Differentiation between transmembrane helices and peripheral helices by the deconvolution of circular dichroism spectra of membrane proteins

KYUSUNG PARK, ANDRÁS PERCZEL, AND GERALD D. FASMAN

Graduate Department of Biochemistry, Brandeis University, Waltham, Massachusetts 02154

(RECEIVED January 31, 1992; REVISED MANUSCRIPT RECEIVED March 20, 1992)

Abstract

The interpretation of the circular dichroism (CD) spectra of proteins to date requires additional secondary structural information of the proteins to be analyzed, such as X-ray or NMR data. Therefore, these methods are inappropriate for a CD database whose secondary structures are unknown, as in the case of the membrane proteins. The convex constraint analysis algorithm (Perczel, A., Hollósi, M., Tusnády, G., & Fasman, G.D., 1991, *Protein Eng.* 4, 669–679), on the other hand, operates only on a collection of spectral data to extract the common spectral components with their spectral weights. The linear combinations of these derived “pure” CD curves can reconstruct the original data set with great accuracy. For a membrane protein data set, the five-component spectra so obtained from the deconvolution consisted of two different types of α helices (the α helix in the soluble domain and the α_T helix, for the transmembrane α helix), a β -pleated sheet, a class C-like spectrum related to β turns, and a spectrum correlated with the unordered conformation. The deconvoluted CD spectrum for the α_T helix was characterized by a positive red-shifted band in the range 195–200 nm (+95,000 deg cm² dmol⁻¹), with the intensity of the negative band at 208 nm being slightly less negative than that of the 222-nm band (–50,000 and –60,000 deg cm² dmol⁻¹, respectively) in comparison with the regular α helix, with a positive band at 190 nm and two negative bands at 208 and 222 nm with magnitudes of +70,000, –30,000, and –30,000 deg cm² dmol⁻¹, respectively.

Keywords: circular dichroism spectra; membrane proteins; secondary structure; transmembrane helices

Conformational studies of membrane proteins lag far behind those of soluble proteins mainly due to the difficulties associated with crystallization of membrane proteins for X-ray diffraction studies and to the restricted movement of the proteins embedded in the membrane for NMR studies (Kühlbrandt, 1988; Smith & Griffin, 1988).

X-ray crystallography is still the only routine method for determining the three-dimensional structures of biological macromolecules at high resolution (Kühlbrandt, 1988) that can be utilized for the comparison of the structures determined by CD deconvolution. Three-dimensional structures of five membrane protein complexes have been determined. These are: the bacterial photosynthetic reaction centers of *Rhodobacter viridis* (*Rb. viridis*) (Deisenhofer et al., 1984, 1985; Deisenhofer & Michel, 1989) and of *Rhodobacter sphaeroides* (Chang et al., 1986; Allen et al., 1986, 1987; Feher et al., 1989), porin from *Rhodobacter capsulatus* (Kreusch et al., 1991; Weiss et al., 1991), photoactive yellow protein (PYP) from the purple photoautotrophic bacterium *Ectothiorhodospira halophila* (McRee et al., 1989), and bacteriorhodopsin from *Halobacterium halobium* (Henderson et al., 1990). The structure of membrane proteins has not been determined by alternate methods, such as NMR, because a membrane-embedded protein, due to restriction of motion that renders anisotropy in dipolar, quadrupolar,

Reprint requests to: G.D. Fasman, Graduate Department of Biochemistry, Brandeis University, Waltham, Massachusetts 02154.

Abbreviations: ATP, adenosine triphosphate; CDTA, cyclohexane diaminotetraacetic acid; C₈E₅, pentaethylene glycol mono-octyl ether (octylpenta-oxyethylene); C₈E_n, octyl(polydisperse)-oligo-oxyethylene (octylpoly-oxyethylene; OPOE); C₈E₄, tetraethylene glycol mono-octyl ether (octyltetra-oxyethylene); C₁₀E₇, decylhepta-oxyethylene (heptaethylene glycol decyl ether); C₁₂E₈, dodecyl-octa-oxyethylene (octaethylene glycol dodecyl ether); DTT, dithiothreitol; EGTA, ethylene glycol tetraacetic acid; LDAO, dimethyl dodecylamine-*N*-oxide; LM, *n*-dodecyl- β -D-maltoside (lauryl maltoside); MES, 2-[*N*-morpholino] ethane sulfonic acid; MOPS, 3-[*N*-morpholino] propane sulfonic acid; OG, 1-*O*-octyl- β -D-glucopyranoside; OPOE, octyl(polydisperse)-oligo-oxyethylene (octylpoly-oxyethylene; C₈E_n); PAGE, polyacrylamide gel electrophoresis; PMSF, phenylmethylsulfonyl fluoride; SDS, sodium dodecyl sulfate; Tween 20, polyoxyethylene sorbitan monolaurate.

and/or chemical shifts in resonances, shows powder patterns in NMR spectroscopy making it harder to interpret the data (Bovey, 1988; Smith & Griffin, 1988). Even solid state NMR, such as cross-polarization magic angle spin NMR, to compensate the anisotropy, has limited success in obtaining structural information for prosthetic groups in membrane proteins, such as retinal in bacteriorhodopsin (Harbison et al., 1985; Smith et al., 1989) and in rhodopsin (Mollevanger et al., 1987; Smith et al., 1987). Although electron microscopy (EM) has provided the overall shape of a protein including the number of helix strands, detailed information of secondary structure is not likely to be easily obtained from EM. Therefore, there is not an adequate conformational database to evaluate the deconvolution results, as was the case for the soluble proteins (Perczel et al., 1992b).

CD spectroscopy has been recognized for its utility and simplicity in operations for conformational studies, even though interpreting the data has not been simple and has been a major controversy concerning the technique. The CD spectrum of each secondary conformation may be obtained on the basis of the CD spectrum either obtained from model polypeptides (Holzwarth & Doty, 1965; Sarkar & Doty, 1966; Greenfield & Fasman, 1969; Woody, 1974; Brahms & Brahms, 1980; Hollósi et al., 1987a,b; Perczel et al., 1991a, 1992a; Perczel & Fasman, 1992) or extracted from a set of protein CD spectra whose X-ray data are available (Saxena & Wetlaufer, 1971; Chen et al., 1974; Chang et al., 1978; Bolotina et al., 1981; Hennessey & Johnson, 1981; Provencher & Glöckner, 1981). In one approach (Saxena & Wetlaufer, 1971; Chen et al., 1974; Chang et al., 1978; Bolotina et al., 1981), the references are the pure component spectra extracted from a set of CD spectra of proteins previously determined by X-ray diffraction (fixed reference methods). In the other approach (Hennessey & Johnson, 1981; Provencher & Glöckner, 1981), a set of the CD spectra of proteins is used as reference (variable reference methods). On the basis of the results of these investigations, the CD spectrum of the α helix consists of a positive peak around 190–195 nm, with magnitudes of about 60,000–80,000 deg cm² dmol⁻¹ (π - π^* transition) and two negative peaks at 208 (π - π^* transition) and 222 nm (n - π^* transition), with the magnitude of $-36,000 \pm 3,000$ deg cm² dmol⁻¹ (Tinoco et al., 1963; Holzwarth & Doty, 1965; Woody & Tinoco, 1967; Greenfield & Fasman, 1969; Saxena & Wetlaufer, 1971; Chen et al., 1974). The CD spectrum of a β sheet has the π - π^* transition around 195–200 nm, has a magnitude of 30,000–50,000 deg cm² dmol⁻¹, and has the n - π^* transition at about 215–220 nm, with a magnitude of $-10,000$ to $-20,000$ deg cm² dmol⁻¹ (Sarkar & Doty, 1966; Greenfield & Fasman, 1969; Saxena & Wetlaufer, 1971), whereas the spectrum generally correlated with the atypical (previously called "random") conformation has a large negative band around 200 nm, has the magnitude of $-20,000$ deg cm² dmol⁻¹, has a small

positive peak or a shoulder with a small negative value at 220 nm (Greenfield & Fasman, 1969; Fulmer, 1979; Yang et al., 1986). For reverse turns, three classes of CD spectra for the different types of turns were theoretically calculated by Woody (1974). The most abundant type (type I or III) of β turn showed a class C CD spectrum (Chou & Fasman, 1977; Hollósi et al., 1987a,b) that resembles the spectrum of the α helix, with smaller magnitudes, whereas the type II turns yielded a class B spectrum with a maximum below 200 nm and a minimum above 220 nm (Hollósi et al., 1987a,b; Perczel et al., 1991a, 1992a; Perczel & Fasman, 1992). However, the CD spectrum of the β turn derived from 18 protein CD spectra by Chang et al. (1978) was a mirror image of the CD spectrum for an α helix.

The CD spectra of membrane proteins often exhibit various degrees of distortions in shapes, intensities, and/or positions of the CD bands, and shifts in cross-over points (Urry & Long, 1980). Such distortions are recognized as the optical artifacts of differential light scattering and differential absorption flattening (Duysens, 1956; Gordon & Holzwarth, 1971; Urry & Long, 1980; Mao & Wallace, 1984; Glaeser & Jap, 1985). The extent of the flattening is a function of the size of the particles and the concentration of the chromophores within a particle (Mao & Wallace, 1984).

The differential light scattering and differential absorption flattening shown in CD spectroscopy (Duysens, 1956; Gordon & Holzwarth, 1971; Urry & Long, 1980; Mao & Wallace, 1984; Glaeser & Jap, 1985) increase the uncertainties of the measured CD spectrum for membrane proteins. However, to avoid the uncertainties arising from the light scattering and absorption flattening, the CD spectra of the membrane proteins reported herein were measured in detergent-solubilized forms, with the single exception of the purple membrane from *H. halobium*.

There have been several experimental as well as theoretical approaches to differentiating the CD bands from the artifacts of the membrane protein CD spectra. One of the most direct experimental approaches is to increase the solid angle of light acceptance in the spectropolarimeter (Chin et al., 1987). Because the differential light scattering originated from the difference in refractive indices between the membrane particles and the media (Schneider & Rosenheck, 1982), increasing the refractive index of the medium by adding high refractive index media, such as glycerol and sucrose, has been an alternative approach for CD measurements in the visible region (Brith-Lindner & Rosenheck, 1977). Unfortunately these media are not transparent enough to be used in the far UV region for conformational investigation of proteins. A drastic measure for compensation of the light scattering is preparation of small unilamellar vesicles in which the size of the liposome is nearly homogeneous and smaller than 100 nm in diameter, with such a small number of the proteins per vesicle that the number of proteins per vesicle approaches

unity (Mao & Wallace, 1984; Glaeser & Jap, 1985). On the other hand, mathematical approaches to reducing the optical artifacts of the membrane protein CD spectra are cumbersome and often time consuming (Urry & Long, 1980; Schneider & Rosenheck, 1982).

Membrane proteins are known to have the ability to tolerate mildly disruptive detergents without loss of activity, depending on the protein as well as the detergent, even though the optimal detergent for a particular protein has to be found empirically (Helenius et al., 1979). For such proteins, CD spectra can be measured in the optimal detergent without the distortion from the differential light scattering and the differential absorption flattening (Gordon & Holzwarth, 1971; Helenius et al., 1979). Due to the numerous artifacts encountered in membrane protein CD spectra, CD studies of membrane proteins have not been widely used. In an effort to alleviate such difficulties by providing a standard for membrane protein CD studies, an extensive study of membrane proteins has been underway in this laboratory, in parallel with utilizing the convex constraint analysis (CCA) deconvolution algorithm (Perczel et al., 1989, 1991b).

