Journal of Peptide Science, Vol. 3, 323–335 (1997)

Design, Synthesis and Conformational Analysis of hGM-CSF(13-31)-Gly-Pro-Gly-(103-116)

N. NOLI[1], M. GURRATH[1], P. ROVERO[2], S. PEGORARO[2], R. P. REVOLTELLA[2], E. SCHIEVANO[1], S. MAMMI[1] and E. PEGGION[2]

[1]University of Padova, Department of Organic Chemistry, Biopolymer Research Center, CNR, Padova, Italy
[2]Institute of Mutagenesis and Differentiation, CNR, Pisa, Italy

Received 15 October 1996
Accepted 12 December 1996

Abstract: On the basis of the X-ray structure and results from structure–activity relationship studies, the following GM–CSF analogue was designed and synthesized by solid-phase methodology: hGM–CSF[13-31]-Gly-Pro-Gly-[103–116]-NH_2. This analogue was constructed to comprise helices A and D of the native hGM–CSF, covalently linked in an antiparallel orientation by the tripeptide spacer Gly-Pro-Gly, which is known as a turn-inducing sequence. The conformational analysis of the analogue by CD spectroscopy revealed an essentially random structure in water, while α-helix formation was observed upon addition of TFE. In 40% TFE the helix content was ∼45%. By two-dimensional NMR experiments in 1:1 water/trifluoroethanol mixture two helical sequences were identified comprising the segments corresponding to helix A and helix D. In addition to medium-range NOESY connectivities, a long-range cross-peak was found involving the leucine residues at positions 13 and 35. Based on the experimentally derived data (54 NOEs), the structure was refined by restrained molecular dynamics simulations over 120 ps at various temperatures. A representative conformation derived from the computer simulation is mainly characterized by two helical segments connected by a loop region. The overall three-dimensional structure of the analogue is comparable to the X-ray structure of hGM–CSF in that helices A and D are oriented in an antiparallel fashion, forming a two α-helix bundle. Nevertheless, there are small differences in the topology of the helices between the solution structure of the designed analogue and the X-ray structure of hGM–CSF. The possible implications of these conformational features at the effects of biological activity are discussed.

J. Pep. Sci. 3: 323–335

No. of Figures: 10. No. of Tables: 5. No. of References: 46

Keywords: hGM-CSF; four-helix bundle; α-helix; NMR-based structure determination; molecular modelling

Abbreviations: CD, circular dichroism; DQF-COSY, double quantum filtered correlation spectroscopy; ES-MS, electrospray mass spectrometry; HBTU, *O*-Benzotriazole-*N,N,N′N′*-tetramethyl-uronium hexafluorophosphate; hGM–CSF, human granulocyte-macrophage colony-stimulating factor; HOB, 1-hydroxybenzotriazole; NMP, *N*-methylpyrrolidone; NOESY, (two-dimensional ^{1}H) nuclear Overhauser enhancement spectroscopy; TFE, 2,2,2-trifluorethanol; TOCSY, (two-dimensional ^{1}H) total correlation spectroscopy; TPPI, time proportional phase incrementation.

Address for correspondence: Prof. Evaristo Peggion, University of Padova, Department of Organic Chemistry, Biopolymer Research Center, CNR Via Marzolo 1, 35131 Padova, Italy.

CCC 1075-2617/97/050323-13 $17.50

INTRODUCTION

Human granulocyte-macrophage colony-stimulating factor (hGM–CSF) is a member of a glycoprotein family that stimulates the growth and differentiation of haematopoietic progenitor cells from various lineages, e.g. granulocytes, macrophages, eosinophils and megakaryotes. Additionally, these proteins are involved in various other biological functions related to differentiation processes of tumour cells lines [1, 2]. Like many other growth factors and

cytokines, GM–CSF belongs to a common protein-fold superfamily comprising the family of four-helix bundles, EGF-like, insulin-like, β-trefoil and cysteine knot proteins [3]. According to Presnell and Cohen [4], four-helix bundles are topologically characterized by (i) the chain connectivity ('over-hand' connections), (ii) the direction of the helices (parallel/antiparallel), and (iii) the overall bundle handedness (right/left handed bundle). Within the family of four-helix bundle proteins GM–CSF belongs to the subclass of the so-called short-chain family [5]. Members of the four-helix bundle cytokine family show low sequence homology but are structurally closely related [4, 6–9]. Their common structural motif consists of a core of four helices arranged in an up–up–down–down mode [6, 10] (Figure 1)

The structure of recombinant hGM–CSF solved by X-ray crystallography [11, 12] is characterized by a left-handed, antiparallel four α-helix bundle, a two-strand antiparallel β-sheet and two disulphide bonds. Helix A comprises residues 13–28, helix B 55–64, helix C 68–87 and helix D 103–116 (Figure 2). Helices A/D and B/C form two pairs of tightly packed two-helix bundles with extended interhelical buried interfaces, thus forming a hydrophobic core in the interior of the overall four-helix bundle. The β-sheet is formed by residues 38–43 and residues 98–102. The four cysteine residues (Cys^{54}/Cys^{96}, Cys^{88}/Cys^{121}) are bridged by two disulphide bonds.

The biological functions of GM–CSF are mediated through binding to specific cell surface receptors [14–17]. The receptor is heterodimeric, consisting of two peptide chains, named α and β of 80 and 120 kDa, respectively. The α chain is cytokine-specific

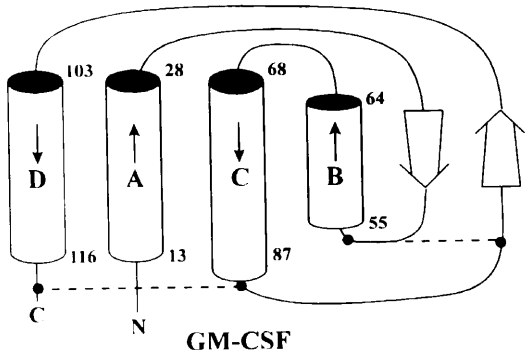

Figure 1 Anatomy of GM–CSF shown in the left handed 'Greek key' representation. The four helices are labelled A–D, in the β-sheet is indicated as two antiparallel arrows. Dashed lines between black circles depict cystine bridges [3].

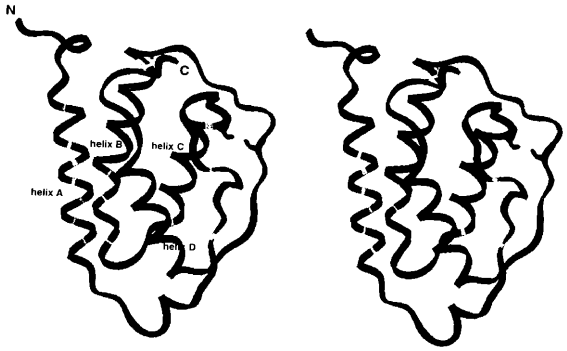

Figure 2 Stereo view of the X-ray structure of rhGM–CSF [11] taken from the Brookhaven Protein Data Bank as 1CSG [13]; only monomer A of the molecule, crystallizing as a dimer, is displayed in ribbon presentation. Cystine bridges are indicated.

