Journal of Chromatography, 540 (1991) 187–198
Elsevier Science Publishers B.V., Amsterdam

CHROM. 22 992

Monitoring of protein conformation by high-performance size-exclusion liquid chromatography and scanning diode array second-derivative UV absorption spectroscopy

CATHLEEN E. ACKLAND, WILLIAM G. BERNDT, JOAN E. FREZZA, BRYAN E. LANDGRAF, KATHERINE W. PRITCHARD and THOMAS L. CIARDELLI*

Department of Pharmacology, Dartmouth Medical School, Hanover, NH 03756 (U.S.A.)

(First received April 10th, 1990; revised manuscript received November 2nd, 1990)

ABSTRACT

Genetic methods now allow the rapid production of mutant proteins for structure–function analysis. To properly interpret any change in biologic activity resulting from modification in primary sequence, it is essential to monitor conformational changes resulting from mutations. Several methods allow low-resolution protein conformational analysis. One method, second-derivative UV absorption spectroscopy, is particularly useful for proteins containing tyrosine and/or tryptophan residues.

Using high-performance size-exclusion liquid chromatography and scanning diode array detection we have demonstrated that it is possible to monitor the degree of aggregation as well as conformational perturbation for a series of interleukin-2 structural mutants. Furthermore, the combination of high-performance liquid chromatography and second-derivative UV absorption spectroscopy avoids a potential artifactual contribution in non-chromatographic analysis due to protein aggregation.

INTRODUCTION

During the past ten years, biochemists have benefited enormously from the application of high-performance liquid chromatography (HPLC) methodology to the separation of proteins. Stationary phases useful for protein separations have been combined with state-of-the-art HPLC instrumentation resulting in increased sensitivity, reproducibility and a considerable savings in time per analysis [1].

Instrumentation and column development has proceeded rapidly, however at the same time, cloning, sequencing and expression of proteins has greatly outpaced the ability of biochemists to perform structure-function analysis. The methods of site-directed [2] and cassette [3] mutagenesis have made single or multiple amino acid residue replacements routine. Thus, even when structure-function studies are limited to a single protein, it is possible to generate hundreds of mutant sequences for analysis.

In order to correctly interpret the consequences of sequence alteration on biologic function in most mutagenesis experiments, it is necessary to distinguish those mutations which directly affect active sites or receptor binding regions from those which indirectly perturb these regions via alteration of conformation. In extreme cases,

0021-9673/91/$03.50 © 1991 Elsevier Science Publishers B.V.

even single amino acid substitutions can have profound effects on folding and stability of proteins [4].

Although high-resolution conformational analysis of proteins requires exacting techniques such as X-ray diffraction or nuclear magnetic resonance, several methods are available which provide lower resolution information. Fluorescence methods, circular dichroism and UV spectroscopy have all proven extremely valuable for protein structural anaysis. These methods are particularly useful if the goal is not the determination of the precise protein structure, but rather, if and how that structure is changed upon mutagenesis.

Derivative UV spectroscopy, because of its ability to reduce overlap of the spectral contributions of aromatic chromophores [5,6], has been successfully applied in the determination of conformationally dependent tyrosine exposure in proteins [7] as well as for monitoring of folding characteristics in mutant proteins [8]. Second-derivative UV spectroscopy has been interfaced with reversed-phase HPLC by using scanning diode array detection to identify the presence or absence of aromatic residues in tryptic digests of proteins prior to microsequencing [9]. This technique can distinguish and identify peptides and proteins differing in aromatic amino acid composition.

In this report we demonstrate that under the appropriate conditions, HPLC with scanning diode array detection and second-derivative UV spectroscopy can also be employed for monitoring conformation of a series of mutant proteins with the same aromatic amino acid composition. Previous reports have also demonstrated that second-derivative UV spectroscopy after hydrophobic interaction chromatography can detect altered protein conformation [10,11], however this technique provides no information on the aggregation state of the protein and its influence on conformation. We have found that the combination of HPLC gel filtration with second derivative spectroscopy can provide an unambiguous indication of the influence of mutations on protein conformation within the limitations of this technique and may often be superior than non-chromatographic spectroscopy alone.

MATERIALS AND METHODS

Chemicals and reagents

Sodium phosphate monobasic, ACS grade was purchased from Fisher Scientific. Guanidine · HCl grade 1 (99.5%) and lysozyme (chicken egg white, grade 1) were purchased from Sigma. Recombinant interleukin-2 and interleukin-2 mutants were prepared and purified as described [12]. Each protein was judged homogeneous by C_{18} and C_3 reversed-phase chromatography and by amino acid analysis (Pico-Tag, Waters). Water was purified by Milli-Q filtration (Millipore) followed by glass distillation.

