Protein Science (1998), 7:975–982. Cambridge University Press. Printed in the USA.

Conformational diversity of acid-denatured cytochrome *c* studied by a matrix analysis of far-UV CD spectra

TAKASHI KONNO

National Institute for Physiological Science, Okazaki 444, Japan

(RECEIVED September 10, 1997; ACCEPTED December 29, 1997)

Abstract

The singular value decomposition (SVD) analysis was applied to a large set of far-ultraviolet circular dichroism (far-UV CD) spectra (100–400 spectra) of horse heart cytochrome *c* (cyt *c*). The spectra were collected at pH 1.7–5.0 in $(NH_4)_2SO_4$, sorbitol and 2,2,2-trifluoroethanol (TFE) solutions. The present purpose is to develop a rigorous matrix method applied to far-UV CD spectra to resolve in details conformational properties of proteins in the non-native (or denatured) regions. The analysis established that three basis spectral components are contained in a data set of difference spectra (referred to the spectrum of the native state) used here. By a further matrix transformation, any observed spectrum could be decomposed into fractions of the native (N), the molten-globule (MG), the highly denatured (D), and the alcohol-induced helical (H) spectral forms. This method could determine fractional transition curves of each conformer as a function of solution conditions, which gave the results consistent with denaturation curves of cyt *c* monitored by other spectroscopic methods. The results in sorbitol solutions, for example, suggested that the preferential hydration effect of the co-solvent stabilizes the MG conformer of cyt *c*. This report has found that the systematic SVD analysis of the far-UV CD spectra is a powerful tool for the conformational analysis of the non-native species of a protein when it is suitably supplemented with other experimental methods.

Keywords: circular dichroism; cytochrome *c*; preferential hydration; singular value decomposition

The folded (or the native) conformer of a protein is in equilibrium with various denatured (or non-native) conformers. Equilibrium constants among them depend strongly on solution conditions (Tanford, 1968). A typical example is found in the case of horse cytochrome *c* (cyt *c*) at acidic pH under various solvent compositions. In water, cyt *c* is in the native state (N conformer) at neutral pH, while change in pH to an extreme low value unfolds the protein into a highly expanded acid-denatured state (D conformer). However, in the presence of salt such as sodium chloride, a major population in the acidic pH is not the D conformer but rather a more compact denatured state with a high content of secondary structures, denoted as the "molten globule" conformer (MG conformer) (Ohgushi & Wada, 1983; Goto et al., 1990a, 1990b; Jeng & Englander, 1991; Kataoka et al., 1993). Investigations on the compact denatured conformer of proteins have been a focus in the field of protein folding and stability (Kuwajima, 1989; Dobson, 1992; Ptitsyn, 1995). In a very high concentration of methanol at low pH, on the other hand, cyt *c* is in a highly expanded conformer similar to D but with an extremely high content of helical structures (H conformer) (Bychkova et al., 1996; Kamatari et al., 1996). The H conformer has been widely observed for various proteins and peptides under water–alcohol mixtures (Conio et al., 1970; Nelson & Kallenbach, 1986; Segawa et al., 1991; Dyson et al., 1992; Buck et al., 1993; Kippen et al., 1994; Shiraki et al., 1995).

A current theoretical view states that the conformational diversity of protein in the non-native region is originated from the complex shape of the free energy landscape of protein (Wolynes, 1995; Dill & Chan, 1997). To support this view experimentally, a comprehensive description on the structure of protein in a full range of the conformational space is necessary. The D, MG, and H conformers and their variants of various proteins have been characterized extensively by many experimental methods (Roder et al., 1988; Hughson et al., 1990; Dill & Shortle, 1991; Jeng & Englander, 1991; Sosnick & Trewhella, 1992; Fan et al., 1993; Alexandrescu et al., 1994; Konno et al., 1995). However, limited sensitivity and resolution of any experimental methods make this problem still largely unsolved.

A possible way to complement this relative lack may be to improve a mathematical method to extract information from an experimental data set. Increase in the size of a data set may also lead to higher sensitivity to detect a small conformational drift of protein as a function of solution conditions. For a slight contribution in this direction, the present study introduces a systematic and extensive method to analyze a large data set of horse heart cyt *c* obtained by the far-ultraviolet circular dichroism spectroscopy (far-UV CD) in various co-solvent solutions. This spectroscopic

Reprint requests to: Takashi Konno, National Institute for Physiological Science, Myodaiji, Okazaki 444, Japan; e-mail: konno@nips.ac.jp.

method is excellent because of its applicability in a very wide range of protein concentration and solution composition. In addition, this measurement can be performed with a relatively short accumulation time and high sensitivity to conformational changes of the peptide-chain backbone. The far-UV CD spectroscopy has long been used for estimating secondary structure contents of protein (Hennessey & Johnson, 1981; Yang et al., 1986; Woody, 1995). The main purpose of this study is to find out how extensively we can extract information contained in the far-UV CD spectra on the conformational change of a protein in the non-native conformational space. For that purpose, we have applied a singular value decomposition (SVD) analysis on a large set of the far-UV CD spectra of cyt *c* (typically more than 200 spectra). Application of SVD to spectroscopic or other data sets of proteins is not at all new, and is even classical (Hofrichter et al., 1983; Henry & Hofrichter, 1992; McPhie & Shrager, 1992; Jones et al., 1993; Ansari et al., 1994; Sreerama & Woody, 1994; Chen et al., 1996; Jager et al., 1997). Here, we use it to analyze far-UV CD spectral changes produced by co-solvents. The target co-solvents, used as examples, are two protein-stabilizing agents, sorbitol and ammonium sulfate, and the data are supplemented with more than 100 spectra in TFE solutions.

