Volume 314, number 3, 435–439 FEBS 11908 December 1992

Secondary structure of human granulocyte colony-stimulating factor derived from NMR spectroscopy

Thomas Zink[a], Alfred Ross[a], Dorothee Ambrosius[b], Rainer Rudolph[b] and Tad A. Holak[a]

[a]*Max-Planck-Institut für Biochemie, D-8033 Martinsried, Germany* and [b]*Boehringer Mannheim GmbH, Biochemical Research Center, D-8122 Penzberg, Germany*

Received 6 November 1992

Recombinant ^{15}N-, ^{13}C-labeled human granulocyte colony-stimulating factor (rh-metG-CSF) has been studied by 2D and 3D NMR using uniformly labeled protein as well as residue-specific ^{15}N-labeled samples. Assignment of the ^{1}H, ^{15}N backbone, and 60% ^{1}H sidechain resonances has enabled the determination of the secondary structure of the protein. The secondary structure is dominated by α-helical regions with four stretches of helices between residues 11–41, 71–95, 102–124 and 144–170.

Nuclear magnetic resonance; Secondary structure; Human granulocyte colony-stimulating factor

1. INTRODUCTION

Human granulocyte colony-stimulating factor (h-G-CSF), a 174-residue glycoprotein (MW = 19 kDa, pI = 6.1), is one of the hematopoietic growth factors that play a major role in the process of blood formation [1]. The protein is produced in monocytes and fibroblasts. It serves three purposes: it causes the proliferation of hematopoietic precursor cells, induces the differentiation of the precursor cells into mature neutrophilic granulocytes, and activates neutrophilic granulocytes. The sugar chain, linked to Thr-133-$^{\gamma}O$, is not required for biological activity [2,3]. Circular dichroism measurements showed that h-G-CSF is largely α-helical with little or no β-structure [4]. Here we describe the secondary structure of rh-metG-CSF, a recombinant h-G-CSF which differs from the native protein at the N-terminal residue (methionine). Sequential assignment with conventional 2D NMR techniques [5] could not be carried out due to severe overlap of resonances in the 2D spectra for a protein of this size. The small variation of the chemical shifts of the α- and NH-protons, a consequence of the predominantly α-helical structure, further complicated the assignment. The sequential assignment of the backbone proton and nitrogen resonances was accomplished with heteronuclear 2D and 3D spectra for both the uniformly ^{15}N-labeled and the residue specific ^{15}N-labeled samples, mostly from the 3D NOESY-HMQC spectra [6–8]. The equivalent TOCSY spectra provided only limited information as most of the expected signals were weak or missing due to short proton T_2-relaxation times for rh-metG-CSF. The linewidths of the proton resonances were typically 30 Hz; slightly but noticeably larger than the expected linewidths for a protein of this size. We also recorded the HNCA triple resonance experiment [9] for a uniformly ^{15}N-,^{13}C-labeled sample. This experiment enabled identification of residues not located in α-helices.

Correspondence address: T.A. Holak, Max-Planck-Institut für Biochemie, D-8033 Martinsried, Germany. Fax: (49) (89) 8578-3777.

2. MATERIALS AND METHODS

The rh-metG-CSF was expressed in the cytoplasm of *E. coli* as insoluble inclusion bodies. The correctly folded protein was extracted from the inclusion bodies by renaturation of the unfolded protein in a 0.8 M arginine buffer after solubilization of the inclusion bodies in 6 M guanidinium chloride [10,11]. The final step of the preparation involved conventional column chromatography.

