

Design and Synthesis of Microporous Dipeptide Structures and Guanidinium-carboxylate-based Organic Supramolecular Materials

Dissertation Submitted and Presented for the Degree of

Philosophiae Doctor

By

Vitthal Narayan Yadav

2013



Department of chemistry

Faculty of Mathematics and Natural Sciences

University of Oslo, Norway

© **Vitthal Narayan Yadav, 2013**

*Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 1432*

ISSN 1501-7710

All rights reserved. No part of this publication may be reproduced or transmitted, in any form or by any means, without permission.

Cover: Inger Sandved Anfinsen.
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Akademia Publishing.
The thesis is produced by Akademia Publishing merely in connection with the thesis defence. Kindly direct all inquiries regarding the thesis to the copyright holder or the unit which grants the doctorate.

Dedicated To

....My Family

Acknowledgement

By immense heart, I thank my principal supervisor Prof. Carl Henrik Görbitz for providing me such a wonderful opportunity, work freedom, encouragements and excellent mentorship during my research at University of Oslo, Norway (2009-2013). I also extend my gratitude and regards to my co-supervisor Asso. Prof. Tore-Bonge Hansen for his helpful suggestions, guidance and co-operation.

The members of the Department of Chemistry have contributed immeasurably to my personal and professional time at the University of Oslo. The following group members have been a source of friendships as well as good advice and collaboration: Åsmund Kaupang, Jens H. A., Tore Erik, Kim, Christian, Masood Kaboli, Martin H., Matthew, Peter, Marianne, Eirin, Marte, Fiona, Mel Siah, Vladimiro, Magnus, Elahe Jafari, Michelle Hanif, Abhijit Khoje, Naresh, Sachin Chavan and Tushar Mahajan.

My special thanks to Frode Rise and Dirk Petersen for their NMR spectrometer related help, and Osamu Sekiguchi for MS measurements. I also would like to acknowledge David Wragg for his frequent X-ray diffractometer associated assistance.

I sincerely thank all friends and families in Oslo with whom I shared memorable moments, celebrated joyful events and offered pleasant and warm hospitality in freezing Norway.

Nevertheless, this tough journey could have not been completed without sacrifice, dedication, profound and meaningful love and moral support from my family. I am gratefully indebted to my grandparents, parents, siblings, extended families and friends. I apologize to Sunita (wife), my little angel Anushka (daughter) and all family members for not being with during this period. I whole heartedly thank you all for invaluable contributions in shaping and guiding my path of education, career and Life.....!



Sincerely Yours

Vitthal Narayan Yadav

(Oslo, Norway, 2013)

Abstract

The basis of supramolecular chemistry is a detailed knowledge of fundamental molecular properties and non-covalent interactions, which is dedicated to the preparation of novel structures as functional materials. Here we discuss the design and synthesis of such new molecular self-assemblies. Based on the structures and properties of these molecules, this dissertation describes the two types of supramolecular structures in two different chapters.

Chapter 1 deals with the design, preparation, characterization and applications of chiral, bio-degradable, guest-specific and environment-friendly nanoporous crystalline dipeptides. In the hydrogen-bonded networks, dipeptides with hydrophobic L-amino acid residues are known to form pores of different diameters, ranging from 3-10 Å. The hydrophobic bulk and orientations of the side chains of these dipeptides provide further scope for structure-based modifications to fine-tune their pore dimensions. Thus, towards the liberation of space occupied by bulky hydrophobic terminals of the amino acid side chains from the periphery of channels, a series of dipeptides from non-proteinogenic and proteinogenic amino acids have been synthesized, crystallized and analyzed by single crystal X-ray diffraction methods. The majority of these dipeptides were obtained as porous structures. As nanoporous materials, a few of these dipeptides were studied for CO₂ and methane gas absorption and selectivity.

Chapter 2 demonstrates the supramolecular synthesis of charge-assisted complexes from binary acid-base components. The well-directed hydrogen bond formation between acids and bases is a very useful tool in designing new supramolecular assemblies. Here we pursued an approach to use 1,5,7 triazabicyclo[4.4.0]-dec-5-ene, a guanidine derivative, with di- or monocarboxylic acids to generate guanidinium-carboxylate complexes. These molecular structures were designed to explore the effects of limited hydrogen-bond forming ability of guanidinium moiety and carboxylate group, and to check the propensity of guanidinium-carboxylate complexes for the inclusion of guest molecules, for instance the water molecules. In fact, these complexes in crystals have interacted with the water molecules or carboxyl groups in the absence of any other potential donors and formed different water networks and 1D molecular pattern, which as organic materials may find various future applications as proton conductors, selective ion channels and gelators etc.

Table of Contents

Acknowledgement	iii
Abstract	iv
Table of contents	v
List of abbreviations	vii
List of Publications.....	viii
Introduction: Supramolecular Chemistry	1
Natural Processes and Supramolecular Chemistry	1
Specific Intermolecular Interactions and Crystal engineering.....	3
Reference List.....	5
Chapter 1: Microporous Dipeptide Structures	9
1.1 Introduction: Amino acids, Peptides and Proteins.....	9
1.1.2 Small Peptides (Natural vs Synthetic Peptides)	10
1.1.3 Nanoporous Self-assembly of Dipeptides and Material Science.....	11
1.1.3.1 Self-Assembly of the Val-Ala Class Dipeptides	12
1.1.3.2 Self-Assembly of Leu-Ser Dipeptide	14
1.1.3.3 Self-Assembly of the Phe-Phe Class Dipeptides	15
1.1.3.4 Side Chain Size, Orientations and Channel Property	15
1.1.4 Rational Strategies for Channel Size Modifications.....	16
1.1.5 Synthetic Targets as Microporous Dipeptides (Summary of Designed Molecules).....	19
1.1.6 Synthesis, Crystallization and Structure Determination of Dipeptides	21
1.2 Results and Discussion: Crystal Structures of Obtained Dipeptides	22
1.2.1 Val-Ala Type Structures.....	23
1.2.2 Structures of Abu-Ser and Pro-Ser	25
1.2.3 Structure of Nva-Phe	25
1.3 Conclusions	25
Reference List.....	26
1.4 Experimental and Characterization Data	30

Chapter 2: Guanidinium-carboxylate Based Supramolecular Complexes.....	44
2.1 The Guanidine Subgroup and Molecular Recognition	44
2.2 Guanidine in Crystal Engineering	45
2.3 Guanidinium-carboxylate Complexes: CSD Survey and Specific Interactions	46
2.4 Effective Strategy to Utilize the Guanidinium Subgroup in Crystalline Materials	48
2.4.1 Selection of Guanidine Derivative.....	48
2.4.2 Di-guanidinium Ligand: Two Better than One.....	50
2.4.3 Organic Frameworks from I	51
2.5 Guanidine Derivative (TBD) and Dicarboxylic Acids in Crystal Engineering.....	53
2.6 Experimental Section	53
2.6.1. Materials and Crystals Synthesis.....	53
2.6.2. Dicarboxylic Acids and TBD Complexes	54
2.6.2.1 Bi-aromatic Dicarboxylic Acids and TBD	54
2.6.2.2 Aliphatic Saturated/Unsaturated Dicarboxylic Acids and TBD.....	55
2.6.2.3 Aromatic Dicarboxylic Acids and TBD.....	56
2.6.3. Mono-carboxylic Acid and TBD.....	56
2.7 Single Crystal X-ray Crystallography	57
2.8 Results and Discussion	58
2.8.1 Pseudopolymorphic Crystalline Complexes of TBD and BPDA.....	59
2.8.2 Reproducibility of Type I Complex and Crystal Structure Transformation.....	60
2.8.3 Di-anionic Complexes from Dicarboxylic Acids and TBD.....	61
2.8.4 Mono-anionic Complexes from TBD and Mono- or Dicarboxylic Acids	61
2.9 Conclusions	62
Reference List.....	62
Appendix: Publications	67

List of Abbreviations

Abu	L-2-aminobutyric acid
aq	aqueous
Boc	<i>tert</i> -butoxycarbonyl (COtC4H9)
Bn	benzyl
BPDA	2,2'-bipyridine-5,5'-dicarboxylic acid
BPHA	biphenyl-4,4'-dicarboxylic acid
bs	broad singlet (NMR signal)
CBz	benzyloxycarbonyl (BnOC=O)
CCD	charge-coupled device
CSD	Cambridge structural database
DCC	dicyclohexylcarbodiimide
DCU	dicyclohexyl urea
DMF	<i>N, N</i> -dimethylformaamide (solvent)
DM	demineralized
DMSO	dimethyl sulphoxide (solvent)
DNA	deoxyribonucleic acid
d	doublet (NMR signal)
1D	one dimensional
2D	two dimensional
3D	three dimensional
δ	NMR chemical shift
ES ^{+/-}	electrospray
Hz	hertz
IPA	2-propanol
K	kelvin
MHz	megahertz
MOFs	metal-organic frameworks
<i>m/z</i>	mass-to-charge ratio
NHS	<i>N</i> -hydroxysuccinimide
NMR	nuclear magnetic resonance
Nva	L-norvaline
RNA	ribonucleic acid
RT	room temperature
TBD	1,5,7-triazabicyclo[4.4.0]-dec-5-ene
THF	tetrahydrofuran

List of Publications

- I)** Porous Organic Materials from Dipeptides with Non-proteinogenic Residues
Vitthal N. Yadav, Carl Henrik Görbitz, Tore-Boge Hansen, Angiolina Comotti and Piero Sozzani, (2013), (Manuscript Submitted to JACS).
- II)** A Water Wire in L-prolyl-L-serine Hydrate
Carl Henrik Görbitz and **Vitthal N. Yadav**, (2013), *Acta. Cryst.* C69, 556-559.
- III)** An Unexpected Tetragonal Unit Cell for N-(L-2-aminobutyl)-L-serine
Carl. Henrik Görbitz and **Vitthal N. Yadav**, (2013), *Acta. Cryst.* C69, 888-891.
- IV)** N-(L-2-Aminopentanoyl)-L-phenylalanine Dihydrate, a Hydrophobic Dipeptide with a Nonproteinogenic Residue
Carl Henrik Görbitz and **Vitthal N. Yadav**, (2013), *Acta. Cryst.* C69, 1067-1069.
- V)** A Supramolecular 2 : 1 Guanidinium-carboxylate Based Building Block for Generation of Water Channels and Clusters in Organic Materials
Vitthal N. Yadav and Carl Henrik Görbitz, (2013), *CrystEngComm*, **15**, 439-442, (*Hot Article, RSC, Nov. 2012).
- VI)** Water of Hydration in 2:1 Hydrogen Bonded Complexes Between 1,5,7 triazabicyclo[4.4.0]-dec-5-ene and Dicarboxylic acids
Vitthal N. Yadav and Carl Henrik Görbitz, (2013), *Crystal Growth & Design*, **13**, 2174-2180.
- VII)** Supramolecular 1D Ribbons from Bicyclic Guanidine Derivative and Di- or Monomonocarboxylic acids
Vitthal N. Yadav and Carl Henrik Görbitz, (2013), *CrystEngComm*, **15**, 7321-7326. (*Hot Article, RSC, Aug. 2013).

Publications Not Listed in this Thesis

- 1)** A supramolecular Ladder-like Network from Trimesic acid and Pyrazine-*N*, *N'*-dioxide
Vitthal N. Yadav and Carl Henrik Görbitz, (2013), submitted, *Acta. Cryst. C*.
- 2)** 1,1'-(4,4'-Bipiperidine-1,1'-di-yl)bis-(2,2,2-trifluoro-ethanone)
Vitthal N. Yadav, Tore Hansen, and Carl Henrik Görbitz, (2011), *Acta. Cryst.* E67, o1691.

Introduction: Supramolecular Chemistry

Supramolecular chemistry is a branch of chemistry devoted to the study of systematic aggregation of two or more chemical moieties through non-covalent forces between them, covering their functions and structure properties. In other words, as stated by Jean Marie Lehn the supramolecular chemistry is “chemistry beyond the molecule”.¹⁻⁴ This means that a better insight into the implicit properties of the covalently bonded atoms at the molecular level is essential in supramolecular chemistry to understand the intermolecular interactions between complementary molecules,⁵ such as the hydrogen-bonded networks formation between the carboxylic acid groups (Fig. 1).

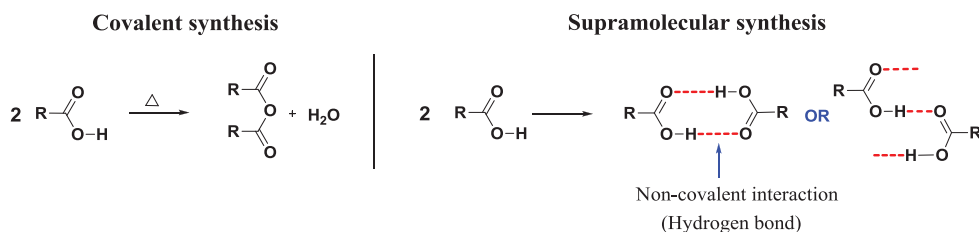


Figure 1: The difference between covalent and supramolecular syntheses.

Natural Processes and Supramolecular Chemistry

In nature, supramolecular chemistry is ubiquitous.^{4,6} Earlier, to facilitate the preliminary understanding of biological transformations through binding of enzyme and substrate, the concept ‘lock-and-key’ was proposed by Emil Fisher (1894).^{7,8} Later with the initial essence of intermolecular interactions, the adducts of receptors and substrates were referred as ‘supermolecules’ (1937).^{9,1} In the late 20th century it was evident from breakthrough discoveries that the weak intermolecular interactions are the responsible forces for self-organization or spatial arrangements of macromolecules and extra- or intracellular biochemical modifications in the biological systems.¹⁰⁻¹² The well known examples of such natural self-assemblies are the pairing between complementary DNA bases i. e. ; Adenine (A)-Thymine (T); and Guanine (G)-Cytosine (C), (Fig. 2) to form a double-stranded DNA (Fig. 2),¹³ ion transportation through the transmembrane channels,^{14,15} DNA-histone protein complex formation,^{16,17} RNA-protein recognition,¹⁸⁻²⁰ enzymatic catalysis,^{21,22} protein-protein interactions,²³ arginine recognition,²⁴ stability/folding of proteins/enzymes by hydration,^{25,26} folding of peptide chain into an α -helix or a β -sheet²⁷ and further into a functional domain, for example, myoglobin, a metalloprotein in all vertebrates (Fig. 3).

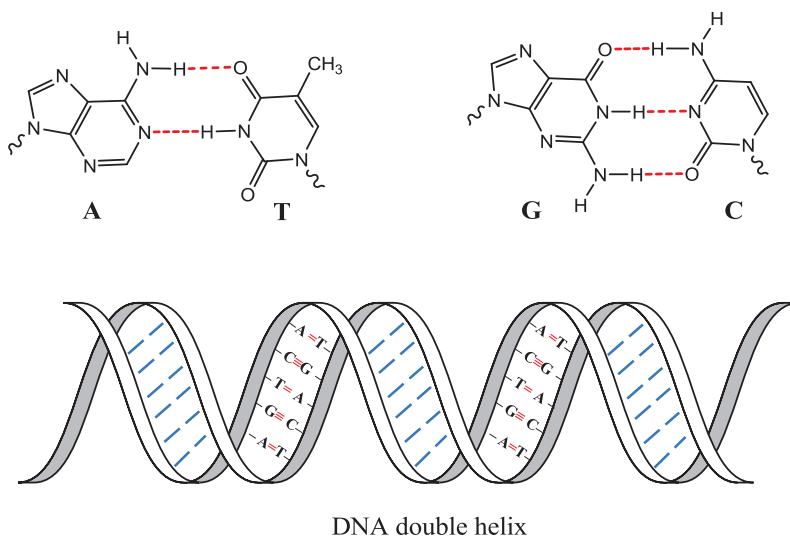


Figure 2: H-bonded DNA base pairs and a double stranded DNA.

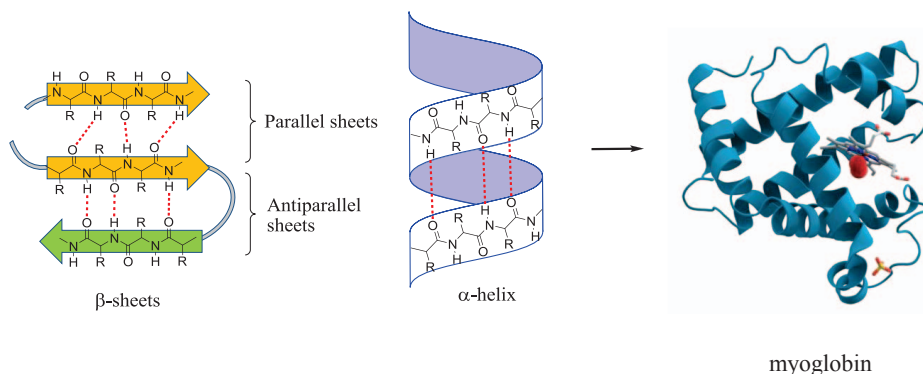


Figure 3: The H-bonding-assisted folding of polypeptide chains into secondary protein structures i. e. α -helix and β -sheets. The hydrogen bonds are shown with the red-colored broken lines. Myoglobin, a functional heme protein, illustration is adapted from reference.²⁸

These observations of precise attraction and binding through non-covalent interactions in nature offered the initial ideas not only to develop the ion-specific artificial macrocyclic receptors^{29,30} and host-guest complexes,³¹ but it has also inspired to design the tactics to prepare nanoscale assemblies from the small peptides³²⁻³⁹ and from various complementary organic building blocks^{40,41} as materials.⁴

Specific Intermolecular Interactions and Crystal Engineering

Self-assembly of molecules, alongside the strong hydrogen bonds (N–H···O, N–H···N, N···H–O, O–H···O, S–H···O and S–H···N etc.) also includes weaker electrostatic or coulombic forces (ion-ion, ion-dipole, and dipole-dipole),^{42,43} co-ordinate interactions (metal-ligand),^{44,45} halogen bonding,^{46,47} π ··· π stacking,⁴⁸⁻⁵¹ cation··· π ,⁵²⁻⁵⁴ anion··· π ,⁵⁵⁻⁵⁸ and C–H··· π ,^{59,60} C–H···O and C–H···N interactions^{11,61-63} and van der Waals forces.^{64,65} However, In the preparation of organic supramolecular structures, hydrogen bond is a dominating and preferred non-covalent interaction due to its specific and directive property in the solid states as well as in solutions.^{4,10,11,66,67}

As shown in Fig. 4, specific functional groups of different molecules form strong hydrogen bonds and generate recurring H-bonded motifs in structures, such patterns are termed as ‘*supramolecular synthons*’.⁴⁰ Formation of these precise and inherent hydrogen-bonded interactions between the complementary functional groups is a well known rational criteria to obtain the new organic supramolecular structures.^{40,68-72}

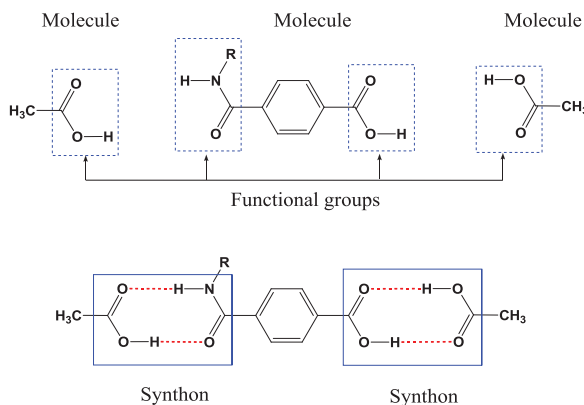


Figure 4: Synthon formation from complementary functional groups.

Surveys of experimentally investigated crystal structures by Cambridge Structural Database (CSD)^{73,74} have showed that the $R_2^2(8)$ motif [coded according to Etter’s nomenclature R = H-bonded ring, (8) = ring of 8 atoms, 2 = number of donors, and 2 = number of acceptors)^{75,76} is one of the most common H-bonded synthons. Such a H-bond pattern in the solid state can be obtained from molecules with a wide range of complementary functional groups such as carboxylic acid, amide^{77,78} amidine,⁷⁹⁻⁸² urea,⁸³⁻⁸⁵ guanidinium-carboxylate/phosphonate/sulfonate/nitro or boronic acid etc. (Fig. 1.2.2, a-l).⁸⁶⁻⁹⁰

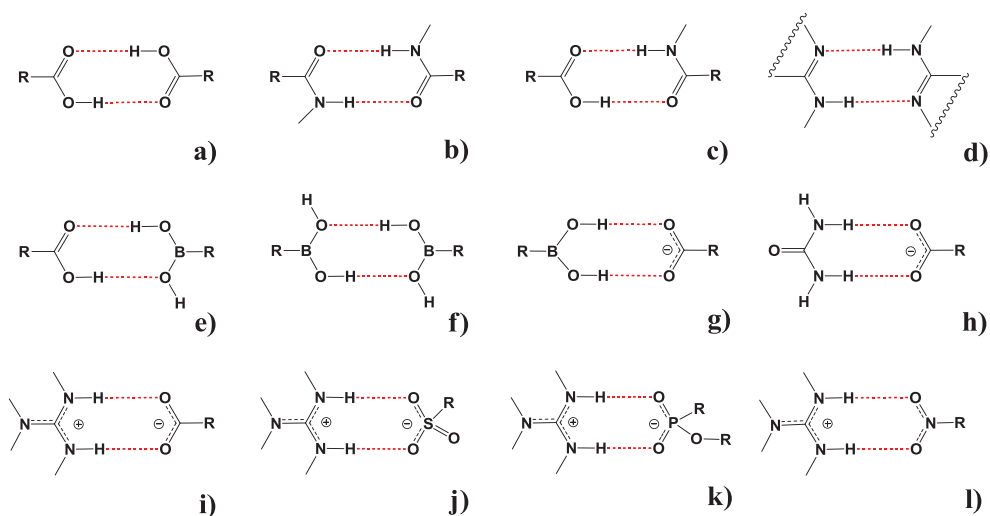


Figure 5: Classical $R_2^2(8)$ supramolecular synthon formation from various functional groups in the solid state, a) a neutral homodimer from carboxylic acid and b) amide groups; c) A neutral heterodimer from carboxylic acid - amide d) amidines e) carboxylic acid - boronic acid, f) boronic acid - boronic acid, g) mixed neutral - anionic, boronic acid - carboxylate, h) urea - carboxylate, e) zwitterionic from guanidinium - carboxylate, f) guanidinium - sulfonate g) guanidinium - phosphate and h) guanidinium - nitro groups.

The $R_2^2(8)$ motif and other various specific H-bonded patterns not only help to design the new supramolecular synthetic strategies, but also offer the ideas to prepare aesthetic and application-oriented new molecular assemblies, which are quite difficult to obtain by traditional-covalent synthesis.^{91,92} The resulting structures with varying properties are vindicated as useful materials in the field of optical activity, conductivity (salt bridges), magnetism, sensors, material science, theoretical science and as artificial receptors in biochemical molecular recognitions etc.^{41,93-95}

Reference List

- (1) Lehn, J.-M. *Angew. Chem. Int. Ed.* **1988**, *27*, 89-112.
- (2) Lehn, J. *Science* **1993**, *260*, 1762-1763.
- (3) Lehn, J. M.; Atwood, J. L.; Davies, J. E. D.; MacNicol, D. D.; Vögtle, F. *Comprehensive Supramolecular Chemistry*; Pergamon: Oxford, 1996.
- (4) Steed, J. W.; Atwood, J. L. *Supramolecular Chemistry*, 2nd ed.; Wiley VCH, 2009.
- (5) Whitesides, G. M.; Boncheva, M. *Proc.Natl. Acad. Sci.* **2002**, *99*, 4769-4774.
- (6) Philp, D.; Stoddart, J. F. *Angew. Chem. Int. Ed.* **1996**, *35*, 1154-1196.
- (7) Fischer, E. *Chem. Ber.* **1894**, *27*, 2985-2993.
- (8) Kunz, H. *Angew. Chem. Int. Ed.* **2002**, *41*, 4439-4451.
- (9) Wolf, K. L.; Wolff, R. *Angew. Chem.* **1949**, *61*, 191-201.
- (10) Steiner, T. *Angew. Chem. Int. Ed.* **2002**, *41*, 48-76.
- (11) Desiraju, G. R.; Steiner, T. *The Weak Hydrogen Bond In Structural Chemistry and Biology*; Oxford University Press: Oxford, 1999.
- (12) Gellman, S. H. *Chem. Rev.* **1997**, *97*, 1231-1232.
- (13) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737-738.
- (14) Bezanilla, F.; Armstrong, C. M. *The Journal of General Physiology* **1972**, *60*, 588-608.
- (15) Hille, B. *Ion Channels of Excitable Membranes*; (Sinauer, Sunderland, Massachusetts), 2001.
- (16) Van Holde, K. E.; Allen, J. R.; Tatchell, K.; Weischet, W. O.; Lohr, D. *Biophys. J.* **1980**, *32*, 271-282.
- (17) Davey, C. A.; Sargent, D. F.; Luger, K.; Maeder, A. W.; Richmond, T. J. *J. Mol. Biol.* **2002**, *319*, 1097-1113.
- (18) Varani, G. *Acc. Chem. Res.* **1997**, *30*, 189-195.
- (19) Chow, C. S.; Bogdan, F. M. *Chem. Rev.* **1997**, *97*, 1489-1514.
- (20) Draper, D. E. *Annu. Rev. Biochem.* **1995**, *64*, 593-620.
- (21) Fersht, A. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*; W. H. Freeman and Co.: New York, 1999.
- (22) Zhang, X.; Houk, K. N. *Acc. Chem. Res.* **2005**, *38*, 379-385.
- (23) Stites, W. E. *Chem. Rev.* **1997**, *97*, 1233-1250.
- (24) Cavarelli, J.; Delagoutte, B.; Eriani, G.; Gangloff, J.; Moras, D. *EMBO J.* **1998**, *17*, 5438-5448.
- (25) Gregory, R. B. *Protein Solvent Interactions*; Dekker: New York, 1995.
- (26) Levy, Y.; Onuchic, J. N. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 389-415.

