Biochemistry

 Volume 23, Number 12 June 5, 1984

Articles

Synthesis and Physicochemical Characterization of Major Fragments of Human Leucocyte Interferon α_1[†]

Thomas Leist and Richard M. Thomas*

ABSTRACT: Four oligopeptides corresponding to overlapping regions of the human leucocyte interferon α_1 (IFN) amino acid sequence have been synthesized by the solid-phase method. The physicochemical properties of these peptides were investigated by analytical ultracentrifugation and circular dichroism, absorbance, and fluorescence spectroscopies. IFNs 24–81 and 111–166 were shown to possess considerable ordered structure in solution. IFN 111–166 appeared to adopt a β-stranded conformation, which could be reversibly unfolded by the addition of urea. Equimolar mixtures of IFNs 24–81 and 111–166 formed a noncovalent complex that was resistant to the formation of the Cys_{29}–Cys_{139} disulfide bond found in native interferon α_1. Predictions of the secondary structure of interferon α_1 were made and the results compared to the measured properties of the fragments. In general, it was found that the prediction methods favored a highly α-helical conformation.

Considerable attention has been focused on the interferon protein family. Most of the extensive information now available on these proteins is concerned with their biological activity and potential therapeutic use as antiviral and antitumour agents (De Maeyer et al., 1981; De Maeyer & Schellekens, 1983). As yet, little is known about the structure and physical chemistry of these molecules other than their primary amino acid sequences (usually inferred from the relevant nucleotide sequences of cloned cDNAs).

Here we report the solid-phase synthesis of four major peptide fragments of human interferon α_1 (IFN),[1] IFN 1–81, IFN 24–81, IFN 71–166, and IFN 111–166, which together span the entire protein and overlap in the central region of the sequence. In a preliminary report (Arnheiter et al., 1981), IFN 111–166 has been shown to elicit antibodies in mice that cross-react with the intact IFN α_1 and IFN α_2 molecules.

The aims of this study are to characterize the physical properties of the fragments and to evaluate the extent to which they are able to form structure in solution. As the three-dimensional structures of the interferons are unknown, attempts have been made to predict the secondary and tertiary structure of the proteins (Sternberg & Cohen, 1982; Zav'yalov & Denesyuk, 1982). We, too, have carried out prediction studies, the results of which, taken in conjunction with the physicochemical data obtained, may provide some information on the sequence regions in intact IFN covered by the synthetic fragments. Tentative predictions of reverse turns were made as it seems probable that these, and other, surface features of the molecule will prove to be important in the antigenicity, receptor binding, and activity of the protein.

Materials and Methods

Chemicals were from Merck (Darmstadt) unless otherwise stated.

Synthesis and Purification. The amino acid sequences of IFNs 1–81, 24–81, 71–166, and 111–166 were deduced from the nucleotide sequence of the cDNA coding for IFN α_1 (Mantei et al., 1980). Peptides were synthesized by the solid-phase method (Merrifield, 1963; Barany & Merrifield, 1980) as described previously (Arnheiter et al., 1981). Boc-protected amino acids were purchased from BaChem AG (Bubendorf, Switzerland). Coupling was monitored by amino acid analysis of samples withdrawn at regular intervals during

[†] From the Biochemisches Institut der Universität Zürich, CH-8057 Zürich, Switzerland. *Received September 14, 1983; revised manuscript received January 3, 1984.* This work was supported in part by a grant from the Schweizerische Nationalfonds (3.214-82).

* Address correspondence to this author at the Institut für Polymere, Eidgenössische Technische Hochschule (ETH), CH-8092 Zürich, Switzerland.

[1] Abbreviations: CD, circular dichroism; GSH and GSSG, reduced and oxidized glutathione; IFN, interferon α_1; TFE, trifluoroethanol; TLC, thin-layer chromatography; Boc, *tert*-butyloxycarbonyl; EDTA, ethylenediaminetetraacetic acid.

0006-2960/84/0423-2541$01.50/0

the synthesis. The sulfhydryl groups of Cys_1, Cys_{29}, Cys_{86}, Cys_{99}, and Cys_{139} were blocked irreversibly by carboxamidomethylation or, reversibly, by forming the mixed disulfide with glutathione (Sigma) in the presence of excess of the reduced form of the reagent (5 mM GSH, 0.1 mM GSSG, pH 5). The stoichiometry of glutathione binding was measured by amino acid analysis. The peptides were purified either as described by Arnheiter et al. (1981) or by gel filtration on Sephacryl S-200 (Pharmacia) in the presence of 4 M urea, 5 mM GSH, and 0.1 mM GSSG in 0.01 M sodium phosphate, 0.09 M NaCl, and 0.001 M EDTA, pH 5 ("phosphate buffer").

Urea and/or glutathione were removed where necessary by dialysis with Spectrapor 3 dialysis membrane (Spectrum Medical Industries) or by desalting on Sephadex G-25. The purity of the products was judged (a) by thin-layer chromatography on silica gel with an eluting solvent of 1-butanol–acetic acid–pyridine–water (10:3:10:12), (b) by polyacrylamide gel electrophoresis in 0.1 M sodium phosphate–6 M urea (Schwarz/Mann, ultrapure), pH 7.2, and (c) by amino acid analysis (Biotronik LC 600-1 analyzer).