For the uniformly ^{15}N-labeled sample, rh-metG-CSF was expressed in the *E. coli* strain TG 1 in a minimal medium [12]. The medium was slightly modified by substituting $(NH_4)_2HPO_4$ for [^{15}N]ammonium chloride (at a concentration of 1 g/l) as the sole source of nitrogen. The purified protein was dialyzed against sodium phosphate (4 mM, pH 4.0) and lyophilized. The samples were typically 1–2 mM in protein and 50–60 mM in sodium phosphate, after addition of 0.5 ml H_2O/D_2O (9:1) or 0.5 ml D_2O to the lyophilized powder, and after adjustment of the pH to 3.5. This low pH value was necessary due to the low solubility of rh-metG-CSF at higher pH values. For the residue specific ^{15}N labeling, the *E. coli* strain C600+ was used in a minimal medium enriched with all 20 amino acids [13]. The ^{15}N-labeled amino acid was added at concentrations between 0.25 and 1.0 g/l, half of which was added to the medium at the beginning of fermentation, the other half on induction. The rh-metG-CSF was labeled with [^{15}N]glycine, [^{15}N]leucine, [^{15}N]valine, [^{15}N]alanine and [^{15}N]lysine. Cross-labeling of serine was achieved by a deliberate omission of serine from the amino acid-enriched medium during labeling with [^{15}N]glycine. The uniformly ^{15}N-,^{13}C-labeled sample was expressed in a minimal medium containing 1 g/l [$^{13}C_6$]glucose as the sole carbon source [12]. All labeled preparations were purified by the procedure outlined above. Table 1 summarizes the various labeled samples prepared and the cross labeling observed [14].

The NMR spectra were recorded on a Bruker AMX600 spectrometer operating at 600.14 MHz for protons. For all experiments the temperature was set to 27°C. Spectra were acquired and processed in the phase sensitive mode using time proportional phase incrementation (TPPI) [15] in all dimensions. Suppression of the water signal was achieved by presaturation or jump–return pulses [16–18]. The mixing time employed for the 3D NOESY-HMQC [6] was set to 120 ms; the MLEV-17 mixing sequence used for the 3D TOCSY-HMQC [6] spectra was set to 40 ms and 70 ms. 2D spectra were processed with UXNMR; the 3D spectra were processed with our in-house software (CC-NMR).

3. RESULTS AND DISCUSSION

The predicted α-helical structure for G-CSF has been confirmed by 2D NOESY spectrum. There are about 200 NH–NH cross peaks (on both sides of the diagonal) in the spectrum. The NH–NH connectivities are characteristic of the α-helical structure [5,19]. Further indication of the helical structure is provided by chemical shift ranges of α-protons (Fig. 1). From the total of 176 $C^{\alpha}H$ resonances that are expected to be present in the spectrum, only five appeared below 4.7 ppm (the chemical shift of water). This upfield shift of 0.2–0.6 ppm relative to random coil values is characteristic of helical structures [20]. In the ^{1}H–^{15}N 3D NOESY-HMQC spectra the 2D overlap was mostly resolved, although some degree of overlap still remained. Fig. 2 shows the 2D NOESY projection of the 3D NOESY-HMQC spectrum. In the 3D TOCSY-HMQC spectrum, the only visible magnetization transfers were from the NH- to the α-protons. The 3D TOCSY-HMQC with a mixing time of 70 ms showed more cross peaks than the same experiment using 40 ms, and also few peaks from the

Table 1

^{15}N labeling of rh-metG-CSF by amino acid type

^{15}N-labeled amino acid in the medium	Cross-labeling observed
[^{15}N]Ala	–
[^{15}N]Gly[a]	Ser
[^{15}N]Gly[b]	Ser
[^{15}N]Leu	–
[^{15}N]Val,[α-^{15}N]Lys	Ala

[a] [^{14}N]Ser added to the medium; low level of cross-labeling.
[b] [^{14}N]Ser omitted from the medium; deliberate high level of cross-labeling.