High-performance liquid chromatography

Separations were performed using a Hewlett-Packard 1090 M HPLC system equipped with a binary pumping system, an automated injection and sampling system, a heated column compartment and a diode array detector and controlled by an HP79994A ChemStation.

All samples were run on a Protein-Pak 125 (Waters) size separation column in

aqueous sodium phosphate (25 mM, pH 4.5) or for lysozyme, phosphate buffer (50 mM, pH 7.0)/KCl (150 mM) using a flow-rate of 0.5 ml/min.

Data collection

UV spectra were acquired on all samples every 2.8 s over the range of 250 to 320 nm. Chromatography was simultaneously monitored at 215, 254 and 280 nm with a reference at 550 nm. Sufficient protein was applied to the column to obtain peaks generating signals greater than 50 mA units at 280 nm for all components present. All data were stored on flexible disks.

Data processing

To obtain a zero-order UV spectrum representative of each peak, individual spectra taken in the region where absorption at 280 nm was greater than 50 mA units were averaged across the region using an averaging function provided with the manufacturers software. Derivative analysis of a given spectrum was performed directly on the unsmoothed data by translocation of a spectrum first 1 nm towards higher wavelength to generate a red spectrum, then 1 nm towards lower wavelength to generate a blue spectrum. Subtraction of the red spectrum from the blue spectrum generated the corresponding derivative spectrum of $\Delta\lambda = 2$ nm. Repeating this process provided second-derivative spectra. The manipulations described were carried out using command language provided in the manufacturers software. It should be noted that the derivative function also provided in the manufacturers software was not adequate for generating second derivative spectra useful for conformational analysis because of the inclusion of a smoothing step which altered fine spectral features.

Circular dichroism

Near-UV circular dichroism (CD) data (250–300 nm) were collected on a Jobin Yvon Mark 5 circular dichrograph calibrated with (+)-10-camphorsulfuric acid using a 1-cm cell. Values were determined every 0.5 nm with a response time of 10 s. Each spectrum represented an average of 3 scans and was corrected for blank cell absorption. Protein samples were run in sodium phosphate (25 mM, pH 4.5) and concentrations were determined by quantitative amino acid analysis from aliquots taken directly from the cell.

RESULTS AND DISCUSSION

Protein preparation

In our investigation of the structure-function relationships of the T-cell growth factor, interleukin-2 (IL-2) [13], we have prepared several IL-2 analogues via cassette mutagenesis. Many of these proteins have amino acid residue changes designed to perturb tertiary conformation [12].

In order to examine the extent of conformational perturbation and the ability of these mutants to refold prior to analysis in a battery of biologic assays, the proteins must first be extracted and purified from the host bacteria. As is the case with most recombinant proteins produced in highly efficient bacterial expression systems, they are stored within the organisms in a reduced, denatured and insoluble form referred to as inclusion bodies. To obtain native-like protein it is necessary to solubilize and refold

the inclusion body protein [14]. Depending upon the conditions employed, this process can lead to the production of protein aggregates. The tendency to aggregate and the nature of the aggregates formed can vary greatly with structural mutations.

In a portion of out IL-2 study we have focused on mutations in a C-terminal helical region, encompassing amino acid residues 117 to 133 of the 133 residue protein. The nature of the sequence modifications for 4 of the mutations is illustrated in Fig. 1. Mutant 1 contained substitutions designed to stabilize helical structure, while mutants 2–4 incorporated prolines within this region to perturb the helical conformation [15]. The mutant proteins were produced via cassette mutagenesis in a custom IL-2 gene and expressed in *Escherichia coli* [12]. After purification to homogeneity as determined by C_3 and C_{18} reversed-phase HPLC, the proteins (upon gel filtration chromatography) were observed to contain significant amounts of aggregates. Since we, as well as others [16] have employed non-chromatographic second-derivative UV absorption to monitor the conformation of IL-2 mutants, we chose to determine whether diode array UV scanning of aggregated and monomeric protein peaks emerging from an HPLC size-exclusion column could be utilized directly to monitor conformation. However, before analysis of IL-2 mutants we established the applicability of the technique using lysozyme, a protein which has been previously studied by second-derivative UV absorption [7].