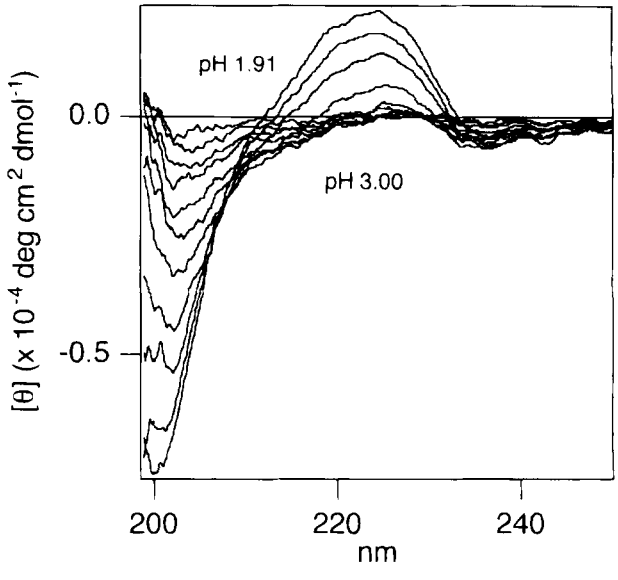

Fig. 1. Difference CD spectra of cyt *c* in 2.0 M sorbitol solution at pH 3.00, 2.79, 2.70, 2.61, 2.49, 2.39, 2.30, 2.20, 2.10, 2.01, and 1.91. The difference spectra were obtained by using a spectrum at pH 4.50 as a reference.

Results and discussion

Minimum number of basis components

The SVD method will extract the essential spectral components of far-UV CD spectra, and the linear combination of the components provides the best least-squares approximation to the spectra. See Materials and methods for further details of the SVD analysis. Far-UV CD spectra in $(NH_4)_2SO_4$, sorbitol and TFE solutions were collected carefully at 199–250 nm and pH 1.7–5.0. The CD spectra of the solution containing no cyt *c* were measured for each solvent condition to subtract carefully contributions from the background. The data matrix A was constructed using the data sets S_1–S_3 listed in Table 1. The spectra in the data set were prepared by taking difference spectra referred to the spectrum of cyt *c* in the native state at pH 4.5–5.0. Figure 1 shows, as an example, the difference CD spectra in 2.0 M sorbitol used in this study.

The SVD analysis was applied to each of the data sets S_1, S_2, and S_3. The largest five singular values w_1–w_5 for each data set are listed in Table 2, which clearly shows that the first two components dominantly contribute to the original spectra. The basis spectra u_1–u_5 for S_1 and S_3 as well as the amplitude vectors v_1–v_5 for S_1 are shown in Figures 2 and 3. We notice that u_1 and u_2 of S_1 are similar to those of S_3, while u_3–u_5 do not show this resemblance among the data sets (Fig. 2). Signal-to-noise ratios of both u and v vectors for S_1, S_2, and S_3 were excellent only for the first two components (Figs. 2 and 3). These results suggest that the first two components contribute to the spectral sets significantly, while contributions from the other components are less important. The variance in the data set unaccounted for by the μ most important basis CD spectra, σ_μ, was calculated by Equation 3 (Table 2), which also supports that the first two components are the most important to describe any experimental spectrum.

The final and the most important procedure to determine the minimum number of components is whether two components are actually enough to totally reconstruct the spectral shape of the data sets. Reconstruction of the spectra was performed using Equation 4 (see Materials and methods). The results have shown that, for S_1 and S_2, the first two components have well reproduced their original data (Fig. 4A). As for S_3, however, reconstruction from the first two components was poorly done in shape, especially at 230–250 nm (Fig. 4B). The reconstructed spectral shape was very much improved when the third compo-

Table 1. *Data sets used for the SVD analysis*

Data set	Total number of spectra
S_1: 0–2.0 M NH_4SO_4	161 spectra
S_2: S_1 + 0.25–2.0 M sorbitol	265 spectra
S_3: S_2 + 5–50% TFE	393 spectra

Table 2. *The largest five singular values w_i and σ_μ (μ = 1–5) for S_1, S_2, and S_3*

	S_1	S_2	S_3
w_1	23.9	38.2	48.1
w_2	12.6	14.1	27.9
w_3	0.86	1.18	2.18
w_4	0.50	0.61	1.27
w_5	0.37	0.45	0.52
σ_1	0.0985	0.0860	0.140
σ_2	0.0101	0.0106	0.0143
σ_3	0.0076	0.0078	0.0093
σ_4	0.0065	0.0069	0.0069
σ_5	0.0059	0.0063	0.0064

