Protein Engineering vol 1 no 5 pp 399–406, 1987

Solution conformations of human growth hormone releasing factor: comparison of the restrained molecular dynamics and distance geometry methods for a system without long-range distance data

Axel T.Brünger*, G.Marius Clore[1], Angela M. Gronenborn[1] and Martin Karplus

Department of Chemistry, Harvard University, Cambridge, MA 02138, USA and [1]Max-Planck-Institut für Biochemie, D-8033 Martinsried bei München, FRG

*Present address Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, 260 Whitney Avenue, New Haven, CT 06511, USA

A series of three-dimensional structures of the 1–29 fragment of human growth hormone releasing factor in trifluoroethanol have been determined by molecular dynamics and distance geometry methods. The resulting structures satisfy information from nuclear Overhauser effect (NOE) distance data and an empirical potential energy function. Although the polypeptide was found to have an ordered structure in all simulations, the NOE data were not sufficient for global convergence to a unique three-dimensional geometry. Several satisfactory structures have been determined, all of which are extended conformations consisting of a short β-strand and two α-helices (residues 6–13 and residues 16–29) connected by short segments of less well defined secondary structure. Because of the lack of NOE data connecting the helix segments, their relative orientation is not uniquely determined.
Key words: nuclear magnetic resonance/nuclear Overhauser enhancement spectroscopy/restrained molecular dynamics/distance geometry

Introduction

The human growth hormone releasing factor (hGHRF) is a polypeptide consisting of 44 amino acids. It stimulates the secretion of growth hormone *in vivo* of all vertebrate species studied to date (Ling *et al* , 1985). Physiological studies have shown that the active core of the molecule is located in the N-terminal region and that the 1–29 fragment of hGHRF retains almost complete activity relative to longer sequences both *in vivo* and *in vitro* (Lance *et al.*, 1984).

To obtain information concerning the solution structure of hGHRF, a biologically active analogue, ^{27}Nle-hGHRF(1–29)NH_2 referred to as fhGHRF in what follows, of the 1–29 fragment of hGHRF was studied by circular dichroism and NMR (Clore *et al.*, 1986a). It was shown that although fhGHRF does not have an ordered structure in water at room temperature, at relatively low concentrations of trifluoroethanol (TFE) in aqueous solution (~30% v/v) considerable α-helical secondary structure is present. By use of two-dimensional NMR, the ^{1}H-NMR spectrum of fhGHRF was completely assigned, and the secondary structure was evaluated following standard procedures (Wüthrich *et al.*, 1984). A qualitative interpretation of the nuclear Overhauser enhancements (NOE) indicated that fhGHRF had two distinct α-helical regions (residues 6–13 and 16–29) with the remaining regions less clearly defined. Because no long-range NOEs were observed, it was inferred that fhGHRF assumes an extended non-globular structure. The detailed nature of the α-helical structure and the three-dimensional conformation of fhGHRF in solution remains to be determined.

In a variety of studies on small globular proteins with measured NMR results or model data for interproton distances (Havel and Wüthrich, 1985; Kline *et al.*, 1986; Brünger *et al.*, 1986; Clore *et al.*, 1986b,c) it was shown that restrained molecular dynamics or distance geometry calculations can be used to determine the three-dimensional structure of the protein with typical accuracy of 2 Å for backbone atoms and 3 Å for side chain atoms. The determination of a unique tertiary fold was dependent on the availability of long-range interproton distances, i.e. distances between residues i, j with $|i-j|>5$. At the present state of structure prediction and energy function calculation, the tertiary structure of a protein cannot be determined without this long-range information.

In the present work we described restrained molecular dynamics and distance geometry calculations with NOE-derived interproton distance restraints for the peptide fhGHRF. Although the protons were fully assigned and 160 NOE interproton distances were available, only one of them involved residues more than five residues apart in sequence. Starting from several different initial structures and different random seeds in the restrained molecular dynamics and distance geometry calculations, an ensemble of structures was obtained with each structure satisfying the NMR information. Thus, the NOE information combined with the stereochemistry, and non-bonding interactions was found to be insufficient to define a unique global structure within experimental error limits. Despite the differences among the structures we show that they have features in common. In particular, the structures are well extended and have significant helical portions.

The present analysis also provides a comparison of the ability of restrained molecular dynamics and distance geometry to sample a large range of molecular conformations.

Materials and methods

A set of 160 approximate interproton distance restraints were derived from pure phase absorption two-dimensional NOE spectroscopy in 30% (v/v) d_3-trifluoroethanol in either D_2O or H_2O; spectra recorded with mixing times of 200 ms and 300 ms were used for the assignment of NOESY cross peaks, whereas those recorded with mixing times of 100 ms were used for the classification of peak intensities (Clore *et al.*, 1986a). The available NOEs comprised only short-range (i.e. $|i-j|\leq 5$ for residues i,j) distances apart from an NOE between Ser-9 and Gly-15. The distances were classified into three ranges, 1.8–2.5 Å, 1.8–3.0 Å and 1.8–4.5 Å, corresponding to strong, medium and weak NOEs respectively. These distance ranges are approximately the same as used in Clore *et al.* (1986b) with the exception that for weak NOEs the lower distance range limit was reduced from 3.0 to 1.8 Å; this takes into account the uncertainties in effective correlation times for the different interproton vectors as well as uncertainties in intensities, in that certain NOEs can be misclassified as weak NOEs without artificially increasing the restraints energy. The use of these distance ranges has been justified

