Chapter 8

Circular-Dichroism and Synchrotron-Radiation Circular-Dichroism Spectroscopy as Tools to Monitor Protein Structure in a Lipid Environment

Koichi Matsuo and Kunihiko Gekko

Abstract

Circular-dichroism (CD) spectroscopy is a powerful tool for the secondary-structure analysis of proteins. The structural information obtained by CD does not have atomic-level resolution (unlike X-ray crystallography and NMR spectroscopy), but it has the great advantage of being applicable to both nonnative and native proteins in a wide range of solution conditions containing lipids and detergents. The development of synchrotron-radiation CD (SRCD) instruments has greatly expanded the utility of this method by extending the spectra to the vacuum-ultraviolet region below 190 nm and producing information that is unobtainable by conventional CD instruments. Combining SRCD data with bioinformatics provides new insight into the conformational changes of proteins in a membrane environment.

Key words: Circular-dichroism spectroscopy, Conformational change, Lipid environment, Membrane protein, Neural network, Secondary-structure analysis, Sequence-based prediction, Synchrotron radiation, Vacuum-ultraviolet circular dichroism

1. Introduction

Knowledge of the structures of membrane proteins is essential to understanding their biological functions such as in intracellular communication and stimulation responses of cells. However, there remain a huge number of membrane proteins whose tertiary structures are unknown due to (1) the technical difficulties of having to crystallize proteins in the membrane environment (or detergent micelle) for X-ray crystallography and (2) only small proteins being analyzable by NMR spectroscopy. Circular-dichroism (CD) spectroscopy, which is a well-established technique in structural biology, does not provide atomic resolution (unlike X-ray crystallography and NMR spectroscopy), but it has been widely used for analyzing the structure and conformational changes of membrane

Jörg H. Kleinschmidt (ed.), *Lipid-Protein Interactions: Methods and Protocols*, Methods in Molecular Biology, vol. 974,
DOI 10.1007/978-1-62703-275-9_8,

proteins (1–6) because CD spectra are very sensitive to local structures and are measurable for any proteins at low concentration under various solvent conditions or in a lipid environment. Proteins exhibit characteristic CD spectra in the visible to vacuum-ultraviolet (VUV) regions. Most folded proteins exhibit significant optical activity in the near-ultraviolet (UV) region (250–300 nm), owing primarily to the presence of aromatic residues and disulfide bonds that sensitively reflect the tertiary structure. In the far-UV region (<250 nm), these spectral characteristics are determined primarily by the secondary structures of protein. The principles, techniques, and applications of CD spectroscopy have been comprehensively reviewed in many books (7–9).

The secondary-structure analysis of proteins by CD spectroscopy has been markedly improved by (1) the development of programs (e.g., DSSP and Xtlsstr) for assigning the secondary structures from atomic coordinates (10, 11), (2) the advancements in software (e.g., CONTIN, SELCON3, and CDSSTR) for analyzing CD spectra (12), and (3) the extension of CD measurements to the VUV region (13). The short-wavelength limit of CD spectroscopy can be extended down to ~160 nm by using synchrotron radiation (SR) as a high-flux source of photons, which yields much more detailed information that cannot be obtained with a conventional CD spectrophotometer (14–16). VUV CD spectroscopy using SR (SRCD) is a powerful tool for estimating not only the contents but also the numbers of α-helix and β-strand segments with high accuracy (17, 18). Furthermore, combining SRCD data with bioinformatics (e.g., neural-network (NN) algorithms) can considerably improve the sequence-based prediction of secondary structures (19), which can also be used to test the predicted tertiary-structure model of a protein. This chapter addresses recent developments in CD and SRCD spectroscopy as tools to monitor protein secondary structures in a lipid environment, focusing on nonintegral membrane proteins.

1.1. Basic Principles

CD is defined as the difference (ΔA) between the absorbance of left- and right-handed circularly polarized light, each of which follows the Beer–Lambert law at a given wavelength $A_L = \varepsilon_L lC$ and $A_R = \varepsilon_R lC$

$$A = A_L - A_R = (\varepsilon_L - \varepsilon_R) lC = \varepsilon lC \tag{1}$$

where ε_L and ε_R are the molar absorptions of the sample for left- and right-handed circularly polarized light, respectively, C is the molar concentration of the sample, and l is the path length of the optical cell (in cm). CD instruments usually represent the CD in $\Delta\varepsilon$ or a molar ellipticity, $[\theta]$, with the latter being widely used for proteins and polypeptides. $[\theta]$ is related to $\Delta\varepsilon$ by

$$\varepsilon = [\theta]/3,298 \tag{2}$$

and is obtained by normalizing the raw ellipticity, θ, with the sample concentration and the path length of the cell:

$$[\theta] = 100\theta/lC \tag{3}$$

For proteins (or other polymers), the molar concentration unit is based on the mean residue weight (MRW), which is the molecular weight of the protein divided by the number of amino acid residues per molecule. Thus, $[\theta]$ can be obtained in the unit of degrees cm^2 $dmol^{-1}$ by

$$[\theta] = \theta \times (MRW)/10cl \tag{4}$$

where c is the concentration of the sample in mg/ml.

Consider the hen-egg lysozyme protein, which has a molecular weight of 14,314 and 129 residues—its MRW value is 111 (i.e., 14,314/129). When θ is 100° at 222 nm, c is 0.1 mg/ml, and l is 1 cm, the value of $[\theta]_{222\ nm}$ is calculated to be 11,100 (i.e., $100 \times 111/(10 \times 0.1 \times 1)$) degrees cm^2 $dmol^{-1}$ from Eq. 4, which corresponds to $\Delta\varepsilon_{222\ nm} = 3.37$ (i.e., 11,100/3,298) from Eq. 2.

2. Materials

2.1. Preparation of Protein Solution

The sample preparation methods are essentially the same in CD and SRCD measurements, although the protein concentration must be adjusted according to the path length of the cell. The most important factor is the protein purity because the presence of impurities in the protein solution will affect the CD spectrum, inevitably leading to incorrect data on the secondary structure. The purity should be at least 95% as judged by mass spectroscopy or sodium dodecyl sulfate– polyacrylamide gel electrophoresis. The protein solution should also be free from aggregates that cause light scattering and abnormal absorption. Aggregates can be removed by centrifugation (at 14,000 rpm (20,400 G) for 20 min) and/or by filtration (0.2 μm Millipore filter).

The use of a solvent with a high absorbance should be avoided, especially when CD spectra are measured down to the VUV region. Chloride and carboxylate ions exhibit strong absorption in the VUV region. Phosphate, acetate, and Tris–HCl buffers are available at high concentrations below 50 mM. HEPES buffer can be used only at concentrations below 5 mM. Chloride ions may be replaced by fluoride or sulfate ions. If the use of chloride ions is unavoidable for a reason such as protein stability or solubility, their concentration should be below 5 mM. Organic solvents such as acetonitrile, ethanol, methanol, and trifluoroethanol, which may be useful for increasing the solubility of membrane proteins, can be used at moderate concentrations in the VUV region. The absorption by buffer or solvent should be confirmed as being sufficiently low by

checking the high-tension voltage of the CD instrument before preparing protein solutions.

2.2. Preparation of Liposomes

While the liposome preparation varies with the type of lipid, the following methods are typically used (6, 20–23):

1. Lipid is dissolved in chloroform (or a chloroform/methanol mixture with a volume ratio of 3:1) and the solvent is evaporated under a stream of nitrogen gas and in a vacuum for at least 3 h. The obtained lipid film is dispersed in buffer and further hydrated by vibrating on a vortex to produce multilamellar vesicles (MLVs).
2. Sonic-disintegration and extrusion methods are often used to prepare small unilamellar vesicles (SUVs) from MLVs. In the sonic-disintegration method, SUVs are prepared by sonicating an MLV solution with a probe sonicator to near optical transparency, and the residual MLVs and the titanium particles released from the probe are removed by centrifugation at 14,000 rpm (20,400 G) for 20 min. The average diameter of the obtained SUVs can be estimated by measuring light scattering (e.g., using the DynaPro device). In the extrusion method, SUVs are prepared by filtering an MLV solution through polycarbonate membrane. The liposome size can be adjusted by using membranes with pore sizes from 1.0 to 0.03 μm (Avanti Polar Lipids). The uniformity of the liposomes can be improved by performing three freeze–thaw cycles.
3. The SUV liposome is mixed with protein in the buffer at a given mole ratio of protein and lipid by vibrating on a vortex, and the mixture is then incubated at room temperature overnight before performing CD measurements.

2.3. Determination of Protein Concentration

Accurate knowledge of the protein concentration is necessary for obtaining reliable CD spectra (Eq. 4) and secondary-structure data of proteins. Amino acid analysis is the preferred method for determining the concentration of proteins with a known sequence. In this method, the protein is completely hydrolyzed into amino acids by heating at 110°C for 24 h in a 6-N HCl solution, and the protein concentration is determined from the amount of abundant amino acids. Unfortunately, this method requires special techniques and is expensive and generally cannot be applied to membrane protein samples containing detergents and phospholipids.

The Bradford (24), Lowry (25), and bicinchoninic acid (26) methods, which are based on the chromogenic reactions of ligands or dyes bound to proteins, are more convenient for determining protein concentrations. However, these methods might not be sufficiently accurate because the response is dependent on the proteins involved and is affected by the coexistence of substances such as phospholipids.

