Nanotubes from hydrophobic dipeptides: pore size regulation through side chain substitution

Carl Henrik Görbitz

Department of Chemistry, University of Oslo, P.O. Box 1033, Blindern, N-0315, Oslo, Norway. E-mail: c.h.gorbitz@kjemi.uio.no; Fax: +47 2285 5441; Tel: +47 2285 5441

Received (in Montpellier, France) 27th May 2003, Accepted 27th August 2003
First published as an Advance Article on the web 30th October 2003

The crystal structures of L-Ala-L-Ile, L-Ile-L-Ala, L-Val-L-Val, L-Val-L-Ile and L-Ile-L-Val are presented. Together with L-Ala-L-Val and L-Val-L-Ala they constitute a unique group of seven isostructural dipeptides with hydrophobic pores and identical three-dimensional hydrogen bond networks. Hydrophobic side chains form the inner surfaces of channels parallel to the hexagonal axes in these structures. By changing the bulk of the side chains the van der Waals' diameter of the channels can be regulated within the interval 3.3 to 5.2 Å.

Introduction

In the crystal structures of peptides the *N*-terminal amino group can donate up to two of its H atoms to main chain carboxylate groups of neighboring peptide molecules. To avoid steric crowding the last amino H atom must find another type of acceptor (at least as long as no *C*- or *N*-terminal Gly residues are involved). Functional groups in the side chains or in co-crystallized solvent molecules normally serve this function. When hydrophobic moieties in the molecules are comparatively small, structures with three-dimensional hydrogen bond networks and small hydrophobic columns are observed, while larger hydrophobic groups usually lead to structures that are divided into hydrophobic and hydrophilic layers.[1] Several solvates of dipeptides with two hydrophobic residues belong to the second group,[2] but when no organic solvent molecules are present new crystal packing patterns are required in order to position three acceptors around each amino group.

An original way of solving the packing problem of hydrophobic dipeptides was provided by the crystal structure of L-Val-L-Ala (**VA**), the first-ever example of nanotube formation by such a small molecule.[3] The tubes are distinctly hydrophobic and are created within a three-dimensional hydrogen bond network. The crystal structure of the retroanalogue, L-Ala-L-Val (**AV**),[4] proved to be very similar, but with a special ability to adapt pore size and in particular pore shape to the nature of the absorbed solvent molecules. Dipeptides with only L-Leu and L-Phe residues, such as L-Phe-L-Phe (**FF**), also form nanotubes,[5] but in a completely different manner, giving one-dimensional hydrogen bond networks, in the shape of cylinders, with hydrophilic tube interiors.

Since the side chains of the dipeptides form the inner surface of the channels in the **VA** class, the inherent layout offers a unique possibility to adjust the nature of the hydrophobic groups without disrupting the hydrogen bond network of the overall structure. This article presents five new **VA** class structures and summarizes the results for the seven isostructural dipeptides.

Experimental

Crystal growth

L-Val-L-Val (**VV**) and L-Ala-L-Ile (**AI**) were obtained from Sigma, while L-Ile-L-Ala (**IA**), L-Val-L-Ile (**VI**) and L-Ile-L-Val (**IV**) were obtained from Bachem. Needle-shaped crystals with diameter >0.1 mm could be grown fairly easily for **VA** and **AV** by vapor diffusion of 2-propanol (and also other solvents) into aqueous solutions of the peptides.[3,4] Initial crystallization attempts for **AI** and **VV** under similar conditions yielded only useless, extremely thin needles. It was subsequently found that the initial solute concentration had a profound effect on crystal size; needles large enough for diffraction experiments were only formed when the concentration was just high enough to give precipitation. For both dipeptides this was achieved at very low concentrations in the range of 0.7–1.0 mg ml^{-1}, corresponding to 0.02–0.03 mg of peptide in 30 μl of water. An increase to 0.05 mg of peptide in 30 μl of water was enough to produce only the already familiar hair-like needles. Crystals of **IA**, **VI** and **IV** were grown by rapid evaporation at 60 °C, after it was discovered that these conditions yielded useful crystals for the very hydrophobic dipeptides in the **FF** class.[5]

