

Interactions faibles protéine – protéine en solution : La malate déshydrogénase halophile

Lionel Costenaro

▶ **To cite this version:**

Lionel Costenaro. Interactions faibles protéine – protéine en solution : La malate déshydrogénase halophile. Biochimie [q-bio.BM]. Université Joseph-Fourier - Grenoble I, 2001. Français. NNT : . tel-00007698

HAL Id: tel-00007698

https://theses.hal.science/tel-00007698

Submitted on 9 Dec 2004

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives| 4.0 International License

Université Joseph Fourier – Grenoble 1
Sciences - Technologie - Médecine

Thèse
Présentée et soutenue publiquement le 5 octobre 2001 par

Lionel COSTENARO

Pour obtenir le titre de
Docteur de l'Université Joseph Fourier – Grenoble 1
(Arrêtés ministériels du 5 juillet 1984 et du 30 mars 1992)

Spécialité : Cristallographie et RMN Biologiques (Physique)

Interactions faibles protéine – protéine en solution : La malate déshydrogénase halophile

---

Composition du jury :

Dr. Christine Ebel	Directeur de Thèse
Pr. Eva Pebay-Peyroula	Présidente
Dr. Madeleine Riès-Kautt	Rapporteur
Dr. Annette Tardieu	Rapporteur
Dr. Stéphane Veesler	Examinateur
Dr. Giuseppe Zaccai	Examinateur

Thèse préparée au Laboratoire de Biophysique Moléculaire
Institut de Biologie Structurale Jean-Pierre Ebel (CEA – CNRS – UJF), Grenoble

Université Joseph Fourier – Grenoble 1
Sciences - Technologie - Médecine

Thèse
Présentée et soutenue publiquement le 5 octobre 2001 par

Lionel COSTENARO

Pour obtenir le titre de
Docteur de l'Université Joseph Fourier – Grenoble 1
(Arrêtés ministériels du 5 juillet 1984 et du 30 mars 1992)

Spécialité : Cristallographie et RMN Biologiques (Physique)

Interactions faibles protéine – protéine en solution : La malate déshydrogénase halophile

---

Composition du jury :

Dr. Christine Ebel	Directeur de Thèse
Pr. Eva Pebay-Peyroula	Présidente
Dr. Madeleine Riès-Kautt	Rapporteur
Dr. Annette Tardieu	Rapporteur
Dr. Stéphane Veesler	Examinateur
Dr. Giuseppe Zaccai	Examinateur

Thèse préparée au Laboratoire de Biophysique Moléculaire
Institut de Biologie Structurale Jean-Pierre Ebel (CEA – CNRS – UJF), Grenoble

© Lionel Costenaro, 2001

Ce travail a été réalisé au sein du Laboratoire de Biophysique Moléculaire, sous la direction de Christine Ebel, à qui j'adresse tous mes remerciements. Je te suis très reconnaissant pour les qualités de ton encadrement et ta disponibilité. Merci pour ton accueil, ton amitié et tes conseils.

Je remercie également Giuseppe Zaccai pour m'avoir accueilli dans son équipe et initié à la biophysique.

Je suis honoré que M. Riès-Kautt et A. Tardieu aient accepté d'être rapporteurs de cette thèse. Je les remercie, ainsi que les autres membres du jury, E. Pebay-Peyroula, S. Veesler et G. Zaccai, d'avoir bien voulu évaluer ce travail et de leurs commentaires enrichissants.

Je suis reconnaissant à L. Belloni pour ces explications et ses programmes sans lesquels une partie de ce travail n'existerait pas. Merci aux " locals " de D22 et D11 de l'ILL et de D24 au Lure, et en particulier à R. May et P. Vachette, pour leurs conseils et leur disponibilité.

A chaque étape de ma thèse, de nombreuses personnes m'ont conseillé, expliqué ou critiqué ; je tiens à les remercier et particulièrement : A. Solovyova, G. Estèves, C. Pélissier pour leur collaboration, leur aide et mes premières expériences d'encadrement ; B. Franzetti, D. Madern et S. De Bonis pour mon initiation à la biochimie ; F. Metoz pour m'avoir fait profité de ses compétences en informatique ; C. Pfister pour ses relectures et conseils avisés.

A Patricia, Uschi, Eric et Frédouille, j'adresse un immense merci pour leur amitié, leur soutien, leur ironie et leurs critiques des courbes qui montent, descendent et dont la description est im… et pour tous les charmants moments passés ensemble. Merci à François et Jean-Sébastien qui se sont trompés de spécialité (ce n'est pas pharmacie) à Isa, Elsa, Stéphane, Florence, Carine, Martin, Moeava et bien d'autres…

Enfin, un grand merci à mes parents et à mes sœurs pour leurs soutiens indéfectibles, malgré l'aspect quelque peu obscur de mon travail.

LC

Sommaire

SYMBOLES ET ABRÉVIATIONS **9**

1. INTRODUCTION **13**

1.1. Pourquoi étudier les interactions protéine – protéine de la malate déshydrogénase de *Haloarcula marismortui* ? **16**
1.1.1. La vie dans des conditions extrêmes 16
1.1.2. Les organismes et les protéines halophiles 17
1.1.3. La Hm MalDH, une malate déshydrogénase halophile 19
1.1.4. Structure cristallographique de la Hm MalDH 20
1.1.5. Stabilité et interactions protéine – solvant de la Hm MalDH 22
1.1.6. Les interactions protéine – protéine de la Hm MalDH 24

1.2. Les interactions intermoléculaires en solutions **24**
1.2.1. Les répulsions de volume exclu 26
1.2.2. Les interactions purement électrostatiques 27
L'interaction coulombienne (charge/charge) 27
L'interaction charge/dipôle permanent 28
L'interaction dipôle/dipôle 29
1.2.3. Les interactions de polarisation 29
L'interaction charge/dipôle induit 30
L'interaction de Debye (dipôle/dipôle induit) 30
Les forces de dispersion 30
1.2.4. Les effets généraux du solvant sur les interactions 31
1.2.5. Les forces de van der Waals 32
La constante de Hamaker 32
La théorie de Lifshitz 32
1.2.6. L'eau : des interactions particulières 33
L'hydratation 34
Les effets hydrophile et hydrophobe 34
1.2.7. Les effets du solvant et des co-solvants sur les interactions colloïde – colloïde 35
La force attractive de déplétion 35
Effets des co-solvants salins 37

1.3. La mesure et la modélisation des interactions intermoléculaires en solution **39**
1.3.1. Le développement du viriel 39
1.3.2. La modélisation des facteurs de structure 40

1.4. Notre démarche pour la caractérisation et les implications des interactions protéine – protéine de la *Hm* MalDH **41**

2. DE LA PURIFICATION DE LA *HM* MALDH AUX INTERACTIONS PROTÉINE – PROTÉINE **43**

2.1. De la cellule aux échantillons **45**
2.1.1. Purification de la Hm MalDH 45

Matériel utilisé 45
Surexpression de la *Hm* MalDH chez *E. coli* 46
Purification de la *Hm* MalDH 46
Mesure de concentration 49
2.1.2. Préparation des échantillons pour les études biophysiques 49
2.1.3. Récupération des échantillons 50

2.2. La diffusion aux petits angles 50
2.2.1. Principes généraux de la diffusion aux petits angles 50
2.2.2. Quelques paramètres structuraux 52
Solution idéale et monodisperse : facteur de forme, masse et rayon de giration de la particule 52
Particules homogènes en interaction : facteur de structure, masse et rayon de giration apparents 53
2.2.3. Diffusion de neutrons aux petits angles 55
Aspects expérimentaux 55
Traitement des données 56
2.2.4. Diffusion de rayons X aux petits angles 57
Aspects expérimentaux 57
Traitement des données 57

2.3. Les interactions protéine – protéine en solution 58
2.3.1. Développement du viriel 58
Détermination du A_2 par diffusion aux petits angles 58
Détermination de A_2 par vitesse de sédimentation 60
2.3.2. Simulations numériques 61
Equations d'Ornstein-Zernike (OZ) et intégrales 62
Principe du calcul numérique 63
Résolution pratique et stratégie ? 64

3. UNE NOUVELLE MÉTHODE DE DÉTERMINATION DU SECOND COEFFICIENT DU VIRIEL 69

3.1. Présentation de l'article 71
Détermination de A_2 par vitesse de sédimentation 72
Validation de la méthode 73

3.2. Article 1 : Non-ideality by sedimentation velocity of halophilic malate dehydrogenase in complex solvents 74

4. INTERACTIONS PROTÉINE – PROTÉINE EN DIFFÉRENTS SELS 89

4.1. Les courbes expérimentales de diffusion 93

4.2. Les seconds coefficients du viriel A_2 96
4.2.1. Evolution de A_2 en fonction de la nature et de la concentration du sel 96
4.2.2. Effet de charge : pH et mutant de la Hm MalDH 98
4.2.3. Effet du D_2O 99

4.3. Modélisation des facteurs de structure 100
4.3.1. Deux potentiels répulsifs suffisent-ils en NaCl, KCl et CsCl ? 101
4.3.2. Un potentiel attractif supplémentaire en $MgCl_2$, $MgSO_4$ et $(NH_4)_2SO_4$ 104
4.3.3. En bref… 108
4.3.4. Limites de la méthode 109

4.4. Discussion **111**
4.4.1. Effet de volume exclu 111
4.4.2. Effet électrostatique 114
4.4.3. Le rôle des interactions protéine – solvant dans les interactions protéine – protéine 117
Second coefficient du viriel et interactions protéine – solvant … 117
… pour la *Hm* MalDH 119
Hydratation et solubilité 121
Effets spécifiques 122
4.4.4. Halophilisme et interactions protéine – protéine 122

5. COMPRÉHENSION DE LA CRISTALLISATION DE LA *HM* MALDH PAR DILUTION **125**

5.1. Présentation de l'article **127**

5.2. Article 2 : Understanding the crystallisation of an acidic protein by dilution in the ternary NaCl – 2-methyl-2,4-pentanediol – H_2O system **130**

6. MODIFICATION DES INTERACTIONS PAR L'AJOUT DE MPD **143**

6.1. Diagrammes de phases des mélanges sel – MPD – H_2O **145**

6.2. Stabilité et solubilité de la *Hm* MalDH dans les mélanges sel – MPD – H_2O **147**
6.2.1. Stabilité 147
6.2.2. Solubilité 149

6.3. Détermination de seconds coefficients du viriel par ultracentrifugation **149**

6.4. Les données expérimentales de diffusion de neutrons **150**
6.4.1. Les courbes de diffusion 150
6.4.2. Stabilité des solutions de Hm MalDH dans les mélanges $(NH_4)_2SO_4$ – MPD – H_2O 152

6.5. Les seconds coefficients du viriel A_2 **156**
6.5.1. Evolution des A_2 en présence de MPD 156
6.5.2. A_2 et cristallisation 158

6.6. Modélisation des facteurs de structure **159**
6.6.1. Mélanges NaCl – MPD – H_2O 159
6.6.2. Mélanges $MgCl_2$ – MPD – H_2O 162
6.6.3. Mélanges $(NH_4)_2SO_4$ – MPD – H_2O 162
6.6.4. MPD et potentiel attractif 164

6.7. Pourquoi le MPD modifie-t-il les comportements de la *Hm* MalDH ? **166**
6.7.1. Les effets du MPD sur la Hm MalDH 166
6.7.2. Remarque préliminaire 168
6.7.3. Cadre de l'analyse 168
6.7.4. MPD et stabilité de la Hm MalDH 169
6.7.5. MPD et interaction protéine – protéine 171
L'attraction induite par le MPD 171
Précipitation de la protéine par le MPD 172

7. CONCLUSION **177**

Une nouvelle méthode de détermination du second coefficient du viriel 179

Interactions protéine – protéine en différents sels 179
Compréhension de la cristallisation de la *Hm* MalDH par dilution 181
Modification des interactions par l'ajout de MPD 182

8. ANNEXES 185

8.1. Constante diélectrique statique 187
8.1.1. Solutions aqueuses de sel 187
8.1.2. Solutions aqueuses de MPD 188
8.1.3. Solutions aqueuses de sel et de MPD 189

8.2. Conditions et résumé des résultats des expériences de diffusion 190

8.3. Second coefficient du viriel et interactions protéine – solvant 193

8.4. Détermination des interactions protéine – solvant 194
8.4.1. Détermination de $(\partial m_3 / \partial m_2)_\mu$ 195
8.4.2. Détermination de N_1 et N_3 195

RÉFÉRENCES BIBLIOGRAPHIQUES 197

Symboles et abréviations

Symboles*

A	Constante d'Hamaker (J)
a	Rayon atomique ou moléculaire (m)
A_2	Second coefficient du viriel ($mol.l.g^{-2}$)
C	Concentration ($M = mol.l^{-1}$)
c	Concentration ($mg.ml^{-1}$)
c(r)	Fonction de corrélation directe
D	Diamètre (Å) ou coefficient de diffusion translationnelle
d	Portée de potentiel (Å)
E	Champ électrique ($V.m^{-1}$)
e	Charge élémentaire (valeur absolue de la charge de l'électron $1.602\ 10^{-19}$ C)
E_i	Energie de première ionisation (J)
f	Coefficient de friction
F	Force (N)
g(r)	Fonction de distribution de paires ou fonction de distribution radiale
h(r)	Fonction de corrélation totale
I	Force ionique ($= 0.5\ \sum c_i Z_i^2$)
I(q,c)	Intensité diffusée
J	Constante d'interaction
k	Constante de Boltzmann ($1.381\ 10^{-23}\ J.K^{-1}$)
k'_s, k_D	Coefficient de dépendance en concentration de s et D
L_B	Longueur de Bjerrum ($= e^2/4\pi\varepsilon_0\varepsilon kT$, ≈ 7.2 Å à 25°C dans l'eau)
M	Masse moléculaire (130552 Da pour la *Hm* MalDH)
M_+	Masse moléculaire d'une protéine de valence z, avec ses z contre-charges
m	Concentration molale ($mol.kg^{-1}$ de solvant principal, l'eau)
$\partial m_3/\partial m_2$	Paramètre d'interaction préférentielle de la protéine avec le sel (mol/mol)
n	Indice de réfraction
N_A	Nombre d'Avogadro ($6.022\ 10^{23}\ mol^{-1}$)
N_i	Paramètre d'interaction avec le solvant : nombre de particules i associées à la protéine (mole de i / mole de protéine)
P(q), I(q,0)	Facteur de forme
pH	Potentiel hydrogène ($= -\log_{10} [H_3O^+]$)
q	Vecteur de diffusion ($Å^{-1}$)
r	Distance (interparticulaire) (m, Å)
R_g	Rayon de giration (Å)
s	Coefficient de sédimentation (S)
S(q,c)	Facteur de structure
T	Température (K, °C)
u	Moment dipolaire (C.m)

* Les grandeurs vectorielles sont en gras.

V	Volume
$\bar{v}$	Volume partiel spécifique ($ml.g^{-1}$)
W, w	Energie libre d'interaction ou potentiel d'interaction (J)
Z, z	Charge (en unité de e, C)
$\varepsilon_{280}^{0.1\%}$	Coefficient d'extinction à 280 nm ($ml.mg^{-1}.cm^{-1}$)
Π	Pression osmotique
α	Polarisabilité ($C^2.m^2.J^{-1}$)
β	1/(kT)
$\delta(x)$	Fonction de Dirac
ε_0	Permittivité du vide ($8.854\ 10^{-12}\ C^2.J^{-1}.m^{-1}$)
ε	Permittivité relative, constante diélectrique statique
κ	Inverse de la longueur de Debye ($= 1/\lambda_D$) ou constante d'écrantage
λ_D	Longueur de Debye (m, Å)
μ	Potentiel chimique
ν	Fréquence
ρ	Densité numérique (cm^{-3})
ρ_{el}	Densité électronique
ρ_n	Densité de longueur de diffusion de neutron
ω	Vitesse angulaire
<<, >>	Très inférieur à, très supérieur à
$\cap$	Intersection
$\otimes$	Produit de convolution
Δx	Erreur sur x
$\approx$	Approximativement égal à
X°	Limite idéale de x, *i.e.* à dilution infinie en protéine (c = 0)
[x]	Concentration de x

Unités

$1\ Å = 10^{-10}$ m
$1\ Da = 1\ g.mol^{-1}$
$1\ eV = 1.602\ 10^{-19}$ J
$1\ kT = 4.045\ 10^{-21}$ J à 293 K (≈20°C)
$1\ M = mol.l^{-1} = 6.022\ 10^{23}$ molécules par l
$1\ S = 10^{-13}$ s (S : Svedberg)

Abréviations

Cste	Constante
DO	Densité optique, absorbance (su)
DTT	Dithio-1,4-thréitol
E. coli	*Escherichia coli*
Hm MalDH	Malate déshydrogénase de *Haloarcula marismortui*
HNC	Hypernetted Chain equation
IPTG	Isopropyl-1-ß-D-thio-1-galactopyranoside

Mes	Acide 2-[N-morpholino]-éthane-sulfonique
MPD	Méthyl-2-pentanediol-2,4
NAD^+	Amide nicotinique-adénine-dinucléotide, forme oxydée
NADH	Amide nicotinique-adénine-dinucléotide, forme réduite
OZ	Equation d'Ornstein-Zernike
rpm	Rotation par minute
TF, ^	Transformée de Fourier
Tris	Tris(hydroxyméthyl)-aminométhane

1. INTRODUCTION

La cellule est un milieu extrêmement concentré où les interactions entre macromolécules, même faibles, jouent un grand rôle. D'une part, les macromolécules doivent rester solubles malgré l'encombrement moléculaire, et d'autre part, il semble maintenant évident que les fonctions biologiques sont assurées par l'intermédiaire de complexes labiles mais spécifiques, de grande taille et associant de nombreuses protéines et parfois des acides nucléiques. L'étude des interactions protéine –protéine même faibles est donc essentielle. Si l'auto-association des protéines en multimères ou l'association d'une protéine et d'un ligand sont couramment étudiées, les interactions faibles protéine – protéine en solution qui modulent les propriétés des solutions protéiques concentrées et leur donnent un caractère non-idéal doivent néanmoins être prises en compte en biologie. Ces interactions, via les propriétés des macromolécules et du solvant, gouvernent la répartition spatiale des macromolécules dans la solution et déterminent donc leur tendance à rester en suspension ou à former une phase distincte amorphe (précipité, agrégat) ou ordonnée (cristal). Elles ont depuis longtemps fait l'objet de nombreuses études dans les domaines de la physique et de la chimie des colloïdes du fait de leur importance dans l'industrie (peinture, agroalimentaire, cosmétique). Un colloïde est une dispersion de particules dont la taille est très supérieure à celle des molécules de solvant, comprise dans une gamme assez large entre 1 nm et 1000 nm, la limite supérieure étant un peu arbitraire, le critère principal étant l'absence de sédimentation des particules sous l'effet de leur propre poids. Il existe des types très différents de colloïdes selon les phases physiques (solide, liquide ou gaz) du composé dispersé et du milieu de dispersion [Gerschel, 1995]. Les protéines peuvent donc être qualifiées de colloïdes. Cette approche " colloïdale " a été utilisée avec succès en biologie pour expliquer la transparence du cristallin de l'œil [Delaye et Tardieu, 1983; Vérétout *et al.*, 1989] puis dans le cadre de l'étude de la cristallogenèse des macromolécules biologiques [Guilloteau, 1991; Ducruix *et al.*, 1996]. Les études des interactions faibles protéine – protéine ont en particulier pour objet la compréhension de ce processus qui représente toujours une étape cruciale et fréquemment difficile pour la détermination des structures tridimensionnelles à haute résolution des macromolécules biologiques par cristallographie.

Le présent travail qui porte sur l'étude des interactions protéine – protéine en solution de la malate déshydrogénase halophile (*Hm* MalDH) s'inscrit dans cette démarche. Il tente d'analyser l'influence des interactions protéine – solvant sur ces interactions interparticulaires avec l'intention d'apporter des éléments de réponse au problème de la cristallogenèse et aussi à l'adaptation moléculaire aux conditions hyper salines.

1.1. Pourquoi étudier les interactions protéine – protéine de la malate déshydrogénase de *Haloarcula marismortui* ?

1.1.1. La vie dans des conditions extrêmes

En dehors des conditions physico-chimiques " normales " (d'un point de vue presque anthropocentrique) la vie existe. Les organismes qui se développent dans ces biotopes " hostiles " sont globalement qualifiés d'extrêmophile (la racine grecque *philos* signifie ami). Ils ont dû s'adapter à des valeurs extrêmes d'un ou de plusieurs paramètres physico-chimiques (Tableau 1.1). On peut citer par exemple la température pour les hyperthermophiles qui croissent de manière optimale au-dessus de 80°C, près des geysers ou près des cheminées volcaniques. A l'inverse, les psychrophiles préfèrent des températures inférieures à 15°C. Les barophiles peuvent supporter des pressions allant jusqu'à 1000 atmosphères dans les grands fonds marins. Les halophiles se développent dans des environnements où la concentration en sels approche de la saturation. Les méthanogènes qui sont des anaérobes se rencontrent dans les aires de décomposition. Les autres paramètres physico-chimiques sont le pH, l'absence d'eau, les radiations ou un environnement chimique particulier (fortes concentrations en métaux, atmosphère de CO_2, …).

Le monde vivant peut être divisé en trois domaines : *Eucarya*, *Bacteria* et *Achaea.* Le troisième, les archaébactéries, a été découvert par l'étude des extrêmophiles qui ont donc bouleversé l'arbre phylogénétique universel [Woese et Fox, 1977]. Les trois domaines du vivant sont représentés dans les divers organismes extrêmophiles. Ils ont cependant en commun le fait d'avoir développé des mécanismes moléculaires d'adaptation aux conditions physico-chimiques particulières. Cette adaptation est la raison principale de leur étude

intensive pour leur intérêt tant industriel qu'académique. Les extrêmophiles et leurs protéines dont le potentiel économique est considérable sont utilisés tels quels ou comme modèles pour modifier des enzymes mésophiles dans le but d'améliorer certaines de leurs propriétés comme leur stabilité à haute température. Les principaux domaines d'application sont l'agriculture, l'industrie chimique (notamment pour les lessives), l'agroalimentaire, l'industrie pharmaceutique. Une récente revue détaille le monde des extrêmophiles et donne de nombreux exemples de leur utilisation [Rothschild et Mancinelli, 2001].

Paramètre	Définition	Type
Température	Croissance > 80°C 60-80°C < 15°C	Hyperthermophile Thermophile Psychrophile
Salinité	NaCl 2-5 M, …	Halophile
pH	> 9 Jusqu'à pH 0	Alcaliphile Acidophile
Pression	Jusqu'à 1000 atmosphères Vide	Barophile -
Dessiccation		Xérophile
Oxygène	Intolérant	Anaérobe
Radiation, gaz (CO_2), forte concentration de métal, …		

Tableau 1.1 : Classification des organismes extrêmophiles selon le paramètre physico-chimique limitant.

Dans cette étude, nous nous intéresserons aux effets des solvants sur les interactions faibles en solution d'une protéine halophile, la malate déshydrogénase halophile que nous allons présenter.

1.1.2. Les organismes et les protéines halophiles

Parmi les extrêmophiles, les organismes halophiles (du grec *halos*, sel et *philos*, ami) prolifèrent dans des biotopes où le sel, essentiellement NaCl, a été concentré par évaporation parfois jusqu'à la limite de sa solubilité (4 à 6 M). On les trouve par exemple dans les grands lacs salés de l'Ouest américain, de la Mer Morte en Israël, du Lac Rose au Sénégal, ou plus proche de nous dans les marais salants français. Ils peuvent représenter l'essentiel des formes

vivantes de ces environnements et la plupart d'entre eux contiennent des pigments qui donne à l'eau une couleur rose, orangée ou pourpre caractéristique (Figure 1.1).

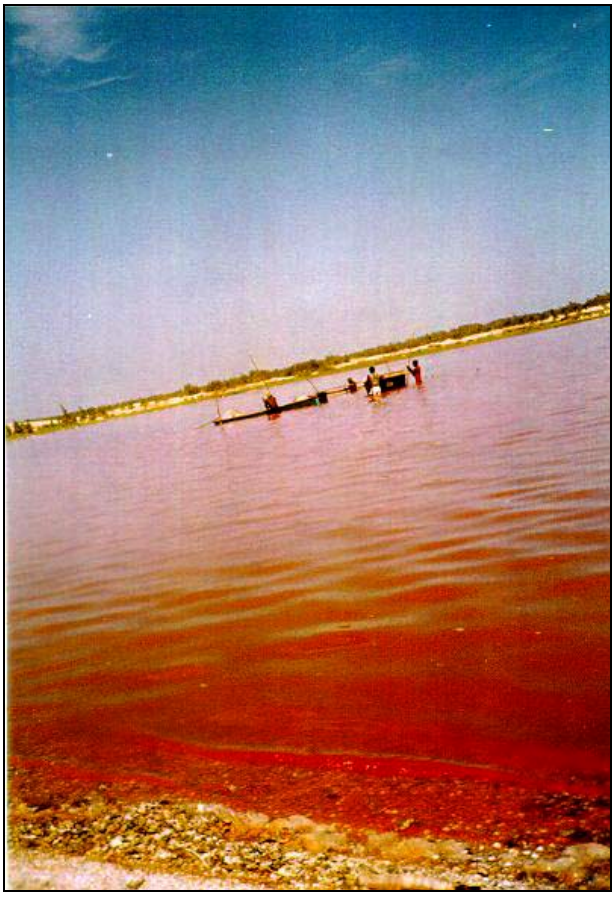


Figure 1.1 : Vue du Lac Rose au Sénégal.

La présence de fortes quantités de sel est nécessaire aux halophiles : ils se développent de façon optimale dans des milieux contenant au moins 2.5 M de NaCl et meurent rapidement en dessous de 1 M. Contrairement aux " halotolérants " qui compensent la forte pression osmotique externe due aux salinités très élevées par une accumulation dans leur cytoplasme de petites molécules, par exemple le glycérol, les halophiles accumulent essentiellement du KCl à des concentrations proche de la saturation ! Toutes les protéines de ces organismes sont donc solubles et fonctionnelles (stables, actives et flexibles) à de telles salinités, mais en plus, elles se dénaturent dès que la concentration en KCl diminue (en dessous de 1-2 M). Elles sont en ce sens qualifiées de protéines halophiles [voir les revues suivantes et leurs références Eisenberg *et al.*, 1992; Madern *et al.*, 2000]. Les mécanismes d'adaptation moléculaire à ces conditions extrêmes ont été étudiés et sont exposés ci-dessous pour la *Hm* MalDH.

1.1.3. La Hm *MalDH, une malate déshydrogénase halophile*

La malate déshydrogénase (*Hm* MalDH) de *Haloarcula marismortui* (une archaébactérie de la Mer Morte), est certainement la plus étudiée de toutes les protéines halophiles. Elle fut purifiée et caractérisée pour la première fois en 1977 [Mevarech *et al*., 1977; Mevarech et Neumann, 1977]. Par la suite, elle a été surexprimée avec succès dans *Escherichia coli* [Cendrin *et al*., 1993]. Il s'agit d'une enzyme cytoplasmique de 303 acides aminés (par monomère) qui est organisée en homo-tétramère de 130 552 Da [Bonneté *et al*., 1993]. Elle contient plus de résidus acides (aspartique et glutamique) et moins de résidus basiques (lysine) que son homologue mésophile, mais aussi plus de petits résidus hydrophobes (Tableau 1.2), comme d'autres protéines halophiles [Madern *et al*., 1995]. Cette accumulation de résidus acides est donc un mécanisme d'adaptation aux fortes concentrations de sels. Son point isoélectrique est à pH 4.5 et elle a 156 charges négatives à pH 8, ce qui traduit bien son caractère acide extrême.

Type	Nombre	Type	Nombre	Type	Nombre	Type	Nombre
A (ALA)	112	G (GLY)	124	M (MET)	16	S (SER)	68
C (CYS)	0	H (HIS)	28	N (ASN)	36	T (THR)	60
D (ASP)	140	I (ILE)	64	P (PRO)	44	V VAL)	120
E (GLU)	108	K (LYS)	32	Q (GLN)	48	W (TRP)	8
F (PHE)	32	L (LEU)	80	R (ARG)	60	Y (TYR)	32

Tableau 1.2 : Composition en acides aminés de la *Hm* MalDH tétramérique.

La *Hm* MalDH fait partie de la famille des déshydrogénases dépendantes du NAD (amide nicotinique-adénine-dinucléotide) et catalyse la réaction d'oxydoréduction transformant le malate en oxaloacétate :

$$\text{Malate} + NAD^+ \quad <\!\!---\!\!> \quad \text{Oxaloacétate} + NADH + H^+.$$

Cette réaction est en particulier la dernière du cycle de Krebs (cycle des acides tricarboxyliques) qui intervient à la fin de la dégradation du glucose en présence d'oxygène.

Contrairement à la plupart des malates déshydrogénases qui sont dimériques, la *Hm* MalDH est tétramérique et possède de plus fortes homologies de séquence et de structure

quaternaire avec les lactates déshydrogénases (LDH). De ce fait, elle appartient à une famille intermédiaire entre les MalDH et les LDH : les " LDH-like MalDH " [Madern, 2000].

1.1.4. Structure cristallographique de la **Hm** *MalDH*

La *Hm* MalDH cristallise dans une solution aqueuse de NaCl et de méthyl-2-pentanediol-2,4 (MPD) selon un protocole original de dilution [Richard *et al.*, 1995]. Nous verrons dans le cadre de cette étude quels sont les déterminants de cette cristallogenèse particulière. Parmi les trois structures cristallographiques tridimensionnelles de protéines halophiles connues (en mai 2001), la structure de la *Hm* MalDH fut la première à être déterminée et la seule à être oligomérique [Dym *et al.*, 1995]. Malgré la faible résolution (3.2 Å) de cette structure, d'importantes caractéristiques – qui pourraient traduire l'adaptation moléculaire aux fortes salinités – ont été mises en évidence. Les hélices α sont stabilisées par la présence de résidus alanines et de résidus acides près de leur extrémité N-terminale. La *Hm* MalDH contient aussi plus de ponts salins que son homologue mésophile. Ils sont localisés aux interfaces entre les monomères et stabilisent la structure quaternaire (Figure 1.2-a). De plus, la protéine est caractérisée par une accumulation exceptionnelle de résidus acides répartis, de manière relativement homogène, sur toute sa surface (en rouge dans les photos b à d), qui est donc extrêmement chargée négativement.

Plus récemment, les structures cristallographiques d'un mutant de la *Hm* MalDH et de la protéine sauvage ont été obtenues à des résolutions de 2.6 et 2.9 Å qui révèlent de nouvelles interactions protéine – solvant en rapport avec l'adaptation halophile [Richard *et al.*, 2000]. L'enzyme peut être considéré comme un dimère de dimères (A-B et D-C sur la Figure 1.2-a) reliés par un large réseau de molécules d'eau structurées dans la cavité centrale et par des ponts salins multiples aux deux extrémités. Ces derniers sont bloqués spécifiquement par des ions Cl^-, ce qui renforce leur stabilité à haut sel. De manière similaire, les ponts salins complexes entre les monomères des dimères sont fermés par des ions Na^+.

Ces analyses révèlent que l'adaptation halophile ne vise pas tant à protéger la protéine des fortes concentrations de sel, mais au contraire à les exploiter. En ce qui concerne les interactions protéine – protéine, il faut aussi noter que la protéine est extrêmement chargée négativement à sa surface et qu'elle est globulaire et peut être raisonnablement considérée comme quasi-sphérique à basse résolution.

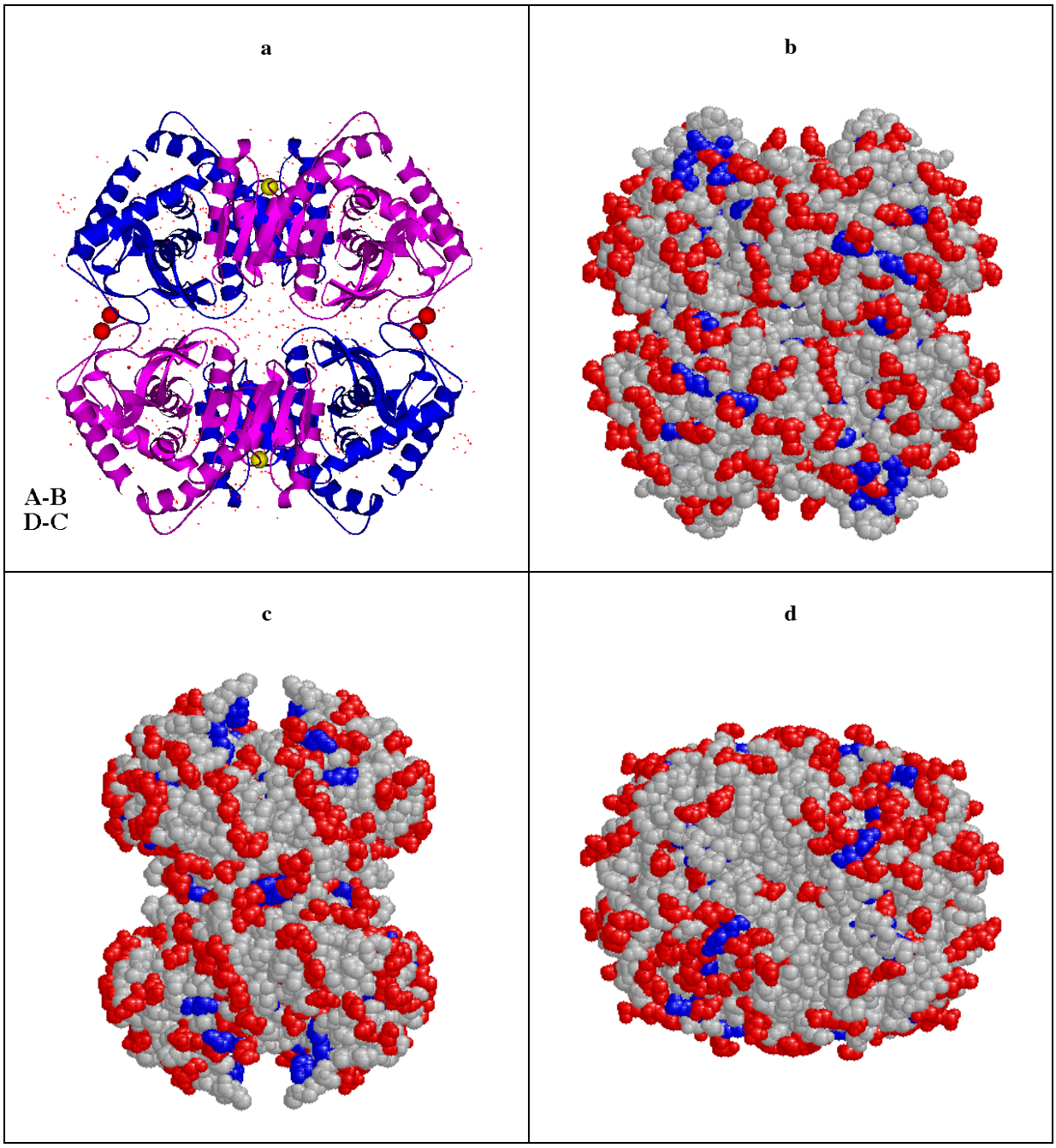


Figure 1.2 : Structure quaternaire de la *Hm* MalDH. **(a)** Représentation en ruban des quatre chaînes polypeptidiques A, B, C et D : les deux ions Na^+ (boules jaunes) participent aux ponts salins entre les monomères des deux dimères A-B et D-C et les quatre ions Cl^- (boules rouges) aux ponts salins entre les deux dimères. Les molécules d'eau sont représentées par les points rouges [Richard *et al*., 2000]. Les dimensions externes du tétramère sont : A-B 77 Å, A-D 80 Å et A-C 90 Å. **(b, c, d)** Les atomes des résidus acides et basiques sont représentés respectivement en rouge et en bleu (vues de face, de profil et de dessus).

1.1.5. Stabilité et interactions protéine – solvant de la Hm *MalDH*

Lorsque la concentration en sel varie, dans la plupart des conditions salines étudiées, l'état de repliement de la *Hm* MalDH évalué par fluorescence ou dichroïsme circulaire évolue de façon similaire à l'activité résiduelle* de l'enzyme, démontrant que l'inactivation et le dépliement de la *Hm* MalDH sont concomitants. D'une manière générale, la *Hm* MalDH se dénature à faible salinité (en dessous de 2 M de NaCl par exemple) et une concentration croissante de sel stabilise sa forme repliée puis la déstabilise parfois (au-dessus de 1.5 M de $MgCl_2$ par exemple). Sa stabilité dépend donc à la fois de la nature et de la concentration du sel. Les effets apparents des anions et des cations s'ajoutent et leur efficacité à maintenir la protéine sous forme native suit l'ordre suivant :

$$Ca^{+} \approx Mg^{2+} > NH_4^{+} \approx Na^{+} > K^{+} > Cs^{+} \text{ et } SO_4^{2-} \approx CH_3CO_2^{-} > Cl^{-},$$

pour la transition à bas sel. L'ordre pour la transition à haut sel est différent pour les cations. Les anions de forte densité de charge sont toujours les plus efficaces à stabiliser la forme repliée de la protéine en accord avec l'ordre de la série d'Hofmeister, alors que les cations de forte densité de charge ne sont stabilisants qu'à faibles concentrations. Ces effets sont liés aux interactions différentes de la protéine dépliée et repliée avec l'eau et les ions du solvant. Des sites de fixations des ions spécifiques à la forme repliée ou dépliée de la protéine sont nécessaires pour expliquer l'ordre des ions [Ebel *et al.*, 1999a].

Contrairement aux études cristallographiques qui ne permettent de détecter que les molécules de solvant immobilisées dans la même position dans toutes les protéines du cristal, des études en solution permettent d'évaluer les interactions protéine – solvant dans leur globalité. Elles utilisent des méthodes biophysiques qui déterminent la masse molaire de la protéine en solution de manière rigoureuse : l'ultracentrifugation analytique et la diffusion aux petits angles. Une protéine est entourée d'une couche de solvatation (eau et co-solvants) dont la composition diffère de celle du solvant de masse (non-perturbé). Les interactions entre la *Hm* MalDH et le solvant ont été étudiées dans différents sels et interprétées dans le cadre du modèle de la particule invariante. Ce modèle est le plus simple et consiste à considérer qu'une mole de protéine est associée à N_1 moles d'eau et N_3 moles de sel, quantités constantes pour

des concentrations variables de sel dans le solvant [Eisenberg, 1976, 1981; Bonneté *et al.*, 1993; Ebel, 1995]. Puisque la particule est définie comme la protéine accompagnée de ses contre-ions, N_3 peut être positif ou négatif. Le dernier cas correspond à la dissociation d'une partie des contre-ions.

La couche de solvatation de la *Hm* MalDH (et d'autres protéines halophiles) est à la fois très riche en eau et en sel. Si son hydratation est similaire à celle des protéines non-halophiles, la quantité de sel qui est associée à la *Hm* MalDH est supérieure d'un ordre de grandeur environ en NaCl et en KCl [Bonneté *et al.*, 1993; Ebel *et al.*, 1999b]. Les interactions avec le solvant dépendent fortement de la nature du sel (Tableau 1.3). Dans le cadre du modèle de la particule invariante, la *Hm* MalDH est associée en $MgCl_2$ à deux fois plus d'eau que dans les autres sels. De plus, si elle est associée à beaucoup de sel en $MgCl_2$ et en NaCl, elle en lie moins en $(NH_4)_2SO_4$ et ses contre-ions se dissocient en CsCl [Ebel, en préparation]. Au lieu de se protéger contre les salinités extrêmes, les protéines halophiles s'associent donc à de grandes quantités de sels qui semblent nécessaires à leur stabilité et expliqueraient leur dénaturation à basse salinité. Nous verrons que cette association est d'une grande importance pour la solubilité des protéines halophiles.

Sel	N_1 (mol/mol)	N_3 (mol/mol)
NaCl	2040	55
CsCl	1900	-20
$(NH_4)_2SO_4$	2100	0
$MgCl_2$	4160	85

Tableau 1.3 : Interactions protéine – solvant de la *Hm* MalDH dans divers sels [d'après Ebel, en préparation]. Dans le cadre du modèle de la particule invariante, une mole de protéine est associée à N_1 moles d'eau et N_3 moles de sel. La particule est définie comme la protéine accompagnée de ses contre-ions.

[*] L'activité résiduelle est l'activité de l'enzyme dans des conditions standard après 24 heures d'incubation dans le solvant d'intérêt.

1.1.6. Les interactions protéine – protéine de la Hm *MalDH*

Les interactions protéine – protéine de la *Hm* MalDH ont été abordées par des expériences de diffusion de neutrons aux petits angles et caractérisées par la détermination du second coefficient du viriel A_2 dans divers solvants salins [Ebel *et al.*, 1999b]. Différentes selon la nature et la concentration du sel, elles sont globalement répulsives en NaCl, en $NaCH_3CO_2$ et en KF ; globalement répulsives en $(NH_4)_2SO_4$ entre 1 et 2 M puis neutres à 3 M ; neutres en dessous de 0.5 M de $MgCl_2$ mais répulsives à plus fortes salinités ; et finalement globalement attractives en NaCl 1 M, méthyl-2-pentanediol-2,4 (MPD) 30%. Cette dernière condition de solvant est d'ailleurs proche des conditions qui conduisent à la cristallisation de la *Hm* MalDH.

Les déterminants de ces interactions variées et notamment de l'attraction à bas $MgCl_2$ ou haut $(NH_4)_2SO_4$ n'ont pas été élucidés. Au vu des interactions particulières de la *Hm* MalDH avec le solvant, ce système apparaît être un bon modèle pour étudier les relations entre les interactions interparticulaires et les interactions protéine – solvant. L'objectif général de cette étude va être de préciser les caractéristiques des interactions faibles protéine – protéine en solution pour la *Hm* MalDH et de les interpréter en considérant les effets du solvant.

Avant de préciser notre démarche, il nous a paru utile de rappeler quelques notions sur les différentes interactions interparticulaires en commençant par les plus élémentaires, pour terminer par les interactions entre grosses particules et les effets induits par les interactions protéine – solvant.

1.2. Les interactions intermoléculaires en solutions *

Depuis la vision poétique et agréablement simple des Grecs d'un monde régi par deux forces : l'amour et la haine, notre compréhension et notre formalisme des phénomènes physiques se sont perfectionnés. D'une manière très générale, dans la physique moderne, les

interactions régissant la matière sont classées en quatre forces fondamentales. Les forces nucléaires faible et forte agissent sur une très courte portée (moins de 10^{-4} Å), entre les neutrons, les protons, les électrons et les particules élémentaires. Elles appartiennent aux domaines de la physique des particules et des hautes énergies et ne nous intéresseront pas ici. Les forces électromagnétique et gravitationnelle ont, quant à elles, une très large gamme de portées d'action : des distances subatomiques à presque l'infini. Elles agissent en plus entre les atomes et les molécules. Les forces intermoléculaires en solution qui nous intéressent font partie des forces électromagnétiques. L'aspect magnétique sera négligé car, pour les systèmes considérés, il reste toujours très faible devant les autres interactions.

On distinguera les interactions répulsives de volume exclu dont l'origine fine est quantique, et les interactions électrostatiques et de polarisation qui proviennent de la nature électrostatique des molécules. Toutes les forces intermoléculaires en solution correspondent à des combinaisons de ces types élémentaires d'interactions. C'est pourquoi nous allons les présenter succinctement, après quelques rappels de notions de base et de terminologie. Cependant, au vu de la complexité des interactions dans les solutions réelles contenant des milliards d'atomes et de molécules, des forces plus complexes intégrant plusieurs de ces interactions élémentaires sont utilisées en pratiques. Elles seront présentées au vu de leur particularité et de leur importance dans les interactions protéine – protéine en solution.

Quelques notions de base et un peu de terminologie

Les interactions intermoléculaires (actions réciproques qu'exercent entre elles deux molécules) sont souvent exprimées en termes de force (**F**), d'énergie potentielle ou de potentiel d'interaction (w). La force (conservative) dérive du potentiel : $\mathbf{F}(\mathbf{r}) = -\mathbf{grad}\ w(\mathbf{r})$ [†]. Elle est dite attractive (réciproquement, répulsive), avec $F(\mathbf{r})$ négatif (positif), lorsqu'elle tend à rapprocher (éloigner) les molécules. Qualifiée de forte ou de faible selon son intensité au contact des particules (par rapport à kT), elle est de faible ou longue portée : l'énergie potentielle correspondante devient comparable à la valeur de référence kT (de l'ordre de l'énergie d'agitation thermique) à une distance r courte ou grande.

Le potentiel d'interaction est appelé potentiel de paires lorsque deux molécules sont considérées, ou potentiel de champ moyen dans le cas général à N corps. Il peut être

[*] Les notions abordées dans ce chapitre sont présentées dans " Liaisons intermoléculaires " de Gerschel [Gerschel, 1995] et " Intermolecular and surface forces " d'Israelachvili [Israelachvili, 1991].

[†] $\mathbf{grad}\ w(\mathbf{r}) = (\partial w/\partial x, \partial w/\partial y, \partial w/\partial z)$, r est la distance interparticulaire.

considéré comme une énergie libre. Nous supposerons par la suite qu'il n'existe pas de direction privilégiée pour les interactions. Le potentiel sera donc à symétrie sphérique et ne dépendra que de la distance entre les molécules : $r = |\mathbf{r}|$. Dans un modèle simple, les particules seront considérées comme des sphères.

1.2.1. Les répulsions de volume exclu

Les forces répulsives de volume exclu, aussi appelées répulsions stériques ou de cœur dur, traduisent l'impossibilité pour les nuages électroniques de deux molécules de s'interpénétrer. Leur interprétation fait appel à la physique quantique (principe de Pauli). Ces répulsions se caractérisent par une très courte portée et une intensité qui augmente très brusquement lorsque deux molécules s'approchent. Elles sont décrites par des formes de potentiels empiriques. Pour des macromolécules globulaires, le potentiel de sphère dure, traduit l'indéformabilité des molécules :

(1.1) $$w_1(r) = \begin{cases} +\infty & \text{pour } r \le D \\ 0 & \text{pour } r > D \end{cases} ,$$

où D est le diamètre de la sphère dure (Figure 1.3).

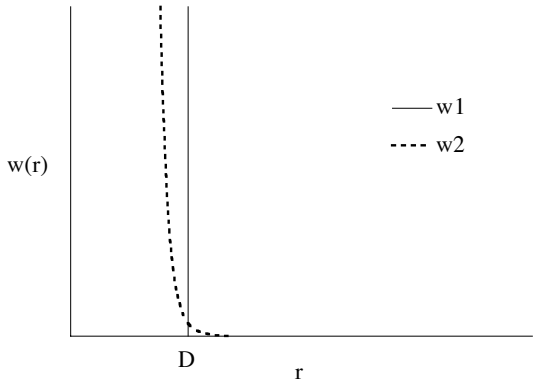


Figure 1.3 : Potentiel de sphère dure (w1) et de loi en puissance (w2) avec n = 12.

Les autres formes de potentiel, moins abruptes, intègrent la compressibilité de la molécule :

(1.2) $$w_2(r) = (D/r)^n \qquad w_3(r) = J e^{-r/D_0} ,$$

où n est un entier généralement entre 9 et 16, J et D_0 des paramètres ajustables. Pour $n = +\infty$, on retrouve le potentiel de sphère dure et pour n = 12, la partie répulsive du potentiel de Lennard-Jones ($w(r) = J_r r^{-12} - J_a r^{-6}$).

1.2.2. Les interactions purement électrostatiques

Les interactions purement électrostatiques ont pour origine la force de Coulomb entre les charges. Ce sont les interactions entre les charges et les dipôles électriques permanents. Elles font intervenir la charge Z (en unité de charge e de l'électron) et le moment dipolaire u de la particule dont le diamètre est D ainsi que la constante diélectrique ε^* et la force ionique $I = 0.5 \sum c_i Z_i^2$ (où c_i et Z_i sont respectivement la concentration et la charge de l'ion i du sel) qui caractérisent le milieu (le solvant) considéré comme continu.

L'interaction coulombienne (charge/charge)

Le potentiel électrostatique entre deux charges Z_1 et Z_2 distantes de r (potentiel de Coulomb), dans un solvant considéré comme un continuum de constante diélectrique ε, est donné par :

(1.3) $$w(r) = Z_1 Z_2 e^2 / 4\pi\varepsilon_0 \varepsilon r \ .$$

Pour des charges de même signe (respectivement de signes opposés), l'énergie d'interaction est positive (négative) et la force répulsive (respectivement attractive).

Cette interaction est la plus forte des interactions entre particules que nous considérerons ici. Elle sera de fait la seule à être explicitement prise en compte pour la caractérisation des interactions entre macromolécules en solution. L'énergie coulombienne de deux ions isolés au contact est comparable à l'énergie d'une liaison covalente. Pour des particules dans le vide, elle a une très longue portée. Toutefois, comme les macro-ions ont en général pour voisins des contre-ions dans une solution, le potentiel est moins intense et sa portée est beaucoup plus faible : il est écranté. La notion d'écrantage a été introduite par Debye et Hückel et découle de l'électroneutralité de la solution. Elle permet de ne plus considérer explicitement les ions du solvant : ils font partie du continuum qu'est le solvant et

[*] La constante diélectrique pondère la permittivité du vide ε_0 pour tenir compte du milieu.

n'interviennent dans l'expression du potentiel interparticulaire que par l'intermédiaire de la fonction d'écrantage ξ(r) [Debye et Hückel, 1923] :

$$\xi(r) = \frac{e^{-\kappa(r-D)}}{(1+\kappa D/2)^2} \qquad (1.4)$$

$$\kappa = \lambda_D^{-1} = \left(4\pi L_B N_A \sum_i c_i Z_i^2\right)^{1/2} = \left(8\pi L_B N_A I\right)^{1/2} .$$

κ est le terme d'écrantage, inverse de la longueur de Debye λ_D (Å), il est lié à la force ionique de la solution I (en M) et à la longueur de Bjerrum L_B. Deux ions monovalents séparés de L_B ressentent une énergie potentielle de kT (ou -kT). L_B vaut $e^2/4\pi\varepsilon_0\varepsilon kT$, soit 7.2 Å à 25°C dans l'eau.

Pour une particule de charge effective Z – tenant compte de la condensation électrostatique à la surface de la particule d'une partie des contre-ions (Z est inférieure à la charge structurale) – le potentiel coulombien écranté s'obtient en utilisant la théorie DLVO (du nom de ses auteurs) [Derjaguin et Landau, 1941; Verwey et Overbeek, 1948] :

$$w(r) = \frac{kT Z^2 L_B}{r} \frac{e^{-\kappa(r-D)}}{(1+\kappa D/2)^2} \qquad r > D . \qquad (1.5)$$

L'interaction charge/dipôle permanent

Les interactions charge/dipôle* permanent dépendent fortement de l'orientation du dipôle **u** par rapport à la charge Z. L'orientation peut être fixe ou le dipôle libre de tourner. Dans un solvant de forte constante diélectrique (ε_{eau} = 80, à 20°C), l'interaction charge/dipôle permanent fixe est généralement négligeable, *i.e.* inférieure à l'agitation thermique (kT). Les dipôles sont donc ordinairement plus ou moins libres de tourner. La valeur moyenne sur les orientations du potentiel d'interaction, calculée avec la distribution de Boltzmann, est :

$$w(r) = -\frac{(Ze)^2 u^2}{6(4\pi\varepsilon_0\varepsilon)^2 kT r^4} \qquad \text{pour} \quad Zeu/4\pi\varepsilon_0\varepsilon r^2 < kT . \qquad (1.6)$$

* Un dipôle électrique (permanent) est un ensemble de deux charges électriques ponctuelles, de même valeur absolue z et de signes opposés, caractérisé par un moment dipolaire **u** = z**l**, où **l** est le vecteur orienté de la charge négative vers la charge positive. Un ensemble de dipôles $\mathbf{u}_i$ d'une molécule par exemple peut être considéré comme un seul dipôle caractérisé par un moment dipolaire $\sum\mathbf{u}_i$. Les molécules dont le moment dipolaire n'est pas nul sont dites polaire. Par la suite, " dipôle " signifiera " dipôle permanent ".

L'interaction dipôle/dipôle

L'interaction dipôle fixe/dipôle fixe est encore plus faible que la précédente (dans le cas d'interaction entre macromolécules) et n'induit pas d'alignement des molécules polaires. Les dipôles sont donc libres de tourner, et une moyenne de Boltzmann sur toutes les orientations conduit au potentiel d'interaction suivant :

$$w(r) = -\frac{u_1^2 u_2^2}{3(4\pi\varepsilon_0\varepsilon)^2 kT r^6} \qquad \text{pour} \quad u_1 u_2 / 4\pi\varepsilon_0\varepsilon r^3 < kT \ . \tag{1.7}$$

Ce potentiel varie en $1/r^6$ est appelé potentiel d'orientation ou de Keesom. C'est l'une des trois interactions qui constituent l'interaction de van der Waals entre les atomes et les molécules (cf. §1.2.5 p.32).

1.2.3. Les interactions de polarisation

Ce sont des interactions électrostatiques qui font intervenir la polarisation des atomes ou des molécules. Les champs électriques $\mathbf{E}^*$ émanant des charges et des dipôles de l'environnement créent des moments dipolaires induits $\mathbf{u}_{ind}$ qui sont proportionnels à la polarisabilité ($\alpha > 0$) de la particule :

$$\mathbf{u}_{ind} = \alpha \mathbf{E} \ . \tag{1.8}$$

La polarisabilité a deux origines. La première est le déplacement, sous l'action d'un champ électrique externe **E**, du nuage électronique des atomes par rapport à leur noyau chargé positivement : on parle de polarisabilité électrique α_0. La seconde est l'orientation préférentielle (selon la direction de **E**) qu'acquiert un dipôle permanent u, libre de tourner. On parle de polarisabilité orientationnelle α_{orient} qui vaut :

$$\alpha_{orient} = u^2 / 3kT \ . \tag{1.9}$$

Tout atome (ou molécule), chargé ou non, est polarisable. Les molécules dites non-polaires ont en réalité un moment dipolaire nul uniquement en moyenne dans le temps.

---

* Le champ électrique est le champ de force associé à une charge électrique unité. Une charge z dans un champ électrique **E** subit une force **F** = z**E**.

L'interaction charge/dipôle induit

L'énergie potentielle d'interaction entre une charge Z et une molécule non-polaire (dipôle induit) de polarisabilité α est :

$$w(r) = -\frac{1}{2}\alpha\left(\frac{Ze}{4\pi\varepsilon_0\varepsilon r^2}\right)^2 . \tag{1.10}$$

L'interaction de Debye (dipôle/dipôle induit)

L'interaction entre un dipôle permanent **u** et une molécule non-polaire n'est pas suffisante pour aligner mutuellement les dipôles : ils sont donc libres de tourner. Le potentiel est donc moyenné sur les orientations :

$$w(r) = -u^2\alpha_0 \Big/ \left(4\pi\varepsilon_0\varepsilon\right)^2 r^6 . \tag{1.11}$$

Cette interaction est souvent appelée interaction de Debye ou d'induction. C'est la seconde contribution aux interactions de van der Waals (cf. §1.2.5 p.32).

Les forces de dispersion

La moyenne dans le temps du moment dipolaire d'une molécule non-polaire est nul, bien qu'il existe un moment dipolaire fini à tout instant, du fait de la position instantanée des électrons autour des protons. Ce dipôle instantané engendre un champ électrique qui polarise les atomes voisins, et induit ainsi des moments dipolaires. Le potentiel correspondant à été déterminé par London en 1930 à l'aide de la théorie des perturbations en mécanique quantique. Pour deux atomes différents, d'énergies de première ionisation* E_{i1} et E_{i2}, le potentiel s'écrit :

$$w(r) = -\frac{3}{2}\frac{\alpha_{01}\alpha_{02}}{\left(4\pi\varepsilon_0\varepsilon\right)^2 r^6}\frac{E_{i1}E_{i2}}{\left(E_{i1}+E_{i2}\right)} . \tag{1.12}$$

Cette interaction est aussi appelée force de London, et constitue la dernière contribution aux forces de van der Waals (cf. §1.2.5 p.32). Elle agit entre tous les atomes ou molécules, même totalement neutres et apolaires. L'interaction entre deux particules est perturbée par les autres

* Energie qu'il faut fournir à un atome pour en extraire un électron.

particules environnantes : le champ électrique émanant d'une particule atteint directement une seconde particule et aussi par " réflexion " sur toutes les autres car elles sont aussi polarisées. On ne peut donc pas simplement sommer ces potentiels de paires pour obtenir l'interaction globale. Ils sont donc non additifs.

1.2.4. Les effets généraux du solvant sur les interactions

Les forces considérées dans cette thèse agissent, en solution, entre des macromolécules identiques qui pourront par la suite être appelées particules ou colloïdes. Les différentes molécules du solvant modifient certaines des interactions précédentes. Quelques effets généraux du solvant sur ces interactions sont soulignés ci-dessous et d'autre plus spécifiques seront abordés dans les paragraphes suivants.

Pour des colloïdes dans un solvant, les interactions colloïde – colloïde sont perturbées par les interactions colloïde – solvant et solvant – solvant. Le colloïde (en plus de son volume) perturbe l'organisation locale du solvant : dans le cadre d'une description du solvant comme un continuum, des nouvelles forces dites de solvatation sont alors à prendre en compte (cf. §1.2.6 & 1.2.7 p.33). Réciproquement, le solvant modifie certaines propriétés du colloïde, comme sa charge (pH, écrantage) ou son moment dipolaire. Une interaction interparticulaire attractive dans le vide ou dans un gaz peut ainsi devenir répulsive dans le solvant.

En ce qui concerne les interactions colloïde – colloïde en solution, deux effet sont importants à noter. D'une part, le solvant réduit la plupart des énergies d'interaction par un facteur ε (par exemple dans l'interaction charge/charge (1.3)) ou ε^2 (dans l'interaction dipôle/dipôle (1.7)) qui n'est pas négligeable : 80 ou 6400 pour l'eau à 20°C et 40 ou 1600 pour une solution aqueuse de NaCl 4 M (cf. annexe 8.1 p.187). D'autre part, comme le colloïde est dans un solvant, il est nécessaire de remplacer la polarisabilité du colloïde par une polarisabilité effective α_{eff}, différence entre celle du colloïde et celle du solvant. Les forces de polarisation (attractives dans le vide) peuvent donc être attractives, nulles voire répulsives (pour des colloïdes différents) selon la polarisabilité des colloïdes et du solvant.

1.2.5. Les forces de van der Waals

Les forces de van der Waals regroupent les forces d'orientation (dipôle/dipôle), d'induction (dipôle/dipôle induit) et de dispersion (dipôle induit/dipôle induit), toutes variant en $1/r^6$. Le potentiel de paires de van der Waals peut s'écrire :

$$w_{vdw}(r) = -J_{vdw}/r^6 = -\left(J_{orient} + J_{ind} + J_{disp}\right)/r^6 \quad , \tag{1.13}$$

et peut être développé en utilisant les équations (1.7), (1.11) et (1.12). Ce potentiel de paires traduit l'interaction entre deux " unités " polaires ou polarisables. Différentes approches permettent d'estimer les forces de van der Waals entre deux particules constituées de nombreuses " unités " polaires ou polarisables.

La constante de Hamaker

Pour des " unités " polaires ou polarisables très diluées, dans les limites de polarisabilités α tendant vers 0 et ε vers 1, les potentiels de paires (1.13) peuvent être considérés comme étant additifs deux à deux. Dans cette hypothèse d'additivité, Hamaker décrit l'interaction entre deux particules sphériques identiques (composées de nombreuses unités polaires), distantes de r par :

$$w(r) = -A\,f(r) \quad , \tag{1.14}$$

où f(r) traduit la taille et la forme des particules, et A (en J ou en unité de kT) est la constante de Hamaker [Hamaker, 1937]. Elle est liée à la densité d'atomes ρ et à J_{vdw} de l'équation (1.13) par $A = \pi^2\rho^2 J_{vdw}$.

La théorie de Lifshitz

A cause des forces de dispersion, les forces de van der Waals ne sont, en général, pas additives. La théorie de Lifshitz traite les particules comme des milieux continus décrits par leur constante diélectrique statique et leur indice de réfraction. Elle permet le calcul de la constante d'Hamaker sans ce problème de la non-additivité [Lifshitz, 1956]. Comme la polarisabilité (et donc l'excès de polarisabilité) dépend des fréquences d'absorption des atomes, la constante d'Hamaker comprend deux termes : le premier, à fréquence nulle, dépend des constantes diélectriques (ε) et résume les interactions de Keesom et de Debye. Le second terme correspond aux interactions de dispersion qui dépendent de la fréquence via les indices

de réfraction (n). Pour deux macromolécules (p) identiques interagissant dans un solvant (s), la constante d'Hamaker calculée selon la théorie de Lifshitz s'écrit :

$$A = A_{\nu=0} + A_{\nu>0} = \frac{3}{4}kT\left(\frac{\varepsilon_p - \varepsilon_s}{\varepsilon_p + \varepsilon_s}\right)^2 + \frac{3E_i}{16\sqrt{2}}\frac{\left(n_p^2 - n_s^2\right)^2}{\left(n_p^2 + n_s^2\right)^{3/2}} \quad . \tag{1.15}$$

Cette constante d'Hamaker – Lifshitz permet de formuler, avec l'équation (1.14), le potentiel de paires des interactions de van der Waals pour la macromolécule.

Le potentiel d'ionisation (E_i = 12 eV), l'indice de réfraction (n_p = 1.519) et la constante diélectrique (ε_p = 2) sont, en première approximation, constants d'une protéine à l'autre [Malfois *et al.*, 1996]. Pour une protéine dans l'eau à 20°C (ε_s = 80, n_s = 1.33), on calcule pour A une valeur de 2.9 kT, avec $A_{\nu=0}$ = 0.68 kT et $A_{\nu>0}$ = 2.22 kT. On constate que les forces de dispersion ($A_{\nu>0}$) prédominent généralement les forces d'orientation et d'induction ($A_{\nu=0}$, toujours inférieur à 3kT/4).

Les forces de van der Waals existent pour toute particule, même totalement neutre. Elles sont toujours attractives (A > 0) pour des particules identiques. Mais elles peuvent être attractives ou répulsives pour deux particules différentes dans un milieu, en fonction des valeurs relatives des ε et n.

1.2.6. L'eau : des interactions particulières

L'eau a été extrêmement étudiée à cause de son omniprésence dans le monde du vivant. Ce solvant s'avère très particulier du fait de ses propriétés physico-chimiques. Il s'agit d'un liquide " structuré " en un réseau connectif en régénération continuelle via des liaisons hydrogènes. Ces dernières sont avec les interactions élémentaires présentées jusqu'ici, à l'origine de deux forces intermoléculaires (les effets hydrophobe et hydrophile) via le phénomène d'hydratation.

La liaison hydrogène est une liaison entre un groupement donneur de proton (SH, OH, NH) et des atomes électronégatifs (O, N, …). Elle est environ 20 fois plus faible qu'une liaison covalente, mais plus stabilisatrice qu'une liaison de van der Waals. Elle met en jeu cinq interactions élémentaires : les forces électrostatiques qui sont dominantes, attractives et fortement directionnelles ; une répulsion de volume exclu ; les interactions de polarisation ; un effet de transfert de charge (*i.e.* une délocalisation intermoléculaire d'un électron, de

manière directionnelle), et les forces de dispersion. Les caractéristiques des molécules impliquées dans une liaison hydrogène sont modifiées. Leur polarité augmente et les longueurs des liaisons intramoléculaires sont plus grandes alors que les distances intermoléculaires sont notablement réduites. Dans l'eau à 25°C, les distances intermoléculaires O–O et O–H sont respectivement de 2.8 et 1.85 Å. Ce processus est directif et coopératif. Les liaisons hydrogènes n'existent que dans des cas limités, mais qui peuvent prendre une très grande importance pratique. C'est le cas pour les macromolécules biologiques (protéines et acides nucléiques) et pour l'eau qui a une coordination tétragonale.

L'hydratation

L'hydratation consiste en une réorganisation des molécules d'eau à proximité de la surface d'une particule dissoute. On parle alors d'eau superficielle ou vicinale par opposition à l'eau de masse (du solvant). Deux sortes d'hydratations : ionique et hydrophobe se distinguent selon l'affinité de la surface avec l'eau.

Dans le cas de l'hydratation ionique, les molécules d'eau s'orientent à proximité des ions sous l'influence des forces coulombiennes, de polarisation et des liaisons hydrogènes. L'environnement local autour des ions se resserre (électrostriction). Pour les surfaces hydrophobes, l'interaction directe avec l'eau est peu importante. Les molécules d'eau ne pouvant développer des liaisons hydrogènes avec les groupements de la surface, elles s'organisent différemment, de manière à optimiser les liaisons hydrogènes entre elles. A ces deux types d'organisation sont respectivement associés les effets hydrophile et hydrophobe.

La réorganisation de l'eau à la surface des macromolécules peut s'observer expérimentalement : par exemple, l'analyse d'expérience de diffusion de rayons X et de neutrons s'accorde avec l'hypothèse d'une première couche d'hydratation 10% plus dense que le solvant de masse pour trois protéines de *E. coli* [Svergun *et al.*, 1998] et pour la *Hm* MalDH [Richard, 1998].

Les effets hydrophile et hydrophobe

Les phénomènes d'hydratation se traduisent par des interactions interparticulaires apparentes : les effets hydrophile et hydrophobe. Ils peuvent être exprimés par une loi en $J \exp\left(-r/\lambda_c\right)$, où la longueur de décroissance caractéristique λ_c est peu sensible à la nature de la surface, contrairement à J qui fixe l'intensité.

L'effet hydrophile apparaît comme une force interparticulaire répulsive ($J > 0$) qui est provoquée par la structure de l'eau au voisinage des particules (ou surfaces) hydrophiles. Pour

les protéines et l'ADN, les valeurs de λ_c sont comprises entre 3.5 et 4 Å [Gerschel, 1995, p. 256-262].

A l'inverse, l'effet hydrophobe apparaît comme une force attractive ($J < 0$, $\lambda_c \approx 10$ Å) entre des particules hydrophobes enveloppées dans leur couche d'hydratation. Elle est beaucoup plus intense et a une plus longue portée que toutes les interactions qui pourraient s'exercer sans la couche d'hydratation. L'effet hydrophobe émanerait de l'orientation préférentielle des molécules d'eau près de la surface hydrophobe qui se propage de proche en proche. Il s'agirait donc plus d'une interaction entre les structures de l'eau que d'une interaction soluté – soluté. Cet effet entropique provoque la coalescence des enveloppes d'hydratation et donc l'agrégation des particules hydrophobes. De nombreux processus naturels et vitaux découlent de cette interaction, comme le repliement des protéines.

1.2.7. Les effets du solvant et des co-solvants sur les interactions colloïde – colloïde

Le solvant des colloïdes est constitué d'un solvant principal (l'eau pour les protéines) auquel s'ajoutent souvent d'autres solutés : du sel ou d'autres espèces chimiques (méthyl-2-pentanediol-2,4 (MPD) dans notre cas). La présence de ces espèces peut induire des attractions ou des répulsions effectives entre protéines qui intègrent plusieurs des interactions " élémentaires " déjà présentées. Comme pour l'eau, les effets des solutés sur les interactions entre macromolécules peuvent résulter d'interactions colloïdes – soluté spécifiques. Par exemple, une attraction colloïde – solvant peut conduire à une répulsion effective colloïde – colloïde à courte portée et une attraction de pontage à distance intermédiaire. A l'opposé, une attraction solvant – solvant ou une répulsion colloïde – solvant ajoutera une composante attractive aux interactions colloïde – colloïde. Nous allons nous intéresser à deux aspects : la force de déplétion et les effets des sels.