- (27) Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walters, P. *The Shape and Structure of Proteins: Molecular Biology of the Cell*, 4th ed.; Garland Science: New York and London, 2002.
- (28) <http://en.wikipedia.org/wiki/Myoglobin>.
- (29) Pedersen, C. J. *Angew. Chem. Int. Ed.* **1988**, *27*, 1021-1027.
- (30) Lehn, J.-M. *Science* **1985**, *227*, 849-856.
- (31) Cram, D. J. *Angew. Chem. Int. Ed.* **1988**, *27*, 1009-1020.
- (32) Gao, X.; Matsui, H. *Advanced Materials* **2005**, *17*, 2037-2050.
- (33) Boyle, A. L.; Woolfson, D. N. In *Supramolecular Chemistry*; John Wiley & Sons, Ltd, 2012.
- (34) Woolfson, D. N.; Ryadnov, M. G. *Curr. Opin. Chem. Biol.* **2006**, *10*, 559-567.
- (35) Hartgerink, J. D.; Clark, T. D.; Ghadiri, M. R. *Chem. Eur. J.* **1998**, *4*, 1367-1372.
- (36) Reches, M.; Gazit, E. *Current Nanoscience*, *2*, 105-111.
- (37) Valery, C.; Artzner, F.; Paternostre, M. *Soft Matter* **2011**, *7*, 9583-9594.
- (38) Jaime Castillo-León; Andersen, K. B.; Svendsen, a. W. E. *Self-Assembled Peptide Nanostructures for Biomedical Applications: Advantages and Challenges, Biomaterials Science and Engineering, Prof. Rosario Pignatello (Ed.), ISBN: 978-953-307-609-6*, (2011).
- (39) Cui, Y.; Kim, S. N.; Naik, R. R.; McAlpine, M. C. *Acc. Chem. Res.* **2012**, *45*, 696-704.
- (40) Desiraju, G. R. *Angew. Chem. Int. Ed.* **1995**, *34*, 2311-2327.
- (41) Desiraju, G. R. *J. Mol. Struct.* **2003**, *656*, 5-15.
- (42) Marcus, Y. In *Ionic Interactions in Natural and Synthetic Macromolecules*; John Wiley & Sons, Inc., 2012; pp. 1-33.
- (43) Schneider, H.-J. In *Ionic Interactions in Natural and Synthetic Macromolecules*; John Wiley & Sons, Inc., 2012; pp. 35-47.
- (44) Lawrance, G. A. *Introduction to Coordination Chemistry*; John Wiley & Sons, 2009.
- (45) Yaghi, O. M.; O'Keeffe, M.; Ockwig, N. W.; Chae, H. K.; Eddaoudi, M.; Kim, J. *Nature* **2003**, *423*, 705-714.
- (46) Rissanen, K. *CrystEngComm* **2008**, *10*, 1107-1113.
- (47) Metrangolo, P.; Neukirch, H.; Pilati, T.; Resnati, G. *Acc. Chem. Res.* **2005**, *38*, 386-395.
- (48) Hunter, C. A.; Sanders, J. K. M. *J. Am. Chem. Soc.* **1990**, *112*, 5525-5534.
- (49) Meyer, E. A.; Castellano, R. K.; Diederich, F. *Angew. Chem. Int. Ed.* **2003**, *42*, 1210-1250.
- (50) Salonen, L. M.; Ellermann, M.; Diederich, F. *Angew. Chem. Int. Ed.* **2011**, *50*, 4808-4842.

- (51) Cockroft, S. L.; Perkins, J.; Zonta, C.; Adams, H.; Spey, S. E.; Low, C. M. R.; Vinter, J. G.; Lawson, K. R.; Urch, C. J.; Hunter, C. A. *Org. Biomol. Chem.* **2007**, *5*, 1062-1080.
- (52) Ma, J. C.; Dougherty, D. A. *Chem. Rev.* **1997**, *97*, 1303-1324.
- (53) Mahadevi, A. S.; Sastry, G. N. *Chem. Rev.* **2012**, *113*, 2100-2138.
- (54) Gallivan, J. P.; Dougherty, D. A. *Proc. Nat. Acad. Sci.* **1999**, *96*, 9459-9464.
- (55) Chifotides, H. T.; Dunbar, K. R. *Acc. Chem. Res.* **2013**, *46*, 894-906.
- (56) Wang, D.-X.; Wang, M.-X. *J. Am. Chem. Soc.* **2012**, *135*, 892-897.
- (57) Robertazzi, A.; Krull, F.; Knapp, E.-W.; Gamez, P. *CrystEngComm* **2011**, *13*, 3293-3300.
- (58) Schottel, B. L.; Chifotides, H. T.; Dunbar, K. R. *Chem. Soc. Rev.* **2008**, *37*, 68-83.
- (59) Tsuzuki, S. *Annual Reports Section "C" (Physical Chemistry)* **2012**, *108*, 69-95.
- (60) Nishio, M. *CrystEngComm* **2004**, *6*, 130-158.
- (61) Desiraju, G. R. *Acc. Chem. Res.* **1996**, *29*, 441-449.
- (62) Desiraju, G. R. *Chem. Commun.* **2005**, 2995-3001.
- (63) Lim, J.; Osowska, K.; Armitage, J. A.; Martin, B. R.; Miljanic, O. S. *CrystEngComm* **2012**, *14*, 6152-6162.
- (64) Parsegian, V. A. *Van der Waals Forces: A Handbook for Biologists, Chemists, Engineers, and Physicists*; Cambridge University Press, 2006.
- (65) Buckingham, A. D.; Fowler, P. W.; Hutson, J. M. *Chem. Rev.* **1988**, *88*, 963-988.
- (66) Jeffrey, G. A. *An Introduction to Hydrogen Bonding* Oxford University Press, Oxford, 1997.
- (67) Oshovsky, G. V.; Reinhoudt, D. N.; Verboom, W. *Angew. Chem. Int. Ed.* **2007**, *46*, 2366-2393.
- (68) Aakeroy, C. B.; Seddon, K. R. *Chem. Soc. Rev.* **1993**, *22*, 397-407.
- (69) Nangia, A.; Desiraju, G. *Supramolecular Synthons and Pattern Recognition: Design of Organic Solids*; Springer Berlin / Heidelberg, 1998; Vol. 198.
- (70) Reddy, L. S.; Babu, N. J.; Nangia, A. *Chem. Commun.* **2006**, 1369-1371.
- (71) Merz, K.; Vasylyeva, V. *CrystEngComm* **2010**, *12*, 3989-4002.
- (72) Reddy, D. S.; Craig, D. C.; Desiraju, G. R. *J. Am. Chem. Soc.* **1996**, *118*, 4090-4093.
- (73) Allen, F. *Acta Crystallogr. Sect. B* **2002**, *58*, 380-388.
- (74) Allen, F. H.; Motherwell, W. D. S. *Acta Crystallogr. Sect. B* **2002**, *58*, 407-422.
- (75) Etter, M. C. *Acc. Chem. Res.* **1990**, *23*, 120-126.
- (76) Etter, M. C.; MacDonald, J. C.; Bernstein, J. *Acta Crystallogr. Sect. B* **1990**, *46*, 256-262.
- (77) Leiserowitz, L. *Acta Crystallogr. Sect. B* **1976**, *32*, 775-802.

- (78) Aakeröy, C. B.; Beatty, A. M.; Helfrich, B. A. *J. Am. Chem. Soc.* **2002**, *124*, 14425-14432.
- (79) Félix, O.; Hosseini, M. W.; De Cian, A.; Fischer, J. *Angew. Chem. Int. Ed.* **1997**, *36*, 102-104.
- (80) Felix, O.; Hosseini, M. W.; De Cian, A.; Fischer, J. *Chem. Commun.* **2000**, 281-282.
- (81) Lie, S.; Maris, T.; Malveau, C.; Beaudoin, D.; Helzy, F.; Wuest, J. D. *Cryst. Growth Des.* **2013**.
- (82) Felix, O.; Wais Hosseini, M.; De Cian, A.; Fischer, J. *New J. Chem.* **1998**, *22*, 1389-1393.
- (83) Xue, F.; Mak, T. C. W. *Acta Crystallogr. Sect. B* **2000**, *56*, 142-154.
- (84) Xue, F.; Mak, T. C. W. *J. Phys. Org. Chem.* **2000**, *13*, 405-414.
- (85) Adarsh, N. N.; Kumar, D. K.; Dastidar, P. *Tetrahedron* **2007**, *63*, 7386-7396.
- (86) H. Allen, F.; D. Samuel Motherwell, W.; R. Raithby, P.; P. Shields, G.; Taylor, R. *New J. Chem.* **1999**, *23*, 25-34.
- (87) Aakeroy, C. B.; Desper, J.; Levin, B. *CrystEngComm* **2005**, *7*.
- (88) Haynes, D. A.; Chisholm, J. A.; Jones, W.; Motherwell, W. D. S. *CrystEngComm* **2004**, *6*, 584-588.
- (89) Dumitrescu, D.; Legrand, Y.-M.; Dumitrascu, F.; Barboiu, M.; van der Lee, A. *Cryst. Growth Des.* **2012**, *12*, 4258-4263.
- (90) Rodríguez-Cuamatzi, P.; Arillo-Flores, O. I.; Bernal-Uruchurtu, M. I.; Höpfl, H. *Cryst. Growth Des.* **2004**, *5*, 167-175.
- (91) Batten, S. R.; Robson, R. *Angew. Chem. Int. Ed.* **1998**, *37*, 1460-1494.
- (92) Beer, P. D.; Gale, P. A. *Angew. Chem. Int. Ed.* **2001**, *40*, 486-516.
- (93) Lüning, U. *Applications of Supramolecular Chemistry*; Edited by Hans-Jörg Schneider, WILEY-VCH, Weinheim, 2013.
- (94) Lüning, U. *Angew. Chem. Int. Ed.* **2013**, *52*, 4724-4724.
- (95) Kirby, J. P.; Roberts, J. A.; Nocera, D. G. *J. Am. Chem. Soc.* **1997**, *119*, 9230-9236.

Chapter 1

Microporous Dipeptide Structures

1.1 Introduction: Amino acids, Peptides and Proteins

In nature there are 20 common amino acids, every single one of them functioning as a basic building block for proteins. Except glycine, all amino acids possess a stereogenic C^α-atom with absolute L-configuration that covalently holds an amino group, a carboxylic acid group and a side chain that has different chemical and physical property such as hydrophobic, polar, acidic, basic or neutral (Fig. 1.1.1).

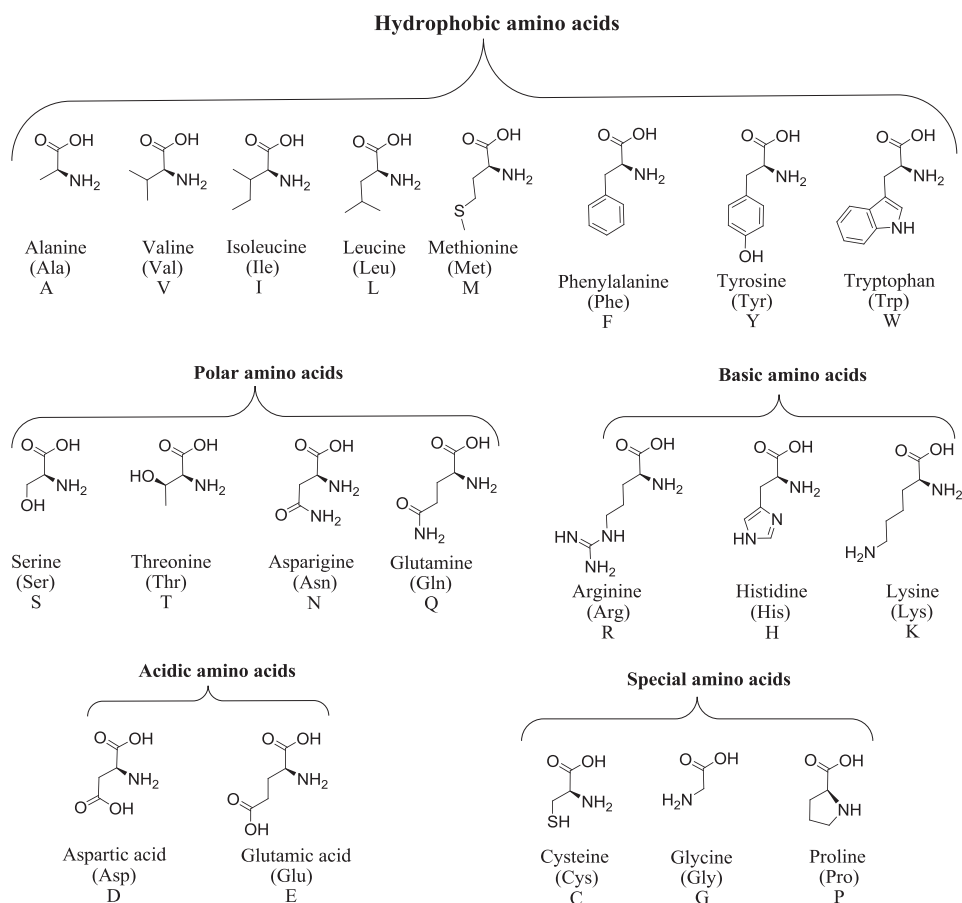


Figure 1.1.1: The structure and chemical properties based classification of 20 natural amino acids shown with three-letter abbreviations and one-letter symbols.

Peptides are chains of amino acids linked together by peptide bonds, and when many amino acids join in a single linear chain it is called a polypeptide. In a peptide chain each amino acid part is referred as a *residue* (Fig. 1.1.2). One or more polypeptide chains can fold through non-covalent interactions into secondary structures, which hierarchically builds a functional compact domain in a native protein or enzyme.¹⁻³ (Fig. 1.1.3, Chapter 1). Proteins are abundantly occurring biopolymers in all cells of living beings and serve a central role in various biological and physiological processes. Distinct proteins with various lengths and diverse structural forms participate in different functions⁴ such as enzymatic catalysis (enzyme that speeds the chemical reaction),^{5,6} in transportation of essential ions or molecules across the cell membrane,^{7,8} in muscular movements as contractile proteins (actin and myosin),⁹ and acts as hormones,¹⁰ neurotransmitters^{11,12} and as antibodies in immune systems,¹³ etc.

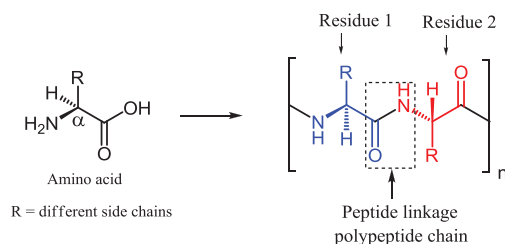


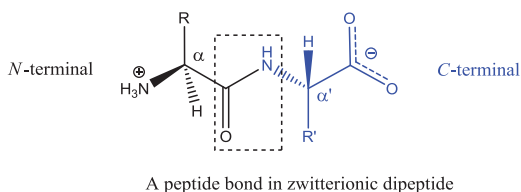
Figure 1.1.2: A schematic representation of amino acid to protein structure build up.

1.1.2 Small Peptides (Natural vs Synthetic Peptides)

Contrary to the biological proteins, small peptides with desired sequence and length can be easily prepared in the laboratory. Such peptide fragments can be used as potential drug candidates.¹⁴⁻¹⁶ Moreover, their natural tendency of self-assembly adds an advantage in ‘bottom-up’ approach¹⁷ (larger assemblies from smaller units, while ‘top-down’ is to carve smaller from larger) to build the efficient and useful nano-scale biomaterials.¹⁸⁻²⁷ Such structures can function as excellent materials for asymmetric catalysts,²⁸ as gelators,^{29,30} nano-vesicles for various bio-medicines,³¹ drug delivery vehicles,³² in regenerative medicines,^{33,34} and more importantly as bio-mimetic devices in clinical-biology.^{25,35-40}

1.1.3 Nanoporous Self-assembly of Dipeptides and Material Science

A dipeptide is the smallest representative molecule of a protein; generally existing in a zwitterionic form, containing a characteristic peptide bond and behaving as a model for precise structure geometries and hydrogen-bonded self-organization of proteins.



Nearly for two decades our research group is engaged in an interdisciplinary field of structural and materials property investigations of small peptides, particularly in zwitterionic dipeptide structures.⁴¹⁻⁷⁰ There have been tremendous outcomes throughout this period, especially from the molecules that are composed of hydrophobic amino acids such as Ala, Val, Leu, Ile, Phe, Met and Trp⁴³ (all amino acids and dipeptides mentioned in this dissertation are of the L-configuration, thus stereochemical indicators are not included). Almost 50% of the investigated dipeptides show porous structures formation,⁴³ the rest generate non-porous networks in the solid states. According to the three-dimensional structures and H-bonding properties,⁴³ the porous dipeptides are divided into the Val-Ala class and the Phe-Phe class. A rare additional example of porous dipeptide includes Ser, a polar residue, in Leu-Ser.⁴⁴ The Val-Ala and Leu-Ser classes produce the hydrophobic micropores, while the Phe-Phe class generates hydrophilic channels in highly stable H-bonded networks (Fig. 1.1.3).⁴³ In the modern nano-science discipline, the porous molecular structures or assemblies are in the demand as futuristic smart materials for growing field of applications.^{71,72} Therefore, we focus here mainly upon studies related to the porous dipeptide structures.

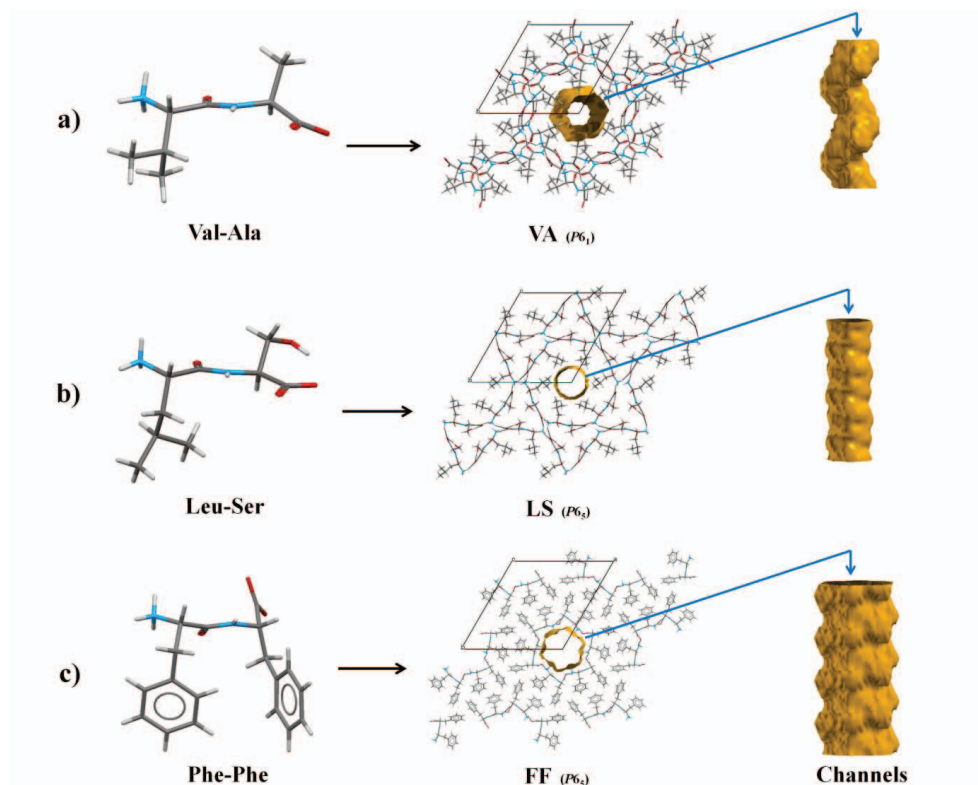
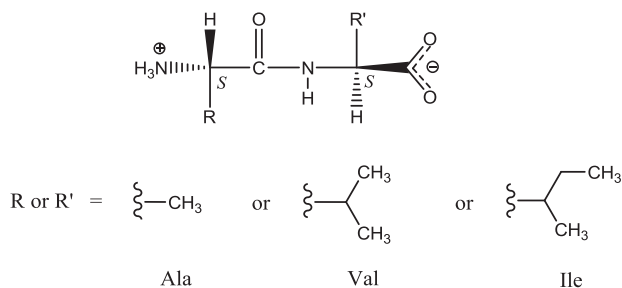


Figure 1.1.3: Representative molecular structures, crystal packing arrangement and channel shapes of nanoporous dipeptide classes, a) Val-Ala [Cambridge Structure Database (CSD v5.34 Nov.2012)⁷³, refcode WIRYEB⁴¹], b) Leu-Ser (JAZBOC)⁴⁴ and d) Phe-Phe (IFABEW).⁴⁵

1.1.3.1 Self-assembly of the Val-Ala Class Dipeptides



The first ever dipeptide structure reported by our group is Val-Ala,⁴¹ which is an interesting chiral small molecule and formed a stable self-assembly with nano-porous structural property. Subsequent investigations of dipeptides predominantly combining Ala, Val and Ile residues

have added numerous members in this series named the Val-Ala class.⁴⁹ To date, this is the major group of porous dipeptides including seven molecules with similar hydrogen-bonding patterns (Fig. 1.1.3.1.a),^{41,49} but different pore dimensions and surface areas due to variable aliphatic side chain lengths and size (Fig. 1.1.3.1.b). Ala-Ala is composed of smallest Ala residue, inspite the similar peptide backbone as Val-Ala it generates a nonporous structure with different crystal packing.⁷⁴ Obviously, the steric crowding of bulky *sec*-butyl side chains of Ile-Ile forces structure to form a densely packed 3D arrangement.⁶³ However, when Ala is coupled with Val or Ile it creates the porous structures (Fig. 1.1.3.1.b).

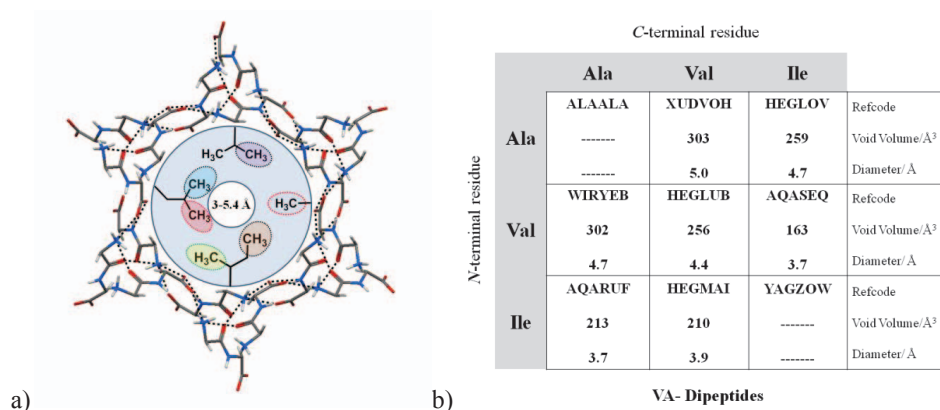


Figure 1.1.3.1: a) A representative hydrogen-bonded peptide backbone of the porous Val-Ala class of dipeptides is shown. The region occupied by hydrophobic amino acid side chains is shown with light-blue shade. b) Dipeptides belonging to a Val-Ala class are listed according to their pore diameters, channel surface areas and CSD refcodes. Ala-Ala and Ile-Ile form non-porous structures.^{63,74} The channel diameters and surface areas were calculated by the program Mercury.⁷⁵

1.1.3.2 Self-assembly of Leu-Ser Dipeptide

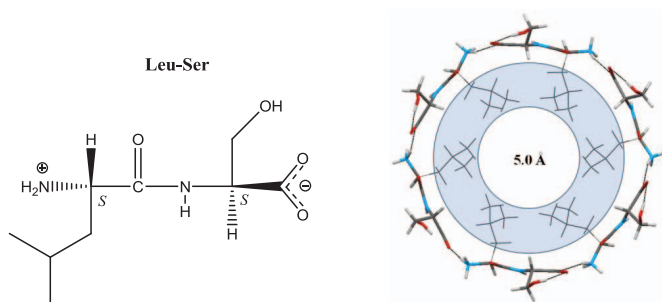


Figure: 1.1.3.2: Molecular structure of Leu-Ser (JAZBOC).⁴⁴ A single channel in Leu-Ser assembly is shown (5.0 Å diameter, unit cell volume = 9.3 %, void volume = 163.5 Å³).⁷⁶ For clarity, Leu side chains aligned around the channel (light blue shade) are drawn as thin black wires and the H-bonded peptide backbone in capped sticks.

Leu-Ser also generates porous structure, but due to involvement of hydroxyl group of Ser residue (Fig. 1.1.3.2a) it acquires a different H-bonding network (Fig. 1.1.3.2b)⁴⁴ than that of the Val-Ala class of molecules (Fig. 1.1.3.1a). In the past, there have been efforts to prepare the additional Leu-Ser type of structures with wider channels by replacing Leu with Val, a residue with one C-atom less than the Leu side chain. The resulting Val-Ser in fact has been found to form a non-porous dihydrate (FOBLUE)⁵⁸ and a porous Val-Ala class-related monoclinic solvate with $Z' = 3$ (CAZGOA)⁵¹. Among the same category of dipeptides, Ala-Ser (LALLSE)⁷⁷ showed a compact 3D H-bonded structure, while the highly hydrophobic *N*-terminal residue generated layered structures [Phe-Ser (MAZYAO), Ile-Ser (MAZXUH) and Met-Ser (MAZYES)].⁵⁶ Thus, the Leu-Ser class of dipeptide has so far exhibited only single porous member, i.e. Leu-Ser itself (Fig. 1.1.3.2).⁴⁴

1.1.3.3 Self-assembly of the Phe-Phe Class Dipeptides

The Phe-Phe class, in addition to Phe-Phe (IFABEW) includes Leu-Leu (IDUZOW), Ile-Leu (ETITUW), Leu-Phe (EDUZUC) and Phe-Leu (IFABAS)⁴⁵ dipeptides and also the additional members, Phe-Trp (GEHTAP)⁷⁰ and Trp-Gly (FULGEY).⁷⁸ The unique characteristic feature of this class is the generation of hydrophilic channels by engaging four to six dipeptide molecules in a H-bonded helical chain in the periphery of channel axis (Fig. 1.1.3.3).^{43,45}

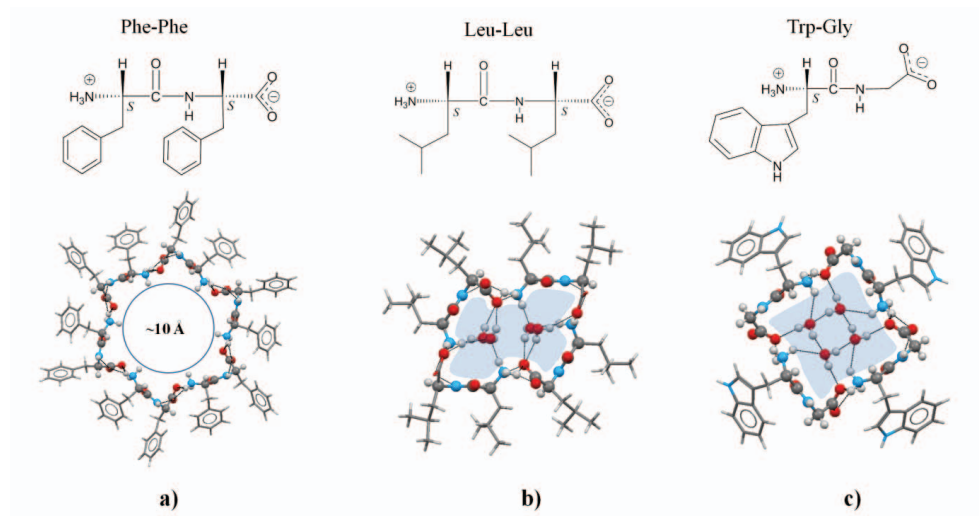
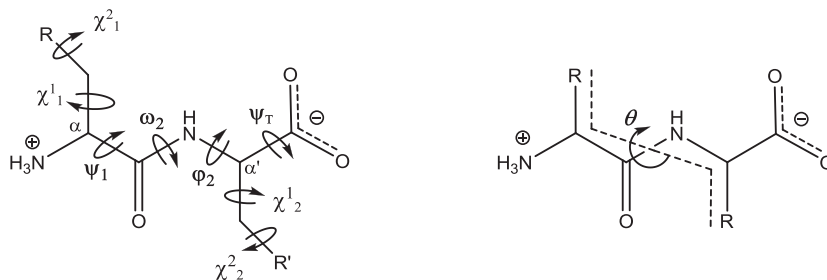


Figure: 1.1.3.3: The representative Phe-Phe class members are shown. a) Phe-Phe (IFABEW)⁴⁵ molecular structure showing an organized chain of six Phe-Phe molecules around the open pore (diameter = ~ 10 Å, unit cell, void volume = 402 Å³). For clarity, the disordered water molecules inside the channel have been omitted. b) Leu-Leu (IDUZOW)⁴⁵ and c) Trp-Gly (FULGEY)⁷⁸ are showing the crystal packings involving four dipeptide molecules around the pore of respective structure. Water filled channels are shown with the light-blue shades. In a-c) the atoms of polar peptide backbones involved in hydrogen bondings are drawn as ball and stick.