The most convenient and practical method is to measure the protein absorbance at 280 nm (A_{280}), which is attributed to absorptions by Trp, Tyr, and Cys residues (27, 28) (see Note 1). This method calculates a protein concentration based on its extinction coefficient (absorbance at $l = 1$ cm and $c = 1$ mg/ml), which is predetermined by standardizing the absorption of a protein solution whose concentration has been determined precisely by another methods such as a dry-weight method. When this value is unknown, the molar extinction coefficient of the protein ($\varepsilon_{\text{protein}}$) can be estimated from those of Trp, Tyr, and Cys as follows:

$$\varepsilon_{protein} = (\varepsilon_{Trp} \times N_{Trp}) + (\varepsilon_{Tyr} \times N_{Tyr}) + (\varepsilon_{Cys} \times N_{Cys}), \quad (5)$$

where ε_{Trp}, ε_{Tyr}, and ε_{Cys} are the molar extinction coefficients of Trp, Tyr, and Cys at 280 nm, respectively, and N_{Trp}, N_{Tyr}, and N_{Cys} are the numbers of corresponding amino acid residues per protein molecule. Since the values of ε_{Trp}, ε_{Tyr}, and ε_{Cys} are known (27), we can calculate $\varepsilon_{\text{protein}}$ and the protein concentration (c_{protein}) using the following equations:

$$\varepsilon_{protein} = 5,690 N_{Trp} + 1,280 N_{Tyr} + 120 N_{Cys} \quad (6)$$

and

$$c_{protein}(mg/ml) = (molecular \times weight) \times (A_{280}/\varepsilon_{protein}). \quad (7)$$

For example, hen-egg lysozyme contains six Trp, three Tyr, and eight Cys residues per molecule (the molecular weight is 14,314). Then, $\varepsilon_{lysozyme} = 5,690 \times 6 + 1,280 \times 3 + 120 \times 8 = 38,940$. If A_{280} is 0.500, $c_{lysozyme} = (0.500/38,940) \times 14,314 = 0.183\, mg/ml$. This absorption method cannot be applied to proteins that do not contain Trp and Tyr residues because the calculated extinction coefficients contain significant errors, as expected from Eq. 6. Any disturbance to the absorbance at 280 nm such as by cofactors and light scattering also introduces large errors into the calculated protein concentration. The contribution of light scattering to the protein absorbance may be estimated by plotting the logarithm of the absorbance against the logarithm of the wavelength above 310 nm and then extrapolating the line down to 280 nm (29).

3. Methods

3.1. CD Instruments

Two types of CD instruments are in use today: (1) conventional CD instruments use a laboratory-based light source and work in the far-UV region down to about 190 nm and (2) SRCD instruments use SR as an intense light source and work in the VUV region down to about 120 nm, and they are designed individually using commercially available components and are set up in the beam lines of SR

facilities. CD instruments using H_2 discharge as the conventional source—which were constructed by some CD experts (30, 31)—are no longer in use.

3.1.1. Conventional CD Instruments

Conventional CD spectrophotometers equipped with xenon lamps are commercially available and are now ubiquitous. These instruments employ similar optical devices but somewhat different data-processing systems. Most optical devices consist of a monochromator, a polarizer (POL), a photoelastic modulator (PEM) for producing right- and left-handed circularly polarized light, and a photomultiplier tube (PMT) for detecting the signal. The CD signal is proportional to the ratio of the intensities of the AC and DC signals on the PMT. Most instruments keep the DC signal constant by controlling the voltage to the PMT. The intensities and signs of the AC signal are detected by a phase-sensitive lock-in amplifier (LIA) and displayed as the CD value or θ in the unit of millidegrees. The sample chamber is purged by nitrogen gas and equipped with a Peltier temperature-control unit. The short-wavelength limit is typically around 180 nm, except in water or in a lipid environment when it is around 190 nm or longer.

Optical cells constructed from fused quartz are often used for conventional CD measurements. Cells with various shapes (square and round) and path lengths (0.1–10 mm) are commercially available. For path lengths of 1 mm or less, demountable cells are recommended because they are relatively easy to fill with the sample solution and clean on the inside.

3.1.2. SRCD Instruments

The SRCD spectrophotometer has been developed worldwide since the first one was constructed in 1980 (32, 33), and proof-of-principle and validation studies have been performed for applications in structural biology. In 2010 there were 14 operational SRCD beam lines, including at the NSLS, ISA, SRS, HiSOR, BSRF, NSRRC, Soleil, TERAS, BESSYII, ANKA, and Diamond synchrotrons (9). These beam lines employ various insertion devices and photon energies (which are three to six orders of magnitude more intense than xenon lamps in the VUV wavelength region around 180 nm). The recent development of SRCD is summarized in reviews (34, 35) and in a book by Wallace and Janes (9). Here we present the optical system of the SRCD spectrophotometer and the optical cell that are operating at Hiroshima Synchrotron Radiation Center (HiSOR). The storage-ring energy of this facility is 0.7 GeV, which is suitable for VUV and soft-X-ray experiments.

Figure 1 shows a block diagram of the SRCD spectrophotometer at HiSOR. All optical devices are set up in two separate chambers: a polarization modulation chamber and a sample chamber. Both chambers are evacuated to 10^{-4} Pa by a turbo molecular pump in order to avoid the absorption of light by air and water

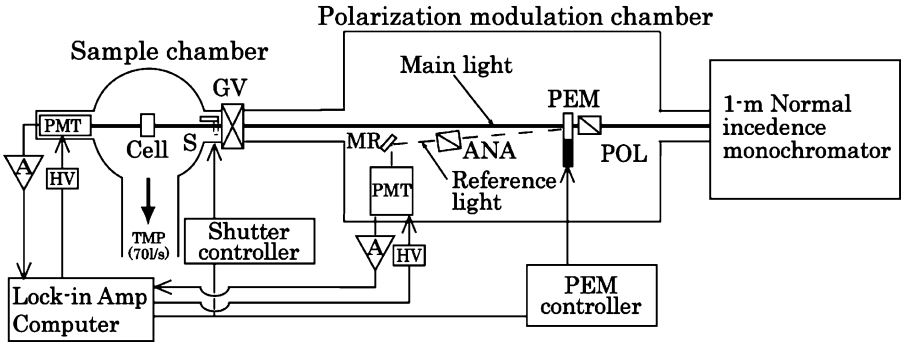

Fig. 1. Block diagram of an SRCD spectrophotometer. *A* preamplifier, *Amp* amplifier, *ANA* analyzer, *GV* gate valve, *HV* high-voltage supply, *MR* mirror, *PEM* photoelastic modulator, *PMT* photomultiplier tube, *POL* polarizer, *S* shutter, *TMP* turbo molecular pump (Adapted from Ref. 16).

vapor; this contrasts with most other SRCD instruments being purged with nitrogen gas. The sample chamber is separated from the polarization modulation chamber by a gate valve, which allows the sample to be exchanged while maintaining a high vacuum in the modulation chamber. The incident light from a monochromator is separated into two orthogonal linearly polarized light beams by an MgF_2 polarizer. Both linearly polarized light beams are modulated into circularly polarized light at 50 kHz by an LiF PEM. The main light beam is directed to the sample cell, and the CD signal is detected using the main PMT, which is covered by an MgF_2 window. The CD signal is rectified and amplified by the LIA and finally recorded on a personal computer. Another light beam, which is detected by a reference PMT, is used as a reference signal to accurately control the PEM and stabilize the LIA. Although it is generally difficult to obtain optimal driving voltages for the PEM due to refractive-index dispersion, this double-beam system—a so-called servo-control system (16, 36, 37)—achieves achromatic modulation and compensation for the thermal drift of the PEM in the VUV region. The offset signal due to the strain birefringence of the MgF_2 window on the main PMT can be minimized by rotating the PMT using a differentially pumping rotary feedthrough apparatus. Damage to the sample by the SR irradiation can be avoided, when necessary, using a shutter that opens automatically only when signals are being accumulated.

Figure 2 shows a schematic of the optical cell for SRCD measurements, which consists of a stainless-steel container with a cylindrical screw and two CaF_2 (or MgF_2) windows that are 20 mm in diameter and 1 mm thick (see Note 2). The optical path length can be adjusted over the range from 1.3 μm (without a spacer) to 100 μm (with donut-shaped Teflon spacers) (see Note 3). The sample solution in the cell is sealed with three fluoride-rubber O-rings (38) (see Note 4). The optical cell is assembled by first setting up two O-rings and a CaF_2 disc on a stainless-steel bottom block.

