

## Diffusion de neutrons aux petits angles : application à l'étude des macromolécules biologiques en solution

D. Lairez<sup>1</sup> et J. Pelta<sup>2</sup>

<sup>1</sup> *Laboratoire Léon Brillouin, CEA Saclay, 91191 Gif-sur-Yvette Cedex, France*  
*e-mail : lairez@cea.fr*

<sup>2</sup> *Groupe Microenvironnements et Comportements Cellulaires, Université de Cergy-Pontoise, 95302 Cergy-Pontoise, France*  
*e-mail : juan.pelta@bio.u-cergy.fr*

**Résumé.** La diffusion de neutrons aux petits angles (DNPA) permet de sonder les propriétés statiques de la matière sur une échelle spatiale allant de quelques dixièmes à quelques dizaines de nanomètres. Cette technique est ainsi particulièrement bien adaptée à l'étude des macromolécules en solution. Il est possible d'accéder à des grandeurs moyennes qui caractérisent la conformation qu'adoptent les macromolécules ou leurs interactions thermodynamiques. Le cours expose à des non spécialistes les différentes grandeurs mesurables et les méthodes à utiliser pour y accéder. En particulier sont abordées: 1) les mesures effectuées dans la limite du vecteur de diffusion nul qui sont liées aux fluctuations de concentration; 2) les notions de facteur de forme, de facteur de structure ... 3) les différentes façon de jouer avec le contraste. Les notions introduites sont illustrées par des exemples didactiques empruntés à la littérature et concernant des macromolécules biologiques en solution.

### Table des matières

<b>Introduction</b>	<b>40</b>
<b>1 Les principes de base</b>	<b>40</b>
1.1 Définitions et notations . . . . .	40
1.2 Diffusion par des objets non ponctuels, facteur de forme . . . . .	41
1.3 Facteur de forme et facteur de structure . . . . .	42
1.4 Diffusion par un milieu homogène . . . . .	43
1.5 Diffusion et fluctuations . . . . .	43
1.6 Espace direct et espace réciproque . . . . .	44
1.7 Diffusion cohérente et incohérente . . . . .	45
1.8 Mélange solvant-soluté, notion de contraste, principe de Babinet . . . . .	46
1.9 Exemple 1 : Compressibilité de la myoglobine . . . . .	47
1.10 Exemple 2 : Interactions, second viriel et cristallisation des protéines . . . . .	48
<b>2 Facteur de forme</b>	<b>49</b>
2.1 Objets ayant des orientations aléatoires . . . . .	49
2.2 Domaine de Guinier, rayon de giration . . . . .	49
2.3 Objets à symétrie sphérique . . . . .	50
2.4 Bâton rigide . . . . .	50
2.5 Chaîne gaussienne . . . . .	51
2.6 Objets autosimilaires : argument d'échelle . . . . .	51
2.7 Diffusion par une interface, loi de Porod . . . . .	52
2.8 Fonction de coupure à grand $q$ . . . . .	53

2.9	Facteurs de formes divers . . . . .	53
2.10	Exemple 3 : Facteurs d'élongation de la traduction . . . . .	54
2.11	Exemple 4 : Repliement d'une protéine non globulaire multidomaine . . . . .	55
2.12	Exemple 5 : Superhélice d'ADN . . . . .	56
<b>3</b>	<b>Variations sur le contraste</b>	<b>56</b>
3.1	Mélange de deux espèces identiques, marquage isotopique . . . . .	57
3.2	Mélange de trois espèces, contraste moyen nul . . . . .	58
3.3	Mélange de trois espèces, variation de contraste . . . . .	58
3.4	Marquage sélectif . . . . .	59
3.5	Variation de contraste par polarisation nucléaire . . . . .	59
3.6	Exemple 6 : Étude d'un chaperon moléculaire, GroEL . . . . .	59
3.7	Exemple 7 : Triangulation et structure quaternaire des sous-unités ribosomales . . . . .	60

## Introduction

Les techniques de diffusion élastique de rayonnement permettent d'accéder aux propriétés statiques de la matière, la façon dont les objets qui la constituent occupent l'espace, s'organisent et interagissent. Le terme de diffusion concerne des techniques sondant ces propriétés à une échelle plus grande que celle des distances interatomiques qui sont le domaine de la diffraction et de la cristallographie. Loin de la résolution atomique, la diffusion de neutrons aux petits angles (DNPA) permet par exemple l'étude de la conformation globale d'objets de grande taille. Contrairement à la microscopie, les informations obtenues sont des grandeurs statistiques moyennes obtenues sur l'ensemble d'un échantillon de taille macroscopique. Ce texte est une introduction à la DNPA appliquée à l'étude des macromolécules biologiques en solution et non une revue exhaustive et détaillée de ces études. Dans la première partie le vocabulaire et les notions de base sont introduites, nous montrerons en particulier l'importance de la notion de fluctuation. La deuxième partie est consacrée à la notion de facteur de forme d'un objet. La troisième partie concerne un aspect spécifique à la DNPA qui est lié à l'origine nucléaire de l'interaction des neutrons avec les atomes et qui permet de jouer avec la notion de contraste. Les différentes notions introduites sont illustrées par des exemples empruntés à la littérature sélectionnés avant tout pour des raisons didactiques et liées à la structure du cours.

## 1. LES PRINCIPES DE BASE

Dans son principe, une expérience de diffusion de neutrons aux petits angles est similaire aux autres expériences de diffusion de rayonnement (rayon X, lumière).

### 1.1 Définitions et notations

Lors d'une expérience de diffusion de rayonnement, l'échantillon étudié est placé sur le trajet d'une onde plane, monochromatique de longueur d'onde  $\lambda$ , dont l'équation de propagation s'écrit  $\psi(x, t) = \psi_0 \cdot e^{i(\omega t - k_0 x)}$ , où  $\omega$  est la pulsation et  $k_0 = 2\pi/\lambda$  la norme du vecteur d'onde. Cette onde peut être vue comme un ensemble de particules (neutrons) se déplaçant à la même vitesse et dans des directions parallèles. Lorsque cette onde interagit avec un atome de l'échantillon, elle est diffusée dans toutes les directions. L'onde diffusée est sphérique. On supposera une **diffusion élastique**, c'est à dire que les ondes incidente et diffusée ont la même longueur d'onde. La probabilité qu'une particule soit diffusée est proportionnelle à l'aire  $\sigma$  d'une surface caractéristique de l'interaction entre l'atome et le rayonnement considéré. Cette aire est la **section efficace de diffusion**, elle correspond à la surface de l'atome "vue"

par ce rayonnement. L'intensité de l'onde diffusée sur  $4\pi$  stéradians (nombre total de particules diffusées par unité de temps) est  $I_{\text{tot}} = \Phi\sigma$ , où  $\Phi$  est le flux de particules incidentes. La section efficace peut s'écrire sous la forme  $\sigma = b^2$ , où  $b$  appelé **longueur de diffusion** de l'atome caractérise la portée des interactions entre les particules et l'atome. Pour un détecteur d'aire  $s$  situé à une distance  $D$  de l'atome diffuseur, l'intensité  $I$  recueillie est proportionnelle au rapport des aires  $\Omega = s/D^2$ , soit l'angle solide formé par le détecteur vu depuis le centre diffuseur:  $I = \Phi\sigma\Omega = \Phi(b/D)^2 s$ . Cette intensité est le carré de l'onde qui s'écrit:

$$\psi = \psi_0 \frac{b}{D} e^{i(\omega t - kD)} \quad (1.1)$$

Dans le cas de la diffusion de lumière ou des rayons X, la longueur de diffusion est une propriété du cortège électronique des atomes. Par contre dans le cas des neutrons, l'interaction est nucléaire. La longueur de diffusion dépend de l'isotope que l'on considère ce qui permet d'utiliser des techniques de marquage (cf. partie 3).

Les ondes diffusées par plusieurs atomes interfèrent. Le résultat de ces interférences sur le détecteur dépend des différences de phase des ondes diffusées. Par rapport à une origine  $O$ , l'onde de vecteur  $\vec{k}_1$  diffusée par un atome situé en  $M$  tel que  $\vec{OM} = \vec{r}$  est déphasée de  $\vec{r}(\vec{k}_1 - \vec{k}_0) = \vec{r} \cdot \vec{q}$ . Le vecteur  $\vec{q} = \vec{k}_1 - \vec{k}_0$  est le **vecteur de diffusion**, sa norme est  $q = \sin(\theta/2)4\pi/\lambda$ , où  $\theta$  est l'angle de diffusion. Si  $\psi(0, t)$  est l'onde diffusée par l'origine, l'onde diffusée par l'atome en  $M$  est

$$\psi(r, t) = \psi(0, t) \cdot e^{i\vec{q} \cdot \vec{r}} \quad (1.2)$$

Le vecteur de diffusion permet de superposer des figures d'interférence obtenues à des angles ou à des longueurs d'onde différents. Sa norme est homogène à l'inverse d'une longueur. C'est à cette longueur à laquelle il est fait référence lorsque l'on parle d'échelle d'observation.

L'onde  $\psi$  diffusée par les  $n$  atomes de l'échantillon est la somme des ondes  $\psi_i$  diffusées par chaque atome. Si le détecteur est situé à une distance beaucoup plus grande que la taille de l'échantillon, les  $n$  atomes sont vus depuis ce détecteur sous le même angle de diffusion. On a:

$$\psi = \frac{\psi_0}{D} \times \sum_{i=1}^n b_i e^{i\vec{q} \cdot \vec{r}_i} \quad (1.3)$$

L'intensité diffusée est  $I = \psi^2 = \psi\psi^* = (\psi_0^2/D^2) \times \sum_i^n \sum_j^n b_i b_j e^{i\vec{q} \cdot (\vec{r}_i - \vec{r}_j)}$ . Pour s'affranchir des dépendances triviales en  $\psi_0^2$  et  $D^2$ , on définit la **section efficace différentielle de diffusion**:

$$\mathcal{S}(q) = \sum_i^n \sum_j^n b_i b_j e^{i\vec{q} \cdot (\vec{r}_i - \vec{r}_j)} \quad (1.4)$$

Lorsque les atomes sont groupés en molécules de petite taille comparée à  $q^{-1}$ , les ondes diffusées par les différents atomes d'une molécule sont en phase et leurs amplitudes s'ajoutent, la molécule est vue comme un point. Dans l'équation 1.4,  $\forall(i, j), \vec{q} \cdot (\vec{r}_i - \vec{r}_j) \rightarrow 0$  et  $\mathcal{S}(q) = (\sum_i^n b_i)^2$ . Une telle molécule se comporte donc comme un **diffuseur élémentaire** dont la longueur de diffusion est la somme des longueurs de diffusion des atomes qui la constituent. Par la suite, nous considérerons les interférences entre les ondes diffusées par ces diffuseurs élémentaires. Dans la pratique ce sont les petites molécules constituant un liquide ou les monomères d'une macromolécule.

## 1.2 Diffusion par des objets non ponctuels, facteur de forme

Soit un ensemble de  $\mathcal{N}$  diffuseurs élémentaires de longueur de diffusion  $b$ , regroupés en  $n$  objets identiques de  $N$  diffuseurs chacun. Dans l'équation 1.4, on peut séparer les termes concernant les

interférences entre diffuseurs appartenant au même objet et ceux concernant les interférences entre diffuseurs d'objet différents:

$$\mathcal{S}(q) = \sum_i^N \sum_j^N b_i b_j e^{i\vec{q} \cdot (\vec{r}_i - \vec{r}_j)} = b^2 \sum_h^n \sum_k^n \sum_i^N \sum_j^N e^{i\vec{q} \cdot (\vec{r}_{ih} - \vec{r}_{jk})} \quad (1.5)$$

Cette expression comporte  $(nN)^2$  termes dont  $nN^2$  tels que  $h = k$  et  $n(n-1)N^2 \approx n^2N^2$  en moyenne identiques tels que  $h \neq k$ :  $\mathcal{S}(q)/b^2 = n \sum_{i_h}^N \sum_{j_h}^N e^{i\vec{q} \cdot (\vec{r}_{ih} - \vec{r}_{jh})} + n^2 \sum_{i_h}^N \sum_{j_k}^N e^{i\vec{q} \cdot (\vec{r}_{ih} - \vec{r}_{jk})}$ . On utilise plutôt des grandeurs normalisées:

$$\frac{\mathcal{S}(q)}{b^2 \times nN^2} = P(q) + nQ(q) \quad \text{avec} \quad \begin{cases} P(q) = \frac{1}{N^2} \sum_{i_h}^N \sum_{j_h}^N e^{i\vec{q} \cdot (\vec{r}_{ih} - \vec{r}_{jh})} \\ Q(q) = \frac{1}{N^2} \sum_{i_h}^N \sum_{j_k}^N e^{i\vec{q} \cdot (\vec{r}_{ih} - \vec{r}_{jk})} \end{cases} \quad (1.6)$$

La grandeur  $Q(q)$  reflète les interférences entre les ondes diffusées par des atomes appartenant à des objets différents (terme "interobjet"), tandis que  $P(q)$  concerne les ondes diffusées par les atomes d'un même objet (terme "intraobjet"). Ce dernier est celui auquel on accède lorsque  $n \rightarrow 0$ , c'est à dire pour un système très dilué.

$$\left. \frac{\mathcal{S}(q)}{b^2 \times nN^2} \right]_{n \rightarrow 0} = P(q) \quad (1.7)$$

$P(q)$  est le **facteur de forme** des objets et contient l'information concernant les corrélations entre les positions des diffuseurs élémentaires d'un même objet. Dans beaucoup de cas, ce facteur de forme dépend des interactions avec les objets voisins. Par exemple une chaîne polymère seule en bon solvant adopte une conformation "gonflée" due aux interactions répulsives de volume exclu qu'exercent les monomères entre eux. Par contre, en présence d'autres chaînes ces interactions sont écrantées et la chaîne adopte une conformation aléatoire Gaussienne. On comprend alors que diluer le système pour mesurer  $P(q)$  n'a pas toujours de sens.