As a way of overcoming the bias effect of the composition of the database, the convex constraint algorithm operates to extract the common components based on the number of the components, P , as an input parameter (Perczel et al., 1991b). As a result, unlike most of the other algorithms, which offer a fixed number of components as an intrinsic parameter (with the exception of the Pancoska & Keiderling [1991] method), the CCA has the flexibility depending on the database to be deconvoluted. For example, when the database is composed of a set of CD spectra of proteins that had high content of α helix and random conformation and was devoid of significant amounts of β -sheet conformation, selecting a lower number of components, e.g., $P = 3$, would have a better probability of extracting the correct components than with a higher number, e.g., $P = 4$ or 5, in which case the algorithm was forced to extract the additional, poorly represented component(s) resulting in unstable components. On the other hand, selecting too large a P value (approaching the number of spectra in the database, N) would result in each spectrum appearing as an independent component with a very low standard deviation. Therefore, selecting the P value corresponding to the inflection point on a plot of standard deviation, $\sigma(P)$, as a function of the number of components, P , would yield the number of common components contained in the CD spectral database, with moderate deviation (Perczel et al., 1991b). At the same time, it may be possible to isolate and differentiate the reference CD spectrum for the transmembrane α helix and the peripheral α helix. This paper reports an extensive effort to measure the CD spectra of a large number of membrane proteins, aiming to distinguish pure CD spectra correlated with the transmembrane helices (α_T helices) from the peripheral helices (α helices).

Results and discussion

The estimation of secondary structures from the CD spectra of membrane proteins was performed using two methods, one (Chang et al., 1978) representing the fixed reference method and the other (Manavalan & Johnson, 1987) variable reference methods, for comparison with the results from the CCA method (Perczel et al., 1989, 1991). The Provencher and Glöckner (1981) analytical method and many other algorithms were omitted because of space constraints; omission does not imply denigration of these methods.

Decomposition of the membrane protein CD spectra by a fixed and a variable reference method

The X-ray study of the reaction center from *Rhodospseudomonas viridis* was reported to have 39% α -helix (23% transmembrane and 16% cytosolic helices) and 7% β -sheet conformation (Deisenhofer et al., 1985; Deisenhofer & Michel, 1989). The Yang CD decomposition method (Chang et al., 1978) yielded 47.2% α -helix and 0% β -sheet conformation (Table 1), and the variable selection (VS) method (Manavalan & Johnson, 1987) yielded a total secondary structural conformation of 139%, 39% of which was in the α helix and 27% in the β sheet (Table 2). The CD spectrum of the reaction center ($\square$ — $\square$ in Fig. 1) could not be compared directly with that reported by Nabedryk et al. (1985), as their spectrum was not converted to molar ellipticity. The CD spectrum for the reaction center from *Rb. sphaeroides* reported herein ($\diamond$ — $\diamond$ in Fig. 1), whose structure was previously determined by X-ray analysis as having 51% α helix (34% transmembrane and 17% cytosolic helices) and 15% β sheet (Allen et al., 1987; Feher et al., 1989), showed that the ellipticity at 195 nm was about twice as large as the magnitude of that of the reaction center from *Rps. viridis*. The decomposition results obtained by both algorithms, the Yang and VS methods, were similar to the X-ray results (Table 1 for the Yang analysis and Table 2 for the VS method). The CD spectra of the purified and the redissolved light harvesting complex (LHC) showed a significant difference in the positive band at 192 nm (Fig. 1), with the magnitude of the purified LHC ($\triangle$ — $\triangle$ in Fig. 1) being about twice as large as that for the LHC redissolved from crystals ($\times$ — $\times$ in Fig. 1). This difference was reflected in the deconvolution results; on the α -helical content by the Yang method, and on the β -sheet content by the VS method. The sum of the secondary structures for redissolved LHC by the VS method was 142%, indicating a failure of the decomposition, except for the α -helical content, whereas the purified LHC gave a sum of 105%. The decomposition of the CD spectrum of the photosystem I ($+$ — $+$ in Fig. 1) by the VS method resulted in the sum of secondary structures of 295% (Table 2), indicating a failure of the analysis.

Table 1. Decomposition of membrane protein CD spectra by Yang's method (Chang et al., 1978)^a

Proteins ^b	α helix	β sheet	β turn	Random
Reaction center, Rv	42.5	0	27.0	30.5
Photosystem I	66.5	0	12.5	21.0
Reaction center, Rs	55.5	5.0	28.0	11.5
LHC, purified	54.0	0	24.0	22.0
LHC, redissolved	44.5	0	25.5	30.0
Antenna complex	67.0	0	11.0	22.0
UCCR, bovine	68.5	5.0	22.5	4.0
NUR	45.0	50.0	5.0	0
UCCR, Nc	39.0	14.5	22.5	24.0
CCO	63.0	0	23.0	14.0
Cytochrome oxidase	65.0	0	24.0	11.0
Porin, Ec, 1st	50.0	20.0	30.0	0
Porin, Ec, 2nd	44.5	25.0	30.5	0
Porin, Rc	18.5	70.0	2.0	9.5
Maltoporin	55.0	5.0	40.0	0
Phosphoporphin	34.5	50.0	14.5	1.0
Rhodopsin	67.0	0	23.5	9.5
Rhodopsin, la	70.0	5.0	23.0	2.0
Purple membrane	76.5	0	23.5	0
Proton ATPase	93.0	0	6.5	0.5
Ca ²⁺ ATPase, SR, 1st	44.5	18.0	20.0	17.5
Ca ²⁺ ATPase, SR, 2nd	26.0	0	51.0	23.0
H,K ATPase	22.0	40.0	13.0	25.0
F ₀ F ₁ ATPase	63.0	0.5	17.0	19.5
NAR	44.5	0	22.5	33.0
Glucose transporter	70.5	0	29.0	0.5
Band 3	50.0	0	25.5	24.5
Band 3, membrane domain	80.0	0	18.0	2.0
Lac permease, preparation A	44.0	55.0	1.0	0
Lac permease, preparation B	13.5	76.0	0	10.5
Colicin A	79.5	8.5	0	12.0
Colicin E1	74.0	4.0	12.0	10.0
Lipoprotein	3.5	77.5	0	19.0
Prostaglandin synthase	18.0	55.0	6.0	21.0

^a CD spectra were measured as described in the Materials and methods. The numbers represent percentages of residues in the conformers. The values in increments of 0.5% are the results of the algorithm used with a resolution of 0.5%.

^b Rv, *Rhodospseudomonas viridis*; Rs, *Rhodobacter sphaeroides*; LHC, light harvesting complex; UCCR, ubiquinol-cytochrome C reductase; NUR, NADH-ubiquinone reductase; Nc, *Neurospora crassa*; CCO, cytochrome C oxidase; Ec, *Escherichia coli*; Rc, *Rhodobacter capsulatus*; la, light-adapted rhodopsin; SR, sarcoplasmic reticulum; NAR, nicotinic acetylcholine receptor.

The CD spectra of the enzymes of electron transporters showed less diversity than those in the group of the photosynthetic complexes (Fig. 2), except the magnitude of the positive CD band at 192 nm for the ubiquinol-cytochrome C reductase (UCCR) from *Neurospora crassa* ($\diamond-\diamond$ in Fig. 2) was about half the value found for the others, and the ratio of the 208- to the 222-nm bands for NADH-ubiquinone reductase (NUR) ($++$ in Fig. 2) was less than those of the others, and the decomposition methods yielded lower contents of the α -helix and higher percentages of the β -sheet conformation (Tables 1, 2). The sum of the secondary structures from the decompo-

Fig. 1. CD spectra of complexes in photosynthesis system: $\square-\square$, reaction center from *Rhodospseudomonas viridis* (1 mg/mL) in 20 mM sodium phosphate, pH 6.5, 0.1% lauryl dimethylamine-*N*-oxide; $++$, photosystem I (0.12 mg/mL) in 5 mM MES, pH 6.4, 25 mM MgSO₄, 2% dodecyl maltoside; $\diamond-\diamond$, reaction center from *Rhodobacter sphaeroides* (0.7 mg/mL) in 15 mM Tris-HCl, pH 8.0, 1 mM EDTA, 0.025% LDAO; $\triangle-\triangle$, light harvesting complex, purified (0.95 mg/mL) in 0.6% nonylglucoside; $\times-\times$, light harvesting complex, redissolved (1 mg/mL) in 0.6% nonylglucoside; $\nabla-\nabla$, B800-850 antenna complex (0.23 mg/mL) in 1% β -octylglucoside. The CD spectra were measured with sensitivity set at 5×10^{-6} to 1×10^{-5} in cylindrical quartz cells with a pathlength of 0.01 cm at 22 °C, unless otherwise indicated. The response time was set at 5 s, with the data acquisition time at 10 s, and the scan ranges were from 184–190 to 260 nm in steps of 0.2 nm. The protein concentrations are noted in parentheses.

sition results from the VS method was 76% for UCCR from bovine heart, 78% for NUR, 103% for UCCR from *N. crassa*, 132% for cytochrome oxidase, and 143% for cytochrome C oxidase (Table 2). Thus only the results of the UCCR from *N. crassa* were decomposed successfully by the VS method.

Two different batches of porin (gene OmpF product) from *Escherichia coli* yielded similar CD spectra (Fig. 3), yet the VS method yielded a large difference in the β -sheet contents, as well as the sums of the secondary structures. The CD spectra of the porins are typical of β -sheet spectra found in soluble proteins, or more accurately, from model polypeptides (Sarkar & Doty, 1966; Greenfield & Fasman, 1969; Saxena & Wetlaufer, 1971). Not many soluble β proteins showed the "typical" CD patterns (Perczel et al., 1992b), however, in accordance with the speculation that the β -sheet conformations in membrane proteins might be more regularly arrayed than those found in soluble proteins (Weiss et al., 1991), more similar spectra were obtained. Only one exception was found in the CD spectrum of the maltoporin (LamB) with a slight positive peak at 230 nm ($\triangle-\triangle$ in Fig. 3), which re-

Table 2. Decomposition of membrane protein CD spectra by variable selection method (Manavalan & Johnson, 1987; Johnson, 1990)^a

Proteins ^b	α helix	Anti- β ^c	Para- β ^d	Turns	Others	Total
Reaction center, Rv	39	20	7	25	48	139
Photosystem I	83	4	27	57	124	295
Reaction center, Rs	43	5	8	16	35	107
LHC, purified	47	3	3	18	34	105
LHC, redissolved	43	12	6	26	55	142
Antenna complex	65	1	5	9	27	108
UCCR, bovine	47	0	9	5	14	76
NUR	20	19	22	0	17	78
UCCR, Nc	31	13	7	15	36	103
CCO	57	14	8	22	41	143
Cytochrome oxidase	57	9	9	19	39	132
Porin, Ec, 1st	22	0	12	0	13	48
Porin, Ec, 2nd	21	16	19	6	42	105
Porin, Rc	12	15	14	7	33	82
Maltoporin	27	22	21	0	34	105
Phosphoporin	17	17	18	6	32	90
Rhodopsin	53	2	9	8	29	100
Rhodopsin, 1a	48	0	9	0	15	71
Purple membrane	53	1	15	8	41	117
Proton ATPase	96	10	11	39	51	207
Ca ²⁺ ATPase, SR, 1st	33	6	9	15	35	98
Ca ²⁺ ATPase, SR, 2nd	—	—	—	—	—	—
H,K ATPase	25	17	5	22	37	105
F ₀ F ₁ ATPase	54	5	6	11	23	99
NAR	47	37	16	31	77	208
Glucose transporter	46	0	10	8	28	92
Band 3	50	5	0	21	28	104
Band 3, Membrane domain	68	1	7	9	19	104
Lac permease, preparation A	—	—	—	—	—	—
Lac permease, preparation B	—	—	—	—	—	—
Colicin A	71	0	5	0	0	76
Colicin E1	—	—	—	—	—	—
Lipoprotein	5	3	1	3	7	19
Prostaglandin synthase	21	12	5	12	25	75

^a CD spectra were measured as described in the Materials and methods. The numbers represent percentages of residues in the conformers. The values with no decimal point are direct results of a conversion from the fraction values with two digits below the decimal point to the percentage values. The error ranges from the output were omitted for brevity.

^b For abbreviations, see footnote in Table 1.

^c Antiparallel β -pleated sheet.

