Crystal structure of a hydrophobic 19-residue peptide helix containing three centrally located D amino acids

Isabella L. Karle*†, Hosahudya N. Gopi‡, and Padmanabhan Balaram‡

*Laboratory for the Structure of Matter, Naval Research Laboratory, Washington, DC 20375-5341; and ‡Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India

Contributed by Isabella L. Karle, September 23, 2003

The design of the synthetic 19-residue peptide Boc-Leu-Aib-Val-Ala-Leu[5]-Aib-Val-D-Ala-D-Leu-Leu[10]-Val-Phe-Val-Aib-D-Val[15]-Leu-Phe-Val-Val-OMe (Aib, α-aminoisobutyric acid; OMe, methyl ester) was intended to produce a crystalline peptide with independent helical and hairpin domains. The design was partially based on an octapeptide with the same sequence as residues 11–18 above, which was shown to fold into a β-hairpin in the crystal. However, the crystal structure of the present peptide provided a surprising result. The conformation is the longest characterized right-handed α-helix, with as many as three internal D residues in the sequence. The completely helical structure was also unexpected, because β-branched residues such as Val have a low propensity for helix formation in proteins. The helical peptides in the present structure assemble to form hydrophobic channels that accommodate five toluene molecules per peptide along the length of the channel. The structural results illustrate the similarity in energetics between helical and β-hairpin conformations for peptides containing Aib residues. The crystallographic parameters for $C_{107}H_{179}N_{19}O_{22}\cdot 3H_2O\cdot 2.5$ toluenes are: space group C2, a = 34.679(3) Å, b = 12.866(1) Å, c = 31.915(3) Å, β = 96.511(8)°, V = 14,148 Å^3, Z = 4, d_{calc} = 1.099 g/cm^3, and agreement factor R_1 = 10.2%.

Stereochemically constrained amino acid residues can be effectively used to nucleate specific polypeptide secondary structures like α-helices and β-hairpins (1). α-Aminoisobutyric acid (Aib) and related α, α-dialkylamino acid residues have been used in the design of well characterized polypeptide helices (2–5). The design of β-hairpins is facilitated by the introduction of centrally positioned D-Pro-Xxx segments into polypeptide sequences (6–8). Several examples of crystalline β-hairpins have been designed (9–11), and the extension of this design strategy to the construction of multistranded β-sheet structures has been reported (12–14). The availability of relatively rigid modules of secondary structures permits the construction of synthetic peptides incorporating multiple structural elements. The crystal structure determination of the 17-residue peptide Boc-Val-Ala-Leu-Aib-Val-Ala-Leu-Gly-Gly-Leu-Phe-Val-D-Pro-Gly-Leu-Phe-Val-OMe (OMe, methyl ester) provides an example of both helical and β-strand segments in a designed synthetic sequence (15). In the Lego set approach to the construction of designed polypeptides, units of secondary structure may be assembled by using linking segments of potentially defined conformational properties. In such an approach, stereochemical control over polypeptide chain folding may be achieved principally through local conformational constraints, imposed by nonbonded interactions; the structural stability of the domains is enhanced by intramolecular hydrogen bonding in the helices and hairpins. In exploring the Lego set approach to *de novo* protein design, we designed and synthesized a 19-residue peptide HBH-19 (Fig. 1), which was anticipated to have well defined helical and hairpin domains. The choice of the sequence Boc-Leu-Aib-Val-Ala-Leu-Aib-Val-D-Ala-D-Leu-Leu-Val-Phe-Val-Aib-D-Val-Leu-Phe-Val-Val-OMe was based on the previously determined structures of the peptide Boc-Leu-Aib-Val-Ala-Leu-Aib-Val-D-Ala-D-Leu-Aib-OMe (LAD 10, a right-handed helix terminating in a Schellman motif) (16) and Boc-Leu-Phe-Val-Aib-D-Ala-Leu-Phe-Val-OMe (UDA8), which yielded a β-hairpin in crystals (17). The structures of the crystallographically characterized fragments corresponding to residues 1–9 and 11–18 of the designed sequence are illustrated in Fig. 1. We describe in this report the crystal structure of HBH-19, which forms a long continuous right-handed helix encompassing the entire length of the sequence despite the presence of as many as three D amino acids, including a centrally placed double-D segment (Fig. 2).

