroïques ont été obtenus à l'aide d'un dichrographe JOBIN-YVON RJ MARK III avec des cuves de 0,01 cm à 1 cm de trajet optique (laboratoire de Biophysique U. 124 INSERM, C.DENIS, analyste).

L'ellipticité molaire $[\theta]$ a été calculée à partir du spectre expérimental par la formule :

$$[\theta] = 3300. \frac{M}{c \cdot d} \cdot \Delta \cdot s \text{ en deg. } \text{dmol}^{-1} \text{ cm}^2$$

M : masse moléculaire moyenne des résidus d'acides aminés

C : concentration en g/l

d : épaisseur de la cuve en cm

Δ : intensité du signal obtenu mesuré en mm.

s : sensibilité de l'appareil

* Résultats et discussion

Les spectres dichroïques de la protéine MCl sont présentés dans la figure IV-6. A la molarité de 50 mM NaF, le spectre dichroïque montre une bande négative au voisinage de 200 nm. L'ellipticité à 222 nm, caractéristique de la structure α , est très faible. Des spectres similaires ont été obtenus avec différentes concentrations de protéine MCl (0,13 à 1,2 mg/ml) ou quand la protéine est dissoute dans de l'eau.

L'influence de la force ionique sur la conformation de la protéine a été étudiée en enregistrant les spectres dichroïques à différentes molarités en fluorure de sodium (Figure IV-6). L'addition de ce sel entraîne une diminution de la bande négative à 200 nm et à partir de 350 mM NaF, cette bande glisse vers 210 nm. Par contre l'ellipticité à 222 nm reste constante même en présence d'autres sels (NaCl, CaCl₂).

Ces résultats indiquent que la protéine MCl ne contient pas ou peu de structure en hélice α . Aucune autre forme structurale n'a pu être mise en évidence à cause des limitations imposées par l'appareil. Toutefois, les variations de spectres observées sous l'action des sels indiquent que la force ionique induit un changement de conformation.

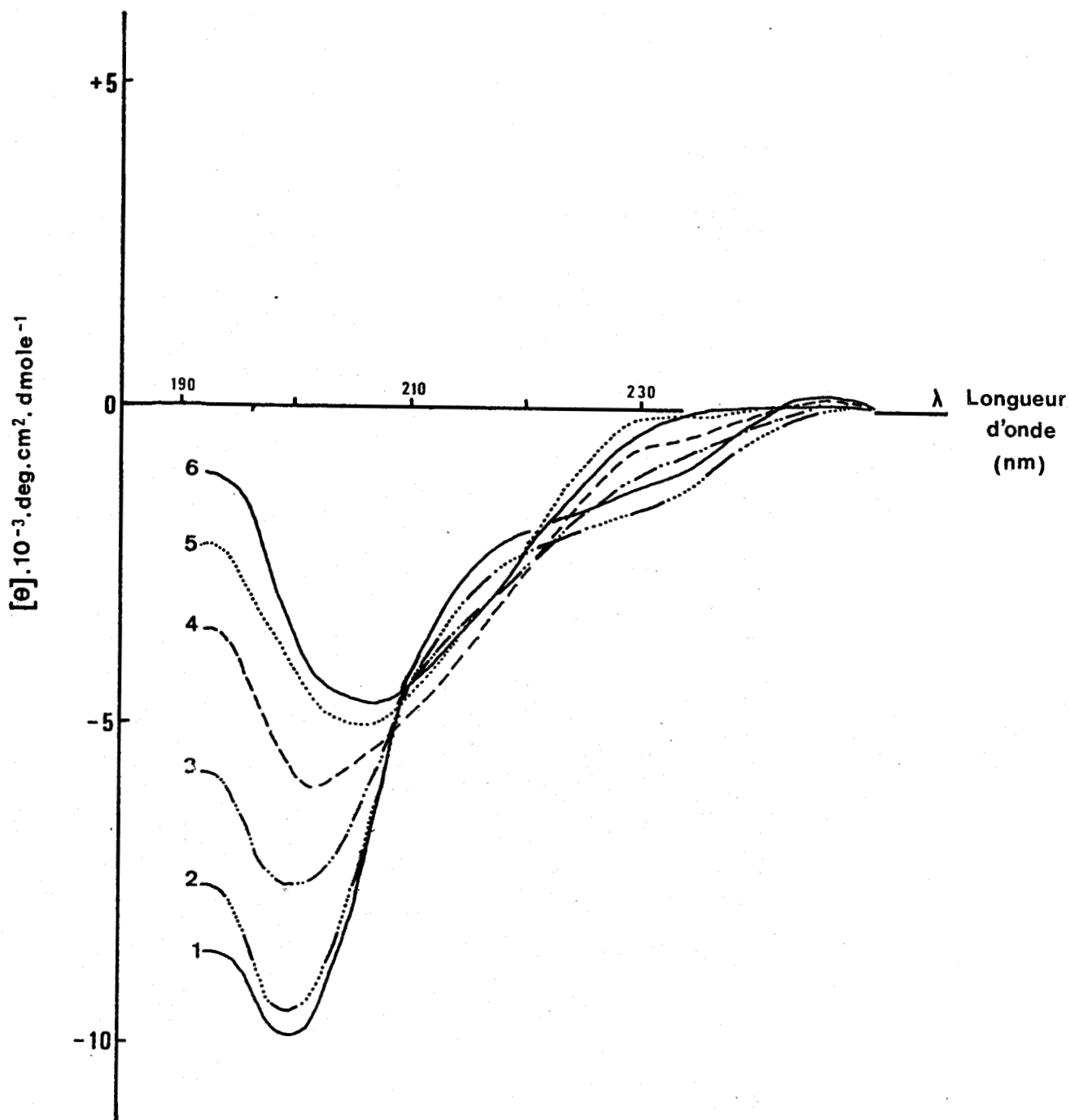


FIGURE IV-6 - Spectres dichroïques, entre 250 et 190 nm, de la protéine MCl (1,2 mg) dissoute dans 1 ml de tampon phosphate de sodium 10 mM pH 6,0 à différentes molarités en NaF : (1) 0 mM, (2) 50 et 100 mM, (3) 250 mM, (4) 350 mM, (5) 450 mM, (6) 550 mM.

Par ailleurs, nous devons garder à l'esprit que le spectre observé représente la somme de toutes les formes structurales de la protéine. Ainsi, on peut supposer que le signal caractéristique de la structure en hélice α peut être masqué par un signal induit par des coudes β . Or, nous avons vu dans les méthodes de prédiction de structure que la partie carboxy-terminale de la protéine serait peu organisée et contiendrait trois coudes β . Cette hypothèse nous a donc conduit à étudier les spectres dichroïques du fragment 1-64 (Figure IV-7). Les spectres de ce peptide sont semblables à ceux de la protéine MCl à une molarité en sel de 50 mM NaF.

Ceci montre que la faible ellipticité de la protéine MCl observée à 222 nm ne résulte pas, d'un masquage du signal de l'hélice α , par le signal produit par les coudes β du tiers C-terminal de la molécule.

Par ailleurs, aucun changement de spectre n'est observé si on augmente la force ionique. En fait, en coupant la protéine au résidu de méthionine en position 64, pour obtenir le fragment 1-64, nous avons amputé la région 55-66 de deux résidus hydrophobes. Ainsi amputée, cette région perd sa structure en hélice α pour laquelle une forte probabilité existait.

Puisqu'on n'observe plus de variation du signal dichroïque pour le fragment 1-64, on peut penser que la région 55-66 est impliquée dans les changements de conformation de la protéine en fonction de la force ionique.

IV - 3-2 SPECTROSCOPIE INFRAROUGE (IR)

L'analyse spectrale IR est basée sur l'interaction entre un rayonnement électromagnétique du domaine infrarouge et la matière. L'énergie apportée par la radiation IR peut exciter les molécules et les faire passer à un niveau d'énergie vibrationnelle plus élevée : il y a alors absorption d'énergie par

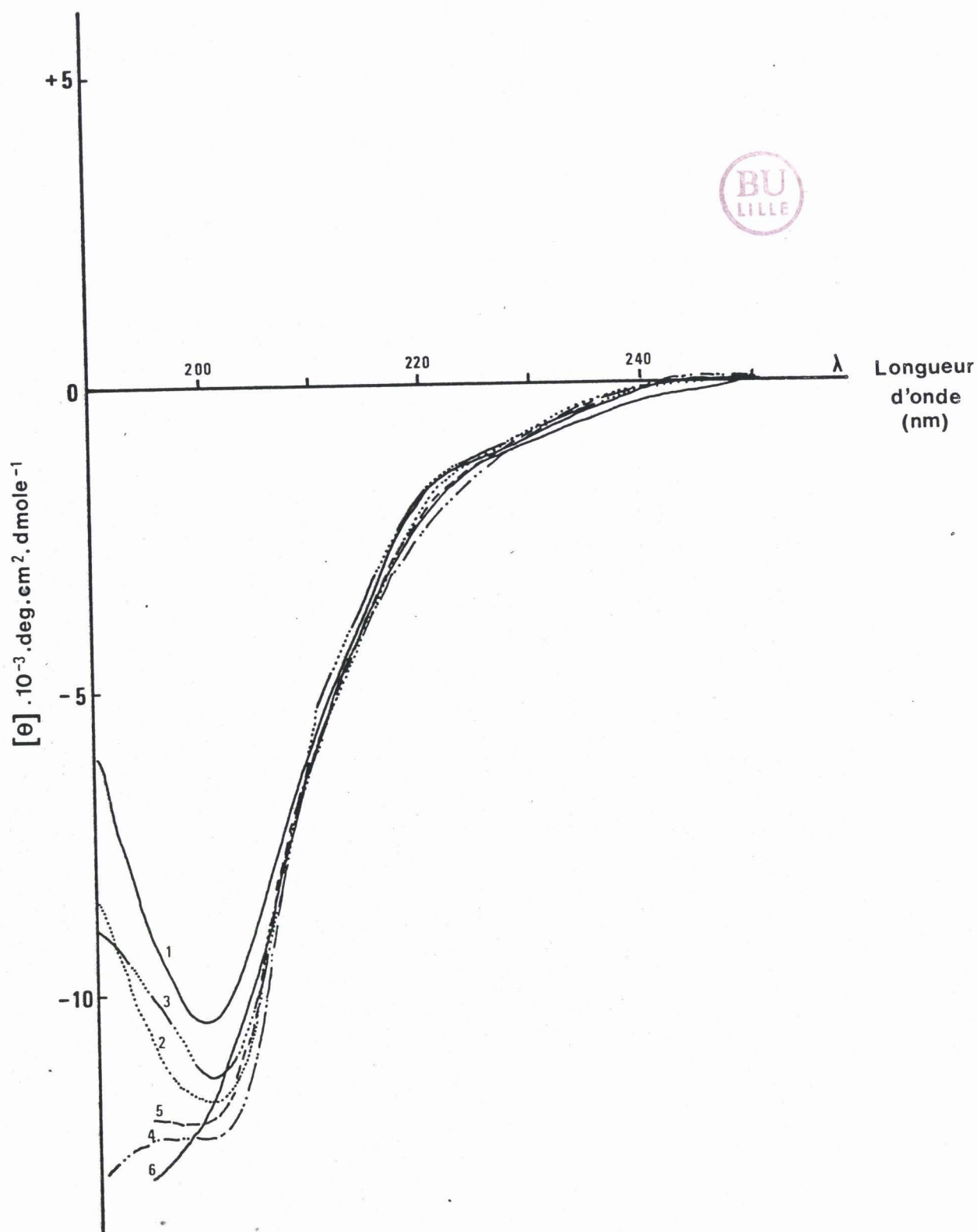


FIGURE IV-7 - Spectres dichroïques, entre 250 et 190 nm du peptide 1-64 (0,28 mg) dissous dans 1 ml de tampon phosphate de sodium 10 mM pH 6,0 à différentes molarités en NaCl : (1) 50 mM, (2) 100 mM, (3) 200 mM, (4) 300 mM, (5) 500 mM, (6) 800 mM.

la matière. Sur les spectres IR obtenus, chaque bande d'absorbance est représentative d'une vibration d'un fragment de la molécule et donc d'une liaison chimique.

Il est possible de caractériser les différentes structures secondaires des protéines grâce à deux bandes d'absorption intenses localisées autour de 1650 et 1550 cm^{-1} correspondant respectivement aux vibrations de l'amide I (-CO-) et de l'amide II (-NH-) de la liaison peptidique (VAN der MAREL et BEUTELSPACHER, 1976).

En ce qui concerne l'absorption due à l'amide I, l'association d'une bande de faible nombre d'ondes (vers 1630 cm^{-1}) à une bande de nombre d'ondes élevée (vers 1680 cm^{-1}) permet de conclure à la présence de conformation en feuillet β . Une bande située à 1650 cm^{-1} sera indicative d'une structure en hélice α tandis qu'une bande à 1658 cm^{-1} dénotera une absence de structuration de la protéine. Cependant en pratique on constate que l'infrarouge permet essentiellement de déterminer la structure β (E.TAILLANDIER, comm.Pers.).

** Matériel et méthodes*

La protéine MCl a été préparée selon la technique décrite p. 46. Après dessalage sur une colonne de Sephadex G-25 (30 x 2,6 cm), la protéine a été dialysée contre de l'eau distillée et lyophilisée deux fois dans D_2O (milieu "transparent" en IR dans le domaine qui nous intéresse).

Les spectres IR de la protéine MCl ont été enregistrés à température ambiante à différentes concentrations en chlorure de sodium à partir d'une solution 5 M de ce sel dans D_2O .

Les spectres ont été obtenus à l'aide d'un spectrophotomètre NICOLET 7000 à transformée de Fourier avec des cellules en fluorure de baryum.