^d Parallel β -pleated sheet.

sulted in a dramatic decrease in the β -sheet content in the decomposition by the Yang method (Table 1). In comparison, Weiss et al. (1991) reported that, by X-ray crystallography, the porin from *Rb. capsulatus* contained 6% α helix, 57% β strand, and 9% reverse turns, and the content of the β -sheet conformation of the *E. coli* outer membrane porin F (Omp F) has been reported to be $65 \pm 8\%$, determined by high angle diffuse X-ray diffraction and by attenuated total reflection infrared spectroscopy (Kleffel et al., 1985), 54–60% β sheet with less than 10% of α helix by Raman spectroscopy (Vogel & Jähnig, 1986), and the α -helical content of maltoporin was reported to contain about 15% with the β -strand content similar to that of OmpF (Vogel & Jähnig, 1986).

Bacteriorhodopsin from *H. halobium* in the purple membrane (PM) is one of the most widely studied membrane proteins by CD spectroscopy (Brith-Lindner & Rosenheck, 1977; Long et al., 1977; Reynolds & Stoekenius, 1977; Dencher & Heyn, 1978; Mao & Wallace, 1984; Glaeser & Jap, 1985; Paul & Rosenbusch, 1985). The resulting spectrum ($\diamond$ — $\diamond$ in Fig. 4) showed the effect of light scattering by having a distorted minima of the CD bands at 208 and 222 nm, and the magnitude of the spectrum was larger than the values reported as a result of sonification in suspension (Long et al., 1977; Mao & Wallace, 1984) but smaller than that of the detergent-solubilized protein (Reynolds & Stoekenius, 1977). The CD spectra of dark-adapted ($\square$ — $\square$ in Fig. 4) and light-adapted

Fig. 2. CD of electron transporters: $\square$ — $\square$, ubiquinol-cytochrome C reductase from bovine heart (1 mg/mL) in 47.5 mM Tris-HCl, pH 7.8, 0.627 M sucrose, 100 mM NaCl, 5 mM dodecanoyl-*N*-methyl glucamide; $+$ — $+$, NADH-ubiquinone reductase (0.8 mg/mL) in 50 mM NaCl, 50 mM Tris-HCl, pH 7.5, 0.1% Triton X-100, 4.75% sucrose; $\diamond$ — $\diamond$, ubiquinol-cytochrome C reductase from *Neurospora crassa* (0.9 mg/mL) in 40 mM Tris-acetate, pH 7.0, 0.05% Triton X-100, 0.2 mM PMSF; $\triangle$ — $\triangle$, cytochrome C oxidase (1.02 mg/mL) in 10 mM sodium phosphate, pH 7.4, 0.5% Tween 20; $\times$ — $\times$, cytochrome oxidase (0.98 mg/mL) in 50 mM sodium phosphate, pH 7.0, 5% glycerol, 0.5% lauryl maltoside. The CD spectra were measured as described in the Materials and methods and in the legend in Figure 1.

Fig. 3. CD spectra of porins: $\square$ — $\square$, porin from *Escherichia coli*, first batch (0.65 mg/mL) in 0.1 M NaCl, 20 mM sodium phosphate at pH 7.2, 1 mM EDTA, 15 mM NaN_3 , 1% OPOE; $+$ — $+$, porin from *E. coli*, second batch (1 mg/mL) in 0.1 M NaCl, 20 mM sodium phosphate, pH 7.0, 0.2 mM EDTA, 1% OPOE, 0.2 mM DTT, 3 mM sodium azide; $\diamond$ — $\diamond$, porin from *Rhodobacter capsulatus* (1 mg/mL) in 0.3 M LiCl, 20 mM Tris-HCl, pH 7.2, 0.6% (w/v) octyltetraoxyethylene, 3 mM sodium azide; $\triangle$ — $\triangle$, maltoporin (1 mg/mL) in 0.1 M NaCl, 20 mM sodium phosphate, pH 7.0, 0.2 mM EDTA, 1% OPOE, 0.2 mM DTT, 3 mM sodium azide; $\times$ — $\times$, phosphoporin (1 mg/mL) in 50 mM NaCl, 20 mM sodium phosphate, pH 7.6, 0.9% β -octylglucoside. The CD spectra were measured as described in the Materials and methods and in the legend to Figure 1.

($+$ — $+$ in Fig. 4) forms of rhodopsin showed little difference, which was not unexpected according to a systematic study of the two forms in various detergents by Rafferty et al. (1977), where the two forms showed differences in CD spectra when a detergent destabilized the protein. Therefore, if a detergent is sufficiently mild that it does not affect the conformational stability, it was expected that the two forms would not show any difference in the CD spectra. However, it would be difficult, if not impossible, to generalize or predict the effect of detergents on a particular protein (Helenius et al., 1979). The decomposition result for PM was that expected, in that the α -helical content was underestimated by the VS method (53%) but overestimated by the Yang method (76.5%), in comparison with 66% α helix determined by EM (Henderson et al., 1990), with small amounts of β strands (16% by VS method, and 0% by Yang method). There are, however, no compatible data for rhodopsin for the evaluation of the decomposition results or the CD spectrum itself, because the CD spectra reported in Rafferty et al. (1977) were not corrected for molecular weight. The α -helical content in this case is higher than reported, which was about 30% (Rafferty et al., 1977).

The Ca^{2+} ATPase from rabbit muscle sarcoplasmic

Fig. 4. CD spectra of rhodopsins: $\square$ — $\square$, rhodopsin from bovine retina (0.72 mg/mL) in 0.15 M NaCl, 15 mM sodium phosphate, pH 6.6, 0.1% lauryl maltoside; $+$ — $+$, light-adapted rhodopsin (0.72 mg/mL) in 0.15 M NaCl, 15 mM sodium phosphate, pH 6.6, 0.1% lauryl maltoside; $\diamond$ — $\diamond$, purple membrane from *Halobacterium halobium* (0.95 mg/mL) in water. The CD spectra were measured as described in the Materials and methods and in the legend to Figure 1.

reticulum (batch #1) was reported to have $52 \pm 8.1\%$ α helix ($22 \pm 1.6\%$ in transmembrane and $30 \pm 6.5\%$ cytosolic helices), $12 \pm 1.5\%$ β strand, and $36 \pm 6.8\%$ in bends and loops by a secondary structure prediction (Brandl et al., 1986). The Yang method for the spectrum ($++$ in Fig. 5) yielded comparable values with 44.5% α helix, 18.0% β strand, 20.0% turns, and 17.5% random conformations (Table 1). The CD spectrum of batch #2 Ca^{2+} ATPase solution has been truncated at 200 nm because the transparency of the solution decreased to an extent that the signal-to-noise ratio was extremely low below 200 nm, due to the high salt in the solution (0.4 M NaCl), in addition to the distorted shape of the spectrum ($\diamond-\diamond$ in Fig. 5). Therefore, results of the deconvolution would not be reliable, so it was excluded for further processing. The magnitude of the CD spectrum of the proton ATPase from yeast plasma membrane ($\square-\square$ in Fig. 5) approached that of 100% α helix, which is reflected in the decomposition results from both the Yang and the VS methods (93% α helix). However, the sum of secondary structures from the VS method was 207% , indicating a failed decomposition. The CD spectra of H,K ATPase

($\triangle-\triangle$ in Fig. 5) and F_0F_1 ATPase from *E. coli* ($\times-\times$ in Fig. 5) differed only in magnitudes, which, again, was reflected in the contents of α helix (Tables 1, 2).

The 208- and 222-nm bands in the CD spectrum of the nicotinic acetylcholine receptor ($\square-\square$ in Fig. 6) were not well resolved, and the ratio of the 192-nm band to the negative bands was approximately 1, so that the spectrum appeared more like that of a β -sheet conformation. However, the Yang method resulted in 0% in β -sheet conformation, whereas the VS method failed, judging by the sum of the secondary structures of 208% . The CD spectrum of the glucose transporter from human erythrocytes ($++$ in Fig. 6) showed distorted minima with a magnitude of about $\frac{3}{4}$ of that measured in membrane vesicles using a scattered light collecting device (Chin et al., 1987). As a result, α -helical contents from the decomposition results were less than previously reported. However, the CD spectra of two different preparations of Lac permease from *E. coli* ($\times-\times$ and $\nabla-\nabla$ in Fig. 6) were abnormal, probably due to artifacts caused by scattered light from residual lipid vesicles, since the preparations had a low

Fig. 5. CD spectra of ATPases: $\square-\square$, H^+ ATPase from yeast plasma membrane (0.7 mg/mL) in 10 mM Tris-HCl, pH 7.1, 1 mM EDTA, 1 mM DTT, 35% (v/v) glycerol; $++$, Ca^{2+} ATPase from rabbit muscle sarcoplasmic reticulum, first batch (0.5 mg/mL) in 50 mM NaCl, 50 mM CaCl_2 , 10 mM Tris-maleate, pH 6.8, 0.1% dodecyl octaoxyethylene, measured at 5.7°C ; $\diamond-\diamond$, Ca^{2+} ATPase from rabbit muscle sarcoplasmic reticulum, second batch (1 mg/mL) in 0.4 M KCl, 50 mM Tris-HCl, pH 8.0, 10% glycerol, 6% (w/v) PEG 1000, 4 mM sodium azide, 2.5% (w/v) trehalose, 10 mM sodium ascorbate, 0.04% dodecylmaltoside; $\triangle-\triangle$, H,K ATPase (0.96 mg/mL) in 20 mM PIPES, pH 6.1, 0.3 mM CDTA, 2 mM ATP, 0.164% C_{12}E_8 , measured at 6.6°C ; $\times-\times$, F_0F_1 ATPase from *Escherichia coli* (0.62 mg/mL) in 50 mM sodium phosphate, pH 7.0, 5% glycerol, 0.5% laurylmaltoside. The CD spectra were measured as described in the Materials and methods and in the legend to Figure 1.

Fig. 6. CD spectra of receptors and transporters: $\square-\square$, nicotinic acetylcholine receptor (0.49 mg/mL) in 50 mM Tris-HCl, pH 7.0, 1 mM EGTA, 2% octylglucoside; $++$, glucose transporter (0.65 mg/mL) in 12.5 mM NaCl, 0.5 mM KCl, 1 mM Tris-HCl, pH 7.4, 0.38 mM CaCl_2 , 0.25 mM MgCl_2 , 1% octylglucoside; $\diamond-\diamond$, Band 3 from human erythrocytes (1.37 mg/mL) in 5 mM sodium phosphate, pH 8.0, 0.1% C_{12}E_8 , 0.2 mM DTT; $\triangle-\triangle$, Band 3, membrane domain (0.21 mg/mL) in 5 mM sodium phosphate, pH 8.0, 0.1% C_{12}E_8 ; $\times-\times$, Lac permease from *Escherichia coli*, preparation A (18 $\mu\text{g}/\text{mL}$) in 50 mM potassium phosphate buffer at pH 6.0 with 1.25% octylglucoside, 10 mM lactose, 1 mM DTT, 0.25 mg/mL *E. coli*-extracted phospholipids; $\nabla-\nabla$, Lac permease from *E. coli*, preparation B (52 $\mu\text{g}/\text{mL}$) in 50 mM potassium phosphate buffer at pH 6.0 with 1.25% octylglucoside, 10 mM lactose, 1 mM DTT, 0.25 mg/mL *E. coli*-extracted phospholipids. The CD spectra were measured as described in the Materials and methods and in the legend to Figure 1.

protein-to-lipid ratio so that the added detergent might have not been adequate to dissolve all the protein. Consequently, the CD spectra of the Lac permease were excluded from further processing. The CD spectra for the intact erythrocyte and membrane domain Band 3 ($\diamond-\diamond$ for intact Band 3 and $\triangle-\triangle$ for membrane domain Band 3 in Fig. 6) were almost identical to that reported for the magnitudes of the 208- and 222-nm bands (within $\pm 5\%$), except the crossover point was blue shifted to 200 nm in the CD spectrum of the intact Band 3 (Oikawa et al., 1985).