1.1.3.4 Side Chain Size, Orientations and Channel Property

From the experimentally well-investigated dipeptide structures one can see the importance of the main chain conformation of dipeptides which is an important factor for their self-assembly and the nature of the pores formed, hydrophobic or hydrophilic. When the side chains acquire positions on opposite sides of the peptide plane (θ and φ both close to 180°) the structure forms a hydrophobic pore (in Val-Ala and Leu-Ser classes).^{49,44} Alternatively, when the side

chains of dipeptide are located on same side of the peptide plane (θ and φ both close to 0°) it forms hydrophilic channel in the structures (FF class)⁴⁵ (Fig. 1.1.3 and Fig. 1.1.3.4).



Scheme 1.1.3 The schematic representations showing the torsion angle symbols used for a dipeptide.

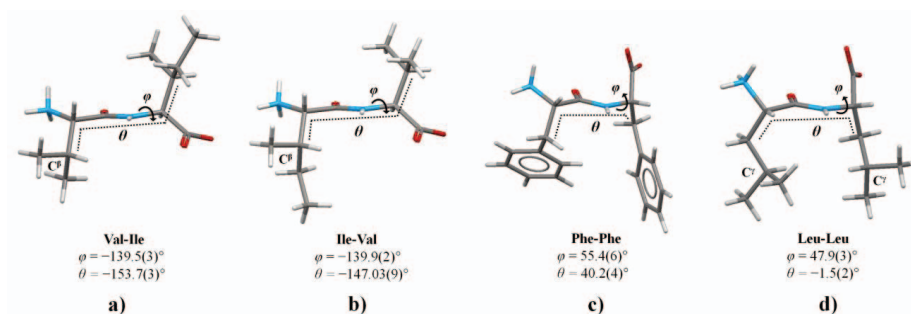


Figure 1.1.3.4: Representative dipeptides with bulkier side chains showing difference in side chain orientations (θ and φ torsions) in a) Val-Ile, b) Ile-Val, molecules of the Val-Ala class, and c) Phe-Phe d) Leu-Leu members of the FF class. a) and b) are hydrophobic pore formers, whereas, c) and d) are hydrophilic channel formers.

1.1.4 Rational Strategies for Channel Size Modifications

In recent years, the microporous Val-Ala and Leu-Ser classes of small-molecular assemblies have been successfully examined as materials for CO₂, methane and hydrogen gas absorption and separation, and also as a catalyst.⁷⁹⁻⁸² Whereas, the Phe-Phe structure (Fig. 1.1.1.3a), due to its stability⁸³ and wide pore dimensions, is an extensively exploited molecule as an advanced material in the diverse nano-technological applications.^{84,21,35,85} Also, Phe-Phe constitutes the recognition motif in Alzheimer's β -amyloid fibrils, thus the self-assembly of FF is also well studied for clinical reasons.^{21,45,86,87}

Molecular structures with nanoporous properties are great in demand for emerging nano-scale applications.^{25,33,35,36,38,72,87-91} Hence, herein we continue our research efforts in this area to investigate the novel structural, specifically, microporous properties of bio-derived small molecules. For this, we decided to design a strategy based on previously studied porous dipeptides and take the advantage from the residues with bulky aliphatic side chains, which organize in the periphery of the channel, mainly in the Val-Ala and Leu-Ser class of structures, as shown in Fig. 1.1.4.1. From this approach we envisaged that increased pore dimensions in such structures could potentially be achieved by replacing bulkier side chains of Val and Ile in VA class, or Leu of LS, with side chains of Abu or Nva (ethyl or n-propyl groups) (Fig. 1.1.4.2).

In addition, although Pro is a cyclic side chain containing amino acid, it possesses three methylene C atoms, still a C atom less than Leu side chain, therefore we planned to include Pro along with Ser in a dipeptide to verify the compatibility and effect of structural diversity in the porous structure formation.

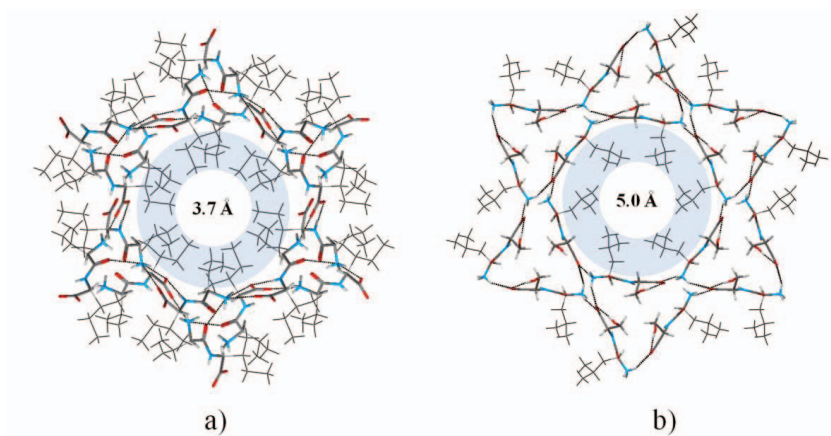


Figure 1.1.4.1 a) Crystal packing of Val-Ile view along the hexagonal axis. Val-Ile is a Val-Ala class member and generates the narrowest pore.⁴³ b) Crystal packing of Leu-Ser. The hydrogen bonded peptide backbones in a) and b) are drawn as capped sticks.

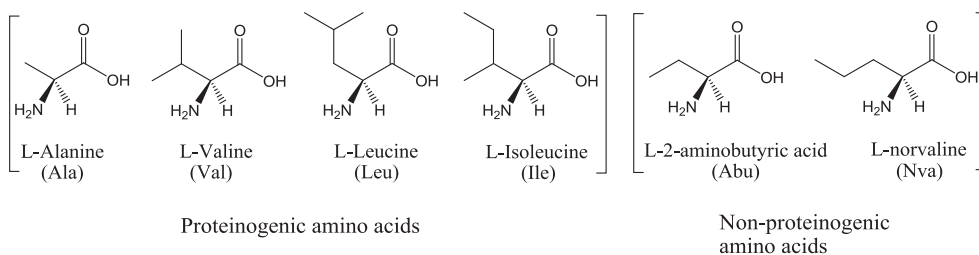


Figure 1.1.4.2: The structures of aliphatic side chain containing proteinogenic and non-proteinogenic amino acids.

Some dipeptide structures due to bulky hydrophobic side chains of both the residues do not leave a space as a pore and instead generate closely-packed structures. e.g. Leu-Ile (ETIWIN)⁵⁰ or (HIZCOJ)⁶² and Ile-Ile (YAGZOW).⁶³ Looking at these molecules we predicted a pore formation if a Leu or Ile or both are replaced by the less bulky ethyl (Abu) or n-propyl (Nva) side chain. Another an interesting fact is that Ala-Ile and Ile-Ala form porous structures, but their structurally isomeric counterparts Ala-Leu (DEZQOO)⁶⁷ and Leu-Ala (RAVMOQ)^{68,92} are nonporous. Hence we thought it would be interesting to check the effect of the combination of the intermediately sized side chain of Abu (ethyl) or Nva (n-propyl) along with the Ile or Leu.

Previously, in the search for larger pores than in the Ala-Val and Val-Ala crystals, the commercially available Ala-Abu and Abu-Ala molecules have been analyzed,^{47,48} but their structures resulted into non-porous molecular networks.

Conversely, the molecules of the Phe-Phe class of dipeptides form hydrophilic pores by pointing hydrophobic side chains away from the channel (Fig. 1.1.3.3), which further limits the pore size fine tuning based on side chain modifications. Also such fabrication in the Phe-Phe (IFABEW), Phe-Trp (GEHTAP) and Trp-Gly (FUGLEY) molecules is almost impossible because of their rigid aromatic side chain properties (Fig. 1.1.3.3a and Fig. 1.1.3.4c). However, the other members in the Phe-Phe class such as Leu-Leu (IDUZOW),⁴⁵ Ile-Leu (ETITUW),⁵⁰ Leu-Phe (IDUZUC)⁴⁵ and Ile-Phe (ETONIK)⁶¹ (Fig. 1.1.4.3) contain residues with aliphatic side chains, thus to some extent we saw room to replace Leu and Ile by Abu or Nva. Therefore, we presumed that these replacements may favor a change in the structures in such a way that the forming network would either; adapt the six molecules in a H-bonded chain to generate a larger pores [(as seen in Phe-Phe (IFABEW))] or the structures with change in their θ and φ torsions, and thus can give rise to hydrophobic channels or the structures with other properties.

According to these strategies for substitutions of bulky amino acids with Abu or Nva one can obtain wider channels with additional surface areas. Incorporation of such fine structural modifications in the Val-Ala, Leu-Ser and Phe-Phe classes presumably may offer numerous assemblies with fine-tuned nanotube diameters for variety of uses, here especially for the CO₂ and methane gas absorption and selectivity studies.

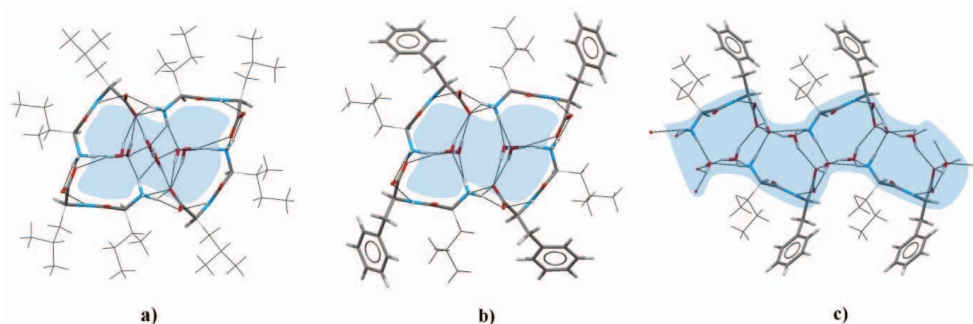


Figure 1.1.4.3: The crystal packing of a) Ile-Leu (ETITUW),⁵⁰ b) Leu-Phe (IDUZUC)⁴⁵ and c) Ile-Phe (ETONIK)⁶¹ molecules. The Leu and Ile side chains have been shown with the black thin wires. Light blue shades in a) and b) represent the water-filled channels, whereas in c) it shows a two-dimensional hydrophilic layer. The hydrogen bonds are shown with the dotted lines.

Here, it is necessary to mention that both Abu and Nva are non-proteinogenic amino acids, but according to their successful use in biological studies they can be considered as biocompatible molecules. A physiological testing showed that Nva has arginase inhibitory property, which regulates the nitric oxide (NO) production in mammals,⁹³ and also acts as an anti-inflammatory agent,⁹⁴ whereas Abu is an active precursor used in anticonvulsant drug molecules such as Brivaracetam,⁹⁵ Levetiracetam⁹⁶ as well as in the antituberculous agent Ethambutol.⁹⁷

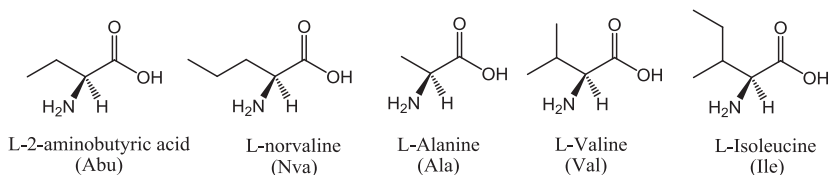
1.1.5 Synthetic Targets as Microporous Dipeptides (Summary of Designed Molecules)

As explained above, this study is a thorough structure-based strategy to obtain different dimensions of nanotubes in dipeptide assemblies. Since the dipeptides composed of Abu or Nva and other standard amino acids are not commercially available, we planned in-house synthesis of such target molecules to demonstrate experimentally the presumptions of various types of pore formation in hydrophobic dipeptides.

1.1.5.1 Val-Ala Class

Table 1.1.5.1: Synthetic targets in the Val-Ala class.

		C-terminal residue				
		Ala	Abu	Nva	Val	Ile
N-terminal residue	Abu	----	Abu-Abu	Abu-Nva	Abu-Val	Abu-Ile
	Nva	Nva-Ala	Nva-Abu	Nva-Nva	Nva-Val	Nva-Ile



1.1.5.2 Leu-Ser Class

Table 1.1.5.2: Synthetic targets in the Leu-Ser class.

		C-terminal residue
		Ser
N-terminal residue	Abu	Abu-Ser
	Nva	Nva-Ser
	Pro	Pro-Ser

L-proline
(Pro)

L-serine
(Ser)

1.1.5.3 Phe-Phe Class

Table 1.1.5.3: Synthetic targets in the Leu-Ser class.

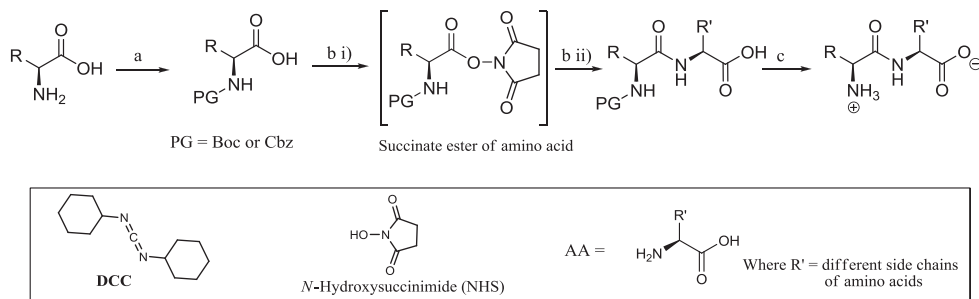
		C-terminal residue	
		Leu	Phe
N-terminal residue	Abu	Abu-Leu	Abu-Phe
	Nva	Nva-Leu	Nva-Phe

L-Leucine
(Leu)

L-phenylalanine
(Phe)

1.1.6 Synthesis, Crystallization and Structure Determination of Dipeptides

There are numerous solid-phase as well as traditional liquid-phase synthetic methods available for the preparation of peptides.^{98,99} For ligation of amino acids here we have used a convenient and efficient solution-phase process by employing *N,N*-dicyclohexylcarbodiimide (DCC) and *N*-hydroxysuccinimide (NHS) (Scheme 1.1.6). These reagents were readily available in the laboratory as left-over chemicals from previous users. Since dicyclohexyl urea (DCU), a by-product formed from DCC after active ester formation of the *N*-terminal residue with NHS, is insoluble in dichloromethane (CH₂Cl₂), it can be filtered off easily. After the peptide coupling reaction, the regenerated and water soluble NHS is also can be readily removed by water washings.¹⁰⁰⁻¹⁰² Thus, we chose to include these reagents in synthetic processes to avoid subsequent laborious purification complications of chemically sensitive peptide molecules. In accordance with the process requirements, the Boc- or Cbz-protected amino acids were used at *N*-terminals. All the non-protected amino acids, *O*-benzyl serine, 10% Pd/C and 4M HCl in dioxane were purchased from Sigma-Aldrich and used as received. For the synthetic and crystallization experimental details see section 1.4



Scheme 1.1.6: General synthetic scheme for the dipeptides in this work.

Reagents and conditions: a) i) (Boc)₂O OR Cbz-Cl b) i) DCC, NHS, CH₂Cl₂, ii) AA, NaHCO₃, H₂O, Acetone, RT; c) i) HCl, RT OR H₂, Pd/C, ii) Aq. NaOH, pH = 7-8.

1.1.6.1 Synthesis of Val-Ala and Phe-Phe Class of Dipeptides

For the synthesis of Val-Ala and Phe-Phe group of molecules (Table 1.1.5.1 and Table 1.1.5.3) where the side chains of amino acids are devoid of any chemically reactive groups, we used common methods.

1.1.6.2 Synthesis of Leu-Ser Class of Dipeptides

The C-terminal residue Ser contains a polar hydroxyl group, thus a benzyl protected serine was used. Also, depending upon the feasibility of the process, the Boc- or Cbz- protected N-terminal residue was employed.

1.1.6.3 Crystal Preparation and Structure Determination

After the synthesis of targeted dipeptide molecules, their crystallization was the main bottleneck in accomplishing this study. Subsequently by trial-and-error of slow evaporation of water/organic solvents (methanol, ethanol, 2-propanol (IPA) etc.) and various solvent diffusion methods for the crystallizations were attempted. We noticed that gel crystallization was the most effective method for small quantities, but large quantities of Nva-Nva, Nva-Val, Nva-Ile and Abu-Ile crystals for CO₂ and methane gas absorption studies were obtained by slow evaporation of water.

Single crystal X-ray diffraction experiments for all crystalline dipeptides were carried out on a Bruker APEXII¹⁰³ CCD detector diffractometer with graphite-monochromatized (Mo K α = 0.71073 Å) radiation at 105 K by continuous flow of liquid nitrogen controlled by an Oxford low-temperature device. Data were collected with an ω scan width of 0.5° at different settings of φ (0°, 90° and 180°) with the detector fixed at $2\theta = -30^\circ$. Data integration/reduction and absorption was carried out by SAINT and SADABS,¹⁰³ respectively and refinement by SHELXTL.¹⁰⁴ All refinement details are described in the respective manuscripts. Molecular and packing illustrations for the thesis and manuscripts were generated by using the program Mercury.⁷⁵

1.2 Results and Discussions : Crystal Structures of Obtained Dipeptides

Val-Ala class: In this category, out of nine synthesized molecules (Table 1.1.5.1), all except Nva-Ala formed crystals of X-ray diffractions quality. All the analyzed structures show a common hexagonal $P6_1$ space group and belong to the porous VA family with different channels sizes and shapes in identical H-bonded networks.

Abu-Ser, Nva-Ser and Pro-Ser (Table 1.1.5.2) were synthesized in substantial quantities. While Abu-Ser and Pro-Ser formed single crystals, Nva-Ser, which was regarded as the most likely additional member of the Leu-Ser class, remained as a powder even after several crystallization attempts by various methods (slow evaporation, solvent diffusion and gel etc).

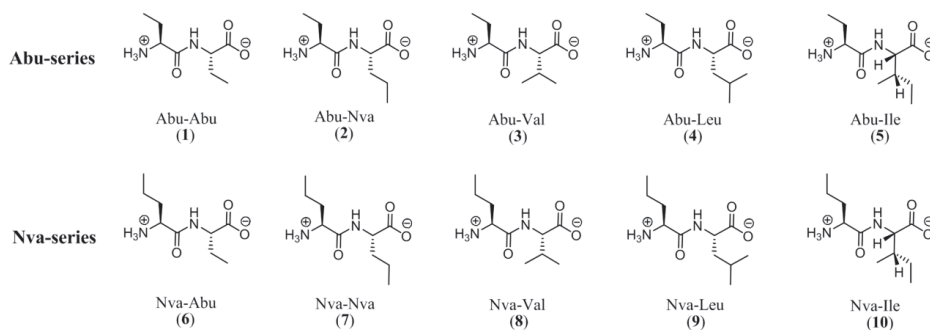
Abu-Leu, Nva-Leu and **Nva-Phe** (Table 1.1.5.2) were obtained as crystalline specimens necessary for X-ray diffraction study, but Abu-Phe did not form crystals. Surprisingly, the structures from X-ray diffraction data revealed that both Abu-Leu and Nva-Leu belong to the Val-Ala class of molecules with equivalent crystallographic properties (hexagonal unit cell, $P6_1$ = space group) (Table 1.2).

Table 1.2: Summary of crystallized dipeptides and their structural properties.

Class	Val-Ala								Leu-Ser		Phe-Phe		
Molecules	Abu- Abu	Abu- Nva	Abu- Val	Abu- Ile	Nva- Abu	Nva- Nva	Nva- Val	Nva- Ile	Abu- Ser	Abu- Pro	Abu- Leu	Nva- Leu	Nva- Phe
Space group	$P6_1$								$I4$	$P2_1$	$P6_1$		$P2_12_12_1$
Unit cell	Hexagonal								Tetra- gonal	Mono- clinic	Hexagonal		Ortho- hombic
Porous/non- porous	Porous								Non-porous		Porous		Non- porous

1.2.1 Val-Ala Type Structures

According to crystal data and three-dimensional structure properties (Table 1.2), we have obtained a total of ten new members of the Val-Ala class. For a convenient description of these molecules in further discussions we renamed them as **1-10**, (Scheme 1.2.2).



Scheme 1.2.2: The Val-Ala class of dipeptides **1-10** in present work

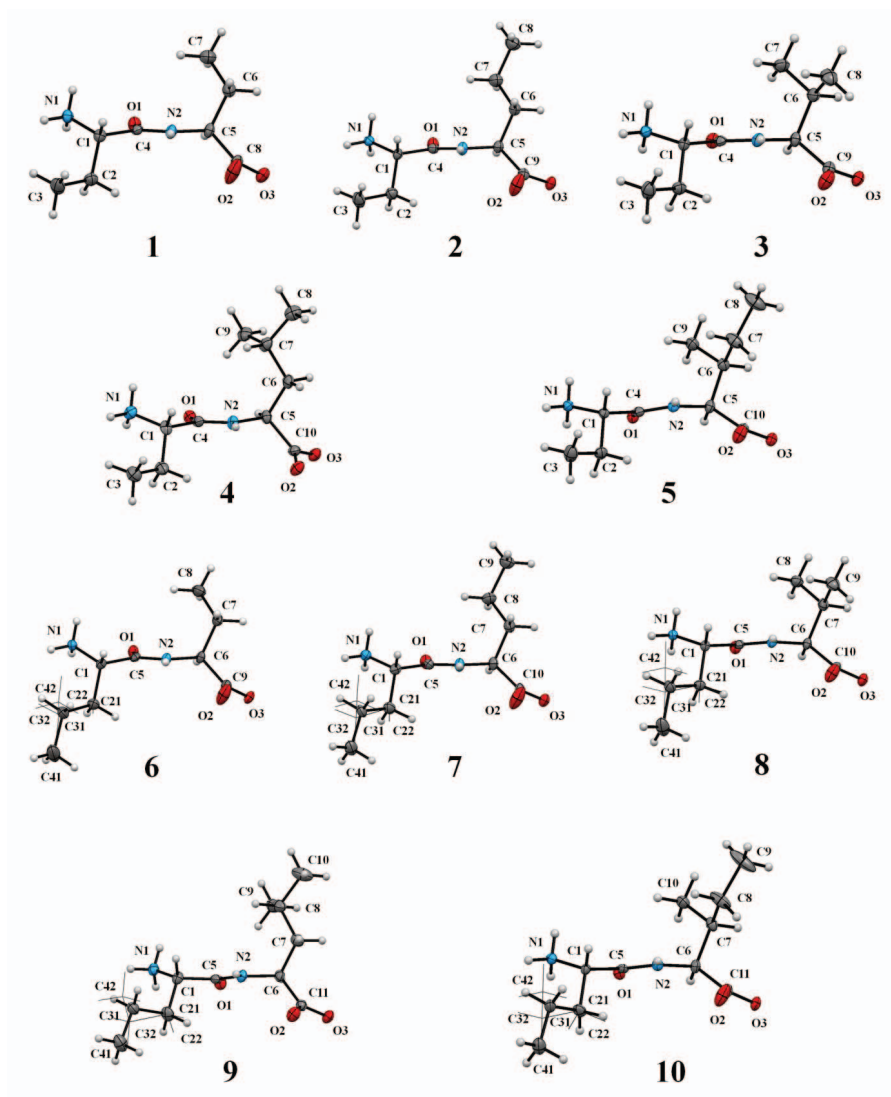
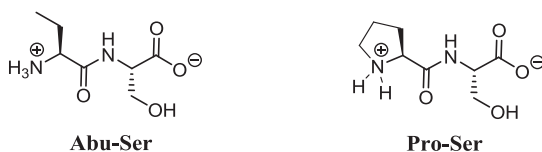


Figure 1.2.2: The atomic labeling schemes of dipeptides **1-10** showing molecular conformations. Thermal ellipsoids are drawn at the 50% probability level. In the Nva series (**6-10**), the *N*-terminal side chain is disordered, and the alternate sites are shown as thin wires.

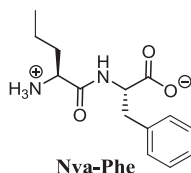
To avoid the repetition, the elaborate details of molecular structures, self-assembly, H-bonding of **1-10**, and CO₂ and methane absorption study results are presented in the **Publication I**.

1.2.2 Structures of Abu-Ser and Pro-Ser



Both Abu-Ser and Pro-Ser form non-porous structures. Abu-Ser yielded an unusual tetragonal $I4$ space group, like the structure of Ala-Ala,⁷⁴ whereas Pro-Ser crystallizes into a chiral monoclinic $P2_1$ space group, which includes an ordered chain of water molecules in the crystal lattice. Details of Abu-Ser and Pro-Ser structures are explained in **Publication II** and **III**, respectively.

1.2.3 Structure of Nva-Phe



Nva-Phe forms a dihydrate in orthorhombic crystals and generates a layered H-bonded network. A complete structural description can be found in **Publication IV**.

1.3 Conclusions

This study constitutes a comprehensive structure based rational design and combinatorial synthetic strategy for detailed investigations of nanoporous materials as well as structural properties of bio-degradable small molecules. Out of a total of 15 synthesized novel dipeptides, 13 were successfully obtained as crystals. All molecules in the Val-Ala class of dipeptides formed nanotubes. This is by far the largest family of nanoporous organic materials known, including 17 different molecules (seven earlier and ten molecules from present study). Such an extensive topology-based approach led to the understanding and confirmation of the optimum micropore formation ability of the dipeptide molecules, especially in the Val-Ala and Leu-Ser classes. The formation of hydrophobic channels in Abu-Leu and Nva-Leu suggest that a small structural change can divert self-assembly into an unexpected very stable open channel formation. Even though Pro-Ser did not form any kind of pores, structurally it was a worthwhile self-assembly due to its inclusion of water wires. This is a one of the bio-

mimicking molecular property showed by Pro-Ser because Pro residue is a member of the significant NPA (Asn-Pro-Ala) motif of aquaporin, a biological water channel.⁷ Besides Ile-Phe (ETONIK), and Val-Phe (MOBYAD)⁶⁹ (Fig. 1.1.4.3 c), the Nva-Phe is an additional new member we found with rare structural property of layer formation. The pore dimensions enhancement upon side chain bulk reduction is not only observed in narrow pore forming members of Val-Ala class, but also this approach generated the porous (Nva-Nva, Abu-Nva, Nva-Abu, Abu-Leu, Nva-Leu, Abu-Ile and Nva-Ile) structures from completely non-porous structures (Ile-Ile or Leu-Leu).

Keywords: hydrophobic dipeptides, nanotubes, bio-degradable material, self-assembly, non-proteinogenic amino acids, hydrophobic channels.

