EurJOC
European Journal of Organic Chemistry

A Journal of
ChemPubSoc Europe

Accepted Article

Title: Structural Insights into the Hydrogen-Bonding and Folding Pattern in Ant-Ant-Pro-Gly Tetrapeptides

Authors: Sachin B Baravkar, Amol S Kotmale, Samir R Shaikh, Rajesh G Gonnade, and Gangadhar J. Sanjayan

To be cited as: *Eur. J. Org. Chem.* 10.1002/ejoc.201700181

Link to VoR: http://dx.doi.org/10.1002/ejoc.201700181

Supported by
ACES Asian Chemical Editorial Society

WILEY-VCH

Structural Insights into the Hydrogen-Bonding and Folding Pattern in Ant-Ant-Pro-Gly Tetrapeptides

Sachin B. Baravkar,[a] Amol S. Kotmale,[b] Samir R. Shaikh,[c] Rajesh G. Gonnade[c] and Gangadhar J. Sanjayan*[a]

Abstract: In this paper, we provide structural insights into the hydrogen-bonding and folding pattern in Ant-Ant-Pro-Gly tetrapeptides (Ant: anthranilic acid; Pro: proline; and Gly: glycine). Comparison of the C-terminal ester and their amide analogs revealed strikingly different H-bonding networks. Whereas the ester analogs displayed an open structure without having terminal H-bonding interactions, the amide analogs showed completely folded structure. Structural details were obtained using a combination of X-ray crystal structure studies and nOe-based MD simulation studies.

Introduction

Folding is a crucial process through which biopolymers fold into their characteristic three-dimensional compact structures mostly because of the presence of a collection of non-covalent forces.[1] Of all non-covalent forces, hydrogen-bonding interaction plays a vital role in the stabilization of secondary structures. Hydrogen-bonding interactions assist biopolymers to adopt a well defined structure - out of millions of other possible conformations and carry out their specific functions.

Foldamers are conformationally ordered synthetic oligomers which fold into a specific conformation akin to bio-macromolecules.[2] Owing to the presence of well-defined three-dimensional structure, foldamers find applications in biomedical science,[3] molecular recognition,[4] catalysis[5] and material chemistry.[6] Hydrogen bonding, along with other conformational constraints such as torsional angle and π-π stacking, plays a pivotal role in imparting folded structure to biopolymers.[7] Peptides having C-terminal esters are sometimes seen devoid of intramolecular hydrogen bonding at the termini (fraying of peptide chains), when compared to their amide counterparts.[8] There are several other contributing factors for peptide chain fraying which include π-π interactions, NH-π interactions and bulkiness of the groups at the termini.[9]

Anthranilic acid (Ant) - a constrained aromatic β-amino acid forms a sheet-like secondary structure in its homo-oligomer **1** (Figure 1) due to the presence of strong intramolecular six-membered (C6) hydrogen bonding between NH and CO group on the individual Ant ring.[10] However, a hybrid (α/β) peptide having Ant and proline (Pro) showed a turn-like conformation characterized by a nine-membered intramolecular hydrogen bonding between the NH of Ant ring and CO group of proline ring. The steric clash between the nearby Ant and Pro groups in the dipeptide **2** (Figure 1) caused the two rings to deviate from planarity, resulting in a folded conformation.[11] It is noteworthy that six-membered (C6) hydrogen bonding was conspicuously absent in **2**. Strangely, introduction of a glycine (Gly) residue into **2**, as shown in tripeptide **3**, resulted in the loss of 9-membered H-bonding, but showed the standard 10-memebered H-bonding (β-turn), expected from Pro-Gly β-turn inducing motif.[12] Intriguingly, introduction of proline as a third residue, as in tripeptide **4**, resulted in the loss of both 9- and 6-membered H-bondings; and only a not-so-common combination of 10- and 11-membered H-bondings was observed.[13]

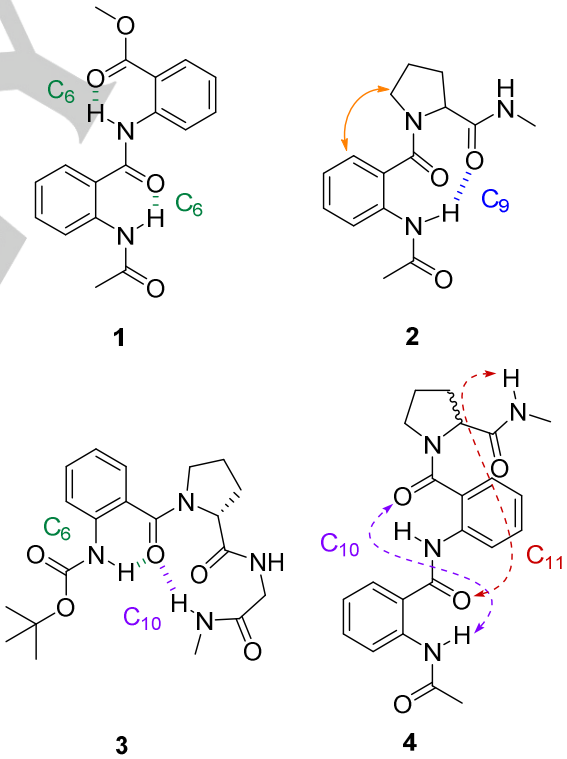


Figure 1. Some typical examples of H-bonding propensities in anthranilic acid (Ant)-containing peptides.