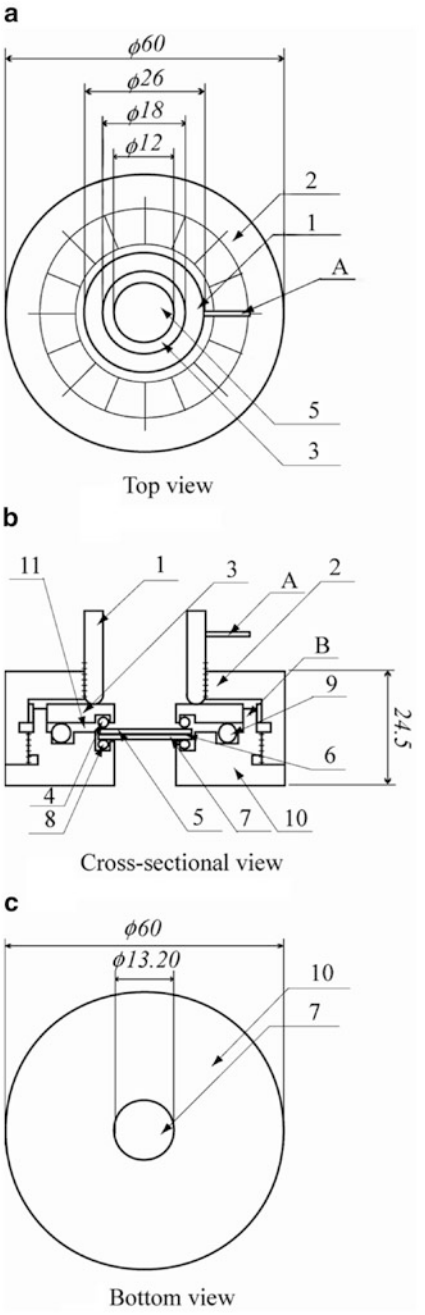

Fig. 2. Block diagram of the optical cell: top (**a**), cross-sectional (**b**), and bottom (**c**) views. *1*, cylindrical screw; *2*, stainless-steel cap; *3*, stainless-steel cover; *4*, *8*, and *9*, fluoride-rubber O-rings; *5* and *7*, CaF_2 discs; *6*, Teflon spacer; *10*, stainless-steel bottom block; *11*, space for overflow; *A*, needle; *B*, disc stopper. Linear dimensions in italics are expressed in millimeters (Adapted from Ref. 38).

A spacer to produce the required path length is then placed on the CaF_2 disc, and the sample solution is mounted at the center of this disc. The sample solution is then covered by another CaF_2 disc and sealed by a stainless-steel cover with an O-ring (see Note 5). Any of the sample that overflows through the spacer is trapped in a separate space. Finally, these assembled components are tightly fixed by a stainless-steel cap and the cell windows are pressed uniformly by turning the needle of a cylindrical screw to a position indicated on the steel cap. The CaF_2 discs and the spacer are fixed by a stopper on the bottom block. These accessories are useful for reassembling the cell with good reproducibility (see Note 6). Before assembling the cell, the CaF_2 discs and the spacer are cleaned with distilled water and ethanol (38). This sample cell is designed so that it can also be used in a conventional CD instrument to enable easy calibration of the path length and the sample concentration. The temperature of the sample is regulated between −20 and 70°C by the Peltier thermoelectric element (see Note 7).

3.1.3. Calibration of CD Instruments

The wavelength and intensity of the spectra must be calibrated in conventional CD and SRCD spectrophotometers using standard materials (see Note 8). For accuracy, calibration should be performed at more than one wavelength in order to confirm the linear wavelength dependence of the intensity over the full wavelength range of a measured protein spectrum. D-(+)-10-camphorsulfonic acid and ammonium camphorsulfonate are typically used as the standards because they have two characteristic CD peaks at 290.5 nm ($\Delta\varepsilon = +2.36$) and 192.5 nm ($\Delta\varepsilon = -4.72$) (39).

3.2. CD Measurements

CD measurements can be made using a conventional instrument following the attached manual describing the adjustable parameters, such as the scan speed, time constant, bandwidth, and accumulation time. The scan speed and time constant are important because the resolution of the CD spectrum is dominantly affected by the combination of these two parameters. A sample with strong absorption often produces a large amount of noise in the CD spectrum. The signal-to-noise (S/N) ratio can be improved by increasing the time constant and accumulation time within a limited range. When the sample solution scatters light, its contribution to the CD spectrum may be reduced by using a short-path-length cell and by positioning the sample cell near the detector of the CD instrument. The CD spectrum of the solvent needs to be measured before performing the sample measurements since even water (which has no chirality) generally exhibits small baseline signals over a wide wavelength range due to the inevitable optical constraints of CD instruments. The true CD spectrum of the sample is then obtained by subtracting the solvent spectrum from the observed sample spectrum.

SRCD measurements are basically the same as conventional CD measurements. Since a short-path-length cell (<50 μm) is used in SRCD measurements to reduce the absorption of the solvent in the VUV region (see Note 9), SRCD measurements require a higher protein concentration than conventional CD measurements. The effects of increased light scattering in the VUV region might be reduced by ensuring that the sample cell and detector (i.e., the PMT) are as close as possible. The S/N ratio can also be improved by performing repeated scans (e.g., 16 accumulations), although long-time data acquisition should be avoided in order to protect the sample from damage induced by SR irradiation when using high-flux beam lines (greater than 2 GeV) (see Note 10).

Accurate knowledge of both the path length of the cell and the protein concentration is needed to obtain reliable CD spectra. Although the path lengths of commercially available cells are indicated by their manufacturers, it is desirable to check their path lengths using the standard sample of known concentration and absorption (or ellipticity) at a given wavelength. When using short-path-length cells (<50 μm) for SRCD measurements, the interference-fringe method using a UV spectrophotometer is useful, although it is only applicable to an empty cell. An alternative method employs confocal laser microscopy to determine the gap between the windows with an accuracy of 0.1 μm for the solution-mounted cell, but this is expensive and requires a high degree of technical skill. A more convenient method is to normalize the CD spectra measured using a short-path-length cell (<50 μm) with those measured using a longer-path-length cell (e.g., 1 mm) of known accuracy in the wavelength region where their spectra overlap. This method is available for very-short-path-length cells, including those without spacers.

3.3. Secondary-Structure Analysis of Proteins

CD spectra of a protein in the far-UV and VUV regions sensitively reflect secondary structures such as α-helices, β-sheets, turns, and unordered structures. Therefore, we can estimate the contents and the numbers of segments of these secondary structures of an unknown protein from its CD spectra using some analytical programs. These secondary-structure data can be further combined with bioinformatics to improve the sequence and tertiary-structure predictions. These advancements in CD spectroscopy are due to developments in analytical programs and the extension of CD measurements to the VUV region. As demonstrated in Fig. 3, the SRCD spectra of proteins and the secondary-structure components (α-helix, β-sheet, turn, and unordered structures) exhibit characteristic peaks below 195 nm, thereby providing much more information on the secondary structures compared with conventional far-UV CD spectra. The secondary-structure analysis by CD spectroscopy mainly consists of the following steps (as depicted by the protocol in Fig. 3):

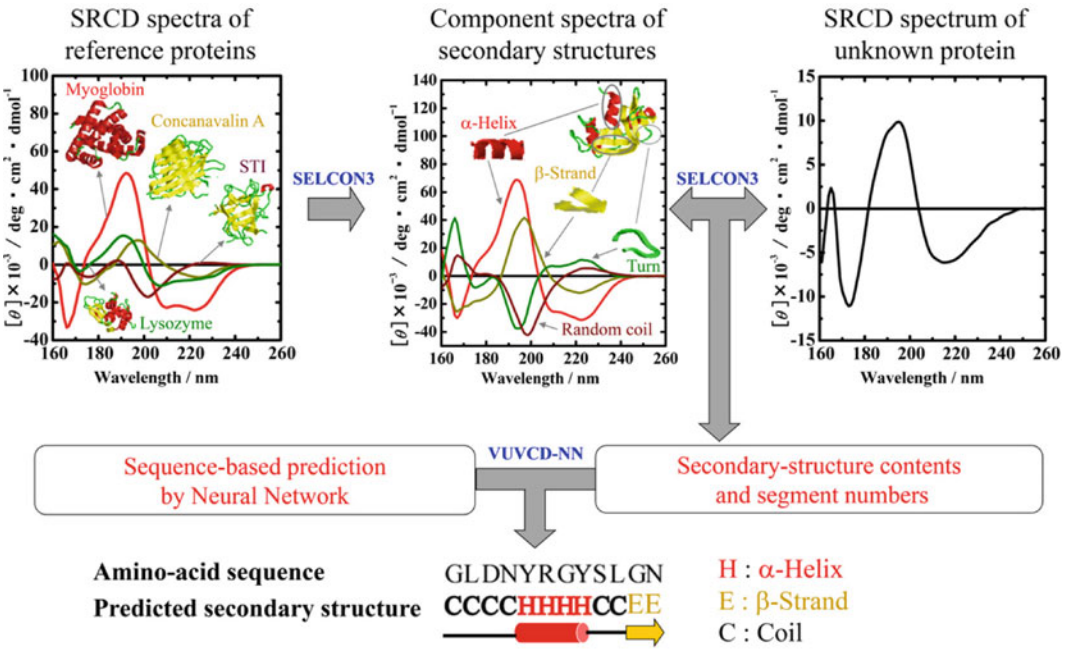

Fig. 3. Scheme for the secondary-structure analysis of proteins by CD spectroscopy.

1. Assign the secondary structures from atomic coordinates using various programs (e.g., DSSP, STRIDE, and Xtlsstr) (10, 11, 40).
2. Estimate the secondary-structure contents using analytical software (e.g., CONTIN, SELCON3, and CDSSTR) (12).
3. Estimate the number of secondary-structure segments (41).
4. Predict the secondary-structure sequences based on bioinformatics (e.g., using an NN) (19).
5. Test the tertiary-structure models predicted by homology modeling or ab initio calculations (22).

Each analytical step is described in more detail in Subheadings 3.3.1–3.3.6.