Table I. Interproton distance restraints used in the restrained molecular dynamics simulations[a]

A. Inter-residue interproton distances

Sequential	NH(i)–NH(i+1)	$C^{\alpha}H(i)$–NH(i+1)	$C^{\beta}H(i)$–NH(i+1)
Y1, A2		S	
A2, D3		S	M
D3, A4		M	
A4, I5	S	M	S
I5, F6	S	M	S
F6, T7	S	W	M
T7, N8	S		W
N8, S9	S	W	W
S9, Y10	M		W
Y10, R11	S	W	M
R11, K12	S	M	M
K12, V13		M	M
V13, L14	S	W	S
L14, G15		S	M
G15, Q16	S		
Q16, L17	S	W	W
L17, S18	S	M	W
S18, A19	S	W	
A19, R20	M		S
R20, K21	S		S
K21, L22			M
L22, L23	S		M
L23, Q24	M		W
Q24, D25	M	M	
D25, I26	S		W
I26, L27	S	W	S
L27, S28	S		M
S28, R29	S		W

NH(i)–NH(j)	NH(i)–$C^{\beta}H(j)$	NH(i)–$C^{\gamma}2H(j)$
I26, S28 W	G15, S9 W	F6, I5 M
		N8, T7 M
		L14, V13 M
		L27, I26 M

$C^{\alpha}H(i)$–HN(j)	$C^{\alpha}H(i)$–$C^{\beta}H(j)$	other
S9, K12 W	Y10, V13 W	I5 –$C\gamma$2H, F6–Cδ1H W
Y10, V13 M	V13, Q16 W	T7–$C\gamma$2H, R11–CδH W
R11, L14 M	Q16, A19 M	Y10–CαH, V13–$C\gamma$1H M
V13, Q16 W	K21, Q24 W	Y10–CδH, L14–CδH M
S18, K21 W	L22, D25 W	Y10–CϵH, L14–CδH M
R20, L23 M		R11–CαH, L14–CδH M
K21, Q24 M		L17–CδH, R20–CδH W
L22, D25 W		I26–$C\gamma$2H, L27–CαH W
L23, I26 W		
Q24, L27 W		

B. Intra-residue interproton distances

NH(i)–$C^{\beta}H(i)$	NH(i)–other	$C^{\alpha}H(i)$–other
A2, A2 S	I5, I5–$C\gamma$1H S	Y1, Y1–CβH W
D3, D3 M	I5, I5–$C\gamma$2H M	I5, I5–$C\gamma$2H M
A4, A4 S	T7, T7–$C\gamma$2H S	T7, T7–CβH M
I5, I5 S	Y10, Y10–Cδ2H W	T7, T7–$C\gamma$2H M
F6, F6 S	R11, R11–$C\gamma$H W	S9, S9–CβH M
T7, T7 M	R11, R11–CδH W	Y10, Y10–CβH S
N8, N8 M	K12, K12–$C\gamma$H W	R11, R11–$C\gamma$H W
S9, S9 W	V13, V13–$C\gamma$1H M	R11, R11–CδH W
Y10, Y10 M	L14, L14–CδH S	K12, K12–CϵH M
R11, R11 S	Q16, Q16–$C\gamma$H W	V13, V13–$C\gamma$1H M

Continued

Table I. Continued

NH(i)–$C^{\beta}H(i)$	NH(i)–other	$C^{\alpha}H(i)$–other
K12, K12 S	L17, L17–CδH M	L13, L13–$C\gamma$2H M
V13, V13 S	R20, R20–$C\gamma$H W	L13, L13–CβH M
L14, L14 S	R20, R20–CδH W	Q16, Q16–CβH M
Q16, Q16 S	L22, L22–CδH W	Q16, Q16–$C\gamma$H W
L17, L17 S	L23, L23–CδH W	S18, S18–CβH M
S18, S18 M	Q24, Q24–$C\gamma$H W	R20, R20–CβH S
A19, A19 S	I26, I26–$C\gamma$2H M	R20, R20–$C\gamma$H W
R20, R20 S	I26, I26–$C\gamma$1H S	R20, R20–CδH M
K21, K21 S	R29, R29–$C\gamma$H W	L22, L22–CδH M
L22, L22 S		I26, I26–$C\gamma$1H S
L23, L23 M		I26, I26–$C\gamma$2H S
Q24, Q24 M		S28, S28–CβH M
D25, D25 M		
I26, I26 S		
L27, L27 S		
S28, S28 S	other	
R29, R29 M	F6–CδH, F6–CβH M	

[a]The interproton distances were classified into three ranges, 1.8–2.5 Å, 1.8–3.0 Å and 1.8–4.5 Å, corresponding to strong ('S'), ('M') and weak ('W') NOEs

Table II. Protocol of the restrained molecular dynamics calculation

Stage	
1	100 cycles conjugent gradient energy minimization, c=1.0 kcal/(mol Å^2); center average[a]
2	0.56 ps molecular dynamics, integration step 0.15 fs, initial velocities at 300 K, velocities checked every 250 steps and if T > 8000 K then scaled by 0.75, initial c=1.0 kcal/(mol Å^2), then doubled every 250 steps but not exceeding c=100.0 kcal/(mol Å^2), center average
3	1.25 ps molecular dynamics, integration step 0.5 fs, initial velocities at 300 K, velocities scaled every 250 steps such that T=300 K, c=100 kcal/(mol Å^2), R-6 average[b].
4	100 cycles conjugent gradient minimization, c=20 kcal/(mol Å^2), R-6 average, ϕ dihedral angle restraints

[a]The 'center average' refers to the $\Sigma\, r_i$ averaging of distances involving unresolvable protons, such as in methyl groups

[b]The 'R-6 average' refers to the $(\Sigma\, r_i^{-6})^{-1/6}$ averaging of distances involving unresolvable protons

in previous work (Braun *et al.*, 1983; Williamson *et al.*, 1985; Clore *et al.*, 1985, 1987). A detailed list of the measured NOEs and the distance range classification is given in Table I. Figure 5 of Clore *et al.* (1986a) summarizes the interresidue NOEs involving NH, $C^{\alpha}H$ and $C^{\beta}H$ protons.