In the present study, we got interested in investigating the conformation of Ant-Ant-Pro-Gly tetrapeptides and consequences on exchanging C-terminus ester with amide.

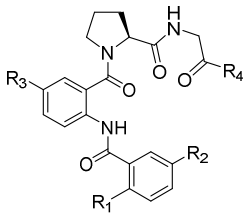
[a] *Division of Organic Chemistry, CSIR-National Chemical Laboratory (CSIR-NCL), Dr. Homi Bhabha Road, Pune 411 008, India. Fax: +91-020-2590-2629;Tel+91-020-2590-2082; E-mail: gj.sanjayan@ncl.res.in*; Homepage: http://nclwebapps.ncl.res.in/gjsanjayan/index.html
[b] *Central NMR facility, CSIR-National Chemical Laboratory (CSIR-NCL), Dr. Homi Bhabha Road, Pune 411 008, India.*
[c] *Center for Materials Characterization, CSIR-National Chemical Laboratory (CSIR-NCL), Dr. Homi Bhabha Road, Pune 411 008, India.*

Supporting information for this article is given via a link at the end of the document. ((Please delete this text if not appropriate))

Results and Discussion

Synthesis:

The syntheses of tetrapeptides **5**-**10** (Figure 2) were carried out using solution-phase peptide coupling reactions (Supporting information, page S3 and S10). After extensive attempts with crystallization trials, we could get the crystals for two peptides (**5** and **10**). On the other hand, conformational studies of other selected four peptides (**6**-**9**) were undertaken by nOe-based MD simulated structures.



Compd	R_1	R_2	R_3	R_4
5	NHAc	iBuO	iBuO	OMe
6	NHAc	H	H	OMe
7	NO_2	H	H	OMe
8	NHAc	H	H	NHMe
9	NO_2	H	H	NHMe
10	H	H	Cl	NHMe

Figure 2. The general molecular structure (left) of tetrapeptides **5**-**10** and a table (right) containing their substitution pattern.

Solution-state conformational analysis of peptides 5-10:

The strength and the nature of hydrogen bonding interactions (intermolecular *vs* intramolecular) in the solution-state can convincingly be ascertained by solvent titration of the compounds (dissolved in non-polar solvent) with polar solvents. DMSO, a polar aprotic solvent, interacts with polar amide NHs present within the molecule, resulting in change in the chemical shift values of amide NHs. However, NHs which are involved in the intramolecular hydrogen bonding are less likely to be affected by the solvent polarity, and hence show negligible change in the chemical shift values. Peptides **6**, **7**, **8** and **9** (5 mM in $CDCl_3$) were titrated gradually against DMSO-d_6. Figure 3 (a), (b), (c) and (d) show the change in the chemical shift values *versus* the volume (in *µL*) of DMSO-d_6 added for the tetrapeptides **6**, **7**, **8** and **9**, respectively. The NHs which showed negligible change in the chemical shift values (≤ 0.41 ppm) were considered to be involved in intramolecular hydrogen bonding.[11] It was also evident that the Gly NHs ($\Delta\delta$ NH_{Gly} > 0.5 ppm) do not participate in intramolecular H-bonding in all the peptides. Further, it became quite clear from the $\Delta\delta NH_{c\text{-terminus}}$ values that the C-terminus NHs of amide analogs **8** and **9**, were involved in the intramolecular H-bonding ($\Delta\delta NH_{c\text{-terminus}}$ ≤ 0.4 ppm).