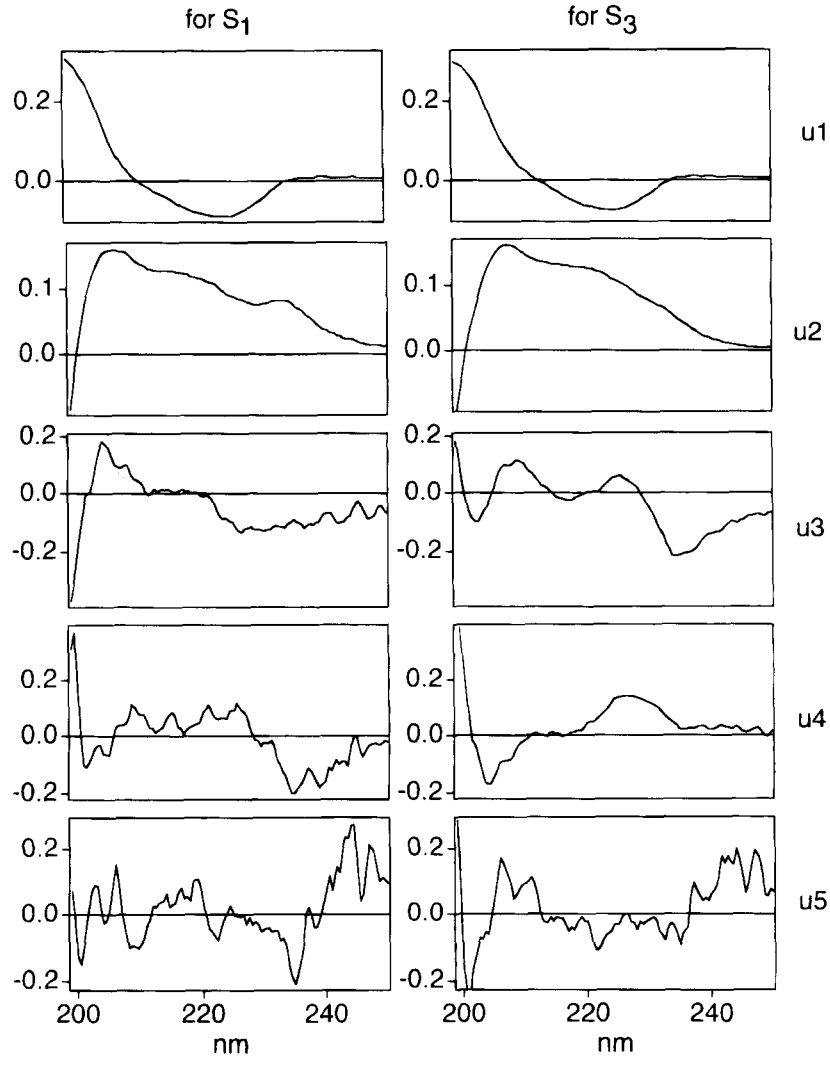

Fig. 2. The basis spectra u_1–u_5 from the SVD analysis of the data set S_1 (left column) and S_3 (right column).

nent was included in the procedure (Fig. 4C). Therefore, we conclude that the number of basis components is two for S_1 and S_2, while three for S_3.

Transformation of basis spectra

A lot of investigations in the literature have identified four typical physical species of cyt *c*, which are the N, MG, D, and H conformers (Ohgushi & Wada, 1983; Goto et al., 1990a, 1990b; Jeng & Englander, 1991; Kataoka et al., 1993; Bychkova et al., 1996; Kamatari et al., 1996). Typical solution conditions for the conformers are listed in Table 3. The difference spectra for D, MG, and H were obtained by referring to the spectrum of the N conformer. Because of the purely mathematical nature of the SVD analysis, any of the *u* basis spectra (Fig. 2) are naturally distinct from the spectra of any physical conformers. The fractional con-

Table 3. *Conditions for typical conformers of cyt c and fractional contribution from u_1–u_3 of S_3*

	u_1	u_2	u_3
D conformer (pH 1.96, water)	−5.59	2.61	0.056
MG conformer (pH 2.53, 0.5 M NaCl)	−1.62	−1.18	0.224
H conformer (pH 3.2, 50% TFE)	−2.60	−4.98	−0.207
N conformer (pH 4.50, water)		(reference to make difference spectra)	

tribution of the u_1–u_3 basis spectra for S_3 (right column of Fig. 2) to the D, MG, and H difference spectra is given in Table 3.

To extract physical meanings from the SVD analysis, it may sometimes be desirable to transform the u basis spectral set to that composed of the spectra of physical conformers. This transformation can be achieved by Equations 6 and 7, and allowed us to decompose any of the CD spectra in S_3 into the fractions of D, MG, and H-form spectra. In the next section, we will show some detailed results of this physically significant decomposition along the pH axis in comparison with the other spectroscopic data, which will support that the procedure works properly. Using the results of the decomposition, fractions of the three spectral forms (D, MG, and H) were averaged through the spectra in S_3 separately for $(NH_4)_2SO_4$, sorbitol, and TFE solutions (Fig. 5). This result clearly indicates that the D and MG forms contribute significantly to the spectra in all the co-solvent solutions while contribution from the H form is significant only for the spectra in TFE solutions. The result has confirmed that the required number of basis spectra for S_1 and S_2 (both are subsets of S_3) is only two, and has excluded the possibility that the H form plays some roles in the spectra in $(NH_4)_2SO_4$ and sorbitol solutions. On the other hand, the requirement of the third basis component in S_3 may reasonably correspond to the existence of the H-form spectra in this data set.

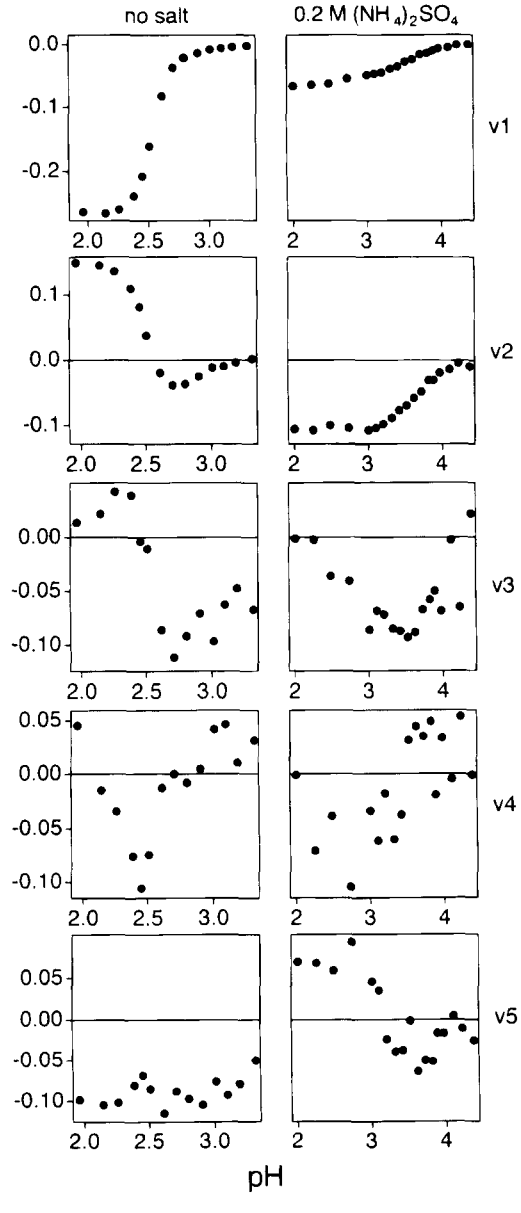

Fig. 3. The amplitude vectors v_1–v_5 from the SVD analysis of the data set S_1. v_1–v_5 in this figure correspond to the data obtained in pH denaturation experiments (large extent of denaturation at low pH on the left, to low extent of denaturation at high pH on the right of each plot) in 0 or 0.2 M $(NH_4)_2SO_4$ conditions (the left or right columns, respectively). The pH range was 3.32–1.96 for no salt, pH 4.36–2.01 for 0.2 M $(NH_4)_2SO_4$.

Some results in water and co-solvent solutions

In this section, we describe the results of the present SVD analysis in comparison with the other spectroscopic data. The data sets S_1 and S_2 were used for the analysis in $(NH_4)_2SO_4$ and sorbitol solutions, respectively, where the spectra were decomposed into the two basic spectra D and MG. The analysis for TFE solutions utilized the data set S_3 together with the D, MG, and H-form spectra. The fraction of the N form is determined by:

$$\text{(fraction of N)} = 1.0 - \text{(fraction of MG)} - \text{(fraction of D)}. \quad (1)$$

We first show results of acid denaturation of cyt c in water. The acid denaturation curves were monitored by $[\theta]_{222}$, $[\theta]_{289}$, and intensity of tryptophan fluorescence at 434 nm (F_{343}) (Fig. 6A), which show that the denaturation transition occurs at around pH 3.0. In the acid denatured state of cyt c at pH 2.0, both secondary and tertiary structures are largely destroyed (monitored by $[\theta]_{222}$ and $[\theta]_{289}$, respectively) and distance between a heme ring and a tryptophan residue is expanded (monitored by F_{343}). Slight difference in the transition zones for $[\theta]_{222}$ and $[\theta]_{289}$ at pH ~3.0 indicates the existence of an equilibrium intermediate (the MG conformer) in this middle range of pH. The fractional transition curves in water given by the SVD analysis (Fig. 6B) are consistent with the results of Figure 6A. Significant loss of the N fraction in Figure 6B by pH reduction occurs at pH less than 3.5, which is the same as the region where $[\theta]_{289}$ deviates from its level in the native state of cyt c (Fig. 6A). On the other hand, the D fraction in Figure 6B comes up at pH less than 2.8 where F_{343} and $[\theta]_{222}$ in Figure 6A start to increase. As expected from Figure 6A, the MG form in Figure 6B appears in the middle range of pH.

The acid-induced fractional transition curves by the SVD analysis in 0.1 M $(NH_4)_2SO_4$ solutions are given in Figure 6D. The fraction of N starts to decrease at pH ~4.5, which is similar to a trend monitored by $[\theta]_{289}$ (Fig. 6C). At pH lower than 2.5, the fraction of MG becomes almost 1 (Fig. 6D). This predominance of the MG form and the lack of the D fraction at low pH are also consistent with no essential change in F_{343} and $[\theta]_{222}$ in the whole pH range (Fig. 6C). This denaturation transition of cyt c from N to MG is common for cyt c in various salt solutions (Goto et al., 1990a, 1990b). Destruction of the N conformer in 0.1 M $(NH_4)_2SO_4$ (Fig. 6C,D) seems to occur at much higher pH than that in water (Fig. 6A,B), indicating that 0.1 M $(NH_4)_2SO_4$ destabilizes the N conformer. At a glance, this may conflict with the well-known protein-stabilization effect of $(NH_4)_2SO_4$. However, we should remember that the preferential hydration effect of sulfate salts is dominant only in a very high concentration range (Arakawa &

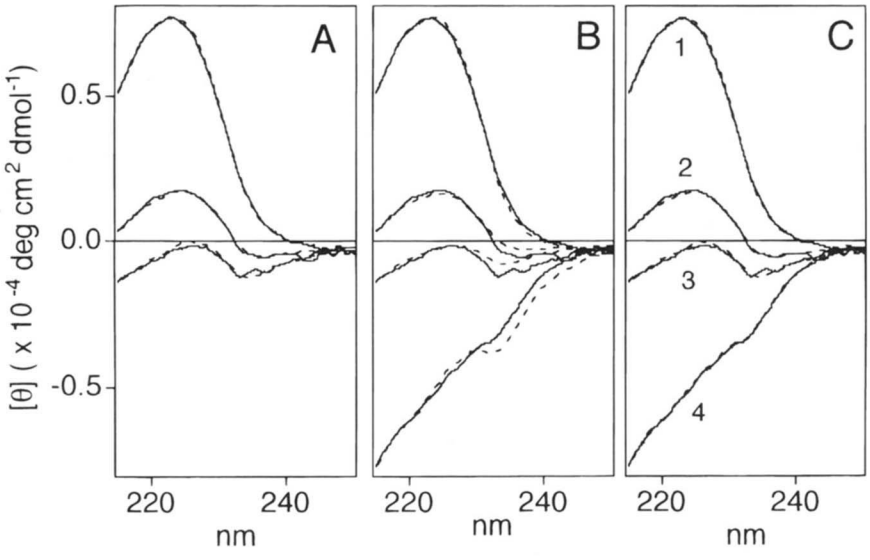

Fig. 4. Reconstruction of spectra in S_2 and S_3 from two or three components of basis spectra. **A:** S_2 using two components. **B:** S_3 using two components. **C:** S_3 using three components. Spectrum 1: pH 1.96, water. Spectrum 2: pH 2.01, 2.0 M sorbitol. Spectrum 3: pH 2.68, 0.1 M $(NH_4)_2SO_4$. Spectrum 4: pH 1.90, 50% TFE. Solid lines: experimental spectra. Dotted lines: spectra reconstructed from the first two or three components of the SVD analysis.

Timasheff, 1984; Timasheff, 1992, 1993). In the lower concentration range of $(NH_4)_2SO_4$ (typically <0.5 M), the salting-in effect is superior to the preferential hydration effect, often resulting in destabilization of the folded state of proteins.