Energy minimization and restrained molecular dynamics calculations were carried out using a combination of the CHARMM empirical potential energy function (Brooks *et al.*, 1983) $E_{\text{empirical}}$ and a restraint energy function E_{NOE}. The total energy of the system is given by

$$E_{\text{total}} = E_{\text{empirical}} + E_{\text{NOE}} \qquad (1)$$

where E_{NOE} describes the NOE distance ranges in the form of square-well potentials with harmonic walls (Clore *et al.*, 1986b); i.e. for a given distance r_{ij} the E_{NOE} is given by

$$E_{\text{NOE}} = \begin{cases} c(r_{ij} - r_{ij}^{U} + 0.2)^2 & \text{for } r_{ij} > r_{ij}^{U} - 0.2 \\ c(r_{ij} - r_{ij}^{L} + 0.2)^2 & \text{for } r_{ij} < r_{ij}^{L} + 0.2 \\ 0 & \text{otherwise} \end{cases} \qquad (2)$$

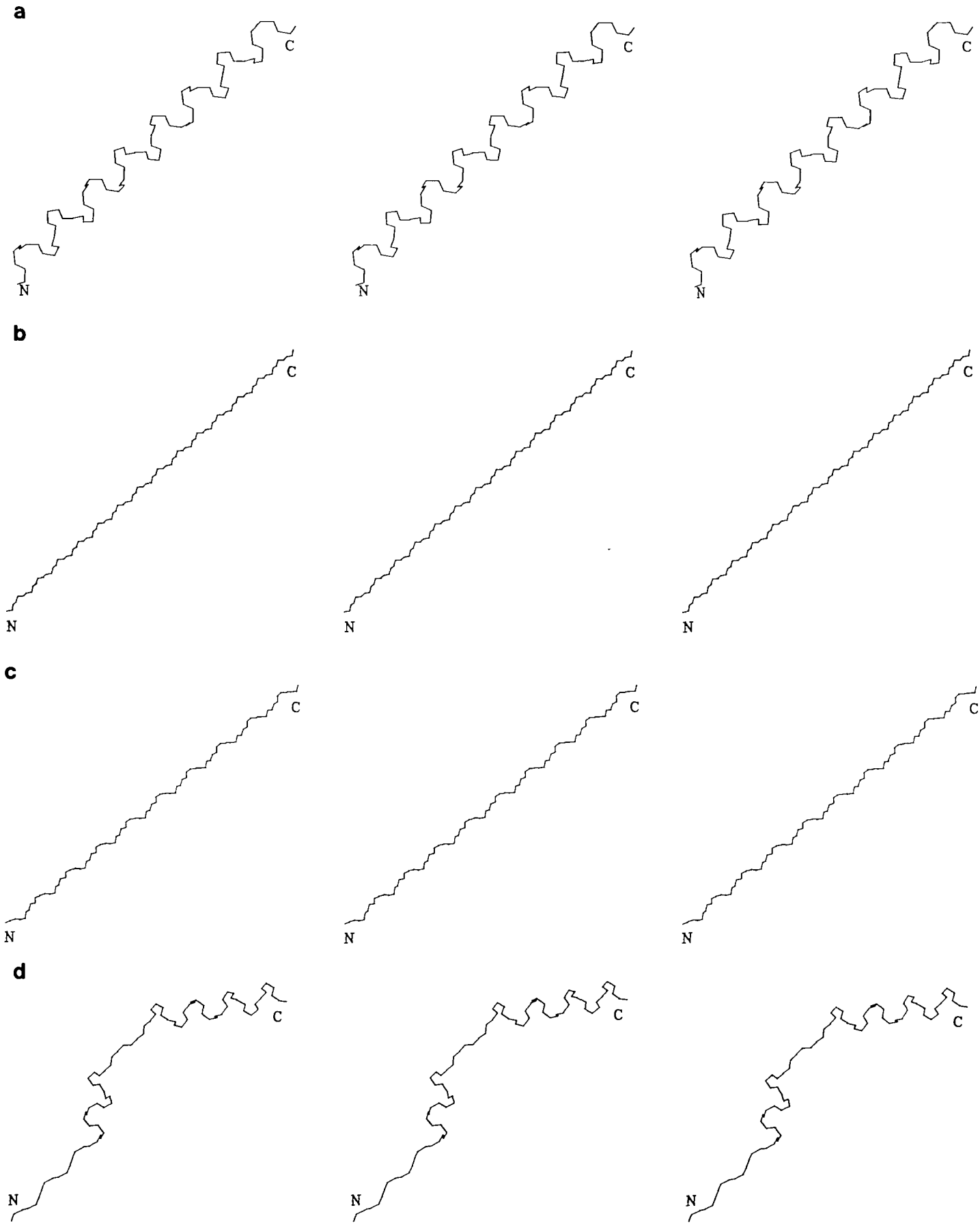

Fig. 1. Backbone (C, C^{α}, N) conformations of initial structures. (**a**) a_{init} (α-helix), (**b**) b_{init} (β-strand), (**c**) p_{init} (polyproline helix) and (**d**) ab_{init} (mixed α/β structure)

