Journal of Biomolecular NMR, 3 (1993) 271–284
ESCOM

J-Bio NMR 120

The three-dimensional structure of acyl-coenzyme A binding protein from bovine liver: Structural refinement using heteronuclear multidimensional NMR spectroscopy

Kim Vilbour Andersen and Flemming M. Poulsen*

Carlsberg Laboratorium, Kemisk Afdeling, Gamle Carlsberg Vej 10, DK-2500 Valby, Copenhagen, Denmark

Received 9 December 1992
Accepted 19 February 1993

Keywords: Acyl-coenzyme A binding protein; Nuclear magnetic resonance spectroscopy; Simulated annealing; Coupling constants; Nuclear Overhauser effects; Diazepam binding inhibitor; Endozepine; ^{15}N

SUMMARY

The 3D structure of bovine recombinant acyl-coenzyme A binding protein has been determined using multidimensional heteronuclear magnetic resonance spectroscopy in a study that combines investigations of ^{15}N-labeled and unlabeled protein. The present structure determination is a refinement of the structure previously determined (Andersen, K.V. and Poulsen, F.M. (1992) *J. Mol. Biol.*, **226**, 1131–1141). It is based on 1096 distance restraints and 124 dihedral angle restraints of which 69 are for ϕ-angles and 8 for chiral centers and 47 for prochiral centers. The new experimental input for the structure determination has provided an increase of 263 distance restraints, 5 ϕ-angle restraints, and 32 χ-angle restraints in 2 chiral centers, and 31 prochiral centers restraining an additional 23 χ^1, 8 χ^2, and 1 χ^3 angles. The increase of 300 distance and dihedral angle restraints representing an additional 30% of input parameters for the structure determination has been shown to be in agreement with the first structure. A set of 29 structures has been calculated and each of the structures has been compared to a mean structure to give an atomic root mean square deviation of 0.44 ± 0.12 Å (1 Å is 0.1 nm) for the backbone atoms C, C^{α}, and N in the four α-helices A1, residues 4–15, A2, residues 21–36, A3, residues 51–62 and A4, residues 65–84. The loop-region of residues Gly^{45}–Lys^{50} could not be defined by the restraints obtained by NMR.

The program PRONTO has been used for the spectrum analysis, assignment of the individual nuclear Overhauser effects, the integration of the cross peaks, and the measurement of the coupling constants. The programs DIANA, X-PLOR and INSIGHT have been used in the structure calculations and evaluations.

INTRODUCTION

In the process of determining the 3D structure of the complex between palmitoyl-coenzyme A and recombinant bovine acyl-coenzyme A binding protein, rACBP, we have recorded spectra of the fully ^{15}N enriched rACBP in complex with the ligand. This led to an increase in the number

*To whom correspondence should be addressed.

of distance and dihedral angle constraints that had previously been available from homonuclear ^{1}H NMR spectroscopy of the protein–ligand complex (Kragelund et al., 1993). It was obvious, therefore, to measure a new *apo-rACBP* structure that could serve as an equivalent reference to the structure of the *holo*-rACBP. The first structure determination of bovine rACBP (Andersen and Poulsen, 1992) was therefore extended to include the many additional data that became available from the spectra of ^{15}N-rACBP in a new structure determination with the purpose of obtaining an even higher degree of refinement of the rACBP structure.

METHODS

The experimental procedures concerning sample preparation and the recording and processing of ^{1}H–^{1}H 2D NMR spectra has been reported in previous papers (Andersen et al., 1991; Andersen and Poulsen, 1992). The preparation of the ^{15}N-rACBP sample was described in Kragelund et al. (1993). Three-dimensional NOESY-HMQC (Marion et al., 1989a; Zuiderweg and Fesik, 1989) and TOCSY-HMQC (Marion et al., 1989b) as well as 2D ^{15}N–^{1}H HSQC spectra were recorded in 90/10 $H_2O/^2H_2O$ at pH 7.0 at 298 K. In the 3D spectra the number of complex points recorded was 128, 32 and 512 in the t_1 (^{1}H), t_2 (^{15}N), and t_3 (^{1}H) dimensions, respectively. Lorentz-to-Gauss weighting and zero filling were employed in all three dimensions. These spectra were recorded and processed according to the States-TPPI scheme (Marion et al., 1989c). For all spectra the ^{1}H spectral width was 12.62 ppm, and where applicable the ^{15}N spectral width was 51.39 ppm. All spectra were recorded on a Bruker 600 MHz AMX spectrometer, equipped with a triple resonance 5-mm probe. The data were processed using the program MNMR and the spectrum analysis was performed with the software package PRONTO (Kjær et al., 1991; PRONTO Software Development and Distribution, Copenhagen, Denmark).

The methods used to determine the structure from NMR data are the hybrid distance geometry/simulated annealing method, followed by restrained molecular dynamics, as was described previously (Ludvigsen and Poulsen, 1992; Andersen and Poulsen, 1992; Kragelund et al., 1993) with references to Ludvigsen et al. (1991) and Montelione et al. (1989) for measurements of homo- and heteronuclear coupling constants, to Williamson et al. (1985) and Clore et al. (1985, 1986) for classification of NOEs and setting of restraints boundaries, to Günthert et al. (1991) for the use of the computer program DIANA, to Clore et al. (1985, 1986) and Nilges et al. (1988) for the use of the simulated annealing protocol, to Brooks et al. (1983) for the CHARMM potentials, and to Brünger (1992) for the use of the program X-PLOR version 3.0. The method involves a three-step calculation strategy. First a series of structures is calculated using the program DIANA. The structures produced by DIANA are refined by simulated annealing followed by restrained molecular dynamics; both simulations are done using the program X-PLOR. The 6.55 ps simulated annealing involves 3.75 ps molecular dynamics at elevated temperatures of 1000 K followed by 2.8 ps slow cooling to 300 K. The force-field used for the simulated annealing is a

---

Abbreviations: ACBP, acyl-coenzyme A binding protein from bovine liver; rACBP, ACBP encoded by the recombinant gene in *E. coli*: ^{15}N-rACBP, fully ^{15}N labelled rACBP; NMR, nuclear magnetic resonance; rmsd, root mean square difference; NOE, nuclear Overhauser effect; TOCSY, total correlation spectroscopy; NOESY, nuclear Overhauser enhancement spectroscopy; HMQC, heteronuclear multiple-quantum correlation; HSQC, heteronuclear single-quantum correlation; TPPI, time-proportional phase incrementation; ppm, parts per million.

simple force-field, only including terms for bonds, angles, planarity of rings and peptide bonds, and a simple repulsive van der Waals term. The final refinement step is 2.0 ps of restrained molecular dynamics using a modified CHARMM force-field including a special hydrogen bond potential (Brünger, 1992) followed by restrained minimization to locate the nearest energy minimum.

RESULTS AND DISCUSSION

From the knowledge of the assignment of the 2D ^1H–^1H spectra, the 3D spectra were easily assigned using the tools in PRONTO. The resonances of the peptide backbone have been completely assigned, with the exception of the first two residues which were not seen in the previous analysis either. Most of the amino acid side chains were assigned with the exception of the outermost parts of some of the long side chains. The ^{15}N side-chain atoms of the tryptophan, glutamate, and asparagine residues were all assigned. The assignments are summarized in Table 1.