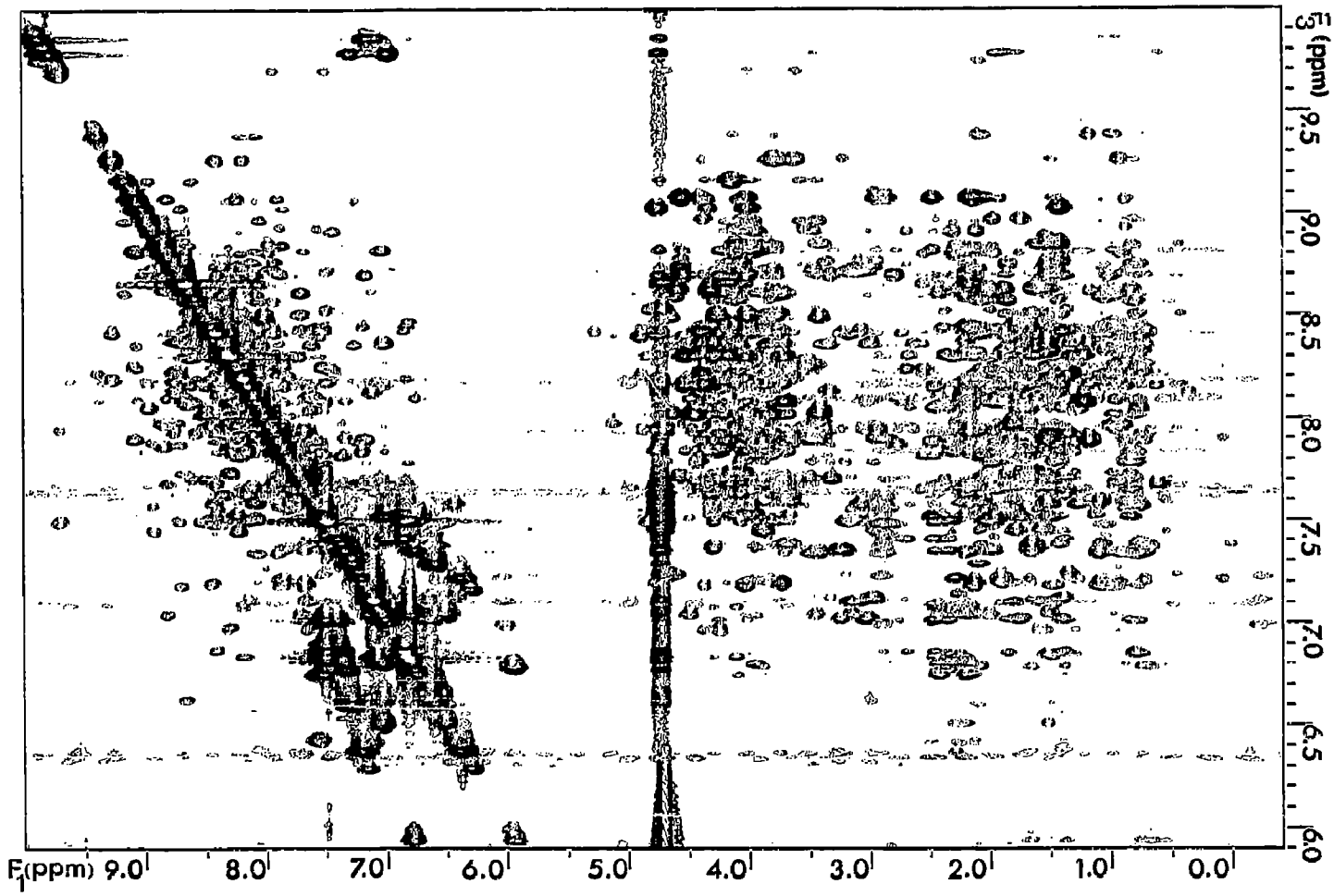

Fig. 1. The 2D-NOESY projection of the 600 MHz 3D NOESY-HMQC spectrum of [^{15}N]rh-metG-CSF.