L'équation 1.7 peut s'écrire sous la forme  $[\mathcal{S}(q)/b^2]_{n \rightarrow 0} = n(N^2 P(q))$ . Cela signifie que: 1) Lorsqu'il n'y a pas de corrélation entre objets, l'intensité diffusée est la somme des  $n$  contributions individuelles de chacun d'eux. Les objets diffusent de façon **incohérente**. 2) La contribution d'un objet est telle que ce sont les amplitudes des ondes diffusées par les diffuseurs élémentaires d'un objet qui s'ajoutent (d'où le facteur  $N^2$ ) et non pas les intensités. Les diffuseurs élémentaires d'un même objet diffusent de façon **cohérente**.

### 1.3 Facteur de forme et facteur de structure

La section efficace différentielle de diffusion peut s'écrire en faisant intervenir les centres de gravité  $h$  et  $k$  des objets:  $\vec{i}\vec{j} = (\vec{i}\vec{h} + \vec{k}\vec{j}) + \vec{h}\vec{k}$ . Considérons des objets de forme quelconque ayant tous la même orientation, ou des objets sphériques. Alors pour chaque point  $j$  d'un objet de centre de gravité  $k$ , il existe un point  $j'$  d'un objet de centre de gravité  $h$  tel que  $\vec{k}\vec{j} = \vec{h}\vec{j}'$ , soit  $\vec{i}\vec{h} + \vec{k}\vec{j} = \vec{i}\vec{j}'$  qui est indépendant de  $h$  et  $k$ . L'équation 1.5 est donc finalement factorisable sous la forme:  $\mathcal{S}(q) = b^2 \sum_h^n \sum_k^n e^{-i\vec{q} \cdot \vec{h}\vec{k}} \times \sum_i^N \sum_{j'}^N e^{-i\vec{q} \cdot \vec{i}\vec{j}'}$ , soit:

$$\frac{\mathcal{S}(q)}{b^2 \times nN^2} = P(q) \times S(q) \quad \text{avec} \quad \begin{cases} P(q) = \frac{1}{N^2} \sum_i^N \sum_{j'}^N e^{-i\vec{q} \cdot \vec{i}\vec{j}'} \\ S(q) = \frac{1}{n} \sum_h^n \sum_k^n e^{-i\vec{q} \cdot \vec{h}\vec{k}} \end{cases} \quad (1.8)$$

Expression dans laquelle on retrouve le facteur de forme  $P(q)$ . La grandeur  $S(q)$  est le **facteur de structure** et contient l'information sur les corrélations entre les positions des centres de gravité des objets.

Réécrivons le facteur de structure en isolant le terme  $h = k$  pour lequel  $e^{-i\vec{q}\cdot\vec{h}\vec{k}} = 1$ , on a  $S(q) = \frac{1}{n} \sum_h^n \left( 1 + \sum_{h \neq k}^n e^{-i\vec{q}\cdot\vec{h}\vec{k}} \right) = 1 + \frac{1}{n} \sum_h^n \sum_{h \neq k}^n e^{-i\vec{q}\cdot\vec{h}\vec{k}}$ , d'où:

$$S(q \rightarrow \infty) = 1 \quad (1.9)$$

Dans la pratique cette limite est atteinte pour des valeurs de  $q$  suffisamment grandes comparées à l'inverse de la taille caractéristique,  $R$ , des objets (cf. exemple 1.10). Il s'ensuit que dans ce domaine de vecteur de diffusion, la mesure n'est sensible qu'au facteur de forme:

$$\frac{\mathcal{S}(q)}{b^2 \times nN^2} \Big|_{qR \gg 1} = P(q) \quad (1.10)$$

Au contraire pour la limite  $q \rightarrow 0$ ,  $P(q \rightarrow 0) = 1$ , la section efficace différentielle s'écrit alors

$$\mathcal{S}(q \rightarrow 0) = (bN)^2 \times nS(q) \quad (1.11)$$

ce qui correspond à la diffusion par  $n$  objets ponctuels de longueur de diffusion  $bN$  chacun.

#### 1.4 Diffusion par un milieu homogène

On considère la longueur de diffusion  $b$  d'un élément de volume  $v$  d'un échantillon. Le rapport  $\rho = b/v$  est appelé **densité de longueur de diffusion**. Un milieu est parfaitement homogène si cette densité est constante dans tout l'échantillon et quelle que soit la dimension de l'élément de volume considéré jusqu'au volume atomique. L'échantillon est un cristal parfait. Dans ce cas, il existe des directions privilégiées, correspondant à une famille de vecteurs de diffusion  $\vec{q}^*$ , pour lesquelles les ondes diffusées par deux plans réticulaires voisins distants de  $d_{ij}$  sont en phase :  $\vec{q}^* \cdot \vec{d}_{ij} = n \cdot 2\pi$ . C'est la loi de Bragg. Elle donne les directions des maximums d'interférence. Du fait de la périodicité du cristal et si celui-ci a une taille infinie, les ondes diffusées par une infinité de plans réticulaires sont en phase et contribuent de façon constructive à l'intensité mesurée pour ces vecteurs de diffusion. Pour une valeur de  $q$  très proche mais différente d'une des valeurs particulières  $q^*$ , il existe une petite différence de phase entre les ondes diffusées par deux plans voisins. Différence qui augmente de proche en proche pour des plans de plus en plus éloignés. Finalement il existe un plan réticulaire dont l'onde diffusée est en opposition de phase avec celle diffusée par le premier. L'onde résultante s'annule. Par contre, si le cristal a une taille finie, sa dimension  $L$  limitera la possibilité précédente. C'est ce qui conditionne la largeur  $\Delta q$  d'un pic de Bragg pour un cristal parfait. Si pour  $qL = n2\pi$  les ondes diffusées sont en phase (maximum d'interférence), alors pour  $(q + \Delta q)L = (n + \frac{1}{2}) 2\pi$  elles sont en opposition de phase (intensité nulle), d'où la largeur du pic  $\Delta q = \pi/L$ . L'intensité mesurée en fonction du vecteur de diffusion se résume donc à des pics dont la largeur est inversement proportionnelle à la dimension du cristal. C'est le domaine de la diffraction. Entre ces pics l'intensité est nulle. C'est le domaine de la diffusion.

**Un milieu homogène ne diffuse pas.**

Il diffracte à  $q = 0$ , mais cette diffraction n'est pas mesurable car elle se confond avec le faisceau transmis à travers l'échantillon. Il diffracte à grand vecteur de diffusion ( $q^{-1}$  de l'ordre de grandeur des distances inter atomiques), mais à cette échelle il n'est plus homogène.

#### 1.5 Diffusion et fluctuations

Soit un gaz de  $n$  diffuseurs de longueur de diffusion  $b$ . Le volume  $V$  est divisé en  $V/v$  volumes élémentaires correspondant à celui d'un diffuseur. La fraction volumique occupée par les diffuseurs est  $x = nv/V$ . A un

instant donné, certains éléments de volume sont vides ( $b_i = 0$ ). Mais du fait de l'agitation thermique, ils sont susceptibles de contenir un diffuseur à un autre instant ( $b_i = b$ ) durant la mesure. La section efficace différentielle est une moyenne sur tous ces états microscopiques possibles par lesquels passe l'échantillon. Chaque élément de volume étant susceptible de diffuser à un moment donné, les interférences sont calculées en effectuant la double somme de l'Eq. 1.4 sur les  $V/v$  éléments de volume:

$$\mathcal{S}(q) = \left\langle \sum_i^{V/v} \sum_j^{V/v} b_i b_j e^{i\vec{q} \cdot (\vec{r}_i - \vec{r}_j)} \right\rangle = \sum_i^{V/v} \sum_j^{V/v} \langle b_i b_j \rangle \times e^{i\vec{q} \cdot (\vec{r}_i - \vec{r}_j)} \quad (1.12)$$

En moyenne la longueur de diffusion d'un élément de volume est  $\langle b \rangle = bx$ . Pour un élément de volume donné, l'écart à cette moyenne est  $\Delta b_i = b_i - \langle b \rangle$ . Par définition  $\Delta b_i$  est nul en moyenne. Le produit  $\langle b_i b_j \rangle = \langle (\Delta b_i + \langle b \rangle) \cdot (\Delta b_j + \langle b \rangle) \rangle$  est donc égal à

$$\langle b_i b_j \rangle = \langle \Delta b_i \Delta b_j \rangle + \langle b \rangle^2 \quad (1.13)$$

Introduit dans l'Eq. 1.12 le second terme de cette somme correspond à la diffusion par un milieu homogène constitué de diffuseurs de longueur de diffusion  $\langle b \rangle$  uniformément répartis. Ce terme est nul et on obtient une expression très générale:

$$\mathcal{S}(q) = \sum_i^{V/v} \sum_j^{V/v} \langle \Delta b_i \Delta b_j \rangle e^{i\vec{q} \cdot (\vec{r}_i - \vec{r}_j)} \quad (1.14)$$

**Seules les fluctuations, spatiales ou temporelles, de longueur de diffusion contribuent à l'intensité diffusée.**

$\langle \Delta b_i \Delta b_j \rangle / b^2 = \langle \Delta x_i \Delta x_j \rangle$  est la **fonction de corrélation de paire** des fluctuations de densité du gaz. Pour un gaz parfait il n'existe aucune corrélation entre la densité d'un élément de volume et celle d'un élément voisin. La longueur de corrélation des fluctuations correspond à la taille d'un diffuseur élémentaire. Pour un élément de volume  $i$  donné, le produit  $\Delta x_i \Delta x_j$  est nul en moyenne excepté pour  $i = j$ . La double somme de l'Eq. 1.14 se résume donc à  $V/v$  termes identiques indépendants de  $q$ :

$$\mathcal{S}(q) = b^2 \frac{V}{v} \langle \Delta x^2 \rangle \quad (1.15)$$

Les fluctuations de la densité de l'élément de volume  $v$  sont liées à la compressibilité isotherme  $\kappa_T = (1/V \cdot dP/d(1/V))^{-1} = v/kT \times \langle \Delta x^2 \rangle / \langle x \rangle^2$ , où  $kT$  est l'énergie thermique. Les diffuseurs élémentaires sont ici des points, la section efficace de diffusion ainsi calculée est donc une grandeur macroscopique qui correspond à la limite  $q \rightarrow 0$ . Finalement on obtient un résultat très général car il s'applique également aux solutions:

$$\frac{\mathcal{S}(q \rightarrow 0)}{V} = b^2 \left( \frac{n}{V} \right)^2 kT \kappa_T \quad (1.16)$$

Pour un gaz parfait, qui correspond à la limite  $n/V \rightarrow 0$ , on a  $\kappa_T = (kTn/V)^{-1}$ , d'où

$$\left. \frac{\mathcal{S}(q \rightarrow 0)/V}{n/V} \right]_{n/V \rightarrow 0} = b^2 \quad (1.17)$$

qui exprime que les intensités diffusées par les  $n$  diffuseurs s'ajoutent. L'expression 1.16 est à comparer à l'équation 1.11, on obtient pour le facteur de structure:  $S(q \rightarrow 0) = (n/V)kT\kappa_T$ .

## 1.6 Espace direct et espace réciproque

Pour des raisons calculatoires, un formalisme de milieu continu est souvent préférable. Pour l'exprimer, on utilise la fraction volumique locale (ou densité locale) en diffuseurs élémentaires:  $x(\vec{r}) = \sum_i^n \Delta(\vec{r} - \vec{r}_i)$ ,

où  $\Delta$  est la fonction de Dirac telle que  $\Delta(0) = 1$  et  $\Delta(\vec{r} \neq 0) = 0$ . L'équation 1.14 devient:

$$\mathcal{S}(q) = b^2 \int_V dr_2 \int_V \langle \Delta x(\vec{r}_1) \Delta x(\vec{r}_2) \rangle e^{i\vec{q} \cdot (\vec{r}_1 - \vec{r}_2)} dr_1 \quad (1.18)$$

Si le milieu est isotrope, la fonction de corrélation de paire  $\langle \Delta x(\vec{r}_1) \Delta x(\vec{r}_2) \rangle$  ne dépend que de la norme de  $\vec{r} = \vec{r}_2 - \vec{r}_1$ , soit:  $\langle \Delta x(\vec{r}_1) \Delta x(\vec{r}_2) \rangle = \langle \Delta x(0) \Delta x(r) \rangle = g(r)$ . Ce changement de variable nous donne:  $\mathcal{S}(q) = b^2 \int_V dr \int_V g(r) e^{-i\vec{q} \cdot \vec{r}} dr$ , soit:

$$\mathcal{S}(q) = b^2 V \int_V g(r) e^{-i\vec{q} \cdot \vec{r}} dr \quad (1.19)$$

Ce qui exprime que l'on mesure la transformée de Fourier de  $g(r)$ . Cette fonction de corrélation de paire représente la probabilité de trouver un diffuseur élémentaire à une distance  $r$  d'une origine choisie au hasard. Souvent on utilise la **fonction de distribution des distances** ou **fonction de distribution radiale** définie par  $p(r) = 4\pi r^2 \times g(r)$ .