				120										130			
IL-2	F	L	N	R	W	I	T	F	C	Q	S	I	I	S	T	L	T
MUTANT 1			Q	K			Q		A		A	L	L	Q	A		K
MUTANT 2			P						A		P						
MUTANT 3			P						A								
MUTANT 4									A		P						

Fig. 1. The amino acid sequence (single letter abbreviation) of the C-terminal α-helix of IL-2 (residues 117–133) and IL-2 mutants 1–4. For the mutant sequences, only the modified positions are indicated.

Gel filtration and second derivative UV analysis

Analysis of the second derivative of the UV absorption spectrum in the aromatic region (250–320 nm) can provide conformational information for proteins. As illustrated in Fig. 2, by determining peak-to-trough values between 283 and 287 nm (a) and between 290.5 and 295 nm (b), the environment of Tyr and Trp residues of a protein can be analyzed.

The relative exposure, α, of Tyr residues to solvent in the native protein can be determined from the ratio of these values, r, (where $r = a/b$) using the following equation [7]:

$$\alpha = (r_n - r_a)/(r_u - r_a) \quad (1)$$

where r_n and r_u represent the ratio of a/b for the native and denatured proteins, respectively and r_a represents the a/b value for a mixture of aromatic amino acids having the same molar ratio as the protein of interest determined in a solvent which is

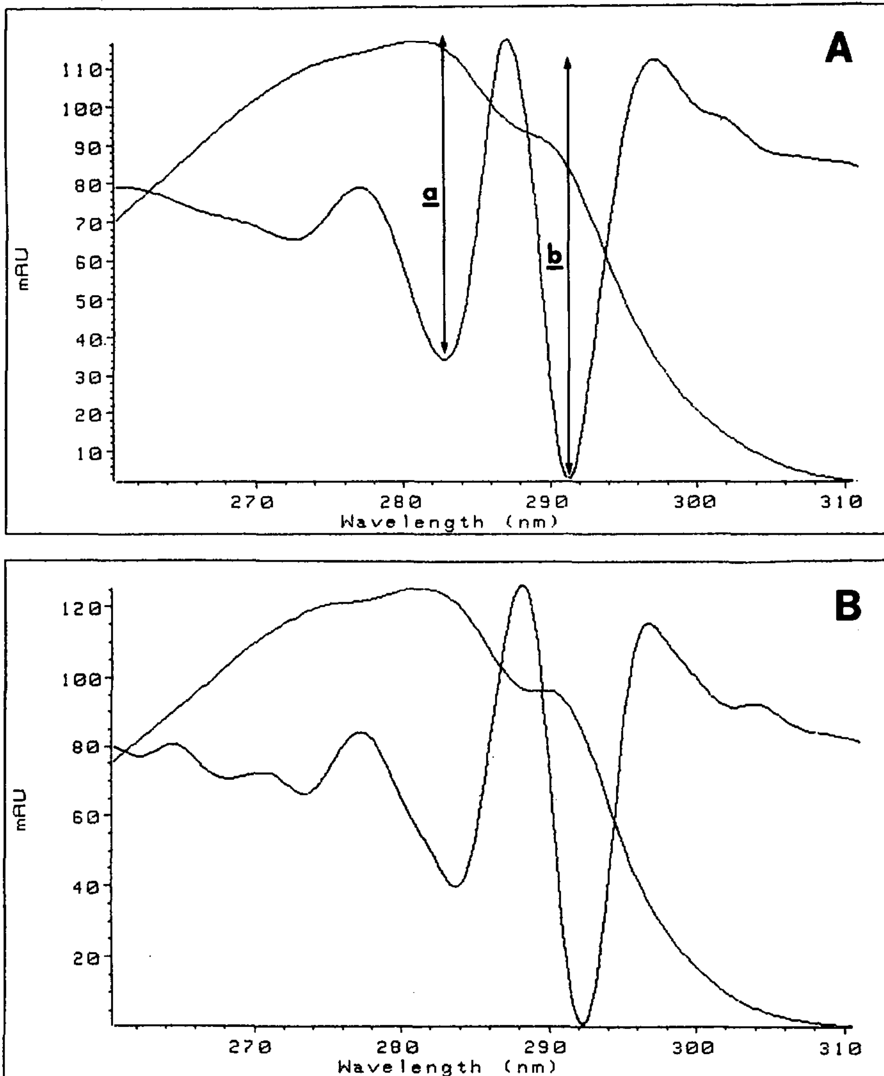

Fig. 2. Zero-order and second-derivative UV absorption spectra of (A) denatured (6 *M* guanidine · HCl in 50 m*M* sodium phosphate, pH 7.0/150 m*M* KCl) and (B) native (50 m*M* sodium phosphate, pH 7.0/150 m*M* KCl) lysozyme. Spectrum A was determined as described in Materials and Methods with the column removed from the instrument while spectrum B was obtained from the monomeric protein peak (retention volume 8.45 ml) emerging from the size-separation column. The trough to peak distances *a* (283–287 nm) and *b* (290.5–295 nm) used to calculate ratio *r* are illustrated in (A).