Table 1. Torsional angles in backbone and side chains in HBH-19

Angle	ϕ	φ	ω	χ_1	χ_2	$C'_{i-1}N_iC_i^\alpha C_i^\beta$ (κ)
Leu-1	−55	−36	−174	−173	−166	180
					+66	
Aib-2	−51	−42	−173			
Val-3	−69	−44	180	−65		168
				168		
Ala-4	−60	−44	178			175
Leu-5	−56	−53	−174	−180	−168	178
					+65	
Aib-6	−56	−44	179			
Val-7	−60	−52	177	−61		171
				174		
D-Ala-8	−53	−45	−179			+77
D-Leu-9	−54	−49	180	+55	+64	+72
					−169	
Leu-10	−62	−42	180	−68	173	177
					−67	
Val-11	−62	−42	178	−52		168
				−178		
Phe-12	−61	−43	179	177	−123	176
					+64	
Val-13	−61	−50	177	168		170
				−69		
Aib-14	−52	−49	−176			
D-Val-15	−56	−52	−178	169		+76
				−47		
Leu-16	−59	−43	−175	−112	−10	175
					−174	
Phe-17	−81	−8	166	−75	−99	147
					+82	
Val-18	−68	−37	176	165		165
				−71		
Val-19	−96	+10	178	+64		137
				−62		

Abbreviations: Boc, *t*-butoxycarbonyl; OMe, methyl ester; Aib, aminoisobutyric acid (U, one-letter code) (dimethyl glycine); α_R, right-handed α-helix.

Data deposition: The atomic coordinates have been deposited in the Cambridge Structural Database, Cambridge Crystallographic Data Centre, www.ccdc.cam.ac.uk (CSD ref. no. 199070).

†To whom correspondence should be addressed. E-mail: williams@harker.nrl.navy.mil.

www.pnas.org/cgi/doi/10.1073/pnas.2336106100

Boc-**Leu-Aib-Val-Ala-Leu-Aib-Val-^{D}Ala-^{D}Leu**-Leu-Val-**Phe-Val-Aib**-^{D}Val-**Leu-Phe-Val**-Val-OMe (**HBH19**)
Boc-**Leu-Aib-Val-Ala-Leu-Aib-Val-^{D}Ala-^{D}Leu**-Aib -OMe (**LAD10**)
Boc- Leu-**Phe-Val-Aib**-^{D}Ala-**Leu-Phe-Val** -OMe (**UDA8**)

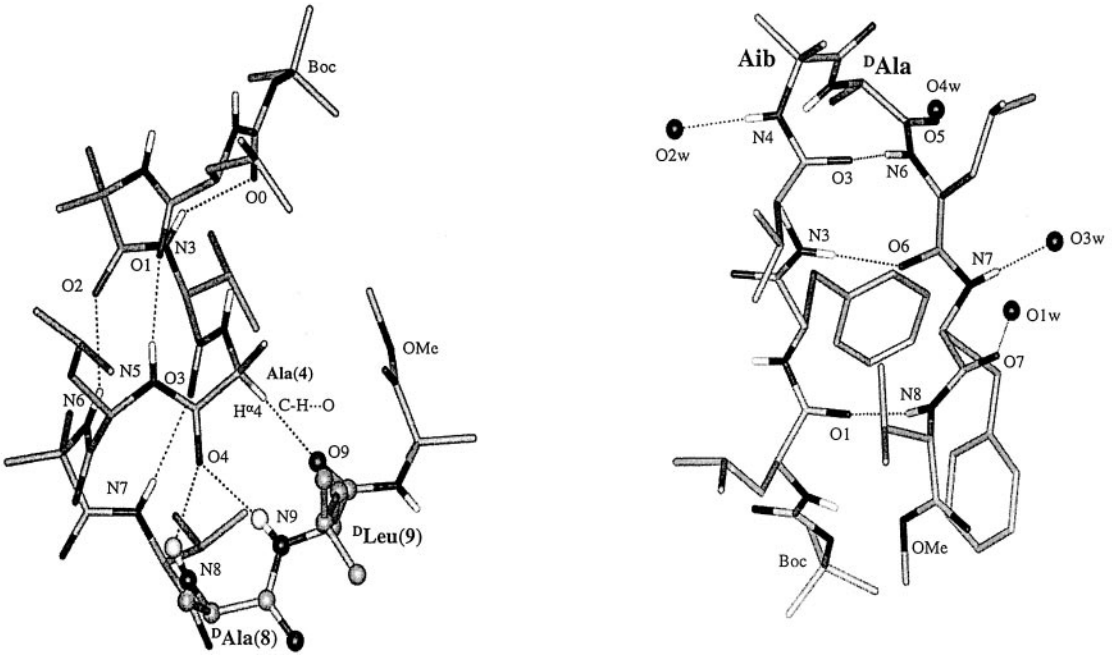

Fig. 1. (*Upper*) Sequences of peptides HBH-19, LAD10, and UDA8. (*Lower*) Molecular conformations in crystals of LAD10 (*Left*) (16) and UDA8 (*Right*) (17).