* Résultats

La figure IV-8 représente les variations de l'intensité des pics à 1617 cm^{-1} et à 1685 cm^{-1} de la bande amide I de la protéine MCl avec différentes concentrations en NaCl. L'intensité de ces pics, caractéristiques des feuillets β anti-parallèles, diminue avec la concentration en sel.

Ce résultat indique la présence de feuillets β anti-parallèles dans la protéine, à faible concentration en sel, et le dépliement de cette structure lorsque la force ionique augmente.

IV - 3-3 FLUORESCENCE

La fluorescence est une lumière émise lorsqu'une molécule passe d'un niveau électronique à un autre de plus faible énergie. Seules certaines molécules, qui en général possèdent des électrons délocalisés, émettent de la fluorescence : on les nomme fluorophores.

Dans les protéines, seuls les résidus de phénylalanine, tyrosine et surtout tryptophane contribuent à la fluorescence. Le noyau indole, et donc le tryptophane sont très sensibles à la polarité du solvant. Par conséquent, le spectre d'émission de fluorescence des résidus de tryptophane permet d'étudier la polarité de leur environnement (LAKOWICZ, 1983).

Puisque la protéine MCl ne contient pas de tyrosine et que la fluorescence due à la phénylalanine peut être éliminée par irradiation sélective, les deux résidus de tryptophane présents dans la molécule aux positions 61 et 74 seront les seuls résidus impliqués dans les phénomènes de fluorescence.

Le tryptophane en position 61 est localisé dans la région structurée 55-66 tandis que l'autre se trouve dans une région où la structure secondaire est rompue par trois résidus de proline (voir figure IV-3). L'étude de la fluorescence des résidus

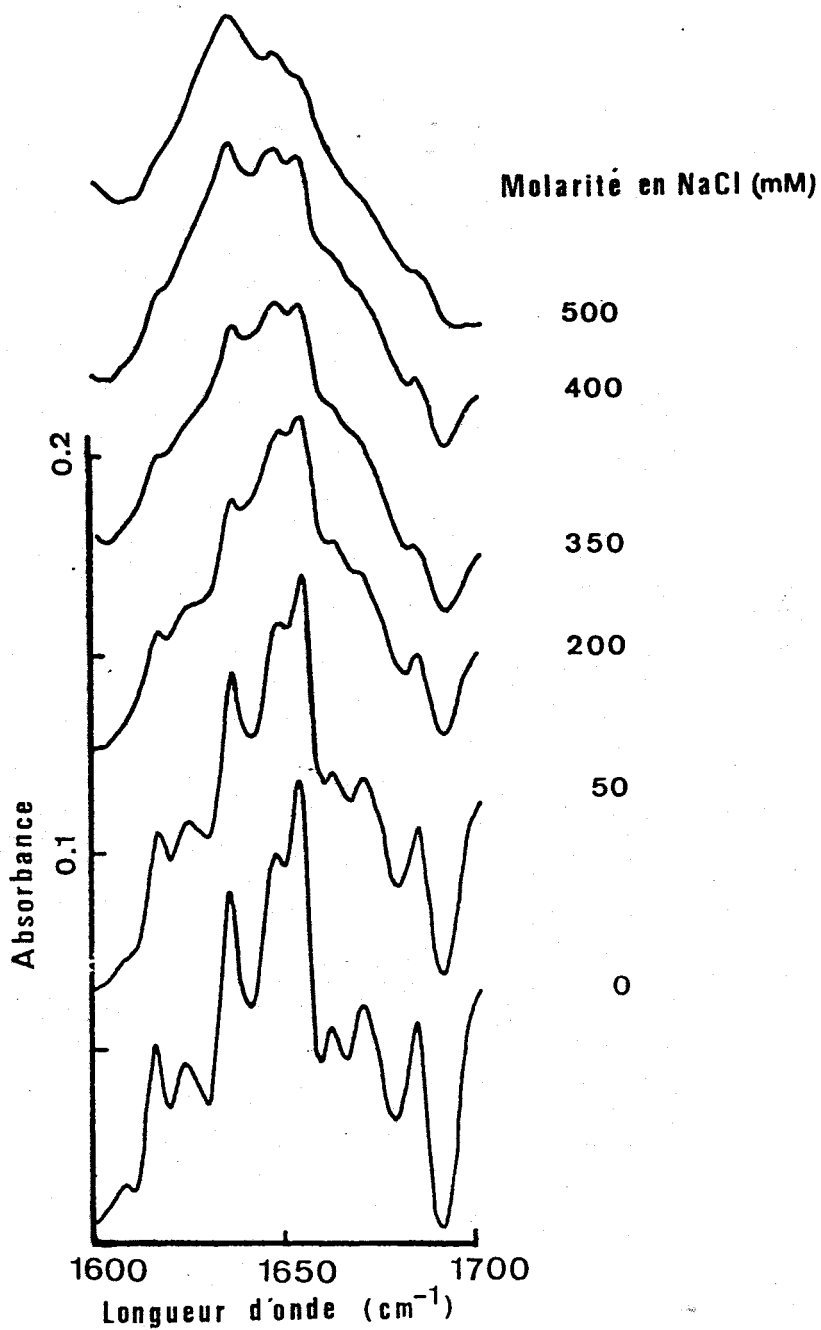


FIGURE IV-8 - Spectres infrarouges de la protéine MCl1 (11 mg/ml) dans D₂O, à différentes molarités en chlorure de sodium.

de tryptophane peut donc fournir des informations sur l'exposition des séquences adjacentes à ces résidus. En d'autres termes, cette étude constitue une approche indirecte de la conformation de ces séquences. En particulier ces études de fluorescence pourraient indiquer si la région structurée 55-66 est impliquée dans les changements conformationnels observés en dichroïsme circulaire et en infrarouge.

** Matériel et méthodes*

La protéine MCl a été préparée selon la méthode décrite p. 46 . Le fragment 1-64 a été obtenu par coupure de la liaison méthionyle 64-65 de la protéine MCl tandis que le fragment 60-93 provient de la digestion de la protéine avec l'endoprotéinase Arg-C. les méthodes pour obtenir et purifier les fragments sont décrites dans la publication n°1. La concentration de la protéine et des fragments a été déterminée par l'analyse de leur composition en acides aminés.

L'intensité de fluorescence a été mesurée sur spectrofluorimètre JOBIN-YVON JY3 en utilisant des cuves de 1 cm de trajet optique. Les échantillons (6 μ M) dissous dans du tampon phosphate de sodium 10 mM pH 6,0 ont été excités à 290 nm, afin d'éviter toute émission de fluorescence due aux résidus de phénylalanine et l'émission de fluorescence a été enregistrée entre 325 et 425 nm.

** Résultats et discussion*

Le maximum d'émission de fluorescence a été observé à 345 nm, cette longueur d'onde a donc été choisie pour étudier la fluorescence de la protéine et des fragments en fonction de la force ionique.

D'après les prédictions de structure, le tryptophane 61 appartient à une région structurée de la protéine : il nous a donc paru intéressant d'en étudier l'environnement plus en détail. Dans ce but, les propriétés de fluorescence de la protéine

MCl ont été comparées à celles des deux fragments contenant la tryptophane 61 en position externe : le fragment 1-64 et le fragment 60-93 qui possède un second résidu de tryptophane en position 74.

Il faut souligner ici que la présence de deux mêmes groupements fluorescents dans la protéine nous a amené à postuler que les deux résidus de tryptophane avaient le même rendement quantique (Q)*

a) Etude de l'accessibilité des résidus de tryptophane

L'oxygène est un inhibiteur efficace de la fluorescence. Pour étudier l'accessibilité des résidus de tryptophane nous avons comparé la fluorescence de la protéine MCl et de ses fragments en présence d'oxygène et en milieu anoxique (bullage d'azote pendant 1 min.).

En présence d'oxygène, l'intensité de fluorescence des trois polypeptides diminue (Figure IV-9) et l'inhibition est plus efficace pour le fragment 60-93 (Figure IV-9b) que pour la protéine MCl (Figure IV-9c). Dans le fragment 60-93, le résidu de tryptophane 61 est très accessible car il est à l'extrémité de la molécule. Ceci explique que sa fluorescence soit très fortement inhibée par l'oxygène. Par contre le fait que la fluorescence de la protéine soit moins inhibée suggère que dans la protéine, le tryptophane 61 est enfoui. Ce résultat est en accord avec les prédictions de structure concernant la séquence 55-66 formant une région structurée hydrophobe qui serait enfouie à l'intérieur de la protéine.

b) Effet de la force ionique

En présence de chlorure de sodium l'intensité de fluorescence de la protéine MCl augmente rapidement jusqu'à 500 mM NaCl puis tend à se stabiliser (Figure IV-10). Par contre l'intensité

*Q = $\frac{\text{nombre de quanta émis par la molécule}}{\text{nombre de quanta absorbé par la molécule}}$

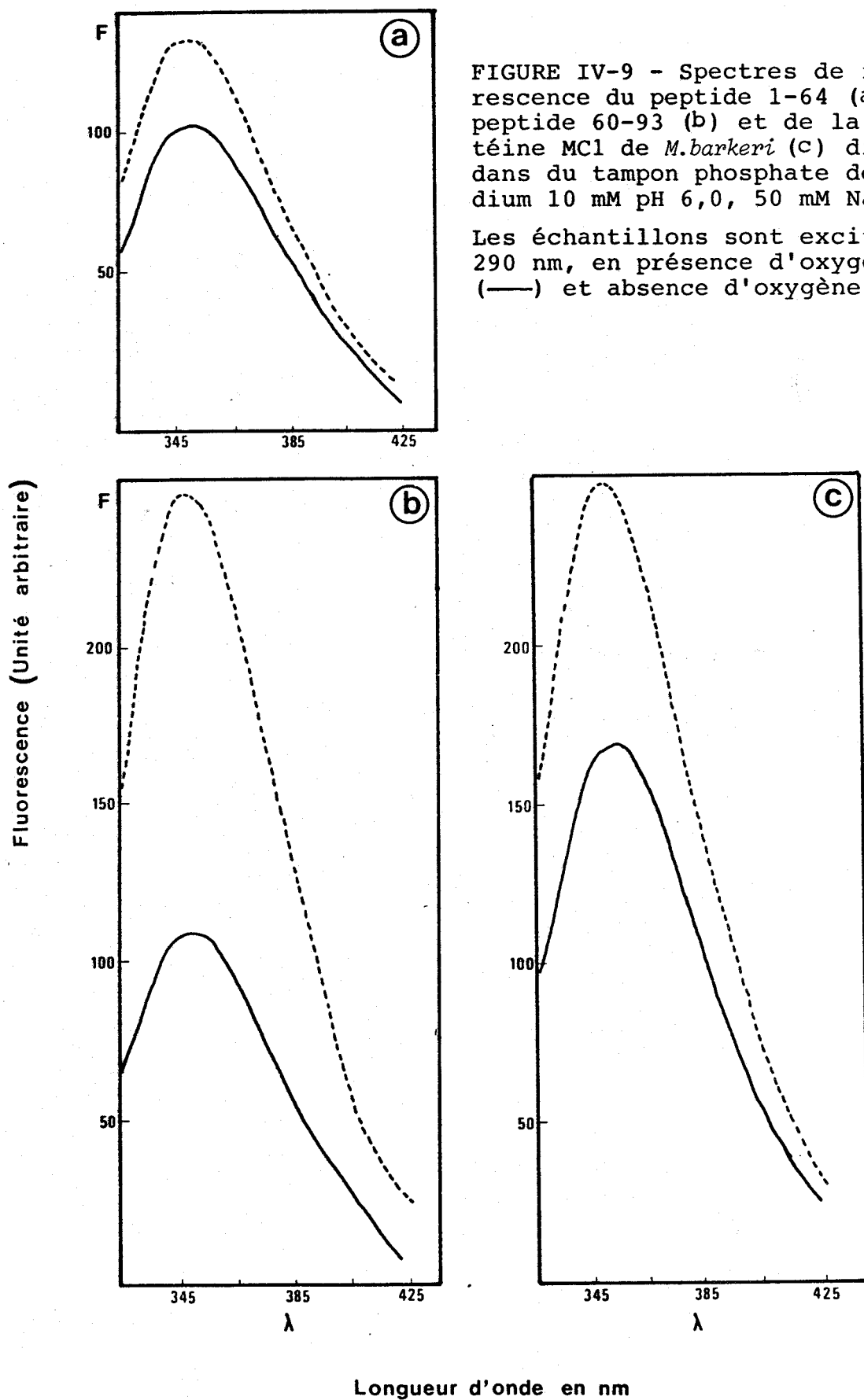


FIGURE IV-10 - Effet de la force ionique sur l'intensité de fluorescence, mesurée à 345 nm, de la protéine MCl de *M. barkeri* (—•—•—), du peptide 60-93 (-●-●-) et du peptide 1-64 (-+--+), après une excitation à 290 nm, en absence d'oxygène et en fonction de la molarité en NaCl.

FIGURE IV-11 - Variation de l'intensité de fluorescence de la protéine MCl de *M. barkeri* excitée à deux longueurs d'onde différentes en fonction de la force ionique. L'intensité de fluorescence F est mesurée à 345 nm et la variation est exprimée en pour cent. La protéine MCl a été excitée à 290 nm (-0-0-) et à 305 nm (-*-*-).