Extraire de la mesure la fonction de corrélation  $g(r)$ , c'est à dire passer d'une représentation dans l'espace réciproque à une représentation dans l'espace réel, est numériquement assez lourd. Il ne suffit pas d'effectuer une transformée de Fourier inverse pour deux raisons. D'une part on mesure rarement la totalité du spectre d'intensité diffusée et d'autre part il est absolument nécessaire de tenir compte de la résolution instrumentale. En DNPA la résolution est généralement une fonction du vecteur de diffusion [1] car le monochromateur est tel que  $\Delta\lambda/\lambda$  est constant, tandis que la géométrie du détecteur et de la collimation sont telles que  $\Delta\theta$  est constant. Tenir compte de cette fonction d'appareil n'est donc pas une simple opération de déconvolution. La méthode la plus employée pour effectuer ces opérations est dite "transformée de Fourier indirecte" [2]. Elle consiste à 1) paramétriser  $g(r)$ ; 2) calculer la section efficace théorique de diffusion en effectuant numériquement la transformée de Fourier de  $g(r)$  pour des valeurs initiales des paramètres; 3) convoluer par la fonction d'appareil; 4) comparer le résultat du calcul à la mesure et optimiser les paramètres afin de minimiser  $\chi^2$ .

### 1.7 Diffusion cohérente et incohérente

Implicitement nous avons considéré jusqu'ici un ensemble d'atomes de longueur de diffusion identique, dans la réalité c'est rarement le cas. Par exemple, le spin 1/2 des neutrons implique que même lorsque les atomes de l'échantillon sont tous identiques mais ont eux-même un spin nucléaire non nul, leur longueur de diffusion dépend de leur état de spin. Ainsi les neutrons voient un mélange de diffuseurs ayant des longueurs de diffusion différentes et dont les positions dans l'espace fluctuent et ne sont pas corrélées. Ces fluctuations de densité de longueur de diffusion contribuent à l'intensité diffusée. La section efficace différentielle de diffusion  $\mathcal{S}(q) = \sum_i^n \sum_j^n \langle b_i b_j \rangle e^{iq(r_i - r_j)}$  s'écrit de façon identique à celle d'un gaz:

$$\mathcal{S}(q) = n (\langle b^2 \rangle - \langle b \rangle^2) + \langle b \rangle^2 \sum_i^n \sum_j^n e^{iq(r_i - r_j)} \quad (1.20)$$

La grandeur  $b_{coh}^2 = \langle b \rangle^2$  est appelée section efficace de diffusion cohérente et  $b_{inc}^2 = \langle b^2 \rangle - \langle b \rangle^2$  section efficace de diffusion incohérente. Cette dernière contribution à l'intensité diffusée est indépendante de  $q$ . On appelle "diffusion incohérente" toutes les contributions à l'intensité diffusée dues à des fluctuations de longueur de corrélation très courte comparée aux longueurs  $q^{-1}$  de la DNPA. C'est par exemple les fluctuations de composition isotopique (un ensemble d'atomes d'une seule espèce chimique peut être un mélange d'isotopes de longueurs de diffusion différentes), ou les fluctuations de densité du solvant (liées à sa compressibilité) dans le cas d'un mélange soluté-solvant. La soustraction de l'ensemble de ces contributions conduit à la **section efficace différentielle de diffusion cohérente**:  $\mathcal{S}_{coh}(q) = \langle b \rangle^2 \sum_i^n \sum_j^n e^{iq(r_i - r_j)}$  C'est uniquement de cette section efficace qu'il s'agira par la suite.

### 1.8 Mélange solvant-soluté, notion de contraste, principe de Babinet

On considère une solution dont le soluté occupe une fraction volumique  $x$ , et dont  $b$ ,  $v$ ,  $\rho = b/v$  sont respectivement la longueur de diffusion, le volume moléculaire et la densité de longueur de diffusion. Notons  $x_0$ ,  $b_0$ ,  $v_0$  et  $\rho_0$  les mêmes grandeurs relatives au solvant. La densité moyenne de longueur de diffusion est  $\langle \rho \rangle = \rho x + \rho_0 x_0$ . Pour un élément de volume  $i$  de l'échantillon, l'écart à cette densité moyenne est  $\Delta \rho_i = (\rho x_i + \rho_0 x_{0i}) - \langle \rho \rangle = \rho \Delta x_i + \rho_0 \Delta x_{0i}$ , où  $\Delta x_i$  représente l'écart à la concentration moyenne en soluté. Si le milieu est incompressible alors  $x_0 = 1 - x$  et  $\Delta x_{0i} = -\Delta x_i$ , d'où  $\Delta \rho_i = (\rho - \rho_0) \Delta x_i$ . La fonction de corrélation des fluctuations de densité de longueur de diffusion est donc:

$$\langle \Delta \rho_i \Delta \rho_j \rangle = (\rho - \rho_0)^2 \langle \Delta x_i \Delta x_j \rangle \quad (1.21)$$

L'équation 1.14 devient

$$\mathcal{S}(q) = v^2 (\rho - \rho_0)^2 \mathcal{S}_{\text{soluté}}(q) \quad \text{avec} \quad \mathcal{S}_{\text{soluté}}(q) = \sum_i^{V/v} \sum_j^{V/v} \langle \Delta x_i \Delta x_j \rangle e^{i\vec{q} \cdot (\vec{r}_i - \vec{r}_j)} \quad (1.22)$$

Le facteur  $v^2 (\rho - \rho_0)^2 = (b - b_0 v/v_0)^2$  est le **facteur de contraste** entre le solvant et le soluté. Il représente le carré de la différence de longueur de diffusion entre un diffuseur élémentaire du soluté et les  $v/v_0$  molécules de solvant qui occupent le même volume. L'équation 1.22 montre que la diffusion est proportionnelle au facteur de contraste et à un terme d'interférence,  $\mathcal{S}_{\text{soluté}}(q)$ , qui ne dépend que du soluté.  $\mathcal{S}_{\text{soluté}}(q)$  est la **fonction de diffusion partielle** du soluté. Le même calcul peut être fait en considérant cette fois le solvant. Ce qui montre que les deux fonctions de diffusion partielles sont identiques:  $\mathcal{S}_{\text{soluté}}(q) = \mathcal{S}_{\text{solvant}}(q)$ . Ce résultat est connu en optique à la suite des travaux de Jacques Babinet (XIX<sup>ième</sup> S.) qui remarqua que les figures de diffraction produites par deux écrans complémentaires sont identiques.

**La diffusion est proportionnelle au facteur de contraste entre le solvant et le soluté.**

**Le terme d'interférence n'est lié qu'aux corrélations des fluctuations de concentration du soluté.**

L'expression 1.22 est analogue à celle obtenue pour un gaz à ceci près que les fluctuations de densité sont remplacées par les fluctuations de concentration du soluté. Par exemple, la section efficace différentielle de diffusion cohérente à  $q \rightarrow 0$  sera proportionnelle à la compressibilité osmotique:  $(Cd\pi/dC)^{-1}$ . Si  $C$  est la concentration en soluté exprimé en masse par unité de volume, l'équation 1.16 devient:

$$\frac{\mathcal{S}(q \rightarrow 0)}{V} = (b - b_0 v/v_0)^2 \left( \frac{C}{M} \right)^2 kT (Cd\pi/dC)^{-1} \quad (1.23)$$

où  $M$  est la masse moléculaire du soluté. A la limite  $C \rightarrow 0$ , la solution est idéale  $(Cd\pi/dC)^{-1} = (kTC/M)^{-1}$ . L'équation 1.17 devient:

$$\left. \frac{\mathcal{S}(q \rightarrow 0)/V}{C} \right]_{C \rightarrow 0} = \left( \frac{b - b_0 v/v_0}{M} \right)^2 M \quad (1.24)$$

Dans le cas où le soluté est une macromolécule faite de  $N$  diffuseurs élémentaires de longueur de diffusion  $b_1$ , masse  $m_1$  et de volume  $v_1$ . Alors  $M = Nm_1$ ,  $v = Nv_1$  et  $b = Nb_1$ . L'équation précédente devient:

$$\left. \frac{\mathcal{S}(q \rightarrow 0)/V}{C} \right]_{C \rightarrow 0} = K^2 M \quad \text{avec} \quad K^2 = \left( \frac{b_1 - b_0 v_1/v_0}{m_1} \right)^2 \quad (1.25)$$

La grandeur  $K^2$  est caractéristique du couple “diffuseur élémentaire/solvant”. Elle est calculée à partir de leurs formules chimiques. L’Eq. 1.25 montre alors la façon dont la mesure permet d’accéder à la masse moléculaire de la macromolécule.

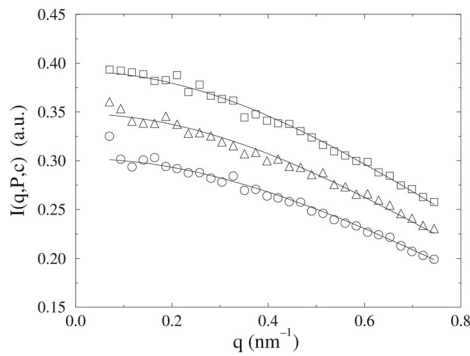
Pour des macromolécules en interaction faible telle qu’en solution diluée, la pression osmotique peut s’écrire sous la forme d’un développement du viriel:  $\pi = kT(n/V)(1 + a_2n/V + \dots)$  où  $a_2$  est le coefficient d’interaction à deux corps. C’est une grandeur homogène à un volume qui pour des sphères dures de rayon  $r$  est égal à  $a_2 = 16\pi r^3/3$  [3]. Il est d’usage d’utiliser la concentration  $C = nM/V$  et le coefficient  $MA_2 = a_2/M$  homogène à l’inverse d’une concentration (concentration d’empilement compact pour des sphères), on obtient pour la compressibilité osmotique  $(Cd\pi/dC)^{-1} = kT(C/M)(1 - 2a_2C/M + \dots)$ . D’où pour l’intensité diffusée:

$$\frac{\mathcal{S}(q \rightarrow 0)/V}{C} = \left[ \frac{\mathcal{S}(q \rightarrow 0)/V}{C} \right]_{C \rightarrow 0} (1 - 2MA_2C + \dots) \quad (1.26)$$

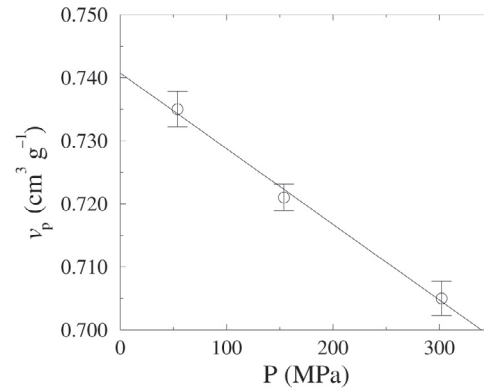
Le coefficient d’interaction à deux corps est donc directement accessible à partir de la dépendance en concentration de l’intensité diffusée.

### 1.9 Exemple 1 : Compressibilité de la myoglobine

Le développement des techniques des hautes pressions hydrostatiques suscite des expériences pour lesquelles la pression devient un paramètre thermodynamique pour l’étude des systèmes biologiques. La diffusion de neutrons est particulièrement bien adaptée à ces études car des matériaux transparents aux neutrons comme l’aluminium ou le quartz autorisent des environnements d’échantillon complexes. Récemment des expériences originales de DNPA ont permis de déterminer la compressibilité isotherme de la myoglobine [4].



**Fig. 1.** Scattering spectra  $I(q, P, c)$  of  $MbN_3$  at  $p^2H$  6.6, as a function of the wave-number transfer  $q$ . The measurements were performed at room temperature. The protein concentration,  $c$ , at atmospheric pressure is  $11.7 \text{ mg}\cdot\text{cm}^{-3}$  and the pressures,  $P$ , are: 54 (○), 154 (△), and 302 (□) MPa. Fits of Eqn 6 to the data are shown as full lines.



**Fig. 7.** Partial specific volume  $v_p(P)$ , of  $MbN_3$  as a function of pressure,  $P$ . The almost linear variation of  $v_p(P)$  with  $P$  allows the isothermal compressibility of the hydrated protein to be computed:  $\kappa_{T,P} = (1.6 \pm 0.1) 10^{-4} \text{ MPa}^{-1}$ .

**Figure 1.** Expériences de DNPA sur des solutions diluées de  $MbN_3$  à différentes pressions hydrostatiques  $P$  [4]. L’augmentation de l’intensité diffusée avec  $P$  est due à l’augmentation du contraste résultant d’une compressibilité isotherme de la protéine hydratée trois fois plus faible que celle de l’eau.