characteristic of the interior of a protein. The value of r_a has been determined for various ratios of Tyr and Trp in ethylene glycol [7] and is independent of the protein being studied. Likewise, r_u, the value of a/b for denatured proteins is dependent only upon the Tyr/Trp ratio in the protein and the denaturant and is equivalent for all proteins having the same Tyr/Trp ratios in the same denaturant [6]. This value,

however, may also be determined directly, for example, from the second-derivative UV absorption spectrum taken on a sample of the protein denatured in 6.0 *M* guanidine · HCl.

Therefore, for lysozyme we calculated the relative exposure of Tyr residues from second-derivative analysis of the UV absorption spectra obtained from peaks obtained by size-exclusion HPLC (Fig. 2). We obtained values of r_n = 0.72 and r_u = 0.76 which corresponded well with those determined non-chromatographically (r_n = 0.74, r_u = 0.79, ref. 7). Likewise, using the value of r_a = 0.56 we obtained a fractional Tyr exposure of 0.80 which compared favorably with the reported value of 0.78 [7]. It should be noted that the spectrum for lysozyme in 6 *M* guanidine · HCl was obtained by removing the column and equilibrating the instrument with the guanidine · HCl containing buffer before injection of a protein sample in the same buffer directly into the detector using the normal chromatographic flow rate. Since r_u is invariant as long as the Tyr/Trp ratio remains the same and aggregation is unlikely to occur in 6 *M* guanidine · HCl, it is not necessary to determine this value chromatographically.

Unlike Tyr, the relative exposure of Trp to solvent does not influence the a/b ratio [7]. However, the magnitude of the trough to peak distance b (290.5–295 nm), due principally to Trp, is reflective of the polarity of the environment of the indole chromophore and has been used to monitor folding dependent conformational changes [17].

Since IL-2 contains 3 Tyr residues and 1 Trp (at positions 31, 45, 107 and 121, respectively), we examined the influence of the mutations shown in Fig. 1 on the second-derivative spectra obtained from proteins after HPLC size separation. Fig. 3 shows both the zero-order and second-derivative spectrum for IL-2 in denatured and native form. From these spectra and using a value for r_a = 0.11 (as determined by Ragone *et al.* [7]) for a Tyr/Trp ratio = 3, we determined r_n = 1.14, r_u = 1.67 and α = 0.68 according to eqn. 1.

These values can be compared to those previously reported for IL-2 using non-chromatographic second-derivative UV spectroscopy (r_n = 1.04, r_u = 1.71 and α = 0.58, [16]. Although the r_u values are similar, the r_n values differ slightly. This could be due to a difference in buffers (the authors employed 100 m*M* sodium acetate, pH 4.6) or the presence of aggregates in their sample. Their analysis was carried out at a protein concentration of 10 μM, however, recent studies [18] have indicated that IL-2 self associates with an apparent dissociation constant (K_d) of 0.6 μM and this self association is accompanied by conformational adjustments which affect Trp fluorescence. Self association would be expected to similarly affect the second derivative UV spectrum and examination of the second derivative spectra of mutants 1–4 after size-separation HPLC support this conclusion.

Fig. 4 shows a composite of the second-derivative spectra obtained for mutants 1–4 compared to natural IL-2. All spectra in Fig. 4 were obtained from monomeric species except for mutant 2 which could only be isolated in aggregated form. Although the presence and quantity of protein aggregates was mutant as well as preparation dependent, monomeric forms always eluted with a retention volume of 8.0 ± 0.05 ml, consistent with the molecular weight of IL-2 (15 500 dalton) and previous column calibration (lysozyme, mol.wt. 14 300 dalton, eluted at 8.45 ml).