Spectral Measurements. CD measurements were performed on a Jasco 40AS recording spectropolarimeter in quartz cuvettes (path length 0.01–1.0 cm). The solvent used was 0.01 M sodium phosphate–0.09 M sodium chloride at pH 5.0 or 7.0. Peptide samples were dissolved in an 8 M solution of urea and then dialyzed extensively against buffer. Solutions were clarified by centrifugation and millipore filtration before use. Protein concentrations were determined by amino acid analysis. Spectra were recorded at 22 °C. They were digitalized manually, and the relevant blank was subtracted. Results were expressed as mean residue ellipticity $[\theta]$. Mean residue molecular weights were calculated from the amino acid compositions. The far-UV CD spectra were analyzed primarily by the method of Provencher and Glöckner (Provencher & Glöckner, 1981; Provencher, 1982). The far-UV CD spectrum of IFN 111–166 in the presence of increasing urea concentration was measured in a 1-cm path-length cuvette at 22 °C. Small aliquots of a solution of 8 M urea in buffer were added to the peptide solution and allowed to equilibrate, and then the spectrum was recorded. Final urea and protein concentrations were determined by weight dilution.

Fluorescence emission spectra were measured on a Perkin-Elmer MPF 2LA spectrofluorometer at 22 °C and absorption spectra on a Hewlett-Packard HP 8450 diode array spectrophotometer, equipped with a thermostated cell holder.

Molecular Weight ($\bar{M}_{w,obsd}$) Determination. Molecular weights were estimated from the results of conventional sedimentation equilibrium experiments in the Beckmann Model E analytical ultracentrifuge (Chervenka, 1969) with UV absorption optics at 280 nm. Values of $\bar{M}_{w,obsd}$ were calculated from a corrected plot of ln c vs. r^2 (Creeth & Pain, 1967). Partial specific volumes were calculated from the amino acid sequences (Cohn & Edsall, 1965). Peptide solutions (~0.2 mM) were dialyzed against phosphate buffer, pH 8.0, overnight, and the dialyzate was used as the reference solution.

Secondary Structure Prediction. A prediction of the secondary structure of IFN α_1 was made with a modification of the method of Argos et al. (1976) and an APL computer program developed by A. Honegger. The methods employed in the construction of the joint prediction histogram were those of Chou & Fasman (1978), Burgess et al. (1974), Garnier et al. (1978), and Tanaka & Scheraga (1976a,b). The hydropathy profile was obtained by the method of Kyte & Doolittle (1982). The span of the "moving window" was set at seven residues (residue i to i + 6), and the moving mean residue hydropathy was attributed to the central residue (i + 3) of each segment.

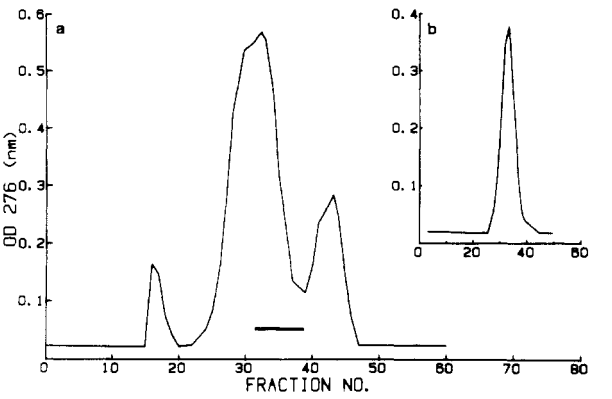

FIGURE 1: (a) Purification of IFN 24–81 on S-200 Sephacryl. The peptide (30 mg cm^{-3}) was dissolved in 10% β-mercaptoethanol–6 M urea in phosphate buffer and incubated overnight at room temperature. The resultant solution was clarified by centrifugation, applied to a preequilibrated, calibrated S-200 Sephacryl column (120 × 1.6 cm) and eluted with 4 M urea, 5 mM GSH, and 0.1 mM GSSG in phosphate buffer, pH 5.0, with an upward flow rate of 15 cm^3 h^{-1}. The elution was monitored by UV absorbance at 276 nm. The fractions indicated by the horizontal bar were pooled, lyophilized, and rechromatographed under the same conditions, the elution characteristics being shown in (b).

Results

Purification. The crude peptide preparations were initially incubated in 10% mercaptoethanol–6 M urea in either 1 M acetic acid or in phosphate buffer for a minimum of 20 h. In all four cases, there was some insoluble material. The precipitates were removed by centrifugation (12000 rpm, 15 min). Increasing the incubation time (up to 72 h) did not improve the yield of solubilized material, suggesting that all soluble species present dissolved readily and that, as the conditions employed were strongly denaturing, the precipitates possibly consisted of cross-linked or chemically modified material. The soluble fraction was then submitted directly to gel filtration. Using Sephadex G-50 and an eluting solvent of 0.1 M acetic acid or using Sephacryl S-200 and eluting with phosphate buffer in the presence of 4 M urea and glutathione produced markedly similar chromatographic patterns. As an example, the elution characteristics of IFN 24–81 are shown in Figure 1.

Taking into account the small differences in molecular weight, the general form of the chromatograms varied little from peptide to peptide. The elution profiles obtained consisted of a small number of overlapping peaks or shoulders corresponding to a molecular weight range of $(IFN_x)_1$ to $(IFN_x)_n$, where n = 3 for IFN 1–81 and n = 2 for IFNs 24–81 and 111–166, little material of higher molecular weight being found. Fractions containing the monomeric form were pooled, lyophilized, dissolved in a small volume of water (<2 cm^3), and rechromatographed under the same conditions as before. In all cases, a single symmetrical peak with an elution volume characteristic of the monomeric molecular weight was obtained (Figure 1b). Fractions containing protein were then used for further experimentation or stored frozen. These products could be dialyzed or desalted (to remove urea/glutathione) with little or no precipitation (depending on the final buffer composition), but subsequent freeze-drying resulted in the formation of largely insoluble products. IFN 71–166 was purified as described previously (Arnheiter et al., 1981).

