STRUCTURAL CHARACTERISATION AND MEMBRANE INSERTION OF M13 PROCOAT, M13 COAT AND Pf3 COAT PROTEINS

THÈSE N° 1468 (1996)

PRÉSENTÉE AU DÉPARTEMENT DE CHIMIE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Margono SOEKARJO

Chimiste, diplômé de l'Université d'Utrecht, Pays-Bas
de nationalité hollandaise

acceptée sur proposition du jury:

Prof. H. Vogel, directeur de thèse
Prof. M. A. Hemminga, corapporteur
Prof. A. Kuhn, corapporteur
Prof. M. Mutter, corapporteur
Prof. U. von Stockar, corapporteur

Lausanne, EPFL
1996

voor mijn ouders

1. ABBREVIATIONS AND SYMBOLS

2YT	doubly concentrated yeast trypton medium
ATR	attenuated total reflection
BP	base peak
c_{as}	concentration of lipid-associated protein molecules
CD	circular dichroism
c_p	total protein concentration
ΔG^o	free energy change
DMPC	1,2-dimyristoyl-sn-3-phosphatidylcholine
DOC	desoxycholate
DPH	1,6-diphenyl-1,3,5-hexatriene
DPPC	1,2-dipalmitoyl-sn-3-phosphatidylcholine
DPPS	1,2-dipalmitoyl-sn-3-phosphatidylserine
ε	molar extinction coefficient
EDTA	ethylene-diaminetetraacetic acid
ES-MS	electrospray mass spectroscopy
ESR	electron spin resonance
F	fluorescence intensity
FD-MF	freeze-dried mixed film reconstitution method
FH	Flory-Huggins
FTIR	Fourier-transform infrared
F_∞	limiting fluorescence intensity
g#p	product of gene #
Γ_{app}	apparent partition coefficient
GdHCl	guanidine hydrochloride
γ_{MF}	mole-fraction partition coefficient
H5	Ser^{-3} -> Phe^{-3} mutation in M13 procoat protein
HPLC	high performance liquid chromatography
κ	orientational factor
L/P	lipid to protein molar ratio
L_α	disordered, or lamellar liquid crystalline phase
L_β	ordered, or lamellar gel phase
m/z	mass to charge ratio
MALDI-MS	matrix assisted laser desorption mass spectroscopy

MD	molecular dynamics
MF	mixed film reconstitution method
MS	mass spectroscopy
MW	molecular weight
NMR	nuclear magnetic resonance
NMWL	nominal molecular weight limit
OM30R	Val^{30} -> Arg^{30} mutation in M13 procoat protein
OPE	organic phase evaporation reconstitution method
PA	phosphatidic acid
PFV	preformed vesicles reconstitution method
POPC	1-palmitoyl-2-oleoyl-sn-3-phosphatidylcholine
POPG	1-palmitoyl-2-oleoyl-sn-3-phosphatidylglycerol
r	molar ratio of lipid-associated protein molecules per lipid molecule.
R_{ATR}	ATR dichroic ratio
R_{ATR}	dichroic ratio
rms	root mean square
RPC	reversed phase chromatography
S	order parameter
S_α	orientational order parameter of the α-helix
SASL	stearic acid spin label
SDS	sodiumdodecylsulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
$S_{M\alpha}$	orientational order parameter of the amide I helix transition moment
S_{mol}	orientational order parameter of the molecular director
ssDNA	single stranded DNA
SUV	small unilamellar vesicles
S_v	vibrational transition moment
TCA	trichloroacetic acid
TFA	trifluoroacetic acid
T_m	lipid phase-transition temperature
TPCK	L-1-tosylamide-2-phenylethyl-chloromethyl-ketone
V_W	molar water volume
w/v	weight per volume
Ψ_o	membrane surface potential
z_{eff}	effective charge

Colofon

This thesis was written in MS Word with an EndNotePlus plug-in module for biblographic reference handling. Figures were digitized on a scanner or by means of a video camera and frame grabber. If necessary contrast and brightness enhancement, recropping and resizing was performed with Adobe Photoshop. Line art in MacDrawPro, graphs and curve-fittng in IgorPro. Most of the work was done on a Macintosh PowerBook 180c.

2. TABLE OF CONTENTS

3. Version Abrégée

Les protéines membranaires remplissent de nombreuses fonctions importantes dans la membrane biologique. Leur incorporation et leur structure, qui sont étroitement liées à leur fonction, ne sont pas encore bien comprises. Nous étudions, à titre d'exemple, trois protéines reconstituées dans des bicouches lipidiques (servant de modèle pour la membrane biologique): les protéines capsidiques (coat) des phages Pf3 et M13 et le précurseur de la dernière, la protéine M13 procoat, qui traverse la membrane deux fois et sert de modèle pour des protéines comportantes plusieurs segments intramembranairs. La Pf3 coat s'insère dans la membrane dans sa forme mature, la M13 au contraire a besoin d'une séquence signal pour s'insérer.

Une nouvelle méthode simple et efficace a été développé pour la purification des protéines M13 procoat, M13 coat and Pf3 coat. Des préparations homogènes ont été obtenues en isopropanol/0.1% TFA, dans lequel la protéine M13 coat est dissoute sous forme monomérique et les autres protéines sous forme de dimères. Les conformations de ces protéines dans des environnements différents ont été déterminées par dichroïsme circulaire et spectroscopie infrarouge. En solution organique, les protéines adoptent une conformation avec un taux moyen d'hélice de 90%. Dans des bicouches lipidiques constituées de phosphatidylcholine et phosphatidylglycerol, le taux moyen d'hélice atteint 50% pour la protéine M13 procoat, 60% pour M13 coat et 75% pour Pf3 coat. Le paramètre d'ordre d'orientation S_α des hélices a été déterminé par des mesures d'infrarouge sur des échantillons orientés dans la gamme de spectre de l'amide I. Les hélices sont orientées de préférence parallèlement à la normale de la membrane, avec $S_\alpha = 0.63$ pour M13 procoat, $S_\alpha = 0.58$ pour Pf3 coat et une valeur nettement plus élevée de $S_\alpha = 0.81$ pour M13.

La topologie des protéines M13 et Pf3 coat dans des vésicules lipidiques a été déterminée par proteolyse limitée, combinée avec de la spectroscopie de masse pour l'analyse des fragments lytiques. La partie C-terminale (à partir du résidu 15 (Pf3) ou 16 (M13) était inaccessible à la protéinase K.

Pour la première fois le changement d'énergie libre standard, $\Delta G^{o\prime}$, d'une protéine qui s'insère dans la membrane à l'aide d'une séquence signal a été déter-

miné, prenant la protéine M13 procoat comme exemple. Le coefficient de partage entre la phase aqueuse et la phase membranaire de vésicules de lipides préformées donne une valeur de $\Gamma = 6.5 \times 10^5\ M^{-1}$, correspondant à une ΔG^o of -10.4 kcal/mole, basé sur des mesures de transfert d'énergie de fluorescence entre le tryptophan intrinsèque et une bicouche lipidique de POPC marquée convenablement. A titre de comparaison, le coefficient de partage de M13 coat est déterminé à la valeur nettement inférieure de $\Gamma = 1 \times 10^5\ M^{-1}$ (ΔG^o = -9.3 kcal/mole). Des expériences de digestion par protéinase K ont été faites, montrant que 20% de la protéine procoat liée aux vésicules lipidiques s'intègre spontanément dans une forme transmembranaire, tandis que 80% reste inséré dans la région interfaciale de la membrane. En combinant ces résultats on peut estimer la limite supérieure pour le changement d'énergie libre de l'insertion transmembranaire de M13 procoat égale à -14,8 kcal/mole. Afin de distinguer la contribution de l'insertion dans la région interfaciale de la membrane de celle de l'insertion transmembranaire, le coefficient de partage de la protéine procoat mutante OM30R (qui contient un acide aminé chargé dans la séquence hydrophobe de la partie mature (échange d'un résidu Val pour un résidu Arg à la position 30)) a été déterminé, donnant $\Gamma = 0.3 \times 10^5\ M^{-1}$ (ΔG^o = -8.6 kcal/mol). Des expériences *in vivo*, publiées antérieurement, ont démontré que la OM30R n'est pas transbahutée à travers la membrane intérieure de *Escherichia coli*, mais reste liée à la surface intérieure. Les résultats présentés indiquent que, malgré le fait que l'insertion dans la région interfaciale de M13 procoat apporte la majeure partie de ΔG^o, c'est le gain d'énergie final de l'interaction des portions hydrophobes de la pré-protéine pliée avec les chaînes lipidiques qui pousse l'insertion transmembranaire de la protéine M13 procoat. Ni la séquence signal, ni la protéine coat mature ne donnent seules ce gain d'énergie libre. Pour les protéines étudiées ici, l'insertion spontanée dans la membrane n'arrive que dans le cas de membranes lipidiques fluides et jamais dans le cas de membranes en phase lipidique cristalline. En plus, utilisant des bicouches lipidiques avec des charges de surface de membrane négatives, il a été démontré que la protéine procoat ainsi que la protéine coat sont attirées électrostatiquement par la surface de la membrane lipidique, même très faiblement, avec des coefficients de partage apparents dans la même ordre de grandeur que pour la membrane lipidique phosphatidylcholine.

Ce travail a donné de nouvelles informations sur la structure et l'insertion des protéines manteau des phages. Les nouvelles méthodes dévelopées pourraient être applicables de manière générale dans l'étude des dynamiques structurelles de l'incorporation des protéines membranaires plus complexes.

4. Summary

Membrane proteins fulfill many central functions in the biological membrane. The insertion process of these proteins and their structure, which are intimately linked to their function, are not yet well understood. As a model we studied three proteins reconstituted into lipid bilayers (which serve as a model for the biological membrane): the major coat proteins of phages Pf3 and M13 and the precursor of the latter, M13 procoat protein, which spans the inner membrane two times and serves as a model for a multi-spanning membrane protein. Whereas Pf3 coat protein can insert into the membrane in its mature form, M13 coat protein needs the signal sequence present in the precursor to insert.

A new, simple and efficient purification method has been developed for M13 procoat, M13 and Pf3 coat proteins. Homogenous preparations were obtained in isopropanol/0.1% TFA, where M13 coat protein is found to be dissolved in a monomeric form, and the two other proteins as dimers. The conformations of these particular proteins in different environments have been determined by circular dichroism and infrared spectroscopy. In organic solvents, the proteins adopt a conformation with an average helix content of 90%. In lipid bilayers composed of phospatidylcholine and phosphatidylglycerol lipids, the average helix content is 50% for M13 procoat protein, 60% for M13 coat protein and 75% for Pf3 coat protein. In order to characterize the orientational distribution the orientational order parameter S_α of the protein helices in planar lipid bilayers have been determined by polarized infrared measurements in the amide I spectral range. The helices of the three proteins are oriented preferentially parallel to the membrane normal, with $S_\alpha = 0.63$ for M13 procoat protein, $S_\alpha = 0.58$ for Pf3 coat protein and a distinctly higher value of $S_\alpha = 0.81$ for M13 coat protein.

The topology of M13 and Pf3 coat proteins was investigated using limited proteolysis in lipid vesicles with proteinase K combined with mass spectroscopy to analyze the lytic fragments. The C-terminal part of both proteins was found to be inaccessible to proteinase K, only 14 (Pf3) or 15 (M13) residues on the N-terminal were protease accessible from the outside of the vesicles.

Furthermore, the standard free energy change, ΔG^o, of a membrane-inserting protein with a leader sequence has been determined experimentally for the first

time, using M13 procoat protein as an example. The partition coefficient for the distribution of this protein between the aqueous phase and the membrane phase of preformed lipid vesicles yielded a value of $\Gamma = 6.5 \times 10^5$ M^{-1}, corresponding to a ΔG^0 of -10.4 kcal/mol, based on measurements of the fluorescence energy transfer between the intrinsic tryptophan of the protein and a suitably labeled lipid membrane of POPC. For comparison, the partition coefficient of the M13 coat protein between the aqueous and the POPC lipid bilayer phase was determined to be distinctly lower: $\Gamma = 1 \times 10^5$ M^{-1} (ΔG^0 = -9.3 kcal/mol). Proteinase K digestion experiments have been performed, showing that 20% of the procoat protein bound to lipid vesicles spontaneously integrate in a transbilayer form, whereas 80% remain inserted in the interfacial membrane region. By taking together these results, an upper limit for the free energy change of the transmembrane insertion of procoat protein was estimated to be -14.8 kcal/mol. In order to distinguish further the contribution arising from insertion of the procoat protein into the membrane interfacial region from that due to transmembrane insertion, the partition coefficient of a mutant procoat protein OM30R (which contains a positively charged amino acid in its mature hydrophobic segment (exchange of a Val to an Arg residue at position 30)) was determined, yielding $\Gamma = 0.3 \times 10^5$ M^{-1} (ΔG^0 = -8.6 kcal/mol). Previously reported *in vivo* experiments have shown that the OM30R mutant protein is not translocated across *Escherichia coli* inner membrane, but only binds to the inner surface. The results presented here indicate that although the insertion of the procoat protein into the interfacial region of the lipid bilayer contributes the major part to ΔG^0, it is the final energy gain of the interaction of the hydrophobic portions of the folded pre-protein with the lipid chains which drives the transmembrane insertion of the M13 procoat protein. Neither the leader sequence nor the mature coat protein alone yields this free energy gain. For the different proteins investigated here, spontaneous membrane insertion occurs only for fluid lipid bilayers, but not for membranes in the crystalline lipid phase. Furthermore, by using lipid bilayers with negative membrane surface charges, it was shown that both procoat and coat proteins are electrostatically attracted to the surface of the lipid membrane, though only to a small extent, with apparent partition coefficients of the same order of magnitude as for the phosphatidylcholine lipid membrane.

This work has yielded new information about the structure and the insertion of the phage coat proteins. The new methods developed in this work might be of general applicability to study the structural dynamics of membrane protein assembly.

5. General Introduction

This thesis contributes to the knowledge of the structure and the insertion of membrane proteins and of the life cycle of filamentous bacteriophages. In this chapter these topics will be introduced to the reader. Furthermore the scope and the organization of this thesis is discussed.

5.1. STRUCTURE AND DYNAMICS OF BIOLOGICAL MEMBRANES

Lipids are the major structural constituents of biological membranes. They arrange themselves in a bilayer structure forming a barrier between two cellular compartments or between a cell and the exterior. A second important component are proteins, which can either be bound to the membrane surface, mainly by electrostatic forces (peripheral membrane proteins), or which can traverse the membrane (integral membrane proteins). Membrane proteins play both a structural and a functional role. They stabilize the membrane and modulate the barrier function in making the membrane selectively permeable to specific substrates. Other proteins play a role in regulating this permeability as a function of cell activity, metabolic state and outside environment or play a role in keeping the membrane energized.

The dielectrical profile of the membrane changes from a charged surface via a polar region to a hydrophobic core. This can be important in "guiding" certain substrates to a specific depth in the membrane and thus restricting them in the plane of the membrane. Specific interaction with another molecule in this restricted layer will be much more efficient since the two molecules will meet by two dimensional diffusion instead of three dimensional diffusion. This phenomenon of "membrane catalysis" by restricting molecular interactions to two dimensions [1,2], might be important for the insertion process of membrane proteins and for enzymatic reactions, e.g. the removal of the signal sequence of M13 procoat protein by the membrane-bound enzyme leader peptidase [3-5]. For M13 procoat protein it is assumed that the first step in the insertion process is an electrostatic binding of its positively charged termini to negatively charged phospholipids (the different stages in the life cycle of phages are described in detail in section 5.3.2). In the assembly process of filamentous bacteriophages the membrane plays

a role in orienting the coat protein molecules in such a way that they can tightly pack in a parallel way (see also section 5.3.2.3).

5.1.1. MOLECULAR INTERACTIONS

Molecules can interact through hydrophobic forces, electrostatic forces, or both. Four classes of intermolecular interaction can be distinguished in a simple model of a biological membrane including only lipids and proteins.

5.1.1.1. Lipid-lipid interaction [6, p. 36]

The biologically most relevant self-assembled lipid structure is that of a bilayer. Most lipids will, upon hydration, spontaneously form multilamellar vesicles consisting of several concentric vesicles with a lamellar organization.

Several energetic contributions make a bilayer organization thermodynamically stable. The lipid tails are aligned, resulting in an optimization of hydrophobic interactions. Only the polar headgroups are exposed to the surrounding aqueous environment. Polar interactions in the head group region are the other factor contributing to the stability of the bilayer.

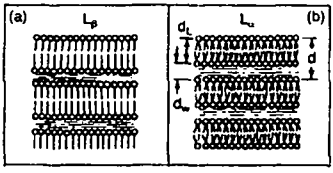

Fig. 5.1. Lipid organization in ordered and in disordered phase. A: disordered or lamellar liquid crystalline phase, L_α. B: ordered or lamellar gel phase, L_β.

The degree of organization in the bilayer differs with its physical state (Fig. 5.1 and 5.2). The major part of the lipids in biological membranes is thought to be in the disordered or lamellar liquid crystalline phase, L_α (Fig. 5.1A). In this phase there is considerable disorder in the acyl chains. Below the phase transition temperature, T_m, there is considerably higher order in the so-called ordered or lamellar gel phase, L_β (Fig 5.1B). In this phase the lipid acyl chains are fully extended in an all-trans conformation (Fig. 5.2A). Hence, the lipids are packed more tightly together (smaller surface area per molecule), the acyl chains are much more ordered (Fig 5.2B) and the bilayer is thicker than in the liquid crystalline phase. Also the positional order of the lipid molecules is higher in the ordered phase.

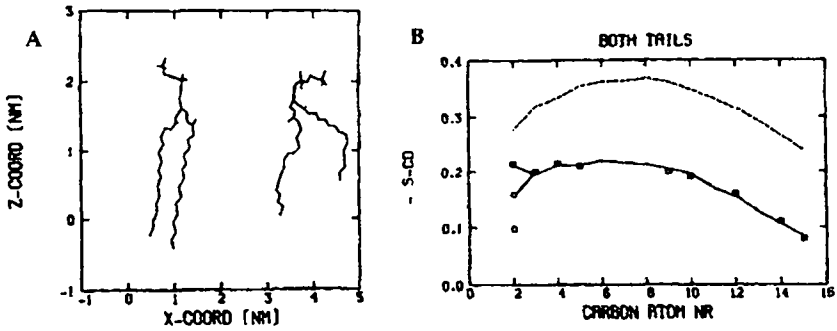

Fig. 5.2. Molecular dynamics calculations of lipid structure in ordered and in disordered phase. A: acyl chain conformation of one molecule at a particular time in ordered Lβ (left) and disordered Lα (right) phase, B: order parameters[1] -S_{CD} of CD_2 groups as a function of the position in palmitoyl lipid chains in the Lα phase (*solid line*) and in the Lβ phase (*dashed line*). Squares represent NMR data from experiments on selectively deuterated DPPC. Figure taken from [7].

The primary function of biological membranes is to form a selective barrier between cellular compartments where incompatible processes take place, or between the cell and the exterior. A lipid bilayer provides an efficient energetic barrier for the passage of water-soluble species because of its hydrophobic core and its two surface layers which present an array of fixed charges and polar groups (see Fig. 5.3).

[1]The formal definition of the order parameter tensor S is: $S_{ij} = 1/2 < 3\cos\Theta_i \cos\Theta_j - \delta_{ij} >$, in which Θ_i represents the angle between the molecular axis i and the bilayer normal (z-axis). The brackets denote an ensemble average. An S_{zz} that equals one indicates full order along the bilayer normal, -1/2 full order perpendicular to the normal, and 0 an isotropic distribution. The order parameter S_{CD} from deuterium NMR is defined along the C-D bond. It can be directly related to the MD order parameters through the expression $S_{CD} = 2/3\ S_{xx} + 1/3\ S_{yy}$.

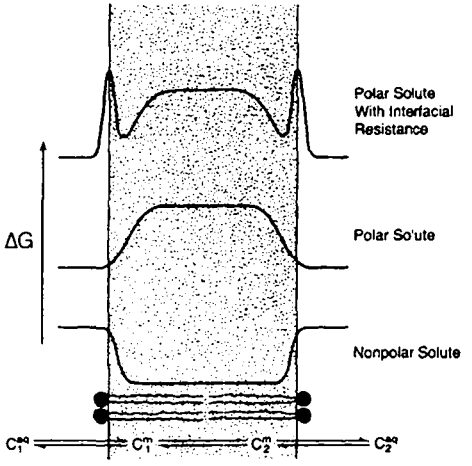

Fig. 5.3. Barrier function of lipid membranes: schematic representation of the shapes of the free energy barriers for polar and non polar solutes crossing a lipid bilayer membrane. The solute concentrations at locations in the membrane (m) and aqueous (aq) phases are also indicated. Figure taken from [6].

5.1.1.2. Lipid-protein hydrophobic interactions

The acyl side chains of lipids are highly hydrophobic. In consequence, those parts of proteins which are embedded in the hydrophobic core of the membrane will have to present a hydrophobic surface to the surrounding lipids in order to be thermodynamically stable. Structural consequences for integral membrane proteins will be discussed in section 5.1.2.

5.1.1.3. Lipid-protein electrostatic interactions

The lipid head groups are hydrophilic moieties containing polar groups (the ester bonds in the sn-1 and 2-position) and fixed charges (phosphates, (*t*-) amines, carboxylates, etc.).

Peripheral proteins can be bound to the membrane surface in different ways: (*i*) via electrostatic forces either to the polar headgroups of the lipids or to the surface-accessible part of integral membrane proteins, which in this case would act as anchor proteins, (*ii*) via hydrophobic anchor(s) in the form of covalently attached fatty acids or lipid molecules (see Fig. 5.4).

Integral membrane proteins may carry charged, or non-charged polar amino acid residues at both ends of the membrane-traversing stretches. These polar parts might interact with the lipid head groups or with other membrane proteins.

Their function could be to keep the protein in a suitable position with respect to the membrane.

5.1.1.4. Protein-protein interactions

Membrane proteins can be comprised of functional oligomeric complexes with other proteins (hetero-oligomers) as for example the acetylcholine receptor AchR [8], or with other molecules of the same kind (homo-oligomers). Examples of the latter are the trimeric *E. coli* porins OmpF and PhoE, membrane channels which play a role in diffusion of small hydrophilic nutrients and waste products over the outer membrane [9]. Both hydrophobic interaction and electrostatic interaction can play a role in subunit interactions.

5.1.2. STRUCTURE OF INTEGRAL MEMBRANE PROTEINS

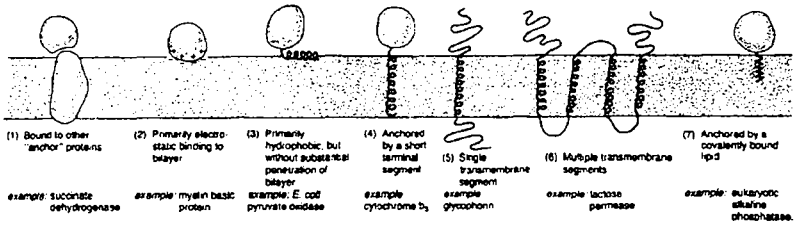

Fig. 5.4. Classification of membrane proteins. Figure taken from [6].

A major goal in modern biology is to establish the relationship between the primary sequence of proteins and their folded structure, which is essential for its proper function. This task might seem simpler in membrane proteins than in water-soluble proteins because of the constraints the membrane imposes on the secondary structure. The dielectrical profile of the membrane (see Fig. 5.3) demands a very specific protein surface structure. On the other hand this special surface makes them more difficult to handle as they have a marked tendency to denature and aggregate when removed from their natural environment during purification. These difficulties are illustrated by the fact that the structure of only a few membrane proteins has been resolved at high resolution (*e.g.*, the photosynthetic reaction center [10,11] , bacteriorhodopsin [12] and porins from the bacterial outer membrane [9,13,14]) in contrast to a multitude of water-soluble proteins.

The most important restriction imposed by the membrane is the fact that the principally polar peptide bonds have to be made hydrophobic, for those parts of the protein backbone to be buried in the hydrophobic core, through involvement in hydrogen bonding structures. This requirement is met both in α-helices and in β-sheets/β-barrels (see Fig. 5.5). Although in some instances β–type structure is present (in the major outer membrane proteins of *E. coli* [9,13,14]), the membrane spanning domains seem to be α-helical in the majority of integral membrane proteins (e.g. all the inner membrane proteins of *E. coli* of which the structure has been resolved) [13]. This might be related to the fact that in an α-helix the requirement of hydrogen bonding is met within the helix itself, whereas a β-strand is only stable through interactions with neighboring β-strands. This can be achieved either in large β-sheets (of which the biological relevance yet has to be proven) where border effects are less important or in β-barrels, which, because of their cylindrical topology, have no borders.

For small, single-spanning membrane proteins like the major coat proteins the only way to form β-type structures which are thermodynamically stable is in large aggregates where the hydrogen bonding requirement is met through intermolecular interaction of residues in neighboring aligned molecules.

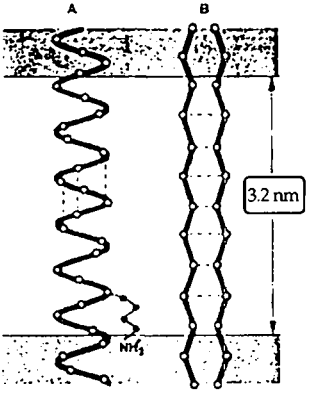

Fig. 5.5. Basic conformations of an integral protein (M13 coat protein) traversing the hydrophobic core of a phospholipid bilayer[2]. (A) α-helix, (B) β-sheet. The open circles schematically represent the C_α-CO-NH moiety of the polypeptide backbone. Within the hydrophobic core of the lipid bilayer the amino acid side chains must be hydrophobic in character. In general, charged side chains should not be in direct contact with the hydrophobic lipid chains, except near the core surface, where a charged residue may have its charged group outside, as illustrated for a lysyl residue. The dashed lines schematically indicate intra-molecular (A) or inter-molecular (B) hydrogen bonds. *The phospholipid head group region is shaded. Figure* from Tanford [16], adapted by Hemminga[17].

Helical membrane-spanning domains generally consist of a stretch of hydrophobic amino acids. Well known examples are bacteriorhodopsin [12,18] and the bacte-

[2]The shaded area represents only the region of fixed charges in the lipid headgroups, not the whole polar interfacial region. See also [15].

rial photosynthetic reaction center [11]. The α-helix provides a structure which has two favorable energetic aspects. The hydrophobic amino acid side chains protrude into the surrounding hydrophobic environment formed by the lipidic fatty acid tails, whereas the backbone is shielded in the core of the helix and stabilized by internal hydrogen bonding [19].

The loops between the membrane-spanning segments may play an important role in the folding of multi-spanning membrane proteins by providing the correct charge distribution needed for vectorial insertion of the protein into the membrane (as formulated in the positive-inside rule of von Heijne [3]), although this could be different for each individual protein. Folding of separate peptides containing the transmembrane segments can be the key to the functional structure of a membrane protein as was shown for bacteriorhodopsin. This protein can be refolded into the native configuration from two independently stable transmembrane helices and the complementary five-helix fragment [20], although it lacks two loop regions in this experiment.

Thus, the formation of helical membrane protein structure can be modeled [20,21] as comprising two separate stages with their own energetic contribution to the stability of the final structure: (a) the formation of individually stable membrane-spanning α-helices, dictated by the low dielectrical constant of the lipid environment; and (b) the translation of the individual helices into a tertiary arrangement without major changes in secondary structure, driven by inter helix interaction [22], the lipophobic and packing effects of the buried helices and the constraints of the extramembraneous portions of the protein. The contact zones between the membrane-spanning parts of the protein may be either specific or unspecific.

Combining these principles with the positive inside rule of von Heijne [3] the topology and folding of most bacterial membrane proteins can be predicted correctly.

The insertion of M13 procoat protein provides us with an attractive and simple model system to investigate the contributions of different transmembrane segments and hydrophilic domains to the insertion process and the stability of integral membrane proteins.

5.2. PROTEIN MEMBRANE INSERTION / EXPORT

Protein integration into or translocation across membranes is a fundamental process in prokaryotic and eukaryotic cells. The biosynthesis of almost all cellular proteins begins in the cytoplasm. During or after synthesis, non-cytoplasmatic proteins are targeted to an appropriate organelle, thereby either inserting into or crossing one or several membranes. Newly synthesized secreted or integral membrane proteins carry specific regions (designated as leader, signal or stop-transfer sequences) that play a functional role in the membrane insertion / transfer process [23,24]. In eukaryotic cells, protein synthesis and membrane transport are often synchronously linked by a complicated signal-receptor mechanism [25,26]. However, small proteins insert into membranes independently of complex translocation machineries [27-29]. This suggests that such small proteins must have the structural properties necessary for targeting and translocation within their primary amino acid sequence.

5.2.1. *E.COLI* EXPORT MACHINERY

The best studied prokaryotic organism is the gram-negative *Escherichia coli* bacteria. In *E. coli* all proteins are synthesized on ribosomes in the cytoplasm. However, a number of them have to be exported from this aqueous compartment to their final functional location in the inner membrane, in the periplasmatic space, in the outer membrane or even in the surrounding environment (e.g. colicins, for review see [30]). In this process, the proteins have to be correctly targeted to one or more membranes, inserted into and translocated across these membranes. The elucidation of which factors are in each of the separate steps of this process is an important theme in molecular biology.

Most precursors of these proteins are synthesized with a N-terminal extension of variable length (about 20 amino acids), the so-called signal or leader sequence which is essential for translocation over the cytoplasmatic membrane [31,32] and which is cleaved after translocation by leader peptidase, a specific protease residing in the cytoplasmatic membrane with its catalytic domain located in the periplasm. This process is very efficient and rapid under normal circumstances. Thus it is nontrivial to isolate these precursor proteins in chemical amounts and to investigate them. M13 procoat protein was the first peptide isolated with an intact leader sequence. In some cases the information necessary for targeting / insertion of the newly synthesized proteins is present in the form of an internal signal sequence which remains in the functional protein [33].

Moreover, the general export pathway for proteins in *E. coli* involves a number of different secretory (Sec) proteins. After or during synthesis on free ribosomes in the cytosol, some proteins which are to be exported form a complex with a chaperone protein, the cytosolic SecB, keeping it in a loosely folded configuration, necessary for translocation. This complex binds to the dimeric functional form of SecA, which is found both in the cytosol and peripheric to the inner membrane, where it may physically interact with two other components of the Sec export pathway, SecE and SecY [34,35]. Apart from its dual affinity to the SecB/precursor complex and to the inner membrane it also contains an ATPase activity which is essential for translocation [36]. Finally, SecE and SecY, two intrinsic membrane proteins, are involved in the translocation step, but their exact function is not yet clear.

Most exported proteins follow this general pathway and are called Sec-dependent. Bacteriophage coat proteins are a notable exception. The fact that apparently all information needed for targeting and translocation is present in their primary sequence makes them an intriguing model molecule to study these processes. It is probably their small size which makes the presence of helper proteins during translocation unnecessary. Interestingly, the translocation process becomes Sec-dependent in mutant procoat proteins where a longer stretch of amino acids was inserted by genetic engeneering between the signal sequence and the mature part of M13 coat protein[37]. Insertion of intermediate stretches into another protein, leader peptidase (Lep), resulted in a loss of translocation efficiency [38], suggesting that the loop region is too long for spontaneous insertion, yet too short for efficient interaction with the Sec components.

5.3. BACTERIOPHAGE STRUCTURE AND LIFE CYCLE

5.3.1. PHAGE STRUCTURE

Filamentous bacteriophages are large nucleoprotein assemblies [39,40], which rely for their reproduction on the protein synthesis machinery of the host bacteria. Their structural and functional data are reviewed in detail elsewhere [41-43]. The main findings will be presented hereafter.

The best studied class of bacteriophages are the Ff coliphages, fd, f1 and M13, which specifically infect *E. coli* male strain by interaction with the F (sex) pili. Their DNA sequences are identical except for a small number of mutations and therefore, they might be considered as slight modifications of basically the same

phage. The major coat protein is identical for fd and f1, whereas M13 coat protein has a single mutation of aspartic acid to asparagine at position 12.

Most data available were obtained for coliphage M13, the nucleotide of which is extensively used as a cloning vector in molecular biology. The following discussion will focus on M13 and the other Ff phages, fd and f1. If not specified, information is referring to this class of phages. Composition and structure of the different phages are compared in TABLE 5.1.

TABLE 5.1. Composition and physical properties of bacteriophages

phages	virion		major coat protein, g8p			DNA			bacterial host [44]
	length [nm]	diameter [nm]	number of subunits	residues per subunit	MW of subunits	Wt % DNA	number of nucleotides	Nucleotides per subunit	
M13 [45]	860	6-10 [1]	2700	50 [48,49]	5262	12	6407	2.4[2]	*E. coli*, male strain
f1 [51,52]			1900[53]	50	5240[53]				*E. coli*, male strain
fd [54,55]	890±30 [56]	5 [57]	2710±110 [56]	50 [49,58]	5240 [58]	12.0±0.3 [59-61]	6408 [62]	2.4	*E. coli*, male strain
Xf	980±40 [56]		3670±190 [56]	44 [63]	4343 [63]	12.6±0.3 [56]		2.0	*Xanthomonas oryzae*[3]
Pf1	1940±90[56]		7620±140 [56]	46 [64]	4609 [64]	6.4±0.3 [56]	6690 ± 450 [65]	1.0 [66]	*Pseudomonas aeruginosa*, strain K
Pf3	720±50 [56]		2630±160[67]	44 [68]	4632 [68]	14±1 [69,70]	5960±90 [68,69]	2.4	*Pseudomonas aeruginosa*, strain PAO1
If1	1300 [71]		-	51 [1]	5278 [44]	-		-	*E. coli*, I-plasmid-containing strains
IKe	1100[71]		-	53 [44]	5695 [44]	-		-	*E. coli*, I-plasmid-containing strains
Phi-LF[72]					4100 [4]		1018		*Xanthomonas-campestris PV campestris*

[1] Depending on the relative humidity [41,46,47]

[2] For M13 it has been shown [50] that 83 % of the viral DNA is less mobile, suggesting that the DNA binding to coat protein actually occurs at the integral ratio of two nucleotide per protein subunit.

[3] A plant bacterium causing rice blight.

[4] Estimated from SDS-PAGE; the major coat protein seems to be synthesized without a cleavable signal sequence like Pf3 coat protein.

Several thousand copies (the exact number for each particular phage is indicated in TABLE 5.I) of the corresponding major coat protein are symmetrically arranged around the average phage axis to form the outer shell of the rod-shaped flexible particle. The inner surface of the protein shell is remarkable for the concentration of Lys residues facing toward the DNA, where they can function to neutralize the DNA negative charge [43]. The single strand of covalently closed DNA is extended lengthwise inside the coat in a hollow core. Fast, isolated motion of the encapsulated nucleotides have been observed by ^{31}P NMR (line shapes and transversal relaxation)[50,73,74]. The DNA of Ff phages, If1, Ike and Xf is wound into two anti parallel, right-handed helices, with the bases towards the center of the structural axis. The DNA of phages Pf1 and Pf3, on the other hand, has an inverted helical structure (I-helix) with the bases pointing outwards and the phosphates inside [41 and refs. therein]. Virion lengths vary from 700 nm for Pf3 up to 2000 nm for Pf1 [44]. The molecules of coat protein overlap to form a closely packed layer and lie parallel to each other with a slight offset (see Fig. 5.7). The axis of the almost entirely helical molecules makes an angle of about 20 degrees with the main axis of the Ff bacteriophages [40,75]. The coat protein subunits of fd were found to be arranged with $C5S_{2.0}$ symmetry that describes a fivefold rotation axis and a twofold screw axis [40,75]. The resulting capsid is a tight, interlocking network of subunits surrounding the cylindrical core containing the DNA (see Fig. 5.6).

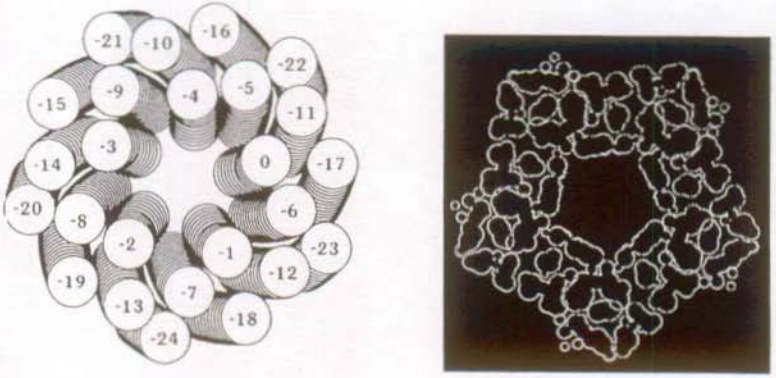

Fig. 5.6. Five-fold symmetry of filamentous phages. Model of the protein coat of the fd virion. Left: slab of about 20Å cut from the array of subunits, with subunits represented as

stacks of disks 9 Å in diameter and 1.5 Å in height. Numbers represent the origins of the cut protein subunits with respect to the reference subunit (marked 0). Right: section perpendicular to the virion axis through the van der Waals outline of the fd protein coat. Figure adapted from[43].

In intact phages the major Ff coat protein is predominantly α-helical as judged by CD [76], laser Raman spectroscopy [44,76,73] and X-ray fiber diffraction [77]. Solid-state ^{15}N-NMR analysis indicates an extended, slewed α-helix from residues 5-50, with a bend in the helix at about residue 30. Residues 1-4 undergo essentially isotropic motions on the kHz timescale and are considered to be non-structured or disordered [78-81]. The same structure was suggested from laser Raman spectroscopy [44] and X-ray fiber diffraction [82]. In the phage these residues are exposed to the external aqueous medium. The lysine residues in the C-terminal region lie in the proximity of the DNA allowing charge interaction of the positively charged ε-amino groups of the protein with the negatively charged phosphate moieties of the DNA backbone. The coat protein goes from the inside to the outside of the protein shell in about 1/6 of a turn.

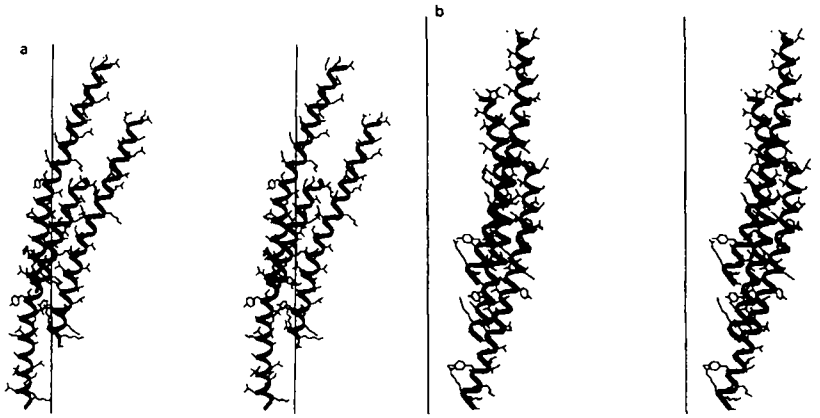

Fig. 5.7. Contacts between neighboring Pf1 coat protein molecules in the phage (stereo pictures). A: view towards the virion axis. B: view perpendicular to the virion. Figure taken from [83].

In the Pf1 phage the structure of the coat protein is known from X-ray diffraction (see Figs. 5.7 and 5.8). There exists a non-helical loop in the middle of the se-

quence, which contrasts with the totally helical structure of the Ff coat proteins in the phage. This loop motive might be the element which gives the related Pf3 coat protein its ability to insert into the inner membrane without a leader sequence.

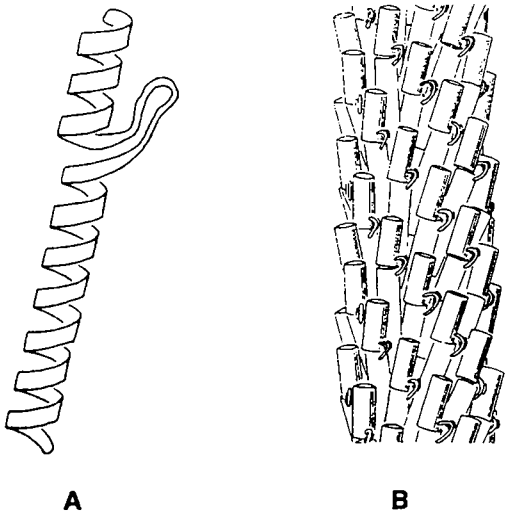

Fig. 5.8. Drawings of A: the Pf1 major coat protein, B: its packing within the protein coat. The coat protein is made up of a 3-turn α-helix and a 6-turn α-helix connected by a non-helical loop. The positioning and conformation of the non-helical loop are unknown, and drawn roughly to what would result in the smoothest surface for the virion. There are approximately 7200 coat proteins in the native Pf1 virus structure, so that the drawing of the virus cylinder corresponds to about 1% of the virus structure. Figure taken from Nambudripad 84

The viral DNA of Ff phages codes for 11 proteins, five of which are structural to the mature phage, whereas the other six gene products are needed for replication and phage assembly, but do not form a structural part of the mature phage. See Fig. 5.9. All gene products are named after the gene that encodes them; g#p stands for the product of gene #, with # ranging from 1 to 10.

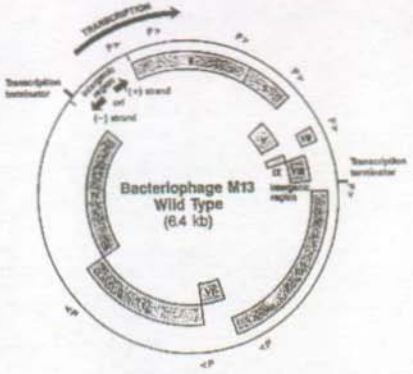

Fig 5.9. Organization of the genome of bacteriophage M13. Recently it has been shown that part of gene I, termed gene I* is also expressed separately, yielding a protein, gp1*, consisting of the C-terminal part of gp1. Both proteins are essential for phage expression [85].

In addition to the major coat protein a few copies of minor coat proteins are located at the ends of the virion. Their copy number per phage is probably five, which is reasonable assuming that they have a function in terminating both ends of the virion, which has a five-fold symmetry as described above.

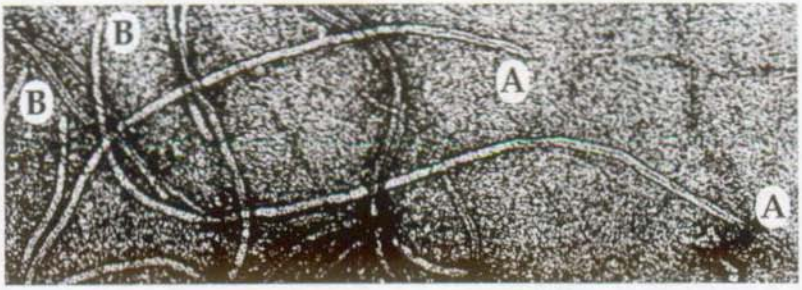

Fig 5.10. Transmission electron micrograph of bacteriophage M13 stained with 2 % uranyl acetate for 20 s. A) Sharp ends, B) Blunt ends. Photo: Henning Stalhberg.

The sharp end of the phage enters the host cell first (proximal end). The sharp form is due to the presence of the absorption complex formed by g3p and g6p. The adsorption protein, g3p, is anchored to the virion by its C-terminal part; its exposed N-terminal part mediates attachment to the tip of the F pili of the host bacterium[86]. Its five copies per phages also have a structural function in ending the virion. Gently disrupting the phage particle with SDS and chloroform yielded isolated complexes of the two proteins [87], called the absorption complex.

G7p and g9p are located on the opposite, distal end of the phage, which is produced first in phage assembly. Their precise function is still unclear.

The other gene products, except for g8p, which is the major coat protein, are necessary for phage replication and assembly and do not form a structural part of the mature phage.

5.3.2. LIFE CYCLE OF M13 BACTERIOPHAGE

Infection: The first step of infection of *E. coli* by M13 bacteriophage is a specific interaction of the absorption protein, g3p, with receptors for small nutrients at the tips of the F-pili. The subsequent uptake mechanism, the details of which are still unclear, leads to release of the viral DNA in the cytosol. The parental M13 coat protein is deposited as an integral membrane protein in the inner membrane where it spans the bilayer with the N-terminus facing the periplasmatic side [88].

Replication of viral DNA: The single-stranded viral DNA (ssDNA) is first converted to a double-stranded replicative form. This form is replicated and progeny single-stranded viral DNA is synthesized.

Expression gene products: Transcription and translation of the replicative form of the viral DNA leads to the expression of 11 gene products, five of which are structural to the mature phage. Studies of gene 8, coding for the major coat protein [89], of its messenger RNA [90], and of its DNA-directed synthesis in vitro [91-93] have shown that it is made as a precursor, termed procoat.

M13 procoat protein is initially synthesized in a soluble form [92] on free ribosomes in the cytosol. In wild-type *E. coli*, M13 coat protein accounts for up to 7 % of the total protein synthesis, i.e. one-quarter of the inner membrane protein [94]. In pulse-chase experiments completely synthesized M13 procoat protein was found in the cytosolic fraction; 50-90% was chased into membrane bound coat protein within a 10-15 s pulse labeling, showing there is no obligate coupling of synthesis of procoat with its insertion across the inner membrane for wild-type phages [95] and amber7 mutants [96]. The amber7 mutant phage has the advantage of slower processing and accumulation of procoat, when the growth temperature is raised from 37° C to 42° C. 40,000 Copies of procoat protein are produced per cell during M13 amber7 virus infection [97], reaching 26 % of the membrane protein synthesis [88].

5.3.2.1. Working model of the membrane incorporation of procoat protein

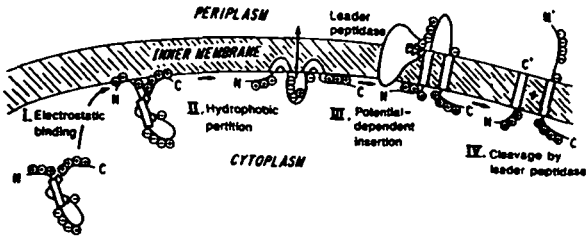

Fig. 5.11. Working model of M13 procoat protein membrane insertion and processing, showing the four principle stages (see text). Figure taken from [98,99].

According to a model of Wickner and co-workers the membrane incorporation process of M13 procoat protein (see Fig. 5.11) can be divided into three different steps: (I) electrostatic binding, (II) hydrophobic partitioning, and (III) translocation of the loop region between the two hydrophobic stretches, resulting in a U-shape conformation with two membrane-spanning segments. Finally, the leader sequence is cleaved off (IV) to yield the mature coat protein.