Les expériences sont réalisées en solution diluée dans l’eau lourde afin de réduire la diffusion incohérente de l’hydrogène  $^1H$ . Des mesures de la section efficace de diffusion cohérente par unité de volume,  $\mathcal{S}(q)/V$  (notée  $I$  par les auteurs), ont été réalisées en fonction du vecteur de diffusion  $q$  pour différentes pressions hydrostatiques  $P$ . Les auteurs exploitent l’idée qu’une différence de compressibilité entre le soluté et le solvant implique que le rapport  $v/v_0$  (cf. Eq. 1.24) est fonction de la pression d’où une

variation de l'intensité diffusée. L'allure des spectres permet une extrapolation à  $q \rightarrow 0$  (cf. section 2.2). L'équation 1.24 s'écrit en fonction des variables qui dépendent de la pression:

$$\left. \frac{\mathcal{S}(q \rightarrow 0)/V}{C(P)} \right]_{C \rightarrow 0} = \left( \frac{b}{M} - \rho_0(P)v'(P) \right)^2 M \quad (1.27)$$

Le volume diffusant  $V$  est constant car délimité par le faisceau de neutrons. La concentration de la solution  $C(P)$  et la densité de longueur de diffusion du solvant  $\rho_0(P)$  dépendent de la pression *via* la compressibilité du solvant qui par ailleurs est connue. La longueur de diffusion  $b$  de la protéine et sa masse molaire  $M$  sont connues et indépendantes de la pression. Le volume spécifique de la protéine  $v'(P) = v(P)/M$  peut donc être déduit de la mesure de  $\mathcal{S}(q \rightarrow 0)$ . Les mesures faites à différentes pressions donnent accès à la compressibilité de la protéine  $\kappa_T = (-v \cdot dP/dv)^{-1}$  de façon très précise. Ce type de mesure est particulièrement intéressant pour la compréhension de la dynamique interne d'une protéine. Par exemple, le coefficient de diffusion associé aux modes coopératifs de cette dynamique interne est proportionnel à  $\kappa_T^{-1}$  qui agit sur les fluctuations de concentration interne comme une force de rappel élastique.

### 1.10 Exemple 2 : Interactions, second viriel et cristallisation des protéines

La cristallisation des protéines est un passage obligé pour la détermination de leurs structures. Toutefois dans le volume, les méthodes permettant de faire croître des monocristaux sont encore très empiriques. Un point clé concerne les conditions physico-chimiques optimales à la nucléation. Dans ce contexte l'étude précise des interactions entre protéines en solution, en fonction du pH ou de la force ionique, est importante. Dans cet exemple [5], les auteurs étudient par diffusion de rayonnement (lumière et neutrons aux petits angles) ces interactions pour deux protéines le lysozyme et le chymotrypsinogène.

Les mesures de diffusion de lumière (Fig. 2 gauche) correspondent à la limites  $q \rightarrow 0$  (typiquement  $qR < 5 \cdot 10^{-2}$  pour une protéine globulaire de rayon  $R$ ). Elles sont une illustration directe de

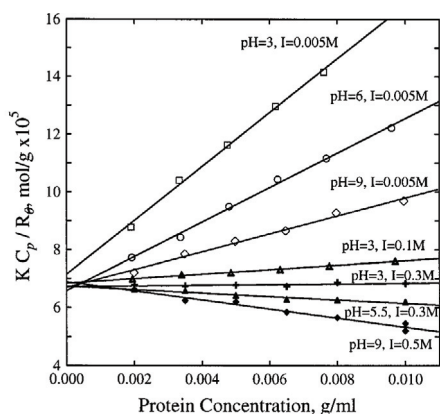


FIGURE 1 Typical plots of lysozyme SLS data according to Eq. 7. The variation of the virial coefficient (given by the slope of the lines) with pH and electrolyte concentration is evident.

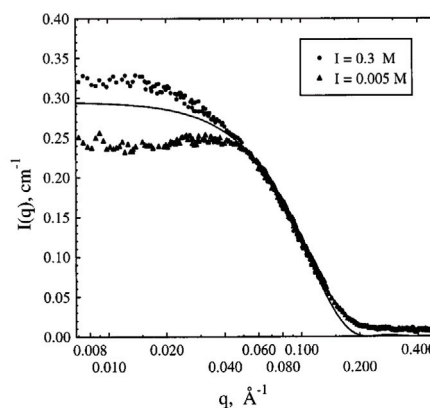


FIGURE 7 Intensity  $I$  versus scattering vector  $q$  for SANS measurements on chymotrypsinogen at pH 3 and two different electrolyte concentrations. The solid curve is the form factor function fitted by approximating the protein shape by a sphere. Positive deviations from this curve in the low  $q$  region indicate attractive interactions and negative deviations indicate repulsions.

**Figure 2.** Expériences de diffusion statique de la lumière (SLS) et de DNPA sur des solutions diluées de protéines [5]. Les expériences de diffusion de lumière correspondent à la limite  $q \rightarrow 0$ . La grandeur représentée, notée  $K C_p / R_\theta$ , est l'inverse de la masse apparente (la masse moléculaire correspond à la limite  $C \rightarrow 0$ ). Dans les notations utilisées dans ce cours  $K = ((b_1 - b_0 v_1 / v_0) / m_1)^2$  est le contraste,  $C_p$  est la concentration en protéines et  $R_\theta$  la section efficace de diffusion cohérente par unité de volume.

l'équation 1.26. La grandeur représentée est l'inverse de la section efficace de diffusion cohérente par unité de volume, de concentration et de contraste. Extrapolée à concentration nulle cette grandeur est égale à l'inverse de la masse des objets en solution (protéines isolées ou agrégats) selon l'Eq. 1.25. La pente des droites est égale à  $2MA_2$ . A pH = 9 par exemple, les interactions répulsives ( $2MA_2 > 0$ ) à faible force ionique (notée  $I$  ici) deviennent attractives à force ionique élevée. La figure de droite correspond à des mesures de DNPA effectuées sur un domaine de vecteur de diffusion plus large autour de  $qR = 1$ . Indépendamment de la forme des courbes qui sera exposée dans la partie suivante, la figure illustre les équations 1.8 à 1.11. A petite valeur du vecteur de diffusion, selon la nature des interactions entre protéines, on observe un excès ou un défaut d'intensité diffusée par rapport à la solution idéale. A grand vecteur de diffusion, l'intensité diffusée est insensible à ces interactions et ne dépend que du facteur de forme des protéines.

## 2. FACTEUR DE FORME

Très souvent l'information clé à laquelle l'expérimentateur souhaite accéder, concerne la structure ou la conformation des objets (qui peuvent être des macromolécules ou des complexes de macromolécules) en solution. Cette information est contenue dans le facteur de forme. Sont examinés ici quelques cas typiques.

### 2.1 Objets ayant des orientations aléatoires

En solutions les objets sont souvent libres de tourner autour de leur centre de masse. Le facteur de forme que l'on mesure est une moyenne sur toutes les orientations possibles de  $\vec{r}$  [6]. Du fait de la symétrie centrale, à chaque  $\vec{r}$  lui correspond son opposé. La partie imaginaire de  $e^{i\vec{q}\cdot\vec{r}}$  est donc nulle en moyenne, d'où  $\langle e^{i\vec{q}\cdot\vec{r}} \rangle = \langle \cos(\vec{q}\cdot\vec{r}) \rangle$ . La densité de probabilité d'une orientation  $\varphi$  donnée pour  $\vec{r}$  est  $p(\varphi) = [2\pi \sin(\varphi)d\varphi]/4\pi = \sin(\varphi)d\varphi/2$ , d'où:  $\langle \cos(\vec{q}\cdot\vec{r}) \rangle = \int \cos(\vec{q}\cdot\vec{r})p(\vec{r})d\vec{r} = \int_0^{\pi/2} \cos(qr \cos(\varphi)) \sin(\varphi)d\varphi$ . Le changement de variable  $x = qr \cos(\varphi)$  conduit à

$$\langle e^{i\vec{q}\cdot\vec{r}} \rangle_{\text{orientations}} = \langle \cos(\vec{q}\cdot\vec{r}) \rangle_{\text{orientations}} = \frac{\sin(qr)}{qr} = j_0(qr) \quad (2.1)$$

$j_0$  est la fonction de Bessel sphérique d'ordre 0. On obtient finalement pour le facteur de forme :

$$P(q) = \frac{1}{N^2} \sum_i^N \sum_j^N \langle e^{i\vec{q}\cdot\vec{r}_{ij}} \rangle = \frac{1}{N^2} \sum_i^N \sum_j^N \frac{\sin(qr_{ij})}{qr_{ij}} \quad (2.2)$$

$r_{ij}$  désigne la norme du vecteur. Le produit  $qr_{ij}$  ne dépend donc que de la distance entre  $i$  et  $j$  et non de leur orientation. L'application immédiate du résultat précédent concerne une paire de diffuseurs élémentaires à une distance  $d$  l'un de l'autre. L'équation 2.2 donne immédiatement:

$$P_{\text{paire}}(q) = \frac{1}{2} \left( 1 + \frac{\sin(qd)}{qd} \right) \quad (2.3)$$

### 2.2 Domaine de Guinier, rayon de giration

On envisage la limite des petites valeurs de  $q$  ( $q^{-1}$  beaucoup plus grand que la taille des objets). Dans le cas où le facteur de forme est une moyenne sur toutes les orientations possibles, le développement  $\sin(qr)/qr = 1 - (qr)^2/6 + \dots$  introduit dans l'Eq. 2.2 conduit à:

$$P_{\text{Guinier}}(q) = 1 - \frac{q^2}{3} \left[ \frac{1}{2N^2} \sum_i^N \sum_j^N r_{ij}^2 \right] + \dots = 1 - \frac{q^2 R_g^2}{3} + \dots \quad (2.4)$$

qui peut s'écrire:

$$P_{\text{Guinier}}(q) \approx e^{-q^2 R_g^2/3} \quad (2.5)$$

avec  $R_g^2 = \frac{1}{2N^2} \sum_i^N \sum_j^N r_{ij}^2$ . Ce résultat est correct quelle que soit la forme des objets et dans un domaine de vecteur de diffusion, appelé domaine de Guinier, tel que  $qR_g < 1$ . La longueur  $R_g$  est le rayon de giration des objets. Il est plus commode de l'exprimer en fonction des distances au centre de gravité pris pour origine. Ainsi  $\sum_i^N \sum_j^N r_{ij}^2 = \sum_i^N \sum_j^N (\vec{r}_i - \vec{r}_j)^2 = \sum_i^N \sum_j^N (\vec{r}_i^2 + \vec{r}_j^2 - 2\vec{r}_i \cdot \vec{r}_j) = N \sum_i^N r_i^2 + N \sum_j^N r_j^2$  car  $\sum_i^N \vec{r}_i = 0$ . Finalement:

$$R_g^2 = \frac{1}{N} \sum_i^N r_i^2 = \langle r_i^2 \rangle \quad (2.6)$$

Le rayon de giration d'un objet est le rayon de la sphère (creuse) qui aurait le même moment d'inertie. Par exemple, dans une boule (pleine) de rayon  $R$ , les diffuseurs élémentaires ne sont pas uniquement situés à la périphérie mais également à des distances plus petites du centre. Le rayon de giration étant la moyenne quadratique des  $r_i$ , est donc plus petit que  $R$ . En remplaçant la somme par une intégrale, on obtient  $R_g^2 = \left[ \int_0^R 4\pi r^2 \times r^2 \times dr \right] / \frac{4}{3}\pi R^3 = \frac{3}{5}R^2$ . Pour un filament rigide de longueur  $L$ , on obtient  $R_g^2 = L^2/12$ .

### 2.3 Objets à symétrie sphérique

Pour un objet à symétrie sphérique l'orientation n'a pas de signification. Par ailleurs, il est commode de placer l'origine au centre de gravité.  $P(q) = \frac{1}{N^2} \sum_i^N \sum_j^N e^{i\vec{q} \cdot \vec{r}_{ij}} = [A(q)]^2$ , avec  $A(q) = \frac{1}{N} \sum_i^N e^{i\vec{q} \cdot \vec{r}_i}$ . Pour une sphère de rayon  $R$ , l'équation 2.1 donne directement  $A(q) = \sin(qR)/qR$ , d'où le facteur de forme:

$$P_{\text{sphère}} = \left[ \frac{\sin(qR)}{qR} \right]^2 = [j_0(qR)]^2 \quad (2.7)$$

Pour un objet volumique dont la densité de diffuseurs élémentaires à une distance  $r$  du centre est  $\rho(r)$ :  $A(q) = \int_0^R j_0(qr)\rho(r)4\pi r^2 dr$ . Pour une boule on obtient:

$$P_{\text{boule}}(q) = \left[ \frac{3}{qR} \times \left( \frac{\sin(qR)}{(qR)^2} - \frac{\cos(qR)}{qR} \right) \right]^2 = \left[ \frac{3j_1(qR)}{qR} \right]^2 \quad (2.8)$$

$j_1$  est la fonction de Bessel sphérique d'ordre 1. De façon générale le facteur de forme d'objets à symétrie sphérique fait intervenir ces fonctions qui sont des combinaisons algébriques des fonctions trigonométriques. Les facteurs de forme correspondant présentent des oscillations dont la période est fonction de  $qR$ . A grand  $q$ , l'enveloppe de ces oscillations décroît comme  $q^{-2}$  pour la sphère et  $q^{-4}$  pour la boule. Dans la pratique, la résolution finie de l'appareil de mesure ou une polydispersité des tailles font que seule cette enveloppe est mesurée lorsque ces oscillations sont très resserrées ( $qR \gg 1$ ). Lorsque ces oscillations sont visibles et que la forme des objets est par ailleurs connue, la taille des objets peut ainsi être mesurée sans faire de mesure dans le domaine de Guinier ( $qR < 1$ ).