X-Ray diffraction

The crystallographic data are reported in Table 1. Reflections were measured on a Siemens SMART 1000 CCD diffractometer using $Mo_{K\alpha}$ radiation ($\lambda = 0.71069$ Å). The data collections with SMART[6] included three sets of exposures with the detector set at $2\theta = 26°$ and crystal-to-detector distance of 5.0 cm. Data integration and cell refinement was carried out by SAINT,[7] while absorption corrections were carried out by SADABS.[8] SHELXTL[9] was used for structure solution by direct methods as well as subsequent full-matrix least-squares refinement on F^2. O, N, and C atoms were refined anisotropically, except for O atoms associated with absorbed solvent water molecules in **AI**, **VV**, **IV** and **VI**, which were refined isotropically. H atoms bonded to N and C^{α}_{1} (*i.e.*, those involved in hydrogen bonds) were also refined isotropically. Other H atoms were placed geometrically and constrained to keep all C–H or N–H distances as well as all C–C–H or C–N–H angles on any one C or N atom the same. Water H atoms could not be located in the electron density maps and were not included in the refinements.

Results and discussion

The crystal structures of **IA**, **AI**, **VV**, **VI** and **IV** are shown in Fig. 1 together with the structures of **VA** and **AV**.[3,4] All bond lengths and bond angles are normal.

Table 1 Crystal data for compounds in this study

Compound	**AI**	**IA**	**VV**	**VI**	**IV**
Chemical formula	$C_9H_{18}N_2O_3{\cdot}H_2O$	$C_9H_{18}N_2O_3$	$C_{10}H_{20}N_2O_3{\cdot}H_2O$	$C_{11}H_{22}N_2O_3{\cdot}0.22H_2O$	$C_{11}H_{22}N_2O_3{\cdot}0.21H_2O$
Formula weight	220.3	202.3	234.3	234.3	234.1
Crystal system	Hexagonal	Hexagonal	Hexagonal	Hexagonal	Hexagonal
Space group	$P6_1$	$P6_1$	$P6_1$	$P6_1$	$P6_1$
T/K	150	105	150	105	105
a/Å	14.2566(6)	14.372(2)	14.6469(7)	14.7403(14)	14.8003(6)
c/Å	10.1535(4)	9.8282(13)	10.2772(5)	10.3140(13)	10.2809(3)
V/Å^3	1787.22(13)	1758.1(4)	1909.40(16)	1940.8(4)	1950.31(13)
Z	6	6	6	6	6
μ/mm^{-1}	0.095	0.086	0.093	0.089	0.087
$N_{measured}$	5302	2213	4377	2717	2985
N_{unique}	1386	1373	1481	1501	1510
$N_{observed}$ $[F^2 > 2\sigma(F^2)]$	1360	1225	1432	1153	1267
R_{int}	0.040	0.041	0.055	0.100	0.047
R $[F^2 > 2\sigma(F^2)]$	0.033	0.028	0.039	0.055	0.033
$wR(F^2)$	0.076	0.070	0.082	0.112	0.082

Molecular conformations

The listing of torsion angles in Table 2 shows that the seven structures can be divided into three subgroups: (A) **AV** and **AI**; (B) **VA** and **IA**; (C) **VV**, **VI** and **IV**. Subgroups A and B can be distinguished mainly by their different φ_2 torsion angles, ~130° and ~152°, respectively, but also at the *C*-terminal carboxylate rotation (ψ_T). Both subgroups A and B have peptide bond rotations around 176°, while in subgroup C this rotation is approximately −177.5°. There are hardly any significant differences between structures in subgroup C, which can clearly be attributed to the tight packing of C^{β} and C^{γ} atoms in the **VV** structure. An extra methyl group may be added at the first residue, giving **IV**, or at the second residue, giving **VI**. In either case a *trans* χ^2 torsion angles makes the terminal methyl group point straight into the channel interior.