3.3.1. Secondary-Structure Assignment from Atomic Coordinates of Proteins

Estimating the secondary structures of an unknown protein requires their component spectra to be extracted from a set of CD spectra of reference proteins with known secondary structures. The secondary structures are assigned from the atomic coordinates of proteins determined by X-ray crystallography or NMR spectroscopy, which are registered in the Protein Data Bank (PDB). DSSP, which is the most widely used assignment program (10), defines the secondary structures based on the pattern of hydrogen-bonded geometrical features and the solvent exposure and thus can assign the α-, 3_{10}-, and π-helices, β-strands, turns, bends, and β-bridges at the residue level (see Note 11). Another general-purpose program

is STRIDE, which assigns the protein secondary structures by combining the hydrogen-bond energy and statistically derived backbone torsion angle (40). This program can determine the same types of secondary structures as determined by the DSSP program, except for bend structures. Both programs are adopted as general assignment methods for protein secondary structures in the PDB. Xtlsstr is a program that determines the secondary structures visually based on two dihedral angles (ζ and τ) and three distances (two hydrogen bonds and one nonhydrogen bond) in amide–amide interactions. This program (unlike STRIDE and DSSP) can assign secondary structures without hydrogen bonds that are outside the normal dihedral angles of the secondary structures in a Ramachandran plot (11), thus providing information about α- and 3_{10}-helices, β-strands, hydrogen-bonded turns, nonhydrogen-bonded turns, and PPII (polyproline type II) helices. Xtlsstr might be useful for CD-spectroscopy-based secondary-structure analysis because CD recognizes not only the pattern of hydrogen bonds but also the dihedral angles of amide groups (11).

3.3.2. Database of Protein CD Spectra

A database of protein CD spectra needs to be available in order to analyze the secondary-structure contents and numbers of segments. Many databases containing a wide range of secondary structures are available on the Internet (e.g., CDPro (http://lamar.colostate.edu/~sreeram/CDPro/) and DichroWeb (http://dichroweb.cryst.bbk.ac.uk/html/home.shtml)) (12, 42), but most of them comprise the CD spectra of soluble proteins, and their utility in analyzing the CD spectra of membrane proteins has been questioned. Park et al. suggested that membrane proteins are immersed in a medium that has a much lower dielectric constant than are soluble proteins, which affects the π–π^* and n–π^* transitions (43). Furthermore, the average chain length of α-helices integrated in plasma membrane (i.e., 22–28 residues) is about twice that of α-helices in soluble proteins. Wallace et al. revealed that soluble and membrane proteins exhibit somewhat different spectral characteristics by comparing the CD spectra of two soluble proteins (myoglobin and concanavalin A) and two membrane proteins (a mechanosensitive channel and a ferric enterobactin receptor) (44). On the other hand, Sreerama and Woody (who developed the CDPro software including seven protein CD databases) found no evidence for wavelength shifts between the CD spectra of soluble and membrane proteins in a comparison of the CD spectra of 30 membrane proteins reported by Park et al. (43) with the SP43 soluble protein reference set in CDPro (45). Thus, while CD spectra might not differ significantly between membrane and soluble proteins, it is known that combining their data sets improves the predictive accuracy of secondary-structure contents and numbers of segments in both protein types (42, 45).

3.3.3. Calculation of Secondary-Structure Contents

Many programs for analyzing secondary-structure contents use a reference CD data set comprising many proteins with known atomic coordinates (46). Typical programs are CONTIN (47), SELCON3 (48), CDSSTR (49), CCA (50), and K2D (51), which employ analytical techniques such as ridge regression (47), variable selection (49), singular value decomposition (48), convex constraint analysis (50), and NN methods (51). These programs generally use the following matrix equation:

$$F = XC \tag{8}$$

where **F** and **C** are matrices consisting of the secondary-structure contents and the CD spectra of n reference proteins, respectively. Transformation matrix **X** is solved using a least-squares method, and then the secondary-structure contents of an unknown protein can be estimated from its CD spectrum. CONTIN, SELCON3, and CDSSTR are freeware programs (http://lamar.colostate.edu/~sreeram/CDPro/) that can provide the secondary-structure fractions under some general restrictions: the sum of the fractions is between 95 and 105%, each fraction is greater than –2.5%, and the root-mean-square deviation between the reconstructed and experimental CD is less than 0.25 $\Delta\varepsilon$. The accuracy of this method can be improved by (1) increasing the number of reference proteins, (2) selecting an appropriate reference set from the large available set of reference proteins, (3) stabilizing the least-squares analysis of **X** for the large set of reference proteins, and (4) extending the short-wavelength limit of the CD spectra. It is the last one of these factors that an SRCD instrument addresses.

The improvement of the secondary-structure estimation was evaluated using the root-mean-square deviation (RMSD) (δ) and Pearson correlation coefficient (r) between X-ray and CD estimates of α-helix, β-strand, turn, and unordered-structure contents. As the short-wavelength limit of CD spectra decreases from 200 to 160 nm, the δ values for α-helices, β-strands, and overall performance decrease and the r values increase up to 0.937 and 0.826 for α-helices and β-strands, respectively, indicating an improved estimation of secondary-structure contents (17, 18). The good performance of SRCD spectroscopy increases the number of secondary-structure components to eight (52), which are useful for estimating the numbers of α-helix and β-strand segments.

3.3.4. Estimation of Numbers of Secondary-Structure Segments

The numbers of α-helix and β-strand segments have been estimated from the CD spectra using two methods. Pancoska et al. used a matrix descriptor of secondary-structure segments for the NN-based analysis of proteins (53). Sreerama et al. estimated these numbers from the distorted residues in α-helices and β-strands, assuming that on average there are four and two distorted residues per α-helix and β-strand, respectively (41). In this

method, the secondary structures are classified into six types: regular α-helix (α_R), distorted α-helix (α_D), regular β-strand (β_R), distorted β-strand (β_D), turn, and unordered structure; in this case matrix **F** in Eq. 8 involves six secondary-structure vectors of *n* reference proteins. The numbers of α-helix and β-strand segments calculated from the contents of α_D and β_D of 31 reference proteins agree well with those obtained from X-ray structures: the *r* values between the SRCD and X-ray estimates for the numbers of α-helix and β-strand segments are 0.954 and 0.849, respectively, which correspond to root-mean-square differences of 2.6 and 4.0. Thus, SRCD spectroscopy provides superior performance compared with conventional CD spectroscopy in estimating not only the secondary-structure contents but also the numbers of secondary-structure segments of proteins (17, 18).

3.3.5. Prediction of Secondary-Structure Sequences

CD spectroscopy in principle yields no information on the sequences of the secondary structures, but sequence-based prediction of α-helices, β-strands, and other structures is possible by exploiting the correlations between the X-ray structures and aminoacid sequences of many proteins without having to perform other experiments. Various types of algorithm (Chou–Fasman, GOR, and Lim methods) have been developed for this purpose (54–56), and their predictive accuracy has been greatly improved by employing various computational techniques, including NNs (57, 58), Profile Network from HeiDelberg (59), discrimination of secondary-structure class (60), Predator (61), and Jpred (62). However, the accuracy of these computational predictions of the secondary-structure contents and numbers of segments is lower than that of SRCD spectroscopy—the accuracy of these sequence-based algorithms should be improved by combining them with the SRCD data. Based on this idea, we developed a method for incorporating SRCD data into the NN algorithm developed by Jones (63) to improve sequence-based prediction (VUVCD–NN method) (19).

The computational protocol for the VUVCD–NN method is shown in Fig. 4. The α-helix and β-strand weights (W_α and W_β) of 20 amino acids are first calculated by the NN method (step 1 in the figure) (see Note 12), and the positions of α-helices and β-strands are then assigned based on the amino acid sequence in descending order of their weights until the determined numbers of α-helix and β-strand residues converge to those estimated in SRCD analysis. Those residues that are assigned as neither α-helices nor β-strands are considered to be other structures (step 2). Overlapping of α-helix and β-strand segments is avoided using the following selection rules: Overlapping regions (residues) are assigned to α-helices for target proteins that have more than twice as many α-helix residues as β-strand residues, while they are assigned to β-strands for target proteins that have more than twice as many β-strand residues as

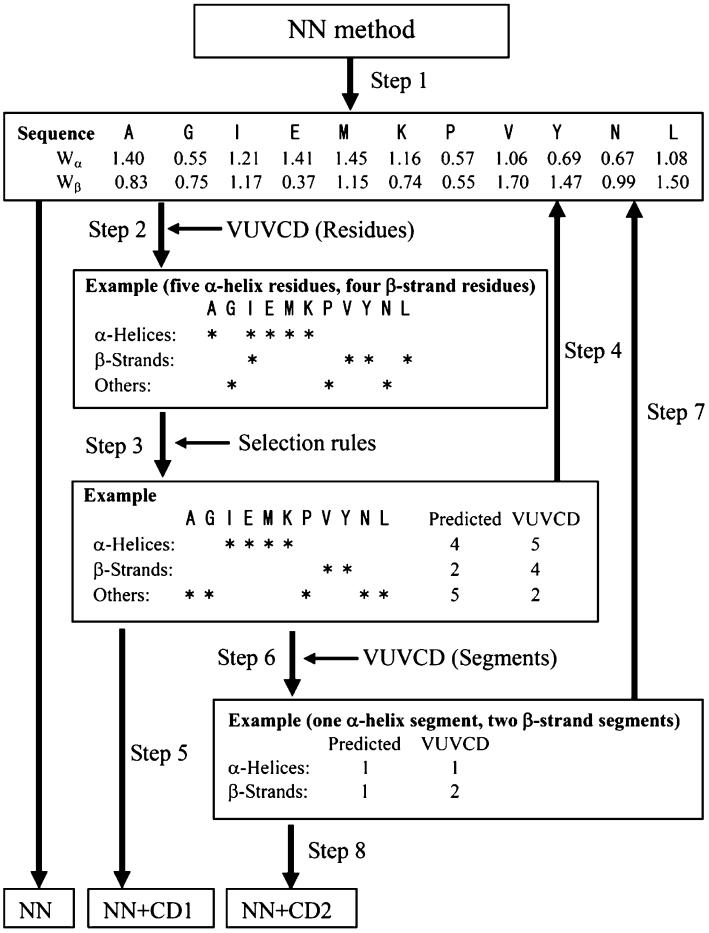

Fig. 4. Computational protocol for combining SRCD data with the NN algorithm (VUVCD–NN method). Details are provided in the text (Adapted from Ref. 19).