La force attractive de déplétion

L'addition de sphères dures (le co-solvant) d'un diamètre D_s grand par rapport à celui du solvant principal (l'eau) à une suspension de sphères dures de diamètre D_p (les protéines) conduit à l'apparition d'une force attractive à courte portée, dite force de déplétion. En effet, les molécules de co-solvant ne pouvant pénétrer la surface des protéines, la sphère de diamètre D_p+D_s (volume V_1 ou V_2 de la Figure 1.4) leur est interdite. Lorsque deux protéines sont à une distance r inférieure à D_s, les molécules de co-solvant sont exclues de la zone de

recouvrement $V_1 \cap V_2$ ou zone de déplétion. La différence de pression osmotique entre le solvant de masse et cette région force sa réduction. Il en résulte une attraction effective entre les deux protéines :

(1.16) $$w_{depletion}(r) = -kTc_s \left(V_1 \cap V_2 \right) .$$

La portée de cette interaction est D_s et son intensité est proportionnelle à la concentration en co-solvant c_s et dépend des tailles des molécules (D_s, D_p par rapport aux molécules de solvant principal) via $V_1 \cap V_2$. Elle s'ajoute au potentiel de volume exclu des protéines, défini par l'équation (1.1). Les molécules de co-solvant sont ici considérées comme diluées et n'interagissant pas entre elles. Lorsque leur concentration augmente, elles interagissent entre elles via leur volume exclu. L'interaction de déplétion devient alors oscillante, alternativement attractive et répulsive : on parle de force de structure.

Pour des solutions de protéine, cette force de déplétion a été mise en évidence en présence de polymères (co-solvant) tels que les polyéthylènes glycols dont la taille est bien supérieure à celle des molécules du solvant [Arakawa et Timasheff, 1985a; Budayova *et al.*, 1999; Finet, 1999].

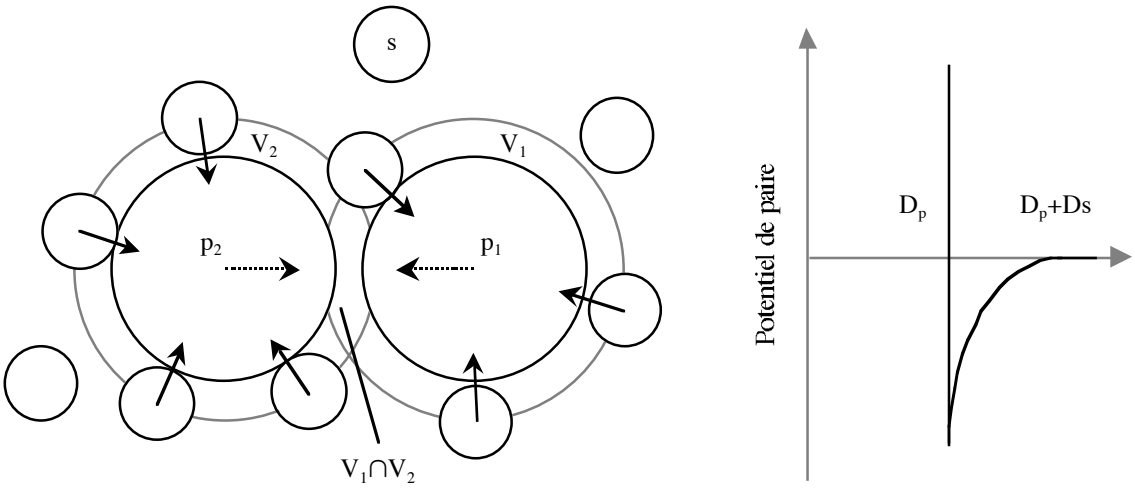


Figure 1.4: La force attractive de déplétion : mécanisme et potentiel de paires associé, pour un co-solvant (s) dilué.

Effets des co-solvants salins

La présence de sel dans une solution de protéine a divers effets sur les interactions protéine – protéine. Ils sont spécifiques (par rapport à la nature de la surface de la protéine) ou non, modifient certaines interactions et peuvent en provoquer d'autres. Nous allons brièvement en présenter quelques aspects.

Tout d'abord, la présence de sel dans le solvant écrante les interactions électrostatiques qui sont ainsi très diminuées à forte salinité. Par ailleurs, les sels modifient aussi la constante diélectrique ε du solvant qui pondère la plupart des forces entre macromolécules en solutions. Elle est par exemple deux fois plus faible en NaCl 4 M que dans l'eau pure (ε = 39 contre 78 à 25°C, cf. annexe 8.1 p.187).

L'augmentation de la tension de surface de l'eau par les sels induit un autre effet non-spécifique. Elle correspond à une accumulation relative de l'eau ou à un appauvrissement relatif du sel aux interfaces. De ce fait, son effet est très proche de la force de déplétion décrite ci-dessus et elle induit une attraction effective entre les macromolécules.

Parallèlement, la surface d'une protéine étant généralement chargée, la répartition des co-ions et contre-ions dans son voisinage diffère de celle du solvant de masse. La protéine est entourée d'une couche de solvatation formée à la fois d'eau (hydratation) et de sel. Sa composition peut varier selon la nature du sel, comme pour la *Hm* MalDH [Ebel *et al.*, 1998; Ebel, en préparation]. La charge et le diamètre effectifs de la protéine peuvent ainsi différer des valeurs structurales et les interactions protéine – protéine en être affectées.

Par ailleurs, selon leur densité de charge, les ions interagissent préférentiellement avec l'eau ou avec un autre ion. Plus un ion est petit et chargé, plus sa liaison avec l'eau sera forte. Ces ions (par exemple Li^+, Na^+) sont dits " kosmotropes ". A l'inverse, les gros ions peu chargés (K^+, Cs^+, Cl^-) fixent faiblement l'eau et sont qualifiés de " chaotropes ". La formation de paires d'ions kosmotrope – kosmotrope et chaotrope – chaotrope est plus favorable que l'hydratation de chaque ion [Collins, 1995]. La formation de telles paires à la surface d'une protéine lui est spécifique et peut modifier les interactions protéine – protéine [Riès-Kautt et Ducruix, 1991].

L'effet des sels sur les interactions protéine – protéine est généralement invoqué sous la dénomination " d'effet Hofmeister ". Hofmeister a classé les anions et les cations selon leur capacité à faire précipiter les protéines du blanc d'œuf de poule [Hofmeister, 1888]. L'effet

des anions qui est plus marqué s'ajoute à celui des cations selon l'ordre (non-exhaustif) suivant pour les protéines chargées négativement :

$$SO_4^- > CH_3CO_2^- > HCO_3^- > Cl^- > SCN^-$$
$$Li^+ > Na^+ \approx K^+ > NH4^+ > Mg^{2+} .$$

Pour les protéines de charge nette positive, la série anionique est inversée. Ces séries, aussi appelées séries lyotropiques, ont par la suite été étendues à d'autres sels, d'autres macromolécules biologiques et d'autres phénomènes (stabilité, solubilité, …) [von Hippel et Schleich, 1969; Collins et Washabaugh, 1985; Riès-Kautt et Ducruix, 1997; Ebel *et al.*, 1999a]. L'ordre relatif des ions semble être globalement respecté pour de nombreuses protéines. Les sels favorisant la stabilité, l'auto-association ou la précipitation (“ salting-out ”) se distinguent des sels dénaturants ou solubilisant (“ salting-in ”). Pour le lysozyme et des α et γ-cristallines dans certaines conditions de solvant, ces effets ont été modélisés par des potentiels attractifs aspécifiques de type Yukawa (cf. (2.23) p.64) ou van der Waals, mais dont la profondeur dépendait du sel et de sa concentration [Finet, 1999; Tardieu *et al.*, 1999].

La *Hm* MalDH est une protéine halophile qui interagit fortement avec les sels présents dans le solvant. Sa couche de solvatation comporte dans certains cas dix fois plus de sel que l'albumine du sérum bovin (BSA) qui est une protéine non-halophile [Bonneté *et al.*, 1993; Ebel *et al.*, 1999b; Ebel, en préparation]. De ce fait, le rôle des interactions protéine – sel dans les interactions protéine – protéine pourrait être accentué pour la *Hm* MalDH.

Finalement, toutes ces interactions sont couplées entre elles de manière complexe et non-additive. La démarche est de regarder les interactions protéine – protéine qui comprennent les effets des interactions protéine – solvant et solvant – solvant. Elles sont généralement définies pour deux colloïdes : on parle de potentiel de paires. Comment alors aborder le problème à N corps qui reflète la réalité des interactions protéine – protéine dans une solution ?

1.3. La mesure et la modélisation des interactions intermoléculaires en solution

Le passage des interactions entre deux particules (transcrites par les potentiels de paires) aux fonctions décrivant les propriétés mesurables d'équilibre et de structure d'une solution (problème à N corps) fait appel à la mécanique statistique des liquides. Différentes méthodes sont possibles. Les techniques de simulation numérique Monte-Carlo ou de dynamique moléculaire donnent des solutions exactes, *i.e.* correspondant au modèle microscopique de départ (w(r)), mais nécessitent d'importants moyens informatiques et des temps de calculs relativement longs. Elles restent, pour l'instant, adaptées à de petits systèmes. La méthode approximative la plus simple consiste en un développement de fonctions thermodynamiques (par exemple la pression osmotique) en puissances croissantes de la concentration en colloïde : on parle de développement du viriel. Généralement, seuls les deux ou trois premiers termes sont conservés. Cette technique analytique décrit correctement les propriétés des solutions à des concentrations modérées de particules, devient rapidement fausse lorsque la concentration augmente. Finalement, la méthode offrant un bon compromis entre précision et efficacité consiste à modéliser la répartition des macromolécules (et donc les facteurs de structure statiques qui sont mesurables) à l'aide des potentiels de paires. Dans cette étude, nous avons utilisé deux de ces approches que nous introduisons brièvement ici.

1.3.1. Le développement du viriel

De manière analogue à l'équation de van der Waals des gaz réels, la pression osmotique (Π) peut être exprimée selon un développement en puissances de la concentration en protéine. Comme les techniques abordées dans cette thèse (diffusion aux petits angles ou ultracentrifugation analytique) permettent d'obtenir une masse moléculaire apparente (M_{app}) reliée à la dérivée $(\partial\Pi/\partial c)_{T,\mu}$ à température T et aux potentiels chimiques μ des espèces diffusibles constants, le développement du viriel est présenté pour cette variable[*] :

[*] La masse moléculaire apparente déterminée par la pression osmotique (M_{app}^{PO}) est développée selon l'équation : $\Pi/(cN_A kT) = 1/M_{app}^{PO} = M^{-1} + A_2 c + A_3 c^2 + \cdots$, les coefficients A_i sont identiques à ceux de l'équation (1.17). Il faut aussi noter que la pression osmotique donne une moyenne en nombre de la masse moléculaire (M_n), alors que la diffusion aux petits angles ou la sédimentation une moyenne en masse (M_m). Pour des molécules monodisperses, $M_n = M_m$.

(1.17) $$\frac{1}{M_{app}} = \frac{1}{N_A kT}\left(\frac{\partial \Pi}{\partial c}\right)_{T,\mu} = \frac{1}{M} + 2A_2 c + 3A_3 c^2 + \cdots \ .$$

L'inverse de la masse moléculaire M^{-1} est le premier coefficient du viriel, A_2 le second (en $mol.l.g^{-2}$) et A_3 le troisième ; c est la concentration massique en protéine ($mg.ml^{-1}$) [Eisenberg, 1976, 1981]. Dans la limite d'une concentration protéique et d'interactions faibles, A_2 est relié au potentiel de paires W(r) par la relation suivante [McMillan et Mayer, 1945; Belloni, 1991] :

(1.18) $$A_2 = 2\pi \frac{N_A}{M^2} \int_0^\infty \left(1 - e^{-W(r)/kT}\right) r^2 dr \ .$$

Le second coefficient du viriel A_2 reflète donc l'ensemble des interactions interparticulaires. Il est positif pour des interactions globalement répulsives et négatif pour une résultante attractive.

Le coefficient A_2 peut être évalué grâce aux différentes techniques qui déterminent la masse molaire apparente de la protéine en solution [Eisenberg, 1976]. Il s'agit de la mesure directe de la pression osmotique, de la diffusion statique de la lumière, de la diffusion aux petits angles des rayons X ou des neutrons et de l'ultracentrifugation analytique. Nous avons principalement utilisé ces deux dernières techniques.

1.3.2. La modélisation des facteurs de structure

Les potentiels de paires (problème à 2 particules) décrits précédemment permettent de calculer, la fonction de distribution de paires g(r) (problème à N particules) grâce à l'équation d'Ornstein-Zernike et d'une équation intégrale [Hansen et McDonald, 1986; Belloni, 1991]. g(r) représente la probabilité de trouver une particule à une distance comprise entre r et r+dr de la particule de référence. Sa transformée de Fourier est le facteur de structure statique en solution S(q,c). Les facteurs de structure ainsi calculés peuvent être comparés à ceux obtenus expérimentalement par la diffusion aux petits angles des rayons X ou des neutrons par des solutions de protéines.

Différents modèles sont utilisés, correspondant à différents degrés de complexité de la description de la solution. Le modèle dit complet décrit toutes les espèces en présence : les colloïdes (chargés), les contre-ions, les co-ions et le solvant lui-même. Une première simplification consiste à faire une moyenne sur les molécules de solvant et de le considérer

comme un continuum diélectrique (modèle primitif). Le modèle le plus simple, dit à un composant, ne traite explicitement que les colloïdes. Nous avons utilisé ce dernier avec des potentiels à symétrie sphérique et relativement simples comparés à la richesse des interactions protéine – protéine. Cette démarche reste donc modeste.

1.4. Notre démarche pour la caractérisation et les implications des interactions protéine – protéine de la *Hm* MalDH

Dans quelle mesure les interactions protéine – solvant influencent-elles les interactions protéine – protéine ? Le présent travail tente de répondre à cette question en mettant en relation ces deux classes d'interactions pour la *Hm* MalDH. Cette enzyme halophile dont la surface est extrêmement chargée négativement est relativement bien caractérisée dans de nombreux solvants salés. Elle montre des solvatations différentes selon la nature du sel, mais qui sont très riches à la fois en eau et en sel par rapport aux protéines non-halophiles. La *Hm* MalDH semble donc être un excellent modèle pour étudier l'influence des interactions protéine – solvant sur les interactions protéine – protéine.

La mesure des interactions interparticulaires nous a permis dans un premier temps de les caractériser. Nous avons ensuite examiné et estimé dans quelle mesure elles sont modifiées ou déterminées par les interactions protéine – solvant. Nous avons pour cela utilisé différentes méthodes d'étude des solutions concentrées en protéine pour la *Hm* MalDH en divers solvants aqueux salins et mixtes sel – méthyl-2-pentanediol-2,4 (MPD).

Ce manuscrit s'articule autour de cinq parties principales. Le prochain chapitre présente la démarche expérimentale et les méthodes d'analyse connues que nous avons utilisées. Il expose dans un premier temps le protocole suivi pour obtenir les importantes quantités de *Hm* MalDH pure nécessaires à cette étude, puis les principes et l'analyse de la diffusion aux petits angles, principale technique utilisée pour les mesures des interactions, et finalement les analyses en termes de second coefficient du viriel A_2 et de modélisation des facteurs de structure.

Une nouvelle méthode de détermination de A_2 est présentée dans le chapitre 3 sous la forme d'un article [Solovyova *et al.*, 2001]. Nous verrons que la modélisation des profils de

vitesse de sédimentation dans une expérience d'ultracentrifugation analytique permet entre autres d'estimer A_2 dans des solvants complexes pour seulement 2 mg de protéine récupérables.

Le chapitre 4 est consacré à l'exploration des interactions protéine – protéine de la *Hm* MalDH dans différents solvants salins. L'évaluation globale de ces interactions par les A_2 et la modélisation des facteurs de structure grâce au potentiel DLVO nous a permis de mettre en évidence et d'estimer, au moins qualitativement, l'importance des interactions protéine – solvant dans les interactions interparticulaires.

Les interactions protéine – protéine permettent une meilleure compréhension des processus qui mènent à la cristallisation des macromolécules biologiques. La *Hm* MalDH cristallise dans une solution aqueuse de NaCl et de méthyl-2-pentanediol-2,4 (MPD) selon un protocole original de dilution [Richard *et al.*, 1995]. Grâce en particulier à l'étude des interactions protéine – protéine dans des solvants ternaires NaCl – MPD – H_2O (chapitre 5), nous avons décrit la cristallogenèse de la *Hm* MalDH qui est associée à un lent conditionnement de la protéine d'interactions interparticulaires répulsives vers des interactions modérément attractives [Costenaro *et al.*, 2001].

Au vu des résultats obtenus dans les mélanges NaCl – MPD – H_2O, il nous a semblé intéressant d'étendre l'étude de ce type de mélange à d'autres sels, principalement $MgCl_2$ et $(NH_4)_2SO_4$, afin de mieux comprendre les modifications des interactions induites par le MPD. Nous mettons en évidence dans le chapitre 6 que le MPD modifie non seulement la stabilité de la *Hm* MalDH, mais aussi ces interactions interparticulaires en ajoutant une composante attractive. Ces effets sont principalement dus à la répulsion du MPD par la surface chargée de la protéine.

Les perspectives ouvertes par ces études seront synthétisées en conclusion. Finalement, des informations complémentaires et le résumé des conditions des expériences de diffusion aux petits angles sont donnés en annexe.

2. DE LA PURIFICATION DE LA *HM* MALDH AUX INTERACTIONS PROTEINE – PROTEINE

2.1. De la cellule aux échantillons

2.1.1. Purification de la Hm MalDH

Une protéine peut être obtenue soit de l'organisme originel soit par l'expression de son gène dans un autre organisme, habituellement la bactérie *Escherichia coli*. La seconde méthode est pour la *Hm* MalDH plus avantageuse en temps et en rendement. Le protocole de surexpression et de purification décrit est celui de F. Cendrin [Cendrin *et al.*, 1993], légèrement modifié.

Matériel utilisé

Solutions :

Milieu LB-amp (Lubria et Broth) : 10 g de peptone de tryptone, 5 g d'extrait de levure, 10 g de NaCl, H_2O qsv 1 l (stérilisé par autoclave) + 200 mg d'ampicilline.
Tampon A : NaCl 0.2 M, Tris 50 mM pH 8, (NADH 0.2 mM).
Tampon B : NaCl 5 M, Tris 50 mM pH 8.
Tampon C : $(NH_4)_2SO_4$ 2 M, Tris 50 mM pH 8.
Tampon D : $(NH_4)_2SO_4$ 3.5 M, Tris 50 mM pH 8.
Tampon E : $(NH_4)_2SO_4$ 2.5 M, Tris 50 mM pH 8.
Tampon F : $(NH_4)_2SO_4$ 1 M, Tris 50 mM pH 8.
Tampon G : $(NH_4)_2SO_4$ 1.3 M, NaCl 2 M, Tris 50 mM pH 8.
Tampon H : NaCl 4 M, Na phosphate 20 mM pH 7.
Tampon I : NaCl 4 M, Na phosphate 300 mM pH 7.
Tampon J : NaCl 4 M, Tris 50 mM pH 8.

Les tampons sont filtrés et dégazés.

Produits :

Colonne Superose® 12 et gels Sepharose® 4B, Sephacryl® S-300 de Pharmacia Biotech.
Gel DEAE cellulose DE52
Gel hydroxyapatite Bio-Gel® HT gel de BIO-RAD Laboratories
Sacs à dialyse Spectra/Por® 4 (12-14 kDa) de Spectrum®.
Centricon® 30, Microcon® 30, membrane YM30 DIAFLO® d'Amicon®.

Surexpression de la *Hm* MalDH chez *E. coli*

La souche bactérienne utilisée est : *E. coli* – HMS 174 (DE3) – pET 11a – *Hm* MalDH. Ces bactéries possèdent le gène de la *Hm* MalDH sous contrôle du promoteur T7, le gène de l'ARN polymérase T7 sous contrôle du promoteur lacUV5 et un gène de résistance à l'ampicilline. La culture de bactéries est réalisée dans du milieu LB-amp (*i.e.* contenant de l'ampicilline) qui permet de sélectionner les cellules surexprimant la *Hm* MalDH. Une colonie de bactéries prélevée sur une boîte de Pétri (LB-amp et Agar) permet d'ensemencer 30 à 100 ml de milieu LB-amp, qui après au minimum une nuit d'incubation à 37°C sous agitation (220 rpm) constitue une pré-culture. Celle-ci est alors diluée 100 fois dans du milieu LB-amp (typiquement 5-7 ml pour 0.5 l dans un erlenmeyer stérile de 2 l). Les bactéries sont cultivées à 37°C sous agitation (200 rpm) et leur multiplication est suivie par mesure de l'absorbance de la solution à 600 nm. Lorsque la densité optique atteint 0.6 (phase exponentielle précédant la phase stationnaire) après environ 2 à 3 heures d'incubation, la surexpression de la *Hm* MalDH est induite par ajout d'IPTG pour obtenir une concentration finale de 1 mM. L'IPTG, inducteur du promoteur lacUV5, déclenche la synthèse de l'ARN polymérase T7 qui est seule capable de transcrire le gène de la *Hm* MalDH dans ce système. La culture induite est alors incubée sous agitation pendant 2 heures, puis centrifugée pendant 15 minutes à 5000 rpm à 4°C. Les culots bactériens sont utilisés immédiatement pour la purification de la *Hm* MalDH ou congelés à -20°C.

Purification de la *Hm* MalDH

Principes

Les protéines halophiles se déplient à faible concentration en sel, leur purification se réalise donc dans des conditions particulières, limitées aux techniques de fractionnement utilisables dans des solutions ultra-salines. Après réactivation de la *Hm* MalDH par augmentation de la concentration en NaCl jusqu'à 4 M, le lysat cellulaire est purifié par centrifugation, puis par précipitation différentielle des protéines de *E. coli* dans du sulfate d'ammonium 2 M puis 2.5 M. Ces trois étapes permettent d'éliminer les membranes et la plupart des protéines de *E. coli*, tout en évitant la précipitation partielle de la *Hm* MalDH et la cristallisation du sodium due à une dialyse directe contre du sulfate d'ammonium 2.5 M. L'échantillon est ensuite déposé sur une colonne de Sepharose® 4B qui, à forte concentration en sulfate d'ammonium, fixe les protéines par précipitation non dénaturante sur ce support (chromatographie d'adsorption). Les protéines se fixent au gel par des liaisons hydrogènes,

des interactions de van der Waals ou hydrophobes. La baisse de la concentration en sulfate d'ammonium permet de resolubiliser de manière différentielle les protéines. Les protéines sont ainsi éluées séquentiellement par un gradient de concentration linéaire décroissant de sulfate d'ammonium. Après une étape de concentration qui peut être réalisée par une colonne de DEAE cellulose (chromatographie échangeuse d'ions), le NADH fixé sur la *Hm* MalDH et les éventuels acides nucléiques parasites sont éliminés par une chromatographie d'affinité sur hydroxyapatite. Une filtration sur gel (chromatographie d'exclusion) permet finalement d'éliminer les agrégats ou les impuretés résiduelles et de vérifier l'état d'oligomérisation de la protéine.

Etapes

Renaturation

Les bactéries resuspendues dans du tampon A (environ 5 ml par litre de culture et 0.1 mg de DNAse) sont lysées par sonication pendant 6 fois 30 secondes (microsonde à 50% d'activité). L'ajout de 4 volumes de tampon B ou de NaCl en poudre par petites quantités et sous agitation permet la renaturation de la *Hm* MalDH. La solution à 4 M NaCl est centrifugée à 20 000 rpm pendant 10 minutes à 4°C dans un rotor Beckman JA20.

Précipitation différentielle

Le surnageant est dialysé pendant au moins 6 heures contre du tampon C changé 2 fois. La solution est centrifugée à 20 000 rpm pendant 10 minutes à 4°C dans un rotor Beckman JA20. Du tampon D ou du sulfate d'ammonium en poudre est ajouté pour augmenter la concentration de 2 M à 2.5 M. La solution est de nouveau centrifugée durant 5 minutes dans les conditions précédentes.

Chromatographie sur Sepharose® 4B

Le surnageant est déposé sur une colonne de Sepharose® 4B (volume de la colonne ~130 ml pour 200 mg de protéine) équilibrée avec du tampon E, à un débit de 60 ml/h. Elle est lavée avec le même tampon (un volume de la colonne) puis les protéines sont éluées avec un gradient négatif de sulfate d'ammonium 2.5 M (1 l de tampon E) à 1 M (1 l de tampon F). Les fractions suffisamment actives (cf. test d'activité) sont rassemblées.

Concentration

La solution pure de *Hm* MalDH est concentrée soit mécaniquement avec des Centricon® 30 ou sur une membrane YM 30, soit par une colonne de DEAE cellulose DE52.

Celle-ci est équilibrée avec du tampon C à un débit de 40 ml/h, puis éluée avec du tampon G ou J.

Chromatographie d'affinité sur hydroxyapatite (si présence de NADH)

Si la solution de *Hm* MalDH contient encore du NADH (pic à 340 nm, $DO_{280}/DO_{260} \neq 2$), elle est dialysée contre du tampon H changé 2 fois afin d'abaisser la concentration en sulfate d'ammonium en dessous de 20 mM. Elle est ensuite déposée sur une colonne d'hydroxyapatite équilibrée dans le même tampon (volume de la colonne : environ 22 ml pour 170 mg de protéine) à un débit de 12 ml/h maximum. La colonne est lavée pour ôter tout le NADH, puis éluée avec du tampon I. Les fractions actives et sans NADH sont rassemblées. Cette chromatographie est entièrement réalisée à température ambiante à cause de la faible solubilité du phosphate de sodium (tampon I).

Chromatographie d'exclusion

Après une dialyse contre du tampon J et une concentration par Microcon® 30 ou Centricon® 30, selon le volume, la solution de *Hm* MalDH est déposée par injection de 200 µl (quantité maximum 20 mg) sur une colonne Superose® 12 équilibrée avec du tampon J au débit de 18 ml/h, ou par injection de 1 ml (40 mg maximum) sur une colonne S 300 (diamètre de 1 cm, hauteur de 100 cm). Les fractions contenant la *Hm* MalDH (DO à 280 nm) sont réunies, concentrées et stockées à 4°C.

Rendement

J'ai généralement effectué les purifications à partir de 20 l de culture bactérienne et obtenu de 7 à 9 mg de protéine par litre de culture, soit 140 à 180 mg.

Test d'activité enzymatique

Au cours des différentes étapes, la présence de la *Hm* MalDH est révélée par des tests d'activité enzymatique et des aliquotes permettent d'estimer le rendement progressif de la purification. La *Hm* MalDH est en effet une enzyme qui catalyse la réaction suivante :

$$\text{Oxalacétate} + \text{NADH} + H^+ \quad <\!\!-\!\!-\!\!-\!\!> \quad \text{Malate} + NAD^+ .$$

Le test d'activité enzymatique consiste à suivre la diminution de la densité optique à 340 nm pendant 15 secondes, traduisant l'oxydation du NADH en NAD car seul le premier présente un pic d'absorption à cette longueur d'onde. La réaction s'effectue à température ambiante, dans une cuve en quartz de trajet optique 1 cm. Quelques µl de solution à tester

sont dilués dans 1 ml de solution de réaction (NaCl 2 M, Tris 50 mM pH 8, NADH 0.2 mM, acide oxalacétique 0.3 mM) fraîchement préparée.

Mesure de concentration

La concentration (c en $mg.ml^{-1}$) en *Hm* MalDH d'une solution est déterminée par mesure son absorbance à 280 nm (DO_{280}), grâce à la loi de Beer-Lambert :

$$DO_{280} = \log_{10} I_0 / I_t = \varepsilon_{280}^{0.1\%} \, c \, l \, , \tag{2.1}$$

où I_0 (I_t) est l'intensité incidente (respectivement transmise), l le trajet optique du faisceau dans la solution (en cm) et $\varepsilon_{280}^{0.1\%}$ le coefficient d'extinction massique ($0.85 \; ml.mg^{-1}.cm^{-1}$) déterminé pour la masse de la protéine sans contre-ions (130552 Da) [Bonneté *et al.*, 1993]. En réalité, on mesure un coefficient d'extinction molaire qui est transformé en coefficient d'extinction massique pour obtenir la concentration massique c (en mg/ml). La linéarité de la loi précédente n'est valable que pour une absorbance inférieure à 1.

La mesure du spectre d'absorbance de 240 à 440 nm permet de vérifier simplement l'intégrité et l'homogénéité de la solution. En effet, une ligne de base non plate au-delà de 330 nm environ est souvent caractéristique de protéines dégradées et agrégées. De plus, un rapport DO_{280}/DO_{260} d'environ 2 indique l'absence de cofacteur (NADH).

2.1.2. Préparation des échantillons pour les études biophysiques

La protéine en NaCl 4 M, Tris 50 mM pH 8 (après filtration sur gel) est concentrée puis dialysée contre le solvant de mesure, en trois bains successifs de volume 50 fois supérieur. La première dialyse est généralement courte (environ une heure) et la dernière dure au moins 12 heures. Typiquement, 1 ml est dialysé contre trois fois 50 ml de solvant, dans des tubes agités par une lente rotation. La solution protéique peut ensuite être diluée par le solvant de dialyse pour obtenir des échantillons à différentes concentrations. La pesée de ces dilutions permet d'obtenir une plus grande précision sur les rapports des concentrations.

Les rayons X étant agressifs pour les protéines (contrairement aux neutrons), 2 mM de DTT sont ajoutés dans le dernier bain de dialyse pour les expériences de diffusion de rayons X. Le DTT prévient en effet la formation des radicaux libres responsables de la dégradation des protéines.

En présence de sel et de MPD, la *Hm* MalDH peut être peu soluble et précipiter (selon sa concentration) en quelques heures, de manière réversible. Dans ces cas, la protéine à forte concentration est dans un premier temps conditionnée dans le sel (où elle reste stable et soluble) par dialyse. Les échantillons sont ensuite réalisés par dilution de cette solution protéique avec des solutions tamponnées de sel, de MPD et d'eau.

2.1.3. Récupération des échantillons

Après chaque expérience, les différents échantillons sont rassemblés et dialysés contre du NaCl 4 M, Tris 50 mM pH 8 (solvant de conservation de la *Hm* MalDH), puis concentrés. Une chromatographie d'exclusion permet, en éliminant les agrégats ou impuretés possibles, de stocker une solution protéique pure pour des expériences futures.

2.2. La diffusion aux petits angles

La diffusion des rayonnements est une technique peu invasive qui permet d'étudier la structure à basse résolution des particules et les corrélations entre particules (et donc les interactions protéine – protéine). Nous avons utilisé deux rayonnements : les rayons X et les neutrons qui permettent d'explorer des distances d'une dizaine à quelques centaines d'angströms. Cette partie présente les aspects généraux de la diffusion aux petits angles en solution et certains paramètres structuraux que l'on peut déduire (communs aux deux rayonnements). Ces notions sont exposées dans de nombreux ouvrages [par exemple Lindner et Zemb, 1991]. Les aspects expérimentaux seront ensuite abordés.

2.2.1. Principes généraux de la diffusion aux petits angles

La diffusion aux petits angles (en réalité aux petits vecteurs de diffusion q, du même ordre de grandeur que l'inverse des distances r à mesurer) est un phénomène principalement élastique : les photons ou neutrons incidents et diffusés ont la même énergie, seule la direction de leur vecteur d'onde change. Les effets inélastiques sont en effet petits dans la zone angulaire étudiée. De plus, seule la diffusion cohérente est considérée, la diffusion

incohérente (*i.e.* sans rapport de phase), plane à tout angle, étant traitée comme un bruit de fond.

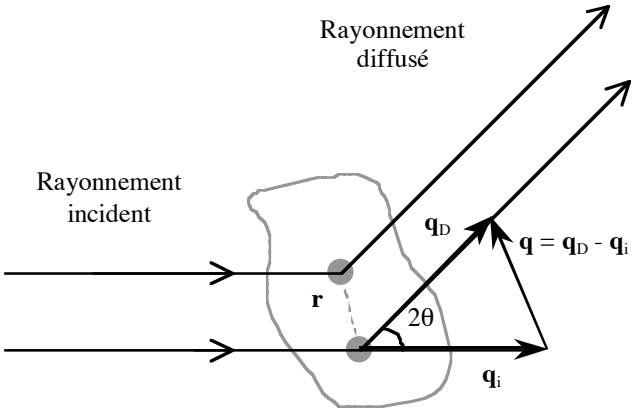


Figure 2.1 : Géométrie de la diffusion élastique : chaque diffuseur élémentaire d'une particule diffuse le rayonnement incident de vecteur d'onde $\mathbf{q}_i$ dans toutes les directions. Les diffusions selon la direction $\mathbf{q}_D$ de deux diffuseurs distants de $\mathbf{r}$ sont déphasées de $\mathbf{qr}$. Pour une diffusion élastique, $q_i = q_D = 2\pi/\lambda$ et $q = 4\pi\sin\theta/\lambda$, où λ est la longueur d'onde et 2θ l'angle de diffusion.

L'amplitude totale diffusée selon la direction $\mathbf{q}_D$ (Figure 2.1) est la somme des amplitudes diffusées par chaque atome (ou diffuseur élémentaire : électron pour les rayons X, noyau pour les neutrons). Comme les distances interatomiques sont petites par rapport aux distances examinées, cette somme discrète peut être remplacée par une intégrale en introduisant la densité de longueur de diffusion (ρ_n ou ρ_{el}[*]), où $\mathbf{q}$ est le vecteur de diffusion ($q = 4\pi\sin\theta/\lambda$) :

(2.2) $$A(\mathbf{q}) = \int \rho_{n,el}(\mathbf{r})\, e^{i\mathbf{qr}} d\mathbf{r} \ .$$

L'intensité diffusée est le produit de $A(\mathbf{q})$ par son complexe conjugué :

(2.3) $$\begin{aligned} I(\mathbf{q}) &= A(\mathbf{q})A^*(\mathbf{q}) = \int p(\mathbf{r})\, e^{i\mathbf{qr}} d\mathbf{r} \\ p(\mathbf{r}) &= \int \rho_{n,el}(\mathbf{r})\, \rho_{n,el}(\mathbf{r}' + \mathbf{r})\, d\mathbf{r}' \end{aligned} \ ,$$

[*] Pour les rayons X la densité de longueur de diffusion est la densité électronique pondérée par la longueur de diffusion de Thomson.

où p(**r**) est la fonction d'autocorrélation de la densité de longueur de diffusion. Cette corrélation à deux points mesure la persistance spatiale de la densité de longueur de diffusion (*i.e.* a-t-on la même densité en **r** et **r**'+**r** ?).

Dans le cas de particules dispersées dans un solvant (s) considéré comme homogène, on utilise généralement dans les équations (2.2) et (2.3) la différence de densité de longueur de diffusion entre les particules et le solvant : $\Delta\rho_{n,el}(\mathbf{r}) = \rho_{n,el}(\mathbf{r}) - \rho_s$. Le carré de cette valeur intervient dans l'intensité diffusée et est appelé le contraste de la particule.

2.2.2. Quelques paramètres structuraux

Solution idéale et monodisperse : facteur de forme, masse et rayon de giration de la particule

Une solution idéale est une solution de particules sans interaction : leur position et orientation ne sont pas corrélées. Le système est isotrope. Une telle solution est approchée par une dilution infinie des particules. L'intégrale de l'équation (2.3) ne porte donc que sur les diffuseurs élémentaires de chaque particule. Si de plus la solution est monodisperse, l'intensité totale diffusée est simplement le produit du nombre de particule N par la diffusion d'une particule $I_1(\mathbf{q})$ moyennée sur toutes les orientations qui est appelée facteur de forme P(q) :

(2.4) $$I(q) = N \langle I_1(\mathbf{q}) \rangle_{orientation} = N\, P(q) \ .$$

La courbe de diffusion I(q) dépend fortement de la forme et de la taille des particules. Toutefois, Guinier souligna qu'à très petits angles (q^{-1} beaucoup plus grand que la taille de la particule), la diffusion dépend de cette taille selon une loi générale indépendante de la forme et de la symétrie de la particule [Guinier et Fournet, 1955] :

(2.5) $$I(q) = I(0) \exp\left(-\frac{1}{3} R_g^2 q^2 \right) \ .$$

Pour des protéines globulaires, cette approximation est valable dans la gamme angulaire vérifiant $R_g q < 1.3$ (plus large que le domaine de Guinier $R_g q << 1$). L'intensité diffusée à q = 0 est proportionnelle à la masse moléculaire M de la protéine [Luzzati et Tardieu, 1980; Jacrot et Zaccai, 1981] :

(2.6) $$I(0) = \text{Cste}\left(\partial\rho_{n,el}/\partial c\right)_{\mu}^{2} \, c\, M \ ,$$

où c est la concentration en protéine (en mg/ml) et $\left(\partial\rho_{n,el}/\partial c\right) = \left(\rho_{n,el} - \rho_s\right)/c$ est l'incrément de densité de longueur de diffusion. Le carré du rayon de giration R_g est :

(2.7) $$R_g^2 = \frac{\int \Delta\rho_{n,el}(\mathbf{r})\, \mathbf{r}^2 d\mathbf{r}}{\int \Delta\rho_{n,el}(\mathbf{r})\, d\mathbf{r}} \ ,$$

où l'intégrale porte sur le volume de la protéine et **r** a pour origine le centre de masse de la protéine. Pour un volume donné, la particule ayant le plus petit R_g est la sphère de rayon $R = (5/3)^{1/2}R_g$.

Les deux paramètres I(0) et R_g peuvent être obtenus par une régression linéaire de ln[I(q)] en fonction de q^2. Il faut cependant bien noter que l'approximation de Guinier n'est valable que pour des particules monodisperses non-corrélées. Ces conditions sont réalisées par une solution à dilution infinie.

Particules homogènes en interaction : facteur de structure, masse et rayon de giration apparents

Pour une solution de particules en interaction, l'intensité diffusée reflète à la fois leur structure et les interactions entre les différentes particules : les corrélations de l'équation (2.3) sont à la fois intra et intermoléculaires. Pour N particules identiques de symétrie sphérique, l'intensité diffusée est le produit du facteur de forme P(q) par un terme d'interférence, le facteur de structure statique S(q,c) (c étant la concentration en particules) :

(2.8) $$I(q,c) = N\, P(q)\, S(q,c) = I(q,0)\, S(q,c) \ .$$

Pour des particules sans interactions (ou suffisamment diluées), S(q) = 1 et l'on retrouve l'équation (2.4). Le facteur de forme est donc proportionnel à l'intensité diffusée dans la limite idéale, pour c tendant vers 0. Le produit (2.8) dans l'espace réciproque correspond au produit de convolution dans l'espace réel d'une particule avec la distribution des centres des particules qui traduit les fluctuations de densité (en particule) de la solution. En introduisant la fonction de distribution de paires g(r) qui représente la probabilité de trouver deux particules séparées par la distance r, on relie S(q) à sa transformée de Fourier par l'équation :

(2.9) $$S(q) = 1 + \rho\int g(r)e^{iqr}dr = 1 + \rho\int\left(g(r) - 1\right)e^{iqr}dr + \rho(2\pi)^3\delta(q) \ ,$$

où ρ est la densité numérique de particule ($\rho = cN_A/M$).

Pour des protéines globulaires (*i.e.* dont aucune des dimensions ne diffère d'un ordre de grandeur des autres) et monodisperses, une solution concentrée est une solution de particules en interactions. Les deux équations précédentes restent utilisables, mais sont des approximations du fait de la non-sphéricité de ces protéines et de leurs interactions. Le calcul du facteur de structure expérimental est présenté par la suite (cf. p.65).

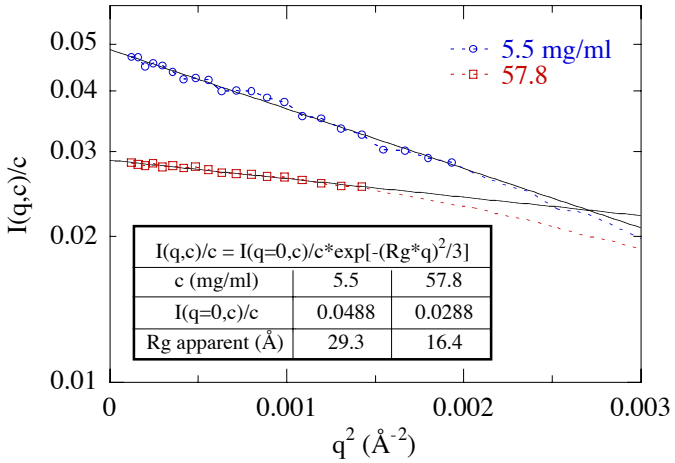


Figure 2.2 : Détermination par l'approximation de Guinier des valeurs apparentes des intensités diffusées à angle nul I(q=0,c)/c et des rayons de giration R_g pour des solutions de *Hm* MalDH à 5.5 et 57.8 mg/ml en $(NH_4)_2SO_4$ 1.5 M (réf. D22 00.10*). Les régressions linéaires de ln[I(q,c)/c)] portent sur la partie des intensités représentées par des ronds ou des carrés, qui correspondent à un domaine angulaire plus restreint que celui respectant la condition $R_g q < 1.3$.

L'approximation de Guinier (2.5) dans le cas de particules en interaction permet d'obtenir, pour la masse et le rayon de giration de la particule, des valeurs apparentes. Elles sont calculées par une régression linéaire de ln[I(q)] dans un domaine angulaire vérifiant $R_g q < 1.3$ et dépendent de la concentration en protéine c. La Figure 2.2 présente un exemple de cette régression linéaire sur des courbes de diffusion de neutrons par des solutions de *Hm* MalDH en sulfate d'ammonium. L'évolution de ces valeurs apparentes en fonction de c révèle le caractère répulsif (I(0,c)/c et R_g diminuent avec c) ou attractif (ils augmentent) des

interactions interparticulaires. De plus, la limite à $c = 0$ du rayon de giration $R_g°$ est caractéristique de la particule " idéale " et permet de vérifier l'état d'agrégation de la protéine : des valeurs de $R_g°$ de l'ordre de 30-32 Å correspondent au tétramère de la *Hm* MalDH [Ebel *et al.*, 1999b].

2.2.3. Diffusion de neutrons aux petits angles

Les expériences de diffusion de neutrons ont été réalisées à l'Institut Laue Langevin (ILL) à Grenoble sur les instruments D11 [www.ill.fr/YellowBook/D11/ et Lindner *et al.*, 1992] et D22 [www.ill.fr/YellowBook/D22/]. Tous deux reçoivent des neutrons issus de sources secondaires froides (bain de $^2H_2$ à 25 K) alimentées par le cœur du réacteur dont la puissance thermique est de 58 MW et qui peut être distant de plus de 100 m de l'échantillon. Ce faisceau polychromatique passe à travers un sélecteur de vitesse (tambour tournant à sillons hélicoïdaux) qui joue le rôle de monochromateur en sélectionnant les neutrons ayant une vitesse correspondant à la longueur d'onde λ choisie. La relation de de Broglie relie en effet la vitesse v d'une particule à sa masse m et à la longueur d'onde associée : $\lambda = h/mv$, où h est la constante de Planck. Le faisceau monochromatique ($\Delta\lambda/\lambda = 0.1$) d'environ 1 cm^2 est collimaté par une série de guides amovibles. L'intensité diffusée par l'échantillon est enregistrée par un détecteur bidimensionnel sous vide (64x64 et 96x96 cm^2 pour respectivement D11 et D22) distant de 1 à 35 m.

Aspects expérimentaux

Afin de couvrir une gamme de q de 0.01 à 0.3 $Å^{-1}$ (environ) essentielle pour obtenir les informations sur les interactions (à " petits angles ") et pour vérifier la soustraction du signal du solvant (" grands angles "), deux distances échantillon – détecteur (E-D) sont nécessaires. Elles ont été choisies afin d'offrir un large recouvrement des deux spectres. A chaque position correspond une longueur de collimation (coll) qui offre le meilleur compromis entre la divergence du faisceau et un flux de neutrons maximum sans trop s'écarter de la zone de réponse linéaire du détecteur. Les valeurs des différents paramètres instrumentaux utilisés sont présentées dans le Tableau 2.1.

---

[*] Les références sont sous le format " instrument année.mois " et permettent une recherche dans la liste des conditions des expériences de diffusion en annexe 8.2 p.190.

Les échantillons sont contenus dans des cuves en quartz de 1 mm de trajet optique et leur volume minimum est de 160 μl. Un porte-échantillon automatisé et thermostaté permet de positionner 22 échantillons à la fois. Selon la concentration de l'échantillon, le temps d'acquisition varie de 7 à 60 min.

Paramètres		D11	D22
Longueur d'onde :	λ (Å) ($\Delta\lambda/\lambda = 0.1$)	6	10
Petits angles :	E-D (m)	4.5	4
	Collimation (m)	5.5	8
	DH (m)	0	0.382
Grands angles :	E-D (m)	1.25	1.4
	Collimation (m)	13	11.2
	DH (m)	0	0.382
Transmission :	E-D (m)	10	12
	Collimation (m)	5.5	11.2
	DH (m)	0	0
Temps mort du détecteur en 10^{-6} s		2.17	2.056
Facteur de correction f		0.8	1

Tableau 2.1 : Paramètres des instruments D11 et D22 (ILL) de diffusion de neutrons aux petits angles (E-D : distance échantillon – détecteur, DH : décalage horizontal du détecteur).

Traitement des données

La procédure de normalisation des données, résumée ci-dessous, est décrite dans les références suivantes [ftp.ill.fr/pub/cs/sans/sas_manual.pdf et Ghosh *et al.*, 1998].

L'intensité enregistrée est moyennée radialement, puis le signal de la cuve vide et le bruit de fond électronique et ambiant (diffusion du cadmium) sont soustraits. L'intensité est alors normalisée par la diffusion (incohérente) de l'eau pure (1 mm) — et donc au flux incident de neutrons — en tenant compte des transmissions et de l'efficacité du détecteur via son temps mort [Lindner, 1998]. Le signal du solvant (traité de la même manière) est soustrait à celui de l'échantillon afin d'obtenir la diffusion des protéines seules. A ce stade, il est souvent nécessaire de soustraire une constante qui traduit une différence de diffusion incohérente entre la solution de protéine et le solvant associé. Elle est déterminée en analysant l'intensité aux plus grandes valeurs de q (0.25-0.3 $Å^{-1}$) qui doit tendre vers 0. La même valeur est obtenue alternativement en appliquant la loi de Porod (*i.e.* l'intensité doit être proportionnelle à q^{-4} dans la zone $2\pi/Rg << q << 2\pi/\lambda$). L'intensité I(q,c) est ensuite divisée

par la concentration en protéine. Afin de pouvoir comparer l'intensité entre différentes expériences, il reste à la normaliser aux transmissions de l'eau (T_{eau}) et de l'échantillon (T_{ech}), à l'épaisseur de l'échantillon (t) et à un facteur (f) de correction de l'anisotropie de la diffusion incohérente de l'eau qui dépend de λ. L'intensité absolue est donc [Jacrot et Zaccai, 1981] :

$$\frac{Iabs}{c} = \frac{1 - T_{eau}}{4\pi f T_{ech} t 10^{-3}} \frac{I}{c} \quad . \tag{2.10}$$

Les valeurs de f sont données dans le Tableau 2.1, et t vaut 0.1 cm.

2.2.4. Diffusion de rayons X aux petits angles

Les expériences de diffusion de rayons X ont été réalisées au Laboratoire pour l'Utilisation du Rayonnement Electromagnétique (LURE) à Orsay, sur l'instrument D24 [www.lure.u-psud.fr/Experiences/DCI/D24.HTM et Depautex *et al.*, 1987; Dubuisson *et al.*, 1997]. Le faisceau de rayons X (rayonnement synchrotron) est rendu monochromatique par réflexion de Bragg sur un monocristal de germanium. Il est focalisé dans le plan horizontal par ce cristal qui est courbé, puis un système de fentes (verticales et horizontales) élimine les diffusions parasites et fournit un faisceau d'environ 1.5 mm^2 (au niveau de l'échantillon). L'ensemble du dispositif est sous vide afin de s'affranchir de la diffusion des rayons X par l'air.

Aspects expérimentaux

Nous avons utilisé une longueur d'onde de 1.488 Å (raie K du nickel) et un détecteur linéaire à sensibilité de position, placé à 1.82 m de l'échantillon qui nous permettaient une acquisition de l'intensité diffusée pour des q de 0.01 à 0.2 $Å^{-1}$. La solution étudiée d'un volume minimum de 80 μl est contenue dans un capillaire de quartz (1 mm de diamètre), thermostaté (parfois avec quelques difficultés ☺). Selon la concentration de l'échantillon, le temps d'acquisition varie de 10 à 30 min.

Traitement des données

L'intensité diffusée est normalisée à l'intensité du faisceau incident, à sa transmission et à l'efficacité du détecteur grâce au signal du noir de carbone qui a une forte diffusion très plate et sert donc de référence [Luzzati et Tardieu, 1980]. Le signal du solvant est ensuite soustrait et les courbes de diffusion sont normalisées à la concentration en protéine.

2.3. Les interactions protéine – protéine en solution

2.3.1. Développement du viriel

Les techniques abordées dans cette thèse (diffusion aux petits angles ou ultracentrifugation analytique) permettent d'obtenir une masse moléculaire apparente (M_{app}) reliée à la dérivée $(\partial\Pi/\partial c)_{T,\mu}$ à température T et aux potentiels chimiques μ des espèces diffusibles constants, le développement du viriel est présenté pour cette variable :

$$\frac{1}{M_{app}} = \frac{1}{N_A kT}\left(\frac{\partial\Pi}{\partial c}\right)_{T,\mu} = \frac{1}{M} + 2A_2 c + 3A_3 c^2 + \cdots . \tag{2.11}$$

L'inverse de la masse moléculaire M^{-1} est le premier coefficient du viriel, A_2 le second (en $mol.l.g^{-2}$) et A_3 le troisième ; c est la concentration massique en protéine ($mg.ml^{-1}$) [Eisenberg, 1976]. Dans la limite d'une concentration protéique et d'interactions faibles, A_2 est relié au potentiel de paires W(r) par la relation suivante [McMillan et Mayer, 1945; Belloni, 1991] :

$$A_2 = 2\pi\frac{N_A}{M^2}\int_0^{\infty}\left(1 - e^{-W(r)/kT}\right)r^2 dr . \tag{2.12}$$

Le second coefficient du viriel A_2 reflète donc l'ensemble des interactions interparticulaires. Il est positif pour des interactions globalement répulsives et négatif pour une résultante attractive.

Détermination du A_2 par diffusion aux petits angles

Dans une expérience de diffusion aux petits angles (de neutrons ou de rayons X), l'intensité à angle nul I(0,c)/c, déterminé par l'approximation de Guinier, est proportionnelle à la masse apparente (cf. (2.6) p.53, pour des particules en interaction) :

$$I(q = 0, c) = Cste\left(\partial\rho_{n,el}/\partial c\right)_{\mu}^{2} c\, M_{app} , \tag{2.13}$$

En utilisant les premiers termes du développement du viriel (2.11), on obtient :

$$\frac{I(0,c)}{c} = Cste\left(\partial\rho_{n,el}/\partial c\right)_{\mu}^{2}\frac{M}{1 + 2A_2 M c} = \frac{I°(0,0)}{1 + 2A_2 M c} , \tag{2.14}$$

où I°(0,0) est la limite de I(0,c)/c lorsque c tend vers 0, soit :

(2.15) $$\frac{c}{I(0,c)} = \frac{1}{I°(0,0)} + \frac{2A_2M}{I°(0,0)}\,c = m1 + m2\,c\ .$$

Une régression linéaire de c/I(0,c) en fonction de c permet donc d'obtenir A_2 qui est lié à la pente (m2) et à l'ordonnée à l'origine (m1) de la droite :

(2.16) $$A_2 = \frac{m2}{2M\,m1} \qquad \Delta A_2 = A_2\left(\frac{\Delta m1}{m1} + \frac{\Delta m2}{m2}\right)\ .$$

La concentration en protéine est déterminée en utilisant un coefficient d'extinction massique $\varepsilon_{280}^{0.1\%}$ déterminé pour la masse de la protéine sans contre-ions (M = 130552 Da) (cf. (2.1) p.49). La protéine en solution est dans une approche thermodynamique rigoureuse définie comme une particule neutre, formée de la protéine et de ses contre-ions. Sa masse M_+ est donc différente selon le solvant : par exemple, 134138 Da en NaCl (protéine avec 156 Na^+) ou 132448 Da en $MgCl_2$ (protéine avec 78 Mg^{2+}). Le A_2 obtenu par l'équation (2.16) est à multiplier par $(M/M_+)^2$ afin de rester cohérent avec cette définition de la protéine en solution.

Trois concentrations c, inférieures à 30 mg/ml (typiquement 7, 14 et 21 mg/ml) ont généralement été utilisées pour obtenir les différents A_2. Dans certains cas (par exemple des interactions fortement attractives), une régression polynomiale du second degré a été utilisée pour obtenir un meilleur accord avec les données expérimentales. Cela correspond à l'ajout d'un terme en c^2 dans l'équation (2.15).

La Figure 2.3 illustre la détermination du second coefficient du viriel par diffusion de neutrons aux petits angles, pour la *Hm* MalDH en $(NH_4)_2SO_4$ 1.5 M. Les I(q,c)/c ont été déterminés par l'approximation de Guinier (cf. Figure 2.2 p.54). Seules les trois premières concentrations en protéine de la gamme mesurée (5 à 58 mg/ml) sont utilisées pour déterminer le A_2 grâce à l'équation (2.16) dont les paramètres sont indiqués en encart dans la Figure 2.3. Le point à plus forte concentration n'est correctement modélisable qu'en ajoutant le terme en c^2 dans le développement du viriel. Les deux ajustements donnent la même valeur pour le A_2 : $3.5 \pm 0.3\ 10^{-8}$ mol.l.g^{-2} (en tenant compte des 156 contre-ions NH_4^+ via M_+ = 133366 Da).

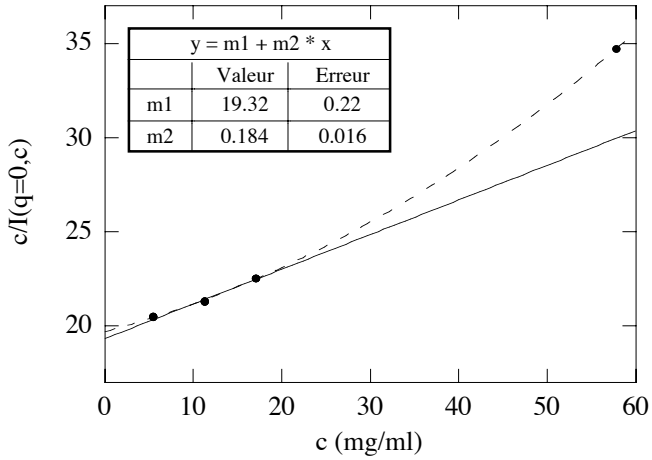


Figure 2.3 : Inverse des intensités diffusées à angle nul I(q=0,c)/c en fonction de la concentration en protéine c, par des solutions de *Hm* MalDH en $(NH_4)_2SO_4$ 1.5 M (réf. D22 00.10). L'encart donne les valeurs des paramètres de la droite modélisant les trois premiers points qui permettent de calculer le second coefficient du viriel A_2.

Détermination de A_2 par vitesse de sédimentation

L'équilibre de sédimentation est une technique d'ultracentrifugation analytique généralement utilisée pour déterminer le second coefficient du viriel A_2. La répartition de la concentration en macromolécule dans la centrifugeuse permet en effet de déterminer la masse moléculaire apparente (M_{app}) grâce à la relation :

$$\frac{d\ln c}{dr^2} = \frac{\omega^2}{2N_A kT}\left(\frac{\partial \rho_m}{\partial c}\right)_{T,\mu} M_{app} \quad , \tag{2.17}$$

où ω est la vitesse angulaire, ρ_m la densité (massique en g/ml) de la solution et r la distance au centre du rotor. M_{app} est rigoureusement liée à $(\partial\Pi/\partial c)$ (cf. (2.11) p.58) et au A_2. Ce dernier est obtenu de l'évolution de M_{app} avec la concentration en protéine. Toutefois, en présence de sel ou de solvant organique à forte concentration, à cause des interactions préférentielles de la protéine avec les espèces du solvant, $(\partial\rho/\partial c)$ (et souvent une ligne de base) doit être déterminé en même temps que A_2, ce qui rend difficile le calcul de ce dernier. Cette technique est donc peu adaptée aux conditions de solvant que nous utilisons.

Nous avons exploré les potentialités d'une autre technique d'ultracentrifugation analytique qui permet aussi de déterminer A_2, il s'agit de la vitesse de sédimentation. Cette

nouvelle technique est présentée plus en détail dans le premier article (cf. §3 p.69, [Solovyova *et al.*, 2001]).

2.3.2. Simulations numériques

Le développement du viriel décrit le comportement asymptotique à faible concentration en protéine des interactions protéine – protéine. Il quantifie la résultante de l'ensemble de ces interactions (attractives et répulsives) en les intégrant sur tout l'espace (cf. (2.12) p. 58). Comme nous l'avons vu précédemment, il est lié à l'intensité diffusée à angle nul I(q=0,c). L'intensité diffusée I(q,c) contient cependant plus d'information puisqu'elle est reliée à la distribution spatiale des particules de protéine dans la solution (cf. (2.9) p.53). L'ensemble de la courbe de diffusion ou plus précisément le facteur de structure S(q,c) pourrait donc potentiellement nous renseigner sur la nature des interactions. Nous avons donc, dans un deuxième temps, abordé les interactions protéine – protéine via la modélisation des facteurs de structure par une méthode de simulation numérique.

Cette méthode relie les potentiels de paires (à symétrie sphérique, donc fonction que de la distance interparticulaire $r = |\mathbf{r}|$) décrits en introduction aux propriétés thermodynamiques et structurales de la solution de protéine. Son principe est de calculer, à partir de ces potentiels (problème à 2 particules), la fonction de distribution de paires g(r) (problème à N particules) grâce à l'équation d'Ornstein-Zernike et d'une équation intégrale [Hansen et McDonald, 1986; Belloni, 1991]. g(r) représente la probabilité de trouver une particule à une distance comprise entre r et r+dr de la particule de référence. La transformée de Fourier de $\delta(r) + \rho g(r)$ étant le facteur de structure $S_{calc}(q,c)$, il peut être comparé au facteur de structure expérimental $S_{exp}(q,c)$. L'intensité diffusée par les protéines et des grandeurs thermodynamiques (compressibilité et pression osmotique) peuvent aussi être déduites. La simulation consiste donc à faire varier les paramètres des potentiels pour obtenir le meilleur accord entre les données calculées et expérimentales.

Les potentiels utilisés supposent une symétrie sphérique et sont relativement simples, comparés à la richesse des interactions protéine – protéine. Cette démarche reste donc modeste.

Le programme que nous avons utilisé a été élaboré par Luc Belloni [Belloni, 1985, 1988]. Son principe et son utilisation sont présentés ci-dessous.

Equations d'Ornstein-Zernike (OZ) et intégrales

L'équation exacte d'Ornstein-Zernike [Hansen et McDonald, 1986] s'écrit :

$$h(\mathbf{r}_{12}) = c(\mathbf{r}_{12}) + \rho \int c(\mathbf{r}_{13})\, h(\mathbf{r}_{32})\, d\mathbf{r}_3 = c(\mathbf{r}_{12}) + \rho\, c \otimes h(\mathbf{r}_{12}) \,, \tag{2.18}$$

où h(**r**) est la fonction de corrélation totale, c(**r**) la fonction de corrélation directe et ⊗ symbolise le produit de convolution. La corrélation totale h entre deux particules 1 et 2 est liée à la fonction de distribution de paires g par la relation h = g – 1. Elle est la somme d'une corrélation directe c et de toutes les corrélations indirectes entre 1 et 2 faisant intervenir un nombre croissant de particules intermédiaires 3. Ceci apparaît clairement lorsque l'on résout formellement l'équation (2.18) par itération :

$$h = c + \rho c \otimes c + \rho^2 c \otimes c \otimes c + \cdots \,. \tag{2.19}$$

L'équation (2.18) s'écrit plus simplement dans l'espace de Fourier (avec ^ symbolisant la transformée de Fourier normalisée par la densité ρ) :

$$\hat{h} = \hat{c} + \hat{c}\hat{h} \qquad \text{soit :} \qquad \hat{h} = \hat{c} / (1 - \hat{c}) \,. \tag{2.20}$$

Comme le facteur de structure S(q) est lié à la transformée de Fourier de g(r) (cf. (2.9) p.53) et donc de h(r), l'équation (2.18) s'écrit* :

$$S(q) = 1 + \hat{h}(q) = 1 / \left(1 - \hat{c}(q)\right) \,. \tag{2.21}$$

L'équation exacte OZ contient deux fonctions inconnues h et c. Il faut donc introduire une autre relation pour clore le système. Il existe différentes relations appelées équations intégrales dérivées de la mécanique statistique des liquides simples qui admettent des solutions analytiques ou numériques et qui sont plus ou moins approximatives. Nous avons utilisé l'équation dite " HyperNetted Chain " (HNC) [Belloni, 1991] :

$$g(r) = \exp\left[-\frac{W(r)}{kT} + h(r) - c(r) \right] \,, \tag{2.22}$$

* Pour être exact, S(q) = TF[δ(r) + ρg(r)] = TF[δ(r) + ρ(h(r) + 1)]. Le terme en δ(q) provenant de la TF de 1 est donc omis dans l'équation (2.21).

où W(r) est le potentiel de paires. Cette équation intégrale non linéaire nécessite une résolution numérique.

Principe du calcul numérique

Chaque fonction est discrétisée sur N points régulièrement espacés r_i. Le choix du pas d'échantillonnage réel Δr fixe le pas dans l'espace réciproque $\Delta q = 2\pi/(N\Delta r)$. Pour une bonne définition des fonctions, Δr doit être très inférieur au diamètre des colloïdes et plus petit que la longueur de Debye (λ_D) déduite de la composition saline du solvant (les paramètres que nous avons utilisés sont présentés par la suite).

Comme l'équation d'OZ s'exprime simplement dans l'espace de Fourier, le calcul est mené dans l'espace de Fourier (q) et dans l'espace réel (r), selon le processus itératif suivant. A partir d'une fonction γ = h - c de départ estimée $\gamma^e(r)$, g(r) est calculée grâce à l'équation HNC (2.22), c(r) est déduite puis c(q) par transformée de Fourier. La relation OZ (2.21) donne S(q) puis γ(q) et $\gamma^s(r)$ par transformée de Fourier inverse.

Le cycle suivant (n+1) est démarré avec $\gamma^e_{n+1} = \alpha\gamma^s_n + (1-\alpha)\gamma^e_n$, où α est le paramètre de mélange, compris entre 0 et 1. La convergence est atteinte lorsque $\left\{\frac{1}{N}\sum_{i=1}^{N}\left(\gamma^s(r_i) - \gamma^e(r_i)\right)^2\right\}^{1/2}$ est inférieure à la précision désirée. Une précision de 10^{-4} (ou de 10^{-10}) est souhaitable lorsque l'on est loin (respectivement proche) d'une transition de phase.

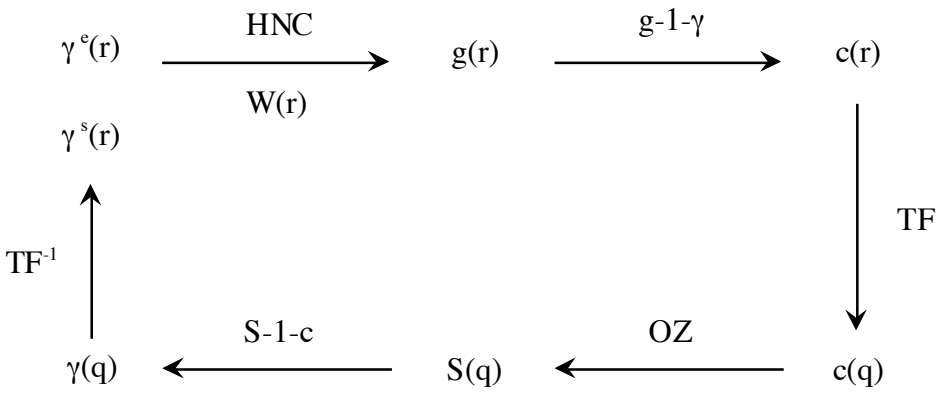


Figure 2.4 : Schéma d'un cycle de simulation numérique.

Résolution pratique et stratégie ?

Le modèle et le potentiel interparticulaire choisis

Nous avons utilisé le modèle dit à un composant où seules les interactions protéine – protéine sont traitées. Il suppose qu'il existe une grande dissymétrie de taille entre la protéine et les autres espèces de la solution, ce qui est notre cas.

Le solvant (eau, sel et co-solvant) n'est représenté que par sa constante diélectrique (ε) et les ions n'interviennent explicitement dans l'expression du potentiel interparticulaire que par l'intermédiaire de la constante d'écrantage ($\kappa = 1/\lambda_D$) via la force ionique (I). La présence de sel ou de co-solvant modifiant la constante diélectrique ε d'une solution, sa valeur n'est donc pas celle de l'eau pure (78.4 à 25°C), mais celle du système eau – sel – co-solvant. Pour chaque condition expérimentale (solvant et température), nous avons calculé cette valeur à partir de la variation de ε(eau) en fonction de la température [voir par exemple Hamer, 1975], modifiée par le sel (type et concentration) [Barthel *et al.*, 1998] et par le co-solvant (MPD) [Travers et Douzou, 1974]. Ces données et leurs obtentions sont détaillées en annexe 8.1 (cf. p.187).

La force ionique est calculée par l'expression qui la définit : $I = 0.5 \sum c_i Z_i^2$, où c_i et Z_i sont respectivement la concentration (en M) et la charge de l'ion i du sel. Ainsi, pour une solution de NaCl 2 M, I vaut 2 M et pour une solution de $MgCl_2$ 1 M, I vaut 3 M.

Dans ce modèle à un composant, le potentiel de paires protéine – protéine (W_{pp}) est un potentiel effectif car moyenné sur les ions. La charge de la protéine (Z) et son diamètre (D) sont des paramètres effectifs qui représentent la protéine après condensation des ions à sa surface : ils peuvent donc différer des valeurs structurales. Le potentiel que nous avons utilisé est le suivant :

$$W_{pp}(r) = \begin{cases} +\infty & \text{pour } r \le D \\ \dfrac{kT\,Z^2\,L_B}{(1+\kappa D/2)^2}\,\dfrac{e^{-\kappa(r-D)}}{r} - J\,D\,\dfrac{e^{-(r-D)/d}}{r} & \text{pour } r > D \end{cases}, \tag{2.23}$$

où $\kappa = (8\pi L_B I)^{1/2}$ et $L_B = e^2/4\pi\varepsilon_0\varepsilon kT$ est la longueur de Bjerrum. Il est la somme de trois potentiels. La première ligne correspond au potentiel de volume exclu dans l'hypothèse d'une sphère dure de diamètre D (cf. (1.1) p.26). Le premier terme de la seconde ligne correspond au potentiel coulombien écranté que nous avons détaillé dans l'introduction (cf. §1.2.2 p.27). Ces deux potentiels sont positifs et traduisent des répulsions interparticulaires. A l'inverse, le

troisième potentiel (2e terme de la seconde ligne) est négatif et traduit une attraction. Il va correspondre à la somme de toutes les contributions attractives possibles (van der Waals, déplétion, chimique, …) et en particulier celles qui sont induites par le solvant. Il prend la forme d'un potentiel de Yukawa décrit par sa profondeur $J > 0$ (en unité de kT) et par sa portée d. La composition de la solution fixe les paramètres : T, ε, I et la concentration en protéine. D, Z, J et d sont les variables ajustables.