For *C*-terminal L-Val and L-Ile residues there is a shift from the favorable *gauche*−/*trans* side chain conformation for subgroup *A* to the more uncommon *gauche*+/*gauche*− conformation for subgroup *C*. Model building subsequently revealed that the *gauche*−/*trans* conformation was incompatible with side chains larger than methyl (L-Ala) or ethyl (L-2-aminobutyric acid; L-Abu) at the first residue. The *gauche*+/*gauche*− conformation is also observed for the *N*-terminal L-Val and L-Ile residues in both subgroup *B* and subgroup *C*.

Crystal structure and hydrogen bond types

The crystal structures of the **VA** class can be modelled as an array of dipeptide left-handed double helices with threefold screw symmetry, linked at the corners as shown in Scheme 1.

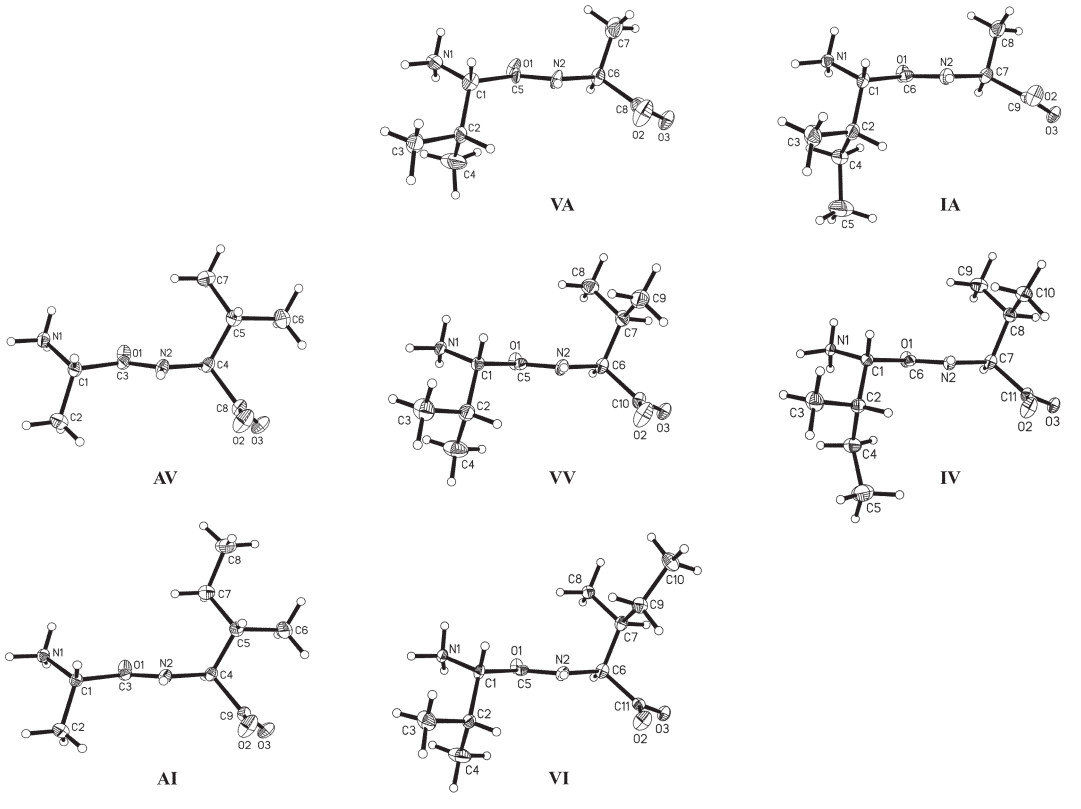


Fig. 1 The molecular structures of **AV** (from ref. 4), **VA** (from ref. 3), **AI**, **VV**, **VI**, **AI** and **IV** (present paper). Displacement ellipsoids are shown at the 50% probability level.

Table 2 Torsion angles (Å) for the compounds in this study

	AV[a]	VA[b]	AI	IA	VV	VI	IV
N_1–C_1^{α}–C'_1–N_2 (ψ_1)	151.01(6)	162.9(6)	153.25(16)	157.91(14)	166.06(19)	165.5(3)	170.68(19)
C_1^{α}–C'_1–N_2–C_2^{α} (ω_1)	175.98(6)	176.0(6)	176.80(15)	175.65(14)	−177.30(18)	−177.6(3)	−177.65(19)
C'_1–N_2–C_2^{α}–C'_2 (φ_2)	−130.56(6)	−150.6(6)	−128.92(17)	−153.76(15)	−140.0(2)	−139.5(3)	−139.9(2)
N_2–C_2^{α}–C'_2–O'_2 (ψ_T)[c]	−45.24(8)	−28.2(2)	−42.5(2)	−31.6(2)	−26.1(3)	−24.5(3)	−25.4(3)
N_1–C_1^{α}–C_1^{β}–$C_1^{\gamma 1}$ ($\chi_1^{1,1}$)	—	53.6(8)	—	48.44(19)	55.4(3)	55.3(4)	57.5(3)
N_1–C_1^{α}–C_1^{β}–$C_1^{\gamma 2}$ ($\chi_1^{1,2}$)	—	−71.5(8)	—	−77.48(18)	−69.7(3)	−70.0(4)	−69.2(2)
C_1^{α}–C_1^{β}–$C_1^{\gamma 1}$–$C_1^{\delta 1}$ (χ_1^2)	—	—	—	164.79(17)	—	—	157.3(2)
N_2–C_2^{α}–C_2^{β}–$C_2^{\gamma 1}$ ($\chi_2^{1,1}$)	−65.79(7)	—	−62.8(2)	—	60.1(3)	63.0(4)	60.2(3)
N_2–C_2^{α}–C_2^{β}–$C_2^{\gamma 2}$ ($\chi_2^{1,2}$)	171.72(6)	—	173.78(18)	—	−64.0(2)	−62.4(4)	−64.3(3)
C_2^{α}–C_2^{β}–$C_2^{\gamma 1}$–$C_2^{\delta 1}$ (χ_2^2)	—	—	173.18(19)	—	—	165.5(4)	—

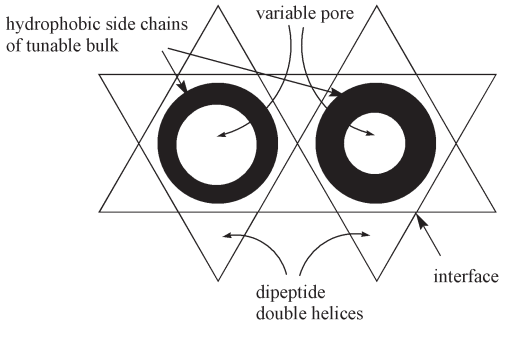
[a] ref. 4. [b] ref. 3. [c] Measured to the O atom giving the smallest positive or negative value.

Channels are created at the hexagonal axes in this pattern. **VV** is shown in Fig. 2 as a representative example of a **VA** class structure.

The detailed view of a pair of neighboring dipeptide helices in Fig. 3 shows all five types of hydrogen bonds in **VA** class structures as listed in Table 3. N1–H2···O3 hydrogen bonds generate the two head-to-tail dipeptide chains that constitute the individual strands of the double helices. The strands are connected by N1–H1···O1 hydrogen bonds across the threefold screw axis as seen in Fig. 2. At the interface between double helices N2–H4 and C1–H11 interact with the carboxylate group of the next peptide down the twofold screw axis, Fig. 3. This bidentate motif was found in only nine other structures in the Cambridge Structural Database (version 5.24, November 2002),[10] eight of them involving H^{α} atoms of Gly residues. By repetition the last hydrogen bond, N1–H3···O2, forms a head-to-tail chain in the shape of a large right-handed helix that defines the circumference of each pore.