Reference List

- (1) Petsko, G.; Ringe, D. *Protein Structure and Function*; New Science Press, 2004.
- (2) Sewald, N.; Jakubke, H.-D. *Peptides: Chemistry and Biology*; Second Edition, Wiley-VCH, 2009.
- (3) Anfinsen, C. B. *Science* **1973**, *181*, 223-230.
- (4) Brocchieri, L.; Karlin, S. *Nucleic Acids Res.* **2005**, *33*, 3390-3400.
- (5) Hengge, A. C.; Stein, R. L. *Biochemistry* **2003**, *43*, 742-747.
- (6) Foigel, A. *Mol. Cell. Biochem.* **2011**, *352*, 87-89.
- (7) Agre, P. *Angew. Chem. Int. Ed.* **2004**, *43*, 4278-4290.
- (8) MacKinnon, R. *Angew. Chem. Int. Ed.* **2004**, *43*, 4265-4277.
- (9) Cooper, G. M. In *The Cell: A Molecular Approach, Chapter: Actin, Myosin, and Cell Movement*; Sunderland (MA): Sinauer Associates, 2000.
- (10) Tager, H. S.; Steiner, D. F. *Annu. Rev. Biochem.* **1974**, *43*, 509-538.
- (11) Cherezov, V.; Rosenbaum, D. M.; Hanson, M. A.; Rasmussen, S. G. F.; Thian, F. S.; Kobilka, T. S.; Choi, H.-J.; Kuhn, P.; Weis, W. I.; Kobilka, B. K.; Stevens, R. C. *Science* **2007**, *318*, 1258-1265.
- (12) Lohse, M.; Benovic, J.; Codina, J.; Caron, M.; Lefkowitz, R. *Science* **1990**, *248*, 1547-1550.
- (13) Janeway, C. A.; P.Travers, J.; Walport, M.; Shlomchik, M. *Immunobiology*; Garland Science: New York, 2001.
- (14) Craik, D. J.; Fairlie, D. P.; Liras, S.; Price, D. *Chem. Biol. Drug Des.* **2013**, *81*, 136-147.
- (15) Vlieghe, P.; Lisowski, V.; Martinez, J.; Khrestchatsky, M. *Drug Discovery Today* **2010**, *15*, 40-56.
- (16) Edwards, C. M. B.; Cohen, M. A.; Bloom, S. R. *QJM* **1999**, *92*, 1-4.
- (17) Zhang, S. *Materials Today* **2003**, *6*, 20-27.
- (18) Gao, X.; Matsui, H. *Advanced Materials* **2005**, *17*, 2037-2050.
- (19) Hartgerink, J. D.; Clark, T. D.; Ghadiri, M. R. *Chem. Eur. J.* **1998**, *4*, 1367-1372.
- (20) Gazit, E. *Chem. Soc. Rev.* **2007**, *36*, 1263-1269.
- (21) Reches, M.; Gazit, E. *Science* **2003**, *300*, 625-627.

- (22) Ghadiri, M. R.; Granja, J. R.; Milligan, R. A.; McRee, D. E.; Khazanovich, N. *Nature* **1993**, *366*, 324-327.
- (23) Reches, M.; Gazit, E. *Nat Nano* **2006**, *1*, 195-200.
- (24) Rabone, J.; Yue, Y.-F.; Chong, S. Y.; Stylianou, K. C.; Bacsá, J.; Bradshaw, D.; Darling, G. R.; Berry, N. G.; Khimiyak, Y. Z.; Ganin, A. Y.; Wiper, P.; Claridge, J. B.; Rosseinsky, M. J. *Science* **2010**, *329*, 1053-1057.
- (25) Jaime Castillo-León; Andersen, K. B.; Svendsen, a. W. E. *Self-Assembled Peptide Nanostructures for Biomedical Applications: Advantages and Challenges, Biomaterials Science and Engineering, Prof. Rosario Pignatello (Ed.), ISBN: 978-953-307-609-6*, (2011).
- (26) Ranganathan, D.; Lakshmi, C.; Karle, I. L. *J. Am. Chem. Soc.* **1999**, *121*, 6103-6107.
- (27) Woolfson, D. N.; Ryadnov, M. G. *Curr. Opin. Chem. Biol.* **2006**, *10*, 559-567.
- (28) Davie, E. A. C.; Mennen, S. M.; Xu, Y.; Miller, S. J. *Chem. Rev.* **2007**, *107*, 5759-5812.
- (29) Gao, Y.; Zhao, F.; Wang, Q.; Zhang, Y.; Xu, B. *Chem. Soc. Rev.* **2010**, *39*, 3425-3433.
- (30) Dasgupta, A.; Mondal, J. H.; Das, D. *RSC Advances* **2013**.
- (31) Naskar, J.; Roy, S.; Joardar, A.; Das, S.; Banerjee, A. *Org. Biomol. Chem.* **2011**, *9*, 6610-6615.
- (32) Panda, J. J.; Kaul, A.; Kumar, S.; Alam, S.; Mishra, A. K.; Kundu, G. C.; Chauhan, V. S. *Nanomedicine* **2013**, 1-16.
- (33) Matson, J. B.; Stupp, S. I. *Chem. Commun.* **2012**, *48*, 26-33.
- (34) Hosseinkhani, H.; Hong, P.-D.; Yu, D.-S. *Chem. Rev.* **2013**.
- (35) Gazit, E. In *NanoBioTechnology*; Shoseyov, O.; Levy, I. Eds.; Humana Press, 2008; pp. 385-395.
- (36) Cui, Y.; Kim, S. N.; Naik, R. R.; McAlpine, M. C. *Acc. Chem. Res.* **2012**, *45*, 696-704.
- (37) Ulijn, R. V.; Smith, A. M. *Chem. Soc. Rev.* **2008**, *37*, 664-675.
- (38) Matson, J. B.; Zha, R. H.; Stupp, S. I. *Current Opinion in Solid State and Materials Science* **2011**, *15*, 225-235.
- (39) Cui, H.; Webber, M. J.; Stupp, S. I. *Peptide Science* **2010**, *94*, 1-18.
- (40) Luo, Z.; Zhang, S. *Chem. Soc. Rev.* **2012**, *41*, 4736-4754.
- (41) Görbitz, C. H.; Gundersen, E. *Acta Crystallogr. Sect. C* **1996**, *52*, 1764-1767.
- (42) Görbitz, C. H.; Torgersen, E. *Acta Crystallogr. Sect. B* **1999**, *55*, 104-113.
- (43) Görbitz, C. H. *Chem. Eur. J.* **2007**, *13*, 1022-1031.
- (44) Görbitz, C. H.; Nilsen, M.; Szeto, K.; Tangen, L. W. *Chem. Commun.* **2005**, 4288-4290.
- (45) Görbitz, C. H. *Chem. Eur. J.* **2001**, *7*, 5153-5159.
- (46) Görbitz, C. H. *Acta Crystallogr. Sect. B* **2002**, *58*, 849-854.
- (47) Görbitz, C. H. *Acta Crystallogr. Sect. C* **2002**, *58*, o533-o536.
- (48) Görbitz, C. H. *Acta Crystallogr. Sect. E* **2005**, *61*, o3735-o3737.
- (49) Görbitz, C. H. *New J. Chem.* **2003**, *27*, 1789-1793.
- (50) Görbitz, C. H. *Acta Crystallogr. Sect. E* **2004**, *60*, o626-o628.
- (51) Görbitz, C. H. *CrystEngComm* **2005**, *7*, 670-673.
- (52) Görbitz, C. H. *Chem. Commun.* **2006**, *0*, 2332-2334.
- (53) Görbitz, C. H. *Acta Crystallogr. Sect. B* **2010**, *66*, 84-93.
- (54) Görbitz, C. H. *Acta Crystallogr. Sect. B* **2010**, *66*, 253-259.

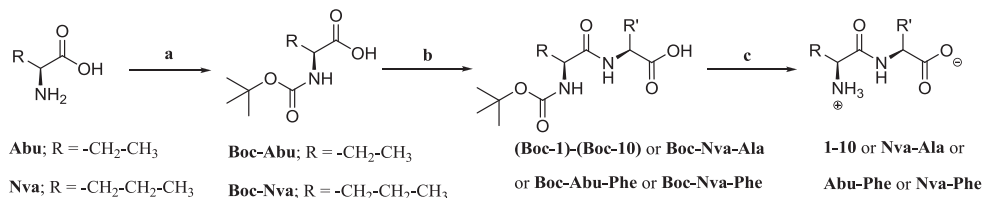
- (55) Görbitz, C. H. *Acta Crystallogr. Sect. C* **2010**, *66*, o531-o534.
- (56) Görbitz, C. H.; Bruvoll, M.; Dizdarevic, S.; Fimland, N.; Hafizovic, J.; Kalfjos, H. T.; Krivokapic, A.; Vestli, K. *Acta Crystallogr. Sect. C* **2006**, *62*, o22-o25.
- (57) Moggach, S. A.; Görbitz, C. H.; Warren, J. E. *CrystEngComm* **2010**, *12*, 2322-2324.
- (58) Johansen, A.; Midtkandal, R.; Roggen, H.; Görbitz, C. H. *Acta Crystallogr. Sect. C* **2005**, *61*, o198-o200.
- (59) Netland, K. A.; Andresen, K.; Görbitz, C. H.; Dalhus, B. *Acta Crystallogr. Sect. E* **2004**, *60*, o951-o953.
- (60) Görbitz, C. H. *Acta Crystallogr. Sect. B* **1999**, *55*, 1090-1098.
- (61) Görbitz, C. H. *Acta Crystallogr. Sect. C* **2004**, *60*, o371-o373.
- (62) Görbitz, C. H.; Rise, F. *J. Pept. Sci.* **2008**, *14*, 210-216.
- (63) Görbitz, C. H. *Acta Crystallogr. Sect. B* **2004**, *60*, 569-577.
- (64) Görbitz, C. H. *Acta Crystallogr. Sect. C* **2000**, *56*, 1496-1498.
- (65) Görbitz, C. H. *Acta Crystallogr. Sect. C* **1999**, *55*, 2171-2177.
- (66) Görbitz, C. H. *Acta Crystallogr. Sect. C* **2003**, *59*, o730-o732.
- (67) Görbitz, C. H. *Acta Crystallogr. Sect. C* **1999**, *55*.
- (68) Görbitz, C. H. *Acta Crystallogr. Sect. C* **1997**, *53*, 736-739.
- (69) Görbitz, C. H. *Acta Crystallogr. Sect. B* **2002**, *58*, 512-518.
- (70) Görbitz, C. H. *Acta Crystallogr. Sect. C* **2006**, *62*, o328-o330.
- (71) Tian, J.; Thallapally, P. K.; McGrail, B. P. *CrystEngComm* **2012**, *14*, 1909-1919.
- (72) Li, Y.; Fu, Z.-Y.; Su, B.-L. *Adv. Funct. Mater.* **2012**, *22*, 4634-4667.
- (73) Allen, F. *Acta Crystallogr. Sect. B* **2002**, *58*, 380-388.
- (74) Fletterick, R. J.; Tsai, C. C.; Hughes, R. E. *J. Phys. Chem.* **1971**, *75*, 918-922.
- (75) Macrae, C. F.; Bruno, I. J.; Chisholm, J. A.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Rodriguez-Monge, L.; Taylor, R.; van de Streek, J.; Wood, P. A. *J. Appl. Crystallogr.* **2008**, *41*, 466-470.
- (76) Soldatov, D. V.; Moudrakovski, I. L.; Grachev, E. V.; Ripmeester, J. A. *J. Am. Chem. Soc.* **2006**, *128*, 6737-6744.
- (77) Jones, P. G.; Falvello, L.; Kennard, O. *Acta Crystallogr. Sect. B* **1978**, *34*, 1939-1942.
- (78) Birkedal, H.; Schwarzenbach, D.; Pattison, P. *Angew. Chem. Int. Ed.* **2002**, *41*, 754-756.
- (79) Comotti, A.; Bracco, S.; Distefano, G.; Sozzani, P. *Chem. Commun.* **2009**, 284-286.
- (80) Comotti, A.; Fraccarollo, A.; Bracco, S.; Beretta, M.; Distefano, G.; Cossi, M.; Marchese, L.; Riccardi, C.; Sozzani, P. *CrystEngComm* **2013**, *15*, 1503-1507.
- (81) Afonso, R. V.; Durão, J.; Mendes, A.; Damas, A. M.; Gales, L. *Angew. Chem. Int. Ed.* **2010**, *49*, 3034-3036.
- (82) Distefano, G.; Comotti, A.; Bracco, S.; Beretta, M.; Sozzani, P. *Angew. Chem. Int. Ed.* **2012**, *51*, 9258-9262, and references cited therein.
- (83) Adler-Abramovich, L.; Reches, M.; Sedman, V. L.; Allen, S.; Tandler, S. J. B.; Gazit, E. *Langmuir* **2006**, *22*, 1313-1320.
- (84) Yan, X.; Zhu, P.; Li, J. *Chem. Soc. Rev.* **2010**, *39*, 1877-1890.
- (85) Yan, X.; Li, J.; Möhwald, H. *Advanced Materials* **2011**, *23*, 2796-2801.
- (86) Reches, M.; Gazit, E. *Current Nanoscience*, *2*, 105-111.

- (87) Xu, K., Characterization and Utilization of Self-Assembled Diphenylalanine Nanotubes: Ph. D. Thesis, University of Nottingham, 2011.
- (88) Zhu, F.; Schulten, K. *Biophys. J.* **2003**, *85*, 236-244.
- (89) Parlett, C. M. A.; Wilson, K.; Lee, A. F. *Chem. Soc. Rev.* **2013**, *42*, 3876-3893.
- (90) Adiga, S.; Curtiss, L.; Elam, J.; Pellin, M.; Shih, C.-C.; Shih, C.-M.; Lin, S.-J.; Su, Y.-Y.; Gittard, S.; Zhang, J.; Narayan, R. *JOM* **2008**, *60*, 26-32.
- (91) Valery, C.; Artzner, F.; Paternostre, M. *Soft Matter* **2011**, *7*, 9583-9594.
- (92) Akazome, M.; Hirabayashi, A.; Takaoka, K.; Nomura, S.; Ogura, K. *Tetrahedron* **2005**, *61*, 1107-1113.
- (93) Chang, C.-I.; Liao, J. C.; Kuo, L. *American Journal of Physiology - Heart and Circulatory Physiology* **1998**, *274*, H342-H348.
- (94) Ming, X.-F.; Rajapakse, A.; Carvas, J.; Ruffieux, J.; Yang, Z. *BMC Cardiovascular Disorders* **2009**, *9*, 12.
- (95) Rogawski, M. A. *Br. J. Pharmacol.* **2008**, *154*, 1555-1557.
- (96) Sanchez, P. E.; Zhu, L.; Verret, L.; Vossel, K. A.; Orr, A. G.; Cirrito, J. R.; Devidze, N.; Ho, K.; Yu, G.-Q.; Palop, J. J.; Mucke, L. *Proc. Natl. Acad. Sci.* **2012**, *109*, E2895-E2903.
- (97) Yendapally, R.; Lee, R. E. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 1607-1611.
- (98) Kimmerlin, T.; Seebach, D. *J. Pept. Res.* **2005**, *65*, 229-260.
- (99) Tsuda, Y.; Okada, Y. In *Amino Acids, Peptides and Proteins in Organic Chemistry*; Wiley-VCH Verlag GmbH & Co. KGaA, 2010; pp. 201-251.
- (100) Anderson, G. W.; Zimmerman, J. E.; Callahan, F. M. *J. Am. Chem. Soc.* **1963**, *85*, 3039-3039.
- (101) Anderson, G. W.; Zimmerman, J. E.; Callahan, F. M. *J. Am. Chem. Soc.* **1964**, *86*, 1839-1842.
- (102) Benoiton, N. L.; Lee, Y. C.; Chen, F. M. F. *Int. J. Pept. Protein Res.* **1993**, *41*, 587-594.
- (103) Bruker APEX2, SAINT-Plus and SADABS (2007), Bruker AXS Inc., Madison, Wisconsin, USA.
- (104) Sheldrick, G. *Acta Crystallogr. Sect. A* **2008**, *64*, 112-122.

1.4 Experimental and Characterization Data

Dipeptides synthesis: By incorporating two non-proteinogenic (Abu and Nva) and various proteinogenic amino acids we have opted for the following synthetic methods (scheme 1 and 2) to prepare the 16 novel dipeptides with substantial yields. The structures of synthesized intermediates and final compounds were confirmed by mass spectrometry (MS), NMR ($^1\text{H-NMR}$ and $^{13}\text{C-NMR}$) and optical rotation (wherever it was necessary). The three-dimensional structures of 13 zwitterionic-crystalline dipeptides have been analyzed by X-ray diffraction methods.

a) Synthesis of Dipeptides from Hydrophobic Amino acids



For **Boc-1** or **1**; R = $-\text{CH}_2\text{-CH}_3$; R' = $-\text{CH}_2\text{-CH}_3$

Boc-2 or **2**; R = $-\text{CH}_2\text{-CH}_3$; R' = $-\text{CH}_2\text{-CH}_2\text{-CH}_3$

Boc-3 or **3**; R = $-\text{CH}_2\text{-CH}_3$; R' = $-\text{CH}(\text{CH}_3)_2$

Boc-4 or **4**; R = $-\text{CH}_2\text{-CH}_3$; R' = $-\text{CH}_2\text{-CH}(\text{CH}_3)_2$

Boc-5 or **5**; R = $-\text{CH}_2\text{-CH}_3$; R' = $-\text{CH}(\text{CH}_3)\text{-CH}_2\text{-CH}_3$

Boc-6 or **6**; R = $-\text{CH}_2\text{-CH}_2\text{-CH}_3$; R' = $-\text{CH}_2\text{-CH}_3$

Boc-7 or **7**; R = $-\text{CH}_2\text{-CH}_2\text{-CH}_3$; R' = $-\text{CH}_2\text{-CH}_2\text{-CH}_3$

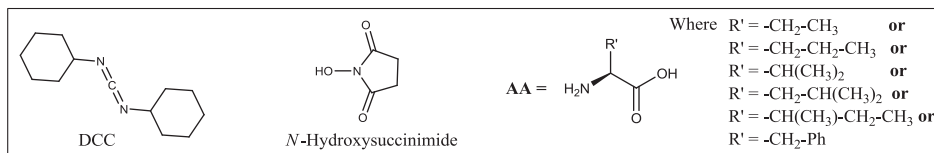
Boc-8 or **8**; R = $-\text{CH}_2\text{-CH}_2\text{-CH}_3$; R' = $-\text{CH}(\text{CH}_3)_2$

Boc-9 or **9**; R = $-\text{CH}_2\text{-CH}_2\text{-CH}_3$; R' = $-\text{CH}_2\text{-CH}(\text{CH}_3)_2$

Boc-10 or **10**; R = $-\text{CH}_2\text{-CH}_2\text{-CH}_3$; R' = $-\text{CH}(\text{CH}_3)\text{-CH}_2\text{-CH}_3$

Nva-Ala or **Boc-Nva-Ala**, R = $-\text{CH}_2\text{-CH}_2\text{-CH}_3$; R' = $-\text{CH}_3$

Abu-Phe or **Boc-Abu-Phe**, R = $-\text{CH}_2\text{-CH}_3$; R' = $-\text{CH}_2\text{-Ph}$; **Nva-Phe** or **Boc-Nva-Phe**, R = $-\text{CH}_2\text{-CH}_2\text{-CH}_3$; R' = $-\text{CH}_2\text{-Ph}$



Scheme 1: Synthesis of dipeptides **1-10**, **Nva-Ala**, **Abu-Phe** and **Nva-Phe**.

Reagents and conditions: a) i) $(\text{Boc})_2\text{O}$, THF, NaOH; b) i) DCC, N-Hydroxysuccinimide, CH_2Cl_2 ,

ii) AA, NaHCO_3 , H_2O , Acetone, RT; c) i) HCl, RT, ii) Aq. NaOH, pH = 7-8.

Synthesis of (*S*)-2-(*tert*-butoxycarbonylamino)butanoic acid (Boc-Abu)

L-2-aminobutyric acid (3.0 g, 29.0 mmol) was added to the chilled ($5\text{-}0\text{ }^\circ\text{C}$) solution of a 1:1 tetrahydrofuran (THF): water mixture (50 mL). To this solution 5 mL aq. NaOH (1.03 g, 2.56 mmol) was added over 10 minutes. At the same temperature, di-*tert*-butyl dicarbonate ($(\text{Boc})_2\text{O}$, 7.7 mL, 35.0 mmol) was added and turbid mixture stirred at room temperature (RT) for 24 h, until it became clear. After the reaction was complete, the THF was removed at $30\text{ }^\circ\text{C}$ under reduced pressure and

remaining aq. residue was washed with diethyl ether (20 ml x 2), and the ether extract discarded. The aq. residue was cooled to 5-0 °C and acidified with saturated KHSO₄ solution (pH 2-3). The acidic aq. mixture was extracted with fresh diethyl ether (30 mL x 3) and the combined ether layers were washed with water and brine solutions (20 mL each) separately and dried over anhydrous MgSO₄ and filtered through Whatmann filter paper. The evaporation of diethylether under vacuum afforded **Boc-Abu**. Yield = 5.5 g (93%).

Optical rotation: $[\alpha]_{589}^{20} = -16.8$ ($c = 1$ g/100 mL, MeOH).

¹H-NMR (300MHz, DMSO-*d*₆), $\delta = 12.45$ (bs, 1H), 7.02-6.99 (d, 1H, $J = 7.4$ Hz), 3.83-3.75 (m, 1H), 1.74-1.51 (m, 2H), 1.38 (s, 9H), 0.90-0.85 (t, 3H, $J = 7.4$ Hz).

¹³C-NMR (75MHz, DMSO-*d*₆), $\delta = 174.07$ (COOH), 155.63 (CO), 77.95 (>C<), 54.91 (CH), 28.24 (3CH₃), 24.17 (CH₂), 10.60 (CH₃).

(S)-2-(tert-butoxycarbonylamino)pentanoic acid (Boc-Nva)

Synthesis of **Boc-Nva** from L-norvaline (3.0 g, 25.6 mmol) and (Boc)₂O (7.0 mL, 30.8 mmol) was achieved according to the procedure reported for Boc-Abu. Yield = 5.3 g (95%).

Optical rotation: $[\alpha]_{589}^{20} = -13.6$ ($c = 1$ g/100 mL, MeOH).

¹H-NMR (300MHz, DMSO-*d*₆), $\delta = 12.39$ (s, 1H), 7.02-7.00 (d, $J = 8.1$ Hz, 1H), 3.90-3.82 (m, 1H), 1.62-1.50 (m, 2H), 1.38-1.30 (m, 11H), 0.88-0.83(t, $J = 7.3$ Hz, 3H).

¹³C-NMR (75MHz, DMSO-*d*₆), $\delta = 174.31$ (COOH), 155.60 (CO), 77.91 (>C<), 53.14 (CH), 32.88 (CH₂), 27.99 (3CH₃), 18.82 (CH₂), 13.48 (CH₃).

(S)-2-((S)-2-(tert-butoxycarbonylamino)butanamido)butanoic acid (Boc-1, Boc-Abu-Abu)

In a 50 mL dry round bottom flask (RBF) under nitrogen atmosphere, 10 mL of dry CH₂Cl₂ (dichloromethane) was charged and cooled to 10 °, then Boc-Abu (0.60 g, 2.95 mmol), dicyclohexylcarbodiimide (DCC) (0.67 g, 3.25 mmol) and *N*-hydroxysuccinimide (NHS, 0.40 g, 3.45 mmol) were added. The reaction mixture was warmed to RT and stirred for 2 h. From reaction suspension, the undissolved byproduct dicyclohexyl urea (DCU) was removed by filtration through a celite bed. The CH₂Cl₂ was removed from the filtrate under reduced pressure to obtain the intermediate Boc-Abu succinate ester (1g), which was then redissolved in acetone (10 mL). In separate flask, L-2-aminobutyric acid (Abu, 0.30 g, 2.95 mmol), NaHCO₃ (0.84 g, 10.0 mmol) and 10 mL of water were charged at RT. To this, the premixed Boc-Abu succinate ester in acetone was added over 15 min and stirred vigorously for 30 min. Acetone from the reaction was evaporated under reduced pressure and residual water was extracted with 20 mL of diethyl ether. The remaining water layer was diluted with the additional 10 mL of water and cooled to 0 °C, and slowly acidified with aq. 1M HCl solution (pH = 2-3). The acidic aq. solution was extracted with diethyl ether (30 mL x 3) and combined extract was washed with brine and water (10 mL each). The ether extract was dried over

anhydrous MgSO₄, filtered, and solvent removed under reduced pressure yielding Boc-Abu-Abu (**Boc-1**). Yield = 0.69 g (72%).

MS Cl⁺ *m/z* (relative intensity in %) = [M+H]⁺ 289 (50), 233 (85), 189 (100), 58 (58)

¹H-NMR (300MHz, CDCl₃-*d*), δ = 9.00 (bs, 1H), 6.90-6.87 (d, *J* = 7.4 Hz, 1H), 5.31-5.28 (m, 1H), 4.59-4.53 (m, 1H), 4.12-3.09 (d, *J* = 8.5 Hz, 1H), 2.03-1.57 (m, 4H), 1.43 (s, 9H), 0.97-0.91 (m, 6H).

¹³C-NMR (75MHz, CDCl₃-*d*), δ = 175.09 (COOH), 172.50 (CON), 156.12 (CO), 74.83 (>C<), 55.97 (CH), 53.46 (CH), 28.42 (3CH₃), 25.76 (CH₂), 25.38 (CH₂), 10.08 (CH₃), 9.57 (CH₃).

(S)-2-((S)-2-(tert-butoxycarbonylamino)butanamido)pentanoic acid (Boc-2, Boc-Abu-Nva)

Boc-2 was synthesized from Boc-Abu (0.60 g, 2.95 mmol) and L-norvaline (0.35 g, 3.0 mmol) according to the procedure reported for Boc-1. Yield = 0.77 g (77%).

MS Cl⁺ *m/z* (relative intensity in %) = [M+H]⁺ 303 (35), 247 (85), 203 (100), 58 (50).

¹H-NMR (300MHz, CDCl₃-*d*) δ = 9.18 (s, 1H), 6.92-6.89 (d, *J* = 7.8 Hz, 1H), 5.37-5.34 (m, 1H), 4.65-4.58 (m, 1H), 4.13-4.10 (d, *J* = 8.1 Hz, 1H), 1.95-1.61 (m, 4H), 1.46-1.33 (m, 11H), 1.01-0.91 (m, 6H).

¹³C-NMR (75MHz, CDCl₃-*d*) δ = 175.52 (COOH), 172.55 (CON), 156.14 (CO), 80.51 (>C<), 55.81 (CH), 52.20 (CH), 34.29 (CH₂), 28.40 (3CH₃), 25.67 (CH₂), 18.64 (CH₂), 13.77 (CH₃), 10.07 (CH₃).

(S)-2-((S)-2-(tert-butoxycarbonylamino)butanamido)-3-methylbutanoic acid (Boc-3, Boc-Abu-Val)

Boc-3 was synthesized from Boc-Abu (0.60 g, 2.95 mmol) and L-valine (0.35 g, 3.0 mmol) according to the procedure reported for Boc-1. Yield = 0.81 g (81%).

MS Cl⁺ *m/z* (relative intensity in %) = [M+H]⁺ 303 (57), 247 (100), 203 (100), 58 (50).

¹H-NMR (300MHz, CDCl₃-*d*) δ = 6.88-6.86 (d, *J* = 8.7 Hz, 1H), 6.48 (bs, 1H), 5.30-5.28 (d, *J* = 5.0 Hz, 1H), 4.61-5.56 (m, 1H), 4.13-4.10 (d, *J* = 8.2 Hz, 1H), 2.30-2.21 (m, 1H), 1.91-1.81 (m, 1H), 1.71-1.61 (m, 1H), 1.45 (s, 9H), 1.02-0.94 (m, 9H).

¹³C-NMR (75MHz, CDCl₃-*d*) δ = 174.90 (COOH), 172.66 (CON), 156.19 (CO), 80.49 (>C<), 57.19 (CH), 56.04 (CH), 31.19 (CH), 28.43 (3CH₃), 25.54 (CH₂), 19.11 (CH₃), 17.70 (CH₃), 10.13 (CH₃).

(S)-2-((S)-2-(tert-butoxycarbonylamino)butanamido)-4-methylpentanoic acid (Boc-4, Boc-Abu-Leu)

Boc-4 was synthesized from Boc-Abu (0.60 g, 2.95 mmol) and L-leucine (0.39 g, 3.00 mmol) according to the procedure reported for Boc-1. Yield = 0.84 g (79%).

MS Cl⁺ *m/z* (relative intensity in %) = [M+H]⁺ 317 (55), 261 (96), 217 (100), 58 (50).

¹H-NMR (300MHz, CDCl₃-*d*) δ = 9.86 (bs, 1H), 6.81-6.79 (d, *J* = 8.7 Hz, 1H), 5.34-5.31 (d, *J* = 8.2 Hz, 1H), 4.64- 4.57 (m, 1H), 4.08-4.06 (d, *J* = 8.5 Hz, 1H), 1.90-1.53 (m, 5H), 1.43 (s, 9H), 0.96-0.92 (m, 9H).

¹³C-NMR (75MHz, CDCl₃-*d*) δ = 176.10 (COOH), 172.65 (CON), 156.15 (CO), 80.51 (>C<), 55.89 (CH), 50.90 (CH), 41.29 (CH₂), 28.40 (3CH₃), 24.91 (CH₂), 23.03 (CH), 21.89 (CH₃), 10.07 (CH₃).

(2*S*, 3*S*)-2-((*S*)-2-(tert-butoxycarbonylamino)butanamido)-3-methylpentanoic acid (Boc-5, Boc-Abu-Ileu)

Boc-5 was synthesized from Boc-Abu (0.58 g, 2.85 mmol) and L-isoleucine (0.38 g, 2.90 mmol) according to the procedure reported for Boc-1. Yield = 0.74 g (81%).

MS Cl⁺ *m/z* (relative intensity in %) = [M+H]⁺ 317 (52), 261 (90), 217 (99), 58 (48).

¹H-NMR (300MHz, CDCl₃-*d*) δ = 9.37 (bs, 1H), 6.89-6.86 (d, *J* = 8.6 Hz, 1H), 5.32-5.29 (d, *J* = 8.6 Hz, 1H), 4.62-4.58 (m, 1H), 4.11-4.06 (m, 1H), 2.03-1.89 (m, 1H), 1.87-1.75 (m, 1H), 1.72-1.56 (m, 1H), 1.53-1.43 (m, 10H), 1.28-1.13 (m, 1H), 0.96-0.89 (m, 9H).

¹³C-NMR (75MHz, CDCl₃-*d*) δ = 175.04 (COOH), 172.52 (CON), 156.16 (CO), 80.49 (>C<), 56.62 (CH), 55.93 (CH), 37.79 (CH), 28.41 (3CH₃), 25.52 (CH₂), 25.05 (CH₂), 15.52 (CH₃), 11.71 (CH₃), 10.11 (CH₃).

(*S*)-2-((*S*)-2-(tert-butoxycarbonylamino)pentanamido)butanoic acid (Boc-6, Boc-Nva-Abu)

Boc-6 was synthesized from Boc-Nva (0.50 g, 2.30 mmol) and L-2-aminobutyric acid (0.24 g, 2.33 mmol) according to the procedure reported for Boc-1. Yield = 0.60 g (84%).

MS Cl⁺ *m/z* (relative intensity in %) = [M+H]⁺ 303 (57), 247 (80), 203 (100), 72 (48).

¹H-NMR (300MHz, DMSO-*d*₆) δ = 12.62 (bs, 1H), 7.88-7.86 (d, *J* = 7.6 Hz, 1H), 6.83-6.80 (d, *J* = 7.6 Hz, 1H), 4.16-4.09 (m, 1H), 3.98-3.90 (m, 1H), 1.78-1.20 (m, 15 H) 0.89-0.82 (m, 6H).

¹³C-NMR (75MHz, DMSO-*d*₆) δ = 173.36 (COOH), 172.25 (CON), 155.26 (CO), 77.98 (>C<), 53.91 (CH), 52.98 (CH), 34.12 (CH₂), 28.19 (3CH₃), 24.43 (CH₂), 18.62 (CH₂), 13.69 (CH₃), 10.08 (CH₃).

(*S*)-2-((*S*)-2-(tert-butoxycarbonylamino)pentanamido)pentanoic acid (Boc-7, Boc-Nva-Nva)

Boc-7 was synthesized from Boc-Nva (0.52 g, 0.24 mmol) and L-norvaline (0.28 g, 0.25 mmol) according to the procedure reported for Boc-1. Yield = 0.59 g (77%).

MS Cl⁺ *m/z* (relative intensity in %) = [M+H]⁺ 317 (43), 261 (70), 217 (100), 72 (61).

¹H-NMR (300MHz, CDCl₃-*d*) δ = 8.68 (bs, 1H), 6.88-6.85 (d, *J* = 8 Hz, 1H), 5.27-5.24 (d, *J* = 8 Hz, 1H), 4.62-4.56 (m, 1H), 4.16-4.14 (d, *J* = 8.1 Hz, 1H), 1.93-1.51(m, 4H), 1.44-1.31 (m, 13H), 0.94-0.89 (t, 6H).

¹³C-NMR (75MHz, CDCl₃-*d*) δ = 175.50 (COOH), 172.70 (CON), 156.11 (CO), 80.55 (>C<), 54.45 (CH), 52.19 (CH), 34.30 (CH₂), 28.40 (3CH₃), 18.96 (CH₂), 18.63 (CH₂), 13.85 (CH₃), 13.78 (CH₃).

(S)-2-((S)-2-(tert-butoxycarbonylamino)pentanamido)-3-methylbutanoic acid (Boc-8, Boc-Nva-Val)

Boc-8 was synthesized from Boc-Nva (0.52 g, 0.24 mmol) and L-valine (0.28 g, 0.25 mmol) according to the procedure reported for Boc-1. Yield = 0.59 g (77%).

MS Cl^+ m/z (relative intensity in %) = $[\text{M}+\text{H}]^+$ 317 (52), 261 (73), 217 (100), 72 (53).

$^1\text{H-NMR}$ (300MHz, CDCl_3 - d) δ = 9.19 (bs, 1H), 6.90-6.87 (d, J = 8.2 Hz, 1H), 5.30-5.28 (d, J = 8.3 Hz, 1H), 4.59-4.55 (m, 1H), 4.20-4.15 (d, J = 8.0 Hz, 1H), 2.26-2.21 (m, 1H), 1.81-1.71 (m, 1H), 1.64-1.51 (m, 1H), 1.44-1.33 (m, 11H), 0.97-0.89 (m, 9H).

$^{13}\text{C-NMR}$ (75MHz, CDCl_3 - d) δ = 174.94 (COOH), 172.85 (CON), 156.16 (CO), 80.56 ($>\text{C}<$), 57.15 (CH), 54.59 (CH), 34.21 (CH), 31.21 (CH_2), 28.42 (3CH_3), 19.11 (CH_3), 19.00 (3CH_3), 17.62 (CH_2), 13.86 (CH_3).

(S)-2-((S)-2-(tert-butoxycarbonylamino)pentanamido)-4-methylpentanoic acid (Boc-9, Boc-Nva-Leu)

Boc-9 was synthesized from Boc-Nva (0.60 g, 0.28 mmol) and L-leucine (0.36 g, 0.28 mmol) according to the procedure reported for Boc-1. Yield = 0.70 g (76%).

MS Cl^+ m/z (relative intensity in %) = $[\text{M}+\text{H}]^+$ 331 (15), 275 (65), 231 (100), 72 (58).

$^1\text{H-NMR}$ (300MHz, CDCl_3 - d) δ = 8.24 (s, 1H), 6.86-6.83 (m, 1H), 5.31-5.29 (d, J = 6.2 Hz, 1H), 4.67-4.60 (m, 1H), 4.17-4.14 (d, J = 8.3 Hz, 1H), 1.85-1.52 (m, 5H), 1.46-1.33 (m, 11H), 0.99-0.91 (m, 9H).

$^{13}\text{C-NMR}$ (75MHz, CDCl_3 - d) δ = 176.06 (COOH), 172.80 (CON), 156.19 (CO), 80.54 ($>\text{C}<$), 54.49 (CH), 50.87 (CH), 41.33 (CH_2), 34.28 (CH), 28.39 (3CH_3), 24.90 (CH_3), 23.03 (CH_3), 21.90 (CH_2), 18.94 (CH_2), 13.84 (CH_3).

(2S, 3S)-2-((S)-2-(tert-butoxycarbonylamino)pentanamido)-3-methylpentanoic acid (Boc-10, Boc-Nva-Ileu)

Boc-10 was synthesized from Boc-Nva (0.50 g, 0.23 mmol) and L-isoleucine (0.30 g, 0.23 mmol) according to the procedure reported for Boc-1. Yield = 0.65 g (85%).

MS ES^+ m/z (relative intensity in %) = $[\text{M}+\text{Na}]^+$ 353 (100), 253 (48), 231 (8).

$^1\text{H-NMR}$ (300MHz, CDCl_3 - d) δ = 9.73 (bs, 1H), 6.89-6.86 (d, J = 7.7 Hz, 1H), 5.27-5.24 (d, J = 7.7 Hz, 1H), 4.63- 4.58 (m, 1H), 4.16-4.14 (d, J = 6.2 Hz, 1H), 2.00-1.89 (m, 1H), 1.83-1.69 (m, 1H), 1.64-1.30 (m, 13H), 1.26-1.13 (m, 1H), 0.95- 0.89 (m, 9H).

$^{13}\text{C-NMR}$ (75MHz, CDCl_3 - d) δ = 175.06 (COOH), 172.68 (CON), 156.12 (CO), 80.51 ($>\text{C}<$), 56.59 (CH), 54.55 (CH), 37.81 (CH), 34.23 (CH_2), 28.41 (3CH_3), 25.01 (CH_2), 18.99 (CH_2), 15.52 (CH_3), 13.86 (CH_3), 11.72 (CH_3).

(S)-2-((S)-2-(tert-butoxycarbonylamino)pentanamido)propanoic acid (Boc-Nva-Ala)

Boc-Abu-Ala was prepared from **Boc-Nva** (0.15 g, 0.7 mmol) and alanine (0.06 g, 0.7 mmol) according to the procedure reported for Boc-1. Yield = 0.14 g (67%).

$^1\text{H-NMR}$ (300MHz, CDCl_3 -*d*) δ = 7.30 (bs, 3H), 6.98-6.96 (d, J = 8.4 Hz, 1H), 5.29-5.26 (d, J = 8.5 Hz, 1H), 4.63-4.54 (m, 1H), 4.19-4.17 (m, 1H), 1.83-1.73 (m, 1H), 1.66-1.53 (m, 1H), 1.48-1.35 (m, 12H), 0.96-0.91 (t, J = 7.3 Hz, 3H).

$^{13}\text{C-NMR}$ (75MHz, CDCl_3 -*d*) δ = 175.62 (COOH), 172.61 (CON), 145.68 (CO), 79.96 (>C<), 48.33 (CH), 34.67 (CH), 28.41 (3CH₃), 18.94 (CH₃), 18.19 (CH₂), 18.15 (CH₂), 13.85 (CH₃).

(S)-2-((S)-2-aminobutanamido)butanoic acid (1, Abu-Abu)

Boc-1 (Boc-Abu-Abu) (0.65 g, 2.56 mmol) was dissolved in 20 mL of dry THF and cooled to 5-0 °C then 2 mL of 4M HCl in dioxane solution was added slowly. This mixture was stirred for 48 h at RT. Completion of reaction was monitored by the thin layer chromatography (TLC, eluent system = 20% MeOH in CH_2Cl_2 followed by ninhydrin staining). *In vacuo* evaporation of THF and dioxane yielded crude **1**. The syrupy crude **1** was dissolved in 10 mL of water and extracted with diethyl ether (20 mL x 2) to remove the unwanted non-polar impurities. The aq. solution was neutralized with 1N NaOH (pH = 7-8). From this, a small amount of solution was used for crystallization (will be described in forthcoming section). The slow evaporation of water from the remaining solution at RT for over 2-3 weeks afforded solid crude product, which was slurried in a minimum volume of water (5 mL) and filtered to remove the inorganic substances, affording the pure product **1** (Yield = 0.33 g, 77%).

MS ES⁺ m/z (relative intensity in %) = [M+H]⁺ 189 (100), 41 (9).

$^1\text{H-NMR}$ (300MHz, $\text{DMSO-}d_6$ + D₂O) δ = 4.14-4.09 (m, 1H), 3.76-3.72 (t, J = 6.2 Hz, 1H), 1.82-1.47 (m, 4H), 0.91-0.86 (m, 6H).

$^{13}\text{C-NMR}$ (75MHz, $\text{DMSO-}d_6$ + D₂O) δ = 173.75 (COO⁻), 169.23 (CON), 54.26 (CH), 53.62 (CH), 25.01 (CH₂), 24.65 (CH₂), 10.96 (CH₃), 9.33 (CH₃).

(S)-2-((S)-2-aminobutanamido)pentanoic acid (2, Abu-Nva)

Compound **2** was synthesized from **Boc-2** (0.75 g, 2.48 mmol) according to the procedure reported for **1**. Yield = 0.38 g (75%).

MS ES⁺ m/z (relative intensity in %) = [M+H]⁺ 203 (100), 142 (7).

$^1\text{H-NMR}$ (300MHz, $\text{DMSO-}d_6$ + D₂O) δ = 4.20-4.15 (m, 1H), 3.74-3.70 (t, J = 6.2 Hz, 1H), 1.82-1.21 (m, 6H), 0.91-0.82 (m, 6H).

$^{13}\text{C-NMR}$ (75MHz, $\text{DMSO-}d_6$ + D₂O) δ = 173.87 (COO⁻), 169.10 (CON), 53.53 (CH), 52.47 (CH), 33.22 (CH₂), 24.93 (CH₂), 19.22 (CH₂), 14.01 (CH₃), 9.27 (CH₃).

(S)-2-((S)-2-aminobutanamido)-3-methylbutanoic acid (3, Abu-Val)

Compound **3** was synthesized from **Boc-3** (0.75 g, 2.48 mmol) according to the procedure reported for **1**. Yield = 0.40 g (80%).

MS ES⁺ m/z = [M+H]⁺ 203.

¹H-NMR (400MHz, DMSO-*d*₆ + D₂O) δ = 8.55-8.53 (d, J = 8.1 Hz, 1H), 8.19 (bs, 3H), 4.16- 4.13 (m, 1H), 3.82-3.79 (t, J = 6.4 Hz, 1H), 2.13-2.04 (m, 1H), 1.81-1.69 (m, 2H), 0.90 (t, J = 7.6 Hz, 9H).

¹³C-NMR (100MHz, DMSO-*d*₆) δ = 173.30 (COO⁻), 169.52 (CON), 58.45 (CH), 53.64 (CH), 30.16 (CH), 25.15 (CH₂), 19.64 (CH₃), 18.57 (CH₃), 9.37 (CH₃).

(S)-2-((S)-2-aminobutanamido)-4-methylpentanoic acid (4, Abu-Leu)

Compound **4** was synthesized from **Boc-4** (0.80 g, 2.53 mmol) according to the procedure reported for **1**. Yield = 0.46 g (83%).

MS ES⁺ m/z = [M+H]⁺ 217.

¹H-NMR (300MHz, DMSO-*d*₆ + D₂O) δ = 4.23-4.18 (m, 1H), 3.72-3.68 (t, J = 6.1 Hz, 1H), 1.79-1.48 (m, 5H), 0.91-0.81 (m, 9 H).

¹³C-NMR (75MHz, DMSO-*d*₆ + D₂O) δ = 174.47 (COO⁻), 169.27 (CON), 53.71 (CH), 51.31 (CH), 40.36 (CH₂), 25.02 (CH₃), 24.99 (CH₃), 23.48 (CH₂), 21.94 (CH), 9.33 (CH₃).

(2S, 3S)-2-((S)-2-aminobutanamido)-3-methylpentanoic acid (5, Abu-Ileu)

Compound **5** was synthesized from **Boc-5** (0.70 g, 2.21) according to the procedure reported for **1**.

Yield = 0.41 g (85%).

MS ES⁺ m/z = [M+H]⁺ 217.

¹H-NMR (300MHz, DMSO-*d*₆) δ = 8.57 (d, J = 7.9 Hz, 1H), 8.18 (bs, 3H), 4.18-4.14 (m, 1H), 3.51-3.54 (m, 1H), 1.83-1.66 (m, 3H), 1.45-1.33 (m, 1H), 1.27-1.12 (m, 1H), 0.90-0.80 (m, 9H).

¹³C-NMR (75MHz, DMSO-*d*₆) δ = 172.94 (COO⁻), 169.20 (CON), 57.19 (CH), 53.40 (CH), 36.55 (CH), 25.13 (CH₂), 24.98 (CH₂), 16.00 (CH₃), 11.79 (CH₃), 9.23 (CH₃).

(S)-2-((S)-2-aminopentanamido)butanoic acid (6, Nva-Abu)

Compound **6** was synthesized from **Boc-6** (0.58 g, 1.92 mmol) according to the procedure reported for **1**. Yield = 0.32 g (82%).

MS ES⁺ m/z = [M+H]⁺ 203.

¹H-NMR (300MHz, DMSO-*d*₆ + D₂O) δ = 3.82-3.77 (m, 2H), 1.74-1.60 (m, 4H), 1.41-1.23 (m, 2H), 0.89-0.82 (m, 6H).

¹³C-NMR (75MHz, DMSO-*d*₆ + D₂O) δ = 167.88 (COO⁻), 167.56 (CON), 54.98 (CH), 53.89 (CH), 35.38 (CH₂), 26.04 (CH₂), 17.52 (CH₂), 13.78 (CH₃), 8.99 (CH₃).

(S)-2-((S)-2-aminopentanamido)pentanoic acid (7, Nva-Nva)

Compound **7** was synthesized from **Boc-7** (0.56 g, 1.77 mmol) according to the procedure reported for **1**. Yield = 0.33 g (86%).

MS ES⁺ m/z = [M+H]⁺ 217.

¹H-NMR (300MHz, DMSO-*d*₆) δ = 8.71-8.69 (d, *J* = 6.2 Hz, 1H), 8.20 (bs, 3H), 4.22-4.15 (m, 1H), 3.79-3.75 (m, 1H), 1.74-1.53 (m, 4H), 1.40-1.24 (m, 4H), 0.88-0.82 (m, 6H).

¹³C-NMR (75MHz, DMSO-*d*₆) δ = 173.52 (COO⁻), 169.11(CON), 52.30 (CH), 52.23 (CH), 33.58 (CH₂), 33.10 (CH₂), 18.95 (CH₂), 17.64 (CH₂), 14.00 (CH₃), 13.78 (CH₃).

(S)-2-((S)-2-aminopentanamido)-3-methylbutanoic acid (8, Nva-Val)

Compound **8** was synthesized from **Boc-8** (0.55 g, 1.74 mmol) according to the procedure reported for **1**. Yield = 0.30 g (80%).

MS ES⁺ m/z = [M+H]⁺ 217.

¹H-NMR (300MHz, DMSO-*d*₆ + D₂O) δ = 4.14-4.09 (m, 1H), 3.87-3.82 (m, 2H), 2.10-1.99 (m, 1H), 1.69-1.62 (m, 2H), 1.35-1.23 (m, 2H), 0.88-0.80 (m, 9H).

¹³C-NMR (75MHz, DMSO-*d*₆ + D₂O) δ = 173.12 (COO⁻), 169.60 (CON), 58.28 (CH), 52.54 (CH), 33.87 (CH₂), 30.14 (CH), 19.58 (CH₃), 18.52 (CH₃), 17.91 (CH₂), 14.20 (CH₃).

(S)-2-((S)-2-aminopentanamido)-4-methylpentanoic acid (9, Nva-Leu)

Compound **9** was synthesized from **Boc-9** (0.65 g, 1.96 mmol) according to the procedure reported for **1**. Yield = 0.36 g (79%).

MS ES⁺ m/z = [M+H]⁺ 231.

¹H-NMR (300MHz, DMSO-*d*₆ + D₂O) δ = 4.23-4.16 (m, 1H), 3.76-3.72 (t, *J* = 6.2 Hz, 1H), 1.70-1.19 (m, 7H), 0.85-0.78 (m, 9H).

¹³C-NMR (75MHz, DMSO-*d*₆ + D₂O) δ = 175.00 (COO⁻), 169.91 (CON), 53.17 (CH), 51.73 (CH), 40.36 (CH₂), 34.09 (CH₂), 25.18 (CH), 23.61 (CH₃), 22.11 (CH₃), 18.17 (CH₂), 14.45 (CH₃).

(2S, 3S)-2-((S)-2-aminopentanamido)-3-methylpentanoic acid (10, Nva-Ileu)

Compound **10** was synthesized from **Boc-10** (0.62 g, 1.87 mmol) according to the procedure reported for **1**. Yield = 0.36 g (83%).

MS ES⁺ m/z = [M+H]⁺ 231.

¹H-NMR (300MHz, DMSO-*d*₆ + D₂O) δ = 4.15-4.12 (m, 1H), 3.84-3.81 (m, 1H), 1.80-1.74 (m, 1H), 1.68-1.62 (m, 2H), 1.40-1.10 (m, 4H), 0.84-0.78 (m, 9H).

¹³C-NMR (75MHz, DMSO-*d*₆ + D₂O) δ = 173.23 (COO⁻), 169.62 (CON), 57.43 (CH), 52.65 (CH), 36.73 (CH), 33.94 (CH₂), 25.34 (CH₂), 17.99 (CH₂), 16.16 (CH₃), 14.27 (CH₃), 11.95 (CH₃).

(S)-2-((S)-2-aminopentanamido)propanoic acid (Nva-Ala)

Nva-Abu was prepared from **Boc-Nva-Ala** (0.06 g, 0.47 mmol) according to the procedure reported for **1**. Yield = 0.7 g (79%).

MS ES⁺ m/z = [M+H]⁺ 189.

¹H-NMR (300MHz, DMSO-*d*₆) δ = 8.82-8.80 (d, J = 6.9 Hz, 1H), 8.33- 8.11 (m, 3H), 4.26-4.16 (m, 1H), 3.77- 3.73 (m, 1H), 1.69-1.63 (m, 2H), 1.32-1.24 (m, 5H), 0.86-81 (t, J = 7.3 Hz, 3H).

¹³C-NMR (75MHz, DMSO-*d*₆) δ = 174.47 (COO⁻), 169.25 (CON), 52.81(CH), 48.59 (CH), 40.41, 33.96 (CH₂), 18.07 (CH₃), 17.58 (CH₂), 14.45 (CH₃).

Crystallization of dipeptides 1-10: X-ray quality crystals of dipeptides **1-10** were obtained by a gel crystallization diffusion method. The neutralized aq. solution (100 μ L) of each dipeptide was mixed with tetramethoxysilane (50 μ L) in different small tubes and shaken vigorously for about 1 min. The tubes were then sealed with parafilm and the seal pierced by a sharp needle to make 3-4 holes. These tubes were placed in the air tight bottles containing acetonitrile (2-3 ml) allowing slow diffusion into the tubes. After 3-4 days the crystals of **1-10** were seen as colourless needles in the respective tubes.

Synthesis of Abu-Phe and Nva-Phe

(S)-2-((S)-2-(tert-butoxycarbonylamino)butanamido)-3-phenylpropanoic acid (Boc-Abu-Phe)

Boc-Abu-Phe was synthesized from **2a** (0.27 g, 1.33 mmol) and L-Phenylalanine (0.22 g, 1.33 mmol) according to the procedure reported for **3a**. Yield = 0.20 g (43%).

CI⁺ m/z (relative intensity in %) = [M+H]⁺ 351 (60), 251 (100).

¹H-NMR (300MHz, CDCl₃-*d*) δ = 7.42 (bs, 2H), 7.32-7.12 (m, 5H), 6.85-6.83 (d, 1H), 5.23-5.20 (d, 1H), 4.14-4.06 (m, 1H), 3.24-3.18 (m, 1H), 3.07-3.00 (m, 1H), 1.85-1.71 (m, 1H), 1.59-1.55 (m, 1H), 1.46 (s, 9H), 0.92-0.87 (t, J = 6.3 Hz, 3H).

¹³C-NMR (75MHz, CDCl₃-*d*) δ = 174.28 (COO⁻), 172.47 (CON), 159.47 (CO), 136.27 (Ph C), 129.81 (Ph C), 128.90 (Ph C), 127.43 (Ph C), 80.50 (>C<), 56.08 (CH), 53.59 (CH), 37.87 (CH₂), 28.69 (3CH₃), 26.06 (CH₂), 10.26 (CH₃).

(S)-2-((S)-2-aminobutanamido)-3-phenylpropanoic acid (Abu-Phe)

Abu-Phe was prepared from Boc-Abu-Phe (0.20 g, 0.57 mmol) according to the procedure reported for **1**. Yield = 0.12 g (84%).

MS ES⁺ m/z = [M+H]⁺ 251.

¹H-NMR (300MHz, DMSO-*d*₆ + D₂O), δ = 7.23-7.15 (m, 5H), 4.42-4.06 (m, 2H), 3.10-3.03 (m, 2H), 2.03-1.90 (m, 1H), 1.66-1.53 (m, 1H), 0.81-0.78 (t, J = 6.2 Hz, 3H).

^{13}C -NMR (75MHz, $\text{DMSO-}d_6 + \text{D}_2\text{O}$), $\delta = 160.70$ (COO^-), 164.25 (CON), 136.23 (Ph C), 130.50 (Ph C), 128.49 (Ph C), 53.76 (CH), 52.56 (CH), 35.35 (CH_2), 29.07 (CH_2), 11.49. (CH_3).

(S)-2-((S)-2-(tert-butoxycarbonylamino)pentanamido)-3-phenylpropanoic acid (Boc-Nva-Phe)

Boc-Nva-Phe was synthesized from **Boc-Nva** (0.11 g, 0.50 mmol) and L-phenylalanine (0.85 g, 0.51 mmol) according to the procedure reported for **Boc-1**. Yield = 0.12 g (63%).

MS $\text{ES}^+ m/z$ (relative intensity in %) = $[\text{M}+\text{H}]^+$ 365 (58), 265 (100).

^1H -NMR (300MHz, CDCl_3-d), $\delta = 7.77$ (bs, 1H), 7.28-7.15 (m, 5H), 6.83-6.81 (d, $J = 7.5$ Hz, 1H), 5.15-5.12 (d, $J = 8.4$ Hz, 1H), 4.87-4.81 (m, 1H), 4.15 (m, 1H) 3.24-3.16 (m, 1H), 3.06-2.99 (m, 1H), 1.75-1.66 (m, 1H), 1.46 (s, 9H), 1.38-1.24 (m, 2H), 0.92-0.87 (t, $J = 7.2$ Hz, 3H)

^{13}C -NMR (75MHz, CDCl_3-d), $\delta = 174.39$ (COO^-), 172.66 (CON), 156.26 (CO), 136.29 (Ph C), 129.83 (Ph C), 128.88 (Ph C), 127.45 (Ph C), 77.61 ($>\text{C}<$), 54.68 (CH), 53.63 (CH), 37.85 (CH_2), 34.82 (CH_2), 28.68 (3CH_3), 19.15 (CH_2), 14.06 (CH_3).

(S)-2-((S)-2-aminopentanamido)-3-phenylpropanoic acid (Nva-Phe)

Nva-Phe was prepared from Boc-Nva-Phe (0.12 g, 0.33 mmol) according to the procedure reported for **1**. Yield = 0.06 g, (70%).

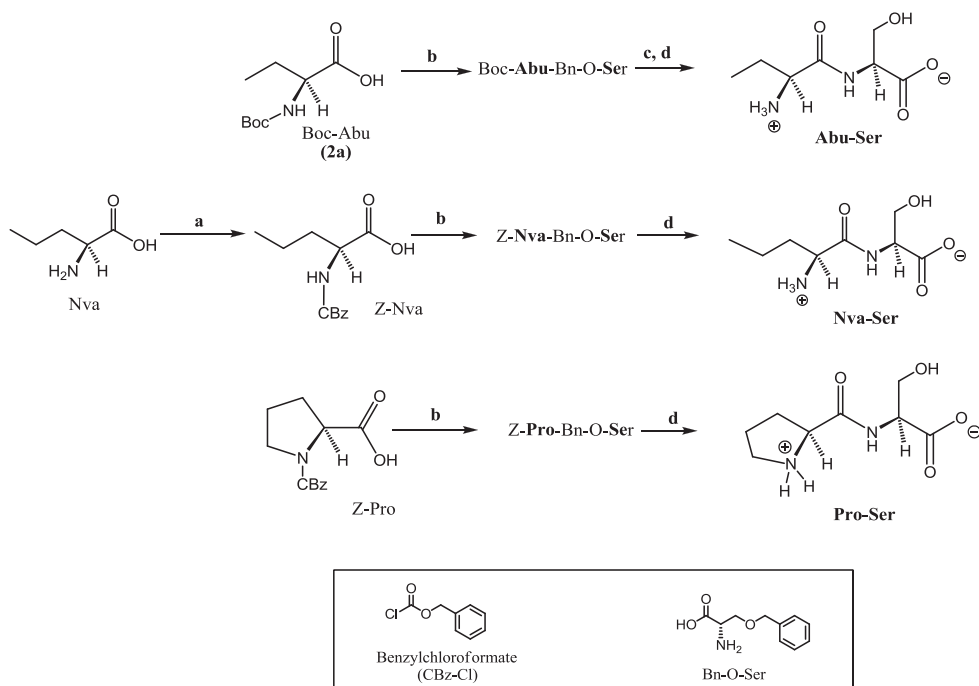
MS $\text{ES}^+ m/z = [\text{M}+\text{H}]^+ 265$.

^1H NMR (300 MHz, $\text{DMSO-}d_6$) $\delta = 8.84$ -8.81 (d, $J = 9$ Hz, 1H), 8.25-8.10 (m, 3H), 7.34-7.14 (m, 5H), 4.48-4.40 (m, 1H), 3.76-3.70 (m, 1H), 3.09-3.03 (m, 1H), 2.96-2.88 (m, 1H), 1.69-1.65 (m, 2H), 1.33-1.27 (m, 2H), 0.86-0.81 (t, $J = 7.3$ Hz, 3H).

^{13}C -NMR (75MHz, $\text{DMSO-}d_6$) $\delta = 173.24$ (COO^-), 169.67 (CON), 137.97 (Ph C), 129.88 (Ph C), 129.19 (Ph C), 127.49 (Ph C), 54.78 (CH), 52.80 (CH), 37.16 (CH_2), 34.04 (CH_2), 18.07 (CH_2), 14.45. (CH_3).

Crystallization: Slow evaporation of water from a neutralized solution at RT over three weeks afforded the colourless crystals of Nva-Phe, but Abu-Phe remained as powder.

b) Dipeptides from Abu, Nva, Pro and Ser amino acids



Scheme-2: Synthetic schemes for the Abu-Ser, Nva-Ser and Pro-Ser molecules.

Reagents and conditions: **a)** Benzylchloroformate, 2N NaOH, H₂O; **b)** i) DCC/ CH₂Cl₂/ *N*-hydroxysuccinimide ii) NaHCO₃, THF/acetone, H₂O, Bn-O-Ser; **c)** THF, HCl; **d)** i) Pd/C, H₂, HCl ii) NaOH.

Synthesis of (*S*)-3-(benzyloxy)-2-((*S*)-2-((*tert*-butoxycarbonyl) amino) butanamido) propanoic acid (Boc-Abu-Bn-O-Ser)

Boc-Abu-Bn-O-Ser was prepared from Boc-L-2-aminobutyric acid (**Boc-Abu**, 0.50 g, 2.46 mmol), NHS (0.31 g, 2.70 mmol) and DCC (0.56 g, 2.70 mmol) according to the procedure reported for **Boc-1**, however, THF was used instead of acetone (as in earlier preparations) for the coupling of *O*-benzyl-L-serine (Bn-O-Ser, 0.48 g, 2.46 mmol) at *C*-terminal. Yield = 0.90 g (95%).

¹H-NMR (200MHz, DMSO-*d*₆), δ = 13.77 (bs, 1 H), 8.02-7.99 (d, *J* = 9.0 Hz, 1H), 7.38-7.25 (m, 5H), 6.86-6.81 (d, *J* = 9.0 Hz, 1H), 4.42-4.54 (m, 3 H) 3.98-3.87 (m, 1 H) 3.79-3.71 (m, 1H), 3.64-3.68 (m, 1 H), 1.73-1.45 (m, 2 H), 1.36 (s, 9 H), 0.87-0.84 (t, *J* = 7.2 Hz, 3H).

¹³C-NMR (50MHz, DMSO-*d*₆), δ = 172.90 (COOH), 172.28 (CO), 156.17 (CO), 138.88 (Ph C), 129.04 (Ph C), 128.34 (Ph C), 128.30 (Ph C), 78.85 (>*C*<), 73.04 (O-CH₂), 70.41 (CH₂-O), 56.30 (CH), 53.14 (CH), 29.05 (3CH₃), 26.16 (-CH₂), 11.18 (-CH₃).

(S)-2-((S)-2-ammoniobutanamido)-3-hydroxypropanoate (Abu-Ser)

Abu-Ser was prepared by dissolving Boc-Abu-Bn-O-Ser (0.70 g) in dry THF (25 mL) and 2 mL of 4 M HCl in dioxane was added slowly at RT. The reaction mixture was stirred at RT for 48 h (TLC, eluent system MeOH:CH₂Cl₂ (1:9) and ninhydrin spray). Evaporation of the reaction solvents under reduced pressure afforded the product (0.50 g) as HCl salt, which was redissolved in fresh THF (30 mL) and 10% Pd/C (0.200 mg) was added. The RBF was then pressurized with a balloon containing hydrogen gas. Reaction mixture was continuously stirred for overnight and monitored by TLC (1:4, MeOH:CH₂Cl₂). Filtration of reaction slurry through the celite pad afforded sticky substance (Pd/C + product), which was rinsed with fresh THF (10 mL). The residual sticky Pd/C was slurried in water (10 mL) and filtered through celite. The filtrate was made basic with 2N NaOH (pH = 7-8).

Crystallization of Abu-Ser

The fine X-ray quality crystals of L-Abu-L-Ser were grown by adding few drops of DMF into the above aq. solution and slow evaporation of solvents at RT afforded substantial amount of crystalline **Abu-Ser** (0.28 g, 78%).

MS ES⁺ m/z = [M+H]⁺ 191.

NMR data ¹H-NMR (200MHz, DMSO-*d*₆ + D₂O), δ = 4.29-4.26 (m, 1 H), 3.82-3.63 (m, 3 H), 1.77 (s, 2 H), 0.94-0.87 (m, 3H);

¹³C-NMR (50MHz, DMSO-*d*₆ + D₂O), δ = 171.20 (COO⁻), 168.43(CO), 60.97(CH), 54.98 (CH₂), 52.88 (CH), 24.30 (CH₂), 8.76 (CH₃).

(S)-2-(benzyloxycarbonylamino)pentanoic acid (Z-Nva)

Z-Nva was prepared from L-norvaline (Nva, 0.70 g, 6.0 mmol) and benzylchloroformate (CBz- or Z-chloride, 1.02 g, 6.0 mmol) as per the literature procedure for Z-D-norvaline.¹ Yield = 1.40 g (79%).

¹H-NMR (300MHz, CDCl₃-*d*), δ = 10.90 (s, 1H), 7.41-7.30 (m, 5H), 5.37-5.34 (d, *J* = 9.2 Hz, 1H), 5.20-5.14 (m, 2H), 4.47-4.42 (m, 1H), 1.88-1.67 (m, 2H), 1.47-1.39 (m, 2H), 0.99-0.94 (t, *J* = 7.3 Hz, 3H).

¹³C-NMR (75MHz, CDCl₃-*d*), δ = 174.91(COOH), 157.03 (CO), 137.88 (Ph C), 129.18 (Ph C), 128.63 (Ph C), 128.54 (Ph C), 66.21 (CH₂), 54.40 (CH), 33.70 (CH₂), 19.61 (CH₂), 14.27 (CH₃).

(S)-3-(benzyloxy)-2-((S)-2-(benzyloxycarbonylamino)pentanamido)propanoic acid (Z-Nva-Bn-O-Ser)

Z-Nva-Bn-O-Ser was prepared from Z-Nva (0.50 g, 2.0 mmol) and (S)-2-amino-3-(benzyloxy)-propanoic acid (Bn-O-Ser, 0.39 g, 2.0 mmol), according to the procedure reported for **Boc-1**.

Yield = 0.75 g (88%).

¹H-NMR (300MHz, DMSO-*d*₆), δ = 12.75 (m, 1H), 8.15 (d, *J* = 7.9 Hz, 1H), 7.47-7.21 (m, 11H), 5.03 (s, 2H), 4.51-4.45 (m, 3H), 4.13-4.09 (m, 1H), 3.76-3.64 (m, 2H), 1.70-1.41 (m, 2H), 1.43-1.20 (m, 2H), 0.85 (t, *J* = 7.2 Hz, 3H).

¹³C-NMR (75MHz, DMSO-*d*₆), δ = 173.02 (COOH), 172.25 (CON), 156.76 (CO), 138.85 (Ph C), 137.91 (Ph C), 129.18 (Ph C), 129.05 (Ph C), 128.60 (Ph C), 128.48 (Ph C), 128.34 (Ph C), 128.32 (Ph C), 73.04 (CH₂), 70.28 (CH₂), 66.19 (CH₂), 55.08 (CH), 53.13 (CH), 35.00 (CH₂), 19.46 (CH₂), 14.49 (CH₃).

(S)-2-((S)-2-aminopentanamido)-3-hydroxypropanoic acid (Nva-Ser)

Z-Nva-Bn-O-Ser (0.46 g, 1.1 mmol) was dissolved in 20 mL of dry THF and cooled between 5 - 0 °C. 10% wet Pd/C (0.2 g) was added to the solution followed by the addition of 4M HCl in dioxane (0.5 mL). The reaction mixture was stirred for 48 h at RT (TLC, MeOH:CH₂Cl₂ (1:4) and ninhydrin staining). THF was decanted and the sticky substance (Pd/C + product) washed with fresh THF (5 mL). The residual Pd/C was then slurried in water (5 mL) to dissolve the water soluble product and filtered through a celite pad. Evaporation of water filtrate was carried out under reduced pressure to yield syrupy product (Nva-Ser.HCl) (0.20 g, 77%). The syrupy product was again dissolved in fresh water and then neutralized the solution with aq. 1N NaOH (pH = 7-8). The basified solution was further used for the crystallization experiments.

MS ES⁺ *m/z* = [M+H]⁺ 205.

¹H-NMR (200MHz, DMSO-*d*₆ + D₂O), δ = 4.33-4.32 (t, *J* = 4.7 Hz, 1H), 3.87-3.64 (m, 3H), 1.75-1.62 (m, 2H), 1.39-1.26 (m, 2H), 0.87-0.82 (t, *J* = 6.3 Hz, 3H)

¹³C-NMR (75MHz, DMSO-*d*₆), δ = 171.12 (COO⁻), 169.56 (CON), 61.88 (O-CH₂), 55.84 (CH), 52.72 (CH), 34.07 (CH₂), 18.21 (CH₂), 14.56 (CH₃).

(S)-1-(benzyloxycarbonyl)pyrrolidine-2-carboxylic acid (Z-Pro-Bn-O-Ser)

Z-Pro-Bn-O-Ser was prepared from Z-proline (0.70 g, 4.00 mmol) and Bn-O-Ser, NHS (0.47 g, 4.00 mmol) according to the procedure reported for **Boc-1** (scheme 1). Yield = 0.74 g (61%).

MS ES⁻ *m/z* (relative intensity in %) = [(2xM)-H]⁻ 851 (52), [M-H]⁻ 425 (100), 317 (60), 165 (8).

¹H-NMR (300MHz, DMSO-*d*₆), δ = 12.78 (bs, 1H), 8.30-8.22 (m, 1H), 7.38-7.26 (m, 10H), 5.12-4.94 (m, 2H), 4.52-4.35 (m, 4H), 3.81-3.32 (m, 4H), 2.22-2.02 (m, 1H), 1.90-1.76 (m, 3H).

¹³C-NMR (75MHz, DMSO-*d*₆), δ = 172.33 (2 CO), 154.64 (Z-CO), 138.82 (Ph C), 137.83 (Ph C), 129.04 (Ph C), 128.31 (Ph C), 127.79 (Ph C), 73.06 (CH₂), 70.37 (CH₂), 66.67 (CH₂), 66.55 (CH), 59.80 (CH), 47.99 (CH₂), 31.95 (CH₂), 23.81 (CH₂).

(S)-3-hydroxy-2-((S)-pyrrolidine-2-carboxamido)propanoic acid (Pro-Ser)

Pro-Ser was prepared from Z-Pro-Bn-O-Ser (0.40 g) by the procedure reported for Nva-Ser.

Yield = 0.17 g (88%).

MS ES⁺ *m/z* (relative intensity in %) = [M+H]⁺ 203 (100), 70 (7).

¹H-NMR (200MHz, DMSO-*d*₆) = δ 10.33-10.15 (m, 1H), 8.99-8.95 (d, 1H, *J* = 8 Hz), 8.49-8.47 (s, 1H), 4.32-4.20 (m, 2H), 3.77-3.62 (m, 2H), 3.22-3.14 (m, 2H), 2.41-2.20 (m, 1H), 1.93-1.72 (m, 3H);

¹³C-NMR (50MHz, DMSO-*d*₆), δ = 171.21 (CO), 168.33 (COO⁻), 60.94 (CH), 58.53 (CH₂), 55.17 (CH), 45.64 (CH₂), 29.72 (CH₂), 23.53 (CH₂).

Crystallization of Pro-Ser

X-ray quality crystals of Pro-Ser were obtained by using the similar gel crystallization method used for the dipeptides **1-10** (see page 38).

Reference:

- (1) Oka, T.; Yasusa, T.; Ando, T.; Watanabe, M.; Yoneda, F.; Ishida, T.; Knoll, J. *Bioorg. Med. Chem.* **2001**, *9*, 1213-1219.

Chapter 2

Guanidinium-carboxylate Based Supramolecular Complexes

2.1 The Guanidine Subgroup and Molecular Recognition

The guanidine moiety found in biomolecules such as arginine, an amino acid residue of various enzymes or proteins, and guanine, a nucleic acid base of RNA and DNA. It has a natural tendency to form charge-assisted hydrogen bonds with the carboxylate or phosphate anions of protein or nucleic acids, respectively.¹⁻⁴ In living beings, such definitive intermolecular interactions are very crucial for the bio-chemical modifications, for example; the nitric oxide synthesis by enzymatic catalysis, electron transfer reactions in respiration process and RNA-protein recognition etc.⁴⁻⁸

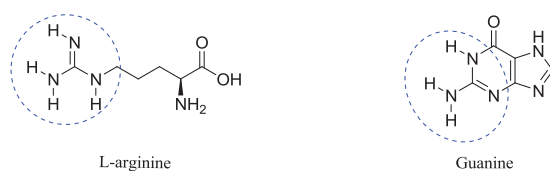


Figure 2.1.1: L-arginine and guanine, guanidine precursors in the biological macromolecules.¹⁻⁴

Inspired by naturally occurring specific interactions there have been numerous research efforts in the field of bio-organic chemistry to prepare abiotic receptors based on the guanidinium structures⁹ (Fig. 2.1.2). As receptors, the cationic guanidiniums were designed to interact non-covalently and selectively with various oxoanions, such as carboxylate, carbonate, sulfonate, phosphonate and nitro groups, to generate host-guest complexes in the solids and solutions.^{4,10-16,17-29}

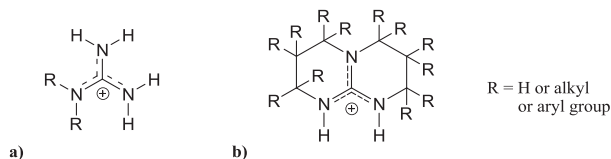
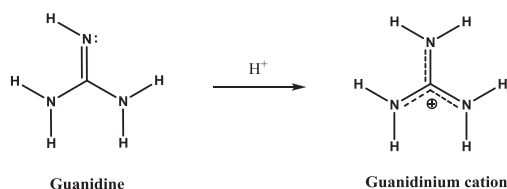


Figure 2.1.2: The abiotic receptor²¹ structures a) a substituted guanidine derivative and b) rigid-bicyclic guanidine system for high selectivity.

2.2 Guanidine in Crystal Engineering



Unsubstituted guanidine is a versatile molecule in organic chemistry, acting as base ($pK_a = 13.6$) or reagent.³⁰ It readily forms a guanidinium cation with various organic protic acids (carboxylic, sulfonic, phosphonic acids and boric acid³¹) or a co-crystal with the nitro compounds.³² The formed cation is stable and remains protonated over a wide pH range. Specially, structural features such as low molecular weight, planar conformation and three-fold molecular symmetry (triangular geometry)³³ have made the guanidinium group a favorite for use in research. It has been extensively used building block in crystal engineering to prepare novel supramolecular assemblies.³⁴⁻⁵¹

The sulfonate (sulfate) and phosphonate (phosphate) groups are the anions with tetrahedral geometry and they form guanidinium complexes with different structural properties.⁵² By contrast, the carboxylate group has planar geometry; moreover, carboxylic acid is a very common functional group in biomolecules and also a well exploited group in the supramolecular chemistry.⁵³ Therefore, the study of carboxylate and guanidinium related complexes was the main topic of this investigation.

Usually, when guanidine abstracts a proton from a carboxylic group, the guanidinium ion interacts in the *syn* direction of the O atoms of the carboxylate group and forms two parallel charge-assisted ($^+N-H \cdots O^-$) hydrogen bonds to generate a stable cyclic association i. e. $R_2^2(8)$ motif [Fig. 2.1.1].

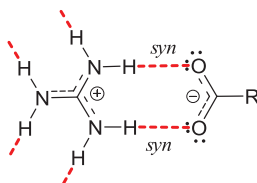


Figure 2.2.1: A guanidinium-carboxylate complex showing formation of the hydrogen-bonded $R_2^2(8)$ motif in *syn* direction of carboxylate O atoms.

2.3 Guanidinium-Carboxylate Complexes: CSD Survey and Specific Interactions

Though the guanidinium-carboxylate interactions are directive, it is very rare to see three carboxyl/carboxylate (or carbonate) counterions interacting evenly with a single guanidinium cation by accepting its six hydrogen atoms to generate three independent $R_2^2(8)$ motifs [as shown in Fig. 2.3.1 (a, b)]. In 66 structures with guanidinium ions and carboxylates retrieved from the CSD, this pattern was found only nine times (Table 2.3). The fraction forming just one $R_2^2(8)$ motif (34 structures) is comparatively higher than two- and three- $R_2^2(8)$ motif formers (Table 2.3). Surprisingly, in spite the high propensity for $R_2^2(8)$ motif formation between guanidinium and carboxylate, several structures do not show these H-bonded interactions at all (Table 2.3) (Fig. 2.3.2 a, b).

Table 2.3: A CSD survey of **guanidinium-carboxylate** organic complexes forming $R_2^2(8)$ motifs (Total 65 hits)

Three, $R_2^2(8)$ motifs (9 structures)	Two, $R_2^2(8)$ motifs (12 structures)	One, $R_2^2(8)$ motif (34 structures)	Zero, $R_2^2(8)$ motif (10 structures)
QIFFIU, QIFFOA ³⁶ , DOTWOZ ⁴⁰ , TAVFOM, TAVFUS ⁵¹ , JOGTIJ, JOGTEF ³¹ , GUANCB ⁵⁴ and TUTGEV ⁵⁵ .	SEWXAU, SEWXEY, YEYWUV ⁴³ , JOGTUV, JOGVAD ³¹ , TUTFIY, TUTFOE, TUTGIZ, TUTGAR ⁵⁵ , KABKEF ⁵⁶ , XUVRUB ⁵⁷ and DUJKID ⁵⁸ .	ANEPAL, ANENOX, ANENUD01, ANENUD, BOJFEM, DADBEQ, DEHSAK10, DUFDIR10, DUFDOX10, DUMPUW, ETIZOW, EVASAW, FEDPUA, FOVMEL, GIHCIK, GUHOXM, GUPICL10, GUPMYL, HUTNEP, IWIRAH, IZIJEH, JEWCUK, NOTPAN, PACDEC, PAGDIG, PUJDAA, TEVRAO, UJOCOM, VOTMEN, XAGFAM, YEYMEN, YEJZAO, YUVRUB and ZAYHEO.	BUDXII, GADOXP10, GUPMEL, KIKKEV, LABZAP, MENPOL, SIDMOH, UFALET ⁵⁹ UJOCUS and UJODAZ.

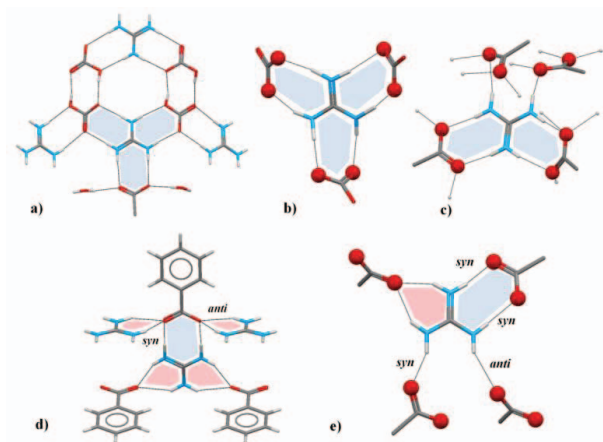


Figure 2.3.1: The representative guanidinium-carboxylate structures showing formation of different interactions; a) and b) show the structures with three $R_2^2(8)$ motifs in a guanidinium-carboxylate recognition complex in *syn* direction of O atoms of the carboxylate groups (Refcodes QIFFIU and GUANCB).^{36,54} c) A guanidinium cation forming two $R_2^2(8)$ motifs while the remaining two H atoms of the guanidinium donor interact with carboxylates in different directions (KABKEF),⁵⁶ d) and e) A guanidinium cation forming a single $R_2^2(8)$ motif, the rest of the guanidinium hydrogen atoms being engaged in bifurcated $R_1^2(6)$ motifs (GIHCIK and DUFDX10).^{60,61} The light blue color represents the $R_2^2(8)$ motifs while light red color represents $R_1^2(6)$ motifs. The hydrogen bonds are shown by dotted lines.

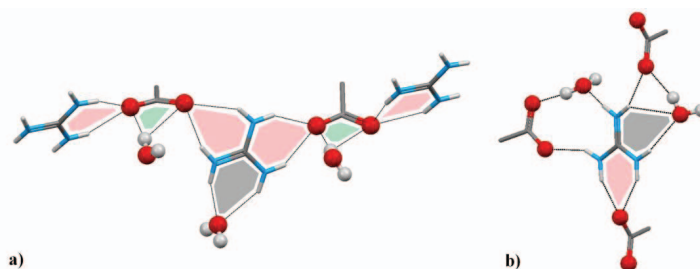


Figure 2.3.2: a) A guanidinium-carboxylate complex without any $R_2^2(8)$ H-bonding motifs (KIKKEV⁶²) instead forms bifurcated $R_1^2(6)$ hydrogen bonds either with the O atom of carboxylate group (light-red color) or with the co-crystallized water molecule (light grey color). b) Random H bonding networks formed by three carboxylates and a guanidinium cation along with co-crystallized water molecules (Refcode; UJODAZ).⁶³

In general, the carboxylate groups in zwitterionic structures easily form more than two hydrogen bonds, such as in small peptides.^{64,65} Similarly, when a $R_2^2(8)$ motif is formed by

one carboxylate group with a guanidinium group, the other carboxylate groups remain in the proximity by accepting a hydrogen atom of guanidinium with a normal or bifurcated manner (Fig. 2.3.1, d). Also, if the mono-, di- or tricarboxylates interact with the guanidinium group through *syn* lone pairs of carboxylate O atoms to form $R_2^2(8)$, concurrently the *anti* lone pairs of carboxylate O atoms become accessible to the available donors (water or guanidinium cation etc) (Fig. 2.3.2). By these multiple hydrogen bonds formation, the guanidinium and carboxylate groups mutually satisfy their capacity of non-covalent interactions in their complexes.

Therefore, in the cases of guanidinium-carboxylates forming two, one and zero $R_2^2(8)$ motifs, we assume that in addition to the unpredictable steric and structural properties of the molecules with carboxylic acid functionality, there could be two obvious reasons for such irregular H-bonding patterns: i) guanidinium ion as a multiple H atom (six H atoms) donor and ii) the additional interactions (more than two) of carboxylate O atoms through the lone pairs of electrons upon charge-assisted complex formation. These observations show that when a guanidinium ion and carboxylate group forms single $R_2^2(8)$ H-bonded motif, their zwitterionic adduct simultaneously interacts with other donors or acceptors and ultimately forms two- or three-dimensional closely packed networks.

2.4 Effective Strategy to Utilize the Guanidinium Subgroup in Crystalline Materials

2.4.1 Selection of Guanidine Derivative

To exploit a guanidinium moiety as a building block in crystal engineering in the present work we directed attention to manipulating its H bond forming capacity. Hence we decided to introduce a guanidine derivative that could interact with just a single carboxylate group to form one $R_2^2(8)$ motif. We considered that these efforts would minimize the structural complications we have been noticed in earlier structures with native guanidine. Furthermore, through this restrictive approach we foresaw the diversion of carboxylate O atom *anti* lone pairs to interact with a single donor at a time. In the absence of any other donors (Fig. 2.4.1.1) co-crystallized water molecules could fill this hole. Thus, the resulting complex would include two water molecules per guanidinium-carboxylate in a crystal. Since water is dynamic and it has the ability to form different clusters, chains and tapes in crystals through co-operative hydrogen bonding, here we expected that the water molecules interacting directly with the guanidinium-carboxylate complex may also interact with additional molecules in the process of network building in crystals. The idea to facilitate water incorporation in crystals of

such complexes was a deliberate move, because the various water networks in different molecular crystals lattices are of the utmost interest to the theoretical, biological and nanomaterial sciences.⁶⁶⁻⁸⁴

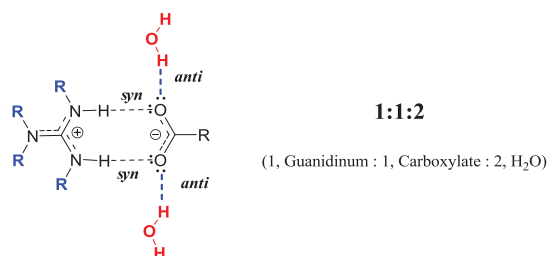


Figure 2.4.1.1: A modified and representative design of the guanidinium-carboxylate complex to be used in this work, where only a pair of H atoms of the guanidinium ion is interacting with the two O atoms of the carboxylate group in *syn* direction and directs the *anti*-lone pairs of carboxylate O atoms to interact freely with the water molecules. Other two pairs of H atoms are blocked by (R) substitutions.

To begin with, the selection of an appropriate guanidine derivative was a major concern before actual demonstrations. Clearly, in the required molecule four out of six guanidinium H atoms should be replaced by alkyl (R) group (R = methyl or ethyl etc) (Fig. 2.4.1.1). Guanidinium derivatives were however not available commercially and also quite challenging to prepare in the laboratory.

A few of di-*N,N'* and tri-*N,N',N''* substituted guanidine derivatives (Fig. 2.4.1.2. b, c and d) were cocrystallized with the carboxylic acids, but the non-planar geometry and accessible H-atoms other than involved in R₂²(8) of cations generate irregular H-bonded networks (HOFDAH, HOFDEL, POVDUZ, SANQUU,⁸⁵⁻⁸⁸ EWODOK, EWODUQ, EWOFAY,⁸⁹ PAXJIJ, YEPBAX, YEPBEB, YEPBIF, YEPBOL^{90,91}). Thus these derivatives were also not the suitable options to incorporate in this design as they did not show the formation of R₂²(8) motifs.

Finally we found that a rigid guanidinium derivative 1, 5, 7-Triazabicyclo[4.4.0]dec-5-ene (**TBD**, pK_a = 13.6) (Fig. 2.4.1.2.e) was an appropriate molecule. It has a guanidine group locked in rigid cyclic conformation, which selectively can donate only two H-atoms in its cationic state. The TBD molecule is often used as a building block in co-ordination chemistry,^{92,93} as a base or catalyst in synthetic organic chemistry,⁹⁴⁻⁹⁷ but it has very rarely

been exploited as a building block in crystal engineering.⁹⁸⁻¹⁰⁵ Hence we decided to explore the versatility of TBD in our supramolecular synthetic strategy.

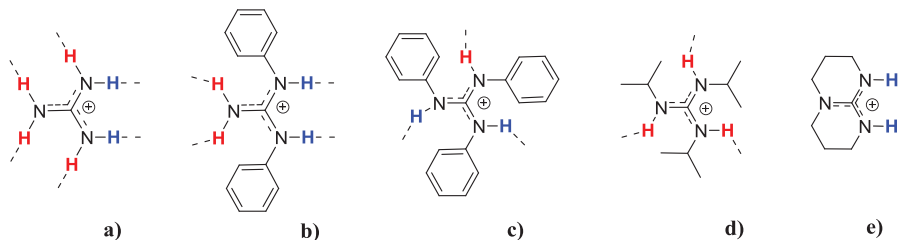


Figure 2.4.1.2: a) A guanidinium cation with six H atoms. b) di-phenyl substituted guanidinium. c) tri-phenyl substituted guanidinium. d) tri-isopropyl guanidinium. e) TBD a rigid guanidinium derivative. In b), c) and d) H atoms not involved in $R_2^2(8)$ motif are shown in red.

2.4.2 Di-guanidinium Ligand: Two Better than One

A compound with more than one guanidinium group in a single molecule such as a hypothetical ditopic-guanidinium system (**I**) could be a linker between two carboxylates with two $R_2^2(8)$ motifs in a guanidinium-carboxylate related complex, thus ultimately could interact with four water molecules (**1:2:4** complex) of hydration in a complex as shown in strategy (A) (Fig. 2.4.2).

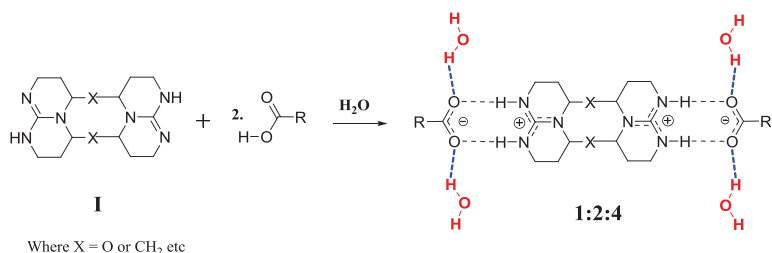


Figure 2.4.2: A strategy (A) showing ditopic guanidinium-carboxylate complex for enhanced hydration in crystals.

2.4.3 Organic Frameworks from I

Presently, the preparations of metal-organic frameworks (MOFs) is a hot topic in chemistry due to their strong, attractive and very useful networks.¹⁰⁶ In MOFs metal ions or metal clusters function as a node and the organic ligands as node connections. By the same principle, di-topic guanidinium **I** could be used effectively as an organic linker with different carboxylic acids (nodes) to obtain linear, 2D and 3D networks as shown in Fig. 2.4.3.

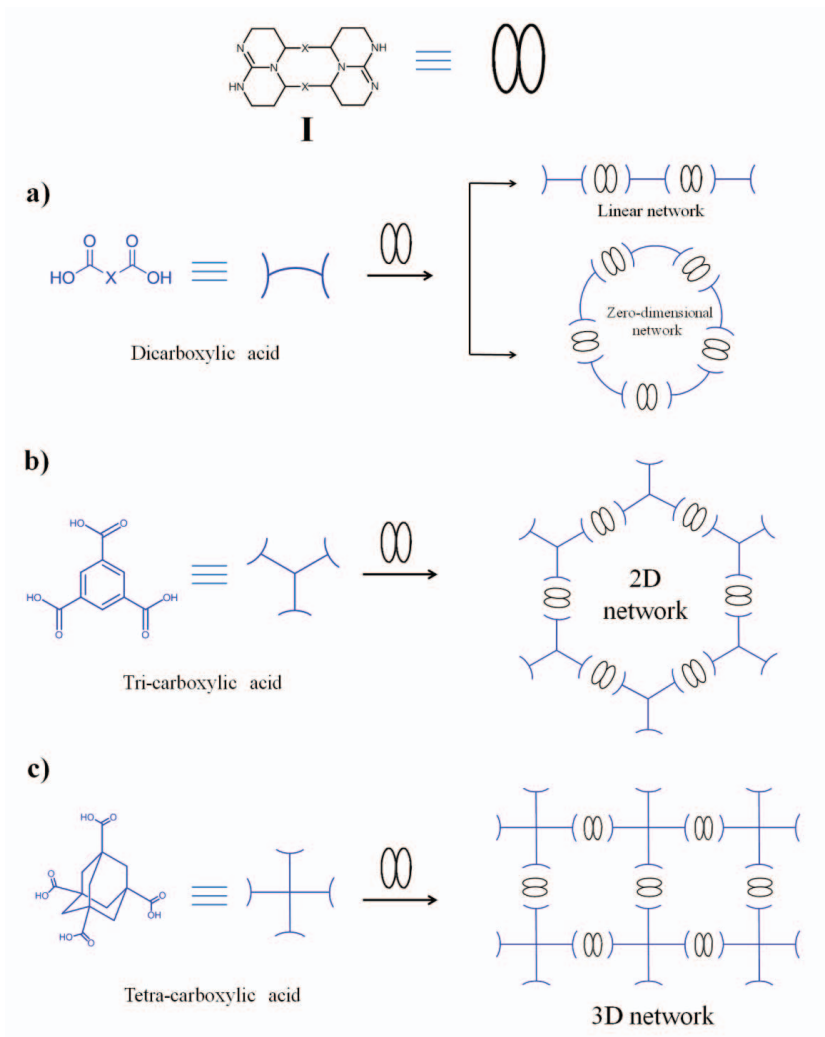
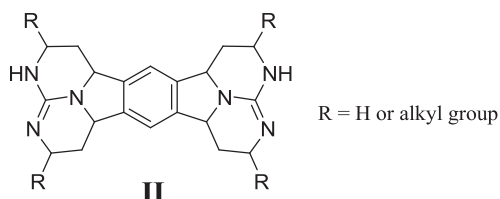


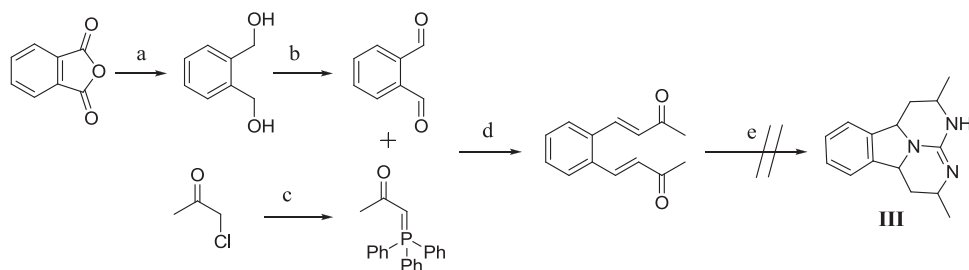
Figure 2.4.3: Designs showing preparation of purely organic a) 1D or 0D, b) 2D and c) 3D frameworks from hypothetical versatile di-topic guanidinium (**I**) and various dicarboxylic acid molecules.

Unfortunately, molecule **I** is neither available commercially nor synthetically easy to prepare, but in the future efforts could be directed towards the synthesis of **I**. Hence, the possibility of employing compound **I** in the present study was ruled out.

Alternatively, we planned to synthesize another molecule **II**, in which the ring linking ($X = -O-$ or $-CH_2-$) between the two rings of **I** is replaced by an aromatic group.



In order to check the feasibility of desired product formation, the following reactions (Fig. 2.4.5) were attempted, unfortunately the crucial cyclization step¹⁰⁷ failed to produce the mono guanidinium derivative **III**, instead yielding the unreacted reagents.



Reagents and conditions: a) LiAlH_4 , THF; b) i) HNO_3 (68%), 80 °C, ii) AcOH, H_2O ; c) PPh_3 , ACN; d) reflux, 20 h; e) i) Guanidine, DMF, 3h, ii) NaBH_4 .

Figure 2.4.5: Reaction scheme for attempted synthesis of molecule **III**.

Therefore, without wasting time on molecule **II**, we abandoned further investigations in this direction, but such types of molecules would have great impact as an organic ligand in the preparation of entirely organic supramolecular materials, hence in the future further synthetic efforts would be worthwhile.

2.5 Guanidine Derivative (TBD) and Dicarboxylic acids in Crystal Engineering

As the bi-guanidine derivatives **I** and **II** were unavailable, we decided to use TBD, with dicarboxylic acids (as linkers between the TBD molecules) in an initial screening to generate two $R_2^2(8)$ in a complex. In fact this strategy (B, Fig. 2.5.1) also provides the same sites of interactions for the equivalent number of water molecules as earlier (A, Fig. 2.4.3), however, here the binding order, orientation and the molecular ratio (**2:1:4**) of the components is different (Fig. 2.5.1). The advantage of strategy B is that availability of a wide variety of dicarboxylic acids as linkers that can produce distinct networks.

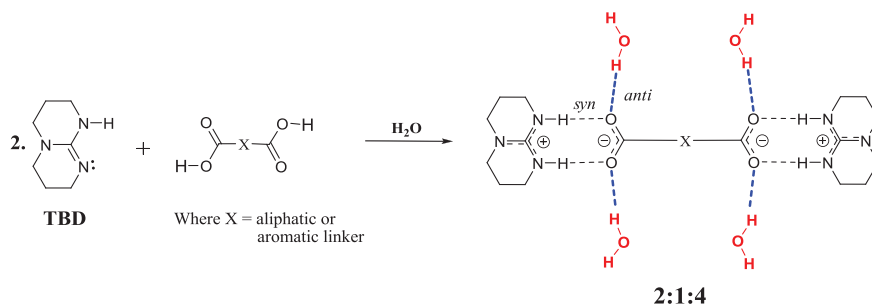


Figure 2.5.1: The alternative strategy B showing interactions of four water molecules with guanidinium-carboxylate and formation of the **2:1:4** complexes.

2.6 Experimental Section

2.6.1 Materials and Crystals Synthesis

In order to validate our hypothesis of the limited H-bonding ability of guanidinium-carboxylate complex and its H bonds formation with the water molecules through *anti* lone pairs of carboxylate O atoms, a series of dicarboxylic acids and a monocarboxylic acid with a special structural feature (such as glycolic acid) have been screened in combination with TBD.

A total of 15 carboxylic acids were purchased from Sigma-Aldrich together with TBD and used without further purification. Structural property-wise these carboxylic acids are different from one another; hence we categorized them into two groups di- and monocarboxylic acids.

2.6.2. Dicarboxylic Acids and TBD Complexes

In this group 13 dicarboxylic acids of different linkers were studied with TBD. On the basis of linker nature, this group is again divided into three subgroups.

2.6.2.1 Bi-aromatic dicarboxylic acids and TBD

2.6.2.2 Aliphatic saturated/unsaturated dicarboxylic acids and TBD

2.6.2.3 Aromatic dicarboxylic acids and TBD

2.6.2.1 Bi-aromatic Dicarboxylic Acids and TBD

Two dicarboxylic acids with extended aromaticity were used: biphenyl-4,4'-dicarboxylic acid (BPHA) and 2,2'-bipyridine-5,5'-dicarboxylic acid (BPDA) (Fig. 2.6.2.1a).

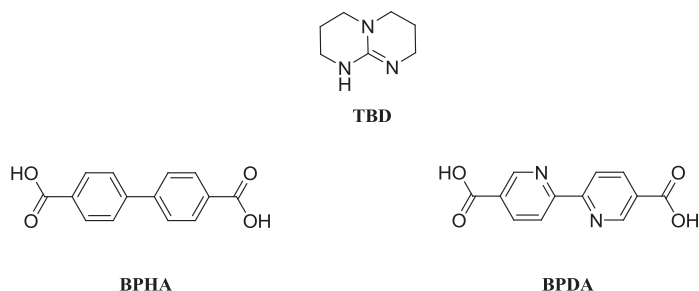


Figure 2.6.2.1a: TBD and di-aromatic carboxylic acids studied here for the 2:1:4 crystalline complex formations.

Synthesis of complex BPHA : 2TBD (BHT): The BPHA (24 mg, 0.1 mmol) was mixed with TBD (28 mg, 0.2 mmol) for proton exchange in water (3 mL) to generate the **2:1:4** solid-state complex (BHT). A few drops of *N,N*-dimethylformamide (DMF) were also added to the mixtures to assure a homogeneous and clear reaction solution. The purpose of DMF addition was also to prevent the complete dehydration during the course of slow solvent evaporation process while obtaining the X-ray diffraction quality single crystals. However, after four weeks, this mixture was afforded an amorphous powder.

Synthesis of complex BPDA : 2TBD (Type I): A 2:1 complex between BPDA and TBD here referred as **Type I** and was prepared using BPDA (24 mg, 0.1 mmol) and TBD (28 mg, 0.2 mmol) by the same method as described above for BPHA and TBD. This complex afforded colorless, fragile and air-sensitive crystals.

Synthesis of complex BPDA:1TBD (This complex also referred as **Type II**): After obtaining **Type I** crystals, another experiment was set to obtain a **2:2:4** complex, an extended complex of TBD and BPDA (Fig. 2.6.2.1b) For this, BPDA (24 mg, 0.1 mmol) and TBD (14 mg, 0.1 mmol) were mixed in water (3 mL) and few drops of DMF at room temperature. Slow evaporation of the solvents afforded colorless crystals of **Type II**. Unlike **Type I**, the **Type II** crystals were robust and stable.

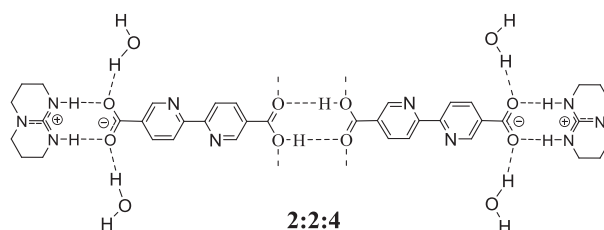


Figure 2.6.2.1b: A schematic representation showing an expected 2:2:4 complex from TBD and BPDA.

2.6.2.2 Aliphatic Saturated/Unsaturated Dicarboxylic Acids and TBD

This subgroup constitutes seven carboxylic acids with different length aliphatic linkers bridging two carboxylic groups. These were either saturated (oxalic, malonic, succinic, glutaric and adipic acid) or unsaturated (maleic and fumaric acids) in nature (Fig. 2.6.2.2).

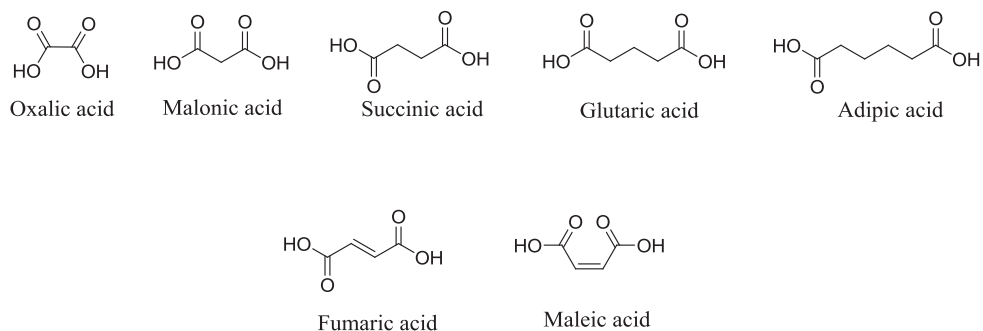


Figure 2.6.2.2: Aliphatic saturated and unsaturated dicarboxylic acids used for 2:1:4 complex formation with TBD.

Preparation of complexes: Each dicarboxylic acid (0.1 mmol) was mixed with TBD (0.2 mmol) and dissolved in 2 mL of water at RT before stirring for 10 min. They were left standing for 3-4 weeks whereupon slow evaporation of solvent afforded the complexes:

Only oxalic acid (**OxT**), glutaric acid (**GuT**), adipic acid (**AdT**) and fumaric acid (**FuT**) resulted into crystals while rest of the acids did not form crystals with TBD.

2.6.2.3 Aromatic Dicarboxylic Acids and TBD

The linker of the dicarboxylic acids in this subgroup is a six-membered aromatic group such as phenyl or pyridine and two carboxyl groups being attached at different positions on the rings with respect to each other. These acids are terephthalic, isophthalic, phthalic and pyridine-3,5-dicarboxylic acid (Fig. 2.6.2.3).

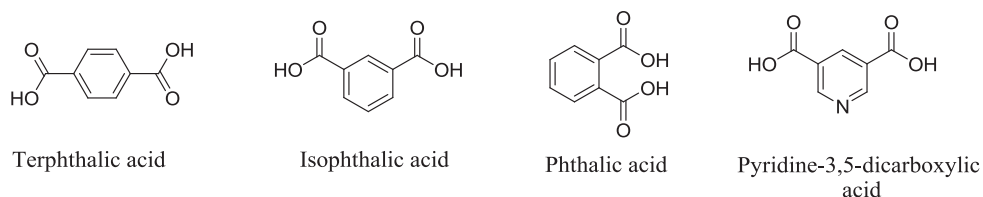


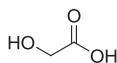
Figure 2.6.2.3: Aromatic dicarboxylic acids tested with TBD for 2:1:4 complex formation.

Preparation of complexes: Each carboxylic acid and TBD in 1:2 stoichiometries were combined in water (2 mL). These mixtures were stirred for 10 min and left standing for 3 - 4 weeks for slow solvent evaporation to allow the crystalline complexes form.

Out of these four combinations, only phthalic acid (**PhT**) and pyridine-3,5-dicarboxylic acid (**PyT**) yielded X-ray diffraction quality crystals.

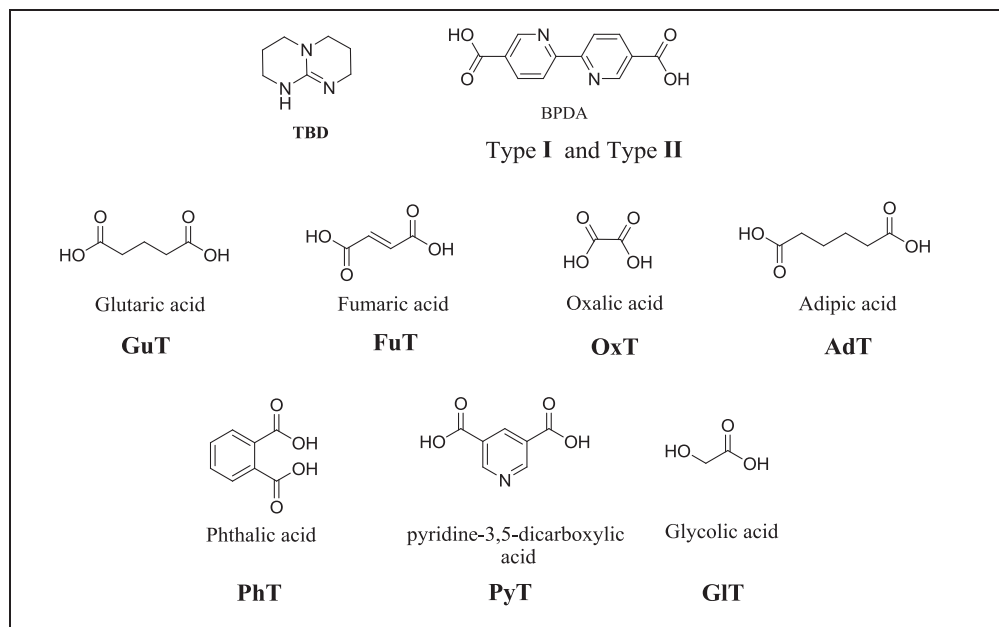
2.6.3. Mono-carboxylic Acid and TBD

Along with the dicarboxylic acids, the complex formation diversity of TBD was also examined in combination with a monocarboxylic acid, glycolic acid. Glycolic acid contains only a single carboxylic acid group, but in the crystallization process two equivalents of TBD were nevertheless added. The purpose of addition of two equivalents of TBD was to investigate the H-bonding property of an alcoholic –OH group in the presence of surplus TBD in the complex.



Glycolic acid

Preparation of Glycolic acid: 2TBD complex (GIT): Glycolic acid (38 mg, 0.5 mmol) and TBD (140 mg, 1 mmol) were mixed in water (2 mL) at room temperature. The crystals of **GIT** were obtained after 3-4 weeks.



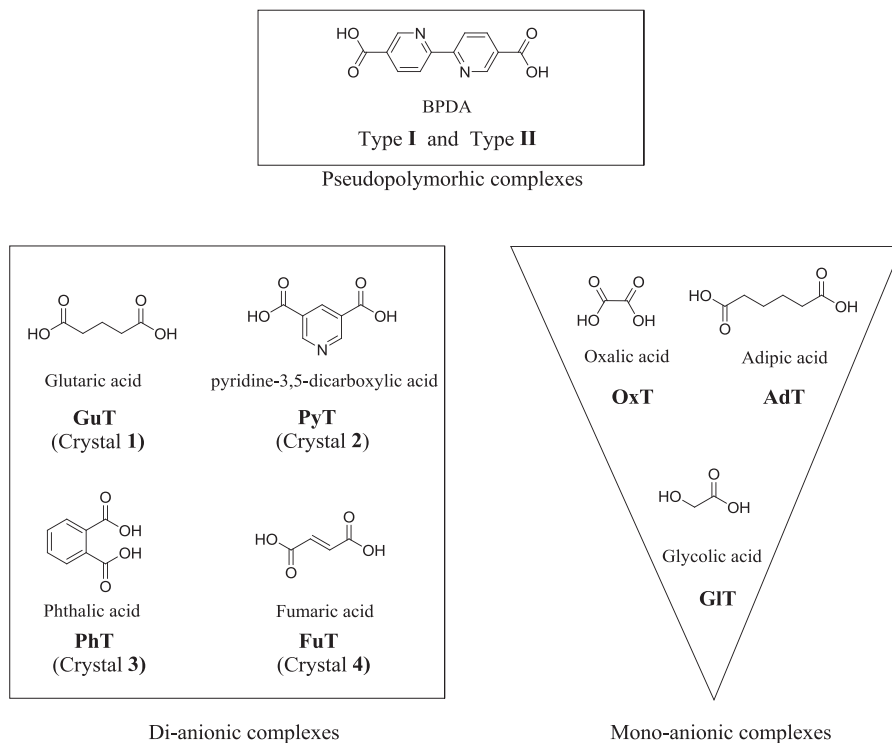
Scheme 2.6.2: Summary of obtained crystalline complexes of TBD with mono- or dicarboxylic acids in this study.

2.7 Single Crystal X-ray Crystallography

A total of nine crystalline complexes were obtained, then examined under microscope and suitable specimens selected for X-ray crystallography. Single crystal X-ray diffraction datasets were collected at 105 K on Bruker-ApexII CCD diffractometer (monochromatized MoK α radiation, $\lambda = 0.71073\text{\AA}$), which is coupled to an Oxford Cryosystem low-temperature device. Data integration reduction and absorption was carried out by SAINT and SADABS, respectively.¹⁰⁸ All structures were solved by direct methods and refined by full matrix least squares based on F^2 using SHELXTL.¹⁰⁹ Molecular and packing diagrams were generated using the program Mercury.¹¹⁰ Detailed data collection, refinements and crystallographic information can be found in respective article.

2.8 Results and Discussion

The crystal structure analysis of the nine complexes showed that several of them share some common structural properties. This allowed us to classify these nine complexes into the three groups below for the convenient descriptions and comparisons.



Scheme 2.8: Classification of the obtained crystalline TBD-carboxylate complexes in the present study.

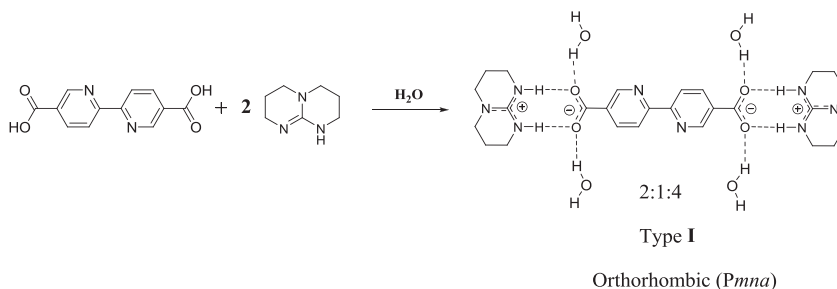
2.8.1 Pseudopolymorphic Crystalline Complexes of TBD and BPDA

The **Type I** and **Type II** crystals were obtained from 2:1 and 1:1 mixtures, respectively, from **TBD** and **BPDA**. Basically these complexes generated a common **2:1:4 (TBD:BPDA:water)** network building unit (Fig. 2.8.1.1), however their arrangement in the corresponding crystals is non-identical, wherein they exist in pseudopolymorphic forms.

Polymorphism: It is the existence of the one chemical substance in more than one crystalline form. These exhibit different physical and chemical properties,^{[111-113](#)} whilst **pseudopolymorphism** defines different crystalline forms of a same substance that differ in stoichiometry or type of the included solvents.^{[114-116](#)}

Detailed description of these complexes can be found in **Publication V**.

Type I:



Type II:

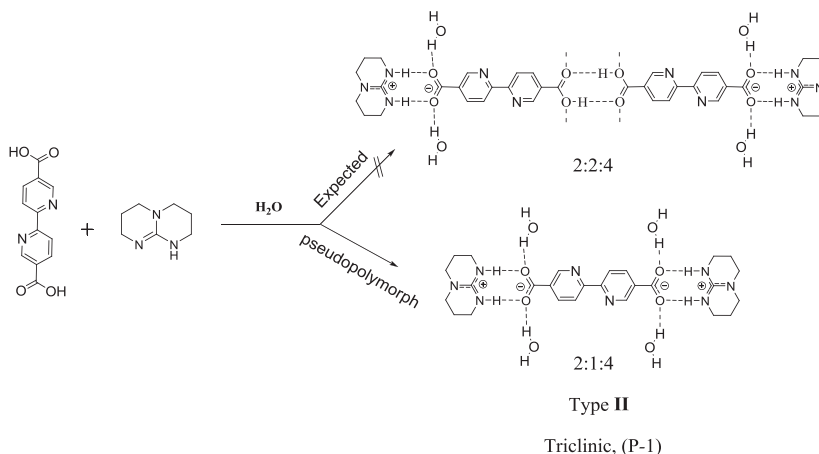


Figure 2.8.1.1: The pseudopolymorphic **Type I** and **Type II** crystals formation from TBD and BPDA.

2.8.2 Reproducibility of Type I Complex and Crystal Structure transformation

The **Type I** crystals were fragile and very sensitive outside of the water or DMF medium, becoming anhydrous within the 30 - 60 min after removed from the mother liquor. When this resulting non-crystalline dehydrated complex was redissolved in minimum amount of water (1 mL) and DMF (2-3 drops), the crystalline **Type I** form reproduced in 2 weeks. This procedure was repeated over two cycles, however, when the similar mixture was processed for the third attempt of determining reproducibility, the crystallization process took longer time (6 weeks) than usually (1-2 weeks) required for **Type I** crystals (Fig. 2.8.2). Surprisingly, the obtained crystals after third attempt were remarkably stable, even when devoid of any solvent medium. The crystal structure unveiled that this stable crystals are in fact **Type II** crystals, with a 1:1 (TBD:BPDA) composition. These observations show that the **Type I** structure has a property of structural transformation in solution.

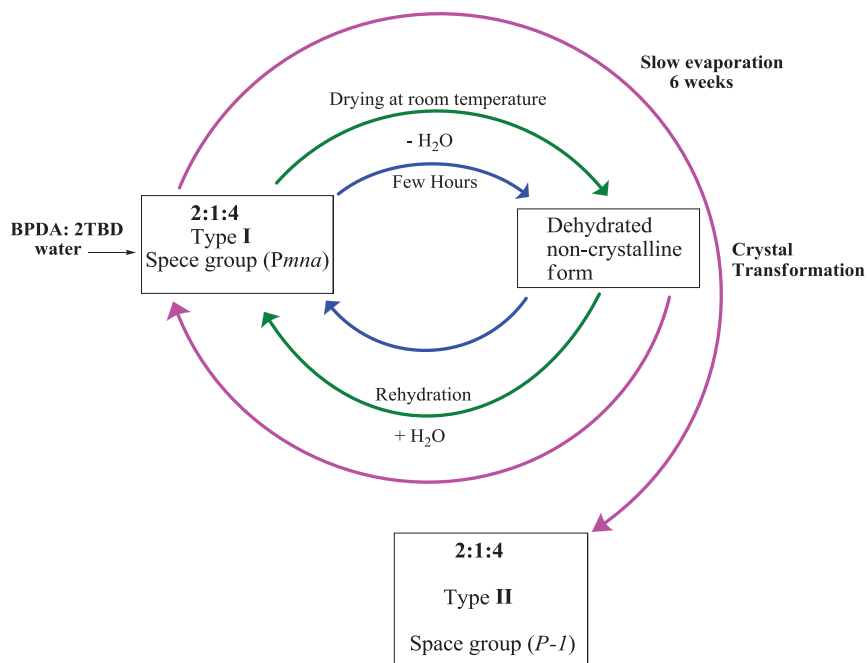


Figure 2.8.2: A schematic representation showing the cycles of the **Type I** crystals reproducibility and crystal transformation.

2.8.3 Di-anionic complexes from Dicarboxylic acids and TBD

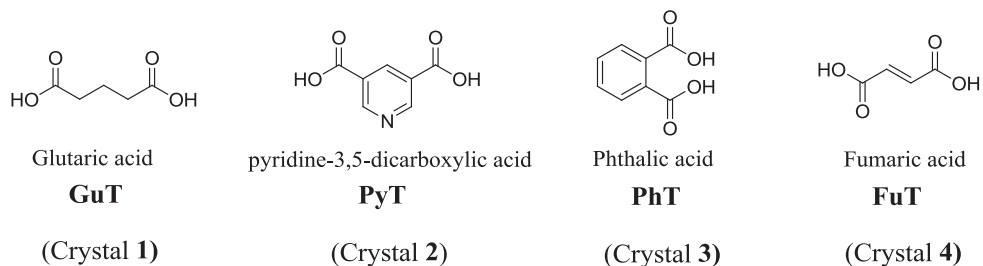


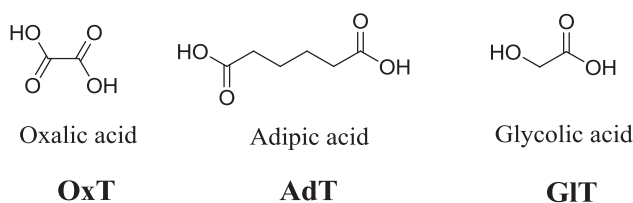
Figure 2.8.3: The dianionic complex forming complexes **GuT**, **PyT**, **PhT** and **FuT** are renamed and referred as the Crystal 1, 2, 3 and 4, respectively, in paper VI.

In the crystalline forms, the glutaric acid, pyridine-3,5-dicarboxylic acid, phthalic acid and fumaric acid with TBD were produced the H-bonded dianionic complexes (Fig. 2.8.3). These crystals show distinct structural arrangements and include a variety of water assemblies H-bonded to the *anti*-lone pairs of O atoms of the carboxylates.

Crystal 1 and 2 were found with the incorporated water tapes that are H-bonded to the guanidinium-carboxylate adduct. In Crystal 3 TBD did not form $R_2^2(8)$ motif with the carboxylates due to the coulombic repulsions of *ortho* dianions and steric crowding of TBD, nevertheless this complex showed a organized water cluster in the network. Uniquely, crystal 4 was anhydrous, as the *anti*-lone pairs of carboxylate O atoms were interacting with weak C-H donors of TBD molecules.

Details of these four structures can be found in **Publication VI**.

2.8.4 Mono-anionic Complexes from TBD and Mono- or Dicarboxylic acids



Oxalic acid (OxT) and adipic acid (AdT), in spite the two carboxylic acid groups, form monoanionic complexes with TBD, while glycolic acid (GIT) generates a monoanionic complex leaving the hydroxyl group to interact with the acceptor carboxylate O atom. The H-

bonded networks of OxT, AdT and GIT form infinite one-dimensional ribbons. Hence these three structures are grouped together and described thoroughly in **Publication VII**.

2.9 Conclusions

This study was a carefully designed and experimentally demonstrated strategy to illustrate the constructive use of TBD for the preparation of organic materials. A total of nine crystalline complexes were obtained and analyzed by X-ray diffraction methods. Out of these, the anticipated interactions of *anti*-lone pairs of carboxylate O atoms with water molecules were observed, and six complexes showed the inclusion of different water networks in their crystals. The **Type I** complex from BPDA and TBD is a very rare example of a purely organic framework with a high degree of hydration in the channels. The crystalline OxT, AdT and GIT complexes did not include a reasonable amount of water within the structures, but according to their 1D structural properties they can be considered as materials, as gelators etc.¹¹⁷ Moreover, as the guanidinium and carboxylate groups are biological functional groups, the complexes studied here may provide some insight into specific interactions of similar type of molecules. Collectively, the complexes from this study may find various future applications as proton conductors, as selective ion channels and as gelators etc.

Reference List

- (1) Frankel, A. D. *Curr. Opin. Struct. Biol.* **2000**, *10*, 332-340.
- (2) Hermann, T.; Patel, D. J. *Science* **2000**, *287*, 820-825.
- (3) Cheng, A. C.; Calabro, V.; Frankel, A. D. *Curr. Opin. Struct. Biol.* **2001**, *11*, 478-484.
- (4) Calnan, B.; Tidor, B.; Biancalana, S.; Hudson, D.; Frankel, A. *Science* **1991**, *252*, 1167-1171.
- (5) Crane, B. R.; Arvai, A. S.; Ghosh, D. K.; Wu, C.; Getzoff, E. D.; Stuehr, D. J.; Tainer, J. A. *Science* **1998**, *279*, 2121-2126.
- (6) Ramirez, B. E.; Malmström, B. G.; Winkler, J. R.; Gray, H. B. *Proc. Natl. Acad. Sci.* **1995**, *92*, 11949-11951.
- (7) Puglisi, J. D.; Chen, L.; Frankel, A. D.; Williamson, J. R. *Proc. Natl. Acad. Sci.* **1993**, *90*, 3680-3684.
- (8) Varani, G. *Acc. Chem. Res.* **1997**, *30*, 189-195.
- (9) Best, M. D.; Tobey, S. L.; Anslyn, E. V. *Coord. Chem. Rev.* **2003**, *240*, 3-15.
- (10) Mulligan, C.; Fischer, M.; Thomas, G. H. *FEMS Microbiol. Rev.* **2011**, *35*, 68-86.
- (11) Sakai, N.; Matile, S. *J. Am. Chem. Soc.* **2003**, *125*, 14348-14356.
- (12) Bass, R. B.; Strop, P.; Barclay, M.; Rees, D. C. *Science* **2002**, *298*, 1582-1587.
- (13) Cunningham, C. N.; Krukenberg, K. A.; Agard, D. A. *J. Biol. Chem.* **2008**, *283*, 21170-21178.
- (14) Saunders, L. R.; Barber, G. N. *The FASEB Journal* **2003**, *17*, 961-983.

- (15) Bond, P. J.; Guy, A. T.; Heron, A. J.; Bayley, H.; Khalid, S. *Biochemistry* **2011**, *50*, 3777-3783.
- (16) Schug, K. A.; Lindner, W. *Chem. Rev.* **2004**, *105*, 67-114.
- (17) Schmidtchen, F. P.; Berger, M. *Chem. Rev.* **1997**, *97*, 1609-1646.
- (18) Stephan, H.; Gloe, K.; Schiessl, P.; Schmidtchen, F. P. *Supramol. Chem.* **1995**, *5*, 273-280.
- (19) Schrader, T. *Chem. Eur. J.* **1997**, *3*, 1537-1541.
- (20) Blondeau, P.; Segura, M.; Perez-Fernandez, R.; de Mendoza, J. *Chem. Soc. Rev.* **2007**, *36*, 198-210.
- (21) Houk, R. J. T.; Tobey, S. L.; Anslyn, E. V. In *Topics in Current Chemistry*; by Stibor& Ivan; Springer Berlin / Heidelberg, Vol. 255, 2005; pp. 199-229
- (22) Jadhav, V. D.; Herdtweck, E.; Schmidtchen, F. P. *Chem. Eur. J.* **2008**, *14*, 6098-6107.
- (23) Schmuck, C. *Coord. Chem. Rev.* **2006**, *250*, 3053-3067.
- (24) Berger, M.; Schmidtchen, F. P. *J. Am. Chem. Soc.* **1999**, *121*, 9986-9993.
- (25) Echavarren, A.; Galan, A.; Lehn, J. M.; De Mendoza, J. *J. Am. Chem. Soc.* **1989**, *111*, 4994-4995.
- (26) Müller, G.; Riede, J.; Schmidtchen, F. P. *Angew. Chem. Int. Ed.* **1988**, *27*, 1516-1518.
- (27) Bell, T. W.; Khasanov, A. B.; Drew, M. G. B.; Filikov, A.; James, T. L. *Angew. Chem. Int. Ed.* **1999**, *38*, 2543-2547.
- (28) Peschke, W.; Schiessl, P.; Schmidtchen, F. P.; Bissinger, P.; Schier, A. *J. Org. Chem.* **1995**, *60*, 1039-1043.
- (29) Fitzmaurice, R. J.; Kyne, G. M.; Douheret, D.; Kilburn, J. D. *J. Chem. Soc., Perkin Trans. 1* **2002**, *0*, 841-864.
- (30) Yamada, T.; Liu, X.; Englert, U.; Yamane, H.; Dronskowski, R. *Chem. Eur. J.* **2009**, *15*, 5651-5655; and references cited therein.
- (31) Han, J.; Yau, C.-W.; Lam, C.-K.; Mak, T. C. W. *J. Am. Chem. Soc.* **2008**, *130*, 10315-10326.
- (32) Lynch, D. E.; McClenaghan, I. *Cryst. Eng.* **2003**, *6*, 99-107.
- (33) Raczyńska, E. D.; Cyrański, M. K.; Gutowski, M.; Rak, J.; Gal, J.-F.; Maria, P.-C.; Darowska, M.; Duczmal, K. *J. Phys. Org. Chem.* **2003**, *16*, 91-106.
- (34) Zyss, J.; Pecaut, J.; Levy, J. P.; Masse, R. *Acta Crystallogr. Sect. B* **1993**, *49*, 334-342.
- (35) Legrand, Y.-M.; van der Lee, A.; Barboiu, M. *Science* **2010**, *329*, 299-302.
- (36) Mak, T. C. W.; Xue, F. *J. Am. Chem. Soc.* **2000**, *122*, 9860-9861.
- (37) Yang, C.; Wong, W.-T. *Chem. Lett.* **2004**, *33*, 856-857.
- (38) Latham, K.; Downs, J. E.; Rix, C. J.; White, J. M. *J. Mol. Struct.* **2011**, *987*, 74-85.
- (39) Soegiarto, A. C.; Yan, W.; Kent, A. D.; Ward, M. D. *J. Mater. Chem.* **2011**, *21*, 2204-2219.
- (40) Xu, W.-Z.; Sun, J.; Huang, Z.-T.; Zheng, Q.-Y. *Chem. Commun.* **2009**, 171-173.
- (41) Yu, Z.-B.; Sun, J.; Huang, Z.-T.; Zheng, Q.-Y. *CrystEngComm* **2011**, *13*, 1287-1290.
- (42) Sopkova-de Oliveira Santos, J.; Montouillout, V.; Fayon, F.; Fernandez, C.; Delain-Bioton, L.; Villemin, D.; Jaffres, P.-A. *New J. Chem.* **2004**, *28*, 1244-1249.
- (43) Videnova-Adrabinska, V.; Obara, E.; Lis, T. *New J. Chem.* **2007**, *31*, 287-295.
- (44) Videnova-Adrabinska, V.; Turowska-Tyrk, I.; Borowiak, T.; Dutkiewicz, G. *New J. Chem.* **2001**, *25*, 1403-1409.
- (45) Horner, M. J.; Holman, K. T.; Ward, M. D. *Angew. Chem. Int. Ed.* **2001**, *40*, 4045-4048.
- (46) Pivovar, A. M.; Holman, K. T.; Ward, M. D. *Chem. Mater.* **2001**, *13*, 3018-3031.

- (47) Schülke, U.; Averbuch-Pouchot, M. T. *Zeitschrift für anorganische und allgemeine Chemie* **1995**, *621*, 1232-1236.
- (48) Adams, J. M.; Small, R. W. H. *Acta Crystallogr. Sect. B* **1976**, *32*, 832-835.
- (49) Adams, J. M.; Ramdas, V. *Acta Crystallogr. Sect. B* **1976**, *32*, 3224-3227.
- (50) Soegiarto, A. C.; Comotti, A.; Ward, M. D. *J. Am. Chem. Soc.* **2010**, *132*, 14603-14616.
- (51) Lam, C.-K.; Xue, F.; Zhang, J.-P.; Chen, X.-M.; Mak, T. C. W. *J. Am. Chem. Soc.* **2005**, *127*, 11536-11537.
- (52) Katayev, E. A.; Ustynyuk, Y. A.; Sessler, J. L. *Coord. Chem. Rev.* **2006**, *250*, 3004-3037.
- (53) Ivasenko, O.; Perepichka, D. F. *Chem. Soc. Rev.* **2011**, *40*, 191-206.
- (54) Adams, J. M.; Small, R. W. H. *Acta Crystallogr. Sect. B* **1974**, *30*, 2191-2193.
- (55) Han, J.; Zang, S.-Q.; Mak, T. C. W. *Chem. Eur. J.* **2010**, *16*, 5078-5088.
- (56) Peng, B.; Peng, Q.; Zhou, W.; Zhou, Z. *Acta Crystallogr. Sect. E* **2010**, *66*, o2679.
- (57) Bayot, D.; Tinant, B.; Mathieu, B.; Declercq, J.-P.; Devillers, M. *Eur. J. Inorg. Chem.* **2003**, *2003*, 737-743.
- (58) Smith, G.; Wermuth, U. D. *Acta Crystallogr. Sect. E* **2009**, *65*, o3220.
- (59) Smith, G.; Wermuth, U. D.; Healy, P. C. *Acta Crystallogr. Sect. E* **2007**, *63*, o3527-o3528.
- (60) Pereira Silva, P. S.; Ramos Silva, M.; Paixao, J. A.; Matos Beja, A. *Acta Crystallogr. Sect. E* **2007**, *63*, o2783.
- (61) Krumbe, W.; Haussühl, S. *Zeits. Kristallogr. - Crystalline Materials* **1987**, *179*, 267-279.
- (62) Najafpour, M. M.; Holynska, M.; Lis, T. *Acta Crystallogr. Sect. E* **2007**, *63*, o3727.
- (63) Smith, G.; Wermuth, U. D. *Acta Crystallogr. Sect. C* **2010**, *66*, o575-o580.
- (64) Görbitz, C. H.; Etter, M. C. *J. Am. Chem. Soc.* **1992**, *114*, 627-631.
- (65) Görbitz, C. H. *Acta Crystallogr. Sect. B* **2010**, *66*, 84-93.
- (66) Levy, Y.; Onuchic, J. N. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 389-415.
- (67) Oshovsky, G. V.; Reinhoudt, D. N.; Verboom, W. *Angew. Chem. Int. Ed.* **2007**, *46*, 2366-2393.
- (68) Hummer, G.; Rasaiah, J. C.; Noworyta, J. P. *Nature* **2001**, *414*, 188-190.
- (69) Sansom, M. S. P.; Biggin, P. C. *Nature* **2001**, *414*, 156-159.
- (70) Thirumalai, D.; Reddy, G.; Straub, J. E. *Acc. Chem. Res.* **2011**, *45*, 83-92.
- (71) Agre, P. *Angew. Chem. Int. Ed.* **2004**, *43*, 4278-4290.
- (72) Barboiu, M. *Angew. Chem. Int. Ed.* **2012**, *51*, 11674-11676.
- (73) Cheruzel, L. E.; Pometun, M. S.; Cecil, M. R.; Mashuta, M. S.; Wittebort, R. J.; Buchanan, R. M. *Angew. Chem. Int. Ed.* **2003**, *42*, 5452-5455.
- (74) Fei, Z.; Zhao, D.; Geldbach, T. J.; Scopelliti, R.; Dyson, P. J.; Antonijevic, S.; Bodenhausen, G. *Angew. Chem. Int. Ed.* **2005**, *44*, 5720-5725.
- (75) Ludwig, R. *Angew. Chem. Int. Ed.* **2001**, *40*, 1808-1827.
- (76) Natarajan, R.; Charmant, J. P. H.; Orpen, A. G.; Davis, A. P. *Angew. Chem. Int. Ed.* **2010**, *49*, 5125-5129.
- (77) Thallapally, P. K.; Lloyd, G. O.; Atwood, J. L.; Barbour, L. J. *Angew. Chem. Int. Ed.* **2005**, *44*, 3848-3851.
- (78) Wang, H.; Xu, X.; Johnson, N. M.; Dandala, N. K. R.; Ji, H. F. *Angew. Chem. Int. Ed.* **2011**, *50*, 12538-12541.

- (79) Teeter, M. M. *Ann. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 577-600.
- (80) Jensen, M. Ø.; Tajkhorshid, E.; Schulten, K. *Biophys. J.* **2003**, *85*, 2884-2899.
- (81) Zhu, F.; Schulten, K. *Biophys. J.* **2003**, *85*, 236-244.
- (82) Pal, S. K.; Zewail, A. H. *Chem. Rev.* **2004**, *104*, 2099-2124.
- (83) Infantes, L.; Motherwell, S. *CrystEngComm* **2002**, *4*, 454-461.
- (84) Liu, B.; Li, X.; Li, B.; Xu, B.; Zhao, Y. *Nano Lett.* **2009**, *9*, 1386-1394.
- (85) Paixao, J. A.; Matos Beja, A.; Ramos Silva, M.; Alte da Veiga, L. *Acta Crystallogr. Sect. C* **1999**, *55*, 1290-1292.
- (86) Paixao, J. A.; Pereira Silva, P. S.; Matos Beja, A.; Ramos Silva, M.; de Matos Gomes, E.; Belsley, M. *Acta Crystallogr. Sect. C* **1999**, *55*, 1287-1290.
- (87) Paixao, J. A.; Pereira Silva, P. S.; Matos Beja, A.; Ramos Silva, M.; Alte da Veiga, L. *Acta Crystallogr. Sect. C* **1998**, *54*, 1484-1486.
- (88) Hu, M.-L.; Wang, W.-D. *Acta Crystallogr. Sect. E* **2005**, *61*, o1408-o1410.
- (89) Silva, P. S. P.; Cardoso, C. u.; Silva, M. R.; Paixão, J. A.; Beja, A. M.; Garcia, M. H.; Lopes, N. *J. Phys. Chem. A* **2010**, *114*, 2607-2617.
- (90) Said, F. F.; Ong, T.-G.; Bazinet, P.; Yap, G. P. A.; Richeson, D. S. *Cryst. Growth Des.* **2006**, *6*, 1848-1857.
- (91) Said, F. F.; Ali, B. F.; Richeson, D.; Korobkov, I. *Acta Crystallogr. Sect. E* **2012**, *68*, o1906.
- (92) Coles, M. P. *Chem. Commun.* **2009**, 3659-3676.
- (93) Oakley, S. H.; Soria, D. B.; Coles, M. P.; Hitchcock, P. B. *Dalton Trans.* **2004**, 537-546.
- (94) Simoni, D.; Rossi, M.; Rondanin, R.; Mazzali, A.; Baruchello, R.; Malagutti, C.; Roberti, M.; Invidiata, F. *P. Org. Lett.* **2000**, *2*, 3765-3768.
- (95) Phillips, D. J.; Graham, A. E. *Synlett* **2010**, 769,773.
- (96) Pratt, R. C.; Lohmeijer, B. G. G.; Long, D. A.; Waymouth, R. M.; Hedrick, J. L. *J. Am. Chem. Soc.* **2006**, *128*, 4556-4557.
- (97) Das Neves Gomes, C.; Jacquet, O.; Villiers, C.; Thuéry, P.; Ephritikhine, M.; Cantat, T. *Angew. Chem. Int. Ed.* **2012**, *51*, 187-190.
- (98) Meetsma, A.; van Aken, E.; Wynberg, H. *Acta Crystallog. Sect. C* **1992**, *48*, 1874-1876.
- (99) Cotton, F. A.; Timmons, D. J. *Polyhedron* **1998**, *17*, 179-184.
- (100) Huczynski, A.; Ratajczak-Sitarz, M.; Katrusiak, A.; Brzezinski, B. *J. Mol. Struct.* **2008**, *888*, 84-91.
- (101) Wild, U.; Roquette, P.; Kaifer, E.; Mautz, J.; Hübner, O.; Wadepohl, H.; Himmel, H.-J. *Eur. J. Inorg. Chem.* **2008**, *2008*, 1248-1257.
- (102) Binkowska, I.; Jarczewski, A.; Katrusiak, A.; Wojciechowski, G.; Brzezinski, B. *J. Mol. Struct.* **2001**, *597*, 101-107.
- (103) van Aken, E.; Wynberg, H.; van Bolhuis, F. *J. Chem. Soc., Chem. Commun.* **1992**, 629-630.
- (104) Ng, S. W.; Naumov, P.; Chantrapromma, S.; Shanmuga Sundara Raj, S.; Fun, H. K.; Ibrahim, A. R.; Wojciechowski, G.; Brzezinski, B. *J. Mol. Struct.* **2001**, *562*, 185-196.
- (105) Wojciechowski, G.; Brzezinski, B.; Naumov, P.; Chantrapromma, S.; Ibrahim, A. R.; Fun, H. K.; Ng, S. W. *J. Mol. Struct.* **2001**, *598*, 153-159.

- (106) Yaghi, O. M.; O'Keeffe, M.; Ockwig, N. W.; Chae, H. K.; Eddaoudi, M.; Kim, J. *Nature* **2003**, *423*, 705-714.
- (107) Murphy, P. J.; Harri Lloyd, W.; Hibbs, D. E.; Hursthouse, M. B.; Abdul Malik, K. M. *Tetrahedron* **1996**, *52*, 8315-8332.
- (108) Bruker APEX2, SAINT-Plus and SADABS (2007), Bruker AXS Inc., Madison, Wisconsin, USA.
- (109) Sheldrick, G. *Acta Crystallogr. Sect. A* **2008**, *64*, 112-122.
- (110) Macrae, C. F.; Bruno, I. J.; Chisholm, J. A.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Rodriguez-Monge, L.; Taylor, R.; van de Streek, J.; Wood, P. A. *J. Appl. Crystallogr.* **2008**, *41*, 466-470.
- (111) Braga, D.; Grepioni, F.; Maini, L.; Polito, M. *Structure & Bonding*; Hosseini, M. W. Ed.; Springer Berlin / Heidelberg, 2009; pp. 25-50.
- (112) Desiraju, G. R. *Science* **1997**, *278*, 404-405.
- (113) Kato, T.; Okamoto, I.; Masu, H.; Katagiri, K.; Tominaga, M.; Yamaguchi, K.; Kagechika, H.; Azumaya, I. *Cryst. Growth Des.* **2008**, *8*, 3871-3877.
- (114) Nangia, A.; R. Desiraju, G. *Chem. Commun.* **1999**, 605-606.
- (115) Jetti, R. K. R.; Boese, R.; Thallapally, P. K.; Desiraju, G. R. *Cryst. Growth Des.* **2003**, *3*, 1033-1040.
- (116) S. Senthil Kumar, V.; S. Kuduva, S.; R. Desiraju, G. *J. Chem. Soc., Perkin Transactions 2* **1999**, 1069-1074.
- (117) Luboradzki, R.; Gronwald, O.; Ikeda, M.; Shinkai, S.; Reinhoudt, D. N. *Tetrahedron* **2000**, *56*, 9595-9599.

V

