Crystal structures and hydrogen bond analysis of five amino acid conjugates of terephthalic and benzene-1,2,3-tricarboxylic acids†

Anirban Karmakar,a,‡ Clive L. Oliver,b Ana E. Platero-Prats,c Elina Laurilaa and Lars Öhrströma

Four linear connecting amino acid derived ligands, 1–4, and one potentially three connecting, 5, were prepared by the reaction of the appropriate terephthaloyl dichloride or benzene-1,3,5-tricarbonyl trichloride with the methyl ester protected amino acid. Amino acids used here were alanine (1, 5), isoleucine (2), leucine (3) and valine (4). Crystalline forms of four amino acid substituted terephthalamides (2,2′-(terephthaloylbis(azanediyl))dipropionic acid dihydrate 1, 2,2′-(terephthaloylbis(azanediyl))bis(3-methylpentanoic acid) monohydrate 2, 2,2′-(terephthaloylbis(azanediyl))bis(4-methylpentanoic acid) dihydrate 3; 2,2′-(terephthaloylbis(azanediyl))bis(3-methylbutanoic acid) dihydrate 4) and one benzene-1,3,5-tricarboxamide molecule (2,2′,2″-((benzene-1,3,5-tricarbonyl)tris(azanediyl)) tripropionic acid hemihydrate 5) were characterised and the single crystal structures were solved. All the compounds form hydrogen bonded 2D and 3D nets. Linear connecting amino acid derivatives can be categorised into three groups depending on the hydrogen bond patterns and final structures. Compounds 1 and 2 form 3D structures but the final structure is different due to the different hydrogen bond synthons. Compounds 3 and 4 are isostructural and form 2D hydrogen bonded structures while 5 forms a hydrogen bonded puck-net. Intermolecular interactions have been analysed with Hirshfeld surfaces and graph set symbols.

Introduction

Understanding the driving forces of the formation of extended assemblies is important for designing molecule based materials, be these for potential electronic, magnetic or porous applications. Often, the idea is to preprogram the building blocks to interact with each other in a certain way to give larger entities with the desired location and geometric relationships between them. In other words, the interaction behaviour is based on the functional groups of the molecules.

Terephthalamides and benzene-1,3,5-tricarboxamides are examples of building blocks used in preparing extended assemblies. These molecules have a benzene ring core and usually contain flexible functional groups. In addition to stronger interactions such as coordination and hydrogen bonds, the role of non-covalent interactions is also important in the assembly of molecules. For example, the benzene ring in these systems offers the possibility of aromatic interactions.

The assembly process of benzene-1,3,5-tricarboxamide derivatives has been studied widely by using not only different techniques such as CD and UV-Vis but also theoretical methods. Terephthalamides and benzene-1,3,5-tricarboxamides have produced nanostructured materials, for instance nano-staircases and fibres are among the reported architectures. Spontaneous formation of hydrogels and their encapsulation and sorption properties of both di- and tri-substituted carboxamides have also been studied. In addition, terephthalamides and benzene-1,3,5-tricarboxamides derivatives have been reported to form liquid columnar crystal phases (LC), preorganized surfactants and multi-component organic cages within dynamic combinatorial libraries.

Introduction of amino acids into terephthalamide and benzene-1,3,5-tricarboxamides offers the possibility of producing enantiopure chiral building blocks and structures.
with potential applications in enantioselective synthesis, separation and detection. The amino acid part can also stabilize the assemblies via hydrogen bonding and moreover provide metal coordination sites.\textsuperscript{5,11,16,17,24–26}

The complexing abilities of amino acid substituted terephthalamide and benzene-1,3,5-tricarboxamides have been studied by preparing network structures with metals such as copper, cobalt, manganese, nickel, cadmium, zinc, lanthanoids, and calcium.\textsuperscript{7,11,17,24–31} These studies have focused on non-chiral glycine (–H) substituted ligands, but alanine (–CH\textsubscript{3}) and phenylalanine (–CH\textsubscript{2}(C\textsubscript{6}H\textsubscript{5})) examples are also known, although reported X-ray structures of such network compounds are rare.\textsuperscript{32,33}

The choice of the ligand and metal is essential in producing such metal–organic frameworks (MOFs)\textsuperscript{34} that can have application in gas sorption, catalysis and molecular recognition.\textsuperscript{27,35} These properties are incorporated into the metal during preparation using different building blocks. Therefore, the design of new materials with desired properties requires understanding and control of the interactions of the building blocks.

We report here the preparation and crystal structures of the amino acid derivatives 1–5 of terephthalamide and benzene-1,3,5-tricarboxamides (Fig. 1). These closely related molecules give structures with differences and similarities that are investigated not only by traditional means, but also graph set and Hirshfeld surface analyses.

**Experimental**

**Materials**

Amino acids L-phenylalanine (≥99%), l-leucine (≥99%), l-isoleucine (≥99%) and dl-valine (≥99%) were produced by Fluka. l-alanine methyl ester hydrochloride (≥99%) was purchased from Aldrich and other amino acid methyl esters were prepared according to the literature.\textsuperscript{40} Triethyl amine (99.5%), terephthaloyl dichloride (99%), benzene-1,3,5-tricarbonyl trichloride (98%) and 2-chloroform (98.6%) were manufactured by Aldrich. Anhydrus magnesium sulfate (98%) and sodium hydroxide (98%) were produced by Scharlau. Hydrochloric acid (≥37%), potassium bromide (99%) and methanol (99.7%) were manufactured by Sigma Aldrich. Dichloromethane (p.a.) was produced by Merck. Dimethylsulfoxide-d\textsubscript{6} (99.8%) was produced by Armar Chemicals.

**X-ray crystal structure determination**

The crystallographic details of compounds 1–5 are summarized in Table 1. The crystals of 1–5 were immersed in cryo-oil, mounted in a Nylon loop, and measured at a temperature of 293(2) K (1, 5) and 173(2) K using an Oxford Cryostream 600 (3) and 700 (2, 4). The X-ray diffraction data were collected using an Xcalibur Sapphire3 diffractometer (1, 5), a Nonius Kappa CCD diffractometer (3) and a Bruker D8 Apex II CCD diffractometer (2, 4) with Mo K\textsubscript{α} radiation (\(\lambda = 0.71073\) Å). For 1 and 5, the CrystAlisPro programme package was used for cell refinement and data reduction,\textsuperscript{41} whilst for 3, DENZO was used, and for 2 and 4 data reduction and cell refinement were performed using SAINT-Plus.\textsuperscript{42}

Space groups were determined from systematic absences using XPREP\textsuperscript{43} and confirmed using the program layer for 2–4.\textsuperscript{44} The structures were solved by direct methods using the SHELXS-97 (ref. 45) program. Semi-empirical or numerical absorption (SCALE3 ABSPACK (1), SADABS\textsuperscript{46} (2–5)) was applied to all of the data. Structural refinement was carried out using SHELXL-97 (1–5).\textsuperscript{47} The refinement procedure by full-matrix least-squares methods, based on \(R^2\) values against all reflections was performed using SHELXL-97, including the anisotropic displacement parameters for all non-H atoms. Hydrogen atoms were placed geometrically, except for O–H, N–H and water hydrogen atoms which were located in difference Fourier maps. All other hydrogen atoms were positioned geometrically and constrained to ride on their parent atoms \(U_{eq} = 1.2U_{eq}(\text{parent atom})\). The crystals of 5 have an asymmetric carbon that is disordered over two sites with equal occupancies. It also contains electron density, interpreted as residual water molecules, that cannot be refined properly. This electron density excess was cleaned using the SQUEEZE routine of PLATON.\textsuperscript{48} CCDC numbers 993036–993040.
Synthesis of dimethyl 2,2′- (terephthaloylbis(azanediyl))dipropionate

L-Alanine methyl ester hydrochloride (1.39 g, 10 mmol) and triethylamine (2.52 g, 25 mmol) were placed in a round bottom flask and then dissolved in dry 20 mL of dichloromethane. Terephthaloyl dichloride (1.11 g, 5 mmol) in 2 ml of dry triethylamine (2.52 g, 25 mmol) were placed in a round bottom flask and then the solution was acidified (pH = 2) with dilute hydrochloric acid solution. The obtained white solid product was filtered and washed with water. (0.9 g, 62%). (Found: C 48.61, H 5.81, N 8.11 calc. for C14H16N2O6 requires C 48.83, H 5.85, N 8.14) v_{\text{max}} (KBr)/cm^{-1} 3394 (bs), 3060 (w), 2982 (w), 2931 (w), 2542 (mb), 2019 (mb), 1732 (s), 1705 (s), 1639 (s), 1542 (s), 1500 (m), 1466 (m), 1406 (w), 1381 (w), 1345 (m), 1323 (m), 1289 (m), 1269 (s), 1240 (s), 1181 (s), 1039 (w), 1017 (w), 945 (w), 869 (s), 829 (m), 743 (s), 646 (m), 596 (m), 555 (w). δ_{1} (400 MHz; DMSO) 12.54 (s, 2H, −COOH), 8.77 (d, 2H, −CH), 7.95 (s, 4H, Bn), 4.42 (m, 2H, −CH), 1.39 (d, 6H, −CH3, J = 8 Hz). δ_{C} (100 MHz; DMSO) 174.5, 165.9, 136.7, 127.8, 52.7, 48.6, 18.6.

Synthesis of 2,2′-(terephthaloylbis(azanediyl))dipropionic acid (1)

Dimethyl 2,2′-(terephthaloylbis(azanediyl))dipropionate (1.68 g, 5 mmol) and sodium hydroxide (0.6 g, 15 mmol) were dissolved in 20 mL of methanol : water mixture (4:1). The reaction mixture was stirred for eight hours at room temperature. Then the solvent was evaporated under reduced pressure. The solid product was dissolved in 10 mL of water and the solution was acidified (pH = 2) with dilute hydrochloric acid solution. The obtained white solid product was filtered and washed with water. (0.9 g, 62%). (Found: C 48.61, H 5.81, N 8.11 calc. for C14H16N2O6 requires C 48.83, H 5.85, N 8.14) ν_{\text{max}} (KBr)/cm^{-1} 3394 (bs), 3060 (w), 2982 (w), 2931 (w), 2542 (mb), 2019 (mb), 1732 (s), 1705 (s), 1639 (s), 1542 (s), 1500 (m), 1466 (m), 1406 (w), 1381 (w), 1345 (m), 1323 (m), 1289 (m), 1269 (s), 1240 (s), 1181 (s), 1039 (w), 1017 (w), 945 (w), 869 (s), 829 (m), 743 (s), 646 (m), 596 (m), 555 (w). δ_{1} (400 MHz; DMSO) 12.54 (s, 2H, −COOH), 8.77 (d, 2H, −CH), 7.95 (s, 4H, Bn), 4.42 (m, 2H, −CH), 1.39 (d, 6H, −CH3, J = 8 Hz). δ_{C} (100 MHz; DMSO) 174.5, 165.9, 136.7, 127.8, 52.7, 48.6, 18.6.
Synthesis of 2,2′-(terephthaloylbis(azonediyl))bis(3-methylpentanoic acid) (2)

2,2′-(Terephthaloylbis(azonediyl))bis(3-methylpentanoic acid) was synthesized following the synthesis of 2,2′-(terephthaloylbis(azonediyl))dipropionic acid, but using dimethyl 2,2′-(terephthaloylbis(azonediyl))dipropionate (2.1 g, 5 mmol) instead of dimethyl 2,2′-(terephthaloylbis(azonediyl))dipropionate. (1.5 g, 77%). [Found: C 55.67, H 7.25, N 6.38 calc. for C_{20}H_{24}N_{2}O_{4} 56.06, H 7.53, N 6.54 νmax (KBr/cm\(^{-1}\) 3367 (bs), 3060 (w), 2962 (s), 2575 (mb), 1728 (s), 1632 (s), 1545 (s), 1499 (m), 1471 (w), 1451 (w), 1387 (w), 1338 (m), 1270 (s), 1252 (s), 1180 (m), 1169 (m), 1089 (w), 1017 (w), 967 (w), 870 (s), 854 (s), 736 (s), 600 (m). δH (400 MHz; DMSO) 12.61 (bs, 2H, -COOH), 8.71 (d, 2H, -NH, J = 8Hz), 7.95 (s, 4H, Bn), 4.44 (m, 2H, -CH), 1.80-1.55 (m, 6H, -CH2 and -CH), 0.91 (d, 6H, -CH3, J = 8Hz), 0.86 (d, 6H, -CH3, J = 8Hz). δC (100 MHz; DMSO) 174.5, 166.2, 136.7, 127.8, 51.4, 24.9, 23.4, 21.6.

Synthesis of dimethyl 2,2′-(terephthaloylbis(azonediyl))bis(4-methylpentanoate)

Dimethyl 2,2′-(terephthaloylbis(azonediyl))bis(4-methylpentanoate) was synthesized following the synthesis of dimethyl 2,2′-(terephthaloylbis(azonediyl))dipropionate, but using L-isoleucine methyl ester hydrochloride (1.8 g, 10 mmol) instead of L-alanine methyl ester hydrochloride. (2.7 mg, 64%). νmax (KBr/cm\(^{-1}\) 3345 (s), 3051 (w), 2966 (s), 2877 (m), 1746 (s), 1638 (s), 1547 (s), 1503 (s), 1459 (m), 1434 (m), 1372 (m), 1342 (s), 1298 (s), 1254 (m), 1201 (s), 1155 (s), 1117 (w), 1089 (w), 1008 (m), 964 (w), 878 (m), 847 (m), 769 (m), 731 (s), 679 (w), 619 (bs), 589 (m), 505 (w). δH (400 MHz; CDCl\(_3\)) 7.86 (s, 4H, Bn), 6.72 (d, 2H, -NH, J = 8Hz), 4.82 (q, 2H, -CH, J = 8Hz), 3.78 (s, 6H, -COOC(CH3)\(_2\)), 2.03 (m, 2H, -CH), 1.53 (m, 2H, -CH2), 1.25 (m, 2H, -CH3), 0.93 (m, 12H, -CH3). δC (100 MHz; CDCl\(_3\)) 172.5, 166.1, 136.9, 127.4, 56.9, 52.3, 38.2, 25.4, 15.5, 11.6. The NMR shifts were in agreement with the literature.16

Synthesis of 2,2′-(terephthaloylbis(azonediyl))bis(4-methylpentanoic acid) (3)

2,2′-(Terephthaloylbis(azonediyl))bis(4-methylpentanoic acid) was synthesized following the synthesis of 2,2′-(terephthaloylbis(azonediyl))dipropionic acid, but using dimethyl 2,2′-(terephthaloylbis(azonediyl))bis(4-methylpentanoate), instead of dimethyl 2,2′-(terephthaloylbis(azonediyl))dipropionate. (1.6 mg, 83%). [Found: C 58.47, H 7.28, N 6.87 calc. for C_{20}H_{28}N_{2}O_{6} 59.53, H 6.64, N 7.69 νmax (KBr/cm\(^{-1}\) 3444 (bs), 3325 (bs), 2967 (s), 2936 (w), 2573 (bm), 1731 (s), 1629 (s), 1537 (s), 1499 (m), 1470 (w), 1394 (w), 1375 (w), 1368 (m), 1232 (s), 1253 (w), 1236 (s), 1193 (s), 1161 (s), 1033 (s), 1017 (w), 977 (m), 915 (m), 865 (s), 815 (s), 733 (s), 636 (m), 576 (w). δH (400 MHz; DMSO) 8.56 (d, 2H, -NH, J = 8Hz), 7.94 (s, 4H, Bn), 4.28 (tr, 2H, -CH, J = 8Hz). δH (100 MHz; DMSO) 173.5, 166.7, 136.9, 127.9, 58.9, 29.9, 19.8, 19.3.

Synthesis of dimethyl 2,2′-(terephthaloylbis(azonediyl))bis(3-methylbutanoate)

Dimethyl 2,2′-(terephthaloylbis(azonediyl))bis(3-methylbutanoate) was synthesized following the synthesis of dimethyl 2,2′-(terephthaloylbis(azonediyl))dipropionate, but using L-valine methyl ester hydrochloride (1.7 g, 10 mmol) instead of L-alanine methyl ester hydrochloride. (2.9 g, 75%). νmax (KBr/cm\(^{-1}\) 3298 (s), 3031 (w), 2965 (s), 2875 (m), 1744 (s), 1642 (s), 1537 (s), 1501 (s), 1468 (m), 1435 (m), 1347 (s), 1321 (s), 1295 (m), 1261 (s), 1201 (s), 1121 (w), 1070 (m), 1020 (s), 1002 (m), 923 (w), 862 (s), 826 (w), 737 (m), 681 (bs), 577 (w). δH (400 MHz; CDCl\(_3\)) 7.86 (d, 4H, Bn, J = 8Hz), 6.66 (2H, -NH), 4.77 (s, 2H, -CH), 3.77 (s, 6H, -COOC(CH3)\(_2\)), 2.27 (m, 2H, -CH), 0.99 (m, 12H, -CH3). δH (100 MHz; CDCl\(_3\)) 172.5, 166.2, 136.9, 127.4, 57.5, 52.4, 31.6, 18.9, 17.9.

Synthesis of 2,2′-(terephthaloylbis(azonediyl))bis(3-methylbutanoic acid) (4)

2,2′-(Terephthaloylbis(azonediyl))bis(3-methylbutanoic acid) was synthesized following the synthesis of 2,2′-(terephthaloylbis(azonediyl))dipropionic acid, but using dimethyl 2,2′-(terephthaloylbis(azonediyl))bis(3-methylbutanoate) (1.9 g, 5 mmol) instead of dimethyl 2,2′-(terephthaloylbis(azonediyl))dipropionate. (1.5 g, 84%). [Found: C 59.15, H 6.49, N 7.55 calc. for C_{20}H_{24}N_{2}O_{4} 59.53, H 6.64, N 7.69 νmax (KBr/cm\(^{-1}\) 3444 (bs), 3325 (bs), 2967 (s), 2936 (w), 2573 (bm), 1731 (s), 1629 (s), 1537 (s), 1499 (m), 1470 (w), 1394 (w), 1375 (w), 1368 (m), 1232 (s), 1253 (w), 1236 (s), 1193 (s), 1161 (s), 1033 (s), 1017 (w), 977 (m), 915 (m), 865 (s), 815 (s), 733 (s), 636 (m), 576 (w). δH (400 MHz; DMSO) 8.56 (d, 2H, -NH, J = 8Hz), 7.94 (s, 4H, Bn), 4.28 (tr, 2H, -CH, J = 8Hz). δH (100 MHz; DMSO) 173.5, 166.7, 136.9, 127.9, 58.9, 29.9, 19.8, 19.3.
Synthesis of 2,2’,2”-(benzene-1,3,5-tricarboxyl)tris(azanediyl)tripropionate (5)

Trimethyl 2,2’,2”-(benzene-1,3,5-tricarboxyl)tris(azanediyl)tripropionate (2.3 g, 5 mmol) and sodium hydroxide (0.6 g, 15 mmol) were dissolved in 20 mL of methanol-water mixture (4:1). The reaction mixture was stirred for eight hours at room temperature. Then the solvent was evaporated under reduced pressure. The solid product was dissolved in 10 mL of water and the solution was acidified (pH = 2) with dilute hydrochloric acid solution. The obtained white solid product was filtered and washed with water. (1.9 mg, 91%). (Found: C 50.00, H 5.13, N 9.72) δ<sub>max</sub> (KBr)/cm<sup>-1</sup> 3393 (vs), 3233 (m), 3066 (s), 2992 (s), 2942 (s), 1722 (s), 1644 (s), 1544 (s), 1458 (s), 1412 (s), 1382 (w), 1289 (w), 1240 (s), 1167 (s), 1053 (w), 966 (w), 922 (w), 836 (w), 741 (m), 692 (w), 660 (s). δ<sub>H</sub> (400 MHz; DMSO) 8.20 (s, 3H, Bn), 7.47 (d, 9H, -CH<sub>3</sub>), 1.43 (d, 9H, -CH<sub>3</sub>). δ<sub>δ</sub>(100 MHz; DMSO) 176.5, 168.3, 134.1, 129.4, 49.4, 15.9. The NMR shifts are in agreement with the literature.24

Results

Synthesis and crystallization

Compounds 1–5 were prepared in two steps at room temperature. In the first step, the protected amino acid methyl ester is introduced to terephthaloyl dichloride and in the next step the amino acid is deprotected (Schemes 1 and 2). The compounds were recrystallized from methanol (1, 2), H<sub>2</sub>O (3, 4) or from D<sub>2</sub>O (5) for single crystal X-ray diffraction studies. 1–3 were obtained enantiomerically pure while racemic valine was used to prepare 4 and 5 was obtained as a racemate.

Structure descriptions

A complete set of ORTEP and packing diagrams can be found in the ESI.†

**Compound 1 dihydrate.** The molecules of compound 1 are connected to each other via hydrogen bonding of N−H⋯O (N(1)−H(1A)⋯O(2) #1 3.132 Å, #1 −x + 1, y, z −1/2, −z + 1/2)

**Compound 2 monohydrate.** Compound 2 also forms a 1D chain structure along the b-axis via hydrogen bonding between the carboxylic group and the ketone group (O(6)−H(6A)⋯O(1) #1 2.579(3) Å, #1 −x + 1, y, z −1/2, −z + 3/2). Lattice water links the chains together, forming a hydrogen bonded 3D net. Lattice water interacts with the amine groups, the ketone groups and the carboxylic groups, and acts as both hydrogen bond donors and acceptors (O(3)−H(3A)⋯O(1) #2 2.618(4) Å, O(W1)−H(2W1)⋯O(2) #3 2.748(3) Å, O(W1)−H(W1)⋯O(4) #4 2.676 (3) Å, N(1)−H(1A)⋯O(W1) #2 2.948(4) Å, #2 x −1, y, z; #3 x + 1, y, −z + 2, −z + 1 + 1/2, −y + 3/2, −z + 1).

**Compound 2 dihydrate.** Compound 2 dihydrate also forms a 1D chain along the b-axis via hydrogen bonding between the carboxylic group and the ketone group (O(6)−H(6A)⋯O(1) #1 2.579(3) Å, #1 −x + 1, y, z −1/2, −z + 3/2). Lattice water links the chains together, forming a hydrogen bonded 3D net. Lattice water interacts with the amine groups, the ketone groups and the carboxylic groups, and acts as both hydrogen bond donors and acceptors (O(3)−H(3A)⋯O(1) #2 2.618(4) Å, O(W1)−H(2W1)⋯O(2) #3 2.748(3) Å, O(W1)−H(W1)⋯O(4) #4 2.676 (3) Å, N(1)−H(1A)⋯O(W1) #2 2.948(4) Å, #2 x −1, y, z; #3 x + 1, y, −z + 2, −z + 1 + 1/2, −y + 3/2, −z + 1).
structure are disordered over two positions and it is therefore evident that the alanine units became racemic during preparation. This compound forms dimeric units via a completely symmetric π–π stacking (3.5 Å). In addition to aromatic interaction, the building blocks of the dimeric unit are hydrogen bonded, as shown in Fig. 2 {N(2)−H(2A)⋯O(2)#1 2.888(4) Å, #1 1/3 + x − y, −1/3 + x, 2/3 − z}. The hydrogen bonding also causes the substituents of the benzene rings to have the same orientation within the dimers. The substituent chains of two stacked benzene rings are anti-eclipsed/staggered. These units further connect via hydrogen bonding of O–H⋯O [O(3)−H(3A)⋯O(1)#2 2.641(4) Å, #2 1/3 − x + y, 2/3 − x, −1/3 + z] to a 3D network.

Discussion

What we want to do here is to perform a structure comparison between our compounds using first the “big picture” methods graph set and Hirshfeld surface analyses and then, when needed, look into the structural details. For more general applicability we also include in our discussion of the compounds the earlier published structures of the glycine terephthalamide derivative 6 (ref. 32) and the glycine benzene-1,3,5-tricarboxamides derivative 7.33

General hydrogen bond analysis by the graph set method

Graph set analysis is a general method to investigate the recurring hydrogen bond motifs and hydrogen bond patterns in discrete and extended assemblies.35,49,50 It does not discern the different atom types such as N or O, but once the criteria for hydrogen bonding are set, the analysis will generate the motifs such as a chains (C) and rings (R) with the number of atoms directly involved in the bonding as well as the number of atoms linking these (in, for example, a ring). Table 2 shows the graph set symbols for 1–7 as calculated using Mercury. We chose to show only the ring-forming motifs as we believe these to be fairly distinct.

We note first that there are no directly hydrogen bonded dimers in the 1–4 compounds, because that would have implied a 2,2 designation, i.e. two donor atoms D and two acceptor atoms H as in the acetic acid dimer with symbol R2,2(8)). Compound 1 forms hydrogen bonds directly to another building block with a loop-back via hydrogen bonded water, the R3,3 assignment. This does not happen for compound 2 that has chains of building block–water–building block instead.

Then we can also see a number of recurring motifs in the different structures, the shorter circuits, i.e. the number in parenthesis, likely the most significant. Thus, the R4,4(13) motif is recurring in 1–4 but is absent in 6 and the R2,2(16) motif is found in both 5 and 7. An illustration of some of these motifs can be found in the ESI.†

Comparison of compounds 1–4, and 6

Before we examine in more detail the structures of 1–4 and 6, a few remarks are needed. These compounds, differing only in the amino acid chain part (R, Fig. 1), all contain three distinct entities, and each of them has different intermolecular “preferences”. Thus we have a flat benzene ring that we often find either π-stacked or with C–H σ–π interactions. Then we have aliphatic groups (except for 6) that will normally prefer to be in the presence of each other, and finally the polar hydrogen bond donors and acceptors that will generally “want” to have all free electron pairs and acidic hydrogens in bonding interactions.

A convenient way to get an overall picture is to use Hirshfeld surfaces. The Hirshfeld surface reveals the intermolecular interactions, and even the weaker secondary interactions such as C–H⋯π, C⋯H and H⋯H contacts.51–53 This surface represents the volume where, inside, electron density is dominated by the atoms in the molecule in focus, while outside, electron density is dominated by all other components of the crystal. Each point on the surface has one closest atom inside and one closest atom outside the surface. We can therefore determine what percentage of the surface that originates from O⋯H, C⋯C or any other close interaction. This could be a crude first way of comparing similar structures and a diagram of these percentages for compounds 1–4§ and 6–7 are found in Fig. 3.

† The disorder of 5 makes a Hirshfeld surface difficult to compare and interpret.

Table 2 Graph set symbols for compounds 1–5

<table>
<thead>
<tr>
<th>Comp.</th>
<th>Graph set symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disubst.</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>R3,3(15); R4,4(13); R4,4(18); R3,4(17); R4,4(21)</td>
</tr>
<tr>
<td>2</td>
<td>R4,4(13); R4,4(18); R3,4(17); R4,4(21)</td>
</tr>
<tr>
<td>3</td>
<td>R4,4(13); R4,4(18); R3,4(17); R4,4(18); R4,4(21); R3,4(23)</td>
</tr>
<tr>
<td>4</td>
<td>R4,4(13); R4,4(18); R3,4(17); R4,4(18); R4,4(21); R3,4(23)</td>
</tr>
<tr>
<td>5</td>
<td>R4,4(18); R2,4(28); R3,4(17); R4,4(21); R3,4(23)</td>
</tr>
<tr>
<td>Trisubst.</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>R2,2(16); R2,2(22); R3,4(23);</td>
</tr>
<tr>
<td>7</td>
<td>R2,2(16); R2,2(22); R4,4(18); R4,4(20)</td>
</tr>
</tbody>
</table>

Fig. 2 Schematic presentation of dimeric units (blue and black) of 5 forming via π–π stacking and hydrogen bonding (dotted line). Graph set symbol R2,2(16).
We immediately observe that the derivatives with longer aliphatic chains, 2–4, have a higher proportion of H⋯H interactions (blue), which is reasonable as these groups, as stated above, will tend to interact with each other, but also, simply, because of the higher percentage of hydrogen in these molecules.\footnote{Possible ways to account for the difference in composition are to weigh the contributions by atom percentages or to calculate interactions only for the part that is common to all molecules.}

Another notable feature is the higher proportion of C⋯O (purple) in the glycine derivative 6 compared especially to 1–2. This stems from the interaction of the carbonyl groups with the flat side of the benzene ring in 6, an arrangement we suggest arises because 6 lacks the CH₃ groups that tend to come close to the aromatic ring in the alanine derivative 1.

One thing that is not picked up so clearly in these plots is the difference between on one side compounds 3 and 4 that form the same hydrogen bonded assemblies in unit cells that are close to identical, and compound 2.

Both structures 3 and 4 contain two lattice water molecules that connect the building blocks to form a 2D hydrogen bonded net. The interactions between the building blocks are all lattice water based. In other words there are no direct intermolecular hydrogen bonds between the actual building blocks. This is in contrast to 2 that has direct hydrogen bonds between the building blocks and only one water molecule.

It is worth noting here that 3 and 4 pack with the aliphatic groups facing each other, as shown in Fig. 4, just as intuition would suggest. This leaves the aromatic ring exposed and instead of CH₃ interactions as in 1 it adopts the carbonyl interaction as in 6.

Another property of 3 and 4 that has some implications is chirality. As racemic valine was used, 4 is a racemic compound, the crystals obtained are of the R,S molecule, and a non-chirodescriptive space group would have been expected. However, solutions in P2₁/c give a significantly higher R-factor than when solved and refined in the space group P2₂₁, and the apparent centre of symmetry is only pseudo. The abnormal anisotropic displacement parameters of the C8-carbon (see Fig. S18\textsuperscript{†}) could be due to some minor R/S disorder present at the chiral carbon, however this could not be satisfactorily modelled.

On the other hand, 3 is enantiomerically pure (some indication of very minor isomerisation can be detected and again, solutions in P2₁/c give a significantly higher R-factor) and having the chirodescriptive space group P2₁ is logical. Somehow there is thus a chiral arrangement of 4. This riddle is solved when one realizes first that the R-side is always facing another R-side and *vice-versa* and that the aliphatic groups meet around a screw axis parallel to the b-axis, thus arranging these groups in a chiral, helical way.

This observation then begs the question why compound 2 does not form the same kind of structure, being an isomer of 3. It is hard to fathom why the same packing and chiral arrangement would not be possible, but we refrain from further speculation, as this may simply be a question of crystallization conditions.

**Comparing compound 5 and compound 7**

The local molecular structure of 5 resembles that of the earlier published glycine derivative 2,2',2″-((benzene-1,3,5-
In contrast to the glycine derivative 7, in compound 5 these dimers are connected to each other via hydrogen bonds and form a 3D pcu-net. Hirshfeld surface analyses of compounds 1–7 indicate that some of the differences between these systems and graph set symbols are also helpful in categorising and discovering differences and similarities.

From a synthesis point of view, for the further use of these ligands in building MOFs, it will be important to have complete control of the stereochemistry and less harsh conditions may be needed.

Acknowledgements

LÖ, EL and AK thank the Swedish Research Council, LÖ and CLO thank the Swedish Research Links program for Southern Africa, AK also expresses his gratitude to the Foundation for Science and Technology (FCT), Portugal, for post-doctoral fellowship (SFRH/BPD/76192/2011), AEEP was supported by the CMCSU project, the Knut and Alice Wallenberg Foundation and the MATsynCELL project, Röntgen–Ångström Cluster.

Notes and references