Stratégie de la modélisation : paramètres ajustables et facteur de structure

Pour modéliser les interactions protéine – protéine, nous partons de la description la plus simple possible du système en termes de potentiels interparticulaires, pour la préciser progressivement afin de rendre compte des données expérimentales. Nous avons en premier lieu considéré le potentiel de volume exclu et donc ajusté le diamètre D de la particule. Puis, pour des conditions globalement répulsives, nous avons ajouté la composante électrostatique et accordé la charge effective Z. Dans le cas d'une modélisation incorrecte et pour les conditions globalement neutres ou attractives, nous avons finalement fixé la valeur de la charge Z et introduit la composante attractive (J, d).

Le choix du type et des caractéristiques du potentiel attractif a été déterminé par les considérations suivantes. Il est décrit que pour des grosses particules sphériques dans des solvants aqueux, le potentiel attractif n'est important que sur des distances de l'ordre de quelques angströms, ce qui détermine sa portée d [Tardieu *et al.*, 1999]. De plus, dès que cette portée reste plus petite (ou du même ordre) que les autres distances caractéristiques (diamètre et longueur de Debye), la forme précise du potentiel attractif n'est pas fondamentale. Nous avons donc choisi pour décrire l'ensemble des interactions attractives une forme de potentiel de type Yukawa de description mathématique simple. Il s'avère que divers couples (J, d) peuvent, dans ces conditions, rendre compte d'un même facteur de structure expérimental [Peyre *et al.*, 1997]. Les résultats des modélisations ne peuvent donc permettre d'assigner à une condition expérimentale de solvant une valeur unique de couple (J, d). Pour rendre compte de l'évolution des potentiels (J, d) en fonction d'une variable caractérisant le solvant (concentration saline, température, …), nous avons donc choisi de ne faire varier que la profondeur J et de fixer la portée d. Il a été montré que seules des portées inférieures à D/12 rendent compte de transition de phases liquide – liquide de solution de protéines [Malfois *et al.*, 1996]. D'autres études sur les interactions interparticulaires ont aussi utilisé des portées de l'ordre du dixième du diamètre de la macromolécule [Peyre *et al.*, 1997; Tardieu *et al.*, 1999].

Compte tenu d'un diamètre structural d'environ 80 Å pour la *Hm* MalDH, nous avons donc choisi de fixer d à 5 Å.

Le programme de Luc Belloni permet de calculer le facteur de structure S(q,c) puis l'intensité I(q,c) diffusée par des particules sphériques, mono ou polydisperses [Belloni, 1985, 1988]. Les intensités diffusées sont le produit de S(q,c) déterminé par les interactions protéine – protéine et du facteur de forme I(q,c=0). Nous avons choisi d'effectuer la modélisation sur les facteurs de structure. En effet, ce choix n'introduit pas de contrainte sur la forme (symétrie et dispersité) de la protéine, via le facteur de forme I(q,c=0). La Figure 2.5-a compare le facteur de forme expérimental issu de la diffusion de neutrons par la *Hm* MalDH en $(NH_4)_2SO_4$ 1.5 M avec celui d'une sphère de 80 Å de diamètre. Dans l'intensité expérimentale, on ne retrouve pas les minima prononcés provenant de la diffusion d'une sphère, même lissée pour tenir compte de la résolution en longueur d'onde ($\Delta\lambda/\lambda$) de l'instrument. Cette différence est observée quel que soit le diamètre de la sphère. Une sphère ne modélise donc pas une molécule de *Hm* MalDH.

Pour chaque condition de solvant, les facteurs de structures expérimentaux S(q,c) sont obtenus par la division des intensités I(q,c)/c par le facteur de forme I(q,0)/c (cf. (2.8) p.53). Nous avons utilisé des facteurs de forme I(q,0)/c extrapolés à concentration nulle de protéine plutôt que les intensités diffusées aux plus faibles concentrations en protéine (quelques mg/ml). En effet, les interactions se manifestent même à de très faibles concentrations en protéine : ceci apparaît clairement dans les expériences de vitesse de sédimentation réalisées à des concentrations relativement faibles (1 à 12 mg/ml) (cf. §3 p.69). Pour chaque valeur de q, les $I(q,c_i)/c_i$ ont été linéairement extrapolés à $c = 0$ sur les trois plus faibles concentrations c_i (inférieures à 30 mg/ml, Figure 2.5-b). Cette extrapolation est relativement aisée pour les petits angles ($q < 0.06$-0.08 Å^{-1}), mais s'avère délicate au-delà car les différences entre les intensités mesurées à plusieurs concentrations c sont très faibles. Au-delà de $q = 0.08$ Å^{-1} – qui correspond à des distances dans l'espace réel ($r = 2\pi/q$) inférieures à environ 80 Å – le facteur de forme prédomine dans l'intensité diffusée : on mesure principalement des corrélations intramoléculaires. Le facteur de structure est indiscernable de 1 et sera finalement peu discriminant pour la modélisation. Nos facteurs de forme I(q,0)/c sont donc l'extrapolation linéaire des I(q,c)/c dont la " queue " (au-delà d'environ $q = 0.08$ Å^{-1}) est l'intensité diffusée à une concentration intermédiaire (environ 20 mg/ml) et ils sont légèrement lissés afin de réduire le bruit. Pour certaines expériences où seulement deux

concentrations ont été enregistrées, le facteur de forme est simplement l'intensité diffusée à plus faible concentration en protéine (c de l'ordre de 5 mg/ml).

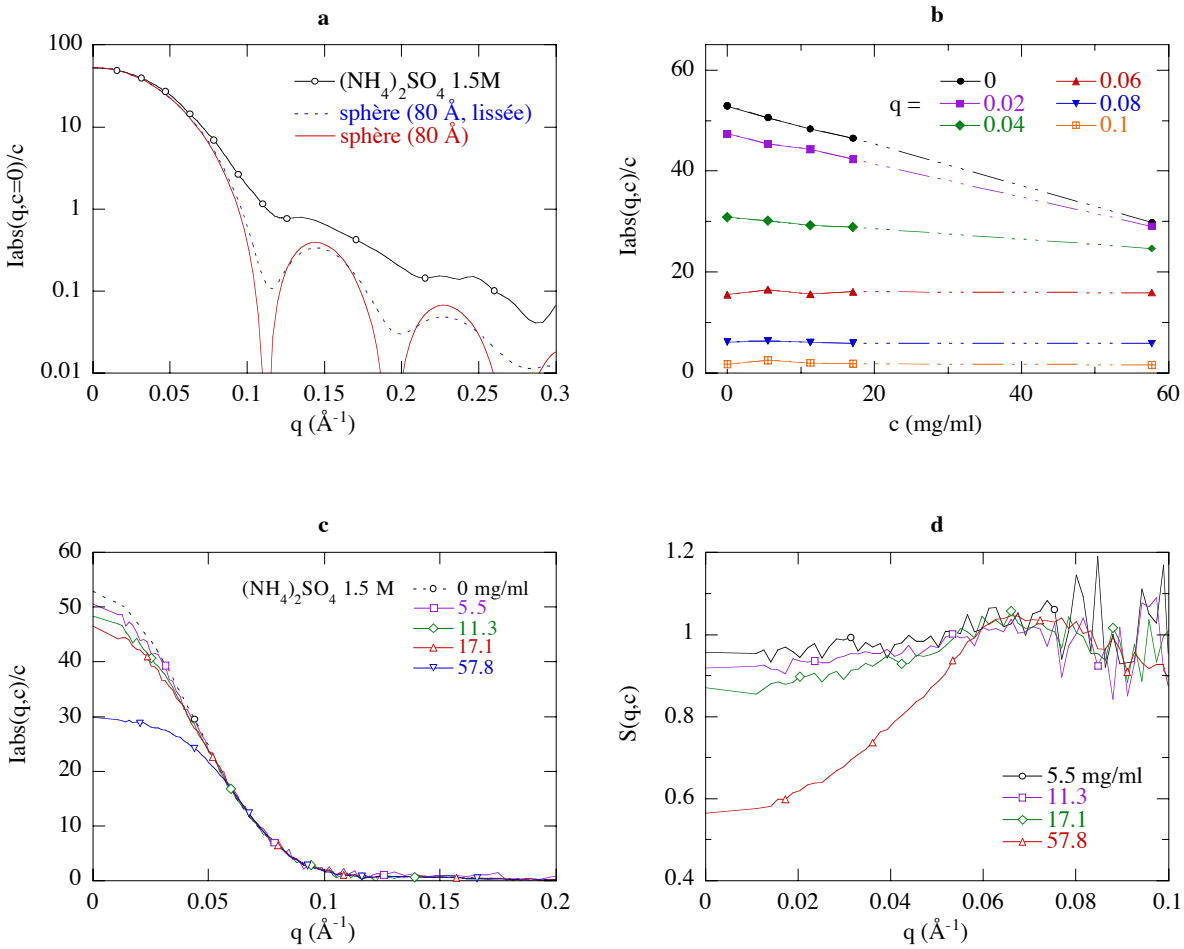


Figure 2.5 : Facteurs de forme et de structure de la *Hm* MalDH en $(NH_4)_2SO_4$ 1.5 M par diffusion de neutrons. **(a)** Le facteur de forme expérimental Iabs(q,c=0)/c est comparé à l'intensité diffusée par une sphère de 80 Å de diamètre. L'intensité dite " lissée " tient compte de la résolution instrumentale ($\Delta\lambda/\lambda = 0.1$). **(b)** L'extrapolation linéaire des Iabs(q,c)/c pour quelques valeurs de q sur les trois plus faibles concentrations en protéine permet de déterminer le facteur de forme Iabs(q,c=0)/c. Ce dernier est présenté avec les courbes de diffusion de neutrons enregistrées en **(c)** et les facteurs de structure correspondants en **(d)**.

La Figure 2.5-c présente les intensités diffusées par les solutions de *Hm* MalDH en $(NH_4)_2SO_4$ 1.5 M, aux différentes concentrations de protéine enregistrées et le facteur de forme extrapolé (courbe à c = 0). Ces courbes divisées par le facteur de forme donnent les facteurs de structure expérimentaux correspondants (Figure 2.5-d). Les S(q,c) restent toujours proches de 1 et sont assez imprécis pour les faibles c. La modélisation se fera donc

principalement sur le $S_{exp}(q,c)$ à la plus forte concentration en protéine car il est plus discriminant.

Les paramètres du calcul numérique

Du fait du caractère halophile de la *Hm* MalDH, les solvants utilisés sont fortement salins : la concentration en sel est fréquemment comprise entre 1 et 4 M. La longueur de Debye λ_D est donc très petite (1-3 Å), ce qui nous oblige à choisir un pas Δr tout aussi petit : de 1 Å. Afin de décrire le plus correctement les fonctions dans l'espace réel, qui doivent être nulle au-delà de la distance $N\Delta r$, et celles dans l'espace réciproque, discrétisées avec le pas $\Delta q = 2\pi/(N\Delta r)$, nous avons utilisé le N le plus grand possible, *i.e.* 2048. Le paramètre de mélange α valait généralement 0.5, mais était diminué à 0.1 ou 0.05 lorsqu'une attraction était ajoutée, afin d'éviter la divergence du processus itératif.

Après la convergence, le facteur de structure calculé $S_{calc}(q,c)$ qui tient compte de la résolution instrumentale ($\Delta\lambda/\lambda$), est comparé à l'expérimental $S_{exp}(q,c)$.

3. UNE NOUVELLE METHODE DE DETERMINATION DU SECOND COEFFICIENT DU VIRIEL

3.1. Présentation de l'article

Bien que l'article " Non-ideality by sedimentation velocity of halophilic malate dehydrogenase in complex solvents " [Solovyova *et al.*, 2001] fasse référence à des résultats présentés plus loin dans ce manuscrit, nous avons préféré l'introduire ici à cause de son caractère méthodologique.

Au cours de l'étude des interactions protéine – protéine, il nous est apparu intéressant de rapidement déterminer l'effet sur le second coefficient du viriel A_2 de variations dans la composition de divers solvants contenant de fortes concentrations de sels ou de solvant organique. Les propriétés statiques et dynamiques des macromolécules en solution sont reliées non seulement aux caractéristiques des particules, mais aussi à leur distribution spatiale qui dépend des interactions interparticulaires. A_2 peut ainsi être évalué grâce aux nombreuses techniques qui déterminent la masse molaire apparente de la protéine en solution [Eisenberg, 1976, 1981]. Cependant, des mesures précises de pression osmotique s'avèrent par exemple difficiles en présence de fortes concentrations de sel ; la diffusion statique de la lumière est très sensible aux hétérogénéités de l'échantillon ; et la diffusion aux petits angles des rayons X et des neutrons nécessite l'utilisation de grands instruments. Dans le domaine de la centrifugation analytique, l'équilibre de sédimentation souffre du fait que la masse molaire en solution (et souvent une ligne de base) doit être déterminée en même temps que A_2 lorsqu'elle diffère de la masse molaire à cause des interactions préférentielles de la macromolécule avec le solvant. Finalement, A_2 peut être déterminé par la combinaison des facteurs de dépendance à la concentration de la macromolécule des coefficients de sédimentation s et de diffusion translationnelle D. Les expériences de vitesse de sédimentation doivent permettre de déterminer ces coefficients s et D qui déterminent ensemble la forme des fronts de sédimentation, et d'en déduire A_2. Cet article présente une nouvelle méthode de détermination de A_2 par modélisation des profils de vitesse de sédimentation. Elle est validée par des mesures effectuées pour des solutions de *Hm* MalDH en NaCl avec ou sans MPD, pour lesquelles les propriétés de la particule idéale et de ses

interactions (s, D et A_2) ont été évaluées indépendamment. Nous rappelons ici très brièvement la méthode et les principaux résultats de cette étude.

Détermination de A_2 par vitesse de sédimentation

L'équation de Lamm du transport exprime l'évolution spatiale et temporelle (t) de la concentration c (mg/ml) d'un soluté homogène (en centrifugation) en fonction des coefficients de sédimentation s et de diffusion D [Lamm, 1929; Tanford, 1961, p.369] :

(3.1) $$\left(\frac{\partial c}{\partial t}\right)_r = \frac{1}{r}\left[\frac{\partial}{\partial r}\left(Dr\frac{\partial c}{\partial r} - s\omega^2 r^2 c\right)\right]_t ,$$

où ω est la vitesse angulaire et r la distance à l'axe de rotation. Le coefficient de sédimentation s est relié à la masse molaire en solution $M(\partial\rho/\partial c)_{T,\mu}$ et au coefficient de friction f par la relation de Svedberg [Svedberg et Pedersen, 1940] :

(3.2) $$s = M\left(\partial\rho/\partial c\right)_{T,\mu}/N_A f ,$$

où $(\partial\rho/\partial c)_{T,\mu}$ est l'incrément de la densité de la solution. Par ailleurs, pour un fluide incompressible, l'équation de Stokes - Einstein relie D à la dérivée de la pression osmotique Π par la relation :

(3.3) $$D = \frac{M}{N_A f}\left(\frac{\partial \Pi}{\partial c}\right)_{T,\mu} \qquad D° = kT/f° .$$

Les coefficients s (ou f) et D sont des fonctions de la concentration en protéine c :

(3.4) $$s/s° = f°/f = 1 - k'_s c + \cdots$$

(3.5) $$D/D° = 1 + k_D c + \cdots ,$$

où ° indique les grandeurs à dilution infinie (*i.e.* dans la limite c = 0) et k'_s, k_D sont simplement les coefficients de dépendance en concentration de s et D. Rappelons que la dérivée de la pression osmotique peut être développée en série de puissance de c (cf. (2.11) p.58) :

(3.6) $$\frac{1}{M_{app}} = \frac{1}{N_A kT}\left(\frac{\partial \Pi}{\partial c}\right)_{T,\mu} = \frac{1}{M} + 2A_2 c + \cdots .$$

En exprimant D/D° de l'équation (3.3) à l'aide de (3.4) et (3.6) et en comparant les termes en c avec l'équation (3.5), nous obtenons la relation suivante qui permet de calculer A_2 à partir de k_D et k'_s :

(3.7) $$k_D \approx 2A_2M - k'_s \ .$$

La non-idéalité (k_D) du coefficient D a donc deux sources : A_2 et k'_s. La première est thermodynamique et conduit à la dépendance de la masse molaire à la concentration c ; la deuxième est hydrodynamique et concerne la dépendance de s et f.

Les profils de sédimentation sont déterminés par les coefficients de sédimentation s et de diffusion D de la macromolécule. Ils sont très sensibles aux interactions protéine – protéine faibles principalement via k'_s et via k_D. Nous avons obtenu ces coefficients par la modélisation directe des profils de sédimentation par les solutions de l'équation de Lamm (3.1) modifiée pour tenir compte de la dépendance linéaire en concentration (programme SEDFIT) [Schuck, 1998]. Une décomposition algébrique systématique du bruit des données expérimentales a été utilisée [Schuck et Demeler, 1999].

Pour chaque solvant, les expériences de vitesse de sédimentation sont réalisées pour trois concentrations initiales de protéine (entre 1 et 12 mg/ml) avec une ultracentrifugeuse analytique Optima XLA-I (Beckman) en utilisant le système de détection par interférence optique. La modélisation est menée d'une part sur les trois jeux de profils de sédimentation pris séparément et en ajustant successivement k'_s puis k_D, et d'autre part sur les trois jeux simultanément.

Validation de la méthode

Cette méthode a dans un premier temps été validée en modélisant des profils de sédimentation simulés pour des solutions de protéines non-idéales, avec différentes valeurs de s et D, et pour des systèmes attractifs ou répulsifs. Les valeurs des coefficients s, D, k_s et k_D trouvés sont en excellent accord avec les valeurs introduites pour la simulation.

Dans un deuxième temps, nous avons appliqué cette modélisation à des solutions de *Hm* MalDH en NaCl avec ou sans MPD, dont une proche d'une séparation de phases. Les A_2 obtenus par les deux types de modélisations sont comparés avec ceux issus d'autres techniques. Ils s'accordent très bien avec les A_2 déterminés par diffusion de neutrons aux

petits angles. Les valeurs obtenues par la diffusion de lumière sont moins précises, du fait de la forte salinité ou de la présence de MPD qui rendent ce type d'expérience plus difficile.

La modélisation des profils de vitesse de sédimentation apparaît comme une technique efficace de détermination du second coefficient du viriel. Elle nécessite, pour un solvant donné, moins de 2 mg de protéine qui sont récupérables et permet d'étudier des solvants complexes. Elle apparaît utile pour la recherche des conditions de cristallisation d'une macromolécule.

3.2. Article 1 : Non-ideality by sedimentation velocity of halophilic malate dehydrogenase in complex solvents[*]

Alexandra Solovyova, Peter Schuck, Lionel Costenaro and Christine Ebel

Biophysical Journal, 81(4) (2001), 1868-80.

[*] Les références citées dans l'article sont listées à la fin de celui-ci, indépendamment de la bibliographie du manuscrit.

1868 Biophysical Journal Volume 81 October 2001 1868–1880

Non-Ideality by Sedimentation Velocity of Halophilic Malate Dehydrogenase in Complex Solvents

Alexandra Solovyova,*† Peter Schuck,‡ Lionel Costenaro,* and Christine Ebel*
*Laboratoire de Biophysique Moléculaire, Institut de Biologie Structurale J. P. Ebel, F-38027 Grenoble, France; ‡Institute for Problems of Cryobiology and Cryomedicine of the National Academy of Sciences of Ukraine, 61015 Kharkov, Ukraine; and ‡Division of Bioengineering and Physical Science, Office of Research Services, National Institutes of Health, Bethesda, Maryland 20892 USA

ABSTRACT We have investigated the potential of sedimentation velocity analytical ultracentrifugation for the measurement of the second virial coefficients of proteins, with the goal of developing a method that allows efficient screening of different solvent conditions. This may be useful for the study of protein crystallization. Macromolecular concentration distributions were modeled using the Lamm equation with the approximation of linear concentration dependencies of the diffusion constant, $D = D°\ (1 + k_D c)$, and the reciprocal sedimentation coefficient $s = s°/(1 + k_s c)$. We have studied model distributions for their information content with respect to the particle and its non-ideal behavior, developed a strategy for their analysis by direct boundary modeling, and applied it to data from sedimentation velocity experiments on halophilic malate dehydrogenase in complex aqueous solvents containing sodium chloride and 2-methyl-2,4-pentanediol, including conditions near phase separation. Using global modeling for three sets of data obtained at three different protein concentrations, very good estimates for k_s and $s°$ and also for $D°$ and the buoyant molar mass were obtained. It was also possible to obtain good estimates for k_D and the second virial coefficients. Modeling of sedimentation velocity profiles with the non-ideal Lamm equation appears as a good technique to investigate weak inter-particle interactions in complex solvents and also to extrapolate the ideal behavior of the particle.

INTRODUCTION

The behavior of macromolecules is termed as ideal in very dilute solutions. In this limit, their positions, orientations, and movements are considered to be non-correlated. The macroscopically measured quantities usually are averages over an ensemble of molecules but nevertheless reflect the properties of individual particles, for example, molar mass and dimensions. In real solutions, however, even weak inter-particle interactions will cause a concentration dependence of the observed properties. Free-particle properties are obtained, therefore, from extrapolation of the data to infinite dilution of the macromolecule, and their concentration dependencies provide information about inter-particle interactions. In the past, the study of such weak inter-particle interactions has been the object of numerous studies because of its practical importance in many processes involving colloids in industry (e.g., paints, foods, and cosmetics). For example, weak interactions determine the tendency for a suspension of particles to remain in solution, to aggregate, to overcome phase separation or to form crystals. Interest has focused more recently on biological macromolecules in concentrated solution, with the aim of understanding their capacity to form crystals, which still represents a crucial and frequently very difficult step for the determination of their three-dimensional structure at high resolution. Likewise, studies on biological macromolecules in concentrated solutions are stimulated by the increasing awareness of the effect of macromolecular crowding on the biological function, for example, of proteins in the cytosol of a cell.

The static and transport properties of macromolecules in solution are related not only to the particle characteristics but also to their spatial distribution, which depends on intermolecular interactions (Pusey and Tough, 1985). The entire inter-particle interactions in solution determine the value of the second virial coefficient A_2, which has positive values for globally repulsive inter-particle interactions and negative values in the case of global attraction. Restricting the consideration to quasi-spherical particles, the second virial coefficient is related to the pair correlation function $g(r)$, which expresses the probability of finding particles of molar mass M at a distance r:

$$A_2 = 2\pi N_A/M^2 \int (1 - g(r))r^2 dr$$

For weakly interacting spheres, the pair correlation function $g(r)$ can be related, in terms of MacMillan-Mayer theory, to the mean particle-particle interaction potential $W(r)$ (MacMillan and Mayer, 1945):

$$g(r) = \exp^{-W(r)/kT}$$

The spatial particle distribution modulates the elastic scattering of the macromolecule in solution. Because $g(r)$ can be extracted from scattering profiles and calculated from pair potentials (Belloni, 1991), scattering techniques have been used to model inter-macromolecular potentials (Malfois et

Received for publication 27 November 2000 and in final form 18 June 2001.

Address reprint requests to Dr. Christine Ebel, Laboratoire de Biophysique Moléculaire, Institut de Biologie Structurale J. P. Ebel, UMR 5075 CEA-CNRS-UJF, 41, rue Jules Horowitz, F-38027 Grenoble Cedex 01, France. Tel.: 33-0-4-38789638; Fax: 33-0-4-38785494; E-mail address: christine.ebel@ibs.fr.

© 2001 by the Biophysical Society

0006-3495/01/10/1868/13 $2.00

al., 1996). Furthermore, the spatial and velocity correlations between particles in solution also determine the dynamic properties of the macromolecules in solution, for example, the collective diffusion coefficient and its concentration dependency as well as the concentration dependence of the friction coefficient. Different theoretical approaches have been described for the concentration dependencies of the friction and diffusion coefficients of diluted hard spheres, interacting through excluded volume effects, Coulomb repulsion, and van der Waals attraction (see references in Pusey and Tough, 1985). In another approach, the role of shape and solvation is emphasized (Rowe, 1992; Harding and Johnson, 1985a). As indicated above, there is a recent renewed interest in the measurement of parameters related to non-ideality. For a number of proteins in solvents known to promote crystallization, A_2 values measured in undersaturated or supersaturated solutions are moderately negative and lie in a fairly narrow range, and an experimental relationship with protein solubility has been established (George and Wilson, 1994; George et al., 1997; Guo et al., 1999; Bonneté et al., 1999).

Halophilic malate dehydrogenase from *Haloarcula marismortui* (*Hm* MalDH) is being studied in our laboratory as a model protein for the effect of the environment on protein folding, stability, solubility, and dynamics. The bacteria grow in media that are nearly saturated with salt, mainly sodium chloride, and their cytoplasm is also nearly saturated with salt, mainly potassium chloride. Halophilic proteins have a high content of acidic amino acid residues, which is related to their instability in the absence of high salt concentrations. Ions of high charge density are the most efficient for stabilizing the protein at low salt (Ebel et al., 1999a). Anion binding was observed on the very acidic protein through high-resolution structure determination by x-ray crystallography (Richard et al., 2000), and salt binding was inferred from solvation studies performed in NaCl and KCl by various complementary methods (Bonneté et al., 1993; Kernel et al., 1999). Recent results showed that the solvation is very different for different salts (Ebel et al., 1998; C. Ebel, M. Pascu, P. Faou, L. Costenaro, G. Zaccai, in preparation). The weak inter-particle interactions in various salts were characterized by measurements of the second virial coefficients (Ebel et al., 1999b). Conditions of high protein solubility were in general, but not always, associated with positive values of A_2. Values for the second virial coefficient decrease slightly with increasing salt concentration, and values close to zero or negative values were associated with precipitation or crystallization conditions. The best crystals are found after a complex dilution process initiated in a biphasic drop in the presence of 2-methyl-2,4-pentanediol (MPD) and NaCl (Richard et al., 1995; Dym et al., 1995). Previously, we have determined the solvent composition in the environment of the protein during the dilution process of the crystallization drop. By measuring the second virial coefficient by neutron scattering, we have found that crystal formation is correlated with a slow evolution from repulsive to attractive protein-protein interactions (Costenaro et al., 2001). Based on the hypothesis that such an evolution may be promoting crystallization more generally, our aim is now to define the determinants of this behavior.

For this purpose, before a more detailed analysis, it would be very useful to be able to rapidly screen a variety of solvents containing high concentrations of inorganic and organic solutes for their effect on the second virial coefficient. The second virial coefficient can be evaluated from a variety of measurements, which are those able to provide the molar mass in a rigorous way (Eisenberg, 1976, 1981). Among these techniques, many have drawbacks in practice. For example, precise osmotic pressure measurements are very difficult to perform in solvents concentrated in salt, light scattering is very sensitive to sample heterogeneity, and neutron or x-ray scattering experiments require large instrument facilities. Sedimentation equilibrium analytical ultracentrifugation suffers from the requirement that in the presence of salts or organic solvents at high concentrations, due to preferential interactions of the protein with solvent components, the buoyant molar mass (and often also a baseline parameter) needs to be determined from the experiment simultaneously together with the second virial coefficient; this makes a precise determination of A_2 difficult. As an alternative ultracentrifugal method, A_2 can be determined, in theory, from the concentration dependence of the sedimentation and diffusion coefficients, both of which determine the shape of sedimentation boundary profiles in velocity ultracentrifugation. It has been shown that the sedimentation profiles are very sensitive to weak interparticle interactions, most notably through the concentration dependency of s. In particular, repulsive interactions produce retardation and sharpening of the sedimentation profiles, as studied in detail by Yphantis and co-workers (Dishon et al. 1967; Fujita, 1975) and others, and a large literature exists on different traditional and more recent descriptions and methods for their analysis (Fujita. 1975; Dishon et al. 1967; van Holde and Weischet 1978; Behlke and Ristau, 1997; Demeler and Saber, 1998).

In the present study, we have investigated whether weak inter-particle interactions and the second virial coefficient could be evaluated from sedimentation velocity experiments. To allow for a large concentration range, we used the Rayleigh interference optical detection and applied direct boundary modeling by Lamm equation solutions (Schuck, 1998), modified to take into account linear dependencies (k_D and k_s) for D and $1/s$, and algebraic systematic noise decomposition (Schuck and Demeler, 1999). First, simulated curves were used to validate a fitting procedure of the parameters characterizing the particle and its non-ideal behavior. Then, the fitting procedure was applied to experimental sedimentation velocity experiments performed on halophilic malate dehydrogenase at different protein con-

centrations. By using samples of only 105 μl, the total amount of protein was less than 2 mg for each of the three solvent conditions studied. The results were compared with those obtained by dynamic light scattering and small-angle neutron-scattering techniques.

EXPERIMENTAL PROCEDURES AND THEORETICAL BACKGROUND

Preparation of the protein and solvents

Hm MalDH was over-expressed in *Escherichia coli* and purified as previously described (Cendrin et al., 1993), with a final step of gel filtration in 4 M NaCl, and stored as a concentrated sample in this solvent at 4°C. The solutions of *Hm* MalDH in MPD and NaCl were prepared by dialysis before the sedimentation experiments or by weight dilutions followed by ultracentrifugation for a few minutes at 175,000 × *g* for the light-scattering experiments, respectively (these experiments were not performed at constant chemical potential of the solvent components but the difference in *D*-values should be very small). All solvents are buffered with Tris-HCl 50 mM, pH 8. Protein concentration was measured in an UV spectrometer using an extinction coefficient at 280 nm of 0.85 cm^2 mg^{-1}. Density measurements of the solvents were performed using a DMA-58 device (Anton PAAR, Graz, Austria).

Sedimentation velocity experiments and numerical procedure of boundary modeling

The sedimentation velocity experiments were carried out in an analytical ultracentrifuge Optima XLA-I (Beckman, Palo Alto, CA), using a four-hole rotor AN-Ti at rotor speeds of 42,000 rpm or 60,000 rpm and at a temperature of 20°C. We used interference optics for the detection of the protein concentration as a function of radial position and time. The *Hm* MalDH and solvent solutions (105 μl) were filled in 3-mm double-sector centerpieces. (In these cells, the total number of fringes is approximately equal to the loading concentration expressed in mg/ml.) In a second set of experiments, ~300 μl of solution was filled in 12-mm double-sector centerpieces. One hundred acquisitions were performed in intervals of 4–5 min. The calculated loading concentrations measured in fringe units resulting from the data analysis were used to convert k_s and k_D from fringe units into mass units.

Theoretical concentration distributions were calculated by finite element solutions of the Lamm equation (Eq. 4 below) with static or moving frame of reference, as described by Claverie (1976) and Schuck (1998), respectively. The concentration dependence of $s(c)$ and $D(c)$ (Eqs. 8 and 9 below) was introduced locally in each time-step, applying a two-step propagation scheme for improved accuracy (Schuck, 1998). We implemented separately the cases for repulsive ($k_s > 0$, $k_D > 0$) and attractive ($k_s < 0$, $k_D < 0$) interactions. The calculated profiles were compared with those described in the literature (Dishon et al., 1967). It was verified that they exhibit the typical sharpening and retardation and that they approach a time-independent boundary shape in highly non-ideal conditions. Furthermore, as a test for their numerical accuracy we applied the transformation of $[\ln(c/c_0) - (1 + k_s c_p) \times \ln(c_p - c/c_0)]$ versus r^2 (for the case of $k_D = 0$). As described by Yphantis and colleagues, for concentration-dependent sedimentation of the form $s(c) = s°/(1 + k_s c)$, this function transforms the theoretically expected (limiting) steady-state boundary shapes into straight lines with slopes equal to $(0.5\omega^2 s°/D°)[k_s c_p/(1 + k_s c_p)]$. As expected, with increasing non-ideality, the transformed boundaries approached straight lines with slopes close to the theoretical value, although at slightly lower precision than as described in Dishon et al. (1967). For the analysis of experimental profiles, a number of data transformation approaches have been developed previously (Dishon et al., 1967; Fujita, 1975). However, we took advantage of the rapid calculation of numerical solutions and employed direct nonlinear regression of the raw data (see below). Systematic noise components in the interference optical data were calculated algebraically as described previously (Schuck and Demeler, 1999). This approach was incorporated into the software SEDFIT, which is available from the authors on request.

For global modeling, an extended version of SEDFIT was used that provides for several data channels, each containing a series of scans of a sedimentation velocity experiment. For global regression, all experiments were given equal weights. The loading concentrations and the systematic noise parameters were optimized locally to each data channel, whereas $s°$, $D°$, k_s, and k_D were optimized globally.

Dynamic light-scattering experiments

Dynamic light-scattering (DLS) was performed using a DynaPro-801 Instrument (Protein Solutions, Charlottesville, VA) and analyzed using the provided Autopro software. A volume of 7 μl was illuminated at $\lambda = 780$ nm by a 25-mW solid-state laser, and the auto-correlation function of the light intensity scattered at $\theta = 90°$ was measured for 3.3 ms in intervals of ~4 μs. The normalized auto-correlation functions $R(t)$ from typically three sets of ~25 measurements were averaged and analyzed in terms of one or two exponential decays of amplitude c_i and decay constant τ_i, and a baseline b:

$$R(t) = b + \sum c_i e^{t/\tau_i} \tag{1}$$

The diffusion coefficients D_i are calculated from the values of the decay constant:

$$D_i = \tau_i / q^2 \tag{2}$$

where q denotes the scattering vector $q = (4\pi/\lambda) \times n \times \sin(\theta/2)$ with the refractive index n. To correct for the temperature effects, the diffusion coefficients D_{20} were calculated considering the temperature dependency of the viscosity η:

$$D_{20} = D/(293.15/T)(\eta/\eta_{20}), \tag{3}$$

with T denoting the temperature of the solution in Kelvin. Tabulated values for the viscosity and refractive index for water were used.

Theoretical background

The Lamm equation expresses the spatial and temporal evolution of the concentration c of a homogeneous solute in the centrifugal field as a function of its sedimentation and diffusion coefficients (s and D) (Lamm, 1929; Tanford, 1961):

$$\partial c/\partial t = 1/r\partial[(D\partial c/\partial r - cs\omega^2 r)r]/\partial r, \tag{4}$$

with t denoting the time, r the distance to the axis of rotation, and ω the angular velocity. The sedimentation coefficient s is related to the buoyant molar mass $M(\partial\rho/\partial c)_{T,\mu}$, and to the friction coefficient f, through the Svedberg equation (Svedberg and Pederson, 1940):

$$s = M(\partial\rho/\partial c)_{T,\mu_S}/N_A f \tag{5}$$

In this expression, N_A is the Avogadro number, and $(\partial\rho/\partial c)_{T,\mu S}$ is the increment of the solution density ρ at constant temperature T and chemical potential μ_S of solvent components, c the protein concentration in mass unit, and M its molar mass. The increment of density is $(1 - \rho^0 \bar{V})$ in one-component solvents, with ρ^0 denoting the solvent density and $\bar{V}$ the partial specific volume of the macromolecule (Eisenberg, 1976). For an incompressible fluid, the Stokes-Einstein relation

$$D = (\partial\Pi/\partial C)_{T,\mu_S}/N_A f \quad (6)$$

(Pusey and Tough, 1985; Schmitz, 1990; Harding and Johnson, 1985a) relates the diffusion coefficient D to the friction coefficient f and the osmotic susceptibility $(\partial\Pi/\partial C)_{T,\mu_S}$ (C denoting the solute concentration in molar units). The osmotic susceptibility can be related to the chemical potential μ of the macromolecule through the identity $(\partial\Pi/\partial C)_{T,\mu S} = c(\partial\mu/\partial c)_{T,\mu_S}$ and approaches RT in the limit of infinite dilution. In this limit, the diffusion coefficient is $D° = RT/N_A f°$, where R is the gas constant, and the friction coefficient can be calculated via the hydrodynamic radius, R_H, as $f° = 6\pi\eta R_H$. Accordingly, $f°$ is related to the mean size and the shape of the macromolecule. The buoyant molar mass can be obtained from the combination of $s°$ and $D°$:

$$s°/D° = (\partial\rho/\partial c)_{T,\mu_S} M/RT \quad (7)$$

In non-ideal solutions, the macromolecule concentration influences s and D through spatial and velocity correlations between particles. For moderate protein concentrations, s, D, and f can be described in the linear approximations:

$$s/s° = f°/f = (1\text{-}k'_s c + \cdots) = 1/(1 + k_S c + \cdots), \quad (8)$$

$$D/D° = (1 + k_D c + \cdots) \quad (9)$$

At not too high protein concentrations, both variants of Eq. 8 are very similar, and $k_s \approx k'_s$. (An alternative formulation on the basis of $s_{w,20}$ and $D_{w,20}$ values, which explicitly contains terms involving partial specific volumes has been discussed by Harding and Johnson (1985a).) Both k_s and k_D can be either positive or negative, with positive values corresponding to conditions of repulsive interactions and negative values to attractive interactions. The osmotic susceptibility $(\partial\Pi/\partial C)_{T,\mu_S}$ can be expanded in power series of the concentration and related to the second virial coefficient A_2 expressed in ml mol g^{-2}:

$$(\partial\Pi/\partial C)_{T,\mu_S} = RT(1 + 2A_2 Mc + \cdots) \quad (10)$$

It should be noted that A_2 is here the same second virial coefficient that can be obtained from the concentration dependence of the apparent molar mass M_{app} in equilibrium sedimentation or in scattering techniques via the relationship $1/M_{app} = 1/M + 2\,A_2 c + \ldots$. Inserting Eqs. 10 and 8 into 6 and comparison of the terms linear in concentration with Eq. 9 leads to

$$k_D \approx 2A_2 M - k_s \approx 2A_2 M - k'_s. \quad (11)$$

This relationship allows the calculation of the second virial coefficient from k_s and k_D. Inspection of Eq. 11 shows two sources of non-ideality determining the values of k_D: thermodynamic non-ideality, which is related to A_2 and leads to the concentration dependency of the molar mass, and hydrodynamic non-ideality, which is related to that of the frictional coefficient and of the sedimentation coefficient. Because it is the difference between two numbers of the same sign and of similar magnitudes (which is expected if A_2 and k_S are linked to the weak inter-particle interactions), k_D can be expected to be smaller than $2A_2M$ or k_S. Further, in general, k_D, k_S, and $2A_2M$ should have the same sign (theoretical exceptions correspond to small values of A_2, with a mixture of attractive and repulsive interactions (Pusey and Tough, 1985)).

RESULTS AND DISCUSSION

Modeling the effect of non-ideality on the sedimentation boundary

A large number of laboratories have studied the effects of concentration dependence of s on the sedimentation profiles; fewer have considered concentration dependence of D. In the present context, we have reexamined the influence of k_s and k_D on the sedimentation profiles to evaluate their relative contributions to deviations from ideal boundaries and to study the inverse problem if parameter values can be extracted by direct boundary modeling of noisy data.

First, we consider a boundary localized nearly in the middle of the centrifuge cell for a solute that is characterized by $s° = 2.5$ S and $D° = 2.5 \times 10^{-7}$ cm^2 s^{-1}, at a concentration of 10 mg/ml and in the absence of concentration dependence of D. In Fig. 1, *A* and *B*, simulated sedimentation profiles and their radial derivatives are compared, in the ideal case ($k_s = 0$, solid line) or with a k_s value of 10 ml/g (dotted line), which grossly corresponds to our experimental values for the repulsive case (see below). The well-known retardation of the boundary associated to its sharpening is apparent, and opposite effects are observed in the case of attractive interactions corresponding to negative values of k_s (data not shown). It is obvious that extracting a D value from the boundary spreading of concentration-dependent sedimentation profiles without taking into account the inter-particle effects would lead to an incorrect description of the macromolecule. Second, we have studied the effect of k_D on the sedimentation boundaries. As can be seen in Fig. 1 *C*, in a 3-mm cell at protein loading concentrations of >10 mg/ml, the effect of k_D values >1 ml/g already exceeds the instrumental noise. However, in real experiments, we would expect significantly affected boundaries only when k_D is greater than 3 ml/g. The deformations due to the concentration dependence of D is exaggerated in the derivative curve of Fig. 1 *D*, by using a k_D value of 100 ml/g, which is one to two orders of magnitude larger than the values to be expected under most experimental conditions. In this case, we observe a significant deviation of the gradient curve from the Gaussian shape, combined with a slight displacement of its maximum value. For negative k_D values corresponding to attractive conditions, sharpening of the boundary is observed (data not shown).

The inspection of these simulated data suggests that $s°$ and k_s values should be easily extracted, even from a single sedimentation velocity experiment, whereas fits of $D°$ and k_D may be more difficult. Because a global nonlinear regression of several sedimentation velocity experiments is very computationally intensive and difficult to initialize, we studied first how much information can be extracted from single-experiment analyses. We simulated a series of concentration distributions of non-ideal sedimentation with $s° = 2$ S, $D° = 1.5 \times 10^{-7}$ cm^2 s^{-1}, combining concentration dependencies of s and D, with $|k_s| = |k_D| = 12.5$ ml/g, and introduced statistical noise of 0.01 fringes. The $s°$ and $D°$ values are close to our experimental values obtained for *Hm* MalDH in viscous solvents containing a high content of salt or organic solvent. The same procedure was applied with $s° = 7$ S and $D° = 5 \times 10^{-7}$ cm^2 s^{-1}, which approximate the values for the same particle in water. To obtain good starting guesses and avoid local minima in the nonlinear regres-

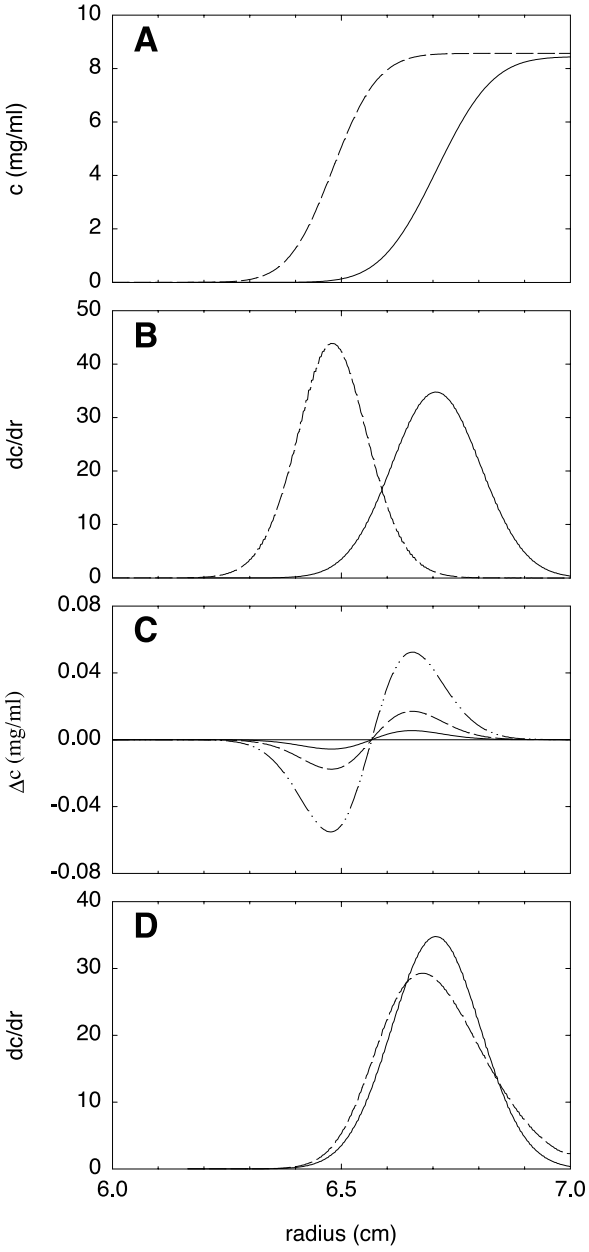


FIGURE 1 Effect of the concentration dependence of sedimentation and diffusion coefficients on sedimentation profiles. Shown are the theoretical concentration distributions of a solute with $s° = 2.5$ S, $D° = 2.5 \times 10^{-7}$ cm^2 s^{-1} and different non-ideality parameters k_s and k_D, calculated for an initially uniform loading concentration of 10 mg/ml, after 17,500 s of sedimentation at a rotor speed of 42,000 rpm. (*A* and *B*) Theoretical boundary profiles (*A*) and radial derivatives (*B*) are compared for ideal solutes (— —) and for concentration-dependent sedimentation with $k_s = 10$ ml/g and $k_D = 0$ (— — —). (*C*) Calculated differences of ideal boundary profiles and those with concentration-dependent diffusion coefficient at $k_s = 0$ and $k_D = 1$ (— —), 3 (— — —), and 10 (— · · —) ml/g. (*D*) Radial derivatives of the sedimentation boundary for $k_s = 0$ and $k_D = 0$ (— —) and $k_D = 100$ ml/g (— — —).

sion, we choose a fitting procedure in three steps. First, the system is modeled as an ideal one, i.e., with the constraint $k_s = k_D = 0$. As a result, apparent values of s and D are obtained. These values are introduced as starting values for the second fit, in which only the concentration dependency of s is considered; i.e., $k_D = 0$, and log k_s is treated as an additional floating parameter, with a starting value of $k_s = 0.01$ ml/g. The third step starts with the values obtained after the second fit, and log k_D is also optimized together with $s°$, $D°$, and k_s, with an initial guess of $k_D = 0.01$ ml/g. This procedure was applied to the case of repulsive interactions (positive coefficients k_s and k_D) and to attractive interactions (negative coefficients). The numerical results of this procedure are given in Table 1. They are qualitatively similar for the two sets of generated data and detailed below for the closest to our experimental system. As can be expected, the first fitting step leads to the apparent s values that are close to the collective calculated s values of 1.78 S and 2.29 S, using Eq. 8, for repulsive and attractive cases, respectively. However, at this stage, the quality of the fit is poor (as judged from both the overall root mean square deviation (rmsd) and the systematic residuals), and the fitted values of the diffusion coefficient do not have a physical meaning. It should be noted that they are not the collective D values of Eq. 9. After the second step, in general, values for $s°$, D, and k_s are obtained that are close to the ones underlying the simulation. The rmsd of the fit is significantly improved, and residuals were nearly randomly distributed. However, we noted that $s°$ and k_s can be slightly correlated; an example of this is the underestimated $s°$ combined with an underestimated k_s shown in Table 1 for the repulsive case. However, this situation was improved after the third step, after which the parameter values converged to those underlying the simulation with a good accuracy. Correspondingly, the boundary profiles were well described, and for the repulsive case, a further significant drop in the rmsd was observed, indicating that the obtained parameter values are well defined.

These results demonstrate that already a single set of sedimentation profiles should contain sufficient information to obtain good estimates of $s°$, $D°$, k_s, and k_D. At first, this may seem surprising, as qualitatively, k_s and k_D modulate the shape of the sedimentation boundary in opposite ways. For example, in the case of repulsion, there is a strong sharpening of the boundary associated with the fact that s decreases with c and a slight broadening of the boundary related to the fact that D increases with c. However, it can be easily demonstrated that there are very significant quantitative differences. If we model the data generated with $k_s = k_D = 12.5$ ml/g with only a concentration dependency of the diffusion coefficient, we obtain a quality of fit similar to the impostor ideal model. Likewise, when we generated data with $k_s = 0$ and $k_D = 12.5$ ml/g, modeling with either the ideal model, or with $k_s \neq 0$ and $k_D = 0$, provides the same quality of fit (rmsd = 0.014 and the same systematic

TABLE 1 Modeling of simulated concentration distributions in non-ideal sedimentation

Input	Fitting steps	s (S)		D (10^{-7} cm^{2}s^{-1})		k_s (ml/g)		k_D (ml/g)		rmsd (fringes)	
		rep	attr	rep	attr	rep	attr	rep	attr	rep	attr
$s°$ = 2 S $D°$ = 1.5 × 10^{-7} cm^{2}s^{-1}, $\|k_s\| = \|k_D\|$ = 12.5 ml/g	Step 1: $k_s = k_D = 0$	1.79	2.26	0.98	2.68					0.037	0.049
	Step 2: $k_s \neq 0$, $k_D = 0$	1.85	2.02	1.61	1.49	4	−11.8			0.028	0.014
	Step 3: $k_s \neq 0$, $k_D \neq 0$	2.00	2.02	1.54	1.55	13.2	−11.9	11.5	−9	0.010	0.011
$s°$ = 7 S $D°$ = 5 × 10^{-7} cm^{2}s^{-1}, $\|k_s\| = \|k_D\|$ = 12.5 ml/g	Step 1: $k_s = k_D = 0$	6.3	7.9	3.0	9.5					0.068	0.078
	Step 2: $k_s \neq 0$, $k_D = 0$	7.1	7.2	5.6	5.7	14	−10			0.012	0.024
	Step 3: $k_s \neq 0$, $k_D \neq 0$	7.0	7.1	5.2	5.7	13	−10	10	−10	0.011	0.018

Concentration distributions were calculated for a rotor speed of 42,000 rpm, based on an initial concentration of 10 mg/ml, with k_s and k_D positive in the repulsive case (rep) and negative in the attractive one (attr). Normally distributed noise of 0.01 was added to the distributions. The fit was performed on 30 or 20 radial profiles taken in time intervals of 500 or 600 s after an initial sedimentation of 10,000 or 600 s, for $s°$ of 2 S and 7 S data, respectively.

distribution of residual), but when we allowed both k_s and k_D to float in modeling these data, the parameters converged to k_s = 0.3 ml/g and k_D = 10 ml/g, similar to the input parameters (rmsd = 0.010 and unsystematic residuals). Thus, when considering the evolution of the sedimentation boundary, k_s and k_D have distinctly different quantitative effects that can be distinguished by direct boundary modeling. The confidence in these parameters can be further improved by global modeling of experiments at different loading concentrations and rotor speeds, as will be shown in the following.

Sedimentation analysis of halophilic malate dehydrogenase in complex solvents

The modeling strategy outlined above was applied to the sedimentation profiles of concentrated halophilic malate dehydrogenase (*Hm* MalDH) at three different protein concentrations and in three different solvents at pH 8. Under the conditions used in these studies, *Hm* MalDH behaved as a stable tetramer, in agreement with previous sedimentation velocity experiments (Hecht and Jaenicke, 1989). Although some dissociation was observed in sedimentation equilibrium at pH 7 (Bonneté et al., 1993), we have seen no evidence of this at pH 8, which is consistent with a clear transition in the protein stability between pH 7 and pH 8 observed earlier (Madern and Zaccai, 1997). Therefore, we can apply the single-component analysis developed above to tetrameric *Hm* MalDH as an experimental test.

In a first approach, 9–10 scans taken in intervals of 40 or 50 min were analyzed. As an example, Fig. 2 shows the experimental and best-fit distributions, and the differences between them (residuals), for a protein loading concentration of 12 mg/ml in 4 M NaCl. Even when neglecting non-ideality (Fig. 2, *A* and *B*), the experimental curves are not too badly described by the model. Considering non-ideality in the second and third fitting step improves, however, the quality of the fit as can be judged from the decreased absolute values and the more random distribution of the residuals (Fig. 2, *C* and *D*). Table 2 gives the detailed

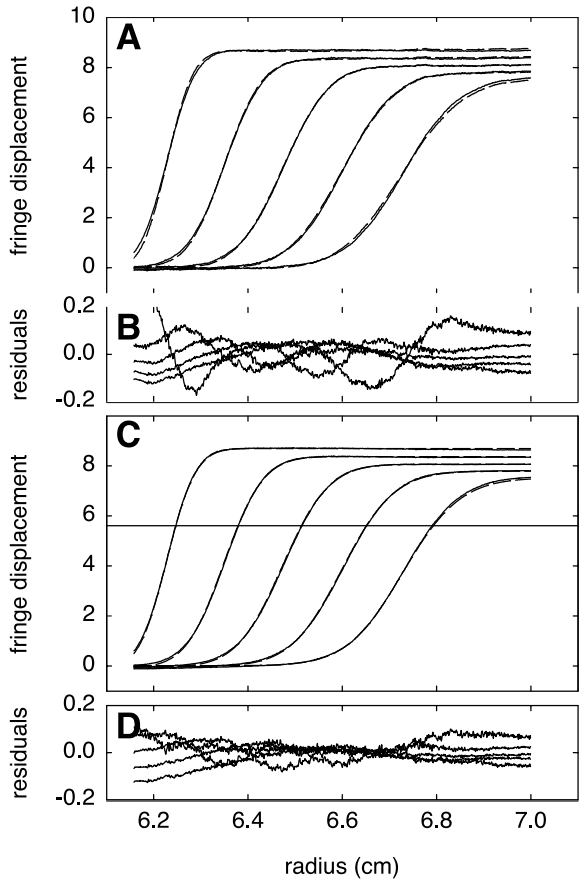


FIGURE 2 Modeling of the sedimentation profiles of *Hm* MalDH at 12.3 mg/ml in 4 M NaCl. (*A* and *C*) Experimental fringe displacement distributions (— —) and best-fit distributions (— — —) from sedimentation velocity performed at 42,000 rpm at 20°C in 3-mm cells; (*B* and *D*) Residuals. Ideal sedimentation modeling ($k_s = k_D = 0$) on *A* and *B* resulted in an s value of 2.06 S and an apparent D-value of 2.0 × 10^{-7} cm^2 s^{-1}, with a rmsd of 0.059 fringe (see text). Non-ideal modeling ($k_s \neq 0$,) on *C* and *D*, resulted in $s°$ = 2.35S, $D°$ = 3.2 × 10^{-7} cm^2 s^{-1}, k_s = 12 ml g^{-1}, and k_D = 1 ml g^{-1} with a rmsd of 0.039 fringe. For clarity, only every second profile that was used in the analysis is shown.

TABLE 2 Modeling of sedimentation velocity profiles of *Hm* MalDH

	Ideal fit				Non-ideal fit with $k_s \neq 0$				Non-ideal fit with $k_s \neq 0$, $k_D \neq 0$				
	c (mg/ml)	s (S)	D_{app} (10^{-7} cm^{2}s^{-1})	rmsd	$s°$ (S)	D (10^{-7} cm^{2}s^{-1})	k_s (ml/g)	rmsd	$s°$ (S)	$D°$ (10^{-7} cm^{2}s^{-1})	k_s (ml/g)	k_D (ml/g)	rmsd
4 M NaCl	12.3	2.06	2.0	0.059	2.37	3.2	13	0.038	2.35	3.2	12	1	0.039
	4.43	2.26	2.1	0.043	2.34	3.1	9	0.033	2.34	3.1	9	2.5	0.033
	2.21	2.36	3.1	0.013	2.40	3.4	11	0.014	2.39	3.2	8	2.5	0.014
Mean					2.37 ± 0.02	3.2 ± 0.1	11 ± 1		2.36 ± 0.01	3.1 ± 0.1	10 ± 2	2 ± 1	
Mean (10 sequential profiles)					*2.35 ± 0.02*	*2.9 ± 0.1*	*8 ± 1.5*		*2.36 ± 0.01*	*2.9 ± 0.1*	*8 ± 2*	*2 ± 1*	
Global modeling									2.42 ± 0.01	3.4 ± 0.1	14 ± 1	0 ± 1	
Global modeling of duplicate experiment									2.43	3.0	12	1	
Global modeling, 12-mm cells									2.41	3.2	9	0	
Global modeling, 12-mm cells, 60,000 rpm									2.41	2.9	11	1	
Global modeling, 12-mm cells, 60,000 rpm and 40,000 rpm									2.42	3.1	11	0	*0.074*
5% MPD + 2 M NaCl	7.29	3.62	2.4	0.034	3.81	3.2	8	0.032	3.81	3.2	8	8	0.032
	3.51	3.72	2.9	0.026	3.76	3.2	4	0.026	3.76	3.2	4	8	0.026
	1.46	3.75	3.1	0.010	3.77	3.2	6	0.010	3.77	3.1	6	8	0.010
Mean					3.78 ± 0.02	3.2 ± 0.1	6 ± 1		3.78 ± 0.03	3.2 ± 0.1	6 ± 2	8 ± 1	
Mean (10 sequential profiles)					*3.8 ± 0.02*	*3.0 ± 0.1*	*7 ± 1*		*3.81 ± 0.03*	*3.0 ± 0.1*	*7 ± 2*	*5 ± 4*	
Global modeling									3.83 ± 0.01	3.2 ± 0.1	9 ± 1	0 ± 5	
30% MPD + 1.5 M NaCl	6.87	1.96	2.3	0.092	1.78	1.3	−15	0.087	1.78	1.3	−15	−6	0.087
	3.39	1.84	1.3	0.062	1.80	1.2	−7	0.062	1.79	1.6	−9	−7	0.062
	1.46	1.82	1.4	0.030	1.79	1.4	−10	0.030	1.79	1.4	−9	−7	0.030
Mean					1.79 ± 0.01	1.3 ± 0.1	−11 ± 2		1.78 ± 0.01	1.4 ± 0.1	−11 ± 2	−7 ± 1	
Mean (10 sequential profiles)					*1.68 ± 0.04*	*0.8 ± 0.2*	*−22 ± 6*		*1.70 ± 0.03*	*0.9 ± 0.1*	*−17 ± 3*	*−2 ± 1*	
Global modeling									*1.72 ± 0.01*	*1.1 ± 0.1*	*−18 ± 1*	*−3 ± 2*	

Experimental fringe displacement profiles of *Hm* MalDH sedimenting in 3 mm cells at 42,000 rpm were analyzed, unless noted otherwise. Time intervals of the scans used for analysis was ~45 min. (regular type) or 4 to 5 min. (italics), respectively. Global modeling was performed using only every 5th or 10th scans for each cell.

outputs obtained after each of the three fitting steps, for three protein concentrations and three different solvents. In 4 M NaCl, fitting with the ideal model (step 1) gives sedimentation coefficients that decrease with increasing protein concentration, a behavior that is commonly observed and typical of overall repulsive interactions. Consistent with the simulation studies described above, the second step reduces the rmsd significantly, except for the more diluted sample. It should be noted that the sedimentation coefficients $s°$ corresponding to infinite dilution, the diffusion constants D, and the k_s values obtained after this second step are similar for each dilution. Allowing for a concentration dependency in D in the third step also provides similar values for k_D for all experiments. However, this last step does not improve the quality of the fit significantly. This indicates that the k_D values could be poorly determined.

Using this approach, a difficulty was encountered in the analysis of complex solvents near the phase separation, such as salt-MPD mixtures at high solute concentrations (Richard et al., 1995; Pittz and Timasheff, 1978). When the solvent composition is close to the coexistence curve between one-liquid and two-liquid phases in the phase diagram, we observed strong distortions in the interference optical signals, in particular, in the proximity of the meniscus, which are probably related to solvent redistribution at the interfaces. These distortions were reproducibly observed in the present study for 1.5 M NaCl, 30% MPD (and were also observed for mixtures of ammonium sulfate and MPD for compositions close to phase separation; data not shown). As a consequence, modeling the complete sedimentation process (Fig. 3, *A* and *B*) does not lead to very good fits, despite the algebraic calculation of the time-invariant and radial-invariant systematic noise components of the interference signal. As an alternative approach, we analyzed the data with more closely spaced scans (4–5-min intervals) after the boundary was localized in the middle of the cell, thereby avoiding the perturbed region near the meniscus. As can be seen in Fig. 3, *C* and *D*, these boundaries were significantly better described by the model. The resulting best-fit parameters with this approach are given in Table 2 (data in italics). For each of the three solvents, similar values for $s°$, $D°$, and k_s were obtained for the three protein concentrations. For 1.5 M NaCl, 30% MPD, the numerical values differ slightly from the analysis of scans with larger time intervals but remain in the same order of magnitude. For the experiments in 4 M NaCl and in 2 M NaCl, 15% MPD, far from phase separation, the best-fit values are very similar to those from larger time intervals. This indicates that this approach can also be successful, even though the improvement of the quality of the fit from step to step is much smaller.

It is clear from inspection of Table 2 that for all of the three investigated solvents, self-consistent results are obtained, with very similar values for the fitted parameters independent of the selection of the boundary profiles and

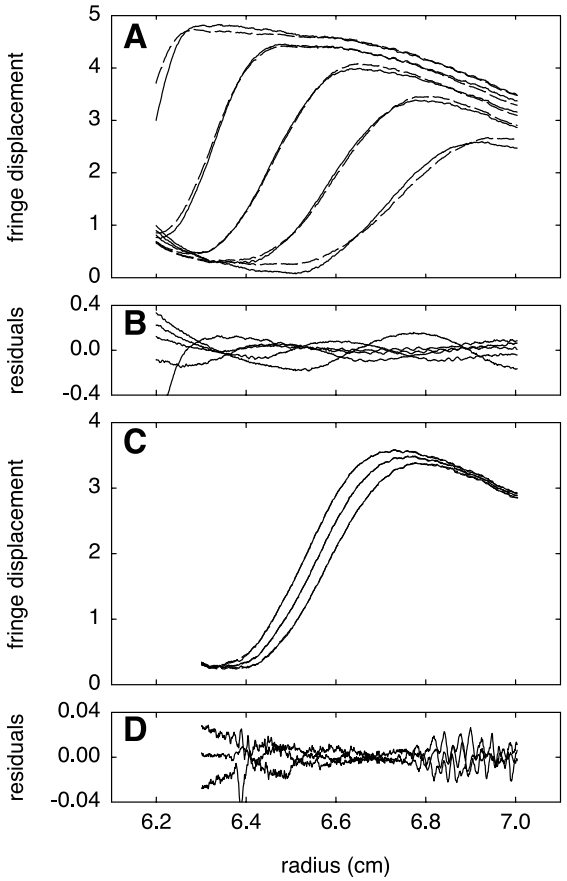


FIGURE 3 Non-ideal modeling of the sedimentation profiles of *Hm* MalDH at 6.9 mg/ml in 1.5 M NaCl, 30% MPD. (*A* and *C*) Experimental fringe displacement distributions (— —) and best-fit distributions of the non-ideal model with $k_s \neq 0$, $k_D \neq 0$ (— — —) from sedimentation velocity performed at 42,000 rpm at 20°C in 3-mm cells; (*B* and *D*) Residuals. On *A* and *B*, the analysis of nine profiles with large time intervals (every second profile shown) that resulted in $s° = 1.78$ S, $D° = 1.3 \times 10^{-7}$ cm^2 s^{-1}, $k_s = -15$ ml g^{-1}, and $k_D = -6$ ml g^{-1} with a rmsd of 0.087 fringe. On *C* and *D*, the analysis of 10 closely spaced scans (showing scans 1, 5, and 9) that resulted in $s° = 1.65$ S, $D° = 0.7 \times 10^{-7}$ cm^2 s^{-1}, $k_s = -23$ ml.g^{-1}, and $k_D = -2$ ml g^{-1} with a rmsd of 0.006 fringe.

loading concentrations. This suggests that the model used for the description of the boundary profile is essentially correct. A global fitting procedure was thus applied to determine more precisely the four global parameters ($s°$, $D°$, k_s, and k_D) characterizing the transport of the non-ideal material.

The same general procedure can be followed as described above. From local ideal modeling of the profiles corresponding to each protein concentration, we can estimate quickly if the inter-particle effects are large, and repulsive or attractive. Further, we can define the starting conditions for non-ideal modeling from the more diluted sample. Then,

TABLE 3 Thermodynamic and hydrodynamic parameters of *Hm* MalDH from modeling of the non-ideal sedimentation compared with other methods

	$s°$ (S)	$D°$ (10^7 cm^{2}s^{-1})	$M(\partial\rho/\partial c)_{T,\mu_S}$ (kg/mol)	$M(\partial\rho/\partial c)_{T,\mu_S}$ in NaCl (kg/mol)	$\rho°$ (g/ml)	k_s (ml/g)	k_D (ml/g)	$2A_2M$ ($k_s + k_D$) (ml/g)	$2A_2M$ (SANS) (ml/g)
4 M NaCl, non-ideal boundary model from	2.41	3.1	21	21	1.153	12	0	13	9
apparent $s(c)$ and $D(c)$	2.41	2.6; 3.1	23			14	3	17	
5% MPD, 2 M NaCl, non-ideal boundary	3.83	3.2	29	32	1.079	9	0	9	7
model from apparent $s(c)$ and $D(c)$	3.8					8			
30% MPD, 1.5 M NaCl, non-ideal boundary	1.72	1.1	37	39	1.047	−18	−3	−21	−26
model from apparent $s(c)$ and $D(c)$	1.73	1.5	28			−13	−2	−15	

Values are given for the direct global boundary modeling (reproduced from Table 2), from the extrapolation of the concentration dependence of s values from a series of traditional sedimentation velocity analyses, and from the concentration dependence of the diffusion coefficients by dynamic light scattering (Fig. 4; or from Pittz and Timasheff, 1978). The two columns for the buoyant molar masses $M(\partial\rho/\partial c)_{T,\mu_S}$ are from combining $s°$ and $D°$ and from combining density increments $(\partial\rho/\partial c)_{T,\mu_S}$ measured by Bonneté et al. (1993) in NaCl solutions with the molar mass of a stable *Hm* MalDH tetramer, respectively (see text). In the two last columns, the second virial coefficient terms $2A_2M$ from boundary modeling are compared with those obtained from SANS by Costenaro et al. (2001), in 4 M NaCl, 5% MPD, 1.5 M NaCl, and 30% MPD, 1.5 M NaCl, respectively.

the quality of the modeling is estimated globally, from the value of the global reduced χ^2. The precision of the global parameters for each set of experimental data was estimated by varying the initial conditions of the fit. The values obtained for $s°$, $D°$, and k_s are similar to those derived from individual data sets. Values for k_D are in general smaller and appear as nearly indistinguishable from instrumental noise, with absolute value of k_D equal or above 3 ml/g.

We also studied the reproducibility of the global modeling results for different experimental conditions. The results obtained with data from three loading concentrations of *Hm* MalDH in 4 M NaCl acquired using either 3- or 12-mm-thick cells, at 42,000 or 60,000 rpm are given in Table 2. Similarly, a global modeling of data from different rotor speeds and loading concentrations was performed. Although these measurements were performed on different *Hm* MalDH preparations, the parameters are very consistent. $D°$ is obtained with a precision of 10%. However, the residuals are not always randomly distributed, as can be seen on Figs. 2 and 3, which could be attributed, in part, to the remaining systematic noise from the data acquisition. On the other hand, it should be considered that the additional constraints imposed by global analysis necessarily leads to an increase in the rms error, and that rms error in the global fit of 0.07 still represents a fit with a relative error of only 0.4%, due to the high loading concentrations. To study the possibility of systematic errors from modeling, we compared the best-fit results with those obtained from other techniques. This comparison is summarized in Table 3.