The CD spectrum of the lipoprotein from the *E. coli* outer membrane ($\diamond-\diamond$ in Fig. 7) had a magnitude of about $\frac{1}{10}$ less than previously reported (Lee et al., 1977). The CD analysis reported 88% α helix and 12% random (Lee et al., 1977). Therefore, it was not surprising to find that the decompositions of the CD spectrum yielded either a very low α -helical content (less than 5% by both Yang and VS algorithms) or a failed analysis in the case of the VS method. Similarly, prostaglandin synthase (PGS) yielded a CD spectrum ($\triangle-\triangle$ in Fig. 7) with a low magnitude and, hence, a low α -helical content (18% by the Yang method and 21% by the VS method; Tables 1, 2). The secondary structure prediction for the thermolytic fragment of colicin A was reported as 50.5% cytosolic α helix and 16.2% transmembrane helix, with little or no

β strand (Parker et al., 1989). The CD spectrum of the peptide ($\square-\square$ in Fig. 7) was almost identical to that previously reported (Pattus et al., 1985; Lakey et al., 1991), and the decomposition results of the spectrum resulted in an α -helical content higher than 70%, with less than 10% of β strand, which coincided with the prediction result. On the other hand, colicin E1 from *E. coli* was reported to have 56–57% α helix in a nonionic detergent by a CD spectrum analysis (Brunden et al., 1984), and 58% α helix and 17% β strand from the secondary structure prediction (Yamada et al., 1982). The decomposition of the CD spectrum of the protein ($++$ in Fig. 7), even though similar to that reported, yielded 74% α helix with a small amount in β strand (4%) by the Yang method. Due to the poor transparency of the solution, the CD spectrum was truncated at 190 nm, so that the VS method was not applied to the spectrum.

From the 30 membrane protein CD spectra analyzed by the VS method (Table 2), 15 spectra failed to yield satisfactory results by the criterion suggested by Johnson and his colleagues (Manavalan & Johnson, 1987; Johnson, 1990), where the sum of the secondary structures should fall in the range of 0.9–1.1 for an analysis to be successful. It is, however, noteworthy that a failed decomposition by the VS method does not mean the measured CD spectrum was in error, although an erroneous spectrum may cause the decomposition failure (Manavalan & Johnson, 1987; Johnson, 1990), but rather one or more of the components that compose the spectrum in question might not be included in the five most important basis spectra obtained from the protein database used in the algorithm (Hennessey & Johnson, 1981), so that a linear combination of the basis spectra failed to regenerate the spectrum to be decomposed. When compared with previously reported spectra—the reaction center from *Rps. viridis* by Nabadryk et al. (1985), bacteriorhodopsin from *H. halobium* by Long et al. (1977) and by Mao and Wallace (1984), rhodopsin by Rafferty et al. (1977), glucose transporter by Chin et al. (1987), intact and membrane domain Band 3 by Reithmeier and his colleagues (Oikawa et al., 1985), *E. coli* outer membrane lipoprotein by Lee et al. (1977), the thermolytic fragment of colicin A by Pattus et al. (1985) and Lakey et al. (1991), and colicin E1 by Brunden et al. (1984)—most of the spectra were similar, if not identical, to the spectra previously reported, with the one exception of the lipoprotein. Therefore, it could be deduced that the higher rate of failure in the decomposition of the CD spectra by the VS method might arise from the “bias effect” of the database. Their CD spectral databases are composed of soluble proteins, which may lack the chromophores found only in membrane proteins, so that the deconvolution based on the soluble protein database would interpret the additional secondary structures as artifacts due to the “biased” database. The chlorophylls etc. are examples of the additional chromophores found in membrane proteins.

Fig. 7. CD spectra of miscellaneous membrane proteins: $\square-\square$, colicin A, thermolytic fragment (0.3 mg/mL) in 0.1 M NaF, 20 mM Tris-HCl, pH 6.8, 1% OPOE; $++$, Colicin E1 (0.84 mg/mL) 0.1 M KCl, 10 mM formate, pH 3.5, 1% octylglucoside; $\diamond-\diamond$, lipoprotein (1 mg/mL) in 0.025% SDS, 0.01 mM $MgCl_2$; $\triangle-\triangle$, prostaglandin synthase (1 mg/mL) in 10 mM TES, pH 7.4, 20 mM Na_2SO_4 , 0.05% (w/v) decylheptaoxyethylene ($C_{10}E_7$), 0.05% sodium azide. The CD spectra were measured as described in the Materials and methods and in the legend to Figure 1.

Deconvolution of the membrane protein CD spectra by the convex constraint algorithm

From the total of 34 CD spectra of 30 membrane proteins, 30 CD spectra (Table 2) have been chosen for deconvolution by the CCA method. Most of the CD spectra of the membrane protein database showed characteristics of the α -helical spectra, except for the group of porins. However, the rotational strengths varied rather dramatically from spectrum to spectrum, so that the Yang method, however reliable for CD analysis of soluble proteins, was not appropriate in that its dependence on the intensity of the positive band below 200 nm interfered with the accuracy of the secondary structural analysis (Table 1). For example, when the CD spectrum of the nicotinic acetylcholine receptor was truncated at 200 nm for the lower limit, CD analysis resulted in 45.0% α helix, 27.5% β sheet, 3.5% β turn, and 24.0% random, compared with 44.5, 0, 25.5, and 33.0% for the same spectrum extended to 190 nm. For porin from *Rb. capsulatus*, the values changed from 18.5, 70.0, 2.0, and 9.5% (α helix, β sheet, β turn, and random, respectively) to 26.5, 15.5, 38.0, and 20.0% by the truncation. Such discrepancies, however, are not uncommon among the CD analysis algorithms based on a linear least-squares method, because the minimization of the root mean square errors is usually done at the expense of the resolution of two neighboring bands of the same sign.

By using the volume minimization algorithm, the convex constraint analysis extracts common components as building blocks from the data matrix, composed of the ellipticity values at a given wavelength, to regenerate the original data matrix, while creating a matrix composed of the weight coefficient for each component for the protein (Perczel et al., 1989, 1991b). As shown previously (Perczel et al., 1992b), there are two complementary steps in assigning the secondary structural component spectra and evaluating the deconvolution results: determination of the number of components by utilizing the standard deviation plot against the number of components, followed by assignment of the component spectra to specific secondary structures. Unlike the soluble protein database, the minimal X-ray database of membrane proteins is not sufficient to be useful. Therefore, the assay of the deconvolution result was performed by theoretical and empirical expectation of the consensus spectra for the secondary structures. First, it was imperative to determine a reasonable value for the number of common components. The task was performed by running the program with various values for the number of components. Because a run with more components would naturally have a higher probability of regeneration of the original data matrix more accurately, it is expected that the standard deviation decreases as the number of components increases. The rate of the change, however, decreases exponentially with inflection, as shown in Figure 8, indicating that after the in-

Fig. 8. The dependency of standard deviation, as a result of the convex constraint analysis, on the number of components for the database of the 30 membrane protein CD spectra shown in Table 3.

flection point the decrease in the standard deviation was insignificant and the number of components at the inflection point would represent most of the common components in the protein CD database (Perczel et al., 1991b). In this case, the inflection point was shown at the number of components, $P = 6$.

The deconvolution component spectra, with $P = 6$, yielded two curves that were not correlated with any of the known secondary structures, judging by their wavelengths and the signs of the bands, as well as the overall shapes. They may originate from any combinations of secondary structural components or from artifacts. Such uncharacteristic spectra from the deconvolution may be the cause of distortions of the other components, such as the spectrum correlated with the unordered conformation, with the blue-shifted π - π^* transitions and two maxima at 208 and 225 nm. Furthermore, two other component spectra, although correlated with the β turn and the β sheet, respectively, are also ambiguous in that the bands corresponding to the π - π^* transitions are shifted (Némethy et al., 1967; Manning et al., 1988). Therefore, the overall effect of the deconvolution into six components is that of overestimating the number of the components.

As the deconvolution result with $P = 6$ is an example of overestimation of the number of components, the results with $P = 4$ and $P = 3$ are examples of underestimations. The result with $P = 4$ depicted the class C CD spectrum for the β turn (Woody, 1974) and the unordered conformation (Greenfield & Fasman, 1969; Fulmer, 1979; Yang et al., 1986) more accurately than the spectra for the α -helix and the β -pleated sheet, whose magnitudes were larger and smaller, respectively, than expected. The

deconvolution into $P = 3$ resulted in CD spectra for the α helix and the β -pleated sheet that were almost identical to the consensus ones (Tinoco et al., 1963; Holzwarth & Doty, 1965; Sarkar & Doty, 1966; Woody & Tinoco, 1967; Greenfield & Fasman, 1969; Saxena & Wetlaufer, 1971; Chen et al., 1974) with the rest of the conformers forming a featureless flat line. When comparing the two different spectra for helices in the results with P set at 4 and 3, the spectrum resulting from $P = 4$ had a red-shifted positive band with the ratio of intensities of the two negative bands ($[\theta]_{208}/[\theta]_{222}$) of less than 1, whereas the helix spectrum in $P = 3$ showed the positive band at 192 nm with the ratio of $[\theta]_{208}/[\theta]_{222}$ approaching 1, as expected for an α helix in the cytosolic domain (Greenfield & Fasman, 1969; Saxena & Wetlaufer, 1971; Chen et al., 1974). As in the case of overestimation, the underestimated results showed one or more component spectra correlating well with corresponding secondary structures, whereas the rest of the component spectra were characterized by highly distorted or featureless shapes. Therefore, despite the standard deviation plot (Fig. 8), the characteristics of the component spectra indicated that the number of common components in the database of the 30 membrane protein CD spectra should be five.

However, the component spectra obtained from the deconvolution with $P = 5$ (Fig. 9) yielded new features: the second component spectrum (+—+) was very similar to the spectrum of the α helix, whereas the third component spectrum ($\diamond-\diamond$), helix-like, had the band corre-

sponding to the $\pi-\pi^*$ transition red-shifted from that of an α helix with a larger magnitude, and was, therefore, assigned as the transmembrane helix (α_T helix). This assignment distinguishes it from the α_{II} helix with a lower value as predicted by Manning et al. (1988). This variation may be due to the longer average chain length of the helical strand that spans the thickness of the membrane, approximately 24 residues, which is about twice as long as the average length found in soluble proteins (Chen et al., 1974; Chou & Fasman, 1974). The fourth component spectra ($\triangle-\triangle$), depicted as the β sheet, showed great variability over the database, with a weak band for the $n-\pi^*$ transition, probably because of the poor representation of the β -pleated sheet in the database (5 out of 30 membrane protein CD spectra contained about 50% β -pleated sheet, so that the average content of the conformation per protein was less than 7%). On the other hand, the first and fifth component spectra ($\square-\square$ and $\times-\times$, respectively) represented, rather accurately, the spectrum of the unordered conformation and the class C spectrum of the β turn, respectively, with the theoretical and empirical expectations (Greenfield & Fasman, 1969; Woody, 1974; Fulmer, 1979; Yang et al., 1986). Even with these minor deviations in some of the spectra, the deconvolution results, with the number of components (P) set at 5, succeeded in isolating the two different types of helices along with random and turn conformations. The deviations shown in some of the component spectra (i.e., α helix and α_T helix) may have originated from the composition of the database itself, because the CD spectra of the majority of the membrane proteins reported herein display the characteristics of the α helix with variations in their magnitudes and positions (Figs. 1–7), thereby overwhelming the database.

Because, as discussed above, there are no substantial data to be used as an external probe for the correlation evaluation, a series of deletion analyses was attempted, so that the strongest contributor to a deviation could be isolated. The deletion assay was applied for the purpose of eliminating the "bias" effect of the membrane protein database. Because there is little detailed conformational information on the proteins used herein, there is no way of determining how well the database represents each secondary structural component. Therefore, by deleting a single CD spectrum, one at a time, we attempted to isolate any redundant factor (a CD spectrum whose secondary structure was well represented in the database). For deletion assays to be systematic, every single spectrum was to be deleted, one at a time, from the database, when a single factor was to be removed, and, in turn, every possible combination of two or more spectra, if two or more factors were to be deleted, and so on, therefore making the task impractical. As a compromise, a single spectrum was removed one at a time from the previous database, that is, a spectrum was removed arbitrarily from the 30-protein database, to form a 29-protein data-

Fig. 9. The component CD spectra deconvoluted from the 30 membrane proteins, shown in Table 3, with the number of the components, P , equal to 5 (for detail, see text): curve 1 ($\square-\square$), unordered conformation; curve 2 (+—+), α helix; curve 3 ($\diamond-\diamond$), α_T helix; curve 4 ($\triangle-\triangle$), β -pleated sheet; curve 5 ($\times-\times$), β turn.

base, from which another spectrum was removed to form a 28-protein database, which was repeated up to a point where there were still substantial factors remaining in the database, while the contributors to the deviations could be identified.