Table I lists the values of a, b, r_n and α for IL-2 and mutants 1–4. Mutant 2,

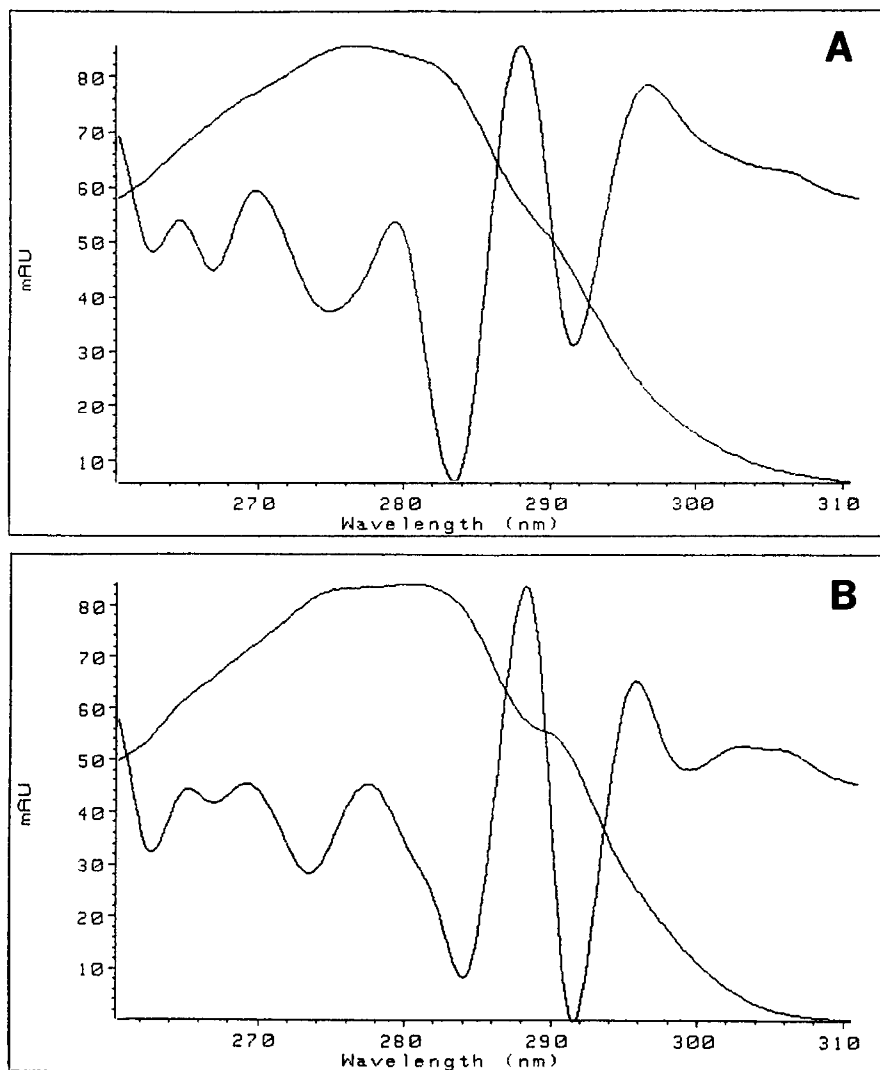

Fig. 3. Zero-order and second-derivative UV absorption spectra of (A) denatured (6 *M* guanidine · HCl in 25 m*M* sodium phosphate, pH 4.5) and (B) native (25 n*M* sodium phosphate, pH 4.5) recombinant IL-2. Spectrum A was determined as described in Materials and Methods with the column removed from the instrument while spectrum B was obtained from the monomeric protein peak (retention volume 8.0 ml) emerging from the size-exclusion column.

having two Pro residues in the C-terminal helix, could only be isolated as an aggregate (retention volume 6.1 ml, approximately 35 000 dalton; the retention volume of glycerol dehydrogenase, 36 100 dalton, was 6.2 ml) and displayed significantly different r_n and α values. Furthermore, when a sample was found to contain both aggregated and monomeric protein, the second-derivative spectra and resulting r_n and α values for the aggregated peak were distinctly different (Fig. 5).