Experimental Methods

Peptide Synthesis. The 19-residue peptide HBH-19 was synthesized by conventional solution phase procedures by using a fragment condensation strategy. *t*-Butyloxycarbonyl (Boc) and methyl ester groups are used for N- and C-terminal protection, respectively. Peptide couplings were mediated by N,N′-dicyclohexylcarbodiimide and 1-hydroxybenzotriazole. The final fragment coupling was achieved by using the segments 1–9 and 10–19. The peptide was purified by medium-pressure liquid chromatography on a reverse-phase C_{18} (10–60 μm) column, followed by HPLC on a C_{18} (5–10 μm) column with methanol–water gradients. The identity of the final peptide was confirmed by matrix-assisted laser desorption ionization mass spectrometry [M_{obs} = 2,106 Da; M_{calc} = 2,104 Da M^{+} Na^{+}] and complete assignment of the 500-MHz ^{1}H NMR spectrum. Crystallization attempts from aqueous methanol proved unsuccessful. Crystals suitable for x-ray diffraction were obtained after dissolving the peptide in a hot methanol–water–toluene mixture, followed by slow evaporation at room temperature.

X-Ray Diffraction. A crystal in the form of a wafer, 0.27 × 0.27 × 0.04 mm in size, was used to collect x-ray diffraction data on a Bruker (Billerica, MA) charge-coupled device diffractometer by using a high-intensity rotating anode (Cu radiation) to a preset limit of 67° for θ (0.83-Å resolution). Due to the very weak reflections at the larger scattering angles and the resulting uncertainty in the intensity measurements, the experimental data were limited to a resolution of 1.0 Å. The structure solution process was not automatic. The inclusion of three D residues in the 19-residue sequence made the prediction of the overall conformation of the backbone uncertain. Accordingly, various models for hairpin turns, 3_{10}-helices, and α-helices were used in vector search procedures for rotation and translation of a known model (18, 19).

A model consisting of an α-helical 15-atom fragment from the central portion of Boc-Val-Ala-Leu-Aib-Val-Ala-Leu-Val-Ala-Leu-Aib-Val-Lac-Leu-OMe (20) yielded one turn of an α-helix in HBH-19 after rotation, translation, and partial structure extension with the use of the tangent formula (21). A slow accretion of atoms in the structure occurred after each of many single cycles of tangent formula extension and difference maps, until 90 atoms were located. During these procedures, the 20 C═O moieties appeared as single elongated peaks. After several isotropic least-squares refinements, the C═O moieties separated into single peaks for the C and O atoms, and additional atoms were located at the flexible N and C termini, as well as some in the flexible side chains of Leu and Phe residues. The complete 148-atom peptide appeared during anisotropic refinement. After the addition of 179 hydrogen atoms in idealized positions and three water molecules, R_1 was still at 12.5%. Large voids were present in the crystal in a direction lateral to the helices. The cross sections of the voids are >10 Å. One toluene molecule occurring with three of its C atoms on a 2-fold rotation axis was clearly visible. Another two toluene molecules that were loosely fitted into spaces between the peptide molecules were discernable. The C atoms in the three toluene molecules were restrained to give ideal bond engths of 1.39 Å in the ring and 1.53 Å to the CH_3 group. For the crystal with a formula containing $C_{107}H_{179}N_{19}O_{22}{\cdot}3H_2O{\cdot}2.5$ toluenes, the crystallographic parameters are: space group C2, cell parameters a = 34.679(3) Å, b = 12.866(1) Å, c = 31.915(3) Å, β = 96.511(8)°, V = 14,148 Å^3, Z = 4, d_{calc} = 1.099 g/cm^3, final R_1 factor = 10.2%.

Details of data collection, coordinates, bond lengths and angles, and anisotropic thermal parameters are deposited in the Cambridge Crystallographic Data Centre, reference no. CCDC 199070.