Selected results of the deletion assays are shown in Figures 10–12. By deleting the CD spectrum of the *E. coli* outer membrane lipoprotein (Fig. 10), the component CD spectra for the α_T helix (+—+), the β -pleated sheet ($\times$ — $\times$), and the β turn ($\diamond$ — $\diamond$) showed almost no change. The magnitude of the CD spectrum for the α helix (Δ — Δ) decreased a little, and the band for the n - π^* transition decreased disproportionately with respect to the minimum at 208 nm. However, the most noticeable change occurred in the CD spectrum of the random conformation, where the negative band of the CD spectrum was shifted upward to a flattened shape, indicating that the CD spectrum of the lipoprotein was required for stabilization of the component spectrum corresponding to the random conformation. However, by deleting the CD spectra of the membrane domain of Band 3 and the prostaglandin synthase (PGS), the distorted spectrum of the β sheet was persistent, whereas other component spectra showed stabilization (similar to those in Fig. 9), including the spectrum of the random conformation. This confirmed that the CD component spectrum for the random conformation was stabilized by including the CD spectrum of the lipoprotein, whereas the contribution of the membrane domain of Band 3 and the PGS as stabilizing factors seemed minor. The deletion of the lipoprotein in addition

to the deletion of the membrane domain of Band 3 and the PGS (Fig. 11) or an additional deletion of the reaction center from *Rps. viridis* resulted in almost identical component curves, indicating that the contribution of the

Fig. 11. The component CD spectra deconvoluted from 27 membrane proteins (*Escherichia coli* outer membrane, lipoprotein membrane domain Band 3 and prostaglandin synthase excluded) into five components (for details, see text): curve 1 ($\square$ — $\square$), α helix; curve 2 (+—+), α_T helix; curve 3 ($\diamond$ — $\diamond$), β -pleated sheet; curve 4 (Δ — Δ), β turn; curve 5 ($\times$ — $\times$), unordered conformation.

Fig. 10. The component CD spectra deconvoluted from 29 membrane proteins (*Escherichia coli* outer membrane lipoprotein excluded) into five components (for details, see text): curve 1 ($\square$ — $\square$), unordered conformation; curve 2 (+—+), α_T helix; curve 3 ($\diamond$ — $\diamond$), β turn; curve 4 (Δ — Δ), α helix; curve 5 ($\times$ — $\times$), β -pleated sheet.

Fig. 12. The component CD spectra deconvoluted from 24 membrane proteins (*Escherichia coli* outer membrane lipoprotein, membrane domain Band 3, prostaglandin synthase, reaction center from *Rhodospseudomonas viridis*, second batch of porin from *E. coli* and purple membrane excluded) into five components (for details, see text): curve 1 ($\square$ — $\square$), α helix; curve 2 (+—+), α_T helix; curve 3 ($\diamond$ — $\diamond$), β -pleated sheet; curve 4 (Δ — Δ), β turn; curve 5 ($\times$ — $\times$), unordered conformation.

CD spectrum of the reaction center was minimal, whereas deletion of the lipoprotein yielded a better resolution for the β strand, which was unusual, considering that the deletion of the spectrum from the 30-protein database (Fig. 10) showed a destabilizing effect on the random conformation. However, further deletions, such as the redissolved light harvesting complex, the purple membrane, the light-adapted rhodopsin, and/or the second batch of porin from *E. coli* (Fig. 12) proved more destabilizing than stabilizing. Furthermore, deleting the B800–850 antenna–pigment complex (∇ – ∇ in Fig. 1) or the photosystem I ($+$ – $+$ in Fig. 1), because they are complexes and not well defined, containing several prosthetic groups that may give CD spectral contributions, destabilized all other components except the α_T helix (data not shown). The deletion assay further proved that the two types of helices were persistent, despite the wide range of magnitudes, whereas the component spectrum of the β -pleated sheet was so unstable that, in some cases, the bands for

the n – π^* transition were shown as troughs between two positive peaks rather than negative bands. This was not surprising considering the low content of the β -sheet conformation in the database. This problem will be persistent in the deconvolution algorithms for the CD spectra of membrane proteins because of the lack of membrane proteins that contain high percentages of the β -sheet conformation.

The weight coefficients for the deconvoluted components using 30 membrane protein CD spectra, with the number of components set to five, with the component spectra represented in Figure 9, are summarized in Table 3. The compositions of the reaction centers from *Rps. viridis* and from *Rb. sphaeroides*, obtained from the deconvolution, were: the complex from *Rps. viridis*, 8.0% α helix, 21.6% α_T helix, 8.9% β sheet, 10.0% β turn, and 51.5% random; and the complex from *Rb. sphaeroides*, 19.6% α helix, 18.9% α_T helix, 24.2% β sheet, 7.7% β turn, and 29.6% random. Analysis of the X-ray

Table 3. Deconvolution of membrane protein CD spectra by convex constraint analysis methods (Perczel et al., 1989, 1991b) of 30 membrane protein CD spectra into five components (Fig. 9)^a

Proteins ^b	First (random)	Second (α helix)	Third (α_T helix)	Fourth (β sheet)	Fifth (β turn)
Reaction center, Rv	52	8	22	9	10
Photosystem I	47	0	51	1	0
Reaction center, Rs	30	20	19	24	8
LHC, purified	39	26	18	14	3
LHC, redissolved	53	12	24	9	2
Antenna complex	30	53	12	2	4
UCCR, bovine	9	36	10	22	23
NUR	0	25	0	27	49
UCCR, Nc	45	20	13	16	6
CCO	31	20	24	13	12
Cytochrome oxidase	27	23	22	16	12
Porin, Ec, 1st	11	36	3	47	3
Porin, Ec, 2nd	14	2	15	45	24
Porin, Rc	35	23	4	30	9
Maltoporin	1	0	15	54	30
Phosphoporin	19	15	8	36	22
Rhodopsin	22	42	14	21	2
Rhodopsin, 1a	10	55	5	28	2
Purple membrane	0	40	19	40	0
Proton ATPase	0	0	44	1	56
Ca ²⁺ ATPase, SR, 1st	34	23	13	20	11
H,K ATPase	53	3	17	15	12
F ₀ F ₁ ATPase	28	41	12	8	12
NAR	52	0	27	0	20
Glucose transporter	11	36	14	32	7
Band 3	48	18	22	10	2
Band 3, membrane domain	13	53	16	18	1
Colicin A	7	78	0	0	14
Lipoprotein	51	25	0	24	0
Prostaglandin synthase	49	19	8	18	6

^a The CD spectra were measured as described in the Materials and methods. The numbers represent percentages of residues in the conformers.

^b For abbreviations, see footnote in Table 1.

data from the reaction centers from *Rps. viridis* gave 16% α helix, 23% α_T helix, and 7% β sheet (Deisenhofer et al., 1985; Deisenhofer & Michel, 1989), and the X-ray data for the reaction center from *Rb. sphaeroides* yielded 17% α helix, 34% α_T helix, and 15% β sheet (Allen et al., 1987; Feher et al., 1989).

The Pearson correlation coefficient for the reaction center of *Rps. viridis* was 0.98, whereas the correlation coefficient for the reaction center of *Rb. sphaeroides* was 0.46 (Table 4). The porin from *Rb. capsulatus*, with the X-ray-determined composition of 6% α helix, 0% α_T helix, 57% β sheet, and 9% reverse turn (Weiss et al., 1991), also showed a reasonable correlation (Pearson correlation coefficient, R , was 0.72; Table 4) with the deconvoluted composition (22.7% α helix, 4.4% α_T helix, 29.7% β sheet, and 8.5% reverse turn). The relatively high correlation (between the deconvoluted composition of porin and the X-ray data) was unexpected because the database had the β -sheet conformer poorly represented and the β sheet was the predominant conformation in the porin. However, perhaps there is a " β_T -sheet" (transmembrane β -sheet) spectrum, which is still to be detected. There remains uncertainty in assigning the α_T helix in the case of Band 3. It was expected that the α_T helical content of the membrane domain Band 3 would be higher than that of the intact Band 3. Instead, most of the helix in the membrane domain Band 3 was deconvoluted to the cytosolic α helix (52.9%), while the α_T helix was decreased to 15.6%, in comparison with $\alpha_T = 22.4\%$ for the intact Band 3 (Table 3). The discrepancy might indi-

cate that the transmembrane helix of the membrane domain Band 3 might not be as stable as that of the intact protein in the detergent-solubilized forms. Another possibility might be that the protease digestion of the detergent-solubilized Band 3 might have produced another fragment of the protein that contained some of the previous α_T helix and the majority of the cytosolic domain, due to the exposure of a new digestion site for the enzyme by the detergent solubilization. The higher percentage of the overall helical content of the membrane domain Band 3 (68.5%), compared to 40.3% in the intact Band 3, may be due to removal of conformational restraints that allow the detergent to increase the helical content. Additional evidence, such as X-ray crystallography or infrared spectroscopy, would be needed to establish the stability of the α_T helix in the membrane domain Band 3. It was, however, obvious that with only three membrane proteins with known secondary structure, evaluation of the results is still tenuous.

The most interesting result is that two types of helices were isolated from the set of CD spectra of membrane proteins, together with the spectra for the β -turn and the β -sheet conformations. The spectrum for the β turn was almost exclusively that of class C according to the definition of Woody (1974). Supportive data are required to confirm the idea that the unique environment of the membrane bilayer coerces the reverse turns to adapt a single conformation, rather than a mixture of conformers as was found for soluble proteins (Woody, 1974). As for the β -sheet CD spectrum, more membrane proteins would be

Table 4. Comparison of the secondary structures determined by X-ray crystallography, the convex constraint algorithm (CCA), the Yang method, and the variable selection (VS) method

Proteins	Methods	α helix	α_T helix	β sheet	β turn	Random	R^a
Reaction center from <i>Rhodopseudomonas viridis</i>	X-ray ^b	16	23	7	—	54	0.98
	CCA ^c	8	22	9	10	52	
	Yang ^d	43	—	0	27	31	
	VS ^e	39	—	27	25	48	
Reaction center from <i>Rhodopseudomonas sphaeroides</i>	X-ray	17	34	15	—	34	0.46
	CCA	20	19	24	8	30	
	Yang	56	—	5	28	12	
	VS	43	—	13	16	35	
Porin from <i>Rhodopseudomonas capsulatus</i> X-ray		6	0	57	9	28	0.72
	CCA	23	4	30	9	35	
	Yang	19	—	70	2	10	
	VS	12	—	29	7	33	

^a Pearson product-moment correlation coefficient.

^b The criteria for secondary structural assignments were not clear. The β turns were not classified in the references (reaction center from *Rps. viridis* [Deisenhofer et al., 1985; Deisenhofer & Michel, 1989], reaction center from *Rps. sphaeroides* [Allen et al., 1987], and porin from *Rps. capsulatus* [Weiss et al., 1991]) so that it has been included in the random conformation except for the porin.

^c The convex constraint analysis (Perczel et al., 1991b). For the Pearson coefficient calculation, the turn was added to the random conformation.

^d The CD spectral analysis program developed by Yang and his colleagues (Chang et al., 1978). Because the transmembrane helix was not segregated by the algorithm, it was included in the α -helix percentage.

^e The variable selection method (Manavalan & Johnson, 1981). Because the transmembrane helix was not segregated by the algorithm, it was included in the α -helix percentage. Also the antiparallel and the parallel β sheets were combined.

needed for better resolution of this conformer by CD spectroscopy. Due to the lack of any means of evaluating the CD spectra or the deconvolution results, the discussion remains open for revision as more detailed studies on the structure of membrane proteins progresses to provide an objective means for evaluation.