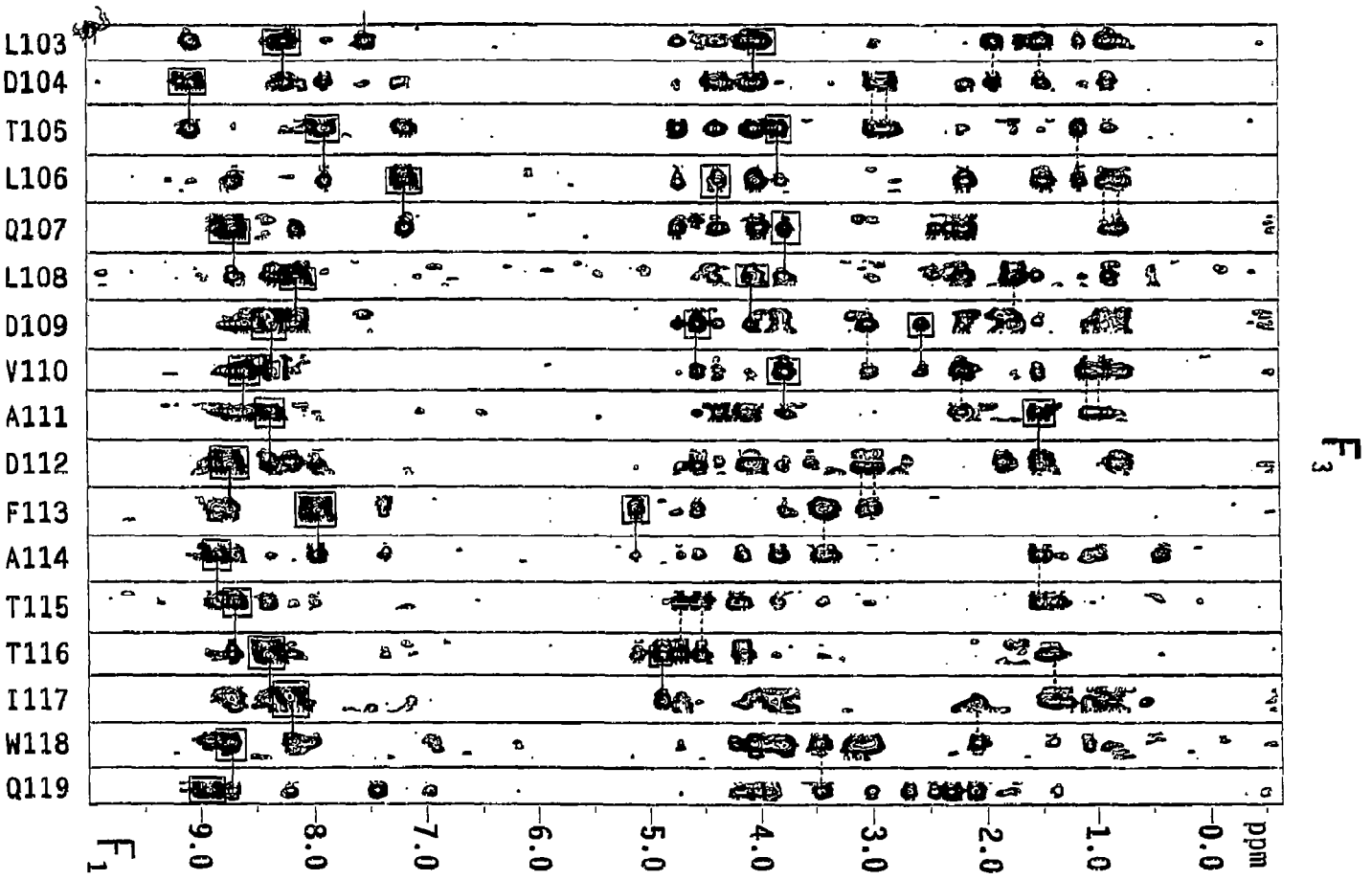

Fig. 2. Strips from the 3D NOESY-HMQC spectrum of [^{15}N]rh-metG-CSF showing resonances of residues L103-Q119 belonging to an α-helix. The strips, 8–11 points (50–66 Hz) wide, are taken from the 3D NOESY-HMQC at the ^{15}N- and ^{15}NH-chemical shift of the residues of interest. Rectangles mark intraresidual peaks identified from the 3D TOCSY-HMQC. Solid lines indicate sequential NH(i)–NH(i−1), NH(i)–C$^{\alpha}$H(i−1), NH(i)–C$^{\beta}$H(i−1) and NH(i)–C$^{\gamma}$(i−1) connectivities. Dashed lines are drawn when the peaks from residue (i−1) could not be identified as intraresidual in the 3D TOCSY-HMQC.