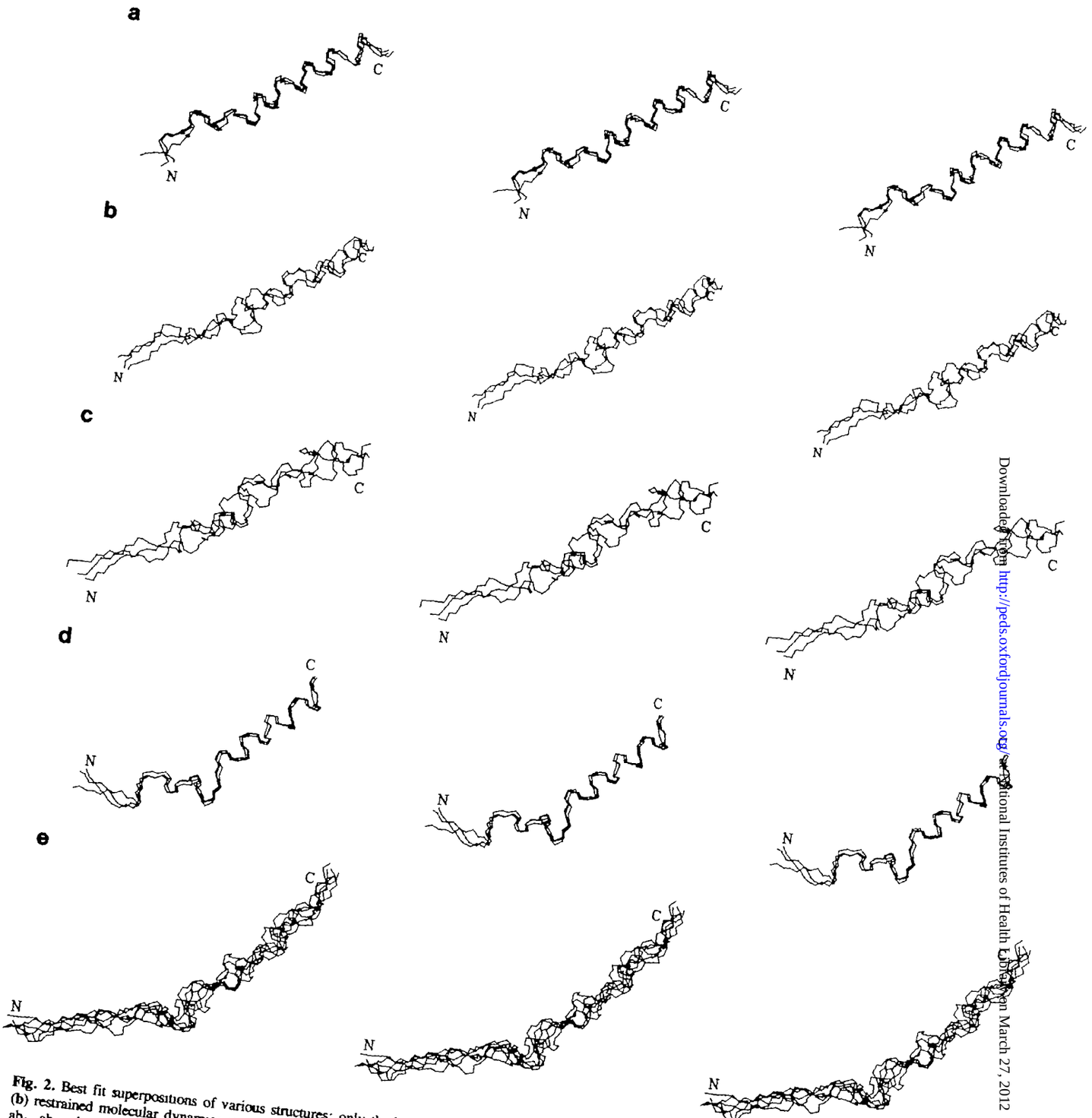

Fig. 2. Best fit superpositions of various structures; only the backbone (C, C^{α}, N) is shown. (**a**) Restrained molecular dynamics structures a_1, a_2, a_3, (**b**) restrained molecular dynamics structures b_1, b_2, b_3, (**c**) restrained molecular dynamics structures p_1, p_2, p_3, (**d**) restrained molecular dynamics structures ab_1, ab_2, ab_3 and (**e**) distance geometry structures $dg_1 - dg_6$. All structures are shown in the same orientation by superimposing them onto structure a_1

where r_{ij}^{U} and r_{ij}^{L} are the upper and lower limits of the target distance range, and c specifies a relative weight factor between the restraint energy and the empirical potential energy.