Conclusions

The object of the studies reported herein was to obtain the component CD spectra corresponding to the secondary structural conformations for soluble as well as membrane proteins. Even though there are several algorithms available for analyzing CD spectra (Saxena & Wetlaufer, 1971; Chang et al., 1978; Brahms & Brahms, 1980; Bolutina et al., 1981; Hennessey & Johnson, 1981; Provencher & Glöckner, 1981; Manavalan & Johnson, 1987), they require the known secondary structures of the proteins in the database (usually determined by X-ray diffraction) and, therefore, compound the errors due to the interpretation of the X-ray data. As yet there is no database whose secondary structures are known for membrane proteins. The convex constraint analysis algorithm, on the other hand, operates only on a collection of spectral data to isolate the common components, linear combinations of which could, in turn, rebuild the original data set. The straightforward and objective way to correlate the deconvoluted component curves with the secondary structures (external evaluation), as shown previously for the soluble proteins (Perczel et al., 1992b), is to utilize external data, such as X-ray crystallography. A more intuitive, and therefore less accurate, application was performed here due to the lack of external data. The most likely number of pure components was determined by the standard deviation plot as a function of the number of components (Fig. 8). With knowledge of the consensus CD spectra for the secondary structures, these components of the deconvoluted component spectra could be assigned to the respective secondary structures. However, the value of the number of secondary structural components for scores of protein is somewhat vague, because there are ambiguities in defining the secondary structures. Such ambiguities, added to the intrinsic uncertainties of the CD spectrum, warrant extreme caution in interpreting the deconvolution results. Despite these shortcomings, which are inherent in all such algorithms, the CCA method has been proven to be successful in isolating component spectra comparable to theoretical expectations (Perczel et al., 1991a, 1992a; Perczel & Fasman, 1992). In addition, the CCA method could be used to deconvolute and obtain the secondary structural component spectra characteristic of membrane proteins, without the secondary structures being defined by other methods, such as X-ray crystallography.

For the membrane protein data set, the five component spectra obtained from the deconvolution consisted of *two*

different types of α helices (the α helix in the soluble domain and the α_T helix, for transmembrane α helix), the β -pleated sheet, the class C spectrum related to the β turn (Woody, 1974), and the unordered conformation. The isolated CD spectrum for the α_T helix was characterized by its positive band red-shifted to the range of 195–200 nm ($95,000 \text{ deg cm}^2 \text{ dmol}^{-1}$), with the intensity of the 208-nm band ($-50,000 \text{ deg cm}^2 \text{ dmol}^{-1}$) slightly less negative than that of the 222-nm band ($-60,000 \text{ deg cm}^2 \text{ dmol}^{-1}$). Unlike the prediction of the α_{II} helix by Manning et al. (1988), however, the rotational strength of the deconvoluted CD spectrum for the α_T helix was larger than the intensity of the helix in the soluble domain. Such an increase in rotational strength could be attributed to the fact that membrane proteins are immersed in a much lower dielectric medium than are soluble proteins, which could affect the $\pi-\pi^*$ and $n-\pi^*$ transitions. In addition, the average chain length of the membrane-spanning α helix (α_T helix) would be between 22 and 28 residues per helical strand, which is about twice the average length of the α helix in soluble proteins, and, according to Chen et al. (1974) and Chou and Fasman (1974), the intensities of CD bands for the helix are dependent on the average number of residues per strand. Therefore, the CD band intensities of the helix in the membrane domain, with its average number of residues of about 24, could be expected to be larger than that of the soluble protein helix. The other component CD spectra deconvoluted from the membrane proteins correlated well with the soluble α -helix, the β -turn, and the random conformations. Such similarity was expected because those secondary structures in membrane proteins are believed to be exposed above and below the membrane layer.

The deconvoluted CD spectra for secondary structures for membrane proteins contained almost all the characteristics expected for these conformations despite the difference in approach for each set of proteins. However, due to the bias effect from the database, some component CD spectra, correlated with conformations that are not well represented in the database, were not well resolved as well as the others. Overcoming the bias effect has been a major task in the deconvolution algorithms and generally was performed by selecting a database representing a wide range of secondary structures. However, this was not always possible due to the rare occurrence of some conformers. The deletion assay, suggested herein, is intended to be used as a stability test for a particular component by deleting a set of spectra from the database, so that the selection of the spectra for deletion was random. More systematic application of the deletion assay may be used as a way to overcome the bias effect.

Materials and methods

The lists of membrane proteins are shown in Tables 1–3, and the experimental procedures regarding the prepa-

ration of the protein samples appear on the August 1992 *Protein Science* Diskette Appendix. The CD spectra of the proteins used herein also appear on the Diskette Appendix. See the printed Appendix at the end of this paper for a listing of the proteins used and their origins.

Purification of cytoplasmic and membrane domains of Band 3

The procedure was a modification of that of Appell and Low (1981) and Oikawa et al. (1985), where the detergent-solubilized Band 3, in place of the membrane-embedded protein, was used for digestion with TPCK-treated α -chymotrypsin or trypsin. The details will be reported elsewhere.

CD measurements

All the CD spectra were measured using a Jovin-Yvon Mark V Autodichrograph at 22 °C in steps of 0.2 nm, unless otherwise indicated, and calibrated with the ellipticity value of a standard solution of *d*-10-camphorsulfonic acid, measured immediately before each spectrum was measured. Each spectrum was the average of five repeated scans. For the membrane proteins, the sensitivity was set at either 5×10^{-6} or 1×10^{-5} , the pathlength 0.01 cm, the response time at 5 s, and the data acquisition time at 10 s in the range of 184–260 nm.

Deconvolutions of the CD spectra

All the CD spectra were analyzed by several algorithms; the Yang method (Chang et al., 1978), the variable selection (VS) method (Manavalan & Johnson, 1987; Johnson, 1990), and the new algorithm, the convex constraint analysis (CCA) (Perczel et al., 1989, 1991b), using an IBM-compatible personal computer with an 80286 microprocessor (Packard Bell) and an 80287 numerical coprocessor (Intel).

The VS method uses two factors (in addition to a mathematical constraint, the orthogonality of the basis spectra)—one, that a protein has positive values in weight coefficient for secondary structures, and the other, that the sum of secondary structure should be 100%—as an evaluating tool for reliability of the deconvolution result (Manavalan & Johnson, 1987), but not as constraints in the algorithm reporting that the unconstrained analysis improved prediction for α -helical content (Johnson, 1990). The Yang method (Chang et al., 1978) uses two constraints. The first constraint is that the sum of the secondary structure should be 100%. The second constraint defines the percentage of each component (weight coefficient) to be positive. The CCA utilizes three constraints. The sum of the secondary structure should be 100%, whereas each component would have a positive value

(weight coefficient). In addition, each coefficient value for a given protein should be placed in a simplex of the *P*-dimensional Euclidean space with the smallest volume (volume minimization), where the value, *P*, is the number of the common components. The CCA method has been put to extensive tests for its utilization and efficiency (Perczel et al., 1992c), even though direct comparison with other methods for the accuracy is not suitable, because each algorithm has its own advantages over others for different purposes.

Only a few membrane proteins have had their secondary structure determined by X-ray crystallography, so that the analysis of the data set of the membrane protein CD spectra was performed differently than for the globular proteins described previously (Perczel et al., 1992b). The secondary structural information was not necessary in applying the algorithm to the set, but the lack of it deprived one of a way to assess the reliability of the result. The basic idea of the approach was from the variable selection method, in that a set (a subset) of CD spectra was deleted systematically from the whole set. The difference in the approach from that of the VS method was that it was to determine how well each deconvoluted conformer was represented in the individual CD spectrum in the database, rather than eliminating a set of CD spectra from the database, containing contributions that are not found in the CD spectrum to be decomposed.

Acknowledgments

This research was generously supported by grants from the National Science Foundation (DMB-9413337) and the U.S. Army Laboratory Command (DAAL03-89-K-0088). We express our gratitude to the many colleagues who generously contributed precious samples of their membrane proteins, which made this study possible.

References

- Allen, J.P., Feher, G., Yeates, T.O., Komiya, H., & Rees, D.C. (1987). Structure of the reaction center from *Rhodobacter sphaeroides* R-26: The protein subunits. *Proc. Natl. Acad. Sci. USA* 84, 6162–6166.
- Allen, J.P., Feher, G., Yeates, T.O., Rees, D.C., Deisenhofer, J., Michel, H., & Huber, R. (1986). Structural homology of reaction centers from *Rhodospseudomonas sphaeroides* and *Rhodospseudomonas viridis* as determined by X-ray diffraction. *Proc. Natl. Acad. Sci. USA* 83, 8589–8593.
- Appell, K.C. & Low, P.S. (1981). Partial structural characterization of cytoplasmic domain of the erythrocyte membrane protein, Band 3. *J. Biol. Chem.* 256, 11104–11111.
- Bolotina, I.A., Chekhov, V.O., Lugauskas, V.Y., & Ptitsyn, O.B. (1981). Determination of the secondary structure of proteins from circular dichroism spectra. III. Protein derived reference spectra for antiparallel and parallel β structures. *Mol. Biol.* 15, 130–137. [Translated from *Molekulyarnaya Biologiya (USSR)* 15, 167–175 (1980).]
- Bovey, F.A. (1988). NMR of solids. In *Nuclear Magnetic Resonance Spectroscopy*, 2nd Ed. (Bovey, F.A., Jelinsky, L., & Mirau, P.A., Eds.), pp. 399–436. Academic Press, New York.
- Brahms, S. & Brahms, J. (1980). Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism. *J. Mol. Biol.* 138, 149–178.