### 2.4 Bâton rigide

On considère [7] un bâton infiniment fin de longueur  $L$ . Soit  $\vec{u}$  un vecteur unitaire colinéaire au bâton. L'amplitude de l'onde diffusée est:  $A(q) = (1/L) \int_{-L/2}^{L/2} e^{i\vec{q} \cdot \vec{r}u} dr$ . Du fait de la symétrie, seul le terme en cosinus est conservé, soit  $A(q) = j_0(\vec{q} \cdot L\vec{u}/2)$ . Le facteur de forme est une moyenne de  $A(q)^2$  sur toutes

les orientations possibles de  $\vec{u}$  (cf. section 2.1):

$$P_{\text{bâton}}(q) = \frac{1}{2} \int_0^\pi \left( j_0 \left( \frac{qL}{2} \cos \varphi \right) \right)^2 \sin \varphi d\varphi$$

Le changement de variable  $x = \cos \varphi$  donne  $P_{\text{bâton}}(q) = \int_0^1 \left( j_0 \left( \frac{qL}{2} x \right) \right)^2 dx$ ; puis  $y = qLx/2$  conduit à:

$$P_{\text{bâton}}(q) = \frac{2}{qL} \int_0^{qL/2} j_0(y)^2 dy \quad (2.9)$$

$\int_0^\infty j_0(y)^2 dy = \pi/2$  d'où le comportement asymptotique:  $P_{\text{bâton}}(qL \gg 1) = (\pi/L)q^{-1}$ .

## 2.5 Chaîne gaussienne

Une chaîne gaussienne (ou idéale) correspond à la trace laissée par une marche aléatoire pour laquelle le déplacement carré moyen  $\langle r^2 \rangle$  est proportionnel au nombre  $n$  de pas de longueur  $a$ :  $\langle r^2 \rangle = a^2 n$ . Pour une telle chaîne, la distance entre extrémités est nulle en moyenne mais elle obéit à une loi de distribution gaussienne:  $p(r) = (2\pi/3\langle r^2 \rangle)^{-3/2} e^{-3r^2/2\langle r^2 \rangle}$ . Dans l'équation 2.1 du facteur de forme, la moyenne doit être réalisée sur cette distribution des distances  $r_{ij}$  [8].

$$P(q) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \left( \int e^{i\vec{q}\cdot\vec{r}_{ij}} p(r_{ij}) d^3 r_{ij} \right) \quad (2.10)$$

Cette moyenne est la transformée de Fourier d'une gaussienne, c'est également une gaussienne de largeur inverse:  $\int e^{i\vec{q}\cdot\vec{r}_{ij}} p(r_{ij}) d^3 r_{ij} = e^{-q^2 \langle r_{ij}^2 \rangle / 6}$ , où  $\langle r_{ij}^2 \rangle$  est une fonction du nombre  $j - i$  de maillons entre ces deux points:  $\langle r_{ij}^2 \rangle = a^2 |i - j|$ .

$$P(q) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N e^{-q^2 a^2 |i-j|/6} = \frac{2}{N^2} \sum_{k=1}^N (N-k) e^{-q^2 a^2 k/6} \quad (2.11)$$

La dernière égalité est obtenue en remarquant que parmi les  $N^2$  paires  $(i, j)$  possibles, il en existe  $2(N-k)$  de  $k = |i - j|$  maillons. En remplaçant la somme par une intégrale:  $P(q) = \frac{2}{N^2} \int_0^N (N-k) e^{-q^2 a^2 k/6} dk = 2 \int_0^1 (1-u) e^{-uq^2 a^2 N/6} du$ , avec  $u = k/N$ . En notant  $x = a^2 N/6$  et en intégrant par parties on obtient finalement la formule de Debye:

$$P(q) = \frac{2}{(q^2 R^2)^2} \left( q^2 R^2 - 1 + e^{-q^2 R^2} \right) \quad (2.12)$$

A  $qR \ll 1$ , le développement  $e^{-x} = 1 - x + x^2/2 - x^3/6 + \dots$  et l'identification avec l'équation 2.4 montre que  $R = R_g$ . A  $qR \gg 1$ , le facteur de forme décroît comme  $P(q) = (qR)^{-2}$ .

## 2.6 Objets autosimilaires : argument d'échelle

Les objets de forme euclidienne ont une masse,  $M$ , qui varie comme la puissance 1, 2 ou 3 de leur taille  $R$ . Pour les objets fractals, cette puissance n'est pas nécessairement entière:

$$M = m_1 \left( \frac{R}{a} \right)^D \quad (2.13)$$

où  $m_1$  et  $a$  sont respectivement la masse et la taille d'un élément constitutif de l'objet.  $D$  est la dimension fractale de cette famille d'objets. Elle caractérise la façon dont ils remplissent l'espace. Les fractals sont le plus souvent autosimilaires, c'est à dire invariant par changement d'échelle. Une fois grossie, une petite partie est statistiquement semblable à l'objet entier. Observons sous différents grossissements la

pelote que forme une chaîne polymère de rayon de giration  $R_g$ . Si  $\mathcal{R}$  est la taille de la zone observable, pour  $R_g/\mathcal{R} < 1$  (petit grossissement), la masse visible,  $m$ , est égale à la masse totale  $M$ . Par contre pour  $R_g/\mathcal{R} > 1$  (fort grossissement) la masse visible décroît. Supposons une loi du type:

$$m\left(\frac{R_g}{\mathcal{R}}\right) = M \cdot f\left(\frac{R_g}{\mathcal{R}}\right) \quad \text{avec} \quad \begin{cases} f(x < 1) = 1 \\ f(x \geq 1) = x^{-\alpha} \end{cases} \quad (2.14)$$

Pour déterminer l'exposant  $\alpha$ , on utilise un argument d'échelle qui postule [9]:

1. qu'une seule longueur est pertinente pour décrire l'objet (ici son rayon  $R_g$ ).
2. qu'à fort grossissement, la mesure est insensible à  $M$  (on ne peut déduire la masse totale de l'objet en n'observant qu'une partie). Soit  $m(R_g/\mathcal{R} > 1) = M^0$ .

La relation 2.13 donne  $\alpha = \mathcal{D}$ . Le même argument est utilisé pour décrire le résultat d'une expérience de diffusion de rayonnement. En solution très diluée, la grandeur physique donnant accès à la masse d'un objet est la section efficace de diffusion cohérente par unité de concentration et de contraste:  $[\mathcal{S}(q)/(CVK^2)]_{C \rightarrow 0} = MP(q)$  (cf. Eq. 1.7). Dans le régime intermédiaire du vecteur de diffusion tel que  $R_g < q^{-1} < a$ , la mesure sonde l'intérieur de l'objet et est sensible à son autosimilarité. L'argument d'échelle postule qu'une seule longueur est pertinente (or  $R_g$  est déjà nécessaire à  $qR_g < 1$ ) et que l'expérience est insensible à  $M$ . Soit:

$$[\mathcal{S}(q)/(CVK^2)]_{C \rightarrow 0} = M \cdot P(qR_g) \quad \text{avec} \quad \begin{cases} P(qR_g < 1) = 1 \\ P(qR_g \geq 1) = (qR_g)^{-\mathcal{D}} \end{cases} \quad (2.15)$$

A  $qR_g > 1$ , On obtient :

$$[\mathcal{S}(q)/(CVK^2)]_{C \rightarrow 0, qR_g > 1} = M \cdot (qR_g)^{-\mathcal{D}} = m_1 \left(\frac{q^{-1}}{a}\right)^{\mathcal{D}} \quad (2.16)$$

Cette relation exprime l'autosimilarité d'un objet seul tandis que la relation 2.13 exprime celle des chaînes entre elles. La dernière formulation de l'Eq. 2.16 montre son sens physique si on la compare à l'Eq. 2.13. L'intensité diffusée à  $qR > 1$  est proportionnelle au nombre,  $g = (q^{-1}/a)^{\mathcal{D}}$ , de diffuseurs élémentaires d'un segment de chaîne de rayon  $q^{-1}$ .

**Les diffuseurs situés dans un volume  $q^{-3}$  diffusent de façon cohérente tandis que ces volumes entre eux diffusent de façon incohérente.**

Cet argument d'échelle permet de retrouver facilement le comportement de l'intensité diffusée pour les objets examinés jusqu'à présent. La dimension fractale d'une chaîne idéale est celle d'une marche aléatoire,  $\mathcal{D} = 2$  et  $P(q) \propto q^{-2}$  (Eq. 2.12). Pour un bâton  $\mathcal{D} = 1$ ,  $P(q) \propto q^{-1}$  (Eq. 2.9). Pour une sphère, à suffisamment grand  $q$  la mesure est insensible à la courbure, on voit un plan soit  $\mathcal{D} = 2$ ,  $P(q) \propto q^{-2}$  (Eq. 2.7). Pour une boule, l'argument est un peu différent, il est développé ci-dessous.

## 2.7 Diffusion par une interface, loi de Porod

Certains objet sont essentiellement caractérisés par une interface entre deux milieux homogènes de longueurs de diffusion,  $b$  et  $b_0$ , différentes. C'est le cas d'objets denses, homogènes et de grande taille par rapport à l'échelle d'observation. Pour obtenir la fonction de diffusion, utilisons [10] un argument d'échelle similaire à celui de la partie 2.6. Partageons l'échantillon en volumes de cohérence  $q^{-3}$ . Seuls les volumes de cohérence situés à l'interface contribuent à l'intensité diffusée, car ceux situés dans le volume correspondent à la diffusion d'un milieu homogène. La section efficace différentielle de diffusion cohérente s'écrit donc

$$\frac{\mathcal{S}(q)}{(b - b_0)^2} = ng^2 \quad (2.17)$$

où  $n$  est le nombre de volumes de cohérence situés à l'interface et  $g$  le nombre de diffuseurs élémentaires dans chacun. Si  $\mathcal{A}$  est l'aire de l'interface, alors  $n \propto (\mathcal{A}/q^{-2})$ . Par ailleurs les deux milieux étant homogènes,  $g$  varie comme le volume  $q^{-3}$ ,  $g = (q^{-1}/a)^3$ . On obtient:

$$\frac{\mathcal{S}(q)}{(b-b_0)^2} \propto \frac{\mathcal{A}}{a^6} \cdot q^{-4} \quad (2.18)$$

Cette relation est connue sous le nom de loi de Porod. Cette forte décroissance de l'intensité diffusée est la signature d'une interface nette entre deux milieux homogènes. On retrouve ce comportement de l'intensité diffusée à  $qR \gg 1$  pour des boules.

## 2.8 Fonction de coupure à grand $q$

Les facteurs de formes envisagés jusqu'ici, l'ont été dans un domaine de vecteur de diffusion pour lequel le diffuseur élémentaire est considéré ponctuel. Par exemple le bâton est vu comme une ligne sans épaisseur. A suffisamment grand  $q$  cette approximation est incorrecte et la mesure devient sensible à la structure locale. Dans le cas particulier où le diffuseur élémentaire a une symétrie sphérique, ou si les diffuseurs élémentaires ont tous la même orientation,  $P(q)$  se factorise de façon analogue à celle de la partie 1.3.

$$P(q) = P_{\text{global}}(q) \times P_{\text{local}}(q) \quad (2.19)$$

où  $P_{\text{global}}(q)$  est le facteur de forme de l'objet si les diffuseurs élémentaires étaient ponctuels et réduit à leurs centres de gravité et  $P_{\text{local}}(q)$  est le facteur de forme du diffuseur élémentaire. Si l'échelle d'observation  $q^{-1}$  reste supérieure à la taille  $a$  du diffuseur élémentaire,  $P_{\text{local}}(q)$  pourra être approximé par la formule de Guinier (Eq. 2.5):

$$P(q) = P_{\text{global}}(q) \times e^{-q^2 a^2/3} \quad (2.20)$$

## 2.9 Facteurs de formes divers

Les calculs qui ont été détaillés plus haut montrent les techniques de base utilisées pour établir le facteur de forme. En biologie, d'autres formes peuvent se rencontrer. Certaines sont listées ci-dessous. Dans chaque cas le résultat est une moyenne sur toutes les orientations.