Comparison of the parameters for the ideal particle the non-ideality coefficients k_s and k_D with those from other experimental approaches

An alternative, more traditional analysis of the sedimentation coefficients is an initial interpretation of the sedimentation data with a model for an ideal apparent particle, followed by the regression of the determined sedimentation coefficients $s(c)$ as a function of concentration according to Eq. 8. Although this approach requires several experiments at different concentrations, it also can estimate the extrapolated sedimentation coefficient at infinite dilution, $s°$, and the non-ideality coefficient k_s. This approach is shown in Fig. 4 *A*. It results in $s°$ values that are very similar to those

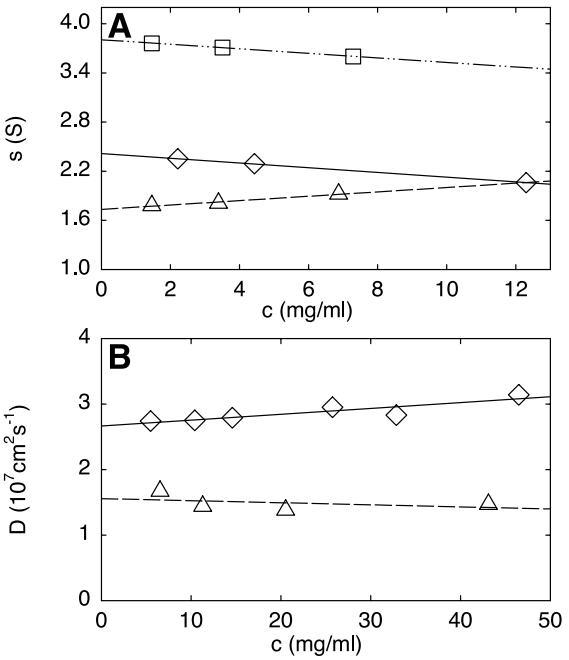


FIGURE 4 Linear regression of the mean sedimentation and diffusion coefficients measured as a function of *Hm* MalDH concentration. (*A*) Sedimentation coefficients s evaluated from modeling of sedimentation profiles as a single ideal component; (*B*) Diffusion coefficients D_{20} from dynamic light scattering. Solvents were 4 M NaCl (◇), 5% MPD + 2 M NaCl (□), and 30% MPD + 1.5 M NaCl △.

from non-ideal boundary modeling and also in k_s values of similar magnitude (see Table 3).

An analogous procedure can be applied to the concentration dependence of the collective diffusion coefficient measured by the exponential decay of the autocorrelation function from DLS (see Experimental Procedures and Theoretical Background above). For weakly interacting particles, the collective diffusion coefficient in DLS is the same as measured by sedimentation experiments if the characteristic distance $2\pi/q$ is much larger than the mean inter-particle distance and if the characteristic decay times of the autocorrelation function τ are much larger than the time scale of the macromolecular interactions τ_I (Pusey and Tough, 1985). These conditions are fulfilled in our experiments. $2\pi/q \approx 3000$ Å, which is large compared with the inter-macromolecular distances (~600 Å and 130 Å for protein concentration of 1 and 100 mg/ml, respectively). For the protein in 4 M NaCl, for example, we measured a characteristic decay time τ of 70 μs, whereas the maximal value for τ_I (estimated as the time needed for a protein to freely diffuse the average inter-macromolecular distance, $L^2/D°$) for a protein concentration of 1 mg/ml in the same solvent would be ~20 μs and much smaller in more concentrated solutions.

Light-scattering experiments were performed for *Hm* MalDH concentrations in the range 5–50 mg/ml in two solvents. The auto-correlation functions were described well by a single-exponential decay in 4 M NaCl and a double-exponential decay in 1.5 M NaCl, 30% MPD (the slower decay with $D \approx 0.4 \times 10^{-7}$ cm^2 s^{-1} is attributed to the solvent fluctuations). Typical baseline values were 1.01–1.02, which is reasonable compared with the theoretical expectation value of unity. Values of $D°$ and k_D were obtained assuming a linear concentration dependency of D according to Eq. 9 (Fig. 4 *B*) and are listed in Table 3. Unfortunately, we found that D values in the mixed NaCl-MPD solvent are not very precise, and large errors are associated with the estimated k_D in the two solvents. However, the k_D values determined by DLS are in qualitative agreement with those determined from non-ideal modeling of the sedimentation boundary profiles, despite the large relative error in both experiments (see above). When comparing the $D°$ values from non-ideal boundary modeling, from DLS and from literature values (Pundak and Eisenberg, 1981), we found similar, although not identical, values. It should be noted, however, that the two types of experiments were not performed on the same samples and that slight variations in the solvent composition are possible that may affect diffusion coefficients.

A further possibility for independent validation of the $s°$ and $D°$ values from non-ideal modeling of the sedimentation boundary represents the comparison of buoyant molar mass values $M(\partial\rho/\partial c)_{T,\mu_S}$, which is determined by the ratio of $s°$ and $D°$ (Eq. 7). In the present study, the most important aspect in confirming the buoyant molar masses is the evaluation of the $D°$ estimate, as this parameter is derived solely from the shape of the sedimentation boundary. In the case of a complex solvent, the density increment $(\partial\rho/\partial c)_{T,\mu_S}$ is determined not only by the solvent density but also by the interactions between the macromolecule and the solvent components (Eisenberg, 1976; Ebel, 1995). We calculated $M(\partial\rho/\partial c)_{T,\mu_S}$ on the basis of tetrameric *Hm* MalDH of molar mass 130.5 kDa from the experimental values of $(\partial\rho/\partial c)_{T,\mu_S}$ measured by Bonneté et al. (1993). Because the $(\partial\rho/\partial c)_{T,\mu_S}$ values were obtained from density measurements on the protein in NaCl solutions, we interpolated the data to solvents of a density equal to the density of our complex solvents used for the sedimentation experiments. This results in buoyant molar mass values that are in good agreement for all three solvents (Table 3), which suggests the reliability of our $D°$ estimate from non-ideal modeling of the boundary shape. Interestingly, for the solvent close to phase separation (30% MPD and 1.5 M NaCl), the determination of $D°$ from the non-ideal sedimentation boundary appears to be superior to the estimate from DLS. However, one important consideration in this respect could be that in sedimentation experiments, both $s°$ and $D°$ are determined strictly in the same buffer, a feature that could be important, in particular, in the context of studies in complex solvents.

By comparing the $(\partial\rho/\partial c)_{T,\mu_S}$ values in NaCl and mixed NaCl-MPD solvents, we assume that the density of the solvated protein, i.e., of its solvation shell, is not much affected by the presence of MPD. This hypothesis is reasonable, even if MDP is excluded from the proximity of the protein, as described (in the absence of salt) for bovine pancreatic ribonuclease A from differential refractometry (Pittz and Timasheff, 1978). Changes in the water and MPD concentrations in the proximity of the protein are not expected to affect significantly the values of the buoyant molar mass of the protein, because the partial specific volumes of water and MPD are very similar. For this reason, and because the buoyant molar masses are estimated here only with a precision of ~10–15%, our measurements will not be discussed in terms of protein-solvent interactions.

From the comparisons above, we can conclude that the non-ideal boundary modeling of sedimentation profiles at high protein concentration in the presence of inter-particle interactions allows a good estimate of the properties of the ideal particle, in terms of both its sedimentation and diffusion coefficient.

Comparison of the second virial coefficients with SANS and with theory

The second virial coefficient can be calculated from k_s and k_D using Eq. 11. The resulting values, expressed as $2A_2M$, are listed in Table 3. They can be compared with the values measured previously from small-angle neutron scattering (SANS) (Ebel et al., 1999b; Costenaro et al., 2001). (The values in the solvent for 5% MPD was reported by SANS

for 1.5 M NaCl, which we expect to be very close to the slightly higher NaCl concentration of 2 M used in the present study.) For each of the solvent conditions, the second virial coefficient values obtained from modeling of the sedimentation boundary profiles and from SANS experiments are of the same order of magnitude and of reasonably good agreement, which seems sufficient for our goal of screening different solvent conditions for approximate values and changes in the sign of the virial coefficient.

In the present study, the values of k_s, k_D, and $2A_2M$ have the same sign and k_D was smaller than k_s. This suggests that a qualitative investigation of the whole particle interactions may be made already from the value of k_s alone. Because the sedimentation profiles are highly disturbed by the concentration dependence of the sedimentation coefficient, the determination of k_s is very robust, in non-ideal boundary modeling or even simply from the evaluation of the apparent s value at different protein concentrations. This simplification could be justified theoretically on the basis that k_D should be small compared with k_s for non-interacting spheres, because the collective diffusion coefficient is determined by partial cancellation of the thermodynamic factor $(\partial\Pi/\partial C)_{T,\mu_s}$ and the hydrodynamic ones (friction coefficient) (see Eqs. 6 and 11). An exception would be the case, for example, of particles interacting through long-range direct interactions, such as charged macromolecules in pure water (Pusey and Tough, 1985; Retailleau et al., 1997). Also, k_D/k_s ratios between 0.3 and 0.6 were reported from experimental studies with turnip-yellow-mosaic virus in non-dissociating solvents (Harding and Johnson, 1985b), and ratios of 1 or larger were estimated for γ-crystallins in attractive or repulsive regimes of inter-particle interactions (Bonneté et al., 1997).

The small absolute values for k_D found in the range 0–3 ml/g from global direct modeling of the sedimentation profiles were confirmed by DLS, in our attractive and repulsive conditions. Clearly, global fitting procedure avoided an overestimation of k_D. As a consequence, a very good correlation is found in Table 3 for k_S, k_D, and $2A_2M_2$ obtained through different techniques. It should be noted that, in 4 M NaCl, the values $k_D = 0$–3 ml/g and $k_s \approx 2A_2M_2 \approx$ 12–13 ml/g are compatible with the theoretical ratios k_D/k_S and $k_D/2A_2M_2$ predicted for non-interacting hard spheres, which are 0.25 and 0.2, respectively (Pusey and Tough, 1985). In this solvent, it is reasonable to consider this protein as a non-interacting hard sphere, because globular *Hm* MalDH is very soluble (above 200 mg/ml) and electrostatic interactions can be neglected in 4 M NaCl. Furthermore, for non-interacting hard spheres, the virial term $2A_2M$ is related to v_s, the swollen specific volume, or R, the corresponding radius according to $2A_2M = 8v_s = (N_A/M)32/3\pi R^3$ (Tanford, 1961). Using our $2A_2M$ values of 13 ml/g or 17 ml/g in 4 M NaCl, we can calculate radii of 44 or 48 Å, which are close or slightly larger than the radius of 40 Å estimated from the largest dimension of the globular *Hm* MalDH in the high-resolution structure (Dym et al., 1995). At higher MPD concentration, finally, we found the negative value of $2A_2M$ of ~ -20 ml/g ($A_2 = -8 \times 10^{-5}$ ml mol g^{-2}) in 30% MPD, 1.5 M NaCl, which is indicative of moderately attractive intermolecular interactions described to be favorable to protein crystallization (George and Wilson, 1994). This is in agreement with the fact that the composition of this solvent is close to that in which our protein crystallizes nicely (Richard et al., 1995; Costenaro et al., 2001).

CONCLUSION

In the present study, we have shown that sedimentation velocity combined with direct boundary modeling using the Lamm equation for non-ideal sedimentation can be used to investigate weak interactions. Although the concept of positive and negative virial coefficients for expressing global attraction and repulsion is not appreciated by many traditional biochemists, this approach can describe both attractive and repulsive interactions, with positive and negative values for k_s, k_D, and A_2, respectively. Many authors describe attractive interactions as specific contacts between well-defined macromolecular species and, accordingly, relate non-ideality only to intermolecular repulsive contributions, i.e., excluded volume and electrostatic repulsion (Behlke and Ristau, 1999). With these assumptions, the second virial coefficient is always positive and can be calculated considering the shape, hydration, and charge of the macromolecules, as was done, for example, considering modeled ellipsoids (Harding et al., 1999). A low or negative value, corresponding to attractive conditions, is then interpreted in terms of macromolecular association related to an equilibrium between different oligomeric forms of the protein. However, in the case of weak attractive interactions, the proposed association process is generally very complex, because it is characterized by a large number of polymeric species, and corresponding association constants can hardly be discriminated (Behlke and Knespel, 1996; Winzor and Wills, 1994; Wills et al., 1980). It is then a fruitful approach to describe weak interactions in undersaturated protein solutions in terms of simple potentials between modeled homogeneous species. For example, x-ray scattering curves for lysozyme in various conditions have been modeled considering three simple spherical potentials, and reasonable parameters for the protein diameter and charge and attractive potential depth and range has allowed the description of macromolecular phase diagrams, pH effects, and nonspecific salt effects (Malfois et al., 1996; Ducruix et al., 1996; Tardieu et al., 1999). In these descriptions, specific solute-solvent interactions have not been modeled. However, more complex potentials, in which, for example, the solvent is not a continuum, or with a more precise description of the protein, are being studied for improving the description of inter-particle interactions and for investigat-

ing the specific effects of salts on crystallogenesis (Georgalis and Saenger, 1999; Piazza, 1999; Belloni, 1991).

The understanding of the effects of solvent composition on global inter-particle interactions is a first step for the understanding of protein crystallization processes in complex solvents. We have shown here that sedimentation velocity can be a good technique for this purpose. The protocol applied uses less than two recoverable milligrams of protein for each solvent condition. Complex solvents can be investigated, containing high concentrations of salt and organic compounds close to phase separation. By global direct boundary modeling of interference optical ultracentrifuge data combined with systematic noise decomposition, one can take advantage of the high sensitivity and the high dynamic range of the detection system. Because $s°$ and $D°$ values are obtained with a relatively good precision, this procedure should be generally useful for the characterization of the molar masses and hydrodynamic properties of asymmetric or charged solutes displaying large non-ideality. As discussed above, both attractive and repulsive interactions can be described, with positive and negative values for k_s, k_D, and A_2, respectively. Importantly, the second virial coefficient can be determined with an accuracy that appears sufficient to screen different solvent conditions for changes of globally repulsive to globally attractive interactions. In the present study, we observed such a change in the sign of the weak interactions for halophilic malate dehydrogenase in mixed NaCl-MPD, which is clearly related to the capacity of the protein either to remain soluble or to crystallize in solution.

REFERENCES

Behlke, J., and A. Knespel. 1996. Observation of precrystallisation aggregation in protein solutions during centrifugation. *J. Crystal Growth*. 158:388–391.

Behlke, J., and O. Ristau. 1997. Molecular mass determination by sedimentation velocity experiments and direct fitting of the concentration profiles. *Biophys. J*. 72:428–34.

Behlke, J., and O. Ristau. 1999. Analysis of the thermodynamic non-ideality of proteins by sedimentation equilibrium experiments. *Biophys. Chem*. 76:13–23.

Belloni, L. 1991. Interacting monodisperse and polydisperse spheres. *In* Neutron, X-Ray and Light Scattering: Introduction to an Investigative Tool for Colloidal and Polymeric Systems. Elsevier Science Publishers, New York. 135–155.

Bonneté, F., C. Ebel, H. Eisenberg, and G. Zaccai. 1993. Biophysical study of halophilic malate dehydrogenase in solution: revised subunit structure and solvent interactions in native and recombinant enzyme. *J. Chem. Soc. Faraday Trans*. 89:2659–2666.

Bonneté, F., S. Finet, and A. Tardieu. 1999. Second virial coefficient: variations with lysozyme crystallization conditions. *J. Crystal Growth*. 196:403–414.

Bonneté, F., M. Malfois, S. Finet, A. Tardieu, S. Lafont, and S. Veesler. 1997. Different tools to study interaction potentials in γ-crystallin solutions: relevance to crystal growth. *Acta Cryst*. D53:438–447.

Cendrin, F., J. Chroboczek, G. Zaccai, H. Eisenberg, and M. Mevarech. 1993. Cloning, sequencing, and expression in *Escherichia coli* of the gene coding for malate dehydrogenase of the extremely halophilic archaebacterium *Haloarcula marismortui*. *Biochemistry*. 32: 4308–4313.

Claverie, J. M. 1976. Sedimentation of generalized systems of interacting particles. III. Concentration-dependent sedimentation and extension to other transport methods. *Biopolymers*. 15:843–57.

Costenaro, L., G. Zaccai, and C. Ebel. 2001. Understanding protein crystallisation by dilution: the ternary NaCl-MPD-H_2O system. *J. Crystal Growth*. 232:102–113.

Demeler, B., and H. Saber. 1998. Determination of molecular parameters by fitting sedimentation data to finite-element solutions of the Lamm equation. *Biophys. J*. 74:444–54.

Dishon, M., G. H. Weiss, and D. A. Yphantis. 1967. Numerical solutions of the Lamm equation. III. Velocity centrifugation. *Biopolymers*. 5:697–713.

Ducruix, A., J. P. Guilloteau, M. Ries-Kautt, and A. Tardieu. 1996. Protein interactions as seen by solution x-ray scattering prior to crystallogenesis. *J. Crystal Growth*. 168:28–39.

Dym, O., M. Mevarech, and J. L. Sussman. 1995. Structural features that stabilize halophilic malate dehydrogenase from an *archaebacterium*. *Science*. 267:1344–1346.

Ebel, C. 1995. Characterisation of the solution structure of halophilic proteins. Analytical centrifugation among complementary techniques (light, neutron and x-ray scattering, density measurements). *Prog. Colloid Polymer Sci*. 99:17–23.

Ebel, C., P. Faou, B. Franzetti, B. Kernel, D. Madern, M. Pascu, C. Pfister, S. Richard, and G. Zaccai. 1998. Molecular interactions in extreme halophiles: the solvation-stabilization hypothesis for halophilic proteins. *In* Microbiology and Biogeochemistry of Hypersaline Environments. CRC Press, Boca Raton, FL. 227–237.

Ebel, C., P. Faou, B. Kernel, and G. Zaccai. 1999a. Relative role of anions and cations in the stabilization of halophilic malate dehydrogenase. *Biochemistry*. 38:9039–9047.

Ebel, C., P. Faou, and G. Zaccai. 1999b. Protein-solvent and weak protein-protein interactions in halophilic malate dehydrogenase. *J. Crystal Growth*. 196:395–402.

Eisenberg, H. 1976. Biological macromolecules and polyelectrolytes in solution. Clarendon Press, Oxford.

Eisenberg, H. 1981. Forward scattering of light, x-rays and neutrons. *Q. Rev. Biophys*. 14:141–172.

Fujita, H. 1975. Foundations of Ultracentrifugal Analysis. J. Wiley and Sons, New York.

Georgalis, Y., and W. Saenger. 1999. Light scattering studies on supersaturated protein solutions. *Sci. Prog*. 82:271–294.

George, A., Y. Chiang, B. Guo, A. Arabshahi, Z. Cai, and W. W. Wilson. 1997. Second virial coefficient as predictor in protein crystal growth. *Methods Enzymol*. 276:100–110.

George, A., and W. W. Wilson. 1994. Predicting protein crystallization from a dilute solution property. *Acta Cryst*. D50:361–365.

Guo, B., S. Kao, H. McDonald, A. Asanov, L. L. Combs, and W. W. Wilson. 1999. Correlation of second virial coefficients and solubilities useful in protein crystal growth. *J. Crystal Growth*. 196:424–433.

Harding, S. E., J. C. Horton, S. Jones, J. M. Thornton, and D. J. Winzor. 1999. COVOL: an interactive program for evaluating second virial coefficients from the triaxial shape or dimensions of rigid macromolecules. *Biophys. J*. 76:2432–2438.

Harding, S. E., and P. Johnson. 1985a. The concentration-dependence of macromolecular parameters. *Biochem. J*. 231:543–7.

Harding, S. E., and P. Johnson. 1985b. Physicochemical studies on turnip-yellow-mosaic virus: homogeneity, relative molecular masses, hydrodynamic radii and concentration-dependence of parameters in non-dissociating solvents. *Biochem J*. 231:549–555.

Hecht, K., and R. Jaenicke. 1989. Malate dehydrogenase from the extreme halophilic archaebacterium *Halobacterium marismortui*: reconstitution of the enzyme after denaturation and dissociation in various denaturants. *Biochemistry*. 28:4979–4985.

Kernel, B., G. Zaccai, and C. Ebel. 1999. Determination of partial molal volumes, and salt and water binding of highly charged biological mac-

romolecules (tRNA, halophilic protein) in multimolar salt solutions. *Prog. Colloid Polymer Sci.* 113:168–175.

Lamm, O. 1929. Die Differentialgleichung der Ultrazentrifugierung. *Ark. Mat. Astr. Fys.* 21B:1–4.

Madern, D., and G. Zaccai. 1997. Stabilisation of halophilic malate dehydrogenase from *Haloarcula marismortui* by divalent cations: effects of temperature, water isotope, cofactor and pH. *Eur. J. Biochem.* 249: 607–611.

Malfois, M., F. Bonneté, L. Belloni, and A. Tardieu. 1996. A model of attractive interactions to account for fluid-fluid phase separation of protein solutions. *J. Chem. Phys.* 105:3290–3300.

MacMillan, W. G., and J. E. Mayer. 1945. The statistical thermodynamics of multicomponent systems. *J. Chem. Phys.* 13:276–305.

Piazza, R. 1999. Interactions in protein solutions near crystallisation: a colloid physics approach. *J. Crystal Growth.* 196:415–423.

Pittz, E. P., and S. N. Timasheff. 1978. Interaction of ribonuclease A with aqueous 2-methyl-2,4-pentanediol at pH 5.8. *Biochemistry.* 17:615–623.

Pundak, S., and H. Eisenberg. 1981. Structure and activity of malate dehydrogenase from the extreme halophilic bacteria of the Dead Sea. I. Conformation and interaction with water and salt between 5 M and 1 M NaCl concentration. *Eur. J. Biochem.* 118:463–470.

Pusey, P. N., and J. A. Tough. 1985. Particle interactions. *In* Dynamic Light Scattering: Applications of Photon Correlation Spectroscopy. Plenum Press, New York. 85–179.

Retailleau, P., M. Ries-Kautt, and A. Ducruix. 1997. No salting-in of lysozyme chloride observed at low ionic strength over a large range of pH. *Biophys. J.* 73:2156–2163.

Richard, S., F. Bonneté, O. Dym, and G. Zaccai. 1995. Protocol 21: the MPD-NaCl-H_2O system for the crystallization of halophilic proteins. *In* Archaea: A Laboratory Manual. Cold Spring Harbor Laboratory Press, New York. 149–153.

Richard, S. B., D. Madern, E. Garcin, and G. Zaccai. 2000. Halophilic adaptation: novel solvent protein interactions observed in the 2.9 and 2.6 Å resolution structures of the wild type and a mutant of malate dehydrogenase from *Haloarcula marismortui*. *Biochemistry.* 39:992–1000.

Rowe, A. J. 1992. The concentration dependence of sedimentation. *In* Analytical Ultracentrifugation in Biochemistry and Polymer science. The Royal Society of Chemistry, Cambridge, UK. 394–406.

Schmitz, K. S. 1990. An Introduction to Dynamic Light Scattering by Macromolecules. Academic Press, San Diego.

Schuck, P. 1998. Sedimentation analysis of noninteracting and self-associating solutes using numerical solutions to the Lamm equation. *Biophys. J.* 75:1503–1512.

Schuck, P., and B. Demeler. 1999. Direct sedimentation analysis of interference optical data in analytical ultracentrifugation. *Biophys. J.* 76: 2288–2296.

Svedberg, T., and K. O. Pederson. 1940. The Ultracentrifuge. Oxford University Press, London.

Tanford, C. 1961. Physical Chemistry of Macromolecules. Wiley, New York.

Tardieu, A., A. Le Verge, M. Malfois, F. Bonneté, S. Finet, M. Riès-Kautt, and L. Belloni. 1999. Proteins in solution: from x-ray scattering intensities to interaction potentials. *J. Crystal Growth.* 196:193–203.

van Holde, K. E., and W. O. Weischet. 1978. Boundary analysis of sedimentation velocity experiments with monodisperse and paucidisperse solutes. *Biopolymers.* 17:1387–1403.

Wills, P. R., L. W. Nichol, and R. J. Siezen. 1980. The indefinite self-association of lysozyme: consideration of composition-dependent activity coefficients. *Biophys. Chem.* 11:71–82.

Winzor, D. J., and P. R. Wills. 1994. Omega analysis and the characterization of solute self-association. *In* Modern Analytical Ultracentrifugation. Birkhäuser, Boston. 66–80.

4. INTERACTIONS PROTEINE – PROTEINE EN DIFFERENTS SELS

Rester active, stable, soluble et flexible en présence de très fortes concentrations de sels sont les défis que doit surmonter une protéine halophile. L'activité de la *Hm* MalDH est en partie conditionnée par sa stabilité qui dépend des interactions de la protéine avec le solvant [Ebel *et al.*, 1999a]. Parallèlement, ces mêmes interactions devraient aussi influencer les propriétés de solubilité de la protéine. Les interactions avec le solvant mesurées pour la *Hm* MalDH native sont très différentes selon la nature du sel. Dans le cadre du modèle de la particule invariante, la protéine en solution est associée à de grandes quantités d'eau et de sel en $MgCl_2$, à deux fois moins d'eau en NaCl et à peu de sel en $(NH_4)_2SO_4$ [Bonneté *et al.*, 1993; Ebel, en préparation]. Les interactions protéine – protéine en solution sont-elles en relation avec ces comportements différents ? Les résultats publiés à ce jour pour la *Hm* MalDH montrent effectivement de grandes différences selon la nature et la concentration des sels qui restent cependant incomprises [Ebel *et al.*, 1999b].

Mais que connaît on des relations entre les interactions protéine – solvant et les interactions protéine – protéine ? Concernant la classification des ions salins par rapport à leur efficacité à faire précipiter les protéines (séries d'Hofmeister) [voir par exemple Riès-Kautt et Ducruix, 1997], il existe de nombreuses études qui expliquent ces relations par différents mécanismes. Un mécanisme proposé est l'hydratation préférentielle de la protéine : elle est entourée d'une couche plus riche en eau que le solvant de masse et a alors tendance à précipiter afin de la réduire. Ce mécanisme peut engendrer une baisse de sa solubilité et dans des conditions favorables conduire à sa cristallisation [Timasheff, 1993]. De manière plus spécifique, la formation de paires d'ions à la surface d'une protéine peut aussi modifier les interactions protéine – protéine [Riès-Kautt et Ducruix, 1991]. L'évolution des interactions faibles protéine – protéine dans différents sels est très clairement révélée par des expériences de diffusion aux petits angles qui permettent d'observer puis de modéliser les facteurs de structure de solutions concentrées en protéine [Tardieu *et al.*, 1992; Ducruix *et al.*, 1996]. Dans certaines conditions de solvant, une composante attractive dépendante du sel et de sa concentration devient nécessaire pour modéliser les interactions protéine – protéine du lysozyme et des α et γ-cristallines [Finet, 1999; Tardieu *et al.*, 1999]. La modélisation ne permet cependant pas de préciser l'origine de cette attraction.

Le travail que nous allons présenter ici a consisté à approfondir la caractérisation des interactions protéine – protéine de la *Hm* MalDH afin d'apprécier dans quelle mesure elles sont influencées par les interactions protéine – solvant. La *Hm* MalDH nous semble en effet être à cette fin un excellent modèle. Cette enzyme halophile dont la surface est extrêmement chargée négativement est en effet relativement bien caractérisée dans de nombreux solvants. Elle montre des solvatations différentes selon la nature du sel, mais qui sont très riches à la fois en eau et en sel par rapport aux protéines non-halophiles. De plus, cette étude pourrait dégager de nouvelles explications à l'adaptation moléculaire aux milieux extrêmement salés.

Les seconds coefficients du viriel A_2 de la *Hm* MalDH ont été décrits pour divers sels [Ebel *et al*., 1999b]. Nous avons confirmé certains résultats et étendu nos investigations à d'autres sels et dans certains cas à d'autres conditions de température et de pH afin de cerner les paramètres physico-chimiques déterminants. Sauf mention contraire, toutes les expériences de diffusion ont été réalisées dans des conditions de solvant, de pH et de température où la *Hm* MalDH est sous forme native, *i.e*. active et stable [Bonneté *et al*., 1994; Madern et Zaccai, 1997; Ebel *et al*., 1999a]. Lorsque la stabilité n'était pas connue, des mesures d'activité résiduelle (activité après 18-24 heures d'incubation) et de fluorimétrie ont été préalablement réalisées selon le protocole décrit dans [Ebel *et al*., 1999a]. L'activité (ou l'inactivité) de l'enzyme est en effet liée à son repliement correct (ou à son dépliement). L'ensemble des conditions étudiées est listés en annexe (cf. §8.2, p.190)[*] .

Nous avons utilisé les techniques de diffusion de neutrons et de rayons X aux petits angles. La simple comparaison des courbes de diffusion (normalisées, notamment à c, cf. §2.2 p.50) obtenues dans un même solvant à différentes concentrations en protéine permet de déduire directement le caractère globalement répulsif ou attractif des interactions protéine – protéine. Les seconds coefficients du viriel A_2 qui quantifient ces interactions sont dérivés de l'approximation de Guinier aux très petits angles (cf. §2.3.1 p.58). De plus, on peut déterminer les facteurs de structure S qui traduisent la distribution spatiale des macromolécules en solution et qui sont informatifs à forte concentration de protéine (plusieurs dizaines de mg/ml). Pour préciser la nature des interactions interparticulaires, la modélisation numérique des S expérimentaux est effectuée en ajustant les paramètres du

[*] Tous les solvants sont tamponnés par du Tris 50 mM, à l'exception des solvants à pH 6-6.5, tamponnés par du Mes 10-50 mM. Par la suite, seul le pH sera précisé et son omission signifiera un pH de 8.

potentiel DLVO utilisé afin que les facteurs de structure calculés les décrivent mieux possible (cf. §2.3.2 p.61). Les déterminants des interactions protéine – protéine seront finalement discutés.

4.1. Les courbes expérimentales de diffusion

Les courbes des intensités diffusées par des solutions de *Hm* MalDH sont présentées en Figure 4.1 pour quelques conditions de solvants étudiés, en fonction de la concentration en protéine (c en mg/ml). Les intensités à q = 0 ont été extrapolées selon l'approximation de Guinier (cf. (2.5) p.52) et les courbes à concentration nulle correspondent aux facteurs de forme extrapolés (cf. Figure 2.5 p.67).

Pour chaque courbe de diffusion, nous avons déterminé en utilisant l'approximation de Guinier, les valeurs apparentes du rayon de giration R_g et de l'intensité diffusée à angle nul I(q=0,c)/c. L'intensité des courbes extrapolées à concentration nulle de protéine varie selon le solvant puisque le contraste de la protéine varie avec la composition du solvant. Pour chaque condition de solvant, la valeur du rayon de giration $R_g°$ extrapolée à concentration nulle de protéine est de l'ordre de 30-32 Å, valeur qui correspond au tétramère de la *Hm* MalDH [Ebel *et al*., 1999b]. De plus toutes les courbes de diffusion (pour un solvant donné) se rejoignent aux grandes valeurs de q : la structure de la protéine reste la même à toutes les concentrations. Nos objets diffusants sont donc bien les molécules de la protéine sous forme native tétramérique et non des agrégats ou des formes dissociées.

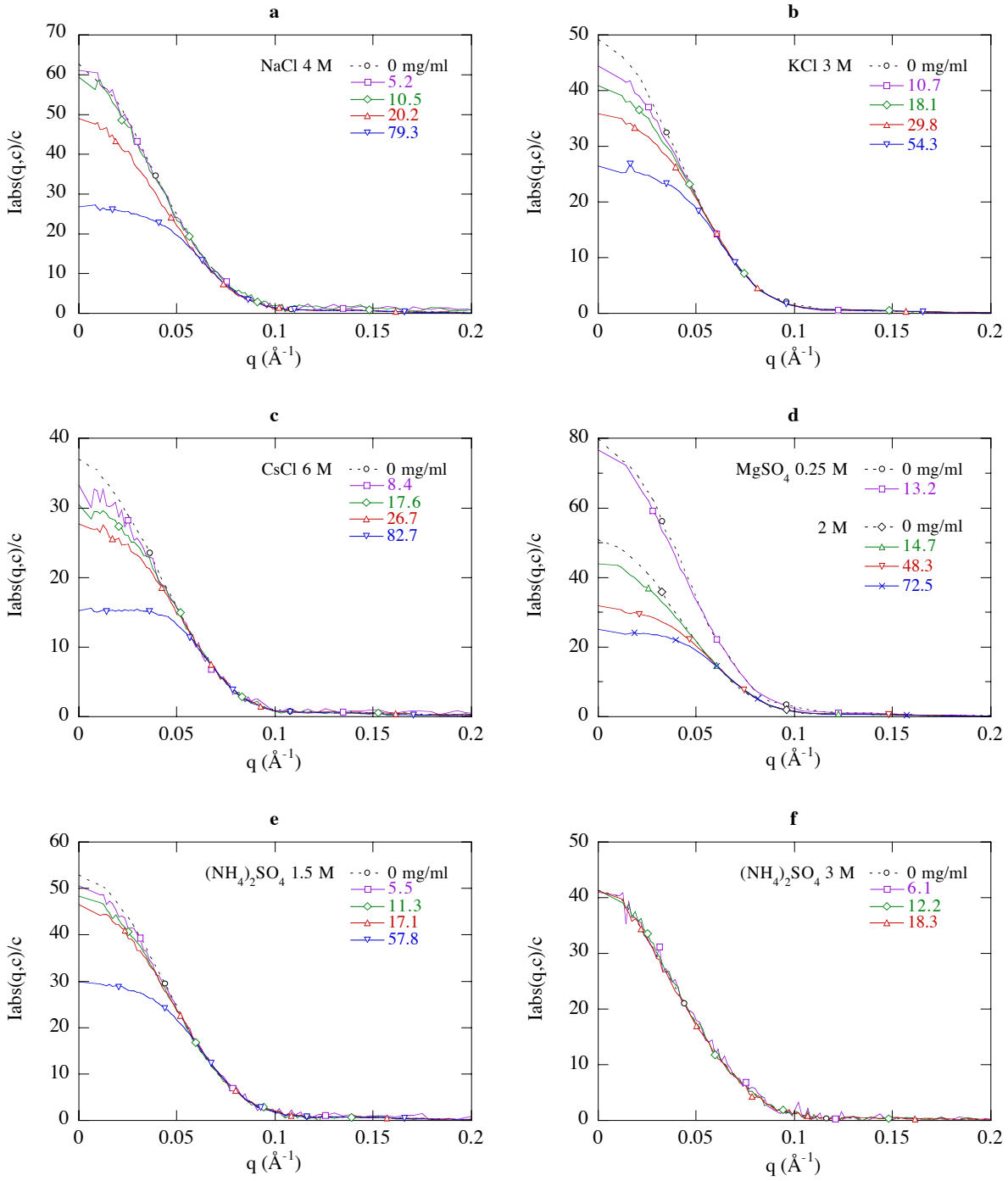


Figure 4.1 : Courbes de diffusion de neutrons par des solutions de *Hm* MalDH en différents sels, en fonction de la concentration en protéine c en mg/ml, à 20-25°C et pH 8-8.2. **(a)** NaCl 4 M (réf. D22 00.4)* ; **(b)** KCl 3 M (réf. D11 98.2) ; **(c)** CsCl 6 M (réf. D22 99.9) ; **(d)** $MgSO_4$ 0.25 et 2 M (réf. D11 98.7) ; **(e-f)** $(NH_4)_2SO_4$ 1.5 et 3 M (réf. D22 00.10). Les courbes à concentration nulle correspondent aux facteurs de forme. Ils ont été déterminés par l'extrapolation linéaire des Iabs(q,c)/c, pour chaque valeur de q, sur les trois plus faibles concentrations en protéine, puis faiblement lissés. Les points à q = 0 sont extrapolés selon l'approximation de Guinier.

* Les références sont sous le format " instrument année.mois " et permettent une recherche dans la liste des conditions des expériences de diffusion en annexe 8.2 p.190.

D'une manière très générale, on constate que les courbes de diffusion normalisées à c dépendent de cette concentration en protéine pour une condition de solvant donné : il ne s'agit donc pas de solutions idéales, mais de solutions où les protéines sont en interactions. Les intensités normalisées à c décroissent aux petits angles lorsque c augmente dans le cas d'interactions protéine – protéine répulsives et inversement augmentent dans le cas d'interactions attractives. On constate ainsi sur la Figure 4.1 (a à e) que les conditions salines NaCl 4 M, KCl 3 M, CsCl 6 M, $MgSO_4$ 2 M et $(NH_4)_2SO_4$ 1.5 M correspondent à des interactions globalement répulsives. Nous verrons que pour NaCl et CsCl elles restent répulsives pour toutes les molarités étudiées.

Les interactions protéine – protéine sont plus diverses en sulfate d'ammonium. Répulsives à 1.5 M, elles deviennent en effet presque nulles à 3 M : les intensités Iabs(q,c)/c évoluent très peu avec c (Figure 4.1-f). Les particules ne semblent plus interagir. La solution peut être alors qualifiée de quasi idéale et le solvant est dit " thêta ". En $MgSO_4$, l'évolution des interactions protéine – protéine en fonction de la concentration en sel est inversée : elles sont répulsives à haut sel (2 M) et globalement neutres à 0.25 M (Figure 4.1-d). A bas sel, les courbes de diffusion évoluent peu avec la concentration en protéine. Le même phénomène avait été observé en $MgCl_2$ [Ebel *et al.*, 1999b] et nous avons confirmé ces résultats : les interactions sont répulsives à 1-1.3 M et neutres à 0.2 M.

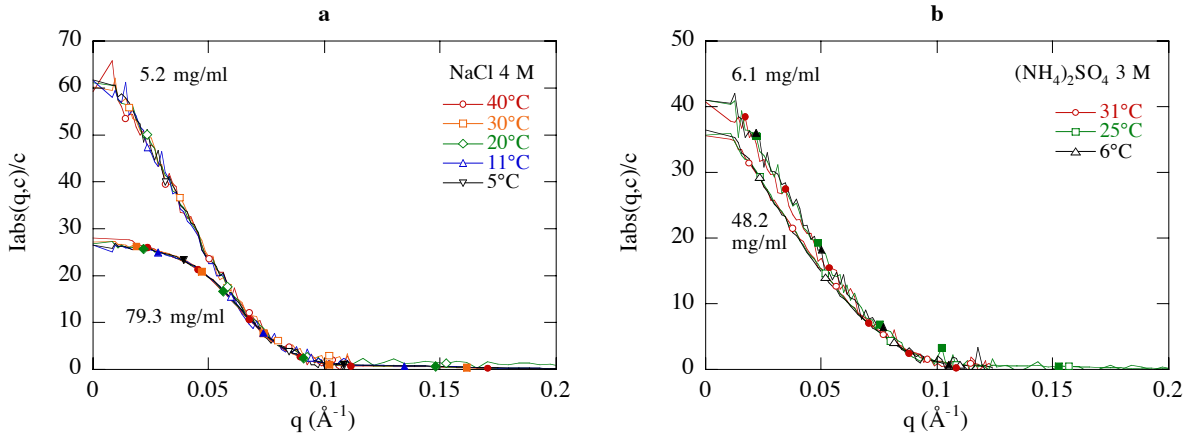


Figure 4.2 : Courbes de diffusion de neutrons par des solutions de *Hm* MalDH en fonction de la température, en : **(a)** NaCl 4 M pH 8 (réf. D22 00.4) ; **(b)** $(NH_4)_2SO_4$ 3 M pH 8 (réf. D22 00.10).

Nous avons effectué quelques mesures en fonction de la température en NaCl 4 M, en $(NH_4)_2SO_4$ 1.5-3 M et en $MgCl_2$ 0.2-1.3 M. Comme le traduit l'évolution faible ou nulle des courbes de diffusion (Figure 4.2), la température n'a pas d'effet en NaCl et très peu en $(NH_4)_2SO_4$, comme en $MgCl_2$ (données non montrées).

4.2. Les seconds coefficients du viriel A_2

L'évolution des courbes d'intensité diffusée par des solutions à différentes concentrations en protéine dans un même solvant permet de qualifier globalement les interactions interparticulaires de répulsives ou d'attractives, mais la comparaison entre diverses conditions de solvant est essentiellement qualitative. Les seconds coefficients du viriel A_2 quantifient l'intensité de ces interactions en les intégrant sur tout l'espace (cf. (2.12) p. 58) et révèlent ainsi les différences entre les conditions. Ils dérivent des régressions linéaires des intensités diffusées à angle nul pour une même condition de solvant comme expliqué précédemment (cf. §2.3.1 p.58). Nous les avons déterminés à partir des trois solutions aux plus faibles concentrations en protéine (inférieures à 30 mg/ml). Rappelons qu'une valeur positive de A_2 correspond à des interactions globalement répulsives et, inversement, une valeur négative à des interactions globalement attractives.

4.2.1. Evolution de A_2 en fonction de la nature et de la concentration du sel

La Figure 4.3 présente les valeurs de A_2 pour la *Hm* MalDH en fonction de la nature du sel et de sa concentration. Les salinités explorées (à l'exception des points isolés) correspondent globalement aux concentrations en sel accessibles, restreintes par la stabilité du sel et de la protéine (excepté $(NH_4)_2SO_4$ 4 M et $MgSO_4$ au-dessus de 2 M) [Ebel *et al.*, 1999a]. Toutes les mesures ont été effectuées à pH 8-8.2 et à 20-25°C ou à 5°C pour les sels signalés d'un astérisque [d'après Ebel *et al.*, 1999b]. Ces dernières valeurs peuvent toutefois être comparées aux autres puisque nous avons vu que l'effet de la température est extrêmement ténu pour ces sels. Nous le constatons d'ailleurs en $(NH_4)_2SO_4$ où les valeurs aux deux températures s'accordent bien compte tenu de leur précision. Cette observation souligne d'ailleurs la reproductibilité des résultats d'une expérience à l'autre, même avec des préparations différentes de protéine. Notons toutefois que les A_2 proches de 0 sont moins bien

définis car ils correspondent à des intensités qui n'évoluent quasiment pas avec la concentration en protéine. L'erreur portée sur la Figure 4.3 déterminée lors de la régression linéaire est sans doute sous-estimée pour ces valeurs, comme en $(NH_4)_2SO_4$ 3 M à 25°C. Nous considérerons par la suite que les deux valeurs de A_2 pour cette condition saline décrivent le même état d'interaction.

On peut distinguer deux classes de solvant. La première est constituée des sels suivants : NaCl, CsCl, KF et $NaCH_3CO_2$ pour lesquels A_2 conserve des valeurs positives proches de 4 10^{-8} mol.l.g^{-2} à toute salinité. Pour ces sels, A_2 diminue d'une manière monotone avec la concentration en sel, en tenant compte de la précision des valeurs.

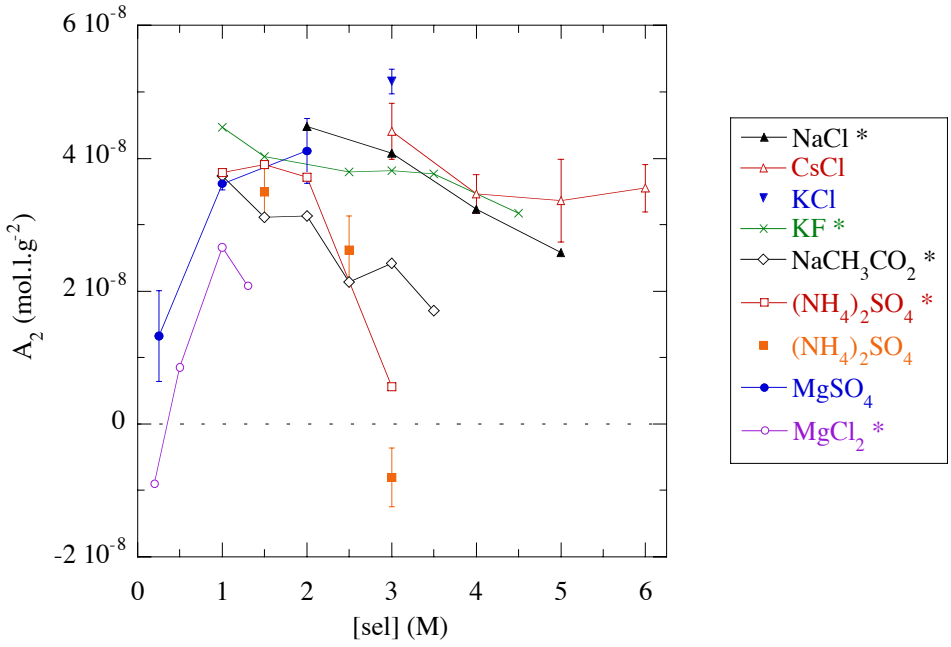


Figure 4.3 : Evolution du second coefficient du viriel A_2 de la *Hm* MalDH en fonction de la nature et de la concentration du sel. Les mesures ont été effectuées à pH 8-8.2 et à 20-25°C (réf. NaCl : D22 98.10, CsCl : D22 99.9, KCl : D11 98.2, $(NH_4)_2SO_4$: D22 00.10, $MgCl_2$ et $MgSO_4$: D11 98.7), ou à 5°C pour les sels marqués d'un astérisque (réf. D11 97.4, d'après [Ebel *et al.*, 1999b]).

En revanche, A_2 évolue fortement en $(NH_4)_2SO_4$, $MgSO_4$ et $MgCl_2$ qui forment la seconde classe de solvants. Il varie de manière non monotone avec la concentration en sel, entre des valeurs nulles ou faiblement négatives et des valeurs positives de l'ordre de 3-4 10^{-8} mol.l.g^{-2}. En $MgCl_2$ et en $MgSO_4$, de façon similaire, A_2 augmente très rapidement pour

se stabiliser à plus forte salinité. A l'opposé, en $(NH_4)_2SO_4$, il est constant entre 1 et 2 M puis diminue subitement au-delà, pour atteindre des valeurs proches de zéro à 3 M. D'ailleurs, la *Hm* MalDH précipite à 100 mg/ml à cette concentration de $(NH_4)_2SO_4$ et à 10 mg/ml à 4 M (ce qui explique l'absence de détermination de A_2 pour cette salinité) [Ebel *et al.*, 1999b].

L'évolution de A_2 en fonction de la salinité peut donc être très différente selon le sel, ce qui traduit clairement une dépendance marquée des interactions protéine – protéine à la composition du solvant.

4.2.2. Effet de charge : pH et mutant de la Hm *MalDH*

Afin d'évaluer l'influence sur les interactions interparticulaires de la charge de la protéine, nous l'avons fait varier en utilisant deux approches. La première, facilement réalisable, est la variation du pH de la solution. Nous avons donc enregistré les courbes de diffusion puis déterminer les A_2 en NaCl 4 M et $MgCl_2$ 0.5 et 1 M à divers pH : 6.5, 8 et 9, pour lesquels la charge structurale calculée de la *Hm* MalDH est environ de -140, -158 et -165 respectivement. Cette gamme de pH est limitée par la plus faible stabilité de la protéine à pH acide [Madern et Zaccai, 1997]. La deuxième approche consistait en l'utilisation du mutant E121K de la *Hm* MalDH où les quatre acides glutamiques 121 du tétramère ont été mutés en lysine [Madern, communication personnelle]. La différence de charge est donc supposée être de +8. Les valeurs de A_2 obtenues pour les différentes conditions sont résumées dans le Tableau 4.1.

Sel	pH 6.5	pH 8-8.2	pH 9
NaCl 4 M	4 ±0.2	3.2 *	3.4 ±0.7
NaCl 4 M (E121K)			4.9 ±0.6
$MgCl_2$ 0.5 M	1.2 ±0.4	0.85 *	
$MgCl_2$ 1 M	3.1 ±0.1	2.7 *	

Tableau 4.1 : Seconds coefficients du viriel A_2 (en 10^{-8} mol.l.g^{-2}) pour la *Hm* MalDH native et un mutant (E121K) en fonction du pH pour différents sels, à 20-25°C (réf : D22 98.10, D11 98.7, D11 98.2) ou à 5°C pour les valeurs marquées d'un astérisque (réf. D11 97.4, d'après [Ebel *et al.*, 1999b]).

Au vu de la précision obtenue, on peut considérer qu'une variation de pH de 6.5 à 8-8.2 ou 9 ne modifie les valeurs de A_2 ni en NaCl 4 M (pour la forme sauvage ou mutée de la

protéine), ni en $MgCl_2$ 0.5 ou 1 M. Dans ces sels, les interactions protéine – protéine ne semblent donc pas être affectées par ces variations de charge de la macromolécule. Cette même variation induit cependant une modification majeure de la stabilité de la *Hm* MalDH : elle est stabilisée à plus bas sel à pH 8 qu'à pH 7 [Madern et Zaccai, 1997]. Cette variation n'impliquerait donc que la structure interne de la protéine.

4.2.3. Effet du D_2O

En diffusion de neutrons, l'eau lourde D_2O (ou ^2H_2O)[*] est souvent utilisée comme solvant principal à cause de la très faible diffusion incohérente de l'atome D par rapport à H ce qui augmente la précision des mesures (pour un même temps de comptage). Nous avons effectué des mesures de diffusion de neutrons pour des solutions de *Hm* MalDH en KCl 3 M en H_2O et en D_2O afin de savoir si cet échange influençait dans notre cas les interactions protéine – protéine. Les valeurs des A_2 obtenus sont de 5.2 (±0.2) 10^{-8} en H_2O et 5.9 (±0.1) 10^{-8} mol.l.g^{-2} en D_2O. Cette différence significative montre que cet échange n'est pas sans conséquence sur les interactions protéine – protéine. Il a été montré que le D_2O modifiait les interactions protéine – protéine et diminuait la solubilité du lysozyme et de l'inhibiteur de la trypsine pancréatique bovine [Gripon *et al*., 1997; Budayova-Spano *et al*., 2000]. La *Hm* MalDH en KCl semble donc être un contre-exemple puisque A_2 – et sans doute la solubilité de la protéine – est plus élevé en D_2O qu'en H_2O. De plus, le D_2O affecte la stabilité des protéines et notamment augmente celle de la *Hm* MalDH à bas sel [Bonneté *et al*., 1994; Madern et Zaccai, 1997]. Comme notre intention est de mettre en relation les interactions interparticulaires et les interactions protéines – solvant qui ont été caractérisées en H_2O, nous avons continué de mener nos expériences de diffusion de neutrons en H_2O, malgré les avantages du D_2O.

[*] ^2H est un isotope de l'atome H. Il possède un neutron supplémentaire et est donc formé d'un proton, d'un neutron et d'un électron.

4.3. Modélisation des facteurs de structure

Le second coefficient du viriel quantifie la résultante de l'ensemble des interactions protéine – protéine (attractives et répulsives) en les intégrant sur tout l'espace. L'intensité diffusée I(q,c) contient cependant plus d'information : via le facteur de structure S(q,c), elle est en effet reliée à la distribution spatiale des particules de protéine dans la solution qui dérive des interactions interparticulaires (cf. (2.9) p.53). Le facteur de structure pourrait donc potentiellement nous renseigner sur la nature des interactions. Nous avons donc, dans un deuxième temps, abordé les interactions protéine – protéine de la *Hm* MalDH par la modélisation des facteurs de structure selon la méthode de simulation numérique décrite au §2.3.2 p.61).

Les facteurs de structures expérimentaux sont (dans la mesure du possible) déterminés en utilisant les facteurs de forme extrapolés à concentration nulle en protéine I(q,c=0) (cf. Figure 2.5 p.67). Nous avons focalisé la modélisation sur les S_{exp} à plus forte concentration en protéine car ils diffèrent le plus de 1 (cas sans interaction) et sont donc plus discriminants. Pour certaines des conditions salines précédentes, nous avons approfondi l'étude pour des concentrations en protéines plus élevées (au-dessus d'environ 20-30 mg/ml) afin de modéliser les S_{exp}. Les résultats sont présentés pour les deux classes de solvants déterminées selon les valeurs des A_2 : NaCl, KCl et CsCl, puis $MgCl_2$, $MgSO_4$ et $(NH_4)_2SO_4$.

La modélisation des facteurs de structure grâce au potentiel DLVO (cf. (2.23) p.64) utilise des paramètres " structuraux " pour décrire les interactions protéine – protéine. Il s'agit du rayon D (en Å) et de la charge effective Z de la protéine (considérée comme quasi-sphérique) qui interviennent dans les potentiels répulsifs, ainsi que la profondeur J (en kT) et la portée d (en Å) de la composante attractive du potentiel. La stratégie que nous avons suivie consiste à partir de la description la plus simple possible du système. Nous avons donc dans un premier temps utilisé les potentiels répulsifs de volume exclu et coulombien[*] et affiné les valeurs des paramètres D et Z afin de rendre compte des facteurs de structures expérimentaux (S_{exp}). Dans le modèle DLVO, cette charge effective qui tient compte de la condensation d'une partie des contre-ions est constante (pour un sel donné). Lorsque la salinité varie, seul le terme d'écrantage est modifié dans le potentiel coulombien (cf. **(1.4)** et **(1.5)** p.28). Dans les cas d'une modélisation inadéquate, nous avons ensuite ajouté le potentiel attractif de type

[*] Les constantes diélectriques utilisées sont détaillées en annexe 8.1 p.187.

Yukawa (J, d) pour lequel nous avons fixé sa portée d à 5 Å et ajusté uniquement J – ce choix est justifié dans la partie méthodologique (cf. p.65).

4.3.1. Deux potentiels répulsifs suffisent-ils en NaCl, KCl et CsCl ?

Les facteurs de structures S(q,c) expérimentaux sont représentés par des " o " dans la Figure 4.4 pour la *Hm* MalDH en NaCl 2-5 M pH 8.2. Nous constatons immédiatement qu'ils sont à l'origine (S(q=0,c)) inférieurs à 1, ce qui est caractéristique d'interactions protéine – protéine répulsives et corrélé aux valeurs positives de A_2. La différence d'amplitude entre les graphiques (a, c) qui correspondent à des concentrations en protéine c proches de 100 mg/ml et les graphiques (b, d) pour lesquels c est de l'ordre de 20 mg/ml illustre bien la dépendance à la concentration en protéine des facteurs de structure et le fait qu'ils soient plus informatifs à forte c et moins précis à faible c.

Aux S(q,c) expérimentaux sont superposés les S(q,c) calculés avec seulement les deux potentiels répulsifs de volume exclu et coulombien dont les valeurs des paramètres D et Z sont inscrits en légende des graphiques de la Figure 4.4. Quoique ces deux potentiels soient très simples et de symétrie sphérique, ils modélisent étonnamment bien les facteurs de structure expérimentaux. Toutefois, la modélisation est peu résolutive quant aux valeurs des paramètres ajustables qui décrivent la protéine.

En faisant varier le diamètre de la particule D de 75 à 85 Å, le meilleur accord avec les S_{exp} est obtenu pour 80 Å, à quelques angströms près. Cette valeur décrit très correctement le comportement des interactions protéine – protéine en NaCl 2 et 4 M (Figure 4.4-a et c) et reste compatible avec les S_{exp} en NaCl 3 et 5 M qui sont moins précis (graphiques b et d).

Les effets de la charge et du volume se manifestent de la même manière dans les S(q), comme on le constate sur les simulations présentées en Figure 4.4-a où les facteurs de structure calculés avec les couples (D, Z) de (75, 156) et (80, 0) sont quasiment identiques. Ceci est lié à la très faible portée de l'interaction électrostatique (λ_D) pour des concentrations élevées de sel : les formes des potentiels coulombien écranté et de volume exclu deviennent peu discernables. La longueur de Debye λ_D (inverse de la constante d'écrantage κ, cf. (2.23) p.64) vaut par exemple 1.8 et 0.9 Å en NaCl 2 et 5 M à 25°C. Par ailleurs, les variations du potentiel électrostatique avec la charge se réduisent avec l'augmentation de la salinité puisqu'il est inversement proportionnel à la force ionique via la constante d'écrantage. On remarque très bien cet effet entre les graphiques (a et c) de la Figure 4.4. En supposant D et Z constant, on déterminerait ainsi plus facilement D à forte salinité et Z à faible salinité.

Néanmoins, même en 2 M de NaCl, nous ne pouvons pas objectivement déterminer la charge effective Z de la protéine (variable qui tient compte de la condensation d'une partie des contre-ions). Les variations induites pour des valeurs de Z de 0 à 156 (valeur équivalant à la charge structurale au pH utilisé) sont effectivement trop faibles et restent de l'ordre de la précision des S_{exp}.

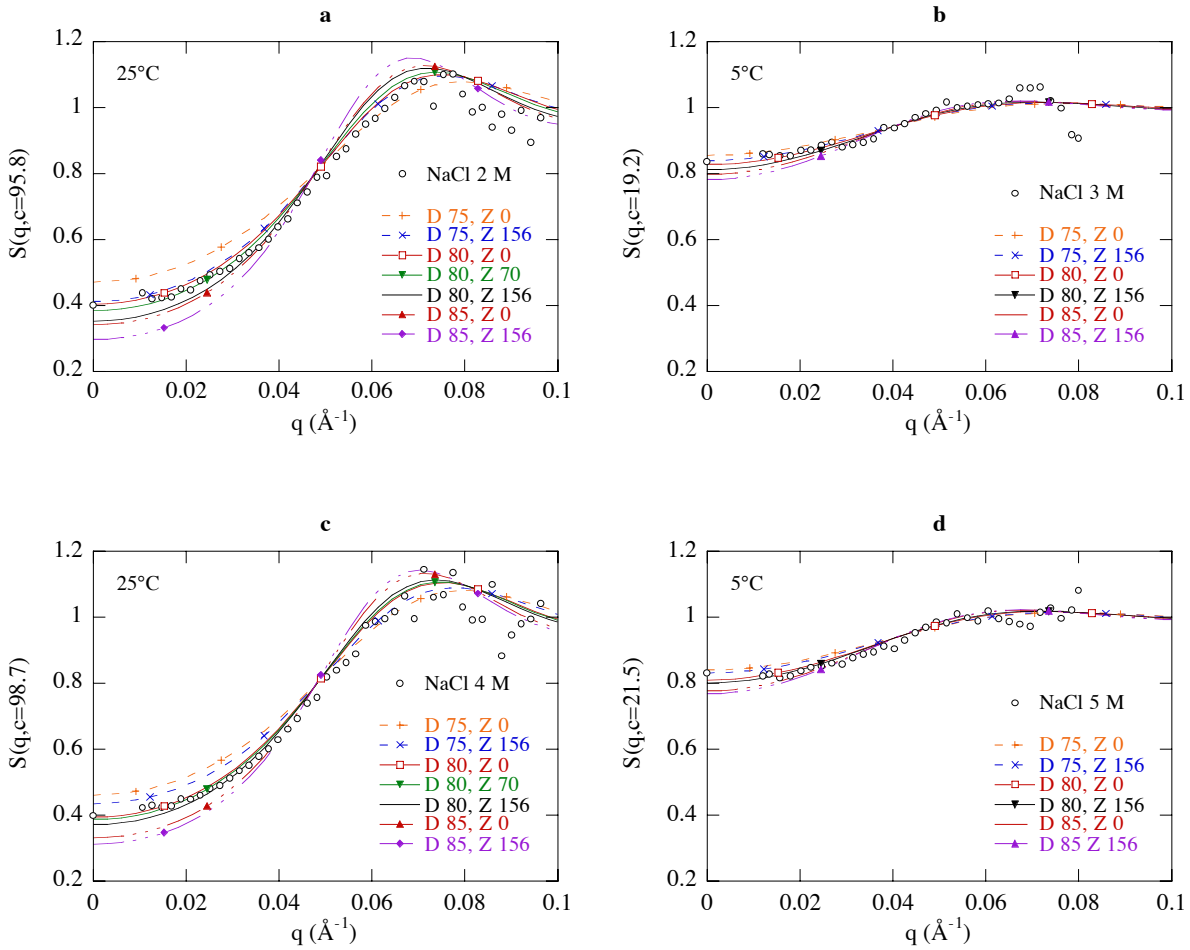


Figure 4.4 : Modélisation des facteurs de structure S(q,c) de la *Hm* MalDH en NaCl pH 8.2 (2 et 4 M, à 25°C : réf. D22 97.8 ; 3 et 5 M, à 5°C : réf. D11 97.4). Les " o " correspondent aux S expérimentaux et les lignes aux S calculés à partir d'un potentiel de volume exclu caractérisé par le diamètre D en Å et d'un potentiel coulombien avec une charge protéique Z. Les concentrations en protéine c sont en mg/ml. En NaCl, la température n'a pas d'effet notable sur les interactions.

Les facteurs de structure de la *Hm* MalDH en CsCl et KCl peuvent également être modélisés avec les deux composante répulsives du potentiel DLVO (Figure 4.5). Les S_{exp} sont cependant moins bien décrits pour les q supérieurs à environ 0.05 $Å^{-1}$, sans doute à cause d'une détermination du facteur de forme de moindre qualité. En CsCl, le diamètre D

correspondant au meilleur ajustement vaut 82 Å comme on peut le voir dans le graphique (d) à forte salinité où les variations des S_{calc} avec la charge sont très faibles. En KCl, un D de 80-82 Å déterminé avec plus de difficultés semble aussi convenir (Figure 4.4-e). Le décalage entre les facteurs de structure expérimentaux et calculés en KCl 3 M à 54.3 mg/ml provient d'une erreur d'acquisition des données pour cet échantillon puisque nous ne retrouvons pas de tel décalage pour une autre concentration en protéine (29.8 mg/ml). Comme précédemment, la charge effective est très difficilement déterminable.

La description très simple que nous avons utilisée permet une modélisation correcte des facteurs de structure, et donc des interactions grâce aux deux potentiels de volume exclu et électrostatique. Nous avons pu ainsi déterminer le diamètre de la particule mais pas sa charge effective à cause des salinités très élevées que nous utilisons. Par ailleurs, il est possible de modéliser les facteurs de structures expérimentaux dans les trois sel précédents par un diamètre plus élevé en ajoutant un potentiel attractif de type Yukawa. Ainsi, en NaCl 4 M, pour un diamètre de 85 Å (au lieu de 80), un potentiel attractif de profondeur 1 kT et de 5 Å de portée permet de rendre compte du facteur de structure. Cependant, l'ajout d'un potentiel attractif aux potentiels de volume exclu et électrostatique n'améliore pas l'accord entre les S expérimentaux et calculé. Nous allons donc considérer que, dans ces solvants, il n'existe pas de potentiel attractif (ou qu'il est trop faible pour que son effet sur les facteurs de structure soit discernable). L'absence d'effet de température en NaCl (entre 5 et 40°C) conforte ce choix puisque, contrairement aux deux autres, le potentiel attractif dépend en effet de la température (nous verrons que l'effet de température est toutefois ténu en conditions modérément attractives).

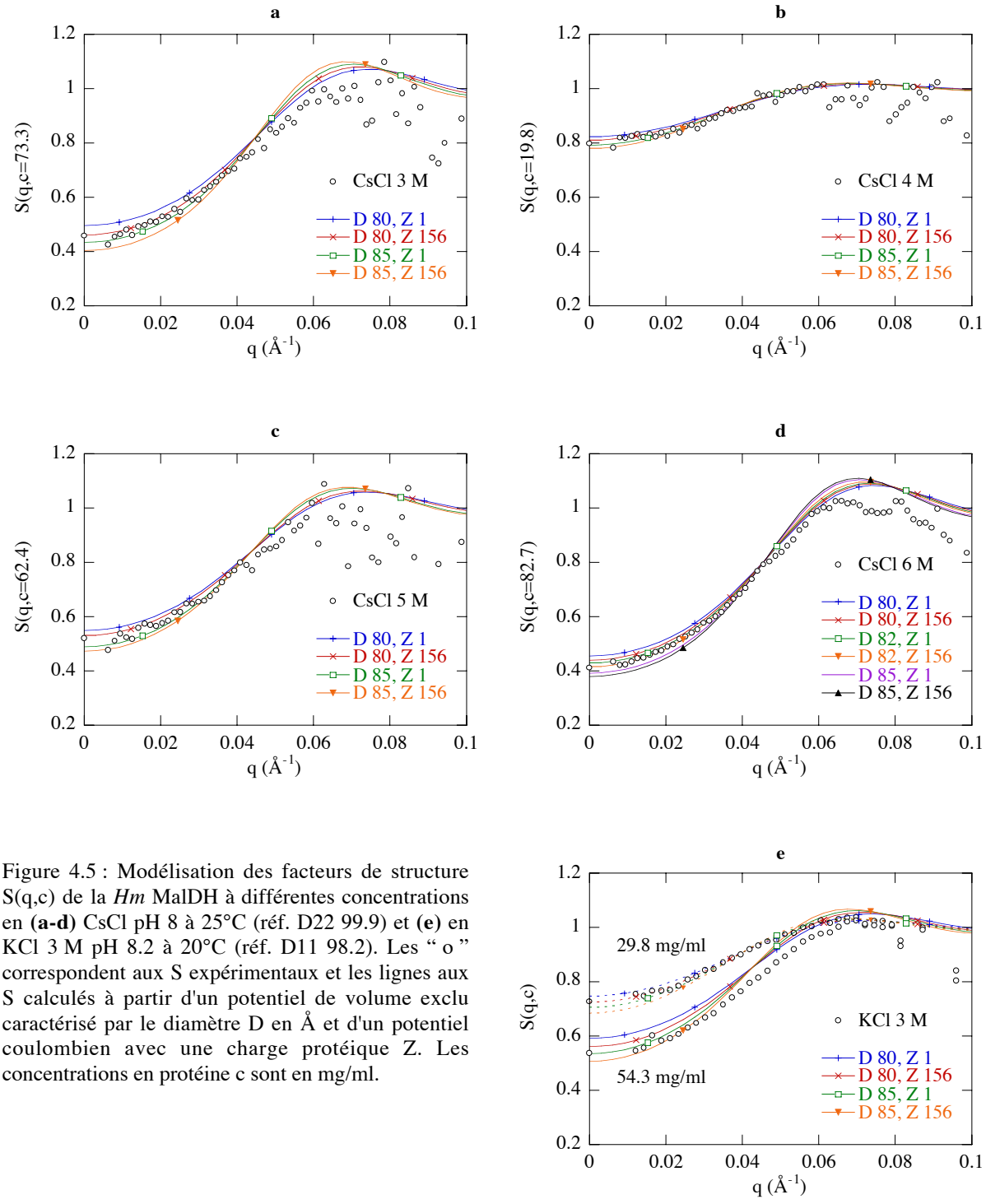


Figure 4.5 : Modélisation des facteurs de structure S(q,c) de la *Hm* MalDH à différentes concentrations en **(a-d)** CsCl pH 8 à 25°C (réf. D22 99.9) et **(e)** en KCl 3 M pH 8.2 à 20°C (réf. D11 98.2). Les " o " correspondent aux S expérimentaux et les lignes aux S calculés à partir d'un potentiel de volume exclu caractérisé par le diamètre D en Å et d'un potentiel coulombien avec une charge protéique Z. Les concentrations en protéine c sont en mg/ml.

4.3.2. Un potentiel attractif supplémentaire en $MgCl_2$, $MgSO_4$ et $(NH_4)_2SO_4$

Pour la seconde classe de solvant formée par $MgCl_2$, $MgSO_4$ et $(NH_4)_2SO_4$, l'évolution des A_2 avec la concentration en sel montre qu'elle a une forte influence sur les interactions

protéine – protéine. Nous avons remarqué que les valeurs les plus fortes de A_2 étaient du même ordre de grandeur que celles obtenues pour les sels précédents (NaCl, KCl, CsCl). Nous avons donc, dans un premier temps, cherché à modéliser les facteurs de structure S_{exp} de la *Hm* MalDH pour ces conditions les plus répulsives (en n'utilisant que les potentiels de volume exclu et électrostatique). Cependant, pour une solution de *Hm* MalDH en $MgCl_2$ 1.3 M, les facteurs de structure calculés grâce aux seuls deux potentiels répulsifs (D et Z) ne rendent pas correctement compte du S expérimental (Figure 4.6-a). Un D aberrant de 70 Å (car inférieur aux dimensions réelles de la protéine) permet de modéliser les valeurs du S_{exp} seulement aux très petites valeurs de q. Nous avons donc ajouté le potentiel attractif de type Yukawa de profondeur J (en kT) et de portée d (de 5 Å) et fait varier D (entre 75 et 85 Å) et J. Comme la modélisation ne permet pas à haut sel de déterminer objectivement une charge entre 1 et 156, nous avons fixé celle-ci à 70. Le meilleur accord est obtenu pour une profondeur du potentiel attractif de 1.6 ±0.1 kT avec un diamètre de 80 Å qui est plus relevant des dimensions de la *Hm* MalDH (Figure 4.6-b). Notons cependant qu'une modification de la charge module faiblement les valeurs de S proches de l'origine et cet effet peu être compensé par une variation de J. Les couples (Z, J) suivants : (1, 1.3), (70, 1.5) et (156, 2.1) donnent ainsi des S_{calc} difficilement discernables.