- Brandl, C.J., deLeon, S., Martin, D.R., & MacLennan, D.H. (1987). Adult forms of the Ca^{2+} ATPase of sarcoplasmic reticulum. *J. Biol. Chem.* 262, 3768-3774.
- Brandl, C.J., Green, N.M., Korczak, B., & MacLennan, D.H. (1986). Two Ca^{2+} ATPase genes: Homologies and mechanistic implications of deduced amino acid sequences. *Cell* 44, 597-607.
- Brith-Lindner, M. & Rosenheck, K. (1977). The circular dichroism of bacteriorhodopsin: Asymmetry and light scattering distortions. *FEBS Lett.* 76, 41-44.
- Brunden, K.R., Uratani, Y., & Cramer, W.A. (1984). Dependence of the conformation of a colicin E1 channel-forming peptide on acidic pH and solvent polarity. *J. Biol. Chem.* 259, 7682-7687.
- Capaldi, R.A. (1990). Structure and function of cytochrome C oxidase. *Annu. Rev. Biochem.* 59, 569-596.
- Casey, J.R., Lieberman, D.M., & Reithmeier, R.A.F. (1989). Purification and characterization of Band 3 protein. *Methods Enzymol.* 173, 494-512.
- Casey, J.R. & Reithmeier, R.A.F. (1991). Analysis of the oligomeric state of Band 3, the anion transport protein of the human erythrocyte membrane, by size exclusion high performance liquid chromatography. *J. Biol. Chem.* 266, 15726-15737.
- Chang, C.-H., Tiede, D., Tang, J., Smith, U., Norris, J., & Schiffer, M. (1986). Structure of *Rhodospseudomonas sphaeroides* R-26 reaction center. *FEBS Lett.* 205, 82-86.
- Chang, C.T., Wu, C.-S.C., & Yang, J.T. (1978). Circular dichroism analysis of protein conformation: Inclusion of the β -turns. *Anal. Biochem.* 91, 13-31.
- Chen, Y.-H., Yang, J.T., & Chau, K.H. (1974). Determination of the helix and β -form of proteins in aqueous solution by circular dichroism. *Biochemistry* 13, 3350-3359.
- Chin, J.J., Jung, E.K.Y., Chen, V., & Jung, C.Y. (1987). Structural basis of human erythrocyte glucose transporter function in proteoliposome vesicles: Circular dichroism measurements. *Proc. Natl. Acad. Sci. USA* 84, 4113-4116.
- Chou, P.Y. & Fasman, G.D. (1974). Prediction of protein conformation. *Biochemistry* 13, 222-245.
- Chou, P.Y. & Fasman, G.D. (1977). β -Turns in proteins. *J. Mol. Biol.* 115, 135-175.
- Cleveland, M.B., Slatin, S., Finkelstein, A., & Levinthal, C. (1983). Structure-function relationships for a voltage-dependent ion channel: Properties of COOH terminal fragments of colicin E1. *Proc. Natl. Acad. Sci. USA* 80, 3706-3710.
- Cogdell, R.J. & Scheer, H. (1985). Circular dichroism of light harvesting complexes from purple photosynthetic bacteria. *Photochem. Photobiol.* 42, 669-678.
- Cogdell, R.J., Woolley, K.J., Ferguson, L.A., & Dawkins, D.J. (1990). Crystallization of purple bacterial antenna complexes. In *Crystallization of Membrane Proteins* (Michel, H., Ed.), pp. 125-136. CRC Press, Boca Raton, Florida.
- Cogdell, R.J., Woolley, K., McKenzie, R.C., Lindsay, J.G., Michel, H., Dobler, J., & Zinth, W. (1985). Crystallization of the B800-850 complex from *Rhodospseudomonas acidophila* strain 7750. In *Antennas and Reaction Centers of Photosynthetic Bacteria* (Michel-Beyerle, M.E., Ed.), pp. 85-87. Springer-Verlag, New York.
- Cross, R.L. (1988). The number of functional catalytic sites on F_1 -ATPases and the effects of quaternary structural asymmetry on their properties. *J. Bioenerg. Biomembr.* 20, 395-405.
- Dankert, J.R., Uratani, Y., Grabau, C., Cramer, W.A., & Hermodson, M. (1982). On a domain structure of colicin E1. *J. Biol. Chem.* 257, 3857-3863.
- Davidson, V.L., Brunden, K.R., Cramer, W.A., & Cohen, F.S. (1984). Studies on the mechanism of action of channel-forming colicins using artificial membranes. *J. Membr. Biol.* 79, 105-118.
- Deisenhofer, J., Epp, O., Miki, K., Huber, R., & Michel, H. (1985). Structure of the protein subunits in the photosynthetic reaction center of *Rhodospseudomonas viridis* at 3 Å resolution. *Nature* 318, 618-624.
- Deisenhofer, J., Epp, O., Miki, K., Huber, R., & Michel, H. (1984). X-ray structure analysis of a membrane protein complex. Electron density map at 3 Å resolution and a model of the chromophores of the photosynthetic reaction center from *Rhodospseudomonas viridis*. *J. Mol. Biol.* 180, 385-398.
- Deisenhofer, J. & Michel, H. (1989). The photosynthetic reaction center from the purple bacterium *Rhodospseudomonas viridis*. *Science* 245, 1463-1473.
- Dencher, N.A. & Heyn, M.P. (1978). Formation and properties of bacteriorhodopsin monomers in the non-ionic detergents octyl- β -D-glucoside and Triton X-100. *FEBS Lett.* 96, 322-326.
- Duysens, L.N.M. (1956). The flattening of the absorption spectrum of suspensions, as compared to that of solutions. *Biochim. Biophys. Acta* 19, 1-12.
- Falson, P., Di Pietro, A., Jault, J.-M., & Gautheron, D.C. (1986). Chemical modification of thiol group of mitochondrial F_1 -ATPase from the yeast *Schizosaccharomyces pombe*. *J. Biol. Chem.* 261, 7151-7159.
- Falson, P., Di Pietro, A., Jault, J.-M., Gautheron, D.C., & Boutry, M. (1989). Purification from a yeast mutant of mitochondrial F_1 with modified β -subunit. High affinity for nucleotides and high negative cooperativity of ATPase activity. *Biochim. Biophys. Acta* 975, 119-126.
- Feher, G., Allen, J.P., Okamura, M.Y., & Rees, D.C. (1989). Structure and function of bacterial photosynthetic reaction centers. *Nature* 339, 111-116.
- Foster, D.L. & Fillingame, R.H. (1979). Energy-transducing H^+ -ATPase of *Escherichia coli*. *J. Biol. Chem.* 254, 8230-8236.
- Fulmer, A.W. (1979). Studies on chromatin reconstitution. Ph.D. Thesis, Brandeis University, Waltham, Massachusetts.
- Garavito, R.M. & Rosenbusch, J.P. (1986). Isolation and crystallization of bacterial porin. *Methods Enzymol.* 125, 309-328.
- Glaeser, R.M. & Jap, B.K. (1985). Absorption flattening in the circular dichroism spectra of small membrane fragments. *Biochemistry* 24, 6398-6401.
- Gordon, D.J. & Holzwarth, G. (1971). Artifacts in the measured optical activity of membrane suspensions. *Arch. Biochem. Biophys.* 142, 481-488.
- Greenfield, N. & Fasman, G.D. (1969). Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry* 8, 4108-4116.
- Harbison, G.S., Smith, S.O., Pardo, J.A., Courtin, J.M.L., Lugtenburg, J., Herzfeld, J., Mathies, R.A., & Griffin, R.G. (1985). Solid-state ^{13}C NMR detection of a perturbed 6-s-trans chromophore in bacteriorhodopsin. *Biochemistry* 24, 6955-6962.
- Helenius, A., McCaslin, D.R., Fries, E., & Tanford, C. (1979). Properties of detergents. *Methods Enzymol.* 56, 734-749.
- Henderson, R., Baldwin, J.M., Ceska, T.A., Zemlin, F., Beckmann, E., & Downing, K.H. (1990). Model for the structure of bacteriorhodopsin based on high-resolution electron cryomicroscopy. *J. Mol. Biol.* 213, 899-929.
- Hennessey, J.P., Jr. & Johnson, W.C., Jr. (1981). Information content in the circular dichroism of proteins. *Biochemistry* 20, 1085-1094.
- Hollósi, M., Kövér, K.E., Holly, S., & Fasman, G.D. (1987a). β -Turns in serine-containing linear and cyclic peptide models. *Biopolymers* 26, 1527-1533.
- Hollósi, M., Kövér, K.E., Holly, S., Radics, L., & Fasman, G.D. (1987b). β -Turns in bridged proline-containing cyclic peptide models. *Biopolymers* 26, 1555-1572.
- Holzwarth, G. & Doty, P. (1965). The ultraviolet circular dichroism of polypeptides. *J. Am. Chem. Soc.* 87, 218-228.
- Johnson, W.C., Jr. (1990). Protein secondary structure and circular dichroism: A practical guide. *Proteins* 7, 205-214.
- Kleffel, B., Garavito, R.M., Baumeister, W., & Rosenbusch, J.P. (1985). Secondary structure of a channel-forming protein: Porin from *E. coli* outer membranes. *EMBO J.* 4, 1589-1592.
- Kreusch, A., Weiss, M.S., Welte, W., Weckesser, J., & Schulz, G.E. (1991). Crystals of an integral membrane protein diffracting to 1.8 Å resolution. *J. Mol. Biol.* 217, 9-10.
- Kühlbrandt, W. (1988). Three-dimensional crystallization of membrane proteins. *Q. Rev. Biophys.* 21(4), 429-477.
- Lakey, J.H., Massotte, D., Heitz, F., Dasseux, J.-L., Faucon, J.-F., Parker, M.W., & Pattus, F. (1991). Membrane insertion of the pore-forming domain of colicin A: A spectroscopic study. *Eur. J. Biochem.* 196, 599-607.
- Lee, N., Cheng, E., & Inouye, M. (1977). Optical properties of an outer membrane lipoprotein from *Escherichia coli*. *Biochim. Biophys. Acta* 465, 650-656.
- Long, M.M., Urry, D.W., & Stoeckenius, W. (1977). Circular dichro-