α-helix residues. For other target proteins, overlapping regions are assigned to either α-helices or β-strands by directly comparing the weights of the overlapping residues. Furthermore, short stretches comprising less than four α-helix residues and less than two β-strand residues are classified as other structures since these short-peptide-length limits are used as the cutoff parameters of secondary structures in the DSSP program (10) (step 3). The procedures in steps 2 and 3 are repeated until the predicted numbers of α-helix and β-strand residues (contents) agree with those estimated in SRCD analysis (step 4). The assignment is complete if the two estimates are consistent or if the sequence alignment minimizes the difference between the two estimates (step 5). Next, the numbers of α-helix and β-strand segments estimated in SRCD analysis are used in the NN calculation (step 6). If the predicted numbers of α-helix and β-strand segments do not agree with those estimated in SRCD analysis, the weights (W_α) of 20 amino acids for

α-helices are multiplied by factors from 0.1 to 2.0 until the predicted numbers of segments converge to those obtained in the SRCD estimates (step 7). The assignment is complete if the two estimates are consistent or if the sequence alignment minimizes the difference between the two estimates (step 8).

The accuracy of predicting sequence-based secondary structures using the VUVCD–NN method can be tested with three indices: Q_3, C_{sr}, and the segment overlap value (*SOV*). Q_3 is the overall percentage of correctly predicted residues for α-helices, β-strands, and other structures (10); C_{sr} is the correlation coefficient between the success rates of different secondary structures, thereby reflecting a higher-ordered predictive accuracy (64); and the *SOV* is a more meaningful measure of the secondary-structure predictive accuracy since it is related to the types and positions of secondary-structure segments, in contrast to a per-residue assignment such as Q_3 (65). These three indices evidently increase by introducing the contents and numbers of segments and by selecting the reference data set obtained from the SRCD spectra down to 160 nm; for example, the Q_3 value of the NN algorithm increased from 70.9 to 74.9%. When combined with the currently best sequence-prediction algorithms, the predictive accuracy could be enhanced to over 80%.

3.3.6. Testing Tertiary-Structure Models

Secondary-structure information can be used to test the tertiary-structure models predicted by homology modeling or ab initio calculations. There are many model-prediction methods, such as the Modeler package (66–68), but it is generally difficult to select the most appropriate model to use. Secondary-structure information obtained by SRCD spectroscopy cannot prove which model is correct, but it can provide a novel criterion for eliminating certain models. Thus, SRCD spectroscopy and VUVCD–NN methods are indispensable tools for analyzing the structure of proteins that are difficult targets in X-ray crystallography and NMR spectroscopy.

3.4. Application to Nonintegral Membrane Proteins

3.4.1. α-Amyloid Peptide βAP

α-Amyloid peptide βAP (residues 1–40) is a major component of Alzheimer's amyloid deposits and presumably exerts its neurotoxic action via interactions with neuronal membranes. Terzi et al. measured the CD spectra of βAP in the absence and presence of 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoglycerol (POPG) at βAP/POPG molar ratios of 1:11, 1:22, 1:55, and 1:110 and observed distinct changes in the spectra, which suggest the presence of an equilibrium conformation between coil and β-structured aggregate at low βAP/POPG ratios and the transition from a β-structure to an α-helix at βAP/POPG ratios higher than 1:55 (1). Bokvist et al. measured the CD spectra of βAP interacting with vesicles for different molar ratios of L-α-dimyristoylphosphatidylcholine (DMPC) to DMPG (L-α-dimyristoylphosphatidylglycerol) (4). Based on the contents of α-helices, β-strands, and unordered

structures, they suggested that there are two different membrane-binding mechanisms of βAP to hinder and accelerate aggregations. Yagi-Utsumi et al. measured the SRCD spectra of βAP down to 175 nm in the absence and presence of monosialotetrahexosylganglioside (GM1) at βAP/GM1 ratios of 1:15 and 1:30 and found that the contents of α-helix, β-strand, turn, and unordered structures depend on this ratio (23). Combining these data with NMR spectroscopy, they suggested that GM1 clusters promote specific βAP–βAP interactions that vary with the sizes and curvatures of the clusters.

3.4.2. α-Synuclein

α-Synuclein (α-syn) plays an important role in lipid transport and synaptic membrane biogenesis. Jo et al. measured the CD spectra of α-syn in various membrane environments: 1-palmitoyl 2-oleoyl phosphatidylcholine (POPC), 1-palmitoyl 2-oleoyl phosphatidylserine (POPS), phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidic acid (PA), phosphatidylinositol, and phosphatidylserine (PS) (2). From the dependence of α-helix content on the type of membranes, they showed that α-syn binds to acidic phospholipid vesicles (PS and PA), with this being significantly augmented by the presence of PE (neutral phospholipid). This result suggests that the interaction of α-syn with cellular membranes in vivo sensitively depends on the lipid compositions of the membranes. Perrin et al. measured the CD spectra of α-syn in various vesicles (PS/PA, PS/PC, POPC/1-palmitoyl 2-oleoyl phosphatidic acid, and POPC/POPS) and found a general correlation between the lipid-induced α-helix contents and the degree of binding to vesicles, which was applied to searching the lipid-binding capacity of the mutants associated with Parkinson's disease (3).

3.4.3. α_1-Acid Glycoprotein

Human α_1-acid glycoprotein (AGP) exhibits characteristic binding abilities to numerous drugs as well as to steroid hormones. The drug-binding capacity of AGP is known to decrease upon interaction with biomembranes, indicating a large conformational change of AGP (20). AGP is a glycoprotein containing five glycan chains comprising about 40% of its total mass (36 kDa), and this has made it a difficult target in X-ray crystallography and NMR spectroscopy. SRCD spectroscopy is highly advantageous for such glycoproteins because it can detect the CD spectra of sugar moieties in the VUV region below 200 nm.

Matsuo et al. measured the SRCD spectra of AGP down to 160 nm in the absence and presence of liposomes of L-α-phosphatidyl-DL-glycerol (PG) (22). As shown in Fig. 5, a liposome environment induces a large change in the CD spectra corresponding to the conformational transition of AGP from a β-strand-rich to α-helix-rich structure, although these spectra are affected by the presence of the glycan moiety. To eliminate this affect, Matsuo et al. measured the SRCD spectra of the constituent sugars—L-fucose, D-mannose,

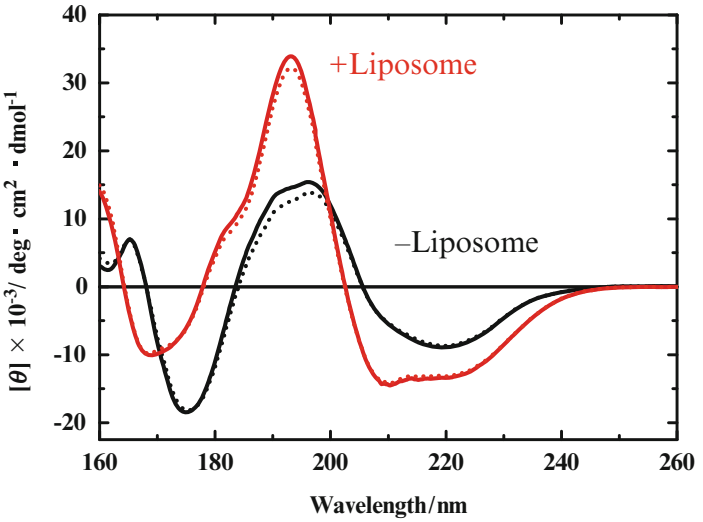

Fig. 5. SRCD spectra of AGP in the absence and presence of liposomes at 25°C. *Dotted lines* show spectra of the protein moiety obtained by subtracting the spectra of the glycan moiety (inset of Fig. 8.6) from the observed AGP spectra (Adapted from Ref. 22).