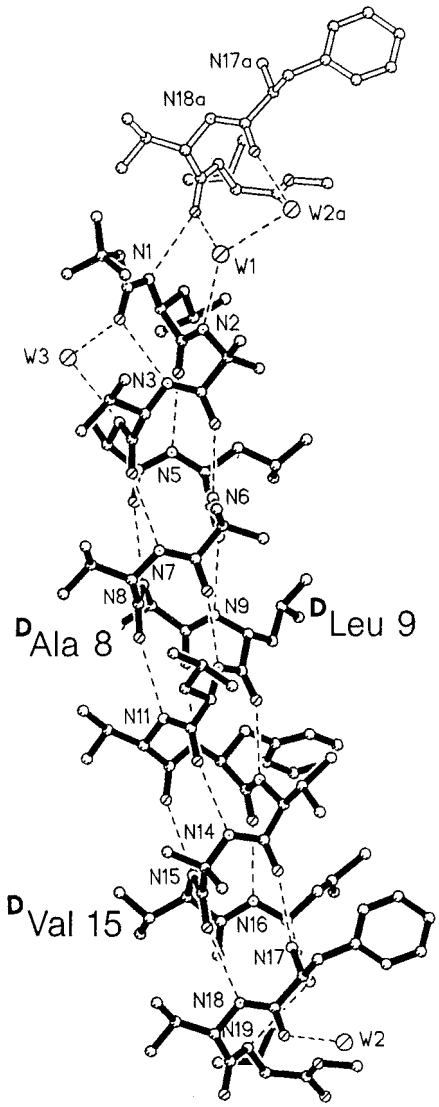

Fig. 2. The α_R formed by the 19-residue peptide that contains three D residues, Boc-LUVALUV-D-A-D-L-LVFVU-D-V-LFVV-OMe. Hydrogen bonds are indicated by dashed lines. The light-colored segment at the top of the diagram, containing the N17a to C20a fragment, shows the head-to-tail connection between helices that stack into continuous columns.

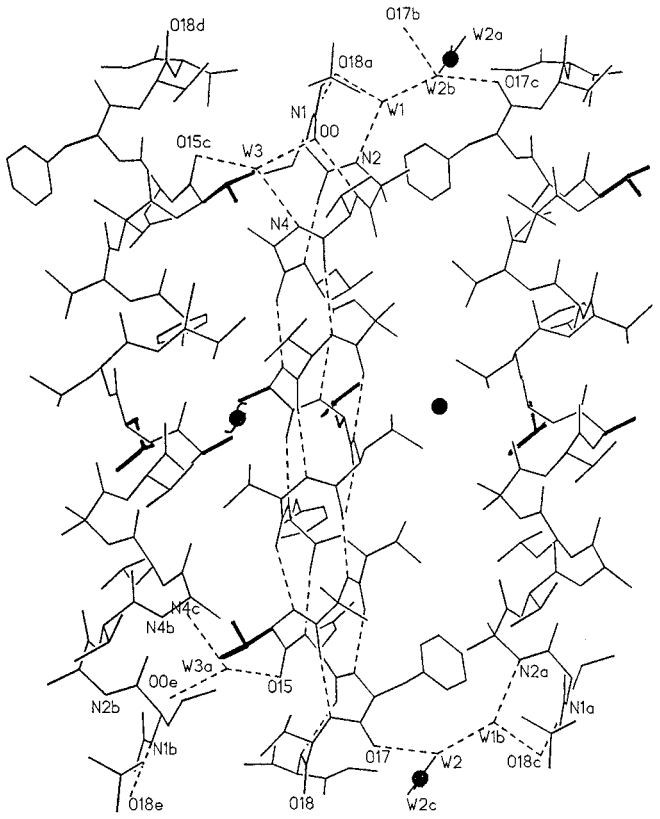

Fig. 3. Three adjacent antiparallel helical molecules in the crystal. The pair on the left is related by a 2-fold screw axis, and the pair on the right is related by a 2-fold rotation axis. Side chains for the D residues are darkened.

Results and Discussion

Fig. 2 shows a view of the molecular conformation of HBH-19, which reveals a complete right-handed α-helix (α_R) over the entire backbone. The backbone and side chain torsion angles and hydrogen bonds are listed in Tables 1 and 2. Remarkably, despite the insertion of the doublets of D-Ala-8 and D-Leu-9 and Aib-14 and D-Val-15 into the sequence, the molecule forms an almost ideal α-helix. There is ample space between the atoms in the side chains of the D residues and atoms in the neighboring side chains of L residues. The closest intramolecular approach is between C9D1 of D-Leu-9 and C5D2 of L-Leu-5 at 3.99 Å, well above minimum values for van der Waals approaches between nonbonded C atoms.