5.3.2.1.1. Electrostatic binding

In a manner independent of the Sec machinery [100 and refs. therein], newly synthesized procoat protein molecules find their way to the inner membrane where they bind electrostatically through interaction of the positive charges located in both the N- and C-terminal regions with negatively charged phospholipid headgroups [101]. *In vitro* procoat protein molecules can still be removed from the membrane in this stage by augmenting the osmotic strength of the surrounding medium.

5.3.2.1.2. Hydrophobic partitioning

After electrostatic binding of procoat protein to the inner membrane the hydrophobic stretches present in the mature part and in the leader sequence partition into the membrane leading to a tightly bound conformation of the protein which can no longer be removed by high salt concentrations.

5.3.2.1.3. Translocation

Translocation of the loop between the two hydrophobic parts of wild-type M13 procoat protein occurs only if the inner membrane of *E. coli* is energized. As the electrical potential is more positive on the outside (periplasmic space) and the loop region has an excess of three negative charges, one could imagine that the driving force of the translocation process could be the movement along the electrochemical gradient of charges (electrophoretic movement) [102]. However, experiments with mutant proteins in which the negative charged amino acids in the loop region were replaced with neutral ones were only slightly affected in their translocation efficiency. Only a mutant with four positive charges inserted showed a more pronounced reduction in translocation rate [100], without however changing the way of insertion. This effect was found to be independent of the membrane potential [103]. The energetic basis for the insertion process is more likely to be a conformational change in the hydrophobic domains [104]. The integrity of both these domains is necessary for correct membrane incorporation of M 13 procoat protein, as was shown by inserting polar residues into either one or both of these stretches [99].

5.3.2.2. Removal of leader sequence

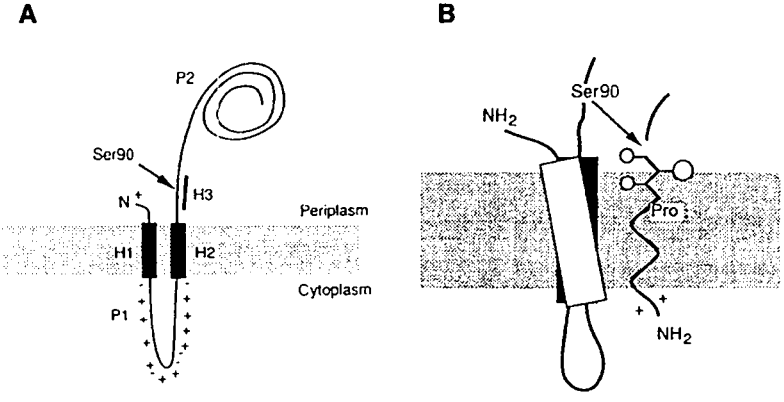

Fig. 5.12. Leader peptidase cleaves off the leader sequence. A: membrane topology of *E.coli* leader peptidase, with hydrophobic (H) and polar (P) regions marked. The position of the catalytically important Ser^{90} is also shown. B: Model for the interaction between *E.coli* leader peptidase and a signal peptide. The precise location of the catalytic site relative to the periplasmatic surface of the membrane is not known. Picture taken from [105].

Translocation of M13 procoat protein is completed by removal of the first 23 amino acids forming the signal sequence by the enzyme leader peptidase [91,106], which has its catalytic site on the periplasmatic side of the inner membrane where it is located (See Fig. 5.12). Per cell there are 500 molecules of leader peptidase [107]. In cells infected with wild type M13 phages procoat is processed with a half time shorter than 10 s [103]. Delay in the conversion of pulse-labeled procoat protein to coat protein was found in an M13 strain with an amber mutation in gene 7 [108], which shows a half-time of about 1 min.

The cleavage of procoat protein by leader peptidase depends on the amino acid sequence of procoat at positions -6 to +1. Of this stretch the amino acids on positions -6, -3 and -1 are critical for the cleavage by leader peptidase (see Fig. 5.12B). The Pro residue at the -6 position is essential for efficient *in vivo* processing, although *in vitro* most -6 mutants can be cleaved by high levels of peptidase. This suggests that the presence of a helix-destabilizing residue allows the peptidase to bind with a higher affinity. Another common motive in signal peptides and thylakoid transfer peptides are two small residues at positions -3 and -1 (Ser and Ala for M13 procoat protein), suggesting a second binding pocket which can accommodate two small residues [105,109]. We have used this motive to our advantage in our research. In order to obtain high yields of unprocessed M13 procoat protein we used the H5 mutant [4]. Here a point mutation at position 163 of M13 gene 8 exchanging guanidine for adenine results in a Ser^{-3} -> Phe^{-3} mutation in M13 procoat protein.

The incorporation process leads to mature coat protein spanning the inner membrane of the host, with the N-terminus on the outside and the C-terminus on the inside of the plasma membrane [88,98,110]. For further details see section 9.1. The cleaved-off leader sequence is digested by host peptidases.

5.3.2.3. Domain formation / extrusion

The membrane-mediated assembly and extrusion process causes no cell lysis. It might involve the membrane potential [111] and phage- and host-encoded accessory transport proteins [112-114].

Mature coat protein molecules assemble in the inner membrane in domains either spontaneously or by electrostatic interaction of the positive charges in the N-terminal region with the negative charge of the progeny viral DNA (this process is not yet well understood). As the DNA passes through the membrane it sheds

the DNA binding protein g5p to which it was complexed in the cytosol and major coat protein molecules are ordered around it in a tube as described in section 5.3.1. The g5p molecules are subsequently reused in the ssDNA trapping process [115]. The phage-encoded outer membrane protein g4p is essential for phage assembly and secretion [114,116]. It can form a homo-multimer, probably composed of 10 to 12 monomers, which are synthesized as soluble preproteins (including a leader sequence) in the cytosol. Membrane association and oligomerization of g4p occurs probably in a correlated process [114]. The multimers may form transient or instable complexes with the inner membrane protein g1p, which is also required for phage assembly [117]. Kazmierczak et. al. [114] propose g4p might form a channel through which the phages pass the outer membrane. The gating of this channel has to be very well controlled in view of the large putative channel diameter (the diameter of M13 phage being at least about 65 Å [47]).

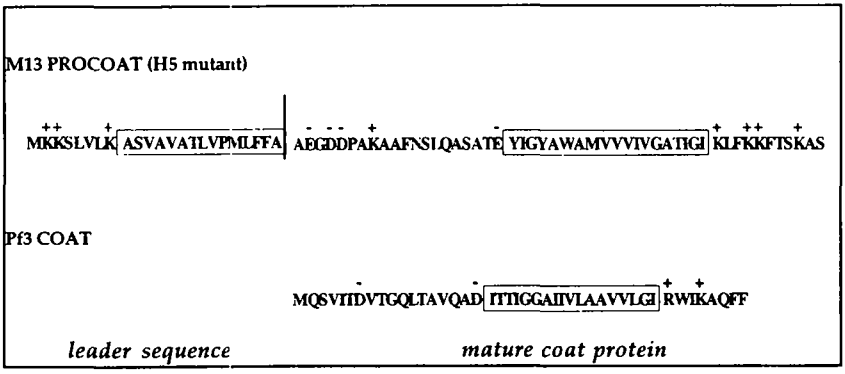

Fig. 5.13. Amino acid sequences of M13 procoat and Pf3 coat proteins. The line indicates the cleavage site of M13 procoat protein by leader peptidase between the leader sequence from amino acid -23 to -1 and the mature coat protein sequence from 1 to 50. The charged amino acids are indicated; the rectangles symbolize the hydrophobic, putative transmembrane domains. The H5 procoat mutant protein used in the present work is derived from the wild type procoat protein by a single amino acid mutation at sequence position -3 with an exchange of Ser to Phe [4].

5.3.3. PEPTIDE DISPLAY ON BACTERIOPHAGES

Recently the filamentous bacteriophages have been applied in various interesting screening techniques.

Phage M13 was used as a vehicle for displaying foreign peptides and proteins, thus providing new tools for the development of therapeutic, diagnostic and technological products [118]. Foreign peptides or proteins can be inserted at the N-terminal region of either the major coat protein, or of the adsorption protein, g3p. In principle one of these inserts could be used to target to specific cell types [119,120], whereas the other insert would function to load the phage with a pharmaceutical. Within this context it is important that neither the viability of the phage, nor the functional structure of the insert is impaired [121-123].

In the field of HIV research promising results were obtained using phages in two fundamentally different ways. The principal neutralizing determinant of the human immunodeficiency virus type 1 (HIV-1), the V3 loop of gp 120 from HIV-1 strain MN was expressed at the N-terminal region of the major coat protein of phage fd, leading to their functional display in multiple copies on the surface of the virion[124] . The other approach concerned the search for a useful vaccination by screening epitope sequences for a suitable HIV-neutralizing monoclonal antibody on a phage display library. The search yielded a four residue core antibody recognition sequence of GPGR flanked by two preferred residues on either side [125].

For this field of application the research of the structure and functioning of bacteriophages is an important prerequisite. Inversely, research on chimeric constructs of phage proteins and foreign peptides or proteins will yield additional information about the normal functioning of phages and their constituting proteins.

5.4. SCOPE OF THIS THESIS

This thesis concerns the fields of protein insertion and structure and dynamics of integral membrane proteins.

The approach chosen is to use model systems of purified small hydrophobic proteins reconstituted into artificial lipid membranes. Particularly, with this approach we are able to study the protein insertion process in a well-defined system with as little components as possible. This allows us to concentrate on the structural properties of the proteins themselves in a lipid bilayer, without the presence of other molecules which might normally interact with the proteins during or after insertion and thus complicate the analysis of the results obtained. The

insertion process and the structure of the inserted proteins are studied with optical methods.

The major coat proteins of filamentous bacteriophages provide very attractive models. They behave during part of their life cycle as integral membrane proteins of the bacterial host cells. Furthermore, *in vivo* they insert into the host's plasma membrane in a manner independent of any proteinaceous compound from the host bacterium. Thus all the information necessary for correct targeting and membrane insertion is present in the primary structure of the proteins themselves, which makes them interesting molecules to study. In addition to this, bacteriophage coat proteins are very suitable from an experimental point of view because they can be easily isolated in mg amounts from their respective phages. Moreover, the genetics, especially of phage M13, which are very well understood, offer numerous possibilities of overproduction and selective mutagenesis.

5.4.1. ORGANIZATION OF THIS THESIS

The purification and characterization the major coat proteins of bacteriophages Pf3 and M13 and the precursor of the latter, M13 procoat protein, are described in Chapter 6.

The proteins are reconstituted into lipid bilayers following several reconstitution protocols without using detergents, two of which are newly developed (Chapter 7).

Their secondary structure was determined in different solvents and reconstituted into lipid bilayers and into detergent micelles (Chapter 8).

The membrane topology of the inserted proteins was studied by the classical method of protease accessibility in a novel combination with amino acid sequencing and mass spectroscopy of the purified resulting fragments (Chapter 9).

The binding of the proteins to lipid membranes is quantified in Chapter 10. This is one of the rare examples where the thermodynamics of membrane insertion of such hydrophobic proteins have been determined.

5.5. LITERATURE

(1) Sargent, D. F.; Schwyzer, R. **1986** *Proc. Natl. Acad. Sci. USA, 83*, 5774-5778.

(2) Sargent, D. F.; Bean, J. W.; Schwyzer, R. **1988** *Biophys. J., 31,* 183-193.
(3) Heijne, G. v. **1983** *Eur. J. Biochem., 133,* 17-21.
(4) Kuhn, A.; Wickner, W. **1985** *J. Biol. Chem., 260,* 15907-15913.
(5) Dierstein, R.; Wickner, W. **1986** *EMBO J., 5,* 427-431.
(6) Gennis, R. B. *Biomembranes, molecular structure and function;* Springer-Verlag: New York, Berlin, Heidelberg, London, Paris, Tokyo, 1989, pp 533.
(7) Egberts, E.; Marrink, S.-J.; Berendsen, J. C. **1994** *Eur. J. Biochem., 22,* 423-436.
(8) Unwin, N. **1993** *Cell, 72,* 31-41.
(9) Cowan, S. W.; Schirmer, T.; Rummel, G.; Steiert, M.; Ghosh, R.; Pauptit, R. A.; Jansonius, J. N.; Rosenbusch, J. P. **1992** *Nature, 358,* 727-733.
(10) Deisenhofer, J.; Epp, O.; Miki, K.; Huber, R.; Michel, H. **1984** *J. Mol. Biol., 180,* 385-398.
(11) Deisenhofer, J.; Epp, O.; Miki, K.; Huber, R.; Michel, H. **1985** *Nature, 318,* 618-624.
(12) Henderson, R.; Baldwin, J. M.; Ceska, T. A.; Zemlin, F.; Beckmann, E.; Downing, K. H. **1990** *J. Mol. Biol., 213,* 899-929.
(13) Popot, J.-L.; Vitry, C. d. **1990** *Annu. Rev. Biophys, Biophys. Chem., 19,* 369-403.
(14) Weiss, M.; Kreusch, S.; Schiltz, E.; Nestel, U.; Welte, W.; Weckesser, J.; Schultz, G. **1991** *FEBS Lett., 280,* 379-382.
(15) White, S. H.; Wimley, W. C. **1994** *Curr. Opp. Struct. Biol., 4,* 79-86.
(16) Tanford, C. *The Hydrophobic Effect: Formation of micelles and biological membranes.;* 2nd ed.; John Wiley: New York, 1980.
(17) Hemminga, M. A.; Sanders, J. C.; Spruijt, R. B. **1992** *Prog. Lipid. Res., 31,* 301-333.
(18) Henderson, R.; Unwin, P. N. T. **1975** *Nature, 257,* 28-32.
(19) Engelman, D.; Steitz, T.; Goldman, A. **1986** *Annu. Rev. Biophys, Biophys. Chem., 15,* 321-353.
(20) Kahn, T. W.; Engelman, D. M. **1992** *Biochemistry, 31,* 6144-6151.
(21) Popot, J.-L.; Engelman, D. **1990** *Biochemistry, 29,* 4031-4037.
(22) Lemmon, M. A.; Engelman, D. M. **1992** *Curr. Opin. Struct. Biol., 2,* 1571-1576.
(23) Blobel, G. **1980** *Proc. Natl. Acad. Sci. USA, 77,* 1496-1500.
(24) Wickner, W. T.; Lodish, H. F. **1985** *Science, 230,* 400-407.
(25) Rapoport, T. **1992** *Science, 258,* 931-936.
(26) Nunari, J.; Walter, P. **1992** *Curr. Opin. Cell Biol., 4,* 573-580.
(27) Wolfe, P. B.; Rice, M.; Wickner, W. **1985** *J. Biol. Chem., 260,* 1836-1841.
(28) Rohrer, J.; Kuhn, A. **1990** *Science, 250,* 1418-1421.

(29) Wiech, W.; Sagstetter, M.; Müller, G.; Zimmermann, R. **1987** *EMBO J., 6,* 1011-1016.

(30) Lazdunski, C. J.; Baty, D.; Geli, V.; Cavard, D.; Morlon, J.; Howard, S. P.; Knibiehler, M.; Chartier, M.; Varenne, S.; Frenette, M.; Dasseux, J.-L.; Pattus, F. **1988** *Biochim. Biophys. Acta, 947,* 445-464.

(31) Duffaud, G. D.; Lehnhardt, S. D.; March, P. E.; Inouye, M. **1985** *Curr. Top. Membr. Transp., 24,* 65-104.

(32) Benson, S. A.; Hall, M. N.; Silhavy, T. J. **1985** *Ann. Rev. Biochem., 54,* 101-134.

(33) Heijne, G. v. **1994** *Annu. Rev. Biophys. Struct., 23,* 167-192.

(34) Driessen, A. J. M. **1993** *Biochemistry, 32,* 13190-13197.

(35) Joly, J. C.; Wickner, W. **1993** *EMBO J., 12,* 255-263.

(36) Lill, R.; Cunningham, K.; Brundage, L. A.; Ito, K.; Oliver, D.; Wickner, W. **1989** *EMBO J., 8,* 961-966.

(37) Kuhn, A. **1988** *Eur. J. Biochemistry, 177,* 267-271.

(38) Anderson, H.; Heijne, G. v. **1993** *EMBO J., 12,* 683-691.

(39) Marvin, D. A.; Hohn, B. **1969** *Bacteriol. Rev., 33,* 172-209.

(40) Makowski, L. In *Macromolecular Assemblies*1984; pp 203-253.

(41) Rashed, I.; Oberer, E. **1986** *Microbiol. Rev., 50,* 401-427.

(42) Model, P.; Russel, M. In *The Bacteriophages;* R. Calendar, Ed.; Plenum Press: New York, London, 1988; Vol. 2; pp 375-456.

(43) Marvin, D. A.; Hale, R. D.; Nave, C.; Citterich, M. H. **1994** *J. Mol. Biol., 235,* 260-286.

(44) Thomas, G. J.; Prescott, B.; Day, L. A. **1983** *J. Mol. Biol., 165,* 321-56.

(45) Hofschneider, P. H. **1963** *Z. Naturforschg. Teil B, 18,* 203-210.

(46) Dunker, A. K.; Klausner, R. D.; Marvin, D. A.; Wiseman, R. L. **1974** *J. Mol. Biol., 81,* 115-117.

(47) Glucksman, M. J.; Bhattarchajee, S.; Makowski, L. **1992** *J. Mol. Biol., 226,* 455-470.

(48) Asbeck, V. F.; Beyreuther, K.; Kohler, H.; von Wettstein, G.; Braunitzer, G. **1969** *Hoppe-Seyler's Z. Physiol. Chemie, 350,* 1047-1066.

(49) Nakashima, Y.; Koningsberg, W. **1974** *J. Mol. Biol, 88,* 596-600.

(50) Magusin, P. C. M. M.; Hemminga, M. A. **1993** *Biophys. J., 64,* 1861-1868.

(51) Zinder, N. R.; Valentine, M. R.; Stoeckenius, W. **1963** *Virology, 20,* 638-640.

(52) Loeb, T. **1960** *Science, 131,* 932-933.

(53) Schroeder, W. A.; Shelton, J. B.; Shelton, J. R. **1969** *Arch. Biochem. Biophys., 130,* 551-556.

(54) Marvin, D.; Hoffman-Berling, H. **1963** *Nature, 197,* 517-518.

(55) Marvin, D.; Hoffman-Berling, H. **1963** *Z. Naturforschg. Teil B, 18,* 884-893.
(56) Day, L. A.; Wiseman, R. L. In *The single-stranded DNA phages;* D. T. Denhardt, D. Dressler and D. S. Ray, Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1978; pp 605-675.
(57) Knippers, R.; Hoffmann-Berling, R. **1966** *J. Mol. Biol, 21,* 281-292.
(58) Snell, D. T.; Offord, R. E. **1972** *Biochem. J., 127,* 167-178.
(59) Newman, J.; Swinney, H. L.; Day, L. A. **1977** *J. Mol. Biol., 116,* 593-603.
(60) Wiseman, R. L.; Day, L. A. **1977** *J. Mol. Biol., 116,* 607-611.
(61) Berkowitz, S. A.; Day, L. A. **1976** *J. Mol. Biol., 102,* 531-547.
(62) Schaller, H.; Beck, E.; Takanami, M. In *The single-stranded DNA phages;* D. T. Denhardt, D. Dressler and D. S. Ray, Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1978; pp 139-163.
(63) Frangione, B.; Nakashima, Y.; Koningsberg, W.; Wiseman, R. L. **1978** *FEBS Lett., 96,* 381-384.
(64) Nakashima, Y.; Wiseman, R. L.; Koningsberg, W.; Marvin, D. A. **1975** *Nature, 253,* 68-70.
(65) Wiseman, R.; Berkowitz, S. A.; Day, L. A. **1976** *J. Mol. Biol., 102,* 549-561.
(66) Kostrikis, L. G.; Liu, D. J.; Day, L. A. **1994** *Biochemistry, 33,* 1694-1703.
(67) Peterson, C.; Dalack, G.; Day, L. A.; Winter, W. T. **1982** *J. Mol. Biol., 162,* 877-881.
(68) Putterman, D. G.; Boyle, P. D.; Yang, H. L.; Gryezan, T. J.; Frangione, B.; Day, L. A. **1984** *Proc. Natl. Acad. Sci. USA, 81,* 699-703.
(69) Berkowitz, S. A.; Day, L. A. **1980** *Biochemistry, 19,* 2696-2702.
(70) Newman, J.; Day, L. A.; Dalack, G. W.; Eden, D. **1982** *Biochemistry, 21,* 3352-3358.
(71) Denhardt, D. T.; Dressler, D.; Ray, D. S., Editors *The single-stranded DNA phages;* Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1978.
(72) Wen, F. S.; Tseng, Y. H. **1994** *J. Gen. Virol., 75,* 15-22.
(73) Magusin, P. C. M. M.; Hemminga, M. A. **1993** *Biophys. J., 64,* 1851-1860.
(74) Magusin, P. C. M. M.; Hemminga, M. A. **1994** *Biophys. J., 66,* 1197-1208.
(75) Banner, D. W.; Nave, C.; Marvin, D. A. **1981** *Nature, 289,* 814-816.
(76) Arnold, G. E.; Day, L. A.; Dunker, A. K. **1992** *Biochemistry, 31,* 7948-7956.
(77) Marvin, D. A.; Pigram, W. J.; Wiseman, R. L.; Wachtel, E.; Marvin, F. J. **1974** *J. Mol. Biol., 88,* 581-600.
(78) Valentine, K. G.; Schneider, D. M.; Leo, G. C.; Colnago, L. A.; Opella, S. J. **1985** *Biophys. J., 49,* 36-38.

(79) Colnago, L. A.; Leo, G. C.; Valentine, K. G.; Opella, S. J. In *Biomolecular stereodynamics, proceedings of the fourth SUNYA conversation in the discipline biomolecular stereodynamics*; R. H. Sarma and M. H. Sarma, Eds.; Adenine: Guilderland, NY, 1986; Vol. 3; pp 147-158.

(80) Colnago, L. A.; Valentine, K. G.; Opella, S. J. **1987** *Biochemistry, 26,* 847-854.

(81) Valentine, K. G. Thesis, University of Pennsilvania, 1987.

(82) Chamberlain, B. K.; Nozaki, Y.; Tanford, C.; Webster, R. E. **1978** *Biochim. Biophys. Acta, 510,* 18-37.

(83) Marvin, D. A. **1990** *Int. J. Biol. Macromol., 12,* 125-138.

(84) Nambudripad, R.; Stark, W.; Opella, S. J.; Mokowski, L. **1991** *Science, 252,* 1305-1308.

(85) Rapoza, M. P.; Webster, R. E. **1995** *J. Mol. Biol., 248,* 627-638.

(86) Armstrong, J.; Perham, R.; Walker, J. **1981** *FEBS Lett., 135,* 167-172.

(87) Gailus, V.; Rasched, I. **1994** *Eur. J. Biochem., 222,* 927-931.

(88) Smilowitz, H.; Carson, J.; Robbins, P. W. **1972** *J. Supramol. Struct., 1,* 8-18.

(89) Schaller, H.; Takanami, M. In *Single-stranded DNA Phages*; D. T. Denhardt, D. T. Dressler and D. Ray, Eds.; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1979.

(90) Sugimoto, K.; Sugisaki, H.; Takanami, M. **1977** *J. Mol. Biol., 110,* 487-507.

(91) Chang, C. N.; Blobel, G.; Model, P. **1978** *Proc. Natl. Acad. Sci. USA, 75,* 361-365.

(92) Wickner, W.; Mandel, G.; Zwizinski, C.; Bates, M.; Killick, T. **1978** *Proc. Natl. Acad. Sci. USA, 75,* 1754-1758.

(93) Konings, R. N. H.; Hulsebos, T.; van den Hondel, C. A. **1975** *J. Virol., 15,* 570-584.

(94) Wickner, W. **1983** *TIBS, March,* 90-94.

(95) Date, T.; Wickner, W. **1981** *J. Virol., 37,* 1087-1089.

(96) Ito, K.; Date, T.; Wickner, W. **1980** *J. Biol. Chem., 255,* 2123-1230.

(97) Zwizinski, C.; Wickner, W. **1982** *EMBO J., 1,* 573-578.

(98) Kuhn, A.; Wickner, W.; Kreil, G. **1986** *Nature, 322,* 335-339.

(99) Kuhn, A.; Kreil, G.; Wickner, W. **1986** *EMBO J., 5,* 3681-3685.

(100) Kuhn, A.; Zhu, H.-Y.; Dalbey, R. E. **1990** *EMBO J., 9,* 2381-2385.

(101) Kusters, R.; Breukink, E.; Gallusser, A.; Kuhn, A.; Kruijff, B. d. **1994** *J. Biol. Chem., 269,* 1560-1563.

(102) Enequist, H. G.; Hirst, T. R.; Harayama, S.; Hardy, S. J. S.; Randall, L. L. **1981** *Eur. J. Biochem., 116,* 227-233.

(103) Cao, G. Q.; Kuhn, A.; Dalbey, R. E. **1995** *EMBO J., 14,* 866-875.

(104) Kuhn, A.; Troschel, D. In *Membrane biogenesis and protein targeting*; W. Neupert and R. Lill, Eds.; Elsevier: New York, 1992; pp 33-47.

(105) Dalbey, R. E.; Heijne, G. v. **1992** *TIBS, 17*, 474-478.

(106) Mandel, G.; Wickner, W. **1979** *Proc. Natl. Acad. Sci. USA, 76*, 236-240.

(107) Wolfe, P.; Silver, P.; Wickner, W. **1982** *J. Biol. Chem., 257*, 7898-7902.

(108) Ito, K.; Mandel, G.; Wickner, W. **1979** *Proc. Natl. Acad. Sci. USA, 76*, 1199-203.

(109) Lee, J. I.; Kuhn, A.; Dalbey, R. E. **1992** *J. Biol. Chem., 267*, 938-943.

(110) Wickner, W. **1975** *Proc. Natl. Acad. Sci. USA, 72*, 4749-4753.

(111) Ng, Y. C.; Dunker, A. K. **1981** *Prog. Clin. Biol. Res., 64*, 467-474.

(112) Russel, M. **1991** *Mol. Microbiol., 5*, 1607-1613.

(113) Guy-Caffey, J. K.; Rapoza, M. P.; Jolley, K. A.; Webster, R. E. **1992** *J. Bacteriol., 174*, 2460-2465.

(114) Kazmierczak, B. I.; Mielke, D. L.; Russel, M.; Model, P. **1994** *J. Mol. Biol., 238*, 187-198.

(115) Pratt, D.; Laus, P.; Griffith, J. **1974** *J. Mol. Biol., 37*, 181-200.

(116) Pratt, D.; Tzagoloff, H.; Erdahl, W. S. **1966** *Virology, 30*, 397-410.

(117) Russel, M. **1993** *J. Mol. Biol., 231*, 689-697.

(118) Makowski, L. **1994** *Curr. Opin. Struct. Biol., 4*, 225-230.

(119) Hart, S. L.; Knight, A. M.; Harbottle, R. P.; Mistry, A.; Hunger, H. D.; Cutler, D. F.; Williamson, R.; Coutelle, C. **1994** *J. Biol. Chem., 269*, 12468-12474.

(120) Eerola, R.; Saviranta, P.; Lilja, H.; Pettersson, K.; Lovgren, T.; Karp, M. **1994** *Biochem. Biophys. Res. Commun., 200*, 1346-1352.

(121) Eroshkin, A. M.; Minenkova, O. O.; Fomin, V. I.; Ivanisenko, V. A.; Ilichev, A. A. **1993** *Mol. Mem. Biol., 27*, 843-849.

(122) Kishchenko, G.; Batliwala, H.; Makowski, L. **1994** *J. Mol. Biol., 241*, 208-213.

(123) Peters, E. A.; Schatz, P. J.; Johnson, S. S.; Dower, W. J. **1994** *J. Bacteriol., 176*, 4296-4305.

(124) Veronese, F. D. M.; Willis, A. E.; Boyerthompson, C.; Appella, E.; Perham, R. N. **1994** *J. Mol. Biol., 243*, 167-172.

(125) Jellis, C. L.; Cradick, T. J.; Rennert, P.; Salinas, P.; Boyd, J.; Amirault, T.; Gray, G. S. **1993** *Gene, 137*, 63-68.

6. ISOLATION AND PURIFICATION

6.1. INTRODUCTION

The fact that M13 procoat protein as well as M13 and Pf3 coat protein behave during certain stages of the assembly process as integral membrane proteins imposes certain constraints for their isolation and purification. In case of the coat proteins using the whole phage as a starting material has the immense advantage that the protein is already over 85% pure with the phage DNA as the only major contaminant, which can be removed in a relatively simple way. Purification of wild-type and mutant M13 procoat proteins and M13 and Pf3 coat protein from bacterial cell membranes is a more complex matter.

Although there already existed in the literature several protocols for the isolation of coat proteins as well as one protocol for the isolation of procoat protein, these methods had severe disadvantages. In most cases there was little control of the physical state of the coat proteins, leading to conflicting results in the biophysical studies based on these protocols. In the case of procoat protein the purity obtained according to the published procedure [1] was insufficient to allow careful biophysical studies. One of the goals at the beginning of this work was to develop highly reproducible, simple and efficient procedures for the preparation of molecularly dissolved coat and procoat proteins.

6.1.1. SPECIFIC PROBLEMS IN THE PURIFICATION OF MEMBRANE PROTEINS [2].

Water soluble peptides can be purified from an initial mixture in solution taking advantage of differences in at least one of four characteristic parameters:

a) molecular weight or size
b) isoelectrical point
c) biological activity or specific affinity to certain substances or classes of substances
d) hydrophobicity

The respective separation techniques to be used are:

a) gel filtration or size exclusion chromatography, preparative SDS PAGE, capillary electrophoresis, FPLC

b) ion exchange chromatography
c) affinity chromatography
d) hydrophobic interaction chromatography or reversed phase chromatography.

For water insoluble proteins such as integral membrane proteins there are a few restrictions to this scheme. Solubilizing agents often are necessary to prevent aggregation or irreversible binding to the column matrix by unspecific interaction of sample molecules and the matrix surface. Most frequently used solubilizing agents include concentrated solutions of urea or guanidine hydrochloride (GdHCl) (acting as chaotropic agents), a multitude of different detergents (acting as micelle forming agents) and organic acids (acting through protonation of prosthetic groups, or as chaotropic agents) and organic solvents. The presence of these agents might exclude several possible routes.

For example, ion exchange chromatography is impossible in the presence of GdHCl or high concentrations of acetic or formic acid, whereas detergents are incompatible with reversed phase chromatography (RPC). Because of the high concentrations needed when using urea or GdHCl, solubility problems might occur at higher fractions of the organic solvent (acetonitril, isopropanol, or a mixture of both solvents).

Only the very inefficient gel filtration is feasible in all solvent systems.

6.1.2. PUBLISHED EXAMPLES OF PURIFICATION PROCEDURES FOR HYDROPHOBIC PEPTIDES

The following examples of purification procedures for hydrophobic peptides have been used as a first approach in developing our purification methods. Their applicability to our system will be discussed in section 6.5.

6.1.2.1. Bacteriorhodopsin fragments.

The following protocol was used to purify two fragments of bacteriorhodopsin, B-1 (residues 1-155) and B-2 (residues 156-248) [3]. It was extremely helpful in developing one of our own purification protocols (see section 6.2.2.2.1).

Purple membrane was digested with $NaBH_4$, collected by centrifugation, washed twice with water and lyophilized. The dry membrane preparation was dissolved in 88% formic acid and diluted in ethanol (provoking hydrolysis of the Schiff's base linking the retinal moiety). The mixture was chromatographed on LH-60 (1.5 x 180 cm) in formic acid/ethanol 3/7 at room temperature. Fractions containing particular fragments were pooled and dialyzed against buffer (0.2 % SDS in 10 mM Na_3PO_4 (pH 6.0), NaN_3 0.025%).

6.1.2.2. M13 procoat protein

M13 procoat protein was first isolated by Wickner and co-workers [1, 4]. The protocol was lengthy and yielded relatively small amounts of protein of only 75 % purity. Although useful for biological experiments, it excluded biophysical experiments. In consequence, for our work we were obliged to develop a new, simple and efficient method, yielding highly pure protein.

Normally, procoat protein is rapidly processed to mature coat protein in infected cells (half-time <10 s [4]). To avoid this problem these authors used M13 virus carrying an amber mutation in gene 7 (M13 amber 7), the processing of which is considerably slower (half-time ≈ 1 min).

After 1h incubation of HJM114 *E. coli* with amber7 M13 virus (infection 1/75) at 37 °C in M9 minimal medium a portion of the culture was pulse-labeled with [^{35}S]-methionine. The label was recovered mainly in the procoat fraction and not in the mature coat. This fact was used to follow the purification process. Pulse-labeling and processing of procoat to coat was stopped with trichloroacetic acid (TCA). The resulting cell precipitate was washed with buffer and solubilized in SDS. The detergent extract was purified during three consecutive runs over Sephacryl S-200 SF superfine. The second and third run were recycling chromatography runs in the course of which the interesting fractions passed three times through the column. After each run procoat containing fractions (as judged by liquid scintillation assay for [^{35}S]-methionine) were concentrated by acetone precipitation. The precipitate was then resolubilized in SDS containing buffer. The estimated yield of this method was 13 % procoat protein with a purity of 75 %. 4 l of culture medium yielded approximately 0.4 mg of protein.

6.2. METHODS AND MATERIALS

6.2.1. GROWTH OF PHAGES, EXPRESSION OF MUTANTS

6.2.1.1. M13 and Pf3 phages

M13 and Pf3 phages were grown and isolated according to published procedures [5,6] as described in detail below. Phages were stored as a stable suspension in 0.1 M Tris-HCl pH 7, 0.1% NaN_3 at 4 °C.

Two media were used during phage preparation: 2YT medium (doubly concentrated yeast trypton medium) consists of 16 g bactotrypton, 10 g bacto yeast extract and 5 g NaCl per liter; TE consists of 10 mM Tris-HCl pH 7.5 and 1 mM EDTA.

A colony of *E. coli* (for M13 phages) or *Pseudomonas aeruginosa* (for Pf3 phages) was picked from a plate and inoculated in 10 ml 2YT medium. This preculture was grown at 37 °C either on a rotary wheel or in a water temperated shaking bath to a final density of 2 x 10^8 cells/ml. 1 l. of 2YT medium was inoculated with this preculture, grown to a density of 10^8 and infected with the appropriate phages at a phage/cell ratio of 1/1. After overnight growth the cells were spun down at 4300 g for 20 min. The supernatant was incubated for 1 h. in ice-cold water in the presence of 0,5 % Sarkosyl, 0,5 M NaCl and 5 % (w/v) PEG6000. The mixture was then centrifuged again at 4300 g for 20 min. The resulting pellet was resuspended in 100 ml 10 mM TE and Sarkosyl, NaCl and PEG6000 were added as above. After centrifugation for 20 min. at 16,000 g the resulting pellet was resuspended in 9.5 ml TE and 8 g CsCl was added, yielding a refractive index of 1.3605. This mixture was centrifuged at 100,000 g, 15 °C, for 72 h. In order to remove contamination from cell debris the best fractions of the gradient (as determined by a spot test) were recentrifuged under the same conditions as described above. The final fractions were dialyzed against TE. The usual titer was 2-4 X 10^{12} phages/ml. This preparation can be further concentrated by centrifugation at 200,000 g for 3 h and carefully removing 5/6 of the volume. The remaining fraction contains 97 % of the phage.

6.2.1.2. Expression of M13 procoat protein in *E. coli*

E. coli HJM 114 (Δ *lacpro*) F' (*lacpro*) and *Pseudomonas aeruginosa* were used to grow M13 and Pf3 phage, respectively. *E. coli* LC 137 (htpR, lon, lac_{am}, trp_{am}, pho_{am}, mal_{am}, rpsL, $supC^{ts}$, tsc::Tn10) was used for procoat expression. Gene VIII of M13 coding for the H5 procoat mutant, which contains a single amino acid mutation at sequence position -3 with an exchange of Ser to Phe, was in the pJF119 HE plasmid [7,8] under the control of an inducible tac-promotor, termed pJQ-8-H5.

6.2.2. PRECIPITATION/EXTRACTION

6.2.2.1. Isolation of M13 and Pf3 coat proteins from phages

Aliquots of phage suspension in buffer were vortexed with an equal volume of chloroform, yielding a white precipitate of phages. After centrifugation for 15 min at 20,000 g (Centrikon T-124, rotor A8-24, Kontron), the phages form a pellet at the water/chloroform interface. The pellet was re-suspended in water, again vortexed with chloroform and centrifuged. This washing step was repeated five times. The ultimate pellet was dissolved in an appropriate amount of isopropanol containing 0.1% TFA (v/v), yielding a solution of the coat protein ranging in concentration from 100 µM - 2 mM (as judged by UV absorbtion spectroscopy), which was isolated as the supernatant in a final centrifugation step, leaving the DNA as a clear pellet.

6.2.2.2. Isolation of M13 procoat protein from *E.coli*.

E. coli LC 137 with plasmid pJQ8-H5 was grown at 30 °C to a density of $4x10^8$ cells per ml in 2 l shaking cultures. IPTG was added at 0.5 mM and the culture was allowed to continue to grow for 4 h under extensive shaking. The cells were then chilled and 200 ml of a 50 mM NaOH solution was added. After lysis, the cell membranes were collected by centrifugation at 2,000 g for 20 min and resuspended in 50 mM Tris-HCl pH 6.8. After sonification at 4 °C for 1 min, the membrane fragments were collected by centrifugation and subsequently sonified at 4 °C for 1 min in order to homogenize the membrane preparations. Membrane fragments were isolated by centrifugation at 20,000 g, followed twice by re-suspension in water and centrifugation. The isolated membranes were then stored at -50 °C prior to protein extraction.

The extraction and purification of procoat protein has succeeded after extensive trials with different extraction methods and HPLC separations. Two protocols have been established.

6.2.2.2.1. Formic acid extraction protocol for M13 procoat purification

After thawing, the membranes were washed once with ethanol and then extracted twice with formic acid/ethanol/isopropanol 2/1/1 (by volume). Between each step the mixture was centrifuged for 10 min at 20,000 g in a Centrikon T-124 centrifuge fitted with a A8-24 Kontron rotor. The extract was purified twice by HPLC on a Hewlett Packard Series 1050 instrument over a Nucleosyl 100-7 C2 reversed-phase column (Macherey-Nagel, Düren, Germany) using a water/isopropanol gradient in 0.1% TFA. The collected peaks were concentrated by ultra-filtration on Millipore Ultrafree-CL Low Binding Cellulose (5,000 NMWL) filter units to a final concentration of 10-20 µM as determined by UV absorbance at 280 nm (see below).

6.2.2.2.2. Chloroform/methanol extraction protocol for M13 procoat purification

Frozen membrane pellets (typically 2.4 g) were thawed and centrifuged for 10 min at 20,000 g (Centrikon T-124 centrifuge, rotor A8-24, Kontron). All subsequent centrifugation steps were performed in the same manner unless stated otherwise. The supernatant was discarded and 10 ml 0.1 M Tris pH 6.8 was added. The mixture was homogenized and recentrifuged. The pH of the supernatant was tested with a pH paper. If the pH was above 8.0 the washing step with Tris was repeated. 10 ml 80 % EtOH was added, the sample was homogenised and centrifugated. The resulting pellet was weighed and 5 ml extracting solvent was added per gram of pellet. The extracting solvent consists of $CHCl_3$/MeOH 1/1, 0.1 M NaAc. The mixture was homogenized and incubated for 10 min at room temperature. The supernatant of a subsequent centrifugation was concentrated over Millipore UFC 30.000 MW cut-off filter units. This filtration step removed most of the dissolved lipids as well as part of the coat protein, if present.

6.2.3. HPLC

High performance liquid chromatography was performed on a Hewlett Packard Series 1050 instrument over Nucleosyl 100-7 C2 reversed-phase columns

(Macherey-Nagel, Düren, Germany). Elution profiles were monitored continuously by measuring the absorbtion at 214 and 280 nm. At the peak maxima an absorbtion spectrum was acquired by the built-in diode array detector, using one of the following gradient systems (TABLE 6.1).

TABLE 6.1. HPLC eluent systems.

	eluents	TFA content
a)	water-acetonitril	1 ‰
b)	water-isopropanol	1 ‰
c)	water-acetonitril/isopropanol 1/1	2 ‰

Gradients were linear with slopes not exceeding 1% changes in solvent composition per minute. Peaks of interest were eluted isocratically at the appropriate eluent composition.

6.3. RESULTS

The results of HPLC purification of the studied proteins is presented hereafter ordered by the eluent system applied, as described in TABLE 6.1.

6.3.1. WATER-ACETONITRIL

A sample of pure M13 coat protein analyzed on a water to acetonitril gradient yielded two different peaks at 47 % and 49 % organic solvent. The first peak was about 4 times as large as the second one, based on the on-line absorbance measured at 280 nm. The coat protein present in both peaks had a different structure as judged by circular dichroism (CD) (see Fig. 6.1 and section 8.2). They might represent two different aggregation forms, as the preparation of major coat protein from phages does not contain significant amounts of other proteins. The high dilution of the samples did not allow for an accurate determination of the concentration by absorbtion spectroscopy, but qualitatively the major peak seemed to correspond to a protein with a high amount of β-strand, whereas the protein eluting in the smaller peak appeared to contain predominantly α-helix. There was no equilibrium between the two forms as judged from control experiments after two days. Reinjection of the isolated peaks yielded elugrams of pure compounds and the CD experimental spectrum of a 1/1 mixture of both

peaks was virtually identical to the one calculated from the spectra of the two peaks measured separately.

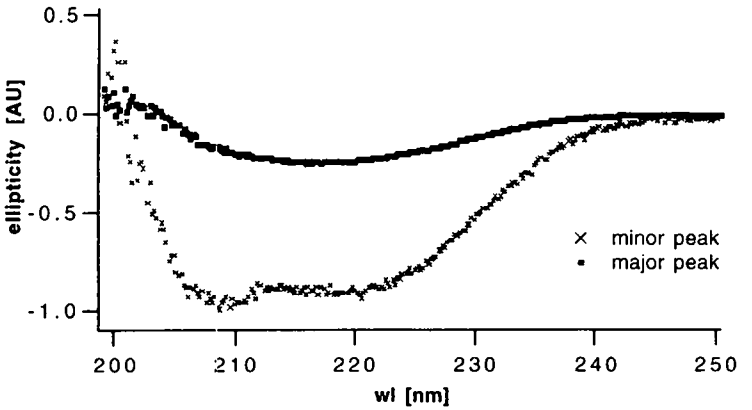

Fig. 6.1. CD spectra of H_2O-acetonitril gradient HPLC fractions. Spectra are corrected for solvent ellipticity by baseline substraction, normalized and corrected for approximate relative concentration by dividing the normalized ellipticity of the major peak by a factor 4.

This phenomenon did not occur to any serious extent with the other solvent systems we subsequently tested. We assume that aggregation processes occur on the column because the eluent system was too hydrophilic. The matter was not investigated any further, as it seemed to have limited relevance to the aims of the present study. Furthermore, this problem was readily overcome in the other purification protocols described below.

6.3.1.1. Water-isopropanol

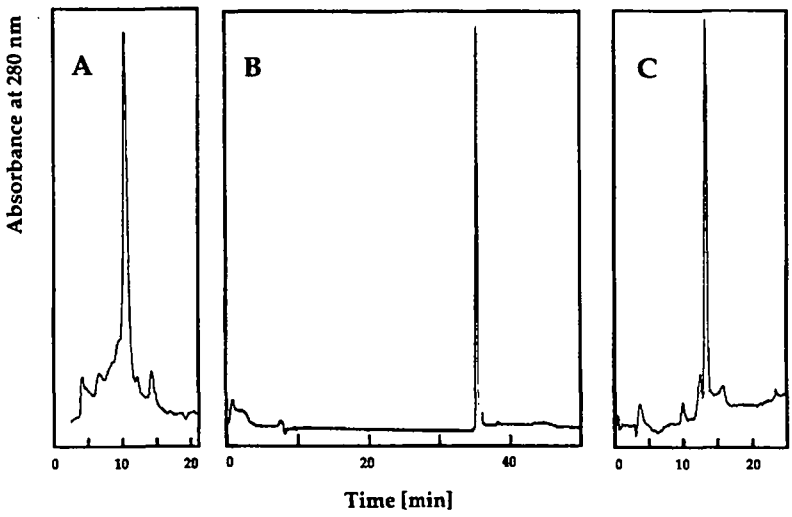

Fig. 6.2. HPLC runs of (A) M13 procoat protein, (B) M13 coat protein, and (C) Pf3 coat protein on a C_2 reversed phase column (125 x 4 mm^2). The proteins were detected at 280 nm. 100 µl of a 0.2 mg/ml procoat protein solution in formic acid/isopropanol = 1/1 (A), 15 µl of a 2.6 mg/ml M13 coat protein solution in isopropanol/0.1% TFA (B) and 20 µl mg/ml of a Pf3 coat protein solution in isopropanol/0.1% TFA were applied to the column and subsequently eluted with mixtures of isopropanol/0.1% TFA and water/0.1% TFA under conditions and of compositions of the organic solvent component as follows: (A) 0-4 min 40%, 4-12 min linear gradient 40-48%, 12-20 min 48%, 20-25 min linear gradient 48-100%; (B) 0-10 min 35-40 %, 10-70 min 40-100%; (C) 0-5 min 40%, 5-45 min 40-100%.

Fig. 6.2 shows HPLC runs of M13 procoat, coat and Pf3 coat proteins obtained with a water-isopropanol gradient. In each case a major protein peak is recorded, both with detection at 280 and 214 nm (not shown), indicating an apparent homogeneous protein preparation with more than 90% purity. Only the volume fractions corresponding to the major protein band were collected and used for further experiments. The minor peak appearing at 15 minutes in the HPLC run of M13 procoat protein shows the same properties on an SDS gel as the major peak and is therefore assigned as a higher molecular aggregate of the procoat protein in the isopropanol water mixture. This aggregate is induced during the HPLC column chromatographic run as rechromatographed pure M13 coat protein often showed a few small side peaks. Also in the case of the Pf3 coat protein the small bands appearing at 10, 13 and 15 minutes in the chromatogram

of Fig. 6.2C correspond, according to SDS gel electrophoresis (results not shown for this specific elution), to the same protein as the major HPLC band.