Cell dimensions and hydrogen bond lengths

Although the hydrogen bond network is essentially the same in all seven structures of the **VA** class, strain obviously becomes more important as the cell dimensions increase from $a = b = 14.1635(8)$ Å and $c = 9.8702(5)$ Å for the acetonitrile solvate of **AV**[4] up to $a = b = 14.8003(6)$ Å and $c = 10.2909(3)$ Å for **IV**. Some of this strain is released by lengthening the hydrogen bonds in the structure. In particular, the N1–H1···O1 hydrogen bond is quite variable, as reflected by the large sample standard deviation in Table 3. The N1···O1 distance is linearly correlated with the length of the a and b axes (and also of the c axis) and varies from 2.668(1) Å for **AV** to 2.816(4) Å for **VI**. The H3···O2 distance of the N1–H3···O2 interaction is virtually constant, while the hydrogen bond angle and thus the N1···O2 distance vary with cell dimensions. Oddly enough, the geometry of the N1–H2···O3 bond is almost invariant in the crystallographic material; no correlation with the cell dimensions is found.



Scheme 1

It has previously been noted that the cell dimensions for **AV** increased when co-crystallized acetonitrile solvent was driven out by drying, meaning that solvent absorption has a contracting effect on the channels.[4] A related effect is probably responsible for the extraordinary decrease in cell dimensions from $a = b = 14.424(4)$ Å and $c = 9.996(6)$ Å for **VA** to $a = b = 14.372(2)$ Å and $c = 9.8282(13)$ Å for the larger **IA** molecule.

Channel dimensions

As the bulk of the side chains increases from **AV** and **VA** (four side chain C atoms) through **AI**, **IA** (five C atoms) and **VV** (six C atoms) to **VI** and **IV** (seven C atoms) the diameter of the channels at the hexagonal axes decreases. The top left image in Fig. 4 shows a space filling model of one of the widest channels, observed in the structure of **AV**.[4] A better impression of the pore is gained by the bottom view of a 3.0 Å thick slice of the structure, cut at an arbitrary point along the channel axis. A5.2 Å sphere can fit in the channel , but it is actually more rectangular in shape with dimensions of 6.0×4.5 Å. The **AI** channel is very similar to the **AV** channel, except that the two protrusions emanating down from the channel core in the slice view in Fig. 4 are filled. The **IA** channel is comparatively flat, with approximate dimensions of 5.0×2.5 Å. The channels of **VV** and **VI** are irregularly shaped, just like the **IV** channel shown in Fig. 4. The van der Waals' diameters can be roughly estimated to be 4.3, 3.3 and 3.3 Å, respectively. All these values should be compared with the 4.6 Å channel diameter of the naturally occurring peptide, antibiotic gramicidin A,[11] the 7–8 Å diameter of the cyclic D,L-octapeptides studied by Ghadiri *et al.*,[12] and the 6.0×2.5 Å rectangular to 10 Å circular channels for dipeptides in the **FF** class.[5,13]

After completing the studies of the seven dipeptides in the VA class, it was questioned whether even larger channels than

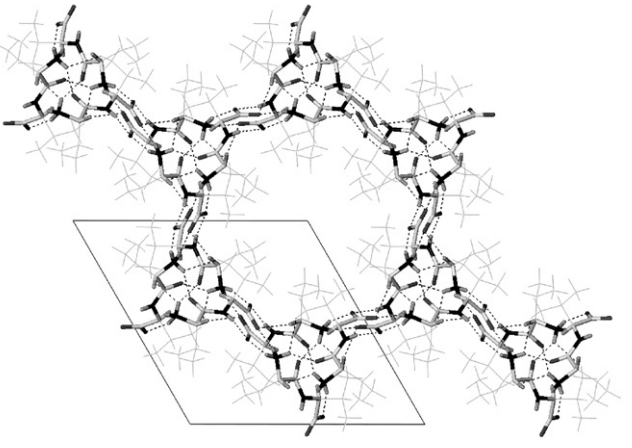


Fig. 2 The crystal structure of **VV** viewed along the c axis. Side chains are shown as thin lines.

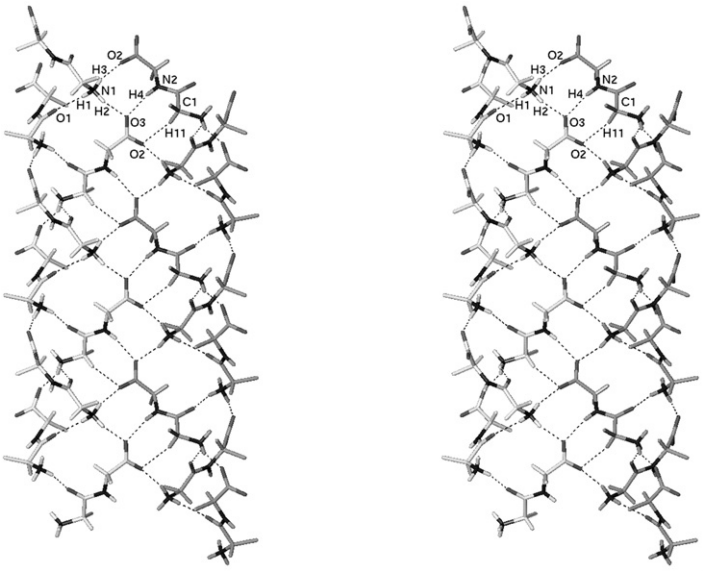


Fig. 3 Stereo view of the hydrogen bond pattern in the **VV** crystal. C and H atoms in one dipeptide double helix are colored white, similar atoms in the other helix are colored grey. For clarity only the C^{β} atoms of the L-Val side chains are shown.

those observed for **VA** and **AV** could be constructed. The structure of L-Ala-L-Ala (**AA**) is known to form a densely packed structure with tetragonal symmetry, while dipeptides with Gly residues, due to their flexibility both in terms of conformational properties and hydrogen bonding associations, seemed to be less likely building blocks for peptides forming **VA** class structures. Accordingly, the use of the non-standard amino acid L-Abu was investigated in the two dipeptides L-Ala-L-Abu (**AB**) and L-Abu-L-Ala (**BA**). The former was found to form a crystal structure very similar to **AA**, while **BA** formed a different type of structure without channels.[14] At the opposite end of the scale, model considerations indicated that, with very small adaptations, there was in fact room for the extra methyl groups in the structure of **IV** (or **VI**) required to create a hexagonal L-Ile-L-Ile (**II**) crystal with slightly narrower channels than observed for **IV** and **VI**. Crystallization and structure determination of **II**, however, revealed a monoclinic structure with tightly packed hydrophobic columns.[15] L-Phe-L-Ala, which also has eight side chain C atoms, forms a closely related structure.[16] In summary, the current experimental range for the van der Waals' diameter of the channels in the **VA** class is 3.3 to 5.2 Å, with **IV**/**VI** and **AV** representing the extreme limits.

