High-resolution solution structure of two members of a conformationally homogeneous combinatorial peptide library based on the classical zinc-finger motif

ARTICLE *in* JOURNAL OF BIOMOLECULAR NMR · AUGUST 1996

Impact Factor: 3.14 · DOI: 10.1007/BF00198138 · Source: PubMed

5 AUTHORS, INCLUDING:

Gaetano Barbato
University of Rome Tor Vergata
59 PUBLICATIONS **2,195** CITATIONS
SEE PROFILE

Daniel Oscar Cicero
University of Rome Tor Vergata
102 PUBLICATIONS **1,556** CITATIONS
SEE PROFILE

Journal of Biomolecular NMR, 8 (1996) 36–48
ESCOM

J-Bio NMR 359

High-resolution solution structure of two members of a conformationally homogeneous combinatorial peptide library based on the classical zinc-finger motif

Gaetano Barbato[a], Daniel O. Cicero[a], Elisabetta Bianchi[b], Antonello Pessi[b] and Renzo Bazzo[a,*]

Departments of [a]Chemistry and [b]Biotechnology, Research Institute of Molecular Biology (IRBM), P. Angeletti, Via Pontina km 30.600, I-00040 Pomezia (Rome), Italy

Received 20 November 1995
Accepted 1 April 1996

Keywords: Zinc-finger domains; Peptide libraries; Distance geometry

Summary

We describe the high-resolution structure by NMR of two peptides that belong to a combinatorial library based on the zinc-finger motif. The library represents, to the best of our knowledge, the first example of a conformationally homogeneous peptide library and was obtained by introducing random residues in five positions of the α-helical portion of a 26-residue 'consensus' peptide (CP1) belonging to the Cys^2-Hys^2 zinc-finger family. The result was shown to be a highly homogeneous α-helical library (Bianchi et al., 1995). The structures of the parent compound (CP1) and of a representative member (CP1m) that was selected by screening the library with a monoclonal antibody are compared in detail as an example of the very high stability of the zinc-finger scaffold upon sequence variability. The two peptides exhibit an extremely high degree of structural similarity. The use of this type of conformationally constrained combinatorial library might represent a step forward in the design of peptidomimetics, as it considerably accelerates the process of the identification of the spatial relationship among the pharmacophoric groups.

Introduction

Much effort is currently devoted to the development of rational ways to convert the information encoded in peptide structures into peptidomimetics: substances with no residual amide bonds which mimic the biological activity of a peptide effector. The jump from a peptide to a peptide mimic is a challenging one, but the continuous progress in structure determination methods, computer-modelling algorithms and synthetic organic chemistry are raising hopes that a general solution may be found. In a recent review, Moore (1994) logically divides the peptide-to-peptidomimetic transition into three stages: (i) identification of the identity of the amino acid side chains that are responsible for the activity ('pharmacophoric groups'); (ii) establishment of the spatial relationship between the pharmacophoric groups ('pharmacophore model'); (iii) peptidomimetic design, i.e. selection of a suitable template upon which the pharmacophoric groups can be mounted in such a way as to reproduce the geometry of the pharmacophore model. The rate-limiting step is so far the second one.

Against this background, major progress would come from extension of the use of selection methods. In this scenario, the target of rational design should be shifted from the crafting of single compounds to the design of populations (Gallop et al., 1994). These populations should be structurally homogeneous with respect to the peptide backbone, i.e. the members of the population should all display the same topography, and the diversity should be high enough to fully explore the binding potential of the represented pharmacophore. As a consequence, ligands selected from such conformationally homogeneous combinatorial peptide libraries (CHCPL) would encode all the information necessary for stage (iii) in the design of a peptide mimic. Overall, the number of pharmacophores

*To whom correspondence should be addressed.
Abbreviations: *t*-Bu, *tert*-butyloxycarbonyl; Fmoc, 9-fluorenylmethoxycarbonyl.

TABLE 1
OVERVIEW OF RESTRAINTS USED IN THE CALCULATION OF THE STRUCTURES AND VIOLATIONS OF THE FINAL STRUCTURES OF THE CONSENSUS PEPTIDES

	CP1	CP1m
Summary of constraints		
NOE		
Intraresidue stereospecific	39	45
(i,i+1)	36	39
(i,i+2)	6	11
(i,i+3)	19	21
(> i+3)	71	72
Total number of NOEs	171	188
Dihedral angles		
ϕ	14°	14°
χ_1	8°	10°
Zinc binding site	7	7
Total number of constraints	200	219
Summary of results		
Number of structures at first DIANA REDAC run	100	100
Number of structures accepted after first run	31	5
Minimum/maximum target function value accepted	2.3/2.9	2.3/2.9
Number of structures at second DIANA REDAC run	100	100
Number of structures accepted after second run	47	16
Minimum/maximum target function value accepted	2.3/2.9	2.2/2.9
Number of structures at third DIANA REDAC run	100	100
Number of structures accepted after third run	58	27
Minimum/maximum target function value accepted	2.2/2.9	2.2/2.9
Summary of violation after rEMs		
Average number of NOE violations > 0.4 Å	1	0
Average number of NOE violations 0.3–0.4 Å	2	1
Average number of NOE violations 0.2–0.3 Å	6	7
Maximum distance violation (Å)	0.45	0.36
Average number of dihedral angle violations	5	4
Maximum dihedral angle violation (°)	22.4	18.6
Average bond length deviation (Å)	0.010	0.012
Maximum bond length deviation (Å)	0.048	0.061
Average valence angle deviation (°)	1.2	1.4
Maximum valence angle deviation (°)	5.3	6.2

corresponding to the various geometries of the peptide backbone is limited to combinations of the basic secondary structure elements, i.e. turns, helices and sheets, and diversity is achieved by decorating these backbones with the appropriate side chains. Ultimately, therefore, the availability of a panel of CHCPL would enable a process that we call selection-driven peptidomimetic design, whereby a pharmacophore model is rapidly derived from screening, and the information is then either directly converted into a first-generation peptidomimetic, or probed and refined by the synthesis of secondary libraries spanning a narrower conformational space.

We have recently described the first example of an α-helical CHCPL (Bianchi et al., 1995), whose potential for the above-mentioned process was demonstrated by the selection of a peptide (named CP1m) which bound a monoclonal IgA reactive against the lipopolysaccharide of the human pathogen *Shigella flexneri*. The library and the parent compound CP1 were examined and compared by CD and Co(II) absorption-spectroscopy methods (Bianchi et al., 1995). The selected ligand CP1m and CP1 produced indistinguishable spectra (data not shown). In preliminary NMR analysis the expected pharmacophoric geometry in the critical helical region of both peptides was observed (Bianchi et al., 1995). Our goal was achieved by exploiting the structural stability and tolerance to sequence variation of the zinc-finger fold. The term zinc-finger is used to describe a highly conserved structural motif, but since its first description in 1985 by Miller et al., the term has been enriched in its meaning, and recently Schwabe and Klug (1994) classified ten different zinc-finger topologies. The parent sequence chosen for our library was a previously described consensus peptide (CP1) with the following sequence: Pro-Tyr-Lys-Cys-Pro-Glu-Cys-Gly-

Lys-Ser-Phe-Ser-Gln-Lys-**Ser-Asp**-Leu-**Val-Lys**-His-Gln-**Arg**-Thr-His-Thr-Gly (Krizek et al., 1991; Kim and Berg, 1993; Bianchi et al., 1995,1996; the residues randomised in the library are in boldface). It belongs to the family referred to as 'a)' or 'classical zinc-finger'; using a classification based on zinc-coordinating amino acids, it is of the Cys^2Hys^2 type, zinc being ligated by two cysteines and two histidines. CP1 was designed by selecting at each position in the sequence the most represented amino acid in a database of 131 zinc-finger domains, and the authors verified that CP1 did indeed bind zinc ions more tightly than any other zinc-finger characterised to date (Krizek et al., 1991); since in this structural motif metal binding and folding are strictly coupled, this was taken as evidence of an unusually stable structure. Although they did not include a detailed three-dimensional structure for CP1, the provided experimental evidence convincingly showed that CP1 adopted the 'classical' fold, with the antiparallel β-hairpin and the α-helix.