6.3.2. WATER-ACETONITRIL/ISOPROPANOL

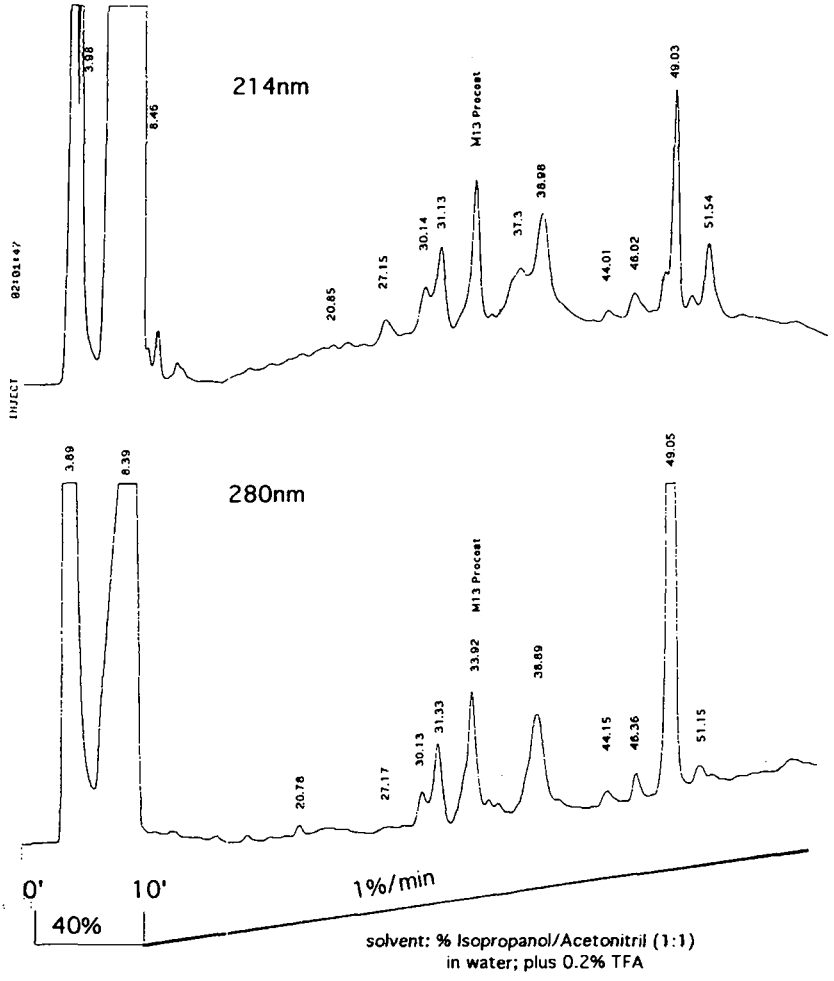

Fig. 6.3. HPLC purification of crude M13 procoat protein extract on a C_2 reversed phase preparative column (250 x 10 mm^2), using a gradient of acetonitril/isopropanol (1/1) in 2 ‰ aqueous TFA. The proteins were detected at 214 nm (top) and at 280 nm (bottom).

Figure 6.3 shows an example HPLC purification of a crude extraction of M13 procoat protein on a water-acetonitril/isopropanol gradient. All fractions were collected separately and analyzed by SDS-PAGE (see section 6.4.1) and immunoblotting (see section 6.4.2) to identify and to determine the purity of the procoat protein containing peak.

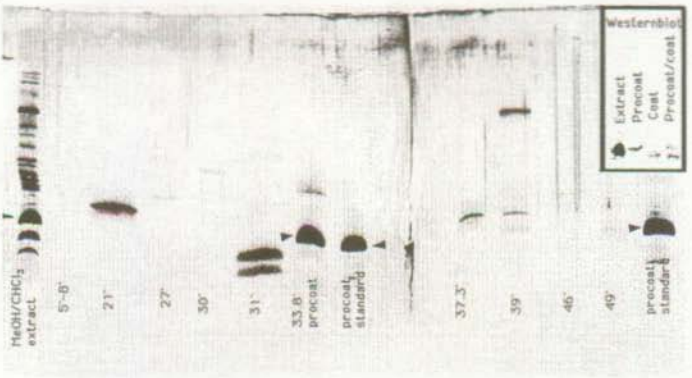

Fig 6.4. SDS-PAGE (silver-stained)and Western Blot (*inset*)analysis of M 13 procoat protein HPLC purification as described in Fig. 6.3. The lower minor band in the procoat standards represents coat protein. The techniques are described in sections 6.4.1 and 6.4.2 respectively.

6.4. CHARACTERIZATION OF PURIFIED PROTEINS

In order to determine the purity and the physical state of the purified proteins we applied various techniques each of which will be described below.

6.4.1. SDS-PAGE

6.4.1.1. Materials and Methods

Aliquots of the protein solutions were dried in Eppendorf tubes in the presence of 10 µl of an aqueous 10% SDS solution under high vacuum (HV) for about four hours. The dried samples were dissolved in 20 µl sample buffer containing β-mercaptoethanol, heated at 95 °C for two minutes and then loaded onto 21 % acrylamide gels containing urea[9], 48 % (w/v) in both stacking and separating gel. Protein bands were visualized by silver staining. Typically a volume corresponding to 100 ng-1 µg of protein was loaded per slot.

6.4.1.2. Results

Figure 6.4 shows an SDS-PAGE gel prepared as described above. It shows the three purified proteins. In the rightmost lane it is clear from the minor band that this preparation of M13 procoat protein contains a small amount of the dimeric form. Both bands were recognized by specific antibodies in an immunoblot assay (results not shown). Similar results were obtained in control experiments in which the samples were not heated prior to being loaded onto the gel.

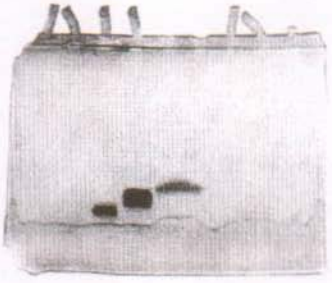

Fig 6.5. SDS-PAGE analysis of purified proteins. From left to right: Pf3 coat protein, M13 coat protein and M 13 procoat protein.

6.4.1.3. Discussion

In contrast to the starting material the purified proteins show only a single band in SDS-PAGE assays (see also figure 6.4 in section 6.3.3). The major part of the proteins loaded onto the gel is in monomeric form, as judged by the positions of the main bands, irrespective of heating of the sample before loading to the gel. This is in contrast to findings of other groups that M13 coat protein is a dimer in SDS (for references and discussion see section 8.1.1.3). However, our preparation differs from that of others: it is the first time that M13 coat protein is obtained in a monomolecular form in organic solution. All the preparations used by other groups in this field utilized detergents in the purification process, which might yield a different structure of the purified proteins [10]. Occasionally, we also found small amounts of dimer in individual preparations. A discussion on the different secondary structures found for coat proteins in relation to the purification and reconstitution methods used is in the introduction to Chapter 8.

6.4.2. IMMUNOBLOTTING

6.4.2.1. Materials and Methods

Protein samples were separated with SDS-PAGE as described above. The protein bands were transferred (1.5 hrs. at 50 V) to protein sequencing membranes PVDF (BioRad, Transfer-Medium Trans-Blot, 0.2 µm, cellulose).

Bands containing M13 coat or procoat protein were detected using polyclonal antibodies to M13 coat as a primary antibody and the Amersham ECL kit (Beaconsfield, U.K.). Horseradish peroxidase-labelled anti-rabbit IgG from Amersham was used as a secondary antibody at a dilution of 1/2000. Extensive blocking of non-specific binding sites with 2% milk powder solution and subsequent washing to remove excess milk was necessary to suppress background due to unspecific binding and to clearly resolve the labeled protein bands. The bound peroxidase is allowed to catalyse a reaction between hydroperoxide and luminol. Bands containing M13 coat or procoat protein therefore specifically emit chemiluminescence of which a contact print (10 s. - 1.5 min. depending on the amount of sample) is made on Kodak IR film.

6.4.2.2. Results

Fig 6.6. Immunoblot. From left to right: membrane extract, purified M13 procoat protein from this extract and M 13 coat protein standard. Due to the inherent low contrast of the IR photos of original immunoblots the background of the scanned image was suppressed using Adobe Photoshop to enhance the readibility of this figure. Pixels shown correspond only to specifically bound secondary antibodies.

Fig 6.6 shows an immunoblot of an *E. coli* membrane extract containing H5 mutant procoat protein, the purified procoat protein from this extract and M13 coat protein. Whereas the purified proteins shown one distinct band, the

membrane extract shows procoat protein at many positions where no procoat is expected.

6.4.2.3. Discussion

The accuracy in assigning proteins bands with methods based on immuno-precipitation is very much dependant on the specificity of the used antibodies. The polyclonal antibodies we used also contained some activity in recognizing *E. coli* lipoprotein. Thus, bands which did not contain procoat protein were neverthelesss detected by our immuno assay. Although it was a good indication for the presence of M13 procoat protein if a band was recognized by the antibodies, it was not enough to identify a procoat band with an acceptable level of confidence. It might seem that the use of monoclonal antibodies could improve the accuracy of this method of analysis. Recently three different monoclonal antibodies against different epitopes of fd coat protein were developed, which might be suitable candidates [11]. However, the activity of these monoclonal antibodies is probably much lower than that of a polyclonal mixture. The advantage of polyclonal antibodies is that they present a whole range of binding activities against different epitopes of the protein present in the band of interest. An approach to purify polyclonal antibodies is to mix them with *E. coli* cell extracts or membranes. The antibodies binding to *E. coli* proteins can then be removed together with the whole cells or cell membranes by centrifugation.

6.4.3. ABSORPTION SPECTROSCOPY

6.4.3.1. Materials and Methods

UV Absorption spectra were collected for each protein preparation on an HP 8452 A diode-array absorption spectrometer (Hewlett-Packart Co., Waldbronn, FRG). Each spectrum was corrected by substraction of a spectrum of the appropriate blank preparation. Peptide concentrations were determined from the baseline corrected spectra by means of the UV absorbance at 280 nm using ε = 8000 $M^{-1}cm^{-1}$ (1 Trp + 2 Tyr) for M13 procoat and coat proteins, and 5600 $M^{-1}cm^{-1}$ (1 Trp) for Pf3 coat protein[12].

6.4.3.2. Results

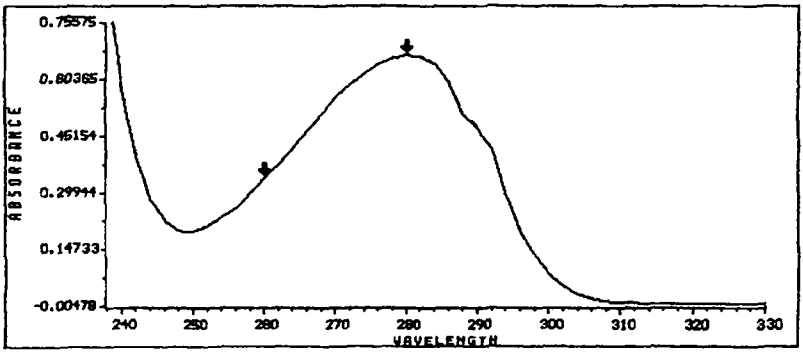

Fig 6.7. UV absorption spectrum of M13 coat protein in organic solution. The absorption ratio 260 nm/280 nm, indicated by the arrows, is used as a measure for the absence of DNA contamination in the purified coat proteins (see text).

Fig. 6.7 shows the spectrum of a solution of M13 coat protein in isopropanol, 0.1 % TFA, measured in a 1 mm quartz cell. The spectrum is baseline corrected by subtracting a spectrum of pure isopropanol, 0.1 % TFA. This baseline solvent spectrum is recorded directly prior to the actual measurement of the protein solution.

From the absorbance at 280 nm a concentration of 750 µM M13 coat protein was calculated. The ratio Abs_{280}/Abs_{260} for this spectrum is equal to 1.9 (see the arrows in Fig 6.7).

6.4.3.3. Discussion

As a general rule the extraction procedure for M13 and Pf3 coat protein from phages yielded preparations free of DNA as judged by the pronounced minimum in the UV absorption spectra at 245 nm (see Fig. 6.7). A generally accepted measure of absence of DNA (which has a strong absorption band centered around 259 nm) is an Abs_{280}/Abs_{260} ratio greater than 1.5 [13]. In our preparations this ratio was between 1.7 and 2.0. A good indication for the absence of larger, light-scattering aggregates in the preparations is the fact that the baseline is straight and virtually zero down to 310 nm.

6.4.4. MASS SPECTROSCOPY

Mass spectroscopy has been used to characterize the intact purified proteins as well as fragments of protease treated reconstituted proteins. This subject is discussed at length in chapter 9, which concentrates on the topology of reconstituted proteins.

In summary, the purified proteins were identified as having the expected molecular mass. No major contaminants could be detected. Pf3 coat protein could be N-formylated due to post translational modification, but data quality was insufficient to be certain. Both M13 coat protein and M13 procoat protein were found not to be formylated.

6.5. DISCUSSION

6.5.1. PURITY AND THE USE OF RPC FOR PURIFICATION OF SMALL MEMBRANE PROTEINS

In view of the large number of components in *E. coli* membranes it is interesting to note that the high purity of the M13 procoat protein preparation was obtained by a single extraction of crude *E. coli* membranes with either a mixture of isopropanol/formic acid = 1/1 or a mixture of $CHCl_3$/MeOH 1/1, 0.1 M NaAc, followed by a repeated twofold HPLC purification on a C_2 reversed-phase column. In order to obtain a pure preparation of M13 and Pf3 coat proteins, only a single extraction of the corresponding phages by isopropanol/0.1% TFA was necessary. Then again, phages only contain one major protein component. The other components which constitute the bacteriophages, the DNA and the minor coat proteins, are much more hydrophilic than the major coat protein and therefore remain in the aqueous medium during the purification process.

Hence, the difficulties in the purification of the major phage coat proteins are more related to their hydrophobicity and tendency to aggregate than to the removal of impurities. This is illustrated by the aggregation phenomena that occur when using the less hydrophobic acetonitril as an organic eluent, indicating a strong interaction between the proteins and the column material. This is also the case when the more hydrophobic C_8/C_{18} column (Vydac) instead of a C_2 column is used: in this case substantial amounts of the protein do not elute at all at normal working conditions, but bind irreversibly to the stationary

phase of the column (results not shown). However, polar interactions might equally be a factor contributing to protein loss and aggregation. The particular proteins contain, in addition to large hydrophobic stretches, also acidic as well as basic domains. The complementary charges in these domains might play a role in aggregation.

The aliphatic chains used in our RPC columns are linked to OH-groups at the surface of the silica which forms the matrix of the solid phase. According to information from the manufacturer Macherey-Nagel, a substantial amount of free, unblocked -OH groups are present on the column material (about 20 % of all the surface -OH groups present is not acylated during the fabrication of this type of RPC column material). At shorter chain lengths, of which C_2 is the extreme case, these free hydrophilic sites might not be sufficiently screened from the liquid phase, making hydrophilic as well as hydrophobic interaction between sample molecules and column material possible.

Another result corroborates this hypothesis. We compared the performance of two types of C_4 columns. A normal reversed phase C_4 Vydac column showed more protein retention and aggregation than the corresponding C_2 Macherey-Nagel column, as expected in view of the comparison between the C_2 and C_8/C_{18} columns. Another C4 column, a Merck LiChroCART 250-4 filled with C_4 LiChrospher 60 RP-select B (5 µm) (E. Merck, Darmstadt, FRG), however, yielded results similar to those obtained with the C_2 column. The Merck column is especially developed for lipid purification and might have a better surface coverage. In view of the fact that an important impurity in our inner membrane extracts are lipids it could be advantageous to use a column developed for lipids. The lipids ideally eluted at the beginning of the chromatography run before the proteins we are interested in. The ideal column material for our application, as yet unavailable, would be C_2 coupled to another matrix, which does not have free -OH groups like silica.

With recent improvements in HPLC technology and the availability of an ever growing range of new column materials its versatility as a purification technique increases. New end-fitting and joint technology as applied in both the Merck manuCART series column holders and the new generation of Macherey-Nagel columns based on the same system overcome many of the mechanical difficulties associated with handling a system with small particles under high pressure. Changing pre columns, a recurrent task with our purifications, becomes a very simple operation, without risk of disturbing the head of the

column with concurrent loss of separation efficiency. Preparative capillary electrophoresis is another technique which will gain importance in protein purification once it penetrates into more research laboratories. In our group its applicability to this specific case is under investigation.

6.5.2. YIELD

Loss of protein present in the raw materials (phages and crude *E. coli* membranes) is mainly due to the marked tendency of the proteins to aggregate. This tendency is due to their hydrophobic nature and the presence of complementary charges. There are two principle ways to increase the yield. The first is to use a richer source. It is not excluded that the M13 gene VIII could be even more strongly (over-)expressed, but it is not clear if this approach would result in higher amounts of extractable procoat protein (in a monomeric or oligomeric form). Part of the procoat protein elutes at the void volume, probably in the form of large aggregates. Higher concentrations of procoat protein in the inner membrane might result in more aggregates. For applications where the presence of detergents is not *a priori* unacceptable, like in high resolution NMR, it might be worthwhile to consider an extraction/purification protocol which includes (deuterated) detergents in order to resolve more of these aggregates. On the other hand, the investment might be larger than the benefits, since the chromatography as we used it functions only in the absence of detergents. For reconstitution experiments it is not an alternative, as it is more elegant to start reconstitution from a well defined molecular solution of purified proteins. Including detergents into the purification protocol would annul this important advantage of our present approach. As discussed in the previous section, the yield will probably also be enhanced when more advanced stationary phases become available. Optimizing our present methods would be more in line with our research strategy.

6.5.3. COMPARISON WITH LITERATURE

As a first approach we have tried to apply to our system gel permeation chromatography over Sephadex LH60 in formic acid/ethanol 3/7, as used in the purification of hydrophobic fragments of bacteriorhodopsin[3], see section 6.1.2.1. However, in our case the separation of mature coat protein and of procoat protein was very poor and it was impossible to separate *E. coli's* lipoprotein from procoat protein. This is in accordance with the limited purity achieved in earlier

attempts of procoat purification based on gel permeation chromatography in SDS [1], see section 6.1.2.2.

The improvements of our methods compared to the previously established method are the following. Whereas Zwizinski et. al. [1] had to repeat their chromatographic purification procedure of M13 procoat protein seven times, in our case it was sufficient to apply a single reversed phase high performance chromatography step. Moreover, we have reached a higher degree of purity, >90 % against 75% (estimated solely from SDS-PAGE results). This makes our preparations suitable for biophysical experiments on the purified compounds in solution and in reconstituted systems. Our purification methods for M13 and Pf3 coat proteins avoid aggregation problems and lengthy dialysis procedures of previous methods (see Chapter 8).

6.6. CONCLUSIONS

The purification protocols that we have developed yield very pure, well defined organic solutions of the three particular proteins. These solutions are suitable for reconstitution and subsequent biophysical studies aiming at elucidation of their structure in different stages of the insertion process into model membranes. A distinct advantage of our method above earlier methods described in literature is the lack of detergents in our preparations. Reconstitution methods with and without detergents will be discussed in Chapter 7. The presence of unblocked silanol groups in silica-based stationary phases for reversed-phase HPLC might aggravate protein aggregation during HPLC purification due to unwanted polar interactions.

6.7. LITERATURE

(1) Zwizinski, C.; Wickner, W. **1982** *EMBO J., 1,* 573-578.
(2) Heukeshoven; Dernick, R. **1988** *Chromatogr., 25,* 230-236.
(3) Liao, K.-S.; Huang, H.; Khorana, G. **1984** *J. Biol. Chem., 259,* 4200-4204.
(4) Ito, K.; Mandel, G.; Wickner, W. **1979** *Proc. Natl. Acad. Sci. USA, 76,* 1199-203.
(5) Wickner, W. **1975** *Proc. Natl. Acad. Sci. USA, 72,* 4749-4753.
(6) Thomas, G. J. J.; Day, L. A. **1981** *Proc. Natl. Acad. Sci. USA, 78,* 2962-2966.
(7) Fürste, J. P.; Pansegrau, W.; Frank, R.; Blocker, H.; Scholz, P.; Bagdasarian, M.; Lanka, E. **1986** *Gene, 48,* 119-131.
(8) Kuhn, A.; Zhu, H.-Y.; Dalbey, R. E. **1990** *EMBO J., 9,* 2381-2385.
(9) Ito, K.; Date, T.; Wickner, W. **1980** *J. Biol. Chem., 255,* 2123-1230.
(10) Hemminga, M. A.; Sanders, J. C.; Spruijt, R. B. **1992** *Prog. Lipid. Res., 31,* 301-333.

(11) Micheel, B.; Heymann, S.; Scharte, G.; Bottger, V.; Vogel, F.; Dubel, S.; Breitling, F.; Little, M.; Behrsing, O. **1994** *J. Immun. Meth., 171,* 103-109.
(12) Wetlaufer, D. B. **1962** *Adv. Protein Chem., 17,* 303-390.
(13) Bayer, R.; Feigenson, G. W. **1985** *Biochim. Biophys. Acta, 815,* 369-379.

7. RECONSTITUTION PROCEDURES OF THE PROTEINS INTO LIPID BILAYERS

7.1. GENERAL STRATEGIES FOR PROTEIN RECONSTITUTION

The surface of integral membrane proteins is partly hydrophobic and partly hydrophilic. The hydrophobic membrane-spanning part of the protein is in contact with the hydrophobic lipid acyl chains in the core of the membrane. Those parts of the protein extending into the aqueous environment are hydrophilic. The challenge in reconstituting these proteins into model membranes consists of matching the environmental dielectrical conditions during the whole process of transferring the protein from its original membrane to the model membrane.

The classical approach is to use detergents in the purification process, yielding mixed protein-detergent micelles in an aqueous environment, in which the detergents take over the role of the lipids to avoid exposure of the hydrophobic parts to the aqueous surroundings. During membrane reconstitution the detergents have to be removed again, after addition of the lipids. The disadvantage of this approach is the difficulty to completely remove the detergents again. The same disadvantage is valid for fatty acids, natural detergents, often used as membrane probes with a variety of different labels. Preferential selectivity for fatty acids (FA) relative to phosphatidylcholine (PC) might be 5-6 orders of magnitude [1]. Consequently, the off-rate for FA bound to membrane protein would be smaller by the same order. Moreover, it has become clear that detergents might influence the structure of membrane proteins. This complicates the interpretation of results of studies where detergents are used in the reconstitution procedure.

In view of these limitations detergents were excluded throughout isolation, purification and reconstitution protocols developed in this work. Organic solvents were used instead of detergents. This was made possible primarily by the development of the protein isolation and purification methods described in Chapter 6, which yielded well-defined protein stock solutions in organic solvents.

7.2. PUBLISHED PROCEDURES FOR M13 COAT PROTEIN RECONSTITUTION

Many different protocols have been published to extract coat proteins from the phage and to reconstitute them into lipid bilayers. These protocols were continuously changed in subsequent publications from the same research groups due to the difficulties associated with the isolation and reconstitution of coat proteins. All methods are based on extraction of the phages with either phenol or detergents. In some cases DNA and /or minor coat proteins are removed, in other cases not.

7.2.1. COSONICATION RECONSTITUTION METHODS

Phages in buffer were dried in the presence of a small amount of chloroform under a nitrogen flow. Lipids or detergent were added in chloroform and, after mixing, the sample was dried again under nitrogen flow, rehydrated with Tris buffer and finally sonicated (10-12 x 4 min. with a tip sonifier at 50 V)[2]. In this method DNA is not removed, excluding most types of biophysical experiments from a technical point of view (with the notable exception of NMR). Moreover, an uncertainty is introduced about a possible influence of the presence of DNA on the structure of the coat proteins.

7.2.2. CHOLATE DILUTION RECONSTITUTION METHOD

Coat protein was extracted from the phages with phenol [3] and dried. Then it was solubilized at a concentration of 2 mg/ml with buffer containing 1% cholate. To this solution a dispersion of lipids was added, yielding a clear solution. After 80-fold dilution with buffer, vesicles formed spontaneously. These vesicles are collected by centrifugation and resuspended in buffer. Residual cholate was found to be as high as 1.7 molecules per molecule coat protein, even after extensive washing for 48h[4-7], or much lower [8]: 1 molecule of cholate per 15-25 molecules of coat protein (Marcus Hemminga, personal communications). Interaction between reconstituted coat protein and cholate renders the analysis of the structural consequences of coat protein-lipid interactions more difficult, since the two types of interactions are inseparable in this system.

7.2.3. CHOLATE DIALYSIS RECONSTITUTION METHOD

Coat protein was extracted from the phages with phenol [3] and dried. Then it was solubilized with buffer containing 2% cholate and 8 M urea.

Cholate and urea were removed by extensive dialysis against a 100 fold excess of Tris buffer for a total of 48 h. with buffer changes every 12 h. [9]. In an earlier protocol 10% methanol was added to the first three dialysis mixtures [10]. The same disadvantages of the presence of residual cholate as mentioned in the previous section are valid.

7.2.4. FREEZE-THAW CHOLATE RECONSTITUTION METHOD

Phages were incubated with cholate and chloroform. Gel filtration over Sephacryl S-200 removed the DNA from the solubilized coat protein. The resulting solution was added to an aqueous dispersion of sonified lipid vesicles. This mixture was frozen in liquid nitrogen, thawed at room temperature and resonicated. Finally cholate was removed by 4 x 12 h. dialysis against buffer. The samples thus prepared equally contain residual cholate.

7.2.5. DIRECT MIXING METHOD

For some NMR studies a lipid dispersion was added to phages in water. This mixture was then sonified with a tip sonifier [11,12]. Incorporation into SDS micelles for NMR purposes can similarly be achieved by incubating phages with SDS at 50°C. In this case sonication is not necessary, but the sample are lyophilized. Prior to use they are rehydrated [13]. For NMR experiments the signal due to the presence of DNA is negligible. The presence of DNA does however exclude most other types of experiments.

7.3. METHODS USED IN THIS WORK

In order to investigate the structure of the particular proteins in a membrane-like environment we have reconstituted them into lipid bilayers using the four protocols described below, all of which start from well-defined organic stock solutions (see Chapter 6 for detail). The first two are standard protocols for peptide reconstitution, the latter two have been developed in order to cope with the especially hydrophobic character of the proteins used. In Fig. 7.1 the different methods we employed are summarized in the form of pictograms. The necessity for thoroughly degassing will be discussed in section 7.4.

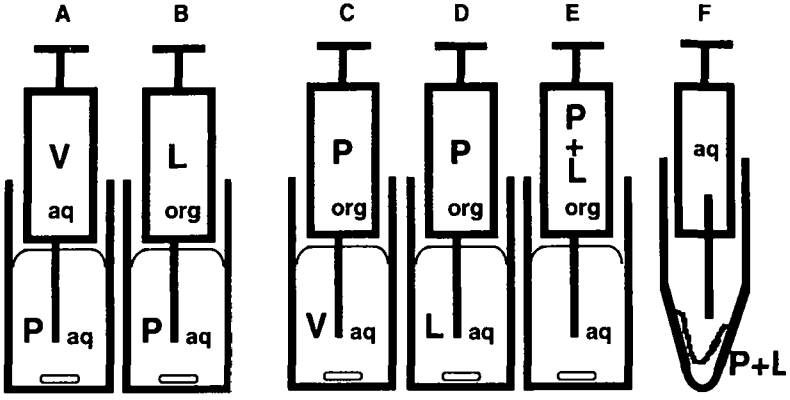

Fig. 7.1. Protein reconstitution methods. Protein is either incubated into sonicated vesicles (V)in aqueous environment (A and B) or incorporated from organic solution (C-F) C: "Preformed Vesicles" method, E: "Organic Phase Evaporation" method, F: "(Freeze-Dried) Mixed Film" method, C, E, F: main reconstitution methods, A, B, D: control experiments. V: sonicated vesicles, L: unsonicated lipid solution or dispersion, P: protein, aq: aqueous buffer, org: organic solvent.

7.3.1. "PREFORMED VESICLES" (PFV)

SUV were formed by sonication in buffer. Peptides in organic solvent were injected to a suspension of these vesicles under stirring. The final amount of organic solvent in the samples was smaller than 0.2 % before the samples were degassed with He. For control experiments slight modifications were introduced into this general scheme. In one type of control experiment the protein was injected under stirring into the aqueous phase first. After 2 minutes incubation time the vesicles were added to an appropriate final concentration. In another control large unilamellar vesicles (LUV) were used instead of SUV. LUV with a diameter of 1000 Å were produced by pressing 5 times freeze-thawed liposomes through a 1000 Å Nucleopore polycarbonate filter (Costar corp., Cambridge, MA) fitted into a stainless steel extruder (The Extruder, Lipex Biomembranes Inc., Vancouver, BC).

7.3.2. "MIXED FILM" (MF)

Appropriate volumes of solutions of lipids (2 mM in chloroform/methanol 1/1 (vol./vol.)) and peptides in organic solvent, were mixed. The solvent was evaporated first under a continuous flow of N_2, then under vacuum. The resulting films were rehydrated with buffer, sonicated until a clear suspension of

small unilamellar vesicles (SUV) was obtained. Finally, the sample was degassed with He.

7.3.3. "ORGANIC PHASE EVAPORATION" (OPE)

Lipids and peptide, both in organic solution, were mixed. This chloroform containing mixture was added to buffer and dispersed through continuous vigorous stirring. The remaining chloroform-containing droplets of organic solvent were evaporated from the water phase by a continuous stream of He. This procedure also yielded a clear suspension of vesicles.

7.3.4. "FREEZE-DRIED MIXED FILMS" (FD-MF)

Lipids and peptide, both in organic solution, were added to a 10 times larger volume of *t*-butanol. This solution was freeze-dried. The resulting powder was rehydrated with buffer, sonicated until a clear suspension of small unilamellar vesicles (SUV) was obtained. Finally, the sample was degassed with He.

7.4. PHOTO OXIDATION OF TRYPTOPHAN RESIDUES

(Pro)coat proteins are slowly oxidized by molecular oxygen, as judged by the appearance of a small peak at M+16 in mass spectroscopy for samples in organic solution which have been kept at 4 °C for several months (results not shown). This oxidation leads to a loss of α-helical structure (and to concomitant aggregation) as judged by circular dichroism.

This phenomenon is dramatically accelerated in oxygen-containing samples which are irradiated by UV light. During fluorescence spectroscopy and circular dichroism measurements the samples are subjected to high power UV irradiation from the spectrometer's lamp. When oxygen is present this is transformed into ozone, which attacks the indole group of Trp-26.

The phenomenon of Trp oxidation by ozone and its influence on the structural stability has been studied in the case of type λ immunoglobulin light chain, ribonuclease T1 and hen egg white lysozyme[14]. It was shown by CD and by fluorescence that these proteins retain their original structure. However their stability towards heat and guanidine hydrochloride (chaotropic agent) decrease dramatically even at small exposure to ozone. The effect is strongest for those Trp residues which are buried and have little solvent accessibility. The enthalpy

and entropy changes for the modified proteins are larger than for the native protein[14].

Ozone oxidation modifies tryptophan residues to N'-formyl-kynurenine (NFK), which can be converted in a controlled manner to kynurenine (Kyn) by freezing in acid[15-20] (see Fig. 7.2). NFK has been shown by X-ray crystallography analysis[21] to have a structure very similar to Trp. The aromatic amide hydrogen is hydrogen bonded to the oxygen at Cβ and the two rings are almost co-planar as they are in Trp. The most important change is an increase in polarity, which explains why modification of those Trp residues which are buried in a hydrophobic environment have the most impact on protein stability.

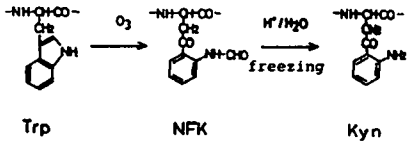

Fig. 7.2. Modification of Trp by ozone oxidation. Figure adapted from [14].

NFK and Kyn have absorption and fluorescence characteristics which differ significantly of those of Trp[19], as is shown in Fig. 7.3. NFK has an emission maximum at 434 nm on excitation at 325 nm, the emission maximum of Kyn is at 480 on excitation at 365 nm. With decreasing solvent polarity both compounds show a blue shift which is linear with respect to the dielectrical constant of the solvent[19].

The degree of oxidation of our proteins preparations was routinely checked by absorption spectrometry. Occasionally we had to discard a preparation of M13 procoat protein, because oxidation had occurred during the extraction / purification procedure, giving rise to the formation of Kyn-M13 procoat protein. Formation of NFK-modified proteins (see Fig. 7.4) and concomitant irreversible aggregation during experiments was a good measure for insufficient degassing with Helium.

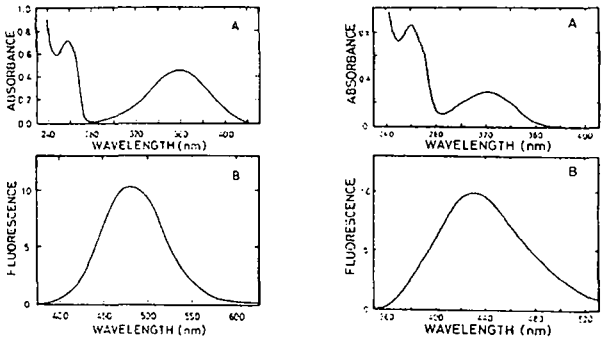

Fig. 7.3. Absorption (A) and fluorescence (B) spectra of Ac-Kyn-NH_2 (left) and Ac-NFK-NH_2 (right) in aqueous buffer. The excitation wavelength was 365 nm for Ac-Kyn-NH_2 and 325 nm for Ac-NFK-NH_2. Neither compound emitted fluorescence on excitation at 280 nm.

Although neither NFK nor Kyn emit fluorescence upon excitation at 280 nm in our oxidized proteins we do see evidence of NFK-fluorescence. In Fig. 7.4 a series of fluorescence spectra is shown for an oxygenated M13 coat protein solution after different UV irradiation times. A decrease in Trp fluorescence is accompanied by an increase in NFK fluorescence, showing a fluorescence energy transfer between Tyr-21 and /or Tyr-24 and NFK-26.

Care has to be taken to remove all oxygen from the samples. In many of our experiments it was observed that proteins containing the oxidized indole moiety are more likely to aggregate. Thorough flushing of buffers and samples with He circumvents this problem.

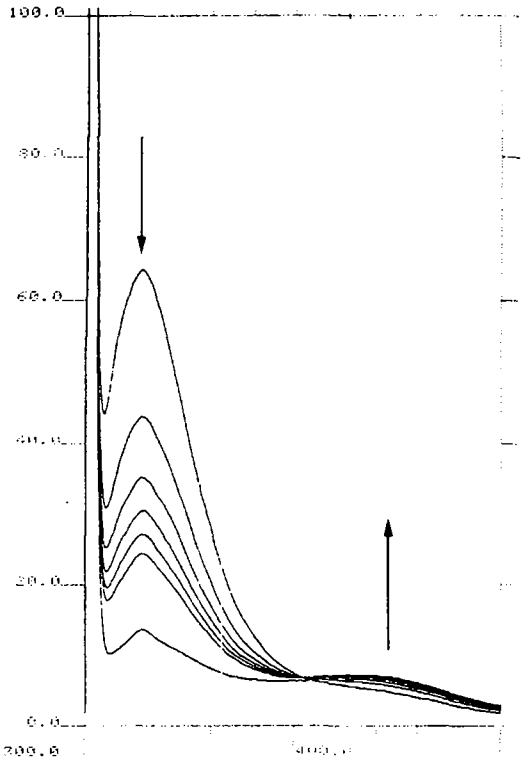

Fig. 7.4. Photo oxidation of tryptophan in M13 coat protein. Fluorescence emission spectra of 10 μM M13 coat protein in oxygenated EtOH/1 ‰ TFA. Spectra are taken of the same sample after different UV irradiation times (0, 15, 30, 45, 60, 75, 195 minutes, left peak, top to bottom). Arrows indicate the direction of change for both two peaks with increasing irradiation times.

Spectra were taken at ambient temperature on a Shimadzu R450 fluorometer, slit width 20 nm during irradiation, 5 nm during measurements. Excitation wavelength was 280 nm. Uncorrected spectra are shown.

7.5. DISCUSSION

Our newly developed protocols have two great advantages over previously published methods.

1) No detergents were employed in the purification/reconstitution process.

2) The reconstitution methods are rapid and simple.

Ad 1: our purification methods (see previous chapter) yield well-defined protein stocks in organic solution. Other groups did not have this advantage and had to resort to the use of detergents. The phenol extraction of coat protein from the phage, which has been widely used, has proven to be a particularly unsuitable for reconstitution endeavors, because it starts reconstitution from coat protein in the solid phase. Once coat protein is solid it is extremely difficult to redissolve it

again. Only by using detergent, and some times by adding urea as a chaotropic agent it was possible to dissolve coat proteins, often in form of aggregates.

Ad 2: previous methods often included extensive dialysis to remove detergent. As described above, it is impossible to completely remove the detergents. As will be pointed out in the next chapter, the affinity of coat proteins for detergents is much higher than for lipids (up to a factor 10^6).

Our reconstitution methods avoid both the exposure of our proteins to the aqueous phase and drying them. Thus we have no need for detergents and we avoid many problems of aggregation.

The fact that our samples are thoroughly degassed is also essential to prevent aggregation during our measurements. In none of the other biophysical publications on coat proteins this phenomenon has been described. Neither has degassing of samples been mentioned in the protocol descriptions. Therefore, it can not be excluded that insufficient degassing of samples is one of the causes of the large aggregation problems other groups have observed (see Chapter 8). Apparently there is no noticeable oxidation in detergent-based systems, possibly because of a smaller solubility of the apolar oxygen than in organic solvents (Marcus Hemminga, personal communications). However, a minute amount of oxidized protein could be sufficient to initiate the formation of β-type aggregates. Oxygen present in an aqueous dispersion of detergents/lipid vesicles would preferentially partition into the hydrophobic acyl chain region and thus be confined to the immediate vicinity of the Trp residues.

7.6. LITERATURE

(1) Marsh, D. **1990** *FEBS Lett., 268,* 371-375.
(2) Chamberlain, B. K.; Nozaki, Y.; Tanford, C.; Webster, R. E. **1978** *Biochim. Biophys. Acta, 510,* 18-37.
(3) Knippers, R.; Hoffmann-Berling, R. **1966** *J. Mol. Biol, 21,* 281-292.
(4) Racker, E.; Chein, T. F.; Kandrach, A. **1975** *FEBS Lett., 57,* 14-18.
(5) Wickner, W. **1977** *Biochemistry, 16,* 254-258.
(6) Kimelman, D.; Tecoma, E. S.; Wolber, P. K.; Hudson, B. S.; Wickner, W. T.; Simoni, R. D. **1979** *Biochemistry, 18,* 5874-5464.
(7) Smith, L. M.; Smith, B. A.; McConnell, H. M. **1979** *Biochemistry, 18,* 2256-2259.
(8) Spruijt, R. B.; Wolfs, C. J. A. M.; Hemminga, A. M. **1989** *Biochemistry, 28,* 9158-9165.

(9) Sanders, J. C.; Ottaviani, M. F.; Hoek, A. v.; Visser, A. J. W. G.; Hemminga, M. A. **1992** *Eur. Biophys. J., 21,* 305-311.

(10) Wolfs, C. J. A. M.; Horvath, L. I.; Marsh, D.; Watts, A.; Hemminga, M. A. **1989** *Biochemistry, 28,* 9995-10001.

(11) Fodor, S. P. A.; Dunker, A. K.; Ng, Y. C.; Carsten, D.; Williams, R. W. In *Bacteriophage Assembly;* M. S. Dubow, Ed.; Alan R. Liss: New York, 1981; pp 441-455.

(12) Leo, G. C.; Colnago, L. A.; Valentine, K. G.; Opella, S. J. **1987** *Biochemistry, 26,* 854-862.

(13) Schiksnis, R. A.; Bogusky, M. J.; Tsang, P.; Opella, S. J. **1988** *J. Mol. Biol., 200,* 741-743.

(14) Okajima, T.; Kawata, Y.; Hamaguchi, K. **1990** *Biochemistry, 29,* 9168-9175.

(15) Kuroda, M.; Sakiyama, F.; Narita, K. **1975** *J. Biochem. (Tokyo), 78,* 641-651.

(16) Tamaoki, H. **1978** *J. Biochem., 83,* 771-781.

(17) Yamasaki, N.; Tsujita, T.; Eto, T.; Masuda, S.; Mizuno, K.; Sakiyama, F. **1979** *J. Biochem., 86,* 1291-1300.

(18) Teshima, K.; Kuramitsu, S.; Hamaguchi, K.; Sakiyama, F.; Mizuno, K.; Yamasaki, N. **1980** *J. Biochem. (Tokyo), 83,* 771-781.

(19) Fukunaga, Y.; Katsuragi, Y.; Izumi, T.; Sakiyama, F. **1982** *J. Biochem. (Tokyo), 92,* 129-141.

(20) Fukunaga, Y.; Tamaoki, H.; Sakiyama, F. N., K. **1982** *J. Biochem. (Tokyo),* 92, 143-153.

(21) Kennard, C. H. L.; Matsura, Y.; Tanaka, N.; Kakudo, M. **1979** *Aust. J. Chem.,* 32, 911-915.

8. SECONDARY STRUCTURE OF RECONSTITUTED PROTEINS [1]

8.1. INTRODUCTION

The major problems imposed in the biophysical work with bacteriophage major coat proteins stem from their tendency to aggregate. The monomeric α-helical coat proteins can reversibly aggregate to α-oligomers containing up to 20 monomers[2,3]. An irreversible conformational change of α-oligomers to β-polymers consisting of up to 200 monomers occurs under certain isolation and reconstitution conditions. Only recently this phenomenon could be controlled and prevented as a fruit of the work described in this thesis and concomitant work of other groups active in this field[1,4]. After the practical part of this work was finished additional other structural data of the coat protein became available [4 and refs. therein], corroborating the findings of our work that the α-helical structure is the most relevant in reconstituted systems. Consensus is that the β-aggregated form is an experimental artifact without significance for the situation *in vivo*.

The status quo at the beginning of this project was an absence of structural data on M13 procoat protein and Pf3 coat protein and a controversy about the structure of the membrane-bound M13 coat protein (see section 8.1.2). In the phage the coat protein clearly is almost entirely α-helical (see section 5.3.1).

8.1.1. STRUCTURE IN AQUEOUS ENVIRONMENT

Although no quantitative data are available, it is assumed that the concentration of (pro-)coat proteins in the cytosol of the bacterial host is extremely low. As a rough estimate the upper limit for the actual cytosolic M13 procoat protein concentration can be calculated to be in the order of 20 fM [2]. The actual procoat

---

[1]Parts of this chapter were published previously in Biochemistry [1].

[2]Per generation (corresponding to about 30 min.) of infected E. coli 300 phages are formed [5], or 10 phages per minute per cell. Conversion of procoat protein to coat protein occurs within 30 s after infection in cells infected with wild-type M13 virus [6]. The conversion has a half time $t_{1/2} < 2$ s. [7].

...cont. on next page

concentration is probably even lower because of two reasons: overestimation in the calculation and co-translational insertion of nascent procoat protein chains. Either during or directly after synthesis the proteins bind with high affinity to the inner membrane. (The affinity for lipid membranes is quantified in Chapter 9.) Thus the structure in water is probably of little biological relevance, because *in vivo* the coat proteins are predominantly inserted into the inner membrane of the host bacteria or form a part of the protein coat in the mature phage.

The aqueous structure is however a factor to be considered in the choice and assessment of results of reconstitution experiments where the exposure to the aqueous environment can be substantial.

In the aqueous environment the major coat proteins are prone to irreversibly aggregation as experienced by all groups active in this field [e.g. 8], yielding an undefined mixture of structures unsuitable for most experiments. Therefore other workers have used a wide variety of detergents in reconstitution schemes (see section 7.1). Our own approach has been to reconstitute from organic solutions and avoiding the aqueous phase and the necessity of using detergents all together.

8.1.2. STRUCTURE IN MEMBRANE MIMETIC ENVIRONMENT

Both detergent and lipidic suspensions have been used to characterize the structure of coat proteins inserted into a membrane. Detergent suspensions have been applied both as a model system to mimic the membrane, e.g. in high resolution solution NMR, which is not feasible in lipid vesicles due to slow tumbling rates, and to avoid exposure of the proteins to the aqueous environment during membrane reconstitution.

In the literature until 1990, when the work described here started, there was a great confusion concerning the protein structures in reconstituted systems due to a lack of appropriate reconstitution methods.

Thus, procoat is immediately processed once it is inserted correctly in the membrane. Neglecting transcription/translation time in the 30 s. lag time procoat for 5 phages is synthesized. Thus a maximum of 5 x 2700 molecules procoat are secreted into the cytosol, which has a volume of about 1.25 nl [5]. This corresponds to a concentration of 18 fM.

In principle there are two possible structures for M13 coat protein in a membrane-spanning conformation and in detergent micelles: it can adopt either an α-helix (either monomer or aggregate) or a β-aggregate. In figure 8.1 both possibilities are depicted in a model of membrane inserted M13 coat protein.

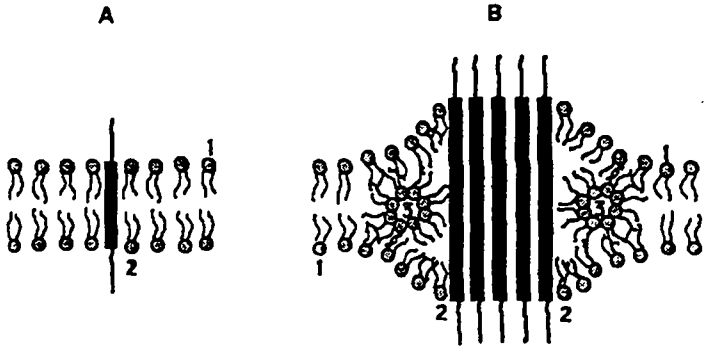

Fig. 8.1. Models of M 13 coat protein in lipidic environment. A: monomeric α-helical conformation, B: β-sheet conformation. As a consequence of the presence of M13 coat protein the lipids can be divided into populations with different motional restrictions: 1) bulk lipid, 2) lipid in contact with protein molecules. In case B at the border of the protein aggregates lipid is present in the form of inverted micelles (3). This structure was postulated to accommodate the increased length of a β-strand coat protein molecule within the membrane. Figure taken from [4].

Most early studies of the isolated or reconstituted forms of bacteriophage coat proteins either describe an aggregated form or an unidentified mixture of monomers and/or different aggregates, which makes it difficult to appropriately evaluate their results [4 and refs therein].

8.1.2.1. Structure in detergent suspensions

The most detailed structural studies on phage coat proteins were obtained by multidimensional NMR studies of SDS-protein micelles. The quality of the data depends to a great extent on an adequate amount of detergents [9,10]. The proteins were found to be predominantly present, but not exclusively, as dimers at detergent to peptide ratios up to D/P=200 [11-13]. A SDS micelle is composed of 60 SDS monomers [11], as compared to 8 monomers per micelle in the case of desoxycholate (DOC) [13]. At higher D/P values the monomeric form seems to be more stable. The occurrence of two sets of NMR signals was at first attributed to two monomeric units constituting an asymmetric dimer [11], before it became clear that they could be explained by a mixture of monomers and dimers [e.g.10].

The asymmetric dimer model proposes that the topology in the coat of the phage, where neighboring coat molecules are parallel, but slightly replaced to one another, remains intact (within the constraint of a possible slight rearrangement) in dimers solubilized from the phage by detergents.