NHs to β-protons. Altogether 105 signals could be identified as intraresidual α-protons, while only 38 β and γ resonances were seen. Therefore the majority of the spin systems were identified with residue specific ^{15}N labeling.

The sequential assignment was started with the NH–NH connectivities. Most of the residues showed strong NH(i)–NH(i−1) and NH(i)–NH(i+1) cross peaks. The directionality of the sequence could often be determined by a C$^{\alpha}$H(i)–NH(i+1) cross peak between the residues involved in the NH(i)–NH(i+1) connectivity. Also, C$^{\beta}$H–NH(i+1) peaks could sometimes be identified and served as further proof for the sequential assignment. The C$^{\alpha}$H–NH(i+3) cross peaks were identified in 90% of residues that are in the helical conformation. By this procedure, protein segments varying in length from three to nine residues were found. All these segments contained the residue specific ^{15}N-labeled amino acids. Fig. 2 shows an example of the sequential assignment in the 3D NOESY-HMQC spectrum for residues Leu-103 to Gln-119. In Fig. 3, a schematic representation of the sequential and medium-range NOEs is displayed. It is clear from the pattern of short- and medium-range NOEs that the secondary structure of rh-metG-CSF in solution is dominated by four helices. The precise definition of the start and the end of helices depends upon criteria used and has to await detailed structure calculations. Taking into account that the length of a helix may vary by one or two residues, the four helices are defined as follows: 11–41, 71–95, 102–124 and 144–170.

Turns, which are also characterized by NH–NH NOEs, can be distinguished from helices by the chemical shift of the α-protons and by characteristic C$^{\alpha}$H(i)–NH(j) connectivities [19]. The chemical shifts of α-protons in turns are similar to random coil values, while in helices an upfield shift is observed. The NH–NH, C$^{\alpha}$H(i)–NH(j) connectivities, and characteristic random coil chemical shifts are observed for residues 66–70, 129–131, 133–137 and 139–143. Hence these residues are likely to be in turns rather than in other non-helical structures. In the ^{15}NH–^{13}C-planes of the 3D triple resonance HNCA experiment [9] two signals are expected