Nous avons donc choisi de fixer le diamètre à 80 Å et la charge à 70 pour les différentes concentrations de $MgCl_2$, afin de pouvoir comparer les meilleures valeurs de J. Un autre exemple de la détermination de J est donné en Figure 4.6-c pour la *Hm* MalDH à 79.6 mg/ml en $MgCl_2$ 0.5 M. L'ensemble des meilleurs facteurs de structure calculés pour les quatre concentrations en $MgCl_2$ aux plus fortes concentrations en protéine est repris dans la Figure 4.6-d. Les S_{calc} calculés à l'aide de ces trois potentiels reproduisent très correctement les S_{exp} aux différentes salinités. L'écart observé à 0.2 M est lié au fait que le facteur de structure étant proche de 1, il est délicat de déduire le S_{exp} des intensités diffusées, d'autant plus que la concentration maximale en protéine est plus faible pour cette condition : 30 mg/ml contre environ 80 mg/ml pour les autres salinités. On constate une augmentation de la composante attractive (J) lorsque la concentration en $MgCl_2$ diminue de 1.3 à 0.2 M.

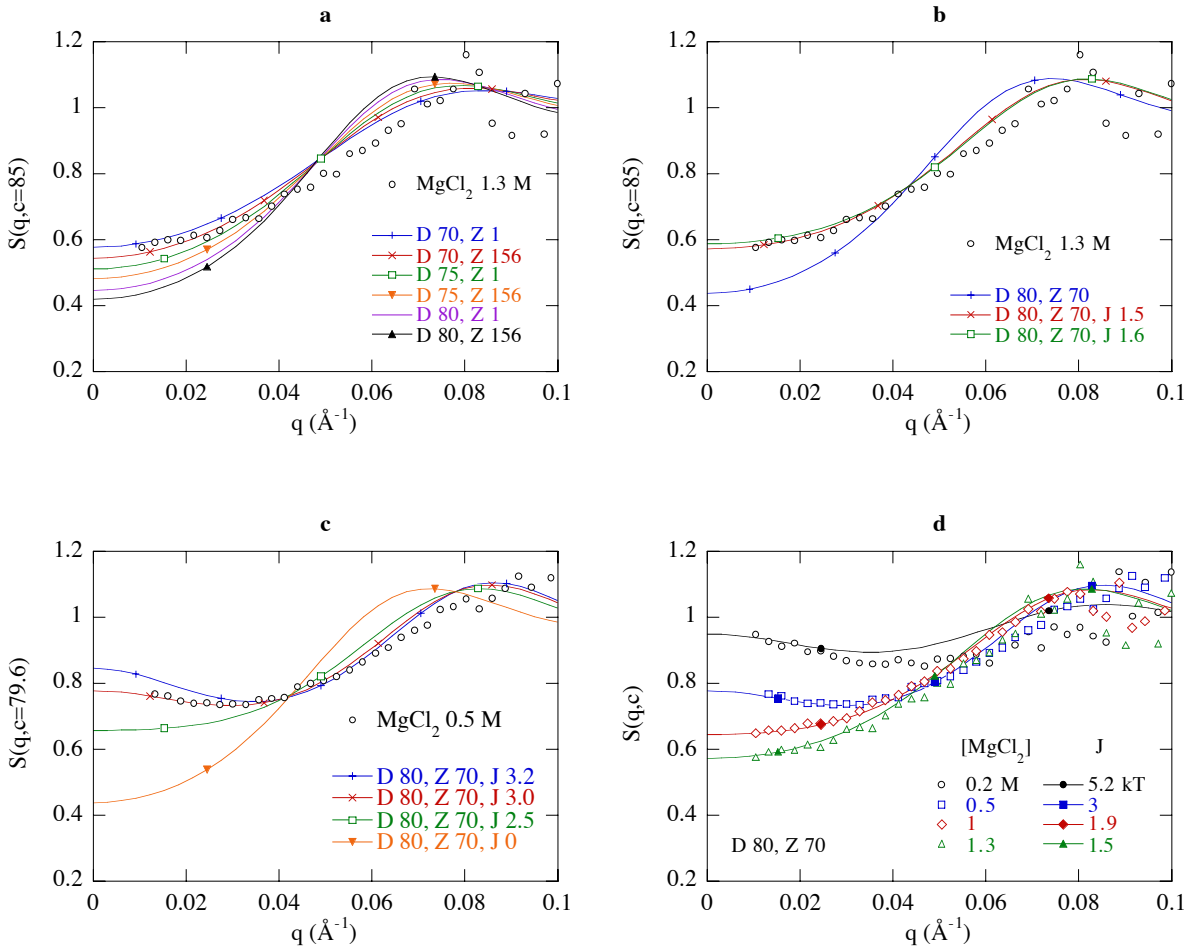


Figure 4.6 : Modélisation des facteurs de structure S(q,c) de la *Hm* MalDH à différentes concentrations de $MgCl_2$ pH 8.2 à 25°C (réf. Lure 98.12). Les symboles creux correspondent aux S expérimentaux et les lignes aux S calculés à partir d'un potentiel de volume exclu caractérisé par le diamètre D en Å, d'un potentiel coulombien avec une charge protéique Z et d'un potentiel attractif de profondeur J en kT et de 5 Å de portée. Les concentrations en protéine c (en mg/ml) sont, pour **(d)**, de 30.9, 79.6, 79.2 et 85 mg/ml à respectivement 0.2, 0.5, 1 et 1.3 M de $MgCl_2$. On ne peut donc pas comparer le S_{exp} à 30.9 mg/ml aux autres, mais seulement les valeurs de J qui ne dépendent pas de la concentration en protéine.

Pour les échantillons précédents, nous avons effectué des mesures en fonction de la température entre 5 et 25°C. Les valeurs J du potentiel attractif utilisé pour modéliser les facteurs de structure expérimentaux à 5°C sont comparés dans le Tableau 4.2 avec celles obtenues à 25°C. Comme nous l'avons remarqué pour les intensités diffusées, la température (entre 5 et 25°C) ne modifie pas significativement les interactions protéine – protéine. On peut toutefois noter la tendance à l'augmentation de la composante attractive lorsque la température baisse. Pour chaque concentration en sel, l'ajustement entre les facteurs de structure expérimentaux et calculés (pour les meilleures valeurs de J) est de qualité identique

à chaque température. Cela signifie que l'ordre de grandeur de la portée d du potentiel attractif que nous avons fixé à 5 Å semble être correct.

[$MgCl_2$] (M)	J (kT) à 5°C	J (kT) à 25°C
0.2	5 ±0.2	5.2 ±0.2
0.5	3.2 ±0.1	3 ±0.1
1	2.1 ±0.1	1.9 ±0.1
1.3	1.8 ±0.1	1.5 ±0.1

Tableau 4.2 : Profondeurs J en kT, à 5 et 25°C, du potentiel attractif de type Yukawa de portée d = 5 Å qui permettent de modéliser les facteurs structure expérimentaux de la *Hm* MalDH en $MgCl_2$ 0.2-1.3 M pH 8.2 (réf. Lure 98.12) avec un potentiel de volume exclu caractérisé par un diamètre D = 80 Å et un potentiel coulombien avec une charge protéique Z = 70.

En $MgSO_4$ 2 M, les interactions protéine – protéine peuvent être décrites grâce au potentiel répulsif de volume exclu (Figure 4.7-a). A cette force ionique (I = 8 M), la modélisation devient très discriminante pour le diamètre et une valeur de 82 ±1 Å est raisonnablement déterminée. Elle reste conforme aux valeurs trouvées précédemment. De nouveau, la charge ne peut être estimée. Lorsque la concentration en sel diminue, on retrouve la même évolution qu'en $MgCl_2$: les interactions deviennent moins répulsives et nous devons introduire une nouvelle composante attractive dont l'intensité est du même ordre de grandeur. La profondeur du potentiel attractif qui permet la meilleure modélisation est ainsi de 4 ±0.2 kT en $MgSO_4$ 0.25 M (Figure 4.7-b).

En sulfate d'ammonium, les interactions protéine – protéine peuvent aussi être considérées comme exclusivement répulsives à 1.5 M et le facteur de structure expérimental est assez bien modélisé avec un diamètre de volume exclu de 80 Å (Figure 4.7-c). A l'inverse des deux sels précédents, les interactions deviennent neutres ou faiblement attractives lorsque la concentration en $(NH_4)_2SO_4$ augmente (graphique d). La profondeur J de la composante attractive supplémentaire varie de 0 à 2.2 kT entre 1.5 et 3 M de sel. Par ailleurs, la température (entre 6 et 31°C) n'ayant pas d'effet notable en $(NH_4)_2SO_4$ 3 M (cf. Figure 4.2 p.95), les facteurs de structure aux différentes températures sont pour chaque salinité indiscernables.

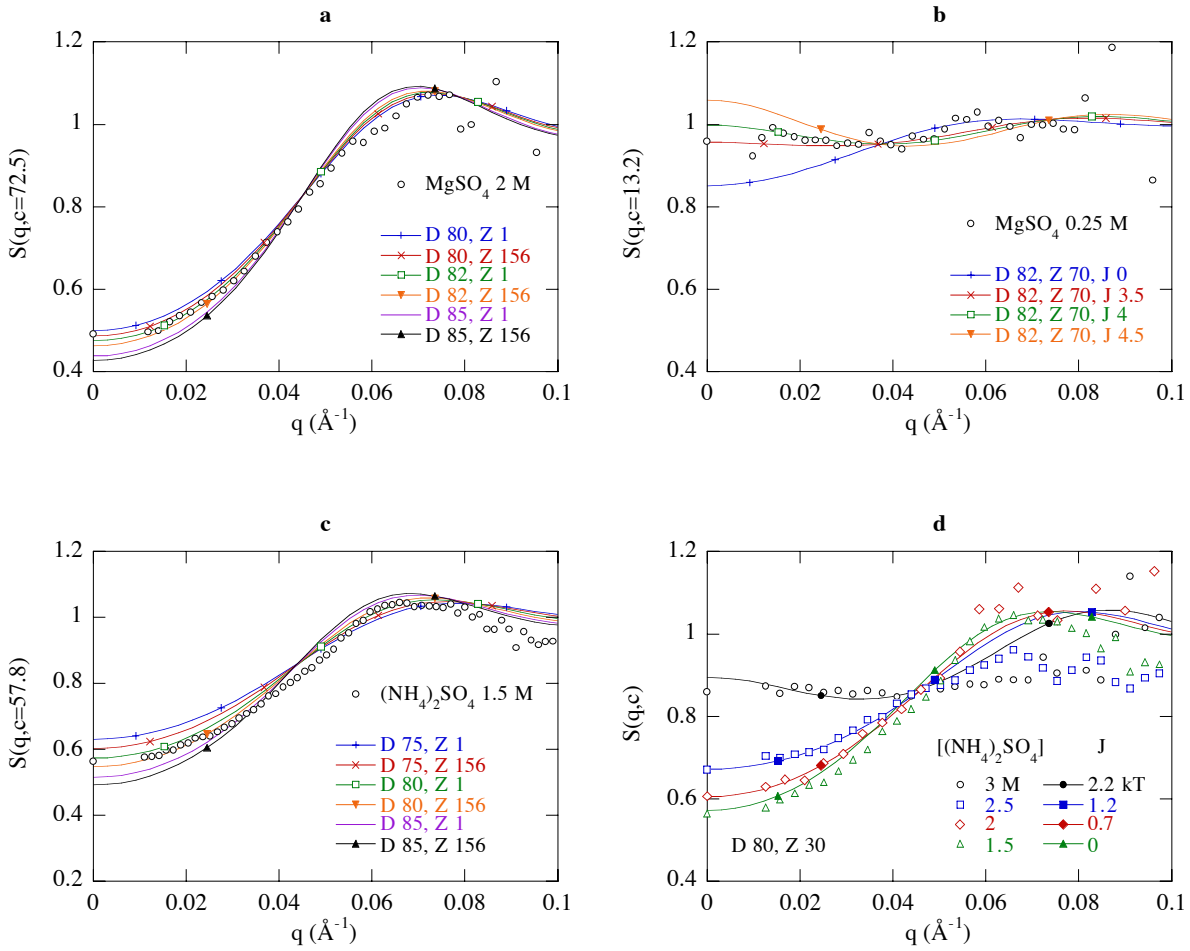


Figure 4.7 : Modélisation des facteurs de structure S(q,c) de la *Hm* MalDH à différentes concentrations **(a-b)** de $MgSO_4$ pH 8.2 à 23°C (réf. D11 98.7) et **(c-d)** de $(NH_4)_2SO_4$ pH 8, à 25°C (réf. D22 00.10, sauf à 2 M : réf. D22 97.8). Les symboles creux correspondent aux S expérimentaux et les lignes aux S calculés à partir d'un potentiel de volume exclu caractérisé par le diamètre D en Å, d'un potentiel coulombien avec une charge protéique Z et d'un potentiel attractif de profondeur J en kT et de 5 Å de portée. Les concentrations en protéine c sont pour **(d)** de 57.8, 61, 58.1, et 48.2 mg/ml à respectivement 1.5, 2, 2.5 et 3 M.

4.3.3. En bref…

Les meilleures valeurs des paramètres des trois potentiels (volume exclu, coulombien et attractif) utilisés pour modéliser les facteurs de structure expérimentaux de la *Hm* MalDH sont résumées dans le Tableau 4.3. Ces potentiels permettent de modéliser correctement les S_{exp} dans la plupart des conditions saline étudiées. Grâce au potentiel de volume exclu, nous avons pu évaluer le diamètre de la protéine en solution à 80-82 Å, valeur qui s'accorde avec sa structure cristallographique tridimensionnelle (cf. Figure 1.2 p.21). La charge de la particule reste toutefois indéfinie puisque les variations des S_{calc} induites par le potentiel

coulombien pour une charge de 1 à 156 sont de l'ordre de la précision des S_{exp}. Finalement, pour les seuls sels qui, à certaines concentrations, induisent des interactions protéine – protéine neutres ou faiblement attractives, nous avons pu modéliser cet effet par l'ajout d'un potentiel attractif de type Yukawa de profondeur J et d'une portée fixée à 5 Å. Les valeurs du paramètre J sont présentées en fonction de la salinité en Figure 4.8. Elles semblent évoluer régulièrement avec la salinité, sauf peut-être pour le $MgCl_2$ 0.2 M.

Sel	D (Å)	Z	J (kT)
NaCl 2-5 M	80 ±2	1-156	0
CsCl 3-6 M	82 ±1	1-156	0
KCl 3 M	82 ±2	1-156	0
$MgCl_2$ 0.2-1.3 M	80 (fixé)	70 (fixé)	5.2 à 1.5 ±0.1-0.2
$MgSO_4$ 0.25-2 M	82 ±1	1-156	4 à 0 ±0.2
$(NH_4)_2SO_4$ 1-3 M	80 ±2	1-156	0 à 2.2 ±0.1-0.2

Tableau 4.3 : Meilleures valeurs des paramètres D, Z et J des potentiels de volume exclu, coulombien et attractif de type Yukawa utilisés pour modéliser les facteurs de structure expérimentaux de la *Hm* MalDH en diverses conditions de sel à pH 8-8.2 et à 20-25°C. Les profondeurs J du potentiel attractif ont été déterminées pour une portée de 5 Å avec D (80 ou 82 Å) et Z (70 ou 30) fixés.

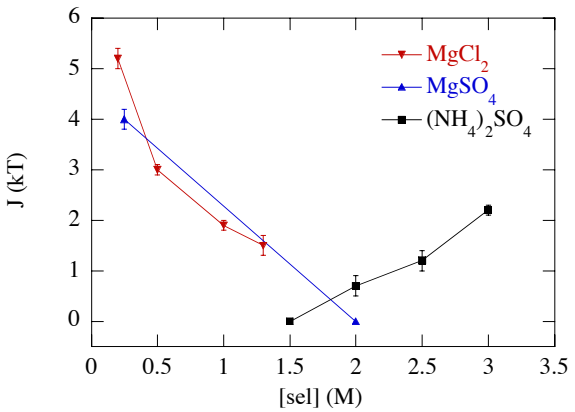


Figure 4.8 : Profondeurs J (en kT) du potentiel attractif de type Yukawa de 5 Å de portée utilisées pour modéliser les facteurs de structure expérimentaux de la *Hm* MalDH en fonction de la concentration en sel, à pH 8-8.2 et à 20-25°C. Les paramètres D et Z des potentiels de volume exclu et coulombien étaient fixés aux valeurs indiquées dans le Tableau 4.3.

4.3.4. Limites de la méthode

Dans la méthode de modélisation des facteurs de structure expérimentaux et donc des interactions protéine – protéine que nous utilisons, trois hypothèses fortes sont posées.

Premièrement, le solvant (y compris les co-solvants) est considéré comme un continuum caractérisé par une constante diélectrique et une force ionique. Les portées des potentiels que nous utilisons sont de l'ordre de 1 à 4 Å pour le potentiel électrostatique écranté (valeurs déterminée par la longueur de Debye) et de 5 Å pour le potentiel attractif (la portée du potentiel de volume exclu est nulle). Ces distances sont comparables aux dimensions des molécules du solvant. L'hypothèse de continuum est alors contestable. Deuxièmement, toutes les interactions sont supposées avoir une symétrie sphérique et le potentiel de volume exclu traite la protéine comme une sphère. Si à basse résolution une protéine globulaire comme la *Hm* MalDH peut être raisonnablement modélisée par une sphère, les autres interactions n'ont pas pour autant une symétrie sphérique. Ainsi, pour une sphère chargée positivement d'un côté et négativement de l'autre, le potentiel électrostatique dépend évidemment de la direction. Cependant, la surface de la *Hm* MalDH est chargée négativement de manière relativement uniforme. Troisièmement, face à la complexité des interactions élémentaires décrites en introduction (cf. §1.2 p.24), les potentiels utilisés sont extrêmement simples, surtout le potentiel de Yukawa qui en réalité est supposé traduire l'ensemble des interactions attractives (van der Waals, exclusion d'un co-solvant, effet hydrophobe, …). Malgré toutes ces contraintes, nos facteurs de structure expérimentaux sont étonnamment bien modélisés, et s'ils ne sont pas exactement reproduits, leur évolution en fonctions des différents paramètres est très correctement décrite.

Cependant, nous avons constaté une autre limite à cette démarche : le système (où D, Z, J et d sont les variables ajustables) apparaît comme sous-déterminé. En effet, dans nos conditions de solvants fortement salins, il est difficile de discerner l'effet du volume exclu de l'effet électrostatique. D'autre part, dès qu'une composante attractive est introduite, plusieurs couples de potentiels répulsifs (somme des deux précédents) et attractif peuvent rendre compte d'un même facteur de structure expérimental. Les portées des différents potentiels étant du même ordre de grandeur explique cette décomposition non-univoque.

A l'exception de l'indétermination de la charge qui est spécifique aux fortes concentrations en sel, les succès et les limites de cette modélisation semblent être communs aux protéines globulaires. En effet, la même démarche a été utilisée pour d'autres protéines comme le lysozyme ou les cristallines et avait aussi permis de rendre compte de leurs interactions interparticulaires de manière satisfaisante [Malfois *et al.*, 1996; Bonneté *et al.*, 1997; Finet, 1999; Tardieu *et al.*, 1999]. Pour le lysozyme, des potentiels de volume exclu et électrostatique associés à une attraction de courte portée permettent de modéliser les facteurs de structure à faible force ionique. Mais pour rendre compte de la spécificité des ions à plus

forte concentration (associé à l'effet d'Hofmeister), une répulsion et une attraction supplémentaires sont nécessaires. Comme dans notre cas, l'évolution des facteurs de structure expérimentaux en fonction de la concentration en sel peut être traduite par une charge protéique variable ou par une composante attractive variable [Tardieu *et al.*, 1999].

La modélisation des facteurs de structure semble donc être, malgré ses hypothèses, bien adaptée aux protéines globulaires. Cependant, l'utilisation de potentiels aspécifiques et génériques ne permet évidemment pas de comprendre l'origine moléculaire des interactions en jeu. Nous allons successivement préciser et discuter des interactions de volume exclu, électrostatique et attractives de la *Hm* MalDH en utilisant les résultats des seconds coefficients du viriel et des modélisations.

4.4. Discussion

4.4.1. Effet de volume exclu

L'effet de volume exclu traduit le simple fait que les différentes molécules de protéine ne peuvent occuper simultanément le même espace, autrement dit, elles ne peuvent pas s'interpénétrer. Il correspond à une répulsion effective des particules et donc à une valeur positive de A_2 qui est proportionnelle au volume V_p de la protéine en solution [Eisenberg, 1976] :

(4.1) $$A_{2\ \text{vol exclu}} = 4\ V_p\ N_A / M_+^2\ .$$

Le volume V_p peut être déterminé en considérant la protéine en solution comme une sphère ou à partir de sa composition.

Dans le premier cas, il est lié au diamètre D de la particule par $V_p = 4\pi(D/2)^3/3$. La modélisation des facteurs de structure de la *Hm* MalDH nous a permis de déterminer selon le sel des valeurs de D de 80 à 82 Å qui correspondent au diamètre que l'on peut déduire de la structure cristallographique de la *Hm* MalDH. Les dimensions maximales du tétramère sont en effet de 77 Å entre les deux monomères d'un dimère, de 80 et 90 Å (en diagonale) entre les

deux dimères (cf. Figure 1.2 p.21). Les seconds coefficients du viriel correspondant aux diamètres D obtenus par la modélisation sont reportés dans la troisième colonne du Tableau 4.4.

Dans le second cas, V_p est le volume de la particule solvatée qui est composée de la protéine, de ses contre-ions et du solvant perturbé ou " lié ". Les interactions entre la *Hm* MalDH et le solvant ont été étudiées dans différents sels et interprétées dans le cadre du modèle de la particule invariante. Ce modèle est le plus simple et consiste à considérer qu'une mole de protéine est associée à N_1 moles d'eau et N_3 moles de sel, quantités constantes pour des concentrations variables de sels dans le solvant [Tardieu *et al.*, 1981; Eisenberg, 1976, 1981; Ebel, 1995]. Puisque la particule est définie comme la protéine accompagnée de ses contre-ions (sa masse est notée M_+), N_3 peut être positif ou négatif. Le dernier cas correspond à la dissociation d'une partie des contre-ions. Dans le cadre de ce modèle, le volume V_p de la protéine solvatée est exprimé par la relation :

$$V_p = \left(\bar{v} M_+ + N_1 \bar{v}_1 M_1 + N_3 \bar{v}_3 M_3 \right) / N_A \quad . \tag{4.2}$$

L'eau est le composant 1, la protéine le 2 (indice omis) et le co-solvant (le sel sans distinction des ions) le 3. Les $\bar{v}_i$ sont les volumes partiels spécifiques, M_i les masses molaires et N_i le nombre de molécules i associées à la protéine (en moles de i par mole de 2). Les quantités N_1 et N_3 ont été déterminées pour la *Hm* MalDH dans divers sels (annexe 8.4 p.194) et sont indiquées dans le Tableau 4.4 [Bonneté *et al.*, 1993; Ebel, en préparation]. Eu égard au faible poids de la couche de solvatation par rapport à celui de la protéine, les mesures expérimentales sont très délicates et les erreurs des paramètres déduits importantes. Elles sont de l'ordre de ±500 mol/mol pour N_1 et de 25-50 % pour N_3. Avec l'équation (4.2), nous déduisons de ces paramètres le volume V_p de la *Hm* MalDH solvatée puis le diamètre D_{solv} et le A_2 de volume exclu correspondants pour une sphère (cf. Tableau 4.4). Sans tenir compte de la solvatation de la protéine (N_1 et N_3 nuls), nous aurions trouvé un diamètre de 68 Å et un A_2 de 2.2 10^{-8} mol.l.g^{-2}. La différence entre cette valeur et celles du tableau traduit toute l'importance de la couche de solvatation de la protéine. Notons aussi que l'hydratation (N_1) est responsable de 80 % de cette différence. Cette remarque nous permet d'étendre l'importance de la couche d'hydratation dans les interactions interparticulaires aux protéines non-halophiles qui ont en effet une hydratation du même ordre de grandeur que celle qui est observée pour la *Hm* MalDH en NaCl [Bonneté *et al.*, 1993].

Sel	Modélisation de S		Composition de la protéine solvatée						
	D (Å)	$A_{2\ vol\ exclu}$ (10^{-8} mol.l.g^{-2})	M_+ (g/mol)	N_1 (mol/mol)	N_3 (mol/mol)	$\tilde{v}_3$ (ml/g)	M_3 (g/mol)	D_{solv} (Å)	$A_{2\ vol\ exclu}$ (10^{-8} mol.l.g^{-2})
NaCl	80	3.6	134138	2040	55	0.38	58	76	3.0
$(NH_4)_2SO_4$	80	3.6	133366	2100	0	0.5	132	75	3.0
$MgCl_2$	80	3.7	132448	4160	85	0.22	95	82	4.0
CsCl	82	3.0	151284	1900	-20	0.27	168	77	2.5

Tableau 4.4 : Seconds coefficients du viriel A_2 correspondant au volume exclu de la *Hm* MalDH pour différents sels. Les A_2 sont calculés avec l'équation (4.1) d'après les diamètres D déterminés par la modélisation des facteurs de structure S(q) (reportés du Tableau 4.3) ou d'après le volume de la protéine solvatée. Dans le cadre d'une particule invariante, la protéine avec ses contre-ions est en solution associée à N_1 molécules d'eau et N_3 molécules de sel [d'après Ebel, en préparation]. Son volume est calculé selon l'équation (4.2) avec un $\tilde{v}_1$ de 1 et un $\tilde{v}_2$ moyen de 0.73 ml/g [Bonneté *et al.*, 1993]. La précision des A_2 est de $\pm 0.2\ 10^{-8}$ mol.l.g^{-2}.

Les A_2 de volume exclu de la *Hm* MalDH déterminés d'après les diamètres issus des modélisations sont du même ordre de grandeur que ceux calculés à partir de la composition de la protéine solvatée, mais ils sont légèrement supérieurs à ces derniers (sauf en $MgCl_2$, où le diamètre D à peu être été sous-estimé) (Tableau 4.4). Dans la deuxième approche, A_2 correspond au volume de la protéine solvatée considérée comme sphérique : il est peut-être sous-estimé. En définitive, il nous apparaît que la précision des diamètres déterminés par la modélisation des facteurs de structure ne nous permet pas d'évaluer une variation du volume de la particule due à son solvant associé selon la nature du sel.

L'effet de volume exclu se traduit pour la *Hm* MalDH par des seconds coefficients du viriel A_2 compris entre 2.5 et 4.0 10^{-8} mol.l.g^{-2} selon le sel, mais constant pour chaque sel. La Figure 4.9 permet aisément de comparer les valeurs issues du diamètre D et du volume de la protéine solvatée à celles des A_2 globaux mesurées. On remarque immédiatement que le terme de volume exclu est responsable d'une grande partie des valeurs positives des A_2 (interactions répulsives). Mais par exemple en NaCl 2-3 M et en CsCl (graphiques a et b), il reste inférieur aux A_2 mesurés. En considérant que nos estimations du volume exclu sont correctes, il subsiste donc une répulsion supplémentaire dont l'origine serait électrostatique. Malgré les fortes concentrations de sel utilisées, la charge de la *Hm* MalDH ne serait pas totalement écrantée. A l'inverse, le A_2 de volume exclu est supérieur aux A_2 mesurés en $MgCl_2$ et au-delà de 2.5 M de $(NH_4)_2SO_4$ (Figure 4.9-c et d) : il existe donc une force

attractive entre les particules. D'une manière générale, le terme de volume exclu ne rend absolument pas compte de la baisse monotone des A_2 expérimentaux avec la salinité et encore moins des brusques variations observées pour certaines conditions.

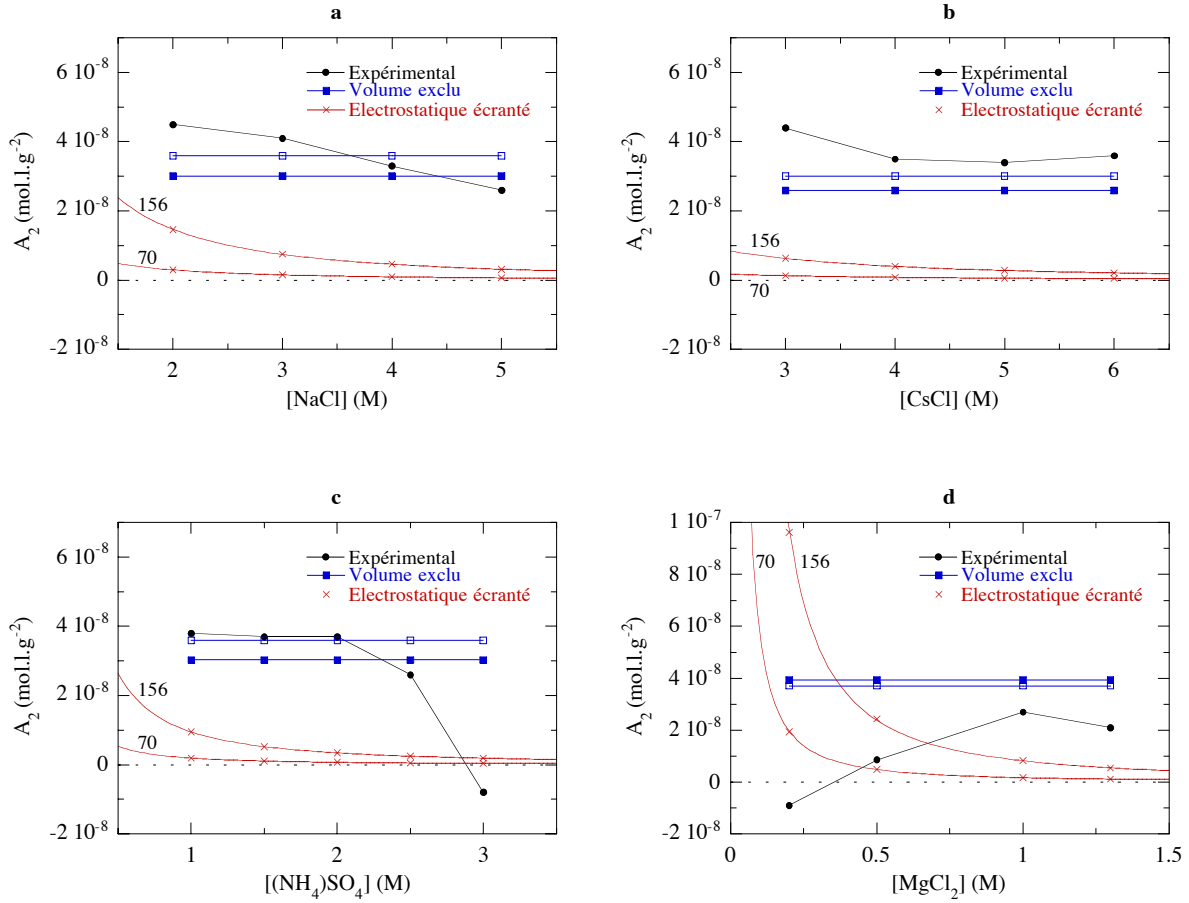


Figure 4.9 : Comparaison des seconds coefficients du viriel de volume exclu et électrostatique écranté aux valeurs expérimentales obtenues pour la *Hm* MalDH en divers sels. Les A_2 de volume exclu issus du diamètre D (□) et du volume de la protéine solvatée (■) sont repris du Tableau 4.4, leur précision est de ±0.2 10^{-8} mol.l.g^{-2}. Les A_2 électrostatiques écrantés ont été calculés selon l'équation (4.4) avec un D de 80 Å et un Z de 70 ou 156 indiqué en regard des courbes. La précision des A_2 expérimentaux est de l'ordre de ±0.4 10^{-8} mol.l.g^{-2}.

4.4.2. Effet électrostatique

La *Hm* MalDH étant une particule chargée (dès lors que nous ne nous plaçons pas au pH correspondant à son point isoélectrique), il existe une interaction électrostatique entre ces particules qui est répulsive puisqu'elles ont la même charge. Cette interaction est traduite par le potentiel coulombien écranté décrit précédemment (cf. (2.23) p.64). Comme le second

coefficient du viriel A_2 est relié au potentiel total W(r) par la relation suivante [McMillan et Mayer, 1945; Belloni, 1991] :

$$A_2 = 2\pi \frac{N_A}{M_+^2} \int_0^\infty \left(1 - e^{-W(r)/kT}\right) r^2 dr \ , \tag{4.3}$$

nous pouvons en déduire la partie électrostatique de A_2. En faisant un développement limité de l'exponentielle* dans (4.3) dont on ne conserve que le terme linéaire, l'intégration conduit au $A_{2\,élec}$ suivant [Harding *et al.*, 1992, p.314] :

$$A_{2\,elec} = \frac{Z^2 (1 + \kappa D)}{4 I M_+^2 (1 + \kappa D/2)^2} \ , \tag{4.4}$$

où Z, D et M_+ sont respectivement la charge (après condensation d'une partie des contre-ions), le diamètre et la masse moléculaire (avec tous les contre-ions) de la protéine, κ la constante d'écrantage (ou $1/\lambda_D$) et I la force ionique exprimée en unité molaire. κ étant proportionnel à $I^{0.5}$, $A_{2\,élec}$ varie approximativement comme $I^{-1.5}$ et diminue donc avec la salinité du solvant, traduisant l'écrantage progressif de la charge protéique.

Nous avons représenté dans la Figure 4.9 l'évolution de $A_{2\,élec}$ avec la salinité exprimée par l'équation (4.4) pour deux valeurs de la charge Z : 70 et 156. Compte tenu de la précision des valeurs, nous ne pouvons faire que des observations semi-quantitatives. La baisse de $A_{2\,élec}$ lorsque la concentration en sel augmente est assez bien corrélée à la faible baisse monotone que nous observons pour les A_2 expérimentaux en NaCl, CsCl et en $(NH_4)_2SO_4$ 1-2 M (graphiques a-c). De plus, on constate que la somme des termes A_2 de volume exclu et électrostatique peut raisonnablement rendre compte des valeurs obtenues en NaCl entre 2 et 4 M, en CsCl et en $(NH_4)_2SO_4$ 1-2 M. Ces deux observations corroborent l'idée d'une particule relativement chargée pour laquelle la répulsion électrostatique n'est pas négligeable, même pour des forces ioniques de l'ordre du molaire.

Avec les A_2 de volume exclu déterminés précédemment (cf. Tableau 4.4 p.113), l'évolution des A_2 expérimentaux est en accord avec une charge supérieure à 70 pour les trois sels, mais il faut remarquer que les effets électrostatiques sont largement écrantés par les

* $\exp(x) = 1 + x + x^2/2! + \ldots \approx 1 + x$ pour $x << 1$. Dans notre cas x est le potentiel électrostatique W(r)/kT qui est bien inférieur à 1 pour $r > D$.

fortes salinités utilisées. Cette forte charge signifierait que relativement peu de contre-ions se condensent à la surface de la protéine pour former des paires d'ions. D'après l'interprétation structurale en termes de particule invariante des paramètres d'interaction avec le solvant, la couche de solvatation de la *Hm* MalDH contiendrait une quantité non négligeable de sel (cf. N_3 dans le Tableau 4.4). Ces deux observations semblent donc être en opposition. Il faut cependant noter que les mesures de solvatation sont statistiques et globales : elles ne discernent pas les cations des anions du sel. Nous ne pouvons donc pas déduire la valeur de charge effective de la protéine.

D'une manière générale, l'équation (4.4) est une approximation qui ne tient pas compte des effets spécifiques des ions. Cette spécificité est d'autant plus importante à forte concentration de sel. En CsCl, les A_2 expérimentaux restent significativement supérieurs à la somme des contributions du volume exclu (déterminée d'après la modélisation ou le volume de la protéine solvatée) et coulombienne, même pour une charge de 156. La nature de Cs^+, cation volumineux de faible densité de charge, pourrait être à l'origine de cette différence. Il est en effet chaotropique, *i.e.* il déstructure l'organisation des molécules d'eau, et devrait donc rester éloigné des carboxylates kosmotropes de la protéine. Collins montre en effet que lorsqu'un anion et un cation sont tous deux kosmotropes ou chaotropes, ils ont tendance à former une paire d'ions, et lorsqu'ils ont des tendances opposées à organiser l'eau (kosmotrope – chaotrope), les ions restent séparés, hydratés en solution [Collins, 1995]. La capacité d'écrantage de Cs^+ à faible distance de la protéine serait donc moindre que celle qui est décrite dans le modèle DLVO, lequel ne tient pas compte de la spécificité de chaque sel. Le A_2 électrostatique serait donc plus élevé et rendrait ainsi mieux compte des valeurs mesurées. Le même effet pourrait expliquer la valeur plus élevée de A_2 en KCl 3 M qu'en NaCl 3 M, 5.2 10^{-8} contre 4.1 10^{-8} mol.l.g^{-2} (Figure 4.3 p.97).

En dehors de ces conditions de sel, il est flagrant sur la Figure 4.9 que la somme des contributions de volume exclu et électrostatique ne rend absolument pas compte des A_2 mesurés en $(NH_4)_2SO_4$ 2.5-3 M ou en $MgCl_2$ 0.2-1.3 M (graphiques c et d). Dans les deux cas, les valeurs expérimentales sont en effet largement inférieures au seul A_2 de volume exclu. De plus, à basses concentrations de $MgCl_2$, le terme électrostatique écranté devient extrêmement important alors que le A_2 expérimental décroît fortement. Comme nous l'avons vus lors de la modélisation des facteurs de structure dans ces conditions salines, une composante attractive dépendante du sel et de sa concentration devient nécessaire pour modéliser les interactions

protéine – protéine. Les mêmes observations ont été faites pour le lysozyme et des α et γ-cristallines dans certaines conditions de solvant [Finet, 1999; Tardieu *et al.*, 1999]. L'origine de cette attraction reste pour l'instant indéterminée.

4.4.3. Le rôle des interactions protéine – solvant dans les interactions protéine – protéine

Second coefficient du viriel et interactions protéine – solvant …

Afin de tenter de comprendre l'origine de l'attraction induite par certains sels, nous allons nous intéresser au rôle des interactions protéine – solvant dans les interactions protéine – protéine. A cette fin, rappelons le développement du viriel de la pression osmotique Π ou de sa dérive $(\partial\Pi/\partial c)_{T,\mu}$ à température et aux potentiels chimiques μ des espèces diffusibles constants :

$$(4.5) \qquad \frac{\Pi}{N_A kT} = c\left(\frac{1}{M} + A_2 c + \cdots\right) \quad \text{et} \quad \frac{1}{N_A kT}\left(\frac{\partial\Pi}{\partial c}\right)_{T,\mu} = \frac{1}{M} + 2A_2 c + \cdots .$$

La pression osmotique d'une solution non idéale à plusieurs composants a été rigoureusement exprimée par Scatchard en 1946, puis démontrée différemment par d'autres auteurs [Scatchard, 1946; Casassa et Eisenberg, 1964; Eisenberg, 1976]. Pour un système à trois composant, outre les termes de volume exclu et électrostatique, il apparaît dans l'expression du second coefficient du viriel un terme qui implique les interactions protéine – solvant (cf. annexe 8.3 p.193). Ce terme découle de la dépendance des potentiels chimiques à toutes les espèces en solutions. Pour l'unité de concentration que nous utilisons (mg/ml), l'expression de A_2 est [Pittz et Timasheff, 1978][*] :

$$(4.6) \qquad A_2 = \frac{1}{2}\frac{V_m^o}{N_A kT\, M_+^2}\left[\left(\frac{\partial\mu_2^{(e)}}{\partial m_2}\right)_{T,P,m} + \left(\frac{\partial\mu_2}{\partial m_3}\right)_{T,P,m}\left(\frac{\partial m_3}{\partial m_2}\right)_{T,P,\mu}\right].$$

[*] Dans leur expression, ces auteurs ont omis le terme $(\partial V_m/\partial m_2)_\mu{}^\circ/V_m{}^\circ$ [équation 2.50 de Eisenberg, 1976] dans le changement d'unité de la concentration (de m vers c). Pour la *Hm* MalDH, il est d'environ 0.5 10^{-8} mol.l.g^{-2}, valeur de l'ordre de la précision de nos A_2. Il est omis dans nos calculs.

$V°_m$ est le volume qui contient 1000 g de solvant principal (l'eau), m_i la concentration molale (ou molalité) de l'espèce i (mol/kg de 1). Les indices m ou μ indiquent la constance des molalités ou des potentiels chimiques de toutes les espèces, exceptée celle qui est utilisée pour la différentiation. $\mu_2^{(e)}$ est l'excès de potentiel chimique défini par :

$$\mu_2 = \mu_2^\circ + N_A kT \ln a_2 = \mu_2^\circ + N_A kT \ln m_2\gamma_2 = \mu_2^\circ + N_A kT \ln m_2 + \mu_2^{(e)} , \tag{4.7}$$

où a_2 est l'activité de la protéine et γ_2 le coefficient d'activité. $\mu_2^{(e)}$ est donc la partie du potentiel chimique lié à γ_2. Le terme $(\partial m_3/\partial m_2)_\mu$ est le paramètre d'interaction préférentielle de la protéine native avec le sel. Il s'agit d'une grandeur thermodynamique mesurable qui représente la quantité hypothétique de sel qu'il faut ajouter ou retrancher à la solution quand on ajoute une mole de protéine, pour maintenir les potentiels chimiques de l'eau μ_1 et du sel μ_3 identiques à ceux du solvant pur [Eisenberg, 1976]. Il traduit l'ensemble des interactions protéine – solvant (comme tout paramètre thermodynamique, il ne dépend pas d'un modèle structural des interactions). Il est lié à l'évolution du potentiel chimique du sel (respectivement de la protéine) avec la concentration en protéine (respectivement en sel) dans la solution :

$$\left(\frac{\partial m_3}{\partial m_2}\right)_{T,P,\mu} = -\left(\frac{\partial \mu_3}{\partial m_2}\right)_{T,P,m} \Big/ \left(\frac{\partial \mu_3}{\partial m_3}\right)_{T,P,m} = +\left(\frac{\partial \mu_2}{\partial m_3}\right)_{T,P,m} \Big/ \left(\frac{\partial \mu_3}{\partial m_3}\right)_{T,P,m} , \tag{4.8}$$

car $(\partial\mu_2/\partial m_3) = -(\partial\mu_3/\partial m_2)$. En remplaçant le facteur $(\partial\mu_2/\partial m_3)$ de l'équation (4.6) par son expression dans (4.8), nous obtenons pour A_2 une expression générale équivalente :

$$A_2 = \frac{1}{2}\frac{V_m^o}{N_A kT\, M_+^2}\left[\left(\frac{\partial \mu_2^{(e)}}{\partial m_2}\right)_{T,P,m} - \left(\frac{\partial \mu_3}{\partial m_3}\right)_{T,P,m}\left(\frac{\partial m_3}{\partial m_2}\right)^2_{T,P,\mu}\right] . \tag{4.9}$$

Dans cette formulation, le premier terme rend compte des interactions entre les protéines et comprend les effets de volume exclu et électrostatique décrit précédemment. En absence d'interaction protéine – protéine attractive, il est positif. Le second terme représente l'effet induit sur les interactions protéine – protéine par les interactions protéine – solvant. Comme $(\partial\mu_3/m_3)$ est positif, il correspond toujours à une contribution négative, quel que soit le signe de $(\partial m_3/\partial m_2)$. Les interactions protéine – solvant ajoutent donc une composante attractive aux interactions protéine – protéine (sauf dans le cas particulier $(\partial m_3/\partial m_2) = 0$). Exprimé différemment, cet effet peut se comprendre intuitivement : dès que la couche de solvatation

de la protéine diffère du solvant libre (à cause d'une accumulation d'eau ou de co-solvant), le solvant déstabilisé tend à repousser les particules de protéine, ce qui correspond à une attraction protéine – protéine.

… pour la *Hm* MalDH

Les paramètres d'interaction préférentielle ($\partial m_3/\partial m_2$) ont été déterminés par Ebel *et al*. pour la *Hm* MalDH native en différents sels (cf. annexe 8.4 p.194) et sont reportés dans le Tableau 4.5 pour NaCl, $(NH_4)_2SO_4$, $MgCl_2$ et CsCl [Ebel, en préparation]. Ils nous permettent d'estimer les contributions aux A_2 des interactions protéine – solvant ($A_{2\ prot\text{-}sel}$). Puisque ($\partial\mu_3/\partial m_3$) varie en première approximation* comme $2/m_3$ pour les sels de type 1:1 et comme $3/m_3$ pour les sels de type 1:2 ou 2:1, une même valeur de ($\partial m_3/\partial m_2$) contribue au A_2 de façon différente selon la composition et la concentration du sel dans le solvant.

On constate rapidement que la somme des trois contributions ($A_{2\ calc}$) rend imparfaitement compte des valeurs expérimentales du second coefficient du viriel ($A_{2\ exp}$). La somme calculée est inférieure à ces derniers dans presque toutes les conditions de sel. Sans exclure complètement une sous-estimation du terme en $\left(\partial\mu_2^{(e)}/\partial m_2\right)$ qui comprend les effets de volume exclu et électrostatique, les paramètres d'interaction préférentielle ($\partial m_3/\partial m_2$) ou leur effet sur A_2 (où ils interviennent portés au carré) semblent être surestimés.

On note cependant une bonne corrélation qualitative entre l'évolution globale des A_2 expérimentaux et des A_2 calculés (Figure 4.10). Les baisses prononcées des $A_{2\ exp}$ sont très nettement corrélées aux écarts prononcés des ($\partial m_3/\partial m_2$) (Tableau 4.5) : aux fortes valeurs négatives de ($\partial m_3/\partial m_2$) en $(NH_4)_2SO_4$ 2.5-3 M ou aux fortes valeurs positives en $MgCl_2$ 0.2-0.5 M correspondent des $A_{2\ exp}$ significativement plus faibles. Nous avons vu dans la Figure 4.9 (p.114) que dans ces conditions, les A_2 expérimentaux étaient très inférieurs à la somme des termes de volume exclu et électrostatique. Parallèlement, la modélisation des facteurs de structure pour ces mêmes conditions nécessite un potentiel attractif de forte profondeur (cf. Figure 4.8 p.109). Ce potentiel attractif traduirait donc principalement l'effet des interactions protéine – solvant et comme les paramètres d'interaction de la protéine avec le sel varie avec ce dernier, il est normal que sa profondeur dépende de la concentration en sel. En CsCl et NaCl 5 M, les fortes valeurs de ($\partial m_3/\partial m_2$) ne sont pas traduites dans les A_2 expérimentaux.

* Pour une solution saline, $(\partial\mu_3/\partial m_3)_{T,P,m} = n\,NAkT\,(1/m_3 + \partial\ln\gamma_\pm/\partial m_3)$ où $\gamma_\pm$ est le coefficient d'activité ionique moyen et n vaut 2 pour un sel 1:1 et 3 pour un sel 1:2 ou 2:1.

Sel	[sel]	$\partial m_3/\partial m_2$	$\Delta(\partial m_3/\partial m_2)$	$\partial \mu_3/\partial m_3$	$V°_m$	$A_{2\ prot\text{-}sel}$	$\Delta(A_{2\ prot\text{-}sel})$	$A_{2\ vol\ exclu}$	$A_{2\ élec}$	$A_{2\ calc}$	$A_{2\ exp}$
	M	mol/mol		$J.mol^{-2}$	l	10^{-8} mol.l.g^{-2}					
NaCl	2	-38	40	1784	1.09	-3.4	8.0	3.3	1.5	1.4	4.5
	3	-54	40	1634	1.10	-6.3	10.3	3.3	0.7	-2.2	4.1
	4	-90	45	1474	1.11	-15.9	18.0	3.3	0.5	-12.2	3.2
	5	-164	45	1325	1.12	-48.2	31.9	3.3	0.3	-44.5	2.6
$(NH_4)_2SO_4$	1	-46	35	6244	1.06	-17.1	26.0	3.3	0.9	-12.8	3.8
	1.5	-65	30	4077	1.10	-23.1	21.3	3.3	0.5	-19.3	3.7
	2	-90	33	2957	1.15	-33.4	24.5	3.3	0.3	-29.8	3.7
	2.5	-118	34	2273	1.20	-46.2	26.6	3.3	0.2	-42.6	2.6
	3	-147	34	1819	1.26	-60.0	27.8	3.3	0.2	-56.5	-0.8
$MgCl_2$	0.2	64.5	20	27737	1.00	-142.6	140.1	3.8	9.6	-129.2	-0.9
	0.5	48	25	14392	1.01	-41.2	53.4	3.8	2.4	-34.9	0.9
	1	23	20	10309	1.02	-6.8	12.4	3.8	0.8	-2.2	2.7
	1.3	-23	25	9217	1.02	-6.2	13.7	3.8	0.5	-1.8	2.1
CsCl	3	-140	60	1393	1.13	-29.2	34.0	2.8	0.6	-25.8	4.4
	4	-185	50	1023	1.19	-39.5	29.3	2.8	0.4	-36.3	3.5
	5	-255	60	799	1.26	-61.9	39.4	2.8	0.3	-58.9	3.4
	6	-295	70	652	1.34	-72.0	44.3	2.8	0.2	-69.0	3.6

Tableau 4.5 : Seconds coefficients du viriel calculés de volume exclu, électrostatique et d'interaction protéine – solvant comparés aux A_2 expérimentaux de la *Hm* MalDH pour différents sels. Les A_2 de volume exclu sont les moyennes des deux valeurs du Tableau 4.4 (p.113). Les A_2 électrostatiques pour une charge Z de 156 sont reportés de la Figure 4.9 (p.114). Les paramètres d'interaction préférentielle avec le sel (∂m3/∂m2) [d'après Ebel, en préparation] permettent le calcul des A_2 d'interaction avec le solvant ($A_{2\ prot\text{-}sel}$) grâce à l'équation (4.9) où $V°_m = (1000+m_3M_3)/\rho°$, $\rho°$ est la densité massique du solvant et ($\partial\mu$3/∂m3) est calculé d'après l'évolution des γ_3 en fonction de m_3 [Zaitsev et Aseev, 1992; Lide et Frederikse, 1996]. $A_{2\ calc}$ est la somme des trois contributions $A_{2\ vol\ exclu}$, $A_{2\ élec}$ et $A_{2\ prot\text{-}sel}$.

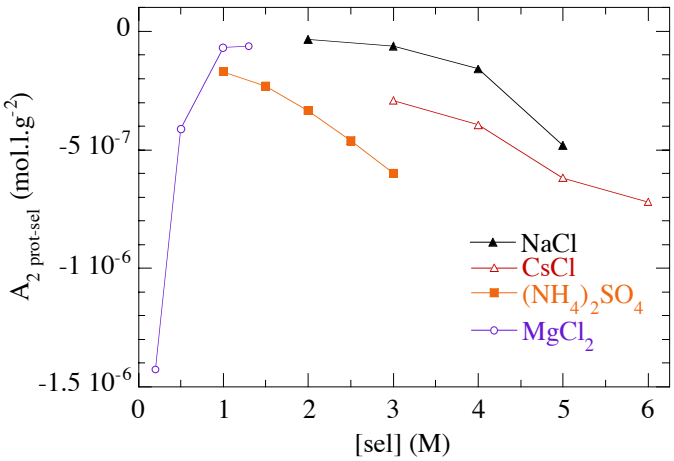


Figure 4.10 : Second coefficient du viriel d'interaction avec le solvant $A_{2\ prot\text{-}sel}$ (cf. Tableau 4.5).

L'écart observé entre les $A_{2\ prot\text{-}sel}$ et les A_2 expérimentaux peut signifier l'existence d'un seuil en dessous duquel les interactions protéine – solvant n'influent pas sur les interactions protéine – protéine. Au-dessus de cette limite, l'accumulation de sel ou d'eau dans la couche de solvatation rend cette dernière trop différente du solvant libre et ajoute une attraction entre les particules de *Hm* MalDH.

Hydratation et solubilité

L'effet des sels sur les interactions protéine – protéine est généralement invoqué sous la dénomination " d'effet Hofmeister ". Les anions et les cations sont classés selon leur capacité à faire précipiter les protéines. L'effet des anions qui est plus marqué s'ajoute à celui des cations selon l'ordre (non-exhaustif) suivant pour les protéines acides (non-halophiles) :

$$SO_4^- > CH_3CO_2^- > HCO_3^- > Cl^- > SCN^-$$
$$Li^+ > Na^+ \approx K^+ > NH4^+ > Mg^{2+} ;$$

pour les protéines basiques, la série anionique est inversée [voir par exemple Riès-Kautt et Ducruix, 1997]. Le fait que les interactions protéine – solvant soient intégrées dans les interactions protéine – protéine via les paramètres d'interaction préférentielle ($\partial m_3/\partial m_2$) traduit l'effet d'Hofmeister et donc la capacité spécifique de chaque sel à faire précipiter une protéine. L'ordre dans les séries d'Hofmeister semble être globalement respecté pour de nombreuses protéines. L'effet d'Hofmeister est donc en premier lieu indépendant de la spécificité de la surface de chaque protéine et doit relever de phénomènes aspécifiques qui engendrent des paramètres d'interaction préférentielle ($\partial m_3/\partial m_2$) ayant le même effet sur les interactions protéine – protéine. Les mécanismes d'exclusion stérique (force de déplétion), d'augmentation de la tension superficielle de l'eau, de répulsion par les charges protéiques de surface qui tous engendrent une hydratation préférentielle de la protéine ont été utilisés pour expliquer l'effet des co-solvants sur la solubilité des protéines [Timasheff et Arakawa, 1988].

L'effet de l'hydratation préférentielle est mentionné pour son influence sur les interactions protéine – protéine pour d'autres co-solvants non-salins. Ainsi, les polyéthylènes glycols dont la taille est bien supérieure à celle des molécules du solvant sont exclus de la surface de la protéine qui est alors préférentiellement hydratée. Il en résulte une attraction entre les particules de protéine connue sous le nom de force de déplétion qui a été mise en évidence pour quelques protéines [Budayova *et al.*, 1999; Finet, 1999]. Par ailleurs, la

présence de méthyl-2-pentanediol-2,4 (MPD) hydrate préférentiellement la ribonucléase A et dans certaines conditions la contraint à cristalliser [Pittz et Timasheff, 1978].

Effets spécifiques

A ces mécanismes non-spécifiques vis-à-vis de la protéine s'ajoutent des effets mettant en jeu la nature de sa surface, comme la liaison de co-solvant. Qualifiés de spécifiques, ils peuvent être responsables des inversions ou modifications locales de l'ordre des séries lyotropiques.

Le $MgCl_2$ est un sel particulier : il se lie à la surface des protéines, tout en augmentant la tension superficielle de l'eau. Ces deux phénomènes ont des effets opposés sur les interactions protéine – protéine : à basse salinité, la solubilité des protéines est ainsi fortement diminuée et la tendance s'inverse à forte concentration de $MgCl_2$ [Arakawa et Timasheff, 1984; Timasheff, 1993].

En dehors de ce sel, l'accumulation de co-solvant salin n'est à notre connaissance pas reconnue pour provoquer des interactions protéine – protéine attractives. La *Hm* MalDH, protéine halophile dont les interactions fortes avec le sel ont été mesurées, nous permet de mettre en évidence ce phénomène. La relation entre les interactions protéine – solvant et les interactions protéine – protéine traduite par l'équation (4.9) (p.118) donne d'ailleurs un argument supplémentaire pour démontrer la spécificité des interactions protéine –sel avec en particulier une accumulation de sel à basses concentrations de $MgCl_2$ ou un appauvrissement à toute concentration de $(NH_4)_2SO_4$. Elle suggère aussi que la composition du solvant perturbé à fortes concentrations de NaCl et de CsCl est peu différente de celle du solvant de masse. Rappelons cependant que les études thermodynamiques permettent de mesurer des paramètres d'interaction préférentielle de la protéine avec le solvant et d'en déduire des implications relatives à sa solubilité, mais en aucun cas d'expliquer l'origine structurale de ces paramètres.

4.4.4. Halophilisme et interactions protéine – protéine

La thermodynamique nous montre que les interactions protéines – solvant sont une composante essentielle déterminant les interactions protéine – protéine. Toute différence de composition du solvant perturbé par la protéine par rapport au solvant de masse induit une attraction protéine – protéine. Pour qu'une protéine soit soluble dans un solvant, elle doit

donc rester le plus " neutre " possible vis-à-vis de celui-ci, *i.e.* elle doit éviter tout enrichissement ou appauvrissement en eau ou en co-solvant. Pour la *Hm* MalDH, les conditions de solvant où les paramètres d'interaction ($\partial m_3/\partial m_2$) sont les plus proches de zéro correspondent à des interactions protéine – protéine globalement répulsives ($A_2 >> 0$) (cf. Tableau 4.5 p.120) et à une forte solubilité de la protéine native. Effectivement, lors de la préparation de certains échantillons, de très fortes concentrations en protéine ont été atteintes sans problème dans les solvants où ($\partial m_3/\partial m_2$) est relativement faible : par exemple, 240 mg/ml en NaCl 4 M pH 8, 200 mg/ml en $MgCl_2$ 0.8 M pH 8 ou 150 mg/ml en $(NH_4)_2SO_4$ 1.5 M. A l'inverse, la protéine précipite à 100 et 10 mg/ml en $(NH_4)_2SO_4$ 3 et 4 M, concentrations pour lesquelles ($\partial m_3/\partial m_2$) est le plus négatif.

Cette étude met donc en évidence la nécessité pour une protéine d'un organisme halophile d'intégrer de fortes quantités de sels dans sa couche de solvatation afin de rester " neutre " par rapport au solvant et donc soluble. L'accumulation de charges de même signe à la surface de la protéine, caractéristique première des protéines halophiles, a dès lors deux avantages : elle augmente la composante répulsive coulombienne entre macromolécules voisines et autorise une solvatation riche en sel. En définitive, la préservation de la grande solubilité des macromolécules qui est nécessaire au bon fonctionnement de l'organisme apparaît comme une contrainte forte et donc déterminante des mécanismes moléculaires d'adaptation à l'environnement.

5. COMPREHENSION DE LA CRISTALLISATION DE LA *HM* MALDH PAR DILUTION

5.1. Présentation de l'article

La *Hm* MalDH a été cristallisée dans une solution aqueuse de NaCl et de méthyl-2-pentanediol-2,4 (MPD) [Dym *et al*., 1995; Richard *et al*., 1995]. La cristallisation est généralement associée à l'augmentation des concentrations de la protéine et de l'agent précipitant, conduisant à une solution sursaturée en protéine. Dans notre protocole, la *Hm* MalDH cristallise par dilution, selon la méthode de diffusion de vapeur. La goutte initiale contient la protéine en présence d'eau, de NaCl et de MPD. Elle est mise en équilibre avec un réservoir ne contenant que du MPD aqueux. La vapeur d'eau diffuse du réservoir vers la goutte et la dilue. Il faut noter que selon la concentration en sel et en MPD, le système NaCl – MPD – H_2O est mono-, bi- ou triphasique. Dans notre protocole, la goutte initiale est biphasique [Richard *et al*., 1995]. Nous avons tout d'abord cerné les zones du diagramme de phases où la protéine est stable, puis celles où sa solubilité est fortement diminuée. Nous avons ensuite évalué le second coefficient du viriel de la *Hm* MalDH dans divers mélanges NaCl – MPD – H_2O monophasiques. Enfin, nous avons décrit le cheminement complexe de la protéine dans le diagramme de phases, lors du processus de cristallisation par dilution. Ces résultats nous permettent de comprendre les raisons de la cristallisation de la *Hm* MalDH. Ils sont brièvement rappelés ci-dessous et développés dans l'article ci-joint : " Understanding the crystallisation of an acidic protein by dilution in the ternary NaCl – 2-methyl-2,4-pentanediol – H_2O system " [Costenaro *et al*., 2001].

La Figure 5.1 comprend l'essentiel de nos résultats.

Dans un premier temps, nous avons déterminé la stabilité apparente de la protéine dans diverses solutions aqueuses monophasiques de NaCl et de MPD en interprétant des mesures d'activité résiduelle (activité enzymatique après 18 heures d'incubation) en termes de protéine repliée ou dénaturée. En NaCl et sans MPD, la *Hm* MalDH est dénaturée en dessous de 2 M (symbole " i " dans la Figure 5.1). L'addition progressive de MPD permet à la protéine de rester active et stable à des concentrations salines de plus en plus faibles : elle est par exemple

toujours active en NaCl 0.3 M avec 30% de MPD*. En présence de NaCl, le MPD est donc un stabilisant de la *Hm* MalDH.

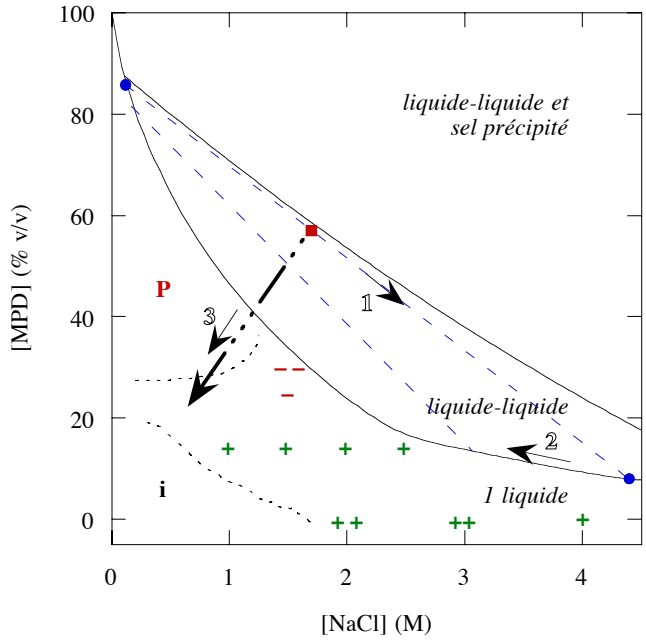


Figure 5.1 : Cheminement de la *Hm* MalDH dans le diagramme de phases NaCl – MPD – H_2O lors de sa cristallisation. Les traits pleins représentent les courbes de coexistence de phases à 22°C [d'après Richard *et al.*, 1995]. Les courbes en pointillés délimitent les zones où la *Hm* MalDH est inactive (i) et où elle précipite (P) à 25.5 mg/ml à température ambiante. Les valeurs expérimentales du second coefficient du viriel A_2 sont indiquées par " ++ " pour des valeurs supérieures à 4 10^{-5} mol.l.g^{-2}, par "+" pour des A_2 compris entre 2 et 4 10^{-5} mol.l.g^{-2} et par " – " et " – – " pour des A_2 d'environ -2 et -10 10^{-5} mol.l.g^{-2}. La goutte initiale (■) est constituée de deux phases (●). Durant la dilution globale de la goutte (flèche grasse hachurée), la composition de la phase contenant la protéine suit le cheminement 1, 2 et 3.

Dans un deuxième temps, des mesures de solubilité apparente à 25.5 mg/ml ont montré que la protéine ne précipite (symbole " P ") qu'en dessous de 1 M de NaCl avec au moins 30% de MPD.

La solubilité, la précipitation et la cristallisation des protéines sont reliées à la valeur et au signe du second coefficient du viriel A_2 [George et Wilson, 1994; Ebel *et al.*, 1999b]. Nous avons donc mesuré les valeurs de A_2 pour la *Hm* MalDH dans de nombreuses solutions monophasiques de NaCl – MPD – H_2O, par diffusion de neutrons aux petits angles. A

* Les concentrations en MPD sont exprimées comme le volume utilisé de MPD pur divisé par le volume final de la solution en pourcentage : % (v/v).

concentration constante de MPD, A_2 varie peu avec la concentration en sel. Par contre pour une concentration en NaCl donnée, il décroît de manière notable lorsque la concentration en MPD augmente, ce qui correspond à une augmentation des attractions protéine – protéine. En NaCl 1.5 M, A_2 évolue de 2.6 10^{-5} mol.ml.g^{-2} en MPD 5%, à -10 10^{-5} mol.ml.g^{-2} en MPD 30%.

Nous avons aussi déterminé les compositions des phases de deux solutions biphasiques de NaCl – MPD – H_2O dont celles de la goutte initiale. Finalement, des mesures d'absorbance pour les deux phases d'une solution biphasique de NaCl et de MPD contenant de la *Hm* MalDH indiquent très clairement que la protéine n'est présente que dans la phase saline. L'ensemble de ces résultats nous permet de décrire le chemin de la cristallisation de la *Hm* MalDH sur le diagramme de phases NaCl – MPD – H_2O.

Au début du processus de dilution globale de la goutte (1 sur la Figure 5.1), la phase fortement saline et pauvre en MPD qui contient la protéine évolue vers une composition moins riche en protéine et en sel, mais plus concentrée en MPD (2). La *Hm* MalDH est ainsi progressivement conditionnée d'interactions protéine – protéine répulsives ($A_2 > 0$) vers des interactions modérément attractives ($A_2 < 0$). La goutte devient ensuite monophasique. La nucléation se produit alors dans une solution homogène où A_2 est raisonnablement extrapolé à -20 10^{-5} mol.ml.g^{-2}. Cette valeur est en accord avec celles qui correspondent aux conditions de cristallisation de nombreuses protéines [George et Wilson, 1994]. L'équilibration se poursuivant, la protéine, le sel et le MPD sont tous trois dilués (3 sur la Figure 5.1). Les interactions protéine – protéine deviennent moins attractives, ce qui devrait conduire la protéine, de conditions propices à la nucléation, vers des conditions avantageant la croissance cristalline et donc favoriser l'apparition d'un nombre limité de cristaux de taille néanmoins raisonnable. Ce cheminement complexe permet de préserver la stabilité de la protéine durant tout le processus.

La *Hm* MalDH a été cristallisée selon le même protocole dans des mélanges de sel – MPD – H_2O avec le KCl et le KF [Richard, 1998]. Le principe de ce protocole pourrait être adapté pour la cristallisation de macromolécules stables et solubles à forte salinité et (ou) peu solubles à bas sel.

5.2. Article 2 : Understanding the crystallisation of an acidic protein by dilution in the ternary NaCl – 2-methyl-2,4-pentanediol – H_2O system*

Lionel Costenaro, Giuseppe Zaccai and Christine Ebel

Journal of Crystal Growth, 232(1-4) (2001), 102-13.

* Les références citées dans l'article sont listées à la fin de celui-ci, indépendamment de la bibliographie du manuscrit.