Torsional angles for the backbone and side chains are listed in Table 1. The average values for residues 1 to 16 are $\phi = -60°$, $\Psi = -46°$, close to the "classical" values of (−57°, −47°) for an α_R formed by L amino acids. For the D residues in the present molecule, D-Ala 8, D-Leu 9, and D-Val 15, the average ϕ, ψ values are (−54°, −49°). The change in sign for the ϕ, ψ values in the D residues is compatible with Ramachandran's original maps showing the allowed regions of torsional ϕ and ψ values (4, 22), for example, where positive ϕ and ψ values for α-helices consisting of L residues are "partially allowed" and, by inference, negative ϕ and ψ values for D residues in α-helices. Experimental observations of the incorporation of single D residues in right-handed helical peptides or, conversely, L residues into left-handed helical segments are available (23, 24). A double D residue segment in a 13-residue right-handed helix has been reported in crystals (25). The present structure is an example of three internal D residues incorporated into a continuous α_R The only indication in Table 1 among the values of torsional angles that D residues are present is the values for the $C'_{i-1}N_i\ C\alpha_i C\beta_i$ (κ) angles. The κ angle assumes values near +75° for the D residues rather than near 180°, as is common for the L residues. Hydrogen bond parameters are listed in Table 2. There are 14 consecutive hydrogen bonds of the 5→1 type, with N–O distances that range from 2.895 to 3.022 Å, except for N9-O5 (2.865 Å) and N11-O7 (3.093 Å) that may have been mildly disturbed by the presence of D residues, D-Ala 8, and D-Leu 9. At the C terminus, one weak 4→1 type hydrogen may be present. Both the N19–O16 and H19–O16 distances are long; there is no other possibility for N19H to participate in a hydrogen bond. Near the C terminus, N1H makes a direct hydrogen bond to O18a of the next molecule in the head-to-tail region of the continuous column of helices. N2H is a donor to a water molecule W1, which in turn is also a donor to O18a and W2a. W2a is also a donor to O17a. W2a and W2b (not shown) are related by a 2-fold rotation axis, and their locations are too close to each other (2.20 Å) for coexistence. Hence, each location can have only one-half occu-

Table 2. Hydrogen bonds in HBH-19 at R_1 = 0.102

Type	Donor	Acceptor	D—A, Å	H—A, Å	Angle C═O···N(or O),°
Head-to-tail	N1	018*	2.947	2.12	155
Peptide–water	N2	W1	2.914	2.03	
4→1	N3	00	3.029	2.24	132
4→1	N4	01	3.000	2.54†	119
5→1	N5	01	2.943	2.03	170
5→1	N6	02	2.958	2.09	148
5→1	N7	03	3.020	2.13	155
5→1	N8	04	3.022	2.14	158
5→1	N9	05	2.865	1.98	159
5→1	N10	06	2.967	2.11	155
5→1	N11	07	3.093	2.22	163
5→1	N12	08	2.975	2.10	161
5→1	N13	09	3.007	2.12	157
5→1	N14	010	2.986	2.09	153
5→1	N15	011	2.895	2.05	151
5→1	N16	012	2.909	2.03	156
5→1	N17	013	3.022	2.15	160
5→1	N18	014	2.998	2.22	161
4→1	N19	016	3.287‡	2.66‡	108
Water–peptide	W1	018*	2.914		
	W1	W2§	3.097		
	W2	017	2.780		
	W2	017¶	2.590		
	W3	O0	2.856		
	W3	O15‖	2.940		
	N4	W3	3.084	2.44†	

*Symmetry equivalent $x, y, -1 + z$

†The donor N4H appears to make a bifurcated hydrogen bond to W3 and O1 with long H to acceptor distances. O1 is also an acceptor for N5H, where normal D—A and H—A distances prevail.

‡The final hydrogen bond in the helix, N19H···O16, has long distances for D—A and H—A. There are no other possibilities for hydrogen bonds with N19H.

§Symmetry equivalent $1-x, y, 1-z$.

¶Symmetry equivalent $1-x, y, 2-z$.

‖Symmetry equivalent $0.5-x, 0.5 + y, 1-z$.

pancy. W3 forms a hydrogen-bonded bridge between N4 and O0 and, in addition, forms a link to carbonyl oxygen O15b in a lateral peptide molecule. Finally, O1 is at a transition point between a short segment of a 3_{10}-helix and the long α-helical portion. All of the hydrogen bonds are inside the extended helical column or three water molecules that link the helices in a head-to-tail fashion.