The protocol for the restrained molecular dynamics calculation is shown in Table II; it consisted of four stages. In the first stage, conjugent gradient minimization is employed to relieve strain in the initial model that could pose a problem when starting the molecular dynamics calculation. During the second stage molecular dynamics was carried out and the temperature of the system was allowed to reach several 1000 K. This is followed by cooling the system to 300 K in stage 3 and finally by conjugent gradient minimization to further optimize the structure. This protocol differs from the one described in earlier work (Brünger *et al.*, 1986; Clore *et al.*, 1986c) essentially in the use of a larger initial value for and a more rapid increase in the scale factor. This results in a reduction of the required computation time. Since the tem-

Table III. R m s differences between target and calculated interproton distances, r m s deviations of bonds and angles from ideality and non-bonded energy[a]

	NOE deviations NH(i)−NH(j) NH(i)−C$^{\alpha}$H(j) (50)	Other (110)	Δ_{bonds} (Å)	Δ_{angles} (degrees)	E_{nbond} (kcal/mol)
A Restrained molecular dynamics structures					
a_1	0 08(0)	0 10(2)	0 019	3 8	−573
a_2	0 10(0)	0 10(2)	0 019	3 9	−586
a_3	0 11(0)	0 10(2)	0 018	3 9	−531
b_1	0 22(4)	0 15(3)	0 024	6 8	−339
b_2	0 25(3)	0 18(7)	0 026	6 3	−328
b_3	0 17(2)	0 20(7)	0 027	7 4	−282
p_1	0 14(1)	0 18(5)	0 025	7 9	−322
p_2	0 22(2)	0 22(10)	0 029	8 0	−243
p_3	0 25(5)	0 22(8)	0 030	8 0	−293
ab_1	0 11(0)	0 17(3)	0 020	4 5	−462
ab_2	0 11(0)	0 17(3)	0 020	4 6	−430
ab_3	0 11(1)	0 16(4)	0 021	4 5	−379
B Distance geometry structures			extended range		
dg_1	0 36(9)	0 26(12)	[0 11(1) 0 05(0)]	0 030	6 1 −334
dg_2	0 39(12)	0 33(17)	[0 13(1) 0 10(1)]	0 029	6 2 −286
dg_3	0 43(12)	0 34(13)	[0 19(3) 0 13(2)]	0 031	6 5 −230
dg_4	0 34(8)	0 24(8)	[0 11(1) 0 05(0)]	0 027	5 3 −364
dg_5	0 34(9)	0 30(13)	[0 09(0) 0 11(2)]	0 027	5 5 −298
dg_6	0 38(6)	0 27(16)	[0 16(2) 0 06(0)]	0 027	5 7 −327

[a]The r m s interproton distance deviations are calculated with respect to the upper and lower limits of the individual distance ranges 1 8−2 5 Å, 1 8−3 0 Å and 1.8−4.5 Å. The numbers in parentheses give the number of distances which violate the distance ranges by more than 0 5 Å In the case of the distance geometry structures the r m s interproton distance deviations and violations are also computed with the extended distance ranges 1 8−3 0 Å, 1 8−3.5 Å and 1 8−5 0 Å (in square brackets) The non-bonded energy E_{nbond} comprises van der Waals and electrostatic energy contributions to the empirical potential energy used in CHARMM (Brooks *et al* , 1983)

perature can become very high (>1000 K) during stage 2, the integration step has been reduced to 0.15 fs. The present protocol is similar to the one used in molecular dynamics refinement of protein crystal structures (Brünger *et al.*, 1987; Brünger, 1987) and ensures that the system can overcome large energy barriers. The macromolecular refinement program X-PLOR (Brünger, 1987) was used to carry out the restrained molecular dynamics calculations on a CRAY-2; the empirical potential energy part of X-PLOR is based on the CHARMM program (Brooks *et al.*, 1984). The program DISGEO (Havel and Wüthrich, 1985; Havel, 1986) was used to carry out distance geometry calculations based on the metric matrix algorithm.

During the fourth stage of the molecular dynamics refinement (Table II) the NOE restraints were augmented by 10 ϕ backbone torsion angle restraints derived from $^3J_{HN\alpha}$ coupling constants measured by double quantum filtered homonuclear correlated spectroscopy (Pardi *et al.*, 1984). The coupling constants involving residues 10, 11, 12, 17, 19, 22, 23, 24, 26, 27 are <5 Hz; hence, the ϕ torsion angles were restrained to the range 0° to −90° by a quartic square-well potential with the scale factor c set to 80 kcal/(mol degrees^4) (Clore *et al.*, 1986a,b). The remaining ϕ backbone torsion angles were restrained to the range 0° to −180° with the exception that no restraint was applied to the ϕ torsion angle of residue Gly-15. This was done since the ϕ backbone torsion angles of all known protein structures fall into the range $0 > \phi > 180°$ except in cases where glycine residues are involved in type-II turns.

Results

Four different initial structures were used for the restrained molecular dynamics calculations (Figure 1): an α-helix ($\phi = -57°$, $\psi = -47°$) referred to as a_{init}, an extended β-strand ($\phi = -139°$, $\psi = 135°$) referred to as b_{init}, a polyproline helix ($\phi = -80°$, $\psi = -150°$) referred to as p_{init}, and a mixed α/β structure referred to as ab_{init} in which residues 1−6, 13−16 are in the form of an extended β-strand and residues 6−13 and 16−29 are α-helical. The latter structure was chosen from a qualitative interpretation of NOE data (Clore *et al.*, 1986b); however, there is direct evidence only for the α-helical regions. The side chains were placed in an extended geometry for all initial structures. For each initial structure three restrained molecular dynamics calculations were carried out using different random number seeds for the assignment of the initial velocities in stage 2 (Table II). This yielded a total of 12 restrained molecular dynamics structures referred to as a_1, a_2, a_3 starting from a_{init}, b_1, b_2, b_3 starting from b_{init}, p_1, p_2, p_3 starting from p_{init}, and ab_1, ab_2, ab_3 starting from ab_{init}.