Here we describe in detail the structures of CP1 and of the IgA-binding zinc-finger selected from the library (CP1m, with the four mutations S15H, D16F, K19Q, R22H). The structures are compared at their atomic level evidencing the high degree of similarity, as well as small alterations in the packing of secondary structure elements.

Materials and Methods

Peptide synthesis and sample preparation

The peptides synthesis was performed on a Zinsser SMPS 350 synthesizer by Fmoc/*t*-Bu chemistry. The purity was 70 to 85% and >95% after HPLC purification. Peptides were characterised by mass spectrometry and amino acid analysis. Full details on the synthesis and purification have already been published (Bianchi et al., 1995).

NMR spectroscopy

All the experiments were recorded at 27 °C, pH 6.6 and 6.3, for CP1 and CP1m, respectively, in H_2O and D_2O on a AMX500 Bruker spectrometer. The following experiments were performed on both CP1 and CP1m: TOCSY with mixing of 60 ms (MLEV17 isotropic mixing), COSY, P.E.COSY, NOESY with 100, 200, and 350 ms mixing time. All the experiments were recorded using TPPI in the incremented dimension. A weak continuous irradiation was used to suppress the intense water signal, except in the NOESY experiments where a combined use of jump-and-return pulses and spin-lock water dephasing were used. All matrices were 256 (t_1) × 1000 (t_2) complex data points, except for the P.E.COSY experiment in which they were acquired with 512 (t_1) × 8000 (t_2) complex data points. The $^3J_{HNH^\alpha}$ and $^3J_{H^\alpha H^\beta}$ coupling constants were obtained from the 1D rows extracted from the P.E.COSY and, where resolved, from the 1D spectrum acquired with 32 000 complex points.

The amide proton exchange rate was evaluated by observing the decreasing intensity of TOCSY HN-H^α cross-peaks, after the fresh preparation of a solution in D_2O, with the same experimental conditions of temperature and pH as the above-mentioned (Marion et al., 1989).

Distance and dihedral angle restraints

The NOEs were divided in three categories: (i) weak; (ii) medium; and (iii) strong, with an allowed upper boundary of 4.5 Å, 3.6 Å and 3.0 Å, respectively, and a lower boundary equivalent to the sum of the van der Waals radii. A total of 132 interresidue distance constraints were included for the structures calculation of CP1, and 143 for CP1m. Furthermore, seven distance constraints were used in order to take into account the geometry of the complexation site, as described in the next section. Calculation of the ϕ and χ_1 dihedral angle restraints from the $^3J_{HNH^\alpha}$ and $^3J_{H^\alpha H^\beta}$ spin–spin coupling constants made use of the program HABAS (Güntert et al., 1989). The NOE intensities were evaluated from the 100-ms NOESY. The set of constraints included 22 dihedral angle restraints for CP1, and 24 for CP1m with an allowance of ± 15°. A summary of the whole set of restraints used, as well as of the obtained results, is given in Table 1.

Structure calculation

The structure calculation was performed by the combined use of the program DIANA (Güntert et al., 1991) and restrained Energy Minimisation (rEM), using the conjugate gradient method and the Kollman force field. The resulting structures were graphically analysed with the SYBYL software package (Tripos, St. Louis, MA).

TABLE 2
CHARACTERISTICS OF THE ZINC COORDINATION SITE

Peptide residue	χ_1 (°)	χ_2 (°)
CP1		
Cys^4	169 ± 2	–
Cys^7	72 ± 3	–
His^{20}	176 ± 7	85 ± 8
His^{24}	−75 ± 17	−54 ± 26
CP1m		
Cys^4	174 ± 2	–
Cys^7	73 ± 3	–
His^{20}	170 ± 4	75 ± 3
His^{24}	−124 ± 13	−50 ± 10

Range of the constraints used for the zinc coordination site: $d_{S^\gamma S^\gamma}$ = 3.47–4.06 Å, $d_{S^\gamma N^\varepsilon}$ = 3.25–3.84 Å, $d_{N^\varepsilon N^\varepsilon}$ = 3.03–3.60 Å. The range of these distances was calculated geometrically imposing two boundary conditions: (i) a tetrahedral zinc coordination, given by a X-Zn-Y angle of 109 ± 10° (Vedani and Huhta, 1990); and (ii) the distance S^γ–Zn is 2.32 ± 0.04 Å and that of N^ε–Zn is 2.04 ± 0.05 Å (Vedani and Huhta, 1990).

TABLE 3
ASSIGNMENTS OF PROTON RESONANCES OF THE CONSENSUS PEPTIDES

Residue	Chemical shift (ppm)						
	HN	H^{α}	H^{β}	H^{γ}	H^{δ}	H^{ε}	H^{ζ}
CP1							
Pro^{1}		4.26	2.30, 1.91	2.16, 2.08	3.75, 3.68		
Tyr^{2}	8.05	4.52	3.00, 2.61	6.92	6.79		
Lys^{3}	8.61	4.90	1.48, 1.43	1.04	1.48	2.80	
Cys^{4}	9.22	4.66	3.36, 2.85				
Pro^{5}		4.36	2.31, 2.98	2.19, 2.02	4.18		
Glu^{6}	8.50	4.16	1.70, 1.28	1.79			
Cys^{7}	7.81	5.03	3.32, 2.75				
Gly^{8}	8.12	4.10, 3.64					
Lys^{9}	7.87	3.88	1.33, 1.18	0.99	1.35	2.83	
Ser^{10}	7.84	5.01	3.38				
Phe^{11}	8.60	4.60	3.38, 2.55	7.15	6.71		6.10
Ser^{12}		4.34	3.95				
Gln^{13}	7.92	4.75	2.12, 1.89	2.32			
Lys^{14}	(8.60)	2.95	1.40, 0.91	0.85	1.05	2.70	
Ser^{15}	7.60	3.83	3.70				
Asp^{16}	6.72	4.30	2.78, 2.63				
Leu^{17}	6.94	3.12	1.96, 1.51	1.13	0.91, 0.85		
Val^{18}	8.00	3.57	1.86	0.84, 0.77			
Lys^{19}	7.25	3.78	1.70	1.42, 1.29	1.58	2.83	
His^{20}	7.55	4.12	2.96, 2.78		6.85	7.88	
Gln^{21}	8.89	3.60	2.20, 2.10	2.65			
Arg^{22}	7.07	4.07	1.74, 1.48	1.58	3.10		
Thr^{23}	7.64	4.03	3.92	1.05			
His^{24}	7.07	4.75	3.20, 3.10		6.55	7.92	
Thr^{25}	7.78	4.30	4.20	1.12			
Gly^{26}	8.23	3.82, 3.70					
CP1m							
Pro^{1}		4.05	2.12, 1.66	1.81, 1.70	3.17		
Tyr^{2}		4.47	2.91, 2.56	6.86	6.75		
Lys^{3}	8.62	4.87	1.45, 1.38	1.00	1.10	2.78	
Cys^{4}	9.20	4.65	3.31, 2.79				
Pro^{5}		4.32	2.27, 1.92	2.01	4.15		
Glu^{6}	8.47	4.12	1.24	1.75, 1.63			
Cys^{7}	7.81	5.05	3.30, 2.73				
Gly^{8}	8.10	4.08, 3.62					
Lys^{9}	7.83	3.84	1.30, 1.10	0.94	1.35	2.78	
Ser^{10}	7.83	4.98	3.35				
Phe^{11}	8.60	4.62	3.20, 2.63	7.10	6.68		6.03
Ser^{12}	7.79	4.27	3.99, 3.89				
Gln^{13}	7.05	4.68	1.85, 0.96	2.30, 2.20			
Lys^{14}	8.58	2.88	1.43, 0.99	0.83	1.36	2.70	
His^{15}	6.33	4.06	2.87, 2.64		6.37	7.22	
Phe^{16}	6.17	4.23	3.19, 2.84	7.05	7.32		7.21
Leu^{17}	6.84	3.13	1.95, 1.14	1.48	0.90, 0.84		
Val^{18}	7.90	3.47	1.84	0.82, 0.73			
Gln^{19}	7.38	3.82	2.20, 1.93	2.32			
His^{20}	7.43	4.17	2.96, 2.72		6.82	7.85	
Gln^{21}	8.68	3.59	2.14, 2.06	2.58			
His^{22}	6.81	4.39	3.17		7.15	8.12, 7.78	
Thr^{23}	7.71	4.00	3.89, 3.95	1.01, 1.11			
His^{24}	7.08	4.73	3.16, 3.08		6.51, 7.04	7.89, 8.16	
Thr^{25}	7.74	4.23	4.16, 4.04	1.08, 1.06			
Gly^{26}	8.22	3.79					

All the calculations were performed on a Sun Sparc2 data station.