The helical columns have a totally hydrophobic surface without any possibility for hydrogen bonds between the columns. Fig. 3 shows three adjacent columns with antiparallel packing. The pair on the left has numerous van der Waals contacts between the ends of the side chains with C–C distances of 3.8–3.9 Å, between the adjacent molecules. For the peptide pair on the right, there is ample space for cocrystallized solvent molecules as shown in Fig. 4. Three toluene molecules, T1, T2, and T3, and their symmetry mates, T2′ and T3′, fill the hydrophobic space quite efficiently. There are appropriate van der Waals contacts between the toluene molecules and the hydrophobic side chains of the peptide helices with C–C distances ranging from 3.54 to 4.02 Å.

A view into the helices shows the formation of the channel by four α-helices (Fig. 5). The channel is in the shape of a straight tube but is irregular in cross section. It appears to be continuous along the *c* axis direction. The walls of the channel are stabilized by van der Waals contacts of 3.7–3.8 Å between the phenyl group of Phe-17 of one helix and atoms C2B and C3G1 of an adjacent

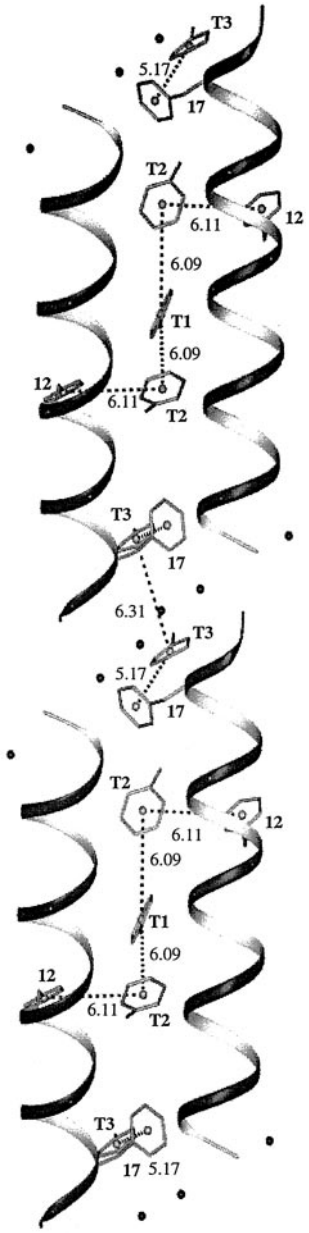

Fig. 4. A view of the toluene channel. The helical peptide molecules are schematically shown as ribbons. The centroid-to-centroid distance of proximal aromatic rings is marked. Toluene molecules are labeled T1, -2, and -3. The numbers 12 and 17 refer to Phe-12 and -17.

helix, as well as similar contacts between C5D1 and C3G2 of other adjacent helices around the channel. Notably, the toluene molecules accommodate their orientations to the irregular shape of the channel. There is no π-stacking or distinct T shaped approaches between the rings of the toluene molecules or with the rings of the peptide helix (Fig. 4). The only close approaches are between C3T of T1 and C14T of T2, at 3.66 Å; between C16T of T2 and C12F of Phe-12, at 3.69 Å; and between C25T of T3 and C17F of Phe-17, at 3.66 Å.

The present example illustrates clearly the surprises that may lie in store in the area of designer peptides. The octapeptide sequence in Boc-Leu-Phe-Val-Aib-D-Ala-Leu-Phe-Val-OMe adopted a β-hairpin conformation, nucleated by an Aib-D-Ala Type I′ β-turn, as anticipated by design (17). In contrast, the closely related segment -Leu-Phe-Val-Aib-D-Val-Leu-Phe-Val- adopts a right-