The fragments had the correct amino acid compositions as determined from the sequence of IFN α_1. IFN 24–81 and IFN

Table I: Molecular Weight Properties of Synthetic Interferon Fragments

peptide	pH	concn (mg cm^{-3})	$\bar{M}_{w,obsd}$[a]	M_c[a]
IFN 1-81	8	0.6	10	9.6
IFN 24-81	8	2.3	6.4	6.6
IFN 24-81	8	0.47	7.1	6.6
IFN 111-166	5	0.6	6.5	6.7

[a] $\bar{M}_{w,obsd}$ is the experimentally determined molecular weight, and M_c is that calculated from the amino acid composition, both values being expressed in kilodaltons.

111–166 both gave single spots on TLC (R_f 0.52 and 0.74, respectively) and single bands on gel electrophoresis (of M_r 6500 and 6000, respectively). The two shorter peptides (IFN 24–81 and IFN 111–166) will be considered first.

Solubility. The peptides were both soluble in 1 M acetic acid and aqueous urea solutions and could be dissolved in phosphate buffer in the 0.1–0.5 mM concentration range required for the physicochemical experiments, forming solutions that were stable for several days.

Molecular Weight Measurements. The observed molecular weights ($\bar{M}_{w,obsd}$) of the purified fragments derived from sedimentation experiments are listed in Table I. The values showed little change over a 5-fold increase in peptide concentration. The observed values are in good agreement with those calculated from the amino acid composition of the peptides. While this is of interest in itself, more importantly, it demonstrates that, under the conditions used for spectral measurements, the peptides were present as monomeric species. This eliminates the possibility that such experiments were performed on disulfide cross-linked or associated forms.

Spectroscopic Measurements. UV absorption and fluorescence spectra were measured primarily to check on the tryptophan content of the peptides. It is well-known that tryptophan is labile under the synthesis conditions unless protective measures are taken. IFN α_1 has two tryptophan residues, Trp_{77} in IFN 24–81 and Trp_{141} in IFN 111–166. Amino acid analysis of IFN 111–166 [hydrolyzed for 24 h in 6 M HCl in the presence of 2.5% (v/v) thioglycolic acid] gave 1 mol of tryptophan/mol of peptide whereas analysis of IFN 24–81 under the same conditions gave only 0.6 mol of tryptophan/mol of peptide.

The absorption, fluorescence, and near-UV CD spectra of IFN 111–166 were typical of a peptide containing both tyrosine and tryptophan. In addition, the absorption spectrum fitted well to that of a model chromophore mixture that consisted of *N*-acetyltryptophanamide and *N*-acetyltyrosinamide in the correct molar ratio (1:3). On the other hand, the spectrum of IFN 24–81 deviated considerably from that of the relevant model (*N*-acetyltryptophanamide to *N*-acetylphenylalaninamide, 1:6), and the fluorescence emission and excitation spectra suggested that more than one fluorescent species were present. In order to try to fit the observed spectrum, a portion of the model mixture was submitted to hydrolysis in 6 M HCl at 110 °C for 4 h in an attempt to simulate, in part, the extreme conditions experienced by the peptide during synthesis.

The absorption spectra of solutions containing varying ratios of hydrolyzed to unhydrolyzed model mixture were measured. A ratio of hydrolyzed to unhydrolyzed model of 2:3 gave the best fit to the observed spectrum, reinforcing the view that about 60% of the tryptophan was present in an unperturbed form. The pathways of tryptophan degradation are complex [e.g., Savige (1975)], and it is impossible to say which of the

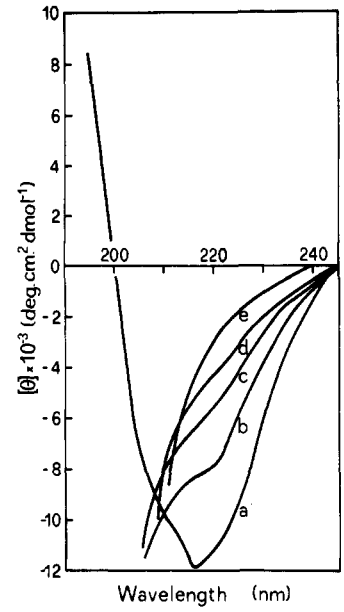

FIGURE 2: Far-UV CD spectrum of IFN 111–166 in (a) aqueous phosphate buffer and in the presence of increasing concentrations of urea: (b) 4.2 M urea, (c) 5 M urea, (d) 6 M urea, and (e) 8 M urea. Experimental details are given in the text.