All attempts to carry out distance geometry calculations with the same distance ranges as used in the restrained molecular dynamics calculations failed due to numerical instabilities during the metrization stage, i.e. certain parameters of the distance geometry calculation produced a numerical overflow. Using somewhat relaxed distance ranges (1.8−3.0 Å, 1.8−3.5 Å and 1.8−5.0 Å for strong, medium and weak NOEs respectively) six out of 20 trials using different initial random substructures were successful. These six distance geometry structures were then subjected to 1500 cycles restrained conjugate gradient minimization with X-PLOR as described in Clore *et al.* (1986b). They are referred to as dg_1 through dg_6.

The resulting structures of the restrained molecular dynamics and distance geometry calculations are shown in Figure 2; the r.m.s. interproton distance deviations, r.m.s. deviations of bonds and angles from ideality, and non-bonding energies of the structures are listed in Table III. A detailed list of interproton distance violations is given in Table IV. It is apparent that convergence to a unique structure was not achieved, e.g. structures a_1 and ab_1 satisfy the NMR information within experimental error limits and have good stereochemistry, yet the structures are clearly different. All structures assume an extended, somewhat banana-shaped conformation without sharp turns even when starting from the slightly kinked structure ab_{init} (Figures 1d and 2d). The structures a_i and ab_i satisfy the NMR interproton distances best and also have good stereochemistry. The other structures (b_i, p_i, dg_i) have more interproton distance violations; these violations are comparable with those obtained in the model study on crambin for some of the structures (Brünger *et al.*, 1986; Clore *et al.*, 1986c). Thus, structures (b_i, p_i, dg_i) can still be considered within the experimental error limits. All structures satisfy the one long-range NOE observed between Ser-9 C$^{\beta}$H and Gly-15 NH.

Table V shows the r.m.s. deviations of the converged structures with respect to the initial structures or the averages of the converged structures; differences between average structures are also included. As expected, the r.m.s. deviations of the converged structures from the initial structures are much larger than the deviations from the averages. However, the distributions around the average structure of restrained molecular dynamics structures

Table IV. Interproton distance violations[a]

a_1		
I5–NH,	I5–Cγ2H	0 66
I26–NH,	I26–Cγ2H	0 69
a_2		
I5–NH,	I5–Cγ2H	0 53
I26–NH,	I26–Cγ2H	0 69
a_3		
I5–NH,	I5–Cγ2H	0 64
I26–NH,	I26–Cγ2H	0 65
b_1		
I5–NH,	F6–NH	0 77
F7–NH,	N8–NH	0 85
I26–NH,	L27–NH	0 61
S9–CαH,	K12–NH	0 52
Y10–CβH,	R11–NH	0 51
R11–CβH,	K12–NH	0 56
I5–CβH,	I5–NH	0 58
b_2		
I5–NH,	F6–NH	0 78
K21–CαH,	Q24–NH	0 60
L22–CαH,	D25–NH	1 05
Y10–CβH,	R11–NH	0 61
K21–CβH,	L22–NH	0 59
L27–CβH,	S28–NH	0 51
L22–CαH,	D25–CβH	0 91
I5–CβH,	I5–NH	0 61
F6–CβH,	F6–NH	0 61
I26–NH,	I26–Cγ2H	0 64
b_3		
I5–NH,	F6–NH	0 64
Y10–CαH,	V13–NH	0 61
F6–CβH,	T7–NH	0 61
V13–CβH,	L14–NH	0 54
A19–CβH,	R20–NH	0.78
R20–CβH,	K21–NH	0 96
F6–CβH,	F6–NH	0 54
T7–CαH,	T7–Cγ2H	0 53
V13–CαH,	V13–Cγ1H	0 55
p_1		
I5–NH,	F6–NH	0 59
K12–CβH,	V13–NH	0 66
L14–CβH,	G15–NH	0 56
R20–CβH,	K21–NH	0 85
L22–CβH,	L23–NH	0 55
I26–NH,	I26–Cγ2H	0 69
p_2		
A4–NH,	I5–NH	0 88
I26–NH,	L27–NH	0 86
F6–CβH,	T7–NH	0 51
Y10–CβH,	R11–NH	0 55
K12–CβH,	V13–NH	0 61
L14–CβH,	G15–NH	0 61
A19–CβH,	R20–NH	0 98
R20–CβH,	K21–NH	0.65
L22–CβH,	L23–NH	0 56
F6–CβH,	F6–NH	0 52
T7–CαH,	T7–Cγ2H	0 55
I26–CαH,	I25–Cγ2H	0 81
p_3		
A4–NH,	I5–NH	0 66
I5–NH,	F6–NH	0 90
S9–NH,	Y10–NH	0 53
I26–NH,	L27–NH	0 68
L23–CαH,	I26–NH	0 89
Y10–CβH,	R11–NH	0 62
K12–CβH,	V13–NH	0 62
L14–CβH,	G15–NH	0 53
A19–CβH,	R20–NH	0 72
R20–CβH,	K21–NH	0.81
F6–CβH,	F6–NH	0.69
T7–CαH,	T7–Cγ2H	0 58
I26–CαH,	I26–Cγ2H	0 83
ab_1		
V13–CβH,	L14–NH	0 97
L14–NH,	V13–Cγ2H	0.73
I26–NH,	I26–Cγ2H	0.68
ab_2		
V13–CβH,	L14–NH	0.97
L14–NH,	V13–Cγ2H	0 76
I26–NH,	I26–Cγ2H	0 70
ab_3		
R11–CαH,	L14–NH	0 52
V13–CβH,	L14–NH	0 80
L14–NH,	V13–Cγ2H	0 54
I26–NH,	I26–Cγ2H	0 68
V13–CαH,	V13–Cγ1H	0 53
dg_1		
I5–NH,	F6–NH	1 10
dg_2		
S9–CαH,	K12–NH	1 05
L14–NH,	V13–Cγ2H	1 04
dg_3		
I5–NH,	F6–NH	1 13
Y10–NH,	R11–NH	1 13
L23–CαH,	I26–NH	1 29
A4–CβH,	I5–NH	1 17
L14–NH,	V13–Cγ2H	1 03
dg_4		
L23–CαH,	I26–NH	1 03
dg_5		
Y10–CαH,	V13–Cγ1H	1 23
V13–NH,	V13–Cγ1H	1 04
dg_6		
A4–NH,	I5–NH	1 14
I5–NH,	F6–NH	1 30