[18]. In contrast to the crystal structure of the human growth hormone in complex with its receptor [19], structural data for the human granulocyte-macrophage colony-stimulating factor/receptor complex (GM–CSF/GM–CSFR) are not yet available. Structure–activity relationship studies revealed the critical domains involved in the cytokine/receptor interactions [20–31]. The GM–CSF binding site to the receptor α-chain, i.e. the cytokine-specific receptor chain that mediates the initial contact, involves parts of helix D and C-terminal amino acids (residues 110–127) [23]. The second binding site to the receptor beta subunit is located within helix A [28, 29]. Since helices A and D seem to be simultaneously involved in the interaction with the α and β receptor chains, an analogue was designed which comprises hGM–CSF [14–32] linked through a spacer unit to hGM–CSF [103–116]. The tripeptide spacer Gly-Pro-Gly contains amino acid residues known as helix-breakers and turn-inducers 32–34 which should allow the hGM–CSF analogue to retain the relative orientation of the helix pair A/D as determined in the four-helix bundle of natural factor. The complete sequence of the designed analogue is the following:

W-E-H-V-N-A-I-Q-E-A-R-R-L-L-N-L-S-R-D-G-P-G-
F-E-S-F-K-E-N-L-K-D-F-L-L-V-NH_2

where the tripeptide spacer is underlined.

In the present work we describe the synthesis and the conformational characterization of this analogue by CD, 2D NMR and restrained molecular dynamics simulations.

MATERIAL AND METHODS

Peptide Synthesis

The peptide was synthesized by the solid-phase method using a Milligen 9050 automatic synthesizer with Fmoc/tBu chemistry and continuous flow technology. The synthesis was performed using 0.4 g of PAL–PEG-PS resin (substitution level: 0.17 mmol/g), in order to obtain a C-terminal amide by acidic cleavage. [PAL; 5-(4-amino-methyl-3,5-dimethoxyphenoxy) valeric acid linker; PEG-PS: polystyrene-supported polyoxyethylene scaffold.] The following side-chain protections were used: tBu for Asp, Glu, Tyr and Ser; Trt for Asn, Gln and His; Pmc for Arg and Boc for Lys and Trp. The following synthetic cycles were used: Fmoc deprotection (20% piperidine in DMF, 5 min at 12 ml/min); DMF washing (12 min at 6 ml/min); coupling (four-fold excess of Fmoc-amino acid, HBTU, HOBt and NMP in DMF; recycle 20 min at 12 ml/min); DMF washing (8 min at 6 ml/min). Cleavage from the resin and side-chain deprotection was achieved by treatment of the dried peptide–resin with 20 ml of a mixture of TFA/thioanisole/ethanedithiol/water/phenol (82.5/2.5/5/5 v/v) for 2 h at room temperature. The crude peptide was precipitated with cold diethylether and lyophilized. Yield: 210 mg (73%, calculated on the resin loading). The crude peptide was analysed by HPLC on a Beckman System Gold apparatus under the following conditions: Vydac C18 column (0.46×15 cm^2); eluent A, 0.1% TFA/water; eluent B, 0.1% TFA/acetonitrile; gradient from 5% to 65% B over 20 min; flow, 1 ml/min; detectionl UV, 210 nm; R_t, 13.9 min; HPLC purity, 57% (expressed as peak height %). The main peak was isolated by preparative HPLC, using a Vydac C18 column (2.2×25 cm^2); eluents A and B as indicated above; gradient from 20% to 44% B over 120 min; flow, 8 ml/min; detectionl UV, 210 nm. The final, purified peptide exhibited the correct amino acid ratios of the acid hydrolysate and an HPLC purity greater than 99%. The peptide was further characterized by ES-MS, yielding the correct molecular weight of 4242 DA.

CD Measurements

CD measurements were carried out on a JASCO model J-600A spectropolarimeter. All data were recorded and processed with a PC. A 2 nm bandwidth was used, with a scan-speed of 10 nm/min and a time constant of 4–8 s. Four to six scans were usually added in order to improve the signal to noise ratio. All measurements were carried out at 25°C, using quartz cells with path-length ranging from 0.001 to 0.1 cm. The spectra are reported in terms of mean residue molar ellipticity $[\Theta]_R$ (deg cm^2/dmol^2). The helix content of the peptide chain was estimated by the amplitude of the CD band at 222 nm according to the method of Fasman *et al.* [35].

NMR Measurements

All NMR experiments were performed at 298 K on a Bruker AM 400. The data were then processed on a Bruker X-32 workstation, with UXNMR software. The experiments were carried out at 1.6 mM peptide in 1:1 (v/v) H_2O/TFE-d_3 at pH 4–5 (uncorrected), using tetramethylsilane (TMS) as an internal standard. The CD spectrum of the solution used for the NMR studies was identical to that of the dilute solution used for CD measurements. Two-dimensional homonuclear ^{1}H-NMR spectra were recorded in the pure absorption mode according to the TPPI method, with 280 experiments of 2K data points. Prior to Fourier transformation the time domain data were multiplied by cosine squared in the t_1 dimension of all spectra, and by Gaussian window functions in the t_2 dimension of TOCSY and NOESY spectra. For DQF-COSY experiments a cosine squared function was used in the t_2 dimension also. The matrix of the data was zero-filled to 2048×1024 real points. Fifth order polynomial baseline correction was performed in the t_2 dimension after Fourier transformation. To perform the suppression of rapid pulsing artefacts appropriate phase cycles were used in the acquisition of DQF-COSY spectra. Two-dimensional homonuclear Hartmann-Hahn spectra were obtained as clean TOCSY. The MLEV-17 spin-lock sequence was applied with trim pulses of 2.5 ms each, using mixing times of 80 ms. The water signal was saturated by irradiation during the relaxation delay and, for NOESY spectra, also during the 200 ms mixing time. Inter-proton distances were calculated from the integrated volumes of non-ambiguous cross-peaks using the Aurelia software.

Modelling

Energy minimizations and molecular dynamics (MD) simulations were carried out with the program DISCOVER, version 2.80, implemented in the BIOSYM software package [36]. MD simulations were performed using the Consistent Valence Force Field

(CVFF) without cross- and Morse terms on a Silicon Graphics personal Iris computer. The program INSIGHT-II, version 2.1.2, was employed for model-building procedures and as a graphic interface. Standard force-field parameters for oligopeptides were taken from the INSIGHT residue library.

After energy minimization for initial relaxation using steepest descent for 250 iterations followed by 250 iterations of conjugate gradients minimization, the structure was further simulated by restrained MD at different temperaturs. Newton's equations of motion were numerically integrated applying the Verlet algorithm with an integration time step of 1 fs. The entire trajectory was updated every 1 ps, yielding an ensemble of 120 distinct conformations, thus covering a total simulation period of 120 ps. A dielectric permittivity of 78.0 was used in order to compensate for the problems generally associated with an overestimate of electrostatic long-range interactions for in vacuo simulations. The distinct distance values derived by volume-integration and calibration of NOE cross-peak intensities were transferred into distance ranges by imposing a tolerance of $\pm$ 10% onto the distances. The distance ranges were used as harmonic distance restraints during the restrained molecular dynamics simulations and subsequent energy minimizations.

RESULTS AND DISCUSSION

CD Results

The CD spectra of the GM–CSF analogue in aqueous solution and in the presence of incresing amounts of TFE are shown in Figure 3. In water the CD pattern is pH-independent in the range 3.9–9.6 (data not shown) and indicates the presence of a random structure with a very small amount of ordered conformation. Upon increasing the TFE content in the solvent mixture, the CD spectrum typical of the helical conformation is formed, converging to a plateau at 50% TFE. Under these conditions, the maximum helix content, estimated from the amplitude of the CD band at 222 nm [35], is around 45%. The plot of the molar ellipticity at 222 nm vs. the percentage of TFE in the solvent mixture is sigmoidal, characteristic of a cooperative conformational transition, the half-transition point occurring at 25% TFE.