ELSEVIER

Journal of Crystal Growth 232 (2001) 102–113

JOURNAL OF CRYSTAL GROWTH

www.elsevier.com/locate/jcrysgro

Understanding the crystallisation of an acidic protein by dilution in the ternary NaCl–2-methyl-2,4-pentanediol–H_2O system

Lionel Costenaro[a], Giuseppe Zaccai[a,b], Christine Ebel[a,*]

[a] *LBM, Institut de Biologie Structurale, UMR 5075 CEA – CNRS UJF, 41 rue Jules Horowitz, 38027 Grenoble Cedex 1, France*
[b] *Institut Laue Langevin, BP 156X, 38042 Grenoble, France*

Abstract

Halophilic malate dehydrogenase is a negatively charged protein that crystallises well in a dilution process following a complex interplay with the three components of NaCl–MPD–H_2O solvents (MPD: 2-methyl-2,4-pentanediol). The process was characterised by measuring the folding state of the protein, its concentration, its apparent solubility and second virial coefficients in various NaCl–MPD–H_2O ratios representative of the phase diagram of the system. The protein crystallises by vapour diffusion between a drop containing protein in aqueous NaCl and MPD and a bath containing a given percentage of MPD in water. The starting drop is bi-phasic, in which the protein is concentrated in the salt-rich phase. This phase evolves towards a solution less concentrated in both NaCl and protein but more concentrated in MPD, before the system is driven to a single-phase region, where crystals are obtained. The protein stability is preserved during its crystallisation. We showed from second virial coefficient measurements that crystal formation is correlated with a slow evolution from repulsive to attractive protein–protein interactions. The end of the vapour diffusion process between the drop and the bath corresponds to a decrease of the attractive interaction, which we suggest favours crystal growth. The application of these results to crystallisation of other charged proteins and nucleic acids is discussed. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: A1. Biocrystallization; A1. Solubility; B1. Biological macromolecules; B1. Organic compounds; B1. Proteins; B1. Salts

1. Introduction

Halophilic archaea that grow in hypersaline environments such as the Great Salt Lake or the Dead Sea, balance the external osmotic pressure by accumulating intracellular KCl close to saturation. Their entire biological machinery must be adapted to high salt concentrations [1]. As an example, halophilic malate dehydrogenase (*Hm* MalDH[1]) from the archaeon *Haloarcula marismortui* is an enzyme that requires molar concentrations of KCl or NaCl to be stable and active [2]. *Hm* MalDH is a 130 kDa homotetramer [3] extremely rich in acidic residues (with an excess

*Corresponding author. Tel.: +33-4-38-78-96-38; fax: +33-4-38-78-54-94.
E-mail address: christine.ebel@ibs.fr (C. Ebel).

[1] Abbreviations: *Hm* MalDH, malate dehydrogenase from the halophilic archaeon *Haloarcula marismortui*; SANS, small angle neutron scattering; SAXS, small angle X-ray scattering; MPD, 2-methyl-2,4-pentanediol; NADH, reduced β-nicotinamide adenine dinucleotide.

0022-0248/01/$ - see front matter © 2001 Elsevier Science B.V. All rights reserved.
PII: S0022-0248(01)01150-2

of 156 negative charges), which are mainly localised on the protein surface [4]. It has been extensively studied by biochemical and biophysical approaches [5]. Its gene has been overexpressed in *Escherichia coli* and the protein, refolded by high salt concentration, is obtained in large amounts [6]. It is a good model for the study of the environment effects on protein folding, stability and macromolecular interactions. Inactivation and unfolding of *Hm* MalDH occur simultaneously. For various salts in a large range of concentrations, anions and cations with the highest charge density are more efficient to stabilise the folded protein. This effect was interpreted in terms of surface tension effects and ion binding [7]. In high NaCl or KCl concentrations, *Hm* MalDH shows remarkable protein–solvent interactions since it binds significantly large amounts of water and salt [3,8]. In recent crystallographic structures of native and mutant *Hm* MalDH, this tetramer appears to be stabilised by ordered water molecule networks and intersubunit complex salt bridges "locked in" by bound chloride and sodium ions [9]. In a further exploration of solvation, the weak protein–protein interactions of *Hm* MalDH have been characterised in various solvents and shown to be attractive in conditions that promote protein crystallisation [10].

Hm MalDH was first crystallised in molar phosphate buffer using the hanging drop method [11], but better results were obtained in aqueous NaCl and 2-methyl-2,4-pentanediol (MPD) mixtures [4]. MPD is a precipitant widely used to crystallise both proteins and nucleic acids (see, for example, Ref. [12]). The crystallisation protocol was studied and described in detail by Richard et al. [13]: *Hm* MalDH crystallises from a bi-phasic solution of protein at 10–20 mg ml^{-1} in 1.7 M NaCl, 57% MPD in vapour equilibrium with water—55–70% MPD using the sitting drop method. Water, the only volatile species, diffuses from the reservoir to the drop, diluting it by approximately a factor 3. By opposition to the classical vapour diffusion protocol used in crystallogenesis, in which the protein and precipitant concentrations are increased leading to a solution supersaturated in protein, this dilution process was named "reverse" vapour diffusion [14]. Many proteins and nucleic acids have been crystallised by a dilution process, in the presence or absence of organic solvent. McPherson has reviewed some of the protein crystals obtained by dialysis against water or less concentrated solvents [15]. As another example, crystals of yeast transfer RNA^{Asp} were obtained from a precipitate equilibrated against water [16]. More recently, T7 RNA polymerase has been crystallised by the "reverse" vapour diffusion method [14]. Until now, the mechanisms of these processes have not been fully investigated.

Pittz and Timasheff [17] studied the solvent interactions of ribonuclease A in MPD aqueous solutions, conditions which stabilise the folded protein and allow it to crystallise. They showed that the protein is preferentially hydrated and interpreted their results in terms of local repulsion of MPD by the surface charges on the protein. This unfavourable thermodynamic interaction would lead to phase separation and induce crystallisation. The same phenomenon occurs when salt is added to aqueous MPD solution and leads to MPD phase separation. Richard et al. [13] established the phase diagram of the NaCl–MPD–H_2O mixture used for the *Hm* MalDH crystallisation. It shows that, at low concentration of MPD or salt, all species are miscible, but that increasing MPD or NaCl leads in a first step to the coexistence of two liquid phases, and in a second step to additional salt precipitation.

George and Wilson [18] have shown that the capacity of proteins to remain soluble at high concentration, to form an amorphous precipitate, or to crystallise, is related to the strength and the sign of the second virial coefficient A_2. This thermodynamical parameter is obtained by small angle neutron scattering (SANS), small angle X-ray scattering (SAXS), light scattering or sedimentation equilibrium [19,20] experiments performed on solutions containing protein at various concentrations. The second virial coefficient reflects all the interparticle interactions: its value is positive for repulsive global protein–protein interactions, zero for ideal solutions and negative for attractive cases, the solvent being referred to, respectively, as a "good", "theta" and "poor" solvent [21]. For a number of proteins in solvents known to promote

crystallisation, the values of the second virial coefficient A_2 are moderately negative and lie within a fairly narrow range: between about -10×10^{-5} and -80×10^{-5} mol ml g^{-2} [18]. In this case, protein–protein interactions are sufficiently attractive to promote nucleation, but not so attractive that an amorphous precipitate occurs. Such precipitating conditions would correspond to larger negative values of A_2. For example, a A_2 value of -470×10^{-5} mol ml g^{-2} was measured for Concanavalin A in conditions producing an amorphous precipitate, to be compared to about -20×10^{-5} mol ml g^{-2} for crystallisation conditions [22]. The correlation between protein solubility and second virial coefficient is confirmed by other studies on lysozyme and ovalbumin [23,24] and on *Hm* MalDH [10].

In the present work, we investigated the *Hm* MalDH crystallisation process by the study of the intermolecular protein–protein interactions in solution. From SANS experiments in solution, we derived the second virial coefficient A_2 of the protein in various solvents. The crystallisation of *Hm* MalDH is described with respect to the NaCl–MPD–H_2O phase diagram.

2. Materials and methods

2.1. Sample preparation

All solutions contained 50 mM Tris-HCl and were buffered at pH 8. MPD concentration is expressed as the volume of pure MPD divided by the final volume of the solution in percentage (% (v/v)), and NaCl concentration on the molar scale (1 M = 1 mol l^{-1}). *Hm* MalDH concentration was determined spectrophotometrically by using an extinction coefficient $E_{280}^{0.1\%}$ of 0.85 ml mg^{-1} cm^{-1} [3].

Hm MalDH, overexpressed in *Escherichia coli*, was purified according to the protocol of Cendrin et al. [6]. It was stored in 4 M NaCl at 4°C. Before the SANS experiments, the protein was passed through a gel filtration column (Sephacryl 300 from Pharmacia) and reconcentrated by ultrafiltration over Centricon 30 membranes (Amicon) to about 240 mg ml^{-1}. It was dialysed at 4°C for two days in 2 M NaCl with three changes of the solvent. The samples were obtained by dilutions (monitored by weighing) of this protein solution with three buffered stock solutions (5 M NaCl, 80% MPD and H_2O) and followed by dilution with the solvents prepared in the same way, in order to obtain samples at various protein concentrations.

2.2. Protein stability in NaCl–MPD aqueous solutions

The apparent stability of the enzyme was monitored after 2 and 18 h of incubation at room temperature, in solvents of various compositions, by a standard activity measurement. The protein solution at 0.03 mg ml^{-1} (10 μl) was diluted in 1 ml of freshly prepared 2 M NaCl, 0.2 mM NADH, and 0.3 mM oxaloacetate in a 1.00 cm quartz cuvette. The rate of NADH oxidation was immediately recorded over a period of 15 s by the decrease of the absorbance (A) at 340 nm as a function of time (t), with a Beckman DU 7400 spectrophotometer. The dA/dt values, corrected for the blank, were compared to those obtained from a sample at the same protein concentration incubated in parallel in 4 M NaCl, solvent in which *Hm* MalDH is fully active and stable [2]. The residual activity measurements were interpreted as reflecting the proportion of folded or unfolded states of the protein.

2.3. Protein precipitation in NaCl–MPD aqueous solutions

A stock solution of *Hm* MalDH at 170 mg ml^{-1} in 2 M NaCl was diluted with appropriate amounts of 5 M NaCl, 80% MPD and 50 mM Tris-HCl pH 8 to obtain samples at a protein concentration of 25.5 mg ml^{-1}, in solvents of various salt and MPD concentrations. Linbro boxes were used to equilibrate 8 μl of these samples against 0.8 ml of solvents of the same composition. The sitting drop technique allowed an easy observation of the drop for months. The visual inspection of the drops did not reveal volume variations, and we assumed there was no water diffusion. The macroscopic aggregates were noted.

2.4. Small angle neutron scattering (SANS) experiments

The scattering curves were recorded at the Institut Laue Langevin (Grenoble, France) on the instrument D11 [25] using 1.00 mm path quartz cuvettes at 20°C, with a neutron wavelength λ of 6 Å ($\Delta\lambda/\lambda = 0.1$). Sample detector distances of 4.5 and 1.25 m with collimations of 5.5 and 13 m, respectively, allowed a scattering vector (Q) range of 0.008–0.3 $Å^{-1}$ ($Q = 4\pi \sin\theta/\lambda$, where θ is half the scattering angle). The data were radially averaged and normalised using the scattering of 1.00 mm of pure water and the sample transmissions [26], taking into account the detector efficiency [27]. The scattering from the empty cuvette, electronic background and solvents was subtracted by using conventional procedures [28]. The scattering levels of the protein and the solvent were checked to be similar at large angles ($Q = 0.28-0.3\,Å^{-1}$) and when necessary they were corrected for small differences.

2.5. Data interpretation of SANS

The Guinier approximation [29] (according to Eq. (1)) allows, for ideal solutions, determination of the radius of gyration Rg of the particle scattering density contrast and extrapolation of the forward intensity $I(Q = 0)$ from a linear fit of $\ln I(Q)$ vs. Q^2 for $RgQ < 1.3$:

$$\ln I(Q) = \ln I(Q = 0) - Rg^2\, Q^2/3. \quad (1)$$

The forward intensity $I(Q = 0)$ is related, through the protein concentration c in mg ml^{-1}, a normalisation factor k, and Avogadro's number N_A, to the molar mass M in g mol^{-1} and to the neutron scattering length density increment, $(\partial\rho_N/\partial c)_\mu$ in cm g^{-1} [20]:

$$I(Q = 0)/c = k(1/N_A)M(\partial\rho_N/\partial c)_\mu^2. \quad (2)$$

$(\partial\rho_N/\partial c)_\mu$ corresponds to the increment of the neutron scattering length density ρ_N produced by the presence of the protein, at constant chemical potential μ of diffusible solutes. It is the contrast term. It is not only related to the solvent and protein characteristics (composition and volume), but also to all perturbations of the solvent due to the protein presence [19,20,30].

For real solutions, Rg in Eq. (1) and M Eq. (2) are to be replaced by apparent values Rg_{app} and M_{app}. A virial expansion of $1/M_{app}$ in powers of protein concentration c yields:

$$(1/M_{app}) = (1/M) + 2A_2c + 3A_3c^2 + o(c^3), \quad (3)$$

where A_2 (in mol ml g^{-2}) is the second virial coefficient. When protein concentration is moderate, the virial coefficients of order higher than 2 can be neglected. Substituting this expansion in Eq. (2) for real solutions leads to:

$$\frac{c}{I(Q = 0)} = \frac{1}{I(0,0)} + \frac{2A_2M}{I(0,0)}c, \quad (4)$$

where $I(0,0)$ (for $c = 0$ and $Q = 0$) is the limit of $I(Q = 0)/c$ at zero protein concentration (ideal solution). The plot of $c/I(Q = 0)$ vs. c can then be fitted by a straight line of slope proportional to A_2.

The Guinier plots were done in a RgQ range between 0.2 and 1.3. The second virial coefficients A_2 were obtained from linear fits of $c/I(Q = 0)$ vs. c with three protein concentrations between 5 and 17 mg ml^{-1} (except for 1.5 M NaCl, 30% MPD where a second-degree polynomial fit with four concentrations up to 62 mg ml^{-1} was performed). A molecular mass M of 130 552 Da was used, i.e. the molecular mass of *Hm* MalDH without its counter-ions. In order to take into account the particle electroneutrality, A_2 values would have to be multiplied by $(M/M_+)^2$, where M_+ is the molecular mass of the protein with its counter-ions ($M_+ = 134\,138$ Da with 156 Na^+).

2.6. Characterisation of the Hm MalDH crystallisation pathway

The starting crystallisation drop (1.7 M NaCl, 57% MPD) corresponds to a two liquid phase solution, in which the protein had to be precisely localised. A bi-phasic solution of 0.6 ml, containing 2 M NaCl, 40% MPD and protein at 0.85 mg ml^{-1} was equilibrated for 2 h at room temperature in a 1.5 ml micro-tube. A volume of 0.1 ml for each phase was taken and diluted with 0.05 ml of water to stabilise the solution. The same was done with a sample of the same composition

but without protein, to provide suitable background. A wavelength absorbance spectrum of each phase was measured using a 50 µl microcuvette of 1.00 cm path length. The protein concentration was obtained from absorbance at 280 nm.

From the NaCl–MPD–H_2O phase diagram [13] and the measurement of the volume of each phase of a bi-phasic mixture, we can estimate their respective compositions. They are given by the intersections of the tie line (passing through the global concentration point) with the coexistence curve between one-liquid and two-liquid phases in the phase diagram (Fig. 3). A tie line is the set of global composition points that gives phases of same compositions. The volumes of these phases are different for each point of the tie line (see, for example, Ref. [31]).

To determine experimentally the tie line, we consider that a bi-phasic solution of volume V^{global} is globally composed of NaCl and MPD at respective concentrations: $c^{\text{global}}_{\text{NaCl}}$ and $c^{\text{global}}_{\text{MPD}}$. It is composed of two phases: phase A rich in MPD and phase B rich in NaCl. Phase A floats on phase B because of the lower density of MPD compared to NaCl aqueous solution. Phases A and B contain both NaCl and MPD, at respective concentrations $c^{\text{A}}_{\text{NaCl}}$ and $c^{\text{A}}_{\text{MPD}}$, and $c^{\text{B}}_{\text{NaCl}}$ and $c^{\text{B}}_{\text{MPD}}$. In each phase, these concentrations are related by the coexistence curve between one-liquid and two-liquid phases. They are calculated by the iterative process described below. The superscript indicates the phase and the subscript, the component considered; (i) is the iteration number. The NaCl and MPD concentrations are related by:

$$c^{\text{A}(i)}_{\text{MPD}} = \frac{c^{\text{global}}_{\text{MPD}} V^{\text{global}} - c^{\text{B}(i\text{-}1)}_{\text{MPD}} V^{\text{B}}}{V^{\text{A}}}, \quad (5)$$

$$c^{\text{B}(i)}_{\text{NaCl}} = \frac{c^{\text{global}}_{\text{NaCl}} V^{\text{global}} - c^{\text{A}(i)}_{\text{NaCl}} V^{\text{A}}}{V^{\text{B}}}, \quad (6)$$

in which V^{A} and V^{B} are the (measured) volumes of phases A and B respectively.

We start the iterative calculations with the hypothesis that there is no MPD in phase B (i.e. $c^{\text{B}(0)}_{\text{MPD}} = 0$), Eq. (5) allows a first estimation of the MPD concentration of the phase A ($c^{\text{A}(i=1)}_{\text{MPD}}$). This concentration, reported on the coexistence curve on the phase diagram, gives the corresponding NaCl concentration of the phase A ($c^{\text{A}(i=1)}_{\text{NaCl}}$). From this value, the NaCl concentration of the phase B ($c^{\text{B}(i=1)}_{\text{NaCl}}$) is calculated with Eq. (6) and gives graphically the corresponding MPD concentration ($c^{\text{B}(i=1)}_{\text{MPD}}$) from the fact that B is also on the coexistence curve. The iterative process goes on with $i = 2$ in Eq. (5). Convergence was obtained after four iterations.

3. Results

3.1. Stability of Hm MalDH in various NaCl–MPD mixtures

The solvents in which the protein was found to be active or non-active are described in Table 1 (respective symbols "a"or "i"). As shown by Ebel et al. [7 and Refs. therein], inactivation and unfolding of *Hm* MalDH occur together. The solutions in which the protein is active can be thus considered as solutions in which the protein is stable. In NaCl pH 8 without MPD, *Hm* MalDH is denaturated below 2 M at 25°C [32]. Our results show that the addition of MPD allows the protein to remain active and stable at progressively lower salt concentrations: for instance, it is still active at 0.3 M NaCl with 30% MPD. In the presence of NaCl, therefore, MPD is a stabiliser for *Hm* MalDH.

3.2. Precipitation of Hm MalDH in various NaCl–MPD mixtures

Hm MalDH solubility was tested for a number of mono-phasic NaCl–MPD aqueous solutions, in which the protein is stable. The protein was diluted at room temperature to a final concentration of 25.5 mg ml^{-1} and we observed if it precipitated or not. In the case of precipitation, we noted that it occurred immediately (less than 10 s) after the dilution of the protein. The soluble or precipitated state was maintained for months. The precipitate could be resolubilised in an active form by dilution in 4 M NaCl. We thus consider this precipitation as related to the auto-association of folded proteins and not to their irreversible denaturation.

Table 1
Stability and apparent solubility of *Hm* MalDH in various ternary NaCl–MPD–H_2O solvents[a]

% (v/v) MPD	NaCl concentration (M)							
	0	0.3	0.5	1	1.5	2	2.5	3
65		P	≈					
60	a		P					
50	a	a P	P	≈				
40	i		P	P	≈			
30		a P	a P	P	S	≈	≈	
15	i		i	a S	a S	S	S	≈
5		i		i	a S	S		S
0	i	i	i	i	i	a S	a S	a S

[a] "a", the protein remains active; "i", it is inactivated after incubation in the solvent at room temperature and pH 8. "P", the protein at 25.5 mg ml^{-1} precipitates or "S" remains soluble; " ≈ " indicates two phases in the solvent at 22°C [13].

These experiments provided a default value for *Hm* MalDH solubility: when it precipitates, its solubility is under 25.5 mg ml^{-1}. Our results show that the folded protein precipitates (symbol "P" in Table 1) at low salt (below about 1 M) and high MPD concentrations (above about 30%), and remains soluble (symbol "S") in other tested conditions. It has to be noted that Ebel et al. [10] were able to obtain supersaturated *Hm* MalDH at high concentration (93.4 mg ml^{-1}) in 1 M NaCl, 30% MPD at 25°C with a sample prepared by extensive dialysis and which led to crystallisation only after some days. This difference is related also to a value of the second virial coefficient A_2 lower by a factor of two or three than the A_2 expected from the series of measurements presented here (see below), and probably linked to differences in the sample preparations.

3.3. Second virial coefficients from SANS experiments

SANS experiments were performed for monophasic NaCl–MPD–H_2O conditions in which *Hm* MalDH is stable and soluble at 25.5 mg ml^{-1} (symbols "a" and "S" in Table 1). Fig. 1 shows typical series of data. The solvent 1.5 M NaCl, 30% MPD (panels a) is attractive for protein–protein interactions since the apparent radius of gyration increases with protein concentration and $c/I(Q=0,c)$ decreases: the corresponding A_2 value is negative. An opposite behaviour occurs in 1 M NaCl, 15% MPD (panels b), corresponding to a repulsive solvent and a positive A_2.

The extrapolation at zero protein concentration of the apparent radius of gyration Rg_{app} for all tested solvent conditions was found to be close to the value of 30 Å corresponding to the tetramer. When NaCl or MPD concentrations were modified, the extrapolation at zero protein concentration of the forward scattered intensity $I(0,0)$ evolved only slightly (10%). These observations support the interpretation that a tetrameric *Hm* MalDH is present in solution whatever the tested solvent composition. This reinforces our interpretation of the evolution of the forward intensity not in terms of discrete association but in terms of second virial coefficients, i.e. weak protein–protein interactions, which contribute more and more with an increase of the protein concentration. This point has already been discussed for lysozyme [33].

The values of the second virial coefficients A_2, calculated with Eq. (4), are shown in Table 2. The A_2 values in 2–5 M NaCl (without MPD) are reported from Ebel et al. [10]. Even if obtained from SANS experiments performed at 5°C, they can be compared with the values of experiments at room temperature because, in absence of MPD, A_2 does not vary significantly in this range of temperature. This was proved by SAXS experiments performed by us at 5°C, 15°C, 25°C, 35°C and 45°C in 4 M NaCl (data not shown), at the "Laboratoire d'Utilisation du Rayonnement

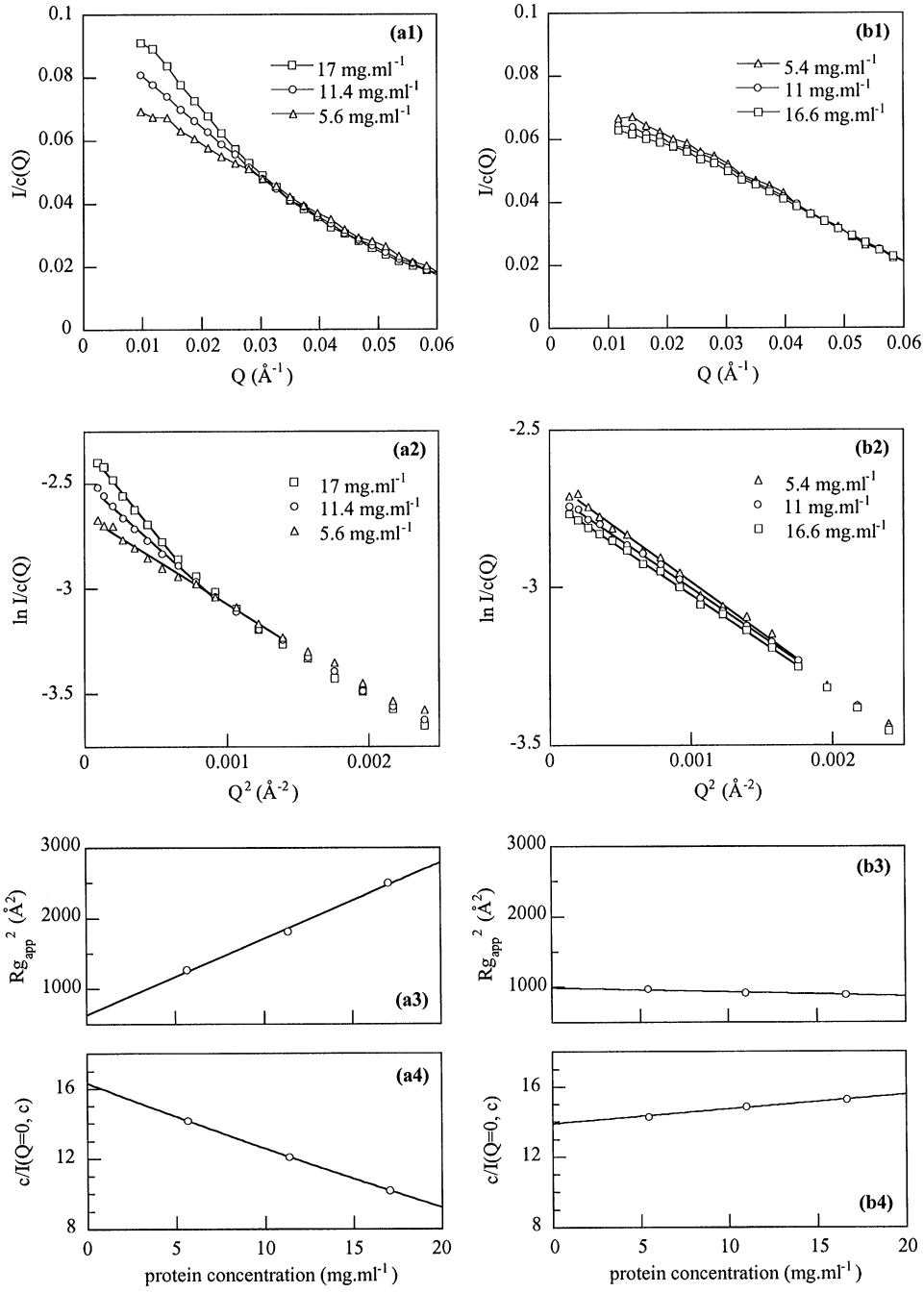


Fig. 1. SANS data for *Hm* MalDH at various protein concentrations in 1.5 M NaCl, 30% MPD (panels a) and in 1 M NaCl, 15% MPD (panels b). Panels a1 and b1 show the scattering intensities $I(Q)/c$ normalised by the protein concentration, and panels a2 and b2 the respective Guinier plots. Panels a3 and b3 show the variations of Rg^2_{app} with the protein concentration c, and panels a4 and b4 the variations of $c/I(Q=0,c)$ with c.

Table 2
Second virial coefficients (in 10^{-5} mol ml g^{-2} with error in brackets) of *Hm* MalDH in various ternary NaCl–MPD–H_2O solvents

% (v/v) MPD	NaCl concentration (M)						
	1	1.5	2	2.5	3	4	5
30		−10.0 (0.1)					
25		−2.3 (0.2)					
20			2.0 (0.1)				
15	2.4 (0.3)	2.1 (0.7)	2.5 (0.4)	2.9 (0.9)			
5		2.6 (0.8)					
0			4.7[a]		4.3[a]	3.4[a]	2.7[a]

[a] Values from Ebel et al. [10], obtained at 5°C.

Electromagnétiqué'' (LURE, Orsay, France), on the D24 beam line [34,35].

At constant MPD concentration, the second virial coefficient does not vary much with the salt concentration. In absence of MPD, A_2 decreases slightly from 4.7×10^{-5} mol ml g^{-2} at 2 M NaCl to 2.7×10^{-5} mol ml g^{-2} at 5 M NaCl. At 15% of MPD, A_2 values are constant within experimental errors for an NaCl range of 1–2.5 M. Note that the excluded volume effect would lead to a value of $A_2 = 4\bar{v}/M = 2.2 \times 10^{-5}$ mol ml g^{-2}, considering a partial specific volume of the protein $\bar{v} = 0.74$ ml g^{-1} [36]. This suggests additional electrostatic repulsion, present even at 2 M salt and decreasing at higher salt concentrations.

Whatever the salt concentration, rising MPD concentration leads to a notable decrease of the second virial coefficient, corresponding to increased protein–protein attractions. The transition between globally repulsive to globally attractive protein–protein interactions, corresponding to a zero value for A_2, would be found at about 20% MPD. At 30% MPD, the interactions are clearly attractive, as seen by a strongly negative value of A_2 (-10×10^{-5} mol ml g^{-2}) or by the protein precipitation (see Table 1).

3.4. Initial crystallisation conditions

The initial conditions of the drop (1.7 M NaCl, 57% MPD) correspond to a bi-phasic solution composed of an ''organic'' phase, rich in MPD and poor in salt, and a ''salt'' phase, poor in MPD. Neglecting the protein effect, the compositions of these two phases were estimated from the visual measurement of their respective volumes (6.3 and 3.7 ml) to be 0.12 M NaCl, 85.8% MPD for the ''organic'' phase; and 4.4 M NaCl, 7.9% MPD for the ''salt'' phase. For a global composition of 1.5 NaCl, 50.3% MPD the compositions of two phases were 0.17 M NaCl, 81.7% MPD (5.4 ml) and 3.1 M NaCl, 13.5% MPD (4.3 ml). These two tie lines are represented on Fig. 3.

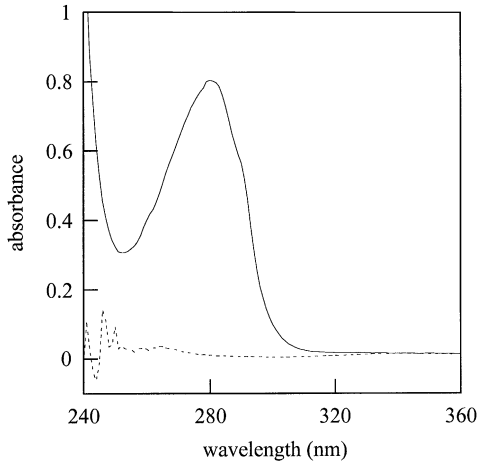


Fig. 2. Wavelength absorbance spectrum of the two phases of a bi-phasic solution of 2 M NaCl, 40% MPD and *Hm* MalDH at 0.85 mg ml^{-1}. The solid curve corresponds to the ''salt'' and the dashed to the ''organic'' phase. The protein remains entirely in the ''salt'' phase.

The spectroscopic measurements performed starting from bi-phasic 2 M NaCl, 40% MPD

solution showed that *Hm* MalDH entirely remains in the "salt" phase. The amount of protein in this phase is identical to the global one. Protein was not detected in the "organic" phase (Fig. 2).

4. Discussion

4.1. Crystallisation pathway during dilution in the ternary NaCl–MPD–H_2O system

In Fig. 3, the pathway of *Hm* MalDH during its crystallisation by dilution in the presence of NaCl and MPD is schematised. The bold dotted arrow corresponds to the evolution of the global composition of the drop during the equilibrium process, leading to a dilution by a factor of about three. It corresponds to the conditions that gave the crystals used for the structure determination of the protein and mutants by X-ray crystallography [4,9]. The starting drop is composed of 1.7 M NaCl and 57% MPD (□ in Fig. 3), and separates into two phases following the tie line that we experimentally determined (dashed line). The composition of each phase is found on the coexistence curve separating the one-liquid and two-liquid zones described by Richard et al. [13] (indicated by ○). From our spectroscopic measurements, we are able to localise the protein as concentrated in the "salt" phase (arrow ①). This behaviour is in agreement with the conclusion of Pittz and Timasheff [17]: charged molecules (i.e. salt and charged protein) repel MPD.

During the "reverse" vapour equilibrium process, water, the only volatile species, diffuses from the reservoir to the drop. The composition of each phase evolves along the coexistence curve. The "salt" phase, containing the protein, evolves towards a solution less concentrated in both NaCl and *Hm* MalDH but more concentrated in MPD (② in Fig. 3). In this second step, the protein submitted to a dilution process experiences an increasing concentration of the precipitant MPD.

When the drop becomes mono-phasic, the equilibrium process dilutes all the species (protein, salt and MPD) (③ in Fig. 3). Suitable crystals are obtained during this process. Advantages of this process are discussed in detail below.

4.2. Preservation of the protein stability during the crystallisation

It is obvious that the protocol of crystallisation by dilution of *Hm* MalDH allows it to avoid the region of the phase diagram where it is inactivated and unstable (symbol "i" in Fig. 3). Furthermore, whereas the salt concentration decreases during the process (②), the *Hm* MalDH stability is preserved by the presence of MPD. The understanding of the stabilisation of the folded form of *Hm* MalDH by MPD in the presence of salt is beyond the objectives of this paper. Solvent perturbations by the folded and unfolded forms are to be taken into account, which means the redistribution of water, salt and MPD. Pittz and Timasheff [17] emphasised the preferential hydra-

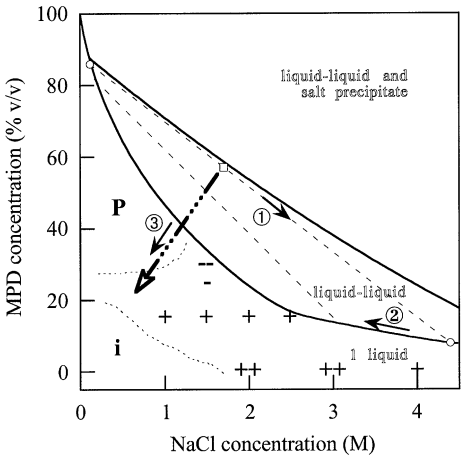


Fig. 3. The crystallisation pathway of *Hm* MalDH during its crystallisation by dilution on the NaCl–MPD–H_2O phase diagram. The solid lines represent the phase coexistence curves at 22°C from Richard et al. [13] and are extrapolated for salt concentrations above 3 M. Dashed lines represent tie lines. Dotted curves delimit the regions where the protein is inactivated and unstable (i), and where it precipitates (P) at 25.5 mg ml^{-1}, at room temperature. The starting drop is composed of 1.7 M NaCl and 57% MPD (□). The compositions of the two phases in the starting drop are given by the intersections of its tie line with the coexistence curve (○). During the global dilution of the drop (bold dashed arrow), the composition of the phase containing the protein follows the pathway ①, ② and ③. The values of the measured second virial coefficient A_2 of *Hm* MalDH are indicated by "++" for values above 4×10^{-5} mol ml g^{-2}, by "+" for a A_2 range of 2–4×10^{-5} mol ml g^{-2} and by "–" and "– –" for values of about -2 and -10×10^{-5} mol ml g^{-2}, respectively.

tion related to the repulsion of MPD by charged moieties. The addition of MPD also lowers the dielectric constant of the solvent [37] from which ion-pair formation is favoured according to the ion-pair theory.

4.3. The second virial coefficient, an actor in crystallogenesis

The first step in the crystallisation process (① in Fig. 3) is the separation of the drop into two phases. *Hm* MalDH is concentrated in a solvent in which the second virial coefficient A_2 is positive, i.e. the protein is highly soluble. During the dilution process of the bi-phasic drop along the coexistence curve (②), *Hm* MalDH is progressively conditioned from repulsive conditions ($A_2 > 0$) to moderately attractive ones ($A_2 < 0$), due to the large increase of MPD concentration. In the same time, the NaCl concentration also decreases largely but the effects of this dilution on protein–protein interactions (that would favour an increase of A_2) are minor compared to those due to MPD increase. When the drop becomes mono-phasic (③), the second virial coefficient has a value, reasonably extrapolated from Table 2, of about -20×10^{-5} mol ml g^{-2}, which is in good agreement with the narrow range found to correspond to crystallisation conditions by George and Wilson [18]. Richard [38] observed that it is necessary to begin with a bi-phasic drop to crystallise *Hm* MalDH. From our results, it appears that it is a way to slowly condition the protein from positive to moderately negative A_2. Crystal formation would be favoured by such a slow evolution from repulsive to attractive protein–protein interactions.

4.4. Nucleation and crystal growth

The nucleation occurs near the zone where the drop becomes mono-phasic and is followed by crystal growth [38]. Under usual crystallisation processes, the increase of precipitant and (or) protein concentrations progressively brings the protein, through a metastable zone, into the nucleation zone, which is not necessarily the optimal zone for crystal growth [39]. In our protocol, *Hm* MalDH is also brought into the nucleation zone by an increase in the precipitant concentration during the second step of the process (② in Fig. 3). But the subsequent dilution of the drop, decreasing both protein and MPD concentrations (③), corresponds to protein–protein interactions that become less attractive (see Table 2). We believe it allows the protein to move from the nucleation zone into one more favourable to crystal growth and favours the production of a limited number of crystals of reasonable size.

4.5. A more general system for crystallisation?

The "reverse" vapour diffusion method with the NaCl–MPD–H_2O system allows *Hm* MalDH to crystallise through a tortuous way in which salt and organic solvent concentrations evolve in opposite directions. Note that dialysis, or other methods have been also used for the same purpose (see, for example, Refs. [15,40]). The phase separation of MPD in the presence of NaCl allows this complicated evolution by the use of a simple experimental protocol. Could it be generalised to other salts and macromolecules?

Richard [38] was able to crystallise *Hm* MalDH with this dilution protocol with KCl and KF instead of NaCl. We have begun the comparison of the behaviour of *Hm* MalDH in various salt–MPD mixtures. Our preliminary results show similar phase diagrams for MPD in presence of CsCl compared to NaCl and KCl. *Hm* MalDH has been described to be soluble and stable in high salt concentrations of all these salts [7]. Its solvation is similar in NaCl and KCl solutions [3]. We found that the protein–protein interactions are characterised by second virial coefficients similar to those already measured in NaCl ($4.9–3.5 \times 10^{-5}$ mol ml g^{-2}) and in 1–4.5 M KF, at 5°C [10]. We measured A_2 values of 5.7×10^{-5} mol ml g^{-2} in 3 M KCl, at 20°C and $5.9–4.8 \times 10^{-5}$ mol ml g^{-2} in 3–6 M CsCl, at 25°C.

This "reverse" vapour diffusion method using ternary salt–MPD–H_2O systems might be useful to crystallise other macromolecules that are stable at high salt concentrations but have low solubility at lower salt, such as halophilic proteins. A number of non-halophilic proteins also show this

behaviour. As an example, the M protein of vesicular stomatitis virus aggregates at low NaCl concentrations [41]. This method could also be tested for the crystallisation of nucleic acids for which alcohol and especially MPD are widely used in the presence of salt [42].

The rational generalisation of this protocol of crystallisation requires the fine understanding of the effect of solvent composition on protein–protein interactions. Concerning the effect of the nature of the salt, our first results on *Hm* MalDH in the presence of ammonium sulphate and MPD (not shown here) show a different behaviour when compared to NaCl, which could be related to the observation of Papanikolau and Kokkinidis [43]. They were able to derive two zones of low protein solubility in aqueous salt–organic solvent mixtures: one for the "salting-in" conditions and one for the "salting-out". In "salting-in" zone, increasing the organic solvent concentration or reducing the salt concentration reduces the macromolecular solubility, as obtained for *Hm* MalDH in NaCl–MPD. For "salting-out" conditions, the decreased solubility arises from decreasing organic solvent and increasing salt concentration. Their interpretation is based on Green's general equation and ion-pair theory: the organic solvent increases electrostatic interactions between macromolecules and decreases polar ones. It does not take explicitly into account the redistribution of water and the organic solvent around the protein suggested by Timasheff and co-workers [17,44]. In the case of charged macromolecules, such as halophilic proteins or nucleic acids, salt and water redistribution depends on the salt nature and concentration [3,8,30,45]. We indeed measured solvent–protein interactions for *Hm* MalDH in the absence or presence of MPD that are very different in NaCl or ammonium sulphate (to be published). With the aim of furthering on understanding of these mechanisms, we are presently extending our investigations. The interpretation of SANS and SAXS curves in terms of protein–protein interaction potentials has recently been proved to allow the description of protein solubility [46–48]. We shall use this approach on *Hm* MalDH in the presence of MPD and salts other than NaCl.

Acknowledgements

We are very grateful to Drs. Patrice Vachette and Françoise Bonneté for their help in SAXS experiments at LURE (Orsay, France). Thanks are due to the staff of the D11 instrument and to the Institut Laue Langevin (Grenoble, France) for providing access to SANS facilities. Acknowledgements are also due to Alexandra Solovyova and Germain Esteves for help in the phase diagram determination of salt–MPD–H_2O mixtures and in the protein purification. We are very grateful to Prof. Claude Pfister for a critical reading of the manuscript.

References

[1] D. Madern, C. Ebel, G. Zaccai, Extremophiles 4 (2000) 91.
[2] M. Mevarech, E. Neumann, Biochemistry 16 (1977) 3786.
[3] F. Bonneté, C. Ebel, H. Eisenberg, G. Zaccai, J. Chem. Soc. Faraday Trans. 89 (1993) 2659.
[4] O. Dym, M. Mevarech, J.L. Sussman, Science 267 (1995) 1344.
[5] F. Bonneté, D. Madern, G. Zaccai, J. Mol. Biol. 244 (1994) 436.
[6] F. Cendrin, J. Chroboczek, G. Zaccai, H. Eisenberg, M. Mevarech, Biochemistry 32 (1993) 4308.
[7] C. Ebel, P. Faou, B. Kernel, G. Zaccai, Biochemistry 38 (1999) 9039.
[8] B. Kernel, G. Zaccai, C. Ebel, Prog. Colloid Polym. Sci. 113 (1999) 168.
[9] S.B. Richard, D. Madern, E. Garcin, G. Zaccai, Biochemistry 39 (2000) 992.
[10] C. Ebel, P. Faou, G. Zaccai, J. Crystal Growth 196 (1999) 395.
[11] M. Harel, M. Shoham, F. Frolow, H. Eisenberg, M. Mevarech, A. Yonath, J. Sussman, J. Mol. Biol. 200 (1988) 609.
[12] Y. Timsit, D. Moras, Methods Enzymol. 211 (1992) 409.
[13] S. Richard, F. Bonneté, O. Dym, G. Zaccai, in: S. Dassarma (Ed.), Archaea: A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York, 1995, p. 149.
[14] D. Jeruzalmi, T.A. Steitz, J. Mol. Biol. 274 (1997) 748.
[15] A. McPherson, Preparation and Analysis of Protein Crystals, 2nd Edition, R.E. Krieger, Malabar, Florida, 1989.
[16] R. Giegé, D. Moras, J.C. Thierry, J. Mol. Biol. 115 (1977) 91.
[17] E.P. Pittz, S.N. Timasheff, Biochemistry 17 (1978) 615.
[18] A. George, W.W. Wilson, Acta Crystallogr. D 50 (1994) 361.
[19] H. Eisenberg, Biological Macromolecules and Polyelectrolytes in Solution, Clarendon Press, Oxford, 1976.

[20] H. Eisenberg, Q. Rev. Biophys. 14 (1981) 141.
[21] W.G. McMillan, J.E. Mayer, J. Chem. Phys. 13 (1945) 276.
[22] V. Mikol, E. Hirsch, R. Giegé, J. Mol. Biol. 213 (1990) 187.
[23] F. Bonneté, S. Finet, A. Tardieu, J. Crystal Growth 196 (1999) 403.
[24] B. Guo, S. Kao, H. McDonald, A. Asanov, L.L. Combs, W.W. Wilson, J. Crystal Growth 196 (1999) 424.
[25] P. Lindner, R.P. May, P.A. Timmins, Physica B 180–181 (1992) 967.
[26] B. Jacrot, G. Zaccai, Biopolymers 20 (1981) 2413.
[27] P. Lindner, Internal publication 98LI12T, Institut Laue Langevin, Grenoble, France, 1998.
[28] R.E. Ghosh, S.U. Egelhaaf, A.R. Rennie, Internal publication 98GH14T, Institut Laue Langevin, Grenoble, France, 1998.
[29] A. Guinier, G. Fournet, Small-angle scattering of X-rays, Wiley, New York, 1955.
[30] C. Ebel, Progr. Colloid Polym. Sci. 99 (1995) 17.
[31] P.W. Atkins, in: Physical chemistry, Oxford University Press, Oxford, 1994, p. 239.
[32] D. Madern, C. Pfister, G. Zaccai, Eur. J. Biochem. 230 (1995) 1088.
[33] M. Muschol, F. Rosenberger, J. Crystal Growth 167 (1996) 738.
[34] C. Depautex, C. Desvignes, P. Leboucher, M. Lemonnier, D. Dagneaux, J.P. Benoit, P. Vachette, Annual activity report, LURE, Orsay, France, 1987.
[35] J.M. Dubuisson, T. Decamps, P. Vachette, J. Appl. Crystallogr. 30 (1997) 49.
[36] C. Tanford, Physical Chemistry of Macromolecules, Wiley, New York, 1961.
[37] F. Travers, P. Douzou, Biochimie 56 (1974) 509.
[38] S. Richard, Ph.D. Thesis, University Joseph Fourier, Grenoble, France, 1998.
[39] M. Riès-Kautt, A. Ducruix, in: A. Ducruix, R. Giegé (Eds.), Crystallization of Nucleic Acids and Proteins: A Practical Approach, IRL Press/Oxford University Press, Oxford, 1992, p. 195.
[40] A. Ducruix, R. Giegé, Crystallization of Nucleic Acids and Proteins: A Practical Approach, IRL Press/Oxford University Press, Oxford, 1992.
[41] Y. Gaudin, A. Barge, C. Ebel, R.W.H. Ruigrok, Virology 206 (1995) 28.
[42] A.-C. Dock-Bregeon, D. Moras, in: A. Ducruix, R. Giegé (Eds.), Crystallization of Nucleic Acids and Proteins: A Practical Approach, IRL Press/Oxford University Press, Oxford, 1992, p. 145.
[43] Y. Papanikolau, M. Kokkinidis, Protein Eng. 10 (1997) 847.
[44] T. Arakawa, S.N. Timasheff, Methods Enzymol. 114 (1985) 49.
[45] H. Eisenberg, in: W. Saenger (Ed.), Landolt-Brnstein: Numerical Data and Functional Relationships in Science and Technology, New Series Biophysics-Nucleic Acids, Springer, Berlin, 1990, p. 257.
[46] M. Malfois, F. Bonneté, L. Belloni, A. Tardieu, J. Chem. Phys. 105 (1996) 3290.
[47] A. Ducruix, J.P. Guilloteau, M. Riès-Kautt, A. Tardieu, J. Crystal Growth 168 (1996) 28.
[48] A. Tardieu, A. Le Verge, M. Malfois, F. Bonneté, S. Finet, M. Riès-Kautt, L. Belloni, J. Crystal Growth 196 (1999) 193.

6. MODIFICATION DES INTERACTIONS PAR L'AJOUT DE MPD

Le méthyl-2-pentanediol-2,4 (MPD) est un solvant organique utilisé comme précipitant dans la cristallisation des macromolécules biologiques, souvent en présence de sel. En présence de NaCl, il a été utilisé avec succès pour obtenir des cristaux de la *Hm* MalDH selon un processus original. Dans la partie §5 (p.125), nous avons expliqué ce processus (dilution d'une goutte biphasique) au vu des effets des mélanges de MPD et de NaCl sur la stabilité, la solubilité et les interactions protéine – protéine de la *Hm* MalDH [Costenaro *et al.*, 2001]. Nous avons constaté que l'ajout de MPD stabilise la protéine à bas sel et introduit une interaction attractive qui s'additionne aux interactions interparticulaires en présence de NaCl seul. Ces effets sont-ils induits uniquement par le MPD ou dépendent-ils aussi de la nature du sel ? Afin d'apporter des éléments de réponse à ces questions, nous avons étudié le comportement de la *Hm* MalDH en présence de MPD et d'autres sels, principalement $MgCl_2$ et $(NH_4)_2SO_4$. Les résultats obtenus pour ces sels seront comparés ceux obtenus en présence de NaCl.

Comme précédemment, nous nous sommes successivement intéressés aux diagrammes de phases sel – MPD – H_2O, à la stabilité et à la solubilité de la *Hm* MalDH dans les mélanges monophasiques, avant d'aborder les interactions protéine – protéine. La détermination des seconds coefficients du viriel par des expériences de vitesse de sédimentation selon la méthode présentée dans le chapitre 3 (p.69) a permis d'orienter nos choix de conditions de solvant pour des expériences de diffusion de neutrons ou de rayons X. Nous avons analysé ces dernières en termes de seconds coefficients du viriel et de modélisations de facteur de structure.

6.1. Diagrammes de phases des mélanges sel – MPD – H_2O

La Figure 6.1-a présente le diagramme de phases NaCl – MPD – H_2O selon une représentation bidimensionnelle en fonction des concentrations en NaCl et en MPD* [d'après

* Les concentrations en MPD sont exprimées comme le volume utilisé de MPD pur divisé par le volume final de la solution en pourcentage : % (v/v).

Richard *et al.*, 1995]. Une solution aqueuse de NaCl et de MPD est monophasique pour de faibles concentrations de sel et de MPD. Lorsque l'on augmente progressivement ces concentrations, la solution se sépare en deux phases liquides, puis le sel précipite, formant la troisième phase.

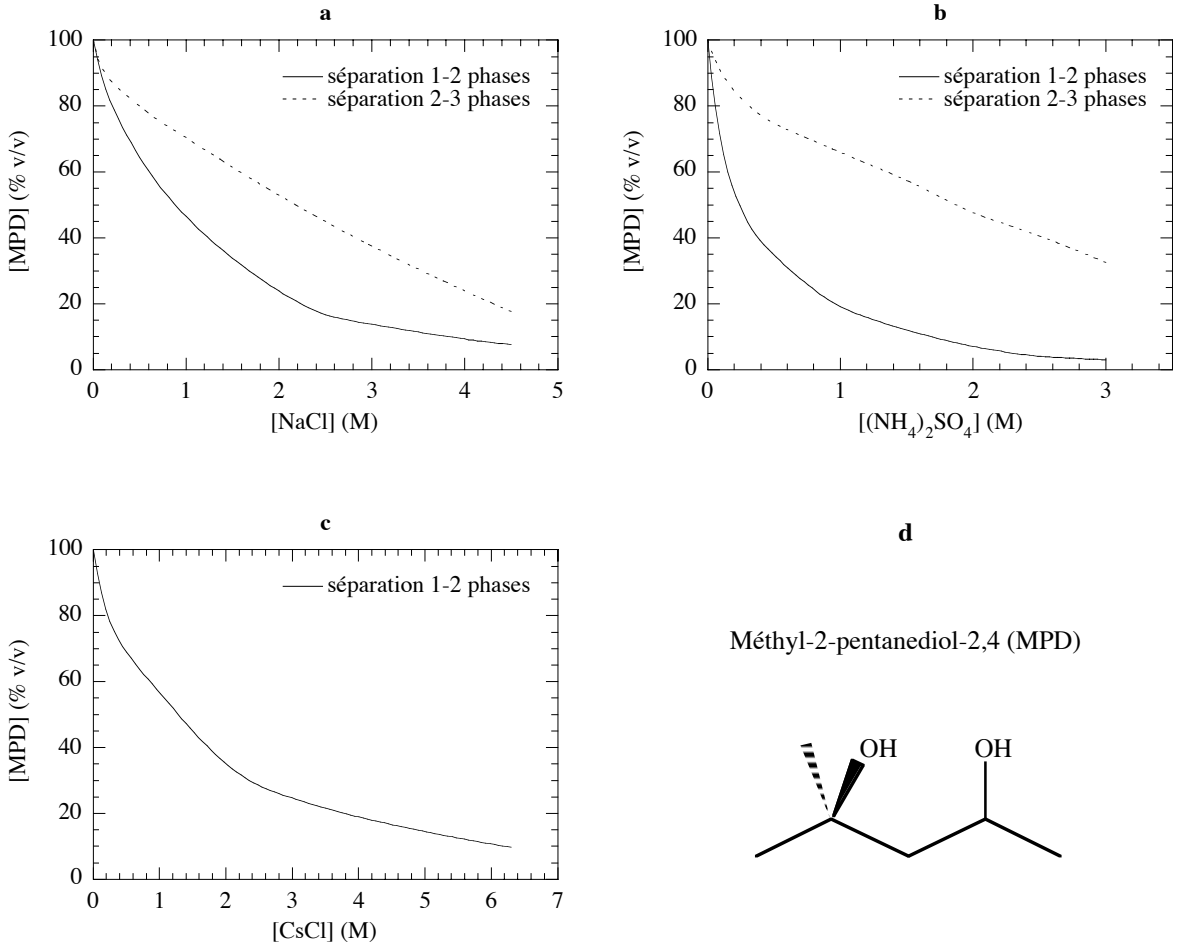


Figure 6.1 : Diagrammes de phases sel – MPD – H_2O : **(a)** NaCl à 22°C [d'après Richard *et al.*, 1995], **(b)** $(NH_4)_2SO_4$ à 4°C et **(c)** CsCl à 4°C (la séparation entre les zones biphasique et triphasique n'a pas été déterminée).

L'évaluation des interactions interparticulaires de la *Hm* MalDH doit être réalisée dans des solutions monophasiques. Nous avons donc partiellement construit les diagrammes de phases des mélanges sel – MPD – H_2O pour le sulfate d'ammonium et le chlorure de césium (Figure 6.1-b et c). La forme générale des courbes de coexistence est très similaire pour les trois sels présentés, mais pour une concentration donnée des trois sels, il faut ajouter des quantités de MPD différentes pour obtenir deux phases.

En plus des sels précédents, nous avons réalisé quelques mélanges $MgCl_2$ – MPD – H_2O. Dans notre domaine d'intérêt ($[MgCl_2] < 1.75$ M et [MPD] < 40%), ils ne montrent qu'une phase liquide et nous n'avons pas précisé la position des courbes de coexistence pour ce sel. Notons par ailleurs que le diagramme de phases KCl – MPD – H_2O est extrêmement similaire à celui obtenu en NaCl [Richard, 1998].

6.2. Stabilité et solubilité de la *Hm* MalDH dans les mélanges sel – MPD – H_2O

6.2.1. Stabilité

Comme l'inactivation et le dépliement de la *Hm* MalDH sont simultanés dans une grande variété de sels [Ebel *et al.*, 1999a], nous admettons que la protéine active correspond à une protéine repliée (et stable). Pour déterminer la stabilité de la *Hm* MalDH dans les mélanges sel – MPD – H_2O monophasiques, nous avons donc mesuré son activité résiduelle[*]. Les résultats, interprétés de manière semi-quantitative en termes de protéine active " a " ou inactive " i ", sont présentés dans la Figure 6.2, sur les diagrammes de phases sel – MPD – H_2O, pour le NaCl, le $MgCl_2$ et le $(NH_4)_2SO_4$.

Dans les trois sels, le MPD modifie la stabilité de la *Hm* MalDH : la gamme de concentrations salines où elle est active [Ebel *et al.*, 1999a] est élargie ou rétrécie pour des teneurs croissantes de MPD. La protéine est par exemple dénaturée en dessous de 2 M de NaCl à 25°C, mais elle est active et stable à plus faible concentration de sel lorsque du MPD est ajouté (Figure 6.2-a). Ainsi, en présence de 30% de MPD, elle reste active à 0.3 M de NaCl. A forte concentration de sel, l'activité et donc la stabilité de la protéine ne sont pas altérées. Des effets qualitativement identiques sont observés pour la protéine en présence de CsCl et de MPD. A l'opposé, si l'ajout de MPD préserve la stabilité de la protéine entre 0.2 et

[*] L'activité résiduelle est l'activité de l'enzyme dans des conditions standard après 24 heures d'incubation dans le solvant d'intérêt. Cette mesure est détaillée dans la partie " Materials and methods " de l'article [Costenaro *et al.*, 2001] inséré en p.130. Des mesures de fluorimétrie ont aussi été réalisées selon le protocole décrit dans l'article [Ebel *et al.*, 1999a].

0.8 M de $MgCl_2$, il la détériore aux fortes salinités (graphique b). Pour le $(NH_4)_2SO_4$ (graphique c), on observe la préservation de l'activité à haut sel en présence de faible concentration de MPD, limitée par la forme du diagramme de phases. Cependant, le MPD déstabilise la protéine aux concentrations intermédiaires de $(NH_4)_2SO_4$, entre 0.5 et 1 M. Une nouvelle zone d'activité à bas sel (environ 0.25 M) apparaît pour de fortes teneurs en MPD. Finalement, en absence de sel, le MPD seul (testé jusqu'à 50%) ne suffit pas à maintenir la protéine active.

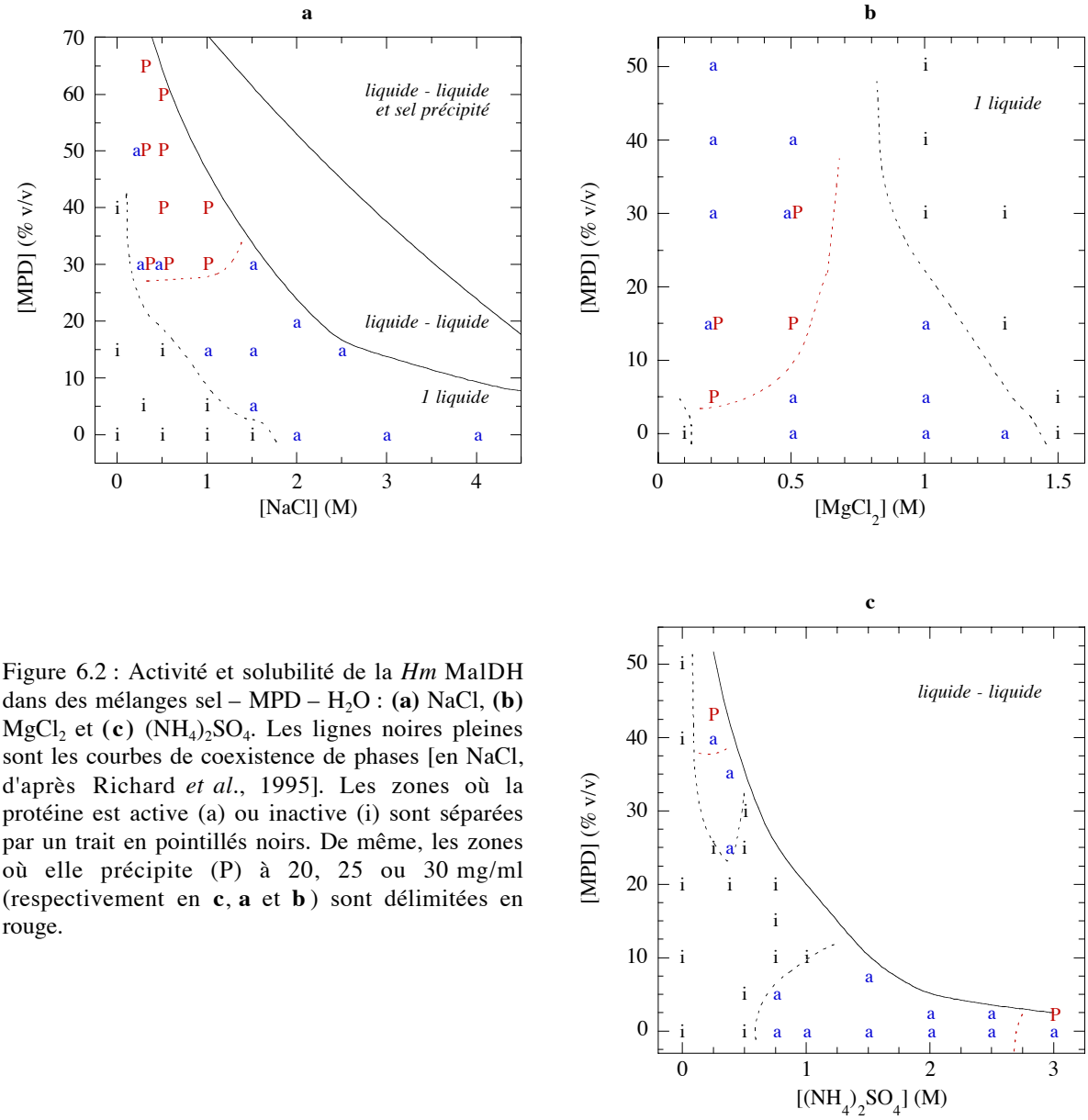


Figure 6.2 : Activité et solubilité de la *Hm* Ma1DH dans des mélanges sel – MPD – H_2O : **(a)** NaCl, **(b)** $MgCl_2$ et **(c)** $(NH_4)_2SO_4$. Les lignes noires pleines sont les courbes de coexistence de phases [en NaCl, d'après Richard *et al.*, 1995]. Les zones où la protéine est active (a) ou inactive (i) sont séparées par un trait en pointillés noirs. De même, les zones où elle précipite (P) à 20, 25 ou 30 mg/ml (respectivement en **c**, **a** et **b**) sont délimitées en rouge.

6.2.2. Solubilité

Avant toute mesure de diffusion de neutrons ou de rayons X, nous nous sommes assurés de la solubilité de la *Hm* MalDH active dans les mélanges sel – MPD – H_2O monophasiques en effectuant des tests ponctuels à environ 25 mg/ml de protéine, en utilisant le protocole décrit dans l'article [Costenaro *et al.*, 2001] inséré en p.130. Ils nous ont permis de déterminer dans quelles conditions de sel et de MPD la protéine précipite (indiquées par le symbole " P " dans la Figure 6.2). Elle demeure soluble (aux concentrations testées) dans le reste des zones d'activité.

Dans les trois sels, la *Hm* MalDH précipite à basse salinité et à forte concentration de MPD. Pour le sulfate d'ammonium, une autre zone de précipitation apparaît à 3 M de sel, qui s'avèrera être plus étendue lorsque les volumes des solutions protéiques seront plus importants que ceux utilisés dans cette approche (cf. §6.4.2 p.152). Ces précipités peuvent êtres re-solubilisés par simple dilution. Ils sont donc réversibles et correspondent à des solutions sursaturées.

6.3. Détermination de seconds coefficients du viriel par ultracentrifugation

Avant d'entreprendre des expériences de diffusion de neutrons en présence de MPD, nous avons abordé les interactions protéine – protéine par des expériences d'ultracentrifugation analytique en vitesse de sédimentation. Ces expériences ont été réalisées pour des mélanges $(NH_4)_2SO_4$ – MPD – H_2O où la protéine est stable et soluble à 20 mg/ml. La modélisation des profils de sédimentation nous a permis d'estimer le second coefficient du viriel A_2 à partir de la dépendance à la concentration en protéine des coefficients de sédimentation s et de diffusion D, selon la méthode présentée dans le premier article [Solovyova *et al.*, 2001] (cf. §3 p.69). Les valeurs de A_2 obtenues sont résumées dans le Tableau 6.1.

Ces résultats nous ont permis d'orienter nos choix de conditions de solvant pour les expériences de diffusion de neutrons à réaliser. Nous avons choisi d'étudier l'influence de la concentration du MPD à 0.25 et 1.5 M de sel et de celle du sel à 2.5% de MPD. Nous verrons

que les valeurs des A_2 ainsi déterminées correspondent à celles obtenues par diffusion de neutrons, mais aussi que la protéine cristallise dans la condition caractérisée par un A_2 faiblement négatif (cf. §6.5.2 p.158).

[MPD] (% v/v)	$[(NH_4)_2SO_4]$ (M)				
	0.25	1	1.5	2	2.5
40	-6				
30	5				
15		4			
2.5		5	6	4	0

Tableau 6.1 : Estimation du second coefficient du viriel A_2 (en 10^{-8} mol.l.g^{-2}) par vitesse de sédimentation.

6.4. Les données expérimentales de diffusion de neutrons

6.4.1. Les courbes de diffusion

L'évolution des courbes de diffusion normalisées (notamment par c) en fonction de la concentration en protéine c montre très clairement que l'ajout de MPD dans le solvant peut parfois modifier fortement les interactions protéine – protéine de la *Hm* MalDH. En NaCl 1.5 M, MPD 5%, elles sont globalement répulsives : l'intensité normalisée à la concentration en protéine c décroît lorsque c augmente (Figure 6.3-a). Rappelons qu'en NaCl 2-5 M, sans MPD, le même type de comportement répulsif a été observé. L'ajout de 30% de MPD modifie drastiquement les interactions qui deviennent extrêmement attractives en NaCl 1.5 M (graphique b). En $(NH_4)_2SO_4$ 0.25 M, on observe une évolution similaire : une augmentation de 10% de la teneur en MPD permet le passage d'interactions globalement répulsives (avec 30% de MPD, Figure 6.3-c) à globalement attractives (avec 40% de MPD, graphique d). Ces effets et leur intensité relative seront mieux révélés par les valeurs du second coefficient du viriel.

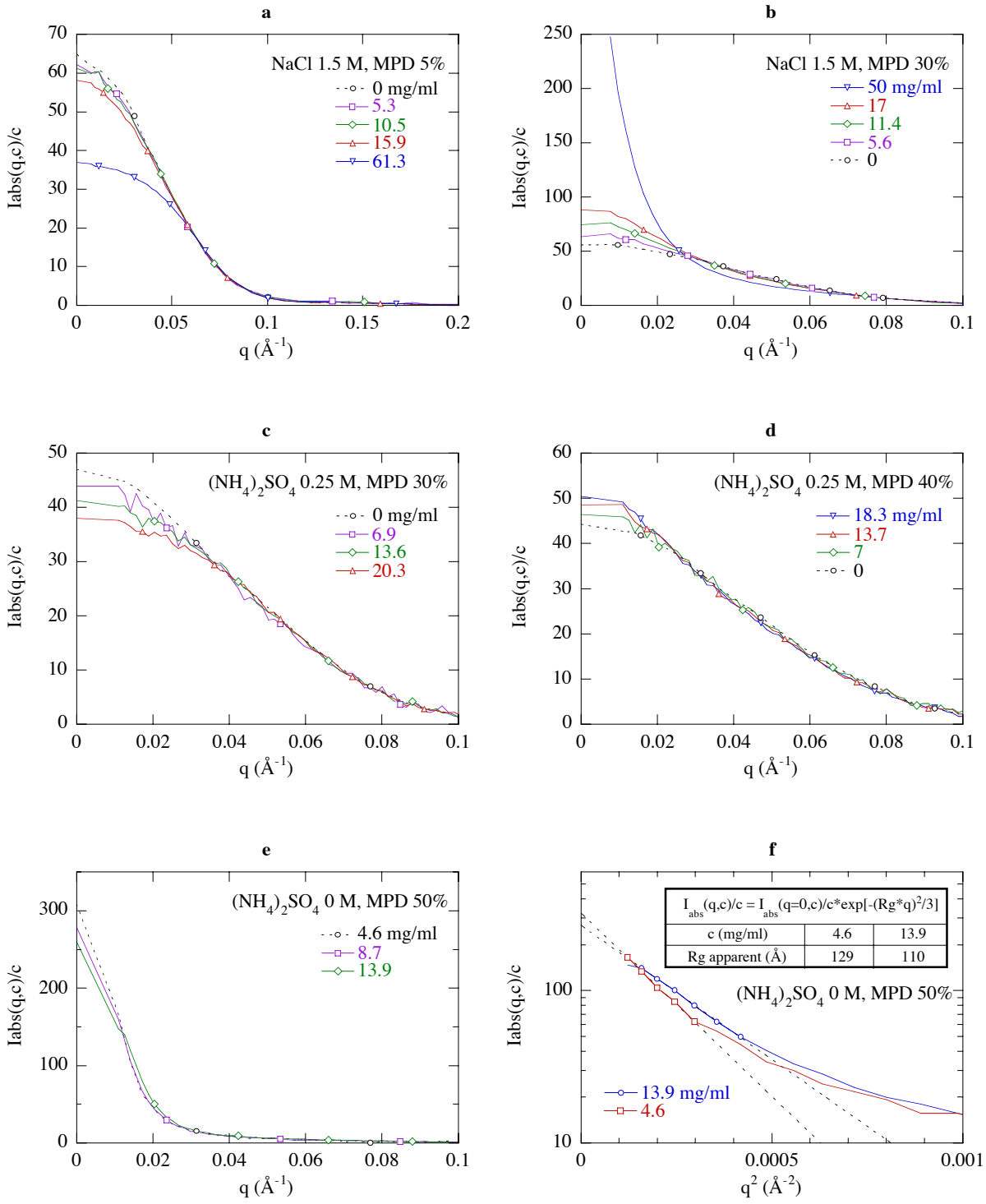


Figure 6.3 : Courbes de diffusion de neutrons par des solutions de *Hm* MalDH en présence de sel et de MPD, en fonction de la concentration en protéine c en mg/ml, à pH 8 et à 20-25°C : **(a-b)** NaCl 1.5 M, MPD 5 et 30% (réf. D11 99.6) ; **(c-d)** $(NH_4)_2SO_4$ 0.25 M, MPD 30 et 40% (réf. D22 00.10). Les courbes à concentration nulle correspondent aux facteurs de forme. Ils ont été déterminés par l'extrapolation linéaire des Iabs(q,c)/c, pour chaque valeur de q, sur les trois plus faibles concentrations en protéine, puis faiblement lissés. Les points à q = 0 sont extrapolés selon l'approximation de Guinier. **(e-f)** Intensité diffusée en $(NH_4)_2SO_4$ 0 M, MPD 50% et rayons de giration apparents correspondants déterminés par l'approximation de Guinier sur la partie des intensités représentées par des ronds ou des carrés.