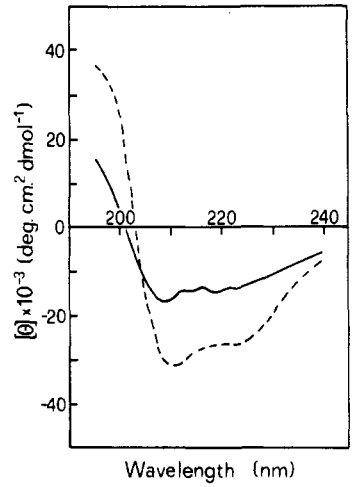

FIGURE 3: Far-UV CD spectrum of IFN 24–81 in phosphate buffer (—) and in 75% (v/v) TFE (– –).

wide range of products might be present in the peptide, although they were assumed to be acid-catalyzed oxidation derivatives. The differences in the tryptophan contents of IFN 24–81 and IFN 111–166 might be explained by the number of synthetic cycles that these residues had to endure following their own insertion into the polypeptide chain: 30 in the case of Trp_{141} in IFN 111–166 but 53 for Trp_{77} in IFN 24–81.

The far-UV CD spectra of the synthetic polypeptides in phosphate buffer were measured in the 190–240-nm spectral region (Figures 2 and 3). The spectra were digitalized manually, without smoothing, at 0.5-nm intervals (center of noise), giving 101 data points per spectrum. The spectra obtained from IFN 24–81 and IFN 111–166 indicated that a considerable amount of ordered secondary structure was present in both the fragments. The spectrum of IFN 111–166 was invariant over a 100-fold change in the peptide concentration and was reminiscent of the spectra of high β-sheet content proteins. The covalent, disulfide-bonded dimer of IFN 111–166 formed under oxidizing conditions and involving Cys_{139} of two monomer molecules had a CD spectrum superimposable on that of the monomeric form, suggesting that

Table II: α-Helical and β-Structure Contents of Interferon α_1 Fragments

peptide	obsd[a]	range[a]	prediction 1[b]	prediction 2[c]
	α-Helical Content (%)			
IFN 1-81	NA[d]		48	39
IFN 24-81	16	16-18	52	35
IFN 71-166	NA[d]		47	39
IFN 111-166	24	23-24	20	21
	β-Structure Content (%)			
IFN 1-81	NA[d]		0	6
IFN 24-81	35	30-42	0	9
IFN 71-166	NA[d]		7	9
IFN 111-166	36	36-36	12	16

[a] Provencher & Glöckner (1981). [b] Method of Argos et al. (1976), this study. [c] Sternberg & Cohen (1982). [d] NA, not analyzed. On visual inspection, the CD spectra of both IFN 1-81 and IFN 71-166 clearly indicated large "random" structural components. Due to the uncertainties in the analysis of such structures, the CONTIN method was not used in these cases.

dimerization produced little if any structural change and that Cys_{139} was located in an accessible or flexible region of the molecule. The CD spectrum of IFN 24–81 had three prominent bands of approximately equal intensities centered on 205, 215, and 220 nm (Figure 3), indicating that the peptide probably contained elements of α- and β-, as well as unordered, structures.

Enhancement of the structural content of IFN 24–81 was observed in the presence of 75% TFE (Figure 3). The aqueous solution far-UV CD spectra of IFN 24–81 and IFN 111–166 were analyzed by using the CONTIN regularization method of Provencher and Glöckner (Provencher & Glöckner, 1981; Provencher, 1982). The various possibilities that the program offers have been examined by Schnarr & Maurizot (1982) in an investigation of the CD of the *lac*-repressor head-piece fragment. With the CONTIN program, it is possible (a) to account for uncertainties in the absolute protein concentration and (b) to differentially weight the data to accommodate increased noise at low wavelengths, and (c) the program generates several different solutions [i.e., with varying α in eq 4 of Provencher & Glöckner (1981)]; only one of these is selected as being the "best fit" although other members of the generated set of solutions may adequately describe the data. Schnarr and Maurizot also tested the effect of varying the low-wavelength cut-off (λ_{min}) of their data and found that raising λ_{min} lowered the β-sheet content. Additionally, they examined the variation of α- and β-structure content as a function of the standard deviation of each of the solutions offered by the program. Here it was found that, while the quality of the fit to the data obviously falls with increasing standard deviation, nevertheless a considerable variation (lowering) in β-structure was possible while the amount of α-helix remained fairly constant.

Analyses of the spectra of the interferon fragments are given in Table II. Peptide concentrations were determined by amino acid analysis (average of two duplicate measurements) and are assumed to be within the accuracy of the analytical method (taken as 5%), and the CONTIN concentration uncertainty option was not employed. Differentially weighting the spectrum (two regions, 190–210 and 211–240 nm) produced little alteration in any of the values (maximum variation ±4%), which probably indicates that the quality of the data was approximately the same over the whole spectral range. Varying λ_{min} had little or no effect and tended to produce small variations in both α- and β-content of the spectra. In practice,

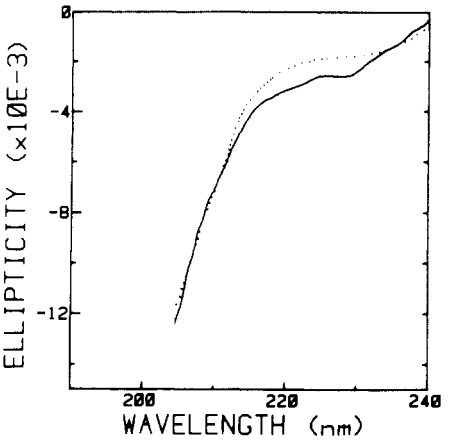

FIGURE 4: Far-UV CD spectra of IFN 71–166 (—) and IFN 1–81 (···) in aqueous phosphate buffer, pH 7.0, 22 °C.

measurements on the peptides could be made down to 185 nm, and the entire range of the analytical package (190–240 nm) was used. α- and β-content both varied as a function of the standard deviation of the computed fit, and in general, acceptable values fell within the ranges given in Table II, which represent the limiting permissible values following the rules formulated by Provencher. The authors of the regularization method stress that little reliance can be placed on the computed values for the remainder and turn configurations.