The spectra do not define a single isodichroic point and indicate that we are not dealing with a simple two-component equilibrium system. The spectra recorded at TFE contents of 20% fit an isodichroic point located slightly below 200 nm, while the spectra recorded at lower TFE contents

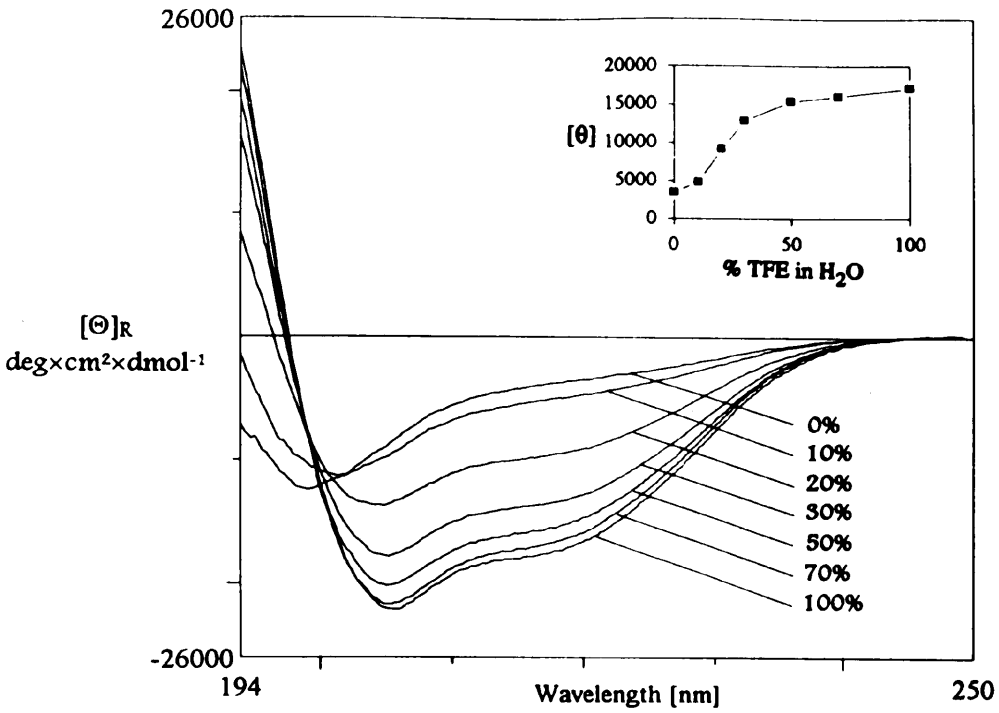

Figure 3 CD spectra of the hGM–CSF analogue in H_2O/TFE solutions at 2.5×10^{-5} M concentration in a 0.1 cm cell. In the inset, the profile of the residue molar ellipticity at 222 nm vs. TFE concentration is also shown.

intercross at 203 nm. These features are not typical of the coil–helix equilibrium and indicate that additional structural elements contribute to the optical activity.

NMR Results

Since the maximum helix content is almost reached in the presence of 50% TFE, NMR studies were carried out in a 1:1 water/TFE mixture at 1.6 mM peptide concentration. The solutions used for NMR experiments exhibited the same CD spectrum of the dilute solutions used for CD measurements. The spin systems of all amino acid residues were identified using standard 2D experiments. As an example, two representative portions of the NOESY spectra are shown in Figures 4 and 5. The assignment of proton resonances is reported in Table 1, and the summary of relevant NOESY connectivities is shown in Figure 6. Sequential and medium-range connectivities and the deviations from the random coil position of the Cα protons (Figure 7) indicate the presence of two helical segments corresponding to the helical domains A and D. Most important, the

Table 1 ^{1}H Chemical Shift Values at 298 K Relative to TMS for hGM-CSF[13]31]G-P-G[103-116]-NH_2 Peptide in H_2O/TFE-d_3

Residue	NH	CαH	CβH	CγH	CδH	Other H
Trp[1]		4.39	3.39,3.49	–	–	Aromatic 7.31,8.54
Glu[2]	8.59	4.47	1.99	2.11,2.40	–	–
His[3]	8.59	4.72	3.27	–	–	–
Val[4]	8.91	4.16	2.15	–	–	–
Asn5	8.37	4.76	2.93	–	–	–
Ala[6]	8.11	4.21	1.55	–	–	–
Ile[7]	7.74	3.89	2.03	1.30	0.99	–
Gln[8]	8.02	4.09	2.23	2.45,2.55	–	–
Glu[9]	8.12	4.04	2.20	2.52	–	–
Ala[10]	8.30	4.11	1.59	–	–	–
Arg[11]	7.92	4.14	2.07	1.70	3.25	NH_{ex} 7.24
Arg[12]	8.25	4.03	1.99	1.73	3.22	NH_{ex} 7.12
Leu[13]	8.56	4.13	1.91	1.65	0.93	–
Leu[14]	8.12	4.3	?		0.95	–
Asn[15]	8.01	4.54	2.92	–	–	–
Leu[16]	8.03	4.27	1.95	1.71	0.94	–
Ser[17]	8.15	4.02?	4.0	–	–	–
Arg[18]	7.74	4.31	1.91	1.70	3.22	NH_{ex} 7.17
Asp[19]	8.09	4.88	2.99,2.93	–	–	–
Gly[20]		4.18	–	–	–	–
Pro[21]	–	4.46	2.35	2.10,2.02	3.80,3.68	
Gly[22]	8.38	3.98	–	–	–	–
Phe[23]	8.12	4.59	3.29		–	–
Glu[24]	8.36	4.18	2.19	2.56	–	–
Ser[25]	8.12	4.02?		–	–	–
Phe[26]	7.98	4.42	3.27	–	–	
Lys[27]	8.12	3.88	1.86	1.54,1.41	1.73	
Glu[28]	8.06	4.25	2.20	2.58	–	–
Asn[29]	7.99	4.62	2.82	–	–	–
Leu[30]	8.01	4.15	1.67	1.55	0.89	–
Lys[31]	7.95	4.02	1.95	1.58,1.49	1.76	
Asp[32]	7.98	4.51	2.83,2.98	–	–	–
Phe[33]	8.07	4.42	3.26,3.31	–	–	
Leu[34]	8.25	4.17			0.95	
Leu[35]	7.91	4.33	1.91	1.83	0.98	
Val[36]E;	7.69	4.091	2.22	1.03	–	–

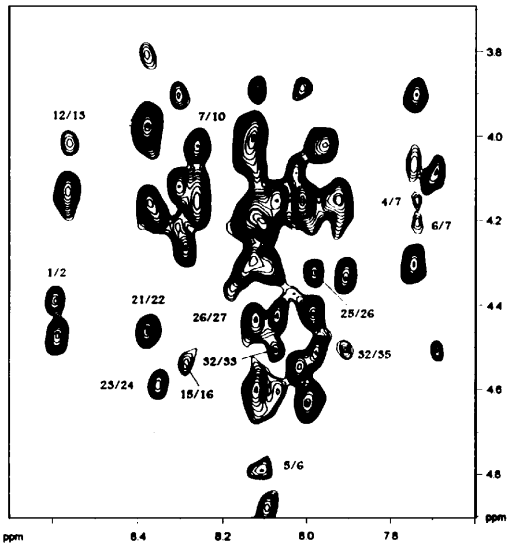

Figure 4 Fingerprint region of a NOESY spectrum (200 ms mixing time) recorded at 1.6 mM peptide concentration. Important NH–CH connectivities are indicated.