It is noteworthy that the dipeptides presented in this paper incorporate, besides L-Ala, only residues with side chain branching at C^{β} (L-Val and L-Ile). Dipeptides with L-Leu or L-Phe residues, including L-Ala-L-Leu,[17] do not form **VA** class crystals, indicating that side chains with branching at C^{γ} cannot be efficiently packed inside the hydrophobic channels. To determine if crystallization in the **VA** class is compatible with dipeptides incorporating *unbranched* side chains (apart from the methyl group of L-Ala), crystallization and structure determinations have been carried out for L-Ala-L-Met (**AM**) and L-Met-L-Ala (**MA**). **AM** formed a monoclinic structure with a chain of water molecules filling a narrow channel parallel to a short crystallographic axis. The structure of **MA**, on the other hand, is very closely related to the **VA** class, but is more complex with $Z' = 7$. The **AM** and **MA** structures will be described in detail elsewhere.[18,19]

Table 3 Average geometric parameters (Å, °)[a] for the hydrogen bonds in the seven isostructural dipeptides of the **VA** class

D–H···A	H···A	D···A	∠D–H···A
N1–H1···O1	1.943(84)	2.738(65)	147.3(48)
N1–H2···O3	1.837(38)	2.716(17)	156.5(44)
N1–H3···O2	1.908(28)	2.780(49)	158.2(50)
N2–H4···O3	2.125(63)	2.971(32)	164.3(45)
C1–H11···O2	2.477(95)	3.418(80)	170.4(69)

[a] Sample standard deviation in parentheses.

Solvent structure

Crystal structure determinations of **VA** and **IA** showed no residual electron from absorbed solvent molecules inside the channels. **VA** crystals grown with alcohols as precipitating agents visibly lose solvent within seconds after being removed from the mother liquor.[3] For **IA** crystals solvent loss probably takes place within a matter of minutes. Drying of **AV** crystals, as discussed above, took place over 2 weeks at 60 °C. In fact, these dipeptides are among the first organic compounds to form crystals that survive complete removal of co-crystallized solvent. For **AI**, **VV**, **VI** and **IV** the number of positions for disordered solvent sites used in the refinement was 5, 7, 2 and 2, respectively. These positions, shown in Fig. 5, complement the pictures of channel shape provided by Fig. 4 and demonstrate clearly the larger diameter of the **AI** and **VV** pores compared to the pores of **VI** and **IV**.

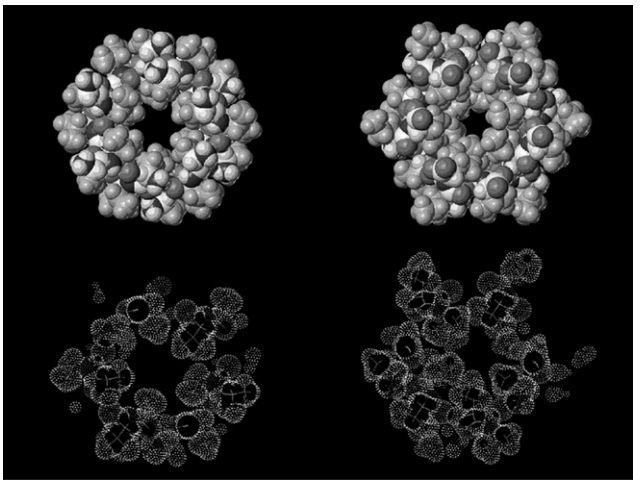


Fig. 4 The channels of **AV** (left) and **IV** (right). The top views show space filling models, the bottom views show 3.0 Å thick slices of the structures as stick drawings with a dotted van der Waals' surface.

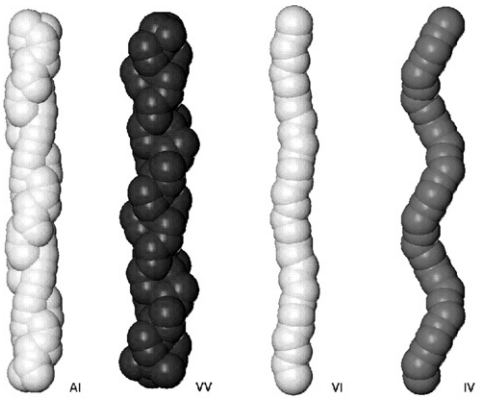


Fig. 5 Refined positions for disordered solvent atoms inside the channels of **AI**, **VV**, **VI** and **IV**.

Conclusion

Together with the previously presented structures of L-Val-L-Ala and L-Ala-L-Val, the structures of L-Ala-L-Ile, L-Ile-L-Ala, L-Val-L-Val, L-Val-L-Ile and L-Ile-L-Val demonstrate a new manner of regulating the pore size of persistent nanotubular crystal packing arrangements. Small pores with a diameter around 3.3 Å allow only relatively restricted transport of water molecules, while larger pores with diameters up to 5.2 Å allow the transport of metal ions and small organic solvent molecules. Future research into this family of compounds will focus on the storage properties and functionalization of the inner surfaces.

Acknowledgements

The purchase of the SMART diffractometer was made possible through support from The Research Council of Norway (NFR).

References

1 C. H. Görbitz and M. C. Etter, *Int. J. Pept. Protein Res.*, 1992, **39**, 93.
2 C. H. Görbitz, *Acta Chem. Scand.*, 1998, **52**, 1343; C. H. Görbitz, *Acta Crystallogr., Sect. C*, 1999, **55**, 2171.
3 C. H. Görbitz and E. Gundersen, *Acta Crystallogr., Sect. C*, 1996, **52**, 1764.
4 C. H. Görbitz, *Acta Crystallogr., Sect. B*, 2002, **58**, 849.
5 C. H. Görbitz, *Chem.-Eur. J.*, 2001, **7**, 5153.
6 *SMART, Version 5.054*, Bruker AXS Inc., Madison, WI, USA, 1998.
7 *SAINT, Version 6.01*, Bruker AXS Inc., Madison, WI, USA, 1998.
8 G. M. Sheldrick, *SADABS, Program for area detector adsorption correction*, Institute for Inorganic Chemistry, University of Göttingen, Germany, 1996.
9 G. M. Sheldrick, *SHELXTL, Version 5.10*, Bruker AXS Inc., Madison, WI, USA, 1998.
10 F. H. Allen and W. D. S. Motherwell, *Acta Crystallogr., Sect. B*, 2002, **58**, 407.
11 B. M. Burkhart, N. Li, D. A. Langs and W. L. Duax, *Proc. Natl. Acad. Sci. USA*, 1998, **95**, 12 950.
12 M. R. Ghadiri, J. R. Granja, R. A. Milligan, D. E. McRee and N. Khazanovich, *Nature (London)*, 1993, **366**, 324; J. D. Hartgerink, T. D. Clark and M. R. Ghadiri, *Chem.-Eur. J.*, 1998, **4**, 1367; D. T. Bong and M. R. Ghadiri, *Angew. Chem., Int. Ed.*, 2001, **40**, 2163; S. Fernandez-Lopez, H.-S. Kim, E. C. Choi, M. Delgado, J. R. Granja, A. Khasanov, K. Kraehenbuehl, G. Long, D. A. Weinberger, K. M. Wilcoxen and M. R. Ghadiri, *Nature (London)*, 2001, **412**, 452.
13 H. Birkedal, D. Schwarzenbach and P. Pattison, *Angew. Chem., Int. Ed.*, 2002, **41**, 754.
14 C. H. Görbitz, *Acta Crystallogr., Sect. C*, 2002, **58**, 533.
15 C. H. Görbitz, in preparation.
16 C. H. Görbitz, *Acta Crystallogr., Sect. C*, 2001, **57**, 575.
17 C. H. Görbitz, *Acta Crystallogr., Sect. C*, 1999, **55**(12) , cif-access *IUC9900149*.
18 C. H. Görbitz, *Acta Crystallogr., Sect. C*, 2003, **59**, submitted.
19 C. H. Görbitz, *Acta Crystallogr., Sect. C*, 2003, **59**, o589–o592.