- ism of biological membranes: Purple membrane of *Halobacterium halobium*. *Biochem. Biophys. Res. Commun.* 75, 725-731.
- MacLennan, D.H., Brandl, C.J., Korczak, B., & Green, N.M. (1985). Amino-acid sequence of a Ca^{2+} - Mg^{2+} -dependent ATPase from rabbit muscle sarcoplasmic reticulum deduced from its complementary DNA sequence. *Nature* 316, 696-700.
- Madden, T.D., Chapman, D., & Quinn, P.J. (1979). Cholesterol modulates activity of calcium dependent-ATPase of the sarcoplasmic reticulum. *Nature* 279, 538-541.
- Manavalan, P. & Johnson, W.C., Jr. (1987). Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. *Anal. Biochem.* 167, 76-85.
- Manning, M.C., Illangasekare, M., & Woody, R.W. (1988). Circular dichroism studies of distorted α -helices, twisted β -sheets, and β -turns. *Biophys. Chem.* 31, 77-86.
- Mao, D. & Wallace, B.A. (1984). Differential light scattering and absorption flattening optical effects are minimal in the circular dichroism spectra of small unilamellar vesicles. *Biochemistry* 23, 2667-2673.
- McRee, D.E., Tainer, J.A., Meyer, T.E., van Beeumen, J., Cusanovich, M., & Getzoff, E.D. (1989). Crystallographic structure of a photoreceptor protein at 2.4 Å resolution. *Proc. Natl. Acad. Sci. USA* 86, 6533-6537.
- Mendel-Hartvig, J. & Capaldi, R.A. (1991). Catalytic site nucleotide and inorganic phosphate dependence of the conformation of the ϵ subunit in *Escherichia coli* adenosinetriphosphatase. *Biochemistry* 30, 1278-1284.
- Molvenger, L.C.P.J., Kentgens, A.P.M., Pardo, J.A., Courtin, J.M.L., Veeman, W.S., Lugtenburg, J., & de Grip, W.J. (1987). High-resolution solid-state ^{13}C -NMR study of carbons C-5 and C-12 of the chromophore of bovine rhodopsin. *Eur. J. Biochem.* 163, 9-14.
- Nabedryk, E., Berger, G., Andrianambinintsoa, S., & Breton, J. (1985). Comparison of α helix orientation in chromatophore, quantasome and reaction center of *Rhodospseudomonas viridis* by circular dichroism and polarized infrared spectroscopy. *Biochim. Biophys. Acta* 809, 271-276.
- Nabedryk, E., Tiede, D.M., Dutton, P.L., & Breton, J. (1982). Conformation and orientation of the protein in the bacterial photosynthetic reaction center. *Biochim. Biophys. Acta* 682, 273-280.
- Némethy, G., Phillips, D.C., Leach, S.J., & Scheraga, H.A. (1967). A second right-handed helical structure with the parameters of the Pauling-Corey α helix. *Nature* 214, 363-365.
- Ohno-Iwashita, Y. & Imahori, K. (1982). Assignment of the functional loci in the colicin E1 molecule by characterization of its proteolytic fragments. *J. Biol. Chem.* 257, 6446-6451.
- Oikawa, K., Lieberman, D.M., & Reithmeier, R.A.F. (1985). Conformation and stability of the anion transport protein of human erythrocyte membranes. *Biochemistry* 24, 2843-2848.
- Pancoska, P. & Keiderling, T.A. (1991). Systematic comparison of statistical analyses of electronic and vibrational circular dichroism for secondary structure prediction of selected proteins. *Biochemistry* 30, 6885-6895.
- Papiz, M.Z., Hawthornthwaite, A.M., Cogdell, R.J., Woolley, K.J., Wightrian, P.A., Ferguson, L.A., & Lindsay, J.G. (1989). Crystallization and characterization of two crystal forms of the B800-850 light harvesting complex from *Rhodospseudomonas acidophila* strain 10050. *J. Mol. Biol.* 209, 833-835.
- Parker, M.W., Pattus, F., Tucker, A.D., & Tsernoglou, D. (1989). Structure of the membrane-pore-forming fragment of colicin A. *Nature* 337, 93-96.
- Pattus, F., Heitz, F., Martinez, C., Provencher, S.W., & Lazdunski, C. (1985). Secondary structure of the pore-forming colicin A and its C-terminal fragment. Experimental fact and structure prediction. *Eur. J. Biochem.* 152, 681-689.
- Paul, C. & Rosenbusch, J.P. (1985). Folding patterns of porin and bacteriorhodopsin. *EMBO J.* 4, 1593-1597.
- Penin, F., Codinot, C., & Gautheron, D.C. (1979). Optimization of the purification of mitochondrial F_1 adenosine triphosphatase. *Biochim. Biophys. Acta* 548, 63-71.
- Perczel, A. & Fasman, G.D. (1992). Quantitative analysis of cyclic β -turn models. The effect of ring stress on β -turn geometries. *Protein Sci.* 1, 378-395.
- Perczel, A., Hollósi, M., Foxman, B.M., & Fasman, G.D. (1991a). Conformational analysis of pseudo cyclic hexapeptides based on quantitative circular dichroism (CD), NOE, and X-ray data. *J. Am. Chem. Soc.* 113, 9772-9784.
- Perczel, A., Hollósi, M., Sandor, P., & Fasman, G.D. (1992a). Mixtures of type I and type II β -turn conformers showing helix-like (class C) circular dichroism spectra. *Int. J. Peptide Protein Res.* (in press).
- Perczel, A., Hollósi, M., Tusnady, G., & Fasman, G.D. (1991b). Convex constraint analysis: A natural deconvolution of circular dichroism curves of proteins. *Protein Eng.* 4, 669-679.
- Perczel, A., Hollósi, M., Tusnady, G., & Fasman, G.D. (1989). Convex constraint decomposition of circular dichroism curves of proteins. *Croatica Chim. Acta* 62, 189-200.
- Perczel, A., Park, K., & Fasman, G.D. (1992b). Deconvolution of the circular dichroism spectra of proteins: The circular dichroism spectra of the antiparallel β -sheet in proteins. *Proteins Struct. Funct. Genet.* 13, 57-69.
- Perczel, A., Park, K., & Fasman, G.D. (1992c). Analysis of the circular dichroism spectrum of proteins using the convex constraint algorithm: A practical guide. *Anal. Biochem.* 203, 83-93.
- Pimplikar, S.W. & Reithmeier, R.A.F. (1986). Affinity chromatography of Band 3, the anion transport protein of erythrocyte membranes. *J. Biol. Chem.* 261, 9770-9778.
- Provencher, S.W. & Glöckner, J. (1981). Estimation of globular protein secondary structure from circular dichroism. *Biochemistry* 20, 33-37.
- Rafferty, C.N., Cassim, J.Y., & McConnell, D.G. (1977). Circular dichroism, optical rotatory dispersion, and absorption studies on the conformation of bovine rhodopsin in situ and solubilized with detergents. *Biophys. Struct. Mechanism* 2, 277-320.
- Rath, P., Bousché, O., Merrill, A.R., Cramer, W.A., & Rothchild, K.J. (1991). Fourier transform infrared evidence for a predominantly α -helical structure of the membrane bound channel forming COOH terminal peptide of colicin E1. *Biophys. J.* 59, 516-522.
- Reynolds, J.A. & Stoekenius, W. (1977). Molecular weight of bacteriorhodopsin solubilized in Triton X-100. *Proc. Natl. Acad. Sci. USA* 74, 2803-2804.
- Sarkar, P.K. & Doty, P. (1966). The optical rotatory properties of the β configuration in polypeptides and proteins. *Proc. Natl. Acad. Sci. USA* 55, 981-989.
- Saxena, V.P. & Wetlaufer, D.B. (1971). A new basis for interpreting the circular dichroic spectra of proteins. *Proc. Natl. Acad. Sci. USA* 68, 969-972.
- Schneider, A.S. & Rosenheck, K. (1982). Circular dichroism of membrane proteins. In *Techniques in Lipid and Membrane Biochemistry*, Part II, B424, pp. 1-26. Elsevier Science Publishers Ireland Ltd.
- Senior, A.E. (1988). ATP synthesis by oxidative phosphorylation. *Physiol. Rev.* 68, 177-231.
- Smith, S.O., Courtin, J., van den Berg, E., Winkel, C., Lugtenburg, J., Herzfeld, J., & Griffin, R.G. (1989) Solid-state ^{13}C NMR of the retinal chromophore in photointermediates of bacteriorhodopsin: Characterization of two forms of M. *Biochemistry* 28, 237-243.
- Smith, S.O. & Griffin, R.G. (1988). High-resolution solid-state NMR of proteins. *Annu. Rev. Phys. Chem.* 39, 511-535.
- Smith, S.O., Palings, I., Copie, V., Raleigh, D.P., Courtin, J., Pardo, J.A., Lugtenburg, J., Mathies, R.A., & Griffin, R.G. (1987). Low-temperature solid-state ^{13}C NMR studies of the retinal chromophore in rhodopsin. *Biochemistry* 26, 1606-1611.
- Tinoco, I., Jr., Woody, R.W., & Bradley, D.F. (1963). Absorption and rotation of light by helical polymers: The effect of chain lengths. *J. Chem. Phys.* 38, 1317-1325.
- Urry, D.W. & Long, M.M. (1980). Ultraviolet absorption, circular dichroism, and optical rotatory dispersion in biomembrane studies. In *Membrane Physiology* (Andreoli, T.E., Hoffman, J.F., & Faneslil, D.D., Eds.), pp. 107-124. Plenum Medical Book Co., New York.
- Vogel, H. & Jähnig, F. (1986). Models for the structure of outer membrane proteins of *Escherichia coli* derived from Raman spectroscopy and prediction methods. *J. Mol. Biol.* 190, 191-199.
- Weiss, M.S., Kreusch, A., Schiltz, E., Nestel, U., Welter, W., Weckesser, J., & Schulz, G.E. (1991). The structure of porin from *Rhodobacter capsulatus* at 1.8 Å resolution. *FEBS Lett.* 280, 379-382.
- Witt, H.T., Rögner, M., Mühlenhoff, U., Witt, I., Hinrichs, W., Saenger, W., Betzel, C., Dauter, Z., & Boekema, E.J. (1990). On isolated complexes of reaction center I and X-ray characterization

- of single crystals. In *Current Research in Photosynthesis*, Vol. II (Baltscheffsky, W., Ed.), pp. 547–554. Kluwer, Dordrecht.
- Woody, R. (1974). Studies of theoretical circular dichroism of polypeptides: Contributions of β turns. In *Peptides, Polypeptides and Proteins* (Blout, E.R., Bovey, F.A., Goodman, M., & Lotan, N., Eds.), pp. 338–350. Wiley, New York.
- Woody, R.W. & Tinoco, I., Jr. (1967). Optical rotation of oriented helices. III. Calculation of rotatory dispersion and circular dichroism of the alpha and 3_{10} helix. *J. Chem. Phys.* 46, 4927–4945.
- Yamada, M., Ebina, Y., Miyata, T., Nakazawa, T., & Nakazawa, A. (1982). Nucleotide sequence of the structural gene for colicin E1 and predicted structure of the protein. *Proc. Natl. Acad. Sci. USA* 79, 2827–2831.
- Yang, J.T., Wu, C.-S.C., & Martinez, H.M. (1986). Calculation of protein conformation from circular dichroism. *Methods Enzymol.* 130, 208–269.
- Yoshikawa, S., Choc, M.G., O'Toole, M.C., & Caughey, W.S. (1977). An infrared study of CO binding to heart cytochrome C oxidase and hemoglobin A. *J. Biol. Chem.* 252, 5498–5508.
- Zhang, Y.-Z., Ewart, G., & Capaldi, R.A. (1991). Topology of subunits of the mammalian cytochrome C-oxidase: Relationship to the assembly of the enzyme complex. *Biochemistry* 30, 3674–3681.

Appendix

The membrane proteins used in this study, and their origins, are listed below.

Complexes of photosynthetic membranes

Reaction center from Rhodospseudomonas viridis: Dr. Hartmut Michel, Max-Planck Institut für Biophysik, Frankfurt, Germany. *Photosystem I*: Dr. H.T. Witt, Fachbereich Physikalische und Angewandte Chemie, Max-Planck Institut für Biophysikalische und Physikalische Chemie, Technische Universität, Berlin, Germany. *Bacterial photosynthetic reaction center from Rhodobacter sphaeroides*: Dr. George Feher, Department of Physics, University of California, San Diego, California. *Light harvesting complex*: Dr. W. Kühlbrandt, European Molecular Biology Laboratory, Heidelberg, Germany. *B 800–850 antenna complex from Rhodospseudomonas acidophila 10050*: Dr. R. Cogdell, Department of Botany, University of Glasgow, Glasgow, Scotland.

Electron transporters

Ubiquinol-cytochrome C reductase from bovine heart: Dr. Chang-An Yu, Department of Biochemistry, Oklahoma Agricultural Experimental Station, Oklahoma State University, Stillwater, Oklahoma. *Mitochondrial NADH-ubiquinone reductase from Neurospora crassa*: Dr. Hanna Weiss, Institut für Biochemie, Heinrich-Heine-Universität, Düsseldorf, Germany. *Ubiquinol-cytochrome C reductase from N. crassa*: Dr. Hanna Weiss, Institut für Biochemie, Heinrich-Heine-Universität, Düsseldorf, Germany. *Cytochrome C oxidase from human erythrocytes*: Drs. W.S. Caughey and R. Woody, Colorado State University, Fort Collins, Colorado. *Cytochrome oxidase from bovine heart*: Dr. Roderick A. Capaldi, Institute of Molecular Biology, University of Oregon, Eugene, Oregon.

Porins

Porin (gene OmpF product) from E. coli: Dr. Jürg P. Rosenbusch, University of Basel, Basel, Switzerland; Drs. A. Tucker

and J.H. Lakey, European Molecular Biology Laboratory, Heidelberg, Germany. *Porin from Rhodobacter capsulatus*: Drs. Andreas Kreusch and Georg Schulz, Institut für Organische Chemie und Biochemie, Albert-Ludwigs-Universität, Freiburg, Germany. *Malto porin (LamB)*: Drs. A. Tucker and J.H. Lakey, European Molecular Biology Laboratory, Heidelberg, Germany. *Phosphoporin (PhoE)*: Drs. A. Tucker and J.H. Lakey, European Molecular Biology Laboratory, Heidelberg, Germany.

Rhodopsins

Rhodopsin from bovine retina: Dr. Daniel Oprian, Department of Biochemistry, Brandeis University, Waltham, Massachusetts. *Purple membrane from Halobacterium halobium*: Dr. Geza Sirokman in Dr. Fasman's Laboratory, Brandeis University, Waltham, Massachusetts.

ATPases

Proton ATPase from yeast plasma membrane: Dr. David S. Perlin, Department of Biochemistry, Public Health Research Institute, New York, New York. *Ca²⁺ ATPase from rabbit muscle sarcoplasmic reticulum*: Dr. Anthony N. Martonosi, Department of Biochemistry and Molecular Biology, Health Science Center, College of Medicine, State University of New York, Syracuse, New York; Dr. Raphael Benner, Institut für Biophysik und Strahlenbiologie, Freiberg, Germany. *H,K dependent ATPase*: Dr. Ed Rabon, Center for Ulcer Research and Education (CURE), Ion Transport Facility, School of Medicine, University of California, Los Angeles, California. *F₀F₁ ATPase from E. coli*: Dr. Roderick A. Capaldi, Institute of Molecular Biology, University of Oregon, Eugene, Oregon.

Receptors and transporters

Nicotinic acetylcholine receptor: Dr. Ferdinand Hucho, Fachbereich Chemie, Institut für Biochemie, Freie Universität, Berlin, Germany. *Glucose transporter from human erythrocytes*: Dr. Chan Y. Jung, Veterans Administration Medical Center, Buffalo, New York. *Band 3 from human erythrocytes*: Dr. Reinhart A.F. Reithmeier, Department of Medicine, Membrane Biology Group, University of Toronto, Toronto, Canada. *Lac permease from E. coli*: Dr. H. Ronald Kaback, Howard Hughes Medical Institute Research Laboratories, Molecular Biology Institute, University of California, Los Angeles, California.

Miscellaneous

Colicin A, thermolytic fragment: Dr. A. Tucker, European Molecular Biology Laboratory, Heidelberg, Germany. *Colicin E1 from E. coli*: Dr. W.A. Cramer, Department of Biological Science, Purdue University, West Lafayette, Indiana. *Lipoprotein from E. coli outer membrane*: Dr. Masayori Inouye, Department of Biochemistry, University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School, Piscataway, New Jersey. *Prostaglandin synthase (PGS)*: Dr. Michael Garavito, Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois.