

First-principles structures for the close-packed and the 7/2 motif of collagen

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2012 EPL 100 28003

(<http://iopscience.iop.org/0295-5075/100/2/28003>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 193.174.53.85

This content was downloaded on 28/02/2014 at 17:52

Please note that [terms and conditions apply](#).

First-principles structures for the close-packed and the 7/2 motif of collagen

KARL J. JALKANEN¹, KASPER OLSEN¹, MICHAELA KNAPP-MOHAMMADY² and JAKOB BOHR¹

¹ DTU Nanotech, Ørstedes Plads, Technical University of Denmark - DK-2800 Kongens Lyngby, Denmark

² Division of Functional Genome Analysis, German Cancer Research Center - D-69120 Heidelberg, Germany

received 12 July 2012; accepted 24 September 2012

published online 31 October 2012

PACS 87.14.em – Biomolecules: types: Fibrils (amyloids, collagen, etc.)

PACS 87.15.ag – Biomolecules: structure and physical properties: Quantum calculations

PACS 87.15.M– – Biomolecules: structure and physical properties: Spectra of biomolecules

Abstract – The newly proposed close-packed motif for collagen and the more established 7/2 structure are investigated and compared. First-principles semi-empirical wave function theory and Kohn-Sham density functional theory are applied in the study of these relatively large and complex structures. The structures are refined, and comparable stability is displayed using these methods. The electronic circular dichroism of the close-packed model is shown to have a significant negative bias and a large signal. This bias is consistent with existing experimental data, though it is not unambiguously possible from the circular dichroism calculations to select one structure over the other. An interesting feature of the close-packed structure is the existence of a central channel. Simulations show that, if hydrogen atoms are placed in the cavity, a chain of molecular hydrogens is formed suggesting a possible biological function for molecular hydrogen.



Copyright © EPLA, 2012

Hitherto, the 10/3 structure suggested in the 1950s by Ramachandran, and Rich and Crick [1–4] and the 7/2 structure suggested in the 1970s by Okuyama *et al.* [5,6] have been the most studied motifs for collagen [7] and collagen-like peptides [8,9]. These two motifs belong to the same symmetry class: three left-handed helical polypeptide chains are supercoiled in a right-handed arrangement. For long helical structures the relatively short range of the strand and the interstrand interactions does not favour commensuration, *i.e.*, that an integer number of residues on the polypeptide strands corresponds to one, or multiple, 2π supercoilings. Hence, in general the structure will not be periodic. For the periodic 10/3 structure it takes ten residues to make three full 2π rotations to complete the ~ 85 Å unit cell, while for the periodic 7/2 structure it takes seven residues of one polypeptide chain to complete two full 2π rotations which complete a unit cell length of about 60 Å. A contributing factor to the original suggestion to change focus from the 10/3 structure to the 7/2 structure was diffraction spots in X-ray patterns corresponding to a longitudinal period of 20 Å which were unaccounted for by the 10/3 structure [10].

In this letter, we consider the close-packed (CP) structure for collagen [11] and subject this motif to first-principles calculations and compare with calculations for

the 7/2 motif. The motivation behind the CP structural motif is the remarkable coincidence that it is optimally packed while at the same time having a vanishing coupling between strain and twist [12]. While the 7/2 and 10/3 structural motifs are supercoiled triple helix structures, the CP structure is akin to that of a rope where the three polypeptide strands are intertwined. The CP geometry is the non-periodic triple helix which optimises the volume fraction, it has a helical pitch of 20 Å.

In high resolution crystallographic studies of peptide-like structures the question of whether the 7/2 or the 10/3 structure describe the data, or not, is not always resolved. *R*-factors indicate that both structures describe features of the diffraction data [8,13–16]. Therefore, the atoms have been assigned differently to the three strands and hence the assignment of the atoms to the individual strands remains uncertain. Perhaps a third structure such as the CP structure would be a better overall description which could account for some of the experimental data not yet accounted for by either the 7/2 or the 10/3 structures.

Density functional theory (DFT) studies of various collagen-like peptides is an active research area [17,18] including the interesting issue of studying biomineralization [19]. For the initial configurations of the first-principles calculations we used the approximate

coordinates for the CP structure obtained by simple geometrical methods [11] and for the 7/2 structure the coordinates obtained from the PDB depository (entry: 1K6F 3[Pro-Pro-Gly]₁₀) [13]. These structures were then truncated to various lengths (3[PPG]₆, 3[PPG]₃ and 3[PPG]₁) and hydrogens were added with the InsightII program from Biosym Technologies, Inc. To treat the effects of an aqueous environment, we have utilised the polarised continuum model (PCM) as implemented in the Gaussian 09 program [20–22].

The semi-empirical wave function theory (WFT) PM6 model of Stewart [23] and the CAM-B3LYP [24,25] Kohn-Sham density functional theory (KS-DFT) exchange correlation functionals have been used in this work. For the larger model, 3(PPG)₆, only the PM6 model was used for geometry optimisation.

Circular dichroism is commonly used to distinguish structures with various degrees of chirality, *e.g.*, α -helices, β -sheets, etc. The electronic circular dichroism (ECD) spectra were simulated at the CAM-B3LYP time-dependent DFT (TD-DFT) level of theory for only the two smaller models, the 3(PPG)₃ and the 3(PPG)₁ systems. The 81 lowest energy singlet transition energies were calculated along with the corresponding dipole strengths and rotational strengths in both the length and velocity formalisms for both 7/2 and CP 3(PPG)₃ structures, and the 27 lowest singlet transition energies were calculated along with the corresponding dipole strengths and rotational strengths in both the length and velocity formalisms for both CP and 7/2 3(PPG)₁ structures.

In fig. 1 the PM6/PCM CP and 7/2 structures for 3(PPG)₆, top and side views are depicted for both systems. As one can see in fig. 1(a), there is a visible open channel in the CP structure surrounded by oxygen atoms. In tables 1 and 2 we present the backbone and side chain torsional angles for all of the residues for the three chains for the CP structures 3(PPG)₃ and 3(PPG)₆. In our analysis, the CP and 7/2 structures have nearly the same energy. The PM6 energy difference between the 7/2 structure and CP structure for collagen is $E(7/2) - E(CP) = -1.739$ eV, or -0.065 eV per residue for 3(PPG)₃, and $E(7/2) - E(CP) = -29.58$ eV, or -0.37 eV per residue for 3(PPG)₆. This higher energy difference for the longer structures reflects the difficulties that the applied methods have in properly optimising long helical structures as the PM6 model is unable to predict the structure of large proteins *de novo* due to local minima [26].

Collagen has also been studied by measurements of circular dichroism (CD) spectra [27–29]. In contrast with the case for β sheets and for α helices the measured spectra of collagen have not yet been satisfactorily reproduced in model calculations [27]. Figure 2 shows the calculated ECD spectrum of the CP and 7/2 motifs along with experimental data. The experimental ECD spectrum is reproduced from ref. [29]. The data shows a characteristic negative peak at around 200 nm and a smaller

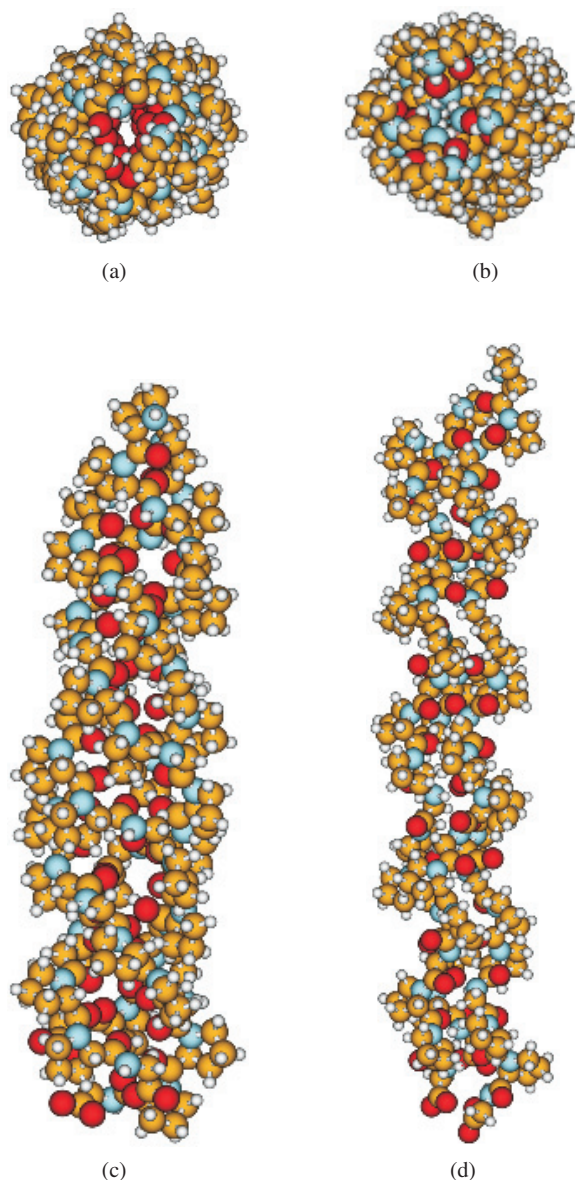


Fig. 1: (Colour on-line) PM6/PCM CP and 7/2 structures for collagen-like peptides 3(PPG)₆. (a) Top view of CP structure; (b) top view of the 7/2 structure, (c) side view of the CP structure, and (d) side view of the 7/2 structure. Colours used for atoms: O (red), N (blue), C (brown) and H (white).

positive peak at 220 nm. For the CAM-B3LYP/KS-TD-DFT/PCM electronic CD (ECD) calculated spectra for the PM6/PCM 3(PPG)₃ structures the semi-empirical PM6 WFT was used for the geometry optimisations and KS-TD-DFT with the CAM-B3LYP exchange correlation (XC) functional was used for the ECD calculations. The EA and ECD spectra were all simulated with FWHM line widths of 15 nm. As one can see by comparison, neither of the two calculations describe the spectra significantly better than the other. The 7/2 model is somewhat better at predicting the positive peak around 220 nm. We have observed that the detailed shape of the calculated

Table 1: PM6/PCM 3[PPG]₃ CP structure: backbone and side chain angles (deg).

Torsion angles	Chain 1	Chain 2	Chain 3
$\phi'(C_\delta NC_\alpha C')$	-62	-60	-63
$\psi_1(NC_\alpha C'N)$	136	141	150
$\omega_1(C_\alpha C'NC_\alpha)$	-176	177	169
$\xi_{1,1}(NC_\alpha C_\beta C_\gamma)$	-7	-14	-15
$\xi_{1,2}(C_\alpha C_\beta C_\gamma C_\delta)$	14	19	22
$\xi_{1,3}(C_\beta C_\gamma C_\delta N)$	-16	-17	-20
$\xi_{1,4}(C_\gamma C_\delta NC_\alpha)$	12	9	11
$\phi_2(C'NC_\alpha C')$	-73	-71	-92
$\psi_2(NC_\alpha C'N)$	148	141	151
$\omega_2(C_\alpha C'NC_\alpha)$	172	179	-176
$\xi_{2,1}(NC_\alpha C_\beta C_\gamma)$	14	14	21
$\xi_{2,2}(C_\alpha C_\beta C_\gamma C_\delta)$	-14	-11	-18
$\xi_{2,3}(C_\beta C_\gamma C_\delta N)$	10	5	8
$\xi_{2,4}(C_\gamma C_\delta NC_\alpha)$	-1	4	6
$\phi_3(C'NC_\alpha C')$	-76	-95	-117
$\psi_{3,1}(NC_\alpha C'O1)$	-59	109	58
$\psi_{3,1}(NC_\alpha C'O2)$	169	179	173

Table 2: PM6/PCM 3[PPG]₆ CP structure: average backbone and side chain angles (deg) for the two middle PPG groups.

Torsion angles	Chain 1	Chain 2	Chain 3
$\phi'(C_\delta NC_\alpha C')$	-21	-23	-27
$\psi_1(NC_\alpha C'N)$	176	180	178
$\omega_1(C_\alpha C'NC_\alpha)$	-176	177	177
$\xi_{1,1}(NC_\alpha C_\beta C_\gamma)$	10	6	12
$\xi_{1,2}(C_\alpha C_\beta C_\gamma C_\delta)$	-32	-28	-33
$\xi_{1,3}(C_\beta C_\gamma C_\delta N)$	41	40	41
$\xi_{1,4}(C_\gamma C_\delta NC_\alpha)$	-36	-38	-35
$\phi_2(C'NC_\alpha C')$	12	7	13
$\psi_2(NC_\alpha C'N)$	146	141	150
$\omega_2(C_\alpha C'NC_\alpha)$	-176	-172	-179
$\xi_{2,1}(NC_\alpha C_\beta C_\gamma)$	-44	-48	-45
$\xi_{2,2}(C_\alpha C_\beta C_\gamma C_\delta)$	27	36	29
$\xi_{2,3}(C_\beta C_\gamma C_\delta N)$	0	-11	-3
$\xi_{2,4}(C_\gamma C_\delta NC_\alpha)$	-27	-18	-25
$\phi_3(C'NC_\alpha C')$	-94	-75	-93
$\psi_{3,1}(NC_\alpha C'O1)$	-130	-124	-126
$\psi_{3,1}(NC_\alpha C'O2)$	-170	-176	-176

spectra is sensitive to the exact values of the refined coordinates and hence we must assume some level of uncertainty simply due to the accuracy of the methods. The integrated intensities of the spectra (ref. [29]), of the CP, and of the 7/2 are -8.4×10^{-2} , -9.2×10^{-2} , and -2.9×10^{-2} degrees·cm³·dmol⁻¹, respectively. Hence, the CP calculation is more in agreement with regards to the negative bias. The reason for this is that the CP structure does not have the right-handed supercoil geometry that reduces the overall left-handed chirality of the 7/2 structure.

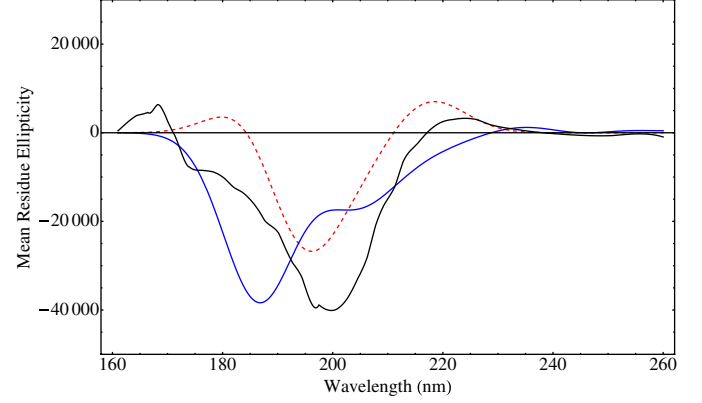


Fig. 2: (Colour on-line) Electronic circular dichroism CAM-B3LYP/PCM for CP (blue curve) and 7/2 (dotted red curve) structures and ECD M06/PCM for CP structure of 3(PPG)₃. The experimental collagen data (black curve) are replotted from ref. [29] after digitising. Mean residue ellipticity (MRE) is measured in units of degrees·cm²·dmol⁻¹.

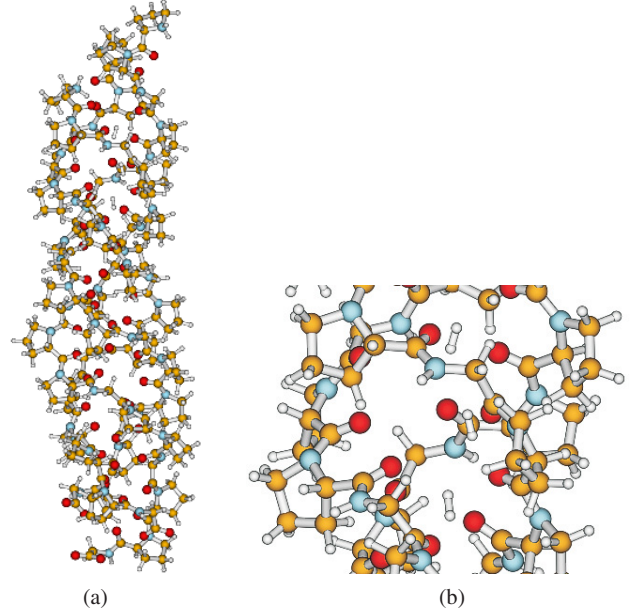


Fig. 3: (Colour on-line) (a) A stick and ball model of the close-packed 3(PPG)₆ structure with four H₂ molecules in the central channel. (b) Expanded view of the part of the structure containing the H₂ molecules. The vertical distance between individual H₂ molecules is about 6 Å.

Contrary to what is the case for the 7/2 structure, the central channel of the CP structure seen in fig. 1(a) opens the possibility for collagen to hold, intercalate, and transport small molecules. In ref. [11], the diameter of the channel was estimated to be about 2 Å. In fig. 3 we depict the minimisation of the 3(PPG)₆ structure with four H₂ molecules in the channel. Eight atomic hydrogens were placed in the channel with an interspacing distance of about 3.5 Å. The purpose was to see if the hydrogens would interact with the oxygens of the collagen strands, or

if they would interact with each other to form molecular hydrogen? The result was that molecular hydrogens were formed. It is a demonstration of the relative stability of the CP structure that it is adaptable to such changes.

In summary, we have reported the result that triple helix structures for collagen that fulfil the close-packing principle are atomically possible and chemically plausible according to the performed minimisation. More investigations are needed to decipher the effects and limitations of the applied minimisation algorithms used. First-principles semi-empirical WFT and KS-DFT calculations of the structures and properties of collagen and compared the close-packed motif (CP) with the supercoiled motif (7/2) have been performed. It is shown that within the method applied both structures behave as stable structures, and both have roughly the same energy. Explicit solvent molecules can change the relative energy of various conformations peptides, and they can stabilise conformers and species that are not stable either in the gas phase or using the simple continuum solvent models like PCM [30]. Complete features of the ECD spectra could not be described by neither of the two calculations although the negative bias is well described by the CP structure. Atomic hydrogens placed in the channel combine to form molecular hydrogen, and may perhaps contribute to enhance the stabilisation of the close-packed structure. The CP structure with its central cavity appears to be a possible hydrogen molecule channel protein as atomic hydrogen can form stable molecular hydrogen in this channel. It is an intriguing possibility that such H₂ molecules will have a biological significance and function.

This work is supported by the Villum Foundation. KJJ wishes to thank DKFZ and DTU for hospitality.

REFERENCES

- [1] RAMACHANDRAN G. N. and KARTHA G., *Nature*, **174** (1954) 269.
- [2] RAMACHANDRAN G. N. and KARTHA G., *Nature*, **176** (1955) 593.
- [3] RICH A. and CRICK F. H. C., *Nature*, **176** (1955) 915.
- [4] RICH A. and CRICK F. H. C., *J. Mol. Biol.*, **3** (1961) 483.
- [5] OKUYAMA K., TANAKA N., ASHIDA T., KAKUDO M., SAKAKIBARA S. and KISHIDA Y., *J. Mol. Biol.*, **72** (1972) 571.
- [6] OKUYAMA K., TAKAYANAGI M., ASHIDA T. and KAKUDO M., *Polym. J.*, **9** (1977) 341.
- [7] BECK K. and BRODSKY B., *J. Struct. Biol.*, **122** (1988) 17.
- [8] BELLA J., EATON M., BRODSKY B. and BERMAN H. M., *Science*, **266** (1994) 75.
- [9] HONGO C., NOGUCHI K., OKUYAMA K., TANAKA Y. and NISHINO N., *J. Biochem.*, **138** (2005) 135.
- [10] OKUYAMA K., NAGARAJAN V. and KAMITORI S., *Proc. Indian Acad. Sci. (Chem. Sci.)*, **111** (1999) 19.
- [11] BOHR J. and OLSEN K., *Theor. Chem. Acc.*, **130** (2011) 1095.
- [12] OLSEN K. and BOHR J., *New J. Phys.*, **14** (2012) 023063.
- [13] BERISIO R., VITAGLIANO L., MAZZARELLA L. and ZAGARI A., *Protein Sci.*, **11** (2002) 262.
- [14] OKUYAMA K., HONGO C., FUKUSHIMA R., WU G., NARITA H., NOGUCHI K., TANAKA Y. and NISHINO N., *Biopolymers*, **76** (2004) 367.
- [15] OKUYAMA K., XU X., IGUCHI M. and NOGUCHI K., *Biopolymers*, **84** (2006) 181.
- [16] OKUYAMA K., WU G., JIRAVANICHANUN N., HONGO C. and NOGUCHI K., *Biopolymers*, **84** (2006) 421.
- [17] TSAI M., XU Y. and DANNENBERG J. J., *J. Am. Chem. Soc.*, **127** (2005) 14130.
- [18] PÁLFI V. K. and PERCZEL A., *J. Comput. Chem.*, **29** (2008) 1374.
- [19] ALMORA-BARRIOS N. and DE LEEUW N. H., *Langmuir*, **26** (2010) 14535.
- [20] FRISCH M. J. *et al.*, Gaussian 09, Gaussian, Inc., Wallingford, CT (2009).
- [21] BARONE V., COSSI M. and TOMASI J., *J. Comput. Chem.*, **19** (1998) 404.
- [22] TOMASI J., MENNUCCI B. and CAMMI R., *Chem. Rev.*, **105** (2005) 2999.
- [23] STEWART J. J. P., *J. Mol. Model.*, **13** (2007) 1173.
- [24] YANAI T., TEW D. and HANDY N., *Chem. Phys. Lett.*, **393** (2004) 51.
- [25] SHCHERBIN D. and RUUD K., *Chem. Phys.*, **349** (2008) 234.
- [26] STEWART J. J. P., *J. Mol. Model.*, **15** (2009) 765.
- [27] BHATNAGAR R. S. and GOUGH C. A., in *Circular Dichroism and the Conformational Analysis of Biomolecules*, edited by FASMAN G. D. (Plenum Press, New York) 1996, pp. 184–199, ISBN 0-306-45142.
- [28] WALLACE B. A. and JANES R. W., *Curr. Opin. Chem. Biol.*, **5** (2001) 567.
- [29] MILES A. J. and WALLACE B. A., *Chem. Soc. Rev.*, **35** (2006) 39.
- [30] JALKANEN K. J., DEGTYARENKO I. M., NIEMINEN R. M., CAO X., NAFIE L. A., ZHU F. and BARRON L. D., *Theor. Chem. Acc.*, **119** (2008) 191.