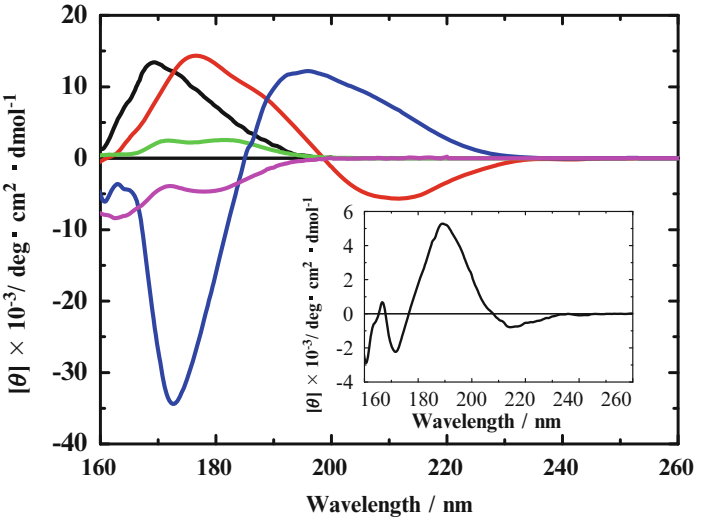

Fig. 6. SRCD spectra of the constituent sugars of AGP in water at 25°C: mannose (*black line*), fucose (*green*), galactose (*pink*), *N*-acetylglucosamine (*red*), and *N*-acetylneuraminic acid (*blue*). The inset shows the SRCD spectrum of the glycan moiety calculated from those of the constituent sugars (Adapted from Ref. 22).

D-galactose, *N*-acetylglucosamine, and *N*-acetylneuraminic acid—and constructed the SRCD spectrum of the glycan moiety by combining these constituent spectra while considering the sugar compositions in the glycan chains (Fig. 6). The SRCD spectra of the protein moiety of AGP, which were obtained by subtracting the glycan

Table 1
Secondary-structure contents and segment numbers of AGP without glycan chains in native (N) and membrane-binding (I) states (22)

State		α-Helices		β-Strands		Turns	Unordered structures
		Content (%)	Number	Content (%)	Number	(%)	(%)
N	–Glycan	14.4	3	37.7	10	19.3	28.8
	Modeler	11.5	3	36.6	10	23.5	28.4
I	–Glycan	49.9	8	5.6	2	19.3	25.4

spectrum from the observed AGP spectra, are shown by dotted lines in Fig. 5, which indicates that the contributions of the glycan chains are small but not negligible in secondary-structure analysis.

The secondary-structure contents and numbers of segments of AGP were estimated using the SELCON3 program from the SRCD spectra of the protein moiety down to 160 nm. As listed in Table 1, AGP consists of 37.7% β-strands (ten segments) and 14.4% α-helices (three segments) in the absence of liposomes (pH 7.4) and of 49.9% α-helices (eight segments) and 5.6% β-strands (two segments) in the presence of liposomes (pH 4.5), confirming the conformational change from a β-strand-rich to an α-helix-rich structure without significant changes in the contents of turns and unordered structures. These secondary-structure data are in good agreement with those calculated from the tertiary-structure model of native AGP predicted by the Modeler package (Fig. 7), confirming the validity of the predicted model.

The positions of secondary structures on the amino acid sequence were predicted using the VUVCD–NN method (see Subheading 3.5). As shown in Fig. 8a, native AGP includes 11 β-strands and 3 α-helices, while liposome-bound AGP consists of 3 α-helices and 2 β-strands. The numbers of segments thus predicted at the sequence level do not necessarily agree with those predicted from the SRCD data, which is due to sequence-alignment minimization being performed in the VUVCD–NN method. Interestingly, most of progesterone binding sites changed to helices or unordered structures in the presence of liposomes, suggesting complete collapses of the stereo-structures of the binding sites. Considering the net charge and hydrophobicity of each helix and the evidence that W25 and W160 interact with liposomes (20), Matsuo et al. proposed the following model for the membrane-binding mechanism of AGP (Fig. 8.8b): (1) positively charged AGP with a ligand approaches the anionic membrane surface via an electrostatic

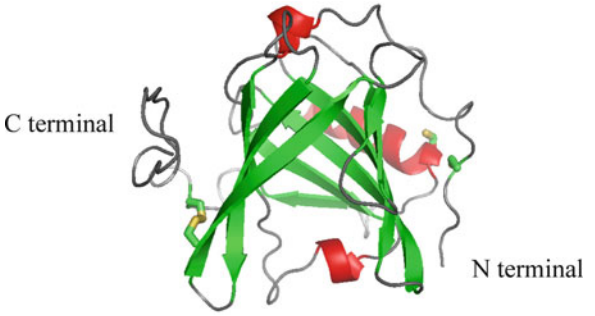

Fig. 7. Ribbon model of the native structure of AGP constructed by the Modeler package: α-helices (*red*), β-strands (*green*), and other structures (*gray*). AGP has a β-barrel-like conformation comprising ten antiparallel β-strands and three α-helices (Adapted from Ref. 22).

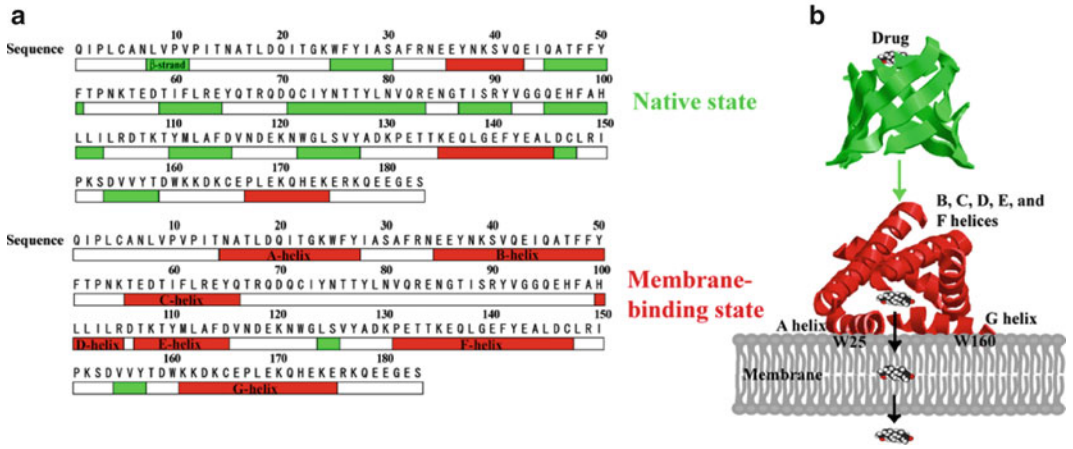

Fig. 8. Conformational change of AGP by interaction with liposomes. (**a**) Sequence-based secondary structures of AGP in the native and membrane-binding states predicted by the VUVCD–NN method. The α-helices, β-strands, and other structures are shown in *red*, *green*, and *white*, respectively. (**b**) Speculative model of the membrane-binding mechanism and structure–function relationship of AGP (Adapted from Ref. 22).

interaction and (2) W25 and W160 are inserted into liposomes via hydrophobic interactions, forming A- and G-helices in contact with the membrane surface and subsequently other helices, accompanied by the release of ligand into the membrane. This model provides the sequence-level information on the membrane-binding mechanism and the structure–function relationship of AGP, which are more generally related to the protein-mediated uptake and membrane-transport mechanisms of hormones and drugs.

3.4.4. Others

Vernier et al. measured the CD spectra of apo-myoglobin in the presence of vesicles (PA and PC) at various pH values to investigate whether the interaction of this protein with the vesicles facilitates the uptake of heme from the outer mitochondrial membrane (69).

On the other hand, Zhang and Keiderling (70) and Zhang et al. (71) measured the CD spectra of β-lactoglobulin at various PG concentrations to reveal a general mechanism of protein–membrane interactions. Arnulphi et al. examined the α-helix contents of human apolipoprotein A-I bound to POPC in the presence of sphingomyelin (SM) and cholesterol (Chol) (72) and found that raft-like lipid domains composed of POPC, SM, and Chol reduce folding of this protein compared with Chol-free model membranes.

3.5. Application to Integral Membrane Proteins

3.5.1. Photoreceptor Cyclic-Nucleotide-Gated Channels

Photoreceptor cyclic-nucleotide-gated (CNG) channels are nonselective cation channels located in the plasma membrane of the outer segments, where certain transmembrane proteins function as important components of the visual transduction system. Matveev et al. applied CD spectroscopy to examine the influence of two mutations (R377W and F488L) of the CNGA3 subunit on the structure and function of the channels (5). These mutations decreased the α-helicity of the protein, while the allosteric transitions by 8-pCPT-cGMP, cGMP, and cAMP were diminished in the two mutants. From these results they suggested that the observed alterations in the local secondary structure exert adverse effects on the channel's activity and cellular processing.

3.5.2. Protein-Dependent Manganese Transporter

MntH (*Escherichia coli* protein-dependent manganese transporter) is an integral transmembrane protein comprising 11 transmembrane segments. Nunuková et al. synthesized three peptides corresponding to the sixth segment among them and measured the CD spectra of these peptides in the presence of $MnCl_2$ and lipid membrane composed of 1,2-diphytanoyl-*sn*-glycero-3- phosphatidylcholine and 1,2-diphytanoyl-*sn*-glycero-3-[phospho-rac-(1-glycerol)] at a molar ratio of 3:1 (6). They found that these peptides interact with the lipid membrane to induce α-helices and that their helicities are further increased by adding $MnCl_2$. They also measured the CD spectra of the peptides in trifluoroethanol solution and assessed their abilities to form manganese-conducting channels.