The CD spectrum of IFN 111–166 was also analyzed by the methods of Chang et al. (1978) and Bolotina et al. (1979), giving values for α-helical content of 21 and 30% and for β-sheet of 30 and 20%, respectively. Figure 2 shows the effect of increasing urea concentration on the far-UV CD spectrum of IFN 111–166. Equilibrium measurements indicated that the conformational transition was broad and noncooperative with a midpoint between 5 and 6 M urea. The transition was complete in 7–8 M urea. No further conformational changes occurred in 6 M guanidinium chloride, where it is normally accepted that proteins adopt a completely random conformation (Tanford, 1968). The reaction was fully reversible (>90%), and the native spectrum was restored upon dilution or dialysis of the denaturant.

The two larger peptides were considerably less soluble than their shorter counterparts, and although IFN 1–81 was soluble in phosphate buffer, it showed a tendency to precipitate after about 48 h. IFN 71–166 was the least soluble in aqueous solution ($<5 \times 10^{-6}$ M) but readily dissolved in high (>40% v/v in water) concentrations of TFE. The CD spectra of the two peptides (Figure 4) indicated that they adopt a largely random configuration in aqueous solution, although shoulders in the 215–230-nm spectral range might be indicative of a small component of organized structure. In TFE solutions, IFN 71–166 became considerably ordered as might be expected given the known structure-promoting properties of this solvent. The spectrum remained the same when the concentration of TFE was increased to 90% v/v.

Structure Predictions for α Interferons. Predictions of both the secondary and tertiary structures of interferon (Sternberg & Cohen, 1982; Zav'yalov & Denesyuk, 1982) have been made. Sternberg & Cohen (1982), using three different methods for predicting the secondary structure (Chou & Fasman, 1978; Garnier et al., 1978; Lim, 1974), proposed a three-dimensional model in which four α-helices (A, residues 12–23; B, residues 54–70; C, residues 79–104; D, residues 141–152) were packed to form a right-handed helical bundle and in which there were two short tracts of β-sheet (residues 30–34 and 123–131, respectively). The structural constraints

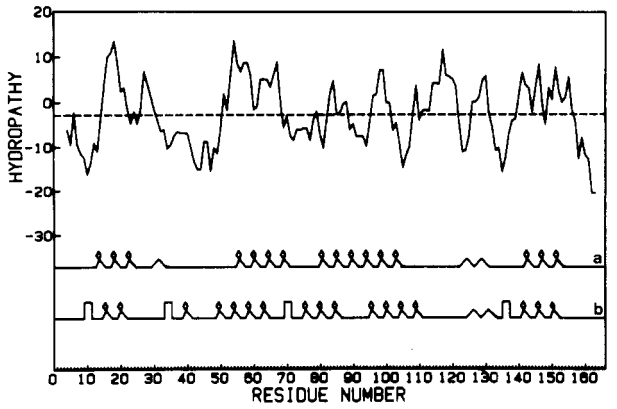

FIGURE 5: Structural predictions for interferon α_1: (top) hydropathy profile prepared according to the method of Kyte & Doolittle (1982); (bottom) secondary structure predictions; (a) that of Sternberg & Cohen (1982); (b) method of Argos et al. (1976), this study. (Looped line) α-helix; (〰) β-strand; (‖) reverse turn; (—) undefined.

imposed on the interferon molecule by the disulfide bonds were met by this model. Another secondary structure prediction from a modification of the method of Argos et al. (1976) has been made in the present study. The results of the predictions are summarized in Figure 5.

A comparison of the results showed broad agreement; all identify helices close to both the C- and N-termini, major internal helical regions within the sequence intervals 48–70 and 74–109, and a β-strand between residues 123 and 131. In order to reinforce the prediction studies, a hydrophobicity plot for interferon was constructed.

Rose & Roy (1980) have studied the correlation between linear-chain hydrophobicity and protein/protein contact density. They found that these two properties correlate well as to the hydrophobicity and main-chain atom temperature factors, the latter being a measure of conformational flexibility. Not only does this method identify the "inside" and "outside" of the protein molecule, but if a hydrophobic cluster model for protein folding is assumed, it may also identify structural segments within the protein. Rose and Roy employed the aqueous to organic phase-transfer free energies for amino acids obtained by Nozaki & Tanford (1971). This method has been refined recently by Kyte & Doolittle (1982), who have constructed a "hydropathicity index", using water vapor transfer free energies and the residue accessibility values determined by Chothia (1976). The method (see Materials and Methods) employs a "sliding window", which examines each segment of the protein sequence, advancing by one residue each cycle. For every segment, the average hydropathicity is calculated and plotted. Following Kyte and Doolittle, the span of the "window" was set at seven residues. The result of this calculation for IFN α_1 is shown in Figure 5. According to Rose & Roy (1980), local maxima in chain hydrophobicity should correspond to areas of incipient structure formation, and as can be seen, with the possible exception of helices 74–85 and 14–22, all α- and β-elements predicted in this study lie largely or wholly within such maxima. Conversely, local minima in the plot should correspond to solvent-exposed regions of the chain. The hydropathicity plot shows absolute minima (apart from the highly hydrophilic C-terminus) at residues 10 and 135, close to tentatively predicted reverse turns at residues 10–13 and 136–139.

Fragment Recombination. The disulfide bond Cys_{29}–Cys_{139} of IFN α_A has been shown to be essential for antiviral activity and gives added stability to the molecule (Wetzel et al., 1983). As a preliminary fragment complementation experiment, the combination of IFN 24–81 and IFN 111–166 via the homologous disulfide bond of IFN α, Cys_{29}–Cys_{139}, was attempted. The peptides were dissolved in equimolar quantities in 4 M urea, 5.0 mM GSH, and 0.1 mM GSSG. After incubation overnight, the urea was removed by gel filtration. The reaction mixture contained two products: one that on gel filtration had a molecular weight corresponding to the monomer (~6500) and the other (about 60% of the total material) had molecular weight of 15000. The high molecular weight fraction, which had an amino acid analysis corresponding to a 1 to 1 mixture of the two fragments, was collected and the GSH to GSSG ratio adjusted to 0.3 mM/0.3 mM. The material was incubated for 20 h, freeze-dried, and dissolved in 4 M urea. On subsequent gel filtration in urea, it was found that all the material was again monomeric. The results of this experiment suggest that IFN 24–81 and IFN 111–166 are capable of forming a noncovalent complex but are resistant to the formation of the disulfide bond between Cys_{29} and Cys_{139}. This result is similar to that found when the reoxidation of whole IFN α_A from the mercaptoethanol-reduced form is attempted (Wetzel et al., 1983).

Biological Activity of Interferon Fragments. None of the interferon fragments described herein had any antiviral activity when assayed as described by Arnheiter et al. (1981). Almost all monoclonal antibodies prepared by Staehelin et al. (1981) against a mixture of leucocyte interferons cross-reacted, to various extents, with IFN 24–81 and IFN 111–166 (T. Staehelin, personal communication).

Discussion

Purity and Homogeneity of the Interferon Fragments. Synthetic peptides, produced by the currently available methods, all suffer, to some degree, from the problem of heterogeneity (Barany & Merrifield, 1980). Possible causes of heterogeneity include (a) incomplete coupling, (b) cross-linking of the peptide chains or side-chain modification, (c) deletions resulting in premature chain termination, and (d) internal deletions. The modified products produced as a result of (b) and (c) can be removed, comparatively simply, by the use of molecular weight dependent purification methods under denaturing (disassociating) conditions. Unfortunately, such methods cannot discriminate between the required product and byproducts lacking one (or few) residue(s) resulting from internal deletion(s) or species in which one (or few) residue(s) are chemically modified. The only method for assessing heterogeneity accurately is quantitative sequence analysis, which is often impracticable. The purity of the synthetic interferon fragments has been established by (a) quantitative amino acid analysis, (b) gel filtration under dissociating conditions, (c) independent molecular weight determination, and (d) amino acid analysis of samples taken at regular intervals during the synthesis of the polypeptides.

Evidence has been presented that the two shorter fragments are capable of considerable self-organization. CD measurements indicated that IFN 111–166 contained approximately 25% α-helix, a value which agreed well with the structural predictions for the corresponding region of the intact interferon molecule (Table II). On the other hand, the value obtained for β-content from CD measurements (35%) differed greatly from the predicted content (12–16%). The discrepancies between both the predictions and between them and the observed structural contents were more marked in the case of IFN 24–81. Two of the possible explanations for these inconsistencies concern the methods of analysis used: inaccuracy in the structural analyses of the CD spectra and inaccuracies in the prediction methods employed.

It has been suggested (Schnarr & Maurizot, 1982) that the

accuracy of the CONTIN method is lower for β-content than for α-helix, although reflexive use of the program to examine individually its own reference spectra produces results that are in striking agreement with the X-ray diffraction data for the same protein. The spectra of the interferon fragments had high aperiodic (random) components, a fact that might in itself interfere with analysis for β-structure as is suggested by a high β-content afforded to IFN 1–81, which was clearly structurally random. Prediction methods, while they are being constantly updated and improved, are not expected to offer anything like total accuracy (Schulz & Schirmer, 1979), but it is clear that they offer better results when constrained either by known physical (structural) properties or by the use of other independent structural variables. The hydropathy index method may be valuable in the reinforcement of turn predictions and also gives a guide as to the probable location of large structural elements although possibly produces ambiguous results in the case of amphipathic α- or β-structures. The hydropathy profiles of the homologous interferons α_2, α_{4B}, α_5, α_6, α_7, and α_8 (P. Dreiding and R. M. Thomas, unpublished results) are strikingly similar to one another and to that of interferon α_1. The conservation of the hydropathy profile throughout this group of proteins in turn suggests the conservation of structural motifs and assists in the prediction of structurally important regions in the molecule. This may be of particular interest in the prediction of reverse turns as they are located at the protein surface (Kuntz, 1972) and may be important for the function of the protein. Having noted these points, it is appropriate to consider the nature of the structure found in the interferon fragments. The ability of protein fragments to fold (or not to fold) has been reviewed recently by Wetlaufer (1981).

In a broad sense, the folding of an incomplete polypeptide chain can have one of three consequences: (1) the peptide is incapable of any folding in the absence of the information provided by the missing region(s) of the polypeptide chain; (2) the fragment folds to a thermodynamically available but nonnative state; (3) the fragment folds to the same conformation that it adopts in the intact molecule, in which case it can be considered as a folding domain (Wetlaufer, 1973). In view of the immunochemical behavior of IFN 111–166 (Arnheiter et al., 1981) and the reversibility of the unfolding of the peptide, it seems probable that this fragment is capable of adopting extensive nativelike conformation and may represent a folding domain of the interferon molecule. If this is the case, it is hard to reconcile the high observed β-content with that given by the prediction methods. Bewley et al. (1982) in an investigation of the CD spectrum of IFN α_A found 40–70% α-helix but negligible β-structure. Approximately 50% of the peptide chain of IFN 24–81 was also structurally ordered. As detailed immunochemical data is not available for this peptide, it is not possible to decide whether the structure present is native or nonnative. Elongation of either of the two shorter peptides formed products that were almost devoid of secondary structure despite the fact that they included the sequence regions responsible for structure seen in shorter fragments. IFN 71–166 had a strong tendency to form a variety of aggregates, and this may partly explain the apparent absence of structural features in the CD spectrum although IFN 1–81, which was soluble in aqueous buffer and both monomeric and reasonably stable, was also randomly composed. The folding of a protein is dictated by the requirement that both residue–residue and residue–solvent interactions are maximized favorably. For the two larger interferon fragments, the probability that all such interactions can be formed may reasonably be supposed to be lower than that in the shorter fragments. This in turn may destabilize structures that can form in the sequence intervals spanned by the shorter fragments.

Acknowledgments

We thank B. Straub and Dr. H. Jäckle for their help with the physical measurements, Dr. S. Provencher for supplying a copy of the CONTIN program, and Dr. B. Gutte for many helpful discussions.

Registry No. IFN 1–81, 89789-12-8; IFN 24–81, 89789-10-6; IFN 71–166, 89789-13-9; IFN 111–166, 89789-11-7.

References

Argos, P., Schwarz, J., & Schwarz, J. (1976) *Biochim. Biophys. Acta 439*, 261–273.

Arnheiter, H., Thomas, R. M., Leist, T., Fountoulakis, M., & Gutte, B. (1981) *Nature (London) 294*, 278–280.

Barany, G., & Merrifield, R. B. (1980) in *The Peptides* (Gross, E., & Meienhofer, J., Eds.) pp 1–284, Academic Press, New York.

Bewley, T. A., Levine, H. L., & Wetzel, R. (1982) *Int. J. Pept. Protein Res. 20*, 93–96.

Bolotina, I. A., Chekhov, V. O., & Lugauskas, V. Y. (1979) *Int. J. Quantum. Chem. 16*, 819–824.

Burgess, A. W., Ponnuswamy, P. K., & Scheraga, H. A. (1974) *Isr. J. Chem. 12*, 483–494.

Chang, C. T., Wu, C.-S. C., & Yang, J. T. (1978) *Anal. Biochem. 91*, 13–31.

Chervenka, C. H. (1969) in *A Manual of Methods for the Analytical Ultracentrifuge*, Beckmann Industries, Palo Alto, CA.

Chothia, C. (1976) *J. Mol. Biol. 105*, 1–14.

Chou, P. Y., & Fasman, G. D. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol. 47*, 45–148.

Cohn, E. J., & Edsall, J. T. (1965) *Proteins, Amino Acids and Peptides as Ions and Dipolar Ions*, Hafner, New York.

Creeth, J. M., & Pain, R. H. (1967) *Prog. Biophys. Mol. Biol. 17*, 217–287.

De Maeyer, E., & Schellekens, H., Eds. (1983) *The Biology of the Interferon System*, Proceedings of the International Meeting on the Biology of the Interferon System, Elsevier, Rotterdam.

De Maeyer, E., Galasso, G., & Schellekens, H., Eds. (1981) *The Biology of the Interferon System*, Proceedings of the International Meeting on the Biology of the Interferon System, Elsevier, Rotterdam.

Garnier, J., Osguthorpe, D. J., & Robson, B. (1978) *J. Mol. Biol. 120*, 97–120.

Kuntz, I. D. (1972) *J. Am. Chem. Soc. 94*, 4009–4012.

Kyte, J., & Doolittle, R. F. (1982) *J. Mol. Biol. 157*, 105–132.

Lim, V. I. (1974) *J. Mol. Biol. 88*, 873–894.

Mantei, N., Schwarzstein, M., Streuli, M., Panem, S., Nagata, S., & Weissmann, C. (1980) *Gene 10*, 1–10.

Merrifield, R. B. (1963) *J. Am. Chem. Soc. 85*, 2149–2154.

Nozaki, T., & Tanford, C. (1971) *J. Biol. Chem. 246*, 2211–2217.

Provencher, S. M. (1982) *Comput. Phys. Commun. 27*, 213–227.

Provencher, S. M., & Glöckner, J. (1981) *Biochemistry 20*, 33–37.

Rose, G. D., & Roy, S. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 4643–4647.

Savige, W. E. (1975) *Aust. J. Chem. 28*, 2275–2287.