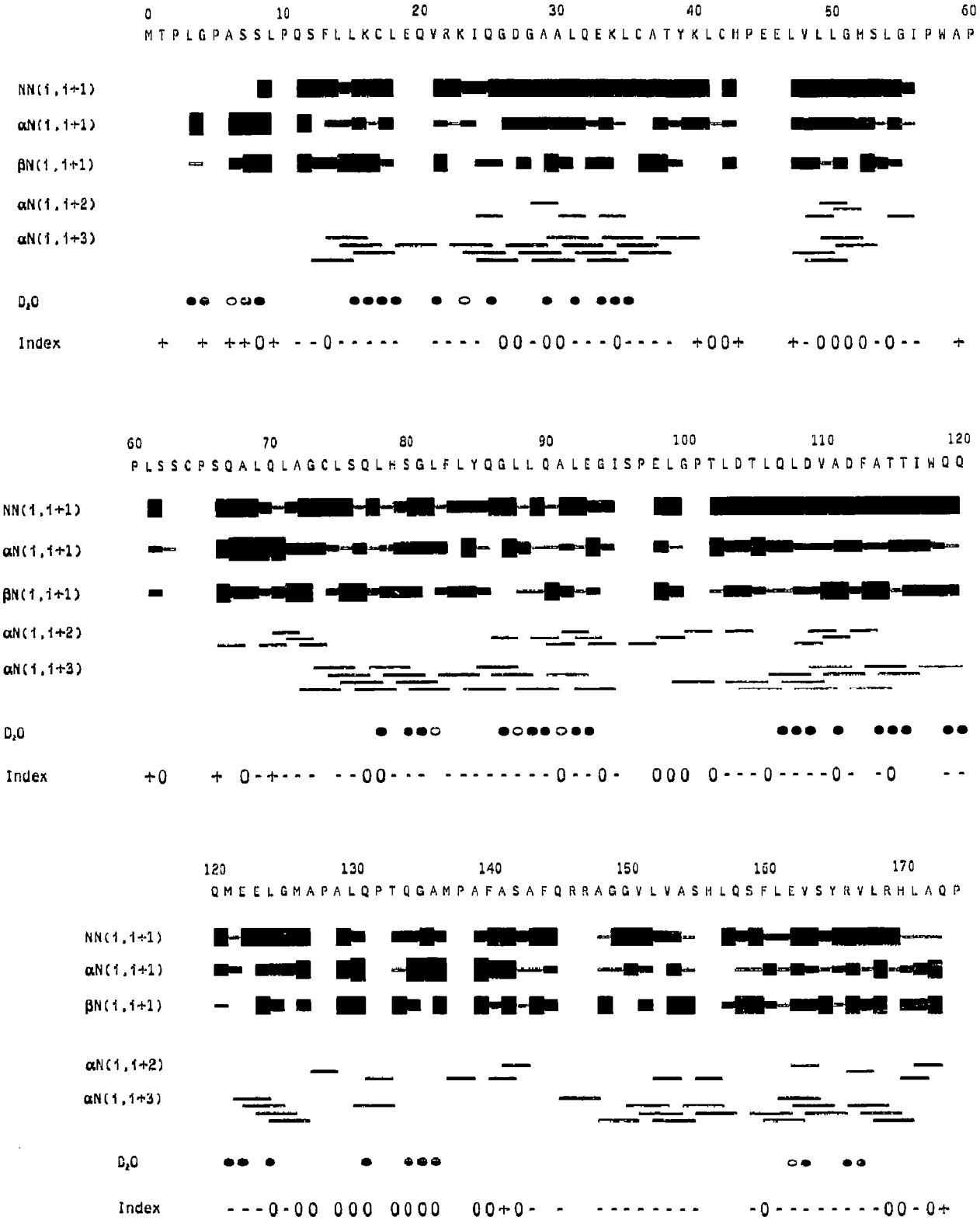

Fig. 3. Summary of the short-range NOEs. The intensities of the NOEs are reflected in the thickness of the bars (except for the $d_{\alpha N}(i,i+2,3)$ connectivities). The $C^{\alpha}H(i)$–$C^{\gamma}H(i-1)$(Pro) NOE is shown along the same line as the $C^{\alpha}H(i)$–NH(i+1) connectivities. Filled bars indicate the data that could be semiquantified, open bars indicate the data that could not be semiquantified due to spectral overlap. Circles indicate residues that show slow NH-exchange and were therefore still seen in the spectra in D_2O after 2 days. In cases where two residues have both identical ^{15}NH- and ^{15}N-chemical shift circles are left open. The bottom row shows an index that is a measure of the $C^{\alpha}H$-shift relative to random coil values, + and − indicating a shift to higher and lower ppm values, respectively. The 0 is given where no or only a small shift was observed.

for a residue in a non-helical conformation: the intraresidual connectivity NH(i)–C^{α}(i) and the interresidual connectivity NH(i)–C^{α}(i−1). Due to the larger heteronuclear one-bond coupling constant ($^1J_{CN}$ = 11 Hz) the intensity of the former signal is expected to be stronger than that of the latter ($^2J_{CN}$ = 7 Hz). This was indeed observed for residues in fragments 3–4 and 6–9. The same pattern was also observed for residues 66–69, 129–130, 133–137 and 139–143 which were already identified as being in turns from the 3D NOESY-HMQC data. In contrast to the predicted coupling patterns of the non-helical amino acids, residues located in helices should exhibit no or only a weak intraresidual NH(i)–C^{α}H(i) peak. This is due to the decreased coupling constants $^1J_{CN}$ and $^2J_{CN}$ which results in a decreased magnetization transfer from the amide proton to the alpha carbons. Indeed, only few NH(i)–C^{α}H(i) peaks were observed for the helical residues in G-CSF.

In this paper we have described the secondary structure of rh-metG-CSF. The highlight of the G-CSF structure is the presence of four helices between residues 11–41, 71–95, 102–124 and 144–170. Turns are located at residues 66–70, 129–131, 133–137 and 139–143. Calculations of the 3D structure will give a more detailed picture of the 3D solution structure of rh-metG-CSF.

Acknowledgements: This work was supported by a research grant from the Bundesministerium für Forschung und Technologie.

REFERENCES

[1] Nomura, H., Imazeki, I., Oheda, M., Kubota, N., Tamura, M., Ono, M., Ueyama, Y. and Asano, S. (1986) EMBO J. 5, 871–876.
[2] Oheda, M., Hasegawa, M., Hattori, K., Kuboniwa, H., Kojima, T., Orita, T., Tomonou, K., Yamazaki, T. and Ochi, N. (1990) J. Biol. Chem. 265, 11432–11435.
[3] Souza, L.M., Boone, T.C., Gabrilove, J., Lai, P., Zsebo, K.M., Murdock, D.C., Chazin, V.R., Bruszewski, J., Lu, H., Chen, K.K., Barendt, J., Platzer, E., Moore, M.A.S., Mertelsmann, R. and Welte, K. (1986) Science 232, 61–65.
[4] Lu, H.S., Bone, T.C., Souza, L.M. and Lai, P. (1989) Arch. Biochem. Biophys. 268, 81–92.
[5] Wüthrich, K. (1986) NMR of Proteins and Nucleic Acids, Wiley, New York.
[6] Fesik, S.W. and Zuiderweg, E.R.P. (1990) Quart. Rev. Biophys. 23, 97–131.
[7] Ikura, M., Kay, L.E. and Bax, A. (1990) Biochemistry 29, 4659–4667.
[8] Montelione, G. and Wagner, G. (1990) J. Magn. Reson. 87, 183–188.
[9] Kay, L.E., Tschudin, R. and Bax, A. (1990) J. Magn. Reson. 89, 496–514.
[10] Rudolph, R. (1990) in: Modern Methods in Protein- and Nucleic Acid Research (Tschesche, H., Ed.) Walter de Gruyter, Berlin, pp. 149–171.
[11] Rudolph, R. et al. (1990) United States Patent 4,933,434.
[12] Riesenberg, D., Menzel, K., Schulz, V., Schumann, K., Veith, G., Zuber, G. and Knorre, W.A. (1990) Appl. Microbiol. Biotechnol. 34, 77–82.
[13] Senn, H., Eugster, G., Otting, G., Suter, F. and Wüthrich, K. (1987) Eur. Biophys. J. 14, 301–306.
[14] Yamazaki, T., Yoshido, M., Kanaya, S. and Nakamura, H. (1991) Biochemistry 30, 6036–6047.
[15] Bodenhausen, G., Vold, R.L. and Vold, R.R. (1980) J. Magn. Reson. 37, 93–106.
[16] Kumar, A., Wagner, G., Ernst, R.R. and Wüthrich, K. (1980) Biochem. Biophys. Res. Commun. 96, 1156–1160.
[17] Hore, P.J. (1983) J. Magn. Reson. 54, 539–550.
[18] Messerle, B.A., Wider, G., Otting, G., Weber, C. and Wüthrich, K. (1989) J. Magn. Reson. 85, 608–613.
[19] Wagner, G., Neuhaus, D., Wörgötter, E., Vasak, M., Kägi, J.M.R. and Wüthrich, K. (1986) J. Mol. Biol. 187, 131–135.
[20] Wishart, D.S., Sykes, B.D. and Richards, F.M. (1991) J. Mol. Biol. 222, 311–333.