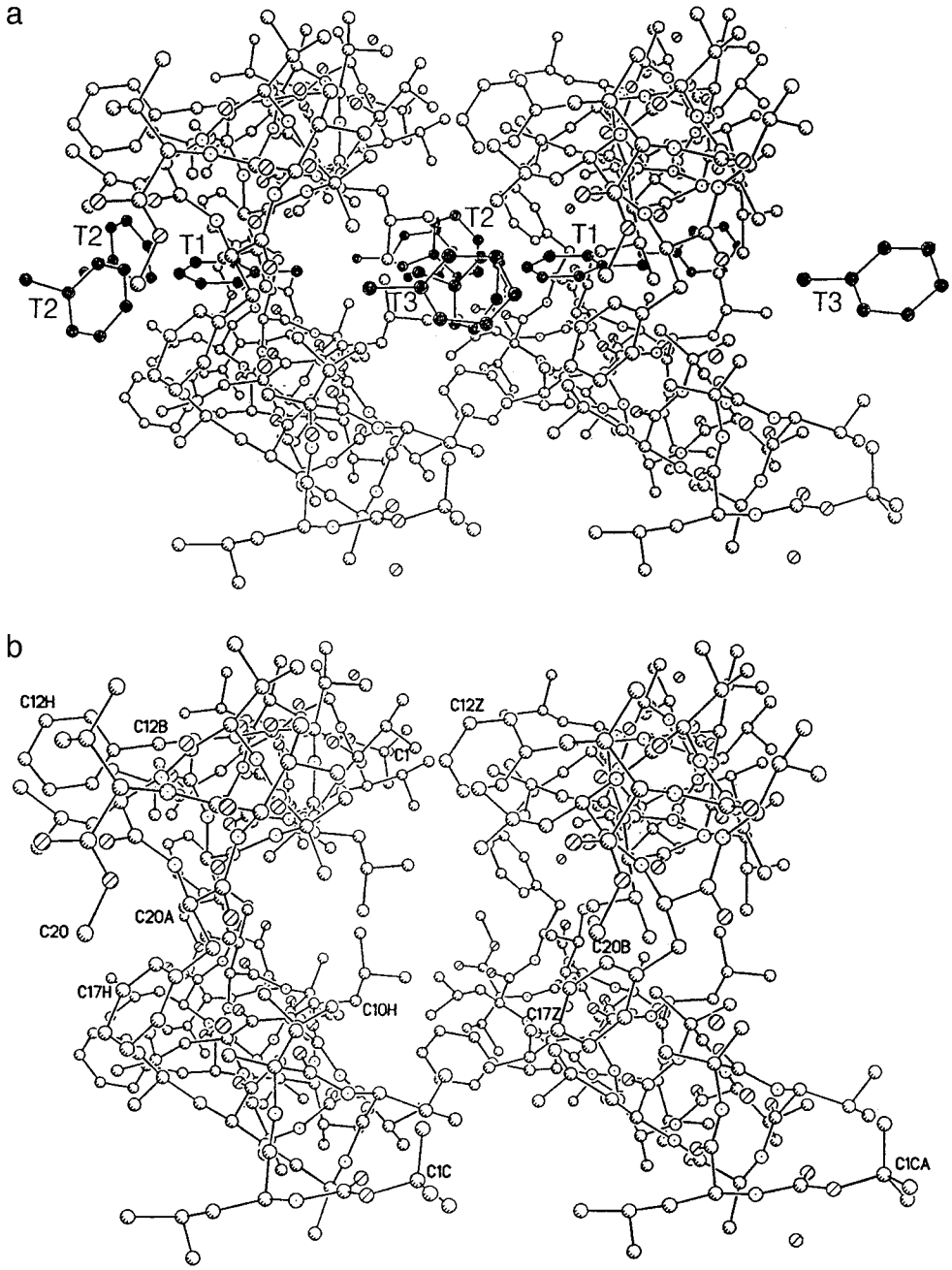

Fig. 5. (*a*) View into four adjacent helical molecules that surround the hydrophobic channel. The *b* axis is horizontal. The toluene molecules are darkened. (*b*) The same view without the toluene molecules.

handed helical conformation in HBH-19. In the peptide Boc-Leu-Aib-Val-Ala-Leu-Aib-Val-D-Ala-D-Leu-Aib-OMe a right-handed helix is formed over the segment 1–7, with D-Ala adopting an α_L conformation resulting in the formation of helix terminating Schellman motif. Further, a peptide chain reversal stabilized by a CH···O bond was observed at the C terminus (16), a feature also observed in protein structures (26). Dramatically, the corresponding segment transforms into a right-handed helix in HBH-19. The energetic penalty for incorporating D-Ala residues into an α_R conformation has been estimated as 1.2 kcal·mol^{-1} (27). Presumably, the formation of a continuous hydrogen bond network offsets this disadvantage. Examination of the structure in the vicinity of the D residues suggests no apparent discomfort in accommodating them in an α_R conformation. The crystal structure of HBH-19 has provided a serendipitous observation of a mode of helix aggregation, which leads to formation of long hydrophobic channels that traverse the length of the crystal. The space between the two adjacent helical columns accommodates a series of five toluene molecules. The three D residues are not oriented close to the channel, suggesting that they do not influence the shape of the channel. It is possible that formation of the channel may be facilitated by aromatic interactions (28) between Phe-17 and the proximate toluene molecule.