3.5.3. Equinatoxin II

Equinatoxin II (EqtII) is a member of the actinoporin family of sea-anemone toxins that function by forming pores in cell membranes via a multistep mechanism. Miles et al. measured the SRCD spectra of EqtII down to 180 nm in the absence and presence of SUVs with different lipid compositions (dioleoylphosphatidylcholine (DOPC), SM, and Chol) (21). All of the SRCD spectra were similar in the far-UV region, but they exhibited large differences in the peak positions and intensities in the VUV region, depending on the lipid composition. The secondary-structure contents of EqtII estimated by SRCD revealed that EqtII is in the "bound" state of high lytic activity in the presence of DOPC/SM vesicles and in the "free" state in the presence of DOPC or DOPC/Chol vesicles, indicating the lipid specificity for structural rearrangements of EqtII associated with toxicity and lysis functions.

3.5.4. Others

SRCD spectroscopy has often been used to monitor structural changes of other integral membrane proteins (73–75). Miles et al. measured the SRCD spectra of pig Na, K-ATPase down to 175 nm in the membrane-bound state at various temperatures and showed that the loss of enzymatic activity is correlated with changes in the protein secondary structures (73). Cronin et al. investigated the effects of deglycosylation of sodium channels on the secondary structures using an SRCD spectrophotometer (74). Furthermore, McKibbin et al. used SRCD spectroscopy to monitor structural changes of rhodopsin and opsin aged for 1–6 h in DMPC/DHPC (L-α-1,2-dihexanoyl-*sn*-glycero-3-phosphocholine) vesicles at room temperature (75).

3.6. Future Prospects

We have demonstrated that CD spectroscopy is a powerful technique for monitoring the conformational changes of nonintegral and integral membrane proteins. Developments in SRCD instruments and in programs for analyzing CD spectra have yielded detailed secondary-structure information—in terms of the contents and numbers of segments, and the amino acid sequences—that cannot be obtained by conventional CD instruments, thereby providing new insight into the structure–function relationships of proteins in a membrane environment. The structures of a huge number of membrane proteins have barely been analyzed by X-ray crystallography or NMR spectroscopy, and hence CD and SRCD spectroscopy will play important roles in the structural biology of membrane proteins.

4. Notes

1. Since a high protein concentration is used in CD measurements, the protein solution must be diluted with solvent to the concentration at which $A_{280} = 0.5 \sim 1.0$. The protein concentration is calculated from the A_{280} value multiplied by the dilution factor determined using the volume or weight of the protein solution and the solvent.
2. Optical crystal discs of CaF_2 and MgF_2 are commercially available. However, the *c*-axis-cut disc should be customized for MgF_2 to eliminate the birefringence effects of the cell, which affect the baseline and CD signal in the VUV region.
3. A homemade die can be used to cut the donut-shaped spacer from a Teflon sheet of the desired thickness (5–100 μm).
4. The O-rings, 4, 8, and 9 in Fig. 2 might be coated with a high-vacuum grease to completely seal the sample solution.

5. The formation of air bubbles in the sample cell can be avoided by using tweezers to carefully and slowly place the cover disc on the surface of the sample solution.
6. Reproducible CD data can be obtained in separate measurements by always fixing the discs at the same position marked on their edge when assembling the optical cell.
7. The temperature of sample solution can be determined with an accuracy of $\pm 1^{\circ}C$ from the calibrated temperature difference between the heat conductor and the cell windows. Water within a sufficiently narrow space does not freeze at subzero temperatures, and hence the sample solution can be cooled down to -17 and $-10^{\circ}C$ without freezing by using cells with path lengths of 25 and 50 μm, respectively. No frost is formed on the cell windows because the cell is kept under vacuum.
8. It is desirable to routinely calibrate the instrument.
9. A 50-μm path length is available for measurements down to 175 nm, but measurements below 175 nm must be carried out without a spacer in order to reduce the effect of light absorption by water.
10. Damage to the sample can be detected by monitoring spectral changes during the measurements (1–3 h).
11. The structures denoted as "H" and "E" in DSSP are designated as α-helices and β-strands, respectively, and 3_{10}-helices ("G"), π-helices ("I"), turns ("T"), bends ("S"), β-bridges ("B"), and blanks ("") are grouped into "others" in the VUVCD–NN method.
12. The training data set is prepared from the protein structures in the PDB using a resolution cutoff of 2.0 Å for X-ray structures and by selecting the subsets of residues with an RMSD for the C_α–C_α distance of less than 1.0 Å for NMR structures. Small proteins comprising less than 30 amino acid residues and proteins containing more than 5% unidentified sequences are not used. Proteins with sequence similarities greater than 25% are eliminated using Hobohm algorithm #1 (76). Finally, a data set of 607 chains is obtained as a training set for the NN algorithm by manually eliminating any chain breaks. The secondary structures of these training proteins and their amino acid sequence information are determined by the PSSM (position-specific scoring matrices) program (63) and used to calculate the weights and biases of 20 amino acids for α-helices and β-strands with 15 window positions and no hidden layer using the public free simulation package Stuttgart Neural Network Simulator (version 4.2) (77) with a momentum term of 0.5 and a learning rate of 0.0001. The threshold criterion for assigning α-helices and β-strands is 0.37.

References

1. Terzi E, Hölzemann G, Seelig J (1997) Interaction of Alzheimer β-amyloid peptide (1–40) with lipid membranes. Biochemistry 36:14845–14852
2. Jo E, McLaurin J, Yip CM, St G-HP, Fraser PE (2000) alpha-Synuclein membrane interactions and lipid specificity. J Biol Chem 275:34328–34334
3. Perrin RJ, Woods WS, Clayton DF, George JM (2000) Interaction of human alpha-synuclein and Parkinson's disease variants with phospholipids. Structural analysis using site-directed mutagenesis. J Biol Chem 275:34393–34398
4. Bokvist M, Lindström F, Watts A, Gröbner G (2004) Two types of Alzheimer's beta-amyloid (1–40) peptide membrane interactions: aggregation preventing transmembrane anchoring versus accelerated surface fibril formation. J Mol Biol 335:1039–1049
5. Matveev AV, Fitzgerald JB, Xu J, Malykhina AP, Rodgers KK, Ding XQ (2010) The disease-causing mutations in the carboxyl terminus of cone cyclic nucleotide-gated channel CNGA3 subunit alter the local secondary structure and interfere with the channel active conformational change. Biochemistry 49:1628–1639
6. Nunuková V, Urbánková E, Jelokhani-Niaraki M, Chaloupka R (2010) Ion channel activity of transmembrane segment 6 of Escherichia coli proton-dependent manganese transporter. Biopolymers 93:718–726
7. Fasman GR (1996) Circular dichroism and the conformational analysis of biomolecules. Plenum, New York
8. Berova N, Nakanishi K, Woody RW (2000) Circular dichroism: principles and applications, 2nd edn. Wiley-VCH, New York
9. Wallace BA, Janes RW (2009) Modern techniques for circular dichroism and synchrotron radiation circular dichroism spectroscopy. IOS, Amsterdam
10. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometric features. Biopolymers 22:2577–2637
11. King SM, Johnson WC (1999) Assigning secondary structure from protein coordinate data. Proteins 35:313–320
12. Sreerama N, Woody RW (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. Anal Biochem 287:252–260
13. Toumadje A, Alcorn SW, Johnson WC Jr (1992) Extending CD spectra of proteins to 168 nm improves the analysis for secondary structures. Anal Biochem 200:321–331
14. Sutherland JC, Emrick A, France LL, Monteleone DC, Trunk J (1992) Circular dichroism user facility at the National Synchrotron Light Source: estimation of protein secondary structure. Biotechniques 13:588–590
15. Wallace BA (2000) Conformational changes by synchrotron radiation circular dichroism spectroscopy. Nat Struct Biol 7:708–709
16. Ojima N, Sakai K, Fukazawa T, Gekko K (2000) Vacuum-ultraviolet circular dichroism spectrophotometer using synchrotron radiation: optical system and off-line performance. Chem Lett 29:832–833
17. Matsuo K, Yonehara R, Gekko K (2004) Secondary-structure analysis of proteins by vacuum-ultraviolet circular dichroism spectroscopy. J Biochem 135:405–411
18. Matsuo K, Yonehara R, Gekko K (2005) Improved estimation of the secondary structures of proteins by vacuum-ultraviolet circular dichroism spectroscopy. J Biochem 138:79–88
19. Matsuo K, Watanabe H, Gekko K (2008) Improved sequence-based prediction of protein secondary structures by combining vacuum-ultraviolet circular dichroism spectroscopy with neural network. Proteins 73:104–112
20. Nishi K, Maruyama T, Halsall HB, Handa T, Otagiri M (2004) Binding of alpha1-acid glycoprotein to membrane results in a unique structural change and ligand release. Biochemistry 43:10513–10519
21. Miles AJ, Drechsler A, Kristan K, Anderluh G, Norton RS, Wallace BA, Separovic F (2008) The effects of lipids on the structure of the eukaryotic cytolysin equinatoxin II: a synchrotron radiation circular dichroism spectroscopic study. Biochim Biophys Acta 1778:2091–2096
22. Matsuo K, Namatame H, Taniguchi M, Gekko K (2009) Membrane-induced conformational change of alpha1-acid glycoprotein characterized by vacuum-ultraviolet circular dichroism spectroscopy. Biochemistry 48:9103–9111
23. Yahi-Utsumi M, Matsuo K, Yanagisawa K, Gekko K, Kato K (2011) Spectroscopic characterization of intermolecular interaction of amyloid beta promoted on GM1 micelles. Int J Alzheimers Dis. doi:10.4061/2011/925073
24. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of