○ *Disque de rayon  $R$*

$$P(q) = \frac{2}{(qR)^2} \left( 1 - \frac{1}{qR} J_1(2qR) \right) \quad (2.21)$$

$J_1$  est la fonction de Bessel d'ordre 1. Si ce disque à une épaisseur  $2L$ , dans le régime intermédiaire de vecteur de diffusion tel que  $1/R < q < 1/L$ , une approximation du facteur de forme est:

$$P(q) = \frac{2}{(qR)^2} e^{-q^2 L^2/3} \quad (2.22)$$

On reconnaît le profil en  $q^{-2}$  caractéristique d'un objet bidimensionnel et la fonction de coupure exponentielle à grand  $q$  (Eq. 2.20).

○ *Cylindre plein de rayon  $R$  et longueur  $L$* . Dans le régime intermédiaire de vecteur de diffusion tel que  $1/L < q < 1/R$ :

$$P(q) = \frac{\pi}{qL} e^{-q^2 R^2/4} \quad (2.23)$$

C'est l'équation d'un bâton avec une fonction de coupure liée à l'épaisseur.

○ *Cylindre creux de longueur infinie de rayon extérieur  $R_1$  et rayon intérieur  $R_2$ .*

$$P(q) = \frac{1}{q} \left( \frac{R_1 J_1(qR_1) - R_2 J_1(qR_2)}{q(R_1^2 - R_2^2)} \right)^2 \quad (2.24)$$

Les exemples pourraient être multipliés. Très souvent, lorsque l'on a affaire à une structure complexe particulière, l'écriture d'une forme analytique est soit impossible, soit très lourde. Il est alors plus intéressant de revenir à l'expression de base du facteur de forme et à calculer numériquement les intégrations nécessaires. En particulier les techniques de simulation de Monte-Carlo sont une façon de réaliser numériquement ces intégrations.

### 2.10 Exemple 3 : Facteurs d'élongation de la traduction

La traduction est la synthèse d'une protéine catalysée par le ribosome à partir de l'information contenue dans la séquence d'un ARN messager. La traduction se déroule en trois phases : l'initiation, l'élongation et la terminaison. Chacune requiert la présence de protéines spécifiques (facteurs protéiques). Les facteurs d'élongation (eEF chez les eucaryotes) interviennent lors de l'allongement des chaînes protéiques en cours de synthèse. Lors de cette phase, un facteur protéique eEF1A catalyse la fixation de l'aminocyl-ARNt sur le ribosome, permettant ainsi l'ajout de l'acide aminé à la chaîne peptidique naissante. Les auteurs de cet exemple [11] étudient la formation du complexe eEF1A à l'ARNt en solution.

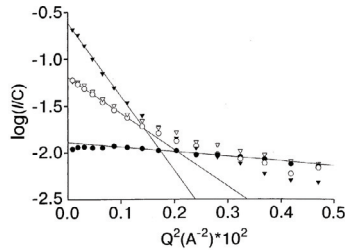


FIGURE 5: Dependence of neutron scattering intensity  $I$  on scattering vector  $Q$  in Guinier coordinates to tRNA (●), eEF1A (○), eEF1A:tRNA = 1:3 (▽), and eEF1A:tRNA=3:1 (▼).

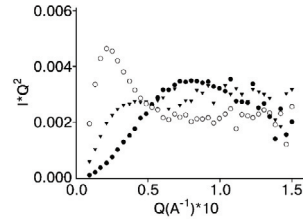


FIGURE 6: Dependence of neutron scattering intensity  $I$  on scattering vector  $Q$  in Kratky coordinates for tRNA (●), eEF1A (▼), and eEF1A:tRNA = 3:1 (○).

**Figure 3.** Expériences de DNPA [11] sur des solutions diluées du facteur d'élongation eEF1A en présence d'ARNt à différentes stœchiométries. La figure de gauche montre l'augmentation de la masse et du rayon des objets en solution lors de la formation du complexe eEF1A/ARNt. Celle de droite met en évidence la structure dense du complexe comparée à celle de la protéine libre.

La figure 3 gauche représente le log de l'intensité diffusée par unité de concentration en fonction de  $q^2$  dans le domaine de Guinier du vecteur de diffusion. La pente des droites est proportionnelle à  $R_g^2/3$  où  $R_g$  est le rayon de giration des objets en solution. L'ordonnée à l'origine est proportionnelle au log de leur masse:  $\log(I/C) = \log(M) - q^2 R_g^2/3 + \dots$ . La figure montre l'augmentation de la taille des objets et de leur masse lorsque la stœchiométrie eEF1A/ARNt est suffisante. L'augmentation de l'ordonnée à l'origine donne accès à la stœchiométrie des complexes (ici 2). La figure de droite est une représentation classique dite de Kratky:  $I(q) \times q^2$  en fonction de  $q$ . Le facteur de forme étant globalement une fonction décroissante, cette représentation permet de mieux mettre en évidence les différents profils. Quelle que soit la forme des objets,  $I(q) \times q^2$  est une fonction croissante dans le domaine de Guinier. Par contre à  $qR_g > 1$  une chaîne gaussienne est telle que  $I(q) \times q^2$  est constant tandis que pour une boule  $I(q) \times q^2$  est décroissant. Ici, la figure montre nettement que eEF1A n'est pas globulaire car l'intensité diffusée ne passe pas le maximum prononcé que l'on attend pour une boule. Par contre la conformation du complexe eEF1A/ARNt est beaucoup plus compacte.

### 2.11 Exemple 4 : Repliement d'une protéine non globulaire multidomaine

La fibronectine est la protéine fonctionnelle majeure de la matrice extracellulaire. Elle est indirectement impliquée dans les processus de cicatrisation, d'angiogénèse et d'invasion cellulaire. C'est une protéine de haute masse moléculaire (5000 acides aminés) constituée de 56 modules, identifiés car résistants à une digestion protéolytique. Ces modules ont chacun une structure globulaire bien déterminée et connue par ailleurs. Grâce à des expériences de DNPA [12], la conformation globale de cette protéine en solution se révèle être précisément celle d'un chapelet aléatoire de 56 globules. Le facteur de forme est calculé à partir de l'Eq. 2.19.  $P_{\text{global}}(q)$  est obtenu par simulation de Monte-Carlo d'une marche aléatoire de 56 pas de longueur  $2a$  et  $P_{\text{local}}(q)$  est le facteur de forme d'une boule de rayon  $a$ . Sans paramètre ajustable le calcul est en bon accord avec l'expérience (Fig. 5 gauche).

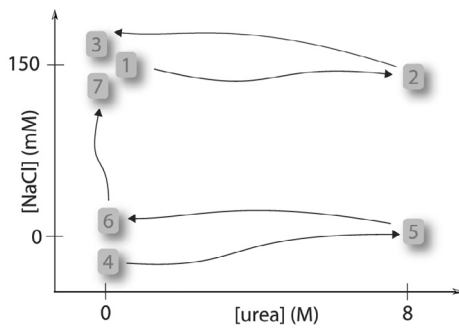


FIGURE 1 The seven samples here studied and their history.

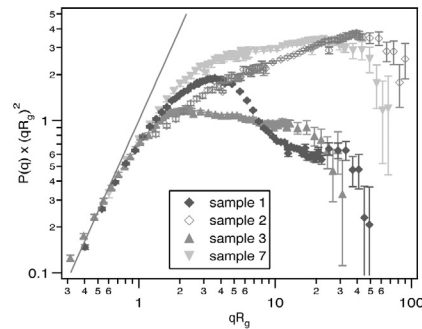


FIGURE 4 Form factors of samples in salt-containing solution in a Kratky representation  $P(q) \times (qR_g)^2$  versus  $qR_g$ . The symbols meaning is the same as in Fig. 2. The straight line corresponds to  $(qR_g)^2$ .

**Figure 4.** La fibronectine est une protéine de grande masse moléculaire (550 kg/mol). Une fois dénaturée par l'urée puis l'agent dénaturant éliminé, la conformation de la chaîne revenue à des conditions physiologiques de salinité, de pH et de concentration en urée dépend du chemin parcouru [13]. Ainsi, les échantillons 1, 3 et 7 présentent des facteurs de forme très différents (figure de droite).

Une fois dénaturée par de l'urée la protéine adopte la structure gonflée d'un polymère en bon solvant ( $D = 1.7$ ) qui est due aux interactions répulsives à courte portée (interactions de volume exclu) entre monomères d'une même chaîne. Le facteur de forme est correctement ajusté par la fonction  $P(q) = (1 + (qR)^2)^{-D/2} \times P_{\text{local}}(qa)$  qui rend compte des comportements asymptotiques à  $qR_g < 1$  et  $qR_g > 1$ . La longueur de persistance d'une chaîne peptidique dépliée est de l'ordre de grandeur de sa dimension radiale  $a$ . Le facteur de forme d'une sphère est donc une bonne approximation de la fonction de coupure à grand  $q$  (Fig. 5 gauche).

Peut-on replier une telle chaîne? Nous avons montré [13] qu'une fois dépliée puis l'agent dénaturant éliminé, la conformation finale de la protéine ramenée à des conditions physico-chimiques physiologiques dépend du chemin parcouru dans le diagramme salinité vs. concentration en urée (Fig. 4). Par exemple, le repliement réalisé en présence de 150 mM de NaCl conduit à la conformation d'une chaîne globulaire à grande échelle mais qui reste localement idéale. Le facteur de forme est bien ajusté par celui d'une chaîne confinée calculé par simulation de Monte-Carlo (Fig. 5 droite). Les résultats de DNPA montrent que la conformation "native" est celle qui possède la plus forte concentration locale, ce qui laisse supposer qu'elle n'est vraisemblablement pas à l'équilibre thermodynamique. Le principe d'Anfinsen, postulant que le milieu physicochimique et l'information contenue dans la séquence primaire sont suffisants au repliement, ne s'applique pas pour ces grandes protéines. Des associations se produisent au sein d'une même molécule entre domaines éloignés dans la séquence primaire empêchant l'individualisation des domaines structuraux.

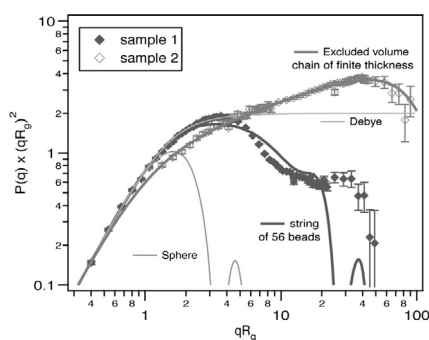


FIGURE 6 Form factor of the native and unfolded fibronectin plotted in a Kratky representation:  $P(q) \times (qR_g)^2$  versus  $qR_g$ . The full lines correspond to the theoretical expectation for a sphere (Eq. 8), a string of 56 beads (Eq. 12), a Gaussian chain with infinitely small monomer (Debye function, Eq. 9), and a swollen chain with a finite thickness (Eq. 13).

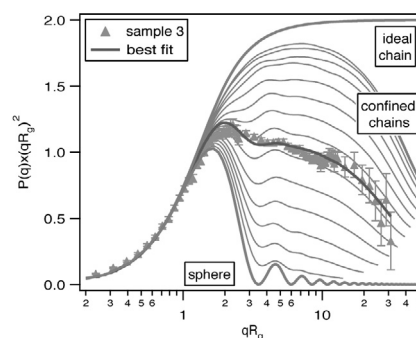


FIGURE 8  $P(q) \times (qR_g)^2$  versus  $qR_g$  for fibronectin refolded in salt-containing solution (data points) compared to confined chains (lines). (Top to bottom), lines correspond to: Debye function (unconfined Gaussian chain with infinitely small monomer); confined chains with radius of confinement,  $R_c$  from 30- to 4-step-length unit; and compact sphere. The line fitting the data corresponds to  $\bar{R}_c = 16 \pm 2$  step-length unit.

**Figure 5.** Le facteur de forme de la fibronectine native est celui d'une chaîne idéale de 56 globules de rayon 2.5 nm [12]. Celui de la chaîne dépliée correspond à une chaîne à volume exclu de 1250 unités de 0.66 nm de rayon. La structure de la chaîne repliée en présence de sel est correctement ajustée par cette chaîne de 1250 segments confinée dans une sphère de 13.9 nm de rayon. La fraction volumique en protéine dans ce globule est de l'ordre de 30%.

## 2.12 Exemple 5 : Superhélice d'ADN

L'ADN se présente sous la forme d'une double hélice dans laquelle les deux chaînes complémentaires s'enroulent autour d'un axe. Sous l'effet de contraintes mécaniques ou sous l'action d'enzymes spécifiques, cette double hélice peut former une hélice d'ordre supérieur (on parle de superhélice ou de surenroulement) d'une façon comparable au fil torsadé d'un téléphone. Le surenroulement est positif ou négatif selon le sens de cette superhélice comparé à celui de la double hélice. Le surenroulement joue un rôle dans les processus de densification de l'ADN mais également dans la régulation de nombreux processus biologiques, comme la réplication, la transcription et la recombinaison, en favorisant ou en s'opposant à la séparation des brins. Dans ce contexte, la compréhension de la structure et de la thermodynamique de cette superhélice est particulièrement importante. Les auteurs de cet exemple [14] étudient par DNPA des solutions d'ADN double brin circulaire (plasmide pUC18). A cette échelle d'observation, l'ADN relaxé (non surenroulé) est perçu comme un bâton rigide (cf. Fig. 6 gauche):  $P_{\text{relaxé}} \propto q^{-1}$ . Lors du surenroulement deux double hélices sont contraintes à se rapprocher. Prenons comme élément de base de la superhélice une paire de segments de double brin en vis-à-vis. A cette échelle d'observation ces paires sont alignées et le facteur de forme de la superhélice se factorise (Eq. 2.19 et 2.3):

$$P_{\text{superhélice}} \propto q^{-1} \times \left( 1 + \frac{\sin(qd)}{qd} \right)$$

où  $d$  est le diamètre de la superhélice. Les mesures réalisées à différentes forces ioniques mettent en évidence un diamètre passant de 16 à 9 nm pour des concentrations respectives en NaCl de 0 et 100 mM. L'écrantage des interactions électrostatiques favorise la densification de l'ADN.