In the case of dimers being solubilized from the phage, as suggested in the model presented above, one of the monomers has to be shifted by 33 residues with respect to its original relative position in the phage. In the whole phage the side-chains of one coat protein molecule interlock with those of neighboring molecules in a conserved pattern [14]. The interlocking is supposed to be slightly flexible as judged by slight changes in phage diameter and length of the virion axial repeat with hydration [15]. Detergent molecules might loosen this packing during phage solubilization, as is probably the case *in vivo* with the membrane lipids. In the inner membrane coat protein molecules are thought to assemble in clusters before phage formation. During the assembly process the molecules have to slide along their neighbors until the final packing is reached. It has not yet been resolved whether the local packing in the membrane might be related to the final packing in the phage or be non-specific [16-18].

Formation of head-to-tails dimers from initially monomeric coat proteins is another model to explain the existence of symmetric dimers. In this case the dimer would be stabilized by intermolecular pairing of complementary charges on the N- and C-termini. This energetically favorable conformation should be considered a reconstitution artifact, because in the inner membrane coat proteins are inserted exclusively in a vectorial way.

NMR studies of coat protein in micelles have indicated a largely α-helical conformation extending through the transmembrane region towards the C-terminus [10,11,19-23]. Recent structural models based on high resolution 2-dimensional NMR suggest that in SDS micelles Pf1 coat and M13 coat protein consist of 2 α-helical segments (residues 6-13 and 19-42 for Pf1 coat protein [24] and residues 6-20 and 24-46 for M13 coat protein [10,25]) connected by a non-helical loop. The longer hydrophobic C-terminal helix seems to be very stable and is proposed to traverse the micelle. The other helix is more structurally labile and is proposed to reside on the outside of the micelle (see Fig. 8.2). However, its position relative to the micellar surface is not expected to be fixed; there is probably significant motion, most likely via the hinge region around residue 22 [10]. The dynamics of the N-terminal part are as yet unclear.

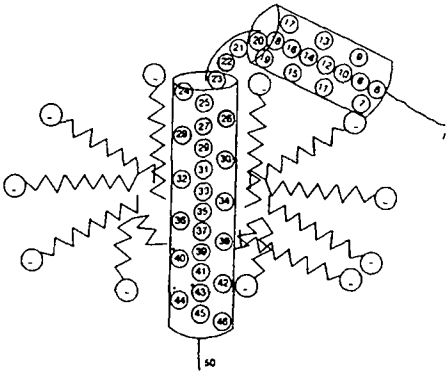

Fig 8.2. Model of M13 coat protein in SDS micelles. All molecules are drawn to scale. The two helices are displayed as cylinders, and the residues are represented as circles. Picture adapted from [25].

The relevance of this model for the structure of M13 coat protein in its membrane-inserted form will be discussed in section 8.4 at the end of this Chapter.

8.1.2.2. Structure in lipidic environment

A wide range of secondary structures of coat proteins in reconstituted membranes has been published over the last two decades, ranging from mainly β to almost completely α-helical conformations [24,26,27]. The first possibility would require a major conformational change upon phage assembly. As aggregation into β-type structures is irreversible *in vitro*, this is unlikely to be relevant for the bulk of the coat proteins in the inner membrane of their host. A hypothetical role in stabilizing contact sites between the inner and the outer membrane through which the newly assembled phages has also become less likely. Recently it has been proposed that two phage-encoded proteins, g1p, residing in the inner membrane, and the outer membrane protein g4p, are instrumental in channeling the completed phage out of the host bacterium [28] (see section 5.3.2.3).

The studies of other groups discussed below are based on two extraction/reconstitution strategies involving the use of phenol and/or detergents, notably cholate and desoxycholate. For NMR experiments phages were often directly extracted with SDS. Because of the limitations and uncertainties of these methods (see also Chapter 7), it is often difficult to assess their relevance to our experiments or the *in vivo* situation.

The first studies [23,26-29], which proposed that the coat protein structure in the membrane differs substantially from that in the phage were based on protein isolation by phenol extraction [12] of phages. Modified versions of this protocol has been instrumental in producing a wide variety of reconstituted aggregates until recent days [e.g. 2,29]. Lately, these aggregates were studied solely because it is interesting from a physical point of view which factors influence the formation of these aggregates and the number of monomers they contain (Ruud Spruijt and Cor Wolfs, personal communication).

Most data describing α-helical structures of coat protein have been acquired with a reconstitution method involving the use of cholate or desoxycholate. The disadvantage of this method is that the detergents can never be completely removed. The affinity of coat protein for detergents is much higher than for lipids, up to a factor 10^6 [30]. Even after 48 h. of dialysis (employed in practically all studies) there still remains 1.7 molecules of cholate per molecule coat protein in the samples, although much lower values have also been found [2]. In one NMR study the vesicle dialysis time was extended to 71 h. In this case a final protein to cholate ratio of 10:1 was obtained [31].

Coat proteins also have a higher affinity for negatively charged phospholipids than for zwitter-ionic lipids. Adding DMPG to DMPC membranes stabilizes smaller α-oligomeric aggregates of M13 coat protein incorporated by the cholate dialysis method [32]. The effect is largest between 0-20 % DMPG and remains constant above 40 % DMPG. This phenomenon coincides with the natural composition of the *E. coli* inner membrane, which contains 19 % (w/w) phosphatidylglycerol (PG) [33]. In a recent biological study [34] it was shown that processing of procoat to mature coat protein was retarded when the level of PG was reduced. In *in vitro* translocation experiments the translocation of procoat and of a SecA-dependent procoat analog could be reversibly slowed down, in a manner proportional to the PG content. The authors suggested a dual role for PG in translocation. Besides membrane binding of SecA this lipid would have a direct interaction with the M13 procoat in translocation across the inner membrane. This interaction might not be specific for PG, but merely reflect the fact that the initial step in the membrane incorporation of M13 procoat protein is mainly an electrostatic binding as shown with differently charged procoat protein mutants in an *in vitro* assay[35]. It has been shown that *in vivo* the cardiolipin (CL) content, normally accounting for 3% of the inner membrane lipids [33], is increased during f1 phage production by a factor 10-20, as a result of both a 3-fold

increase in CL synthesis and a decrease in CL turnover. The PG content remained virtually constant. This seems to suggest a specific association of the coat protein with CL [36]. The role of PG in this respect could be to provide an adequate surface charge density for an initial electrostatic binding. It has been shown by ESR that incorporation of M13 coat protein into lipid vesicles is facilitated by negatively charged phospholipids [32,37].

The occurrence of β-type structure and of α-oligomerization, which can be seen as the first step in the aggregation process, are thought to be promoted at low L/P ratios in reconstituted systems when contacts between inserted coat proteins molecules occur [32].

α-Oligomers can contain up to about 20 coat protein monomers. At higher aggregation numbers only β-polymeric aggregates are present. Around these β-polymeric aggregates (20-200 monomers, depending on experimental conditions) non-bilayer lipid structures are formed [38]. α-Oligomers also immobilize part of the available lipids, as is the case with β-polymeric coat [32,37], but the off-rate of the lipids in contact with the α-helical form is four times larger than in the case of β-type structure [39].

100 ps long molecular dynamics simulations in vacuum from minimized initial conformations (monomer *vs.* dimer and α helix *vs.* β-strand) of M 13 coat protein were performed, where the lipid bilayer was modeled by a hydrophobic potential. It was shown that a U-shaped twisted β-sheet is much more flexible than the α-helix, as monitored by the root mean square (rms) fluctuations of the Cα atoms. An α-helix monomer is more stable than a U-shaped β-strand. When going from monomers to dimers the energy difference between both structures decreases from 266 kJ/mol to 148 kJ/mol. This difference is thought to further decrease at higher aggregation numbers [40]. At higher aggregation numbers the β-structure could be more stable. This is consistent with the fact that β-aggregates consist of about 20-200 monomers and that aggregation is irreversible.

Above L/P = 12 the α-helical monomeric state is conserved in DOPC as judged by time-resolved fluorescence and ESR [41]. An estimated theoretical lower limit can be calculated under which monomers can no longer exist [4]. Assuming that an α-helix has the same diameter as a lipid molecule a hexagonal packing of proteins and lipids could exist in the membrane plane in which each isotropically dispersed coat protein monomer binds four lipids (see Fig. 8.3). When less lipid is

available protein-protein contacts have to occur. ESR measurements have indicated that below L/P= about 12 oligomerization already plays a role [42].

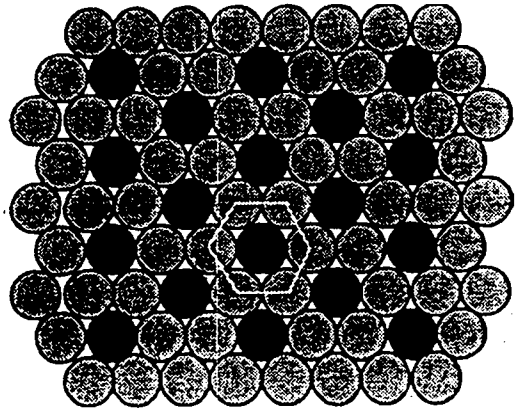

Fig. 8.3. Schematic illustration of a lipid-protein system with the maximum amount of protein without protein-protein contacts. It is assumed that the lipid molecules are in a hexagonal packing and that the bilayer structure is not modified by the presence of the protein. The protein is indicated with a black circle and the lipids are presented by gray circles. The L/P ratio is 4, as may be deduced from the white hexagon. Figure taken *from* [4].

For M13 procoat protein there are no biophysical data from other groups, except for one NMR study on synthetic M13 procoat protein [43]. Based on results for 14 residues (see TABLE 8.1) two rigid stretches are proposed in the hydrophobic domains in the leader sequence and in the mature part of the protein. The region around the cleavage site and both termini are labile.

TABLE 8.1. Dynamics of synthetic M13 procoat protein[22].

Residue	Dynamics	
	Solution NMR	Solid-state NMR
S(-20)	mobile	mobile
L(-19)	rigid	mobile
V(-18)	rigid	rigid
A(-15)	rigid	rigid
A(-12)	rigid	---
V(-11)	rigid	---
L(-4)	rigid	mobile
F(-2)	mobile	---
A(-1)	mobile	mobile
G(+3)	mobile	mobile
G(+23)	rigid	rigid
A(+25)	rigid	rigid
F(+45)	rigid	---
A(+49)	mobile	---

The results for the mature part coincide with results of solid-state NMR studies addressing the structure of coat proteins in the membrane. In DMPC multilamellar vesicles fd coat protein is rigid except for residues 1-4. In addition residues 45-50 are mobile, as judged by ^{2}H NMR and ^{15}N NMR experiments [22]. Furthermore, many of the side chains of residues in the immobilized part of the backbone undergo large amplitude jump motions [21,22,39,44-46]. These motions are 180 degree ring flips about Cβ-Cγ for Phe and Tyr and hops about tetrahedral or near-tetrahedral carbon sites in aliphatic side chains[22].

The α-oligomer state and monomeric state of coat protein cause a long-range effect of reduction of dynamics and increasing of lipid order [41 and refs therein]. This effect could be due to the rigidity of the transmembrane α-helix.

The importance of packing between proteins and their surrounding medium could also explain the high affinity of M13 coat protein for detergents and its preference for unordered bilayers. M13 coat protein is excluded from Ca^{2+}-induced gel-phase regions in PS/PC LUVs. Using a PC-bound quencher 25±5 times more M13 coat protein was found present in the liquid crystalline phase than in the gel phase [8].

8.2. CIRCULAR DICHROISM

8.2.1. INTRODUCTION

Circular dichroism (CD) is a spectroscopic technique measuring, as a function of wavelength, the difference of absorption of left-handed and right-handed circularly polarized light.

The CD signal is caused by the intrinsic asymmetry or the asymmetric arrangement of the chromophores present in the sample. These chromophores can be used as reporter groups for (parts of) the overall structure of the molecule under the experimental conditions we have chosen. The technique is extremely suitable to determine the conformation of macromolecules such as polynucleotides, DNA, polypeptides and proteins.

Although all amino acids except for glycine contain at least one asymmetric carbon atom (L or D configuration), most of them only display small CD effects. It is the conformation of the protein or polypeptide backbone, that is the spatial asymmetric arrangement of peptide units, which gives rise to their characteristic CD spectra.

The dichroic behavior of peptides in the far UV up to 250 nm is governed almost exclusively by the conformation of the peptide backbone. Several common motives in protein folding (α-helix, β-sheet, β-turns and random coil) give rise to specific contributions to the CD spectra. Analysis of the CD spectrum yields information about the overall percentage of these motives present.

The electrons forming the peptide bond are to some extent delocalized over three atoms: the peptide nitrogen, carbon and oxygen. The lowest energy electronic transition for a peptide is the n -> π^* transition (i.e. promoting an electron from the non-bonding oxygen 2p orbital to the anti-bonding π^* orbital) with a spectroscopic absorption band typically observed at 210-220 nm with low intensity. The π° -> π^* transitions (exiting electrons from the π° to the anti-bonding π^* orbitals) are responsible for the strong absorption near 190 nm and for a shoulder near 205 nm.

8.2.2. METHODS AND MATERIALS

8.2.2.2. Sample preparation

Peptide concentrations were determined by UV absorbance at 280 nm using ε = 8000 $M^{-1}cm^{-1}$ for M13 procoat and coat proteins, and 5600 $M^{-1}cm^{-1}$ for Pf3 coat protein (see section 6.4.3). Peptides were incorporated into vesicles as described in section 7.3. If not otherwise stated the buffer used for sample preparation was 1 mM Tris-HCl, pH 7.4, 0.1 mM EDTA.

8.2.2.3. Experimental procedures

Circular dichroism experiments were performed on an AVIV circular dichroism spectrometer model 62DS (AVIV Corp., New Jersey, USA). Unless stated otherwise, spectra between 186 and 260 nm of degassed, stirred 1 μM peptide solutions were acquired in quartz cuvettes of 0.01, 0.1 or 1 cm pathlength at 25°C. Blanks (organic solvent, buffer with/without lipids) were routinely recorded and subtracted from the original spectra. Buffer was 1 mM Tris-HCl pH 7.4, 0.1 mM EDTA.

8.2.2.4. Analysis

Depending on the secondary structure in which a particular peptide group is embedded, the electronic transitions presented in the previous section contribute in very different ways to the overall CD of the macromolecule. Figure 8.3 depicts CD spectra of the synthetic polypeptide poly-L-lysine in α-helical, β-sheet, and random-coil forms. This example was chosen, because polylysine can be easily transformed into the three different structures by changing the pH and/or temperature. The curves are remarkably different, and suggest that circular dichroism might be a powerful tool for analyzing the secondary structure in proteins. Several schemes for this analysis have been proposed based on curves similar to those in Figure 8.4, obtained with synthetic polypeptides in "pure" conformations, or upon attempts to deduce spectra from CD spectra of which the secondary structure is known from X-ray crystallography. Such efforts have been only moderately successful, for a number of reasons. First, some structural features, such as β-turns, cannot be obtained in a "pure" form in polypeptides; one cannot have a synthetic polypeptide which is *all* β-turn. Second, the CD spectrum of β-sheet seems to differ considerably depending on the polypeptide

used. That is, a typical spectrum of β-sheets does not exist. Also, α-helices are often distorted from an ideal conformation. Finally, CD spectra depend on the length of a certain peptide structural element as well as on the environment in the protein, which cannot be readily be accommodated in such an analysis.

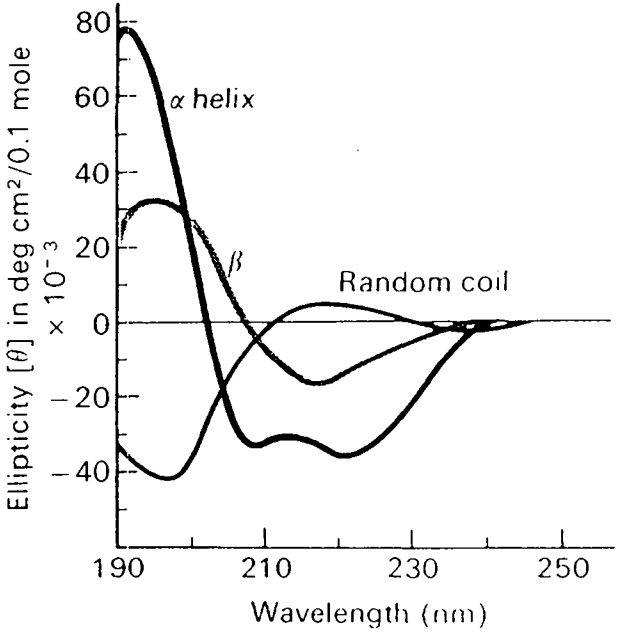

Figure 8.4. CD of a polypeptide (poly-L-lysine) in various conformations . After N. Greenfield and G. D. Fasman [47].

Therefore, more empirical methods have been developed to relate CD spectra to protein secondary structure. We employ the following method, based on the work of Hennesy and Johnson [48,49].

33 Data points are used to describe the measured CD spectrum between 186 and 250 nm. These points are fitted by a superposition of the spectra of proteins with a known three-dimensional structure (obtained by X-ray diffraction). The set of reference proteins is transformed to a set of so-called orthogonal eigenspectra which can be combined to generate the original data within experimental error. The fit determines the weight of the individual reference proteins in the measured spectrum. Knowing the mean secondary structure of the reference

proteins, we can derive the mean secondary structure of the protein under study as percentages of the following four structure classes: α-helix (H) [further distinguished as ordered (O) and disordered (D) helix], β-strand (S), β-turn (T), and undefined (U), which are designated to the corresponding X-ray structures according to Levitt and Greer [50].

The principle of the mathematical formalism is as follows.

The CD spectrum of each reference protein is digitized in the range of 186-250 nm at 2 nm increments yielding 33 data points. The digitized CD spectrum, **c**, of the protein to be analyzed can be fitted as the linear combination of the N reference proteins yielding

$$\mathbf{R}\,\mathbf{x} = \mathbf{c} \tag{8.1}$$

where **c** represents a vector of 33 data points. In our particular application we use N=15 reference protein CD spectra which are combined as a 15x33 data matrix **R**, composed of the digitized spectra in columns, with each column corresponding to a different reference protein. The coefficient vector **x** determines the contribution of each reference spectrum in the fit to the spectrum **c**. First a symmetric square matrix is constructed by multiplying **R** by its transpose $\mathbf{R}^T$. Diagonalization of this square matrix produces a matrix of 15 **eigenvectors**, **U**, and a diagonal matrix of 15 **eigenvalues**, **S**, such that

$$\mathbf{U}^T(\mathbf{R}^T\mathbf{R})\,\mathbf{U} = \mathbf{S} \tag{8.2}$$

The evaluation of the coefficient vector **x** is straightforward. Defining

$$\mathbf{C} = \mathbf{R}^T\mathbf{R} \tag{8.3}$$

we can rewrite eq. (8.1) as

$$\mathbf{C}\,\mathbf{x} = \mathbf{R}^T\mathbf{c} \tag{8.4}$$

Defining furthermore

$$\mathbf{R}^T\mathbf{c} = \mathbf{d} \tag{8.5}$$

it follows from eq. (8.4)

$$\mathbf{C}\,\mathbf{x} = \mathbf{d}$$

$$\mathbf{U}\,\mathbf{C}\,\mathbf{U}^{T}\,\mathbf{U}\,\mathbf{x} = \mathbf{U}\,\mathbf{d} \qquad (8.6)$$

With

$$\mathbf{p} = \mathbf{U}\,\mathbf{x}$$

$$\mathbf{g} = \mathbf{U}\,\mathbf{d} \qquad (8.7)$$

we obtain

$$\mathbf{S}\,\mathbf{p} = \mathbf{g} \qquad (8.8)$$

or

$$\mathbf{x} = \mathbf{U}^{T}\,\mathbf{p} \qquad (8.9)$$

The particular calculation is performed as follows. According to eq. (8.7) the individual elements are

$$\Sigma_v\, U_{iv}\, b_v = g_i \qquad (8.10)$$

and from eq. (8.8) it follows

$$p_i = g_i / s_i \qquad (8.11)$$

The coefficients of **x** are calculated according to eq. (8.9) as

$$x_i = \Sigma_v\, U^{T}_{iv}\, p_v = \Sigma_v\, U_{vi}\, p_v \qquad (8.12)$$

The main point is now that the coefficients x_i are correlated to the secondary structure of the protein under investigation in the following manner

$$\mathbf{f} = \mathbf{F}\,\mathbf{x} \qquad (8.13)$$

with the elements

$$f_r = \Sigma_i\, F_{ri}\, x_i \qquad (8.14)$$

The vector **f** defines the different percentages of the 5 different secondary structure classes (O, D, S, T, U) of the investigated protein. The 5x15 structure matrix **F**

defines the 5 structure classes of the 15 reference proteins used in the reference CD spectra.

Based on a statistical error analysis it can be shown that for the actual calculations only the 5 most important (i.e. those with the five highest eigenvalues) eigenvectors in **U** are necessary to reproduce the CD spectra under investigation within experimental error. Therefore within the calculation procedure we use a truncated set of basis vectors.

The fitting procedures were rewritten in Mathematica by Dr. Pierre Infelta for convenience of data handling and ease and speed of fitting (see Appendix A). Fitting was done either on a Macintosh cx or on a Quadra 700.

8.2.3. RESULTS

8.2.3.1. Secondary structure in different environments

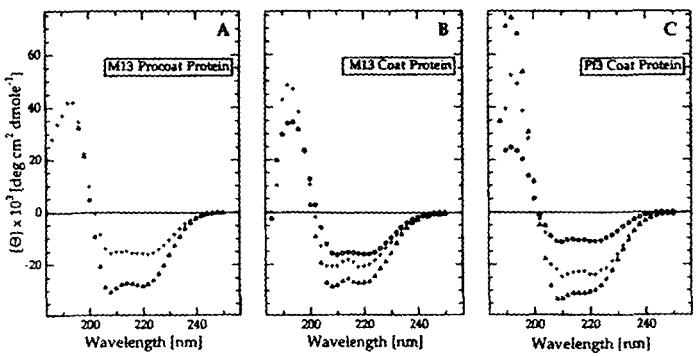

Fig. 8.5. CD spectra of (A) M13 procoat protein, (B) M13 coat protein, and (C) Pf3 coat protein, each protein dissolved in isopropanol (Δ), incorporated in membranes of POPC/POPG lipid vesicles at a lipid to protein molar ratio of 100, 1 mM Tris-HCl pH 7.4 (+), and dissolved in buffer, 1 mM Tris-HCl, pH 7.4 (o). Protein concentration was 0.1 mg/ml, temperature 25 °C.

Fig. 8.5A shows the circular dichroism spectra of M13 procoat protein dissolved in isopropanol/0.1% TFA and incorporated in lipid vesicles of POPC/POPG at a molar ratio of 1:1. In isopropanol, the average helix content of procoat protein is $[\alpha] = 90 \pm 5\%$. Upon incorporation into lipid membranes (mixed film reconstitution method) a considerable decrease of the helix content was observed

with [α] = 50%. The helix content for procoat protein in SDS micelles is comparable to that in lipid membranes of POPC/POPG. However, in pure POPC membranes procoat protein adopts a conformation with a substantially lower helix content, indicating that an electrostatic interaction between procoat protein and the negative membrane surface charges of POPG plays also a role in the formation of the membrane incorporated protein structure. The secondary structures of the three proteins in the different environments as determined by CD are summarized in Table 8.2.

Table 8.2. Helix content derived from CD spectra of the M13 procoat, M13 coat and Pf3 coat proteins in different environments.

	M13 procoat	M13 coat	Pf3 coat
	% α-helix		
isopropanol, 0.1% TFA	90	90	90
1mM Tris-HCl, pH 7.4		50	40
0.5% SDS in 1mM Tris-HCl, pH 7.4	45	55	65
POPC/POPG=1/1 vesicles	50	60	75
POPC vesicles		45	40

Fig. 8.5B and Fig. 8.5C show the CD spectra of M13 coat and Pf3 coat proteins, respectively, each in isopropanol/0.1% TFA, in water and in lipid membranes (mixed film reconstitution method). The average protein secondary structure for the two coat proteins is [α] = 90±5%, similar to the case of procoat protein. Upon incorporation into lipid vesicles the α-helix content decreases to [α] = 70% in the case of Pf3 and to [α] = 60% for M13 coat protein. It should be noted that the membrane incorporation of the three different proteins according to the different reconstitution procedures described in Chapter 7 yielded within experimental error identical results. As in the case of procoat protein, the secondary structure of the two coat proteins in SDS micelles is comparable to that in POPC/POPG membranes. The predominantly α-helical structure is even preserved in an aqueous protein solution with [α] = 40% for Pf3 and [α] = 50% for coat protein.

8.2.3.2. Temperature dependence

The stability of the secondary structure of the three proteins in lipid membranes was investigated at different temperatures by CD measurements. The mean residual ellipticities at 222 nm, $[\Theta]_{222}$, for the 3 proteins were measured between 20 and 80 °C as shown in Fig. 8.6. For M13 coat protein, $[\Theta]_{222}$ is constant in the whole investigated temperature range, while for M13 procoat and Pf3 coat proteins the $[\Theta]_{222}$ values decrease continuously with increasing temperature. Above 60 °C, samples with procoat protein did not reach a stable CD signal. The $[\Theta]_{222}$ value is, to a good approximation, proportional to the α-helix content of the corresponding protein. According to these experimental results, the α-helical structures of M13 coat protein in lipid membranes are considerably more stable than the corresponding helices of M13 procoat and Pf3 proteins.

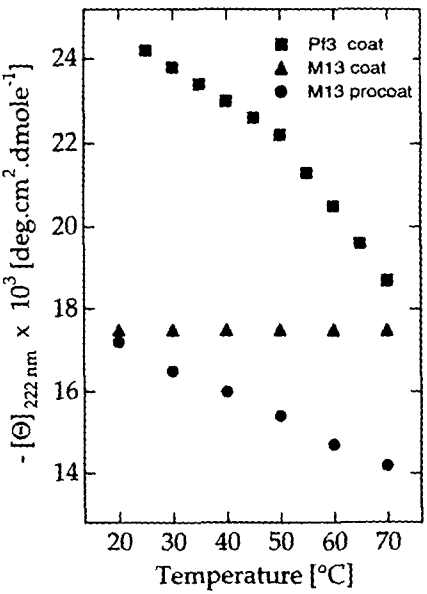

Fig. 8.6. The molar ellipticities at 222 nm, $[\Theta_{222}]$, of M13 procoat protein (•), M13 coat protein (▲), and Pf3 coat protein (■) at different temperatures. Each protein was incorporated in membranes of POPC/POPG lipid vesicles at a lipid to protein molar ratio of 100. Protein concentration was 1 μM, buffer 1 mM Tris-HCl pH 7.4, pathlength of the optical cell 1 cm.

In all experiments thoroughly degassing the samples is an absolute requirement to obtain stable signals in our system (see Chapter 7 for detail). Insufficient degassing leads inevitably to protein aggregation: after an initial irradiation with the 450 W UV lamp used in the CD spectrometer there is a gradual decrease in

CD signal and an increase in turbidity of the sample. This process continues after irradiation has stopped, but does not occur in thoroughly degassed samples.

8.3. IR

8.3.1. INTRODUCTION

The individual absorption bands measured in an infrared spectrum reflect vibrational transitions in the sample molecules. In proteins and polypeptide vibrational transitions of the amide moieties of the peptide backbone are of particular interest, because they can yield information on the secondary structure of the protein. In this context it is important that β-structures can be determined with a much higher liability as compared to CD measurements.

The amide I band in the region of 1600-1700 cm^{-1} is the most sensitive vibrational band for investigating the average protein secondary structure. It originates primarily from the C=O stretching vibrations of peptide amide groups coupled to the in plane N-H bending and C-N stretching modes[51]. The frequency of this vibration depends on the nature of the hydrogen bonds the C=O and N-H moieties of the peptide bond are involved in, and therefore in turn the frequency of the amide I mode reflects the particular secondary structure adopted by the polypeptide backbone. Although, in principle, it would be possible to determine the average protein conformation and to distinguish the different structural elements from their particular amide I vibrational frequencies, this strategy is not straightforward in practice because the amide I band of proteins normally appears as a relatively structureless, broad band in the IR. The analysis will be discussed in the next section. It is based on the same principle as the analysis of CD spectra discussed in the previous section. The spectrum to be analyzed is fitted by a linear combination of a set of IR spectra of reference proteins of known three-dimensional structure.

Due to the anisotropic nature of the vibrational absorption transition moments, it is possible to determine their orientation within a macromolecule such as a polypeptide, if the sample molecules can be macroscopically oriented with respect to the probing infrared radiation. In our particular case, information about the orientation of membrane inserted proteins can be obtained from the attenuated total reflection (ATR) technique. The uniaxially oriented model membranes are brought in direct contact with the reflection plate which must be

of higher refractive index than the surrounding medium. For the following discussion we refer to Fig. 8.7.

Working above the critical angle, the infrared beam is totally reflected within the crystal. From the Fresnell's equations it follows, that an electromagnetic field exists in the rarer medium beyond the reflecting interface, even under conditions of total reflection. This field exhibits the frequency of the incoming light, but the amplitude decreases exponentially with distance z from the surface according to $E(z) = E(0) \exp(-z/d_p)$. The penetration depth d_p is in the order of the infrared wavelength in the reflection element. The use of polarized light is of considerable help for the study of oriented planar lipid membranes. The incident plane-polarized light of parallel orientation ($E_{\parallel}$) is composed of the E_x and E_z components with respect to the plate-fixed coordinate system, whereas the incident perpendicular polarized light ($E_{\perp}$) corresponds only to the E_y component. The orientational distribution of infrared active transition moments are determined by measuring the dichroic ratio R which is defined as the integrated absorption coefficient $A_{\parallel}$ and $A_{\perp}$ with respect to the parallel ($\parallel$) and perpendicular ($\perp$) polarized incident light. Quantitative analysis of the polarized ATR spectra will be discussed in the next section.

8.3.2. METHODS AND MATERIALS

8.3.2.1. Materials

Reference proteins for secondary structure evaluation of IR spectra were from Sigma Chemicals (Eugene, OR, USA). They comprised hemoglobin, myoglobin, papain, prealbumin, cytochrome c, alcohol dehydrogenase, concanavalin A, lysozyme, pepsin, protease, chymotrypsin and triose-phosphate isomerase.

8.3.2.2. Experimental procedures

FTIR measurements were performed using a Bomem model 110 Michelson interferometer (Montreal, Canada) at a resolution of 4 cm^{-1}, equipped with a narrow band, liquid N_2 cooled HgCdTe detector. For each experiment, 200-1000 scans were accumulated, Fourier-transformed, and triangularly apodized. Some spectra were splined with a third order polynomial up to a degradation factor of 4. For attenuated total reflection (ATR) experiments, we used an overhead ATR unit (Specac, Kent, England) made with a trapezoidal-shaped germanium crystal

(50 × 10 × 3 mm^3), an angle of incidence of 45° (Fig. 8.7) and a grid polarizer of KRS5. Spectra of the plain ATR-crystal were recorded for each polarization and subtracted from the respective membrane spectra. The germanium plate was cleaned with chloroform/methanol 1/1 (vol). Planar multi-lamellar lipid membranes with or without peptides were spread from organic solvent on the crystal. Typically 5 µg peptide were used and mixed with the corresponding amount of lipid for a defined lipid-to-protein molar ratio L/P. This corresponds to an average number of bilayers of 1 and 25 for L/P = 2 and 50, respectively.

Three different sample preparations were used; dry peptide-lipid films, the same films under a saturated D_2O vapor atmosphere and films that were hydrated directly with D_2O. Absorption spectra were calculated against blanks (germanium plate with or without D_2O).

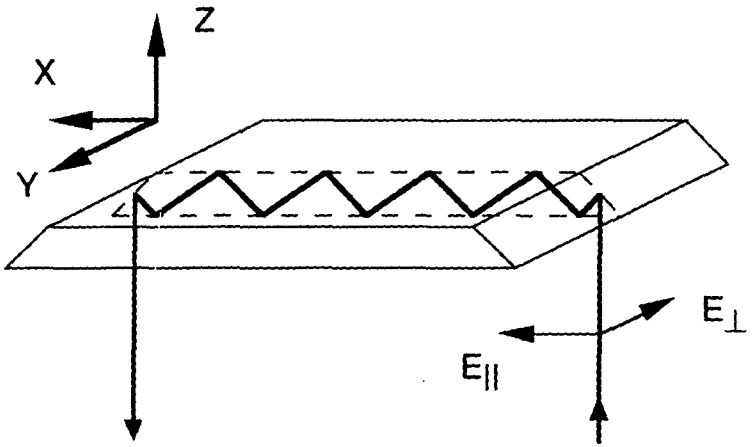

Fig. 8.7. Schematic representation of the ATR-crystal used for polarized infrared measurements of planar lipid bilayers. $E_{||}$ and $E_{\perp}$ define the electric field vectors of the incident polarized light. The coordinate system on the left define the orientations of the electric field vectors E_x, E_y and E_z of the evanescent light wave in the supported, planar lipid film, which is formed directly on the surface of the ATR crystal.

8.3.2.3. Analysis

Quantitative analysis of the polarized ATR spectra was performed as described by Fringeli and Günthard [52]. The dichroic ratio is defined as the ratio of the

absorption polarized parallel to the plane of incidence, $A_{||}$, to that polarized perpendicular, $A_{\perp}$, as $R_{ATR} = A_{||}/A_{\perp}$. For an axially-symmetric molecule with an axially-symmetric distribution of the molecular director around the membrane normal by an angle Θ, the orientational order parameter of the molecular director is defined as $S_{mol} = <3\cos^2\Theta - 1>/2$, describing the space and time averaged fluctuations of the molecular director around membrane normal. Unfortunately S_{mol} is experimentally not directly accessible. Only the orientational order parameter S_v of a particular vibrational transition moment **M** can be determined from the dichroic ratio of the corresponding spectral band according to Fraser and McRae [53] as

$$S_v = [E_x^2 - R_{ATR}E_y^2 + E_z^2] / [E_x^2 - R_{ATR}E_y^2 - 2E_z^2] \qquad (8.15).$$

E_x^2, E_y^2, E_z^2 are the mean square electric fields of the incoming, propagating light wave at the ATR-crystal (Fig. 8.7). In order to calculate the molecular order parameter from S_v, one has to take into account that in general the spectroscopically detected transition moment **M** is oriented at an angle Θ_M from the molecular director. The distribution of **M** around the average molecular director is characterized by $S_M = <3\cos^2\Theta_M - 1>/2$. Furthermore, the orientation of the individual membrane fragments in a multilamellar sample each may not be perfect, which can be described by an orientational order parameter of the mosaic spread, $S_m = <3\cos^2\Theta_m - 1>/2$. Here we assume the membrane normals of the individual fragments to be distributed axial-symmetrically around the laboratory fixed axis z. With this model, the vibrational order parameter is in fact a product of three different order parameters [54],

$$S_v = S_{mol}\, S_M\, S_m \qquad (8.16).$$

In the case of an α-helix, the amide I band appears at around 1657 cm^{-1}. Because, vibrational transitions of non-regular (sometimes called random) structures occur in the same wavenumber range [55], the dichroic ratio of the 1657 cm^{-1} feature is actually a superposition of two different bands, that of the α-helix and that of non-regular structures. This has to be taken further into account for the evaluation of the order parameter of an α-helix from the 1657 cm^{-1} band using the relation $R_{ATR} = [x_\alpha A_{||\alpha} + (1-x_\alpha)A_{||U}] / [x_\alpha A_{\perp\alpha} + (1-x_\alpha) A_{\perp U}]$, where x_α is the fraction of the α-helix band of the 1657 cm^{-1} feature and the indices α, U correspond to the helix and the unordered components, respectively. The orientational order parameter of the 1657 cm^{-1} band is composed of the

orientational order parameter of the α-helix vibration, $S_{v\alpha}$, and that of the non-regular structure transition, S_{vU}, as

$$S_v(1657) = x_\alpha S_{v\alpha} + (1-x_\alpha)S_{vU}$$

$$S_{v\alpha} = S_\alpha S_{M\alpha} S_m \quad \text{and} \quad S_{vU} = S_U S_{MU} S_m \qquad (8.17).$$

Because by definition $S_{vU} = 0$, eq. (8.15) changes for the α-helix to

$$S_{v\alpha} = [E_x^2 - R_{ATR}E_y^2 + E_z^2] / x_\alpha[E_x^2 - R_{ATR}E_y^2 - 2E_z^2] \qquad (8.18).$$

To be able to evaluate the orientational order parameter of the α-helix ($S_{mol} = S_\alpha$), the orientational order parameter of the amide I helix transition moment, $S_{M\alpha}$, with respect to the helix axis has to be known. Two extreme values are reported in the literature with $S_{M\alpha} = 0.41$ [56] and $S_{M\alpha} = 0.60$ [54]. The value of Tsuboi [56] refers to crystals of poly-benzylglutamate. The reported $S_{M\alpha}$ of Rothschild and Clark [54] was determined for bacteriorhodopsin in macroscopically oriented multilayers of purple membranes. Here we use neither of the two values, but calculate from the original polarization measurements of Rothschild and Clark the presently most reliable value of $S_{M\alpha}$ in view of the recently published high resolution structure of bacteriorhodopsin [57]. Applying this three-dimensional structure with the particular tilt angles of the 7 transmembrane helices and the circa 10 helical residues lying flat on the membrane plane, one calculates an average helix order parameter of bacteriorhodopsin $S_\alpha = 0.83$. Taking this value, together with the experimental value of $S_{v\alpha} = 0.488 \pm 0.06$ [54], one obtains from eq. (8.17) the product $S_{M\alpha}S_m = 0.59 \pm 0.07$. In the context of the present work, we will use this value for the product of the two order parameters to calculate the molecular order parameters of the procoat and coat protein helices. Thereby we assume that the mosaic spread of our oriented membranes is comparable to that of the purple membranes in the case of Rothschild and Clark [54].

The mean square electric fields E_x^2, E_y^2, E_z^2 of the incoming, propagating light wave at the ATR-crystal/lipid/air or ATR-crystal/lipid/water interface can be calculated with the corresponding refractive indices of Ge ($n = 4.0$), lipid film ($n = 1.40$), and air ($n = 1.00$) according to the equations given by Fringeli and Günthard [52]: in the amide I band region for the lipid/air interface $E_x^2 = 1.99$, $E_y^2 = 2.13$ and $E_z^2 = 0.59$ and for the membrane-D_2O interface $E_x^2 = 1.97$, $E_y^2 = 2.24$ and $E_z^2 = 1.99$.

The 1500-1800 cm^{-1} region of an infrared spectra was analyzed as a sum of Gaussian/Lorentzian curves. A linear least-squares fit routine was used where consecutively the amplitudes, band positions, halfwidths and Gaussian/Lorentzian composition were optimized. The number and starting positions of each band incorporated in the fit were deduced from second derivative and deconvoluted spectra.

Spectra of peptide/lipid vesicles in water were acquired using an IR cuvette with CaF_2 windows and a pathlength of less than 6 μm. In order to obtain pure protein spectra, the spectra of water and lipids, respectively, were subtracted from the corresponding original spectra. The amide I bands (1600-1700 cm^{-1}) were analyzed against the 12 reference proteins cited above, using a method based on Hennesy and Johnson's procedure described for CD evaluations (see section 8.2.2.3 earlier in this Chapter). The method applied here is comparable with that of Lee et al. [58], as is the correlation between the protein secondary structures obtained by FTIR and X-ray diffraction.

8.3.3. RESULTS

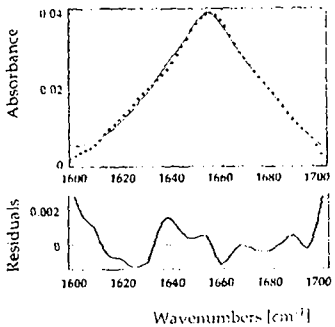

Fig. 8.8. FTIR spectrum of the amide I region of M13 coat protein in an aqueous dispersion of POPC/POPG membranes at 25 °C. Molar lipid to protein ratio was 80. Buffer and lipid spectral components were subtracted. Points represent experimental values, full line corresponds to the fitted spectrum by a superposition of reference spectra. Residuals were calculated as the difference between the fitted and experimental spectra.

8.3.3.1. FTIR Spectroscopy of M13 Coat Protein in Lipid Vesicles

Fig. 8.8 shows the FTIR spectrum in the amide I band of M13 coat protein in an aqueous dispersion of POPC/POPG vesicles. The maximum of the relatively structureless amide I band at 1655 cm^{-1} indicates a predominantly α-helical secondary structure for the membrane bound coat protein. An analysis of the amide I band by a fit with reference protein spectra determines the protein

secondary structure as 57% α-helix, 13% β-structure, 16% turns and 14% other. The α-helix content is in excellent agreement with that determined by CD spectroscopy.

8.3.3.2. ATR-FTIR Spectroscopy of Planar Membranes

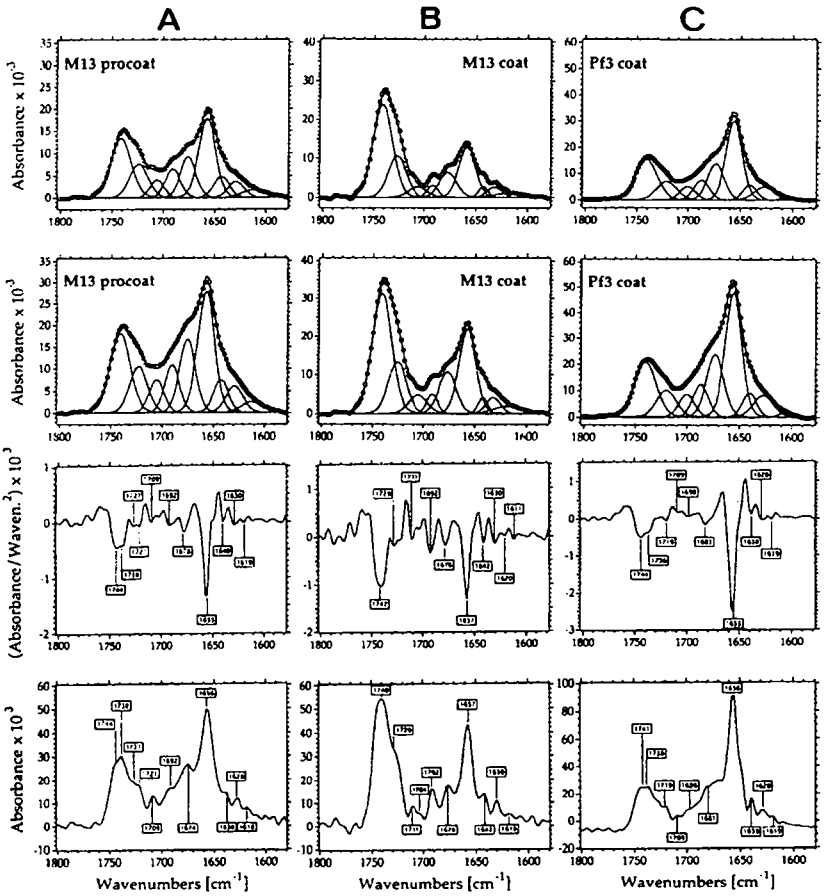

Fig. 8.9. ATR-FTIR of air dried POPC/POPG membranes at 25 °C containing (A) M13 procoat protein at a molar lipid to protein ratio L/P = 40, (B) M13 coat protein at L/P= 80, and (C) Pf3 coat protein at L/P = 20. Shown for each case are the protein spectra for perpendicular polarization (first row, counted from top) and for parallel polarization (second row) together with the best fitted individual bands. Furthermore are shown the second derivative (third row) and the deconvolution (lowest row) of the parallel polarized protein spectra.

The orientational order parameters of the different secondary structure elements of the proteins in lipid membranes have been determined by polarized ATR-FTIR measurements. Planar multilayers of protein/lipid membranes were spread on the surface of a Ge ATR-plate and the infrared spectra were recorded with parallel and perpendicular polarized light, respectively.

Figs. 8.9A, B, C are representative examples of the parallel polarized component spectra of M13 procoat, M13 coat and Pf3 coat proteins in planar, air dried POPC/POPG membranes. In order to resolve the position of the overlapping vibrational components, the second derivative and the deconvoluted spectra are also shown for each particular protein spectrum. Minima in a second derivative spectrum correspond to maxima or shoulders in the original spectrum.

The amide I bands of the three proteins are similar both in the overall shape as well as for the individual resolved spectral components, which are listed in Table 8.3. Subsequently the most prominent spectral components were used to fit the parallel-and perpendicular-polarized FTIR spectra in the 1600-1800 cm^{-1} region, in order to determine the exact values of the polarization of the different spectral transitions. The results are included in Fig. 8.9 for the three proteins. Table 8.3 summarizes the relevant polarization values for all resolved amide I components of the three different proteins, as well as the corresponding calculated order parameters S_v of the vibrational transitions.

The different bands in the amide I region can be assigned according to numerous published studies [51,55,59-61]. The amide I region is dominated by a band at 1657 cm^{-1} which can be assigned to α-helical peptide conformations. However, non-regular structures also appear in the same wavenumber range. The minor bands at 1620 and 1630 cm^{-1} may indicate the presence of β-structures. The small band at 1640 cm^{-1} could be assigned to turns and/or β-structures. The assignment of the bands at 1675 and 1690 cm^{-1} could arise from turns and bends as well as from coupled β-structures. As in the case of CD spectroscopy, different membrane incorporation procedures (see section 8.3.2, 'Materials and Methods' for detail) yielded practically identical protein IR spectra. However, if the proteins are stored over weeks at room temperature in the isopropanol/0.1% TFA solution, the membrane incorporation results in a total different protein conformation, now showing one predominant band at 1630 cm^{-1} (spectrum not shown). Obviously, under this condition the membrane incorporated proteins adopt a preferential β-structure. This fact was also reflected in the CD spectra of old protein solutions (results not shown).

The important finding from the curve fitting in Fig. 8.9 is that the α-helix band at 1657 cm^{-1} is clearly separated from the other spectral features. ATR-FTIR spectra were recorded with membrane preparations of different molar lipid to peptide ratios L/P. The dichroic ratio of the α-helix band of the three proteins is increasing with increasing L/P in the supported lipid membranes, reaching a limiting saturation value at L/P = 20-40 of R_{ATR} = 1.56 for M13 procoat protein, R_{ATR} = 1.71 for M13 coat protein and R_{ATR} = 1.59 for Pf3 coat protein. Using the relations between R_{ATR}, S_v and S_α in eqs. (8.17) and (8.18) one calculates the orientational order parameter of the average α-helix axis for procoat protein as S_α = 0.56-0.71, for M13 coat protein as S_α = 0.73-0.92 and for Pf3 coat protein as S_α = 0.52-0.65, (the two values are due to the uncertainty of $S_{M\alpha}S_m$ = 0.52-0.66 in eq. (8.17); see section 8.3.2, 'Materials and Methods' for detail). This result indicates that the helices are oriented preferentially parallel to the membrane normal.