An extended structure represented the starting point for both CP1 and CP1m. Initially 100 structures per molecule were generated by DIANA, including the REDAC routine (Güntert and Wüthrich, 1991). After three rounds

of calculations, sets of 58 and 27 structures were selected for CP1 and CP1m, respectively. Then, out of these sets 20 structures having the lowest value of the target function were chosen for further refinement, which at this stage included all the NOEs, ϕ and χ_1 dihedral angles restraints (8 and 10 β-methylene groups were stereospecifically assigned for CP1 and CP1m, respectively), using dummy atoms for the aromatic side chains of Tyr^2 and Phe^{11}, and for the methyl groups of Leu^{17} and Val^{18}.

Since the histidine imidazole ring has the ability to coordinate the zinc ion through either N^{δ} or N^{ε} atoms and the exact coordination was not known a priori, the initial calculations were carried out without explicit constraints for the zinc coordination site, but with the χ_1 dihedral angle constraints for both histidine and cysteine residues and the NOE constraints involving the imidazole protons of both histidine residues. These constraints were sufficient to uniquely orient the ring and showed that both histidine residues coordinate the metal through the N^{ε} atom.

The set of structures calculated by DIANA was then further refined by rEM, now including also the stereospecific assignment of methyl groups and seven additional constraints specific for the geometry of the coordination site, applying a 10 times higher force-field constant in the latter case. These constraints comprised distances between N^{ε}–S^{γ}, N^{ε}–N^{ε} and S^{γ}–S^{γ} for all the atoms involved in the coordination (Table 2). Upper and lower distances were determined by imposing two boundary conditions: (i) a tetrahedrical geometry of coordination for the zinc atom, and thus an angle X-Zn-Y of 109° with an allowance of ±10° (Vedani and Huhta, 1990); and (ii) a S^{γ}–Zn distance of 2.32±0.04 Å and a N^{ε}–Zn distance of 2.04±0.05 Å (Vedani and Huhta, 1990). The rEM was performed with 400 conjugate gradient iterations using the Kollman force field.

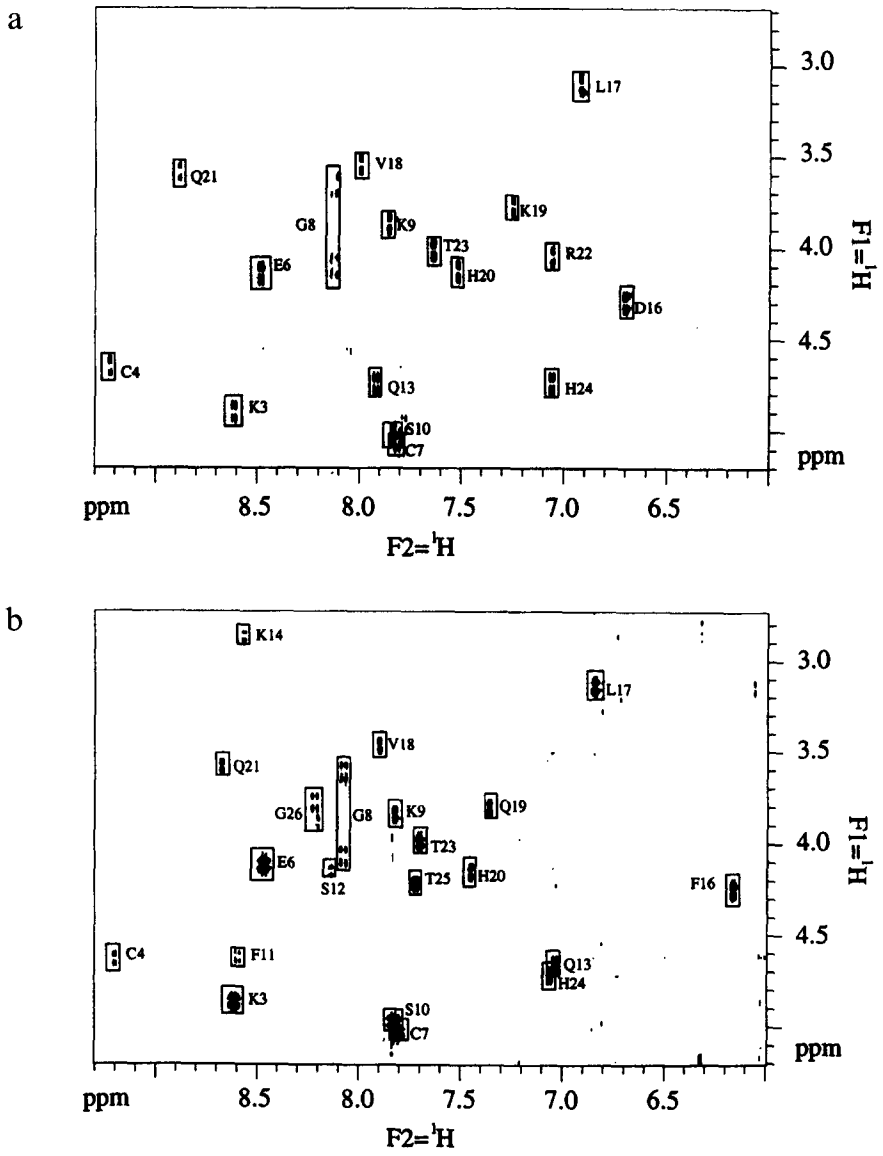

Fig. 1. Finger-print region of the COSY experiment for the CP1 (a) and CP1m (b) molecules. Signal assignment is indicated.

Results

NMR spectra

The assignment of the proton resonances for both CP1 and CP1m was achieved using the method illustrated by Wüthrich (1986) and is reported in Table 3. Lee et al. (1992) identified a number of key chemical shifts that reflect the tertiary structure and are conserved among the zinc-finger family of peptides. In our case, all the key positions agree well with the proposed upfield or downfield shift in both molecules (Table 4).

In the finger print of both molecules shown in Fig. 1 the most striking feature is that significant variation for HN and H^{α} chemical shifts only involve the four mutated residues. The $\Delta\delta$ of HN and H^{α} is on average less than 0.05 ppm (Table 3), already a strong indication that the two structures are very similar. In Fig. 2 the amide region of a 200-ms NOESY spectrum is shown for CP1 (a) and CP1m (b), indicating the entire pattern of NOE contacts characterising the α-helices. Figure 3 reports sequential (i,i+1) and medium-range (i,i+3) NOEs and also $^3J_{HNH^{\alpha}}$ coupling values and amide proton exchange rates. At the bottom of the figure the secondary structure elements are indicated.

The overall structure of both peptides is described by a short antiparallel β-hairpin involving residues 2–11 and an α-helix comprising residues 16–22. Characteristic for the β-element are the strong Lys^3_{α}-Ser^{10}_{α} and the $HNCys^4$-Ser^{10}_{α} NOE contacts typical of antiparallel strands. While the $HNPhe^{11}$-Lys^3_{α} NOE is not observable due to resonance degeneracy, the strong H^{α}_i-HN_{i+1} contacts are evidence of an extended conformation, which is also confirmed by the value of the coupling constants $^3J_{HNH^{\alpha}}$. The slow exchange of the amide proton of Cys^4 can be attributed to hydrogen bonding with the CO of Lys^9, which confirms the β-structure. It is noteworthy that in both molecules these structural elements are virtually identical.