Par ailleurs, les échantillons à 4.5-14 mg/ml de *Hm* MalDH en MPD 50% et sans sel ont des diffusions différentes des autres conditions. Elles sont très intenses aux très petits angles et correspondent à un rayon de giration extrapolé à concentration nulle $R_g°$ de l'ordre de 140 Å (Figure 6.3-e et f). Ces deux observations sont caractéristiques de particules plus grosses que le tétramère de la *Hm* MalDH. Aucune trace d'agrégation n'était cependant visible (aspect limpide et spectre d'absorbance plat au-delà de 340 nm), mais les échantillons étaient totalement inactifs. La protéine semble donc être dépliée et solubilisée par le MPD à cette forte concentration. Cet effet du MPD est connu et vient de son caractère hydrophobe qui lui permet d'interagir avec la chaîne principale des protéines [Arakawa *et al.*, 1990].

6.4.2. Stabilité des solutions de **Hm** *MalDH dans les mélanges $(NH_4)_2SO_4$ – MPD – H_2O*

En général, du fait de la sélection des solvants, les solutions de *Hm* MalDH dans les mélanges sel – MPD – H_2O étudiés par diffusion aux petits angles sont stables. Nous avons toutefois constaté de manière macroscopique deux types de mélanges qui font exception. Le premier correspond à des solutions sursaturées qui ont conduit à la cristallisation de la *Hm* MalDH. Il s'agit des mélanges NaCl – MPD – H_2O et $(NH_4)_2SO_4$ – MPD – H_2O à basse concentration de sel (cf. §5 p.125 et §6.5.2 p.158). Le second type est formé des mélanges $(NH_4)_2SO_4$ – MPD – H_2O à plus forte salinité (au-dessus de 1.5 M) dans lesquels nous avons observé des troubles détaillés ci-dessous.

Lors de la préparation des échantillons et lors des expériences de diffusion de neutrons, nous avons observé la formation de troubles blanchâtres macroscopiques dans certaines solutions de *Hm* MalDH en $(NH_4)_2SO_4$ 1.5 à 3 M, en présence de 0 à 7.5% de MPD. La Figure 6.4 présente les échantillons dans les cuves utilisées pour les expériences de diffusion. La première ligne de cette figure montre que les solutions protéiques en $(NH_4)_2SO_4$ 1.5 M sont limpides à toutes les concentrations en *Hm* MalDH étudiées. Pour des salinités identiques ou supérieures, lorsque les concentrations en MPD et en *Hm* MalDH augmentent, il apparaît de petits flocons, qui peuvent sédimenter avec le temps, puis un gel blanchâtre qui occupe la totalité de la solution (deuxième à quatrième lignes). L'ampleur des troubles augmentant avec les concentrations du sel, du MPD et de la protéine, il s'agit d'un phénomène synergique entre les trois composants. Notons que dans ces solvants, la protéine est toutefois

active et ne semblait pas précipiter à 20 mg/ml (cf. Figure 6.2-c p.148). S'agit-il alors d'une contamination ou sommes nous en présence de formes solubles métastables (non observées précédemment du fait des très petits volumes utilisés) ?

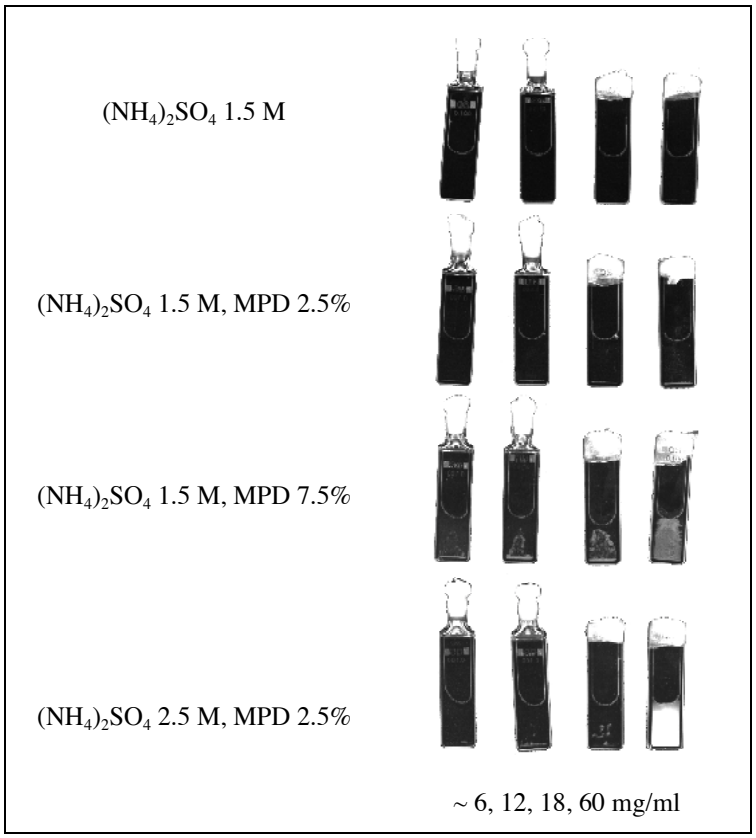


Figure 6.4 : Séries de cuves d'échantillon de *Hm* MalDH en sulfate d'ammonium et MPD pour la diffusion de neutrons (réf : D22 00.10). Les cuves (1.2 cm de largeur) ont été scannées sur un fond noir afin d'obtenir un meilleur contraste. Lorsque la concentration du sel ou du MPD augmente, un gel blanchâtre apparaît proportionnellement à la concentration en protéine.

Les échantillons en question ont été préparés deux fois avec des solvants de provenance différente, les mêmes troubles se sont produits et les courbes de diffusions étaient reproductibles, ce qui exclut donc une contamination accidentelle.

Par ailleurs, les mesures d'activité résiduelle (activité après 24 heures d'incubation) et de fluorimétrie préalablement réalisées ont montré que dans ces conditions la *Hm* MalDH était active et stable. Le $(NH_4)_2SO_4$ seul (au-dessus de 1.5 M) est d'ailleurs un sel très stabilisant de la forme repliée de la protéine [Ebel *et al.*, 1999a]. Bien que ces tests soient réalisés à de faibles concentrations en protéine, ils devraient réfuter l'hypothèse d'une dénaturation de la protéine, puisqu'une concentration protéique plus élevée est généralement une source de " stabilité " de la forme repliée. Sans toutefois définitivement écarter une

dénaturation partielle de la protéine ou un changement de conformation, les troubles et les gels blanchâtres observés semblent raisonnablement liés à un problème de solubilité de la forme native.

L'existence de plusieurs phases dans des mélanges sel – MPD – H_2O qui traduit l'instabilité de ces solutions est décrite comme étant due à la répulsion des charges électrostatiques par le MPD [Pittz et Timasheff, 1978]. Nous avons donc préparé des solutions de $(NH_4)_2SO_4$ 2 M, MPD 2.5-5% contenant de l'acétate d'ammonium ou de l'acide aspartique. Les concentrations de ces espèces additives ont été choisies pour mimer l'augmentation de la concentration en charge induite par la protéine. Nous sommes allés jusqu'à des concentrations équivalentes à 84 mg/ml de *Hm* MalDH, soit 0.1 M d'acétate d'ammonium. Aucun trouble ou précipité n'a été observé. Leur présence dans nos échantillons n'est donc pas attribuable à la seule présence de charges supplémentaires dans les solutions.

L'aspect visuel de nos échantillons est similaire à des agrégats ou à des gels (Figure 6.4). Sous l'action d'un potentiel attractif, des colloïdes (les protéines) peuvent adhérer les uns aux autres en formant, selon leur concentration, des agrégats voir des structures plus larges ou gels qui peuvent être partiellement ordonnés [Poon *et al.*, 1996]. En accord avec cette hypothèse, le profil de diffusion de neutrons de l'échantillon de *Hm* MalDH à 60 mg/ml en $(NH_4)_2SO_4$ 2.5 M, MPD 2.5% présente un pic de corrélation très net et reproductible dont la position ($q = 0.076$ Å^{-1}) correspond à une distance réelle de 83 ±1 Å (Figure 6.6-a). Cette distance est similaire à la distance des tétramères de *Hm* MalDH dans le cristal. Elle est donc compatible avec des protéines repliées juxtaposées. En échelle semi-logarithmique, les courbes Iabs(q,c)/c sont décalées les unes des autres en fonction du temps, ce qui traduit une baisse de la concentration des protéines en solution alors que l'intensité du pic augmente. Ceci est raisonnable si l'on considère qu'elles sont progressivement intégrées dans ce gel. Cet échantillon n'étant pas en équilibre thermodynamique, nous ne l'avons pas utilisé dans les analyses suivantes. Par contre, les courbes à plus faible concentration en protéine ne présentent pas de tel pic, ni de décalage en échelle semi-logarithmique : elles nous ont donc permis de déterminer un second coefficient du viriel. Un pic plus large et correspondant à une distance d'environ 75 Å a aussi été enregistré pour un échantillon à 22 mg/ml de protéine, plus concentré en sel : $(NH_4)_2SO_4$ 3 M, MPD 2.5% (Figure 6.6-b). Ce solvant correspondrait alors à une interaction attractive plus intense.

Les troubles blanchâtres (agrégats et gels) observés dans ces solutions de *Hm* MalDH en $(NH_4)_2SO_4$ – MPD – H_2O relèveraient donc raisonnablement d'un problème de solubilité de la protéine et seraient dus à des interactions protéine – protéine attractives.

Du point de vue de l'analyse des courbes de diffusion, la présence de ces agrégats en petites proportions est relativement peu gênante : leur dimension étant macroscopique, ils ne diffusent pas dans la même gamme de q que les protéines solubles. Pour les échantillons les plus troubles, leur présence ne se traduit que par quelques points aux très petits angles omis pour les extrapolations à q = 0 (en grisé dans la Figure 6.6-a). Ces points ont cependant compromis la précision de certaines valeurs du second coefficient du viriel A_2 et ont empêché leur détermination en $(NH_4)_2SO_4$ 3 M, MPD 2.5% (Figure 6.6-b) et en $(NH_4)_2SO_4$ 1.5 M, MPD 7.5% où les agrégats sont relativement importants aux concentrations protéiques utilisées pour le calcul de A_2 ($5 < c < 30$ mg/ml).

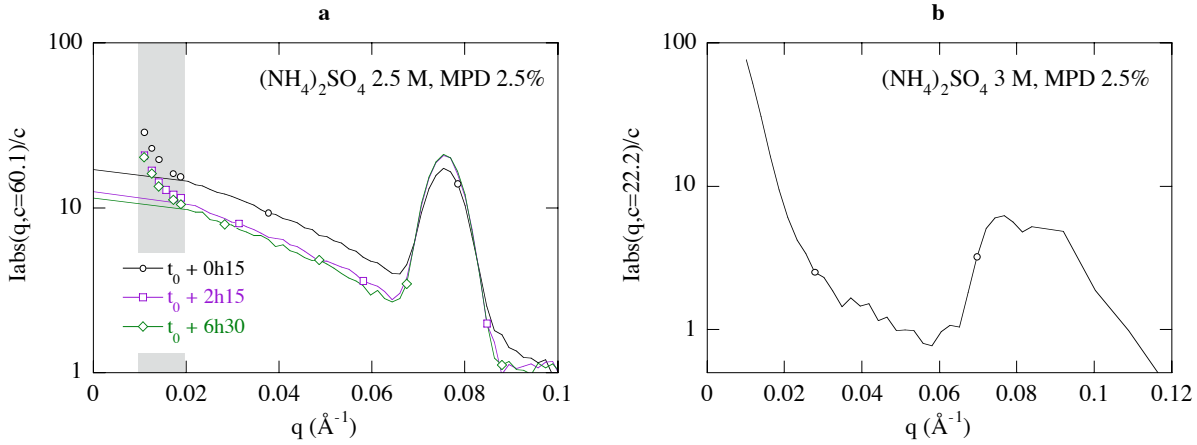


Figure 6.6 : Courbe de diffusion de neutrons aux petits angles par des solutions de *Hm* MalDH en sulfate d'ammonium et MPD. **(a)** Evolution en fonction du temps de l'intensité diffusée (échelle semi-logarithmique) par une solution de *Hm* MalDH à 60.1 mg/ml, en $(NH_4)_2SO_4$ 2.5 M, MPD 2.5% (réf. D22 00.10). L'échantillon a été préparé à l'instant t_0. Les données en grisé n'ont pas été utilisées pour les extrapolations à q = 0. **(b)** Intensité diffusée par une solution de *Hm* MalDH à 22.2 mg/ml, en $(NH_4)_2SO_4$ 3 M, MPD 2.5% (réf. D11 00.8).

6.5. Les seconds coefficients du viriel A_2

6.5.1. Evolution des A_2 en présence de MPD

Les valeurs du second coefficient du viriel A_2 de la *Hm* MalDH dans des mélanges sel – MPD – eau, déterminés par diffusion de neutrons (cf. p.58), sont présentées dans le Tableau 6.3 pour les trois sels étudiés : NaCl, $MgCl_2$ et $(NH_4)_2SO_4$, en fonction des concentrations des deux co-solvants.

Les valeurs déterminées par Ebel *et al.* [Ebel *et al.*, 1999b] sans MPD et à 5°C (indiquées par un *) sont raisonnablement comparables aux valeurs obtenues (ou que l'on obtiendrait) à 25°C. Nous avons en effet vu (cf. §4.2.1 p.96) qu'une telle variation de température ne modifiait pas significativement les valeurs positives de A_2, compte tenu de leur erreur d'environ $\pm 0.2\ 10^{-8}$ mol.l.g^{-2}, et que légèrement les valeurs proche de zéro. Dans le Tableau 6.3-c, les A_2 en $(NH_4)_2SO_4$ 1.5 et 3 M, MPD 0% sont des moyennes des valeurs à 5 et 25°C : respectivement 3.9, 3.5 et 0.6, $-0.8\ 10^{-8}$ mol.l.g^{-2}.

Les valeurs du second coefficient du viriel A_2 indiquent de manière très évidente que le MPD modifie globalement les interactions protéine – protéine. A salinité constante, une augmentation de la concentration du MPD induit une baisse notable de A_2 qui traduit une attraction interparticulaire plus intense. Pour certaines salinités, A_2 atteint des valeurs négatives. Le MPD peut finalement provoquer la précipitation de la *Hm* MalDH à des concentrations en protéine de 20 à 30 mg/ml (symbole " P " dans le Tableau 6.3 et la Figure 6.2 p.148) : ces conditions de solvant sont considérées comme des conditions fortement attractives. Ainsi, en NaCl 1.5 M, les valeurs de A_2 baissent de 2.5 à $-9.4\ 10^{-8}$ mol.l.g^{-2} lorsque la teneur en MPD passe de 5 à 30%. En $MgCl_2$, on observe la même tendance : l'ajout de 15% de MPD fait diminuer A_2 de 2.7 à $-0.6\ 10^{-8}$ mol.l.g^{-2} à 1 M de sel, et provoque la précipitation de la *Hm* MalDH à 30 mg/ml en 0.5 M de sel. En $(NH_4)_2SO_4$ 0.25 M, A_2 chute radicalement lorsque la concentration en MPD varie de 30 à 40% et au-delà, la protéine précipite à 20 mg/ml.

a

[MPD] (% v/v)	[NaCl] (M)					
	1	1.5	2	2.5	3	4
30	P	-9.4 (0.1)				
25		-2.2 (0.2)				
20			1.9 (0.1)			
15	2.3 (0.3)	2.0 (0.7)	2.3 (0.4)	2.7 (0.9)		
5		2.5 (0.8)				
0			4.5*		4.1*	3.2*

b

[MPD] (% v/v)	[$MgCl_2$] (M)			
	0.2	0.5	1	1.3
15	P	P	-0.6 (0.4)	
5	P	-0.8 (0.4)	1.4 (0.2)	
0	-0.9*	0.8*	2.7*	2.1*

c

[MPD] (% v/v)	[$(NH_4)_2SO_4$] (M)					
	0.25	1	1.5	2	2.5	3
40	-2.9 (0.4)					
30	4.8 (0.5)					
5			1.3			
2.5			3.5 (0.5)	3.9 (0.5)	-0.3 (0.9)	P
0		3.8*	3.7 (0.4)	3.7*	2.6 (0.5)	-0.1 (0.7)

Tableau 6.3 : Seconds coefficients du viriel (A_2 en 10^{-8} mol.l.g^{-2}) de la *Hm* MalDH dans des mélanges sel – MPD – eau à pH 8-8.2 et à 20-25°C : **(a)** NaCl (réf. D11 99.6), **(b)** $MgCl_2$ (réf. D11 99.6) et **(c)** $(NH_4)_2SO_4$ (réf. D22 00.10). Les valeurs marquées d'un astérisque ont été déterminées à 5°C (réf. D11 97.4, d'après [Ebel *et al.*, 1999b]). Les erreurs sont indiquées entre parenthèses. Le symbole " P " indique que la protéine précipite à 25, 30 ou 20 mg/ml, respectivement en **a**, **b** et **c** (ils ne sont pas tous indiqués, cf. Figure 6.2 p.148).

A l'inverse, pour une concentration constante et modérée de MPD, l'évolution de A_2 avec la concentration en sel semble le plus souvent reproduire la tendance observée sans MPD, mais de manière décalée. Les variations de A_2 avec la concentration en NaCl sont similaires à 0 et 15% de MPD, avec peu d'évolution. En $MgCl_2$, l'augmentation de A_2 avec la salinité que l'on observe en absence de MPD se retrouve pour 5% de MPD. En $(NH_4)_2SO_4$ avec 2.5% de MPD, A_2 est relativement constant jusqu'à 2 M puis décroît rapidement vers des valeurs négatives. Cette brusque baisse est observée, en absence de MPD, pour une

concentration en sel supérieure à 2.5 M. La présence du MPD semble donc décaler cette transition vers une concentration saline plus faible.

6.5.2. A_2 et cristallisation

Dans le système NaCl – MPD – H_2O, la cristallisation de la *Hm* MalDH se produit à faible concentration de sel et à forte concentration de MPD. Elle est corrélée à une lente évolution de A_2 de valeurs positives vers des valeurs négatives de l'ordre de -10 10^{-8} mol.l.g^{-2}, comme nous l'avons décrit dans l'article [Costenaro *et al.*, 2001] (cf. §5.1 p.127).

De manière très intéressante, lors de l'acquisition des données de diffusion, l'échantillon à 20.5 mg/ml en $(NH_4)_2SO_4$ 0.25 M, MPD 40%, donc correspondant aussi à une condition de faible salinité et fortement concentrée en MPD, présentait des points brillants (absents pour les concentrations protéiques inférieures). Ils ont grossi en des cristaux (aiguilles et parallélépipèdes) d'environ 1 mm dans leur plus grande dimension en deux semaines (Figure 6.7). Le solvant $(NH_4)_2SO_4$ 0.25 M, MPD 40% est donc une nouvelle condition de cristallisation de la *Hm* MalDH.

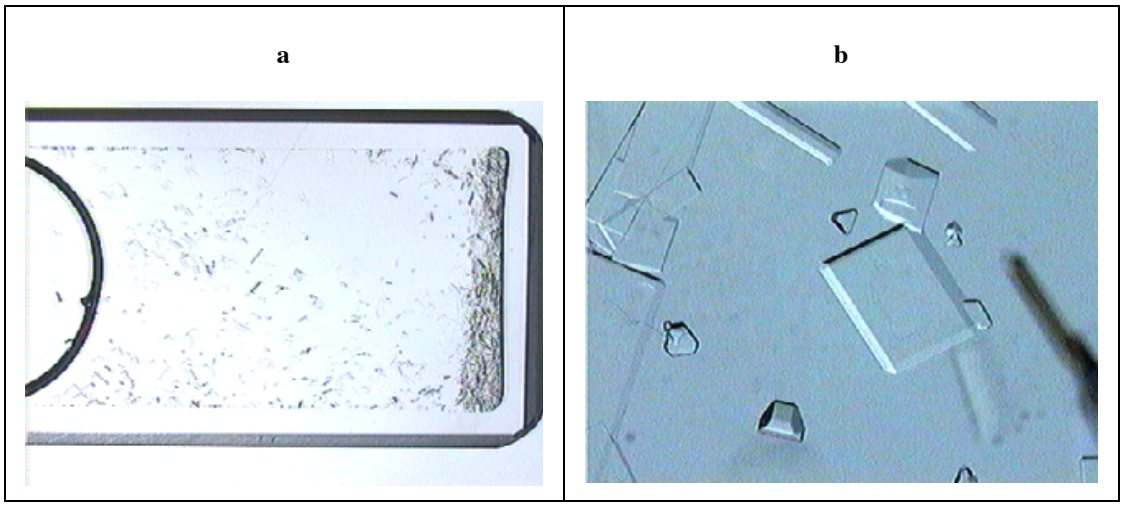


Figure 6.7 : **(a)** Cristaux de *Hm* MalDH obtenus dans la cuve utilisée pour la diffusion de neutrons, deux semaines après l'expérience, en $(NH_4)_2SO_4$ 0.25 M, MPD 40% à une concentration en protéine de 20.5 mg/ml. **(b)** Grossissement : la dimension maximale des cristaux est d'environ 1 mm.

La courbe de diffusion I/c enregistrée environ 5 heures après la préparation de l'échantillon et la mesure de sa concentration en protéine (c = 20.5 mg/ml) présente en échelle semi-logarithmique un décalage constant avec les courbes aux autres concentrations. La

concentration de protéine en solution a donc évolué : elle est estimée à 18 mg/ml soit 88% de la concentration initiale. Nous n'avons pas observé de pic de Bragg lié à la présence des microcristaux. Pour cette condition de solvant, nous avons pu déterminer un A_2 de -3 10-8 mol.l.g^{-2} avec les échantillons plus dilués et en corrigeant la concentration en protéine (à 18 mg/ml) pour l'échantillon générant des cristaux (cf. Tableau 6.3-c p.157).

6.6. Modélisation des facteurs de structure

Les seconds coefficients du viriel indiquent très clairement que l'ajout de MPD introduit une attraction supplémentaire. Les interactions protéine – protéine devenant globalement attractives, le facteur de structure à l'origine sera supérieur à 1. Nous avons donc, à partir des paramètres des potentiels utilisés pour modéliser les facteurs de structure expérimentaux sans MPD, ajouté un potentiel attractif de type Yukawa ou renforcé celui déjà présent. Comme expliqué dans la partie méthodologique (cf. p.65) et observé pour les solvants sans MPD (cf. §4.3.4 p.109), plusieurs couples de potentiel attractif – défini par sa profondeur J (en kT) et sa portée d – et de potentiels répulsifs – de volume exclu et électrostatique définis par le diamètre D et la charge Z de la particule en solution – permettent de rendre compte du même facteur de structure expérimental. Pour modéliser les facteurs de structure S dans les solvants sel – MPD – H_2O, nous avons donc conservé le diamètre déterminé en absence de MPD et fixé la valeur de la charge (pour nos solutions, elle ne peut être objectivement déterminée) et la portée du potentiel attractif à 5 Å. La profondeur J est le paramètre ajusté.

Par ailleurs, rappelons que les facteurs de structure expérimentaux ont été obtenus en divisant l'intensité diffusée expérimentale (normalisée par la concentration protéique c) I(q,c)/c par le facteur de forme I(q,c=0)/c déterminé par l'extrapolation à concentration nulle des I(q,c)/c ou, à défaut, l'intensité diffusée à la plus basse concentration en protéine c (cf. Figure 2.5 p.67). Les intensités sont issues de la diffusion de neutrons ou de rayons X.

6.6.1. Mélanges NaCl – MPD – H_2O

En absence de MPD, la modélisation des facteurs de structure expérimentaux de la *Hm* MalDH en NaCl 2-5 M nous a permis d'estimer le diamètre de la protéine en solution à

80 Å, mais pas de préciser sa charge Z comprise entre 0 et 156. En présence de MPD, nous avons donc conservé ce diamètre et fixé Z à 70, puis progressivement accentué la profondeur J du potentiel attractif afin rendre le mieux possible compte des facteurs de structure expérimentaux. Pour les différents mélanges de NaCl et de MPD où les seconds coefficients du viriel ont été déterminés, nous avions enregistré la diffusion d'échantillons plus concentrés en protéine (environ 60 mg/ml pour la plupart des conditions de solvant) ce qui nous a permis de modéliser les facteurs de structure correspondants (Figure 6.8-a-d).

Les graphiques (a) et (b) présentent des exemples de ces modélisations et permettent d'apprécier, outre la qualité de l'accord entre les S calculés et expérimentaux, la précision des valeurs déterminées pour J qui est généralement de ±0.1 kT. Sur la Figure 6.8-c, nous avons reporté les facteurs de structures expérimentaux et leur modélisation en NaCl 1.5 M, pour des concentrations en protéine du même ordre (50-68 mg/ml) et pour des concentrations croissantes de MPD. On constate très clairement que l'augmentation de la concentration en MPD peut être traduite par une plus grande profondeur J du potentiel attractif. Nous déterminons une valeur de J nulle en absence de MPD qui s'accroît à 4.38 kT pour 25% de MPD. A 50 mg/ml de *Hm* MalDH en NaCl 1.5 M, MPD 30%, qui est la condition la plus attractive enregistrée, le S calculé avec une profondeur du potentiel attractif de 4.55 kT n'est pas suffisamment élevé aux petits angles (Figure 6.8-c). Le S_{exp} correspond donc à un comportement encore plus attractif. Malgré toutes les précautions prises (paramètre de mélange très faible, incrémentation très lente des valeurs de J successives), le calcul diverge pour un J plus élevé, même de 0.01 kT. Une modélisation est cependant possible à une concentration en protéine plus faible (17 mg/ml) et converge vers un J de 4.95 ±0.05 kT comme le montre la Figure 6.8-b.

En NaCl 2 M, on modélise très correctement les facteurs de structure obtenus en présence de concentrations variables de MPD (Figure 6.8-d). On retrouve pour les interactions interparticulaires de la *Hm* MalDH la même tendance qu'en NaCl 1.5 M, avec un J qui augmente de 0 à 2.8 kT entre 0 et 20% de MPD.

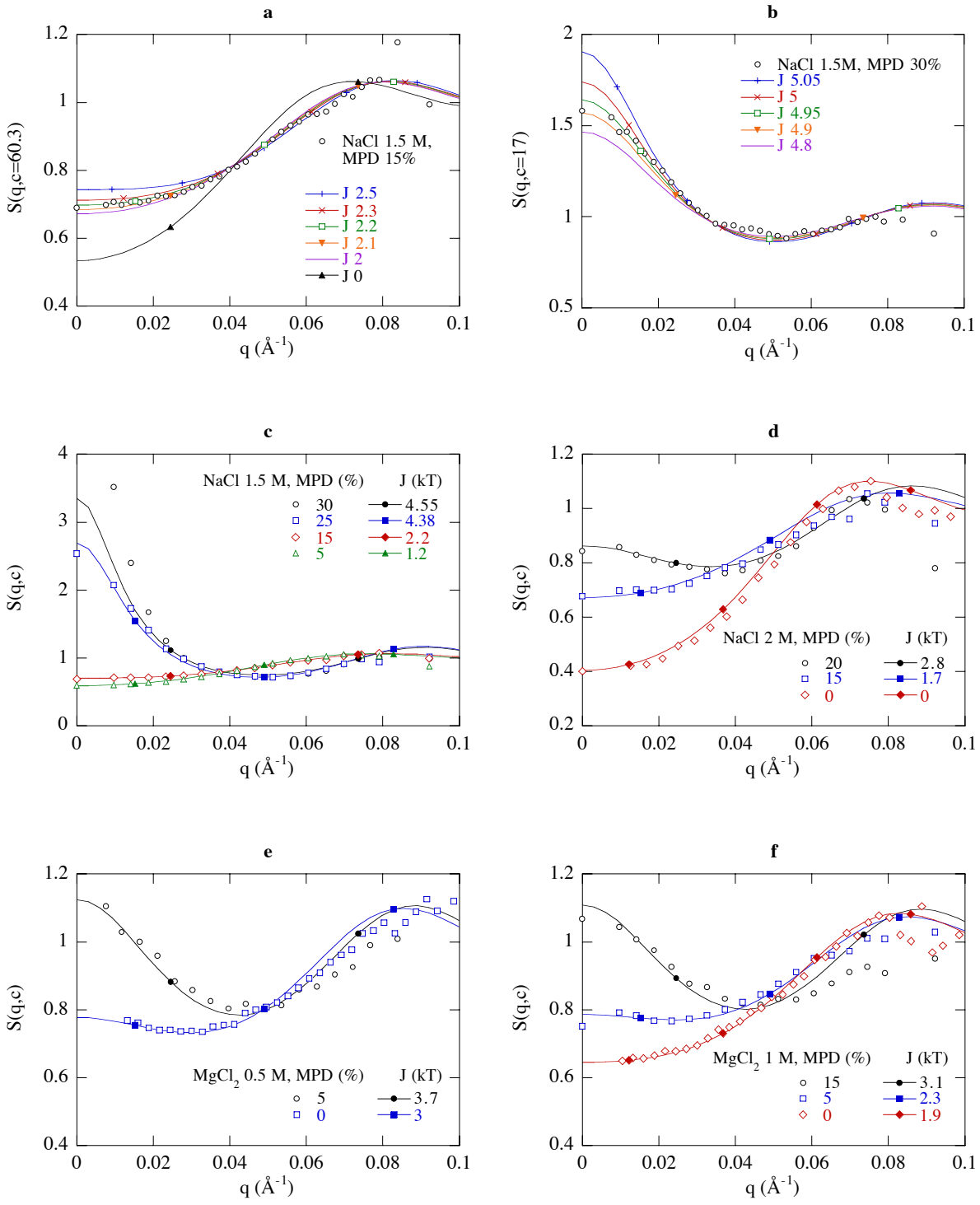


Figure 6.8 : Modélisation des facteurs de structure S(q,c) de la *Hm* MalDH en NaCl et $MgCl_2$ en présence de MPD, à pH 8-8.2 et à 20-25°C (réf. D11 99.6, sauf NaCl 2 M : D22 97.8, MgCl2 0.5-1 M : Lure 98.12). Les symboles creux correspondent aux S expérimentaux et les lignes aux S calculés à partir d'un potentiel de volume exclu caractérisé par le diamètre D de 80 Å, d'un potentiel coulombien avec une charge protéique Z de 70 et d'un potentiel attractif de profondeur J en kT et de 5 Å de portée. Les concentrations en protéine c en mg/ml sont (de haut en bas) : **(c)** 50, 67.9, 60.3 et 61.3 ; **(d)** 65.6, 59.1 et 95.8 ; **(e)** 66 et 79.6 **(f)** 61.2, 64.5 et 79.2.

Par ailleurs, nous avons effectué des mesures en fonction de la température (de 5 à 40°C) pour deux concentrations de *Hm* MalDH (environ 62 et 4 mg/ml) en NaCl 1.5 et 2.5 M en présence de 15% de MPD (réf. D22 00.4). La modélisation des facteurs de structure obtenus donne des valeurs de J qui diminuent lorsque la température augmente de 5 à 40°C : de 3.3 à 2.6 kT pour 1.5 M de NaCl, et de 1.8 à 1.6 kT pour 2.5 M. Pour les différentes températures, les facteurs de structure expérimentaux sont convenablement reproduits par les facteurs de structure calculés. Cela signifie que l'ordre de grandeur de la portée d du potentiel attractif que nous avons fixé à 5 Å semble être correct. Néanmoins, vu les faibles écarts que nous obtenons en fonction de la température et la précision de nos facteurs de structure expérimentaux, nous ne pouvons pas préciser la fiabilité de cette portée en dessous de quelques angströms.

6.6.2. Mélanges $MgCl_2$ – MPD – H_2O

Dans les mélanges $MgCl_2$ – MPD – H_2O pour lesquels des A_2 ont été déterminés, nous avons modélisé les facteurs de structure de la *Hm* MalDH obtenus pour des concentrations en protéine d'environ 60 mg/ml. Nous avons comme en absence de MPD fixé le diamètre D à 80 Å et la charge Z à 70. Les données expérimentales sont modélisées de manière satisfaisante avec un potentiel attractif de portée constante de 5 Å et de profondeur J ajustée. En absence de MPD, ce potentiel n'est pas nul, et l'ajout de MPD correspond encore à une augmentation de sa profondeur J. Elle croît de 3 à 3.7 kT entre 0 et 5% de MPD en $MgCl_2$ 0.5 M et de 1.9 à 3.1 kT entre 0 et 15% de MPD en $MgCl_2$ 1 M (Figure 6.8-e et f). Une tendance similaire à celle que nous avons observée pour le NaCl est donc mise en évidence pour le $MgCl_2$.

6.6.3. Mélanges $(NH_4)_2SO_4$ – MPD – H_2O

A l'inverse des deux sels précédents, nous n'avons pas pu modéliser les facteurs de structure expérimentaux S_{exp} de la *Hm* MalDH dans tous les mélanges $(NH_4)_2SO_4$ – MPD – H_2O pour lesquels des A_2 ont été déterminés. Plusieurs contraintes sont à l'origine de cette limitation. La présence des troubles et des gels blanchâtres que nous avons observés en présence de MPD, principalement aux plus fortes concentrations en protéine étudiées (50-80 mg/ml), a fortement augmenté l'imprécision des S_{exp} près de l'origine qui est la zone la plus sensible aux différentes interactions, empêchant leur modélisation. De plus, pour les

conditions de solvants suivantes : $(NH_4)_2SO_4$ 2.5 M, MPD 2.5% et $(NH_4)_2SO_4$ 3 M, MPD 2.5%, nous avons observé des pics de corrélation dans les courbes de diffusion aux plus fortes concentrations en protéines, respectivement 60.1 et 22.2 mg/ml (cf. Figure 6.6 p.155). Ces conditions sortent du cadre de notre modélisation puisque les potentiels que nous utilisons ne peuvent en aucun cas rendre compte d'agrégats qui impliquent des contacts entre protéines. En définitive, en présence de 2.5% de MPD pour des salinités supérieures à 2 M et pour la condition $(NH_4)_2SO_4$ 1.5 M, MPD 5%, nous sommes restreint aux facteurs de structure à concentration en protéine intermédiaire (environ 20 mg/ml). Ils restent toutefois trop proches de 1 et ne peuvent être correctement ni déduits des intensités diffusées, ni modélisés, malgré l'intérêt de ces conditions.

En dépit des difficultés rencontrées à forte salinité, nous avons pu convenablement modéliser le facteur de structure de la *Hm* MalDH en $(NH_4)_2SO_4$ 1.5 M, MPD 2.5% (Figure 6.9-a). Nous avons, comme en absence de MPD, déterminé un diamètre de la particule en solution de 80 Å, sans pouvoir préciser sa charge Z. Pour la même teneur en MPD et en 2 M de sel, le facteur de structure est, malgré la faible concentration en protéine, compatible avec un diamètre de 80 Å, une charge de 30 fixés et un potentiel attractif de 5 Å de portée et d'une profondeur J d'environ 0.5 kT (Figure 6.9-b). En $(NH_4)_2SO_4$ 2.5 M, MPD 2.5%, on trouve qualitativement une profondeur J de l'ordre de 1 à 2 kT. Ces valeurs à 2.5% de MPD sont similaires à celles que nous avons déterminées en absence de MPD.

En $(NH_4)_2SO_4$ 0.25 M, MPD 40%, le facteur de structure est encore qualitativement compatible avec celui d'une particule de 80 Å de diamètre (Figure 6.9-c). Il est cependant erroné pour des q compris entre 0.3 et 0.5 $Å^{-1}$ du fait d'une mauvaise extrapolation du facteur de forme causé par la faible précision des intensités diffusées. La condition qui a donné naissance à des cristaux de *Hm* MalDH, *i.e.* $(NH_4)_2SO_4$ 0.25 M, MPD 40%, est quant à elle convenablement modélisée (Figure 6.9-d). Pour un diamètre D de 80 Å, une charge Z de 30 et une portée du potentiel attractif de 5 Å, nous trouvons qu'une profondeur J de 3.7 ±0.1 kT est nécessaire pour rendre compte du facteur de structure expérimental. De nouveau, une condition favorable à la cristallisation coïncide avec l'existence d'un potentiel faiblement attractif.

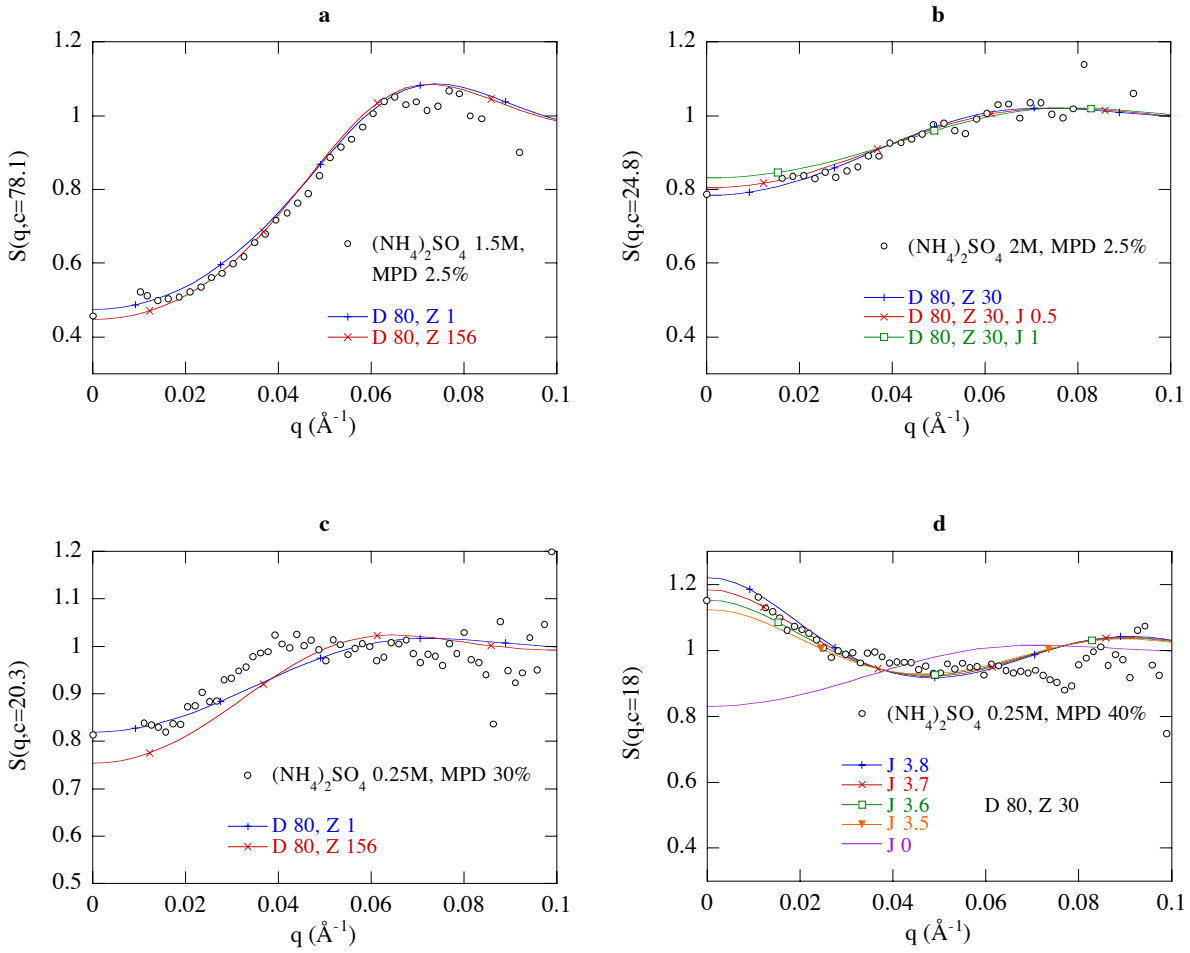


Figure 6.9 : Modélisation des facteurs de structure S(q,c) de la *Hm* MalDH en $(NH_4)_2SO_4$ en présence de MPD, à pH 8 et à 25°C (réf. D11 00.8 et D22 00.10). Les symboles creux correspondent aux S expérimentaux et les lignes aux S calculés à partir d'un potentiel de volume exclu caractérisé par le diamètre D de 80 Å, d'un potentiel coulombien avec une charge protéique Z et d'un potentiel attractif de profondeur J en kT et de 5 Å de portée. Les concentrations en protéine c sont en mg/ml.

6.6.4. MPD et potentiel attractif

Les facteurs de structure expérimentaux de la *Hm* MalDH en divers mélanges sel – MPD – H_2O peuvent être (dans la plupart des cas) correctement modélisés grâce aux trois potentiels génériques de volume exclu, électrostatique et attractif de type Yukawa. En particulier, ce dernier traduit très bien l'attraction supplémentaire que nous avons observée en présence de MPD. La Figure 6.10 résume les différentes valeurs de la profondeur J du potentiel attractif que nous avons déterminée par la modélisation des facteurs de structure de la *Hm* MalDH dans les mélanges sel – MPD – H_2O. Les séries à salinité constante sont

limitées soit par la stabilité ou la solubilité de la protéine, soit par la démixtion des solvants. Dans tous les cas, J augmente avec la concentration du solvant organique (graphique a). Le MPD ajoute ou renforce donc la composante attractive des interactions protéine – protéine. Cette augmentation est pour chaque concentration de sel proportionnelle à la teneur du solvant en MPD.

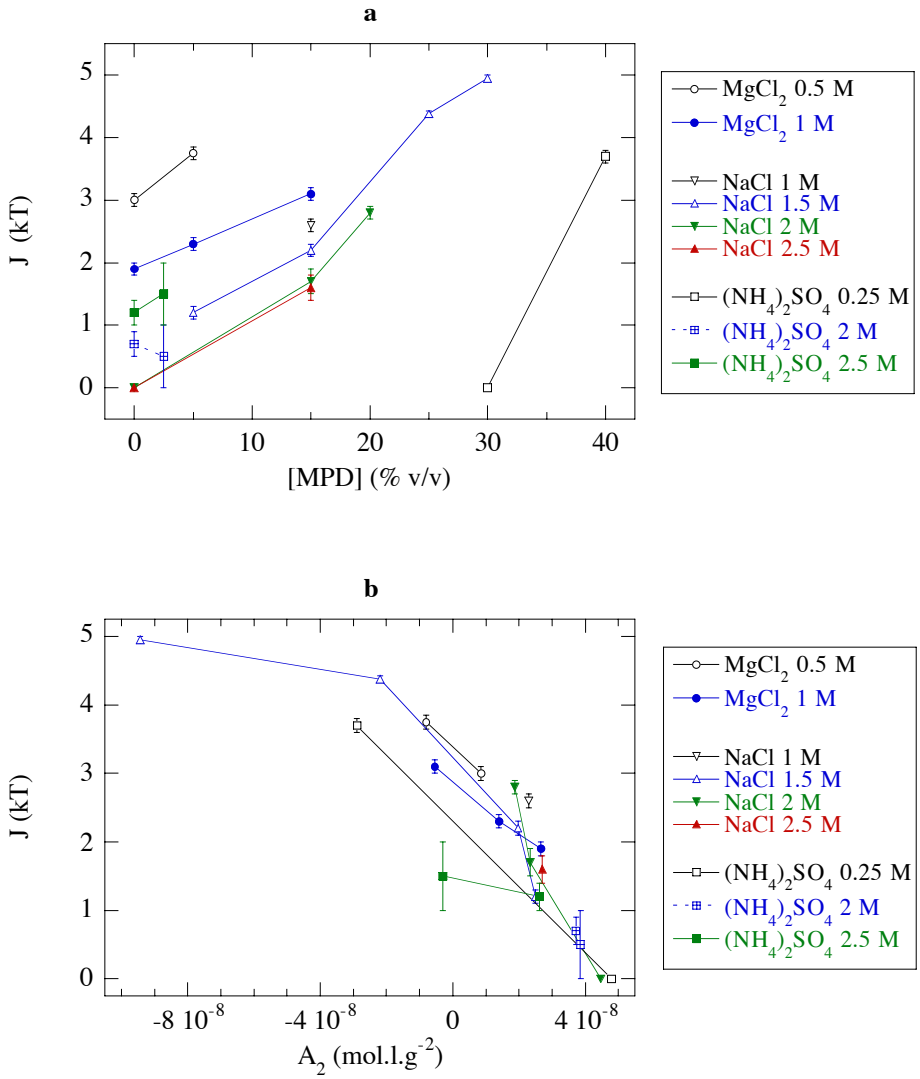


Figure 6.10 : Profondeurs J (en kT) du potentiel attractif de type Yukawa de 5 Å de portée utilisé pour modéliser les facteurs de structure expérimentaux de la *Hm* MalDH dans les mélanges sel – MPD – H_2O, à pH 8-8.2 et à 20-25°C : **(a)** en fonction de la concentration en MPD, **(b)** en fonction du second coefficient du viriel A_2 (cf. Tableau 6.3 p.157). Le diamètre D du potentiel de volume exclu utilisé est de 80 Å et la charge Z de la particule en solution est de 70 en NaCl et $MgCl_2$ et de 30 en $(NH_4)_2SO_4$.

En $MgCl_2$, la croissance des valeurs de J en fonction de la concentration en MPD est plus marquée à 0.5 M qu'à 1 M. La même tendance s'observe en $(NH_4)_2SO_4$ en comparant les

évolutions de J à 0.25 M et à 1.5-2.5 M. La croissance de J avec la teneur en MPD est similaire en NaCl 1.5 et 2 M et semble très légèrement moins prononcée à 2.5 M. L'intensité de l'effet du MPD dépend donc de la concentration du sel : elle est plus marquée à faible salinité.

La Figure 6.10-b nous permet de comparer l'évolution du second coefficient du viriel A_2 et de la profondeur J du potentiel attractif à salinité constante pour les différents mélanges sel – MPD –H_2O étudiés. Quelles que soient la nature du sel et sa concentration, on constate que les deux paramètres sont très bien corrélés : une baisse de A_2 se traduit par une augmentation de J.

6.7. Pourquoi le MPD modifie-t-il les comportements de la *Hm* MalDH ?

6.7.1. Les effets du MPD sur la Hm *MalDH*

L'influence du MPD sur les seconds coefficients du viriel A_2 de la *Hm* MalDH (active) en solutions salines est présentée pour les trois sel étudiés : NaCl, $MgCl_2$ et $(NH_4)_2SO_4$ dans la Figure 6.12, sur les diagrammes de phases sel – MPD – eau en fonction des concentrations des co-solvants. Les résultats semi-quantitatifs de stabilité (" a " protéine active, " i " inactive) et de solubilité (" P " protéine précipitée selon le sel à 20, 25 ou 30 mg/ml) y sont aussi rappelés. Cette figure et la Figure 6.10 nous permettent de dégager les principaux effets du MPD que nous avons observés.

Pour les différents sels étudiés, les diagrammes de phases sel – MPD – H_2O obtenus sont similaires, seules les positions des séparations entre les zones du diagramme étant décalées. Le MPD stabilise la *Hm* MalDH en NaCl et à bas $MgCl_2$, bien qu'il ne se suffise pas en lui-même. Son effet sur la stabilité est cependant plus complexe en $(NH_4)_2SO_4$. Pour les trois sels, la baisse progressive des valeurs du second coefficient du viriel avec l'ajout de MPD et l'augmentation correspondante de la profondeur J du potentiel attractif utilisé pour modéliser les facteurs de structure traduisent l'addition d'une interaction protéine – protéine attractive du fait de la présence du MPD. Cette attraction supplémentaire conduit, à forte

concentration de MPD et à basse salinité, à une faible solubilité de la *Hm* MalDH qui peut précipiter et dans certaines conditions (NaCl et $(NH_4)_2SO_4$) cristalliser. De plus, le MPD étend la zone de précipitation de la protéine observée à forte concentration de $(NH_4)_2SO_4$.

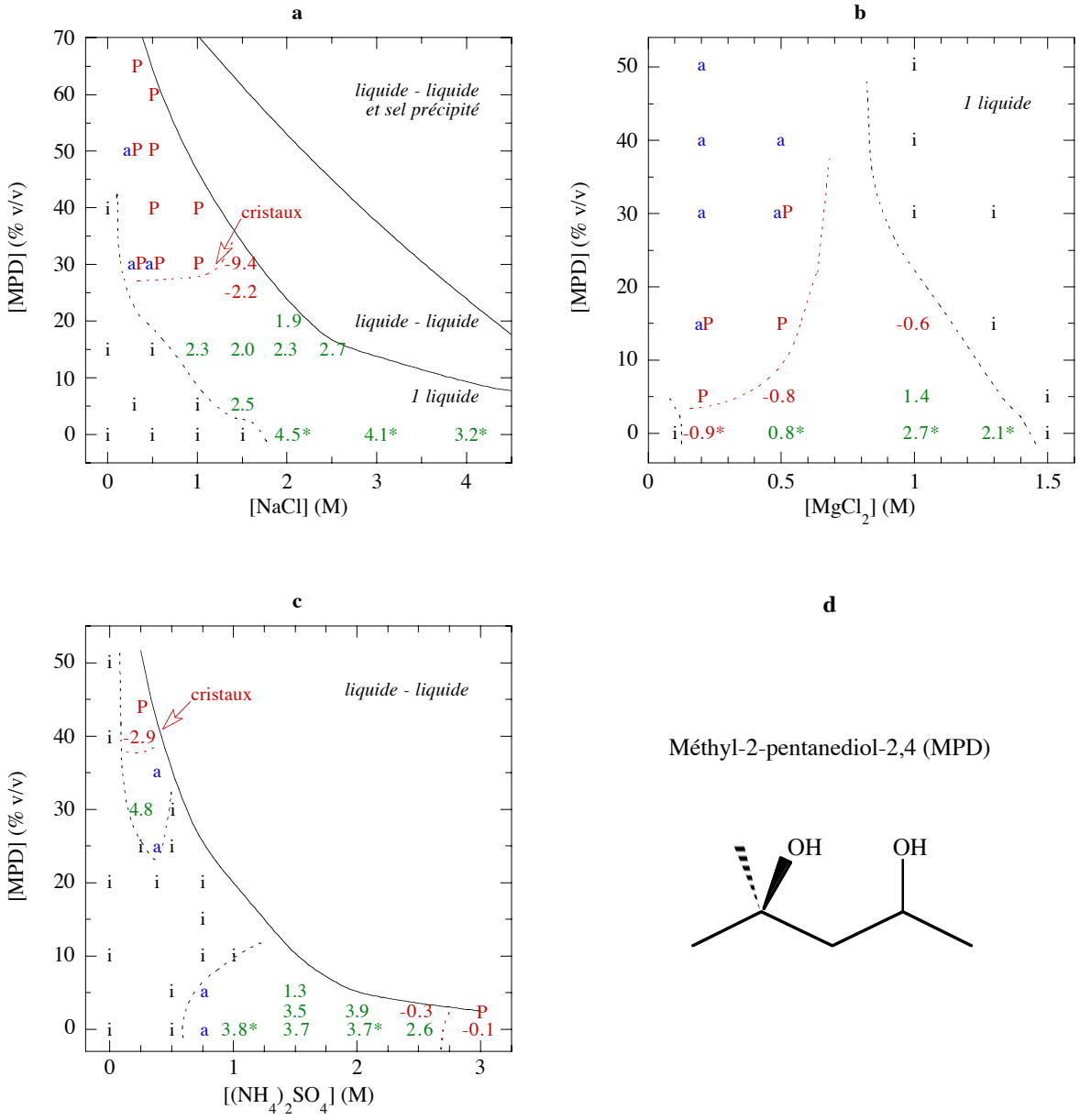


Figure 6.12 : Activité, solubilité et seconds coefficients du viriel (A_2) de la *Hm* MalDH dans des mélanges sel – MPD – eau. Les lignes pleines noires sont les courbes de coexistence de phases [en NaCl, d'après Richard *et al.*, 1995]. Les zones où la protéine est active (a) ou inactive (i) sont séparées par un trait en pointillés noirs. De même, les zones où elle précipite (P) à 20, 25 ou 30 mg/ml (respectivement en **c**, **a** et **b**) sont délimitées en rouge. Les nombres correspondent aux valeurs de A_2 en 10^{-8} mol.l.g^{-2} à pH 8-8.2 et à 20-25°C (réf. NaCl et $MgCl_2$: D11 99.6, $(NH_4)_2SO_4$: D22 00.10) ou à 5°C pour les valeurs marquées d'un astérisque (réf. D11 97.4, d'après [Ebel *et al.*, 1999b]). La protéine est active lorsqu'une valeur de A_2 a été déterminée.

6.7.2. Remarque préliminaire

Le MPD est un agent précipitant largement utilisé dans la cristallisation des macromolécules biologiques. Son effet sur leur solubilité est (parfois) attribué à la forte baisse de la constante diélectrique du solvant qu'il induit. Certes, ce solvant organique réduit fortement ce paramètre et donc les interactions électrostatiques répulsives favorisant ainsi la précipitation des protéines. Mais il faut comparer cette baisse à celle induite par la présence de sel dans le solvant. Ainsi, si la constante diélectrique de l'eau (à 20°C) baisse de 80 à 54 pour 50% de MPD, la même variation est obtenue pour 2.3 M de NaCl, 2.7 M de KCl ou 1.1 M de $MgCl_2$ (cf. annexe 8.1 p.187). Les effets du MPD sur les interactions protéine – protéine doivent donc avoir d'autres origines.

6.7.3. Cadre de l'analyse

En absence de MPD, nous avons vu que les interactions protéine – protéine de la *Hm* MalDH avaient deux origines : d'une part les interactions " directes " entre les protéines principalement via leur volume (effet de volume exclu) et, dans une moindre mesure, via leurs interactions électrostatiques, et d'autre part, les interactions protéine – solvant qui au-delà d'un seuil peuvent induire une attraction. Pour quelles raisons l'ajout de MPD perturbe-t-il ces interactions ? Rappelons que le MPD est un solvant organique hydrophile de par ses deux groupements hydroxyles, mais aussi fortement hydrophobe à cause de ses groupements hydrocarbonés (Figure 6.12-d). Le MPD peut donc favorablement interagir avec les chaînes latérales hydrophobes, ce qui lui confère un caractère déstabilisant de la structure repliée de la protéine. Par ailleurs, comme il diminue la tension superficielle de l'eau, on peut donc penser que sa présence va contribuer à la baisse de l'hydratation préférentielle des surfaces. Cependant, Pittz et Timasheff ont décrit les interactions de la ribonucléase A (protéine non-halophile) avec le solvant dans des solutions aqueuses de MPD sans sel. Ils montrent que cette protéine est, en présence de MPD, préférentiellement hydratée et interprètent leurs résultats en termes de répulsion locale du MPD par les charges de surface de la protéine. Cette interaction thermodynamique défavorable conduit à une séparation de phases : la cristallisation de la protéine [Pittz et Timasheff, 1978]. De manière similaire, ces auteurs notent que le MPD provoque la démixtion de solutions salines variées comme nous l'avons observé. Ils en concluent que les solutions aqueuses de MPD sont fortement déstabilisées par le contact avec des molécules chargées.

Ce comportement pourrait-il alors expliquer les effets du MPD sur la stabilité de la *Hm* MalDH et sur ses interactions protéine – protéine ?

Nos solutions de protéine dans des mélanges sel – MPD – H_2O sont des systèmes à quatre composants dont la description est plus complexe que les systèmes protéine – MPD – H_2O et protéine – sel – H_2O. Dans un système à trois composants, Casassa et Eisenberg ont montré que les interactions protéine – protéine pouvaient être décrites comme si le système n'avait que deux composants (protéine – solvant) en utilisant néanmoins des paramètres thermodynamiques d'interaction préférentielle (avec l'eau ou alternativement avec le co-solvant) pour tenir compte des perturbations induites par les interactions avec le solvant [Casassa et Eisenberg, 1964; Eisenberg, 1976]. Nous avons nous-même utilisé cette description pour la *Hm* MalDH dans des solutions salines (cf. (4.9) p.118), comme Pittz et Timasheff l'ont fait pour la ribonucléase A dans des solutions aqueuses de MPD [Pittz et Timasheff, 1978]. L'extension de cette description à nos systèmes à quatre composants fait intervenir non seulement les termes protéine – sel et protéine – MPD pour la *Hm* MalDH, mais aussi un terme supplémentaire sel – MPD (cf. annexe 8.3 p.193). Malheureusement, nous ne connaissons pas ces interactions. Nous allons cependant utiliser les résultats obtenus pour les systèmes ternaires pour tenter d'expliquer qualitativement les effets que nous observons.

6.7.4. MPD et stabilité de la Hm *MalDH*

La stabilité d'une protéine résulte d'un équilibre complexe entre sa forme native repliée et une forme partiellement ou totalement dépliée. La surface exposée au solvant n'est pas la même dans les deux situations. Les interactions de la protéine avec les composants du solvant peuvent donc différer entre les deux états. Nous avons interprété les résultats de stabilité et d'interaction avec le solvant par différents effets. Le NaCl et le $MgCl_2$ (à faible concentration) ont un effet stabilisant sur la *Hm* MalDH car ils se lient préférentiellement à sa forme repliée. Le $MgCl_2$ à forte concentration déstabilise la protéine car il se lie préférentiellement à la forme dépliée. Par contre, le $(NH_4)_2SO_4$ aurait un effet stabilisant en favorisant l'accumulation d'eau à la surface de la protéine qui se replie donc pour réduire sa surface exposée. Cet effet est en partie lié à l'augmentation de la tension superficielle de l'eau par ce sel [Ebel *et al.*, 1998; Ebel *et al.*, 1999a]. Gardant ces faits en mémoire, nous allons voir comment le MPD peut stabiliser ou déstabiliser la *Hm* MalDH.

Dans nos mélanges sel – MPD – H_2O, le MPD devrait fortement repousser les ions du sel à cause de son interaction défavorable avec les molécules chargées. De ce fait, il favoriserait leur accumulation près de la protéine qui localement se trouverait dans une solution en apparence plus salée : l'effet stabilisant ou déstabilisant des ions du solvant qui interagissent spécifiquement avec la protéine serait renforcé. Ceci expliquerait l'effet stabilisant du MPD que nous observons en dessous de 2 M de NaCl et le fait qu'il ne perturbe pas la stabilité de la *Hm* MalDH en $MgCl_2$ 0.2 ou 0.5 M (cf. Figure 6.12-a et b). A l'opposé, la protéine est inactive au-dessus de 1.5 M de $MgCl_2$, sans MPD, et son ajout déplace progressivement cette transition vers de plus basses concentrations en sel : la protéine est par exemple dénaturée en $MgCl_2$ 1 M, MPD 30%. Nous pouvons raisonnablement proposer que, en facilitant l'accumulation de sel près de la protéine, le MPD favorise donc dans ce cas sa dénaturation.

Ce mécanisme de répulsion du sel par le MPD ne permet pas d'assurer la stabilité de la *Hm* MalDH à très faible salinité ou en absence de sel. Effectivement, elle est inactive en absence de sel jusqu'à au moins 50% de MPD. Au contraire, on sait que les molécules de MPD de par leur caractère hydrophobe peuvent interagir avec les résidus non-polaires exposés par la protéine dénaturée [Arakawa *et al.*, 1990]. On peut penser que pour la *Hm* MalDH cette interaction sera facilitée par la charge surfacique plus faible de la forme dépliée de la protéine. En conséquence de ces interactions, la forme dénaturée devrait être relativement soluble. Nous avons observé le résultat de ce mécanisme en MPD 50% et sans sel où nous avons déterminé un rayon de giration de l'ordre de 130 Å, valeur compatible avec une protéine partiellement dépliée (cf. Figure 6.3-e et f p.151).

Dans les mélanges $(NH_4)_2SO_4$ – MPD – H_2O, l'effet de MPD sur la stabilité de la *Hm* MalDH est plus complexe. Un enrichissement en ions au voisinage de la protéine, induit par le MPD, pourrait rendre compte de la zone de stabilité que nous avons déterminée à faible concentration de sel (entre environ 0.25 et 0.5 M) au-dessus de 25% de MPD. Cependant, il ne peut expliquer l'effet déstabilisant du MPD aux salinités intermédiaires. En effet, la *Hm* MalDH en $(NH_4)_2SO_4$ 0.75 M est active à 0% de MPD et inactive dès 10% (cf. Figure 6.12-c). Dans cette zone, le sulfate d'ammonium est décrit pour stabiliser la *Hm* MalDH par un mécanisme d'hydratation non-spécifique, notamment en augmentant la tension superficielle de l'eau. Notons que le MPD a un effet contraire : il diminue très efficacement la tension de surface de l'eau [Kita *et al.*, 1994]. Il pourrait ainsi contrecarrer l'effet du $(NH_4)_2SO_4$ et déstabiliser la protéine.

Ces arguments – certes spéculatifs – permettent d'apporter une explication qualitative des différents comportements de stabilité de la *Hm* MalDH que nous avons observés. Leur diversité provient de la nature complexe du MPD : hydrophile – hydrophobe – et surtout repoussant les charges. L'équilibre subtil entre la forme repliée et dénaturée de la protéine fait en effet intervenir les interactions entre ces deux formes et les différents composants du solvant qui peuvent être spécifiques ou non et avoir une action sur l'ensemble de la protéine ou sur des régions très localisées. L'importance relative de chaque interaction dépend des conditions de solvant, d'où la complexité des effets du MPD en présence de sels sur la stabilité de la protéine.

6.7.5. MPD et interaction protéine – protéine

A l'opposé de cette complexité, le MPD n'a qu'un effet sur les interactions protéine – protéine : il introduit une attraction. L'ajout de MPD induit dans un premier temps une baisse des valeurs du second coefficient du viriel A_2 et une augmentation corrélée de la profondeur J du potentiel attractif utilisé pour modéliser les facteurs de structure. Dans un deuxième temps, l'ajout de MPD peut provoquer la précipitation de la protéine.

L'attraction induite par le MPD

Pittz et Timasheff ont montré que le MPD était rejeté par les molécules chargées, à savoir les ions salins ou la ribonucléase A [Pittz et Timasheff, 1978]. De la même manière, le MPD devrait être très fortement repoussé par la surface de la *Hm* MalDH qui est extrêmement chargée négativement. Cette répulsion correspond à une attraction effective entre les protéines, comme pour la force de déplétion provoquée par l'exclusion stérique d'un co-solvant (cf. Figure 1.4 p.36). L'origine de notre attraction diffère cependant d'une exclusion stérique : elle proviendrait dans notre cas de la nature chargée de la surface de la protéine. Nous avons remarqué que les valeurs de J augmentaient plus rapidement à faible salinité (cf. Figure 6.10 p.165). Pour une concentration donnée de MPD, le potentiel attractif est donc plus intense lorsque le solvant est moins salé : par exemple, en MPD 15%, J décroît de 2.6 à 1.6 kT lorsque la concentration en NaCl augmente de 1 à 2.5 M. La même observation peut être faite entre le $MgCl_2$ 0.5 et 1 M pour 5% de MPD, même en ne considérant que la partie de J due à la présence du MPD (en $MgCl_2$ seul, J n'est pas nul). Ces observations corroborent notre interprétation qualitative de l'effet du MPD sur les interactions protéine – protéine en

termes de répulsion du solvant organique par la surface chargée de la *Hm* MalDH. En effet, du fait de l'écrantage des forces électrostatiques par le sel, la *Hm* MalDH semble moins chargée à forte salinité : elle repousse donc moins efficacement le MPD. Ainsi, à faible concentration en sel, la charge effective élevée de la protéine correspond à une attraction intense et, inversement, à forte salinité une charge effective amoindrie engendre une attraction plus faible et donc une profondeur J moindre. Par contre, la décroissance de l'attraction avec la salinité n'apparaît pas forcément dans les valeurs du second coefficient du viriel A_2 puisque les deux effets (attraction causée par le MPD et répulsion électrostatique) conduisent à des valeurs opposées pour A_2 (respectivement négative et positive) qui peuvent se compenser. En définitive, l'origine de l'attraction induite par le MPD est la répulsion de ce co-solvant par la surface chargée de la *Hm* MalDH.

Cette interprétation expliquerait l'ensemble de nos résultats de modélisation et nos valeurs de A_2, à l'exception d'une condition de solvant : $(NH_4)_2SO_4$ 0.25 M, MPD 30%. Elle correspond en effet à un A_2 de 4.8 10^{-8} mol.l.g^{-2} qui semble être bien plus élevé que celui attendu en présence de 30% de MPD (cf. Figure 6.12-c p.167). De plus, la modélisation (qualitative) du facteur de structure expérimental correspondant ne semble pas nécessiter de potentiel attractif (cf. Figure 6.9-c p.164). A faible salinité, le terme électrostatique écranté (répulsif) des interactions protéine – protéine est évidemment plus important : il pourrait alors expliquer cette forte valeur de A_2.

Précipitation de la protéine par le MPD

La solubilité d'une protéine est reliée à ses interactions interparticulaires qui comprennent ses interactions avec le solvant. En effet, la pression osmotique Π peut s'exprimer d'une part selon un développement du viriel (cf. (1.17) p.40) et d'autre part en fonction du potentiel chimique μ_1 du solvant principal et de son volume molaire V_1 :

$$(6.1) \qquad \Pi / \left(N_A kT \right) = c \left(M^{-1} + A_2 c + \cdots \right) = - \left(\mu_1 - \mu_1^{\circ} \right) / \left(N_A kT V_1 \right) .$$

Pour les valeurs négatives de $\left(\mu_1 - \mu_1^{\circ} \right)$, la solution est stable. Inversement, pour des valeurs positives, le système est instable et il en résulte une séparation de phases. La limite entre les deux états correspond à une concentration c en protéine qui est définie comme la solubilité de la protéine [Arakawa et Timasheff, 1985b]. L'équation précédente relie donc la solubilité d'une protéine à ses interactions interparticulaires exprimées par le second coefficient du

viriel A_2. En se limitant au terme du second ordre du développement, on constate qu'une précipitation n'est possible que pour les valeurs négatives de A_2. Une très belle corrélation entre des A_2 négatifs et la solubilité de la protéine a été expérimentalement obtenue pour notamment pour le lysozyme dans des sels variés [Bonneté *et al.*, 1999; Guo *et al.*, 1999].

Dans nos mélanges sel – MPD – H_2O, le second coefficient du viriel de la *Hm* MalDH baisse proportionnellement à la quantité de MPD ajoutée et devient négatif (lorsque la démixtion du solvant ne limite pas la concentration en MPD). Au-delà, on constate chaque fois la précipitation de la protéine pour des concentrations de l'ordre de 20-30 mg/ml (Figure 6.12 p.167). L'expression (6.1) qui relie la solubilité aux valeurs négatives de A_2 explique le comportement que nous observons. Toutefois, elle n'explique pas que la protéine précipite sous forme ordonnée dans le cas de la cristallisation ou sous forme amorphe.

Les conditions de cristallisation des protéines correspondent à des interactions interparticulaires moyennement attractives comme l'ont montré George et Wilson pour de nombreuses protéines par mesure de A_2. A l'inverse, de très fortes interactions attractives conduisent à des précipités amorphes [George et Wilson, 1994]. D'après l'intensification de l'attraction que nous observons pour des teneurs croissantes de MPD, on peut raisonnablement supposer que la précipitation de la *Hm* MalDH à faible salinité, au-dessus de 5, 30 et 40% de MPD pour respectivement $MgCl_2$, NaCl et $(NH_4)_2SO_4$ est due à de fortes interactions attractives. D'autre part, les conditions de cristallisation de la *Hm* MalDH correspondent effectivement à des interactions moyennement attractives. En effet, des cristaux de protéine sont obtenus en $(NH_4)_2SO_4$ 0.25 M, MPD 40% et pour des solvants proches du mélange NaCl 1.5 M, MPD 30% qui correspondent à des A_2 respectifs de -2.9 et -9.4 10^{-8} mol.l.g^{-2} (cf. Figure 6.12 p.167).

En revanche, nous retrouvons des valeurs faiblement négatives pour A_2 dans une autre zone du digramme de phases $(NH_4)_2SO_4$ – MPD – H_2O : vers 2.5 M de sel. Toutefois, comme mentionné dans le §6.4.2 (p.152) nous n'avons observé que des précipités ou des gels blanchâtres dans ces conditions de solvant. Des interactions faiblement attractives mais suffisamment pour provoquer une séparation de phases ne conduisent pas nécessairement à la cristallisation de la protéine. D'autres paramètres sont donc importants.

Les cristaux que nous avons obtenus ne sont pas apparus immédiatement et leur croissance s'est étalée sur plusieurs semaines. A l'inverse, les précipités sont tous apparus en quelques minutes. La cinétique de la séparation de phases détermine donc son résultat (cristal

ou précipité amorphe) ce qui est bien reconnu [Ducruix et Giegé, 1992]. Ainsi, une organisation locale et suffisamment lente des protéines en solution permettrait l'apparition d'un ordre cristallin via des contacts précis alors qu'une organisation globale et rapide donnerait lieu à des agrégats ou des gels moins ordonnés qui ferait alors intervenir des contacts plus aléatoires entre les protéines.

Du point de vue des potentiels interparticulaires, cette notion cinétique se traduit par la présence ou non d'une barrière de potentiel, *i.e.* lorsque la distance interparticulaire diminue le potentiel commence par augmenter (situation répulsive) avant de diminuer (situation attractive). Les particules qui ont une énergie cinétique suffisante peuvent franchir cette barrière et se retrouvent piégées dans le puit de potentiel. En revanche, sans barrière de potentiel toutes les particules peuvent se retrouver piégées. Cette différence explique la relative lenteur de la précipitation lorsqu'une barrière de potentiel existe. Cette existence dépend à la fois des portées des potentiels attractif et répulsif et de leur intensité.

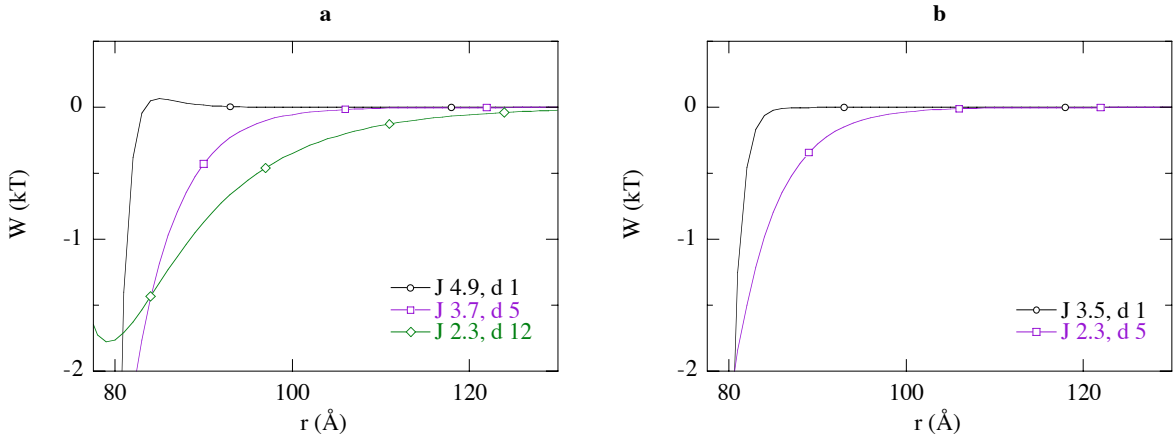


Figure 6.14 : Résultante des potentiels électrostatique et attractif de type Yukawa utilisés pour modéliser les facteurs de structure expérimentaux de la *Hm* MalDH en **(a)** $(NH_4)_2SO_4$ 0.25 M, MPD 40% et **(b)** $(NH_4)_2SO_4$ 3 M, MPD 0% (réf. D22 00.10). Le potentiel de volume exclu n'est pas pris en compte, il ajouterait simplement un terme infiniment positif en dessous de 80 Å et nul au-dessus pour un diamètre D de la particule de 80 Å. La charge Z de la protéine est de 30. La profondeur J du potentiel attractif est exprimée en kT et sa portée d en Å.

Les paramètres (Z, J) que nous avons utilisés lors de la modélisation des facteurs de structure correspondant aux conditions de cristallisation de la *Hm* MalDH n'engendrent pas de barrière de potentiel pour une portée du potentiel attractif fixée à 5 Å (cf. p.65). Cependant, en faisant varie d (J étant ajusté pour rendre compte des facteurs de structure), on constate que, pour un d de 1 Å, une petite barrière de potentielle apparaît pour la condition de solvant

conduisant à la cristallisation de la *Hm* MalDH (Figure 6.14-a) mais pas pour la condition entraînant sa précipitation sous forme de gel (graphique b). En supposant que la présence d'une barrière de potentiel soit nécessaire à la cristallisation, notre choix de fixer d à 5 Å serait donc inadapté. Toutefois, une différence entre 1 et 5 Å pour d donne lieu à des facteurs de structure calculés qui sont difficilement discernables et dont l'écart est bien inférieur à la précision de nos mesures.

En définitive, pour les mélanges sel – MPD – H_2O que nous avons étudiés, le MPD modifie les interactions faibles protéine – protéine de la *Hm* MalDH en leur ajoutant une composante attractive quel que soit le sel. Cette attraction peut être attribuée à la répulsion des molécules de MPD par la surface chargée de la protéine. On constate donc une nouvelle fois que les interactions protéine – solvant contribuent largement aux interactions protéine – protéine. De plus, l'attraction ajoutée par le MPD peut, en présence de concentrations relativement faibles de sels variés (NaCl, KCl ou $(NH_4)_2SO_4$)), induire la cristallisation la *Hm* MalDH.

7. CONCLUSION

Dans quelle mesure les interactions protéine – solvant influencent-elles les interactions faibles protéine – protéine en solution ?

Nous avons, dans cette étude, tenté de répondre à cette question en mettant en relation ces deux classes d'interactions pour la malate déshydrogénase halophile de *Haloarcula marismortui* (*Hm* MalDH). Cette enzyme bien caractérisée dans de nombreux solvants nous semblait en effet être un excellent modèle du fait de ses solvatations très riches à la fois en eau et en sel, mais diverses selon la nature du sel. Nous avons donc étudié les interactions interparticulaires de la *Hm* MalDH dans différents solvants aqueux salins et mixtes sel – méthyl-2-pentanediol-2,4 (MPD).

Une nouvelle méthode de détermination du second coefficient du viriel

Nous avons développé une nouvelle méthode de détermination du second coefficient du viriel A_2 (qui reflète l'ensemble des interactions faibles protéine – protéine) par ultracentrifugation analytique. Il s'agit de la modélisation des profils de vitesse de sédimentation en tenant compte de la dépendance à la concentration protéique des coefficients de diffusion translationnelle et de sédimentation. Notre procédure permet de caractériser la masse molaire et les propriétés hydrodynamiques des macromolécules même très chargées ou présentant un comportement fortement non idéal. Elle permet d'étudier des solvants complexes contenant de fortes concentrations de sels et de solvant organique, avec seulement 2 mg récupérables de protéine par condition. La précision des A_2 ainsi déterminés est suffisante pour discriminer des conditions de solvant globalement répulsives ou attractives. La modélisation des profils de vitesse de sédimentation apparaît donc comme une bonne technique de recherche de conditions de cristallisation d'une protéine [Solovyova *et al.*, 2001].

Interactions protéine – protéine en différents sels

Dans notre exploration des interactions protéine – protéine de la *Hm* MalDH en différents solvants salins aqueux, nous avons confirmé les résultats publiés à ce jour [Ebel *et al.*, 1999b] et étendu la caractérisation de ces interactions à d'autres sels. La détermination des seconds coefficients du viriel A_2 par diffusion de neutrons aux petits angles nous a permis de

mettre en évidence deux classes de solvant. La première est constituée des sels suivant : NaCl, CsCl, KF et $NaCH_3CO_2$ pour lesquels A_2 conserve des valeurs positives comprises entre 2 et 4 10^{-8} mol.l.g^{-2} qui diminuent faiblement et de manière monotone avec la salinité. En revanche, A_2 évolue fortement en $(NH_4)_2SO_4$, $MgSO_4$ et $MgCl_2$ qui forment la seconde classe de solvants. Pour ces sels, il varie de manière non monotone avec la salinité, entre des valeurs nulles ou faiblement négatives (à basse ou forte salinité) et des valeurs positives de l'ordre de 3 à 4 10^{-8} mol.l.g^{-2}.

La mesure des facteurs de structure expérimentaux S_{exp} par diffusion de neutrons ou de rayons X aux petits angles nous a ensuite permis de les modéliser correctement dans la plupart de ces solvants, grâce à trois potentiels à symétrie sphérique et malgré leur simplicité. Du potentiel de volume exclu, nous avons déduit pour la *Hm* MalDH en solution un diamètre D de 80 à 82 Å, valeur qui s'accorde avec sa structure cristallographique tridimensionnelle. Par contre, la charge de la particule en solution reste indéfinie (entre 1 et 156) à cause des fortes salinités de nos solvants qui limitent fortement les effets du potentiel électrostatique écranté sur les facteurs de structure. Finalement, pour les seuls sels qui, à certaines concentrations, induisent des interactions protéine – protéine neutres ou faiblement attractives (deuxième classe de solvants), nous avons pu modéliser cet effet par l'ajout d'un potentiel attractif de type Yukawa d'une portée fixée à 5 Å et dont la profondeur J évolue régulièrement avec la salinité.

De l'analyse de l'ensemble de ces résultats, nous avons dégagé les déterminants des interactions protéine – protéine de la *Hm* MalDH dans ces solvants salins. Les interactions interparticulaires répulsives sont en grande partie dues au terme de volume exclu et dans une moindre mesure au terme électrostatique. La somme de ces deux termes peut raisonnablement rendre compte des interactions protéine – protéine pour la première classe de solvants. La *Hm* MalDH en solution serait donc une particule relativement chargée pour laquelle la répulsion électrostatique n'est pas négligeable, même pour des forces ioniques de l'ordre du molaire. Par contre, pour la seconde classe de solvants, une attraction dont l'origine n'était pas déterminée est nécessaire. Ce terme attractif provient des interactions protéine – solvant qui sont transcrites par les paramètres d'interaction préférentielle de la protéine avec le sel ($\partial m_3/\partial m_2$). Pour la *Hm* MalDH, nous avons en effet montré en $(NH_4)_2SO_4$ et en $MgCl_2$ que l'évolution des ($\partial m_3/\partial m_2$) avec la salinité traduit qualitativement l'évolution de la composante attractive des interactions protéine – protéine. De plus, les baisses prononcées des A_2 expérimentaux sont très nettement corrélées aux écarts prononcés des ($\partial m_3/\partial m_2$). Toutefois,

l'effet des interactions protéine – solvant sur les interactions protéine – protéine est trop important. Les paramètres d'interaction préférentielle ($\partial m_3/\partial m_2$) ou leur effet sur A_2 semblent donc être surestimés. Cet écart peut aussi signifier l'existence d'un seuil en dessous duquel les interactions protéine – solvant n'influent pas sur les interactions protéine – protéine.

Le fait que les interactions protéine – solvant soient intégrées dans les interactions protéine – protéine via les paramètres d'interaction préférentielle ($\partial m_3/\partial m_2$) traduit l'effet d'Hofmeister et donc la capacité spécifique de chaque sel à faire précipiter une protéine. Les mécanismes d'exclusion stérique, d'augmentation de la tension superficielle de l'eau, de répulsion par les charges protéiques de surface qui tous engendrent une hydratation préférentielle de la protéine ont été utilisés pour expliquer l'effet des co-solvants sur la solubilité des protéines [Timasheff et Arakawa, 1988]. En dehors du $MgCl_2$, l'accumulation de co-solvant salin n'est à notre connaissance pas reconnue pour provoquer des interactions protéine – protéine attractives. La *Hm* MalDH, protéine halophile dont les interactions fortes avec le sel ont été mesurées, nous permet de mettre en évidence ce phénomène. Pour confirmer expérimentalement cette observation, il faudrait étendre cette étude non seulement à d'autres sels pour la *Hm* MalDH, mais aussi à d'autres protéines halophiles.

Finalement, cette étude met en évidence la nécessité pour une protéine d'un organisme halophile d'intégrer de fortes quantités de sel dans sa couche d'hydratation afin de rester " neutre " par rapport au solvant et donc soluble. L'accumulation de charges de même signe à la surface de la protéine a dès lors deux effets avantageux : il augmente la composante répulsive coulombienne entre macromolécules voisines et autorise une solvatation riche en sel. En définitive, la préservation de la grande solubilité des macromolécules qui est nécessaire au bon fonctionnement de l'organisme apparaît comme une contrainte forte et donc déterminante des mécanismes moléculaires d'adaptation à l'environnement.

Compréhension de la cristallisation de la *Hm* MalDH par dilution

La *Hm* MalDH a été cristallisée dans une solution aqueuse de NaCl et de méthyl-2-pentanediol-2,4 (MPD) selon la méthode de diffusion de vapeur, mais suivant un processus original de dilution [Richard *et al.*, 1995]. Selon la concentration en sel et en MPD, le système NaCl – MPD – H_2O est mono-, bi- ou triphasique. La caractérisation de l'état de repliement de la protéine et de la composition de la goutte initiale de cristallisation, la détermination de la solubilité apparente de la protéine et de ces interactions protéine – protéine (A_2) par diffusion de neutrons aux petits angles dans différents mélanges

monophasiques de NaCl – MPD – H_2O nous ont permis de décrire entièrement ce processus et de comprendre les raisons de la cristallisation de la *Hm* MalDH [Costenaro *et al.*, 2001].

La goutte initiale qui est biphasique contient la protéine en présence d'eau, de NaCl et de MPD. Elle est mise en équilibre avec un réservoir ne contenant que du MPD aqueux. La vapeur d'eau diffuse du réservoir vers la goutte et la dilue. Au début du processus de dilution globale de la goutte, la phase fortement saline et pauvre en MPD qui contient toute la protéine évolue vers une composition moins riche en protéine et en sel, mais plus concentrée en MPD. La *Hm* MalDH est ainsi progressivement conditionnée d'interactions protéine – protéine répulsives ($A_2 > 0$) vers des interactions modérément attractives ($A_2 < 0$). La goutte devient ensuite monophasique. La nucléation se produit alors dans une solution homogène où le A_2 est raisonnablement extrapolé à -20 10^{-5} mol.ml.g^{-2}. Cette valeur est en accord avec celles qui correspondent aux conditions de cristallisation de nombreuses protéines [George et Wilson, 1994]. L'équilibration se poursuivant, la protéine, le sel et le MPD sont tous trois dilués. Les interactions protéine – protéine deviennent moins attractives, ce qui devrait conduire la protéine, de conditions propices à la nucléation, vers des conditions avantageant la croissance cristalline et donc favoriser l'apparition d'un nombre limité de cristaux de taille néanmoins raisonnable. Ce cheminement complexe permet de préserver la stabilité de la protéine durant tout le processus, même pour des concentrations en NaCl inférieures à 2 M (où en absence de MPD la protéine se dénature).

La *Hm* MalDH a été cristallisée selon le même protocole dans des mélanges de sel – MPD – H_2O avec le KCl et le KF [Richard, 1998]. Le principe de ce protocole pourrait être adapté pour la cristallisation de macromolécules stables et solubles à forte salinité et (ou) peu solubles à bas sel.

Modification des interactions par l'ajout de MPD

Au vu des résultats obtenus dans les mélanges NaCl – MPD – H_2O, nous avons étendu l'étude de ce type de mélange à d'autres sels, principalement $MgCl_2$ et $(NH_4)_2SO_4$, afin de mieux comprendre les modifications des interactions protéine – protéine induites par le MPD. Il faut noter qu'une solution protéique dans ces solvants est un système à quatre composant qui est plus complexe que les systèmes ternaires usuels.

Pour les différents sels étudiés (NaCl, KCl, $(NH_4)_2SO_4$ et CsCl), les diagrammes de phases sel – MPD – H_2O obtenus sont similaires. Seules les positions des courbes de coexistence séparant les zones du diagramme sont décalées : selon le sel, il faut ajouter des quantités de MPD différentes pour obtenir deux phases.

Dans les trois sels NaCl, $MgCl_2$ et $(NH_4)_2SO_4$, la présence de MPD modifie la stabilité de la *Hm* MalDH. A basse salinité, le MPD permet à la protéine halophile de rester active et donc correctement repliée pour des concentrations en sel où normalement elle se dénature en absence de ce solvant organique. Le MPD est donc un stabilisant de la protéine native à basse salinité. Par ailleurs, il n'altère pas sa stabilité aux fortes concentrations de NaCl ou de $(NH_4)_2SO_4$. Par contre, aux fortes concentrations de $MgCl_2$, la protéine se dénature et le MPD favorise ce phénomène. L'ensemble de ces résultats peut s'expliquer par la répulsion du MPD par les espèces chargées, un mécanisme proposé pour expliquer les effets de ce soluté (sans sel) sur les interactions protéine – solvant et protéine – protéine de la ribonucléase A [Pittz et Timasheff, 1978] : le MPD devrait repousser les ions du sel, favorisant ainsi leur accumulation près de la protéine qui localement se trouverait dans une solution en apparence plus salée. Toutefois, aux concentrations intermédiaires de $(NH_4)_2SO_4$, le MPD déstabilise la *Hm* MalDH. Ce comportement en contradiction avec la supposition précédente pourrait provenir de l'augmentation de la tension superficielle de l'eau par le MPD. Finalement, le MPD seul ne suffit pas à stabiliser la protéine, mais au contraire favorise sa forme dépliée en la solubilisant. Ces arguments – certes spéculatifs – permettent d'apporter une explication qualitative aux différents comportements de stabilité de la *Hm* MalDH que nous avons observés. Leur diversité provient à la fois de la nature complexe du MPD : hydrophile – hydrophobe – et surtout repoussant les charges, et de l'équilibre subtil entre la forme repliée et dénaturée de la protéine.

A l'opposé de cette complexité, le MPD n'a, pour les sels que nous avons étudiés, qu'un effet sur les interactions protéine – protéine : il introduit une attraction. L'ajout de MPD induit en effet une baisse des valeurs du second coefficient du viriel A_2 qui se traduit par une augmentation de la profondeur J du potentiel attractif utilisé pour modéliser les facteurs de structure. L'ajout de quantités plus importantes de MPD, en combinaison avec certaines concentrations de certains sels peut conduire à une baisse notable de la solubilité de la *Hm* MalDH. Elle précipite de manière amorphe, sous la forme agrégats ou de gel partiellement ordonnés en présence de fortes concentrations de $(NH_4)_2SO_4$, ou sous la forme de cristaux à basse concentration de NaCl ou de $(NH_4)_2SO_4$. Nos résultats sont en accord avec la répulsion du MPD par la surface chargée de la *Hm* MalDH comme source de l'attraction induite par ce co-solvant. Ils contribuent à une meilleure compréhension des causes de la cristallisation de la *Hm* MalDH et pourraient s'avérer intéressants pour la cristallogenèse d'autres macromolécules en présence de sel et de MPD.

Les interactions protéine – solvant font effectivement non seulement partie intégrante des interactions protéine – protéine, mais elles apparaissent aussi comme une contribution substantielle qui prend toute son importance dans la cristallogenèse ou l'adaptation moléculaire à l'halophilie.

8. ANNEXES

8.1. Constante diélectrique statique

Cette annexe détaille d'après les données de la littérature l'estimation des constantes diélectriques statiques ε des solutions utilisées, leur valeur est indiquée dans l'annexe 8.2.

8.1.1. Solutions aqueuses de sel

La constante diélectrique statique de l'eau ε(eau) est modifiée par la présence de sel dissout. Selon la nature et la concentration (c en M) du sel, la constante diélectrique de la solution saline aqueuse ε(sel) est donnée par la relation [Barthel *et al.*, 1998] :

(8.1) $$\varepsilon(\text{sel}) = \varepsilon(\text{eau}) + a_1 c^{0.5} + a_2 c + a_4 c^2 + a_6 c^3 \ .$$

Les valeurs des coefficients a_i, pour les sels utilisés dans cette thèse, sont présentées dans le Tableau 8.1.

Sel	a1	Δa1	a2	Δa2	a4	Δa4	a6	Δa6	c_{max} (M)
NaCl	0	0	-13.8	0.3	0.98	0.1	0	0	4
KCl	0	0	-12.4	0.3	0.95	0.1	0	0	4
MgCl2	0	0	-34	1.1	10.9	1.2	-1.5	0.3	3.6
MgSO4	14.7	1.4	0	0	-23.4	5.8	7.6	4.5	1.1
CsCl	0	0	-10.7	0.7	0.83	0.17	0	0	4.5

Tableau 8.1 : Coefficients du développement (8.1) de la constante diélectrique statique ε(sel) de solutions salines aqueuses en fonction de la concentration en sel et leurs erreurs Δa_i [d'après la table 7.6 de Barthel *et al.*, 1998]. La dernière colonne indique les concentrations maximales utilisées pour déterminer les paramètres a_i.

Pour les solutions salines aqueuses que nous avons utilisées, ε(sel) est estimé grâce à l'équation (8.1). Selon la température T (en °C), les valeurs de ε(eau) sont déterminées en appliquant la loi d'Akerlov :

(8.2) $$\log_{10} \varepsilon = a - bT \ ,$$

où a = 1.944 et b = 0.001989 proviennent des valeurs de ε(eau) du Tableau 8.2.

N'ayant trouvé aucune référence pour le sulfate d'ammonium, nous avons pour ce sel utilisé la valeur de ε(eau). Cette substitution est de peu d'influence dans nos modélisations puisque les valeurs calculées de S(q=0) différent d'environ 2 % pour des ε de 80 et de 10.

T (°C)	ε (eau)	T (°C)	ε (eau)
0	87.90	50	69.88
5	85.90	55	68.30
10	83.95	60	66.76
15	82.04	65	65.25
18	80.93	70	63.78
20	80.18	75	62.34
25	78.36	80	60.93
30	76.58	85	59.55
35	74.85	90	58.20
38	73.83	95	56.88
40	73.15	100	55.58
45	71.50		

Tableau 8.2 : Constante diélectrique statique ε de l'eau en fonction de la température T en °C [d'après Hamer, 1975].

8.1.2. Solutions aqueuses de MPD

La présence de MPD modifie aussi la constante diélectrique statique de l'eau ε(eau). Travers et Douzou ont mesuré la constante diélectrique statique de solutions aqueuses de MPD, ε(MPD), à diverses températures et concentrations de MPD (cf. Tableau 8.3) [Travers et Douzou, 1974; Douzou *et al.*, 1976].

Nous avons donc déterminé ε(MPD) pour nos solutions en utilisant la loi d'Akerlov (8.2) pour les valeurs des paramètres a et b indiquées dans le Tableau 8.3. Lorsque notre concentration en MPD n'était pas tabulée, nous avons interpolé linéairement la valeur de ε(MPD) entre deux concentrations de MPD.

[MPD] (% v/v)	T (°C)					a	b x 10^3
	20	10	0	-10	-20		
0	80.4	84.2	88.1			1.945	2
10	75.4	79.2	83.2			1.92	2.13
20	70.8	74.7	78.8			1.895	2.33
30	65.7	69.7	73.5	78		1.866	2.48
40	59.6	63.7	67.3	71.3	75	1.829	2.5
50	53.7	57	60.5	64.5	68.3	1.783	2.6
60	47.7	50.6	53.5	56.9	60.5	1.728	2.59
70	41.3	43.8	46.8	49.4	52.6	1.668	2.6
80	35.6	37.8	40.3	42.9	45.3	1.605	2.71
90	30.6	32.6	34.7	36.9	39.1	1.54	2.72
100	26.3	28	29.9	31.9	31	1.476	2.8

Tableau 8.3 : Constante diélectrique statique de solutions aqueuses de MPD, ε(MPD), en fonction de la température (T en °C) et de la concentration en MPD (en % v/v). L'évolution de ε(MPD) en fonction de T obéit à la loi d'Akerlof (8.2) : $\log_{10} \varepsilon = a - bT$ [d'après Travers et Douzou, 1974; Douzou *et al*., 1976].

8.1.3. Solutions aqueuses de sel et de MPD

L'influence simultanée du sel et du MPD sur la constante diélectrique statique de solutions sel – MPD – H_2O n'est pas tabulée. Nous avons donc considéré une telle solution ternaire comme un mélange de deux solutions binaires : sel – H_2O et MPD – H_2O dont les constantes diélectriques respectives ε(sel) et ε(MPD) sont déterminées comme précédemment. La constante diélectrique ε de la solution ternaire est la somme pondérée de ε(sel) et ε(MPD). Pour chaque solution sel – MPD – H_2O, nous avons fait trois estimations avec trois mélanges différents et pris la moyenne des trois ε obtenus.

Pour le sulfate d'ammonium, nous avons utilisé les valeurs de ε(MPD) correspondantes.

8.2. Conditions et résumé des résultats des expériences de diffusion

Le tableau suivant résume les conditions et les résultats des expériences de diffusion de neutrons ou de rayons X citées dans ce manuscrit.

Remarques

Tous les solvants sont tamponnés par du Tris 50 mM, à l'exception des solvants à pH 6.5, tamponnés par du Mes 10-50 mM.

Les valeurs de la constante diélectrique ε du solvant tient compte de la présence du sel – sauf en $(NH_4)_2SO_4$ – et du MPD. Lorsque les expériences ont été réalisées à plusieurs températures, ε est donnée pour la température en gras.

Les expériences dont la référence est " D11 97.4 " ou " CE " ont été réalisées par C. Ebel [Ebel *et al.*, 1999b].

Solvant	[sel] (M)	T (°C)	pH	ε	[*Hm* MalDH] (mg/ml)	Référence	A_2 (10^{-8} mol.l.g^{-2})	Modélisation des S(q,c) Meilleures valeurs D (Å) Z J (kT)	Commentaires
NaCl	2	5	8.2	62	17.6 11.8 6.1	D11 97.4	4.5	D 80 Z 0-156	
NaCl	3	5	8.2	53	19.2 12.9 6.4	D11 97.4	4.1	D 80 Z 0-156	
NaCl	4	5	8.2	46	20.7 14 6.5	D11 97.4	3.2	D 80 Z 0-156	
NaCl	5	5	8.2	41	21.5 14.3 7.1	D11 97.4	2.6	D 80 Z 0-156	
NaCl	2	25		55	95.8 9.7	D22 97.8		D 80 Z 0-156	
NaCl	4	25		39	98.7 10.3	D22 97.8		D 80 Z 0-156	
NaCl	4	25	6.5	39	77.5 22.1 14.8 7	D22 98.10	4	D 80 Z 0-156	
NaCl	4	25	9	39	73.1 20.6 14.2 5.9	D22 98.10	3.4	D 80 Z 0-156	
NaCl	4	5 11 **20** 30 40	8	41	79.3 20.2 10.5 5.2	D22 00.4	3.9	D 80 Z 0-156 S identique pour toute T°	
NaCl	4	25	9	39	19.4 10.3 6.4	D22 98.10	4.9		Mutant E121K
NaCl, MPD 5%	1.5	20	8	60	61 15.9 10.5 5.3	D11 99.6	2.5	D 80 Z 70 J1.2	
NaCl, MPD 15%	1	20	8	61	63 16.6 11 5.4	D11 99.6	2.3	D 80 Z 70 J 2.6	
NaCl, MPD 15%	1.5	20	8	55	60 16.4 10.7 5.3	D11 99.6	2	D 80 Z 70 J 2.2	
NaCl, MPD 15%	2	20	8	51	59 16.3 10.8 5.4	D11 99.6	2.3	D 80 Z 70 J 1.7	
NaCl, MPD 15%	2.5	20	8	46	64 16.5 11.2 5.5	D11 99.6	2.7	D 80 Z 70 J 1.6	
NaCl, MPD 20%	2	20	8	48	66 15.4 10.2 5.3	D11 99.6	1.9	S: D80 Z70 J 2.8	

Solvant	[sel] (M)	T (°C)	pH	ε	[*Hm* MalDH] (mg/ml)	Référence	A_2 (10^{-8} mol.l.g^{-2})	Modélisation des S(q,c) Meilleures valeurs D (Å) Z J (kT)	Commentaires
NaCl, MPD 25%	1.5	20	8	50	68 15.2 10 5	D11 99.6	-2.2	D 80 Z 70 J 4.38	
NaCl, MPD 30%	1.5	20	8	47	50 17 11.4 5.6	D11 99.6	-9.4	D 80 Z 70 J 4.95 pour 17 mg/ml (diverge pour 50 mg/ml)	Polynôme 2e degré pour A_2
NaCl, MPD 15%	1.5	5 11 **20** 30 40	8	55	61.9 4.4	D22 00.4		D 80 Z 70 J3.3 à 2.6 pour T° de 5 à 40°C	
NaCl, MPD 15%	2.5	5 11 **20**	8	46	63.3 4.4	D22 00.4		D 80 Z 70 J 1.6-1.8 T° rien	
$NaCH_3CO_2$	1	5	8.2			CE	3.7		
$NaCH_3CO_2$	1.5	5	8.2			CE	3.1		
$NaCH_3CO_2$	2	5	8.2			CE	3.1		
$NaCH_3CO_2$	2.5	5	8.2			CE	2.1		
$NaCH_3CO_2$	3	5	8.2			CE	2.4		
$NaCH_3CO_2$	3.5	5	8.2			CE	1.7		
$(NH_4)_2SO_4$	1	5	8.2	86	18.8 12.7 6.1	D11 97.4	3.8	D 80 Z 0-156	
$(NH_4)_2SO_4$	1.5	5	8.2	86	20.5 13.7 6.8	D11 97.4	3.9	D 80 Z 0-156	
$(NH_4)_2SO_4$	2	5	8.2	86	23.7 14.1 7.7 4.9	D11 97.4	3.7	D 80 Z 0-156	
$(NH_4)_2SO_4$	3	5	8.2	86	30.6 13.6 7.6 3.7	D11 97.4	0.56	D 80 Z 30 J 2.5	
$(NH_4)_2SO_4$	1.5	25	8	78	57.8 17.1 11.3 5.5	D22 00.10	3.5	D 80 Z 0-156	
$(NH_4)_2SO_4$	2.5	25	8	78	58.1 16.6 11.3 5.5	D22 00.10	2.6	D 80 Z 30 J 1.2	
$(NH_4)_2SO_4$	3	6 **25** 31	8	78	48.2 18.3 12.2 6.1	D22 00.10	-0.8	D 80 Z 30 J 2.3 T° rien	Faible précipité blanc
$(NH_4)_2SO_4$, MPD 2.5%	1.5	25	8	78	59.8 17.2 11.7 5.7	D22 00.10	2.9	D 80 Z 0-156	Précipité blanc
$(NH_4)_2SO_4$, MPD 2.5%	2.5	6 **25** 31	8	78	60.1 18.1 11.8 5.5	D22 00.10	-1.2	## (pic corrélation puis S=1)	Précipité blanc Pic de corrélation
$(NH_4)_2SO_4$, MPD 7.5%	1.5	25	8	75	60.8 17.9 11.7 5.9	D22 00.10	4 #	D 80 Z 0-156	Précipité blanc
$(NH_4)_2SO_4$, MPD 30%	0.25	25	8	64	20.3 13.6 6.9	D22 00.10	4.8	D 80 Z 0-156	
$(NH_4)_2SO_4$, MPD 40%	0.25	6 **25** 31	8	58	18 13.7 7	D22 00.10	-2.9	D 80 Z 30 J 3.7 T° rien	Cristaux à 20.5 mg/ml
$(NH_4)_2SO_4$, MPD 50%	0	6 **25** 31	8	52	13.9 8.7 4.6	D22 00.10	##		Rg 108-125 Å
$(NH_4)_2SO_4$, MPD 2.5%	1.5	25	8	78	78.1 22.8 15 8	D11 00.8	4	D 80 Z 0-156	Précipité blanc
$(NH_4)_2SO_4$, MPD 2.5%	2	15 **25** 40	8	78	24.8 16.6 7.9	D11 00.8	3.9	D 80 Z 30 J 0.5	Précipité blanc
$(NH_4)_2SO_4$, MPD 2.5%	2.5	15 **25** 40	8	78	63.7 23.1 15.4 7.4	D11 00.8	0.7	D 80 Z 30 J 1-2 à 23 mg/ml, T° rien	Précipité blanc Pic de corrélation
$(NH_4)_2SO_4$, MPD 2.5%	3	15 **25** 40	8	78	22.2 7	D11 00.8	##		Précipité blanc Pic de corrélation
$MgCl_2$	0.2	5	8.2	80	14.7 9.3 4.9	D11 97.4	-0.9		
$MgCl_2$	0.5	5	8.2	71	17.2 11.6 6	D11 97.4	0.85		

Solvant	[sel] (M)	T (°C)	pH	ε	[*Hm* MalDH] (mg/ml)	Référence	A_2 (10^{-8} mol.l.g^{-2})	Modélisation des S(q,c) Meilleures valeurs D (Å) Z J (kT)	Commentaires
$MgCl_2$	1	5	8.2	61	20.7 14.6 2.8	D11 97.4	2.7		
$MgCl_2$	1.3	5	8.2	57	22.4 15 2.3	D11 97.4	2.1		
$MgCl_2$	0.5	23	6.5	65	13.1 8.8 4.3	D11 98.7	1.2		
$MgCl_2$	1	23	6.5	54	14.7 9.8 4.9	D11 98.7	3.1		
$MgCl_2$	0.2	5 **25**	8.2	72	30.9 3	Lure 98.12		D 87 Z 70 J 5 à 4.9 pour T° de 5 à 25°C	
$MgCl_2$	0.5	5 15 **25**	8.2	64	79.6 7.5	Lure 98.12		D 80 Z 70 J 3.2 à 3 pour T° de 5 à 25°C	
$MgCl_2$	1	5 15 **25**	8.2	54	79.2 7.8	Lure 98.12		D 80 Z 70 J 2 à 1.9 pour T° de 5 à 25°C	
$MgCl_2$	1.3	5 10 15 20 **25**	8.2	49	85 8	Lure 98.12		D 80 Z 70 J 1.8 à 1.5 pour T° de 5 à 25°C	
$MgCl_2$, MPD 5%	0.5	20	8	64	66 15.6 10.3 5.3	D11 99.6	-0.8	D 80 Z 70 J 3.75	
$MgCl_2$, MPD 5%	1	20	8	54	65 17.1 11.2 5.6	D11 99.6	1.4	D 80 Z 70 J 2.3	
$MgCl_2$, MPD 15%	1	20	8	51	61 16.9 9.5 4.6	D11 99.6	-0.5	D 80 Z 70 J 3.1	
$MgSO_4$	0.25	23	8.2	85	13.2 8.8 4.4	D11 98.7	1.3	D82 Z70 J 4	
$MgSO_4$	1	23	8.2	78	14.6 9.7 4.7	D11 98.7	3.6	D 82 Z 1-156	
$MgSO_4$	2	23	8.2	67	73 48 (24) 14.7 10.8 5	D11 98.7	4.1	D 82 Z 1-156	
KCl H_2O	3	20	8.2	52	54.3 29.8 18.1 10.7	D11 98.2	5.2	D 82 Z 1-156	
KCl D_2O	3	20	8.3		63 29 20.8 12.2	D11 98.2	5.9		
KF	1	5	8.2			CE	4.5		
KF	1.5	5	8.2			CE	4		
KF	2.5	5	8.2			CE	3.8		
KF	3	5	8.2			CE	3.8		
KF	3.5	5	8.2			CE	3.8		
KF	4.5	5	8.2			CE	3.2		
CsCl	3	24.4	8	54	73.3 23.7 7.2	D22 99.9	4.4	D 82 Z 1-156	Polynôme 2e degré pour A_2
CsCl	4	24.4	8	49	19.8 10.9 4.4	D22 99.9	3.5	D 82 Z 1-156	
CsCl	5	24.4	8	46	62.4 13.9 9.6 4.5	D22 99.9	3.4	D 82 Z 1-156	
CsCl	6	24.4	8	44	82.7 26.7 17.6 8.4	D22 99.9	3.6	D 82 Z 1-156	

8.3. Second coefficient du viriel et interactions protéine – solvant

Pour un système à un composant macromoléculaire non-diffusible (2 : la protéine) dans un solvant principal (1 : généralement l'eau) avec des co-solvants (J : le sel, le MPD, …), l'expression du second coefficient du viriel B (les concentrations m étant exprimées en molalité) est donnée par l'équation (2.49) de [Casassa et Eisenberg, 1964] :

$$2B \approx \sum_i \frac{\nu_{2i}^2}{m_i} + \beta_{22} + \sum_J a_{2J} \frac{A_{2J}}{A_{22}} = a_{22} - \frac{1}{m_2} + \sum_J a_{2J} \frac{A_{2J}}{A_{22}} \,. \tag{8.3}$$

Le potentiel chimique μ_J du composant J qui inclut ν_{iJ} moles de l'espèce i (par exemple les contre-ions d'une protéine) est :

$$\mu_J = \mu_J^o + RT \ln a_J = \mu_J^o + RT \sum_i \nu_{iJ} \ln m_i + RT\beta_J \,,$$

où $RT\beta_J$ est l'excès de potentiel chimique ou RT fois le logarithme du coefficient d'activité.

Les paramètres a_{JK} sont définis par : $a_{JK} = a_{kJ} = \left(\frac{\partial \ln a_J}{\partial m_K} \right)_{P,T,m} = \left(\frac{\partial \ln a_K}{\partial m_J} \right)_{P,T,m}$, où l'indice m signifie la constance de toutes les concentrations sauf celle qui est utilisée pour la différentiation.

A_{JK} est le cofacteur de l'élément a_{JK} du déterminant $|a_{JK}|$ pour J, K = 2, …, N soit $(-1)^{J+K}$ fois le déterminant formé en enlevant la ligne et la colonne se croisant en a_{JK},

$$|a_{JK}| = \begin{vmatrix} a_{22} & a_{23} & \cdots \\ a_{32} & \ddots & \\ \vdots & & \end{vmatrix} \,.$$

Les deux premiers termes de l'expression de B (8.3) rendent compte des interactions entre les macromolécules seulement et la dernière somme des interactions de la macromolécule avec le solvant.

Pour trois composants : eau (1), protéine (2) et sel (3), l'équation (8.3) se réduit à :

$$2B^w \approx a_{22} - \frac{1}{m_2} - \frac{a_{23}^2}{a_{33}} \,. \tag{8.4}$$

Pour quatre composants : eau (1), protéine (2), sel (3) et MPD (4), l'équation (8.3) devient :

$$2B \approx a_{22} - \frac{1}{m_2} + \left[a_{23} \frac{A_{23}}{A_{22}} + a_{24} \frac{A_{24}}{A_{22}} \right] .$$

Avec $A_{23} = -\left(a_{32}a_{44} - a_{42}a_{34}\right)$
$A_{22} = +\left(a_{33}a_{44} - a_{43}a_{34}\right)$ et comme $a_{JK} = a_{KJ}$, on obtient :
$A_{24} = +\left(a_{32}a_{43} - a_{42}a_{33}\right)$

$$2B \approx a_{22} - \frac{1}{m_2} + \left[\frac{-a_{23}^2 a_{44} + 2a_{23}a_{24}a_{34} - a_{24}^2 a_{33}}{a_{33}a_{44} - a_{34}^2} \right]$$

$$= a_{22} - \frac{1}{m_2} + \left[\frac{-a_{23}^2 a_{44} + 2a_{23}a_{24}a_{34} - a_{24}^2 a_{33} + \frac{a_{23}^2}{a_{33}}\left(a_{33}a_{44} - a_{34}^2\right)}{a_{33}a_{44} - a_{34}^2} - \frac{a_{23}^2\left(a_{33}a_{44} - a_{34}^2\right)}{a_{33}\left(a_{33}a_{44} - a_{34}^2\right)} \right],$$

$$= a_{22} - \frac{1}{m_2} - \frac{a_{23}^2}{a_{33}} + \left[\frac{-a_{23}^2 a_{44} + 2a_{23}a_{24}a_{34} - a_{24}^2 a_{33} + a_{23}^2 a_{44} - a_{23}^2 a_{34}^2 / a_{33}}{\left(a_{33}a_{44} - a_{34}^2\right)} \right]$$

$$= a_{22} - \frac{1}{m_2} - \frac{a_{23}^2}{a_{33}} + \left[\frac{2a_{23}a_{24}a_{33}a_{34} - a_{24}^2 a_{33}^2 - a_{23}^2 a_{34}^2}{a_{33}\left(a_{33}a_{44} - a_{34}^2\right)} \right]$$

soit:

(8.5) $$2B = a_{22} - \frac{1}{m_2} - \frac{a_{23}^2}{a_{33}} - \left[\frac{\left(a_{23}a_{34} - a_{24}a_{33}\right)^2}{a_{33}\left(a_{33}a_{44} - a_{34}^2\right)} \right] .$$

8.4. Détermination des interactions protéine – solvant

Les paramètres d'interaction de la *Hm* MalDH avec le solvant ont été déterminés par Ebel *et al.* à partir de mesures de densité massique et de diffusion de neutrons, selon la méthode décrite ci-dessous [Kernel *et al.*, 1999; Ebel, en préparation].

8.4.1. Détermination de $(\partial m_3 / \partial m_2)_\mu$

L'incrément de densité massique $(\partial\rho_m/\partial C_2)_\mu$ et l'incrément de densité de longueur de diffusion de neutron $(\partial\rho_n/\partial C_2)_\mu$ à potentiel chimique μ des espèces diffusibles constant peuvent s'exprimer en fonction du paramètre d'interaction préférentielle de la protéine avec le sel $(\partial m_3/\partial m_2)_\mu$ par [Eisenberg, 1976, 1981] :

(6) $$\left(\partial\rho_m / \partial C_2\right)_\mu = \left(M_+ - \bar{v}_2\rho_{ms}\right) + \left(\partial m_3 / \partial m_2\right)_\mu \cdot \left(M_3 - \bar{v}_3\rho_{ms}\right)$$

(7) $$\left(\partial\rho_n / \partial C_2\right)_\mu = \left(b_2 - \bar{v}_2\rho_{ns}\right) + \left(\partial m_3 / \partial m_2\right)_\mu \cdot \left(b_3 - \bar{v}_3\rho_{ns}\right) ,$$

où $\bar{v}_i$ est le volume partiel spécifique du composant i (ici en ml/mol) et b_i sa densité de longueur de diffusion (cm/mol). L'eau, la protéine (avec la totalité de ses contre-ions) et le sel sont respectivement les composants 1, 2 et 3 ; le solvant est noté s. Pour la diffusion de rayons X ou de lumière, des équations similaires peuvent exprimer l'incrément de densité électronique ou d'indice de réfraction.

$\bar{v}_2$ et $(\partial m_3/\partial m_2)_\mu$ sont les deux seules inconnues des équations (6) et (7), les autres paramètres étant mesurés expérimentalement ($(\partial\rho_{m,n}/\partial C_2)$, ρ_{ms}), connus (M_+, M_3, $\bar{v}_3$) ou déterminés d'après la composition du composant (b_2, b_3, ρ_{ns}). Chacune des équations (6) et (7) permet de calculer un $(\partial m_3/\partial m_2)_\mu$ à partir d'une valeur quelconque de $\bar{v}_2$. Ce système de deux équations peut donc être résolu graphiquement en reportant les valeurs des $(\partial m_3/\partial m_2)_\mu$ calculées avec l'équation (6) ou (7) pour un ensemble de valeurs de $\bar{v}_2$: la solution est donnée par l'intersection des deux droites obtenues.

8.4.2. Détermination de N_1 et N_3

Les paramètres d'interaction préférentielle avec le solvant peuvent être interprétés selon une approche structurale : à une mole de protéine sont associées N_1 moles d'eau et N_3 moles de sel. $(\partial m_3/\partial m_2)_\mu$ est alors relié à la proportion de sel dans le solvant (m_3/m_1) par les paramètres d'interaction N_1 et N_3 avec le solvant [Eisenberg, 1976, 1981] :

(8) $$\left(\partial m_3 / \partial m_2\right)_\mu = N_3 - N_1\left(m_3 / m_1\right) .$$

Lorsque $(\partial m_3/\partial m_2)_\mu$ varie linéairement avec (m_3/m_1), la particule peut être décrite comme une particule invariante par rapport à la concentration en sel du solvant, avec N_1 et N_3 constants [Tardieu *et al.*, 1981].

REFERENCES BIBLIOGRAPHIQUES

Arakawa T., Bhat R. et Timasheff S.N. (1990). Why preferential hydration does not always stabilize the native structure of globular protein. *Biochem.* 29, 1924-31.

Arakawa T. et Timasheff S.N. (1984). Mechanism of protein salting in and salting out by divalent cation salts: balance between hydration and salt binding. *Biochemistry* 23(25), 5912-23.

Arakawa T. et Timasheff S.N. (1985a). Mechanism of poly(ethylene glycol) interaction with proteins. *Biochemistry* 24(24), 6756-62.

Arakawa T. et Timasheff S.N. (1985b). Theory of protein solubility. *Methods Enzymol.* 114, 49-77.

Barthel J.M.G., Krienke H. et Kunz W. (1998). Solvent and solution permittivities of electrolyte solutions. Dans *Physical chemistry of electrolyte solutions. Modern aspects*, Vol. 5, p. 381-84. Springer, Steinkopff Darmstadt.

Belloni L. (1985). A hypernetted chain study of highly asymmetrical polyelectrolytes. *Chem. Phys.* 99(1), 43-54.

Belloni L. (1988). Self-consistent integral equation applied to the highly charged primitive model. *J. Chem. Phys.* 88(8), 5143-48.

Belloni L. (1991). Interacting monodisperse and polydisperse spheres. Dans *Neutron, x-ray and light scattering : introduction to an investigative tool for colloidal and polymeric systems : proceedings of the European Workshop on Neutron, X-Ray and Light Scattering as an Investigative Tool for Colloidal and Polymeric Systems, Bombannes, France, 27 May-2June, 1990* (Lindner P. et Zemb T., éds.), p. 135-55. North-Holland, Amsterdam.

Bonneté F., Ebel C., Eisenberg H. et Zaccai G. (1993). Biophysical study of halophilic malate dehydrogenase in solution: revised subunit structure and solvent interactions in native and recombinant enzyme. *J. Chem. Soc. Faraday Trans.* 89(15), 2659-66.

Bonneté F., Finet S. et Tardieu A. (1999). Second virial coefficient: variations with lysozyme crystallization conditions. *J. Crystal Growth* 196, 403-14.

Bonneté F., Madern D. et Zaccai G. (1994). Stability against denaturation mechanisms in halophilic malate dehydrogenase "adapt" to solvent conditions. *J. Mol. Biol.* 244(4), 436-47.

Bonneté F., Malfois M., Finet S., Tardieu A., Lafont S. et Veesler S. (1997). Different tools to study interaction potentials in γ-crystallin solutions: relevance to crystal growth. *Acta Cryst.* D53, 438-47.

Budayova M., Bonneté F., Tardieu A. et Vachette P. (1999). Interactions in solution of a large oligomeric protein. *J. Crystal Growth* 196, 210-19.

Budayova-Spano M., Lafont S., Astier J.P., Ebel C. et Veesler S. (2000). Comparison of solubility and interactions of aprotinin (BPTI) solutions in H_2O and D_2O. *J. Crystal Growth* 217, 311-19.

Casassa F.E. et Eisenberg H. (1964). Thermodynamic analysis of multicomponent solutions. *Adv. Protein Chem.* 19, 287-395.

Cendrin F., Chroboczek J., Zaccai G., Eisenberg H. et Mevarech M. (1993). Cloning, sequencing, and expression in Escherichia coli of the gene coding for malate dehydrogenase of the extremely halophilic archaebacterium *Haloarcula marismortui*. *Biochemistry* 32(16), 4308-13.

Collins K.D. (1995). Sticky ions in biological systems. *Proc. Natl. Acad. Sci. USA* 92, 5553-57.

Collins K.D. et Washabaugh M.W. (1985). The Hofmeister effect and the behaviour of water at interfaces. *Q. Rev. Biophys.* **18**(4), 323-422.

Costenaro L., Zaccai G. et Ebel C. (2001). Understanding the crystallisation of an acidic protein by dilution in the ternary NaCl - 2-methyl-2,4-pentanediol - H_2O system. *J. Crystal Growth* 232(1-4), 102-13.

Debye P. et Hückel E. (1923). *Physik Z.* 24, 185.

Delaye M. et Tardieu A. (1983). Short-range order of crystallin proteins accounts for eye lens transparency. *Nature* 302(5907), 415-17.

Depautex C., Desvignes C., Leboucher P., Lemonnier M., Dagneaux D., Benoit J.P. et Vachette P. (1987). The small angle X-ray scattering instrument D24. LURE, Orsay (France).

Derjaguin B.V. et Landau L. (1941). *Acta Physicochim. URSS* 14, 633-62.

Douzou P., Hui Bon Hoa G., Maurel P. et Travers F. (1976). Physical chemical data for mixed solvents used in low temperature biochemistry. Dans *Handbook of biochemistry and molecular biologie* 3 édit. (Fasman G.D., éd.), Vol. Physical and chemical data vol1, p. 520.

Dubuisson J.M., Decamps T. et Vachette P. (1997). Improved signal-to-background ratio in small angle X-ray scattering experiments with synchrotron radiation using an evacuated cell for solutions. *J. Appl. Cryst.* 30, 49-54.

Ducruix A. et Giegé R. (1992). *Crystallization of nucleic acids and proteins: a practical approach*. The practical approach series (Rickwood D. et Hames B.D., Eds.), IRL Press/Oxford University Press, Oxford.

Ducruix A., Guilloteau J.P., Riès-Kautt M. et Tardieu A. (1996). Protein interactions as seen by solution X-ray scattering prior to crystallogenesis. *J. Crystal Growth* 168, 28-39.

Dym O., Mevarech M. et Sussman J.L. (1995). Structural features that stabilize halophilic malate dehydrogenase from an *archaebacterium*. *Science* 267, 1344-46.

Ebel C. (1995). Characterisation of the solution structure of halophilic proteins. Analytical centrifugation among complementary techniques (light, neutron and X-ray scattering, density measurements). *Progr. Colloid Polym. Sci.* 99, 17-23.

Ebel C. (en préparation). Solvatation.

Ebel C., Faou P., Franzetti B., Kernel B., Madern D., Pascu M., Pfister C., Richard S. et Zaccai G. (1998). Molecular interactions in extreme halophiles. The solvation-stabilisation hypothesis for halophilic proteins. Dans *Microbiology and Biogeochemistry of Hypersaline Environments* (Oren A., éd.), p. 227-37. CRC press.

Ebel C., Faou P., Kernel B. et Zaccai G. (1999a). Relative role of anions and cations in the stabilisation of halophilic malate dehydrogenase. *Biochem.* 38(28), 9039-47.

Ebel C., Faou P. et Zaccai G. (1999b). Protein-solvent and weak protein-protein interactions in halophilic malate dehydrogenase. *J. Crysr. Growth* 196, 395-402.

Eisenberg H. (1976). *Biological macromolecules and polyelectrolytes in solution.* Monographs on physical biochemistry (Harrington W.F. et Peacocke A.R., Eds.), Clarendon Press, Oxford.

Eisenberg H. (1981). Forward scattering of light, X-rays and neutrons. *Q. Rev. Biophys.* 14(2), 141-72.

Eisenberg H., Mevarech M. et Zaccai G. (1992). Biochemical, structural, and molecular genetic aspects of halophilism. *Adv. Protein Chem.* 43, 1-61.

Finet S. (1999). Interactions entre protéines en solution : étude par diffusion des rayons X aux petits angles du lysozyme et des protéines du cristallin ; application à la cristallisation. Thèse de Doctorat, Paris VI, Paris.

George A. et Wilson W.W. (1994). Predicting protein crystallization from a dilute solution property. *Acta Cryst.* D50, 361-65.

Gerschel A. (1995). *Liaisons intermoléculaires : Les forces en jeu dans la matière condensée.* Savoirs Actuels, InterEditions / CNRS Editions, Paris.

Ghosh R.E., Egelhaaf S.U. et Rennie A.R. (1998). A computing guide for small-angle scattering experiments. Institut Laue Langevin, Grenoble, France. 98GH14T.

Gripon C., Legrand L., Rosenman I., Vidal O., Robert M.C. et Boué F. (1997). Lysozyme solubility in H2O and D2O solutions: a simple relationship. *J. Crystal Growth* 177, 238-47.

Guilloteau J.P. (1991). Cristallogenèse des macromolécules biologiques: influence de la nature des sels et de la température sur la solubilité du Lysozyme et sur les interactions protéines-protéines. Thèse de Doctorat, Louis Pasteur, Strasbourg, France.

Guinier A. et Fournet G. (1955). *Small-angle scattering of X-rays*. Structure of matter series, Wiley, New York.

Guo B., Kao S., McDonald H., Asanov A., Combs L.L. et Wilson W.W. (1999). Correlation of second virial coefficients and solubilities useful in protein crystal growth. *J. Crystal Growth* 196, 424-33.

Hamaker H.C. (1937). The London-van der Waals attraction between spherical particules. *Physica* 4(10), 1058-72.

Hamer W.J. (1975). Dielectric constant of water. Dans *Handbook of chemistry and physics* 56 édit. (Weast R.C., éd.), p. E61. CRC Press, Cleveland (USA).

Hansen J.P. et McDonald I.R. (1986). *Theory of simple liquids*. 2nd édit, Academic Press, London.

Harding S.E., Rowe A.J. et Horton J.C., Eds. (1992). Analytical ultracentrifugation in biochemistry and polymer science. Cambridge (UK): The Royal Society of Chemistry.

Hofmeister F. (1888). Zur lehre von der wirkung der salze. *Arch. Exp. Pathol. Pharmakol.* **24**, 247-60.

Israelachvili J.N. (1991). *Intermolecular and surface forces*. 2 édit, Academic Press, London.

Jacrot B. et Zaccai G. (1981). Determination of molecular weight by neutron scattering. *Biopolymers* 20, 2413-26.

Kernel B., Zaccai G. et Ebel C. (1999). Determination of partial molal volumes, and salt and water binding of highly charged biological macromolecules (tRNA, halophilic protein) in multimolar salt solutions. *Prog. Colloid Polym. Sci.* **113**, 168-75.

Kita Y., Arakawa T., Lin T.-Y. et Timasheff S.N. (1994). Contribution of surface free energy perturbation to protein-solvent interactions. *Biochem.* 33(50), 15178-89.

Lamm O. (1929). Die Differentialgleichung der Ultrazentrifugierung. *Ark. Mat. Astr. Fys* 21B(2), 1-4.

Lide D.R. et Frederikse H.P.R., Eds. (1996). CRC Handbook of chemistry and physics. 76 édit. Boca Raton: CRC Press.

Lifshitz E.M. (1956). *Soviet Phys. JETP* 2, 73-83.

Lindner P. (1998). Deadtime of the BF_3-detector at D11. Institut Laue Langevin, Grenoble (France). ILL98/LI 12 T. September 23.

Lindner P., May R.P. et Timmins P.A. (1992). Upgrading of the SANS instrument D11 at the ILL. *Physica B* 180 & 181(Part B), 967-72.

Lindner P. et Zemb T., Eds. (1991). Neutron, x-ray and light scattering : introduction to an investigative tool for colloidal and polymeric systems : proceedings of the European Workshop on Neutron, X-Ray and Light Scattering as an Investigative Tool for

Colloidal and Polymeric Systems, Bombannes, France, 27 May-2June, 1990. Delta series. Amsterdam: North-Holland.

Luzzati V. et Tardieu A. (1980). Recent developments in solution x-ray scattering. *Ann. Rev. Biophys. Bioeng.* 9, 1-29.

Madern D. (2000). Etude des relations structure-fonction, du repliement et de l'évolution au sein de la super-famille des malates et lactates déshydrogénases. Thèse de Doctorat, Université Joseph Fourier, Grenoble, France.

Madern D., Ebel C. et Zaccai G. (2000). Halophilic adaptation of enzymes. *Extremophiles* 4(2), 91-98.

Madern D., Pfister C. et Zaccai G. (1995). Mutation at a single acidic amino acid enhances the halophilic behaviour of malate dehydrogenase from *Haloarcula marismortui* in physiological salts. *Eur. J. Biochem.* 230(3), 1088-95.

Madern D. et Zaccai G. (1997). Stabilisation of halophilic malate dehydrogenase from Haloarcula marismortui by divalent cations -- effects of temperature, water isotope, cofactor and pH. *Eur. J. Biochem.* 249(2), 607-11.

Malfois M., Bonneté F., Belloni L. et Tardieu A. (1996). A model of attractive interactions to account for fluid-fluid phase separation of protein solutions. *J. Chem. Phys.* 105(8), 3290-300.

McMillan W.G. et Mayer J.E. (1945). The statistical thermodynamics of multicomponent systems. *J. Chem. Phys.* 13(7), 276-305.

Mevarech M., Eisenberg H. et Neumann E. (1977). Malate dehydrogenase isolated from extremely halophilic bacteria of the Dead Sea. 1. Purification and molecular characterization. *Biochem.* 16(17), 3781-85.

Mevarech M. et Neumann E. (1977). Malate dehydrogenase isolated from extremely halophilic bacteria of the Dead Sea. 2. Effect of salt on catalytic activity and structure. *Biochem.* 16, 3786-92.

Peyre V., Spalla O., Belloni L. et Nabavi M. (1997). Stability of a Nanometric Zirconia Colloidal Dispersion under Compression: Effect of Surface Complexation by Acetylacetone. *J. Colloid. Interface Sci.* 187(1), 184-200.

Pittz E.P. et Timasheff S.N. (1978). Interaction of ribonuclease A with aqueous 2-methyl-2,4-pentanediol at pH 5.8. *Biochem.* 17(4), 615-23.

Poon W., Pusey P. et Lekkerkerker H. (1996). Colloids in suspense. *Physics World* april, 27-32.

Richard S. (1998). Etude de la solvatation d'une protéine halophile, la malate déshydrogénase de *Haloarcula marismortui*, par cristallographie des rayons X et diffusion centrale. Thèse de Doctorat, Université Joseph Fourier, Grenoble, France.

Richard S., Bonneté F., Dym O. et Zaccai G. (1995). Protocol 21: The MPD-NaCl-H2O system for the crystallization of halophilic proteins. Dans *Archaea: a laboratory*

manual (Dassarma S., éd.), p. 149-53. Cold Spring Harbor Laboratory Press, New York.

Richard S.B., Madern D., Garcin E. et Zaccai G. (2000). Halophilic adaptation: novel solvent protein interactions observed in the 2.9 and 2.6 Å resolution structures of the wild type and a mutant of malate dehydrogenase from *Haloarcula marismortui*. *Biochem.* 39(5), 992-1000.

Riès-Kautt M. et Ducruix A. (1991). Crystallization of basic proteins by ion pairing. *J. Crystal Growth* 110, 20-25.

Riès-Kautt M. et Ducruix A. (1997). Inference drawn from physico-chemical studies of crystallogenesis and precrystalline state. *Methods Enzymol.* 276, 23-59.

Rothschild L.J. et Mancinelli R.L. (2001). Life in extreme environments. *Nature* 409(6823), 1092-101.

Scatchard G. (1946). Physical chemistry of protein solutions. I. Derivation of the equations for the osmotic pressure. *J. Am. Chem. Soc.* 68, 2315-19.

Schuck P. (1998). Sedimentation analysis of noninteracting and self-associating solutes using numerical solutions to the Lamm equation. *Biophys. J.* 75(3), 1503-12.

Schuck P. et Demeler B. (1999). Direct sedimentation analysis of interference optical data in analytical ultracentrifugation. *Biophys. J.* 76(4), 2288-96.

Solovyova A., Schuck P., Costenaro L. et Ebel C. (2001). Non-ideality by sedimentation velocity of halophilic malate dehydrogenase in complex solvents. *Biophys. J.* 81(4), 1868-80.

Svedberg T. et Pedersen K.O. (1940). *The ultracentrifuge*, The Clarendon Press, Oxford (UK).

Svergun D.I., Richard S., Koch M.H.J., Sayers Z., Kuprin S. et Zaccai G. (1998). Protein hydration in solution: experimental observation by X-ray and neutron scattering. *Proc. Natl. Acad. Sci. USA* 95(5), 2267-72.

Tanford C. (1961). *Physical Chemistry of Macromolecules*, Wiley, New York.

Tardieu A., Le Verge A., Malfois M., Bonneté F., Finet S., Riès-Kautt M. et Belloni L. (1999). Proteins in solution: from X-ray scattering intensities to interaction potentials. *J. Crystal Growth* 196, 193-203.

Tardieu A., Vachette P., Gulik A. et Le Maire M. (1981). Biological macromolecules in solvents of variable density: characterization by sedimentation equilibrium, densimetry, and X-ray forward scattering and an application to the 50S ribosomal subunit from Escherichia coli. *Biochem.* 20(15), 4399-406.

Tardieu A., Vérétout F., Krop B. et Slingsby C. (1992). Protein interactions in the calf eye lens: interactions between β-crystallins are repulsive whereas in γ-crystallins they are attractive. *Eur. Biophys. J.* 21(1), 1-12.

Timasheff S.N. (1993). The control of protein stability and association by weak interactions with water: how do solvents affect these processes? *Annu. Rev. Biophys. Biomol. Struct.* 22, 67-97.

Timasheff S.N. et Arakawa T. (1988). Mechanism of protein precipitation and stabilization by co-solvents. *J. Crystal Growth* 90, 39-46.

Travers F. et Douzou P. (1974). Dielectric constant of mixed solvents used for a low temperature biochemistry. *Biochimie* 56(4), 509-14.

Vérétout F., Delaye M. et Tardieu A. (1989). Molecular basis of eye lens transparency. Osmotic pressure and X-ray analysis of alpha-crystallin solutions. *J. Mol. Biol.* 205(4), 713-28.

Verwey E.J.W. et Overbeek J.T.G. (1948). *Theory of stability of lyophobic colloids*, Elsevier, Amsterdam.

von Hippel P.H. et Schleich T. (1969). The effects of neutral salts on the structure and conformational stability of macromolecules in solution. Dans *Structure and stability of biological macromolecules* (Timasheff S.N. et Fasman G.D., éds.), p. 417-574. Dekker, New York.

Woese C.R. et Fox G.E. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. USA* 74(11), 5088-90.

Zaitsev I.D. et Aseev G.G., Eds. (1992). Properties of aqueous solutions of electrolytes. Boca Raton: CRC Press.

Résumé

La cellule est un milieu très concentré où les interactions entre macromolécules, même faibles, jouent un grand rôle dans leur solubilité et dans l'assemblage des complexes labiles assurant les fonctions biologiques. Les interactions faibles entre protéines déterminant la non-idéalité de leurs solutions vont aussi gouverner leur cristallisation.

Dans quelle mesure les interactions protéine – solvant influencent-elles les interactions protéine – protéine ? Nous avons mis en relation ces deux types d'interactions pour la malate déshydrogénase (*Hm* MalDH) de *Haloarcula marismortui*, protéine halophile très acide qui a des solvatations variées et très riches en eau et en sel.

Nous avons développé une nouvelle méthode de détermination du second coefficient du viriel A_2 par la modélisation des profils de vitesse de sédimentation en ultracentrifugation analytique, qui permet l'étude de solvants complexes.

Les interactions protéine – protéine de la *Hm* MalDH en divers sels ont été caractérisées par diffusion de neutrons ou de rayons X aux petits angles. Les A_2 et les facteurs de structure en solution ont été modélisés par des potentiels d'interaction de type DLVO. Les interactions répulsives sont principalement dues au terme de volume exclu et dans une moindre mesure au terme électrostatique. Les interactions attractives sont qualitativement corrélées à des valeurs positives ou négatives des paramètres d'interaction préférentielle avec le sel. Ces résultats permettent d'expliquer l'adaptation moléculaire des protéines halophiles qui doivent ainsi avoir une solvatation riche en sel pour rester soluble à haut sel.

La cristallisation par dilution de la *Hm* MalDH dans des mélanges sel – MPD (méthyl-2-pentanediol-2,4) résulte d'une lente évolution des interactions protéine – protéine, de répulsives à modérément attractives. Le MPD modifie les interactions protéine – protéine en divers sels en ajoutant une attraction qui est liée à la répulsion du MPD par les charges de la protéine.

Mots Clés : interactions protéine – protéine, solvatation, malate déshydrogénase, halophile, second coefficient du viriel, facteur de structure, MPD, cristallisation.

Abstract

Weak protein – protein interactions in solution : The halophilic malate dehydrogenase.

The cell is a crowded environment in which interactions between macromolecules, even weak interactions, determine their solubility and the assembly of the labile complexes that operate biological functions. The weak interactions between proteins in solution that define their non-ideal behaviour also control their crystallisation.

To which extent do protein – solvent interactions influence protein – protein interactions ? In the present thesis, these two kinds of interaction have been put in relation for the malate dehydrogenase (*Hm* MalDH) from *Haloarcula marismortui*. This halophilic protein is very acidic and is solvated by large amounts of water and salt.

We have developed a new method for the determination of the second virial coefficient A_2 by modelling the sedimentation velocity profiles in analytical ultracentrifugation. It allows to study complex solvent compositions with small amounts of protein.

The protein – protein interactions of *Hm* MalDH in various salt solvents were characterised by small-angle scattering of neutrons or X-rays. The A_2 and the solution structure factors were analysed with DLVO-like interaction potentials. The repulsive interactions arise mainly from the excluded-volume term and to a lesser extent from the electrostatic term. The attractive interactions appear to be qualitatively correlated to positive or negative values of the preferential salt biding parameters. These results also allow to explain the molecular adaptation of halophilic proteins that require a solvation rich in salt in order to remain soluble.

The crystallisation by dilution of *Hm* MalDH in various mixture of salt – MPD (2-methy-2,4-pentanediol) is correlated with a slow evolution from repulsive to moderately attractive protein – protein interactions. MPD modifies the protein – protein interactions by adding an attractive term that arises from the repulsion of MPD by the charges on the protein.

Keywords : protein – protein interaction, solvation, malate dehydrogenase, halophilic, second virial coefficient, structure factor, MPD, crystallisation.