TABLE 8.3. Assignments, dichroic ratios R_{ATR} and concomitant orientational order parameters S_{vibr} of the vibrational transitions of M13 procoat, M13 coat and Pf3 coat proteins in POPC/POPG air-dried, supported planar lipid bilayers

		R_{ATR}[3] (S_v)[4]					
Band position[5] [cm^{-1}]	Assignment[6]	M13 procoat		M13 coat		Pf3 coat	
1608-1615	β-strand	1.43	(0.21)	1.32	(0.12)	1.79	(0.41)
1628-1631	β-strand	1.63	(0.34)	1.49	(0.25)	1.64	(0.35)
1640-1644	turn/β-strand	1.54	(0.28)	1.50	(0.26)	1.61	(0.32)
1657	α-helix/ non-regular	1.56	(0.37)	1.71	(0.47)	1.59	(0.35)
1675-1676	turn/bend	1.83	(0.43)	1.55	(0.29)	1.71	(0.38)
1688-1690	turn/bend	1.67	(0.36)	1.53	(0.28)	1.63	(0.34)
1702-1705	turn/bend	1.84	(0.43)	1.61	(0.32)	1.66	(0.35)
1722-1723	C=O lipid	1.38	(0.17)	1.21	(0)	1.44	(0.22)
1741	C=O lipid	1.36	(0.15)	1.31	(0.11)	1.35	(0.14)

Upon hydrating the planar membranes with D_2O water-vapor, the amide bond protons, which are accessible to water, exchange to deuterons. In consequence, the deuterated peptide bonds exhibit an altered amide I spectrum. This is shown in Fig. 8.10 for the case of M13 coat protein in POPC/POPG membranes. There is no significant difference of the amide I maximum in H_2O (1657 cm^{-1}) and in D_2O (1656 cm^{-1}). Similar observations were made with the M13 procoat and the Pf3 coat proteins in POPC/POPG membranes as well as for the three proteins with bulk D_2O. According to literature values, a totally deuterated α-helix polypeptide

[3]Values calculated from areas of the different band components obtained by fits to the spectra in Figure 8.9. Typical deviations between R_{ATR} values of different samples at lipid saturation are ±0.02 for the 1657 cm^{-1} band, ±0.07 for the 1608-1615 cm^{-1} band, and ±0.05 for the other features.

[4]S_v calculated according to eq. (8.18) for the bands at 1657 cm^{-1} (x_α=0.8 for M13 procoat and coat proteins and 0.9 for Pf3 coat protein) and according to eq. (8.15) for all other bands.

[5]Indicated is the variation of the position of the bands resolved in Fig.8.8 between the different protein spectra.

[6]According to Surewicz & Mantsch[55] and Byler & Susi[60] for protein amide I bands, and Blume et al.[62] for lipid bands.

backbone, such as in hemoglobin, myoglobin or cytochrome c would give rise to an amide I band at 1650-1651 cm^{-1} [60].

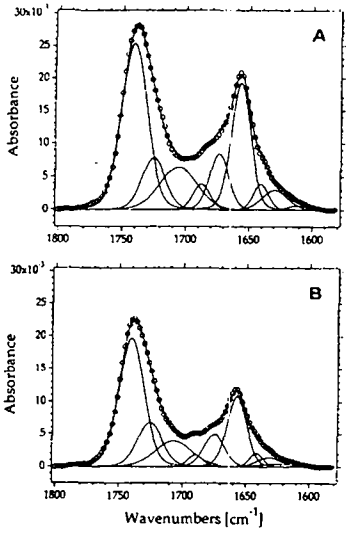

Fig. 8.10. The parallel (A) and perpendicular (B) polarized component of the amide I band of M13 coat protein in POPC/POPG membranes after hydration with D_2O vapor at 25 °C together with the best fitted individual bands.

The results for the three membrane incorporated proteins in the present case indicate that the α-helices are not, or only partially accessible to water within the time range of equilibration (hrs). Compared to the amide I band spectra of the H_2O suspended membranes dried in air, the spectra of the amide I bands of the D_2O hydrated membranes show a substantial intensity decrease between 1670-1700 cm^{-1} and a concomitant increase in the 1620-1645 cm^{-1} range. The relative intensity of the band at 1641 cm^{-1} to that at 1656 cm^{-1} is higher in D_2O than in air/H_2O for all three proteins. A similar observation was made by others with water soluble proteins [63]. In the case of H_2O hydrated proteins, the spectral features of the non-regular protein structures appear around 1657 cm^{-1} together with the vibrational bands of an α-helix. When the corresponding peptide bonds are deuterated, the non-regular protein structure bands are shifted towards 1640-1645 cm^{-1} [60]. Therefore in the deuterated protein spectra of Fig. 8.10 the band observed at 1656 cm^{-1} is a pure "helix" band, free of other structure components. Interestingly, the dichroic ratio is measured to be R_{ATR} = 1.73 in the H_2O/air spectra and R_{ATR} = 1.93 in the deuterated membranes. This corresponds to nearly

identical order parameters of the helix amide I transition moment, namely $S_{v\alpha}$ = 0.48 for H_2O and $S_{v\alpha}$ = 0.46 for D_2O.

Due to severe overlap, the dichroic ratios of the other bands in the amide I region are less well defined than those of the α-helix bands. Small changes in the baseline or in the ratio between Gaussian and Lorentzian line shapes in the curve fitting influence the line-shapes of the non-helix bands more than those of the 1657 cm^{-1} feature. The orientational order parameters of the β-strands as calculated from the dichroic ratios of the β-structure bands at 1620 and 1630 cm^{-1} are calculated to range between S_β = -0.56 and -0.70 (Table 8.3). The physically reasonable lower limit of an orientational order parameter is -0.5. In spite of the uncertainties, the determined S_β values indicate that putative β-strands of the three membrane incorporated proteins are oriented parallel to the membrane plane. A similar conclusion can be drawn from the dichroic ratios of the 1675 and 1690 cm^{-1} band if they indeed correspond to β-structures. If these bands, however, correlate with turn structures, it would be difficult to translate the dichroic ratios to a particular membrane orientation of the turns, because a variety of different turn structures are known, which at present cannot definitely be distinguished by the corresponding infrared spectrum.

The vibrational bands at 1722 and 1741 cm^{-1} of the lipid carbonyl bonds in pure planar multibilayers of POPC/POPG show a polarization ratio of R_{ATR} = 1.15±0.05 (spectra not shown) which corresponds to an orientational order parameter of the vibrational transition of S_v = -0.08±0.08. Similar values have been measured on POPC/POPG multilayers [64]. The values are in agreement with NMR studies of macroscopically aligned multilayers of egg yolk phosphatidylcholine indicating that the lipid carbonyl groups are oriented close to the "magic angle" of 54.7° with respect to the membrane normal which would yield a value of 0 for the order parameter of the carbonyl groups [65]. In the presence of membrane-inserted (pro)coat proteins the dichroic ratios of the lipid carbonyl bands range from R_{ATR} = 1.21 to 1.44. The corresponding vibrational order parameters are S_v = 0 to 0.22.

8.4. DISCUSSION

The preparation methods (see Chapter 6) and the protein reconstitution protocols (see Chapter 7) developed for this work are suited for investigation of the structure and dynamics of the whole membrane insertion and translocation

processes *in vitro*. According to the results of CD and FTIR measurements (sections 8.2 and 8.3, respectively), some general conclusions can be drawn on the structural properties of the three proteins investigated here.

Compared to the nearly totally helical structure in isopropanol/0.1% TFA, there is a considerable decrease of the helix content when the proteins are incorporated into lipid membranes. Because different membrane reconstitution procedures of the proteins result in nearly identical CD and IR spectra, we assume the presence of a single protein population within the lipid membrane. Hence, according to the data in Table 8.2, the helix structure of membrane-inserted M13 procoat protein comprises of 37 amino acid residues, that of M13 coat protein 29-31 residues and that of Pf3 coat protein 33 residues. ATR-FTIR measurements have shown for all three proteins that the helical parts are on average oriented preferentially parallel to the membrane normal, with a substantial higher value of the helix orientational order parameter S_α for M13 coat protein ($S_\alpha = 0.81$) than for both M13 procoat and Pf3 coat proteins ($S_\alpha = 0.64$ and 0.58, respectively). Because the order parameters are derived from ATR-FTIR measurements of air-dried membranes, they reflect primarily the static orientational distribution of the protein helices in a relatively rigid bilayer. Although the polarization measurements in principle do not contain information on the dynamics of protein structure fluctuations it is reasonable to expect that in a fluid lipid bilayer the protein helices perform orientational fluctuations around the membrane normal, with internal structural fluctuations within helix polypeptide backbone superimposed. It was shown elsewhere that for α-helical polypeptides in general and for membrane inserted helices in particular, such structural fluctuations are largest at the helix ends [66,67]. The orientational helix order parameters determined in our case therefore indicate that the mobility (orientational fluctuations of the whole molecule and/or internal structural fluctuations) of the M13 procoat and the Pf3 coat proteins in lipid bilayers are larger than that of M13 coat protein. This behavior corroborates the observation that the thermal stability of M13 coat protein is considerably higher than that of M13 procoat protein and Pf3 coat protein in lipid membranes. It is tempting to speculate that the relative high mobility of M13 procoat and Pf3 coat proteins is correlated with their ability to insert spontaneously into membranes within the natural membrane translocation process. The stability of M13 coat protein at higher temperatures corroborates with earlier experiments measuring amide proton exchange rates with steady-state saturation transfer ^{1}H NMR in SDS micelles [68]. In these experiments it was found that 15-20 amides were very stable at 55°C,

whereas 30 amides were 10^5-fold retarded at 24 °C. These authors, however, did have problems with aggregation at high temperatures and 2 mM protein concentration. This might be caused by an increase in hydrophobic interaction and a decrease in hydrogen bond strengths at high temperatures.

A 100 ps molecular dynamics simulation of M13 coat protein monomers and a head to tail dimeric complex was performed in vacuum where the lipid bilayer was taken into account by a hydrophobic potential. During the simulation the nearly totally helical conformation was stable, showing a bend near Gly 38. Furthermore, the largest mobility was observed near the helix ends where the rms fluctuations of the Cα atoms were maximal. Although this study gives interesting information on the relative stability of the helix along the protein sequence, the suggested model has to be revised in the view of our present work, which shows that nearly half of the membrane incorporated protein is non-helical.

CD and FTIR experiments of the membrane-incorporated proteins reveal the presence of non-helical structures which could formally be assigned to β-strands and turns. However, the interpretation of the spectra in favor of such structures is not as straightforward and unequivocal as in the case of the helical structures. Whether or not the spectral features e.g. in the protein FTIR spectra of Fig. 8.9, listed in Table 8.3, are real β-structures or turns, is difficult to decide. It is interesting to note in this respect that the FTIR spectra of highly helical proteins, such as hemoglobin, myoglobin, cytochrome c and ferritin, dissolved in D_2O, show amide I bands around 1630 and 1670 cm^{-1} which in principle could be assigned to β-structures [60]. However, because according to X-ray structure analysis neither of these particular proteins show β-structures in the generally defined sense, those bands have been assigned to the short extended chains connecting the helical segments [60]. For these proteins, the helix connecting segments consist of 2-6 residues each, which are neither bent into turns nor form sheets. For comparison, our membrane inserted (pro)coat proteins also show bands at 1626 and 1675 cm^{-1} in the deuterated form (Fig. 8.10). Another example is bacteriorhodopsin. Infrared spectra of this intrinsic membrane protein have shown a prominent amide I band at 1660-1663 cm^{-1} representing membrane spanning helices, as well as bands at 1630-1640 and 1684 cm^{-1} which were assigned to β-structures [69,70]. Obviously, this assignment to β-structures has to be revised, in this particular case, according to the recently published three-dimensional structure of bacteriorhodopsin[57]. Based on high resolution electron

diffraction experiments, a structural model of bacteriorhodopsin was presented showing a highly helical membrane protein without β-structures. The bands at 1630-1640 cm^{-1} and 1684 cm^{-1} observed in the IR spectra of bacteriorhodopsin therefore represent likely the loop structures connecting the different transmembrane helices. If β-structures are really present in the case of our coat and procoat proteins, they would be oriented preferentially parallel to the membrane plane as indicated by the dichroic ratios of the corresponding bands (Table 8.3). However, the sometimes rather high dichroic ratio of the different non-helical amide I bands and the concomitant unrealistically low orientational order parameters of the putative β-strands (S_β range from -0.55 to -0.7) argue against a classical β-structure.

Even in aqueous solution, the (pro)coat proteins form a preferential α-helix conformation. Time-resolved fluorescence anisotropy measurements have been performed on these preparations (unpublished results) showing that the proteins form low molecular aggregates composed of about 10 monomers in a buffer of low ionic strength. These low molecular aggregates show a tendency to form higher molecular aggregates with a half-time of several hours. The time course of the second aggregation step is faster in buffer solutions of high ionic strength. The formation of a water soluble form of M13 coat protein was recently reported by others [3].

Both our stock solutions in organic solvent and our reconstituted proteins are stable for several days in the α-helical conformation at low salt when kept at 4° C. In the work of Hemminga and co-workers it was shown [3] that cholate purified M13 coat protein is stabilized in the α-helical conformation by high salt, whereas at low salt concentration irreversible aggregation occurred. At low salt concentration the α-oligomeric form was only found to be maintained in phospholipid systems containing at least one unsaturated acyl chain [2,3]. These differences with our system show once more there is a fundamental difference between our purification/reconstitution methods and those involving detergents.

The following question arises; how can these results be translated into a molecular model of the membrane inserted proteins? Biological studies have shown that both M13 and Pf3 coat proteins span the biological membrane after the membrane insertion process [71]. From our biophysical data the following can be concluded. As has been discussed previously the presently available structural data of membrane proteins indicate that the membrane-traversing protein

8.5. LITERATURE

(1) Thiaudière, E.; Soekarjo, M.; Kuchinka, E.; Kuhn, A.; Vogel, H. **1993** *Biochemistry, 32,* 12186-12196.

(2) Spruijt, R. B.; Wolfs, C. J. A. M.; Hemminga, A. M. **1989** *Biochemistry, 28,* 9158-9165.

(3) Spruijt, R. B.; Hemminga, M. A. **1991** *Biochemistry, 30,* 11147-11154.

(4) Hemminga, M. A.; Sanders, J. C.; Spruijt, R. B. **1992** *Prog. Lipid. Res., 31,* 301-333.

(5) Alberts, B.; Bray, D.; Lewis, J.; Raff, M.; Roberts, K.; Watson, J. D. *Molecular biology of the cell;* third ed.; Garland Publishing, Inc.: New York, London, 1994.

(6) Ito, K.; Mandel, G.; Wickner, W. **1979** *Proc. Natl. Acad. Sci. USA, 76,* 1199-203.

(7) Model, P.; Russel, M.; Boeke, J. D. In *Bacteriophage Assembly;* Alan R. Liss, Inc.: New York, 1981; pp 389-400.

(8) Florine, K. I.; Feigenson, G. W. **1987** *Biochemistry, 26,* 2978-2983.

(9) McDonnell, P. A.; Opella, S. J. **1993** *J. Magn. Res. Series B, 102,* 120-125.

(10) Ven, F. J. M. v. d.; Vanos, J. W. M.; Aelen, J. M. A.; Wymenga, S. S.; Remerowski, M. L.; Konings, R. N. H.; Hilbers, C. W. **1993** *Biochemistry, 32,* 8322-8328.

(11) Henry, G. D.; Sykes, B. D. **1990** *J. Mol. Biol., 212,* 11-14.

(12) Knippers, R.; Hoffmann-Berling, R. **1966** *J. Mol. Biol, 21,* 281-292.

(13) Makino, S.; Woolford, J. L.; Tanford, C.; Webster, R. E. **1975** *J. Biol. Chem., 250,* 4327-4332.

(14) Marvin, D. A.; Hale, R. D.; Nave, C.; Citterich, M. H. **1994** *J. Mol. Biol., 235,* 260-286.

(15) Dunker, A. K.; Klausner, R. D.; Marvin, D. A.; Wiseman, R. L. **1974** *J. Mol. Biol., 81,* 115-117.

(16) Dunker, A. K.; Ensign, L. D.; Arnold, G. E.; Roberts, L. M. **1991** *FEBS Lett., 292,* 271-274.

(17) Marvin, D. A.; Wachtel, E. J. **1976** *Phil. Trans. Roy. Soc. ser. B, 276,* 81-98.

(18) Marvin, D. A. **1989** *Int. J. Biol. Macromol., 11,* 159-163.

(19) Henry, G. D.; Weiner, J. H.; Sykes, B. D. **1986** *Biochemistry, 25,* 590-598.

(20) Henry, G. D.; Weiner, J. H.; Sykes, B. D. **1987** *Biochemistry, 26,* 3626-3634.

(21) Colnago, L. A.; Valentine, K. G.; Opella, S. J. **1987** *Biochemistry, 26,* 847-854.

(22) Leo, G. C.; Colnago, L. A.; Valentine, K. G.; Opella, S. J. **1987** *Biochemistry, 26,* 854-862.

(23) Boguski, M. J.; Leo, G. C.; Opella, S. J. **1988** *Proteins, 4*, 123-130.

(24) Shon, K. J.; Kim, Y.; Colnago, L. A.; Opella, S. J. **1991** *Science, 252*, 1303-1305.

(25) Papavoine, C. H. M.; Konings, R. N. H.; Hilbers, C. W.; Ven, F. J. M. v. d. **1994** *Biochemistry, 33*, 12990-12997.

(26) Nozaki, Y.; Chamberlain, B. K.; Webster, R. E.; Tanford, C. **1976** *Nature, 259*, 335-337.

(27) Fodor, S. P. A.; Dunker, A. K.; Ng, Y. C.; Carsten, D.; Williams, R. W. In *Bacteriophage Assembly*; M. S. Dubow, Ed.; Alan R. Liss: New York, 1981; pp 441-455.

(28) Kazmierczak, B. I.; Mielke, D. L.; Russel, M.; Model, P. **1994** *J. Mol. Biol., 238*, 187-198.

(29) Jongh, H. H. J.; Hemminga, M. A.; Marsh, D. **1990** *Biochim. Biophys. Acta, 1024*, 82-88.

(30) Marsh, D. **1990** *FEBS Lett., 268*, 371-375.

(31) Dettman, H. D.; Weiner, J. H.; Sykes, B. D. **1984** *Biochemistry, 23*, 705-712.

(32) Wolfs, C. J. A. M.; Horvath, L. I.; Marsh, D.; Watts, A.; Hemminga, M. A. **1989** *Biochemistry, 28*, 9995-10001.

(33) Burnell, E.; Alphen, L. v.; Verkleij, A.; Kruijff, B. d. **1980** *Biochim. Biophys. Acta, 597*, 492-501.

(34) Kusters, R.; Breukink, E.; Gallusser, A.; Kuhn, A.; Kruijff, B. d. **1994** *J. Biol. Chem., 269*, 1560-1563.

(35) Galluser, A.; Kuhn, A. **1990** *EMBO J., 9*, 2723-2729.

(36) Chamberlain, B. K.; Webster, R. E. **1976** *J. Biol. Chem., 251*, 7739-7745.

(37) Datema, K. P.; Wolfs, C. J. A. M.; Marsh, D.; Watts, A.; Hemminga, M. A. **1987** *Biochemistry, 26*, 7571-7574.

(38) Sanders, J. C.; Poile, T. W.; Wolfs, C. J. A. M.; Hemminga, M. A. **1992** *Biochim. Biophys. Acta, 1110*, 218-224.

(39) Colnago, L. A.; Leo, G. C.; Valentine, K. G.; Opella, S. J. In *Biomolecular stereodynamics, proceedings of the fourth SUNYA conversation in the discipline biomolecular stereodynamics*; R. H. Sarma and M. H. Sarma, Eds.; Adenine: Guilderland, NY, 1986; Vol. 3; pp 147-158.

(40) Sanders, J. C.; Nuland, N. A. J.; Edholm, O.; Hemminga, M. A. **1991** *Biophys. Chem., 41*, 193-200.

(41) Sanders, J. C.; Ottaviani, M. F.; Hoek, A. v.; Visser, A. J. W. G.; Hemminga, M. A. **1992** *Eur. Biophys. J., 21*, 305-311.

(42) Peelen, S. J. C. J.; Sanders, J. C.; Hemminga, M. A.; Marsh, D. **1992** *Biochemistry, 31*, 2670-2677.

(43) McDonnell, P. A. Ph. D. Thesis, University of Pennsylvania, 1992.

segments which are in contact with the hydrophobic part of a lipid bilayer fold in general as α-helices or β-strands, in order to saturate hydrogen bonds [57,72-74]. In the case of the investigated coat proteins, it is reasonable to assume that the central hydrophobic parts traverse the lipid bilayer as an α-helix. This would also explain the low accessibility for water of the M13 coat protein helices in POPC/POPG membranes as measured by FTIR with H_2O-D_2O exchange experiments, as is the case for Pf1 coat protein [75]. Our results are supported by NMR hydrogen exchange measurements on SDS solubilized M13 coat protein. There the highest exchange rates were measured at the N- and C-termini , while the lowest exchange rates were found for the central hydrophobic segment of the M13 coat protein [76]. Whether the predicted transmembrane helix extends in form of a single helix over 30 or 33 amino acid residues for M13 coat and Pf3 coat proteins, respectively, cannot be decided from our spectral data.

Our models of the structure of the membrane inserted coat proteins are in agreement with recently published models derived from NMR measurements of M13 coat protein in SDS micelles [10,11,25] .

Based on further results these models have been extended for (pro)coat proteins reconstituted into lipid membranes (see Fig. 8.11). The structure in SDS is generally thought to be very close to that in a lipid membrane, although SDS is also criticized as a bad detergent, which poorly covers the surface of M13 coat (Marcus Hemminga, personal communications). Both for M13 coat protein and Pf1 coat protein it is assumed that the membrane spanning helix is rigid and that on the N-terminal side another helix is formed which lies flat on the membrane or micelle [10,24,25], but which is much less stable, and might be unfolded part of the time. The model proposed in Fig. 8.11 is, however, severely critized (see legend). The structure and dynamics of N-terminal part of the reconstituted proteins still need further investigation.

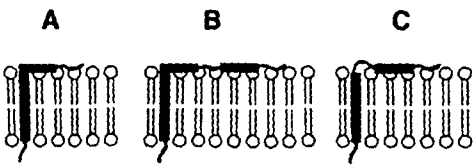

Fig. 8.11. Proposed models for membrane topology of A: fd coat protein, B: M13 procoat protein, C: Pf1 coat protein. Figure taken from [77]. The proposed structure for M13 procoat

protein is based on more sketchy information [43] (see the introduction to this Chapter) and shows a reconstituted conformation which is not in accordance with the in vivo topology as judged by biological studies (see Chapter 5). Another criticism is that Opella and co-workers assume that the secondary structure is almost completely α-helical. They cannot exclude some degree of β-aggregation in their samples. Trans-membrane β-strands would also show H-bonds parallel to the membrane surface, just as the postulated helix on the membrane surface would.

The difference between M13 coat and Pf1 coat protein is that only the first seems to have a hinge about Ile-22, whereas the latter has an unordered loop (between Thr-13 and Met-19) connecting both helices, as is the case in the virion [78], (see Fig. 5.8). The dynamics of the N-terminal part of the mature coat proteins is as yet unknown. In molecular dynamics simulations starting from a totally α-helical conformation either a final structure with a 90 degree bend can be found (for Pf1 coat) [79] , or a totally α-helical, almost straight structure (for M13 coat) [40]. This also points to a large mobility of the N-terminal part of the proteins, which could, *in vivo*, play a role in the phage assembly process.

In the case of a structural model of membrane inserted M13 procoat protein, our data can be discussed with similar arguments as in the case of the two coat proteins. Biological studies indicate that procoat protein inserts into membranes in form of a loop structure (see Chapter 5 for more detail). Because the ATR-FTIR experiments have shown that β-structures, if really present, would be oriented parallel to the membrane plane, the only candidates for the membrane-spanning protein parts are the α-helical segments. Due to the loop structure at least two helix segments must exist in a procoat molecule, most reasonably the hydrophobic segments in the leader and in the mature coat sequence. The number of 37 amino acid residues in a procoat molecule which according to CD spectra form the helical part fits nicely to this idea. As a consequence of this model, the helical parts in M13 coat protein extend differently than the helical part of the coat sequence in procoat protein.

Our data on the structure of M13 procoat and Pf3 coat proteins are the first in this field. On the other hand, the investigations presented in this work have clarified the long dispute on the helix structure of M13 coat protein in lipid membranes (50-90%) [24,26,27]. Our data for the secondary structure of M13 coat protein have been corroborated by similar results achieved with improved protocols based on phenol extraction and cholate reconstitution by others [80].

(44) Gall, C. M.; DiVerdi, J. A.; Opella, S. J. **1981** *JACS, 103,* 5039-5043.
(45) Gall, C. M.; Cross, T. A.; Diverdi, J. A.; Opella, S. J. **1982** *Proc. Natl. Acad. Sci. USA, 79,* 101-105.
(46) Valentine, K. G.; Schneider, D. M.; Leo, G. C.; Colnago, L. A.; Opella, S. J. **1985** *Biophys. J., 49,* 36-38.
(47) Greenfield, N.; Fasman, G. D. **1969** *Biochemistry, 8,* 4108-4116.
(48) Hennesy, J. P.; Johnson, W. C. J. **1981** *Biochemistry, 20,* 1085-1094.
(49) Vogel, H. **1987** *Biochemistry, 26,* 4562-4572.
(50) Levitt, M.; Greer, J. **1977** *J. Mol. Biol., 114,* 181-293.
(51) Krimm, S.; Bandekar, I. **1986** *Adv. Protein Chem., 38,* 183-364.
(52) Fringeli, U. P.; Günthard, H. H. In *Membrane Spectroscopy;* E. Grell, Ed.; Springer-Verlag: Berlin, 1981; pp 270-332.
(53) Fraser, R. D. B.; MacRae, T. P. *Conformation in fibrous proteins and related synthetic peptides;* Academic Press: New York, 1973.
(54) Rothschildt, K. J.; Clark, N. A. **1979** *Biophys. J., 25,* 473-488.
(55) Surewicz, W. K.; Mantsch, H. H. **1988** *Biophys. Biochem. Acta, 952,* 115-130.
(56) Tsuboi, M. **1962** *J. Polymer sci., 59,* 139-153.
(57) Henderson, R.; Baldwin, J. M.; Ceska, T. A.; Zemlin, F.; Beckmann, E.; Downing, K. H. **1990** *J. Mol. Biol., 213,* 899-929.
(58) Lee, D. C.; Haris, P. I.; Chapman, D.; Mitchell, R. C. **1990** *Biochemistry, 29,* 9185-9193.
(59) Haris, I. P.; Lee, D. C.; Chapman, D. **1986** *Biochim. Biophys. Acta, 874,* 255-265.
(60) Byler, D. M.; Suzi, H. **1986** *Biopolymers, 25,* 469-487.
(61) Surewicz, W. K.; Moscarello, M. A.; Mantsch, H. H. **1987** *J. Biol. Chem., 262,* 8598-8602.
(62) Blume, A.; Hübner, W.; Messner, G. **1988** *Biochemistry, 27,* 8239-8249.
(63) Yang, P. W.; Mantsch, H. H.; Arrondo, J. L. R.; Saint-Girons, I.; Guillou, Y.; Cohen, G. N.; Bârzu, O. **1987** *Biochemistry, 26,* 2706-2711.
(64) Frey, S.; Tamm, L. **1991** *Biophys. J., 60,* 922-930.
(65) Braach-Makvytis, V. L. B.; Cornell, B. A. **1988** *Biophys. J., 53,* 839-843.
(66) Vogel, H.; Nilsson, L.; Rigler, R.; Voges, K. P.; Jung, G. **1988** *Proc. Natl. Acad. Sci. USA, 85,* 5076-5071.
(67) Daggett, V.; Levitt, M. **1992** *J. Mol. Biol., 223,* 1121-1138.
(68) O'Neill, D. J.; Sykes, B. D. **1988** *Biochemistry, 27,* 2753-2762.
(69) Jap, B. K.; Maestre, M. F.; Hayward, S. B.; Glaeser, R. M. **1983** *Biophys. J., 43,* 81-89.

(70) Lee, D. C.; Hayward, J. A.; Restall, C. J.; Chapman, D. **1985** *Biochemistry, 24,* 4364-4373.

(71) Kuhn, A.; Troschel, D. In *Membrane biogenesis and protein targeting;* W. Neupert and R. Lill, Eds.; Elsevier: New York, 1992; pp 33-47.

(72) Deisenhofer, J.; Epp, O.; Miki, K.; Huber, R.; Michel, H. **1985** *Nature, 318,* 618-624.

(73) Weiss, M.; Kreusch, S.; Schiltz, E.; Nestel, U.; Welte, W.; Weckesser, J.; Schultz, G. **1991** *FEBS Lett., 280,* 379-382.

(74) Vogel, H. **1992** *Quart. Rev. Biophys., 25,* 433-457.

(75) Azpiazu, I.; Gomez-Fernandez, J. C.; Chapman, D. **1993** *Biochemistry, 32,* 10720-10726.

(76) Molday, R. S.; Englander, S. W.; Kallen, R. G. **1972** *Biochemistry, 11,* 150-159.

(77) Opella, S. J.; McDonnell, P. A. In *NMR of proteins;* G. M. Clore and A. M. Gronenborn, Eds.; Macmillan: London, *1993;* pp 159-189.

(78) Nambudripad, R.; Stark, W.; Opella, S. J.; Mokowski, L. **1991** *Science, 252,* 1305-1308.

(79) Tobias, D. J.; Klein, M. L.; Opella, S. J. **1993** *Biophys. J., 64,* 670-675.

(80) Sanders, J. C.; Haris, P. I.; Chapman, D.; Otto, C.; Hemminga, M. A. **1993** *Biochemistry, 32,* 12446-12454.

9. TOPOLOGY OF RECONSTITUTED PROTEINS

9.1. INTRODUCTION

A central question in the assembly pathway of bacteriophage M13 (Ff) is how the major coat proteins are located in the inner membrane of their host during the replicative phase, i.e., which parts of the membrane inserted protein are buried in the hydrophobic core of the membrane and which parts protrude into the cytoplasm and into the periplasmic space.

Limited proteolysis is a traditional and efficient technique in analyzing the structure and topology of protein complexes which has been used for over 45 years [1]. It is a classical method to investigate the topology of membrane-bound proteins to determine the extent of their protease accessibility either in the native membranes or reconstituted into lipid vesicles. Using different proteases a different cleavage patterns will be obtained, since each protease has its own specificity. Proteinase K and pronase [1] are relatively unspecific proteases. Chymotrypsin typically hydrolyses peptide bonds on the C-terminal side of tyrosine, phenylalanine, tryptophan, and to some extent, leucine and methionine; trypsin cleaves after lysine and arginine residues. Cyanogen bromide cleaves specifically after methionine residues (see TABLE 9.1).

[1]Pronase is a mixture of different proteases.

TABLE 9.1. Protease specificity

protease	hydrolysis on the C-terminal side of :
proteinase K	hydrophobic residues
pronase	unspecific
chymotrypsin	Tyr, Phe, Trp, (less efficiently: Leu, Met)
trypsin	Lys, Arg
CNBr	Met

In order to identify the resulting fragments after proteolysis the classical approach is to incorporate different labeled amino acids at strategic positions in the protein sequence by exchanging one or more of the amino acids for their radioactively labeled counterparts in the growth medium of the bacteria that produce the protein of interest. In all of the studies published on Ff coat proteins two different labels were introduced into each protein molecule, a tritium labeled amino acid marker in one or both of the hydrophilic terminal stretches (Pro, Lys, Phe or Leu) and another marker in the hydrophobic core ([^{14}C]-Tyr or [^{14}C]- or [^{35}S]-Met), yielding sensitive probes for determining the accessibility of certain portions of the coat protein to the proteolytic enzymes added to the outside after reconstitution. After digestion the fragments were analyzed on SDS-PAGE with autoradiography. Comparing the hydrolysis products of differently labeled proteins then gives an indication about which part of the coat protein is protected by the bilayer [e.g. 2-4].

In Fig. 9.1 three possible membrane topologies of coat protein (or of any other single spanning membrane protein) are depicted. In principle a protein molecule with a single hydrophobic stretch is assumed to either traverse the bilayer (Fig. 9.1A), or to be in a surface bound conformation on either side of the membrane (Fig. 9.1B and C). Comparison of label elution patterns of coat proteins fragments containing different combinations of labels enables an approximate determination of the accessible cleavage sites and thus of that part inserted in and protected by the hydrophobic part of the biological membrane, micelle or reconstituted lipid bilayer.

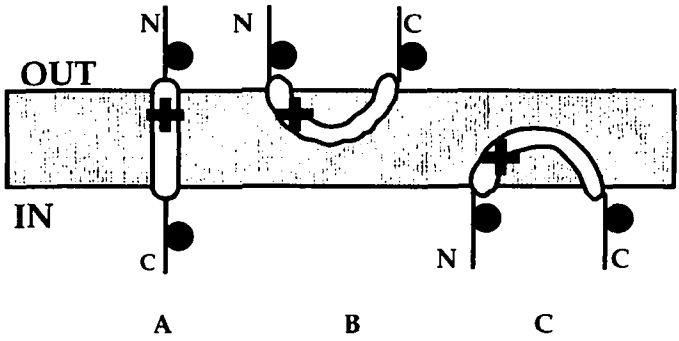

Fig 9.1. Membrane topologies of coat proteins. • : ^{3}H- labeled residue in the hydrophilic part of the protein; + : ^{14}C or ^{35}S labeled residue in the hydrophobic part of the protein. (A) transmembrane configuration, (B) and (C) surface bound configurations. In the case of a surface bound configuration without hydrophobic partitioning (not shown) the whole protein is thought to be digested and no large fragments containing either kind of label are expected.

9.2. REVIEW OF PREVIOUS STUDIES ON MEMBRANE AND MICELLE TOPOLOGY OF M13 AND RELATED COAT PROTEINS

9.2.1. *IN VIVO* TOPOLOGY

The accessibility of the coat protein antigenic site (residues 1 to 8) was determined quantitavely monitoring the specific binding of ^{125}I-labeled, affinity purified antibody to coat protein in the following systems: osmotically shocked *E.coli* cells, spheroplasts (right-side-out inner membrane vesicles) and sonicated spheroplasts (inside-out inner membrane vesicles) [5,6].

Both parental and newly synthesized M13 coat protein were found to be located in the cytoplasmic membrane of *E. coli* with the negatively charged N-terminus exposed on the outside and the positively charged C-terminus on the inside of the membrane, even though they insert into the membrane from opposite directions [5-10].

The parental coat protein is later incorporated dispersively into progeny virus [11]. However, biological recycling of old coat protein might be adventitious and

insignificant compared to the bulk of the newly synthesized protein [12]. Moreover, it has been observed [13,14] that parental coat is included into new phages with a half-time of 1 to 2 hours, whereas newly synthesized coat is incorporated with a half-time of a few minutes.

The newly synthesized coat originates from the cytosol, the parental coat protein from the external medium. However, the two insertion processes also differ in a more fundamental way. The newly synthesized procoat contains a leader peptide which is essential for correct insertion, whereas the parental coat is part of the phage structure, which interacts in an asymmetric way with the target membrane during host infection. Furthermore, the viral DNA is transported over the host membrane into or out of the bacteriophage in a very controlled manner without lysing the host's inner membrane. It is probable that the parental coat protein inserts into the membrane in a concerted and asymmetric way with the de-capsication and entrance of the viral DNA. The encapsulation and budding process requires some additional phage encoded protein which are not present during infection (see Chapter 5, section 5.3.2.3). Infection with phages and assembly/extrusion of newly formed phages could be processes which proceed in opposite directions, but share some fundamental characteristics. The fact that both parental and newly formed coat proteins are found to be included in new phages is in accordance with this idea.

In conclusion, it is evident that in nature there exist two fundamentally different ways of reaching the same membrane inserted topology for M13 coat protein.

9.2.2. TOPOLOGY AFTER RECONSTITUTION

All previous protease digestion studies of M13 (Ff) coat proteins reconstituted into lipid bilayers were performed in the presence of detergents. It has been shown that even after extensive dialysis (4 x 12 h.) not all cholate can be removed, resulting in a final cholate/coat protein molar ratio of at least 1.7 [2]. Moreover, it has been shown by ESR [16], that M13 coat protein preferentially interacts with a fatty acid probe (14-SASL) in mixed bilayers of DMPC and PA. This result was corroborated in another study by the same group [17].

Therefore, it is likely that the results of previous protease digestion studies should be interpreted as valid for a tightly bound coat protein detergent complex. The digestion protocols used are all based on an early study of Woolford and Webster [3], who used a 3-15 fold (by mass) excess of protease and an incubation

overnight at 37°C with vigorous shaking. For each experiment the digestion products were characterized after size exclusion chromatography in the presence of detergent or in formic acid/ethanol 70/30 [2].

In all studies the hydrophobic core of the protein (residues 21-39) was found to be protected from protease by detergent or lipids, or because of aggregation. On the other hand, several enzymes were found to be able to hydrolyze certain bonds in the hydrophilic segments of the coat protein.

9.2.2.1. F1 coat in deoxycholate micelles

F1 coat protein binds to desoxycholate at a detergent to coat protein ratio of 16 [3], yielding normal sized micelles. In this system trypsin was 90 % efficient in cutting at Lys-8, but only 10 % efficient at cutting at Lys-43 and/or Lys-44 and not at all at Lys-40, directly adjacent to the hydrophobic stretch. Inversely, α-chymotrypsin could cleave at Leu-41 and/or Phe-42 (90-95% efficiency), but only with a 10% efficiency at Phe-11 and not at Leu-14 or at Tyr-21 or -24. This result implies that both termini are accessible in micelles, as expected. However the difference in action between chymotrypsin and trypsin shows that a failure to hydrolyze a specific amide bond with one protease does not necessarily mean this bond is not accessible at all to protease attack.

9.2.2.2. M13 coat protein reconstituted into lipid vesicles by cholate dilution

A conformation of M13 coat protein with both termini exposed to the exterior of DPPC or DPPS vesicles could be induced at temperatures well below the lipid-phase transition temperature (T_m) [4]. Around T_m limited cleavage by chymotrypsin (through addition of TPCK) occurs at Phe-11, as judged by the disappearance of the tritium label in Lys-8.

Bayer and Feigenson [2] incorporated M13 coat protein into egg PC and egg PC / egg PA vesicles by means of different reconstitution protocols involving the use of cholate. Their main problem was one of aggregation of coat protein with itself and with protease.

In all of their digestion experiments 20-50 % of the protein was recovered after chromatography as an aggregate peak at the void volume. This peak could not be resolved into separate aggregates of whole coat protein molecules and (hydrophobic) fragments. At high coat protein concentrations (0.5-2.5 mg/ml)

and high α-chymotrypsin concentrations fragments associated with denatured α-chymotrypsin were also found at the void volume. Moreover, fragments with the N-terminus cleaved off eluted at about the same volume as fragments with both ends digested, whereas fragments lacking the C-terminus eluted directly after intact coat protein.

Almost all molecules were found to span the bilayer with the N-terminus facing outwards [2]. No significant amount of coat protein was found inserted in the U-shape conformation with both its ends accessible to protease. 4 to 12 % (depending on the reconstitution protocol) of the molecules were spanning the membrane with the C-terminus pointing outwards. The pronase protected core was found to be Lys-8 to Lys-43 or smaller. Chymotrypsin could cut at both C- and N-terminus, the resulting chromatographic peak might correspond to a dimer of the fragment consisting of the residues 25-42. Cleavage at the C-terminus was much less efficient than cleavage at the N-terminus; using several carboxypeptidases, no conditions could be found under which the C-terminus was digested more rapidly than the release of Pro-8.

In view of the aggregation phenomena described before it cannot be excluded that a great part of the protein was not inserted correctly into the membrane.

9.3. MASS SPECTROSCOPY ON INTACT PROTEINS AND THEIR PROTEOLYTIC DIGESTS

9.3.1. INTRODUCTION

The mass of increasingly large molecules can be accurately measured by means of mass spectroscopy. The latest developments have made it possible to measure proteins and protein fragments. In some cases it even has become possible to measure in the presence of species which up until recently would exclude proper mass determination, like detergents or lipids. This opens the possibility of determining the sequence of protein fragments remaining in a reconstituted system after enzymatic digestion (as will be shown in this chapter for our proteins), as well as characterizing the intact purified proteins in organic solution.

9.3.2. METHODS AND MATERIALS [18]

Two basic strategies have been used to create gaseous protein ions from our samples. These ions are subsequently separated by mass to charge ratio m/z in the analyzer of the mass spectrometer.

9.3.2.1. Electrospray mass spectroscopy (ES-MS)

Electrospray Mass Spectroscopy (ES-MS) allows mass determination of intact proteins. Proteins with a mass of up to 150 kDa have been measured. Multiply charged ions (as opposed to other desorption techniques) are produced with great ease directly in solution. In principle all sample molecules are ionized. This greatly enhances the sensitivity of the technique. On the other hand adducts are formed with water molecules and/or one or more salt ions present in the sample, which can lead to a multitude of peaks for a single sample molecule. Most gaseous proteinaceous ions carry on average about 1 charge per 2000 Da due to (de)protonation. In our case this gives rise to several peaks carrying three to six charges. Taking two or more peaks as a basis, assuming that they represent differently charged ions of the same species, the mass of this species is calculated and the spectrum is checked for additional peaks belonging to the same basis. After transferring the m/z spectrum to a mass spectrum the region around the main peak can be used to determine purity of the protein and the occurrence of mutations in its sequence and/or post translational modifications.

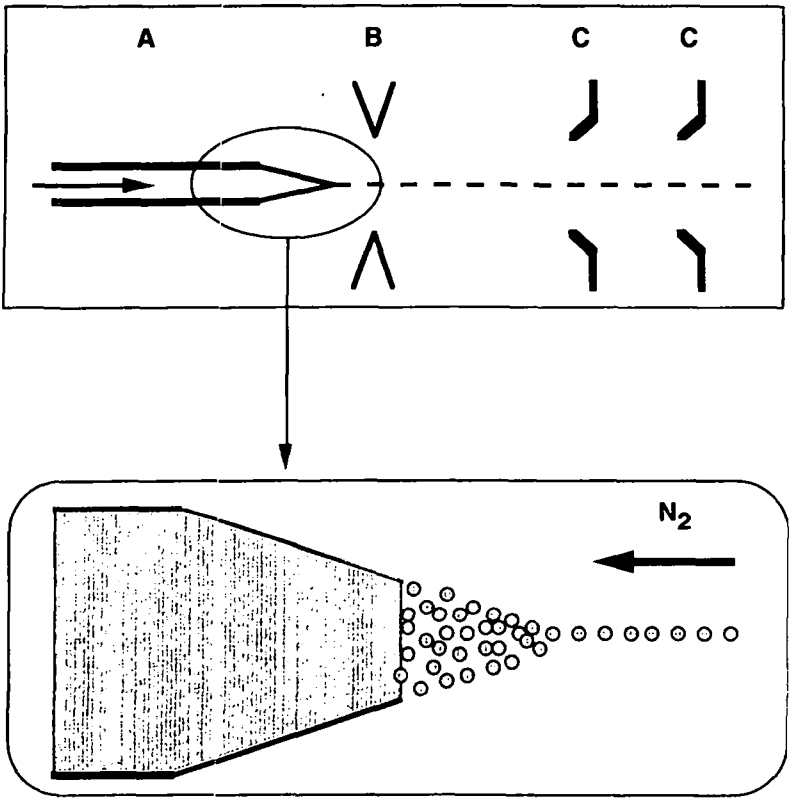

Fig. 9.2. Principle of electrospray mass spectroscopy (ES-MS): (A) injection capillary, (B) circular electrode causing the point effect visualized in the enlargement, (C) deflection screens. See text for more details.

In the most widely used positive mode a solution of sample molecules in water or in a mixture of water and organic solution is injected into the mass spectrometer (see Fig. 9.2) at a high positive electrical potential (typically 4000 V). A subsequent circular electrode held at a lower positive potential (for example 1000 V) accelerates the sample and creates the so called *point effect*. The liquid stream is focused into a point by electrical field in the vicinity of the electrode. This causes the stream to explode into small droplets, each ideally containing a single ionized protein and solvent molecules. These droplets are further accelerated and pass through to deflection shields to yield a parallel stream of droplets. The solvent molecules are evaporated by a counter stream of nitrogen. The resulting

gaseous sample molecules then enter the analytical part of the mass spectrometer.

In our experiments organic stock solutions of M13 and Pf3 coat proteins and of M13 procoat protein (800 μM, 50 μM and 30 μM, respectively) were diluted 1/50 into the solvent system, water/acetonitrile 1/1, 1 % formic acid. A 10 μl sample loop was used. At a solvent flow rate of 4 μl/min. 10-15 sweeps could be performed over the region of interest for our expected peaks, m/z 800 - 1600, the latter mass being the detection limit. The results of all sweeps was summed before analysis.

9.3.2.2. Matrix assisted laser desorption mass spectroscopy (MALDI MS)

Matrix Assisted Laser Desorption Mass Spectroscopy (MALDI-MS) takes advantage of a short pulse of photons for desorption and ionization. The sample is mixed with approximately 1000-fold excess of matrix molecules in organic solution. A drop of this mixture is applied to the sample holder and dried at room temperature. The sample is irradiated by a UV laser which excites the matrix molecules. The excited matrix molecules partly transfer their energy to the sample molecules. Gaseous ions of both sample and matrix molecules are formed and accelerated by an electrical field into the time of flight analyzer of the mass spectrometer. Most ions carry either a single or a double charge and are therefore detected at mass/charge ratios m/z equal to or half a large as their molecular weight MW. The mass range, sensitivity and tolerance of MALDI-MS are unsurpassed. Non-covalently linked subunits may be desorbed as one unit.

Enzymatic digestion of reconstituted proteins: 1.85 nmole protein was incorporated into lipid vesicles using the "Mixed Film" method described in Chapter 7. The molar lipid/protein ratio was approximately L/P=40 for all experiments. 2 and 4 % by weight relative to the substrate of proteinase K was added to the protein/vesicle preparation and the sample was incubated for 1 hour or longer. MALDI-MS spectra were taken of digested samples and of reference samples containing the reconstituted proteins in the absence of proteinase K. Peptide fragments were isolated by HPLC as described below, characterized by MALDI-MS and sequenced by Edman degradation as described below.