Schnarr, G., & Maurizot, J.-C. (1982) *Eur. J. Biochem. 128*, 515–520.
Schulz, G., & Schirmer, R. H. (1979) *Principles of Protein Structure*, Springer-Verlag, New York.
Staehelin, T., Durrer, B., Schmidt, J., Takacs, B., Stocker, J., Miggiano, V., Stähli, C., Rubinstein, M., Levy, W. P., Hershberg, R., & Pestka, S. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 1848–1852.
Sternberg, M. J. E., & Cohen, F. E. (1982) *Int. J. Biol. Macromol. 4*, 137–144.
Tanaka, S., & Scheraga, H. A. (1976a) *Macromolecules 9*, 159–167.
Tanaka, S., & Scheraga, H. A. (1976b) *Macromolecules 9*, 812–833.
Tanford, C. (1968) *Adv. Protein Chem. 23*, 121–282.
Wetlaufer, D. B. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 697–701.
Wetlaufer, D. B. (1981) *Adv. Protein Chem. 34*, 61–92.
Wetzel, R., Johnston, P. D., & Czarniecki, C. W. (1983) in *The Biology of the Interferon System* (De Maeyer, E., & Schellekens, H., Eds.) pp 101–112, Elsevier, Rotterdam.
Zav'yalov, R. P., & Denesyuk, A. I. (1982) *Immunol. Lett. 4*, 4–14.

Syntheses of Biotinylated and Dethiobiotinylated Insulins†

Klaus Hofmann,* Wei Jun Zhang,‡ Hana Romovacek, Frances M. Finn, Aksel A. Bothner-By, and P. K. Mishra

ABSTRACT: The 600-MHz proton spectrum of dethiobiotin (prepared from *d*-biotin with Raney nickel) was measured in order to gain information pertaining to its stereochemical homogeneity. The spectrum demonstrated clearly that the material is a 6:1 mixture of two stereoisomers. The cis compound, corresponding to the stereochemistry of *d*-biotin, is the major isomer. Two biotinyl- and two dethiobiotinylinsulins were prepared in which the distance between the biotins and insulin was varied by interposition of spacer arms. The synthesis of these compounds involved repeated *N*-hydroxysuccinimido ester condensations. Biotin *N*-hydroxysuccinimido ester, dethiobiotin *N*-hydroxysuccinimido ester, 6-aminohexanoic acid, and *N*-[3-[(3-aminopropyl)carboxyamino]propyl]succinamic acid *N-tert*-butyl ester served as the building blocks for the spacers. The latter compound was prepared from *N*-[3-[(3-aminopropyl)amino]propyl]succinamic acid sulfate by the use of a selective amino-protecting method based on the differential stability toward acid of citraconyl and *tert*-butoxycarbonyl amino-protecting groups. The structure of *N*-[3-[(3-aminopropyl)amino]propyl]succinamic acid sulfate was established unequivocally by X-ray diffraction. The attachment of the biotinylated spacers to the insulin was exclusively at the $N^{\alpha,B1}$ position. Homogeneity of the final products as well as of the intermediates used in their synthesis was established by thin-layer chromatography, by high-pressure liquid chromatography, and in most instances by elemental analysis. The ratio of 6-aminohexanoic acid to lysine in hydrolysates of the insulin derivatives was in agreement with theory. The insulin derivatives were required for a study on the effect of avidin on their ability to interact with insulin receptors on rat epididymal adipocytes, which is described in the accompanying paper.

In a recent paper (Hofmann et al., 1982), we reported that the attachment of insulin to the carboxyl group of biotin exerts a rather dramatic effect on its ability to bind to succinoylavidin. The half-time for dissociation of biotin from avidin is approximately 200 days (Green, 1975); we have determined the half-time for dissociation of biotin from succinoylavidin as approximately 127 days (Finn et al., 1980). Most surprising was the observation that the avidin and succinoylavidin complexes of $N^{\alpha,B1}$-biotinylinsulin dissociate rather rapidly with a half-time for dissociation of approximately 3 h. The attachment of the bulky insulin molecule to biotin appears to exert a steric impediment to avidin binding. As a consequence, it became of interest to have available a series of biotinylated insulins in which the distance between biotin and insulin is systematically increased by insertion of spacers.

Avidin interferes with the ability of biotinylinsulin to stimulate lipogenesis in rat epididymal adipocytes but has no effect on insulin-stimulated lipogenesis (Hofmann et al., 1977). May et al. (1978) described similar observations with $N^{\epsilon,B29}$-biotinylinsulin. This finding suggests that avidin weakens the ability of biotinylinsulin to interact with insulin receptors on the adipocyte, and it seems logical to assume that this interference should be inversely proportional to the distance between the biotin and insulin in the analogue. In the present paper, we describe the synthesis of four insulin derivatives (Figure 1) in which spacers of increasing chain length separate the biotin or dethiobiotin from the insulin. An accompanying paper explores the dissociation behavior of succinoylavidin and streptavidin complexes of these compounds and relates their ability to stimulate glucose oxidation in rat epididymal adipocytes in the presence or absence of avidin, succinoylavidin, and streptavidin.

Experimental Procedures

Materials and Methods

Biotin was a gift from Dr. W. E. Scott of Hoffmann-La Roche Inc., Nutley, NJ, and Boc_2-insulin[1] (bovine) was a gift

† From the Protein Research Laboratory and the Department of Medicine, University of Pittsburgh, School of Medicine, Pittsburgh, Pennsylvania 15261 (K.H., W.J.Z., H.R., and F.M.F.), and the Department of Chemistry, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213 (A.A.B. and P.K.M.). *Received September 14, 1983.* Supported by National Institutes of Health Grants AM 21292, AM 16532, and RR 00292 and the Juvenile Diabetes Foundation International. A preliminary account of this work was presented at the 21st Symposium on Peptide Chemistry in Tsukuba, Japan, Oct 26, 1983.
‡ Present address: Institute of Organic Chemistry, Academia Sinica, Shanghai, The People's Republic of China.

0006-2960/84/0423-2547$01.50/0