[a]Listed are the distances which violate the distance ranges listed in Table I by more than δ with respect to the upper limits In the case of the restrained molecular dynamics structures Δ has been set to 0.5 Å; in the case of the distance geometry structures (dg_1–dg_6) Δ has been set to 1 0 Å

starting from the same initial geometry and using different initial random number seeds (part B of Table V) correlate with the shift between the initial structures and the corresponding average structure after restrained molecular dynamics (part A of Table V); i.e. the more the structures move during restrained molecular dynamics, the larger is the r.m.s. distribution of structures around their mean. Figure 2 and Table V also indicate that the global distribution of the distance geometry structures is significantly smaller compared with all restrained molecular dynamics structures combined despite the fact that the actual distance ranges used for the distance geometry calculations were slightly larger. None of the distance geometry structures have purely α-helical segments (Figure 2e), and they exhibit more interproton distance violations than the restrained molecular structures. The distance geometry structures also exhibit a larger distribution of local conformations than the restrained molecular dynamics structures; this is probably due to the neglect of dihedral angles and non-bonded interactions in the distance geometry method. The conformation of the distance geometry structures appears to be closest to the restrained molecular dynamics structures b_i starting from an extended β-strand (Table V).

To distinguish between local and global convergence, atomic r.m.s. differences were averaged over 3, 5 and 11 residues in Figure 3 for the restrained molecular dynamics structures. It

Table V. Atomic r m.s. differences for backbone (N, C, C^{α}) atoms[a]

A. Average r.m.s difference to initial structures[b]	
a_i versus a_{init}	1.7 ± 0.07
b_i versus b_{init}	14 5 ± 0 05
p_i versus p_{init}	12 7 ± 0 6
ab_i versus ab_{init}	4.2 ± 0 1
B Average r m s difference to average structure[c]	
a_i versus A	0 7 ± 0.1
b_i versus B	1 9 ± 0.1
p_i versus P	2 2 ± 0.4
ab_i versus AB	0.5 ± 0.07
RMD_i versus RMD	2 7 ± 0 4
DG_i versus DG	1 8 ± 0 1
C Differences of average structures	
A versus B	3 5
A versus P	3 3
A versus AB	3 4
A versus DG	5 8
B versus P	2 2
B versus AB	4 1
B versus DG	3 5
P versus AB	3 8
P versus DG	4 2
AB versus DG	4 9

[a]The average structures were computed by superimposing the individual structures onto a reference structure through the least-squares method (Kabsch, 1976) and then averaging the coordinates The average of structures a_i is denoted A, the reference structure was a_1 The corresponding notation for the average structures of b_i, p_i, ab_i and dg_i is used RMD is the average structure obtained from all restrained molecular dynamics structures with a_1 used as the reference structure

[b]The symbols a_i versus a_{init} designate the average of the r m s differences of the individual structures a_i from a_{init}, the other symbols have corresponding meanings

[c]The symbols a_i versus A designate the average of the r.m s differences of the individual structures a_i from the average structure A, the other symbols have corresponding meanings

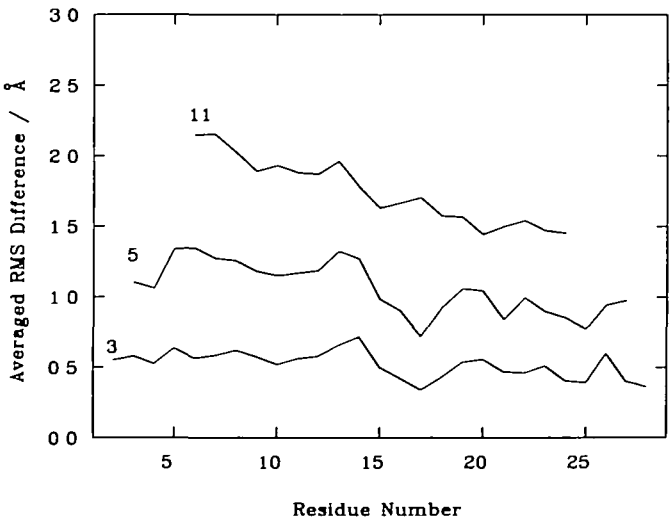

Fig. 3. Averages of atomic r m s differences of the backbone (C, C^{α}, N) atoms of the restrained molecular dynamics structures about their mean The r m.s. differences are averaged over 3, 5 or 11 residue segments of the individual restrained molecular dynamics structures, for each residue range the segments were superimposed onto a_1 with the least-squares method (Kabsch, 1976) The values are indicated at the central residue of the corresponding residue range

appears that local convergence within 0.75 Å around the mean is achieved when comparing only three residues at a time. The increasing r.m.s. differences for the five-residue ranges and 11-residue range clearly reflect the global divergence of the structures. The large differences in overall conformation are located around residue 6 and residue 13. Residues 16–29 exhibit the least deviations and assume an α-helical conformation with distortions present in some of the structures (Figure 2); residues 1–4 assume a β-strand conformation and residues 6–13 assume distorted helical conformations in all cases.