As a result of the new analysis five additional ϕ-angle restraints were added to the list of dihedral angle restraints to make a total of 69 ϕ-angles determined for the calculations. The majority of these, 60, were in the α-helical range (−17° to −97°), the remaining 9 ϕ-angle restraints were in the range of −80° to −160°; these are summarized in Table 1 and Fig. 3A. The stereospecific assignments have been extended significantly in comparison to the previous calculation and these are summarized in Table 1. The conformation in 8 chiral centers and 47 prochiral centers was determined, resulting in the definition of χ^1-angle restraints for 46 residues and χ^2-angle restraints for 8 residues as summarized in Table 1 and Fig. 3C, and one χ^3-angle restraint for Arg^{43} of 180° ± 60°.

The analysis of 2D and 3D NOESY spectra has led to an almost complete assignment of the spectra. Of the assigned NOEs, 1096 were identified as non-trivial by DIANA (833 NOEs in the previous calculation) and subsequently used as distance restraints according to their intensities as 118 strong, 304 medium and 674 weak NOEs (61, 213 and 559, respectively, in the previous calculation). The distributions with respect to intra-residue, sequential ($i,j|i-j|=1$), short range ($i,j|1<|i-j|\leq 5$), and long range ($i,j|i-j|>1$) NOEs were 200, 313, 267 and 316, respectively, and the distribution by residue of these is shown Fig. 1.

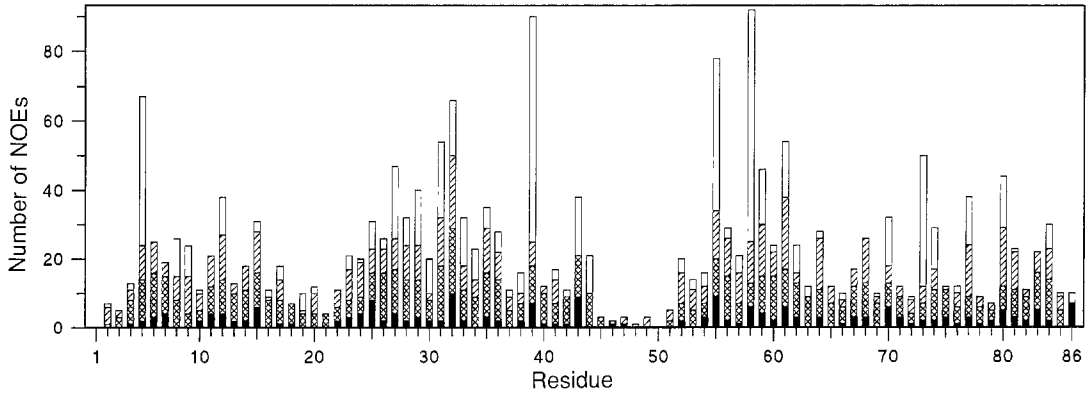

Fig. 1. The distribution of identified and non-trivial NOEs by residue. The columns showing the number of intra-residue NOEs are solid bars, sequential NOEs ($i,j|i-j|=1$) are crossed bars, short-range NOEs ($i,j|i-j|\leq 5$) are hatched bars, and long-range NOEs ($i,j|i-j|>5$) are open bars.

TABLE 1
^{1}H AND ^{15}N CHEMICAL SHIFTS IN rACBP AT pH 7.0 AND 298 K[a]

Residue	N	H^N	H^α	H^β [b]	Others[b]	$^3J_{H^NH^\alpha}$ ϕ, χ^1, χ^2 [c]
Ser[1]			4.08	4.18, 4.03		(-,-,-)
Gln[2]					2.46 (H^γ); 8.25, 6.57 ($H^{\varepsilon 2}$); 111.9 (N^ε)	(-,-,-,-)
Ala[3]	118.0	8.55	4.29	1.47		(<5 Hz, −57°)
Glu[4]	115.6	8.05	4.07	2.13^3, 2.30^2	2.36, 2.45 (H^γ)	(5 Hz, −57°,-,-)
Phe[5]	121.2	8.40	4.26	3.38^3, 3.10^2	7.20 (H^δ); 6.49 (H^ε); 6.37 (H^ζ)	(4 Hz, −57°, 180°)
Asp[6]	116.9	8.89	4 27	2.68^3, 2.76^2		(4 Hz, −57°, −60°)
Lys[7]	118.0	7.76	4.17	2.01^3, 1.91^2	1.38 (H^γ)	(4 Hz, −57°, 180°,-)
Ala[8]	121.3	8.06	4.33	1.70		(4 Hz, −57°)
Ala[9]	115.6	8.40	4.82	1.22		(-,-)
Glu[10]	114.8	7.43	4.30	2.20, 2.20	2.35, 2.45 (H^γ)	(7 Hz,-,-,-)
Glu[11]	116.8	8.50	3.95	1.70^3, 2.27^2	2.45, 2.71 (H^γ)	(<5 Hz, −57°, −60°,-)
Val[12]	113.5	8.14	3.65	2.22	0.93^3, 1.04^2 (H^γ)	(2 Hz, −57°, 60°)
Lys[13]	114.9	7.10	3.99	1.47, 1.60	1.29 (H^γ)	(6 Hz, −57°,-,-)
His[14]	113.3	7.96	4.76	3.02, 3.56	7.35 ($H^{\delta 2}$); 8.13 ($H^{\varepsilon 1}$)	(9 Hz, −120°, −120°)
Leu[15]	116.4	6.91	4.70	1.58^3, 1.78^2	2.10 (H^γ); 1.02^1, 0.88^2 (H^δ)	(2 Hz, −57°, −60°, 180°)
Lys[16]	117.9	8.47	3.96	1.94		(3 Hz, −57°,-,-)
Thr[17]	109.2	7.11	4.29	3.81	1.15 ($H^{\gamma 2}$)	(9 Hz, −120°, −60°)
Lys[18]	119.5	7.88	4.27	1.67		(3 Hz, −57°,-,-)
Pro[19]			3.88	1.69, 2.35	1.52, 1.78 (H^γ); 3.43, 4.04 (H^δ)	(-,-,−30°,-)
Ala[20]	119.1	8.18	4.37	1.57		(6 Hz, −57°)
Asp[21]	120.6	9.11	4.32	2.69		(3 Hz, −57°,-)
Glu[22]	114.7	9.58	4.14	2.13	2.42 (H^γ)	(3 Hz, −57°, -,-)
Glu[23]	117.2	7.09	4.22	1.84^3, 2.71^2	2.24, 2.46 (H^γ)	(7 Hz, -, −60°, -)
Met[24]	115.5	8.30	4.43	2.24, 2.24	2.46, 2.74 (H^γ); 2.05 (H^ε)	(4 Hz,-,-,-)
Leu[25]	118.9	8.57	4.28	1.56^3, 2.06^2	1.88 (H^γ); 0.99^1, 0.87^2 (H^δ)	(5 Hz, −57°, −60°, 180°)
Phe[26]	120.4	7.86	4.36	3.64^3, 3.33^2	7.09 (H^δ); 7.01 (H^ε); 6.88 (H^ζ)	(2 Hz, −57°, 180°)
Ile[27]	118.8	8.41	4.09	2.50	1.39, 1.78 ($H^{\gamma 1}$); 1.11 ($H^{\gamma 2}$); 0.51 (H^δ)	(6 Hz, −57°, −60°,-)
Tyr[28]	117.3	8.93	4.26	3.02^3, 3.43^2	7.04 (H^δ); 6.87 (H^ε)	(2 Hz, −57°, 180°)
Ser[29]	110.0	8.45	3.32	4.00^3, 3.70^2	4 99 ($H^{\gamma 1}$)	(4 Hz, −57°, 60°)
His[30]	116.4	8.18	3.85	3.43^3, 3.17^2		(5 Hz, −57°, −60°)
Tyr[31]	118.9	8.67	3.67	3.06^3, 3.47^2	6.52 (H^δ); 6.57 (H^ε)	(3 Hz, −57°, 180°)
Lys[32]	116.0	7.65	3.20	-0.66^3, 0.15^2	0.54^3, -0.04^2 (H^γ); 0.79, 1.13 (H^δ); 2.48 (H^ε)	(3 Hz, −57°, −60°, −60°)
Gln[33]	118.3	7.94	4.13	1.42	2.38 (H^γ); 108.4 (N^ε)	(2 Hz, −57°,-,-)
Ala[34]	111.6	7.60	4.00	0.85		(<5 Hz, −57°)
Thr[35]	101.9	7.13	4.09	3.90	4.46 ($H^{\gamma 1}$); 0.72 ($H^{\gamma 2}$)	(10 Hz, −120°, 60°)
Val[36]	121.3	8.54	3.91	2.21	1.06^1, 1.00^2 (H^γ)	(8 Hz, −120°, 180°)
Gly[37]	107.2	7.79	4.33, 4.00			(-,-)
Asp[38]	117.8	8.14	4.48	2.44^3, 2.50^2		(2 Hz, −57°, −60°)
Ile[39]	119.7	8.23	2.26	−0.56	-0.51^3, 0.07^2 ($H^{\gamma 1}$); 0.38 ($H^{\gamma 2}$); −0.80 (H^δ)	(<5 Hz, −57°, 180°, 180°)
Asn[40]	118.9	6.66	4.78	2.58, 3.12	6.46, 7.54 (H^δ); 105.8 (N^δ)	(10 Hz, −120°,-)
Thr[41]	106.1	7.12	4.75	4.51	1.41 ($H^{\gamma 2}$)	(10 Hz, −120°, 60°)
Glu[42]	120.3	8.72	4.34	1.87	2.26 (H^γ)	(6 Hz,-,-,-)
Arg[43]	123.7	8.15	2.38	0.91^3, -0.14^2	0.62^3, 1.27^2 (H^γ); 2.83^3, 2.61^2 (H^δ)	(2 Hz, −57°, 180°, 180°)
Pro[44]			4.33	1.81, 2.52	1.62, 1.86 (H^γ); 2.89^3, 3.65^2 (H^δ)	(-,-,−30°,-)
Gly[45]	103.0	8.32	4.01, 3.75			(-,-)
Met[46]	113.2	8.10	3.97	2.01, 2.08	2.57 (H^γ)	(<5 Hz, −57°, -,-)
Leu[47]	114.2	8.19	4.16	1.63	1.60 (H^γ); 0.90, 0.93 (H^δ)	(6 Hz, −57°, -,-)

TABLE 1 (continued)

Residue	N	H^N	H^α	H^β [b]	Others[b]	$^3J_{H^NH^\alpha}$ ϕ, χ^1, χ^2 [c]
Asp^{48}	117.4	7.53	4.78	3.00^3, 2.41^2		(-,-,-)
Phe^{49}	119.7	7.73	4.18	3.12, 3.25	7.34 (H^δ); 7.47 (H^ε); 7.40 (H^ζ)	(2 Hz,-,-)
Lys^{50}	117.9	8.83	4.15	1.79		(<5 Hz, −57°,-,-)
Gly^{51}	106.1	8.19	3.91, 3.83			(-,-)
Lys^{52}	121.3	8.73	3.91	1.86	1.33 (H^γ); 1.52 (H^δ)	(4 Hz, −57°,-,-)
Ala^{53}	118.1	7.69	4.34	1.58		(4 Hz, −57°)
Lys^{54}	115.6	7.60	3.82	1.68	0.63, 1.54 (H^γ); 1 00, 1.17 (H^δ)	(4 Hz, −57°,-,-)
Trp^{55}	118.2	8.10	4.34	3.52^3, 3.30^2	7.09 ($H^{\delta1}$); 9.22 ($H^{\varepsilon1}$); 8.03 ($H^{\varepsilon3}$); 7.01 ($H^{\eta2}$); 7.36 (H^{ζ}); 7.13 ($H^{\zeta3}$); 128.6 (N^ε)	(5 Hz, −57°, 180°)
Asp^{56}	115.8	9.38	4.54	2.87^3, 2.96^2		(2 Hz, −57°, −60°)
Ala^{57}	118.8	7.91	4.28	1.73		(3 Hz, −57°)
Trp^{58}	117.3	8.07	4.64	3.59^3, 3.04^2	7.51 ($H^{\delta1}$); 9.93 ($H^{\varepsilon1}$); 7.59 ($H^{\varepsilon3}$); 7.17 ($H^{\eta2}$); 7.09 ($H^{\zeta2}$); 7.14 ($H^{\zeta3}$); 126.9 (N^ε)	(3 Hz, −57°, 180°)
Asn^{59}	119.6	9.35	3.75	2.57^3, 1.45^2	6.61, 7.82 (H^δ); 111.4 (N^δ)	(5 Hz, −57°, 180°)
Glu^{60}	113.3	7.72	4.01	2.16		(4 Hz, −57°,-,-)
Leu^{61}	112.5	7.49	4.26	1.20^3, 2.00^2	1.76 (H^γ); 0.01^1, 0.52^2 (H^δ)	(8 Hz, −120°, −60°, 180°)
Lys^{62}	121.3	7.21	4.13	1.62^3, 1.85^2	1.31 (H^γ)	(4 Hz, −57°, 180°,-)
Gly^{63}	114.2	10.03	4.40, 3.78			(-,-)
Thr^{64}	117.2	8.31	4.33	4.06	5.90 ($H^{\gamma1}$); 1.49 ($H^{\gamma2}$)	(6 Hz,-,−60°)
Ser^{65}	124.5	9.89	4.36	4.18, 4.39		(3 Hz, −57°,-)
Lys^{66}	120.4	8.92	3.91	1.78		(3 Hz, −57°,-,-)
Glu^{67}	114.0	8.81	3.95	1.97^3, 2.05^2	2.44 (H^γ)	(3 Hz, −57°, −60°,-)
Asp^{68}	118.9	8.00	4.55	2.66^3, 2.86^2		(6 Hz, −57°, −60°)
Ala^{69}	121.4	8.52	4.30	1.43		(2 Hz, −57°)
Met^{70}	114.1	8.60	3.82	1.88^3, 2.17^2	2.43^3, 1.40^2 (H^γ); 1.71 (H^ε)	(2 Hz, −57°, −60°, 180°)
Lys^{71}	115.8	7.78	4.09	2.03	1.61 (H^γ)	(3 Hz, −57°,-,-)
Ala^{72}	117.2	7.87	4.31	1.61		(4 Hz, −57°)
Tyr^{73}	117.3	8.53	3.93	3.79^3, 3.17^2	7.02 (H^δ); 6.77 (H^ε)	(<5 Hz, −57°, 180°)
Ile^{74}	117.3	8.72	3.42	2.03	1.09, 2.30 ($H^{\gamma1}$); 1.07 ($H^{\gamma2}$)	(2 Hz, −57°, −60°,-)
Asp^{75}	115.6	8.38	4.37	2.64^3, 2.77^2		(3 Hz, −57°, −60°)
Lys^{76}	118.7	8.04	4.11	1.89^3, 1.56^2		(<5 Hz, −57°, 180°,-)
Val^{77}	118.9	8.60	3.32	2.18	1.01^1, 0.59^2 (H^γ)	(3 Hz, −57°, 180°)
Glu^{78}	115.7	8.32	4.02	2.22		(3 Hz, −57°,-,-)
Glu^{79}	118.2	7.77	4.00	2.37^3, 2.29^2	1.99 (H^γ)	(4 Hz, −57°, −60°,-)
Leu^{80}	117.5	8.92	4.05	1.61^3, 2.11^2	2.17 (H^γ); 0.78^1, 0.89^2 (H^δ)	(6 Hz, −57°, −60°, 180°)
Lys^{81}	117.4	9.00	3.70	1.88^3, 1.76^2	1.25 (H^γ)	(3 Hz, −57°, 180°,-)
Lys^{82}	115.6	7.37	4.01	1.42, 1.89	1.61 (H^γ)	(5 Hz, −57°,-,-)
Lys^{83}	115.7	7.78	3.88	1.19^3, 1.46^2	0.31, 1.02 (H^γ)	(5 Hz, −57°, 180°,-)
Tyr^{84}	112.4	8.62	4.75	3.41^3, 2.93^2	7.13 (H^δ); 6.87 (H^ε)	(9 Hz, −120°, −60°)
Gly^{85}	106.1	7.60	4.35, 4.00			(-,-)
Ile^{86}	118.8	7.60	4.38	1.78	1.52^3, 1.16^2 ($H^{\gamma1}$); 0.86 ($H^{\gamma2}$); 0.89 (H^δ)	(8 Hz, −120°, 180°,-)

[a] Chemical shifts are accurate within ± 0.01 ppm (^1H) and ± 0.1 ppm (^{15}N). ^1H chemical shifts are measured in parts per million relative to the methyl group proton resonance of 3-(trimethylsilyl)-propionic-d_4 acid and ^{15}N chemical shifts relative to the nitrogen resonance of NH_3.

[b] Stereospecific assignment is indicated by the superscripts 3 and 2 as recommended by the IUPAC-IUB commission (IUPAC-IUB, 1970).

[c] Dihedral angle restraints are listed as middle value, ϕ-angles were restrained by ± 40°, χ-angles were restrained by ± 60° except $His^{14}\chi^2$ which was restrained by ± 120°, and prolines χ^1 which were restrained by ± 15°.

A total of 60 structures were calculated. Those structures having no distance restraint violation larger than 0.4 Å, and having no dihedral angle restraint violation larger than 10° were selected for further presentation. A total of 29 structures were within these limits. The structures were aligned with respect to the C, N, and C^{α} of the individual residues and compared to a geometric average structure. The average rmsd for the backbone atoms of the individual residues is shown in Fig. 2A and similarly the rmsd values for the heavy atoms in all the residues are compared in Fig. 2B. A comparison of the local geometry was performed by analyzing the three dihedral angles ϕ, ψ and χ^1 in the sample of structures (Fig. 3). The ϕ- and ψ-angles are all, except one, in the expected regions. One unrestrained ϕ-angle of Glu^{42} has an average positive ϕ-angle of 175° ± 63°. The NMR data for this residue are sparse, there are only a few NOEs, and the $^3J_{H^N H^{\alpha}}$ were measured as being 6 Hz, indicating averaging due to mobility. The NMR data do not suggest that the ϕ-angle is positive, rather that the residue is mobile or close to a mobile environment in the structure. Finally the contributions to the total energy of the 29 structures were measured; these are summarized in Table 2.

The structure determination of rACBP has been obtained with an input of 1096 distance and 124 dihedral angle restraints. In the final set of 29 structures there were 1.2 ± 0.7 distance

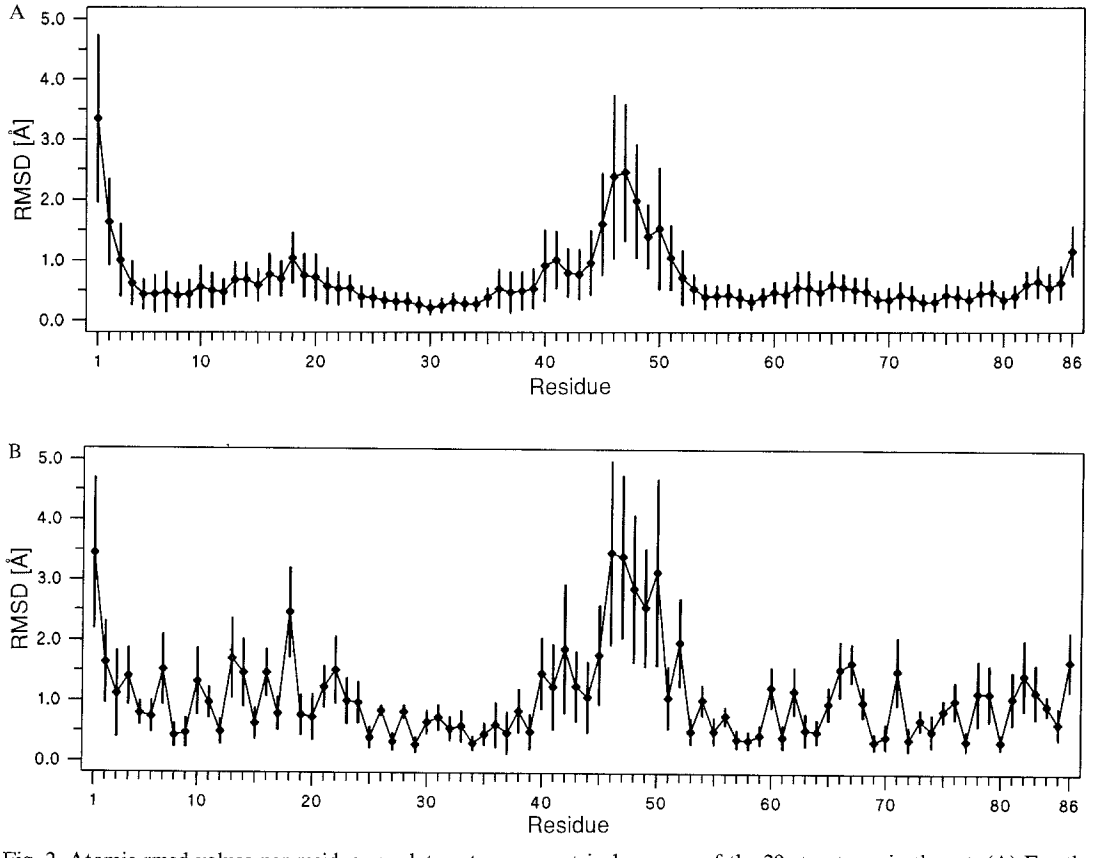

Fig. 2. Atomic rmsd values per residue in relation to a geometrical average of the 29 structures in the set. (A) For the peptide backbone C, N, C^{α} atoms in each residue. (B) For all non-hydrogen atoms in each residue.

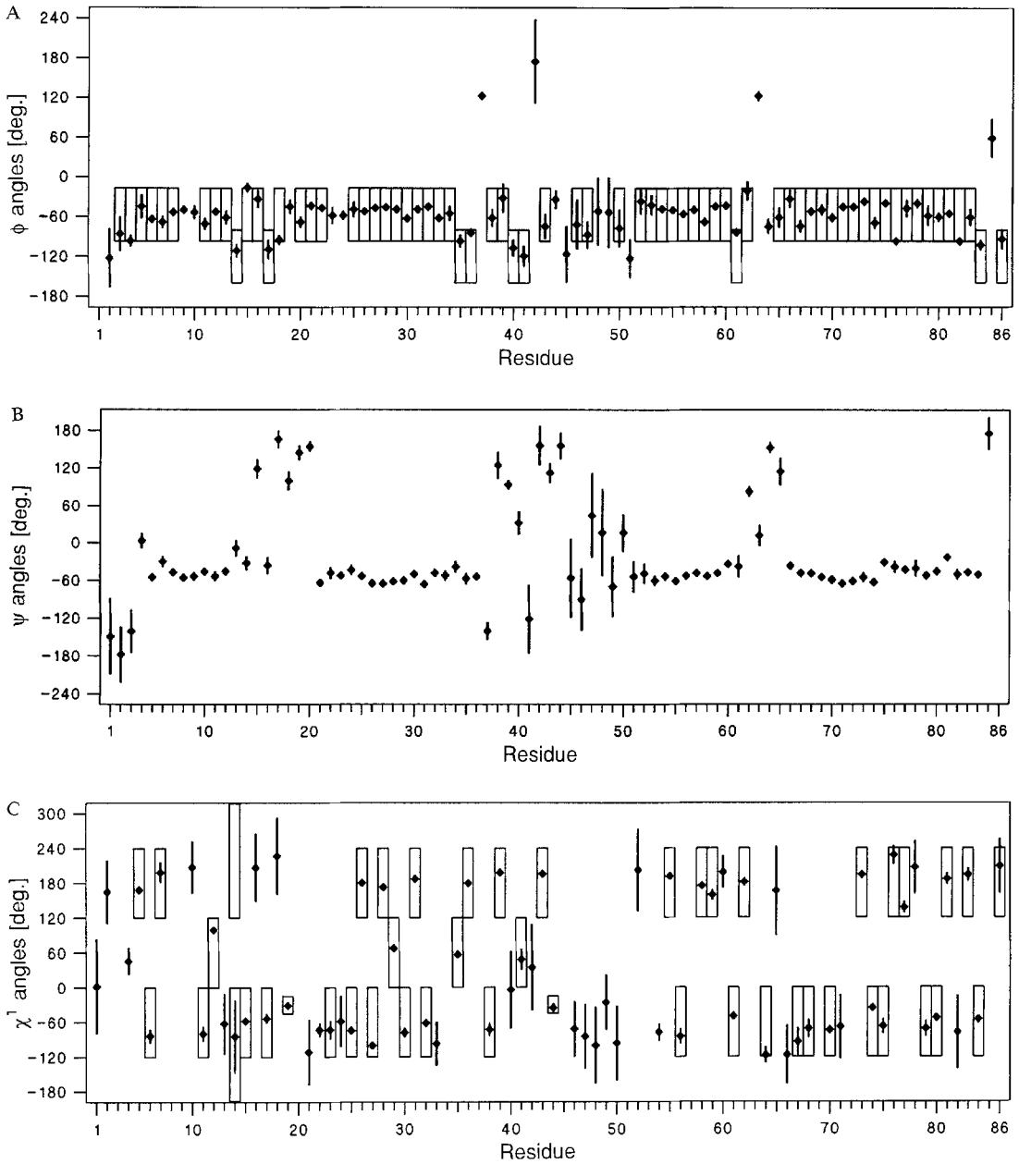

Fig. 3. Comparison of the φ-, ψ-, and χ^1-angles in each residue in the set of 29 structures. (A) φ-angles, (B) ψ-angles and (C) χ^1-angles. The vertical bars show the angular standard deviation. The open boxes indicate the dihedral angle restraint limits used for the angle.

restraints/structure violated by more than 0.3 Å, and 1.2 ± 1.0 dihedral angle restraints/structure violated by more than 5°. The rmsd between the calculated and the target inter-proton distance restraints was 0.044 ± 0.002 Å. The atomic rmsd when compared to the geometric average

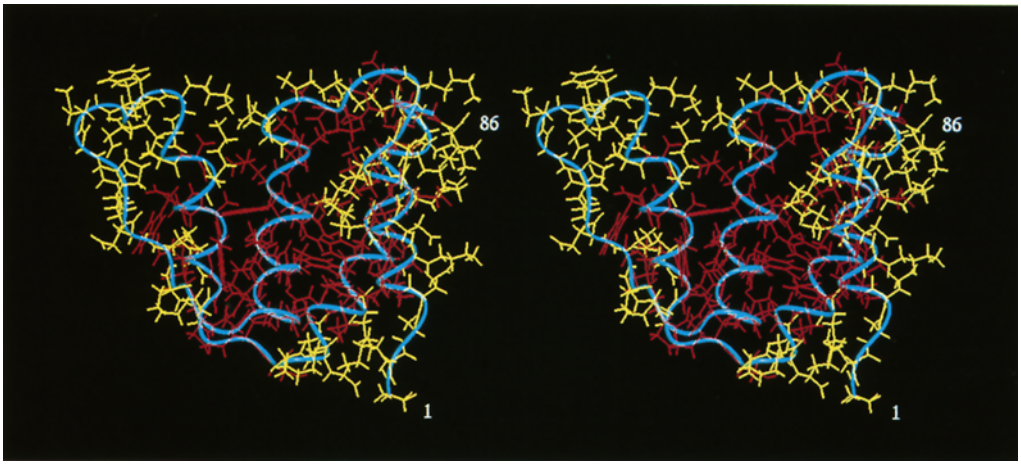

Fig. 4. Stereoview of the lowest-energy structure of rACBP displaying all non-hydrogen atoms and a smooth line through the back-bone atoms. The residues displayed in red have an atomic rmsd of their non-hydrogen atoms smaller than 1.0 Å.

structure was 0.44 ± 0.11 Å for the backbone atoms in the four α-helices and 0.91 ± 0.20 Å for the backbone atoms in all residues (Fig. 4) (0.47 ± 0.09 Å and 1.04 ± 0.18 Å, respectively, in the previous calculations).

The number of structural restraints used in the structure determination is large given the size of the protein. However, the distribution of these on residues, as shown in Fig. 1, is uneven with a considerable number of restraints observed for a relatively small proportion of residues, typically of hydrophobic residues in the core of the protein. This has been important for the definition of the global fold of the protein because it has fixed the orientation of the four α-helices with respect to each other. The rather high quality of the structure, with the exception though of the six residues, Gly^{45}–Lys^{50}, is at the limit of what can be obtained by NMR spectroscopy without using double labeling of the protein with ^{13}C and ^{15}N. The atomic coordinates of the set of 29

TABLE 2
ENERGIES AND STRUCTURAL DEVIATIONS FOR 29 STRUCTURES OF rACBP

	Energy term									
	Total	Bond	Angle	Dihedral	Improper	van der Waals	Electro-static	Hydrogen bond	NOE restraints	Torsion restraints
(Number):		(1405)	(2550)	(628)	(340)				(1096)	(124)
Energy (kcal/mol)	−747 ± 54	56 ± 4	432 ± 16	474 ± 12	9.2 ± 1.1	−299 ± 11	−1385 ± 47	−134 ± 7	96 ± 7	2.7 ± 1.1
Structural deviations		0.0151 Å ± 0.0006 Å	3.46° ± 0.06°		0.42° ± 0.03°					

Energies calculated with the CHARMM (Brooks et al., 1983) potential. Electrostatic potentials were calculated with a distance-dependent dielectric force constant (Gelin, 1976). Force constants for distance restraints and torsion angle restraints are 40 kcal mol^{-1} $Å^{-2}$ and 80 kcal mol^{-1} $radian^{-2}$, respectively. Numbers of bonds, angles, dihedrals, impropers, distance restraints and torsion restraints are given in parentheses.

TABLE 3
HYDROGEN BONDS IDENTIFIED IN THE SET OF 29 STRUCTURES

Donor	Acceptor	$d_{D\text{-}A}$ (Å)	$a_{D\text{-}H\text{-}A}$ (deg)	E (kcal/mol)	Secondary structure
Lys^{7} N	Glu^{4} O	2.75 ± 0.47	115 ± 10	−0.013 ± 0.013	3_{10}
Ala^{8} N	Glu^{4} O	2.69 ± 0.06	155 ± 10	−1.6 ± 0.8	α
Ala^{9} N	Phe^{5} O	2.74 ± 0.03	174 ± 3	−2.8 ± 0.2	α
Glu^{10} N	Asp^{6} O	2.74 ± 0.07	160 ± 9	−1.6 ± 0.8	α
Glu^{11} N	Lys^{7} O	2.82 ± 0.13	161 ± 7	−2.3 ± 0.6	α
Val^{12} N	Ala^{8} O	2.75 ± 0.04	169 ± 6	−2.9 ± 0.4	α
Lys^{13} N	Ala^{9} O	3.01 ± 0.22	162 ± 7	−2.1 ± 0.7	α
His^{14} N	Glu^{11} O	2.87 ± 0.13	150 ± 8	−0.5 ± 0.4	3_{10}
Leu^{15} N	Glu^{11} O	3.10 ± 0.26	150 ± 12	−1.7 ± 0.7	α
Met^{24} N	Ala^{20} O	2.93 ± 0.09	171 ± 4	−3.1 ± 0.3	α
Leu^{25} N	Asp^{21} O	2.67 ± 0.05	163 ± 11	−2.3 ± 0.7	α
Phe^{26} N	Glu^{22} O	3.03 ± 0.37	163 ± 9	−2.1 ± 0.9	α
Ile^{27} N	Glu^{23} O	3.34 ± 0.43	167 ± 5	−2.2 ± 0.7	α
Tyr^{28} N	Met^{24} O	2.73 ± 0.04	162 ± 6	−2.3 ± 0.6	α
Ser^{29} N	Leu^{25} O	2.70 ± 0.03	163 ± 4	−2.2 ± 0.3	α
His^{30} N	Phe^{26} O	2.96 ± 0.09	165 ± 4	−2.7 ± 0.3	α
Tyr^{31} N	Ile^{27} O	2.63 ± 0.04	159 ± 9	−2.0 ± 0.7	α
Lys^{32} N	Tyr^{28} O	2.73 ± 0.03	158 ± 6	−2.1 ± 0.5	α
Gln^{33} N	Ser^{29} O	2.93 ± 0.09	160 ± 6	−1.9 ± 0.5	α
Ala^{34} N	His^{30} O	2.71 ± 0.05	168 ± 7	−2.7 ± 0.5	α
Thr^{35} N	Tyr^{31} O	2.75 ± 0.04	157 ± 10	−2.3 ± 0.7	α
Val^{36} N	Lys^{32} O	2.76 ± 0.08	129 ± 5	−0.27 ± 0.14	α
Gly^{37} N	Lys^{32} O	2.78 ± 0.05	135 ± 25	−2.4 ± 0.3	π
Trp^{55} N	Gly^{51} O	3.52 ± 0.53	160 ± 10	−1.3 ± 1.1	α
Asp^{56} N	Lys^{52} O	2.74 ± 0.08	165 ± 9	−2.3 ± 0.8	α
Ala^{57} N	Ala^{53} O	2.72 ± 0.03	171 ± 5	−3.0 ± 0.3	α
Trp^{58} N	Lys^{54} O	2.91 ± 0.11	168 ± 6	−2.7 ± 0.4	α
Asn^{59} N	Trp^{55} O	2.69 ± 0.02	173 ± 3	−2.85 ± 0.14	α
Glu^{60} N	Asp^{56} O	2.77 ± 0.05	167 ± 6	−2.8 ± 0.4	α
Leu^{61} N	Trp^{58} O	2.74 ± 0.06	144 ± 8	−0.5 ± 0.3	3_{10}
Lys^{62} N	Asn^{58} O	3.34 ± 0.28	152 ± 20	−1.5 ± 1.1	α
Thr^{64} N	Leu^{61} O	2.68 ± 0.09	116 ± 17	−0.2 ± 0.4	T II^{β}
Asp^{68} N	Ser^{65} O	2.67 ± 0.11	116 ± 17	−0.3 ± 0.7	3_{10}
Ala^{69} N	Ser^{65} O	3.14 ± 0.40	169 ± 7	−2.5 ± 1.0	α
Met^{70} N	Lys^{66} O	3.25 ± 0.41	159 ± 8	−1.3 ± 0.8	α
Lys^{71} N	Glu^{67} O	2.79 ± 0.12	161 ± 10	−2.5 ± 0.8	α
Ala^{72} N	Asp^{68} O	2.76 ± 0.06	151 ± 7	−1.6 ± 0.6	α
Tyr^{73} N	Ala^{69} O	2.84 ± 0.11	160 ± 5	−1.8 ± 0.5	α
Ile^{74} N	Met^{70} O	2.98 ± 0.13	165 ± 7	−2.7 ± 0.6	α
Asp^{75} N	Lys^{71} O	2.85 ± 0.13	151 ± 6	−1.9 ± 0.5	α
Lys^{76} N	Ala^{72} O	2.67 ± 0.04	138 ± 5	−0.9 ± 0.3	α
Val^{77} N	Tyr^{73} O	2.83 ± 0.14	149 ± 14	−1.2 ± 0.8	α

TABLE 3 (continued)

Donor	Acceptor	$d_{D\text{-}A}$ (Å)	$a_{D\text{-}H\text{-}A}$ (deg)	E (kcal/mol)	Secondary structure
Glu^{78} N	Ile^{74} O	3.19 ± 0.22	166 ± 3	−2.8 ± 0.3	α
Glu^{79} N	Lys^{76} O	2.85 ± 0.24	144 ± 19	−0.9 ± 0.7	3_{10}
Leu^{80} N	Lys^{76} O	3.89 ± 0.57	149 ± 5	−0.7 ± 0.6	α
Lys^{81} N	Val^{77} O	2.78 ± 0.06	173 ± 4	−2.7 ± 0.3	α
Lys^{82} N	Glu^{78} O	3.10 ± 0.43	124 ± 7	−0.2 ± 0.2	α
Lys^{83} N	Glu^{79} O	2.73 ± 0.07	166 ± 7	−2.5 ± 0.5	α
Tyr^{84} N	Leu^{80} O	2.70 ± 0.06	142 ± 5	−0.7 ± 0.3	α
Gly^{85} N	Leu^{80} O	2.89 ± 0.14	164 ± 6	−1.9 ± 0.4	π
Lys^{16} N	Gly^{85} O	2.77 ± 0.13	135 ± 25	−1.0 ± 0.9	
Gln^{2} $N^{\varepsilon 2}$	Asp^{6} O^{δ}	3.02 ± 0.60	156 ± 14	−1.8 ± 1.0	
His^{14} $N^{\delta 1}$	Glu^{10} O	3.14 ± 0.80	139 ± 16	−0.7 ± 0.8	
Thr^{17} O^{γ}	Tyr^{84} O	3.39 ± 0.62	133 ± 53	−0.9 ± 0.8	
Ser^{29} O^{γ}	Leu^{25} O	2.70 ± 0.10	141 ± 5	−0.5 ± 0.2	
Ala^{20} N	Tyr^{84} O^{η}	3.14 ± 0.40	149 ± 15	−1.8 ± 1.0	
Tyr^{31} O^{ζ}	Asp^{6} O^{δ}	2.88 ± 0.04	175 ± 3	−3.7 ± 0.6	
Gln^{33} $N^{\varepsilon 2}$	Asp^{38} O^{δ}	3.10 ± 0.48	158 ± 8	−1.3 ± 0.6	
Val^{36} N	Thr^{35} O^{γ}	2.58 ± 0.05	125 ± 9	−0.4 ± 0.3	
Glu^{42} N	Thr^{41} O^{γ}	2.94 ± 0.61	112 ± 30	−0.5 ± 0.7	
Trp^{58} N^{ε}	Gly^{37} O	2.78 ± 0.21	154 ± 9	−1.9 ± 0.6	
Thr^{64} N	Gln^{33} $O^{\varepsilon 1}$	3.16 ± 0.80	144 ± 22	−1.3 ± 1.2	
Ser^{65} O^{γ}	Asp^{68} O^{δ}	3.38 ± 0.69	118 ± 40	−0.5 ± 0.7	
Tyr^{84} O^{ζ}	Glu^{23} O^{δ}	2.90 ± 0.53	147 ± 34	−0.9 ± 1.2	

The average donor-acceptor distance and donor-hydrogen-acceptor angle are listed with their standard deviations. For each hydrogen bond the average energy of the 29 structures is listed. These were obtained using the standard X-PLOR hydrogen-bond energy term. The notation for the secondary structure responsible for the hydrogen bond is: α, α-helix; 3_{10}, 3_{10}-helix; π, π-helix; T II^{β}; type II^{β} turn.

refined structures as well as the distance and dihedral angle constraints used for the calculations have been deposited in the Protein Data Bank under accession code 2ABD. The first structures (Andersen and Poulsen, 1992) are deposited in the Protein Data Bank and have accession code 1ABD.

The four α-helices already described in the first rACBP structure paper (Andersen and Poulsen, 1992) are the scaffold of the structure (Fig. 4). These helices have been evaluated carefully by examinations of the donor acceptor distances and the donor-hydrogen-acceptor angle in the hydrogen bonds as well as the energy associated with these hydrogen bonds, as calculated by X-PLOR. The results of this analysis are listed in Table 3. All four helices in the structure were seen to have minor irregularities and the definition of the beginning and the end of the helices was not always clear. The start of helix A1 is clearly defined at Glu^{4}, where the helix starts with a bifurcated hydrogen bond from Glu^{4} O to the N-H atoms of Lys^{7} and Ala^{8}. However, at the end the regular hydrogen bond His^{14} N-H to Glu^{10} O is not formed. Instead there is a 3_{10}-helix hydrogen bond from His^{14} N-H to Glu^{11} O, creating an irregularity in the α-helix and leaving

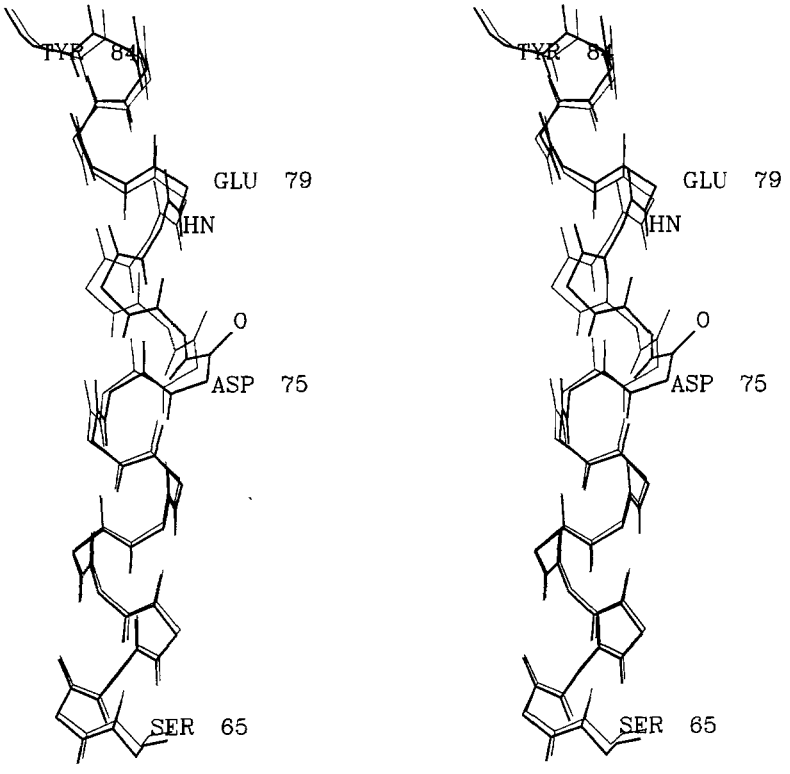

Fig. 5. Stereoview of helix A4 in the 'kinked' rACBP structures (thick line) and the rACBP structures with ideal α-helices (thin line); the C, N, O, C^{α}, and H^{N} atoms of the average structures of the two sets are shown.

Glu^{10} O as an unused hydrogen bond acceptor and Glu^{11} O as acceptor in a bifurcated hydrogen bond arrangement with the two donors His^{14} N-H and Leu^{15} N-H. These findings are partly a consequence of the experimental observation that His^{14} has a large 9 Hz $^{3}J_{H^{N}H^{\alpha}}$ coupling constant. Also helix A2 is well defined, however, at the end of A2 there is a weak hydrogen bond, due to unfavorable angle geometry from Val^{36} N-H to Lys^{32} O. Here it is noted though that the end of the helix seems to be associated with a π-helix type hydrogen bond from Gly^{37} N-H to Lys^{32} O that is part of a bifurcated hydrogen bond arrangement involving also the regular α-helix hydrogen bond from Val^{36} N-H to Lys^{32} O. The last two residues of A2 Ala^{35} and Val^{36} both have large $^{3}J_{H^{N}H^{\alpha}}$ coupling constants. The first well-defined residue of helix A3 is Lys^{52}, however, a significant number of the structures have the helix starting at Gly^{51}. Helix A3 ends at Lys^{62}, but as was the case in helix A1, the last but one hydrogen bond between Leu^{61} N-H and Ala^{57} O is missing, and there is instead a 3_{10}-helix hydrogen bond from Leu^{61} N-H to Trp^{58} O creating an irregularity also at the end of this helix. Also in A3 the experimental data suggest some irregularities in the C-terminal. Here it is the large $^{3}J_{H^{N}H^{\alpha}}$ coupling constant observed for Leu^{61} as was seen for His^{14} in helix A1. The N-terminal of helix A4 is unambiguously well defined from Glu^{67}, however, the two preceding residues, Ser^{65} and Lys^{66}, form a regular α-helix hydrogen bond in a majority of the structures, and as in A1 the first residue of the helix is part of the bifurcated hydrogen bond from

the N-H atoms of Asp[68] and Ala[69] to the O of Ser[65]. Hence the calculated structures suggest that the helix may start at Ser[65]. Helix A4 with its 20 residues is the longest of the four α-helices in rACBP. The helix is bending inward, slightly, with the result that the hydrogen bonds on the outside are slightly weaker, or even missing. Glu[79] N-H to Asp[75] O is missing in all structures, and instead a 3_{10}-helix-type hydrogen bond is formed between Glu[79] N-H and Lys[76] O, leaving Asp[75] O in the helix without a hydrogen bond partner. The hydrogen bonds Leu[80] N-H to Lys[76] O and Tyr[84] N-H to Leu[80] O are weak, and the hydrogen bond Lys[82] N-H to Glu[78] is very weak. There is no evidence in the experimental data for this break in the helix; the coupling constants are all as expected for an α-helix, the sequential NOEs in this part of the sequence are typical for an α-helix, and the amide hydrogen exchange is slow for all residues in the sequence around the break. The C-terminal part of helix A4 ends at Tyr[84] with an extension to Gly[85] forming a π-helix-type hydrogen bond in an arrangement identical to helix A2. Also here the last hydrogen bond from Tyr[84] N-H to Leu[80] O is quite weak. Experimentally this helix-termination was associated with a large $^3J_{H^NH^\alpha}$ for the C-terminal residue Tyr[84]. Again very similar to the observation that was made for the C-terminal residue of A2.

In the first structure determination three turns were identified, two of them (T1 and T2) are now included in helix A3, and the third turn now called T1 (Leu[61]–Thr[64]) is a type II_β turn, as in the previous structure. The slow hydrogen exchange of the amide hydrogen observed for Thr[64] (Andersen et al., 1991) is in agreement with this finding, although the hydrogen-bond energy is low. Apart from the hydrogen bonds in the helical peptide backbone, only one other backbone hydrogen bond was present in the calculated structures, Lys[16] N-H to Gly[85] O, forming a connection between the residue following the C-terminal of helix A1 and the residue following the C-terminal of helix A4. In addition to these there are a number backbone to side chain and side chain to side chain hydrogen bonds that make contact between the helices; these are summarized in Table 3.

We have examined the four helices for structural ideality. Essentially, all the requirements for α-helix configurations were observed in the NMR data both with respect to the sequential NOEs expected and with respect to the $^3J_{H^NH^\alpha}$ coupling constants. In a few cases, though, we observed non-ideal parameters for α-helices. The coupling constants for His[14], in helix A1, Thr[35] and Val[36] in helix A2, Leu[61] in helix A3, and Tyr[84] in helix A4 were all measured to be larger than 7 Hz, suggesting the φ-angle in these residues to be outside the typical range for residues in α-helices. Accordingly the restraint limits for the φ-angle in the corresponding residues were set to cover the range from −80° to −160° in the structure calculations. In the resulting structures three of the five φ-angles fall clearly outside the range of typical helices (Fig. 3A). Nevertheless, the backbones of the four residues remain an integral part of their respective helices with negative ψ-angles (Fig. 3B), but the hydrogen-bond pattern is obscured in some cases. Two of the five φ-angles (Val[36] and Leu[61]) are close to violation in the structures (Fig. 3A) but there was no reason to exclude these constraints from the calculations. We note that these experimental deviations from normality in the helices all occurred in the very C-terminal part of the respective helices.

The structures have been reexamined with a particular emphasis to evaluate the 'kink' in helix A4. A new set of structures was calculated where the hydrogen bonds in the helices were included as distance restraints as if they were regular α-helices, and helix A4 was treated as a complete α-helix including the missing hydrogen bond. The resulting structures were of similar quality as those presented here in terms of energies and restraint violations. Analysis of sequential distances

in the region around the missing hydrogen bond in the 'kinked' structures showed that the distance between Asp^{75} H^{α} and Glu^{79} H^{N} should have been determined as a weak $d_{\alpha N}(i, i+4)$ sequential NOE in the NMR data to give a distance restraint that ensures the formation of the missing hydrogen bond (Fig. 5). In the 'kinked' structures this distance is 6.2 ± 0.5 Å. The NOESY spectrum recorded with a mixing time of 200 ms has indeed a weak peak for these two protons; however, it has been disregarded due to spin diffusion. Therefore, there is no unambiguous experimental evidence to justify the existence of a hydrogen bond that could straighten the 'kink'. However, no such 'kink' was observed in *holo*-rACBP (Kragelund et al., 1993), without making the structures significantly different in this region.

The present refinement of the ACBP structure was initiated as a necessary step for the determination of the 3D structure of the complex between the protein and the ligand. The new structure is the platform for a comparison of the protein in the *apo*- and the *holo*-form, and as such it has been important to determine the 3D structure with precisely the same technology as applied in the study of the complex (Kragelund et al., 1993). The experimentally defined input for the structure calculation has been increased by almost 30% as compared to the first set. In itself this has not led to a significant redefinition of the structure. In outline the refined structure is very similar to the first 3D structure of rACBP; however, a closer examination reveals that the structure has indeed a better definition than before, and especially the extensive stereospecific assignments have resulted in a better definition of the side chains.

ACKNOWLEDGEMENTS

The ACBP, rACBP and the ^{15}N-enriched rACBP were prepared by Dr. Jens Knudsen, Birthe B. Kragelund and colleagues. We are grateful for their collaboration, help and many useful discussions. We wish to thank Jens Chr. Madsen, Mogens Kjær and Svend Ludvigsen for useful discussions and help with the data recording and the computational aspects of the work. Flemming M. Poulsen is member of the Danish Protein Engineering Centre (PERC).

REFERENCES

Andersen, K.V. and Poulsen, F.M. (1992) *J. Mol. Biol.*, **226**, 1131–1141.

Andersen, K.V., Ludvigsen, S., Mandrup, S., Knudsen, J. and Poulsen, F.M. (1991) *Biochemistry*, **30**, 10654–10663.

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

Brünger, A.T. (1992) *The X-PLOR Software*, Yale University and Harvard University, U.S.A.

Clore, G.M., Gronenborn, A.M., Brünger, A.T. and Karplus, M. (1985) *J. Mol. Biol.*, **186**, 435–455.

Clore, G.M., Brünger, A.T., Karplus, M. and Gronenborn, A.M. (1986) *J. Mol Biol.*, **191**, 523–551.

Gelin, B.R. (1976) Ph.D. Thesis, Harvard University.

Günthert, P., Braun, W. and Wüthrich, K. (1991) *J Mol. Biol.*, **217**, 517–530.

Kjær, M., Andersen, K.V., Shen, H., Ludvigsen, S., Windekilde, D., Sørensen, B. and Poulsen, F.M. (1991) In *NATO ASI series A, Vol. 225* (Eds., Hoch, J.C., Poulsen, F.M. and Redfield, C.) Plenum, New York, pp. 291–302.

Kragelund, B.B., Andersen, K.V., Madsen, J.C., Knudsen, J. and Poulsen, F.M. (1993) *J. Mol. Biol.*, **230**, 1260–1277.

Ludvigsen, S. and Poulsen, F.M. (1992) *Biochemistry*, **31**, 8783–8789.

Ludvigsen, S., Andersen, K.V. and Poulsen, F.M. (1991) *J. Mol. Biol.*, **217**, 731–736.

Marion, D., Kay, L.E., Sparks, S.W., Torchia, D.A. and Bax, A. (1989a) *J. Am. Chem. Soc.*, **111**, 1515–1517.

Marion, D., Driscoll, P.C., Kay, L.E., Wingfield, P T., Bax, A., Gronenborn, A.M. and Clore, G.M. (1989b) *Biochemistry*, **28**, 6150–6156.

Marion. D., Ikura, M., Tschudin, R. and Bax, A (1989c) *J. Magn. Reson* , **85**, 393–399.
Montelione, G.T., Winkler, M.E., Rauenbühler, P. and Wagner, G. (1989) *J Magn. Reson.*, **82**, 198
Nilges, M., Clore, G.M. and Gronenborn, A.M. (1988) *FEBS Lett.*, **229**, 317–324.
Williamson, M.P., Havel, T.F. and Wüthrich, K. (1985) *J Mol. Biol.*, **182**, 295–315.
Zuiderweg, E.R.P. and Fesik, S.W. (1989) *Biochemistry*, **28**, 2387–2391.