Peptide purification: The peptides generated by enzymatic digestion were separated by HPLC on a Nucleosyl 300-10-C4 column thermostated at 60°C. For

M13 coat protein a gradient of 90 % acetonitrile/0.08 % TFA in 0.1 % TFA (aq) was used. For Pf3 coat protein the organic phase used in the gradient was 80 % isopropanol/20 % acetonitrile/0.05 % TFA. The HPLC system consisted of two LKB 2150 pumps (Phamacia Biotechnology, Sweden) controlled by a PC and a LKB variable wavelength detector operated at either 214 or 280 nm.

Edman degradation: The intact protein and selected peptide fragments were sequenced by Edman degradation using a Knauer model 810 pulsed-liquid sequencer (Knauer, Berlin). A polybrene coated polyvinylfluoride (PVDF) membrane was immersed in the HPLC fraction containing the peptide to be sequenced. The Eppendorf cup was agitated for some hours before drying the fraction with the membrane in the speed-vac. Degradation and conversion was performed as described by the supplier. Identification of the liberated PTH-amino acid derivatives was performed by HPLC over a Knauer narrow bore column (250 x 2 mm) for separation of PTH-amino acids derivatives.

9.4. RESULTS

9.4.1. Characterization of intact proteins

9.4.1.1. Electrospray

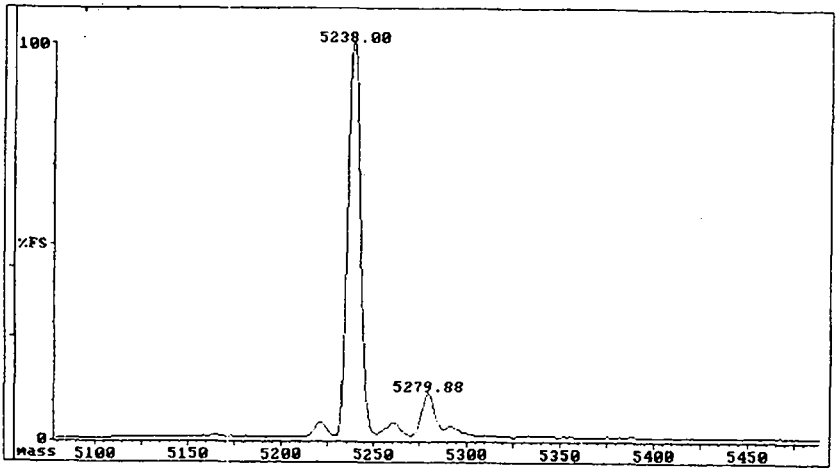

Fig. 9.3. ES-MS spectrum of M 13 coat protein. Transferring of the m/z spectrum to a mass spectrum was based on three values for the base peak at MW = 5238.0 ± 0.3 Da.

The base peak (BP) of the spectrum of M 13 coat in Fig. 9.3 is found exactly at the expected value. The small peaks are probably artifacts. the peaks at BP - 16.5 and +54 are commonly found in all ES-MS spectra taken of our three proteins. The peak at BP + 23 could represent a Na^+ adduct and the peak at BP + 42 an adduct of acetonitrile.

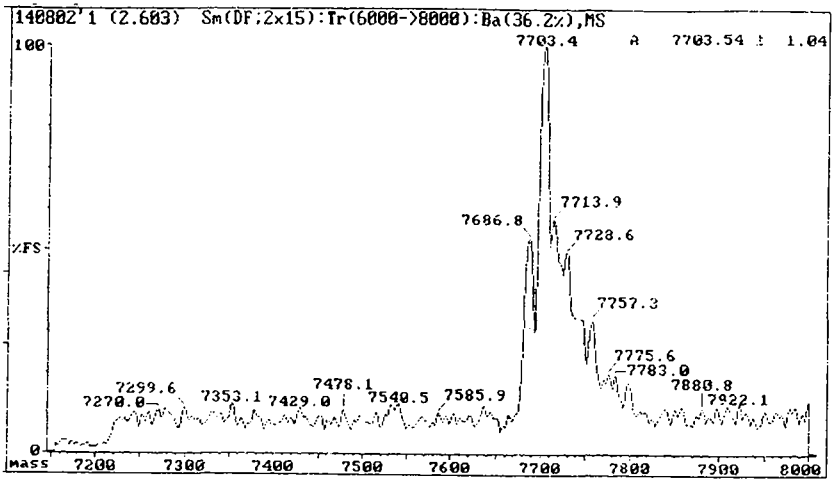

Fig. 9.4. ES-MS spectrum of M 13 procoat protein. Transferring of the m/z spectrum to a mass spectrum was based on three values for the base peak at MW = 7703.4 ± 1.3 Da.

The base peak of the spectrum shown in Fig. 9.4 is thought to represent an adduct of water to H5 procoat protein (calculated MW 7685). The signal to noise ratio is less advantageous than in the spectrum shown for M13 coat protein (Fig. 9.3) because of the lower concentration of the protein stock solution (30 μM compared to 800 μM in the case of M13 coat protein).

The low signal to noise ratio in the spectra of Pf3 coat protein was insufficient to find a base of several differently charged molecular peaks permitting to transfer the m/z spectrum to a mass spectrum. Thus the mass determination is only based on one peak. This protonated base peak was found (results not shown) at m/z = 1552.5. Assuming this peak represents the molecular species with a threefold charge the MW would be 4654. The calculated MW of Pf3 coat protein being 4630.5 the peak could, within the experimental error represent the

formylated form of Pf3 coat protein. Formylation of Pf3 would reduce its charge, bringing it to the detection limit. This is in accordance with the low signal to noise ratio of mass spectra of Pf3 compared to those of M13 procoat.

9.4.1.2 Matrix Assisted Laser Desorption Mass Spectroscopy

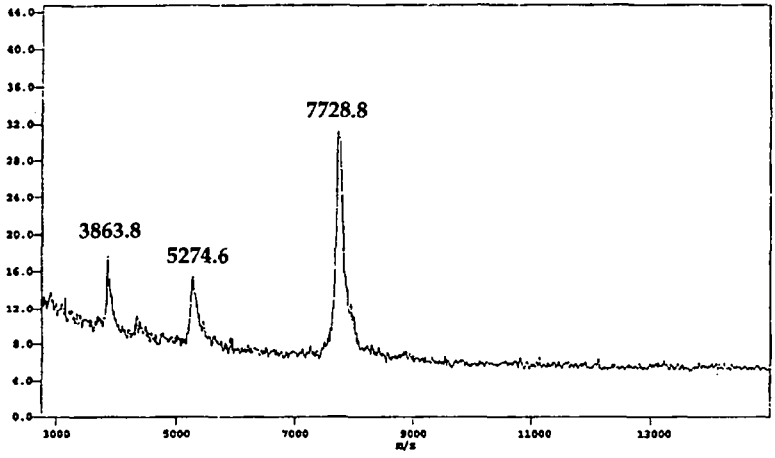

Fig. 9.5. MALDI-MS spectrum of a mixture of M13 coat protein and M13 procoat protein in organic solution.

MALDI-MS of our proteins in organic solution leads to spectra with a high signal to noise ratio clearly demonstrating the purity of our samples. In Fig. 9.5 a spectrum of a solution of M13 procoat protein is shown to which M13 coat protein is added as an internal standard to calibrate the system. Its molecular mass has been determined by ES-MS to be 5238. Assuming a similar overestimation in the corresponding spectral range, the mass of M13 procoat protein, estimated from the peaks at 3863.8 (doubly charged molecular ion) and the peak at 7728.8 (singly charged molecular ion), can be calculated to be 7675, 10 Da below the calculated value of 7685. The underestimation is most probably due to the calibration of the instrument. It could in principle also be possible that adduct formation plays a role. In this case adducts with either K^+ (MW 39) or with acetonitrile MW 41) could explain the difference in mass.

A spectrum of whole M13 phages (see Fig. 9.6) did yield the correct mass. Adducts could be formed to residues in the mature part of M13 coat which are inaccessible in the whole phage.

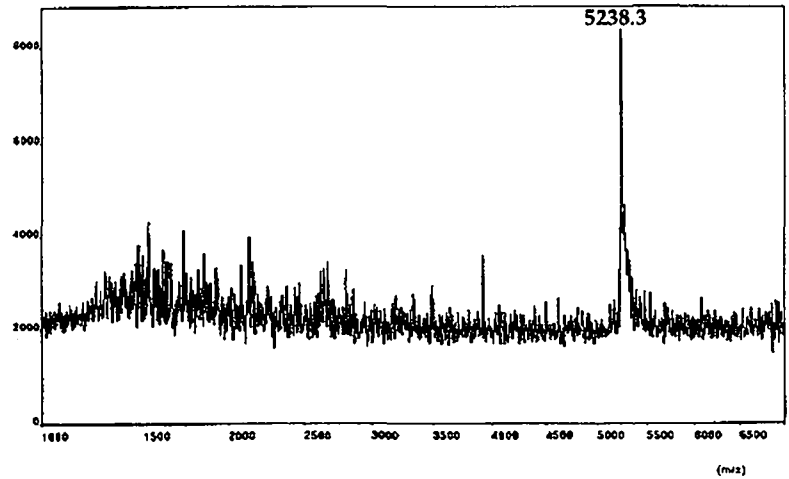

Fig. 9.6. MALDI-MS spectrum of intact M13 phages. The BP represents M13 coat protein in its monomeric form. No higher aggregates were detected.

It can clearly be seen from the different signal to noise ratios in figures 9.5 and 9.6 that the efficiency of creating gaseous protein ions greatly depends on the starting material. Although the amount of protein in the whole phages is much larger than in the experiment with the organic solution the signal is much lower. The same holds for reconstituted proteins as can be seen in the next section.

9.4.2. Digestion of reconstituted M13 coat protein.

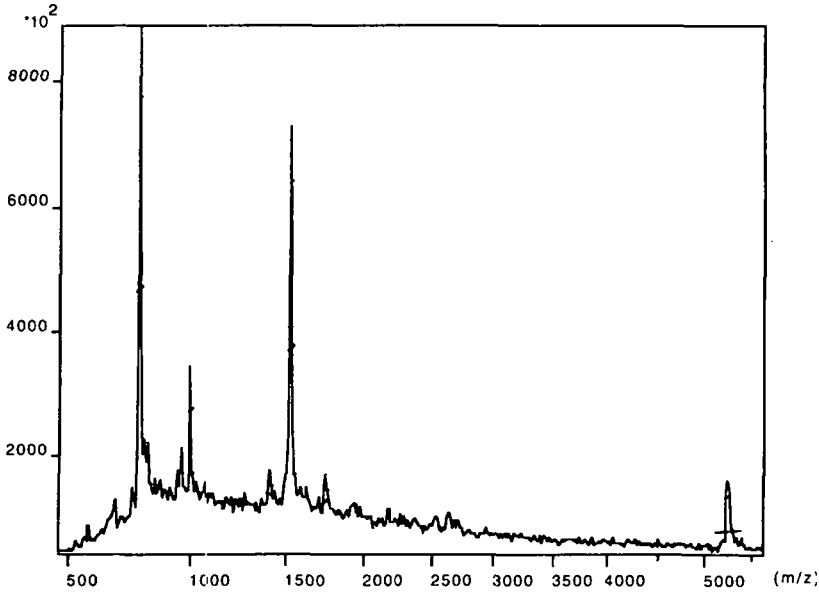

Fig. 9.7. MALDI-MS spectrum of M13 coat protein incorporated into POPG vesicles at a lipid to peptide ratio L/P =40.

Fig 9.7 shows a MALDI-MS spectrum of M13 coat protein incorporated into POPG (MW 760.084) vesicles. It shows that it is possible to obtain mass spectroscopic information directly from a sample containing a large excess of lipids. The peaks at m/z = 763.4 and m/z = 1525.4 are the lipid monomer and dimer peaks. Calibration on the lipid monomer and dimer peaks yielded a mass 5239.1. Edman degradation of the HPLC purified compound gave the following partial sequence:

AEGDDPAKAAFNSLQASATE.

After digestion with 3.8 % (w/w) Proteinase K one large and several smaller additional peaks appear as can be seen in Fig. 9.8, which represents part of a MALDI-MS spectrum taken directly of the digestion mixture.

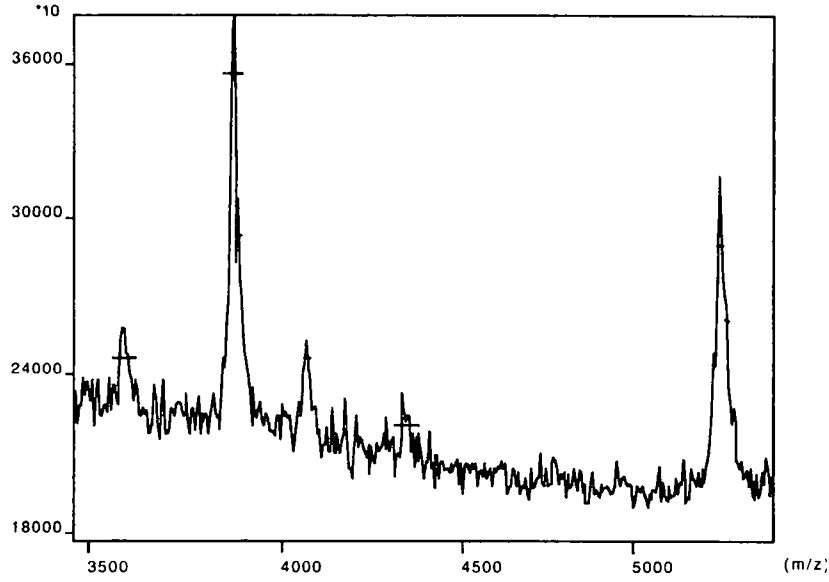

Fig. 9.8. Detail of a MALDI-MS spectrum of a digestion mixture of M13 coat protein reconstituted into POPG vesicles (see Fig. 9.7) and 3.8 % (w/w) proteinase K.

The intact protein appears at 5255.6 Da (16.4 Da above the expected mass) and the major digestion fragment at 3864.1 Da. If one assumes a similar mass error for the fragment as for the intact protein and searches the sequence of M13 coat for a sequence giving rise to a fragment of 3848 Da only two fragments can be found: residues 4-40 (3843.4 Da) and residues 15-50 ((3851.6 Da). In the HPLC elugram of the digestion one additional large peak appeared with a mass of approximately 3860 Da as confirmed by MALDI-MS. Its partial sequence as determined by Edman degradation was :

xAxATEYIGYAWAMVVVIVx.

This confirms that the main digestion product of M13 coat protein after treatment with proteinase K is the fragment from residue 15 to residue 50.

If the mass spectrum in Fig. 9.8 is recalibrated using the masses of the intact M13 coat protein and the fragment 15-50 one can try to identify the minor digestion

products by assigning the masses of the smaller peaks to other stretches in the protein sequence.

The peak at m/z =4049.9 can be unambiguously designed to residues 13-50 (4051.8 Da), whereas the most plausible possibilities for the peaks at m/z = 3572.7 and the one at m/z = 4319.5 are respectively residues 18-50 (3565.3 Da) and residues 11-50 (4313.1 Da). The first two minor peaks could be identified in the HPLC elugram of the digestion mixture, but no partial sequence could be obtained.

In a control experiment the digestion mixture was cooled immediately after addition of proteinase K and sonicated to provoke digestion from the inside of the vesicles. However, no substantial difference in the results could be observed.

9.4.3. Digestion of reconstituted Pf3 coat protein.

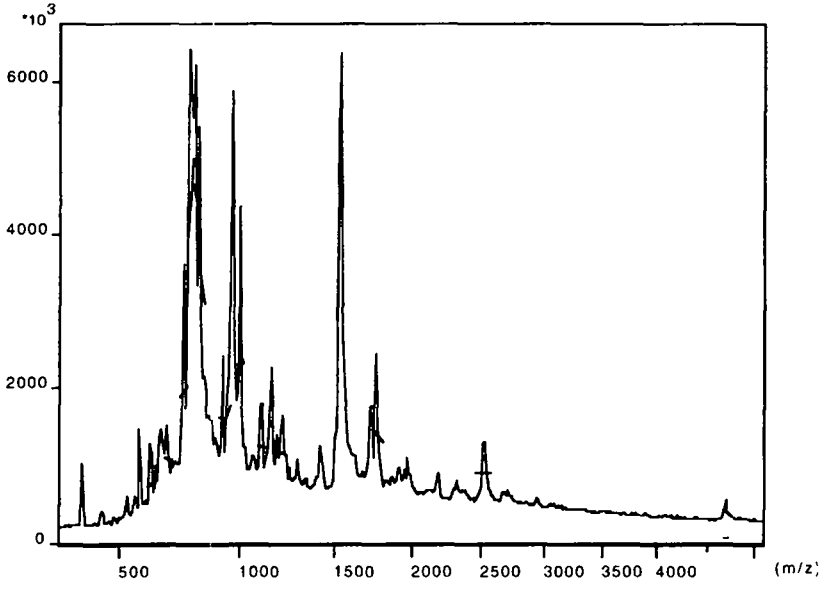

Fig. 9.9 Detail of a MALDI-MS spectrum of Pf3 coat protein reconstituted into POPG vesicles at a lipid to peptide ratio L/P = 40.

Fig. 9.9 shows a MALDI-MS spectrum of Pf3 coat protein reconstituted into POPG vesicles. The protein peak appears at m/z = 4676.6 Da (if the spectrum is recalibrated on the lipid peaks at m/z = 4673.7).

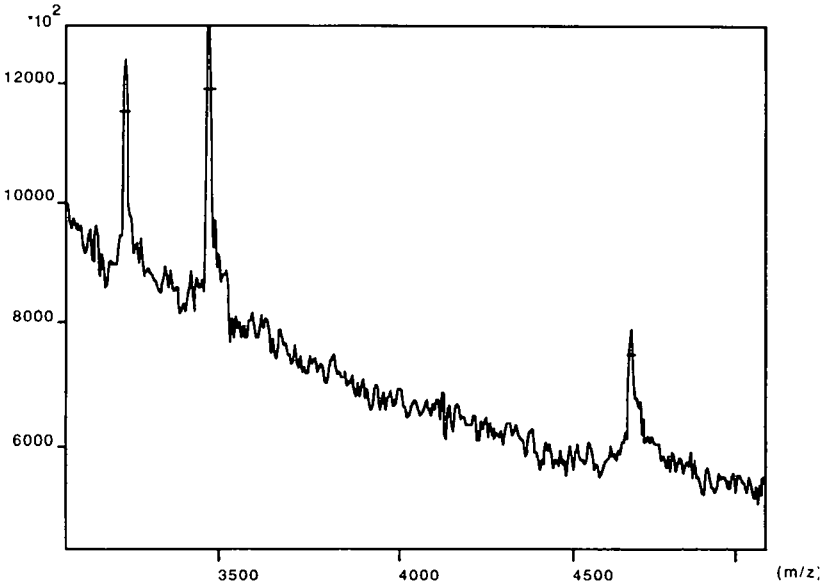

Fig. 9.10. Detail of a MALDI-MS spectrum of a digestion mixture of Pf3 coat protein reconstituted into POPG vesicles (see Fig. 9.9) and 3.8 % (w/w) proteinase K.

After digestion of reconstituted Pf3 coat protein two new peaks appeared (see Fig. 9.10) at m/z = 3470.3 and at m/z = 3256.1. From the subsequent HPLC purification two peaks which were not present in the undigested sample were isolated which appeared in MALDI-MS at m/z 3471 and at m/z 3257. A partial sequence was found by Edman degradation and the fragments were attributed as follows:

compound	mass	partial sequence	attributed sequence
1	3257 Da	AVQADITTIGG	residues 14 - 44 (3257.0 Da)
2	3470 Da	LTAVQADITTIGG	residues 12-44 (3471.2 Da)

The inaccessibility of the C-terminus could be due to a vectorial trans-membrane incorporation of Pf3 coat protein with the C-terminus residing on the inside of the vesicles. In order to test this hypothesis the sample was resonicated in the presence of proteinase K in an attempt to digest the membrane incorporated protein from the inside to of the vesicle. This control experiment yielded no relevant difference with respect to the original experiment.

The fact that the mass of the two digestion products of Pf3 coat protein is found to be in excellent agreement with the expected masses of the attributed protein segments leaves a question regarding the mass of the intact protein. The negatively charged C-terminal part which is truncated by proteinase K is found to be about 40 Da higher than expected. This can be due to three causes: an error in the sequence, a modification of the protein or a high affinity to e.g. K^+. More experiments are needed to settle this question.

9.5. DISCUSSION AND CONCLUSIONS

The C-termini of M13 coat protein and Pf3 coat protein were found to be inaccessible to digestion by proteinase K when they were reconstituted into POPG vesicles using the "mixed film" method (see Chapter 7). Control experiments in which proteinase K was also present inside the vesicles during digestion did not yield different results than when the protease was only added to the outside of the vesicles. This leaves open a few possibilities. The sonication may not have been successful in introducing proteinase K to the lumen of the vesicles, it might have been inactive in the lumen because of steric hindrance, or the C-terminus of the coat proteins is to close to the membrane to be accessible to protease K. Additional experiments at higher protein concentrations and with different enzymes are needed to settle this point. From the present experiments it could not be concluded if the proteins were inserted in a vectorial way into the bilayer or if the C-terminal part of the proteins was inaccessible to proteinase K for other reasons. In any case the N-terminal part of the protein is accessible to proteinase K digestion and thus probably protrudes into the aqueous phase.

Reconstituted M13 coat protein in our system is preferentially hydrolyzed by proteinase K on the C-terminal side of residue 14. Minor peptides have been identified where hydrolysis took place at the C-terminal side of residues 11, 13 and 18. Pf3 coat protein was preferentially cut after residue 12 and 14.

The experiments with Pf3 coat protein are not conclusive. The intact protein could be formylated (M +28) according to ES-MS results or very sensible to K^+ binding (M+39) according to the results of the digestion experiments followed by MALDI-MS. New experiments are possible because recently the expression of Pf3 coat protein in *E. coli* has made a new source accessible (D. Kiefer, unpublished results).

Subsequent experiments with M13 procoat protein and with (synthetic) leader peptide will yield the necessary additional information to determine the topology of both hydrophobic stretches of M13 procoat protein when reconstituted into lipid vesicles.

The present results have shown that MALDI-MS can be used directly on samples of reconstituted proteins. The presence of a large excess of lipid molecules results a in a less advantageous signal to noise ratio, but not to such an extent that analysis is make impossible. On the contrary, the presence of lipids in the samples give rise to additional peaks which can be used as an internal reference to calibrate the m/z spectra.

Analysis of digestion products using a combination of MALDI-MS, HPLC and peptide sequencing leads to results of higher precision in less time than with the traditional techniques involving selective radio-labeling. It is unnecessary to use protease at stoichiometric amounts and during prolonged digestion times as used in previous classical digestion studies.

9.6. LITERATURE

(1) Porter, R. R. **1959** *Biochem. J., 73,* 119-126.
(2) Bayer, R.; Feigenson, G. W. **1985** *Biochim. Biophys. Acta, 815,* 369-379.
(3) Woolford, J. L. J.; Webster, R. E. **1975** *J. Biol. Chem., 250,* 4333-4339.
(4) Wickner, W. **1977** *Biochemistry, 16,* 254-258.
(5) Wickner, W. **1975** *Proc. Natl. Acad. Sci. USA, 72,* 4749-4753.
(6) Wickner, W. **1976** *Proc. Natl. Acad. Sci. USA, 73,* 1159-1163.
(7) Smilowitz, H.; Carson, J.; Robbins, P. W. **1972** *J. Supramol. Struct., 1,* 8-18.
(8) Chamberlain, B. K.; Nozaki, Y.; Tanford, C.; Webster, R. E. **1978** *Biochim. Biophys. Acta, 510,* 18-37.
(9) Ohkawa, I.; Webster, R. **1981** *J. Biol. Chem, 256,* 9951-9958.
(10) Kuhn, A.; Wickner, W.; Kreil, G. **1986** *Nature, 322,* 335-339.

(11) Smilowitz, H. **1974** *J. Virol., 13,* 94-99.

(12) Rashed, I.; Oberer, E. **1986** *Microbiol. Rev., 50,* 401-427.

(13) Bayer, M. E. **1968** *J. Gen. Microbiol., 395-404,*

(14) Cavalieri, S. J.; Goldthwait, D. A.; Neet, K. E. **1976** *J. Mol. Biol., 102,* 713-722.

(15) Marvin, D. A.; Wachtel, F. J. **1975** *Nature, 253,* 19-23.

(16) Datema, K. P.; Spruijt, R. B.; Wolfs, C. J. A. M.; Hemminga, M. A. **1988** *Biochim. Biophys. Acta, 944,* 507-515.

(17) Datema, K. P.; Wolfs, C. J. A. M.; Marsh, D.; Watts, A.; Hemminga, M. A. **1987** *Biochemistry, 26,* 7571-7574.

(18) Roepstorff, P. **1992** *J. Mass Spectrom. Ion Processes, 118/119,* 789-809.

10. THERMODYNAMICS OF MEMBRANE INCORPORATION[1]

10.1. INTRODUCTION

The insertion of newly synthesized membrane proteins into lipid bilayers is in many cases directed by N-terminal signal (leader) sequences, that are removed from the mature part of the protein after insertion (for reviews see [2,3]). Complicated protein insertion and translocation machineries have been characterized genetically and biochemically in prokaryotic and eukaryotic cells. However, there are several examples of relatively small proteins which insert into membranes independently of the translocation machinery. One classical example is the coat protein of the filamentous phage M13 [4,5].

The M13 coat protein is synthesized with a 23 amino acid long signal sequence in the cytoplasm of *E. coli* and is then inserted into the plasma membrane. The membrane insertion process of this so-called M13 procoat protein has been extensively studied genetically and biochemically [6]. There are experimental indications that the binding of the protein to the negatively charged membrane surface occurs via the positively charged amino acid residues at the amino and carboxy termini of the procoat protein [7]. The subsequent partitioning of the two hydrophobic regions into the membrane results in a U-like configuration. This *in vivo* insertion process requires a transmembrane potential for the wild type procoat protein [8]. Biophysical investigations have shown that at least part of the predicted transmembrane regions adopt an α-helical conformation [9]. The topology of this intermediate has been investigated using procoat proteins with either N-terminal or C-terminal extensions [10,11]. It was also found that the N- and C-termini do not leave the cytoplasm during the insertion process and only the negatively charged central region traverses the membrane in accordance with the "positive-inside" rule [12]. After membrane insertion, the signal sequence is cleaved off by the *E. coli* leader peptidase which has its enzymatic activity at the periplasmic face of the cytoplasmic membrane [13,14].

---

[1]Parts of this chapter are published concommitantly in Biochemistry[1]

The membrane insertion process of the procoat protein is apparently independent of any "helper " protein complex, such as *E. coli* translocase and other factors like the ffh protein (Kuhn, unpublished results). In this context it is interesting to note that newly synthesized M13 procoat protein is capable of insertion into liposomes from the aqueous phase [15] suggesting that the insertion process is mainly driven by hydrophobic interactions. This makes quantitative investigation of the membrane insertion of this protein attractive. We determined experimentally the free energy change ΔG^o occurring during the spontaneous transfer of the procoat protein from the aqueous to the lipid bilayer phase. These data serve as a basis for a thermodynamic model describing the membrane insertion of procoat protein. In order to distinguish the energetic contributions which arise from the leader sequence and from the mature protein part, parallel experiments were also performed with the coat protein as well as with the procoat mutant protein OM30R. The OM30R mutant differs from the wild type procoat protein in having a Val -> Arg mutation at position +30, inserting a positive charge into the hydrophobic stretch of the mature part of the protein. This mutant protein is known from *in vivo* experiments to interact with the membrane, but not to be translocated across the *E. coli* plasma membrane. The interest in the present context is to learn whether or not this changed biological property is correlated with a change of the partitioning into lipid bilayer membranes. Furthermore, electrostatic contributions of the procoat protein interaction with membranes were taken into consideration by investigating the protein insertion into membranes with different membrane surface charge densities.

The experiments were performed in a reconstituted system using purified proteins which, after mixing, spontaneously bind to unilamellar lipid vesicles. The process of protein incorporation into the lipid membranes was then observed by measuring the fluorescence energy transfer (FET) between the (pro-) coat protein's tryptophan and tyrosine residues, which act as fluorescence donors, and a diphenylhexatriene (DPH) moiety coupled to phosphatidylcholine molecules in the bilayer membrane which acts as a fluorescence acceptor. The Förster distance of this couple was determined to be 4 nm [16], i.e. all protein molecules which are within this distance or closer to the membrane will lead to FET. Instead of the wild type M13 procoat protein we used in the course of this work the non-cleavable procoat mutant protein H5 (with a Phe at position -3) which inserts into the lipid bilayer as the wild type protein but can be isolated by a simple procedure in chemical quantities as was shown elsewhere[9].

10.2. THEORY

10.2.1. Insertion models

In order to quantify the membrane affinity of proteins we need to choose a model representing the process of membrane association to which we then fit our data. Three models for membrane association of proteins will be discussed here, which view the membrane in a fundamentally different way, respectively as a collection of binding sites, a two dimensional surface, and a three dimensional phase in its own right. We have chosen the last model to analyze our data.

The simplest model is that of a protein associating to a binding site consisting of a defined number of lipids. These binding sites are considered to be independent and equivalent. Criticism to this model concentrates on the fact that the membrane is a dynamic entity and that there are no distinct, independent binding sites.

In the Langmuir model both membranes/micelles and proteins are considered to present an accessible surface to which the other species can adsorb. The model implies that a constant membrane surface is occupied by a protein molecule irregardless of the degree of occupation of the membrane surface. This is probably an oversimplification.

10.2.2. A partitioning model between the aqueous and the lipidic phase

This model, which we have chosen to analyze our data, rejects the notion of binding sites and the idea of a membrane as a surface. Instead it considers the membrane as a separate phase. Fixation of proteins to the membrane is viewed in the simplest variant of the model as an ideal partitioning between the aqueous and the lipidic phase. The model can be complicated if necessary to accommodate aggregation phenomena in both aqueous and lipidic phase. It has already been used to quantitatively evaluate the association to lipid membranes of several amphipatic polypeptides. For review see Pawlak et al [17] and references there in.

In this work we use the formalism of Schwarz et al. [18]. An apparent partition coefficient Γ_{app} is defined as

$$r \cdot f(r) = \Gamma_{app} \cdot c_f \qquad (10.1)$$

with

$$r = c_{as} / c_l \qquad (10.2)$$

being the molar ratio of lipid-associated protein molecules (concentration c_{as}) per lipid molecule (c_l). $f(r)$, the activity coefficient as a function of r, may describe deviations from ideal partitioning ($f = 1$) due to protein-protein interaction at the water-membrane interface or in the lipid bilayer. The molar fraction of lipid associated protein, c_{as}/c_p, was measured by applying

$$c_{as} / c_p = F / F_\infty \qquad (10.3)$$

where

$$c_p = c_{as} + c_f \qquad (10.4)$$

is the total concentration of protein, F can stand for any signal used to monitor the association of protein to membranes . Here it specifically represents the FET induced increase in DPH fluorescence (see next section) and F_∞ the corresponding limiting fluorescence intensity at high L/P ratios for the case where all protein molecules are bound to lipid membranes. By using eq (10.4) it is possible to determine c_f for each known value of c_{as}. The apparent partition coefficient can be determined according to eq. (10.1) from the association isotherm of r versus the free protein concentration c_f if $f(r)$ is known. For the ideal case of $f(r) = 1$ a simple linear relation exists between r and Γ_{app}.

10.2.3. Trp-DPH Fluorescence Energy Transfer

Fluorescence energy transfer (FET) is the excited state energy transfer from a donor D to an acceptor A without mediation of a photon. It is a result of dipole-dipole coupling. The rate of transfer is dependent on the extent of overlap of the emission spectrum of the donor and the excitation spectrum of the acceptor (see Fig. 10.2) and on the relative orientation of the transition dipoles.

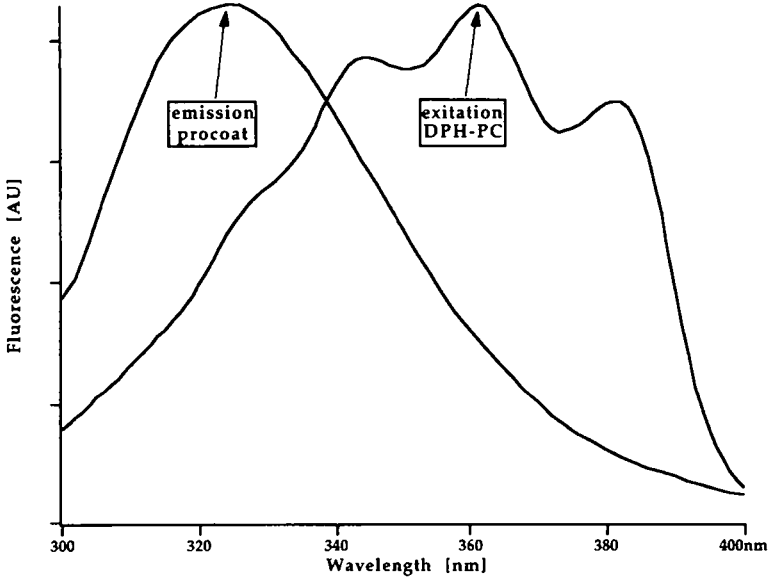

Fig. 10.2. Tyr/Trp emission spectrum of M13 procoat protein (excitation at 280 nm) and diphenylhexatriene (DPH) excitation spectrum of DPH-PC (emission at 430 nm). Due to the overlap between the two spectra fluorescence energy transfer occurs between the Tyr and Trp residues of vesicle-associated proteins which act as donors and lipid-anchored DPH moieties which act as acceptors.

In our case we use FET in the following way. The process of protein incorporation into the lipid membranes was observed by measuring the FET between the (pro-)coat protein's tryptophan and tyrosine residues, which act as fluorescence donors, and a diphenylhexatriene (DPH) moiety coupled to phosphatidylcholine molecules in the bilayer membrane which acts as a fluorescence acceptor. The Förster distance of this couple was determined to be 4 nm [16]. In the present case of procoat and coat proteins, which both contain one Trp and two Tyr residues in close proximity, we estimated from the optical properties a formal Förster distance of 5-6 nm for the coat or procoat proteins and DPH-PC in lipid bilayers assuming an orientational factor of $\kappa^2 = 2/3$ in the standard equations [19]. However, the exact Förster distance is not important in the present case because we use FET only to distinguish between membrane-bound and 'free', non-bound protein molecules. Membrane-bound proteins includes membrane insertion

into the interfacial region and the transmembrane insertion. The present fluorescence assay does not allow to distinguish directly between the two cases.

10.3. RESULTS

10.3.1. FET: TRP-DPH EXPERIMENTS

Fig. 10.2 shows a fluorescence emission spectrum of procoat protein arising from the intrinsic single tryptophan and the two tyrosine residues, together with an excitation spectrum of DPH-PC in POPC lipid vesicles. The intrinsic fluorescence spectrum of the coat protein under identical experimental conditions is comparable to that of procoat protein, since the three chromophores mentioned above are located in the mature part of the protein. The overlap of both spectra makes DPH-PC doped membranes ideally suited for the detection of the membrane binding of coat and procoat proteins by FET by measuring either the concomitant decrease of the intrinsic protein fluorescence or the corresponding increase of the DPH fluorescence.

A typical example of such an experiment is demonstrated in Fig. 10.3. Here a small volume of procoat protein, dissolved at relatively high concentrations in isopropanol (50 μM - 300 μM), was diluted into various samples comprising a large volume of an aqueous dispersion of different concentrations of SUV. The SUV consisted of POPC doped with DPH-PC. There are the following important results to be mentioned in particular: (i) At a given protein concentration, an increase of the lipid vesicle concentration both reduces the intrinsic protein fluorescence intensity and increases the DPH fluorescence intensity, as is shown in Fig. 3. (ii) In the presence of 2 mol % DPH-PC the Trp/Tyr fluorescence intensity of procoat protein (and similarly of coat protein) in the membrane-bound state, i.e. at high lipid-to-protein ratios, is less than 10 % of the intensity when the proteins are bound to pure POPC membrane vesicles without DPH-PC. This considerable quenching of the Trp/Tyr fluorescence due to FET is in good agreement with theoretical considerations of Kwong-Keung and Stryer [20] and Wolber and Hudson [21]. These authors have shown that the fluorescence transfer efficiency between a donor-acceptor pair in a lipid bilayer depends only on the membrane surface density of the acceptor and is independent of the surface density of the donor, i.e. of the protein bound to the lipid membrane. Adapting their considerations to our present case for a Förster distance of 5-6 nm and an

acceptor density of 2 mol % in the membrane, a FET quenching of more than 90% is calculated.

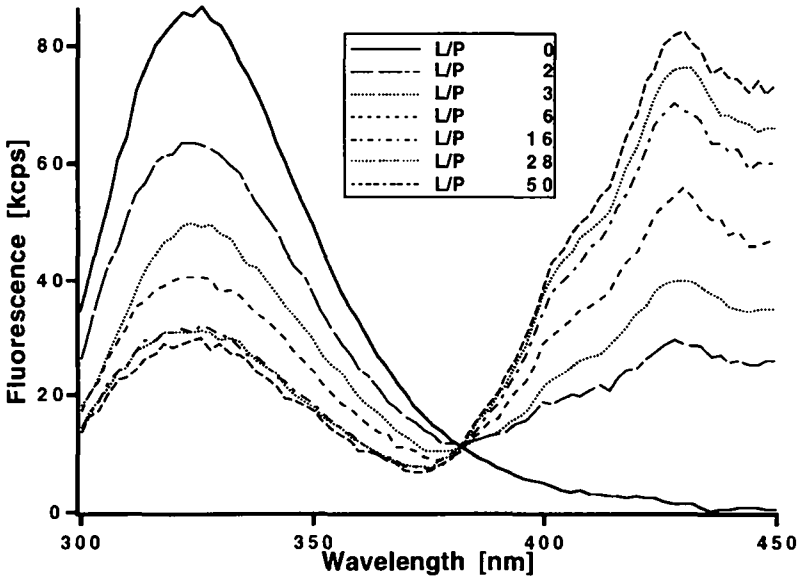

Fig. 10.3. Emission spectra of M13 procoat protein (0.5 μM in 1 mM Tris buffer, pH 7.4, 23 °C) exited at 280 nm in the presence of POPC vesicles doped with 2 mol % DPH-PC at increasing lipid to protein ratios (L/P). All spectra are corrected for DPH fluorescence in absence of protein by blank subtraction. Thus DPH fluorescence in the samples containing lipid vesicles is due only to FET between the Tyr/Trp residues of vesicle associated proteins and lipid-bound DPH moieties.

The fluorescence data are evaluated in Fig. 10.4 in form of a graph showing c_{as}/c_p as a function of molecular lipid-to-protein ratio, L/P.[2] The titration experiments demonstrate, firstly, that the procoat protein binds to POPC lipid membrane vesicles and, secondly, that the changes of the corresponding fluorescence signals continuously approximate a limiting value at high L/P ratios which reflect the decrease of the concentration of free procoat protein in the aqueous phase of the

[2] A prerequisite to determine the amount of membrane associated proteins cas by FET measurements according to eq. (10.3) is a linear dependence between the FET induced increase of the DPH fluorescence F and c_{as}. This was tested to be actually fulfilled by measuring the fluorescence signal F for coat protein bound to lipid vesicles (lipid/protein molar ratio of 300) at 0.5, 1, 1.5 and 2 mM protein concentration.

corresponding samples. As outlined in section 10.1.1, such a titration curve can be described by a thermodynamic equilibrium partitioning of the protein between the aqueous phase and the lipid bilayer. The solid line in the insert of Fig. 10.4 (plot of r versus c_f) corresponds to a fit of the experimental data to eq. (10.1) with an apparent partition coefficient $\Gamma_{app} = 6.5 \times 10^5\ M^{-1}$, the activity coefficient being taken to be $f(r) = 1$; we believe that to be a reasonable approximation, although cooperative binding effects cannot be totally excluded due to data scattering at higher values of c_f.

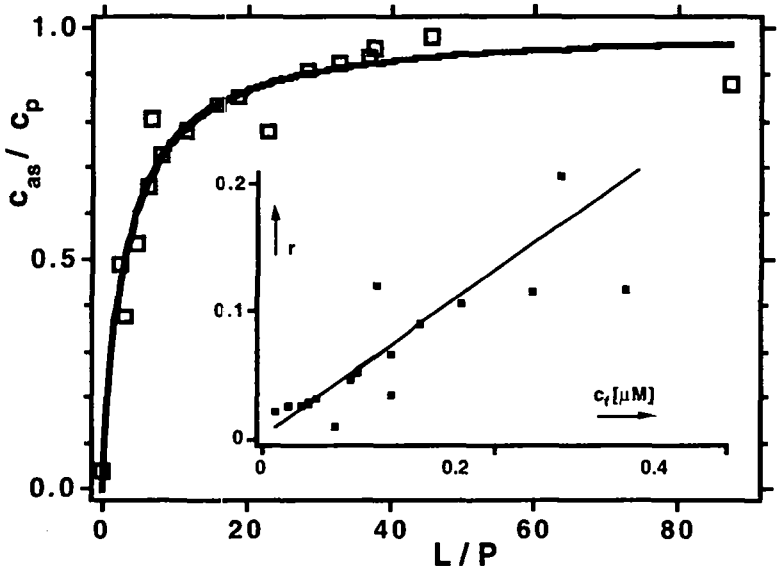

Fig 10.4. The fraction of procoat protein associated to POPC lipid vesicles relative to the total peptide concentration, c_{as}/c_p, plotted as a function of increasing lipid-to-protein ratios L/P. c_{as} was determined from the FET induced increase of DPH-PC fluorescence of the data shown and described in Fig. 10.3, according to eqs (10.3) and (10.4). The insert shows a plot of r, the molar ratio of lipid-associated protein molecules per lipid molecule, as a function of the free protein concentration c_f. The straight line corresponds to a least squares linear fit of an ideal protein partitioning into the lipid bilayers according to eq (10.1), yielding a partition coefficient $\Gamma = 6.5 \times 10^5\ M^{-1}$.

For the association of M13 coat protein to POPC vesicles, performed under identical conditions as for M13 procoat protein, we found a distinctly lower value for the apparent partition coefficient, $\Gamma_{app} = 1.0 \times 10^5\ M^{-1}$ (Fig. 10.5). Again the binding isotherm can be described using an activity coefficient of $f(r) = 1$.

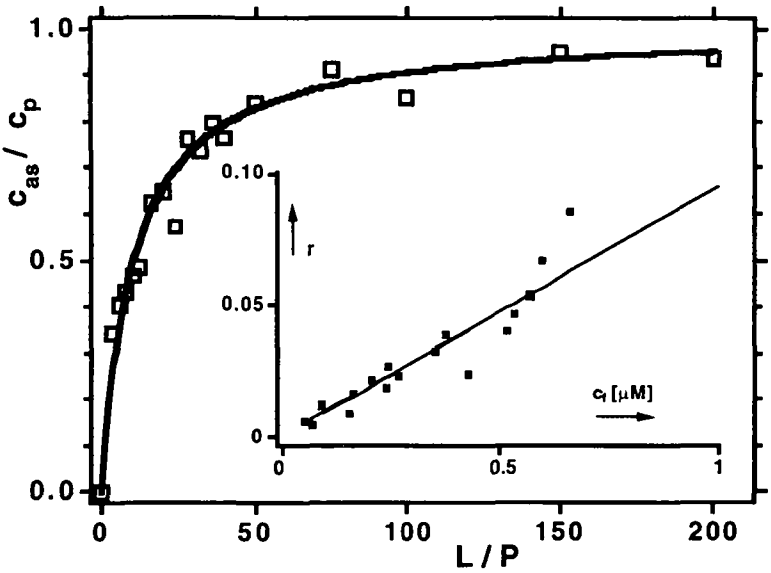

Fig 10.5. Association of M13 coat protein with POPC vesicles containing 2 mol % DPH-PC, determined by FET as described for M13 procoat protein in figures 10.2, 10.3 and 10.4. Experimental conditions: 1 μM coat protein in 1 mM Tris buffer, pH 7.4, 25 °C, excitation at 280 nm. Fitting as shown in the insert yielded a partition coefficient $\Gamma = 1.0 \times 10^5\ M^{-1}$.

In order to investigate whether the vesicle size influences the partition coefficients of either M13 procoat or coat protein, experiments were also performed using extruded vesicles with an average diameter of 1000 Å. No difference in membrane association of either protein was observed when compared with the experiments with sonicated SUVs with an average diameter of about 30 nm.

For comparison, a different reconstitution protocol was used; instead of injecting the proteins into a dispersion of preformed vesicles, the proteins were co-solubilized with the lipids in organic solvent. This solution was injected into an excess volume of aqueous buffer, thereby spontaneously forming lipid bilayers. It is assumed that, during this process, the hydrophobic protein segments are simultaneously integrated into the growing lipid bilayers (see below). Although a clear binding of the proteins could be detected by this assay, the high turbidity of the thus-formed vesicle dispersions as well as the necessity to work with an extra

lipid blank sample led to relatively large uncertainties in the data related to the final partition isotherms. This made it impossible to obtain a reliable partition coefficient for comparison with corresponding data from the preformed lipid vesicle assay.

Additional titration experiments with M13 coat and procoat proteins were performed to gain information about the influence of membrane surface charges and the physical state of the membrane on the association of M13 coat protein to lipid membranes. The results are summarized in Table 10.I.

TABLE 10.1. Apparent partition coefficients Γ_{app} and values for the free energy of transfer ΔG^o of coat and procoat proteins from water to different lipid membranes at 23 °C[a)]

Protein	Lipid	$\Gamma_{app} \times 10^{-5}\ M^{-1}$	$-\Delta G^o$ [kcal/mol]
H5 procoat protein	POPC	6.5	10.4
	POPC/POPG 1/1	>10	>10.7
OM30R procoat protein	POPC	0.3	8.6
M13 coat protein	POPC	1.0	9.3
	POPC/POPG 1/1	3.5	10.1
	DPPC	no spontaneous insertion	

a) The partition coefficients Γ_{app} are defined using the protein/lipid molar ratio as the concentration variable in the membrane. Following the considerations of Tanford [22], the standard free energy change upon transfer of the protein from water to the lipid bilayer is calculated as $\Delta G^o = -RT \ln\gamma_{MF}$ using the mole-fraction partition coefficient γ_{MF} which for sufficiently high dilution of the proteins in the membranes is given by $\gamma_{MF} = \Gamma_{app}/V_W$ with the partial molar volume of water

V_W = 18 ml/mol.[3] This form has been recently applied for the calculation of the partitioning of small peptides to lipid bilayers [26].

Interestingly, the apparent binding of M13 coat protein distinctly increased if lipid membranes with negative surface charges were used. In the present work the negative surface charges were provided by inclusion of POPG in the applied lipid vesicles at a molar composition of POPC/POPG = 1/1.