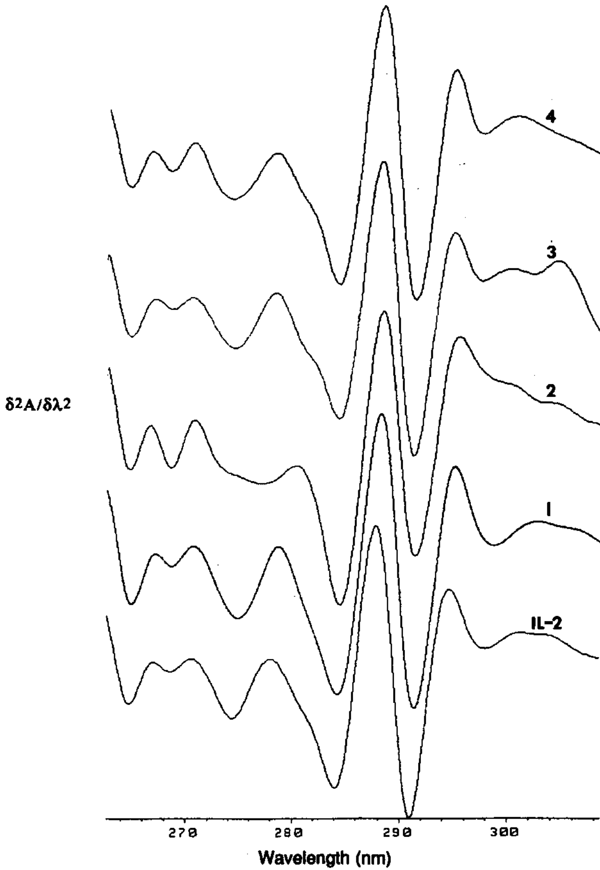

Fig. 4. Composite of the second-derivative UV absorption spectra obtained after HPLC size-exclusion chromatography for native IL-2 and mutants 1–4. All spectra were obtained from protein monomer peaks except for mutant 2 which could only be isolated in aggregated form (retention volume 6.1 ml). Conditions are described in Materials and Methods.

Finally, the parameter *b*, when normalized to equivalent protein concentration, suggested that the polarity of the Trp environment was affected by the nature of the mutations introduced. This was consistent with the fact that the single Trp residue resides at position 121, within the helix investigated. Again with mutant 2, the value for *b* was the lowest observed (except for denatured IL-2 itself) indicating increased polarity of the indole environment (a decrease in *b* is also observed in the aggregate spectrum in Fig. 5). This is consistent with the increased quenchability of Trp fluorescence upon aggregation observed with IL-2 itself [18]. Furthermore, when compared with the near-UV circular dichroism signal at 295 nm (a monitor of the asymmetry of the Trp environment) (Table II), the same trend was observed as for parameter *b*. For the monomeric proteins, mutant 3 displayed the smallest dichroic signals indicating the least asymmetry, while mutants 1 and 4 gave [θ] values (mean residue ellipticity) similar in magnitude but opposite in sign. This suggests that

TABLE I

RESULTS OF SECOND-DERIVATIVE UV ABSORPTION ANALYSIS FOR IL-2 AND MUTANTS 1–4

Protein	a^a	b^a	$r_n{}^b$	α^c
IL-2[d]	5.8	3.5	$r_u = 1.67 \pm 0.02$ (2)	
IL-2	9.0	8.0	1.14 ± 0.03 (4)	0.68
1	7.8	6.7	1.16 ± 0.01 (4)	0.69
2[e]	6.4	4.8	1.33 ± 0.06 (3)	0.79
3	7.1	6.2	1.15 ± 0.01 (3)	0.69
4	9.4	7.8	1.21 ± 0.03 (3)	0.73

[a] Normalized to equivalent protein concentrations.

[b] Values are the mean of the number of independent determinations indicated in ( ) ± the standard error of the mean.

[c] Determined from eqn. 1.

[d] IL-2 in 6 *M* guanidine hydrochloride/25 m*M* phosphate, pH 4.5.

[e] Aggregated protein.