TABLE 4
CONSERVED CHEMICAL SHIFTS CHARACTERISTICS FOR A $Cx_2Cx_{12}Hx_3H$ ZINC-FINGER SUBCLASS

Residue	Proton	Chemical shift (ppm)			
		CP1[a]		CP1m[a]	
		δ	Δδ	δ	Δδ
Cys^4	HN	9.22	0.91	9.20	0.89
Glu^6	HN	8.50	0.13	8.47	0.10
Ser^{10}	H^{α}	4.98	0.48	4.98	0.98
Phe^{11}	H^{ζ}	6.03	−1.31	6.03	−1.31
Lys^{14}	H^{α}	2.88	−1.48	2.88	−1.48
Leu^{17}	HN	6.94	−1.58	6.84	−1.68
	H^{α}	3.13	−1.25	3.13	−1.25
His^{20}	H^{δ}	6.82	−0.32	6.82	−0.32
His^{24}	H^{δ}	6.51	−0.63	6.51	−0.63

Values were taken from Lee et al. (1992)
[a] The Δδ values have been evaluated with respect to the values published by Wüthrich (1986).

The α-helix is well defined (Fig. 3) by the complete pattern of the HN_i-HN_{i+1}, H^{α}_i-HN_{i+3} and H^{α}_i-H^{β}_{i+3} NOE contacts for all the residues up to residue 22. Moreover, $^3J_{HNH^{\alpha}}$ coupling constant values for residues 16–22 were all less than 5.0 Hz and the pattern of slow-exchange amide protons also is compatible with at least five (CP1) and four (CP1m) hydrogen bonds in the helix. However, there is no evidence that the helix extends in both molecules beyond residue His^{22}. At that point we observe a slow conformational equilibrium involving residues 23–25,

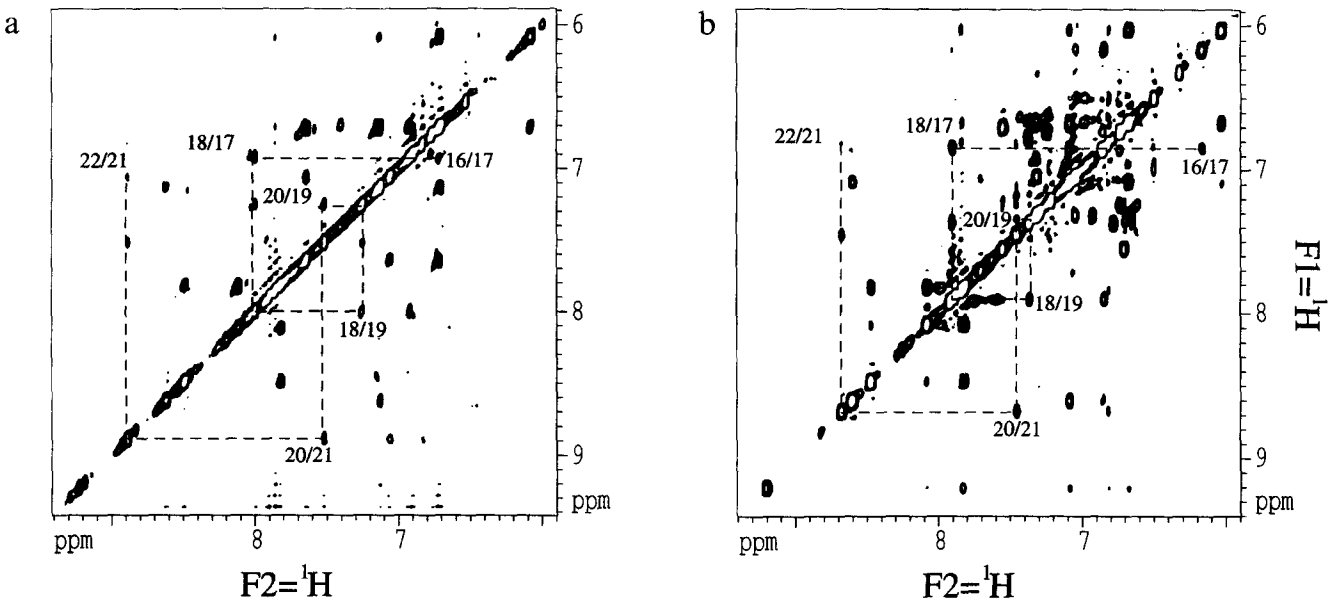

Fig. 2. Amide region of the NOESY experiment for CP1 (a) and CP1m (b) molecules. The experiment was performed with a jump and return excitation pulse. The mixing time used is 200 ms. NOE contacts for the helical elements are indicated by a dashed line.

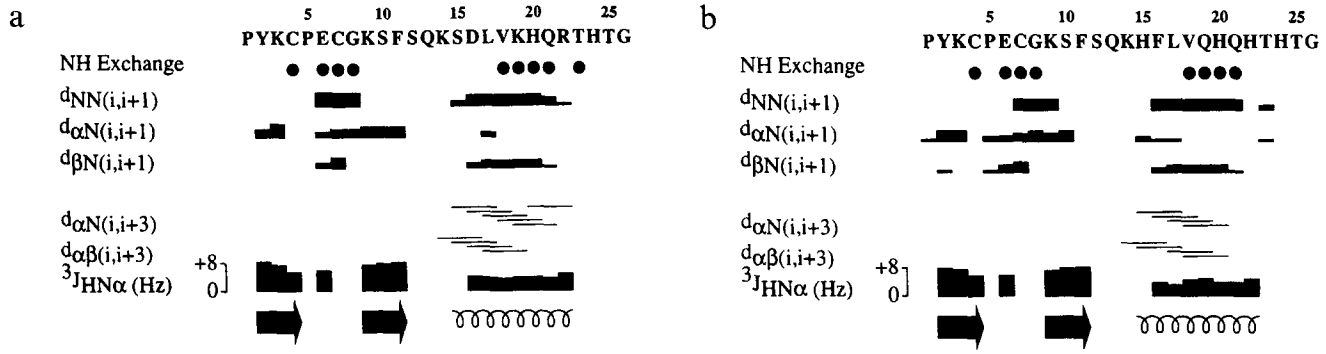

Fig. 3. Schematic representation for CP1 (a) and CP1m (b) of: (i,i+1) and (i,i+3) NOEs, amide proton exchange rates and $^3J_{HN\alpha}$ coupling constants. The characterization of the secondary structure elements is reported at the bottom of the figure as arrows (β-strands) and α-helices.

in the CP1m molecule, between a major and a minor conformation. In fact, two different resonances are found for the methyl group of Thr[23] and Thr[25] and for the imidazole group protons of His[22] and His[24], whereas no evidence for different backbone resonances are observed. The imidazole H^{ε}, for both histidine residues (His[22] and His[24]) and the H^{δ} of His[24], exhibit slow-exchange peaks in the TOCSY and NOESY experiments acquired in D_2O (Fig. 4). The major conformation of the His[24] side chain (the one with the upfield proton H^{δ} and H^{ε} signals) shows an appreciable lifetime, since it gives rise to NOE contacts that are fully compatible with the expected geometry for the complexation site of the zinc ion. By contrast, the alternative resonance position does not result in any detectable NOE contact. In the following, all the structure calculations and discussions will therefore implicitly assume the major conformation.

A further structural element characterising both molecules is the presence of the so-called rubredoxin-like knuckle in the turn Pro[5]-Gly[8] (Schwabe and Klug, 1994) confirmed by the NOE contacts observed between the HN of Cys[4] and those of residues Gly[8] and Lys[9], between the HN of Glu[6] and the HN and H^{β} of Cys[7], between the HN of Gly[8] and the HN of Lys[9]. Moreover, as expected for this type of turn the amide protons of Cys[4], Cys[7] and Gly[8] exhibit a slow exchange with the solvent, confirming the presence of hydrogen bonds.

The short loop connecting the β-structure to the α-helix is poorly defined. The peptide backbone in this region is exposed to the solvent, to the extent that the Ser[12] amide proton is not observed being in very fast exchange with water. However, the definition of this loop is better in CP1m than in CP1 due to several long-range NOE contacts involving their side chains.