In order to investigate the role of the phase state of the lipids in the membrane, experiments were performed with vesicles composed of DPPC, which show an ordered <-> fluid phase transition at T_m = 41 °C. In contrast to the fluid bilayers of POPC, at room temperature DPPC membranes are in an ordered state in which the lipid acyl chains form a densely packed two-dimensional crystal (see Figs. 5.1 and 5.2). Under these conditions one would expect the membrane insertion of proteins to be considerably hindered. Indeed, no association with preformed DPPC vesicles below the phase transition temperature could be measured. In a control experiment a mixed solution of M13 coat protein, DPPC and DPH-PC in organic solvent was injected into buffer. During this procedure the protein is thought to incorporate in the membrane at the moment of bilayer formation, and thus does not need to penetrate the bilayer from the water phase. This experiment yielded vesicles in which the M13 coat protein was incorporated into the lipid bilayer, as judged by a marked decrease in the Trp fluorescence and a

---

[3] At present there are several published propositions for the calculation of the thermodynamic parameters of membrane partitioning processes. Tanford [22] suggested to use mole fraction (MF) based partition coefficents γ_{MF} to calculate the free energy change $\Delta G_o = -RT \ln \gamma_{MF}$ for the transfer of a substance from water to a lipid bilayer. In this way the cratic contribution arising from the entropy of mixing in the partition equilibrium is considered as in the case of partioning between two bulk phases. However, recent papers suggested that Flory-Huggins-corrected (FH) volume fraction units are more appropriate [23,24]:
$\Delta G_o = -RT \ln \gamma_V + RT V_S (1/V_L - 1/V_W)$.
Here γ_V is the volume-fraction partition coefficient and V_S, V_L, V_W are the partial molar volumes of solute, lipid bilayer, and water, respectively. As pointed out by White and Wimley [24] the critical point in membrane partioning is how to consider the cratic entropy contributions. MF units assume that solute and solvent molecules have the same molecular volumes, which certainly is not the case for the incorporation of proteins into a lipid bilayer membrane. The Flory-Huggins correction term takes this difference into account. In the case of membrane insertion of proteins the FH correction term is considerably higher than expected from MF units. For example, in the case of a 19 amino acid long hydrophobic α-helix inserted into a lipid bilayer, the second term in the FH calculation is about 50 kcal/mol which is about 5 times larger than the MF term. The applicability of the FH formalism for protein insertion into membranes is still under debate [24,25]. In our present work we calculate ΔG_o values on the basis of simple MF units. This, however, does not influence the general conclusions drawn from our results.

concomitant increase in the DPH fluorescence which was comparable to the changes observed in analogous experiments performed with POPC vesicles as reported before. These results show that M13 coat protein needs a fluid bilayer to be able to spontaneously insert into from the aqueous phase.

The topology of the membrane-bound proteins was analyzed by protease digestion experiments. The proteolytic fragments were separated by SDS-PAGE and probed for the presence of the intact mature N-terminal coat sequence by binding of an antibody against this protein sequence, followed by an immunoblot assay. After binding of procoat protein to preformed vesicles, proteinase K was added to the outside. We found that 20% of the protein was protected from degradation (data not shown). When the vesicles were first dissolved by a detergent solution of 1% octylglucoside and subsequently treated with proteinase, the procoat protein was entirely digested. We conclude from this experiment that a proportion of the inserted procoat protein is transmembrane (i.e. with the mature coat N-terminus in the lumen of the vesicles) with the remainder not being translocated (i.e. with the mature N-terminal segment located on the vesicle surface and thus accessible to the proteinase K). In the case of coat protein bound to preformed lipid vesicles, the N-termini of all molecules were accessible to proteinase K digestion from the outside. We cannot exclude a membrane translocation of the C-termini, although it appears to be unlikely that four positive lysine residues would cross the lipid bilayer.[4] Interestingly, in the case of adding the coat or the procoat protein during membrane formation (mixed injection method) the mature N-terminal segments of about 80% of both membrane incorporated proteins were protected from proteolysis, indicating an oriented transmembrane insertion.

Using OM30R protein, the procoat mutant protein containing a positively charged amino acid (position 26) in the center of the hydrophobic segment (amino acids 21-39) of the mature region, it is possible to identify the contribution of this particular segment to the overall free energy change of the M13 procoat protein interaction with the lipid membranes. The partition coefficient of OM30R procoat mutant protein was $\Gamma_{app} = 2.6 \times 10^4\ M^{-1}$, more than an order of magnitude lower than that of H5 procoat protein.

[4] A detailed mass spectroscopic investigation of the proteolytically digested proteins is presently being performed in our laboratories. Preliminary experiments indicate that even in detergent solution, the hydrophobic central segment in the sequence of the coat protein is protected against proteolysis.

10.3.2. CD

The reconstitution method of rehydrating freeze-dried mixed films (see section 7.3.2.) made it possible to measure binding isotherms with CD, although the difference in CD signal between the membrane bound proteins and the proteins in aqueous solution is relatively small compared to the difference observed in the FET experiments in the previous section. The form of the CD spectra obtained was that of the lipid bound conformation of the protein (see section 8.2.3) independent of the lipid to protein ratio L/P in the range of L/P 8-200, implying that only lipid-bound protein is measured in this type of experiment. The remaining protein appears to remain aggregated at the walls of the eppendorf cups in which the samples were prepared as judged from SDS-PAGE analysis of the recipients (see Fig. 10.6). Thus it is not transferred to the CD cuvette.

Fig. 10.6. SDS-PAGE analysis of unbound ('free') M 13 coat protein in binding isotherm using the FD-MF reconstitution protocol. Left most lane: M 13 coat protein standard; other lanes, from left to right: preparations at descending L/P ratios.

The signal to noise ratio is very unfavorable compared to that found in the FET experiments with the Trp-DPH-PC couple described in section 10.3.1.3 . No attempt at a quantitative analysis has been made since this would require a considerable amount of protein and experiment time in order to achieve satisfactory results. At a qualitative level the results are similar to those obtained in the FET experiments. In contrast to incorporation methods where protein solutions are injected and kept in solution, in the FD-MF method the sample is rehydrated from a mixed dried powder. Apparently protein which is not partitioned into the lipid bilayer which is formed during hydration does not go into solution. This makes it in principle possible to distinguish between bound and unbound ('free') proteins without the requirement of an additional probe molecule as is the case in the FET experiments. The relative low data quality could be explained by irreproducibility of the rehydration process during sample preparation.

10.4. DISCUSSION

In the present work the membrane insertion process of the M13 procoat protein is analyzed thermodynamically, focusing on three aspects: The spontaneous

membrane insertion of the procoat and coat proteins, the effect of membrane surface charges and, finally, the effect of point mutations in the mature coat protein segment.

10.4.1. THE SPONTANEOUS MEMBRANE INSERTION OF M13 (PRO-)COAT PROTEINS.

Both proteins associate to lipid vesicles comprising bilayers in a fluid lipid state but not to those which are composed of ordered lipid membranes such as DPPC at room temperature. Interestingly, the binding of both proteins to fluid lipid membranes can be described within the experimental accuracy by an ideal partitioning of the particular protein between the aqueous and the membrane phase. Neither aggregation nor electrostatic protein-protein interaction phenomena are observed as is shown by the corresponding linear partition isotherms obtained by FET measurements. The partition coefficients for pure POPC membranes are clearly higher for procoat than for coat proteins under identical conditions, but still of the same order of magnitude.

For the determination of the partition coefficients in the present paper, a given protein was mixed with lipid vesicles with which it spontaneously associates. In principle, the protein might either integrate only into the interfacial region of the membrane, or insert into the lipid bilayer in a transmembrane conformation. A further mechanistic interpretation of the partition coefficients demands the knowledge of the structure of the membrane-assembled (pro-)coat proteins obtained under the present experimental conditions.

With this in mind, the topology of the procoat protein bound to preformed vesicles was analyzed by the subsequent addition of proteinase K to the outside. It was found that 20% of the procoat protein was protected by the vesicles and could only be digested when detergent was added prior to the proteinase treatment. This strongly suggests that a fraction of the procoat protein molecules inserts into the bilayer by forming two transmembrane segments, with the loop region translocated across the bilayer. The translocation efficiency of 20% corresponds to earlier *in vitro* procoat expression experiments where the insertion into inverted *E. coli* membrane vesicles was tested [15,27-29]. Our finding in this respect is important because it demonstrates that our isolated H5 procoat mutant protein behaves in the reconstitution experiments as does the wild-type protein.

Biochemical experiments have shown that the membrane insertion of the procoat protein requires the electrochemical membrane potential [30]. Recently, it was demonstrated that the electrochemical potential promotes the membrane translocation if negatively charged residues are present in the translocated protein domain [31]. In the absence of the electrochemical potential, however, the translocation of the wild-type procoat protein was only partially inhibited, suggesting that there is no absolute requirement for a membrane potential. This might explain why procoat protein can to a certain extent spontaneously integrate into and translocate a lipid vesicle membrane. In addition, the membrane insertion in lipid vesicles might be different from the *in vivo* situation.

In the case of coat protein, the digestion experiments performed with proteinase K show that the N-termini of all protein molecules are accessible, indicating an insertion of coat protein only into the membrane interfacial region (if one excludes the energetically unfavorable membrane translocation of the positively charged C-terminus). In contrast, the hydrophobic parts of the coat and the procoat proteins span the lipid bilayer in a transmembrane helical conformation when the proteins are reconstituted into artificial lipid bilayers by the mixed injection method; this is the conclusion drawn from our present digestion experiments and recent biophysical investigations [9].

Obviously, the two different reconstitution protocols applied in the present work (protein binding to preformed vesicles versus protein binding during vesicle formation) result in differing membrane insertion of the two proteins. It is difficult to predict the partition coefficients for the two cases. The contribution of the hydrophobic effect to the standard free energy change ΔG^o which occurs during the transfer of a protein from water to the membrane phase (either by insertion into the membrane interfacial region or into a transbilayer form), is directly proportional to the surface of the hydrophobic part of the protein in contact with the lipid membrane.[5] In the present work, one has to compare the situation of the

[5] At present there are several published propositions for the calculation of the thermodynamic parameters of membrane partitioning processes. Tanford [22] suggested to use mole fraction (MF) based partition coefficients γ_{MF} to calculate the free energy change $\Delta Go = -RT \ln \gamma_{MF}$ for the transfer of a substance from water to a lipid bilayer. In this way the cratic contribution arising from the entropy of mixing in the partition equilibrium is considered as in the case of partitioning between two bulk phases. However, recent papers suggested that Flory-Huggins-corrected (FH) volume fraction units are more appropriate [23,24]:
$\Delta Go = -RT \ln \gamma_V + RT V_S (1/V_L - 1/V_W)$.

(pro-)coat proteins inserted into the interfacial region of a lipid bilayer with the transbilayer configuration of procoat protein. As discussed in detail by White and colleagues [24,32], the interfacial region of a lipid bilayer, which actually accounts for about 50% of the total time-averaged thickness of the membrane, is large enough to accommodate a helical protein oriented parallel to the membrane plane.

As described in detail elsewhere [33-36] the formation of a hydrophobic transmembrane helix can be thermodynamically described by the standard free energy change, ΔG^o, for the insertion of this α-helix into a lipid bilayer. The overall driving force for the insertion is the hydrophobic effect. But, in addition, unfavorable terms must be considered, for example the fact that part of the helix surface is polar. Even in the case where only hydrophobic amino acid side chains are involved, negative entropic terms arise from the restricted mobility of the polypeptide in the membrane and the effect of disturbing the lipid bilayer itself.

Taking the arguments from the work of Jähnig [34] an estimated value of $-\Delta G^o$ for the partitioning of a hydrophobic α-helix of 15 (procoat leader helix segment) or 19 (mature coat helix segment) amino acids would be between 10 and 20 kcal/mol for the leader sequence helix and 12 - 24 kcal/mol for the coat protein helix, depending on the applied scale for the strength of the hydrophobic effect. Our values of $-\Delta G^o$ are the first experimental data for transmembrane proteins that have a signal sequence, and they are distinctly lower than the proposed estimates.

In order to compare these values with our actual experimental data, we have furthermore to consider the following. If the procoat protein spans the lipid bilayer with two helices, which due to their close neighborhood are already associated,

---

Here gv is the volume-fraction partition coefficient and V_S, V_L, V_W are the partial molar volumes of solute, lipid bilayer, and water, respectively. As pointed out by White and Wimley [24] the critical point in membrane partitioning is how to consider the cratic entropy contributions. MF units assume that solute and solvent molecules have the same molecular volumes, which certainly is not the case for the incorporation of proteins into a lipid bilayer membrane. The Flory-Huggins correction term takes this difference into account. In the case of membrane insertion of proteins the FH correction term is considerably higher than expected from MF units. For example, in the case of a 19 amino acid long hydrophobic a-helix inserted into a lipid bilayer, the second term in the FH calculation is about 50 kcal/mol which is about 5 times larger than the MF term. The applicability of the FH formalism for protein insertion into membranes is still under debate [24,25]. In our present paper we calculate ΔGo values on the basis of simple MF units which, however, does not influence the general conclusions drawn from our results.

then the actual hydrophobic surface of the helices in contact with the lipid hydrocarbon core of the membrane is only about 30-40% larger than in the case of a single, transbilayer helix. In this respect, the structure of the membrane-inserted part of the procoat protein seems to be similar to that of leader peptidase, where about 30% of the surface area of two adjacent transmembrane helices are in direct contact, leaving 70 % of the total helix surface area to interact with the lipid hydrocarbon chains [37]. In addition, it should be considered that for procoat protein, the hydrophobic mature part and the leader sequence contribute differently to ΔG^o of the membrane insertion process. This is actually demonstrated by preliminary experiments performed with chemically synthesized M13 leader peptide. Surprisingly, the peptide does not incorporate to a detectable extent into electrically neutral POPC membranes but shows only electrostatic binding to the surface of negatively charged lipid membranes of POPC/POPG (see also below). A detailed analysis of this effect will be published in a forthcoming paper.

On the basis of these considerations, it is now possible to estimate the $(\Delta G^o)_{tm}$, the standard free energy change for the transmembrane insertion of procoat protein. Here we assume that the experimentally determined ΔG^o of -10.4 kcal/mol for the binding of procoat protein to POPC vesicles is composed of 20% by transmembrane insertion and 80% by interfacial membrane insertion. For the latter case, we furthermore assume no contribution from the leader sequence, i.e. we can adopt the experimental value of ΔG^o = -9.3 kcal/mol for the coat protein binding to the POPC vesicles. Simple algebra yields $(\Delta G^o)_{tm}$ = -14.8 kcal/mol. (A realistic uncertainty of 20±5% in the fraction of membrane-translocated procoat protein would give a range of -13.7 to -16.6 kcal/mol for $(\Delta G^o)_{tm}$. This error estimation implicitly implies isotropic motion of the donor transition moment relative to that of the acceptor.). This estimated value of $(\Delta G^o)_{tm}$ corresponds to an upper limit because a possible contribution of the leader sequence to the membrane binding would necessarily decrease $(\Delta G^o)_{tm}$. If one furthermore takes into account the shielding effect of neighboring transmembrane helices as discussed above, the $(\Delta G^o)_{tm}$ of an individual coat protein would be about -12 kcal/mole, i.e. an additional contribution of only -2.7 kcal/mol for the transmembrane insertion compared to the interfacial insertion. Our results imply that the major part of the free energy change during membrane interaction occurs already at the membrane interface. It is the additional energy gain from the interaction of the hydrophobic portion of the folded pre-protein with the lipid chains which drives the final transmembrane insertion of the M13 procoat protein. Neither the leader sequence nor the mature coat protein alone

yields this free energy gain. This model is in accord with propositions made by Jacobs and White [32] for the insertion of proteins into the membrane interfacial region, based on binding experiments of tripeptides to lipid bilayers. Our values of $-\Delta G^o$ are the first experimental data for transmembrane proteins that have a signal sequence, and they are on the lower limit of the proposed estimates of Engelman & Steitz [33] and Jähnig [34].

In this context it is quite interesting to compare our data with published ones of other membrane-associated proteins. Our experimentally determined partition coefficients Γ for the M13 coat and the H5 procoat proteins are more than one to two orders of magnitude larger than those published for small (about 20-30 amino acid), amphipatic membrane-active polypeptides which show typical values in the range of $\Gamma = 10^3\text{-}10^4\ M^{-1}$ corresponding to $-\Delta G^o$ = 6.5-7.9 kcal mol^{-1} [18,38-40].[6] Recently, the membrane association of a semisynthetic protein was investigated [42]: The water-soluble bovine pancreatic trypsin inhibitor was attached to a hydrophobic Ala_{20} peptide segment, which in the membrane-bound form was assumed to adopt a transmembrane helix. The partition coefficient of this artificial protein to lipid vesicles was determined as $\Gamma = 2\text{-}3x10^3\ M^{-1}$ (value recalculated in order to have the identical dimensions as the data in the present work). A considerably higher value of $\Gamma = 1.5x10^9\ M^{-1}$ ($-\Delta G^o$ = 15 kcal mol^{-1}) was determined for template-assembled melittin, which is supposed to span the lipid membrane by four helical segments [17]. The partitioning of procoat protein to lipid bilayers is distinctly higher than that of simple amphipathic helical polypeptides and lower than that of a four transmembrane helix protein.

It is interesting to note in this context a recent paper by Lee and Manoil [43]. The authors showed that mutations eliminating the protein export function of a membrane-spanning sequence can be explained by a model where a minimum hydrophobicity is required for membrane insertion.

[6] The partition coefficients of Jones and Gierasch [41] cannot directly be compared to our values because these authors apply a different description of the binding of the polypeptides to lipid membranes. Their values of partition coefficients, even for polypeptides with only 12 consecutive hydrophobic amino acids, are orders of magnitude higher than all other published values for either amphipathic or hydrophobic polypeptides of comparable structure. We assume that this is due to the fact that the evaluation procedure of Jones and Gierasch does not allow to distinguish between electrostatic binding and hydrophobic partitioning of the polypeptides.

10.4.2. DEPENDENCE OF THE MEMBRANE INSERTION ON THE NEGATIVE MEMBRANE SURFACE CHARGES.

The apparent partition coefficient for coat and procoat proteins is larger when net negative charges are introduced to the surface of the membrane as shown for mixed POPC/POPG membranes, but it is still of the same order of magnitude as for POPC membranes. This can be explained qualitatively by a preferential electrostatic interaction between the positively charged amino acid residues of the protein and the negative charges of POPG. Such an interaction would lead to an increased concentration of the corresponding protein near the membrane surface. For a qualitative discussion of this electrostatic effect, we assume that the intrinsic coefficient Γ for the protein partitioning to POPC and mixed POPC/POPG membranes is identical. In the presence of a surface potential Ψ_o, induced by negatively charged lipids, the increase (decrease) of the free protein concentration at the membrane surface is determined by the protein's effective positive (negative) net charge. For POPC/POPG = 1/1 (mol/mol) mixed vesicles, the membrane surface potential, Ψ_o, exerted by the negative charges of the POPG molecules can be calculated for the experimental buffer conditions (1 mM Tris-HCl) using the Gouy-Chapman theory [44,45] which gives $\Psi_o \approx -200$ mV.

It is the Ψ_o of the membrane and the effective positive charge z_{eff} of the protein which determine the concentration of the protein at the membrane surface. Under these considerations the following relationship between the apparent and the intrinsic partition coefficients of the protein holds [46 and refs therein]

$$\Gamma_{app} = \Gamma \exp(-z_{eff}\, e_o\, \Psi_o / kT) \qquad (10.5)$$

with e_o the elementary electrical charge and k the Boltzman constant. The measured Γ_{app} yields a formal (effective) positive charge of $z_{eff} \approx 0.2$ for the M13 coat protein. The protein has one positive and 4 negative charges at the hydrophilic N-terminal segment and a further 4 positive charges at the hydrophilic C-terminal segment. It should be noted that, according to eq (10.5), 1 effective positive charge on the protein increases the partitioning by a factor of $\Gamma_{app}/\Gamma \approx 3 \times 10^3$. Although, formally, only one positive excess charge is present in the coat protein at neutral pH, the asymmetrical distribution of the negative and positive charges creates a rather high charge density at the hydrophilic protein termini which might play a role in the organization of the protein on and in the lipid bilayer. The present experiments indicate that only 0.2 positive charges are effective dur-

ing the membrane association of M13 coat protein. There are at least two possible explanations why the electrostatic effects are small in our experiments. Firstly, in the case of relative large proteins the finite size of the molecule might reduce the formal charges considerably as has been shown theoretically [47] and experimentally for other polypeptides [38] and for synthetic proteins [17 and refs therein]. Secondly, the M13 coat protein might form, both in water and in the lipid membrane bound state, a head-to-tail dimer matching the negative and positive charges at the opposite termini. The existence of a protein dimer would result in linear association isotherms (Fig. 10.4) as for monomers.

Indeed, stable dimers of M13 coat protein as well as certain mutant coat proteins have recently been observed by Deber et al. [48]. Whether such structures are present in our reconstitution system will be a topic of future investigation. At least under *in vivo* conditions the aggregation of coat proteins in the membrane must occur during the phage assembly process [49,50].

The influence of electrical surface charges in the case of membrane insertion of procoat protein is less clear than for coat protein, but the situation seems to be similar to that of the coat protein. The partition coefficient of procoat protein for POPC/POPG membranes is higher than 10^6 M^{-1} compared to 6.5×10^5 M^{-1} in POPC which corresponds to a $z_{eff} \geq 0.06$. Unfortunately, it was not possible under the present experimental conditions to determine ΔG^o more precisely because preparations at low L/P molar ratios always resulted in highly turbid samples. Similar difficulties arose if buffers at higher ionic strength than 5 mM were used, both for procoat and for coat proteins, because, under such conditions, the proteins start to aggregate in the aqueous phase.

In contrast to this relatively small electrostatic effect in the case of M13 (pro-)coat proteins, large contributions of charged residues to membrane binding were recently observed in the case of *E. coli* LamB signal sequence peptides [41]. It is interesting to note in this context that preliminary measurements with synthetic M13 leader sequence peptide performed in our laboratory show a strong electrostatic binding to negatively charged lipid membrane surfaces but only weak interactions to electrically neutral POPC membranes. This finding is consistent with the observation that the arginine mutation in the hydrophobic segment of the leader sequence has only a weak effect in the transmembrane insertion of procoat protein [51]. Further, the minor contribution of the leader sequence is also evident if one compares the conformation and solubility of the leader sequence peptide (-1

to -23) with those of M13 coat protein. It has been shown [52] that synthetic M13 leader sequence peptide has a remarkable solubility in water, up to 1 mg/ml. At a concentration of 4 x 10^{-5} M in 0.02 M phosphate buffer of pH 2.8, the M13 leader sequence peptide adopts a non-regular structure, according to CD measurements. In contrast, M13 coat protein is much less soluble in water, and even at a concentration of 10^{-7} M in aqueous buffer, the protein shows a considerable helix content of 50 % [9]. Taken together, these findings indicate that the isolated signal peptide behaves differently from the entire procoat protein. In consequence one should be extremely cautious to draw conclusions on biological functions of newly synthesized proteins simply based on results obtained from experiments with isolated leader peptides. The properties of an entire pre-protein seem to be not simply a sum of the properties of the mature protein and the leader sequence itself.

10.4.3. MEMBRANE ASSOCIATION OF OM30R MUTANT PROCOAT PROTEIN.

The OM30R mutant protein with an additional positively charged amino acid residue within the hydrophobic mature coat protein segment, shows a distinctly lower partitioning (Γ_{app} = 0.3x10^5 M^{-1}; ΔG^o = -8.6 kcal/mol) for POPC membranes than the H5 procoat protein (Γ_{app} = 6.5x10^5 M^{-1}; ΔG^o = -10.4 kcal/mol). Clearly, the hydrophobic mature part plays an important role for insertion of the procoat protein into the lipid bilayer. Yet, the affinity to lipid bilayers of the OM30R mutant protein is still relatively high . This is in total agreement with *in vivo* experiments from which it is known that the OM30R mutant protein is not translocated across the *E. coli* plasma membrane, but still interacts with and binds to these membranes [51]. According to the arguments presented before, i.e. that the leader sequence does not play an important role for the interfacial membrane insertion, one should compare the ΔG^o values of OM30R and coat protein. Clearly, a positive charge in the central hydrophobic segment decreases the affinity for interfacial membrane insertion, although only to a relatively small extent. Therefore, the major barrier for preventing a transmembrane structure in the case of OM30 seems to be the energetically unfavorable insertion of a positive charge into the center of a lipid bilayer.

10.5. CONCLUSIONS

We believe that the presented procedure might be generally applicable to probe the membrane insertion process of hydrophobic proteins into lipid bilayers. The

thermodynamic data obtained are independent of the particular structure of the protein at or in the lipid membrane. They hold for a partitioning of a protein to the polar lipid headgroup region, which has been shown to establish about 50% of a lipid bilayer [24], as well as for a real transmembrane insertion. In the case of M13 procoat protein, future experiments are necessary to reveal details of how an electrochemical membrane potential influences the membrane insertion, helix formation and membrane translocation of the protein.

Partitioning on the membrane interfacial region might be also an important first step of membrane interaction of proteins which are imported into or translocated across membranes by more complex signal-receptor mechanism. A first partitioning of such a protein to the membrane surface region would increase the efficiency to recognize a corresponding receptor and translocation machinery on the membrane level as was proposed as a possible model for hormone-receptor interactions in the case of membrane neuroreceptors [53,54].

10.6 LITERATURE

(1) Soekarjo, M.; Eisenhawer, M.; Kuhn, A.; Vogel, H. **1996** *Biochemistry, in press,*

(2) Heijne, G. v. **1994** *Annu. Rev. Biophys. Struct., 23,* 167-192.

(3) Wickner, W. **1988** *Biochemistry, 27,* 1082-1093.

(4) Kuhn, A. **1995** *Microbiol. Rev., in press,*

(5) Broomesmith, J. K.; Gnaneshan, S.; Hunt, L. A.; Mehraeinghomi, F.; Hashemzadehbonehi, L.; Tadayyon, M.; Hennessey, E. S. **1994** *Mol. Mem. Biol., 11,* 3-8.

(6) Kuhn, A.; Troschel, D. In *Membrane biogenesis and protein targeting;* W. Neupert and R. Lill, Eds.; Elsevier: New York, 1992; pp 33-47.

(7) Galluser, A.; Kuhn, A. **1990** *EMBO J., 9,* 2723-2729.

(8) Kuhn, A.; Wickner, W. **1985** *J. Biol. Chem., 260,* 15907-15913.

(9) Thiaudière, E.; Soekarjo, M.; Kuchinka, E.; Kuhn, A.; Vogel, H. **1993** *Biochemistry, 32,* 12186-12196.

(10) Kuhn, A.; Wickner, W.; Kreil, G. **1986** *Nature, 322,* 335-339.

(11) Kuhn, A.; Kreil, G.; Wickner, W. **1987** *EMBO J., 6,* 501-505.

(12) Heijne, G. v. **1992** *J. Mol. Biol., 225,* 487-494.

(13) Dalbey, R. E. **1991** *Mol. Microbiol., 5,* 2855-2860.

(14) Dalbey, R. E.; Heijne, G. v. **1992** *TIBS, 17,* 474-478.

(15) Geller, B. L.; Wickner, W. **1985** *J. Mol. Biol., 260,* 13281-13285.

(16) Le Doan, T.; Tagasuki, M.; Aragon, I.; Boudet, G.; Montenay-Garestier, T.; Helene, C. **1983** *Biochim. Biophys. Acta, 735,* 259-270.

(17) Pawlak, M.; Meseth, U.; Dhanapal, B.; Mutter, M.; Vogel, H. **1994** *Protein Sci., 3,* 1788-1805.

(18) Schwarz, G.; Stankowski, S.; Rizzo, V. **1986** *Biochim. Biophys. Acta, 1110,* 97-104.

(19) Förster, T. **1949** *Z. Naturforschg., 4a,* 321-327.

(20) Kwong-Keung, B.; Stryer, L. **1978** *Biochemistry, 17,* 5241-5248.

(21) Wolber, P. K.; Hudson, B. S. **1979** *Biophys. J., 26,*

(22) Tanford, C. *The Hydrophobic Effect: Formation of micelles and biological membranes.*; 2nd ed.; John Wiley: New York, 1980.

(23) Sharp, K. A.; Nicholls, A.; Friedman, R.; Honig, B. **1991** *Biochemistry, 30,* 9686-9697.

(24) White, S. H.; Wimley, W. C. **1994** *Curr. Opp. Struct. Biol., 4,* 79-86.

(25) Holtzer, A. **1992** *Biochemistry, 32,* 711-715.

(26) Terzi, E.; Hölzemann, G.; Seelig, J. **1994** *Biochemistry, 33,* 7434-7441.

(27) Wickner, W.; Mandel, G.; Zwizinski, C.; Bates, M.; Killick, T. **1978** *Proc. Natl. Acad. Sci. USA, 75,* 1754-1758.

(28) Silver, P.; Watts, C.; Wickner, W. **1981** *Cell, 25,* 341-345.

(29) Ohno-Iwashita, Y.; Wickner, W. **1983** *J. Biol. Chem., 258,* 1895-1900.

(30) Date, T.; Goodman, J. M.; Wickner, W. **1980** *Proc. Natl. Acad. Sci. USA, 77,* 4669-4673.

(31) Cao, G. Q.; Kuhn, A.; Dalbey, R. E. **1995** *EMBO J., 14,* 866-875.

(32) Jacobs, R. E.; White, S. H. B., 3421-3437. **1989** *Biochemistry, 28,* 3421-3437.

(33) Engelman, D.; Steitz, T. A. **1981** *Cell, 23,* 411-422.

(34) Jähnig, F. **1983** *Proc. Natl. Acad. Sci. USA, 80,* 3691-3695.

(35) Engelman, D.; Steitz, T.; Goldman, A. **1986** *Annu. Rev. Biophys, Biophys. Chem., 15,* 321-353.

(36) Lemmon, M. A.; Engelman, D. M. **1994** *Quart. Rev. Biophys., 27,* 157-218.

(37) Whitley, P.; Nilsson, L.; Heijne, G. v. **1993** *Biochemistry, 32,*

(38) Schwarz, G.; Beschiaschvili, G. **1989** *Biochim. Biophys. Acta, 979,* 82-90.

(39) Peled, H.; Shai, Y. **1993** *Biochemistry, 32,* 7879-7885.

(40) Ben-Efraim, I.; Strahilevitz, J.; Bach, D.; Shai, Y. **1994** *Biochemistry, 33,* 6966-6973.

(41) Jones, J. D.; Gierasch, L. M. **1994** *Biophys. J., 67,* 1546-1561.

(42) Moll, T. S.; Thompson, T. E. **1994** *Biochemistry, 33,* 15469-15482.

(43) Lee, E.; Manoil, C. **1994** *J. Biol. Chem., 269,* 28822-28829.

(44) Träuble, H.; Teubner, M.; Woolley, P.; Eibl, H. **1976** *Biophys. Chem., 4,* 319-342.

(45) McLaughlin, S. **1977** *Curr. Top. Membr. Transp., 9,* 71-144.

(46) Roise, D. In *Thermodynamics of membrane receptors and channels;* M. B. Jackson, Ed.; CRC Press: Florida, 1993; pp 81-126.

(47) Stankowski, S.; Pawlak, M.; Kaisheva, E.; Robert, C. H.; Schwarz, G. **1991** *Biochim. Biophys. Acta, 1069,* 77-86.

(48) Deber, C. M.; Khan, A. R.; Li, Z.; Joensson, C.; Glibowicka, M.; Wang, J. **1993** *Proc. Natl. Acad. Sci. USA, 90,* 11648-11652.

(49) Marvin, D. A. **1989** *Int. J. Biol. Macromol., 11,* 159-163.

(50) Russel, M. **1991** *Mol. Microbiol., 5,* 1607-1613.

(51) Kuhn, A.; Kreil, G.; Wickner, W. **1986** *EMBO J., 5,* 3681-3685.

(52) Shinnar, A. E.; Kaisar, E. T. **1984** *J. Am. Chem. Soc., 106,* 5006-5007.

(53) Schwyzer, R. **1977** *Ann. N. Y. Acad. Sci., 297,* 3-26.

(54) Sargent, D. F.; Schwyzer, R. **1986** *Proc. Natl. Acad. Sci. USA, 83,* 5774-5778.

11. CONCLUDING REMARKS

M13 procoat protein has proven to be a simple model system to study the membrane insertion of precursor proteins. The lipid interaction of procoat protein was compared with those of the mature major coat proteins of bacteriophages M13 and Pf3. M13 coat protein is translocation incompetent *in vivo* , whereas Pf3 coat protein, which is synthesized without a leader sequence, is, *in vivo*, translocation competent. The comparison of these three proteins made it possible to correlate the functional differences of the proteins with certain structural features.

This thesis describes several new methods and applications, generally simplifying existing methods, as well as making them more powerful and suitable for our system.

The purification and reconstitution procedures use organic solvents instead of detergents. They have two advantages over existing methods: they are rapid and circumvent aggregation artifacts, which have severely limited previously existing methods during the past 10 years or more.

CD and IR measurements have established the secondary structure of the proteins in organic solution (>90% α-helical, as in the phage), in aqueous solution and in reconstituted lipid membranes (50-75% α-helical). The results obtained ended the reigning confusion about the secondary structure of the membrane incorporated proteins. Using ATR-FTIR, it was established that the trans-membrane helices of all three proteins are located preferentially along the membrane normal. M13 coat protein is more rigidly fixed in this position than the other two proteins. This fact coincides with a higher thermal stability of M13 coat proteins witnessed by CD and correlates with the difference in translocation competence *in vivo*.

The topology of the membrane inserted proteins has been monitored by limited enzymatic digestion. The proteolytic fragments were separated by HPLC and analyzed by mass spectroscopy and Edman degradation. This combination of methods is novel for membrane proteins inserted into lipid bilayers. It yields direct information about the site of proteolysis and is far less demanding than

the traditional method of radio labeling and separation by SDS-PAGE. 20% of M13 procoat protein was found to insert in a transmembrane conformation when added to preformed lipid vesicles lacking a membrane potential. This correlates with the findings that *in vivo* the membrane potential is required for translocation. The established protocols could be generally applicable to the study of membrane proteins.

For the first time the membrane incorporation of a membrane protein containing a leader sequence has been quantified by FET between the intrinsic Trp/Tyr of the proteins and a lipid bound acceptor, DPH. The main portion of the free energy gain ΔG_0 results from incorporation of the proteins into the membrane interface. The translocation of the central polar domain is mainly driven by the membrane incorporation of the hydrophobic stretch in the mature part of M13 procoat protein.

The present thesis provided the basis for future investigation: refinements of our knowledge of the structure and dynamics of the different membrane associated conformations of M13 procoat protein can be expected from ongoing work in the group of Prof. Vogel and associated laboratories. Chemical modification of a series of Cys-mutant proteins from the laboratory of Prof. Kuhn are the basis of a systematic FET study using different intramolecular donor-acceptor couples. This study should provide information about the distance and the distance distribution between specific residues situated in the two hydrophobic parts of M13 procoat protein. Together with data from high resolution NMR this should provide more precise information for the molecular dynamics calculations, which are currently performed on the basis of the secondary structure and membrane topology information which have become available through the present study. The accumulated methods and data could be applicable to the study of more complex membrane proteins.

12. GENERAL BIBLIOGRAPHY

Abe, H.; Wagner, S. J. **1995** *Photochem. Photobiol.* , *61*, 402-409.

Alberts, B.; Bray, D.; Lewis, J.; Raff, M.; Roberts, K.; Watson, J. D. *Molecular biology of the cell;* third ed.; Garland Publishing, Inc.: New York, London, 1994.

Alberts, B.; Bray, D.; Lewis, J.; Raff, M.; Roberts, K.; Watson, J. D. *Molecular biology of the cell;* 3rd ed.; Garland Publishing, Inc.: New York, 1994.

Anderson, H.; Heijne, G. v. **1993** *EMBO J.* , *12*, 683-691.

Armstrong, J.; Perham, R.; Walker, J. **1981** *FEBS Lett.* , *135*, 167-172.

Arnold, G. E.; Day, L. A.; Dunker, A. K. **1992** *Biochemistry* , *31*, 7948-7956.

Asakura, T.; Yeo, J. H.; Ando, I. **1994** *Polymer J.* , *26*, 229-233.

Asbeck, V. F.; Beyreuther, K.; Kohler, H.; von Wettstein, G.; Braunitzer, G. **1969** *Hoppe-Seyler's Z. Physiol. Chemie* , *350*, 1047-1066.

Azpiazu, I.; Gomez-Fernandez, J. C.; Chapman, D. **1993** *Biochemistry* , *32*, 10720-10726.

Banner, D. W.; Nave, C.; Marvin, D. A. **1981** *Nature* , *289*, 814-816.

Bayer, M. E. **1968** *J. Gen. Microbiol.* , *395-404*,

Bayer, R.; Feigenson, G. W. **1985** *Biochim. Biophys. Acta* , *815*, 369-379.

Ben-Efraim, I.; Strahilevitz, J.; Bach, D.; Shai, Y. **1994** *Biochemistry* , *33*, 6966-6973.

Benson, S. A.; Hall, M. N.; Silhavy, T. J. **1985** *Ann. Rev. Biochem.* , *54*, 101-134.

Berkowitz, S. A.; Day, L. A. **1976** *J. Mol. Biol.* , *102*, 531-547.

Berkowitz, S. A.; Day, L. A. **1980** *Biochemistry* , *19*, 2696-2702.

Bhardwaj, D.; Singh, S. S.; Abrol, S.; Chaudhary, V. K. **1995** *J. Immunol. Meth.* , *179*, 165-175.

Blobel, G. **1980** *Proc. Natl. Acad. Sci. USA* , *77*, 1496-1500.

Blume, A.; Hübner, W.; Messner, G. **1988** *Biochemistry* , *27*, 8239-8249.

Boguski, M. J.; Leo, G. C.; Opella, S. J. **1988** *Proteins* , *4*, 123-130.

Braach-Makvytis, V. L. B.; Cornell, B. A. **1988** *Biophys. J.* , *53*, 839-843.

Brissette, J. L.; Russel, M. **1990** *J. Mol. Biol.* , *211*, 565-580.

Broomesmith, J. K.; Gnaneshan, S.; Hunt, L. A.; Mehraeinghomi, F.; Hashemzadehbonehi, L.; Tadayyon, M.; Hennessey, E. S. **1994** *Mol. Mem. Biol.* , *11*, 3-8.

Burnell, E.; Alphen, L. v.; Verkleij, A.; Kruijff, B. d. **1980** *Biochim. Biophys. Acta* , *597*, 492-501.

Byler, D. M.; Suzi, H. **1986** *Biopolymers* , *25*, 469-487.

Cao, G. Q.; Dalbey, R. E. **1994** *EMBO J.* , *13*, 4662-4669.

Cao, G. Q.; Kuhn, A.; Dalbey, R. E. **1995** *EMBO J.* , *14*, 866-875.

Cavalieri, S. J.; Goldthwait, D. A.; Neet, K. E. **1976** *J. Mol. Biol.* , *102*, 713-722.

Chamberlain, B. K.; Nozaki, Y.; Tanford, C.; Webster, R. E. **1978** *Biochim. Biophys. Acta* , *510*, 18-37.

Chamberlain, B. K.; Webster, R. E. **1976** *J. Biol. Chem.* , *251*, 7739-7745.

Chang, C. N.; Blobel, G.; Model, P. **1978** *Proc. Natl. Acad. Sci. USA* , *75*, 361-365.

Clack, B., A.; Gray, Donald, M. **1992** *Biopolymers* , *32*, 795-810.

Colnago, L. A.; Leo, G. C.; Valentine, K. G.; Opella, S. J. In *Biomolecular stereodynamics, proceedings of the fourth SUNYA conversation in the discipline biomolecular stereodynamics*; R. H. Sarma and M. H. Sarma, Eds.; Adenine: Guilderland, NY, 1986; Vol. 3; pp 147-158.

Colnago, L. A.; Valentine, K. G.; Opella, S. J. **1987** *Biochemistry* , *26*, 847-854.

Corey, D. R.; Shiau, A. K.; Yang, Q.; Janowski, B. A.; Craik, C. S. **1993** *Gene* , *128*, 129-134.

Cowan, S. W.; Schirmer, T.; Rummel, G.; Steiert, M.; Ghosh, R.; Pauptit, R. A.; Jansonius, J. N.; Rosenbusch, J. P. **1992** *Nature* , *358*, 727-733.

Cowgill, R. W. **1963** *Arch. Biochem. Biophys.* , *100*, 36-44.

Cross, T. S.; Opella, S. J. **1985** *J. Mol. Biol.* , *182*, 367-381.

Daggett, V.; Levitt, M. **1992** *J. Mol. Biol.* , *223*, 1121-1138.

Dalbey, R. E. **1991** *Mol. Microbiol.* , *5*, 2855-2860.

Dalbey, R. E.; Heijne, G. v. **1992** *TIBS* , *17*, 474-478.

Date, T.; Goodman, J. M.; Wickner, W. **1980** *Proc. Natl. Acad. Sci. USA* , *77*, 4669-4673.

Date, T.; Wickner, W. **1981** *J. Virol. , 37,* 1087-1089.
Datema, K. P.; Spruijt, R. B.; Wolfs, C. J. A. M.; Hemminga, M. A. **1988** *Biochim. Biophys. Acta , 944,* 507-515.
Datema, K. P.; van Boxtel, B. J. H.; Hemminga, M. A. **1988** *J. Magn. Reson. , 77,* 372-376.
Datema, K. P.; Visser, A. J. W. G.; van Hoek, A.; Wolfs, C. J. A. M.; Spruijt, J. A. M. **1987** *Biochemistry , 26,* 6145-6152.
Datema, K. P.; Wolfs, C. J. A. M.; Marsh, D.; Watts, A.; Hemminga, M. A. **1987** *Biochemistry , 26,* 7571-7574.
Day, L. A. **1969** *J. Mol. Biol. , 39,* 265-277.
Day, L. A.; Wiseman, R. L. In *The single-stranded DNA phages*; D. T. Denhardt, D. Dressler and D. S. Ray, Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1978; pp 605-675.
Day, L. A.; Wiseman, R. L.; Marzec, C. J. **1979** *Nucleic Acids Res. , 7,* 1393-1403.
Deber, C. M.; Khan, A. R.; Li, Z.; Joensson, C.; Glibowicka, M.; Wang, J. **1993** *Proc. Natl. Acad. Sci. USA , 90,* 11648-11652.
Deber, C. M.; Li, Z.; Joensson, C.; Glibowicka, M.; Xu, G. **1992** *J. Biol. Chem. , 267,* 5296-5300.
Deisenhofer, J.; Epp, O.; Miki, K.; Huber, R.; Michel, H. **1984** *J. Mol. Biol. , 180,* 385-398.
Deisenhofer, J.; Epp, O.; Miki, K.; Huber, R.; Michel, H. **1985** *Nature , 318,* 618-624.
Denhardt, D. T.; Dressler, D.; Ray, D. S., Editors *The single-stranded DNA phages*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1978.
Dettman, H. D.; Weiner, J. H.; Sykes, B. D. **1984** *Biochemistry , 23,* 705-712.
Dibble, A. R. G.; Yeager, M. D.; Feigenson, G. W. **1993** *Biochim. Biophys. Acta , 1153,* 155-162.
Dierstein, R.; Wickner, W. **1986** *EMBO J. , 5,* 427-431.
Driessen, A. J. M. **1993** *Biochemistry , 32,* 13190-13197.
Duffaud, G. D.; Lehnhardt. S. D.; March, P. E.; Inouye, M. **1985** *Curr. Top. Membr. Transp. , 24,* 65-104.
Duggan, D. E.; Bowman, R. L.; Brodie, B. B.; Udenfriend, S. **1957** *Arch. Biochem. Biophys. , 68,* 1-14.
Dunker, A. K.; Ensign, L. D.; Arnold, G. E.; Roberts, L. M. **1991** *FEBS Lett. , 292,* 271-274.
Dunker, A. K.; Klausner, R. D.; Marvin, D. A.; Wiseman, R. L. **1974** *J. Mol. Biol. , 81,* 115-117.
Eerola, R.; Saviranta, P.; Lilja, H.; Pettersson, K.; Lovgren, T.; Karp, M. **1994** *Biochem. Biophys. Res. Commun. , 200,* 1346-1352.
Egberts, E.; Marrink, S.-J.; Berendsen, J. C. **1994** *Eur. J. Biochem. , 22,* 423-436.
Enequist, H. G.; Hirst, T. R.; Harayama, S.; Hardy, S. J. S.; Randall, L. L. **1981** *Eur. J. Biochem. , 116,* 227-233.
Engelman, D.; Steitz, T.; Goldman, A. **1986** *Annu. Rev. Biophys, Biophys. Chem. , 15,* 321-353.
Engelman, D.; Steitz, T. A. **1981** *Cell , 23,* 411-422.
Englander, S. W.; Downer, N. R.; Teitelbaum, H. **1972** *Annu. Rev. Biochem. , 41,* 903-924.
Eroshkin, A. M.; Minenkova, O. O.; Fomin, V. I.; Ivanisenko, V. A.; Ilichev, A. A. **1993** *Mol. Mem. Biol. , 27,* 843-849.
Florine, K. I.; Feigenson, G. W. **1987** *Biochemistry , 26,* 2978-2983.
Fodor, S. P. A.; Dunker, A. K.; Ng, Y. C.; Carsten, D.; Williams, R. W. In *Bacteriophage Assembly*; M. S. Dubow, Ed.; Alan R. Liss: New York, 1981; pp 441-455.
Förster, T. **1949** *Z. Naturforschg. , 4a,* 321-327.
Frangione, B.; Nakashima, Y.; Koningsberg, W.; Wiseman, R. L. **1978** *FEBS Lett. , 96,* 381-384.
Franklin, J. C.; Ellena, J. F.; Jayasinghe, S.; Kelsh, L. P.; Cafiso, D. S. **1994** *Biochemistry , 33,* 4036-4045.
Fraser, R. D. B.; MacRae, T. P. *Conformation in fibrous proteins and related synthetic peptides*; Academic Press: New York, 1973.
Frey, S.; Tamm, L. **1991** *Biophys. J. , 60,* 922-930.
Fringeli, U. P.; Günthard; H. H. In *Membrane Spectroscopy*; E. Grell, Ed.; Springer-Verlag: Berlin, 1981; pp 270-332.
Fukunaga, Y.; Katsuragi, Y.; Izumi, T.; Sakiyama, F. **1982** *J. Biochem. (Tokyo) , 92,* 129-141.
Fukunaga, Y.; Tamaoki, H.; Sakiyama, F. N., K. **1982** *J. Biochem. (Tokyo) , 92,* 143-153.
Fung, B. K.-K.; Stryer, L. **1978** *Biochemistry , 17,* 5241-5248.
Fürste, J. P.; Pansegrau, W.; Frank, R.; Blocker, H.; Scholz, P.; Bagdasarian, M.; Lanka, E. **1986** *Gene , 48,* 119-131.
Gailus, V.; Rasched, I. **1994** *Eur. J. Biochem. , 222,* 927-931.
Gall, C. M.; Cross, T. A.; Diverdi, J. A.; Opella, S. J. **1982** *Proc. Natl. Acad. Sci. USA , 79,* 101-105.