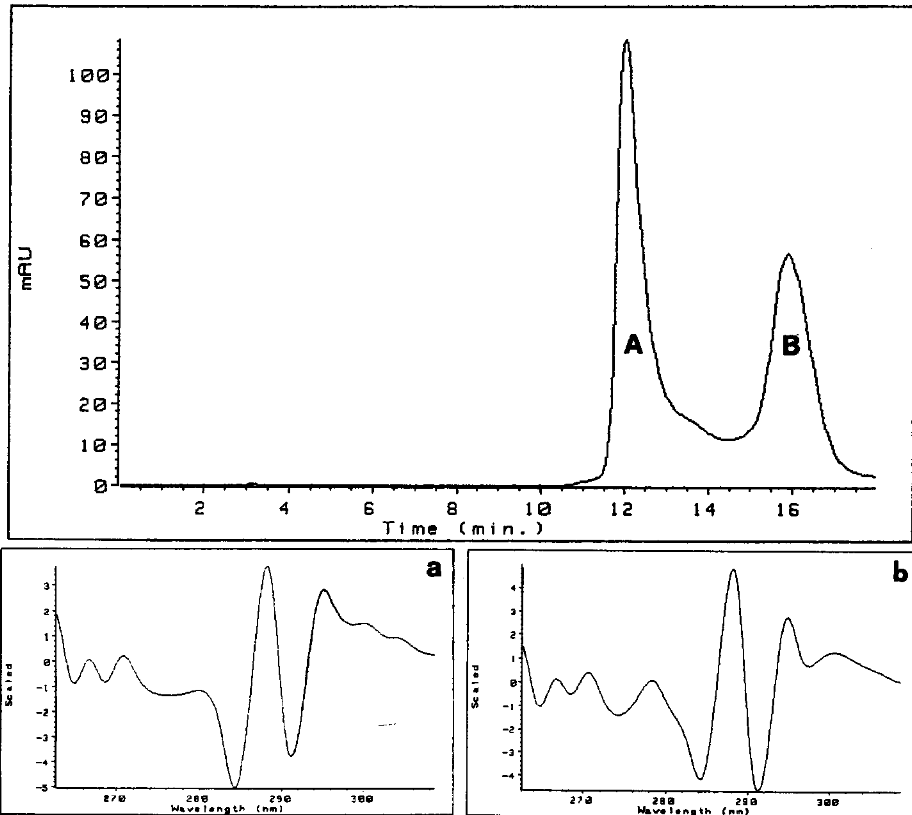

Fig. 5. Above, size-separation chromatogram for mutant 4 illustrating aggregated (A) and monomeric (B) protein components. Chromatograms were obtained using a Waters Protein-Pak 125 size-exclusion column as described in Materials and Methods. Below, the second-order UV absorption spectra (normalized) obtained from the aggregated (a) and monomeric (b) protein peaks above. The values of $a = 6.3$, $b = 4.7$, $r_n = 1.35$ and $\alpha = 0.79$ were obtained from the aggregate spectrum (a) while the monomeric spectrum (b) provided values of $a = 9.4$, $b = 7.8$, $r_n = 1.21$ and $\alpha = 0.73$.

TABLE II

NEAR-UV CIRCULAR DICHROISM ANALYSIS (θ AT 295 nm) FOR IL-2 AND MUTANTS 1–4

Protein	θ (295 nm) (degrees cm^2 $dmol^{-1}$)
IL-2	88.2
1	116.7
2[a]	−268.0
3	− 59.5
4	−136.3

[a] Aggregated protein.

although the nature of the asymmetry is different, the magnitudes of θ for both mutants are comparable to the value determined for IL-2. These findings clearly substantiate that the conformation of the protein surrounding the Trp side chain has been influenced by the some of the mutations and indicate that the values obtained for parameter b (Table I) are a useful probe for sensing this structural alteration.

CONCLUSIONS

We have employed HPLC size-exclusion chromatography and diode array detection to provide information on both the state of aggregation and the conformation for a series of IL-2 related mutant proteins. By analysis of the second derivative of the UV absorption spectrum obtained from peaks emerging from the column, we were able to monitor the relative solvent exposure of Tyr residues and the polarity of the Trp environment. The average solvent exposure for the Tyr residues was approximately 2/3 for IL-2 itself and this value was relatively insensitive to the nature of the mutations made in the C-terminal helix. This suggests that the solvent exposure of the Tyr residues was not significantly altered. The environment of Trp 121 in IL-2 was relatively non-polar, consistent with our previous finding that is fluorescence was not susceptible to iodide quenching [19] and therefore the indole group is buried within the protein core. However, second-derivative analysis suggested that the polarity of the Trp environment was affected by mutations in the helix in which it resides confirming conformational perturbation of this region. It should be noted that conformational alterations which have no net effect on Tyr exposure or Trp polarity would be invisible to this technique. Nevertheless, by using a combination of HPLC and second derivative UV spectroscopy we were able to rapidly analyze a series of IL-2 mutants and confirm that none were significantly denatured as a result of the mutations that were introduced. Therefore, any dramatic change in the observed biologic activities cannot be attributed to grossly misfolded or denatured protein. The conformational alteration that was detected, specifically increased polarity of the Trp indole chromophore, was consistent with the nature and position of the mutations which were introduced. This increased polarity can be attributed either to a change in the surrounding protein environment [17] or increased solvent accessibility [20].