Description and comparison of the three-dimensional structure of CP1 and CP1m

In Table 1 we report the results obtained by the

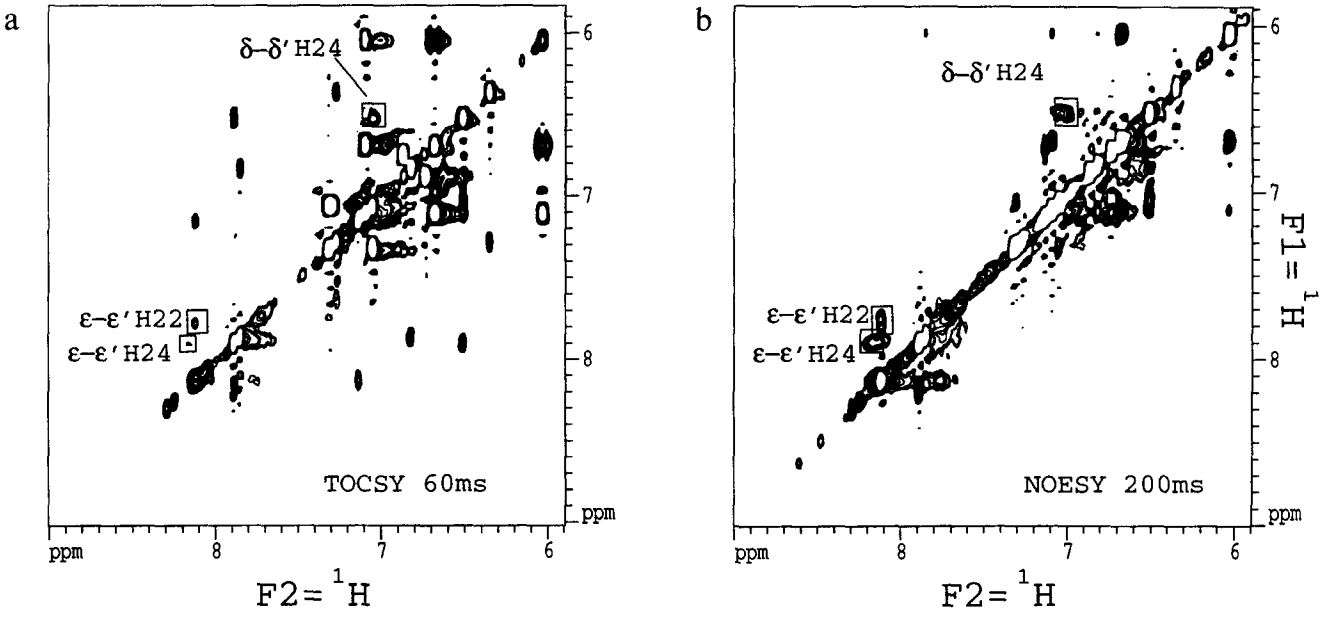

Fig. 4. Exchange cross-peaks for the histidine 22 and 24 residues of CP1m in the (a) TOCSY (60 ms) and (b) NOESY (200 ms) experiments acquired in D_2O. The exchange signals are identified with the apostrophe for the minor conformation.

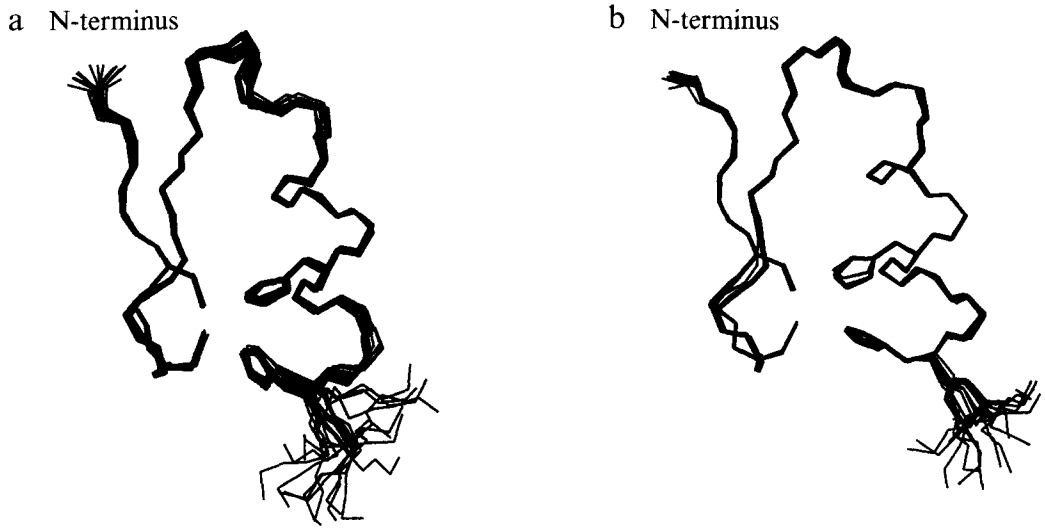

Fig. 5. Best superposition of the backbone heavy atoms for residues 2–11 and 16–22; the side chains of the residues Cys^4, Cys^7, His^{20} and His^{24} involved in the zinc-ion coordination shown. The 20 structures reported for CP1 (a) and CP1m (b) are the final result after the rEM.