Discussion

fhGHRF is an example of a polypeptide where the information about interproton distance restraints obtained from NMR is not sufficient to determine a unique structure. In contrast to NMR studies of small globular proteins such as crambin (Brünger *et al.*, 1986) and purothionin (Clore *et al.*, 1986b) no NOE data about the tertiary structure of fhGHRF was available. The reason for the absence of observable tertiary NOEs could be either that the structures are extended or that the fhGHRF structure in solution is an average over a set of distinct folded conformations.

The restrained molecular dynamics and distance geometry calculations show that while convergence is achieved locally, several different global conformations appear to be possible. All conformations found are extended, banana-shaped structures without sharp turns. The structures consist of a short β-strand and two α-helices connected by short segments of less well defined secondary structure. The two helices are defined best and the greatest variability occurs in the connecting segments. The helices found here agree with those deduced from the earlier qualitative analysis of NOE data (Clore *et al.*, 1986a). Although none of the structures found had sharp kinks in the region 13–16, their existence cannot be excluded; long molecular dynamics simulations that could lead to folded structures were not performed.

Restrained molecular dynamics starting from several different initial structures appears better suited than distance geometry for exploring possible conformations in cases where the NMR information is incomplete. The distance geometry structures of fhGHRF appeared to be biased towards structures that were obtained by restrained molecular dynamics when starting from an extended β-strand. Restrained molecular dynamics with other initial conditions (e.g. α-helix) produced structures that were not obtained by the distance geometry calculations. From the distance geometry calculations alone one might have concluded that there is a unique solution conformation of fhGHRF.

Acknowledgements

This work was supported by the National Science Foundation (A T B and M.K) and the Max-Planck-Institut Gesellschaft and grant No. 321/4003/0318909A from the Bundesministerium für Forschung und Technologie (G M.C. and A M.G) Computer time on the CRAY-2 at the Minnesota Supercomputer Center was supplied by a grant from the Office of Advanced Scientific Computing of the National Science Foundation. We thank F.A Momany for discussions and Polygen Corporation (Waltham, USA) for supplying fhGHRF

References

Braun,W., Wider,G , Lee,K H and Wüthrich,K (1983) *J Mol. Biol.*, **169**, 921–948.

Brooks,B R., Bruccoleri,R E , Olafson,B D , States,D J., Swaminathan,S and Karplus,M (1983) *J. Comput. Chem.*, **4**, 187–217

Brünger,A.T , Clore,G.M., Gronenborn,A.M. and Karplus,M (1986) *Proc. Natl. Acad. Sci USA*, **83**, 3801–3805.

Brünger,A.T. (1987) In Isaacs,N. (ed), *Methods and Applications in Crystallographic Computing* Proceedings of the International School of Crystallographic Computing, Adelaide, Oxford Press, Oxford, in press.

Brünger,A T., Kuriyan,J and Karplus,M. (1987) *Science*, **235**, 458–460
Clore,G M , Gronenborn,A M., Brünger,A.T and Karplus,M (1985) *J. Mol Biol.*, **186**, 435–455
Clore,G.M., Martin,S.R and Gronenborn,A M. (1986a) *J. Mol. Biol.*, **191**, 553–561
Clore,G.M., Nilges,M., Sukumaran,D K., Brünger,A T , Karplus,M and Gronenborn,A M (1986b) *EMBO J* , **5**, 2729–2735
Clore,G.M , Brünger,A T., Karplus,M and Gronenborn,A M (1986c) *J. Mol Biol.*, **191**, 523–551.
Clore,G M , Sukumaran,D.K., Nilges,M and Gronenborn,A M (1987) *Biochemistry*, **26**, 1732–1745
Havel,T F. and Wüthrich,K (1985) *J Mol. Biol* , **182**, 281–294
Havel,T F. (1986) *DISGEO, Quantum Chemistry Program Exchange Program.* No 507, Indiana University
Kabsch,W (1976) *Acta Crystallogr Sect. A*, **32**, 922–923
Kline,A D., Braun,W. and Wüthrich,K (1986) *J. Mol. Biol.*, **189**, 377–382
Lance,V A., Murphy,W A , Sueiras-Dias,J and Coy,D H (1984) *Biochem Biophys Res. Commun* , **119**, 265–270
Ling,N., Zeytin,F , Bohlen,P., Esch,F., Brazeau,P , Wehrenberg,W B., Baird,A. and Guillemin,R (1985) *Annu Rev. Biochem.*, **54**, 403–423
Pardi,A , Billeter,M and Wüthrich,K (1984) *J Mol Biol* , **180**, 741–751
Williamson,M.P , Havel,T F and Wüthrich,K (1985) *J. Mol Biol* , **182**, 295–315
Wüthrich,K , Billeter,M and Braun,W (1984) *J Mol Biol* , **180**, 715–740

Received on July 6, 1987, revised on October 2, 1987