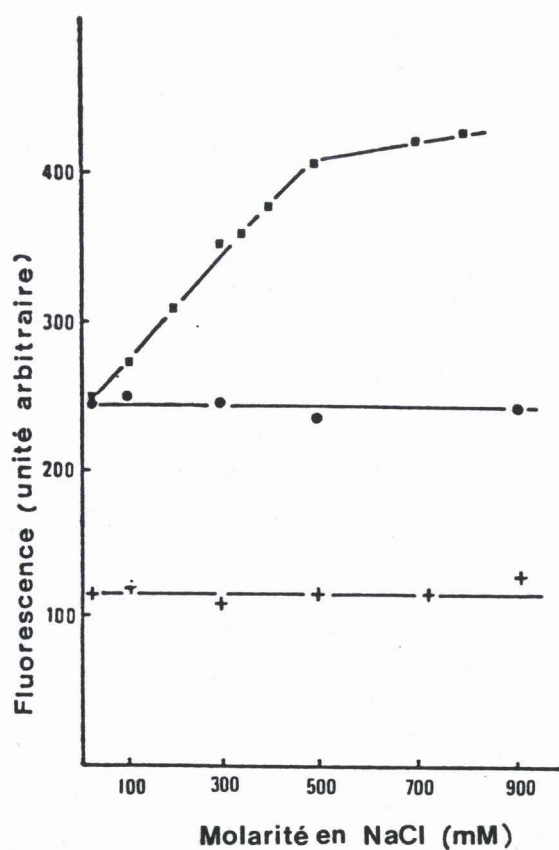


FIGURE IV-10

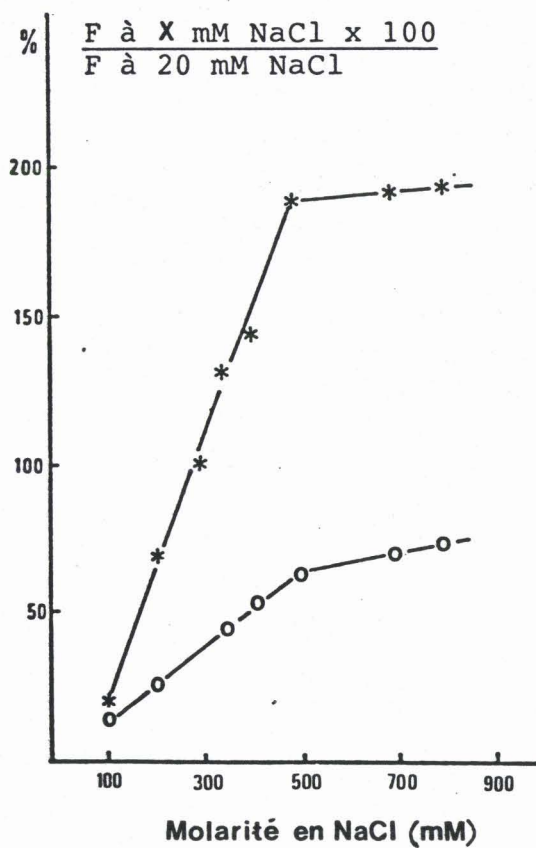


FIGURE IV-11

de fluorescence du fragment 1-64 ne varie pas : en effet, seul le tryptophane de ce peptide est en position externe et sa fluorescence ne dépend donc pas de la conformation du peptide.

L'intensité de fluorescence du fragment 60-93 est également indépendante de la concentration en NaCl. Ce résultat permet de conclure que la variation de fluorescence observée, pour la protéine, sous l'effet de la force ionique n'est pas due au sel mais résulte bien des changements de conformation.

Le tryptophane 61 qui serait enfoui à faible concentration en sel, deviendrait exposé lorsque la force ionique augmente.

c) Etude du transfert d'énergie

La protéine MCl contient deux groupements fluorescents (les résidus de tryptophane 61 et 74) observables dans les conditions de cette étude. La fluorescence observée provient donc de leur fluorescence propre et probablement du transfert d'énergie entre ces deux groupements.

Sachant que seule la fluorescence propre est détectée après excitation à 305 nm (LAKOWICZ, 1983), l'émission de fluorescence (E) peut être définie, après excitation à 305 nm comme $E = E_0$ et après excitation à 290 nm comme $E = E_0 + E_t$ où E_0 représente la fluorescence propre des résidus de tryptophane et E_t l'énergie de transfert entre les deux résidus.

Ceci nous a conduit à comparer les variations de fluorescence de la protéine en fonction de la force ionique après excitation à 305 nm et après excitation à 290 nm. Les résultats sont présentés figure IV-11.

La variation de l'intensité de fluorescence de la protéine en fonction de la concentration en sel, est supérieure après excitation à 305 nm, ce qui indique que le transfert d'énergie diminue lorsque la force ionique augmente. Selon l'équation suivante $R_{AD} = R_0(1/E_t - 1)^{1/6}$ (R_0 dépend des propriétés spectrales du groupe fluorescent), la distance R_{AD} entre deux groupes

fluorescents augmente quand l'énergie de transfert diminue. Ceci indique que les deux résidus de tryptophane s'éloignent l'un de l'autre quand la concentration en chlorure de sodium augmente.

En d'autres termes, il y aurait, sous l'action de la force ionique, un déploiement de la protéine ou au moins de la région contenant les deux résidus de tryptophanes.

d) Etude de l'exposition des résidus de tryptophanes par dichroïsme circulaire

Les études de fluorescence ont montré que le tryptophane (Trp⁶¹) de la protéine MCl se dégage lorsque la force ionique augmente.

L'étude des spectres dichroïques de la protéine MCl enregistrés entre 250 et 400 nm (Figure IV-12) nous ont permis de confirmer ce résultat. En effet, la bande dichroïque à 280 nm qui est attribuée aux résidus de tryptophane augmente avec la molarité en NaCl.

e) Conclusion

L'ensemble de ces résultats montre qu'à faible force ionique, le tryptophane en position 61 est enfoui dans la protéine. Il est situé dans une région très probablement structurée (55-66) et enfouie dans la protéine. Lorsque la force ionique augmente, la protéine se déplie et le tryptophane devient exposé.

IV - 4 COMPARAISON DE LA PROTÉINE MCl AVEC LA PROTÉINE II DE DIFFÉRENTS MICRO-ORGANISMES ET AVEC L'HISTONE H4 DE THYMUS DE VEAU, À L'AIDE DE LA MÉTHODE HCA

Les études de structure primaire ont montré que la protéine MCl ne présente pas d'homologie de séquence en acides aminés ni avec la protéine II d'eubactéries ni avec les histones. La

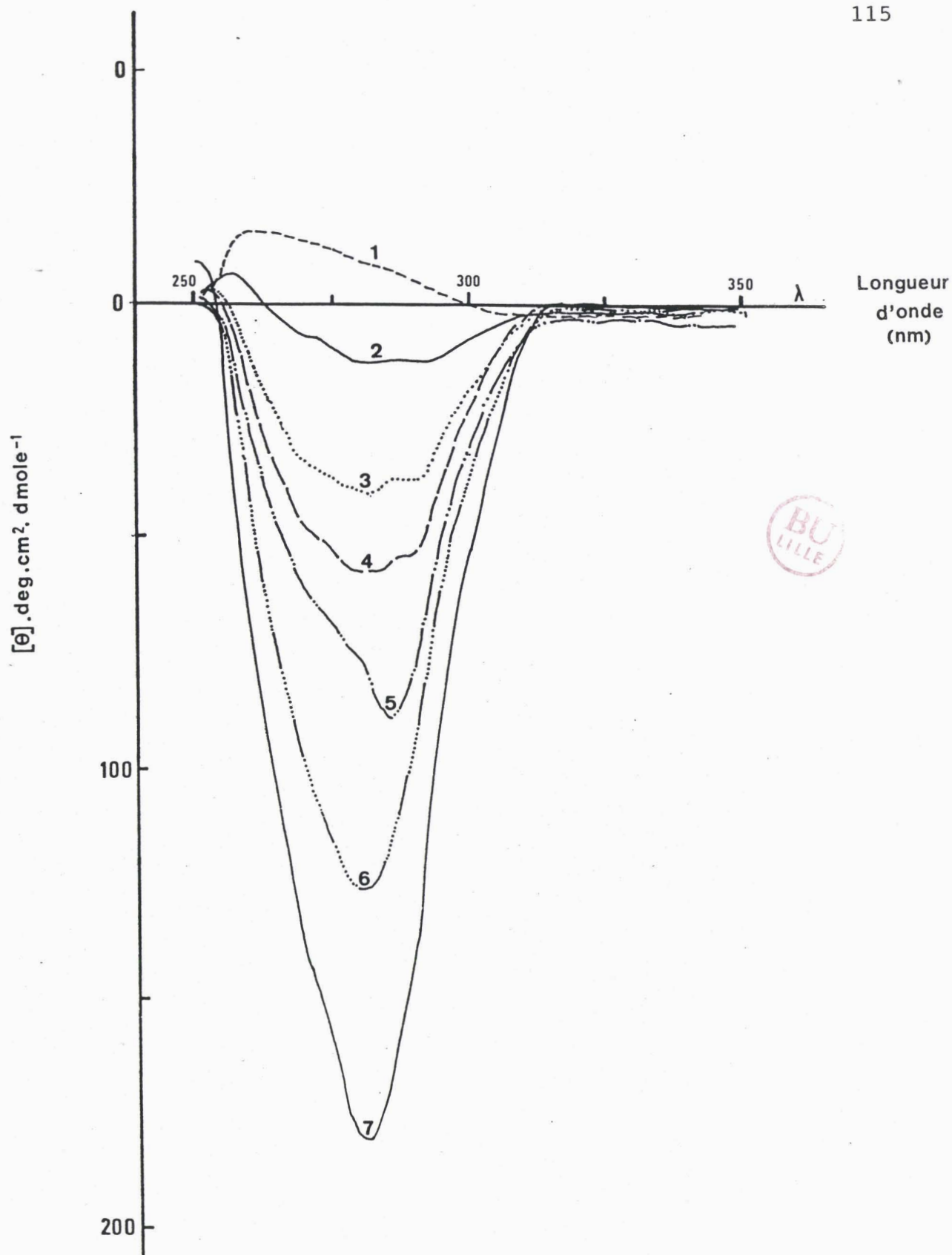


FIGURE IV-12 - Etude de l'exposition des résidus de tryptophane de la protéine MCl de *M. barkeri* (1,2 mg) dissous dans du tampon phosphate de sodium 10 mM pH 6,0 en fonction de la molarité en NaCl : (1) 50 mM. (2) 200 mM ; (3) 300 mM ; (4) 350 mM ; (5) 400 mM, (6) 500 mM ; (7) 700 et 900 mM.

méthode HCA de GABORIAUD et coll. qui permet une perception immédiate des zones structurées correspondant à des amas hydrophobes, est une méthode efficace pour retrouver des parentés entre des protéines qui, malgré une fonction analogue, présentent de faibles homologues de structure primaire. Nous l'avons donc appliquée pour comparer la protéine MCl de *M. barkeri* à la protéine II de différents micro-organismes et à l'histone H4 de thymus de veau.

Ces études comparatives s'effectuent selon le protocole suivant :

- une similitude dans la forme, la taille et l'orientation des amas hydrophobes est d'abord recherchée.
- les séquences sont alignées en se référant aux résidus de proline et de glycine et en recherchant un maximum d'homologie.

Parmi les protéines II dont la séquence en acides aminés est connue (TANAKA et coll., 1984), la protéine II de *Bacillus stearothermophilus* a pu être cristallisée et sa structure tridimensionnelle déterminée (Figure IV-13). La moitié N-terminale de la protéine consiste en deux hélices α (α_1 , résidus 3-13 ; α_2 , résidus 21-37) connectées par une large boucle (résidus 14-20). La moitié C-terminale comprend principalement 3 segments

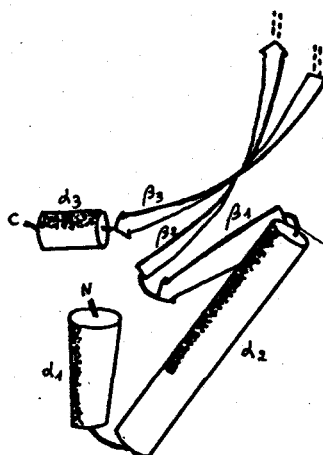


FIGURE IV-13 - Structure du monomère de la protéine II.
(TANAKA et coll., 1984)

en feuillets β antiparallèles (β_1 , résidus 40-44 ; β_2 , résidus 48-51, β_3 , résidus 78-83), les segments β_1 et β_2 sont reliés par un coude très court constitué par la séquence Gly-Phe-Gly qui est hautement conservée. Par ailleurs les segments β_2 et β_3 sont reliés par une séquence en structure étendue (résidus 52-77) qui comprend deux résidus de proline (Pro⁶³ et Pro⁷²) très bien conservés et la région 52-63 qui interagit avec l'ADN. L'extrémité C-terminale forme une hélice α_3 (résidus 84-89). La représentation HCA de cette protéine est illustrée dans la figure IV-14 en comparaison avec la protéine II isolée d'autres micro-organismes et avec la protéine MCl de *M.barkeri*. On observe pour la protéine II de *B.stearothermophilus* une bonne correspondance entre les résultats de la cristallographie et ceux obtenus avec la méthode HCA. Sur la base de cette correspondance, des segments structuraux ont été délimités par des lignes verticales. Par exemple, l'amas hydrophobe (résidus 42 à 50) de la protéine II de *B.stearothermophilus* a été coupé en son milieu compte tenu des résultats cristallographiques et de la présence d'un résidu de glycine en position 46.

Quand on compare la protéine MCl avec les protéines II, on n'observe pas d'homologies dans la disposition des amas hydrophobes. Cependant en alignant ces amas et surtout les résidus de proline et de glycine, quelques similitudes peuvent être observées entre la protéine MCl et les protéines II. En effet, la protéine MCl contient un segment très riche en résidus de proline (résidus 68 à 82) qui peut être aligné avec le segment riche en proline (63-77) des protéines II. Ce segment bien conservé fait partie de la séquence en structure étendue de la protéine II dans laquelle est situé le site d'interaction de l'ADN.

De plus les amas C5 et C6 en feuillet β de la protéine MCl présentent des similitudes avec les segments de structure β situés de part et d'autre du résidu de proline en position 77 dans la protéine II de *Clostridium pasteurianum* et de *Thermoplasma acidophilum*.