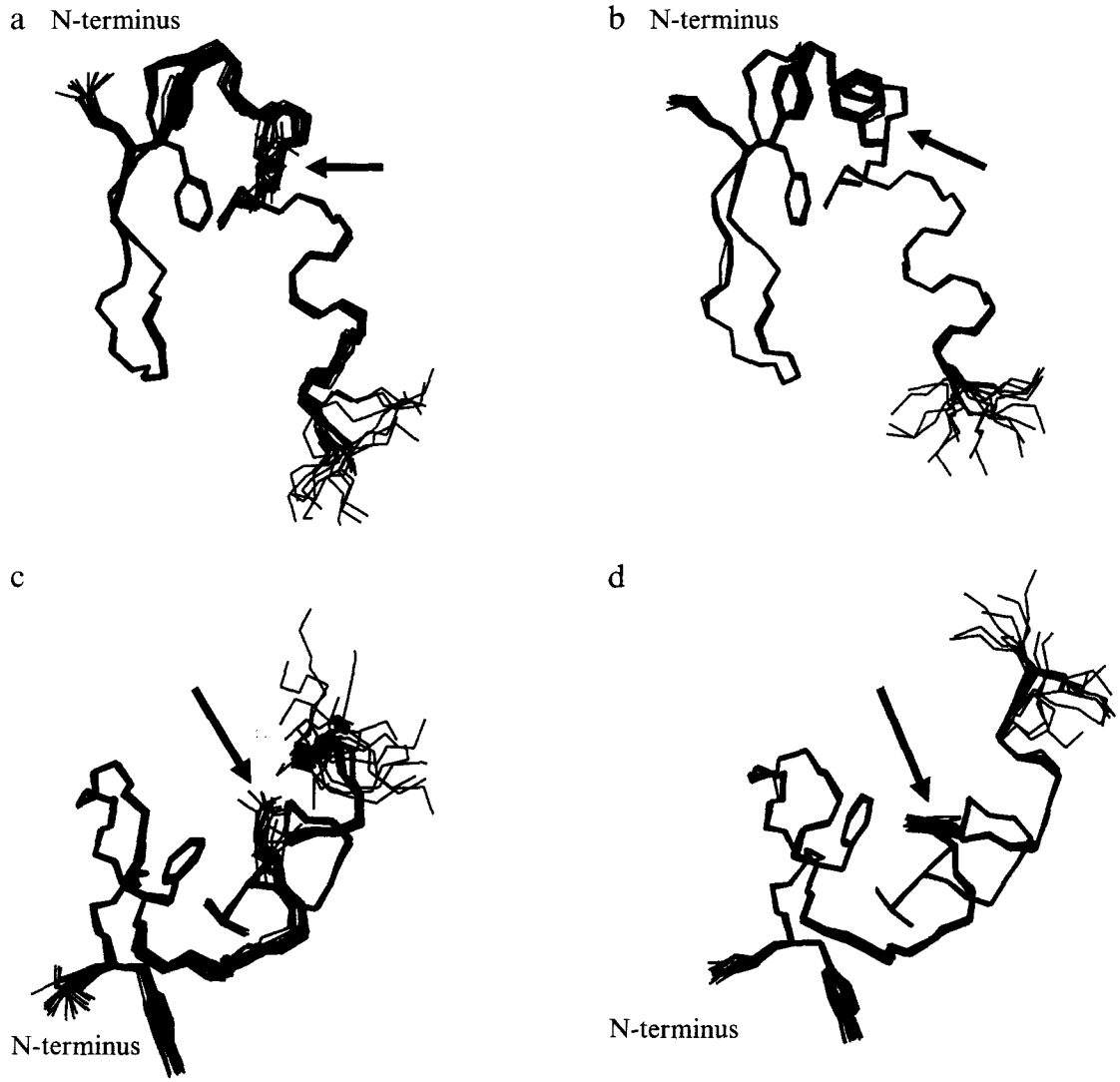

Fig. 6. Best superposition of the backbone heavy atoms for residues 2–11 and 16–22 plus the side chains of residues forming the hydrophobic core. Two different views are given for CP1 (a,c) and CP1m (b,d). Arrows indicates the position of the side chain for the mutated residue D16F.

TABLE 5
STATISTICAL EVALUATIONS OF THE SET OF STRUCTURES FOR CP1 AND CP1m

Residues	CP1				CP1m			
	Rmsd (Å)		<SD ϕ>	<SD ψ>	Rmsd (Å)		<SD ϕ>	<SD ψ>
	Backbone	All			Backbone	All		
2–11, 16–22	0.22 ± 0.05	1.70 ± 0.22	7.9	8.3	0.15 ± 0.05	1.15 ± 0.17	4.2	5.4
2–22	0.33 ± 0.08	1.80 ± 0.26			0.19 ± 0.05	1.28 ± 0.16		

DIANA REDAC routine and by rEM, and the correspondent superposition of the backbone heavy atoms for 20 structures of CP1 (a) and of CP1m (b) optimised for residues 2–11 and 16–22 (Fig. 5).

The characterisation of the zinc binding site is given in terms of the χ_1 and χ_2 dihedral angles of the residues involved in the formation of the complex (Table 2). A previous analysis by Chakrabarti (1989,1990) of cysteine and histidine residues involved in metal coordination has shown the statistical distribution of χ_1 and χ_2 angles of the residues involved in the metal binding (for cysteine, χ_2 was defined ad hoc as the torsion angle C^{α},C^{β},S^{γ},Zn). The distribution turned out to be trimodal: ordering the families of χ_1/χ_2 per number of occurrences we have (I) 60/270;

Fig. 7. Best superposition of backbone heavy atoms for residues 16–22, shown in yellow (CP1) and red (CP1m) and backbone atoms of residues 14–24 and side chains of the positions that were mutated in the library (S15H, D16F, K19Q, R22H).

TABLE 6
α-β PACKING ANGLE[a]

	α-β Packing angle (°)
CP1	18.1 ± 3.0
CP1m	28.5 ± 2.0
Xfin	19.1 ± 3.7
HEBP single zinc finger	21.3 ± 2.4
HEBP double zinc finger	27.9 ± 2.6, 11.4 ± 3.1

[a] This angle has been evaluated using vectorial geometry: the helix axis is represented by the normalized vector sum of the HN and OC vectors which are part of the helix (residues 16–22 for HN and 16–19 for OC), while the plane has been found using a minimisation routine that finds the plane that best fits the coordinates of the C^{α} of the six residues involved in the β-strands (2–4 and 9–11).

(II) 180/180; and (III) 180/90; with a spread of about ±30°. For both CP1 and CP1m Cys^4 belongs to the II/III family and Cys^7 to family I.

For histidine residues the distribution is more complex, described by seven χ_1/χ_2 families, again ordered per number of occurrences: (I) 270/90; (II) 270/180; (III) 270/270; (IV) 180/90; (V) 180/180; (VI) 60/270; and (VII) 90/90, also with a spread of about ±30°. In both CP1 and CP1m His^{20} can be included in family IV and His^{24} of family III; for CP1m the His^{24} χ_1 falls on the very edge of the family distribution (Table 2).

As it is evident from the structural statistics of the single set of structures and the comparison between the CP1 and CP1m lower-energy structures (Table 5) for both molecules the backbone appears to be very well characterised and conserved: for residues 2–11 and 16–22 the total average angular standard-deviation difference of the (φ,ψ) backbone torsion angles are (7.9,8.3) and (4.8,5.4) for CP1 and CP1m, respectively. The poorer characterisation of the backbone residues 12–14 and 23–26 is mainly due to a lack of informative NOEs involving their backbone protons. All the average atomic rmsd distributions were obtained considering pairs of structures and averaging over all possible pairs. For the backbone of residues 2–11 plus 16–22 the mean rmsd value is 0.22 ± 0.05 Å for CP1 and 0.15 ± 0.05 Å for CP1m. These values substantially increase when the less-defined region 12–14 is also considered. It is interesting that when side chains are included and the comparison is extended to almost the entire molecule (residues 2–22), the average atomic rmsd distribution is 1.80 ± 0.26 Å and 1.28 ± 0.15 Å for CP1 and CP1m, respectively, thereby evidencing that the difference of 19 constraints between the two structures mainly involves side chains.

The average value of the angle defined by the axis of the α-helix and the plane of the β-sheet is 18.1° and 29.5° for CP1 and CP1m, respectively. The evaluation of the angle is presented in Table 6. This angle allows a quantitative description of the difference in molecular packing, since it depends on the tertiary structure. The comparison with other natural zinc fingers for which a structure is available in the Brookhaven database is also reported in Table 6. In the Discussion these data will be commented on.

The localisation of the hydrophobic core, constituted by the invariant residues Tyr^2, Phe^{11} and Leu^{17}, is shown in Fig. 6 in two different views for CP1 (a,c) and for CP1m (b,d), highlighting also the mutation D16F that slightly affects the relative positions of the core residues.

A direct comparison of the lower-energy structures for CP1 and CP1m brings about an average atomic rmsd distribution of the backbone atoms of 0.68 Å, for residues 2–11 and 16–22 (Table 7). However, comparison of the single structural elements (residues 2–11 yields 0.71 Å and residues 16–22 gives 0.37 Å) indicates that the degree of local similarity is higher for the helical portion. The most dissimilar portion of the two molecules is located at the C-terminus, as evidenced (Table 8) from the average values of the φ and ψ dihedral angles for residues 22–24. They reflect a situation where the helix extends up to residue 22 for both molecules, and terminates in a wider turn in CP1m than in CP1, placing the imidazole ring of His^{24} in a slightly different position (Fig. 5). The portion of the helix comprising the mutations between CP1 and CP1m (residues 14–23) is very similar for the two molecules; a superposition of the backbone heavy atoms for residues 16–22 is shown in Fig. 7.

In the cumulative Ramachandran plot of the CP1 (a) and CP1m (b) set of structures (Fig. 8), residues falling in an energetically unfavourable area include Gly^8 and the terminal residues. The latter are ill-defined for lack of constraints, whereas Gly^8 has positive main-chain dihedral angles because of its position in the hairpin turn. We also note that the so-called 'rubredoxin-like knuckle' is indeed

TABLE 7
AVERAGE ATOMIC RMSD DISTRIBUTIONS OF CP1 AND CP1m[a]

Residues	Rmsd of backbone (Å)	Rmsd of backbone + specific side chains (Å)	Rmsd of all (Å)
2–11, 16–22	0.68		1.41
2–11	0.71		1.45
16–22	0.37		1.01[b]
2–11, 16–22, Phe^{11}, Leu^{17}, Tyr^2		0.80	
2–11, 16–22, Cys^4, Cys^7, His^{20}, His^{24}		0.72	

[a] Comparison held on the lower-energy structures for both CP1 and CP1m.
[b] This value has been calculated taking into account only the conserved residues.

very similar for both structures to that of rubredoxin itself. The atomic rmsd distributions of the backbone atoms for residues 5–8 (CP1 and CP1m), as compared to residues 7–10 of rubredoxin (Frey et al., 1987), which form one of the typical loops, are 0.73 Å and 0.67 Å for CP1 and CP1m, respectively. The two peptides then fall in the category of the 'full-knuckle zinc fingers' (Schwabe and Klug, 1994).