The advantage of diode array detection coupled to chromatographic separation

for this type of analysis was evident when the second-derivative UV absorption characteristics of peaks representing protein aggregates were examined. In all cases the spectra were distinctly different from those obtained from monomeric peaks. Furthermore, since we averaged several spectra to obtain a single spectrum for second-derivative analysis, the homogeneity within a single peak could be assessed. As long as the $A_{280\ nm}$ value remained above 50 mA units, we observed no significant differences between any single spectrum and the averaged spectrum for monomers. This was not the case for peaks representing aggregated proteins. Differences were detected in spectra obtained at various positions within the peak suggesting that aggregation was not always homogeneous.

For IL-2 mutants as with many other recombinant proteins, the tendency to aggregate is dependent upon the type of mutation, method of preparation and purification, concentration, conditions and length of storage. Since it is important to assess the degree of aggregation prior to any type of biophysical or biologic analysis, size-exclusion chromatography can be a useful technique. We believe that it is also critical to determine any conformational alterations resulting from sequence mutations in order to correctly interpret biologic results. The combination of HPLC size-exclusion chromatography and scanning diode array detection allow one to obtain information on both aggregation state and conformation in a single step, and at the same time, avoid artifactual effects introduced by aggregation in non-chromatographic second-derivative UV absorption analysis.

ACKNOWLEDGEMENTS

The authors would like to express their appreciation to Robert Guiffre of Hewlett-Packard Corp., Paramus, NJ, U.S.A., for his assistance with data analysis and for his comments regarding this manuscript. Supported by the Institute of Allergy and Infectious Disease, National Institutes of Health, Bethesda, MD, U.S.A.

REFERENCES

1 K. J. Wilson, in T.E. Hugli (Editor), *Techniques in Protein Chemistry*, Academic Press, San Diego, pCA, 1989, pp. 337–339.
2 T. A. Kunkel, J. D. Roberts and R. A. Zakour, *Methods Enzymol.*, 154 (1987) 367.
3 J. F. Reidhaar-Olson and R. T. Sauer, *Science (Washington, D.C.)*, 241 (1988) 53.
4 C. W. Matthews, *Methods Enzymol.*, 154 (1988) 498.
5 T. Ichikawa and H. Terada, *Chem. Pharm. Bull.*, 29 (1981) 438.
6 L. Servillo, G. Colonna, C. Balestrieri, R. Ragone and G. Irace, *Anal. Biochem.*, 126 (1982) 251.
7 R. Ragne, G. Colonna, C. Balstrieri, L. Servilo and G. Irace, *Biochemistry*, 23 (1984) 1871.
8 D. M. Brems, S. M. Plaisted, H. A. Hard and C-S. C. Tomich, *Proc. Natl. Acad. Sci. U.S.A.*, 85 (1988) 3367.
9 B. Grego, E. C. Nice and R. J. Simpson, *J. Chromatogr.*, 352 (1986) 359.
10 W. Shiaw-Lin, K. Benedek and B. L. Karger, *J. Chromatogr.*, 359 (1986) 3.
11 W. Shiaw-Lin, A. Figueroa and B. L. Karger, *J. Chromatogr.*, 371 (1986) 3.
12 B. Landgraf, D. P. Williams, J. R. Murphy, K. A. Smith and T. L. Ciardelli, *Proteins: Struct., Func. Gen.*, 9 (1991) in press.
13 K. A. Smith, *Science (Washington, D.C.)*, 240 (1988) 1169.
14 M. P. Weir and J. Sparks, *Biochem. J.*, 245 (1987) 85.
15 D. J. Barlow and J. M. Thornton, *J. Mol. Biol.*, 201 (1988) 601.
16 M. P. Weir, M. A. Chaplin, D. M. Wallace, C. W. Dykes and A. N. Hobden, *Biochemistry*, 27 (1988) 6883.

17 H. A. Havel, E. W. Kauffman, S. M. Plaisted and D. N. Brems, *Biochemistry*, 25 (1986) 6533.
18 J. D. Fleischmann, D. Wentworth, F. Valencic, A. L. Imbembo and K. A. Koehler, *Biochem. Biophys. Res. Commun.*, 152 (1988) 879.
19 F. E. Cohen, P. A. Kosen, I. D. Kuntz, L. B. Epstein, T. L. Ciardelli and K. A. Smith, *Science (Washington, D.C.)*, 34 (1986) 349.
20 H. Terada, Y. Noue and T. Ichikawa, *Chem. Pharm. Bull.*, 32 (1984) 585.