FIGURE IV-14 - Comparaison à l'aide de la méthode HCA de la protéine MCl de *M.barkeri* avec la protéine II des micro-organismes suivants : B.st, *Bacillus stearothermophilus* ; Ecl, *Escherichia coli* (HU-1) ; T.a, *Thermoplasma acidophilum* ; C.p, *Clostridium pasteurianum* (TANAKA et coll., 1984).

Les lignes verticales indiquent les correspondances possibles entre les séquences. Les amas hydrophobes sont entourés par une ligne continue. Les zones colorées en vert correspondent à une structure en hélice α et les zones colorées en rose à une structure en feuillet β .

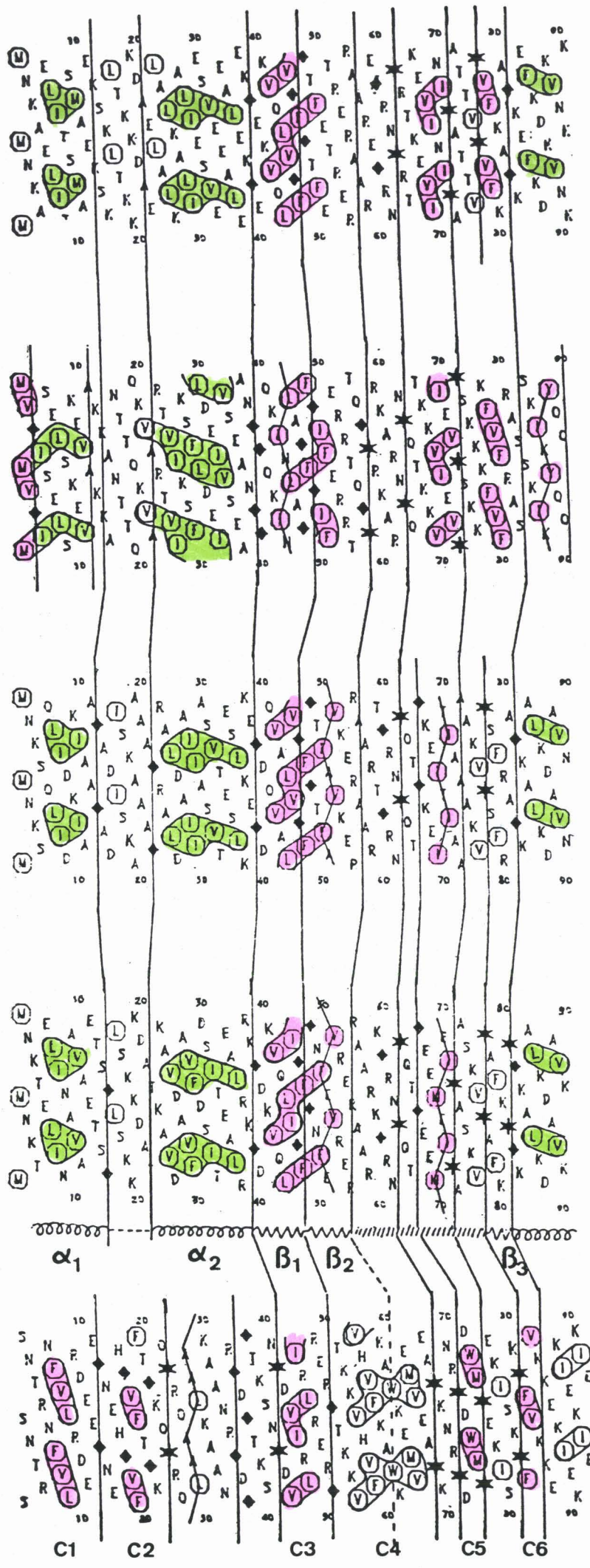
C.p

T.a

E.c1

B.st

MC1



Il semble donc que par rapport à la protéine II, la partie C-terminale de la protéine MCl soit mieux conservée que la partie N-terminale.

Par contre, la comparaison du diagramme HCA de la protéine MCl avec celui de l'histone H4 de thymus de veau ne montre aucune homologie structurale (Figure IV-15).

IV - 5 DISCUSSION

Les études en spectroscopie infrarouge et en dichroïsme circulaire ont montré la présence de structure en feuillets β antiparallèles et la quasi absence de structure en hélice α dans la protéine MCl.

Ces observations concordent avec les données fournies par les méthodes de prédiction de structure qui indiquent que la protéine MCl serait peu organisée. Elle contiendrait en effet 16% de structure en hélice α , 15% de structure en feuillet β et 30% de coudes β . Ce faible taux d'organisation pourrait expliquer que la protéine MCl se comporte comme un monomère puisque, généralement, les molécules présentant un taux important de structure organisée existent plutôt à l'état de polymère. Tel est le cas de l'octamère des histones H2A, H2B, H3, H4 et du dimère de la protéine II d'eubactéries.

Par ailleurs, toutes les méthodes spectroscopiques (infrarouge, dichroïsme circulaire, fluorescence) ont montré que l'augmentation de la force ionique induit des changements de conformation dans la protéine MCl. En spectroscopie infrarouge, on observe une disparition des feuillets β antiparallèles. En dichroïsme circulaire, les résultats obtenus pour la protéine MCl et le fragment 1-64 permettent de supposer que les changements conformationnels observés mettent en jeu la région structurée de la protéine (résidus 55 à 66). De plus les études de fluorescence montrent que le tryptophane en position 61 est au coeur d'une région structurée enfouie dans la protéine.

L'ensemble de ces résultats permet d'émettre quelques suggestions concernant la structure secondaire de la protéine MCl.

Notons d'abord que les spectres infrarouges ont été obtenus en utilisant une concentration en protéine (11 mg/ml) beaucoup plus importante que celle utilisée pour les études en dichroïsme circulaire et en fluorescence. On pourrait donc penser que la structure β antiparallèle, observée à faible force ionique, serait due à l'agrégation entre les molécules. Cependant les études en fluorescence réalisées à une concentration en protéine beaucoup plus faible ($\approx 100 \mu\text{g/ml}$) montrent, comme la spectroscopie infrarouge, que la protéine MCl se déplie quand la force ionique augmente. Pour cette raison et aussi parce que la protéine se comporte en monomère, nous pensons que les feuillets β antiparallèles constituent un arrangement intrachaine, ce qui est concevable avec le nombre important de coudes β ou de boucles dans la protéine.

A partir de là, nous pouvons admettre qu'à faible force ionique, des liaisons ioniques stabilisent les structures en feuillets β antiparallèles. Quand la force ionique augmente, les liaisons ioniques sont rompues provoquant ainsi la déstabilisation des zones structurées et le dépliement de la molécule.

Par ailleurs, la structure secondaire de la protéine MCl de *M. barkeri* diffère de celle de la protéine II d'eubactéries et des histones : la protéine II et les histones ont un taux d'hélicité plus élevé que celui de la protéine MCl. Ainsi la protéine II de *B. stearothermophilus* a un taux d'hélicité voisin de 40 % tandis que l'histone H4 a un taux d'hélicité de l'ordre de 20 %.

On observe également une différence entre les changements conformationnels induits par les sels pour la protéine MCl et ceux observés pour les histones ou les protéines II. En effet, l'augmentation de la force ionique chez les histones, favorise la formation de structure secondaire et particulièrement la formation de structure en hélice α (D'ANNA et ISENBURG, 1972 ; BRADBURY et coll., 1975 ; LEWIS et coll., 1975). De même, KHANAKA (1984) a montré que la protéine II d'eubactéries acquiert une conformation plus hélicoïdale quand la force ionique augmente.

CHAPITRE V

ETUDE ULTRASTRUCTURALE DU COMPLEXE DESOXYRIBONUCLEOPROTEIQUE

DE L'ARCHAEBACTERIE *METHANOSARCINA BARKERI*

INTRODUCTION

En 1985, F. CHARTIER a montré que la protéine MCl représente le composant protéique majeur de l'extrait acido-soluble du complexe nucléoprotéique de *M. barkeri*.

Pour confirmer que la protéine MCl est associée à l'ADN *in vivo*, nous avons utilisé une autre approche. La protéine MCl a été localisée sur l'ADN en microscopie électronique par immunomarquage sur des coupes de cellules de *Methanosarcina barkeri* et sur de la chromatine étalée. En outre, afin de rechercher si des protéines de caractère acido-insoluble étaient présentes dans le complexe désoxyribonucléoprotéique de la bactérie, nous avons déterminé la composition en protéines totales de ce complexe.

La protéine MCl a des caractéristiques de structures primaire et secondaire différentes de celles des histones et de la protéine II ; néanmoins la méthode comparative HCA a permis de relever une certaine homologie entre la protéine MCl et d'une part les protéines II de *Clostridium pasteurianum* et de *Thermoplasma acidophilum* et d'autre part de la protéine II d'eubactéries, au niveau de deux régions en feuillet β (amas C5 (Trp⁷⁴-Met⁷⁵) et amas C6 (Phe⁸³-Val⁸⁴)) et de la région riche en proline (séquence 68 à 82) qui correspond au bras par lequel la protéine II interagit avec l'ADN.

Nous avons donc recherché si la protéine MCl est capable d'organiser l'ADN de *M. barkeri* en unités structurales répétitives similaires à celles décrites dans le nucléoïde d'eubactéries et de *Thermoplasma acidophilum*.

Ces travaux réalisés en collaboration avec le Docteur Gérard PRENSIER (U. 42 INSERM, Villeneuve d'Ascq) pour toutes les études en microscopie électronique, font l'objet d'une publication soumise à Canadian Journal of Microbiology.

THE DEOXYRIBONUCLEOPROTEIN COMPLEX FROM THE
ARCHAEBACTERIUM METHANOSARCINA BARKERI: CHARACTERIZATION AND
ULTRASTRUCTURAL LOCALIZATION OF THE CHROMOSOMAL PROTEIN MCI

Imbert Marlène¹, Laine Bernard¹, Prensier Gérard^{2*}, Touzel Jean-Pierre³ and
Sautière Pierre¹.

1 : Unité Associée CNRS 409, I.R.C.L., Place de Verdun -
59045 Lille, FRANCE

2 : INSERM U. 42, 369, rue Jules Guesde -
59651 Villeneuve d'Ascq, FRANCE

3 : INRA, Station de Technologie Alimentaire, B.P. 39 -
59651 Villeneuve d'Ascq, FRANCE

* Present address : Bâtiment de Zoologie, Université Blaise Pascal, Complexe
scientifique des Cezeaux, B.P. 45 - 63170 AUBIERE

Correspondence and proofs should be sent to : Dr Bernard LAINE
Unité Associée CNRS 409, I.R.C.L., Place de Verdun -
59045 Lille, FRANCE Tél. : 20.95.43.59

ABBREVIATIONS : EDTA, ethylene diamine tetraacetic acid ;

Protein MC1, Methanogen chromosomal protein 1 ;

SDS, Sodium Dodecyl Sulphate.

Enzymes Nuclease from staphylococcus aureus (EC 3.1.31.1)

Ribonuclease A from bovine pancreas (EC 3.1.27.5)

SUMMARY :

The Methanosarcina barkeri protein MCI, a basic polypeptide of 93 amino acid residues, has been localized in the DNA-rich areas on immunolabelled bacterial cryosections revealed with the protein A-gold method. This localization strongly supports the previous studies concerning the association of the protein MCI with the DNA *in vivo*. In the M. barkeri deoxyribonucleoprotein complex the protein MCI accounts for about 90 % of the total proteins and the protein MCI-to-DNA ratio (by weight) is equal to 0.11. Staphylococcal nuclease digestion data and microscopic observations have failed to reveal the presence of repeating structural units and suggest that the protein MCI is unevenly distributed along the DNA.

INTRODUCTION :

The eukaryotic chromatin is organized in repeating structural units called nucleosomes which exhibit a beads -on-a- string appearance upon electron microscopy and can be released by a brief digestion of chromatin with staphylococcal nuclease. These elementary structural units are constituted of about 200 base pairs of DNA, two molecules of each of the histones H2A, H2B, H3 and H4 and one molecule of histone H1 (McGhee and Felsenfeld 1980). The composition of the chromatin of several eubacteria and archaeobacteria has been determined (Varshavsky et al. 1977 ; Yamazaki et al. 1984, Searcy and Stein 1980 and Green et al. 1983). The structural organization of the prokaryote chromatin is much more unstable and is not so well defined as that of eukaryotes ; however the isolation of DNA fragments protected from staphylococcal nuclease digestion suggests the presence of repeating structural units in several microorganisms. In Escherichia coli, the DNA fragments are about 120 base pairs long and are associated with one molecule of each of the proteins B1 and B2 (also called HU) (Varshavsky et al. 1977). The nucleoprotein particles found in Thermoplasma acidophilum are globular and are composed of 40 base pairs of DNA bound to four molecules of protein HTa (Searcy and Stein 1980). Protected DNA fragments were also found in Sulfolobus acidocaldarius but the amount of protein associated to them was not determined (Green et al. 1983). Up to now the structural organization of chromatin in methanogens and halophiles remains unknown ; these two groups of microorganisms constitute among the archaeobacteria a branch well separated from that formed by the thermophilic sulfur-dependent archaeobacteria (Woese and Olsen 1986).

The protein MC1 is the major component found in the acid extract

isolated from the Methanosarcina barkeri nucleoprotein complex (Chartier et al. 1985). The protein MC1 is a basic polypeptide of 93 amino acid residues (M_r 10 757) of which the amino acid sequence is not homologous to that of any prokaryotic or eukaryotic chromosomal proteins described so far (Laine et al. 1986). For this reason, the name HMB originally given to the protein MC1 to follow the nomenclature used for chromosomal proteins from eubacteria, now appears to be inadequate and confusing.

This paper deals with the immunocytochemical localization of DNA and protein MC1 in the archaebacterium Methanosarcina barkeri and describes the main features of the deoxyribonucleoprotein complex by means of two convergent approaches i.e. biochemistry and electron microscopy observations.

MATERIALS AND METHODS

Methanosarcina barkeri MS (DSM 800) was grown as indicated by Touzel and Albagnac (1983). Nuclease from Staphylococcus aureus was purchased from SERVA and ribonuclease A from bovine pancreas was purchased from SIGMA. Diphenylamine was from MERCK and BCA protein assay reagent was from PIERCE. Sera from patients with systemic lupus erythematosus, taken as sources of anti-DNA antibodies, were obtained from Pr Dequiedt (Service de Néphrologie, Hôpital Calmette, LILLE, France). These sera exhibit an affinity for a double-stranded DNA-cellulose column.

Antisera anti-protein MC1

The protein MC1 was prepared by affinity chromatography on a DNA-cellulose column as described by Chartier et al. (1985) and the contaminants were discarded by precipitation with ammonium sulphate at

80 % saturation. The protein MCI obtained pure in the supernatant was dialyzed against 50 mM sodium chloride, 10 mM sodium phosphate pH 6.0. Antibodies directed against protein MCI were raised in rabbits by subcutaneous injections of 100 µg protein in Freund's complete adjuvant (CFA) every week during 3 weeks followed, one and two months later, by an intramuscular injection of 200 µg protein in CFA. Sera obtained 15 days after the last injection were used for immunological studies. The antibodies assays had to be performed by the dot-immunobinding technique (Hawkes et al. 1982) because the anti-protein MCI antibodies produced were not precipitating.

Electron microscopy

Whole cells of M. barkeri and pelleted deoxyribonucleoprotein were fixed for 1 h in a mixture of 4 % formaldehyde and 0.1 % glutaraldehyde in phosphate-buffered saline (PBS ; 50 mM phosphate buffer pH 7.5, 150 mM sodium chloride) and then treated according to Moreau et al. (1986). For immunolabelling of protein MCI, grids were incubated for 1 h at room temperature in the immunoserum anti-protein MCI diluted to 1/10 with PBS. The revelation was performed with protein A-colloidal gold (5 nm-diameter). For immunolabelling of DNA, grids were incubated as above in the sera from patients with systemic lupus erythematosus diluted to 1/10 with PBS. The revelation was performed with the protein A-colloidal gold technique and alternatively with goat anti-human IgG-gold (25 nm-diameter). For double immunolabelling, protein MCI and DNA were labelled as indicated above and revealed with protein A-colloidal gold (5 nm-diameter) and goat anti-human IgG-gold (25 nm-diameter) respectively.

Negative staining of immunolabelled spread deoxyribonucleoprotein was performed on parlodion coated grids with 3 % aqueous uranyl acetate.

Electron microscopy was performed with a Hitachi HUI2A electron microscope.

Staphylococcal nuclease digestion of the deoxyribonucleoprotein complex

The deoxyribonucleoprotein complex was prepared as described by Chartier et al. (manuscript in preparation). In a first step the cellular debris and polysomes were removed by centrifugation, then the deoxyribonucleoprotein complex was pelleted and purified by gel filtration chromatography on a Sepharose 4B column after treatment with ribonuclease A. The deoxyribonucleoprotein complex was dialyzed against 100 mM sodium chloride, 10 mM Tris-HCl, pH 7.6 and diluted to an absorbance of $A_{260\text{ nm}}^{1\text{ cm}} = 2.0$. The digestion was carried out at 37°C with 20 U/ml of staphylococcal nuclease in the presence of 5 mM calcium chloride at 37°C. At various intervals, aliquots were taken off and adjusted to 10 mM disodium-EDTA to stop the reaction. The undigested DNA was precipitated in 0.5 M perchloric acid at 0°C for 30 min and centrifuged at 2000 x g for 15 min. The concentration of the acid-soluble DNA was measured from the absorbance at 260 nm of the supernatant and the amount of the acid-insoluble DNA was calculated and expressed as percent undigested.

DNA and protein assays in the deoxyribonucleoprotein complex

The deoxyribonucleoprotein was precipitated in 0.5 M perchloric acid for 30 min at 0°C and centrifuged at 2000 x g for 15 min to discard free nucleotides. This procedure was repeated twice and the pelleted DNA was assayed by the diphenylamine procedure (Burton 1968). Protein concentration was determined by the Lowry method modified by Smith et al. (1985). Calf thymus histone H4 was used as calibration protein standard.

Electrophoretic procedure

Sodium dodecyl sulphate polyacrylamide gel electrophoresis was

performed according to Laemmli (1970). Gel was stained and destained as described by Stephano et al. (1986). The gel was scanned with a Vernon PHI 4 densitometer.

RESULTS

The knowledge of the protein composition of the deoxyribonucleo-protein complex is a prerequisite to the understanding of the organization of the chromosomal DNA. Chartier et al. (1985) showed that the protein MCI is the major component of the acid extract of the M. barkeri nucleoprotein complex. Electron microscopy observations of M. barkeri cryosections allowed us to confirm the presence of the protein MCI on DNA. The intracellular localization of DNA and protein was performed using an immunolabelling method with the protein A-gold method. In order to test the efficiency of this method, single labelling of protein (Fig. 1a) and DNA (Fig. 1b) was performed on cryosections of the pelleted deoxyribonucleoprotein. In both cases, gold particles were located in dense areas which correspond to the fibrillar structures of the deoxyribonucleoprotein and the number of the gold particles was higher in the darkest areas. It must be emphasized that the background was negligible. Due to unsatisfactory results with double immunolabelling of protein and DNA, we chose to label these components on separate cryosections. The results of DNA immunolabelling on the M. barkeri cryosections showed local accumulation of DNA in the dense areas which were evenly distributed throughout the cell (Fig. 1d) thus suggesting that the DNA is not localized in a restricted region. On the cryosection presented in the figure 1c, the immunolabelling of protein MCI was localized in the dense areas where DNA has been evidenced.

The finding of acid-insoluble proteins associated to the chromosomal DNA in several archaeobacteria (Green et al. 1983, Thomm et al. 1982) led us to analyze the total protein composition of the M. barkeri

deoxyribonucleoprotein complex. The electrophoretic pattern obtained in this analysis is presented in the figure 2. As determined by scanning densitometry, the protein MCI accounted for about 90 % of the total proteins associated to the DNA whereas the other proteins were found in very low amounts (about 2 % for each protein species). Therefore the protein MCI appeared to be the only DNA partner which may have a structural role in the chromosome of M. barkeri. The isolated nucleoprotein exhibited a protein MCI-to-DNA ratio of 0.11 (by weight). Taking into account the molecular weight of the protein MCI, this value corresponds to 1 molecule of protein MCI per 150 base pairs of DNA.

The organization of the chromatin can be studied by convergent approaches : i, electron microscopy on spread chromatin ; ii, digestion of the purified chromatin with staphylococcal nuclease in order to investigate if repeating structural units resistant to the enzyme and reminiscent to nucleosomes are present in the chromatin.

The M. barkeri deoxyribonucleoprotein complex was incubated with staphylococcal nuclease (20 U/100 μ g of DNA in 1 ml), an enzyme-to-substrate ratio used in the study of other archaeobacteria chromatin (Searcy and Stein 1980, Green et al. 1983). The time-course digestion curves of M. barkeri deoxyribonucleoprotein and of deproteinized DNA were similar (Fig. 3), demonstrating that the protein MCI was unable to protect DNA against staphylococcal nuclease.

When revealed by negative staining, the deoxyribonucleoprotein preparations appeared under the electron microscope as fibres without globular structure resembling eukaryotic nucleosomes (Fig. 1h). This observation together with the results of the nuclease digestion of the deoxyribonucleoprotein complex indicates a lack of organization of the DNA in M. barkeri similar to the nucleosomal DNA repeats observed in eukaryotes.

The arrangement of protein MCI on DNA was studied by immuno-

electron microscopy. Preparations of spread deoxyribonucleoprotein were separately labelled with anti-sera anti-protein MCl (Fig. 1e) and with sera containing anti-DNA antibodies (Fig. 1f) and then negatively stained. In these figures, gold particles were found in dense regions corresponding to deoxyribonucleoprotein. Moreover in the preparations used to reveal the protein, the density of gold particles throughout a same dense area was not uniform : some dense regions were better labelled than others. These observations suggest that protein MCl is unevenly distributed along the DNA and that some regions of DNA are probably naked. Similar results were obtained with a double immunolabelling to reveal both DNA and protein MCl : on a same fibre of deoxyribonucleoprotein, labelling of protein MCl and DNA appeared irregularly distributed (Fig. 1g).

DISCUSSION

The evidence of the association of the protein MCl with DNA in Methanosarcina barkeri, previously established by Chartier et al. (1985) has been reinforced by an immunocytochemical localization of DNA and protein MCl in the bacterial cell. DNA-rich areas appeared to be evenly distributed throughout the bacteria and not restricted in some large regions as observed in E. coli (Bjornsti et al. 1986). The protein MCl represents 90 % of the proteins in the M. barkeri deoxyribonucleoprotein complex. In this complex any acid insoluble protein was found in sufficient amount to play a structural role and particularly the acid-insoluble protein of apparent M_r 14 300 present in the chromatin of Methanobacterium thermoautotrophicum (Thomm et al. 1982). On the other hand, these authors did not find any acid soluble protein homologous to the protein MCl. The protein MCl has been found in other strains within the genus Methanosarcina (Chartier et al. 1986) but the presence of this protein has not been reported yet in other methanogenic bacteria. At the present time, it is therefore not possible to determine if the differences

in protein composition of the chromatin between M. barkeri and M. thermoautotrophicum are related to their affiliation to different orders.

The biochemical data and the observations in electron microscopy clearly showed the lack of nucleosomal structure in M. barkeri chromatin and did not allow us to evidence any other kind of repetitive structure in this bacterium. However it must be kept in mind that such putative structures may be extremely labile as are the beaded fibres visualized in E. coli by Griffith (1976). In any case, the rate of the nuclease digestion of the complex is consistent with the probable presence of long naked stretches of DNA : this was suggested by the observation of large regions of the deoxyribonucleoprotein complex in which the protein MCl was not revealed, and which is accounted for by the low protein-to-DNA ratio in the complex. The value of this ratio was found to be 0.11. However, it must be noticed that the cells were harvested in the late stationary phase of growth when the production of methane was decreasing. Presently, we do not know whether the amount of protein MCl is constant or is dependent of the physiological state of the bacterium.

The protein composition of the chromatin in prokaryotes exhibits a large qualitative and quantitative diversity. The protein HTa is the single abundant protein associated with DNA in Thermoplasma acidophilum (Searcy and Stein 1980) whereas two chromosomal proteins are encountered in Sulfolobus acidocaldarius (proteins HSa and NHSa) (Green et al. 1983) and in Escherichia coli (proteins B1 and B2) (Varshavsky et al. 1977). None of these proteins can be related to each other except the proteins HTa and B2 which exhibit amino acid sequence homologies (Laine et al. 1983). When referring to the amount of DNA in the purified deoxyribonucleoprotein, the protein MCl is as abundant as the E. coli protein B2 but is about twice less abundant than each of the proteins HSa and NHSa in S. acidocaldarius and five times less abundant than the

protein HTa in T. acidophilum.

Such a diversity in protein composition accounts for the differences observed in the chromatin structure among prokaryotes. However the chromatin features in M. barkeri are not far from those of E. coli and S. acidocaldarius. The E. coli chromatin is completely digested with the staphylococcal nuclease within a short time (see Green et al. 1983) and, in fact, the 120 base pairs DNA fragments described by Varshavsky (1977) were released at the early stage of nuclease digestion (when 29 % of the DNA is digested). DNA fragments of size ranging from 50 to 220 base pairs were also released from the M. barkeri nucleoprotein complex after 10 min digestion (when 75 % of the DNA was digested) but no fragments could be detected at longer digestion times (data not shown). In S. acidocaldarius only a small fraction (5 %) of the DNA is resistant to the nuclease and the protected DNA fragments exhibit a broad size distribution (with an average of 25 base pairs) contrary to those released by nuclease digestion of eukaryotic chromatin. Only the repeating structural units of the T. acidophilum DNA particles seem to be intermediate between those found in eukaryotes and those found in the other microorganisms cited above. From the point of view of chromatin organization, M. barkeri is closer to eubacteria and S. acidocaldarius than to eukaryotes.

Acknowledgements

The authors are grateful to M. Lestiennes for editorial assistance. This work was supported by grants from the Centre National de la Recherche Scientifique, from the Pôle Régional des Anaérobies de la Région Nord-Pas-de-Calais and from the Université de LILLE II.

Bjornsti, M.A., Hobot, J.A., Kelus, A.S., Villiger, W. and Kellenberger, E. 1986. New electron microscopic data on the structure of the nucleoid and their functional consequences, p. 64-81. In Gualerzi, C.O. and Pon, C.L. (eds), *Bacterial chromatin*. Springer-Verlag, Heidelberg, Federal Republic of Germany.

Burton, K. 1968. Determination of DNA concentration with diphenylamine. In *Methods Enzymol.* 12 : 163-166.

Chartier, F., Laine, B., Sautière, P., Touzel, J.P. and Albagnac, G. 1985. Characterization of the chromosomal protein HMb isolated from Methanosarcina barkeri, *FEBS. Lett.* 183 : 119-123.

Chartier, F., Imbert, M., Laine, B., Sautière, P., and Touzel, J.P. 1986. Characteristics of a chromosomal protein isolated from different species of Methanosarcinaceae : implication in phylogeny. p. 242-244. In Dubourguier, H.C. et al. (eds) *Biology of anaerobic bacteria*. Elsevier Science Publishers B.V., Amsterdam, The Netherlands.

Green, G.R., Searcy, D.G. and DeLange, R.J. 1983. Histone-like protein in the archaeobacterium Sulfolobus acidocaldarius. *Biochim. Biophys. Acta* 741 : 251-257.

Griffith, J.D. 1976. Visualization of prokaryotic DNA in a regulatory condensed chromatin-like fiber. *Proc. Natl. Acad. Sci. USA.* 73 : 563-567.

Hawkes, R., Niday, E. and Gordon, J. 1982. Dot-immunobinding assay for monoclonal and other antibodies. *Anal. Biochem.* 119 : 142-147.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227 : 680-685.

Laine, B., Bélaïche, D., Khanaka, H. and Sautière, P. 1983. Primary structure of the DNA-binding protein HRm from Rhizobium meliloti. *Eur. J. Biochem.* 131 : 325-331.

Laine, B., Chartier, F., Imbert, M., Lewis, R. and Sautière, P. 1986. Primary structure of the chromosomal protein HMb from the archaeobacterium Methanosarcina barkeri. *Eur. J. Biochem.* 161 : 681-687.

McGhee, J.D. and Felsenfeld, G. 1980. Nucleosome structure. *Annu. Rev. Biochem.* 49 : 1115-1156.

Moreau, S., Prensier, G., Maalla, J. and Fortier, B. 1986. Identification of distinct accumulation sites of 4-aminoquinoline in chloroquine sensitive and resistant Plasmodium berghei strains. *Eur. J. Cell. Biol.* 42 : 207-210.

Searcy, D.G. and Stein, D.B. 1980. Nucleoprotein subunit structure in an unusual prokaryotic organism : Thermoplasma acidophilum. *Biochim. Biophys. Acta* 609 : 180-195.

Smith, P.K., Krohn, R.I., Hermanson, G.T., Mallia, A.K., Gartner, F.H., Provenzano, M.D., Fujimoto, E.K., Goete, N.M., Olson, B.J. and Klenk, D.C. 1985. Measurement of protein using bicinchoninic acid. *Anal. Biochem.* 150 : 76-85.

Stephano, J.L., Gould, M. and Rojas-Galicia, L. 1986. Advantages of picrate fixation for staining polypeptides in polyacrylamide gels. *Anal. Biochem.* 152 : 308-313.

Thomm, M., Stetter, K.O. and Zillig, W. 1982. Histone-like proteins in eu- and archaeobacteria. *Zbl. Bakt. Hyg., I. Abt. Orig. C* 3 : 128-139.

Touzel, J.P. and Albagnac, G. 1983. Isolation and characterization of Methanococcus mazei strain MC3. *FEMS Lett.* 16 : 241-245.

Varshavsky, A.J., Nedospasov, S.A., Bakayev, V.V., Bakayeva, T.G. and Georgiev, G.P. 1977. Histone-like proteins in the purified Escherichia coli deoxyribonucleoprotein. *Nucleic Acids Res.* 4 : 2725-2745.

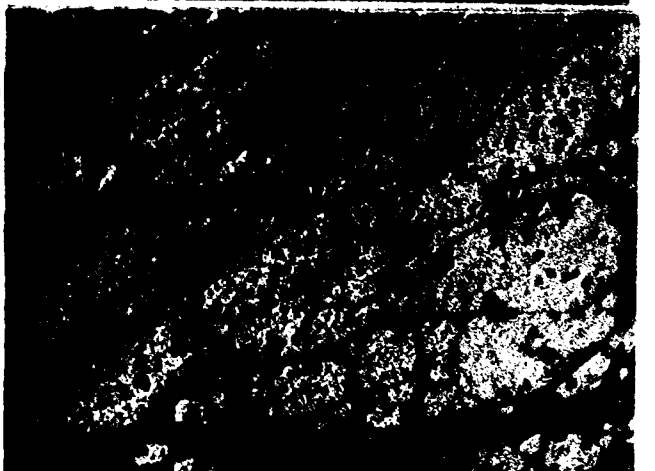
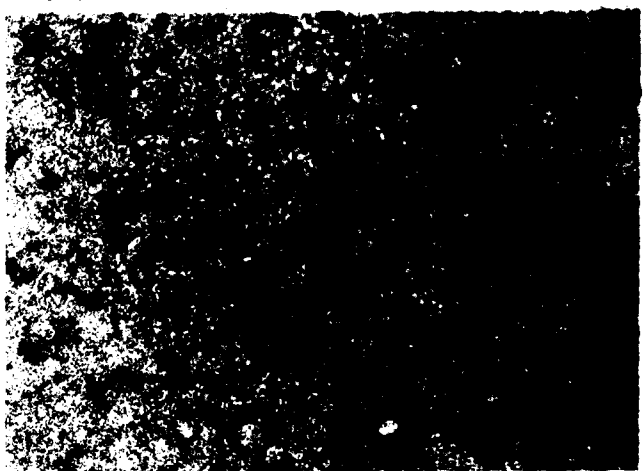
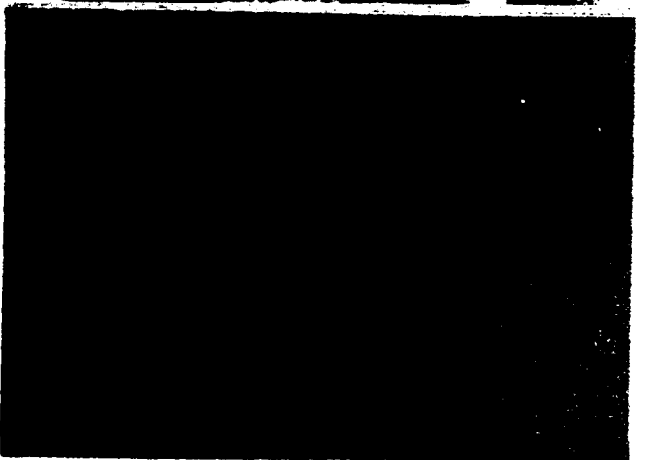
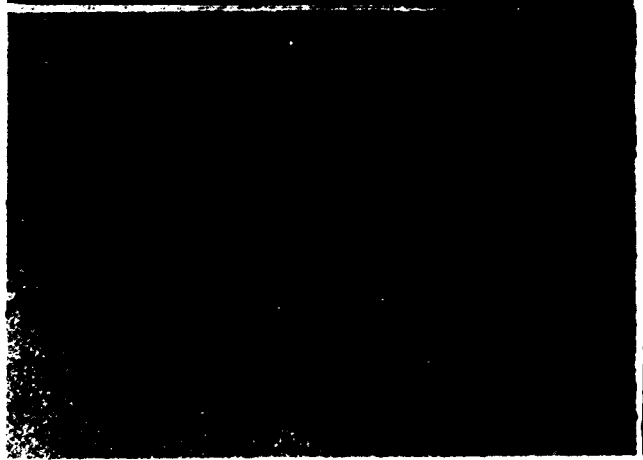
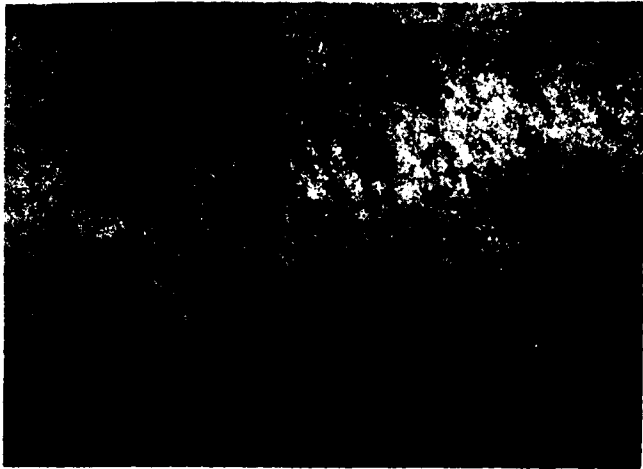
Woese, C.R. and Olsen, G.J. 1986. Archaeobacterial phylogeny : perspectives on the Urkingdoms. *System. Appl. Microbiol.* 7, 161-177.

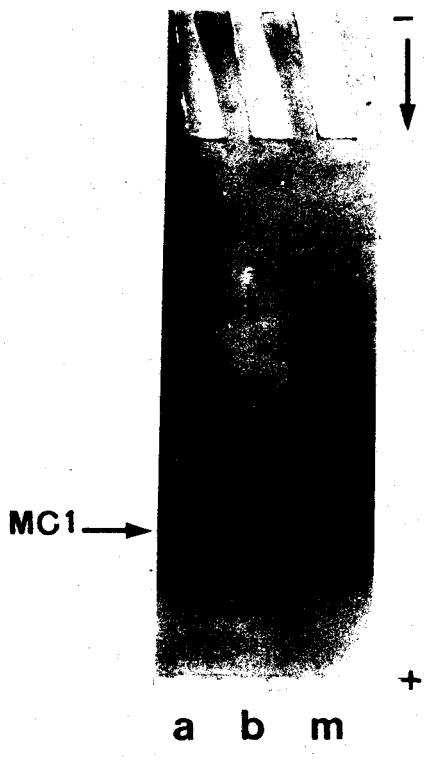
Yamazaki, K, Nagata, A., Kano, Y. and Imamoto, F. 1984. Isolation and characterization of nucleoid proteins from Escherichia coli. *Mol. Gen. Genet.* 196 : 217-224.

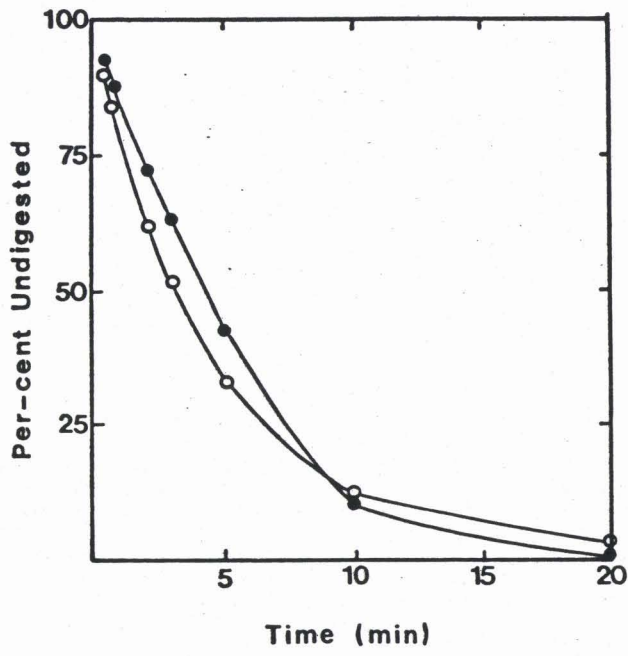
Figure 1 : Immunolocalization of protein MC1 and DNA from Methanosarcina barkeri. -a,b, cryosections of pelleted deoxyribonucleoprotein complex. -c,d, cryosections of whole cells of Methanosarcina barkeri. -e, f, g, h, negative staining of spread chromatin. -a,c,e, immunolabelling of protein MC1 with anti-serum containing anti-protein MC1 antibodies revealed with protein A-gold (5 nm-diameter). -b, d, immunolabelling of DNA with serum containing anti-DNA antibodies (see methods) revealed with protein A-gold (5 nm-diameter). -f, immunolabelling of DNA with serum containing anti-DNA antibodies revealed with goat anti-human IgG-gold (25 nm-diameter). -g, double immunolabelling : protein MC1 and DNA were labelled as indicated above and revealed with protein A-gold (5 nm-diameter) and goat anti-human IgG-gold (25 nm-diameter) respectively. Bars = 0,2 μ m.

Figure 2 : Qualitative analysis of the protein components of the Methanosarcina barkeri deoxyribonucleoprotein complex by SDS-polyacrylamide gel electrophoresis. a, M. barkeri deoxyribonucleoprotein complex ; b, crude protein MC1 obtained by DNA-cellulose chromatography. m, M_r markers : phosphorylase b (M_r 94 000), bovine serum albumin (M_r 67 000), ovalbumin (M_r 43 000), carbonic anhydrase (M_r 30 000), soybean trypsin inhibitor (M_r 20 100), α -lactalbumin (M_r 14 400). Samples (7 μ g) were treated with 5 % SDS, 2 % 2-mercaptoethanol for 3 min at 100°C and run at 40 mA for 150 min on a 5-30% polyacrylamide gradient gel. Electrode buffer (Tris-glycine, pH 8.3) and gel buffer (Tris-HCl, pH 8.9) were made 0.1 % SDS. Gel was stained and destained according to Stephano et al. (1986).

Figure 3 : Digestion with staphylococcal nuclease (20 U/ml) of the M. barkeri deoxyribonucleoprotein complex (●—●) and of chicken erythrocyte DNA (○—○). Results are expressed as per-cent of DNA remaining undigested.







DISCUSSION

La protéine MCl et l'ADN ont été révélés dans les mêmes régions denses aux électrons sur les coupes de *Methanosarcina barkeri* et sur les structures fibrillaires de la chromatine étalée. Ce résultat confirme que la protéine MCl est associée à l'ADN *in vivo*.

La protéine MCl, qui représente 90 % des protéines du complexe désoxyribonucléoprotéique, apparaît être la seule protéine susceptible de jouer un rôle structural dans ce complexe.

Cependant, contrairement aux histones et à la protéine II de l'archaebactérie *T. acidophilum*, la protéine MCl ne protège pas l'ADN de la digestion par la nucléase staphylococcale et ne forme pas de structures globulaires sur les fibres de chromatine. Il faut cependant souligner qu'avec un rapport pondéral protéine MCl/ADN égal à 0,11, la protéine MCl est beaucoup moins abondante que les histones et que la protéine II de *T. acidophilum*. La protéine MCl ne peut donc organiser qu'une faible partie de l'ADN. Nos observations en microscopie électronique suggèrent d'ailleurs que des régions de l'ADN sont nues.

L'organisation de la chromatine de *M. barkeri* est radicalement différente de celle des eucaryotes et de *T. acidophilum*. Elle ressemblerait plutôt à celle des eubactéries : chez *E. coli*, le rapport protéine II/ADN est égal à 0,1, l'ADN est totalement digéré par la nucléase staphylococcale et la structure du nucléoïde est très instable.

CONCLUSIONS ET PERSPECTIVES

Notre travail a porté sur l'étude de la protéine chromosomale MCl isolée d'archaebactéries du genre *Methanosarcina* : *Methanosarcina barkeri* MS (souche de référence) et *Methanosarcina* sp MST-A1.

Nous avons contribué à la détermination de la séquence en acides aminés de la protéine MCl de *M. barkeri*. Cette protéine basique de 93 résidus, ne présente aucune homologie de structure primaire avec des protéines chromosomales d'eubactéries (protéine II) et d'eucaryotes (histones).

Dans le cadre d'études comparatives de la protéine MCl isolée de différentes souches de *Methanosarcinaceae*, nous avons établi la séquence en acides aminés de la protéine MCl de *Methanosarcina* sp MST-A1. Nos résultats ont permis de différencier deux souches très voisines : *Methanosarcina* sp MST-A1 et *Methanosarcina* sp CHTI-55 : on peut apprécier ainsi l'apport de la séquence peptidique à la taxonomie.

Nous avons mis au point une méthode de préparation de la protéine MCl, afin d'aborder l'étude de sa conformation. La protéine MCl est principalement constituée de segments en feuillets β et se désorganise lorsque la force ionique augmente. Ces propriétés confirment que la protéine MCl diffère des histones et de la protéine II : ces protéines sont plus riches en hélice α qu'en feuillet β et acquièrent davantage de structure ordonnée en milieu de force ionique élevée.

La protéine MCl a été localisée sur l'ADN par des observations en microscopie électronique, en utilisant des techniques d'immunomarquage. Par ailleurs, l'étude du complexe désoxyribonucléoprotéique a révélé l'absence de structures répétitives analogues aux nucléosomes des eucaryotes. La protéine ne serait pas distribuée uniformément le long de l'ADN.

La protéine MCl interagit bien avec l'ADN mais le rôle structural qu'elle pourrait avoir est très certainement différent de celui des histones ou de la protéine II d'*E. coli*.

Des études d'interaction de la protéine MCl avec l'ADN devraient permettre de préciser ce rôle. Il faudrait déterminer la conformation de la protéine en interaction avec l'ADN.

La place du résidu de tryptophane (Trp⁶¹) au coeur d'une région bien structurée suggère que ce résidu pourrait jouer un rôle essentiel. Il serait donc intéressant de voir s'il correspond au point de fixation de la protéine MCl sur l'ADN. Ce résidu de tryptophane est entouré de deux résidus de lysine (séquence Lys⁵⁹-Ala-Tryp-Lys⁶²) qui pourraient donner lieu à des liaisons électrostatiques avec les groupes phosphate de l'ADN. Le tryptophane pourrait alors s'insérer entre les bases de l'ADN. Ce type d'interaction correspondrait au modèle proposé par RAJESWARI et coll (1987).

APPENDICE TECHNIQUE

I SOUCHES BACTÉRIENNES ET CULTURE DES BACTÉRIES

La souche *Methanosarcina barkeri* souche MS (DSM 800) a été fournie par le Docteur J.G.ZEIKUS (University of Wisconsin Madison USA). Les bactéries sont cultivées à 35°C sur acétate dans le milieu BCYT (TOUZEL et ALBAGNAC, 1983) dont la composition est indiquée dans l'annexe 1, milieu de base contenant de l'extrait de levure, des dérivés carbonés et de la trypticase. Les cellules sont récoltées en phase post-exponentielle.

La souche *Methanosarcina* sp souche MST-A1 (DSM 2905) a été isolée à l'INRA dans un digesteur anaérobie. Les bactéries sont cultivées à 55°C sur le même milieu que *Methanosarcina barkeri* MS.

II PRÉPARATION DE LA PROTÉINE MCl

La protéine MCl est préparée à partir d'un homogénat bactérien par chromatographie d'affinité sur une colonne d'ADN couplé à la cellulose.

II - 1 PRÉPARATION DE L'HOMOGENAT BACTERIEN

Les bactéries (60 g) sont homogénéisées à +4°C dans 60 ml de tampon P-0 contenant du DFP 0,1 mM à l'aide d'un homogénéiseur Waring Blendor. Le tampon P est un tampon phosphate de sodium 10 mM, EDTA 1 mM, pH 6,8 contenant ou non du chlorure de sodium : la molarité en NaCl exprimée en mmoles est indiquée par le chiffre qui suit la lettre P. Les bactéries sont

"cassées" dans une presse de FRENCH à une pression de 1400 bars (20.000 psi). La molarité en NaCl de la suspension est amenée à 2 M avec du tampon P-4000 afin de dissocier le complexe nucléoprotéique. L'homogénat est ensuite centrifugé à 8000 x g pendant 1 heure pour éliminer les débris membranaires. Le surnageant est additionné de polyéthylène glycol 6000 (30 p.100, poids/volume) contenant du NaCl 2 M pour obtenir une concentration finale de 10 p.100 (poids/volume) en polyéthylène glycol. Après agitation douce pendant 30 minutes, la suspension est centrifugée 1 heure à 8000 x g pour éliminer les acides nucléiques qui ont précipité. Le surnageant est dialysé contre un tampon P-50 contenant du 2-mercaptoéthanol 1 mM pour amener la molarité en NaCl à 50 mM.

II - 2 CHROMATOGRAPHIE SUR COLONNE D'ADN COUPLE A LA CELLULOSE

II - 2-1 Préparation de la colonne d'ADN

La colonne d'ADN (ADN bicaténaire) a été préparée par la méthode de REICHART (1972).

** Préparation de la cellulose*

La cellulose Whatman CF 11 (100 g) est mise en suspension dans deux litres d'eau distillée. Les particules les plus fines sont éliminées par aspiration du surnageant après décantation de la suspension de cellulose. Cette opération est répétée deux fois après resuspension de la cellulose dans l'eau distillée et deux fois après resuspension dans un litre de mélange éthanol-eau (1:1 ; v/v). La cellulose est ensuite resuspendue dans l'éthanol absolu que l'on élimine ensuite par filtration sur un Büchner garni de papier Whatman 3MM. La cellulose est lavée deux fois avec un litre d'eau distillée. La

cellulose est traitée successivement par de la soude 0,5 M et par de l'acide chlorhydrique 0,5 M. Après chaque traitement, la cellulose est lavée avec de l'eau distillée et les solvants sont éliminés par filtration sur un Büchner. Après le traitement par la solution acide, le pH de l'eau de lavage doit être supérieur à 5. La cellulose est séchée pendant 18 heures sous courant d'air froid, puis lyophilisée pendant 48 heures.

* Couplage de l'ADN à la cellulose

Deux colonnes d'ADN ont été réalisées : une colonne avec de l'ADN d'*E.coli* et une colonne avec de l'ADN de *Rhizobium meliloti*. 200 mg d'ADN d'*E.coli* préparé par la méthode de MARMUR (1961) ou 200 mg d'ADN de *Rhizobium meliloti* préparé par la méthode de MIURA (1967) sont dissous dans 100 ml de tampon Tris-HCl 10 mM ajusté à pH 8,0 et contenant de l'EDTA 1 mM. Pour obtenir une bonne solubilisation, la solution d'ADN est maintenue sous agitation très douce pendant 18 heures. L'absorbance à 260 nm (A_{260}°) d'une partie aliquotée de cette solution diluée au centième est mesurée, ce qui permettra de calculer le rendement du couplage de l'ADN à la cellulose.

On ajoute la cellulose purifiée par petites quantités et sous agitation (1 g de cellulose pour 3 ml d'une solution d'ADN à 2 mg/ml). La "pâte" d'ADN-cellulose obtenue est séchée pendant la nuit sous un courant d'air froid, puis est lyophilisée pendant 48 heures.

L'ADN-cellulose est broyée dans un mortier afin d'obtenir une poudre fine, puis est resuspendue dans 400 ml d'éthanol absolu. La suspension d'ADN-cellulose est placée par fractions de 50 ml dans un cristallisateur dont la base a une surface de 125 cm² et est maintenu sous agitation pour être irradiée par une lampe UV avec une énergie par unité de surface de 100.000 ergs/mm².

L'ADN-cellulose est lavée avec de l'éthanol absolu, puis avec un litre de tampon Tris-HCl 10 mM pH 8,0 contenant de l'EDTA 1 mM. Le solvant est ensuite éliminé par filtration sur un Büchner.

L'absorbance à 260 nm (A_{260}^1) du filtrat est mesurée. La quantité d'ADN fixé sur la cellulose est calculée à partir de la différence ($A_{260}^0 \times 10$) - (A_{260}^1), sachant que 1 mg d'ADN a une absorbance de 0,02 unité de densité optique. Nous avons obtenu un rendement de fixation d'ADN sur la cellulose de 70 p.100.

II - 2-2 Conditions utilisées pour la chromatographie sur colonne d'ADN-cellulose

La chromatographie est effectuée à +4°C. Le dialysat de l'homogénat bactérien (~ 600 ml) clarifié par centrifugation pendant 45 minutes à 8000 x g est déposé sur la colonne d'ADN-cellulose (2,5 x 20 cm), équilibrée dans le tampon P-50. Après le dépôt, la colonne est lavée par le tampon d'équilibration jusqu'à ce que la densité optique de l'éluat revienne à la ligne de base. La colonne est ensuite éluée par le tampon P dont on augmente la molarité en NaCl comme il est indiqué dans l'exposé des résultats.

III CONCENTRATION DE LA PROTÉINE MCl

La concentration de la protéine MCl à l'état pur ou dans la fraction 3 éluée de la colonne d'ADN a été effectuée à 4°C par chromatographie d'échange d'ions sur une colonne de carboxyméthyl-TSK (CM-TSK).

La fraction protéique 3b est dialysée contre le tampon d'équilibration de la colonne, le tampon CM-50, en présence de DFP 0,05 mM. Le tampon CM est un tampon phosphate de sodium 10 mM pH 6,0 dont la molarité en NaCl exprimée en mM est indiquée par le chiffre qui suit les lettres CM. Le dialysat clarifié par centrifugation est déposé sur la colonne. Après le dépôt, la colonne (1,6 x 3 cm) est lavée par un volume de tampon d'équilibration. Le matériel retenu est alors élué par le tampon CM-1000. Après retour de la densité optique de l'éluat à la ligne de base, la colonne est lavée avec un volume d'HCl 0,02 M pour éluer tout matériel pouvant encore être fixé. La colonne est alors lavée abondamment par le tampon CM-50. Lorsque la colonne n'est pas utilisée, de l'azide de sodium est ajouté au tampon CM-50 à raison de 0,2 g/l afin d'éviter d'éventuelles contaminations bactériennes.

Après plusieurs chromatographies, la colonne est régénérée. La régénération s'effectue par passage d'un demi-volume d'HCl 0,5 M puis par deux volumes de tampon CM-1000. Elle est immédiatement suivie d'un rééquilibrage dans le tampon CM-50.

Pour la concentration de la protéine MCl à l'état pur, nous utilisons un gel nouvellement régénéré et nous faisons varier le volume du gel dans la colonne en fonction de la quantité de protéine à concentrer.

IV DOSAGE DE L'ADN ET DES PROTÉINES DANS LE COMPLEXE DÉSOXYRIBONUCLÉOPROTÉIQUE ISOLÉ DE *M. BARKERI*

IV - 1 DOSAGE DE L'ADN

La quantité d'ADN est déterminée dans la complexe désoxyribonucléoprotéique selon la méthode de BURTON (1968). Ce complexe a été purifié selon la technique de CHARTIER et coll. (en

préparation). Dans un premier temps, les débris cellulaires et les polysomes sont éliminés par ultracentrifugation. Le complexe désoxyribonucléoprotéique est alors concentré et purifié par chromatographie de gel filtration sur une colonne de Sepharose 4B après traitement avec la ribonucléase A.

Le dosage de l'ADN s'effectue en trois étapes :

1 - Élimination des nucléotides libres

L'échantillon à doser est lavé trois fois avec un volume d' HClO_4 1 M pendant 30 minutes à 0°C . Après centrifugation pendant 15 minutes à $2000 \times g$, le surnageant est éliminé.

2 - Hydrolyse de l'ADN

L'ADN est hydrolysé dans le culot résiduel, à 70°C pendant 15 minutes avec HClO_4 0,5 M sous agitation. L'échantillon est ensuite centrifugé à $2000 \times g$ pendant 15 minutes et le surnageant est récupéré. On répète l'opération une seconde fois et les deux surnageants sont réunis pour le dosage.

3 - Dosage proprement dit

* *Réactif à la diphénylalanine*

La diphénylalanine (MERCK) est dissoute dans 100 ml d'acide acétique glacial auquel on ajoute 1,5 ml d'acide sulfurique concentré.

* *Solution mère standard d'ADN*

De l'ADN d'érythrocyte de poulet est dissous à raison de 0,4 mg/ml à 70°C pendant 15 minutes dans l'acide perchlorique 0,5 M.

Une gamme étalon est réalisée en effectuant une série de dilutions de la solution mère d'ADN dans de l'acide perchlorique 0,5 M. 2 ml de réactif à la diphénylalanine sont ajoutés

à 1 ml de chaque solution de la gamme étalon ou à 1 ml de l'échantillon à doser. Après une incubation à 30°C pendant une nuit la coloration bleue résultante est lue à 600 nm au spectrophotomètre en prenant comme témoin blanc l'acide perchlorique 0,5 M.

IV - 2 DOSAGE DES PROTEINES

La quantité de protéine est déterminée dans le complexe désoxyribonucléoprotéique selon la méthode de SMITH et coll. (1985). Le principe est celui de la méthode de Lowry modifiée: en milieu alcalin les protéines réagissent avec les ions Cu^{2+} pour produire des ions Cu^{+1} . Le réactif BCA* réagit avec les ions Cu^{+1} pour former une coloration pourpre intense.

Une gamme étalon est réalisée en diluant une protéine standard (histone H4 de thymus de veau) dans le même solvant que l'échantillon de protéine à doser. A un volume de 100 μl de chaque solution de la gamme étalon, de l'échantillon de protéine à tester et du solvant de dilution (blanc) on ajoute 2 ml de réactif BCA. Le mélange est agité vigoureusement. Après une incubation à 60°C pendant 30 minutes et refroidissement à température ambiante, l'absorbance de chaque tube est lue à 562 nm contre l'eau dans un intervalle de temps inférieur à 10 minutes car la coloration continue à se développer. L'absorbance du blanc est soustraite de chaque valeur d'absorbance et la concentration en protéine de l'échantillon à tester est alors déterminée à partir de la courbe étalon.

* BCA : Bicinchoninic Acid.

V ISOÉLECTROFOCALISATION

L'isoélectrofocalisation est réalisée sur plaque de gel de polyacrylamide à 7,5 % d'acrylamide contenant 4,25 % de glycérol et 6 % d'un mélange d'ampholytes (LKB) formant un gradient de pH 3,5 à 10. Le gel est coulé entre deux plaques de verre (110x100x1 mm). L'électrophorèse est effectuée à l'aide d'un appareil LKB 2117 MULTIPHOR sous une puissance constante de 4 watts après une pré-électrophorèse de 30 min. dans les mêmes conditions. Le contact du gel avec les électrodes est réalisé avec des mèches épaisses imprégnées de soude 1 M à la cathode et d'acide phosphorique 1 M à l'anode. Au cours de la migration (2 h environ), l'ampérage décroît jusqu'à atteindre une valeur minimale constante ce qui indique que les échantillons sont focalisés à leur point isoélectrique (pHi).

La fixation des échantillons dans le gel est effectuée par une solution aqueuse d'acide trichloracétique à 12,5 % et d'acide sulfosalicylique à 3,4 % sous agitation pendant 1h30. Cette solution permet également d'éliminer les ampholytes. Le gel est ensuite coloré pendant 12 h par une solution de bleu de Coomassie (R-250) à 0,04 % dans le mélange solvant méthanol-eau-acide acétique (40:50:10, v/v) préparé extemporanément. La décoloration est effectuée par le mélange solvant. Des témoins de pHi connus sont utilisés pour établir une courbe de valeurs de pHi en fonction de la distance de migration.

VI DÉTERMINATION DE LA STRUCTURE QUATÉNAIRE DE LA PROTÉINE MCI PAR CHROMATOGRAPHIE DE GEL FILTRATION

La chromatographie s'effectue sur une colonne de Sephadex G-75 superfin (90x1,6 cm) éluée par du tampon phosphate 10 mM pH 6,0 contenant du chlorure de sodium 20 mM ou 400 mM.

Une courbe d'étalonnage est réalisée en reportant le K_{av}^* de protéines standard en fonction du logarithme de leur masse moléculaire relative. Le K_{av} est calculé à partir du volume d'élution (v_e) de chaque protéine standard. Le volume total de la colonne (v_t) correspond au volume d'élution du chlorhydrate de guanidine alors que le volume mort (v_o) correspond au volume du bleu dextran 2000.



$$* K_{av} = \frac{v_e - v_o}{v_t - v_o}$$

REFERENCES BIBLIOGRAPHIQUES

BIBLIOGRAPHIE

- ALBERTS, B. and HERRICK, G. (1971)
DNA-cellulose chromatography.
in Methods Enzymol. (COLOWICK, S.P. and KAPLAN, N.O., eds)
21, 198-217, Academic Press, New-York.
- AUBERT, J.P. (1981)
Le coude β dans les glycoprotéines : organisation spatiale
et N-glycosylation *in vitro* de peptides synthétiques.
Thèse d'Etat de l'Université de Lille I.
- BRADBURY, E.M., CARY, P.D., CRANE-ROBINSON, C., RATTLE, H.W.E.
BOUBLICK, M. and SAUTIERE, P. (1975)
Conformations and interactions of histone H2A (F2A2, ALK)
Biochem. 14, 1876-1885.
- BRAHMS, S. and BRAHMS, J. (1980)
Determination of protein secondary structure in solution by
vacuum ultraviolet circular dichroism.
J. Mol. Biol. 138, 149-178.
- BURTON, K. (1968)
Determination of DNA concentration with diphenylalanine.
in Methods Enzymol. 12, 163-166.
- CHANG, C.T., WU, C.S.C. and YANG, J.T. (1978)
Circular dichroic analysis of protein conformation : inclu-
sion of the β -turns.
Anal. Biochem. 91, 13-31.
- CHARTIER, F., LAINE, B., SAUTIERE, P., TOUZEL, J.P. and ALBAGNAC, G.
(1985)
Characterization of the chromosomal protein HMB isolated
from *Methanosarcina barkeri*.
FEBS Lett. 183, 119-123.
- CHOU, P.Y., FASMAN, G.D. (1978)
Empirical predictions of protein conformation.
Ann. Rev. Biochem. 47, 251-276.
- CID, H., BUNSTER, M., ARRIAGADA, E. and CAMPOS, M. (1982)
Prediction of secondary structure of proteins by means of
hydrophobicity profiles.
FEBS Lett. 150, 247-254.

- D'ANNA, J.A. and ISENBERG, I. (1972)
Fluorescence anisotropy and circular dichroism study of conformational changes in Histones IIB2.
Biochem. 11, 4017-4025.
- EISENBERG, D., WEISS, R.M. and TERWILLINGER, T.C. (1984)
The hydrophobic moment detects periodicity in protein hydrophobicity.
Proc.Natl.Acad.Sci.USA 81, 140-144.
- GABORIAUD, C., BISSERY, V., BENCHETRIT, T. and MORNON, J.P. (1987)
Hydrophobic cluster analysis, an efficient new way to compare and analyse amino acid sequence.
FEBS Lett. 224, 149-155.
- GARNIER, J., OSGUTHORPE, D.J. and ROBSON, B. (1978)
Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins.
J.Mol.Biol. 120, 97-120.
- KHANAKA, H. (1984)
Etude de la protéine chromosomale de type HU chez différentes espèces de Rhizobiaceae. Caractérisation, études structurales, interaction avec le DNA.
Thèse d'Etat de l'Université de Lille I.
- LAKOWICZ, J.R. (1983)
Principles of fluorescence spectroscopy.
Plenum Press, New York
- LAEMMLI, U.K. (1970)
Cleavage of structural proteins during the assembly of the head of bacteriophage T4.
Nature 227, 680-685.
- LEWIS, P.N., BRADBURY, E.M. and CRANE-ROBINSON, C. (1975)
Ionic strength induced structure in histone H4 and its fragments.
Biochem. 14, 3391-3400.*
- LIM, V.I. (1974)
Structural principles of the globular organization of protein chains. A stereochemical theory of globular protein secondary structure.
J.Mol.Biol. 88, 857-872.
- LIM, V.I. (1974)
Algorithms for prediction of α -helical and β -structural regions in globular proteins.
J.Mol.Biol. 88, 873-894.
- MANAVALAN, P., MITTELSTAEDT, D.M., SCHIMERLIK, M.I. and JOHNSON, W.C. (1986)
Conformational analysis of thymidylate synthase from amino acid sequence and circular dichroism.
Biochem. 25, 6650-6655.

- MARMUR, J. (1961)
A procedure for the isolation of deoxyribonucleoproteic acid from micro-organisms.
J.Mol.Biol. 3, 208-218.
- MILNER-WHITE, E.J. and POET, R. (1987)
Loops, bulges, turns and hairpins in proteins.
TIBS 12, 189-192.
- MIURA, K.I. (1967)
Preparation of bacterial DNA by the phenol pH 9 RNase method. in Methods Enzymol. (GROSSMAN, L., MOLDAVE, K., eds) 12, 543-545, Academic Press, New York.
- PANYIM, S. and CHALKLEY, R. (1969)
High resolution acrylamide gel electrophoresis of histones
Arch.Biochem.Biophys. 130, 337-346.
- REICHART, L. (1972)
Ph.D. Dissertation, Dept of Biochemistry Stanford University USA.
- RAJESWARI, M.R., MONTENAY-GARESTIER, T. and HELENE, C. (1987)
Does tryptophan intercalate in DNA ? A comparative study of peptide binding to alternating and nonalternating A-T sequences.
Biochem. 26, 6825-6831.
- SETTE, A., DORIA, G. and ADORINI, L. (1986)
A microcomputer program for hydrophilicity and amphipathicity analysis of proteins antigens.
Mol.Immunol. 23, 807-810.
- SMITH, P.K., KROHN, R.I., HERMANSON, G.T., MALLIA, A.K., GARTNER, F.H., PROVENZANO, M.D., FUJIMOTO, E.K., GOEKE, N.M., OLSON, B.J. and KLENK, D.C. (1985)
Measurement of protein using bicinchoninic acid.
Anal.Biochem. 150, 76-85.
- STEPHANO, J.L., GOULD, M. and ROJAS-GALICIA, L. (1986)
Advantages of picrate fixation for staining polypeptides in polyacrylamide gels.
Anal.Biochem. 152, 308-313.
- STERNBERG, M.J.E. (1986)
Prediction of protein structure from amino acid sequence.
Anti-cancer Drug Design. 1, 169-178.
- TANAKA, I., APPELT, K., DIJK, J., WHITE, S.W. and WILSON, K.S. (1984)
3-A resolution structure of a protein with histone-like properties in prokaryotes.
Nature 310, 376-381.

- TOUZEL, J.P. and ALBAGNAC, G. (1983)
Isolation and characterization of *Methanococcus mazei* strain MC3.
FEMS Lett. 16, 241-245.
- VAN der MAREL, H.W. and BEUTELSPACHER, H. (1976)
Atlas of infrared spectroscopy of clay minerals and their admixtures.
Elsevier, Sc. Publ. Comp. New York
- WU, T.T. and KABAT, E.A. (1971)
Proc. Natl. Acad. Sci. USA 68, 1501.

ANNEXE 1

Composition du milieu BCYT pH 7

H ₂ O bidistillée	800 ml
Solution macrominérale	50 ml
Solution vitamine	10 ml
Trace minérale	10 ml
NH ₄ Cl	1 g
Extrait de levure	0,5 g
Trypticase	0,5 g
Résazurine à 0,2 %	1 ml
KHCO ₃	2,1 g

Ajuster à pH 7. Répartir en flacons sous N₂/CO₂

Solution trace minérale

Titriplex I acide nitriloacétique	12,8 g/l
FeCl ₃ 6H ₂ O	1,35 g/l
MnCl ₂ 4H ₂ O	0,1 g/l
CoCl ₂ 6H ₂ O	0,024 g/l
CaCl ₂ 2H ₂ O	0,1 g/l
ZnCl ₂	0,1 g/l
CuCl ₂ 2H ₂ O	0,025 g/l
H ₃ BO ₃	0,01 g/l
Na ₂ MoO ₄ 2H ₂ O	0,024 g/l
NaCl	1 g/l
NiCl ₂ 6H ₂ O	0,12 g/l
Na ₂ SeO ₃ 5H ₂ O	0,026 g/l

Solution de vitamines

D(+) biotine	2 mg/l
Acide folique	2 mg/l
HCl-pyridoxal	10 mg/l
Riboflavine	5 mg/l
Acide nicotinique	5 mg/l
HCl-thiamine	5 mg/l
D(+) pantothénate de calcium	5 mg/l
Vitamine B12	0,1 mg/l
Acide 4-aminobenzoïque	5 mg/l
Acide DL- α -lipoïque	5 mg/l

Solution macrominérale

KH ₂ PO ₄	6 g/l
NaCl	12 g/l
MgCl ₂ 6H ₂ O	2 g/l
CaCl ₂ 2H ₂ O	1,6 g/l

Résumé

La découverte d'une troisième lignée d'êtres vivants (les archaebactéries) a ouvert une nouvelle perspective sur l'origine et l'évolution de la vie. L'étude des protéines chromosomales des bactéries méthanogènes qui forment avec les bactéries halophiles et thermoacidophiles la lignée archaebactéries constitue un apport original à la phylogénie permettant d'éclairer les relations entre procaryotes et eucaryotes, notamment au niveau de l'organisation de la chromatine.

La structure primaire de la protéine chromosomale MC1 isolée de deux bactéries méthanogènes : *Methanosarcina barkeri* MS et *Methanosarcina* sp MST-A1 a été déterminée. Cette protéine basique de 93 résidus d'acides aminés est caractérisée par un nombre élevé de résidus chargés qui sont distribués tout le long de la chaîne polypeptidique.

Pour étudier la conformation de la protéine MC1, une modification de la méthode de préparation a été apportée afin d'obtenir la protéine dans des conditions non dénaturantes. Cette étude a été effectuée à l'aide de méthodes spectroscopiques - dichroïsme circulaire, infrarouge, fluorescence - et de méthodes de prédiction de structure secondaire. La protéine MC1 est principalement constituée de segments en feuillets β et se désorganise lorsque la force ionique augmente.

La protéine MC1 ne présente aucune homologie de structure primaire ou secondaire avec les protéines chromosomales d'eubactéries (protéine II) et d'eucaryotes (histones).

Par des observations en microscopie électronique, la protéine MC1 a été localisée sur l'ADN en utilisant des techniques d'immunomarquage. L'étude du complexe désoxyribonucléoprotéique a révélé l'absence de structures répétitives analogues aux nucléosomes des eucaryotes et suggère que la protéine MC1 n'est pas distribuée uniformément le long de l'ADN.

Mots clefs : Archaebactéries - Bactéries méthanogènes - Protéine chromosomale - Structure primaire - Conformation - Complexe désoxyribonucléoprotéique - Microscopie électronique.