## 3. VARIATIONS SUR LE CONTRASTE

La spécificité de la diffusion de neutrons réside dans la possibilité de jouer sur le contraste entre les différents constituants d'un système complexe. Diverses méthodes permettent alors d'accéder aux différentes fonctions de corrélation permettant de décrire pleinement le système. Ces méthodes ont

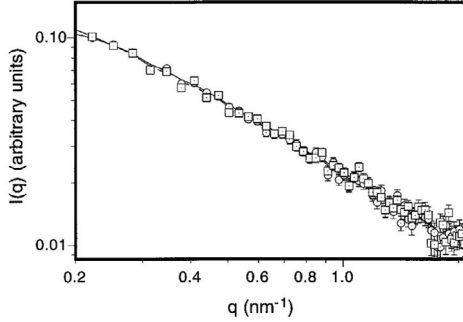


FIGURE 4 Measured scattering intensity  $I(q)$  of pUC18, relaxed, in  $H_2O$ , 10 mM Tris, at 0 mM (O) and 100 mM (□) NaCl concentrations.

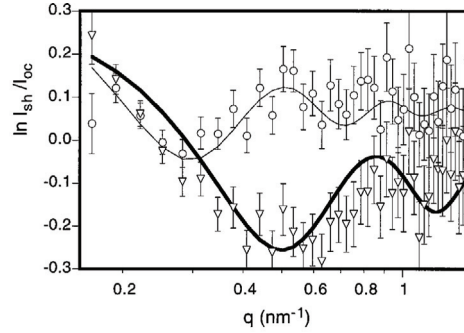


FIGURE 5 Ratio of the measured scattering intensities  $I(q)$  of pUC18, supercoiled, in  $H_2O$ , 10 mM Tris, at 0 mM (O), and 100 mM (∇)  $Na^+$  concentrations and relaxed pUC18 DNA at 100 mM  $Na^+$  concentration. The solid lines are the scattering form factors of a pair of point scatterers at a distance  $r = 16.0$  nm (thin line) and  $r = 9.0$  nm (thick line).

**Figure 6.** Expériences de DNPA [14] sur des solutions d’ADN double brin circulaire (plasmide pUC18). A gauche est tracé le spectre de la chaîne relaxée (sans sur-enroulement). A cette échelle  $q^{-1}$ , le facteur de forme est celui d’un bâton rigide (cf. partie 2.4) (longueur de persistance de la chaîne  $\approx 50$  nm). A droite est tracé le rapport du spectre de la forme sur-enroulée à celui de la forme relaxée. Les courbes sont ajustées par la fonction de l’Eq. 2.3. Le diamètre de la superhélice est inversement proportionnel à la période des oscillations. Il décroît avec l’augmentation de la force ionique.

largement fait leur preuve en physique des polymères et de la matière molle. En biologie l’exploitation de ces possibilités n’en est qu’à ses balbutiements du fait des difficultés d’obtenir des macromolécules biologiques dont les protons non labiles ont été substitués par des deutériums. Les progrès en génomique et en microbiologie devraient faciliter l’obtention de protéines recombinantes ainsi marquées et généraliser leur usage (cf. la communication de P. Timmins). Nous présentons ici certaines des possibilités offertes par ces méthodes de variation du contraste ainsi que quelques exemples de leur utilisation en biologie.

### 3.1 Mélange de deux espèces identiques, marquage isotopique

Un mélange incompressible d’objets identiques ne diffuse pas. Un des intérêts de la diffusion de neutrons est que même dans ce cas, il est possible de créer un contraste par le biais de la substitution isotopique des protons ( $^1H$  ou H) par des deutériums ( $^2H$  ou D). En effet, en diffusion de neutrons la longueur de diffusion d’un proton est égale à  $-3.74$  fm tandis que celle d’un deutérium est égale à  $6.67$  fm. Imaginons un mélange de molécules chimiquement identiques, mais dont une fraction  $x_D$ , de la population aurait subi la substitution de ses protons par des deutériums, ces objets sont dits “marqués”. Soit  $b_D$  et  $b_H$  les longueurs de diffusion respectives des diffuseurs élémentaires appartenant aux objet marqués et non marqués. A la suite de la partie 1.8, la section efficace de diffusion cohérente est  $\mathcal{S}(q) = (b_H - b_D)^2 \mathcal{S}_{DD}$ , où  $\mathcal{S}_{DD}$  est la fonction de diffusion partielle des objets marqués. Cette fonction s’écrit de façon similaire à l’équation 1.6 à l’aide d’un terme d’interférence “intra-objet”,  $P(q)$ , et d’un terme “inter-objet”,  $Q(q)$ . Si  $n$  est le nombre total d’objets et  $N$  le nombre de diffuseurs élémentaires par objet (nombre identique pour les objets marqués ou non), on obtient:

$$\mathcal{S}(q) = (b_H - b_D)^2 \times [x_D n N^2 P(q) + x_D^2 n^2 N^2 Q(q)] \quad (3.1)$$

En supposant que  $P(q)$  et  $Q(q)$  ne dépendent pas du marquage, on peut écrire la même relation en considérant la fraction  $(1 - x_D)$  d’objets non marqués :  $\mathcal{S}(q) = (b_H - b_D)^2 \times \mathcal{S}_{HH} = (b_H - b_D)^2 \times [(1 - x_D) n N^2 P(q) + (1 - x_D)^2 n^2 N^2 Q(q)]$ . L’égalité avec l’équation 3.1 donne:

$$P(q) = -n Q(q) \quad (3.2)$$

relation qui permet de simplifier l'équation 3.1:

$$\mathcal{S}(q) = (b_H - b_D)^2 \times x_D(1 - x_D) \times nN^2 P(q) \quad (3.3)$$

Dans ce cas, par le marquage des molécules on accède directement au facteur de forme des objets, quelle que soit la fraction d'objets marqués (contrairement à la méthode de dilution proposée dans la partie 1.2, Eq. 1.7). L'intensité diffusée est maximale pour  $x_D = 0.5$ .

### 3.2 Mélange de trois espèces, contraste moyen nul

On imagine comme précédemment un mélange d'objets identiques dont une certaine fraction  $x_D$  est marquée, mais on introduit maintenant un solvant. Si l'on considère un élément de volume, sa longueur de diffusion peut fluctuer de deux façons différentes qui n'ont aucune corrélation entre elles et qui par conséquent ont des contributions à l'intensité diffusée qui s'ajoutent. Pour une concentration donnée en objets, la fraction d'objets marqués peut varier. Par ailleurs, pour une fraction donnée d'objets marqués, la concentration totale en objets peut varier. La première contribution est identique au cas envisagé dans la partie 3.1 (Eq. 3.3), tandis que la seconde correspond au cas général d'une fluctuation de concentration d'objets (cf. partie 1.8, Eq. 1.22) dont la longueur de diffusion moyenne est  $\langle b \rangle = x_D b_D + (1 - x_D) b_H$ . On écrit:

$$\begin{aligned} \mathcal{S}(q) = & [(b_H - b_D)^2 \times x_D(1 - x_D) \times nN^2 P(q)] \\ & + (\langle b \rangle - b_0)^2 [nN^2 P(q) + n^2 N^2 Q(q)] \end{aligned} \quad (3.4)$$

Il est donc possible de trouver des conditions de contraste telles que  $\langle b \rangle - b_0$  soit nul, notamment en utilisant un mélange de solvant marqué et de solvant non marqué. On mesure alors directement le facteur de forme des objets sans avoir à les diluer. Des mesures dans d'autres conditions de contraste donnent accès à la fonction de diffusion inter-objet  $Q(q)$  ou éventuellement au facteur de structure  $S(q)$  si les objets sont sphériques (ou orientés).

### 3.3 Mélange de trois espèces, variation de contraste

Pour une solution d'objets constitués de diffuseurs élémentaires A et B appartenant à deux espèces différentes, en toute généralité ces objets peuvent avoir des structures ou des formes très différentes contrairement au cas précédent (partie 3.2). Afin de reprendre le calcul de la partie 1.8, notons  $x$  les fractions volumiques et  $\rho$  les densités de longueur de diffusion assortis des indices A et B pour chacune des espèces et de l'indice 0 pour le solvant. La densité moyenne de longueur de diffusion est  $\rho = \rho_A x_A + \rho_B x_B + \rho_0 x_0$ . Pour un élément de volume  $i$ , l'écart à cette moyenne est:  $\Delta\rho_i = \rho_A \Delta x_{Ai} + \rho_B \Delta x_{Bi} + \rho_0 \Delta x_{0i}$ . L'incompressibilité du milieu donne:  $\Delta x_{0i} = -\Delta(x_{Ai} + x_{Bi})$ , d'où  $\Delta\rho_i = (\rho_A - \rho_0) \Delta x_{Ai} + (\rho_B - \rho_0) \Delta x_{0i}$ . L'équation 1.21 devient:

$$\begin{aligned} \langle \Delta\rho_i \Delta\rho_j \rangle = & (\rho_A - \rho_0)^2 \langle \Delta x_{Ai} \Delta x_{Aj} \rangle + (\rho_B - \rho_0)^2 \langle \Delta x_{Bi} \Delta x_{Bj} \rangle \\ & + 2(\rho_A - \rho_0)(\rho_B - \rho_0) \langle \Delta x_{Ai} \Delta x_{Bj} \rangle \end{aligned} \quad (3.5)$$

Finalement la section efficace de diffusion (Eq. 1.22) s'écrit:

$$\mathcal{S}(q) = v^2 [(\rho_A - \rho_0)^2 \mathcal{S}_{AA}(q) + (\rho_B - \rho_0)^2 \mathcal{S}_{BB}(q) + 2(\rho_A - \rho_0)(\rho_B - \rho_0) \mathcal{S}_{AB}(q)] \quad (3.6)$$

où  $v$  est l'unité de volume. Il est possible de faire varier  $\rho_0$  en utilisant un mélange H<sub>2</sub>O/D<sub>2</sub>O par exemple. Trois mesures réalisées avec des contrastes différents permettent alors de déduire les fonctions de diffusion  $\mathcal{S}_{AA}$ ,  $\mathcal{S}_{BB}$  et  $\mathcal{S}_{AB}$ . Dans la pratique, une solution intéressante consiste à éteindre complètement le contraste, alternativement, pour chacune des deux espèces. Par exemple si un mélange de solvant permet d'obtenir que  $(\rho_B - \rho_0)$  soit nul, la section efficace différentielle de diffusion n'est sensible qu'à la fonction de diffusion de l'espèce A.

### 3.4 Marquage sélectif

Soit un complexe supramoléculaire constitué de  $n$  sous-ensembles. Si le marquage sélectif de deux sous-ensembles 1 et 2 est réalisable, il est alors possible à partir de quatre expériences différentes d'accéder à la fonction de diffusion partielle  $\mathcal{S}_{12}(q)$  [15]. De façon très générale la section efficace de diffusion s'écrit:  $\mathcal{S} = \sum_i^n \sum_j^n b_i b_j \mathcal{S}_{ij}$  où  $\mathcal{S}_{ij}$  est la fonction de diffusion partielle entre les sous-ensembles  $i$  et  $j$  et  $b_i, b_j$  leur longueur de diffusion. En isolant les termes qui concernent les deux sous-ensembles qui nous intéressent, on obtient la section efficace de diffusion en fonction de  $b_1$  et  $b_2$ :

$$\mathcal{S}(b_1, b_2) = 2b_1 b_2 \mathcal{S}_{12} + b_1^2 \mathcal{S}_{11} + b_2^2 \mathcal{S}_{22} + 2b_1 \sum_{i=3}^n b_i \mathcal{S}_{1i} + 2b_2 \sum_{i=3}^n b_i \mathcal{S}_{2i} + \sum_{i=3}^n \sum_{j=3}^n b_i b_j \mathcal{S}_{ij}$$

Soit  $b_{1D}$  et  $b_{1H}$  les longueurs de diffusion respectives du sous-ensemble 1 marqué et non-marqué, et  $b_{2D}, b_{2H}$  les mêmes grandeurs relatives au sous-ensemble 2. La première expérience consiste à marquer les sous-ensembles 1 et 2 et à mesurer  $\mathcal{S}(b_{1D}, b_{2D})$ , la seconde à marquer uniquement 1 et à mesurer  $\mathcal{S}(b_{1D}, b_{2H})$ , la troisième à marquer uniquement 2 et à mesurer  $\mathcal{S}(b_{1H}, b_{2D})$  et la quatrième à n'en marquer aucune et à mesurer  $\mathcal{S}(b_{1H}, b_{2H})$ . On montre facilement que:

$$\mathcal{S}(b_{1D}, b_{2D}) + \mathcal{S}(b_{1H}, b_{2H}) - \mathcal{S}(b_{1D}, b_{2H}) - \mathcal{S}(b_{1H}, b_{2D}) = 2(b_{1D} - b_{1H})(b_{2D} - b_{2H})\mathcal{S}_{12}(q) \quad (3.7)$$

On accède ainsi à la fonction de diffusion croisée des deux sous-ensembles qui permet en particulier de déterminer la distance qui les sépare.