We thank S. Aravinda for help with generating figures. This research was funded by a Program Support Grant from the Department of Biotechnology, Government of India (Bangalore, India). The work at the Naval Research Laboratory was supported by National Institutes of Health Grant GM30902 and the Office of Naval Research.

1. Venkataraman, J., Shankaramma, S. C. & Balaram, P. (2001) *Chem. Rev.* **101,** 3131–3152.
2. Karle, I. L. & Balaram, P. (1990) *Biochemistry* **29,** 6747–6756.
3. Balaram, P. (1992) *Curr. Opin. Struct. Biol.* **2,** 845–851.
4. Kaul, R & Balaram, P. (1999) *Bioorg. Med. Chem.* **7,** 105–117.
5. Silva, R. A. G., Yasui, S. C., Kubelka, J., Formaggio, F., Toniolo, C. & Keiderling, T. (2002) *Biopolymers* **65,** 229–243.
6. Awasthi, S. K., Raghothama, S. & Balaram, P. (1995) *Biochem. Biophys. Res. Commun.* **216,** 375–381.
7. Haque, T. S., Little, J. C. & Gellman, S. H. (1996) *J. Am. Chem. Soc.* **118,** 6975–6985.
8. Gellman, S. H. (1998) *Curr. Opin. Chem. Biol.* **2,** 717–725.
9. Karle, I. L., Awasthi, S. K. & Balaram, P. (1996) *Proc. Natl. Acad. Sci. USA.* **93,** 8189–8193.
10. Das, C., Naganagowda, G. A., Karle, I. L. & Balaram, P. (2001) *Biopolymers* **58,** 335–346.
11. Karle, I. L., Gopi, H. N. & Balaram, P. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 3716–3719.
12. Das, C., Raghothama, S. & Balaram, P. (1998)) *J. Am. Chem. Soc.* **120,** 5812–5813.
13. Das, C. & Balaram, P. (1999) *J. Chem. Soc. Chem. Commun.* 967–968.
14. Venkataraman, J., Naganagowda, G. A. & Balaram, P. (2002) *J. Am. Chem. Soc.* **124,** 4987–4994.
15. Karle, I. L., Das, C. & Balaram, P. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 3034–3037.
16. Aravinda, S., Shamala, N., Pramanik, A., Das, C. & Balaram, P. (2000) *Biochem. Biophys. Res. Commun.* **273,** 933–936.
17. Aravinda, S., Shamala, N., Rajkishore, R., Gopi, H. N. & Balaram, P. (2002) *Angew. Chem. Int. Ed.* **48,** 3863–3865.
18. Nordman, C. E. & Nakatsu, K. (1963) *J. Am. Chem. Soc.* **85,** 353–354.
19. Egert, E. & Sheldrick, G. M. (1985) *Acta Crystallogr. A* **41,** 262–268.
20. Karle, I. L., Das, C. & Balaram, P. (2001) *Biopolymers* **59,** 276–289.
21. Karle, J. (1968) *Acta Crystallogr. B* **24,** 182–186.
22. Ramachandran, G. S. & Sasisekharan, V. (1968) *Adv. Protein Chem.* **23,** 283–437.
23. Karle, I. L. (2001) *Biopolymers Peptide Sci.* **60,** 351–365.
24. Aravinda, S., Shamala, N., Bandyopadhyay, A. & Balaram, P. (2003) *J. Am. Chem. Soc.*, in press.
25. Aravinda, S., Shamala, N., Desiraju, S. & Balaram, P. (2002) *J. Chem. Soc. Chem. Commun.* 2454–2455.
26. Madan Babu, M., Singh, S. K. & Balaram, P. (2002) *J. Mol. Biol.* **322,** 871–880.
27. Fairman, R., Antony-Cahill, S. J. & DeGrado, W. F. (1992) *J. Am. Chem. Soc.* **114,** 5458–5459.
28. Aravinda, S., Shamala, N., Das, C., Sriranjini, A., Karle, I. L. & Balaram, P. (2003) *J. Am. Chem. Soc.* **125,** 465–475.