Gall, C. M.; DiVerdi, J. A.; Opella, S. J. **1981** *JACS , 103,* 5039-5043.
Galluser, A.; Kuhn, A. **1990** *EMBO J. , 9,* 2723-2729.
Geller, B. L.; Wickner, W. **1985** *J. Mol. Biol. , 260,* 13281-13285.
Gennis, R. B. *Biomembranes, molecular structure and function*; Springer-Verlag: New York, Berlin, Heidelberg, London, Paris, Tokyo, 1989, pp 533.
Glucksman, M. J.; Bhattarchajee, S.; Makowski, L. **1992** *J. Mol. Biol. , 226,* 455-470.
Goodman, J. M.; Watts, C.; Wickner, W. **1981** *Cell , 24,* 341-345.
Greenfield, N.; Fasman, G. D. **1969** *Biochemistry , 8,* 4108-4116.
Guy-Caffey, J. K.; Rapoza, M. P.; Jolley, K. A.; Webster, R. E. **1992** *J. Bacteriol. , 174,* 2460-2465.
Hagen, D. S.; Weiner, J. H.; Sykes, B. D. **1978** *Biochemistry , 17,* 3860-3866.
Haris, I. P.; Lee, D. C.; Chapman, D. **1986** *Biochim. Biophys. Acta , 874,* 255-265.
Hart, S. L.; Knight, A. M.; Harbottle, R. P.; Mistry, A.; Hunger, H. D.; Cutler, D. F.; Williamson, R.; Coutelle, C. **1994** *J. Biol. Chem. , 269,* 12468-12474.
Heijne, G. v. **1983** *Eur. J. Biochem. , 133,* 17-21.
Heijne, G. v. **1986** *EMBO J. , 5,* 3021-3027.
Heijne, G. v. **1992** *J. Mol. Biol. , 225,* 487-494.
Heijne, G. v. **1994** *Annu. Rev. Biophys. Struct. , 23,* 167-192.
Heijne, G. v.; Gavel, Y. **1988** *Eur. J. Biochem. , 174,* 671-678.
Helenius, A.; McCaslin, D. R.; Fries, D. R.; Tanford, C. **1979** *Meth. Enzymology , 63,* 734-749.
Hemminga, M. A.; Sanders, J. C.; Spruijt, R. B. **1992** *Prog. Lipid. Res. , 31,* 301-333.
Henderson, R.; Baldwin, J. M.; Ceska, T. A.; Zemlin, F.; Beckmann, E.; Downing, K. H. **1990** *J. Mol. Biol. , 213,* 899-929.
Henderson, R.; Unwin, P. N. T. **1975** *Nature , 257,* 28-32.
Hennesy, J. P.; Johnson, W. C. J. **1981** *Biochemistry , 20,* 1085-1094.
Henry, G. D.; Sykes, B. D. **1990** *J. Mol. Biol. , 212,* 11-14.
Henry, G. D.; Sykes, B. D. **1993** *J. Magn. Res. Series B , 102,* 193-200.
Henry, G. D.; Weiner, J. H.; Sykes, B. D. **1986** *Biochemistry , 25,* 590-598.
Henry, G. D.; Weiner, J. H.; Sykes, B. D. **1987** *Biochemistry , 26,* 3626-3634.
Heukeshoven; Dernick, R. **1988** *Chromatogr. , 25,* 230-236.
Hofschneider, P. H. **1963** *Z. Naturforschg. Teil B , 18,* 203-210.
Hogrefe, H. H.; Amberg, J. R.; Hay, B. N.; Sorge, J. A.; Shopes, B. **1993** *Gene , 137,* 85-91.
Holtzer, A. **1992** *Biochemistry , 32,* 711-715.
Houserscott, F.; Baer, M. L.; Liem, K. F.; Cai, J. M.; Gehrke, L. **1994** *J. Virol. , 68,* 2194-2205.
Ito, K.; Date, T.; Wickner, W. **1980** *J. Biol. Chem. , 255,* 2123-1230.
Ito, K.; Mandel, G.; Wickner, W. **1979** *Proc. Natl. Acad. Sci. USA , 76,* 1199-203.
Jacobs, R. E.; White, S. H. B., 3421-3437. **1989** *Biochemistry , 28,* 3421-3437.
Jähnig, F. **1983** *Proc. Natl. Acad. Sci. USA , 80,* 3691-3695.
Jamieson, A. C.; Kim, S. H.; Wells, J. A. **1994** *Biochemistry , 33,* 5689-5695.
Jap, B. K.; Maestre, M. F.; Hayward, S. B.; Glaeser, R. M. **1983** *Biophys. J. , 43,* 81-89.
Jellis, C. L.; Cradick, T. J.; Rennert, P.; Salinas, P.; Boyd, J.; Amirault, T.; Gray, G. S. **1993** *Gene , 137,* 63-68.
Johnson, I. D.; Hudson, B. S. **1989** *Biochemistry , 29,* 6392-6400.
Joly, J. C.; Wickner, W. **1993** *EMBO J. , 12,* 255-263.
Jones, J. D.; Gierasch, L. M. **1994** *Biophys. J. , 67,* 1546-1561.
Jongh, H. H. J.; Hemminga, M. A.; Marsh, D. **1990** *Biochim. Biophys. Acta , 1024,* 82-88.
Jost, P. C.; Griffith, O. H.; Capaldi, R. A.; Vanderkooi, G. **1973** *Proc. Natl. Acad. Sci. USA , 70,* 480-484.
Kahn, T. W.; Engelman, D. M. **1992** *Biochemistry , 31,* 6144-6151.
Kasahara, M.; Hinckle, P. **1977** *JBC , 252,* 7384-90.
Kazmierczak, B. I.; Mielke, D. L.; Russel, M.; Model, P. **1994** *J. Mol. Biol. , 238,* 187-198.
Kennard, C. H. L.; Matsura, Y.; Tanaka, N.; Kakudo, M. **1979** *Aust. J. Chem. , 32,* 911-915.
Khan, A. R.; Deber, C. M. **1995** *Biochem. Biophys. Res. Commun. , 206,* 230-237.
Kimelman, D.; Tecoma, E. S.; Wolber, P. K.; Hudson, B. S.; Wickner, W. T.; Simoni, R. D. **1979** *Biochemistry , 18,* 5874-5464.
Kishchenko, G.; Batliwala, H.; Makowski, L. **1994** *J. Mol. Biol. , 241,* 208-213.
Knippers, R.; Hoffmann-Berling, R. **1966** *J. Mol. Biol , 21,* 281-292.
Kolner, D. E.; Guilfoyle, R. A.; Smith, L. M. **1994** *DNA Sequence , 4,* 253-257.
Konings, R. N. H.; Hulsebos, T.; van den Hondel, C. A. **1975** *J. Virol. , 15,* 570-584.

Kostrikis, L. G.; Liu, D. J.; Day, L. A. **1994** *Biochemistry , 33,* 1694-1703.
Krimm, S.; Bandekar, I. **1986** *Adv. Protein Chem. , 38,* 183-364.
Kuhn, A. **1987** *Science , 238,* 1413-1415.
Kuhn, A. **1988** *Eur. J. Biochemistry , 177,* 267-271.
Kuhn, A. **1995** *Microbiol. Rev. , in press,*
Kuhn, A.; Kreil, G.; Wickner, W. **1986** *EMBO J. , 5,* 3681-3685.
Kuhn, A.; Kreil, G.; Wickner, W. **1987** *EMBO J. , 6,* 501-505.
Kuhn, A.; Troschel, D. In *Membrane biogenesis and protein targeting*; W. Neupert and R. Lill, Eds.; Elsevier: New York, 1992; pp 33-47.
Kuhn, A.; Wickner, W. **1985** *J. Biol. Chem. , 260,* 15914-15918.
Kuhn, A.; Wickner, W. **1985** *J. Biol. Chem. , 260,* 15907-15913.
Kuhn, A.; Wickner, W.; Kreil, G. **1986** *Nature , 322,* 335-339.
Kuhn, A.; Zhu, H.-Y.; Dalbey, R. E. **1990** *EMBO J. , 9,* 2381-2385.
Kuroda, M.; Sakiyama, F.; Narita, K. **1975** *J. Biochem. (Tokyo) , 78,* 641-651.
Kusters, R.; Breukink, E.; Gallusser, A.; Kuhn, A.; Kruijff, B. d. **1994** *J. Biol. Chem. , 269,* 1560-1563.
Kwong-Keung, B.; Stryer, L. **1978** *Biochemistry , 17,* 5241-5248.
Laforet, G. A.; Kaiser, E. T.; Kendall, D. A. **1989** *J. Biol. Chem. , 264,* 14478-14485.
Lazdunski, C. J.; Baty, D.; Geli, V.; Cavard, D.; Morlon, J.; Howard, S. P.; Knibiehler, M.; Chartier, M.; Varenne, S.; Frenette, M.; Dasseux, J.-L.; Pattus, F. **1988** *Biochim. Biophys. Acta , 947,* 445-464.
Le Doan, T.; Tagasuki, M.; Aragon, I.; Boudet, G.; Montenay-Garestier, T.; Helene, C. **1983** *Biochim. Biophys. Acta , 735,* 259-270.
Lee, D. C.; Haris, P. I.; Chapman, D.; Mitchell, R. C. **1990** *Biochemistry , 29,* 9185-9193.
Lee, D. C.; Hayward, J. A.; Restall, C. J.; Chapman, D. **1985** *Biochemistry , 24,* 4364-4373.
Lee, E.; Manoil, C. **1994** *J. Biol. Chem. , 269,* 28822-28829.
Lee, J. I.; Kuhn, A.; Dalbey, R. E. **1992** *J. Biol. Chem. , 267,* 938-943.
Lee, K. C.; Cross, T. A. **1994** *Biophys. J. , 66,* 1380-1387.
Lemaster, D. M. **1994** *Progr. Nucl. Magn. Res. , 26,* 371-419.
Lemmon, M. A.; Engelman, D. M. **1992** *Curr. Opin. Struct. Biol. , 2,* 1571-1576.
Lemmon, M. A.; Engelman, D. M. **1994** *Quart. Rev. Biophys. , 27,* 157-218.
Lemmon, M. A.; Flanagan, J. M.; Hunt, J. F.; Adair, B. D.; Bormann, B.-J.; Dempsey, C. E.; Engelman, D. E. **1992** *J. Biol. Chem. , 267,* 7683-7689.
Lemmon, M. A.; Flanagan, J. M.; Treutlein, H. R.; Zhang, J.; Engelman, D. M. **1992** *Biochemistry , 31,* 12719-12725.
Leo, G. C.; Colnago, L. A.; Valentine, K. G.; Opella, S. J. **1987** *Biochemistry , 26,* 854-862.
Levitt, M.; Greer, J. **1977** *J. Mol. Biol. , 114,* 181-293.
Li, Z. M.; Glibowicka, M.; Joensson, C.; Deber, C. M. **1993** *J. Biol. Chem. , 268,* 4584-4587.
Li, Z. M.; Khan, A.; Deber, C. M. **1993** *FASEB J. , 7,* A1067-A1067.
Liao, K.-S.; Huang, H.; Khorana, G. **1984** *J. Biol. Chem. , 259,* 4200-4204.
Lill, R.; Cunningham, K.; Brundage, L. A.; Ito, K.; Oliver, D.; Wickner, W. **1989** *EMBO J. , 8,* 961-966.
Liu, D. J.; Day, L. A. **1994** *Science , 265,* 671-674.
Loeb, T. **1960** *Science , 131,* 932-933.
Magusin, P. C. M. M.; Hemminga, M. A. **1993** *Biophys. J. , 64,* 1861-1868.
Magusin, P. C. M. M.; Hemminga, M. A. **1993** *Biophys. J. , 64,* 1851-1860.
Magusin, P. C. M. M.; Hemminga, M. A. **1994** *Biophys. J. , 66,* 1197-1208.
Makino, S.; Reynolds, J. A.; Tanford, C. **1978** *J. Biol. Chem. , 248,* 4926.
Makino, S.; Woolford, J. L.; Tanford, C.; Webster, R. E. **1975** *J. Biol. Chem. , 250,* 4327-4332.
Makowski, L. In *Macromolecular Assemblies*1984; pp 203-253.
Makowski, L. **1992** *J. Mol. Biol. , 228,* 885-892.
Makowski, L. **1994** *Curr. Opin. Struct. Biol. , 4,* 225-230.
Makowski, L.; Caspar, D. L. D. **1981** *J. Mol. Biol. , 145,* 611-617.
Mandel, G.; Wickner, W. **1979** *Proc. Natl. Acad. Sci. USA , 76,* 236-240.
Marsh, D. **1990** *FEBS Lett. , 268,* 371-375.
Marvin, D.; Hoffman-Berling, H. **1963** *Z. Naturforschg. Teil B , 18,* 884-893.
Marvin, D.; Hoffman-Berling, H. **1963** *Nature , 197,* 517-518.
Marvin, D. A. **1989** *Int. J. Biol. Macromol. , 11,* 159-163.
Marvin, D. A. **1990** *Int. J. Biol. Macromol. , 12,* 125-138.

Marvin, D. A.; Hale, R. D.; Nave, C.; Citterich, M. H. **1994** *J. Mol. Biol.* , *235*, 260-286.
Marvin, D. A.; Hohn, B. **1969** *Bacteriol. Rev.* , *33*, 172-209.
Marvin, D. A.; Pigram, W. J.; Wiseman, R. L.; Wachtel, E.; Marvin, F. J. **1974** *J. Mol. Biol.* , *88*, 581-600.
Marvin, D. A.; Wachtel, E. J. **1976** *Phil. Trans. Roy. Soc. ser. B* , *276*, 81-98.
Marvin, D. A.; Wachtel, F. J. **1975** *Nature* , *253*, 19-23.
McDonnell, P. A. Ph. D. Thesis, University of Pennsylvania, 1992.
McDonnell, P. A.; Opella, S. J. **1993** *J. Magn. Res. Series B* , *102*, 120-125.
McLaughlin, S. **1977** *Curr. Top. Membr. Transp.* , *9*, 71-144.
Micheel, B.; Heymann, S.; Scharte, G.; Bottger, V.; Vogel, F.; Dubel, S.; Breitling, F.; Little, M.; Behrsing, O. **1994** *J. Immun. Meth.* , *171*, 103-109.
Model, P.; Russel, M. In *The Bacteriophages*; R. Calendar, Ed.; Plenum Press: New York, London, 1988; Vol. 2; pp 375-456.
Model, P.; Russel, M.; Boeke, J. D. In *Bacteriophage Assembly*; Alan R. Liss, Inc.: New York, 1981; pp 389-400.
Molday, R. S.; Englander, S. W.; Kallen, R. G. **1972** *Biochemistry* , *11*, 150-159.
Moll, T. S.; Thompson, T. E. **1994** *Biochemistry* , *33*, 15469-15482.
Nakashima, Y.; Koningsberg, W. **1974** *J. Mol. Biol* , *88*, 596-600.
Nakashima, Y.; Koningsberg, W. H. **1980** *J. Mol. Biol.* , *138*, 49-501.
Nakashima, Y.; Wiseman, R. L.; Koningsberg, W.; Marvin, D. A. **1975** *Nature* , *253*, 68-70.
Nambudripad, R.; Stark, W.; Opella, S. J.; Mokowski, L. **1991** *Science* , *252*, 1305-1308.
Nave, C.; Brown, R. S.; Fowler, A. G.; Ladner, J. E.; Marvin, D. A.; Provencher, S. W.; Tsugita, A. **1981** *J. Mol. Biol.* , *149*, 675-707.
Newman, J.; Day, L. A.; Dalack, G. W.; Eden, D. **1982** *Biochemistry* , *21*, 3352-3358.
Newman, J.; Swinney, H. L.; Day, L. A. **1977** *J. Mol. Biol.* , *116*, 593-603.
Ng, Y. C.; Dunker, A. K. **1981** *Prog. Clin. Biol. Res.* , *64*, 467-474.
Nozaki, Y.; Chamberlain, B. K.; Webster, R. E.; Tanford, C. **1976** *Nature* , *259*, 335-337.
Nozaki, Y.; Reynolds, J. A.; Tanford, C. **1978** *Biochemistry* , *17*, 1239-1246.
Nunari, J.; Walter, P. **1992** *Curr. Opin. Cell Biol.* , *4*, 573-580.
O'Neill, D. J.; Sykes, B. D. **1988** *Biochemistry* , *27*, 2753-2762.
O'Neill, J. D.; Sykes, B. D. **1989** *Biochemistry* , *28*, 6736-6745.
Ohkawa, I.; Webster, R. **1981** *J. Biol. Chem* , *256*, 9951-9958.
Ohno-Iwashita, Y.; Wickner, W. **1983** *J. Biol. Chem.* , *258*, 1895-1900.
Okajima, T.; Kawata, Y.; Hamaguchi, K. **1990** *Biochemistry* , *29*, 9168-9175.
Opella, S. J.; McDonnell, P. A. In *NMR of proteins*; G. M. Clore and A. M. Gronenborn, Eds.; Macmillan: London, 1993; pp 159-189.
Opella, S. J.; Stewart, G. L.; Valentine, K. G. **1987** *Quart. Rev. Biophys.* , *19*, 7-49.
Ottaviani, M. F.; Huinink, H.; Sanders, J. C.; Hemminga, M. A. **1993** *Biochim. Biophys. Acta* , 1152, 171-176.
Overman, S. A.; Aubrey, K. L.; Vispo, N. S.; Cesareni, G.; Thomas, G. J. **1994** *Biochemistry* , *33*, 1037-1042.
Overman, S. A.; Thomas, G. J. **1994** *Biophys. J.* , *66*, A394-A394.
Papavoine, C. H. M.; Konings, R. N. H.; Hilbers, C. W.; Ven, F. J. M. v. d. **1994** *Biochemistry* , *33*, 12990-12997.
Patchornik, A.; Lawson, W. B.; Witkop, B. **1958** *J. Am. Chem. Soc.* , *80*, 4747-4748.
Pawlak, M.; Kuhn, A.; Vogel, H. **1994** *Biochemistry* , *33*, 283-290.
Pawlak, M.; Meseth, U.; Dhanapal, B.; Mutter, M.; Vogel, H. **1994** *Protein Sci.* , *3*, 1788-1805.
Peelen, S. J. C. J.; Sanders, J. C.; Hemminga, M. A.; Marsh, D. **1992** *Biochemistry* , *31*, 2670-2677.
Peled, H.; Shai, Y. **1993** *Biochemistry* , *32*, 7879-7885.
Peng, K.; Visser, A. J. W. G.; Hoek, A. v.; Wolfs, C. J. A. M.; Hemminga, M. A. **1990** *Eur. Biophys. J.* , *18*, 285-293.
Peters, E. A.; Schatz, P. J.; Johnson, S. S.; Dower, W. J. **1994** *J. Bacteriol.* , *176*, 4296-4305.
Peterson, C.; Dalack, G.; Day, L. A.; Winter, W. T. **1982** *J. Mol. Biol.* , *162*, 877-881.
Popot, J.-L.; Engelman, D. **1990** *Biochemistry* , *29*, 4031-4037.
Popot, J.-L.; Vitry, C. d. **1990** *Annu. Rev. Biophys, Biophys. Chem.* , *19*, 369-403.
Porter, R. R. **1959** *Biochem. J.* , *73*, 119-126.
Pratt, D.; Laus, P.; Griffith, J. **1974** *J. Mol. Biol.* , *37*, 181-200.
Pratt, D.; Tzagoloff, H.; Erdahl, W. S. **1966** *Virology* , *30*, 397-410.

Prevelige, P. E.; King, J.; Silva, J. L. **1994** *Biophys. J. , 66,* 1631-1641.
Prevelige, P. E.; Silva, J. **1994** *Biophys. J. , 66,* A134-A134.
Putterman, D. G.; Boyle, P. D.; Yang, H. L.; Gryezan, T. J.; Frangione, B.; Day, L. A. **1984** *Proc. Natl. Acad. Sci. USA , 81,* 699-703.
Racker, E.; Chein, T. F.; Kandrach, A. **1975** *FEBS Lett. , 57,* 14-18.
Rapoport, T. **1992** *Science , 258,* 931-936.
Rapoza, M. P.; Webster, R. E. **1995** *J. Mol. Biol. , 248,* 627-638.
Rashed, I.; Oberer, E. **1986** *Microbiol. Rev. , 50,* 401-427.
Reusken, C. B. E. M.; Neeleman, L.; Bol, J. F. **1994** *Nucl. Acids. Res. , 22,* 1346-1353.
Roberts, L. M.; Dunker, A. K. **1993** *Biochemistry , 32,* 10479-10488.
Roepstorff, P. **1992** *J. Mass Spectrom. Ion Processes , 118/119,* 789-809.
Rohrer, J.; Kuhn, A. **1990** *Science , 250,* 1418-1421.
Roise, D. **1992** *Biophysics , 89,* 608-612.
Roise, D. In *Thermodynamics of membrane receptors and channels;* M. B. Jackson, Ed.; CRC Press: Florida, 1993; pp 81-126.
Rossomando; Zinder **1968** *J. Mol. Biol. , 36,* 387-99.
Rothschildt, K. J.; Clark, N. A. **1979** *Biophys. J. , 25,* 473-488.
Russel, M. **1991** *Mol. Microbiol. , 5,* 1607-1613.
Russel, M. **1993** *J. Mol. Biol. , 231,* 689-697.
Russel, M.; Model, P. **1981** *Proc. Natl. Acad. Sci. USA , 78,* 1717-1721.
Russel, M.; Model, P. **1982** *Cell , 28,* 177-84.
Sagstetter, M.; Zimmermann, R. **1988** *Bioc. Biop. R. , 153,* 498-502.
Sanders, J. C.; Haris, P. I.; Chapman, D.; Otto, C.; Hemminga, M. A. **1993** *Biochemistry , 32,* 12446-12454.
Sanders, J. C.; Nuland, N. A. J.; Edholm, O.; Hemminga, M. A. **1991** *Biophys. Chem. , 41,* 193-200.
Sanders, J. C.; Ottaviani, M. F.; Hoek, A. v.; Visser, A. J. W. G.; Hemminga, M. A. **1992** *Eur. Biophys. J. , 21,* 305-311.
Sanders, J. C.; Poile, T. W.; Spruijt, R. B.; Nuland, N. A. J. v.; Watts, A.; Hemminga, M. A. **1991** *Biochim. Biophys. Acta , 1006,* 102-108.
Sanders, J. C.; Poile, T. W.; Wolfs, C. J. A. M.; Hemminga, M. A. **1992** *Biochim. Biophys. Acta , 1110,* 218-224.
Sargent, D. F.; Bean, J. W.; Schwyzer, R. **1988** *Biophys. J. , 31,* 183-193.
Sargent, D. F.; Schwyzer, R. **1986** *Proc. Natl. Acad. Sci. USA , 83,* 5774-5778.
Schaller, H.; Beck, E.; Takanami, M. In *The single-stranded DNA phages;* D. T. Denhardt, D. Dressler and D. S. Ray, Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1978; pp 139-163.
Schaller, H.; Takanami, M. In *Single-stranded DNA Phages;* D. T. Denhardt, D. T. Dressler and D. Ray, Eds.; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1979.
Schiksnis, R. A.; Boguski, M. J.; Opella, S. J. **1988** *J. Mol. Biol. , 200,* 741-743.
Schiksnis, R. A.; Boguski, M. J.; Tsang, P.; Opella, S. J. **1987** *Biochemistry , 26,* 1373-1381.
Schiksnis, R. A.; Bogusky, M. J.; Tsang, P.; Opella, S. J. **1988** *J. Mol. Biol. , 200,* 741-743.
Schroeder, W. A.; Shelton, J. B.; Shelton, J. R. **1969** *Arch. Biochem. Biophys. , 130,* 551-556.
Schwarz, G.; Beschiaschvili, G. **1989** *Biochim. Biophys. Acta , 979,* 82-90.
Schwarz, G.; Stankowski, S.; Rizzo, V. **1986** *Biochim. Biophys. Acta , 861,* 141-151.
Schwarz, G.; Stankowski, S.; Rizzo, V. **1986** *Biochim. Biophys. Acta , 1110,* 97-104.
Schwind, P.; Kramer, H.; Kremser, A.; Ramsberger, U.; Rasched, I. **1992** *Eur. J. Biochem. , 210,* 431-436.
Schwyzer, R. **1977** *Ann. N. Y. Acad. Sci. , 297,* 3-26.
Seigneuret, M.; Kainosho, M. **1993** *FEBS Lett. , 327,* 7-12.
Sharp, K. A.; Nicholls, A.; Friedman, R.; Honig, B. **1991** *Biochemistry , 30,* 9686-9697.
Shinnar, A. E.; Kaisar, E. T. **1984** *J. Am. Chem. Soc. , 106,* 5006-5007.
Shon, K. J.; Kim, Y.; Colnago, L. A.; Opella, S. J. **1991** *Science , 252,* 1303-1305.
Silver, P.; Watts, C.; Wickner, W. **1981** *Cell , 25,* 341-345.
Smilowitz, H. **1974** *J. Virol. , 13,* 94-99.
Smilowitz, H.; Carson, J.; Robbins, P. W. **1972** *J. Supramol. Struct. , 1,* 8-18.
Smith, L. M.; Smith, B. A.; McConnell, H. M. **1979** *Biochemistry , 18,* 2256-2259.
Snell, D. T.; Offord, R. E. **1972** *Biochem. J. , 127,* 167-178.
Soekarjo, M.; Eisenhawer, M.; Kuhn, A.; Vogel, H. **1996** *Biochemistry , in press,*

Spande, T. F.; Witkop, B. **1967** *Methods Enzymol. , 11,* 506-522.
Spruijt, R. B.; Hemminga, M. A. **1991** *Biochemistry , 30,* 11147-11154.
Spruijt, R. B.; Wolfs, C. J. A. M.; Hemminga, A. M. **1989** *Biochemistry , 28,* 9158-9165.
Spyracopoulos, L.; O'Neil, J. D. **1994** *J. Am. Chem. Soc. , 116,* 1395-1402.
Stankowski, S.; Pawlak, M.; Kaisheva, E.; Robert, C. H.; Schwarz, G. **1991** *Biochim. Biophys. Acta , 1069,* 77-86.
Stubbs, C. D.; Kinosita, K.; Munkonge, F.; Quinn, P. J.; Ikegami, A. **1984** *Biochim. Biophys. Acta , 775,* 374-380.
Sugimoto, K.; Sugisaki, H.; Takanami, M. **1977** *J. Mol. Biol. , 110,* 487-507.
Surewicz, W. K.; Mantsch, H. H. **1988** *Biophys. Biochem. Acta , 952,* 115-130.
Surewicz, W. K.; Moscarello, M. A.; Mantsch, H. H. **1987** *J. Biol. Chem. , 262,* 8598-8602.
Symmons, M. F.; Welsh, L. C.; Nave, C.; Marvin, D. A.; Perham, R. N. **1995** *J. Mol. Biol. , 245,* 86-91.
Tamaoki, H. **1978** *J. Biochem. , 83,* 771-781.
Tamm, L. K.; Bartoldus, I. **1988** *Biochemistry , 27,* 7453-7458.
Tanford, C. *The Hydrophobic Effect: Formation of micelles and biological membranes.*; 2nd ed.; John Wiley: New York, 1980.
Terzi, E.; Hölzemann, G.; Seelig, J. **1994** *Biochemistry , 33,* 7434-7441.
Teshima, K. **1975** *J. Biochem. (Tokyo) , 78,* 641-651.
Teshima, K.; Kuramitsu, S.; Hamaguchi, K.; Sakiyama, F.; Mizuno, K.; Yamasaki, N. **1980** *J. Biochem. (Tokyo) , 83,* 771-781.
Thiaudière, E.; Soekarjo, M.; Kuchinka, E.; Kuhn, A.; Vogel, H. **1993** *Biochemistry , 32,* 12186-12196.
Thomas, G. J.; Prescott, B.; Day, L. A. **1983** *J. Mol. Biol. , 165,* 321-56.
Thomas, G. J. J.; Day, L. A. **1981** *Proc. Natl. Acad. Sci. USA , 78,* 2962-2966.
Thornton, K.; Gorenstein, D. G. **1994** *Biochemistry , 33,* 3532-3539.
Thornton, K.; Wang, Y.; Weiner, H.; Gorenstein, D. G. **1993** *J. Biol. Chem. , 268,* 19906-19914.
Tobias, D. J.; Klein, M. L.; Opella, S. J. **1993** *Biophys. J. , 64,* 670-675.
Träuble, H.; Teubner, M.; Woolley, P.; Eibl, H. **1976** *Biophys. Chem. , 4,* 319-342.
Tsuboi, M. **1962** *J. Polymer sci. , 59,* 139-153.
Turner, R. J.; Weiner, J. H. **1993** *Biochim. Biophys. Acta , 1202,* 161-168.
Unwin, N. **1993** *Cell , 72,* 31-41.
Valentine, K. G. Thesis, University of Pennsilvania, 1987.
Valentine, K. G.; Schneider, D. M.; Leo, G. C.; Colnago, L. A.; Opella, S. J. **1985** *Biophys. J. , 49,* 36-38.
Vaz, W. L. C.; Kaufman, K.; Nicksch, A. **1977** *Anal. Biochem. , 83,* 385-393.
Ven, F. J. M. v. d.; Vanos, J. W. M.; Aelen, J. M. A.; Wymenga, S. S.; Remerowski, M. L.; Konings, R. N. H.; Hilbers, C. W. **1993** *Biochemistry , 32,* 8322-8328.
Veronese, F. D. M.; Willis, A. E.; Boyerthompson, C.; Appella, E.; Perham, R. N. **1994** *J. Mol. Biol. , 243,* 167-172.
Vogel, H. **1987** *Biochemistry , 26,* 4562-4572.
Vogel, H. **1992** *Quart. Rev. Biophys. , 25,* 433-457.
Vogel, H.; Nilsson, L.; Rigler, R.; Voges, K. P.; Jung, G. **1988** *Proc. Natl. Acad. Sci. USA , 85,* 5076-5071.
Webster, R. E.; Lopez, J. In *Virus Structure and Assembly*; S. Casjens, Ed.; Jones and Bartlett: Boston (MA), Portola Valley (CA), 1985; pp 235-268.
Weezenbeek, P. v.; Hulzebos, T.; Schoenmakers, J. **1980** *Gene , 11,* 129-148.
Weiss, M.; Kreusch, S.; Schiltz, E.; Nestel, U.; Welte, W.; Weckesser, J.; Schultz, G. **1991** *FEBS Lett. , 280,* 379-382.
Weiss, M.; Wacker, T.; Weckesser, J.; Welte, W.; Schulz, G. **1990** *FEBS Lett. , 267,* 268-272.
Wen, F. S.; Tseng, Y. H. **1994** *J. Gen. Virol. , 75,* 15-22.
Wetlaufer, D. B. **1962** *Adv. Protein Chem. , 17,* 303-390.
White, S. H.; Wimley, W. C. **1994** *Curr. Opp. Struct. Biol. , 4,* 79-86.
Whitley, P.; Nilsson, L.; Heijne, G. v. **1993** *Biochemistry , 32,*
Wickner, W. **1975** *Proc. Natl. Acad. Sci. USA , 72,* 4749-4753.
Wickner, W. **1976** *Proc. Natl. Acad. Sci. USA , 73,* 1159-1163.
Wickner, W. **1977** *Biochemistry , 16,* 254-258.
Wickner, W. **1979** *Annu. Rev. Biochem. , 48,* 23-45.
Wickner, W. **1983** *TIBS , March,* 90-94.

Wickner, W. **1988** *Biochemistry , 27,* 1082-1093.
Wickner, W.; Mandel, G.; Zwizinski, C.; Bates, M.; Killick, T. **1978** *Proc. Natl. Acad. Sci. USA , 75,* 1754-1758.
Wickner, W. T.; Lodish, H. F. **1985** *Science , 230,* 400-407.
Wiech, W.; Sagstetter, M.; Müller, G.; Zimmermann, R. **1987** *EMBO J. , 6,* 1011-1016.
Williams, R.; Dunker, A. K. **1977** *J. Biol. Chem. , 252,* 6253-6255.
Wilson, M. L.; Dahlquist, F. W. **1985** *Biochemistry , 24,* 1920-1928.
Wiseman, R.; Berkowitz, S. A.; Day, L. A. **1976** *J. Mol. Biol. , 102,* 549-561.
Wiseman, R. L.; Day, L. A. **1977** *J. Mol. Biol. , 116,* 607-611.
Wolber, P. K.; Hudson, B. S. **1979** *Biophys. J. , 26,*
Wolfe, P.; Silver, P.; Wickner, W. **1982** *J. Biol. Chem. , 257,* 7898-7902.
Wolfe, P. B.; Rice, M.; Wickner, W. **1985** *J. Biol. Chem. , 260,* 1836-1841.
Wolfs, C. J. A. M.; Horvath, L. I.; Marsh, D.; Watts, A.; Hemminga, M. A. **1989** *Biochemistry , 28,* 9995-10001.
Woolford, J. L. J.; Webster, R. E. **1975** *J. Biol. Chem. , 250,* 4333-4339.
Wu, J.; Gorenstein, D. G. **1993** *J. Am. Chem. Soc. , 115,* 6843-6850.
Yamasaki, N.; Tsujita, T.; Eto, T.; Masuda, S.; Mizuno, K.; Sakiyama, F. **1979** *J. Biochem. , 86,* 1291-1300.
Yang, P. W.; Mantsch, H. H.; Arrondo, J. L. R.; Saint-Girons, I.; Guillou, Y.; Cohen, G. N.; Bârzu, O. **1987** *Biochemistry , 26,* 2706-2711.
Zimmermann, R.; Sagstetter, M.; Lewis, M. J.; Pelham, H. R. B. **1988** *EMBO J. , 7,* 2875-2880.
Zinder, N. R.; Valentine, M. R.; Stoeckenius, W. **1963** *Virology , 20,* 638-640.
Zwizinski, C.; Wickner, W. **1982** *EMBO J. , 1,* 573-578.

APPENDIX A. CD FIT PROGRAM

The following is a rendering of the program "protein character fit", written by Dr. Pierre Infelta for this project. Version 3/4/92. The program analyzes CD spectra between 186 and 250 nm according to a modified procedure of Hennesy and Johnson (see Chapter 8).

The program was written in Mathematica and is run by us on a Macintosh cx or a Quadra 700. It reads in directly, through a universal disk drive, data files created by the DOS program running our AVIV 60DS circular dichroism spectrometer. It has tremendously alleviated the task of analyzing our CD data and we are very grateful to its author.

PROGRAM DESCRIPTION

The data is assumed to be on a Mac readable diskette under ASCII form. The diskette can be formatted either for Mac or for DOS. The data consists of data pairs (wavelength <tab> ellipticity <cr>). It is delimited on both ends by _data_ (underlines inclusives).

If the file is typed using Word, store the input data as text only and include _data_ at the beginning and the end of the file.

The reference set of data is made of 15 proteins for which the spectrum have been measured and of which the three dimensional structure is know from high-resolution x-ray diffraction experiments. The reference spectra are in file refprot1-cd33 and the corresponding structures in file strucref1. The data is from 186 to 250 nm.

The program requires a certain number of interactive inputs.

First, you must type the name of the data file (when requested) datafilename. You may have to direct mathematica to it, especially if the file resides on diskette. The routine readfile[filename] returns the measured ellipticity (measin), the corresponding wavelength (wavelin), the minimum and maximum wawelength (wavemin, wavemax)where the measurement was performed.

The rawdata is presented with the pertinent wavelength information to decide in what zone the data should be fitted (lambdamin, lambdamax). One must also input a correction factor corr to take into account concentration, cell thickness. ([Θ] = θ x 100/cnl)

You obtain then a display of cross references indicating the filenames and other parameters, as well as the result of the fit. Only five characters are kept corresponding to the 5 largest eigenvalues of the base system created from the reference protein spectra.

spec is the experimental spectrum to be analyzed.
strucmat is the structure matrix of the reference proteins(strucref). (numprot columns).
ctrans has numprot columns.
ctransneeded is a sub matrix of crtrans reduced to comprise only the appropriate wavelengths.
q==r verifies that to 4 significant digits the eigen vector are in the same order as the eigenvalues.
vectordtronk is the truncated matrix of of the eigen vectors. xv represents the components of the unknown sample in the base of the reference proteins.
strucresul is the proportions of the various structures present.

We also print a normalized result such that the sum of the structures be 1. We plot the residuals (residu) which correspond to the difference between the experimental spectrum expespec (only in wavelength domain considered) and calcspec. These curves are displayed on a sinlge graph. Selecting the appropriate result, one can print the selection which fits on one A4 page.

The results can be archived in Word where you may add some finishing touch if necessary, such as some experimetal conditions or bold faces printing, improvements on the tables of results etc.

garde is the number of vectors we want to keep. It must be smaller than the actual number of eigenvectors. The eigen vectors corresponding to the largest eigenvalues(absolute value) are kept.
valord and vectord are the ordered eigenvalues and eigenvectors.
Function decrord[ ] sorts a matrix of eigenvalues in decreasing order of absolute values, and the corresponding eigen vectors follow the sorted matrix. (The number of eigen values and eigen vectors should be the same. The program does not check this fact.)

The first parameter is the list of eigenvalues (1 dimensional table). The second parameter is the matrix of eigenvectors. The third parameter is the number of values and vectors one wants to keep.

Function readfile[] reads the entire file in a single string, locates the data of interest, suppresses linefeeds , and returns are replaced by spaces. the resulting string is used as a stream to read the interleaved data. The raw data is then plotted and the function returns all of the needed data. (wavelini, ymeasin, wavemin, wavemax)

PROGRAM BODY

```
$Path=Join[$Path,{FileNames["*",Directory[]
,2]}];
(*======order the list val, track the order to
vectors vec, keep only veals and vecs ======*)
decrord[val_,vec_,keep_] := (valtem=val;
vectem=vec;
valord={};
vectord={};
fin=Length[val] - keep;
dum=Abs[valtem];
While [Length[dum] > fin,{m=Max[dum];
pos=Position[dum,m];
AppendTo[valord,Part[valtem,Flatten[pos[[1]]
]]]; (*Print [valord];*)
valtem=Drop[valtem,Flatten[pos[[1]]]];
AppendTo[vectord,Part[vectem,Flatten[pos[[1]]
]]]; (*Print [vectord];*)
vectem=Drop[vectem,Flatten[pos[[1]]]];
(*Print ["================="];*);
dum=Abs[valtem];}];
res=Flatten[{valord,vectord},2];
Return[res])
(*=== read the data file and display it ====*)
readfile[filename_] :=(
in=OpenRead[filename];
rawdat=Read[in,Record,RecordSeparators ->
{}];
limits=Flatten[StringPosition[rawdat,"_data
_"]];
wanted=Part[limits,{2,3}] +{1,-1};
interes=StringTake[rawdat,wanted];
interes=StringReplace[interes,{"\r"->"","\n"-
>" "}];
ini=StringToStream[interes];
data=Sort[Partition[ReadList[ini,Number],2]];
temp=Transpose[data];
wavelini=temp[[1]];
wavemin=Min[wavelini];
wavemax=Max[wavelini];
ymeasin=0.001 temp[[2]];
ListPlot[Transpose[{wavelini,ymeasin}],
Frame ->True,
FrameLabel->{FontForm["l",{"Symbol",14}],
FontForm["Q",{"Symbol",14}]},
DefaultFont->{"Palatino-Roman",10},
RotateLabel->False,
PlotStyle->{PointSize[0.015]}];
data={wavelini,ymeasin,wavemin,wavemax};
Print [MatrixForm[{{"Original File","min
lambda",
"max lambda"},
{filename,wavemin,wavemax}}]];
Return [data])
(* =====
data that stays the same for all cases*
======*)
fichier="refprot1-cd33";
struc="strucref1";
numprot=15;
garde=5;
(*lecture des références*)
c=ReadList[fichier,Number];
ctrans=3.3 Partition[c,numprot];
waveref=Table[lamb,{lamb,186.,250.,2.}];
strucre=ReadList[struc,Number];
strucmat=Partition[strucre,numprot];
(*=====Start of the main program===*)
datafilename=InputString["Type DOS data
Filename"];
readfile [datafilename];
(*==========Select the analysis zone and
the corresponding data==========*)
dataintemp=InputString["type in Lambdamin,
Lambdamax for analysis"];
corr=Input["spectrum correction factor to get
values in [theta]"];
dataintemp=StringReplace[dataintemp,{","-
>" "}];
in=StringToStream[dataintemp];
datain=Flatten[ReadList[in,{Number,Number}
]];
lambdamin=Ceiling[.5*datain[[1]]]*2.;
lambdamax=Floor[.5*datain[[2]]]*2.;
startpoint=Position[wavelini,lambdamin];
endpoint=Position[wavelini,lambdamax];
wanted=Flatten[{startpoint,endpoint}];
```

```
wavetofit=Take[wavelini,wanted];
thetatofit=Take[ymeasin,wanted];
Print [" "];
Print ["Analysis from ",First[wavetofit],
 " to ",Last[wavetofit]," nm"];

spec= corr thetatofit;
startref=Position[waveref,lambdamin];
endref=Position[waveref,lambdamax];
needed=Flatten[{startref,endref}];
ctransneed= Take[ctrans,needed];
cnormalneed=Transpose[ctransneed];

acarr=cnormalneed.ctransneed;
{valeurs,vecteurs}=Eigensystem[acarr];
MatrixForm[valeurs];
MatrixForm[vecteurs];
valsdiagon=DiagonalMatrix[valeurs];
q = SetPrecision[Transpose[vecteurs]
.valsdiagon ,4];
r = SetPrecision[acarr.Transpose[vecteurs],4];
ord=(q == r);
decrord [valeurs,vecteurs,numprot];
MatrixForm[valord=Flatten[valord,1]];
vectord=Flatten[vectord,1];
d=cnormalneed.Transpose[spec] ;
g=vectord.d ;
p=Inverse[DiagonalMatrix[valord]].g;
vectordtronk=Take[vectord,garde];
ptronk=Take[p,garde];
xv=Transpose[vectordtronk].ptronk;
Print ["Fit from File ", datafilename," ",ord];
Print ["Spectrum correction factor ",corr];
Print[" "];
strucresul=strucmat.xv;
sig=Sum[strucresul[[i]],{i,5}];
fac = 1./sig;
strucresulnorm=fac strucresul;
textestr={"Ordered-Helix", "Disordered-
Helix", "Beta-Strand",
  "Turns", "Undefined"};
printform=TableForm[SetPrecision[Transpose[{
textestr,
strucresul,strucresulnorm}],4],TableSpacing-
>{0,4}];
Print [printform];
Print [" "];
Print ["The sum of the structure is ",sig];
calc=ctransneed.xv;
calcspec=Transpose[{wavetofit,calc}];
expespec= Transpose[{wavetofit,spec}];
residu = Transpose[{wavetofit,calc -spec}];
plot1=ListPlot[calcspec,PlotJoined->True,
Frame ->True,
FrameLabel->{FontForm["I",{"Symbol",14}],
FontForm["Q",{"Symbol",14}]},
DefaultFont->{"Palatino-Roman",10},
DisplayFunction->Identity,
RotateLabel -> False};
plot2=ListPlot[expespec,
Frame ->True,
FrameLabel->{FontForm["I",{"Symbol",14}],
FontForm["Q",{"Symbol",14}]},
DefaultFont->{"Palatino-Roman",10},
DisplayFunction->Identity,
RotateLabel -> False};
plot3=ListPlot[residu,
PlotJoined->True,
Frame->True,FrameLabel-
>{FontForm["I",{"Symbol",14}],
FontForm["Q",{"Symbol",14}]},
PlotStyle->{Thickness[.006],GrayLevel[.5]},
DefaultFont->{"Palatino-Roman",10},
RotateLabel -> False};
Show[plot1,plot2,plot3,DisplayFunction-
>$DisplayFunction]
(* =======End of program body ========*)
```

Plots are rendered in post-script. Selecting the area containing the plots and the fit results permits a partial print-out of a one page report on the performed fit.

POSTSCRIPT

The largest and broadest contribution during this thesis was due to Prof. Horst Vogel who followed and guided me enthousiastically with an open and critical mind and who, in spite of his many other obligations, always seemed to be there when it mattered, for discussions or joint experiments until either one of us was to exhausted to continue. A Eric Thiaudière je dois beaucoup pour son amitié et humour et pour ses leçons de français réel et de biophysique.

This thesis would never have come to be without the many people I have had the privilige to interact with during my stay in Lausanne. I welcome this opportunity to cherish our encounter. Your presence in my life have brought an enrichment to me on a professional as well as personal level. The intricate nature of all these interactions makes it difficult if not impossible to enumerate all of you and to differentiate the exact nature of your influence on me and on my work.

A short and incomplete gallery of honor:

bureaucracy:	Mme Gourdou
discussions:	all@ICP4, Andreas Kuhn, Marcus Hemminga, Ruud Spruijt, Cor Wolfs
electron microscopy:	Henning Stahlberg
corrections de français :	Marie-Lorraine Bontron, Eric Thiaudière
fluorescence:	Robin Humphry-Baker
Macintosh:	Pierre Infelta, Marie-Lorraine Bontron
mass spectroscopy:	Sophie Haebel, Prof. Promé, Katharina Servies
molecular biology:	Andreas Kuhn and co-workers
monolayers and insults:	Mila Bontcheva
tango and other moral support:	Marjon Reinders, Isabelle Spirig, Marie-Lorraine Bontron
technical support:	atelier mécanique, atelier électronique
unix, internet:	Shawn Koppenhoefer

CURRICULUM VITAE

The author of this thesis was born on January 27th, 1963 in Utrecht, The Netherlands. After graduating from the *Stedelijk Gymnasium* in Utrecht in 1981 he started studying chemistry at the University of Utrecht in the same year. He graduated on August 26th, 1989 (*doctoraalexamen oude stijl*, equivalent with M. Sc.) with a major in biochemistry and didactics of chemistry (including the teacher's certificate) as a subsidiary subject. On August 1st, 1990 he started working as a scientific assistant at the EPFL in the group of Prof Vogel. On April 27th, 1993 he was accepted as a doctoral student. His investigations under the supervision of Prof. Vogel have led to the present thesis. Presently he is working as an internet specialist in The Netherlands in the function of editor-in-chief Internet for the VSNU, the Association of Dutch Universities and as architect of the first digital province, Holland-Net.