Beaucoup d'autres possibilités sont envisageables sur ce thème général de la variation du contraste entre les différents constituant de systèmes complexes. Une revue plus détaillée de ces possibilités est donnée ref. [16].

### 3.5 Variation de contraste par polarisation nucléaire

Le neutron possède un moment cinétique de rotation sur lui-même appelé spin nucléaire  $\vec{s}$ . La portée de ses interactions avec un isotope ayant lui-même un spin,  $\vec{p}$ , non nul dépend du moment cinétique total du neutron et de cet isotope. La longueur de diffusion de cet isotope s'écrit de façon générale [17]:

$$b = b_0 + b_n \vec{s} \cdot \vec{p} \quad (3.8)$$

$b_0$  est la longueur de diffusion indépendante du spin dont il a été question jusqu'à présent. Le terme dépendant du spin est très grand dans le cas du proton pour lequel  $b_n = 58.2$  fm, ce qui est à l'origine de la principale source de diffusion incohérente en DNPA "classique" (cf. partie 1.7). Par contre, en utilisant un faisceau de neutrons polarisés ( $\langle \vec{s} \rangle \neq 0$ ) ainsi qu'un échantillon lui-même polarisé ( $\langle \vec{p} \rangle \neq 0$ ), il est possible de varier de façon continue et importante les densités de longueur de diffusion des molécules d'un échantillon. La polarisation à pratiquement 100% d'un faisceau de neutrons est une technique courante. Celle d'un échantillon se fait à basse température ( $\lesssim 1$  K), sous champ magnétique (quelques T) et en introduisant dans l'échantillon une faible quantité (quelques mM) de centres paramagnétiques. Les méthodes de variation de contraste (cf. partie 3.3) ou de marquage sélectif (cf. partie 3.4) pourront être utilisées en évitant (ou tout au moins en limitant le nombre) le marquage isotopique des molécules. Par ailleurs, cette technique de variation de contraste par polarisation nucléaire élimine certaines erreurs systématiques dues à la préparation d'échantillons différents et n'est limitée que par les erreurs statistiques de comptage des neutrons.

### 3.6 Exemple 6 : Étude d'un chaperon moléculaire, GroEL

Une fois synthétisées, les protéines doivent se replier dans l'espace et acquérir la structure unique propre à leur fonction. La compréhension de ce repliement est l'un des problèmes fondamentaux de

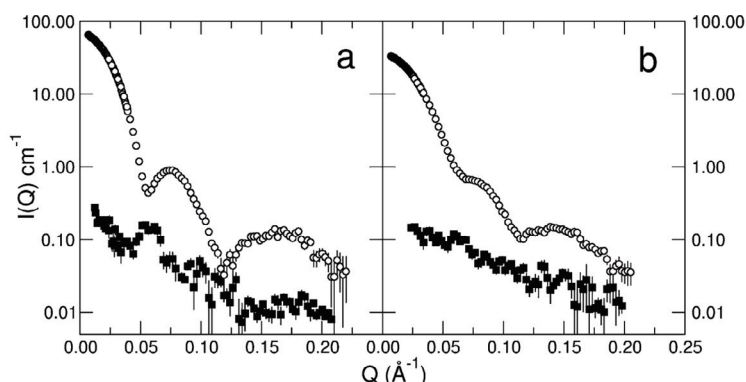


Fig. 6. Scattered intensities of the GroEL ( $I_{EL}(Q)$ ) (○) and dPJ9 ( $I_{dPJ9}(Q)$ ) (■) components as bound in (a) the GroEL/dPJ9 complex and (b) the srGroEL/dPJ9 complex.

**Figure 7.** Etude par DNPA [18] du complexe que forme un chaperon moléculaire GroEL (et son mutant simple anneau srGroEL) avec un substrat polypeptidique dénaturé, la subtilisine, obtenu sous une forme deutériée (dPJ9). La méthode de variation de contraste permet d'extraire les fonctions de diffusion dues aux autocorrélations de chaque constituant du complexe.

la biologie actuelle. *In vivo*, il est souvent assisté par des molécules dites "chaperons" (molécules qui sont elles-même des protéines). Les données structurales concernant ces molécules sont capitales pour comprendre leur fonctionnement. GroEL est un chaperon d'*E. Coli* formé de 14 sous-unités identiques qui s'assemblent en deux anneaux superposés. Les auteurs [18] étudient par DNPA le complexe que forme ce chaperon (ainsi qu'un mutant simple anneau, srGroEL) avec un substrat polypeptidique dénaturé la subtilisine PJ9. Pour ces expériences, le substrat PJ9 est totalement deutérié tandis que GroEL est naturel. La méthode de variation de contraste en utilisant différent mélange  $H_2O/D_2O$  permet d'extraire de la fonction de diffusion du complexe PJ9/GroEL, la fonction de diffusion partielle du substrat PJ9, celle du chaperon GroEL ainsi que le terme de corrélation croisée PJ9/GroEL. Sur la Fig. 7 sont représentées les fonctions de diffusion dues aux autocorrélations de chaque constituant. Les fonctions de diffusion partielles de GroEL et srGroEL correspondent à celle d'un cylindre creux de longueur finie. Chaque anneau de GroEL fixe une molécule de substrat à l'une des extrémités du cylindre. Ceci est à l'origine d'un pic de corrélation sur la fonction de diffusion partielle de PJ9 (figure gauche) qui est absent de la fonction de diffusion mesurée dans le complexe impliquant srGroEL (figure droite). La distance déduite de la position de ce pic mesure la façon dont le substrat pénètre à l'intérieur du cylindre que forme GroEL.

### 3.7 Exemple 7 : Triangulation et structure quaternaire des sous-unités ribosomales

Le ribosome est la machinerie cellulaire responsable de la traduction de l'information génétique en protéine. C'est un complexe macromoléculaire formé de plusieurs chaînes d'ARNr et d'un très grand nombre de protéines (jusqu'à 82 chez les eucaryotes). Comme très souvent en biologie, les données structurales sont ici primordiales à une compréhension fonctionnelle. Le ribosome est formé de 2 sous-unités identifiées par leur coefficient de sédimentation égaux respectivement à 30 et 50S (Svedberg) chez les procaryotes.

En utilisant un marquage sélectif de chacune des 21 protéines de la sous-unité 30S d'*E. Coli*, l'équipe de Brookhaven [19] a pu mesurer de façon systématique le minimum de distances entre paires (cf. Eq. 3.7) permettant par triangulation de déterminer les positions relatives de ces protéines dans la sous-unité ribosomale. Ce travail de longue haleine a pu aboutir à une véritable cartographie de sa structure quaternaire (cf. Fig. 8). Plus récemment une étude de la sous-unité 50S a été

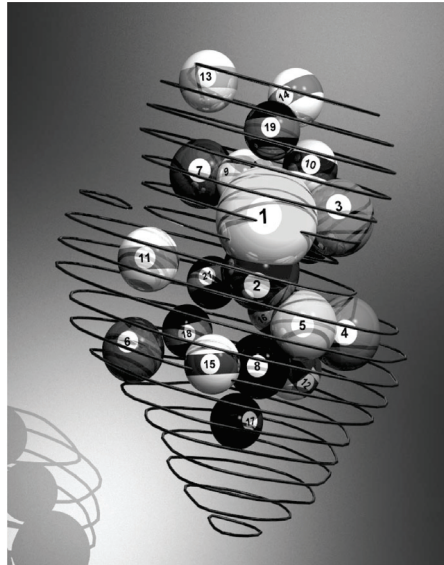


Figure 2 The placement of proteins within the 30S subunit. Proteins are shown as spheres the volumes of which are proportional to molecular weight. Their placement was determined by neutron scattering (10). The protein array is superimposed on the outline of the 30S subunit in a manner that maximizes its overlap with epitope positions determined by immunoelectron microscopy. This figure was provided by Malcolm Capel.

**Figure 8.** Structure tridimensionnelle de la sous-unité 30S du ribosome d'*E. Coli* obtenue par DNPA et marquage sélectif [19]. Figure reproduite d'après la réf. [20].

entreprise par l'équipe de Geesthacht [21]. Pour limiter le nombre de marquages à réaliser, les auteurs utilisent le marquage sélectif en liaison avec la variation de contraste par polarisation nucléaire (cf. partie 3.5).

Nous remercions Sophie Combet-Jeancenel pour sa relecture critique et attentive.

## Références

- [1] D. Lairez. Résolution d'un spectromètre de diffusion de neutrons aux petits angles. *J. Physique IV France*, 9: 67–81, 1999.
- [2] O. Glatter. A new method for the evaluation of small angle scattering data. *J. Appl. Cryst.*, 10: 415–421, 1977.
- [3] J.-P. Hansen and I. McDonald. *Theory of simple liquids*. Academic Press London, 1976.
- [4] C. Loupiac, M. Bonetti, S. Pin, and P. Calmettes. High-pressure effects on horse heart metmyoglobin studied by small-angle neutron scattering. *Eur. J. Biochem.*, 269: 4731–4737, 2002.
- [5] O.D. Velev, E.W. Kaler, and A.M. Lenhoff. Protein interactions in solution characterized by light and neutron scattering: Comparison of lysozyme and chymotrypsinogen. *Biophys. J.*, 75: 2682–2697, 1998.
- [6] A. Guinier and G. Fournet. *Small-angle scattering of X-rays*. J. Wiley & sons, 1955.
- [7] B.J. Berne and R. Pecora. *Dynamic light scattering*. J. Wiley & Sons, New York, 1976.
- [8] J.C. Higgins and H.C. Benoît. *Polymers and neutron scattering*. Clarendon Press, 1994.
- [9] P.-G. de Gennes. *Scaling concepts in polymer physics*. Cornell Univ. Press, 1996.
- [10] L. Auvray and P. Auroy. Scattering by interfaces. In P. Linder and Th. Zemb, editors, *Neutron, X-ray and light scattering.*, pages 199–221. North Holland Publishing, 1994.

- [11] T.V. Budkevich, A. A. Timchenko, E.I. Tiktopulo, B.S. Negrutskii, V.F. Shalak, Z.M. Petrushenko, V.L. Aksenov, R. Willumeit, J. Kohlbrecher, N. Serdyuk, and A.V. El'skaya. Extended conformation of mammalian translation elongation factor 1a in solution. *Biochemistry*, 41: 15342–15349, 2002.
- [12] J. Pelta, H. Berry, G.C. Fadda, E. Pauthe, and D. Lairez. Statistical conformation of human plasma fibronectin. *Biochemistry*, 39: 5146–5154, 2000.
- [13] D. Lairez, E. Pauthe, and J. Pelta. Refolding of a high molecular weight protein: salt effect on collapse. *Biophys. J.*, 84: 3904–3916, 2003.
- [14] M. Hammermann, N. Brun, K.V. Klenin, R. May, K. Tóth, and J. Langowski. Salt-dependent dna superhelix diameter studied by small angle neutron scattering measurements and monte carlo simulations. *Biophys. J.*, 75: 3057–3063, 1998.
- [15] R. May. Label triangulation. In P. Linder and Th. Zemb, editors, *Neutron, X-ray and light scattering.*, pages 119–133. North Holland Publishing, 1994.
- [16] J.-P. Cotton. Variations on contrast in sans: determination of self and distinct correlation functions. *Adv. Colloid Interf. Sci.*, 69: 1–29, 1996.
- [17] H. Glättli. Polarized nuclei and spin-dependent scattering. *Physica B*, 297: 273–276, 2001.
- [18] S. Krueger, S.K. Gregurick, J. Zondlo, and E. Eisenstein. Interaction of groel and groel/groes complexes with a nonnative subtilisin variant: a small-angle neutron scattering study. *J. Struct. Biol.*, 141: 240–258, 2003.
- [19] M.S. Capel, D.M. Engelman, B.R. Freeborn, M. Kjeldgaard, J.A. Langer, V. Ramakrishnan, D.G. Schindler, D.K. Schneider, B.P. Schoenborn, I.-Y. Sillers, S. Yabuki, and P.B. Moore. A complete mapping of the proteins in the small ribosomal subunit of escherichia coli. *Science*, 238: 1403–1406, 1987.
- [20] P.B. Moore. The three dimensional structure of the ribosome and its components. *Annu. Rev. Biophys. Biomol. Struct.*, 27: 35–58, 1998.
- [21] R. Willumeit, G. Diedrich, S. Forthmann, J. Beckmann, R.P. May, H.B. Stuhmann, and K.H. Nierhaus. Mapping proteins of the 50s subunit from escherichia coli ribosomes. *Biochem. Biophys. Acta*, 1520: 7–20, 2001.