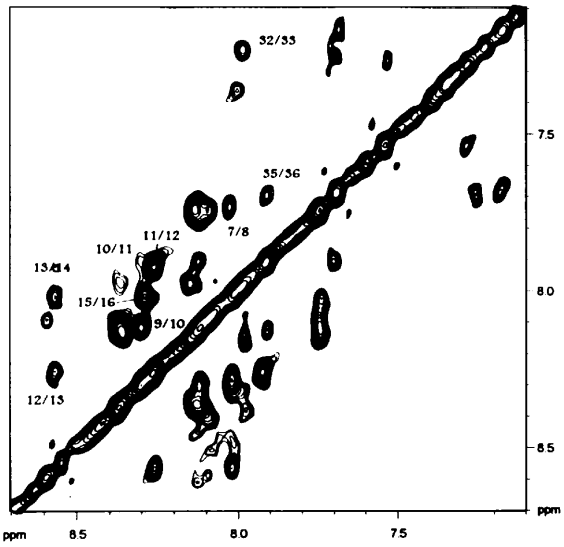

Figure 5 Amide region of the NOESY spectrum; sequential NH–NH connectivities are indicated.

long-range cross-peak connecting L^{13} and L^{35} indicates that the two helical segments must be in close proximity forming a hairpin structure.

Modelling

In order to get a better insight into the conformation of the analogue, structural refinement was carried out by restrained molecular dynamics simulations. The starting structure was constructed by extracting the relevant structural parts of the truncated hGM–CSF derivative from the X-ray structure of the native protein [12], retrieved from the Brookhaven Protein Data Bank (entry code: ICSG) [13]. Residues A13–A31 ('A' symbolizes the correspondence to monomer A of the crystallographically determined dimeric structure) belonging to helix A were linked by the tripeptide spacer unit Gly-Pro-Gly to the second segment, i.e. A103–A116 (helix D). Initially, only the local conformation of the bridging tripeptide and the flanking residues were relaxed in order to reduce the internal strain caused by the manual model-building procedure. While the backbone atoms of residues A13–A30 and A104–A116 were fixed to their X-ray coordinates, the spacer Gly-Pro-Gly together with residues A31 and A103 (notation from the X-ray structure) were allowed to move freely. The resulting

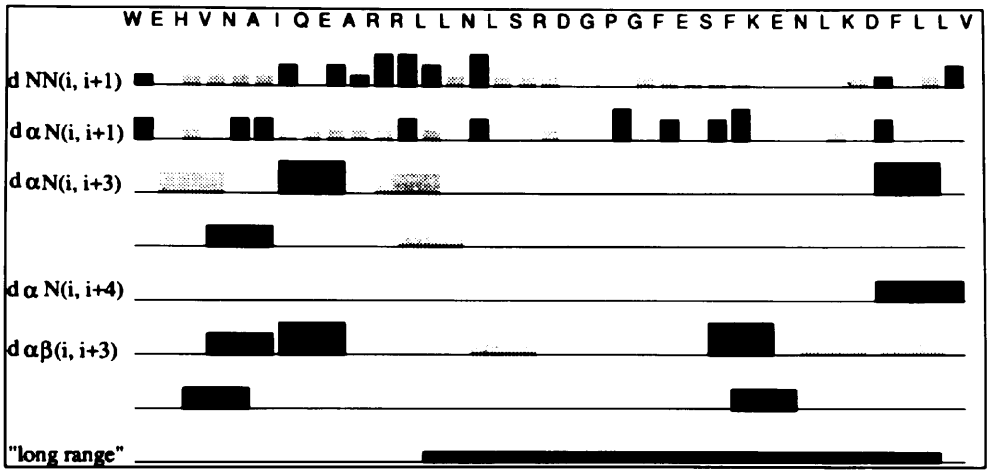

Figure 6 Summary of NOESY connectivities of the hGM–CFF analogue in 1:1 H_2O/TFE solution at 1.6 mM peptide concentration. Shaded areas indicate signal overlap.

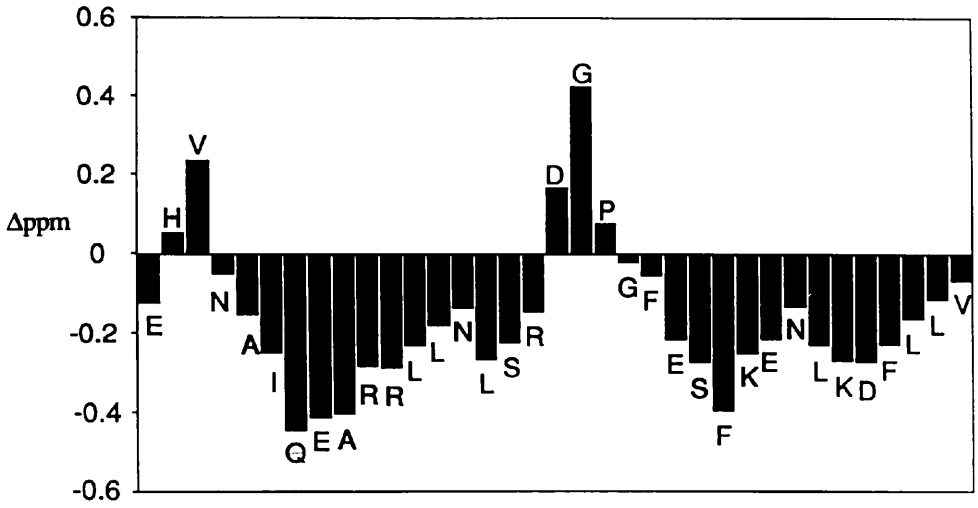

Figure 7 Difference between experimental C^{α} protons chemical shifts of the analogue in a 1:1 H_2O/TFE solution, and typical values for the random coil conformation, anlalysed according to the method of Pastore and Saudek [37]; cluster of five residues.

starting structure for restrained MD is shown in Figure 8(a). After the initial energy minimization of the starting structure, the NMR-derived distance restraints were introduced during the 120 ps MD simulation by gradually up-scaling the harmonic force constant to the target value of $K_{NOE} =$ 10,000 kcal/mol/Å. Simultaneously, the temperature was decreased from 600 K to 300 K. In total, 54 NOE-derived distances were introduced as restraints. During the first part of the simulation ϕ and ψ dihedral angles of the helical residues were forced to the ideal values ($-57°$, $-47°$) and the peptide bonds were kept at trans configuration ($\omega = +180°$) applying an harmonic dihedral forcing potential of $K_{\theta} = 100.0$ kcal/mol/rad^2. Within the last 30 ps of the simulation, ϕ and ψ dihedral restraints were removed.