Discussion

The aim of this study was to investigate in greater detail to what extent mutations introduced in the helical segment of a classical zinc finger were tolerated without substantially altering the structure. Previously reported evidence based on the CD analysis of peptide mixtures had shown a pronounced tolerance in all but the first randomised position (residue 16), where introduction of hydrophobic/aromatic residues induced a dramatic change in the CD spectrum. This tolerance was even more noteworthy, since the structural stability of this naturally evolved motif was probed with an expanded set of amino acids, including several noncoded residues (Bianchi et al., 1995,1996). Of course, when applied to a population (each peptide mixture included on average 2×10^5 peptides), a feature like a conserved CD spectrum could not be taken as an absolute guarantee against the existence, within the population, of peptides with an altered structure. The CD evidence should be rather viewed as a strong indication for a conservation of the expected topology, to be confirmed by further spectroscopic analysis. Here we show confirmatory evidence for a ligand selected from the zinc-finger library, by comparing the structures of the selected ligand (CP1m) and the parent peptide (CP1) by NMR spectroscopy.

TABLE 8
AVERAGE BACKBONE DIHEDRAL ANGLES ϕ AND ψ FOR THE RESIDUES 22–24 OF CP1 AND CP1m

Molecule	Residue	ϕ (°)	ψ (°)
CP1	22	−61 ± 9	−33 ± 8
	23	−70 ± 12	−28 ± 16
	24	−62 ± 15	110 ± 79
CP1m	22	−64 ± 7	−46 ± 8
	23	−81 ± 7	−36 ± 16
	24	−1 ± 24	72 ± 61

The NMR characterisation immediately revealed that no major changes had occurred, since no significant variations in the chemical shifts of the backbone residues were noted. A detailed comparison first revealed that the topology is invariant. Both molecules fall in the same zinc-finger family, the 'full-knuckle' subfamily. There are no significant differences between the two structures. The use of the zinc-finger topology as a template for a conformationally constrained helical library therefore proved to be successful.

As expected, minor structural differences can still be observed. The mutations introduced in the helix (four out of seven residues) affect the tertiary packing, as indicated by the α-β packing angle: 18.1° and 28.5° for CP1 and CP1m, respectively. The former value of α-β packing angle is very similar to that obtained for Xfin (19.1°) (Lee et al., 1989), and the single zinc-finger domain of the Human Enhancer Binding Protein (HEBP; 21.3°) (Omichinski et al., 1990), whereas the latter is more similar to

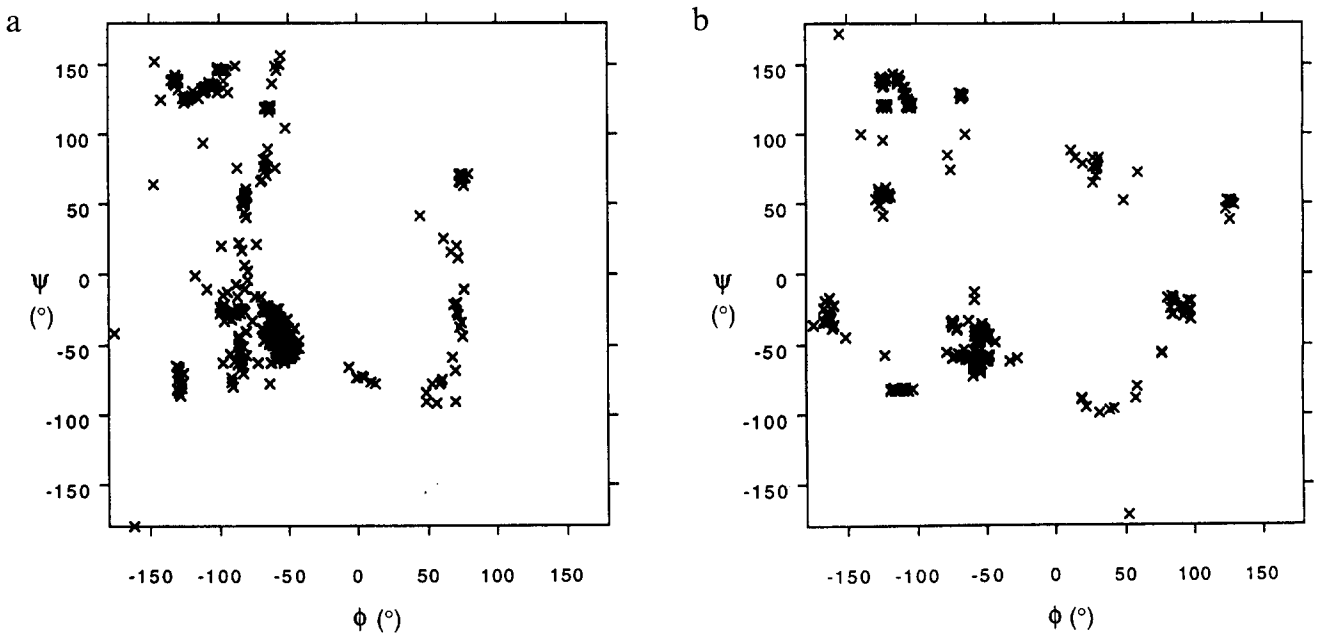

Fig. 8. Cumulative Ramachandran plot for CP1 (a) and CP1m (b) of all the 20 structures selected after restrained energy minimization.

that obtained from the first domain of the double zinc-finger domain of the HEBP (27.9°) (Omichinski et al., 1992). This difference, which will be discussed in detail, is caused mainly by two factors: (i) the mutation D16F; (ii) a slightly different position of the His^{24} imidazole ring in the zinc binding site.

The presence of the Phe^{16} ring slightly affects the relative position of the hydrophobic core with Leu^{17} methyl groups held about in the same position, while the Phe^{11} aromatic ring slightly turns and slides toward the Phe^{16} ring (Figs. 6a–d), resulting in NOE contacts between the aromatic rings. Other consequences are the stabilisation of the conformation of residues 12–14 in CP1m as compared to CP1, which allows the observation of new NOEs among the Phe^{16} aromatic moiety and Gln^{13} side chain and a slight shift and turn also of the Tyr^{2} aromatic ring, positioned on average in front of the backbone of residues 13 and 14. This different position is reflected in the differences of chemical shifts of amide, α-, β- and γ-protons of Gln^{13}, and mainly of β- and δ-protons of Lys^{14}. Moreover, in both molecules the side-chain methylene groups of Lys^{9} pack against the hydrophobic core constituted by Phe^{11} and Leu^{17}, leaving the charged group completely exposed. While in CP1 NOE contacts up to the δ-methylene group of Lys^{9} are observed, in CP1m the ε-methylene group gives an additional NOE with the Phe^{11} aromatic ring, indicating a restricted mobility of the ε-NH_3^+ group itself.

Another difference concerns the terminal residues (22–26) involved in zinc coordination. While in CP1 there is evidence for only one main conformation, in CP1m we observe a slow conformational-exchange equilibrium involving the zinc coordination site. A similar phenomenon has already been qualitatively described by Kochoyan et al. (1991) and Xu et al. (1991), and in a more quantitative approach taken into account by Clore et al. (1991). In the latter case, the Cys^{2} His^{2} single zinc finger of the Human Enhancer Binding Protein was analysed. They assumed the existence of one major and two minor conformations in order to fit the experimental data. In our case, for CP1m, the hypothesis of one major and one minor conformation accounts for the data with the major conformation giving rise to several self-consistent NOEs, involving residues 22–24. By contrast, the minor conformation does not show any observable NOEs and was not investigated any further. If we consider only the major conformation, the difference in the terminal part of the helix between CP1 and CP1m does not seem to affect the overall geometry of the zinc binding site. In fact, the main difference between the CP1 and CP1m structures is observed in the terminal part of the helix (residues 22–24) having a wider turn in CP1m as compared to CP1, reflected by the average corresponding values of ϕ and ψ values. Consequently in CP1, the imidazole ring of His^{24} participates in the zinc coordination site by slightly changing its χ_1 and χ_2 dihedral angles. This effect is illustrated by the i,i+3 contacts between residues 20 and 23 absent in CP1m, thus indicating the helix to be somewhat destabilised in the latter molecule. The difference in the exchange regime of the Thr^{23} amide proton further supports this interpretation, being fast for CP1m and slow for CP1.