protein-dye binding. Anal Biochem 72:248–254

25. Lowry OH, Rosebrough NJ, Farr AL, Randall RJ (1951) Protein measurement with the folin phenol reagent. J Biol Chem 193:265–275
26. Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, Fujimoto EK, Goeke NM, Olson BJ, Klenk DC (1985) Measurement of protein using bicinchoninic acid. Anal Biochem 150:76–85
27. Edelhoch H (1967) Spectroscopic determination of tryptophan and tyrosine in proteins. Biochemistry 6:1948–1954
28. Gill SC, von Hippel PH (1989) Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem 182:319–326
29. Wallace BA, Mao D (1984) Circular dichroism analyses of membrane proteins: an examination of differential light scattering and absorption flattening effects in large membrane vesicles and membrane sheets. Anal Biochem 142:317–328
30. Johnson CW Jr (1978) Circular dichroism spectroscopy and the vacuum ultraviolet region. Annu Rev Phys Chem 29:93–114
31. Pysh ES (1976) Optical activity in the vacuum ultraviolet. Annu Rev Biophys Bioeng 5:63–75
32. Snyder PA, Rowe EM (1980) The first use of synchrotron radiation for vacuum ultraviolet circular dichroism measurements. Nucl Instrum Methods 172:345–349
33. Sutherland JC, Keck PC, Griffin KP, Takacs PZ (1982) Simultaneous measurement of absorption and circular dichroism in a synchrotron spectrometer. Nucl Instrum Methods 195:375–379
34. Wallace BA (2009) Protein characterisation by synchrotron radiation circular dichroism spectroscopy. Q Rev Biophys 42:317–370
35. Wallace BA, Gekko K, Hoffmann SV, Lin YH, Sutherland JC, Tao J, Wien F, Janes RW (2010) Synchrotron radiation circular dichroism (SRCD) spectroscopy: an emerging method in structural biology for examining protein conformations and protein interactions. Nucl Instr Methods Phys Res A 649:177–178
36. Matsuo K, Matsushima Y, Fukuyama T, Senba S, Gekko K (2002) Vacuum-ultraviolet circular dichroism of amino acids as revealed by synchrotron radiation spectrophotometer. Chem Lett 31:826–827
37. Matsuo K, Fukuyama T, Yonehara R, Namatame H, Taniguchi M, Gekko K (2005) Vacuum-ultraviolet circular dichroism spectrophotometer using synchrotron radiation. J Electron Spectrosc Relat Phenom 144–147:1023–1025
38. Matsuo K, Sakai K, Matsushima Y, Fukuyama T, Gekko K (2003) Optical cell with a temperature-control unit for a vacuum-ultraviolet circular dichroism spectrophotometer. Anal Sci 19:129–132
39. Takakuwa T, Konno T, Meguro H (1985) A new standard substance for calibration of circular dichroism: ammonium *d*-10-camphorsulfonate. Anal Sci 1:215–218
40. Frishman D, Argos P (1995) Knowledge-based protein secondary structure assignment. Proteins 23:566–579
41. Sreerama N, Venyaminov SY, Woody RW (1999) Estimation of the number of alpha-helical and beta-strand segments in proteins using circular dichroism spectroscopy. Protein Sci 8:370–380
42. Abdul-Gader A, Miles AJ, Wallace BA (2011) A reference dataset for the analyses of membrane protein secondary structures and transmembrane residues using circular dichroism spectroscopy. Bioinformatics 27:1630–1636
43. Park K, Perczel A, Fasman GD (1992) Differentiation between transmembrane helices and peripheral helices by the deconvolution of circular dichroism spectra of membrane proteins. Protein Sci 1:1032–1049
44. Wallace BA, Lees JG, Orry AJ, Lobley A, Janes RW (2003) Analyses of circular dichroism spectra of membrane proteins. Protein Sci 12:875–884
45. Sreerama N, Woody RW (2004) On the analysis of membrane protein circular dichroism spectra. Protein Sci 13:100–112
46. Woody RW (1995) Circular dichroism. Methods Enzymol 246:34–71
47. Provencher SW, Glöckner J (1981) Estimation of globular protein secondary structure from circular dichroism. Biochemistry 20:33–37
48. Sreerama N, Woody RW (1993) A self-consistent method for the analysis of protein secondary structure from circular dichroism. Anal Biochem 209:32–44
49. Johnson WC (1999) Analyzing protein circular dichroism spectra for accurate secondary structures. Proteins 35:307–312
50. Perczel A, Hollósi M, Tusnády G, Fasman GD (1991) Convex constraint analysis: a natural deconvolution of circular dichroism curves of proteins. Protein Eng 4:669–679
51. Böhm G, Muhr R, Jaenicke R (1992) Quantitative analysis of protein far UV circular dichroism spectra by neural networks. Protein Eng 5:191–195

52. Klose DP, Wallace BA, Janes RW (2010) 2Struc: the secondary structure server. Bioinformatics 26:2624–2625
53. Pancoska P, Janota V, Keiderling TA (1999) Novel matrix descriptor for secondary structure segments in proteins: demonstration of predictability from circular dichroism spectra. Anal Biochem 267:72–83
54. Chou PY, Fasman GD (1974) Conformational parameters for amino acids in helical, beta-sheet, and random coil regions calculated from proteins. Biochemistry 13:211–222
55. Garnier J, Osguthorpe DJ, Robson B (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J Mol Biol 120:97–120
56. Lim VI (1974) Structural principles of the globular organization of protein chains. A stereochemical theory of globular protein secondary structure. J Mol Biol 88:857–872
57. Qian N, Sejnowski TJ (1988) Predicting the secondary structure of globular proteins using neural network models. J Mol Biol 202:865–884
58. Holley LH, Karplus M (1989) Protein secondary structure prediction with a neural network. Proc Natl Acad Sci U S A 86:152–156
59. Rost B, Sander C (1993) Prediction of protein secondary structure at better than 70% accuracy. J Mol Biol 232:584–599
60. King RD, Sternberg MJ (1996) Identification and application of the concepts important for accurate and reliable protein secondary structure prediction. Protein Sci 5:2298–2310
61. Frishman D, Argos P (1996) Incorporation of non-local interactions in protein secondary structure prediction from the amino acid sequence. Protein Eng 9:133–142
62. Cuff JA, Barton GJ (1999) Evaluation and improvement of multiple sequence methods for protein secondary structure prediction. Proteins 34:508–519
63. Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol 292:195–202
64. Matthews BW (1975) Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochim Biophys Acta 405:442–451
65. Zemla A, Venclovas C, Fidelis K, Rost B (1999) A modified definition of Sov, a segment-based measure for protein secondary structure prediction assessment. Proteins 34:220–223
66. Martí-Renom MA, Stuart AC, Fiser A, Sánchez R, Melo F, Sali A (2000) Comparative protein structure modeling of genes and genomes. Annu Rev Biophys Biomol Struct 29:291–325
67. Simons KT, Bonneau R, Ruczinski I, Baker D (1999) Ab initio protein structure prediction of CASP III targets using ROSETTA. Proteins 37:171–176
68. Takada S (2001) Protein folding simulation with solvent-induced force field: folding pathway ensemble of three-helix-bundle proteins. Proteins 42:85–98
69. Vernier G, Chenal A, Vitrac H, Barumandzadhe R, Montagner C, Forge V (2007) Interactions of apomyoglobin with membranes: Mechanisms and effects on heme uptake. Protein Sci 16:391–400
70. Zhang X, Keiderling TA (2006) Lipid-induced conformational transitions of beta-lactoglobulin. Biochemistry 45:8444–8452
71. Zhang X, Ge N, Keiderling TA (2007) Electrostatic and hydrophobic interactions governing the interaction and binding of beta-lactoglobulin to membranes. Biochemistry 46:5252–5260
72. Arnulphi C, Sánchez SA, Tricerri MA, Gratton E, Jonas A (2005) Interaction of human apolipoprotein A-I with model membranes exhibiting lipid domains. Biophys J 89:285–295
73. Miles AJ, Wallace BA, Esmann M (2011) Correlation of structural and functional thermal stability of the integral membrane protein Na, K-ATPase. Biochim Biophys Acta 1808:2573–2580
74. Cronin NB, O'Reilly A, Duclohier H, Wallace BA (2005) Effects of deglycosylation of sodium channels on their structure and function. Biochemistry 44:441–449
75. McKibbin C, Farmer NA, Jeans C, Reeves PJ, Khorana GH, Wallace BA, Edwards PC, Villa C, Booth PJ (2007) Opsin stability and folding: modulation by phospholipid bicelles. J Mol Biol 374:1319–1332
76. Hobohm U, Scharf M, Schneider R, Sander C (1992) Selection of representative protein data sets. Protein Sci 1:409–417
77. Zell A, Mamier G, Vogt M, Mache N, Hubner R, Doring S, Herrmann KW, Soyez T, Schmalzl M, Sommer T, Hatzigeorgiou A, Posselt D, Schreiner T, Kett B, Clemente G, Wieland J (1995) Stuttgart Neural Network Simulator, Version 4.2. University of Stuttgart, Stuttgart, Germany. http://www.ra.cs.uni-tuebingen.de/downloads/SNNS/