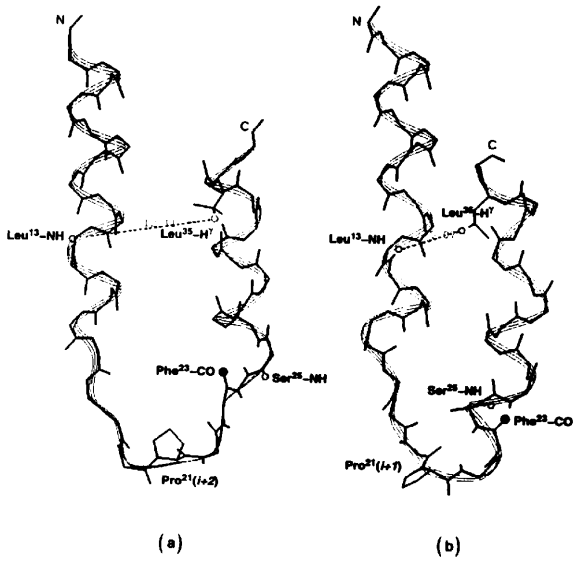

Figure 8 Comparison of selected structural features within the starting structure (a) and the averaged and minimized conformation (b). The inter-helix NMR-derived distance restraint (Leu^{13}-NH/LEU^{35}-H^{γ}). the position of Pro^{21} and atoms potentially forming the H-bond of the γ_i turn (Phe^{23}-Glu^{24}-Ser^{25}) are indicated.

The conformation obtained after averaging over the last 30 ps of the trajectory followed by energy minimization is shown in Figure 8(b). The backbone conformation in the final, refined structure is characterized by two α-helices comprising residues Glu^2 to Asn^{15} and Phe^{26} to Lys^{35}. These two helical segments are connected by a disordered loop region which allows the original, antiparallel orientation of the two helices to be maintained. The dihedral angles of the final refined structure and of the X-ray structure (compared to the values for the α-helix [38] are reported in Table 2. The restraint violations of the final structure and of the crystallographic structure are summarized in Table 3. The average restraint violation of the final structure was found to be 0.41 Å. In total, 13 out of 52 NOE-derived distance restraints violate the applied potential-free distance range. A violation exceeding 0.1 Å was found for seven distance restraints, while two distance restraints are found outside the 0.25 Å threshold; only one distance restraint deviates by more than 0.50 Å from the target value.

Table 2. Comparison of Dihedral Angles ϕ and ψ Within the Final Refined Structure and the X-ray Structure

Final structure	ϕ (°)	ψ (°)	X-ray structure	ϕ [°]	ψ [°]
Trp^{1}	–	– 20.18	Trp^{A13}	– 58.55	– 37.34
Glu^{2}	– 64.50	– 55.31	Glu^{A14}	– 74.23	– 29.58
His^{3}	– 75.56	– 47.12	His^{A15}	– 70.20	– 42.26
Val^{4}	– 58.77	– 52.27	Val^{A16}	– 68.89	– 41.92
Asn^{5}	– 57.27	– 37.24	Asn^{A17}	– 67.20	– 33.22
Ala^{6}	– 69.09	– 42.10	Ala^{A18}	– 71.12	– 35.97
Ile^{7}	– 70.75	– 42.95	Ile^{A19}	– 61.37	– 45.88
Gln^{8}	– 62.23	– 39.72	Gln^{A20}	– 64.88	– 36.71
Glu^{9}	– 68.99	– 38.76	Glu^{A21}	– 73.30	– 39.44
Ala^{10}	– 57.62	– 35.00	Ala^{A22}	– 57.26	– 49.63
Ala^{11}	– 63.23	– 41.60	Arg^{A23}	– 62.39	– 37.65
Arg^{12}	– 62.39	– 19.70	Arg^{A24}	– 62.72	– 60.41
Leu^{13}	– 67.91	– 50.75	Leu^{A25}	– 53.51	– 33.62
Leu^{14}	– 75.73	– 37.49	Leu^{A26}	– 70.55	– 53.63
Asn^{15}	– 70.69	– 46.91	Asn^{A27}	– 64.31	– 11.66
Leu^{16}	– 79.97	– 55.19	Leu^{A28}	– 91.18	7.92
Ser^{17}	– 87.54	111.92	Ser^{29}	– 70.08	112.92
Arg^{18}	– 132.10	105.83	Arg^{A30}	– 99.96	130.60
Asp^{19}	– 103.47	106.66	Asp^{A31}	– 136.29	147.57
Gly^{20}	117.95	88.69	Gly^{20a}	– 60.28	– 36.43
Pro^{21}	– 72.33	117.41	Pro^{21a}	– 83.73	– 174.47
Gly^{22}	132.37	– 77.49	Gly^{22a}	– 83.58	– 81.17
Phe^{23}	– 70.90	– 50.55	Phe^{A103}	– 56.97	– 41.00
Glu^{24}	– 86.64	79.01	Glu^{A104}	– 56.69	– 26.42
Ser^{25}	– 160.65	– 66.16	Ser^{A105}	– 95.55	– 25.24
Phe^{26}	– 65.07	– 43.54	Phe^{A106}	– 55.68	– 50.23
Lys^{27}	– 66.41	– 46.13	Lys^{A107}	– 60.71	– 46.84
Glu^{28}	– 58.40	– 43.54	Glu^{A108}	– 64.89	– 31.18
Asn^{29}	– 66.58	– 46.32	Asn^{A109}	– 61.78	– 46.92
Leu^{30}	– 60.00	– 37.78	Leu^{A110}	– 59.04	– 49.84
Lys^{31}	– 70.30	– 43.22	Lys^{A111}	– 64.13	– 26.38
Asp^{32}	– 27.47	– 53.52	Asp^{A112}	– 69.09	– 38.41
Phe^{33}	– 54.12	– 40.11	Phe^{A113}	– 63.06	– 56.03
Leu^{34}	– 63.97	– 32.18	Leu^{A114}	– 57.09	– 45.02
Leu^{35}	– 63.84	– 52.01	Leu^{A115}	– 67.79	– 0.13
Val^{36}	– 87.24	–	Val^{A116}	– 130.94	– 5.73

[a]Values taken from starting structure.

The important inter-helix NOE-derived distance between Leu^{13}-NH and Leu^{25}-Hγ (5.08 Å) deviates from the corresponding part of the X-ray structure by more than 7 Å. This particular NOE-derived distance accounts for a small difference in the relative orientation of the two helices when the NMR-derived conformation is compared to the structure of the native template protein (Figure 9). In fact, superposition of the two structures reveals that helix D in the solution structure is rotationally shifted with respect to the corresponding helix in the X-ray structure of the native protein. A rotation of about 50° (y-axis) of helix D in the crystal structure diminishes the indicative distance Leu^{A25}-NH/Leu^{A115}-Hγ from original 12.17 Å to 6.8 Å, i.e. close to the value found in the refined structure (Leu^{13}-NH/Leu^{35}-Hγ 5.08 Å). In the starting structure Pro^{21} occupies the $i+2$ position of the βI-like turn (Figure 8(a) and (b). In the NMR-derived conformation Pro^{21} is shifted upstream by one position, thus occupying the $i+1$ position of a βII-like turn. These conformations cannot be described strictly as regular β-turns,

Table 3 Restraint Violations of the Final Structure Compared with the X-ray Structure[a]

$Atom_I$	$Atom_j$	$_r$low	$_r$up	Final structure	Violation	X-ray structure	Violation
His^3-NH	Val^4-NH	2.98	3.64	2.78	0.20	2.54	0.44
Ala^{*6}-NH	Ile^7-NH	2.07	2.53	2.61	0.08	2.75	0.22
Ile^7-NH	Gln^8-NH	2.50	3.05	2.68	–	2.59	–
Glu^9-NH	Ala^{10}-NH	2.21	2.70	2.57	–	2.66	–
Arg^{11}-NH	Arg^{12}-NH	2.17	2.66	2.70	0.04	2.79	0.13
Arg^{12}-NH	Leu^{13}-NH	2.51	3.07	2.94	–	2.89	–
Leu^{13}-NH	Leu^{14}-NH	2.58	3.15	2.66	–	2.80	–
Asn^{15}-NH	Leu^{16}-NH	2.33	2.84	2.72	–	2.87	0.03
Phe^{26}-NH	Lys^{27}-NH	2.41	2.94	2.63	–	2.83	–
Leu^{35}-NH	Val^{36}-NH	2.60	3.18	2.80	–	2.92	–
Asn^{29}-NH	Asn^{29}-H^{α}	2.40	2.93	2.92	–	2.89	–
Arg^{11}-NH	$Arg^{11}H^{\alpha}$	2.32	2.84	2.95	0.11	2.90	0.06
Leu^{35}-NH	Leu^{35}-H^{α}	2.60	3.18	2.95	–	2.94	–
Leu^{16}-NH	Leu^{16}-H^{α}	2.60	3.17	3.01	–	3.05	–
Asn^{15}-NH	Asn^{15}-H^{α}	2.42	2.95	2.95	–	2.92	–
Asp^{19}-NH	Asp^{19}-H^{α}	2.63	3.22	3.06	–	3.03	–
Asp^{32}-NH	Asp^{32}-H^{α}	2.49	3.04	2.60	–	2.97	–
Phe^{33}-NH	Phe^{33}-H^{α}	2.49	3.04	2.83	–	2.91	–
Phe^{23}-NH	Phe^{23}-H^{α}	2.64	3.23	2.94	–	2.93	–
Val^4-NH	Val^4-H^{α}	2.17	2.65	2.87	0.22	2.93	0.28
Leu^{14}-NH	Leu^{14}-H^{α}	2.16	2.64	2.96	0.32	2.92	0.28
Leu^{30}-NH	Leu^{30}-H^{α}	2.28	2.79	2.92	0.13	2.83	0.04
Phe^{26}-NH	Phe^{26}-H^{α}	2.54	3.10	2.97	–	2.88	–
Ile^7-NH	Ile^7-H^{α}	2.64	3.23	2.97	–	2.94	–
Val^{36}-NH	Val^{36}-H^{α}	2.55	3.11	3.01	–	3.03	–
Arg^{18}-NH	Arg^{18}-H^{α}	2.47	3.02	3.07	0.05	3.08	0.06
Lys^{31}-NH	Lys^{31}-H^{α}	2.42	2.96	2.90	–	2.90	–
Ala^{10}-NH	Ala^{10}-H^{α}	2.27	2.78	2.85	0.07	2.88	0.10
Arg^{12}-NH	Arg^{12}-H^{α}	2.56	3.13	2.91	–	2.92	–
Leu^{13}-NH	Leu^{13}-H^{α}	2.49	3.04	2.98	–	2.85	–
Gln^8-NH	Gln^8-H^{α}	2.67	3.26	2.94	–	2.88	–
Lys^{27}-NH	Lys^{27}-H^{α}	2.79	3.41	2.94	–	2.88	–
Pro^{21}-H^{α}	Gly^{22}-NH	2.70	3.30	2.10	0.60	–	–
Trp^1-H^{α}	Glu^2-NH	3.06	3.74	3.37	–	3.59	–
Phe^{23}-H^{α}	Glu^{24}-NH	2.94	3.60	3.58	–	3.54	–
Asp^{32}-H^{α}	Phe^{33}-NH	3.31	4.05	3.63	–	3.57	–
Asn^{15}-H^{α}	Leu^{16}-NH	3.10	3.78	3.60	–	3.37	–
Asn^5-H^{α}	Ala^6-NH	3.20	3.91	3.54	–	3.56	–
Phe^{26}-H^{α}	Lys^{27}-NH	2.79	3.41	3.56	0.15	3.64	0.23
Ser^{25}-H^{α}	Phe^{26}-NH	2.88	3.52	3.61	0.09	3.55	0.03
Ala^6-H^{α}	Ile^7-NH	3.14	3.84	3.60	–	3.66	–
Arg^{12}-H^{α}	Leu^{13}-NH	2.91	3.55	3.46	–	3.63	0.08
Val^4-H^{α}	Ile^7-NH	3.22	3.93	3.72	–	3.28	–
Ile^7-H^{α}	Ala^{10}-NH	2.91	3.55	3.46	–	3.47	–
Asp^{32}-H^{α}	Leu^{35}-NH	3.17	3.88	3.72	–	3.60	–
His^3-H^{α}	Ala^6-H^{β}	3.14	4.38	3.56	–	3.63	–
Ile^7-H^{α}	Ala^{10}-H^{β}	2.63	4.31	3.36	–	3.38	–
Ser^{25}-H^{α}	Glu^{28}-H^{β}	2.38	4.43	3.78	–	3.82	–
Phe^{26}-H^{α}	Asn^{29}-H^{β}	3.14	4.28	3.48	–	2.83	0.31
Asn^{29}-H^{α}	Asp^{32}-H^{β}	3.40	4.56	4.23	–	3.93	–
Asp^{32}-H^{α}	Leu^{35}-H^{β}	3.14	4.28	4.08	–	3.82	–
Leu^{13}-NH	Leu^{35}-H^{γ}	3.00	5.00	5.08	0.08	12.17	7.17

[a]Distance values are given in Å

Table 4 Identification of Turn-like Structural Elements Within the Starting Structure and the Final Refined Structure

Central residue(s)	ϕ_{i+1}(°)	ψ_{i+1}(°)	ϕ_{i+2}(°)	ψ_{i+2}(°)	Turn type
Starting structure					
Starting structure					
$Gly^{20}(i+1) - Pro^{21}(i+2)$	− 60.28	− 36.43	− 83.73	− 174.47	βI
Final refined structure					
$Pro^{21}(i+1) - Gly^{22}(i+2)$	− 72.33	117.41	132.37	− 77.49	βII[a,β]
$Glu^{24}(i+1)$	− 86.64	79.01			γi[b]

[a] ψ_{i+2} does not meet the characteristical value of 0° for β-turns; no H-bonds are observed.
[b] The characteristic H-bond from Phe^{23}-CO (*i*) to Ser^{25}-NH ($i+2$) could be detected.

because in both cases the characteristic H-bond involving CO^{i}/NH^{i+3} is not formed, nor does ψ_{i+2} display the required torsional state of 0°. The βII turn formed by Gly^{20}(i)-$Pro^{21}(i+1)$-$Gly^{22}(i+2)$-$Phe^{23}(i+3)$ is followed by a γ_i turn comprising Phe^{23}(i)-$Glu^{24}(i+1)$-$Ser^{25}(i+2)$. The central Glu^{24} residue shows typical dihedral angles of the γ_i turn type as well as the characteristic H-bond between Phe^{23}(i)-CO and $Ser^{25}(i+2)$-NH (Table 4). The formation of the γ_i turn indicates that during the simulation the disordered loop region adopts a more structured conformation. In contrast to that effect, the helical stretches are interrupted at Arg^{12} and Asp^{32}, respectively. Within the second helix this causes a slight kink with respect to the helical axis.

Barlow and Thornton [8] have observed that α-helices in globular proteins may have a kink of 26°. Within the hGM–CSF analogue, helix 1 shows a kink of 22°, and helix 2 of 23° (Table 5). The central hinge is located at position 12 (Arg^{12}) within helix 1 and Asp^{32} within helix 2, respectively.

A further structural feature observed in coiled coil proteins is the inter-helical crossing angle, a pseudo-dihedral angle between the axes of two helices with a typical value of 28° in fibrous proteins [4, 7]. In globular proteins the crossover angles may vary [39], but are usually found at 20°, or − 50° [40, 41]. In the NMR-derived structure the value of the pseudo-dihedral angle is 14.6°, very close to the value of 20° observed in the crystal structure of hGM–CSF [12]. Thus, the introduction of the tripeptide linker does not substantially affect the interactions between helices A and D observed in the native structure. Helix–helix interactions in the refined structure are not mainly dominated by the classical packing of apolar residues at the helical interface [41]. Residues in helix 1, potentially available for the interaction with helix 2, are indicated in Figure 10. Only 6 of the 12 residues oriented towards the helical interface show hydrophobic characteristics

Table 5 Summary of Helix Characteristics in the Final Refined Structure

Helical kink	Crossing angle	Closest C^{α}/C^{α} distances (Å)
Helix 1 (A): 22°	14.6°	Ile^{7}/Val^{36} 5.16
Helix 2 (D): 23°		Ala^{10}/Leu^{35} 5.85
		Arg^{11}/Asp^{32} 5.37

Helical segments: Glu^{2}-Arg^{11}, Leu^{13}-Asn^{15} (helix A); Phe^{26}-Lys^{31}, Phe^{33}-Leu^{35} (helix D)
Residues located at the inter-helix interface: His^{3}, Val^{4}, Ile^{7}, Gln^{8}, Ala^{10}, Arg^{11}, Leu^{14}, Asn^{15}, Glu^{28}, Asp^{32}, Key^{35}, Val^{36}

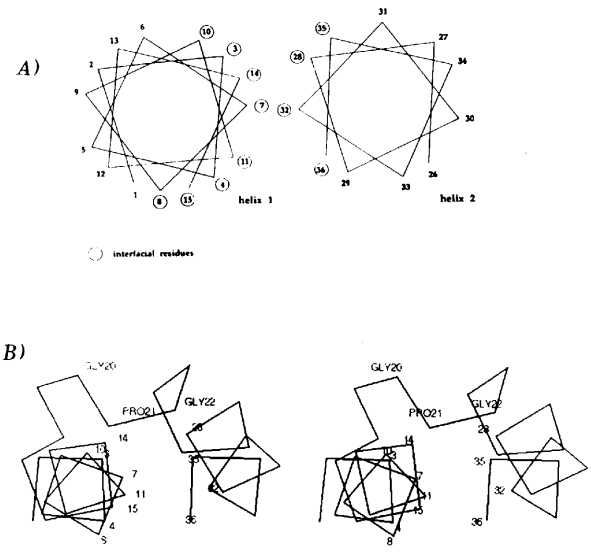

Figure 10 Presentation of residues oriented towards the helix interface. (A) Relative orientation of the helical wheels of antiparallel helix 1 (top view) and helix 2 (bottom view). (B) Stereoview of *csf_avem*. Only H^{α} atoms are shown; interfacial residues are indicated.

(Table 5). Short C^{α}/C^{α} distances are observed for Ile^{7}/Val^{36}, Ala^{10}/Leu^{35} and Arg^{11}/Asp^{32} (Table 5). The first two pairs of amino acids show clearly hydrophobic interactions between their side chains, while in the third pair the side chains are not oriented towards each other in order to form a salt bridge. The antiparallel orientation, which is favoured packing mode for α-helices in proteins [9, 42–45], is preinduced by the artificially introduced covalent linkage.

CONCLUSIONS

The results presented in this paper, based upon CD, NMR and NOE-restrained molecular dynamics calculations, indicate that the designed synthetic analogue in a 1:1 TFE/water mixture assumes a three-dimensional structure in which helices A and D are oriented in a very similar manner as in the crystal structure of hGM–CSF. In our previous paper it was shown that peptide fragments corresponding to the sequences of helices A, B, C and D of hGM–CSF are able to fold into the helical conformation in a 1:1 TFE/water mixture [46]. However, none of the isolated helical fragments interacts with the receptor system. In contrast, preliminary binding assays *in vitro* indicated that the analogue investigated in the present paper, comprising helices A and D, does bind to the receptor. Thus, the conformation preference of the analogue in a 1:1 TFE/water mixture, with helix packing very close to that of the native structure, is of biological relevance. A full characterization of the analogue in terms of biological activity is in progress and will be reported elsewhere.

Acknowledgements

This work was carried out with the financial support of Consiglio Nazionale delle Richerche (CNR). Thanks are due to Dr M. Hamdan, Glaxo-Wellcome Research Laboratories of Verona (Italy), for the mass spectrometry analysis. The technical assistance of Mr Silvio Da Rin Fioretto is gratefully acknowledged.

REFERENCES

1. M. Tomonaga, D. W. Golde and J. C. Gasson (1986). Biosynthetic (recombinant) human granulocyte-macrophage colony-stimulating factor: Effect on normal bone marrow and leukemia cell lines. *Blood 67*, 31–36.
2. G. Morstyn and A. W. Burgess (1988). Hemopoietic growth factors: A review. *Cancer Res. 48*, 5624–5637.
3. S. R. Sprang and J. F. Bazan (1993). Cytokine structural taxonomy and mechanisms of receptor engagement. *Curr Opin. Struct. Biol. 3*, 815–827.
4. S. R. Presnell and F. E. Cohen (1989). Topological description of four-helix bundles. *Proc. Natl Acad. Sci. USA 86*, 6592–6596.
5. D. A. Rozwarski, A. M. Gronenbron, G. M. Clore, J. F. Bazan, A. Bohm, A. Wlodawer, M. Hatada and P. A. Karplus (1994). Structural comparison among the short-chain helical cytokines. *Structure 2*, 159–173.
6. H. R. Mott and I. D. Campbell (1995). Four-helix bundle growth factors and their receptors: Protein–protein interactions. *Curr. Opin. Struct. Biol. 5*, 114–121.
7. P. C. Weber and F. R. Salemme (1980). Structural and functional diversity in 4-α-helical proteins. *Nature 287*, 82–84.
8. D. J. Barlow and J. M. Thornton (1988). Helix geometry in proteins. *J. Mol. Biol. 201*, 601–619.
9. K. C. Chou, G. M. Maggiora, G. Némethy and H. A. Scheraga (1988). Energetics of the structure of the four-α-helix bundle in proteins. *Proc. Natl Acad Sci. USA 85*, 4295–4299.
10. D. A. D. Parry, E. Minasian and S. J. Leach (1988), Conformational homologies among cytokines: interleukins and colony-stimulating factors. *J. Mol. Recogn. 1*, 107–110.
11. K. Diederichs, S. Jacques, T. Boone and P. A. Karplus (1991). Low-resolution structure of recombinant human granulocyte-macrophage colony-stimulating factor. *J. Mol. Biol. 221*, 55–60.
12. M. R. Walter, W. J. Cook, S. E. Ealick, T. L. Nagabhushan, P. P. Trotta and C. E. Bugg (1992). Three-dimensional structure of recombinant human granulocyte-macrophage colony-stimulating factor. *J. Mol Biol. 24*, 1075–1085.
13. F. C. Bernstein, T. F. Koetzle, G. J. B. Williams, E. F. Meyer Jr, M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi and M. Tasumi (1977). The protein data bank: A computer-based archival file for macromolecular structures. *J. Mol. Biol. 112*, 535–543.
14. J. C. Gasson, S. E. Kaufman, R. H. Weisbart, M. Tomonaga and D. W. Golde (1986). High-affinity binding of granulocyte-macrophage colony-stimulating factor to normal and leukemic human myeloid cells. *Proc. Natl Acad. Sci. USA 83*, 669–673.
15. J. DiPersio, P. Billing, S. Kaufman, P. Eghtesady, R. E. Williams and J. C. Gasson (1988). Characterization of the human granulocyte-macrophage colony-stimulating factor receptor. *J. Biol. Chem. 263*, 1834–1841.
16. S. A. Cannistra, P. Groshek, R. Garlick, J. Miller and J. D. Griffin (1990). Regulation of surface expression of the granulocyte-macrophage colony-stimulating factor receptor in normal human myeloid cells. *Proc. Natl Acad. Sci. USA 87*, 93–97.

17. J. F. Bazan (1990). Structural design and molecular evolution of cytokine receptor superfamily. *Proc. Natl Acad. Sci. USA 87*, 6934–6938.
18. N. Sato and A. Miyajima (1994). Multimeric cytokine receptors: Common versus specific functions. *Curr. Opin. Cell Biol. 6*, 174–179.
19. A. M. De Vos, M. Ultsch and A. A. Kosiakoff (1992). Human growth hormone and extracellular domain of its receptor: Crystal structure of the complex. *Science 255*, 306–312.
20. I. Clark-Lewis, A. F. Lopez, L. B. To, M. A. Vadas, J. W. Schrader, L. E. Hood and S. B. Kent (1988). Structure–function studies of human granulocyte-macrophage colony-stimulating factor: Identification of residues required for activity. *J. Immunol. 141*, 881–889.
21. K. Kaushansky, S. G. Shoemaker, S. Alfaro and C. Brown (1989). Hematopoietic activity of granulocyte/macrophage colony-stimulating factor is dependent upon two distinct regions of the molecule: functional analysis based upon the activities of interspecies hybrid growth factors. *Proc. Natl Acad. Sci, USA 86*, 1213–1217.
22. C. B. Brown, C. E. Hart, D. M. Curtis, M. C. Bailey and K. Kaushansky (1990). Two neutralizing monoclonal antibodies against human granulocyte-macrophage colony-stimulating factor recognize the receptor binding domain of the molecule. *J. Immunol. 144*, 2184–2189.
23. G. F. Seelig, W. W. Prosise, J. E. Scheffler, T. L. Nagabhushan and P. P. Trotta (1990). Evidence for a direct involvement of the carboxyl terminus of human granulocyte-macrophage colony-stimulating factor in receptor binding. *J. Cell Biochem. 14C*, 246 (Suppl.).
24. K. Diederichs, T. Boone and P. A. Karplus (1991). Novel fold and putative receptor binding site of granulocyte-macrophage colony-stimulating factor. *Science 254*, 1779–1782.
25. A. B. Shanafelt, K. E. Johnson and R. A. Kastelein (1991). Identification of critical amino acid residues in human and mouse granulocyte-macrophage colony-stimulating factor by scanning-deletion analysis. *J. Biol. Chem. 266*, 13804–13810.
26. Y. Kanakura, S. A. Cannistra, C. B. Brown, M. Nakamura, G. F. Seelig, W. W. Prosise, J. C. Hawkins, K. Kaushansky and J. D. Griffin (1991). Identification of functionally distinct domains of human granulocyte-macrophage colony-stimulating factor using monoclonal antibodies. *Blood 77*, 1033–1043.
27. H. J. Meropol, S. W. Altmann, A. B. Shanafelt, R. A. Kastelein, G. D. Johnson and M. B. Prystowsky (1992). Requirement of hydrophilic amino-terminal rsidues for granulocyte-macrophage colony-stimulating factor bioactivity and receptor binding. *J. Biol. Chem. 267*. 14266–14269.
28. G. F. Seelig, W. W. Prosise and J. E. Scheffler (1994). A role for the carboxyl terminus of human granulocyte-macrophage colony-stimulating factor in the binding of ligand to the alpha subunit of the high affinity receptor. *J. Biol. Chem. 269*, 5548–5553.
29. S. C. Barry, C. J. Bagley, J. Philips, M. Dottore, B. Cambareri, P. Moretti, R. Dandrea, G. J. Goodall, M. F. Shannon, M. A. Vadas and A. F. Lopez (1994). Two contiguous residues in human interleukin 3, Asp-21 and Glu-22, selectively interact with the alpha and beta chains of its receptor and participates in function. *J. Biol. Chem. 269*, 8488–8492.
30. P. Lock, D. Metcalf and N. A. Nicola (1994). Histidine-367 of the human common beta chain of the receptor is critical for high-affinity binding of human granulocyte-macrophage colony-stimulating factor. *Proc. Natl Acad. Sci USA 91*, 252–256.
31. C. Monfardini, T. Kieber-Emmons, D. Voet, A. P. Godillot, D. B. Weiner and W. V. Williams (1996). Rational design of granulocyte-macrophage colony-stimulating factor antagonist peptides. *J. Biol. Chem. 271*, 2966–2971.
32. J. S. Richardson and D. C. Richardson (1988). Amino acid preferences for specific locations at the ends of α-helices. *Science 240*, 1648–1652.
33. P. Y. Chou and G. D. Fasman (1978). Empirical prediction of protein conformation. *Annu. Rev. Biochem. 47*, 251–276.
34. W. F. DeGrado, Z. R. Wasserman and J. D. Lear (1989). Protein design, a minimalist approach. *Science 243*, 622–628.
35. N. Greenfield and G. D. Fasman (1996). Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry 8*, 4108–4116.
36. P. Daubner-Osguthorpe, V. A. Roberts, D. J. Osguthorpe, J. Wolff, M. Genes and A. T. Hagler (1988). Structure and energetics of ligand binding to proteins: *Escherichia coli* dehydrofolate reductase-trimethoprim, a drug-receptor system. *Proteins Struct. Function 4*, 31–47.
37. A. Pastore and V. Saudek (1990). The relationship between chemical shift and secondary structure in proteins. *J. Magn. Reson. 90*, 165–176.
38. J. S. Richardson (1981). The anatomy and taxonomy of protein structure. *Adv. Protein Chem. 34*, 168–339.
39. C. Cohen and D. A. D. Parry (1990). α-Helical coiled coils and bundles: How to design an α-helical protein. *Proteins: Struct. Function Gen. 7*, 1–15.
40. F. H. C. Crick (1953). The packing of α-helices: Simple coiled-coils. *Acta Crystallogr. 6*, 689–697.
41. A. G. Murrain and A. V. Finkelstein (1988). General architecture of the α-helical globul. *J. Mol. Biol. 204*, 749–769.
42. W. G. J. Hol (1985). The role of the α-helix dipole in protein function and structure. *Prog. Biophys. Mol. Biol. 45*, 149–195.
43. W. G. J. Hol, L. M. Halie and C. Sander (1981). Dipoles of the α-helix and β-sheet: Their role in protein folding. *Nature 294*, 532–536.

44. K. R. Shoemaker, P. S. Kim, E. J. York, J. M. Stewart and R. L. Baldwin (1987). Tests of the helix dipole model for stabilization of α-helices. *Nature 326*, 563–567.
45. M. Levitt and C. Chothia (1976). Structural patterns in globular proteins. *Nature 261*, 552–558.
46. S. Fiori, S. Mammi, E. Peggion, P. Rovero, S. Pegoraro and R. P. Revoltella (1997). Conformational of four peptides corresponding to the α-helical segments of human GM–CSF. *J. Pept. Sci. 3*, 000–000.