The difference observed for the conformation of the terminal residues in our opinion could only partially be explained by the difference in pH between the two samples, 6.6 for CP1 and 6.3 for CP1m. In fact this difference should only produce a slightly different degree of protonation of His^{24}, since the pKa measured by Krizek et al. (1991) for CP1 is around 5.3. In the destabilisation of the terminal part of CP1m a major role could be played by the presence of His^{22} estimated to be approximately 50% charged at pH 6.3, which replaces an arginine residue (Arg^{22}) in CP1 whose charge is independent of the pH. The different side chains for position 22, arginine in CP1 and histidine in CP1m, can also differently affect the conformation or the stability by a direct interaction with the so-called hydrophobic core. Although this is plausible, no evidence of such interaction is present in the NMR data, due to the absence of structurally relevant NOEs among the above-mentioned side chains and other non-consecutive residues.

As far as the β-structure (residues 2–11 comprising the β-strands and the rubredoxin-like loop) is concerned, the differences are mainly due to the packing of Glu^{6}, Lys^{9} and Phe^{11} side chains. For Glu^{6}, the NOEs observed to the His^{24} imidazole ring are sensitive to a change in conformation of the latter residue, thus reflecting the widening of the turn involving residues 22–24 in CP1m. All these factors together influence the β-structure in the loop at position of Glu^{6} and in the second strand (Lys^{9} and Phe^{11}), and are responsible for the higher difference in the atomic root mean square observed in the comparison of the corresponding structural elements between CP1 and CP1m.

Overall, the data presented in this paper confirm previous findings that in zinc fingers the local geometrical features of the secondary structure elements are very tolerant to mutations. With regard to future application of our zinc-finger library, the remarkable structural invariance shown here suggests that a very rapid assessment of the preservation of the pharmacophore, even for a large panel of selected ligands, is reasonably possible, requiring (together with CD) solely the measurements of the backbone proton resonances. Moreover, regarding the design of new CHCPL, this work confirms that CD analysis of peptide mixtures, limited as it may appear, is of great practical value: in fact from the direct comparison of the spectra of the model compound and of the library it is possible to quickly assess the necessity of the redesign or as a first approximation the success in obtaining the model topography.

NMR analysis of zinc-finger peptides like those coming

from a combinatorial library, may shed light on the intrinsic 'folding code' of this structural motif. In natural zinc fingers some of the positions which have been randomised in our library (amino acids 16 and 19) are used to bind DNA (Pavletich and Pabo, 1991). In the natural sequence pool, therefore, we will find residues meeting two conditions: (i) ability of interaction with DNA; and (ii) compatibility with maintenance of the fold. At variance, analysis of our artificial sequence pool is only addressing the latter condition. For example, we found that in the first randomised position in the library (residue 15, starting point of the helix) a phenylalanine residue is not allowed, while this is not observed for, e.g., tyrosine (Bianchi et al., 1995a,b). It is tempting to speculate that certain hydrophobic residues at this position are not compatible with the correct packing of the critical, highly conserved hydrophobic triad (Krizek et al., 1991; Michael et al., 1992). To probe this hypothesis and to study in detail the structural alterations responsible for the alterations in the CD curve and Co(II) absorption spectrum, we are now examining by NMR individual zinc-finger peptides which display these spectral features.

Acknowledgements

We are grateful to Dr. Maria Nicotra and Dr. Andrew Wallace for experimental assistance and stimulating discussions, to Dr. Armin Lahm and Dr. Anna Tramontano for a critical reading of the paper, and to Manuela Emili for art work. We also acknowledge the continuous support of Prof. R. Cortese.

References

Bianchi, E., Folgori, A., Wallace, A., Nicotra, M., Acali, S., Phalipon, A., Barbato, G., Bazzo, R., Cortese, R., Felici, F. and Pessi, A. (1995) *J. Mol. Biol.*, **247**, 154–160.

Bianchi, E., Folgori, A., Wallace, A., Nicotra, M., Barbato, G., Bazzo, R., Cortese, R., Felici, F. and Pessi, A. (1996) In *Peptides: Chemistry, Structure and Biology* (Proceedings of the 14th American Peptide Symposium, 1995), Kaumaya, P.T.P. and Hodges, R.S. (Eds), Mayflower Scientific Ltd., Kingswinford, U.K., pp. 284–286.

Chakrabarti, P. (1989) *Biochemistry*, **28**, 6081–6085.

Chakrabarti, P. (1990) *Protein Eng.*, **4**, 57–63.

Clore, G.M., Omichinski, J.G. and Gronenborn, A.M. (1991) *J. Am. Chem. Soc.*, **113**, 4350–4351.

Frey, M., Sieker, L., Payan, F., Haser, R., Bruschi, M., Pepe, G. and La Gall, J. (1987) *J. Mol. Biol.*, **197**, 525–541.

Gallop, M.A., Barrett, R.W., Dower, W.J., Fodor, S.P.A. and Gordon, E.M. (1994) *J. Math. Chem.*, **37**, 1233–1251.

Güntert, P., Braun, W., Billeter, M. and Wüthrich, K. (1989) *J. Am. Chem. Soc.*, **111**, 3997–4004.

Güntert, P. and Wüthrich, K. (1991) *J. Biomol. NMR*, **1**, 447–456.

Güntert, P., Qian, Y.Q., Otting, G., Muller, M., Gehring, W. and Wüthrich, K. (1991) *J. Mol. Biol.*, **217**, 531–540.

Kim, C.A. and Berg, J.M. (1993) *Nature*, **362**, 267–270.

Kochoyan, M., Havel, T.F., Nguyen, D.T., Dahl, C.E., Keutmann, H.T. and Weiss, M.A. (1991) *Biochemistry*, **30**, 3371–3386.

Krizek, B.A., Amann, B.T., Kilfoil, V.J., Merckle, D.L. and Berg, J.M. (1991) *J. Am. Chem. Soc.*, **113**, 4518–4523.

Lee, M.S., Gippert, G.P., Soman, K.V., Case, D.A. and Wright, P.E. (1989) *Science*, **245**, 635–637.

Lee, M.S., Mortshire-Smith, R.J. and Wrigth, P.E. (1992) *FEBS Lett.*, **309**, 29–32.

Marion, D., Ikura, M., Tschudin, R. and Bax, A. (1989) *J. Magn. Reson.*, **85**, 393–399.

Michael, S.F., Kilfoil, V.J., Schmidt, M.H., Amann, B.T. and Berg, J.M. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 4796–4800.

Miller, J., McLachlan, A.D. and Klug, A. (1985) *EMBO J.*, **4**, 1609–1614.

Moore, G.J. (1994) *Trends Pharmacol. Sci.*, **15**, 124–129.

Omichinski, J.G., Clore, M.G., Appella, E., Sagakuchi, K. and Gronenborn, A.M. (1990) *Biochemistry*, **29**, 9324–9334.

Omichinski, J.G., Clore, G.M., Robien, M., Sakaguchi, K., Appella, E. and Gronenborn, A.M. (1992) *Biochemistry*, **31**, 3907–3917.

Pavletich, N.P. and Pabo, C.O. (1991) *Science*, **252**, 809–817.

Schwabe, J.W.R. and Klug, A. (1994) *Nat. Struct. Biol.*, **1**, 345–349.

Vedani, A. and Huhta, D.W. (1990) *J. Am. Chem. Soc.*, **112**, 4759–4767.

Xu, R.X., Horvarth, S.J. and Klevit, R.E. (1991) *Biochemistry*, **30**, 3365–3371.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY.