Modular Design of Synthetic Protein Mimics. Characterization of the Helical Conformation of a 13-Residue Peptide in Crystals†

Isabella L. Karle,*,‡ Judith L. Flippen-Anderson,‡ K. Uma,§ and P. Balaram*,§

Laboratory for the Structure of Matter, Naval Research Laboratory, Washington, D.C. 20375, and Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India

Received February 14, 1989; Revised Manuscript Received April 24, 1989

ABSTRACT: The incorporation of α-aminoisobutyryl (Aib) residues into peptide sequences facilitates helical folding. Aib-containing sequences have been chosen for the design of rigid helical segments in a modular approach to the construction of a synthetic protein mimic. The helical conformation of the synthetic peptide Boc-Aib-(Val-Ala-Leu-Aib)$_3$-OMe in crystals is established by X-ray diffraction. The 13-residue apolar peptide adopts a helical form in the crystal with seven α-type hydrogen bonds in the middle and 3_{10}-type hydrogen bonds at either end. The helices stack in columns, zigzag rather than linear, by means of direct NH···OC head to tail hydrogen bonds. Leucyl side chains are extended on one side of the helix and valyl side chains on the other side. Water molecules form hydrogen bonds with several backbone carbonyl oxygens that also participate in α-helix hydrogen bonds. There is no apparent distortion of the helix caused by hydration. The space group is $P2_12_12_1$, with a = 9.964 (3) Å, b = 20.117 (3) Å, c = 39.311 (6) Å, Z = 4, and d_x = 1.127 g/cm^3 for $C_{64}H_{106}N_{13}O_{16}\cdot 1.33H_2O$. The final agreement factor R was 0.089 for 3667 data observed $>3\sigma(F)$ with a resolution of 0.9 Å.

The de novo design of synthetic polypeptides mimicking the folding patterns of proteins is a subject of intense current interest (Eisenberg et al., 1986; Ho & DeGrado, 1987; Regan & DeGrado, 1988; Mutter et al., 1988; Mutter, 1988; Richardson & Richardson, 1987; Osterman & Kaiser, 1985). A modular approach to the construction of a "synthetic miniprotein" is being undertaken by developing stereochemically rigid synthetic peptide models for secondary structural elements like $3_{10}/\alpha$-helices, β-sheets and β-turns (Balaram, 1984). Conformational constraints have been introduced into peptide sequences by the incorporation of α,α-dialkylated amino acids (Marshall et al., 1976), of which α-aminoisobutyric acid (Aib) is the prototype (Prasad & Balaram, 1984), and by the use of disulfide cross-links (Ravi & Balaram, 1984; Kishore et al., 1985, 1987; Karle et al., 1988a, 1989b). Aib residues have been shown to strongly promote helical folding in acyclic peptides (Nagaraj & Balaram, 1981; Prasad & Balaram, 1984; Toniolo et al., 1983). The precise helix type, 3_{10}, α, or mixed $3_{10}/\alpha$, depends on the nature of the sequence, Aib content, and environmental effects (Bosch et al., 1985a,b; Karle et al., 1986, 1988b,c, 1989c). In this paper, we describe the helical conformation in crystals of the 13-residue peptide Boc-Aib-(Val-Ala-Leu-Aib)$_3$-OMe, which illustrates the utility of Aib residues in the construction of relatively long helical segments. The crystal structure also provides further information on the hydration of peptide helices, a subject of interest following the serendipitous observation of a hydrophilic aqueous channel formed by association of hydrophobic helices in the crystal structure of Boc-Aib-(Ala-Leu-Aib)$_3$-OMe (Karle et al., 1988d). Such aggregation patterns are of relevance in considering models for the structure of transmembrane channels formed by hydrophobic peptides like alamethicin and related fungal antibiotics (Fox & Richards, 1982; Mathew & Balaram, 1983a,b; Menestrina et al., 1986).

Experimental Procedures

Boc-Aib-(Val-Ala-Leu-Aib)$_3$-OMe was synthesized by conventional solution-phase procedures, by a fragment condensation approach, using dicyclohexylcarbodiimide and 1-hydroxybenzotriazole in dichloromethane or dimethylformamide (Balaram et al., 1986). The final peptide was purified by high-performance liquid chromatography on a reverse-phase Lichrosorb RP-18 column (4 mm × 250 mm, particle size 10 μm) with a methanol–water gradient (80–90%, 10 min; 90–95%, 10 min; flow rate 0.8 mL min^{-1}). Repetitive injections of ~1 mg per run were carried out. Crystals were grown by evaporation from a CH_3OH solution in the form of plates. A dry crystal of size 0.9 × 0.4 × 0.15 mm was used to collect X-ray diffraction data on an automated four-circle diffractometer using Cu Kα radiation and a graphite monochromator (λ = 1.54178 Å). The θ–2θ scan technique was used with a scan of 2.0° + $2\theta(\alpha_1)$ – $2\theta(\alpha_2)$, a variable scan rate between 6 and 15°/min, depending upon the diffracted intensity, and $2\theta_{max}$ = 112°, for a total of 6135 unique reflections and 3667 reflections observed $>3\sigma(F)$ used for refinement. Three reflections, 0,0,12, 060, and 201, monitored after every 100 measurements, remained constant within 5% during the data collection. Lorentz and polarization corrections were applied to the data. The cell parameters are a = 9.964 (3) Å, b = 20.117 (3) Å, and c = 39.311 (6) Å for space group $P2_12_12_1$. The calculated density for Z = 4, V = 7879.7 Å^3, and mol wt 1313.6 + 24.0 for $C_{64}H_{106}N_{13}O_{16}\cdot 1.33H_2O$ is 1.127 g cm^{-3}. The structure was solved by a vector search procedure in the PATSEE computer program (Egert & Sheldrick, 1985) contained in the SHELX84 package of programs. (MicroVAX version of SHELXTL system of programs, Nicolet Analytical Instruments, Madison, WI.) The model used consisted of a 28-atom fragment of the Boc-Aib-(Val-Ala-Leu-Aib)$_2$-OMe structure (Karle et al., 1988c). After rotation and translation to the correct position, the remainder of the atoms in the molecule were found with the partial structure procedure (Karle, 1968).

Full-matrix, anisotropic least-squares refinement was performed on the C, N, and O atoms before hydrogen atoms were

† Supported in part by National Institutes of Health Grant GM30902 and in part by a grant from the Department of Science and Technology, India. K.U. is the recipient of a fellowship from the Council of Scientific and Industrial Research, India.

* Address correspondence to this author.
‡ Naval Research Laboratory.
§ Indian Institute of Science.

0006-2960/89/0428-6696$01.50/0

Table I: Atomic Coordinates ($\times 10^4$) and Equivalent Isotropic Displacement Parameters ($Å^2 \times 10^3$)

	x	y	z	U(eq)[a]		x	y	z	U(eq)[a]
O	−3153 (11)	3820 (5)	−1458 (2)	107 (5)	C′(7)	1955 (14)	5502 (6)	1298 (3)	63 (5)
C(1)	−4373 (15)	4146 (8)	−1349 (4)	93 (7)	O(7)	1967 (9)	5614 (4)	1608 (2)	72 (3)
C(2)	−5343 (14)	4000 (8)	−1634 (4)	187 (10)	$C^{\beta}(7)$	−219 (13)	5149 (7)	1046 (3)	90 (6)
C(3)	−4913 (14)	3827 (7)	−1038 (4)	119 (8)	N(8)	2943 (12)	5193 (5)	1135 (2)	63 (4)
C(4)	−4168 (15)	4885 (7)	−1310 (3)	111 (7)	$C^{\alpha}(8)$	4133 (14)	5001 (6)	1314 (3)	66 (5)
C′(0)	−2062 (17)	3836 (7)	−1248 (4)	83 (7)	C′(8)	4910 (13)	5606 (6)	1452 (3)	61 (5)
O(0)	−1915 (9)	4136 (4)	−977 (2)	78 (3)	O(8)	5336 (8)	5620 (3)	1745 (2)	62 (3)
N(1)	−1018 (13)	3464 (4)	−1397 (2)	70 (4)	$C^{\beta}(8)$	5016 (13)	4591 (6)	1083 (3)	69 (5)
$C^{\alpha}(1)$	273 (15)	3327 (6)	−1233 (3)	59 (5)	$C^{\gamma}(8)$	6327 (16)	4340 (6)	1232 (4)	88 (6)
C′(1)	852 (13)	3979 (5)	−1076 (3)	58 (5)	$C^{\delta_1}(8)$	7071 (15)	3958 (6)	961 (3)	133 (8)
O(1)	1404 (9)	3996 (4)	−803 (2)	71 (3)	$C^{\delta_2}(8)$	6135 (15)	3935 (6)	1533 (3)	102 (6)
$C^{\beta_1}(1)$	1140 (12)	3019 (5)	−1484 (2)	65 (5)	N(9)	5104 (10)	6123 (5)	1230 (2)	63 (4)
$C^{\beta_2}(1)$	−27 (16)	2850 (5)	−940 (3)	104 (6)	$C^{\alpha}(9)$	5810 (15)	6738 (6)	1322 (3)	64 (5)
N(2)	803 (9)	4498 (4)	−1292 (2)	51 (3)	C′(9)	5118 (15)	7039 (6)	1642 (3)	61 (6)
$C^{\alpha}(2)$	1477 (15)	5128 (6)	−1208 (3)	79 (6)	O(9)	5760 (9)	7253 (4)	1875 (2)	76 (3)
C′(2)	948 (18)	5389 (7)	−866 (4)	77 (7)	$C^{\beta_1}(9)$	5624 (14)	7206 (5)	1026 (2)	96 (6)
O(2)	1590 (11)	5738 (5)	−679 (2)	121 (5)	$C^{\beta_2}(9)$	7272 (13)	6606 (6)	1396 (3)	78 (6)
$C^{\beta}(2)$	1452 (18)	5588 (6)	−1501 (3)	93 (7)	N(10)	3784 (11)	7072 (4)	1639 (2)	60 (4)
$C^{\gamma_1}(2)$	46 (16)	5881 (6)	−1567 (3)	109 (7)	$C^{\alpha}(10)$	3049 (14)	7384 (6)	1917 (3)	66 (5)
$C^{\gamma_2}(2)$	2453 (17)	6118 (7)	−1448 (3)	134 (9)	C′(10)	3139 (14)	6967 (6)	2229 (3)	66 (5)
N(3)	−295 (16)	5254 (5)	−806 (3)	88 (6)	O(10)	3272 (10)	7221 (4)	2513 (2)	84 (4)
$C^{\alpha}(3)$	−804 (19)	5464 (8)	−490 (4)	118 (9)	$C^{\beta}(10)$	1631 (16)	7551 (9)	1833 (5)	142 (9)
C′(3)	−417 (19)	5121 (8)	−169 (4)	103 (8)	$C^{\gamma_1}(10)$	1498 (18)	8077 (9)	1555 (4)	160 (9)
O(3)	−296 (11)	5453 (4)	100 (2)	101 (4)	$C^{\gamma_2}(10)$[b]	494 (24)	7653 (13)	2084 (4)	76 (11)
$C^{\beta}(3)$	−2150 (20)	5739 (13)	−461 (5)	374 (25)	$C^{\gamma_3}(10)$[b]	425 (23)	7094 (12)	1858 (9)	105 (14)
N(4)	80 (12)	4537 (5)	−205 (2)	86 (5)	N(11)	3147 (11)	6294 (4)	2208 (2)	61 (4)
$C^{\alpha}(4)$	760 (20)	4206 (7)	81 (3)	100 (8)	$C^{\alpha}(11)$	3261 (14)	5866 (5)	2510 (3)	64 (5)
C′(4)	2073 (21)	4549 (7)	166 (3)	89 (7)	C′(11)	4601 (14)	5855 (6)	2679 (3)	58 (5)
O(4)	2419 (11)	4609 (3)	459 (2)	95 (4)	O(11)	4683 (10)	5661 (4)	2972 (2)	88 (4)
$C^{\beta}(4)$	897 (19)	3465 (7)	−4 (3)	122 (8)	$C^{\beta}(11)$	2853 (14)	5172 (5)	2416 (3)	88 (6)
$C^{\gamma}(4)$	1456 (34)	3025 (9)	245 (4)	211 (16)	N(12)	5636 (11)	6055 (4)	2494 (2)	53 (4)
$C^{\delta_1}(4)$	611 (25)	3044 (8)	577 (4)	250 (17)	$C^{\alpha}(12)$	6985 (14)	6085 (5)	2618 (3)	62 (5)
$C^{\delta_2}(4)$	1494 (26)	2336 (8)	128 (4)	226 (15)	C′(12)	7435 (15)	6814 (7)	2694 (3)	57 (5)
N(5)	2806 (16)	4815 (6)	−88 (3)	97 (6)	O(12)	8610 (10)	6912 (4)	2777 (2)	84 (4)
$C^{\alpha}(5)$	4001 (18)	5206 (9)	−36 (4)	97 (8)	$C^{\beta}(12)$	8015 (14)	5722 (5)	2398 (3)	77 (5)
C′(5)	3757 (20)	5747 (7)	213 (4)	86 (7)	$C^{\gamma}(12)$	7740 (18)	5014 (7)	2350 (3)	96 (7)
O(5)	4504 (11)	5896 (5)	438 (2)	105 (4)	$C^{\delta_1}(12)$	7794 (21)	4626 (5)	2662 (3)	170 (11)
$C^{\beta_1}(5)$	4448 (17)	5497 (9)	−377 (3)	153 (9)	$C^{\delta_2}(12)$	8729 (17)	4753 (7)	2104 (3)	147 (9)
$C^{\beta_2}(5)$	5152 (16)	4728 (8)	100 (3)	128 (8)	N(13)	6505 (11)	7284 (4)	2659 (2)	62 (4)
N(6)	2663 (14)	6108 (6)	144 (2)	78 (5)	$C^{\alpha}(13)$	6854 (16)	7990 (5)	2702 (3)	65 (5)
$C^{\alpha}(6)$	2296 (16)	6662 (7)	341 (3)	80 (6)	C′(13)	7359 (15)	8087 (6)	3068 (3)	67 (6)
C′(6)	1890 (13)	6449 (5)	725 (3)	66 (5)	O(13)	6961 (10)	7772 (4)	3301 (2)	90 (4)
O(6)	2345 (10)	6775 (3)	967 (2)	90 (4)	$C^{\beta_1}(13)$	5561 (15)	8358 (6)	2668 (3)	113 (7)
$C^{\beta}(6)$	1118 (17)	7047 (6)	192 (3)	88 (6)	$C^{\beta_2}(13)$	7842 (14)	8209 (6)	2425 (3)	94 (6)
$C^{\gamma_1}(6)$	584 (17)	7587 (7)	420 (3)	149 (9)	O(14)	8217 (11)	8570 (4)	3083 (2)	105 (4)
$C^{\gamma_2}(6)$	1463 (19)	7375 (7)	−138 (3)	155 (9)	C(14)	8714 (16)	8734 (6)	3411 (3)	116 (7)
N(7)	1108 (9)	5922 (3)	747 (2)	61 (4)	W(1)	8905 (13)	2679 (5)	2003 (2)	155 (6)
$C^{\alpha}(7)$	750 (13)	5718 (5)	1086 (3)	56 (5)	W(2)[c]	7585 (37)	6011 (16)	528 (7)	165 (19)

[a] Equivalent isotropic U defined as one-third of the trace of the orthogonalized U_{ij} tensor. [b] Occupancy 0.50. [c] Occupancy 0.33.

added in idealized positions and allowed to ride with the C or N atom to which each was bonded. The thermal factor for the hydrogen atoms was fixed at $U_{iso} = 0.125$ by the computer program. The least-squares refinement was executed in alternating blocks having residues 1–6 (418 parameters) in one block and residues 7–13 (449 parameters) in the other block. Three positions were found for the C^{γ} atoms of the valyl[10] residue with occupancies of 1.0, 0.5, and 0.5. Water molecule W(1) was found at full occupancy, while the occupancy of W(2) was determined to be ~0.33. The final agreement factors were $R = 0.089$ and $R_W = 0.080$ for 3667 data of $>3\sigma(F)$, where $w = 1/[\sigma^2(|F|) + 0.0004(F_o)^2]$. The final difference map had residual peaks with maximum values of +0.37 and −0.34 e/Å^3.

Fractional coordinates for C, N, and O atoms are listed in Table I with the atoms in partially occupied sites specially designated.

Results and Discussion

Conformation of Molecule. A diagram of the peptide molecule, drawn by computer using the experimentally determined coordinates, is shown in Figure 1. An alternate position for one of the C^{γ} atoms in the side chain of Val[10] is not shown. All three possible conformations occur for the C^{γ} atoms in Val[10], that is, g^+, g^-, and t, although one C^{γ} atom is always at the g^+ position while the other C^{γ} atom occupies either the g^- or t position with ~50% occupancy in each. Furthermore, the C^{β} atom of residue Ala[3] has a very large anisotropic thermal component, $U_{22} = 0.85$ Å^2, in the direction perpendicular to the view in Figure 1. The remainder of the hydrocarbon side chains are reasonably rigid in a single conformation. Figure 2 shows a stereo diagram of the molecule.

Bond lengths and angles have values consistent with those usually observed in peptides (Benedetti, 1977), although some bond lengths in the disordered regions have been fixed to idealized values.

The backbone is a typical helix containing six α-helix-type hydrogen bonds in the center portion and one 3_{10}-helix-type hydrogen bond at the C-terminus and two at the N-terminus, Figure 1 and Table II. The dotted lines between N(5) and O(2) and between N(6) and O(2) represent two possible hydrogen bonds with geometries not consistent with usual parameters. This region of the helix is at the boundary between the 4 → 1 and 5 → 1 hydrogen bonds. In the N(6)H···

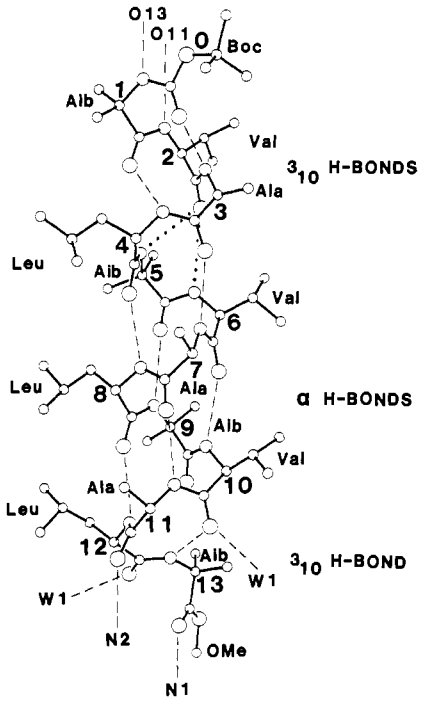

FIGURE 1: View along the $3_{10}/\alpha$-helix of Boc-Aib-(Val-Ala-Leu-Aib)$_3$-OMe, drawn by computer using experimentally determined coordinates. The C^α atoms are labeled 1–13. The number 0 is placed at the O atom of the Boc group. Dashed lines indicate the hydrogen bonds (head to tail, 3_{10}-type, and α-helix type). Dotted lines indicate a 3_{10}-type hydrogen bond with poor geometry and a possible α-type hydrogen bond with a long N···O distance, 3.49 Å.

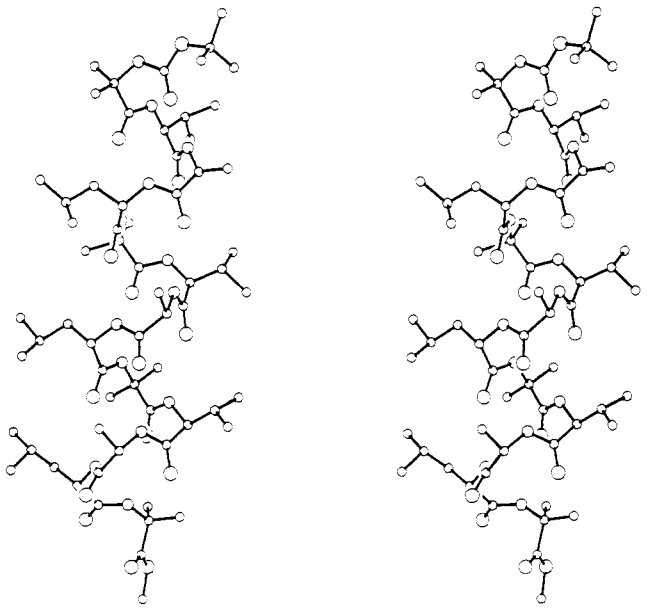

FIGURE 2: Stereo diagram of molecule in same orientation as in Figure 1.

O(2)C′(2) moiety, the hydrogen atom is properly directed toward the carbonyl group for a 5 → 1 type bond; however, the N(6)···O(2) distance is 3.49 Å. In the N(5)H···O(2)C′(2) moiety, the N(5)···O(2) distance is 3.21 Å, a value sometimes observed for hydrogen bonds in structures of apolar helices; however, the NH bond is directed away from O(2) so that the H(5)···O(2) distance is too long at 2.64 Å and the N(5)H···O(2) angle is a mere 118°.

An examination of the conformational angles listed in Table III shows that the values for the rotations about the bonds to $C^\alpha(3)$ are larger for $|\phi|$ and smaller for $|\psi|$ than those for the remainder of the first 11 residues. The somewhat extreme ϕ and ψ values for residue 3 reflect the difficulties for the formation of good 4 → 1 or 5 → 1 hydrogen bonds by N(5) and N(6).

A comparison of the conformation of Boc-Aib-(Val-Ala-Leu-Aib)$_2$-OMe (Karle et al., 1988c) with that of Boc-

Table II: Hydrogen Bonds

	donor[a]	acceptor[a]	D–A (Å)	H–A (Å)	NH···O angle (deg)	N···OC angle (deg)
head to tail	N(1)	O(13)[b]	2.913	2.30	121	152
	N(2)	O(11)[b]	2.951	2.07	151	162
intrahelical	N(3)	O(0)	2.848	2.05	140	131
3_{10}-type	N(4)	O(1)	2.905	2.16	133	122
(4 → 1)	N(5)	O(2)[f]	3.210	2.64	118	108
α-helix (5 → 1)	N(6)	O(2)[f]	3.488	2.58	157	149
	N(7)	O(3)	3.050	2.12	162	155
	N(8)	O(4)	2.958	2.03	161	161
	N(9)	O(5)	3.202	2.62	166	151
	N(10)	O(6)	3.064	2.21	159	159
	N(11)	O(7)	2.969	2.08	153	150
	N(12)	O(8)	3.087	2.35	133	159
3_{10}-type	N(13)	O(10)[g]	3.286	2.36	161	106
hydration	W(1)[c]	O(10)	3.034			
	W(1)[d]	O(12)	3.047			
	W(2)[e]	O(3)	2.937			
	W(2)	O(5)	3.076			

[a] Symmetry equivalent of atoms listed in Table I. [b] $^1/_2 - x$, $1 - y$, $-^1/_2 + z$. [c] $1 - x$, $^1/_2 + y$, $^1/_2 - z$. [d] $2 - x$, $^1/_2 + y$, $^1/_2 - z$. [e] $-1 + x, y, z$. [f] The boundary between 4 → 1 and 5 → 1 hydrogen bonds contains two hydrogen bonds to O(2) with distorted parameters. The N(5)···O(2) bond has a small NH···O angle; the N(6)···O(2) bond has a long N···O distance. [g] A possible N(13)H···O(9) hydrogen bond (5 → 1 type) has D–A = 3.178 Å, H–A = 2.87 Å, NH···O angle = 100°, and N···OC angle = 151°; the middle two parameters have values outside those usually ascribed to hydrogen bonds.

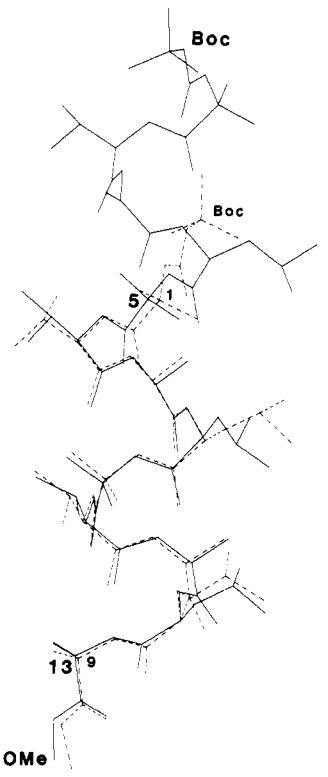

FIGURE 3: Superpositions of Boc-Aib-(Val-Ala-Leu-Aib)$_2$-OMe (dashed lines) and Boc-Aib-(Val-Ala-Leu-Aib)$_3$-OMe (solid lines) by a least-squares fit of backbone atoms. The shorter peptide molecule fits the longer peptide molecule best when the OMe termini are superimposed, rather than when the Boc termini are superimposed.

Table III: Torsional Angles (deg)[a]

	Aib[1]	Val[2]	Ala[3]	Leu[4]	Aib[5]	Val[6]	Ala[7]	Leu[8]	Aib[9]	Val[10]	Ala[11]	Leu[12]	Aib[13]
ϕ(N–C$^\alpha$)	−47[b]	−57	−73	−69	−51	−66	−63	−64	−55	−70	−71	−105	−61
ψ(C$^\alpha$–C′)	−47	−33	−18	−35	−50	−44	−35	−47	−47	−36	−20	5	151[d]
ω(C′–N)	−171	178	167	174	−177	179	177	179	−176	179	179	−176	177[e]
χ^1(C$^\alpha$–C$^\beta$)		72		−176		172		180		85[c]		−59	
		−163				−66				158[c]			
										−65			
χ^2(C$^\beta$–C$^\gamma$)				59				58				−64	
				179				179				174	

[a] The torsion angles for rotation about bonds of the peptide backbone (ϕ, ψ, ω) and about bonds of the amino acid side chains (χ) follow the conventions suggested by the IUPAC–IUB Commission on Biochemical Nomenclature (1970). For a right handed α-helix idealized values of ϕ and ψ are −65° and −41° (Chothia, 1984). For a right-handed 3_{10}-helix, idealized values of ϕ and ψ are −60° and −30°. Estimated SD ~ 1.0°. [b] C′(0)N(1)C$^\alpha$(1)C′(1). [c] Disordered; occupancy 0.5. [d] N(13)C$^\alpha$(13)C′(13)O(OMe). [e] C$^\alpha$(13)C′(13)O(OMe)C(OMe).

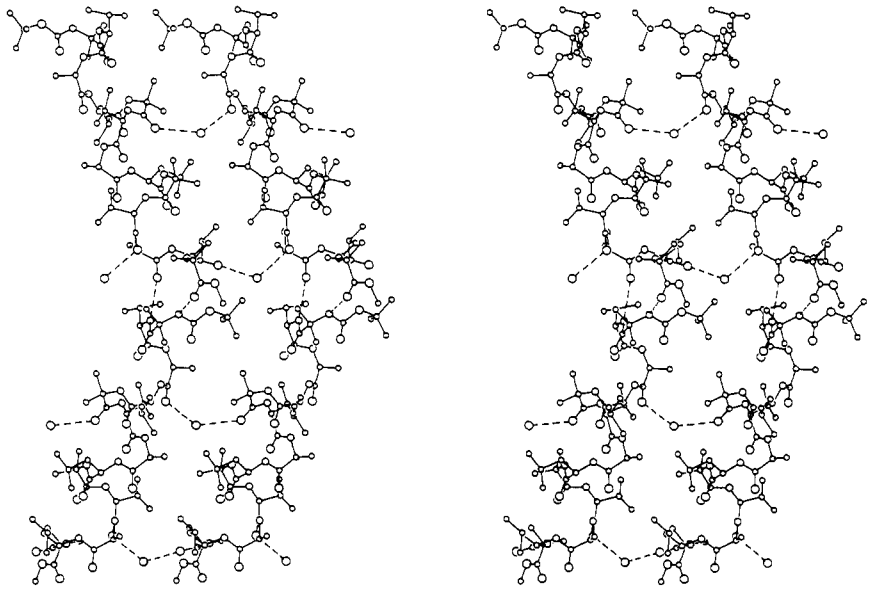

FIGURE 4: Stereo diagram of four neighboring peptide molecules in the crystal. In the horizontal direction, the molecules are related by one cell translation along the *a* axis. In the vertical direction the molecules are related by a 2-fold screw axis. Not shown is a layer of molecules above or below (perpendicular to the view) that is related by another 2-fold screw axis. The molecules not shown have helix directions antiparallel to those shown. Head to tail hydrogen bonds between upper and lower peptide molecules are indicated by dashed lines. Water molecule W(1) is also shown between peptide molecules (at the center and bottom of the diagram) with hydrogen bonds to O(10) and O(12) of neighboring peptide helices. Water molecule W(2) is shown between peptide molecules (at $^1/_4$ and $^3/_4$ height of diagram) with hydrogen bonds to O(3) and O(5) of neighboring peptide helices.

Aib-(Val-Ala-Leu-Aib)$_3$-OMe is shown in Figure 3. The backbone of residues 1–9 of the shorter molecule was fitted by least squares to the backbone of residues 5–13 of the longer molecule. If the first residue of the shorter molecule is disregarded, the average deviation between the backbones of the two molecules is 0.21 Å. The fit of the side chains is also remarkedly close except for Leu[4] of the shorter molecule. In that molecule, the side chain of Leu[4] is disordered among several conformations. For the comparison in Figure 3, the atoms in the conformation with the larger occupancy were used.

Hydration of Backbone. Three different modes of hydrating the backbones of apolar helical peptides have been observed: (a) hydrogen bonding to the carbonyl oxygen left unsatisfied at the bend in the helix caused by a prolyl residue (Karle et al., 1987); (b) the insertion of a water molecule directly into the backbone between a carbonyl and an amide moiety (Karle et al., 1988d, 1989a; Satyshur et al., 1988); (c) an additional lateral hydrogen bond to a carbonyl oxygen in a helix that is accessible to a water molecule by virtue of having only residues with small side groups nearby (Karle et al., 1988c).

In the present structure, a peak in the difference map, which has been labeled as an O of water molecule W(2), has remained stable during the least-squares refinement, and its occupancy has been determined to be ~0.33. This water molecule resides in a void between parallel peptide molecules related by one unit cell translation in the *a* direction, Figure 4. The distances of 2.94 and 3.08 Å between W(2) and O(3) and between W(2) and O(5) of neighboring molecules strongly suggest hydrogen bonding to carbonyl oxygens that are already involved in hydrogen bonding within the helix. The C$_i$═O$_i$···W angles for the two bonds are 128° and 136°, respectively. The next-nearest approach to W(2) is 3.47 Å to C$^\beta$(7). Neither carbonyl oxygen O(3) nor carbonyl oxygen O(5) is shielded by hydrophobic side groups, and each is readily accessible by a water molecule. The W(2)···O(5) hydration in the present 13-residue peptide is entirely comparable in position and approach to the W(3)···O(1) hydration in the nine-residue peptide made up of the same repeating tetrad of residues.

Near the head to tail region, there is a fully occupied water site, W(1). Water W(1) forms only lateral hydrogen bonds to carbonyl oxygens O(10) and O(12) that connect peptide molecules in the *a* direction, perpendicular to the columns. Carbonyl O(10) also forms a 3_{10}-helix-type bond with N(13).

Head to Tail Hydrogen Bonding. The apolar helical peptides containing 9–16 residues studied thus far in this laboratory (Karle et al., 1987, 1988b,c,d, 1989c) form infinite columns by head to tail hydrogen bonding. Backbone amides

N(1)H, N(2)H, and sometimes N(3)H are involved in direct hydrogen bonding with carbonyl O atoms of the molecule above or in indirect hydrogen bonding through water bridges or other cocrystallized solvent molecules such as methanol or 2-propanol. In the present structure, Figure 4, there are two direct head to tail hydrogen bonds, N(1)H···O(13) and N-(2)H···O(11).

The columns formed by the peptide helices are not straight but form a zigzag pattern through the crystal in this structure. In the similar nine-residue peptide, the columns are quite straight.

Aggregation of Helices. Each of the four different residues in this peptide repeats by $i + 4$. In figure 1 the three Leu residues appear on the left of the helix, the three Val residues are on the right, the Ala side chains point forward, and the Aib side chains point toward the rear, except for Aib[1]. It was thought that regularity in the distribution of the side chains may induce some regularity in intermolecular association such as interdigitation. No such regularity has been observed. For the most part, the nearest approach between side chains in neighboring molecules is ~4.2 Å. A layer of peptide molecules shown in Figure 4 shows two water molecules, placed between pairs of parallel helices, that join the helices by hydrogen bonding into an infinite sheet.

Above and below the layer shown in Figure 4 are layers related by a 2-fold screw so that the helical axes are directed antiparallel to those shown. Such antiparallel association of helices does not occur in the analogous nine-residue peptide (Karle et al., 1988c) in which the helix direction is parallel for all the peptide molecules in a crystal. In the apolar helical peptides studied in this laboratory, the packing of helices has been parallel more often than antiparallel, contrary to expectations based on calculations of the electrostatic interactions between helix macrodipoles in crystalline arrays (Hol & de Maeyer, 1984). Thus far, no particular packing motifs have predominated, and no selectivity for particular hydrocarbon side chains has been demonstrated.

Implications for Peptide Design. The conformation of the 13-residue peptide Boc-Aib-(Val-Ala-Leu-Aib)$_3$-OMe in crystals clearly supports the conclusion that peptide sequences containing 25–30% Aib residues overwhelmingly favor helical conformations. Helical rods containing three to four turns of $3_{10}/\alpha$-helix can be readily constructed with Aib-containing sequences. The excellent solubility properties of completely apolar sequences, containing a few Aib residues, in organic solvents permit their conformational characterization in solution by NMR (Balaram et al., 1986; K. Uma, unpublished results). Their crystallizability also facilitates unequivocal determination of their conformation in the solid state by X-ray diffraction. Crystallization may, in fact, be a consequence of appreciable conformational homogeneity in solution, suggesting that such sequences may be ideal candidates for design and construction of stereochemically rigid modules in a "molecular Meccano (or Lego) set" approach to the synthesis of protein mimics. The use of apolar (hydrophobic) sequences permits characterization of conformations in poorly solvating organic solvents, under conditions where peptide folding is largely controlled by nonbonded interactions, intramolecular hydrogen bonding, and electrostatic effects. Such a situation is less complex and more tractable at the present stage of synthetic design than studies in aqueous solutions, where hydrophobic effects generally play a major, but incompletely understood, role in dictating the nature of polypeptide folding (Creighton, 1985). Studies presently under way are focused on the construction of two-helix motifs incorporating helix-breaking linker sequences, with the eventual aim of piecing together a hydrophobic four-helix bundle.

Supplementary Material Available

Five tables listing atomic coordinates and equivalent isotropic displacement parameters (Table I), bond lengths (Table II), bond angles (Table III), anisotropic displacement parameters (Table IV), and hydrogen atom coordinates and isotropic displacement parameters (Table V) (9 pages); table of observed and calculated structure factors (22 pages). Ordering information is given on any current masthead page. Supplementary material consisting of bond lengths, bond angles, anisotropic thermal parameters, and coordinates for hydrogen atoms will be deposited with the Cambridge Crystallographic Data File.

Registry No. Boc-Aib-(Val-Ala-Leu-Aib)$_3$-OMe, 121268-38-0.

References

Balaram, P. (1984) *Proc.—Indian Acad. Sci., Chem. Sci. 93*, 703–717.

Balaram, H., Sukumar, M., & Balaram, P. (1986) *Biopolymers 25*, 2209–2223.

Benedetti, E. (1977) in *Peptides. Proceedings of the Fifth American Peptide Symposium* (Goodman, M., & Meienhofer, J., Eds.) pp 257–279, Wiley, New York.

Bosch, R., Jung, G., Schmitt, H., & Winter, W. (1985a) *Biopolymers 24*, 961–978.

Bosch, R., Jung, G., Schmitt, H., & Winter, W. (1985b) *Biopolyers 24*, 979–999.

Chothia, C. (1984) *Annu. Rev. Biochem. 53*, 537–572.

Creighton, T. E. (1985) *J. Phys. Chem. 89*, 2452–2459.

Egert, E., & Sheldrick, G. M. (1985) *Acta Crysallogr. A41*, 262–268.

Eisenberg, D., Wilcox, W., Eshita, S. M., Pryciak, P. M., Ho, S. P., & DeGrado, W. F. (1986) *Proteins 1*, 16–22.

Fox, R. O., Jr., & Richards, F. M. (1982) *Nature (London) 300*, 325–330.

Ho, S. P., & DeGrado, W. F. (1987) *J. Am. Chem. Soc. 109*, 6751–6758.

Hol, W. G. J., & deMaeyer, M. C. H. (1984) *Biopolymers 23*, 809–817.

IUPAC–IUB Commission on Biochemical Nomenclature (1970) *Biochemistry 9*, 3471–3479.

Karle, J. (1968) *Acta Crystallogr. B24*, 182–186.

Karle, I. L., Sukumar, M., & Balaram, P. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9284–9288.

Karle, I. L., Flippen-Anderson, J. L., Sukumar, M., & Balaram, P. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5087–5091.

Karle, I. L., Kishore, R., Raghothama, S., & Balaram, P. (1988a) *J. Am. Chem. Soc. 110*, 1958–1963.

Karle, I. L., Flippen-Anderson, J. L., Sukumar, M., & Balaram, P. (1988b) *Int. J. Pept. Protein Res. 31*, 567–576.

Karle, I. L., Flippen-Anerson, J. L., Uma, K., & Balaram, P. (1988c) *Int. J. Pept. Protein Res. 32*, 536–543.

Karle, I. L., Flippen-Anderson, J. L., Uma, K., & Balaram, P. (1988d) *Proc. Natl. Acad. Sci. U.S.A. 85*, 299–303.

Karle, I. L., Flippen-Anderson, J. L., Uma, K., & Balaram, P. (1989a) *Biopolymers 28*, 773–781.

Karle, I. L., Flippen-Anderson, J. L., Kishore, R., & Balaram, P. (1989b) *Int. J. Pept. Protein Res.* (in press).

Karle, I. L., Flippen-Anderson, J. L., Uma, K., Balaram, H., & Balaram, P. (1989c) *Proc. Natl. Acad. Sci. U.S.A. 86*, 765–769.

Kishore, R., Kumar, A., & Balaram, P. (1985) *J. Am. Chem. Soc. 107*, 8019–8023.
Kishore, R., Raghothama, S., & Balaram, P. (1987) *Biopolymers 26*, 873–891.
Marshall, G. R., Bosshard, H. E., Kendrick, N. C. E., Turk, J., Balasubramanian, T. M., Cobb, S. M. H., Moore, M., Leduc, L., & Needleman, P. (1976) in *Peptides 1976* (Loffet, A., Ed.) pp 361–369, Editions de l'Universitē de Bruxelles, Brussels, Belgium.
Mathew, M. K., & Balaram, P. (1983a) *FEBS Lett. 157*, 1–5.
Mathew, M. K., & Balaram, P. (1983b) *Mol. Cell. Biochem. 50*, 47–64.
Menestrina, G., Voges, K. P., Jung, G., & Boheim, G. (1986) *J. Membr. Biol. 93*, 111–132.
Mutter, M. (1988) *Trends Biochem. Sci. (Pers. Ed.) 13*, 260–265.
Mutter, M., Altmann, K.-H., Tuchscherer, G., & Vuilleumier, S. (1988) *Tetrahedron 44*, 771–785.
Nagaraj, R., & Balaram, P. (1981) *Acc. Chem. Res. 14*, 356–362.
Osterman, D. G, & Kaiser, E. T. (1985) *J. Cell. Biochem. 29*, 57–72.
Prasad, B. V. V., & Balaram, P. (1984) *CRC Crit. Rev. Biochem. 16*, 307–348.
Ravi, A., & Balaram, P. (1984) *Tetrahedron 40*, 2577–2583.
Regan, L., & Degrado, W. F. (1988) *Science 241*, 976–978.
Richardson, J. S., & Richardson, D. C. (1987) in *Protein Engineering* (Oxender, D. L., & Fox, C. F., Eds.) pp 149–163, Alan R. Liss, New York.
Satyshur, K. A., Rao, S. T., Pyzalska, D., Drendel, W., Greaser, M., & Sundaralingam, M. (1988) *J. Biol. Chem. 263*, 1628–1647.
Toniolo, C., Bonora, G. M. Bavoso, A., Benedetti, E., Di Blasio, B., Pavone, V., & Pedone, C. (1983) *Biopolymers 22*, 205–215.

Direct Observation of the Titration of Substrate Carbonyl Groups in the Active Site of α-Chymotrypsin by Resonance Raman Spectroscopy†

Peter J. Tonge and Paul R. Carey*
Division of Biological Sciences, National Research Council of Canada, Ottawa, Ontario, Canada K1A 0R6
Received March 30, 1989

ABSTRACT: By use of resonance Raman (RR) spectroscopy, the population of the reactive carbonyl group in active acylchymotrypsins has been characterized and correlated with acyl-enzyme reactivity. RR spectra have been obtained, with a flow system and 324- and 337.5-nm excitation, at low and active pH for six acylchymotrypsins, viz., (indoleacryloyl)-, (4-amino-3-nitrocinnamoyl)-, (furylacryloyl)-, [(5-ethylfuryl)-acryloyl]-, (thienylacryloyl)-, and [(5-methylthienyl)acryloyl]chymotrypsin. These acyl-enzymes represent a 100-fold range of deacylation rate constants. Good RR spectral quality has enabled us to obtain the vibrational spectrum of the carbonyl group at low and active pH in each acyl-enzyme. The measured pK_a of the spectroscopic changes in the carbonyl region is identical with that for the deacylation kinetics, showing that the RR carbonyl features reflect the ionization state of His-57. A carbonyl population has been observed in the active acyl-enzymes in which the carbonyl oxygen atom of the reactive acyl linkage is hydrogen-bonded in the active site. The proportion of this hydrogen-bonded population, with respect to other observed non-hydrogen-bonded species, together with the degree of polarization of the carbonyl bond, as monitored by $\nu_{C=O}$, has been correlated with the deacylation rate constants of the acyl-enzymes. It is proposed that the hydrogen-bonded carbonyl species is located at or near the oxyanion hole and represents the ground state from which deacylation occurs. An increase in the proportion of the hydrogen-bonded population and an increase in polarization of the carbonyl bond result in an increase in deacylation rate constant. For the first time we have direct RR data on the catalytically transformed region which relate to the difference in acyl-enzyme reactivity.

In order to completely describe the reaction catalyzed by serine proteases, it is important to understand the chemistry of the acyl-enzyme carbonyl group. Here we show that resonance Raman (RR) spectroscopy is a good probe of the structure and environment of the acyl carbonyl group in *O*-(arylacryloyl)chymotrypsins. The RR carbonyl stretching band profile is used to monitor changes in the carbonyl group as the catalytic mechanism is activated. Using this approach, we identify an "active" carbonyl population in which the acyl carbonyl oxygen atom is hydrogen-bonded in the oxyanion hole. The quantity of this hydrogen-bonded species as well as its degree of polarization can then be related to the relative rates of deacylation of each acyl-enzyme.

Direct vibrational spectroscopic analysis of the reactive carbonyl group is limited to intermediates which accumulate on the reaction pathway. During the hydrolysis of substrates by serine proteases, a covalently bound acyl-enzyme intermediate is formed. In general, for ester substrates acylation (acyl-enzyme formation) is much faster than deacylation (acyl-enzyme breakdown), and the acyl-enzyme intermediate accumulates (Kraut, 1977). In order to obtain the vibrational spectrum of the acyl-enzyme carbonyl by RR spectroscopy, a chromophoric acyl group is required whose electronic absorption spectrum includes a contribution from the carbonyl group. In the present study we use β-substituted arylacryloyl compounds, which commonly absorb in the 300–400-nm region. This class of compounds has been used extensively as

† Issued as NRCC Publication No. 30446.
* To whom correspondence should be addressed.

0006-2960/89/0428-6701$01.50/0