In 2.0 M sorbitol, the pH-induced changes in $[\theta]_{222}$, $[\theta]_{289}$, and F_{343} occur significantly but to lesser extents as compared to the case of water (Fig. 6E). Because of this half-finished co-solvent effect, determination of phase properties of cyt *c* in sorbitol solutions by these simple spectroscopic probes is not very easy. On the other hand the fractional analysis by the SVD method is straightforward (Fig. 6F). The SVD results suggest that sorbitol somehow stabilizes the MG conformer compared to the D conformer at low pH (Fig. 6F). The larger $[\theta]_{222}$ and the smaller F_{343} at pH 2.0 in 2.0 M observed in sorbitol compared to water (Fig. 6E compared with Fig. 6A) are consistent with this SVD result. One of the most plausible mechanisms to stabilize the MG conformer by sorbitol is the preferential hydration effect, which is the same mechanism as that which stabilizes the N conformer of proteins (Timasheff, 1992, 1993). The effects of sorbitol support the view that an equilibrium between D and MG conformers of cyt *c* is determined by the electrostatic repulsive force in balance with the hydrophobic attractive force strengthened by the preferential hydration (Dill & Shortle, 1991; Goto & Nishikiori, 1991; Dill et al., 1995).

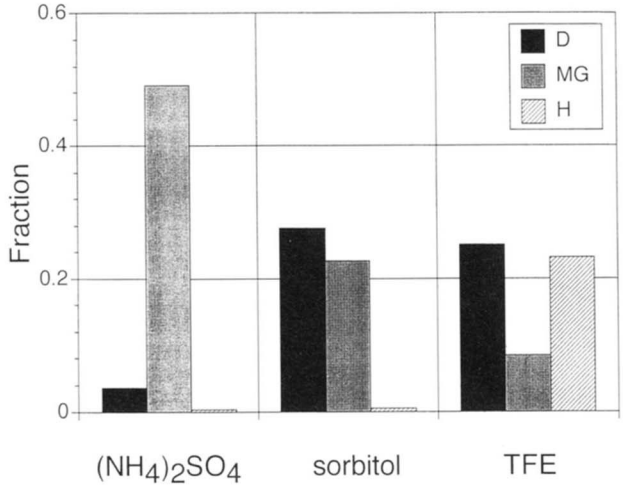

Fig. 5. Fractions of D, MG, and H forms in the spectra of different co-solvent solutions. The values were obtained by the SVD analysis of the data set S_3 (see text). Estimated fraction values of the three spectral forms (D, MG, and H) were averaged through the spectra in S_3 separately for $(NH_4)_2SO_4$, sorbitol and TFE solutions. The values in the figure were obtained by averaging 161 spectra for $(NH_4)_2SO_4$ (0–2 M, pH 4.5–1.7), 118 spectra for sorbitol (0–2 M, pH 3.42–1.79), and 128 spectra for TFE solutions (0–50%, pH 4.2–1.81).

Another typical example of the SVD analysis is shown in Figure 7 for TFE-induced denaturation at pH 3.2. In the middle ranges of TFE concentration, the MG form comes up while in the higher alcohol concentration, the H form becomes dominant. This figure qualitatively gives the same alcohol-denaturation pattern as that in methanol solutions at pH 3.0 monitored by a multi-probe experiments of Kamatari et al. (1996).

Concluding remarks

The SVD analysis of the far-UV CD spectra was able to determine the fractions of the D, MG, and H conformers of cyt *c* in the non-native solution conditions, consistently with the other spectroscopic signs (Fig. 6). In other words, the far-UV CD spectrum may contain enough information to determine the phase properties of cyt *c*. Advantages of the systematic SVD method was demonstrated in various points. For example, the SVD method can give hints for the relative importance of each physical conformer of cyt *c* in a certain co-solvent condition only using the far-UV CD spectra. Because of this advantage, we could find that the H con-

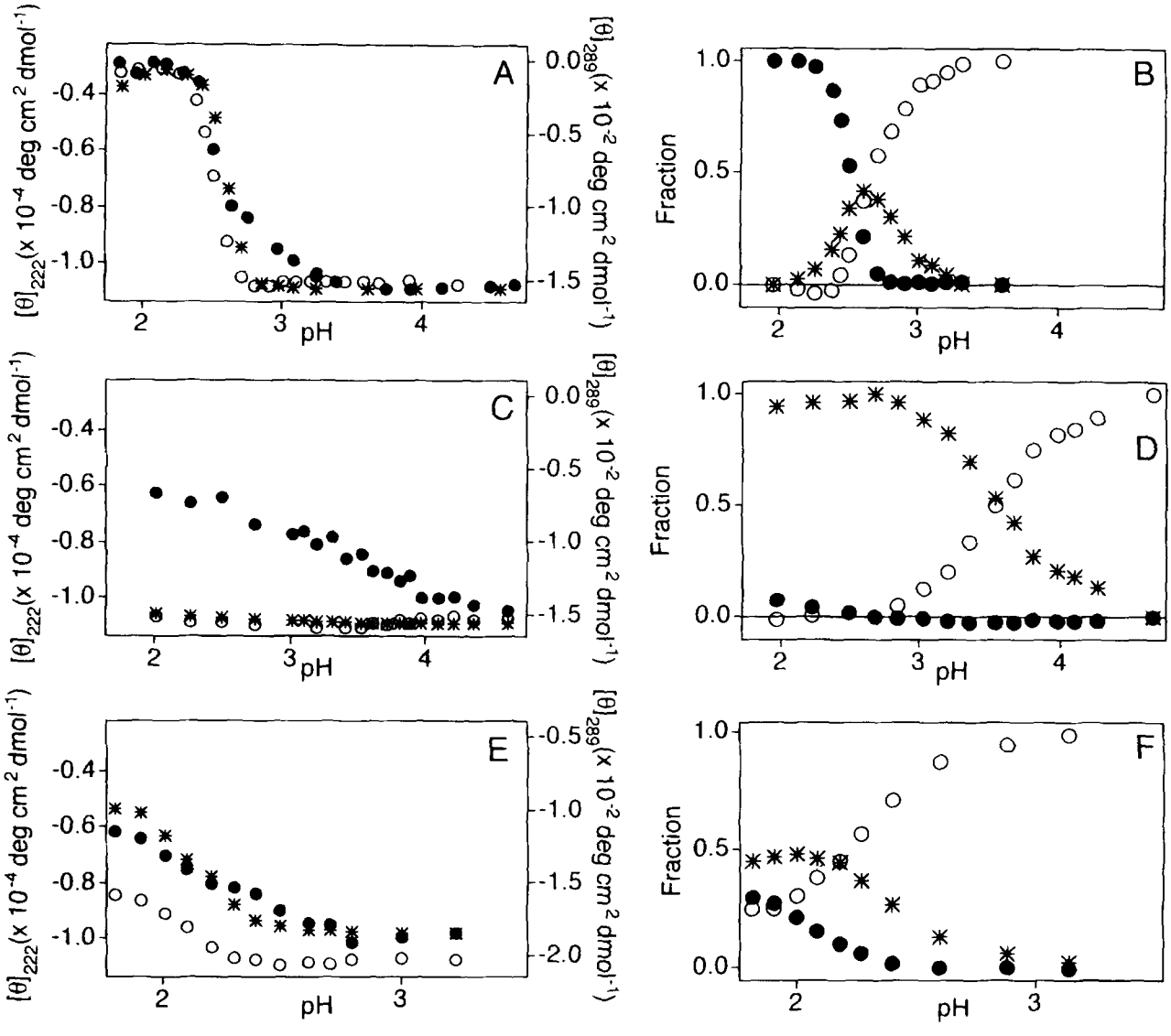

Fig. 6. **A,C,E:** Acid denaturation curves of $[\theta]_{222}$ (open circles), $[\theta]_{289}$ (filled circles), and F_{343} (*) in water. **A:** 0.1 M $(NH_4)_2SO_4$ (**C:**), and 2.0 M sorbitol (**E**). F_{343} is in an arbitrary unit. **B,D,F:** pH-induced fractional change of spectral forms determined by the SVD analysis in water (**B**), 01 M $(NH_4)_2SO_4$ (**D**), and 2.0 M sorbitol (**F**). N form (open circles), D form (closed circles), MG form (*).

former is probably absent in sorbitol or $(NH_4)_2SO_4$ solutions (Fig. 5).

In conclusion, this work has demonstrated the potency of the systematic SVD analysis of a large set of the far-UV CD spectra as a powerful method to extract information on the conformational diversity of the non-native species of protein. It is not our intention, however, to claim that the present method is exclusively better than the conventional way by a simple but complementary usage of many spectroscopic probes. A disadvantage of the far-UV CD spectroscopy resides in the difficulty in extracting information about any specific tertiary interactions in a protein chain. Accordingly, a limitation of the present method is clearly in the detailed characterization of any specific conformations of a protein. Because the conformational properties of the non-native state of proteins are extremely complicated, one should always consider using all the possible experimental methods. We expect that combination

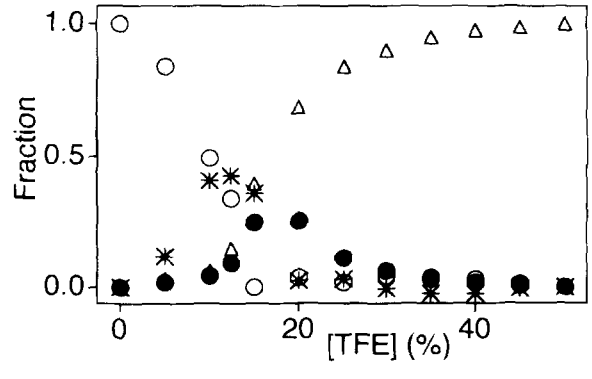

Fig. 7. Fraction of four spectral forms estimated by the SVD analysis as functions of TFE concentration at pH 3.2. N form (open circles), D form (closed circles), MG form (*) and H form (open triangle).

of the present method with the multi-probe approaches will provide an improved way to detect the conformational diversities of non-native proteins.

Materials and methods

Materials

Horse heart cyt *c* (type VI; Sigma) was purified by gel-filtration chromatography (Sephadex G-25; Pharmacia). The homogeneity of the purified protein was checked by electrophoresis. Oxidization with potassium ferricyanide ($K_3Fe(CN)_6$) was not done to avoid contamination of $K_3Fe(CN)_6$ to protein samples. No spectroscopic sign detected any fractions of the ferrous-form cyt *c* in our samples. Sorbitol, 2,2,2-trifluoroethanol (TFE), sodium chloride, and ammonium sulfate were of reagent grade and were purchased from Nacalai tesque Co. Ltd. (Japan). Co-solvent–containing solutions were freshly prepared for each experiment with a volumetric adjustment of concentrations. The cyt *c* solutions contain no buffer, pH being adjusted carefully with small amounts of HCl. The concentrations of cyt *c* were determined photometrically, using an extinction coefficient $E_{409nm} = 1.06 \times 10^5\ M^{-1}\ cm^{-1}$ (Babul & Stellwagen, 1972), and was kept at 16 μM throughout.

Spectroscopic measurements

CD spectra were measured with a Jasco J-720 WI spectropolarimeter using a quartz cell with a light path of 1 or 10 mm. Temperature of the solutions was maintained at 25 °C with a Jasco thermo-electric apparatus within ±0.1 °C. The spectra of cyt *c* were typically obtained with scan rates of 10, 20, 50, and 100 nm/min. Essentially the same spectra were obtained with either of these scan rates in our sample solutions. The spectra taken with a scan rate of 50 nm/min were used for the data sets in Table 1. The data were sampled in 0.5 nm interval. Data from 16 scans were accumulated to obtain a single spectrum. The bandwidth for the measurements was 1 nm. Fluorescence spectra were measured with a F4500 fluorometer (Hitachi, Japan) using a quartz cell with a light path of 10 mm. Temperature of solutions was maintained at 25 °C, with a thermostatically controlled water bath within ±0.1 °C.

SVD analysis of CD spectra

The matrix method has been used successfully in spectroscopic studies of DNA and proteins (McMullen et al., 1967; Hennessey & Johnson, 1981; Henry & Hofrichter, 1992). The singular value decomposition (SVD) analysis has been examined in the present study because this method is known to be robust and numerically stable among some mathematically equivalent strategies (Henry & Hofrichter, 1992). In its application here, we started with a $M \times N$ data matrix, A, which contains a far-UV CD spectrum of cyt *c* in each of N columns. M is the number of data points in a single spectrum. M and N are typically 101 and 200–400, respectively. In case $M < N$, M is increased to N by filling zero to the $(N–M)$ additional rows.

The analysis of A is based on the SVD theorem of linear algebra;

$$A = UWV^{T}, \tag{2}$$

where U is an $M \times N$ column-orthogonal matrix, and V is an $N \times N$ orthogonal matrix. U and V fulfill the condition;

$$U^{T}U = V^{T}V = I, \tag{3}$$

where I is the $N \times N$ identity matrix. W is a diagonal matrix with positive or zero elements (singular values). Columns of U and V are the left and right singular vectors of A, respectively. The singular values, denoted by w_k ($k = 1, \ldots, N$), were ordered (along with the corresponding columns of U and V) so that $w_1 \geq w_2 \geq \ldots \geq w_N \geq 0$. The linear combination composed of the first k columns of U, along with the corresponding columns or elements of V and W, provides the best least-squares approximation to the matrix A having a rank k by

$$A \approx U'W'(V')^{T} \tag{4}$$

where U', W', and V' are submatrices of U, W, and V, respectively, corresponding to the first k basis components.

The relative importance of each column of U, as a component of the original protein CD spectra, can be estimated by the corresponding eigenvalue w_i. Within the frame work of the eigenvector method of multicomponent analysis, the variance in the data set unaccounted for by the μ most important basis CD spectra is given by

$$\sigma_{\mu}^{2} = \frac{1}{M(N-\mu)} \sum_{i=\mu+1}^{N} w_i^2. \tag{5}$$

Here σ is the standard deviation, M is the number of data points in each original vector, μ is the number of basis vectors in the truncated set used for reconstruction. In our present application, it is important to evaluate the minimum number of μ required for satisfactory reconstruction of all the original spectra. For the purpose, not only the σ value but also shape of the column vectors of U and V have been considered.

The basis spectra of columns of U do not, however, correspond to the spectra of any physical conformers of cyt *c*. A submatrix U' composed of the first k basis spectra is transformed to a $M \times k$ matrix, P, composed of k physically significant spectra by:

$$P = U'W'Z, \tag{6}$$

where Z is a $k \times k$ transformation matrix. The easiest way to calculate elements of the Z matrix is to add the k spectra for the physical conformers to the original data matrix A. In this case, Z^{T} is a submatrix of V composed of column vectors corresponding to these physically significant spectra. Now Equation 4 is rewritten in the form:

$$A = PZ^{-1}(V')^{T}. \tag{7}$$

Elements of the matrix $Z^{-1}(V')^{T}$ give estimates for contribution of each spectrum of the physical conformers to any spectra in A.

Numerical calculations for Equations 2–7 were performed on a VT-Alpha 500AXP workstation (Visual Technology, Japan). The program was originally made by the author using the C program language, partly using subroutines in *Numerical Recipes* (Press et al., 1986). The results of SVD were always tested by inserting the estimates of U, V, and W into Equations 2 and 3.

Acknowledgments

We thank O. Ichikawa for his encouragement and Dr. K. Imoto for his kind help on biochemical experiments throughout this study. We would like to

acknowledge to Dr. K. Nagayama, Dr. K. Kanaori, Dr. Y. O. Kamatari, and Dr. N. Tanaka for stimulating discussions with the author.

References

Alexandrescu AT, Ng YL, Dobson, CM. 1994. Characterization of a trifluoroethanol-induced partially folded state of α-lactalbumin. *J Mol Biol 235*:587–599.

Ansari A, Jones CM, Henry ER, Hofrichter J, Eaton WA. 1994. Conformational relaxation and ligand binding in myoglobin. *Biochemistry 33*:5128–5145.

Arakawa T, Timasheff SN. 1984. Mechanism of protein salting in and salting out by divalent cation salts: Balance between hydration and salt binding. *Biochemistry 23*:5912–5923.

Babul J, Stellwagen E. 1972. Participation of the protein ligands in the folding of cytochrome *c*. *Biochemistry 11*:1195–1200.

Buck M, Radford SE, Dobson CM. 1993. A partially folded state of hen egg white lysozyme in trifluoroethanol: Structural characterization and implications for protein folding. *Biochemistry 32*:667–678.

Bychkova VE, Dujsekina AE, Klenin SI, Tiktopulo EI, Uversky VN, Ptitsyn OB. 1996. Molten globule-like state of cytochrome *c* under conditions simulating those near the membrane surface. *Biochemistry 35*:6058–6063.

Chen L, Hodgson KO, Doniach S. 1996. A lysozyme folding intermediate revealed by solution X-ray scattering. *J Mol Biol 261*:658–671.

Conio G, Patrone E, Brighetti S. 1970. The effect of aliphatic alcohols on the helix-coil transition of poly-l-ornithine and poly-l-glutamic acid. *J Biol Chem 245*:3335–3340.

Dill KA, Shortle D. 1991. Denatured states of proteins. *Annu Rev Biochem 60*:795–825.

Dill KA, Bromberg S, Yue K, Fiebig KM, Yee DP, Thomas PD, Chan HS. 1995. Principles of protein folding—A perspective from simple exact models. *Protein Sci 4*:561–602.

Dill KA, Chan HS. 1997. From Levinthal to pathways to funnels. *Nat Struct Biol 4*:10–19.

Dobson, CM. 1992. Unfolded proteins, compact states and molten globules. *Curr Opin Struct Biol 2*:6–12.

Dyson HJ, Merutka G, Waltho JP, Lerner RA, Wright PE. 1992. Folding of peptide fragments comprising the complete sequence of proteins. Models for initiation of protein folding. I. Myohemerythrin. *J Mol Biol 226*:795–817.

Fan P, Bracken AK, Baum J. 1993. Structural characterization of monellin in the alcohol-denatured state by NMR: Evidence for β-sheet to α-helix conversion. *Biochemistry 32*:1573–1583.

Goto Y, Calciano LJ, Fink AL. 1990a. Acid-induced folding of proteins. *Proc Natl Acad Sci USA 87*:573–577.

Goto Y, Takahashi N, Fink AL. 1990b. Mechanism of acid-induced folding of proteins. *Biochemistry 29*:3480–3488.

Goto Y, Nishikiori S. 1991. Role of electrostatic repulsion in the acid molten globule of cytochrome *c*. *J Mol Biol 222*:679–686.

Hennessey JP Jr, Johnson WC Jr. 1981. Information content in the circular dichroism of protein. *Biochemistry 20*:1085–1094.

Henry ER, Hofrichter J. 1992. Singular value decomposition: Application to analysis of experimental data. *Methods Enzymol 210*:129–192.

Hofrichter J, Sommer JH, Henry ER, Eaton WA. 1983. Nanosecond absorption spectroscopy of hemoglobin: Elementary processes in kinetic cooperativity. *Proc Natl Acad Sci USA 80*:2235–2239.

Hughson FM, Write PE, Baldwin RL. 1990. Structural characterization of a partly folded apomyoglobin intermediate. *Science 249*:1544–1548.

Jager S, Lewis JW, Zvyaga TA, Szundi I, Sakmar TP, Kliger DS. 1997. Time-resolved spectroscopy of the early photolysis intermediates of rhodopsin Schiff base counterion mutants. *Biochemistry 36*:1999–2009.

Jeng MF, Englander SW. 1991. Stable submolecular folding units in non-compact form of cytochrome *c*. *J Mol Biol 221*:1045–1061.

Jones CM, Henry ER, Hu Y, Chan C-K, Luck SD, Bhuyan A, Roder H, Hofrichter J, Eaton WA. 1993. Fast events in protein folding initiated by nanosecond laser photolysis. *Proc Natl Acad Sci USA 90*:1860–11864.

Kamatari YO, Konno T, Kataoka M, Akasaka K. 1996. Methanol-induced compact and expanded denatured conformations of cytochrome *c*. *J Mol Biol 259*:512–523.

Kataoka M, Hagihara Y, Mihara K, Goto Y. 1993. Molten globule of cytochrome *c* studied by small angle x-ray scattering. *J Mol Biol 229*:591–596.

Kippen AD, Sancho J, Fersht AR. 1994. Folding of barnase in parts. *Biochemistry 33*:3778–3786.

Konno T, Kataoka M, Kamatari Y, Kanaori K, Nosaka A, Akasaka K. 1995. Solution x-ray scattering analysis of cold-, heat-, and urea-denatured states of a protein, *Streptomyces* subtilisin inhibitor. *J Mol Biol 251*:95–103.

Kuwajima K. 1989. The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins Struct Funct Genet 6*:87–103.

McMullen DW, Jaskunas SR, Tinoco I Jr. 1967. Application of matrix rank analysis to the optical rotatory dispersion of TMV RNA. *Biopolymers 5*:589–613.

McPhie P, Shrager RI. 1992. An investigation of the thermal unfolding of swine pepsinogen using circular dichroism. *Arch Biochem Biophys 293*:46–53.

Nelson JW, Kallenbach NR. 1986. Stabilization of the ribonuclease S-peptide α-helix by trifluoroethanol. *Proteins Struct Funct Genet 1*:211–217.

Ohgushi M, Wada A. 1983. 'Molten-globule state': A compact form of globular proteins with mobile side-chains. *FEBS Lett 164*:21–24.

Press WH, Flannery BP, Teukolsky SA, Vetteling WT. 1986. *Numerical Recipes: The art of scientific computing*. Cambridge: Cambridge Univ. Press.

Ptitsyn OB. 1995. Structures of folding intermediates. *Curr Opin Struct Biol 5*:74–78.

Roder H, Elöve GA, Englander SW. 1988. Structural characterization of folding intermediates in cytochrome *c* by H-exchange labeling and proton NMR. *Nature 355*:700–704.

Segawa S, Fukuno T, Fujiwara K, Noda Y. 1991. Local structure in unfolded lysozyme and correlation with secondary structures in the native conformation: Helix-forming or -breaking propensity of peptide segments. *Biopolymers 31*:497–509.

Shiraki K, Nishikawa K, Goto Y. 1995. Trifluoroethanol-induced stabilization of the α-helical structure of β-lactoglobulin: Implication for non-hierarchical protein folding. *J Mol Biol 245*:180–194.

Sosnick TR, Trewhella J. 1992. Denatured states of ribonuclease A have compact dimensions and residual secondary structure. *Biochemistry 31*:8329–8335.

Sreerama N, Woody RW. 1994. Protein secondary structure from circular dichroism spectroscopy. Combining variable selection principle and cluster analysis with neural network, ridge regression and self-consistent methods. *J Mol Biol 242*:497–507.

Tanford C. 1968. Protein denaturation. *Adv Protein Chem 23*:121–217.

Timasheff SN. 1992. Water as ligand: Preferential binding and exclusion of denaturants in protein unfolding. *Biochemistry 31*:9857–9864.

Timasheff SN. 1993. The control of protein stability and association by weak interactions with water: How do solvents affect these processes? *Annu Rev Biophys Biomol Struct 22*:67–97.

Wolynes PG, Onuchic JN, Thirumalai. 1995. Navigating the folding routes. *Science 267*:1619–1620.

Woody RW. 1995. Circular dichroism. *Methods Enzymol 246*:34–71.

Yang J-T, Wu C-SC, Martinez HM. 1986. Calculation of protein conformation from circular dichroism. *Methods Enzymol 130*:208–269.