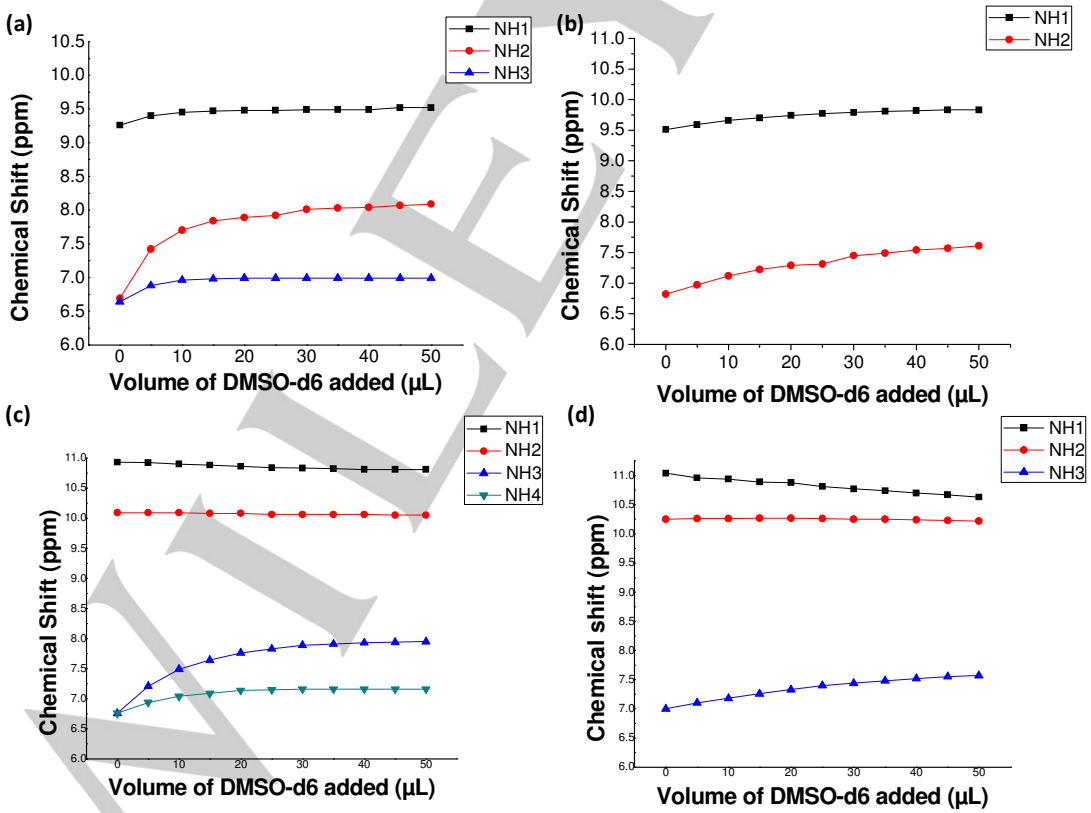


Figure 3. DMSO-*d6* titration studies of the peptides **6** (a), **7** (b), **8** (c) and **9** (d) in $CDCl_3$.

In order to investigate the solution-state conformational studies of the peptides, we carried out 2D NMR studies and the signal assignments were carried out by COSY, TOCSY, HSQC, HMBC and NOESY experiments for all the analogs **5**-**10**. The characteristic strong nOe interactions such as NH1/C8H, NH1/C15H and NH2/C15H suggested that peptide **5** [Figure 4 (a)] showed a similar conformation in the solution-state, as seen in its crystal structure [Figure 6 (a)]. It is noteworthy that the ester analogs **6** and **7** show similar nOe interactions along with some additional nOe interactions as shown in Figure 4 (b, c).

Amide analog **8** [Figure 5 (a)] showed similar nOe interactions, like its ester analog **6**, except showing additional long range inter-residual nOe interactions such as NH2/NH4 and C24H/C5H due to the close proximity of the folded termini as a result of 10-membered H-bond between C-terminus amide NH and CO group of Ant2. 2D NOESY spectrum of **9** also showed similar nOe interactions [Figure 5 (b)]. The presence of NH1/NH3 and C22H/C3H nOe interactions between the aromatic ring (Ant1) and C-terminus methyl group further supported the folded conformation.

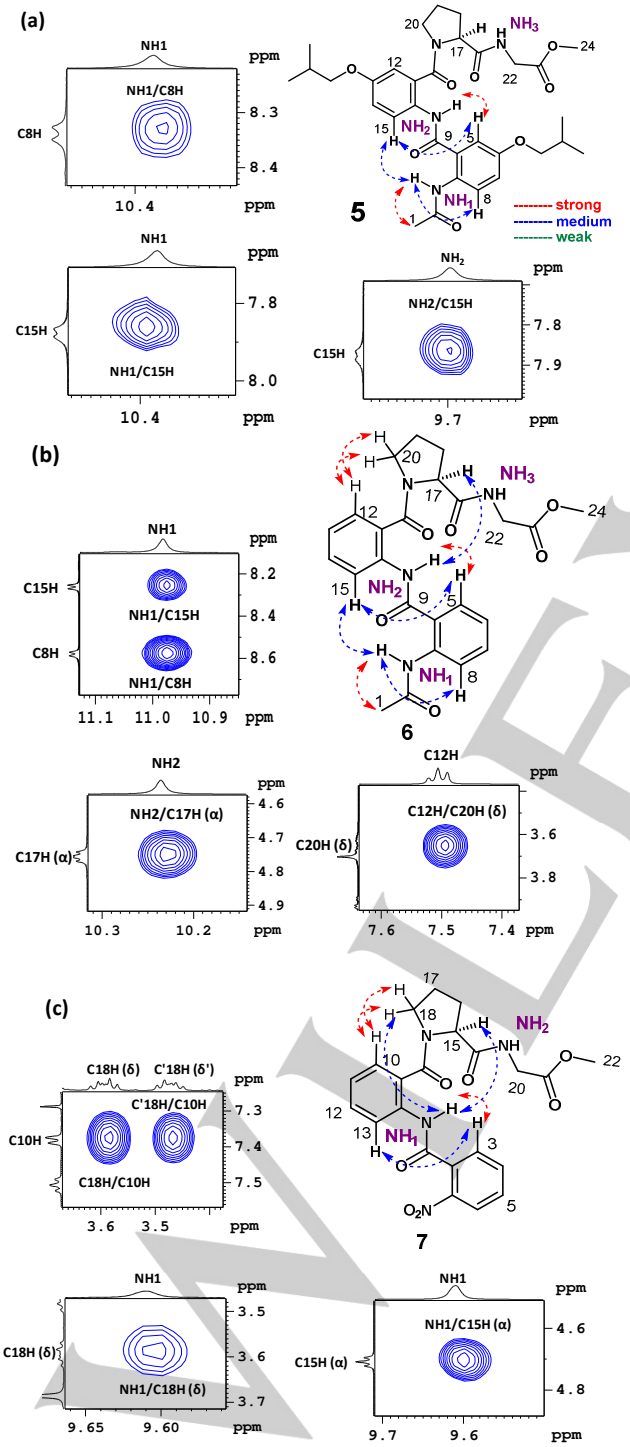


Figure 4. Selected nOe excerpts for C-terminus ester analogs **5** (a), **6** (b) and **7** (c).

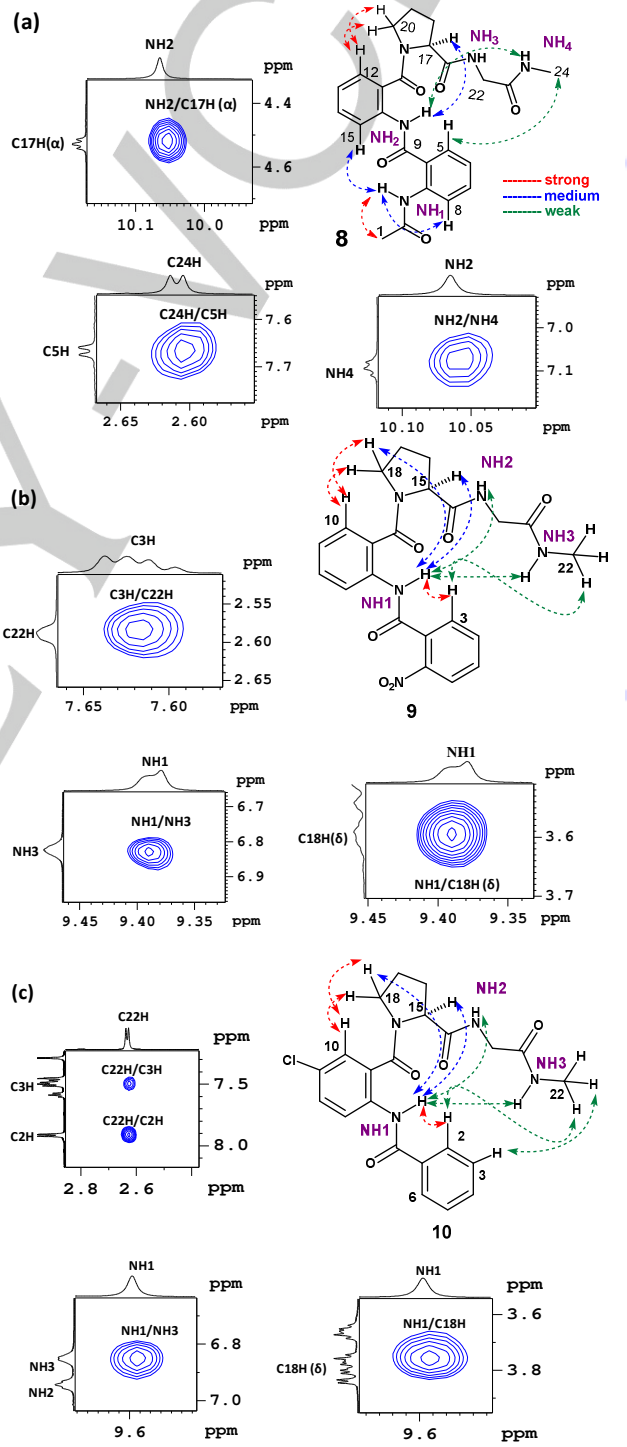


Figure 5. Selected nOe excerpts for the amide analogs **8** (a), **9** (b) and **10** (c).

The folded conformation observed in the solid-state structure of peptide **10** [Figure 7 (c)] was clearly reflected in its solution-state structure as well, as evident from characteristic nOes (NH1/NH3, C22H/C2H and C22H/C3H) [Figure 5 (c)].

Structural studies of C-terminus ester analogs 5-7:

The single crystal X-ray structure analysis of compound **5**[14] is shown below [Figure 6 (a)]. The amide NHs present on both the Ant rings are involved in strong 6-membered H-bonding (C6) with the corresponding carbonyl groups of Ant; with the distance of 1.9 Å (NH of Ant1) and 2.0 Å (NH of Ant2), respectively. The NH of Gly4 is directing outwards without participating in intramolecular H-bonding and the ester group at the C-terminus is found to be fraying away, presumably due to the repulsive interactions giving rise to an extended open structure. The MD simulated structure obtained by using 2D NMR studies for peptide **6** [Figure 6 (b)] showcases similar structural trend seen for peptide **5** with both Ant NHs forming intramolecular 6-membered H-bond with CO group of Ant1 and Ant2. The C-terminus ester group is seen to be fraying away as in peptide **5**, affording an open structure. Ester **7** [Figure 6 (c)] with a N-terminus NO_2 group showed only one 6-membered H-bond, (Ant2), as expected. It is noteworthy that the C-9 turn, usually found in Ant-Pro amides,[8, 11] was conspicuously absent in all these C-terminus esters **5**-**7**.

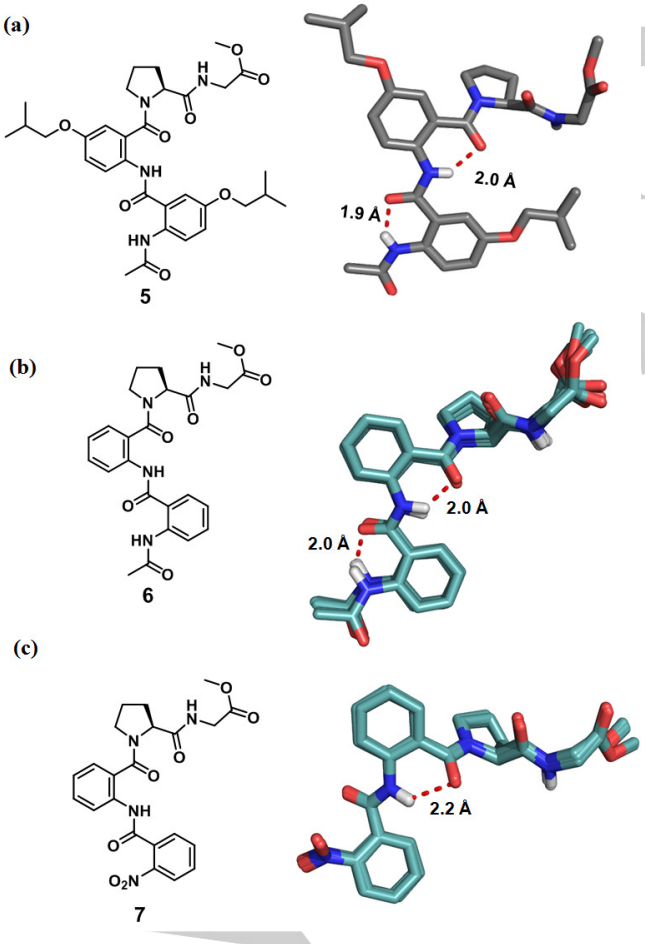


Figure 6. Molecular structures (left) and corresponding crystal (**5**) and nOe-based MD simulated 20 minimum energy structures **6**, **7** (right) of Ant-Ant-Pro-Gly C-terminus esters.

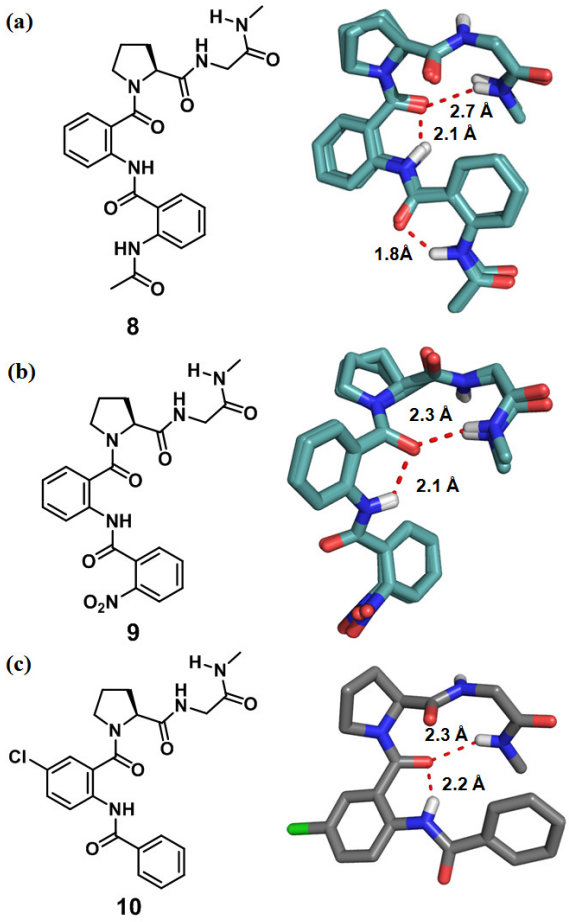


Figure 7. Molecular structures (left) and corresponding crystal (**10**) and nOe-based MD simulated structures of **8**, **9** (right) of Ant-Ant-Pro-Gly C-terminus amides.

Structural studies of C-terminus amide analogs 8-10:

The unpredictable H-bonding features in the Ant-Pro peptide series prompted us to undertake the conformational studies of their corresponding amide analogs. Unlike the ester analogs **5**-**7**, C-terminal amide analogs **8**-**10** featured different structural architecture, essentially dictated by the C-terminus Pro-Gly amide motif. The nOe-based MD simulated 20 minimum energy structures for peptide **8** [Figure 7 (a)] and **9** [Figure 7 (b)] showed completely folded conformation. Folded structure in amide **8** was observed owing to the presence of concurrent display of 6-membered (involving Ant) and a 10-membered H-bonding pattern (emanating from the Pro-Gly amide motif). Similar structural feature was observed for amide **9** (N-teminus NO_2 group) with the presence of 6- and 10-membered H-bonding pattern giving folded conformation as seen in Figure 7 (b). The crystal structure of peptide **10**[14] [Figure 7 (c)] showed a similar folded structure due to the presence of concerted H-bonding formed in the backbone (6-membered H-bonding involving Ant and a 10-membered H-bonding pattern involving Pro-Gly amide motif) with the H-bonding distance of 2.2 Å and 2.3 Å for 6- and 10-membered H-bond, respectively. Notably, C-9 turn, usually found in Ant-Pro amides, was conspicuously absent in all these peptides as well.

Conclusions

In conclusion, extensive structural studies of Ant-Ant-Pro-Gly tetrapeptides provided insights into their folding pattern. Comparison of their conformations suggested that the C-terminus ester analogs display an open (extended) structure having only 6-membered H-bonding network in the backbone, emanating from Ant residues. However, amide analogs namely **8**, **9** and **10** displayed fully folded conformation featuring 10-membered intramolecular H-bonding pa ttern – dictated by the Pro-Gly motif. Strikingly, the C-9 turn, usually found in Ant-Pro amides, was conspicuously absent in all these peptides. Our studies have provided unexpected structural insights into the folding pattern of Ant-Ant-Pro-Gly tetrapeptides, which will have a bearing in the development of *de novo* peptide sequences featuring Ant; a work which is currently underway.

Experimental Section

Crystal data of Compound 5:

Single crystals of **5** were obtained by slow evaporation of the solution of ethyl acetate and pet ether (1:5). $C_{32}H_{42}N_4O_3$, M = 610.69, colorless plate, 0.18 x 0.10 x 0.07 mm^3, orthorhombic, space group $P2_12_12_1$, a = 4.86100(10) Å, b = 23.7384(5) Å, c = 26.9033(6) Å, V = 3104.44(11) Å^3, Z = 4, T = 100 K, $2\theta_{max}$ = 5000°, D_{calc} (g cm^{-3}) = 1.307, $F(000)$ = 1304, μ (mm^{-1}) = 0.094, 40847 reflections collected, 5422 unique reflections (R_{int}=0.0292), 5312 observed ($I > 2\sigma(I)$) reflections, multi-scan absorption correction, T_{min} = 0.983, T_{max} = 0.993, 403 refined parameters,, S = 1.089, $R1$ = 0.0262, $wR2$ = 0.0632 (all data R = 0.0269, $wR2$ = 0.0636), maximum and minimum residual electron densities; $\Delta\rho_{max}$ = 0.170, $\Delta\rho_{min}$= -0.161 (eÅ^{-3}).

Crystal data of Compound 10:

Single crystals of 10 were obtained by slow evaporation of the solution of DCM and pet ether (1:4). $C_{22}H_{23}Cl_1N_4O_4$, M = 442.89, colorless plate, 0.40 x 0.33 x 0.19 mm^3, monoclinic, space group $P2_1$, a = 5.7341(3) Å, b = 17.8299(8) Å, c = 10.6686(5) Å, β = 100.7760(10)°, V = 1071.51(9) Å^3, Z = 2, T = 100 K, $2\theta_{max}$ = 56.00°, D_{calc} (g cm^{-3}) = 1.373, $F(000)$ = 464, μ (mm^{-1}) = 0.215, 13808 reflections collected, 5182 unique reflections (R_{int}=0.0141), 5095 observed ($I > 2\sigma(I)$) reflections, multi-scan absorption correction, T_{min} = 0.919, T_{max} = 0.960, 293 refined parameters,, S = 1.042, $R1$ = 0.0253, $wR2$ = 0.0633 (all data R = 0.0258, $wR2$ = 0.0637), maximum and minimum residual electron densities; $\Delta\rho_{max}$ = 0.259, $\Delta\rho_{min}$= -0.158 (eÅ^{-3}).

The synthesis and characterization of all new compounds are described in supporting information.

The syntheses of the peptides **5**-**10** were carried out by using solution-phase peptide coupling reactions. The crystal structure analysis of compound **5** was done on Bruker D8 VENTURE Kappa Duo PHOTON II CPAD diffractometer. Crystal structure analysis of compound **10** was carried out using Bruker SMART APEX II CCD diffractometer. Solution-state 2D NMR studies were carried out in $CDCl_3$ on AV 500 MHz Bruker NMR spectrometer. Molecular dynamics NMR-based structures of peptides **6**, **7**, **8** and **9** were derived from distant constraints from nOe cross peaks by using macromodel-v10.8 from Schrödinger software. For more details, see supporting information.

Acknowledgements

S.B.B, A.S.K and S.R.S are thankful to CSIR, New Delhi for research fellowships. GJS thanks CSIR-SSB project: SSB-000726 for financial support.

Keywords: H-bonding • Folding • chain-fraying • β-turn • MD simulation

[1] (a) L. M. Salonen, M. Ellermann, F. Diederich, *Angew. Chem.***2011**, *50*, 4908-4944; *Angew. Chem. Int. Ed.* **2011**, *50*, 4808-4842. (b) G J. Rose, P. J. Fleming, J. R. Banvar, A. Maritan, *Proc. Natl. Acad. Sci.***2006**, *103*, 16623-16633. (c) C. B. Anfinsen, H. A. Scheraga, *Adv. Protein Chem.* **1975**, *29*, 205-300. (d) C. B. Anfinsen, *Science* **1973**, *181*, 223-230.

[2] (a) C. M. Lombardo, G. W. Collie, K. Pulka-Ziach, F. Rosu, V. Gabelica, C. D. Mackereth, G. Guichard, *J. Am. Chem. Soc.* **2016**, *138*, 10522-10530. (b) T. Hjelmgaard, O. Roy, L. Nauton, M. El-Ghozzi, D. Avignant, C. Didierjean, C. Taillefumier, S. Faure, *Chem. Commun.* **2014**, *50*, 3564-3567. (c) J. Fremaux, B. Kauffmann, G. Guichard, *J. Org. Chem.* **2014**, *79*, 5494-5502. (d) N. Castellucci, C. Tomasini, *Chem. Soc. Rev.* **2013**, *42*, 156-172. (e) E. D. Santis, T. Hjelmgaard, C. Caumes, S. Faure, B. D. Alexander, S. J. Holder, G. Siligardi, C. Taillefumier, A. A. Edwards, *Org. Biomol. Chem.* **2012**, *10*, 1108-1122. (f) C. Caumes, O. Roy, S. Faure, C. Taillefumier, *J. Am. Chem. Soc.* **2012**, *134*, 9553-9556. (g) K. Basuroy, B. Dinesh, N. Shamala, P. Balaram, *Angew. Chem.* **2012**, *124*, 8866-8869; *Angew. Chem., Int. Ed.* **2012**, *51*, 8736-8739. (h) T. A. Martinek, F. Fulop, *Chem. Soc. Rev.* **2012**, *41*, 687-702. (i) G. Guichard, I. Huc, *Chem. Commun.* **2011**, *47*, 5933-5941. (j) M. I. Simone, A. A. Edwards, G E. Tranter, G. W. J. Fleet, *Amino Acids* **2011**, *41*, 643-661. (k) Y. D. Wu, S. H. Gellman, *Acc. Chem. Res.*, **2008**, *41*, 1231-1232. (l) C. M. Goodman, S. Choi, S. Shandler, W. F. DeGrado, *Nat. Chem Biol.*, **2007**, *3*, 252-262. (m) D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes, J. S. Moore, *Chem. Rev.* **2001**, *101*, 3893-4012. (n) S. H. Gellman, *Acc. Chem. Res.* **1998**, *31*, 173-180.

[3] (a) G. N. Tew, R. W. Scott, M. L. Klein, W. F. DeGrado, *Acc. Chem. Res.*, **2009**, *43*, 30-39. (b) P. I. Arvidsson, J. Frackenpohl, N. S. Ryder, B. Liechty, F. Petersen, H. Zimmermann, G. P. Camenisch, R. Woessner, D. Seebach, *ChemBiochem* **2001**, *2*, 771–773. (c) Y. Hamuro, J. P. Schneider, W. F. DeGrado, *J. Am. Chem. Soc.* **1999**, *121*, 12200-12201.

[4] (a) J. Garric, J.M. Leger, I. Huc, *Angew. Chem.* **2005**, *44*, 1990-1994; *Angew. Chem. Int. Ed.* **2005**, *44*, 1954-1958. (b) J. D. Sadowsky, M. A. Schmitt, H. S. Lee, N. Umezawa, S. M. Wang, Y Tomita, S. H. Gellman, *J. Am. Chem. Soc.* **2005**, *127*, 11966-11968. (c) J. T. Ernst, J. Becerril, H. S. Park, H. Yin, A. D. Hamilton, *Angew. Chem.* **2003**, *115*, 553-557; *Angew. Chem. Int. Ed.,* 2003, *42*, 535-539. (d) I. Huc, V. Maurizot, H. Gornitzka, J.M. Leger, *Chem. Commun.* **2002**, *6*, 578-579.

[5] (a) M. M. Muller, M. A. Windsor, W. C. Pomerantz, S. H. Gellman, D. Hilver, *Angew. Chem.* **2009**, *121*, 940-943; *Angew. Chem. Int. Ed.* **2009**, *48*, 922-925. (b) K. W. Fiori, A. L. A. Puchlopek, S. J. Miller, *Nat. Chem,* **2009**, *1*, 630-634.

[6] (a) Q. Gan, Y. Ferrand, C. Bao, B. Kauffman, A. Grelard, H. Jaing, I. Huc, *Science* **2011**, *331*, 1172-1175. (b) G. T. Wang, X. Zhao, Z. T. Li, *Tetrahedron* **2011**, *67*, 48-57. (c) S. Kwon, A. Jeon, S. H. Yoo, I. S. Chung, H. S. Lee, *Angew. Chem. Int. Ed.* **2010**, *122*, 8223-8227; *Angew. Chem. Int. Ed.* **2010**, *49*, 8232-36.

[7] C. R. Jones, G. D. Pantos, A. J. Morrison, M. D. Smith, *Angew. Chem.* **2009**, *121*, 7527-7530; *Angew. Chem. Int. Ed,* **2009**, *48*, 7391-7394.

[8] K. N. Vijayadas, R. V. Nair, R. L. Gawade, A. S. Kotmale, P. Prabhakaran, R. G. Gonnade, V. G. Puranik, P. R. Rajamohanan, G. J. Sanjayan, *Org. Biomol. Chem.* **2013**, *11*, 8348-8356.

[9] (a) R. Mahalakshmi, A. Sengupta, S. Raghothama, N. Shamala, P. Balaram, *Biopolymers* **2007**, *88*, 36-54. (b) V. V. Harini, S. Aravinda, R. Rai, N. Shamala, P. Balaram, *Chem. Eur. J.*, **2005**, *11*, 3609-3620. (c) S. Aravinda, V. V. Harini, N. Shamala, C. Das, P. Balaram, *Biochemistry* **2004**, *43*, 1832-1846. (d) A. L. Rucker, T. P. Creamer, *Pept. Sci.* **2002**, *11*, 980-985.

[10] Y. Hamuro, S. J. Geib, A. D. Hamilton, *J. Am. Chem. Soc.* **1996**, *118*, 7529-7541.

[11] P. Prabhakaran, S. S. Kale, V. G. Puranik, P. R. Rajamohanan, O. Chetina, J. A. K. Howard, H. J. Hofmann and G. J. Sanjayan *J. Am. Chem. Soc.*, **2008**, *130*, 17743-17754.

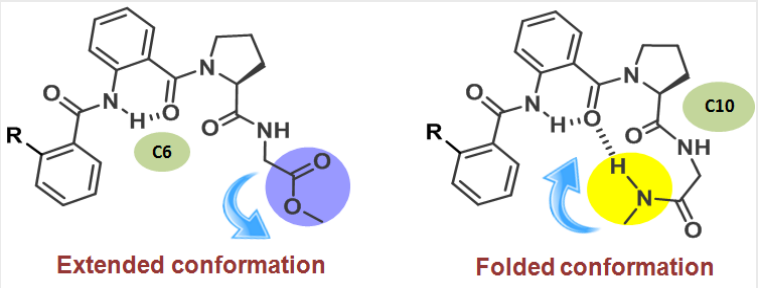
[12] (a) S. R. Raghothama, S. K. Awasthi, P. Balaram, *J. Chem. Soc., Perkin Trans. 2,* **1998**, *1*, 137-144. (b) I. L. Karle, S. K. Awasthi, P. Balaram, *Proc. Natl. Acad. Sci. U.S.A.*, **1996**, *93*, 8189-8193.

[13] V. V. E. Ramesh, G. Priya, A. S. Kotmale, R. G. Gonnade, P. R. Rajamohanan, G. J. Sanjayan, *Chem. Commun.* **2012**, *48*, 11205-11207.

[14] CCDC-1524537 (for peptide **5**) and CCDC-1524536 (for peptide **10**) contain the supplementary crystallographic data for this paper. These data can be obtained free of charge from Data Centre via www.ccdc.cam.ac.uk/data_request/cif.

Entry for the Table of Contents (Please choose one layout)

FULL PAPER



Key Topic*

*Sachin B. Baravkar, Amol S. Kotmale, Samir R. Shaikh, Rajesh G. Gonnade and Gangadhar J. Sanjayan **

Page No. – Page No.

Structural Insights into the Hydrogen-Bonding and Folding Pattern in Ant-Ant-Pro-Gly Tetrapeptides

Synthesis and conformational analysis of peptides containing C-terminal ester and amide was undertaken. Conformational analysis was carried out by using 2D NMR spectroscopy and X-ray crystallographic studies which proved that ester analogs displayed an extended conformation, whereas, amide analogs demonstrated folded conformation owing to the presence of C-terminus β-turn.