Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Two crystal modifications of (Pro-Pro-Gly)$_4$-Hyp-Hyp-Gly-(Pro-Pro-Gly)$_4$ reveal the puckering preference of Hyp(X) in the Hyp(X):Hyp(Y) and Hyp(X):Pro(Y) stacking pairs in collagen helices

Kenji Okuyama,[a]* Tatsuya Morimoto,[a] Hirotaka Narita,[a] Tatsuya Kawaguchi,[a] Kazunori Mizuno,[b] Hans Peter Bächinger,[b,c] Guanghan Wu[d] and Keiichi Noguchi[d]

[a]Department of Macromolecular Science, Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043, Japan, [b]Shriners Hospital for Children, Research Department, Portland, Oregon 97239, USA, [c]Department of Biochemistry and Molecular Biology, Oregon Health and Science University, Portland, Oregon 97239, USA, and [d]Graduate School of Technology, Tokyo University of Agriculture and Technology, Koganei, Tokyo 184-8588, Japan

Correspondence e-mail: okuyamak@chem.sci.osaka-u.ac.jp

Received 2 October 2009
Accepted 5 November 2009

PDB References:
ppg9-OOG_T, 3a08;
ppg9-OOG_H, 3a19.

Two crystal modifications of a collagen model peptide, (Pro-Pro-Gly)$_4$-Hyp-Hyp-Gly-(Pro-Pro-Gly)$_4$ [where Hyp is (4*R*,2*S*)-L-hydroxyproline], showed very similar unit-cell parameters and belonged to the same space group $P2_1$. Both crystals exhibited pseudo-merohedral twinning. The main difference was in their molecular-packing arrangements. One modification showed pseudo-hexagonal packing, while the other showed pseudo-tetragonal packing. Despite their different packing arrangements, no significant differences were observed in the hydration states of these modifications. The peptide in the pseudo-tetragonal crystal showed a cyclic fluctuation of helical twists with a period of 20 Å, while that in the pseudo-hexagonal crystal did not. In these modifications, the puckering conformations of four of the 12 Hyp residues at the X position of the Hyp(X)-Hyp(Y)-Gly sequence were in the opposite conformations to the previous hypothesis that Hyp(X) residues involved in Hyp(X):Hyp(Y) and Hyp(X):Pro(Y) stacking pairs prefer up-puckering and down-puckering conformations, respectively. Detailed investigation of the molecular interactions between Hyp(X) and adjacent molecules revealed that these opposite conformations appeared because the puckering conformation, which follows the hypothesis, is subject to steric hindrance from the adjacent molecule.

1. Introduction

Collagen is the most abundant protein in animals. In both vertebrates and invertebrates collagens share the same characteristic repeating Gly-Xaa-Yaa tripeptide units which are essential for formation of the collagen triple helix. Because of this strict sequence constraint, together with their high content of imino acids (~20%), collagens adopt a structurally unique triple-helical conformation consisting of three identical or different α-chains (Ramachandran & Kartha, 1955). Two helical symmetries, 7/2-helix and 10/3-helix, have been proposed as an average triple-helical conformation and these have been accepted for more than 30 years (Okuyama *et al.*, 1977; Rich & Crick, 1955). Nonetheless, even a recent fibre-diffraction study was unable to reach a definitive conclusion because of the scarcity of diffraction data from native collagen (Okuyama, Xu *et al.*, 2006). To date, the average helical symmetry of native collagen has not been revealed from its X-ray diffraction data alone. In contrast, single-crystal structure analyses of collagen model peptides revealed a tangible conclusion. The overall molecular structures of most the peptide crystals were very close to the ideal 7/2-helical model (Okuyama, Wu *et al.*, 2006), while no such evidence was

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the last shell.

Peptide	ppg9-OOG_H	ppg9-OOG_T	
		This work	Previous work
Data collection			
Facility	PF BL6A	PF BL6A	
Data-collection device	ADSC Quantum 4	ADSC Quantum 4	
Data-collection temperature (K)	95	95	
Wavelength (Å)	0.978	0.978	
Resolution (Å)	50.0–1.55 (1.61–1.54)	26.6–1.22 (1.26–1.22)	
No. of unique reflections	14850	32577	
Completeness (%)	99.5 (99.9)	98.7 (99.8)	
R_{merge}	0.08 (0.35)	0.08 (0.29)	
Redundancy	7.0 (7.2)	3.5 (3.8)	
$\langle I/\sigma(I)\rangle$	6.4	4.5	
Space group	$P2_1$	$P2_1$	
Unit-cell parameters			
a (Å)	23.65	25.99	
b (Å)	26.93	26.67	
c (Å)	79.66	79.84	
β (°)	89.93	90.03	
Structure refinement			
Resolution range (Å)	20.0–1.55	8.0–1.22	
Data cutoff for refinement ($\sigma\|F\|$)	1.0	1.0	
No. of reflections for refinement	14097	30417	
No. of reflections for R_{free}	736	1627	
Twin fraction (initial/last)	0.39/0.42	0.47/0.49	
Twin law	$h, -k, -l$	$h, -k, -l$	
R	0.166	0.139	0.193
R_{free}	0.204	0.216	0.255
No. of refined parameters	4666	9886	10425
Max./min. peaks in D maps (e Å^{-3})	0.31/−0.30	0.35/−0.25	0.41/−0.36
No. of peptide non-H atoms	928	859	895
No. of water sites	237	239	263
R.m.s.d. from standard geometry			
Bonds (Å)	0.008	0.012	0.010
Angles (Å)	0.023	0.030	0.024
Planes (Å^3)	0.027	0.031	0.028
Atomic displacement parameters (Å^2)			
Peptide atoms	16.0	16.5	16.6
Water O atoms	22.7	25.8	26.4
PDB entry	3a19	3a08	2d3h

observed for the 10/3-helical model. Only peptide segments in which imino-acid residues were absent from more than three consecutive triplets were found to have relaxed conformations that were close to the 10/3-helix (Boudko *et al.*, 2008; Kramer *et al.*, 2001). However, these relaxed conformations only existed when there were stable 7/2-helical regions in the same peptide molecule. Details of structural studies of the collagen helix have recently been summarized (Okuyama, 2008).

The collagen helix is stabilized by hydrogen bonds between the N—H of a Gly residue and the C═O of the Xaa residue in the adjacent strand. It is also stabilized by hydroxyproline residues. The helix–coil transition temperature (T_m) of collagen depends on the hydroxyproline content (Burjanadze, 1979). In vertebrate collagen, (4*R*,2*S*)-L-hydroxyproline (Hyp) residues only occupy the *Y* position. Thus, the T_m of (Pro-Hyp-Gly)$_{10}$ is higher than that of (Pro-Pro-Gly)$_{10}$ by 30 K, while (Hyp-Pro-Gly)$_{10}$ cannot form a triple-helical conformation under the same aqueous conditions (Inouye *et al.*, 1982). Therefore, for the last two decades it has been believed that Hyp in the *Y* position stabilizes the triple-helical conformation, whereas Hyp in the *X* position destabilizes it. In order to understand the position-sensitive stabilization and destabilization induced by Hyp, the propensity-based hypothesis was proposed (Vitagliano *et al.*, 2001). This hypothesis is based on the intrinsic preference of Hyp for the up-puckering conformation and the positional preference that imino acids located in the *X* and *Y* positions in the collagen helix generally prefer down-puckering and up-puckering conformations, respectively. This hypothesis nicely explains the stabilizing and destabilizing effects of Hyp in the *Y* and *X* positions, respectively. However, after the proposal of this hypothesis it emerged that peptides with Hyp-Hyp-Gly (Berisio *et al.*, 2004; Mizuno *et al.*, 2004; Persikov *et al.*, 2003) or Hyp-Thr-Gly (Bann & Bächinger, 2000) sequences stabilize the triple-helical structure. Specifically, the T_m of Ac-(Gly-Hyp-Hyp)$_{10}$-NH$_2$ is slightly higher than that of Ac-(Gly-Pro-Hyp)$_{10}$-NH$_2$ (Mizuno *et al.*, 2004).

Based on the results of molecular-mechanics calculations (Bhatnagar *et al.*, 1988), it has been suggested that contacts within the Pro(*X*):Pro(*Y*) stacking pair contribute to the stability of the triple-helical conformation. In a study of (Pro-Pro-Gly)$_9$, the importance of van der Waals interactions between Pro(*X*) and Pro(*Y*) in the adjacent strand to the stabilization of the collagen helix was pointed out and referred to as the 'Pro:Pro van der Waals interaction' (Hongo *et al.*, 2005). Intramolecular Pro(*X*):Pro(*Y*) stacking interactions are most effective when Pro(*X*) and Pro(*Y*) adopt down-puckering and up-puckering conformations, respectively, which agrees with the propensity-based hypothesis. A previous paper (Okuyama *et al.*, 2009) reported the puckering propensity of Hyp in the *X* position in the host–guest peptides (Pro-Pro-Gly)$_4$-Xaa-Yaa-Gly-(Pro-Pro-Gly)$_4$ (hereafter referred to as ppg9-*XY*G), where (Xaa, Yaa) = (Pro, Hyp), (Hyp, Pro), (Pro, alloHyp) or (Hyp, Hyp) and alloHyp denotes (4*S*,2*S*)-L-hydroxyproline. The result, in summary, was that Hyp(*X*) residues involved in Hyp(*X*):Pro(*Y*) stacking pairs prefer the down-puckering conformation, while Hyp(*X*) residues involved in Hyp(*X*):Hyp(*Y*) stacking pairs prefer the up-puckering conformation.

In this study, we report the crystal structures of two modifications of the host–guest peptide (Pro-Pro-Gly)$_4$-Hyp-Hyp-Gly-(Pro-Pro-Gly)$_4$ (ppg9-OOG). One modification adopted a pseudo-hexagonally packed structure (ppg9-OOG_H), while the other was pseudo-tetragonally packed (ppg9-OOG_T). Although the structure of the latter has been reported in a previous paper (Okuyama *et al.*, 2009), we recently noticed

that these crystals were twinned. Therefore, we refined the structure of ppg9-OOG_T under twinning conditions, which resulted in a considerable decrease in the R and $R_{\rm free}$ values. The detailed structures of two modifications of ppg9-OOG provided further confirmation of our previous conclusion concerning the conformational preference of Hyp(X) in Hyp(X):Hyp(Y) and Hyp(X):Pro(Y) stacking pairs.

2. Materials and methods

2.1. Crystallization and data collection

Experimental details of the synthesis, purification and characterization of the ppg9-OOG peptide have been described in a previous paper (Okuyama *et al.*, 2009). Crystallization conditions for ppg9-OOG_T were also described in the same paper. Those for ppg9-OOG_H were exactly the same as those for ppg9-OOG_T. In fact, single crystals of both modifications were found in the same drop. Although we obtained a reasonable crystal structure for ppg9-OOG_T (Okuyama *et al.*, 2009), we could not decrease the R and $R_{\rm free}$ factors of the ppg9-OOG_H crystal until we introduced the twin fraction and twin law into the refinement calculations.

Both diffraction data sets were obtained at 95 K on BL6A at the Photon Factory. For the analysis of ppg9-OOG_T, we used the previously reported diffraction data set (Okuyama *et al.*, 2009). The diffraction images of the ppg9-OOG_H crystal were processed using *HKL*-2000 (Otwinowski & Minor, 1997). Both crystals exhibited pseudo-merohedral twinning, with twin fractions of 0.389 for ppg9-OOG_H and 0.467 for ppg9-OOG_T estimated from the H plot by *phenix.xtriage* (Adams *et al.*, 2002). The twin law was shown to be (h, $-k$, $-l$) in both cases. X-ray data-collection statistics are summarized in Table 1.

2.2. Structure determination and refinement

We previously reported the crystal structure of ppg9-OOG_T, with R and $R_{\rm free}$ values of 0.193 and 0.255, respectively, at 1.22 Å resolution (Okuyama *et al.*, 2009), in which the molecular-replacement and refinement calculations were carried out using the twinned data (Table 1). This peptide structure was used as a starting model in the present study, together with the same twinned reflection data set. The structure was refined by *SHELXL* (Sheldrick, 2008) using the twin fraction given above under the twinning operator (h, $-k$, $-l$), resulting in a substantial improvement in R and $R_{\rm free}$ (to 0.139 and 0.216, respectively).

The initial phase of the ppg9-OOG_H peptide was obtained by *Phaser* (McCoy *et al.*, 2007) using the triple-helical structure of (Pro-Pro-Gly)$_9$ (PDB code 2cuo; Hongo *et al.*, 2005) as a probe and the detwinned data set. Next, the obtained structure was refined against the twinned data using *REFMAC*5 (Murshudov *et al.*, 1997) and *SHELXL* under the twinning operator (h, $-k$, $-l$).

The refinement statistics are listed in Table 1. During structural analyses of both crystals, *Coot* (Emsley & Cowtan, 2004) was used for visualization of the structures and figures were generated using *MolFeat* (FiatLux Co., Tokyo). Coordinates and structure factors have been deposited in the Protein Data Bank under accession codes 3a08 (ppg9-OOG_T) and 3a19 (ppg9-OOG_H).

3. Results and discussion

3.1. Pseudo-merohedrally twinned crystals of ppg9-*XYG*

A twinned crystal, which is often apparently one crystal, is obtained when two (or more) crystals intergrow in different relative orientations. In some cases the presence of two diffraction patterns can be recognized from the beginning since they are not precisely superposed. In other cases (merohedral twinning) the twin domains are superimposable in three dimensions and thus are not evidenced in the diffraction pattern (Yeates, 1997). Although merohedral twinning can only occur in high-symmetry space groups in the tetragonal, trigonal, hexagonal and cubic systems, pseudo-merohedral twinning is possible in lower space groups in the case of fortuitous unit-cell geometry. For example, when $\beta = 90^\circ$ in a monoclinic space group pseudo-merohedral twinning can emulate an orthorhombic space group, as observed in this study.

The apparent space group of (Pro-Pro-Gly)$_9$ (ppg9-PPG) crystal is $P2_12_12_1$, but the actual space group is $P2_1$. The reason that the crystal belongs to this space group, as opposed to the space group $P2_12_12_1$ of (Pro-Pro-Gly)$_{10}$ (Berisio *et al.*, 2002) and the subcell structures of (Pro-Pro-Gly)$_n$, where n = 9, 10 (Hongo *et al.*, 2001; Vitagliano *et al.*, 2001), has been discussed in detail by Hongo *et al.* (2005). In practice, structural analysis of ppg9-PPG in space group $P2_12_12_1$ resulted in high R and $R_{\rm free}$ values, while analysis in space group $P2_1$ gave reasonable R and $R_{\rm free}$ values. In the same way, the space groups of most ppg9-*XYG* peptide crystals, including the two modifications examined in this study, were found to be $P2_1$ despite the apparent $P2_12_12_1$ symmetry. Since their apparent space group was $P2_12_12_1$, it is quite understandable that these modifications possessed pseudo-2_1 symmetry along the a and c axes. As we refined the ppg9-*XYG* peptides in the correct space group, we obtained essentially correct structures with slightly higher R and $R_{\rm free}$ values. For example, the R and $R_{\rm free}$ values of ppg9-OOG_T were 0.193 and 0.255, respectively, in the previous analysis (Okuyama *et al.*, 2009). Therefore, we did not suspect the possibility of crystal twinning of the ppg9-*XYG* peptide crystals until recently. However, in the recent structural analysis of the (Pro-Pro-Gly)$_4$-Hyp-Val-Gly-(Pro-Pro-Gly)$_4$ peptide (PDB code 3a0m; K. Okuyama, T. Morimoto, K. Mizuno & H. P. Bächinger, unpublished work) we found the crystal to be twinned. This finding triggered many other findings of twinned ppg9-*XYG* crystals. We have now found that most ppg9-*XYG* crystals, including the present two crystal modifications, are twinned. The final refined twinning fractions were 0.42 for ppg9-OOG_H and 0.49 for ppg9-OOG_T.

3.2. Peptide main-chain conformation

Similar to the (Pro-Pro-Gly)$_9$ (ppg9-PPG) crystal (Hongo *et al.*, 2005), the asymmetric units of both the ppg9-OOG_H and

the ppg9-OOG_T crystals contained two molecules: *ABC* and *DEF* molecules. Here, *A*, *B*, *C*, *D*, *E* and *F* denote peptide strands with the same amino-acid sequence. The four independent molecular conformations in the two crystal modifications were very similar to each other and essentially the same as that of ppg9-PPG, which adopted a triple-helical structure very close to the ideal 7/2-helical model for collagen (Okuyama *et al.*, 1977; Okuyama, Xu *et al.*, 2006). The root-mean-square deviations (r.m.s.d.s) between two independent molecules were calculated using the main-chain atoms (N, C^{α}, C) of 57 residues from Pro5 to Pro23. The r.m.s.d.s for two molecules in the same asymmetric unit were 0.28 for ppg9-OOG_H and 0.16 Å^2 for ppg9-OOG_T. The r.m.s.d.s. for two molecules in different crystal modifications were less than 0.37 Å^2. These peptide conformations provided inter-chain hydrogen bonds between N—H of Gly in one chain and the carbonyl O atom in the Xaa residue of the adjacent chain. Average hydrogen-bond lengths (the distance between the N and O atoms) were 2.96 Å in the ppg9-OOG_H peptide structure and 2.94 Å in the ppg9-OOG_T peptide structure, which were close to that found in ppg9-PPG (2.93 Å). The presence of rather long hydrogen bonds between N—H and C═O is one of the features of collagen helices. A superposed stereoscopic view of the *ABC* molecules in the ppg9-OOG_H and ppg9-OOG_T crystals is shown in Fig. 1.

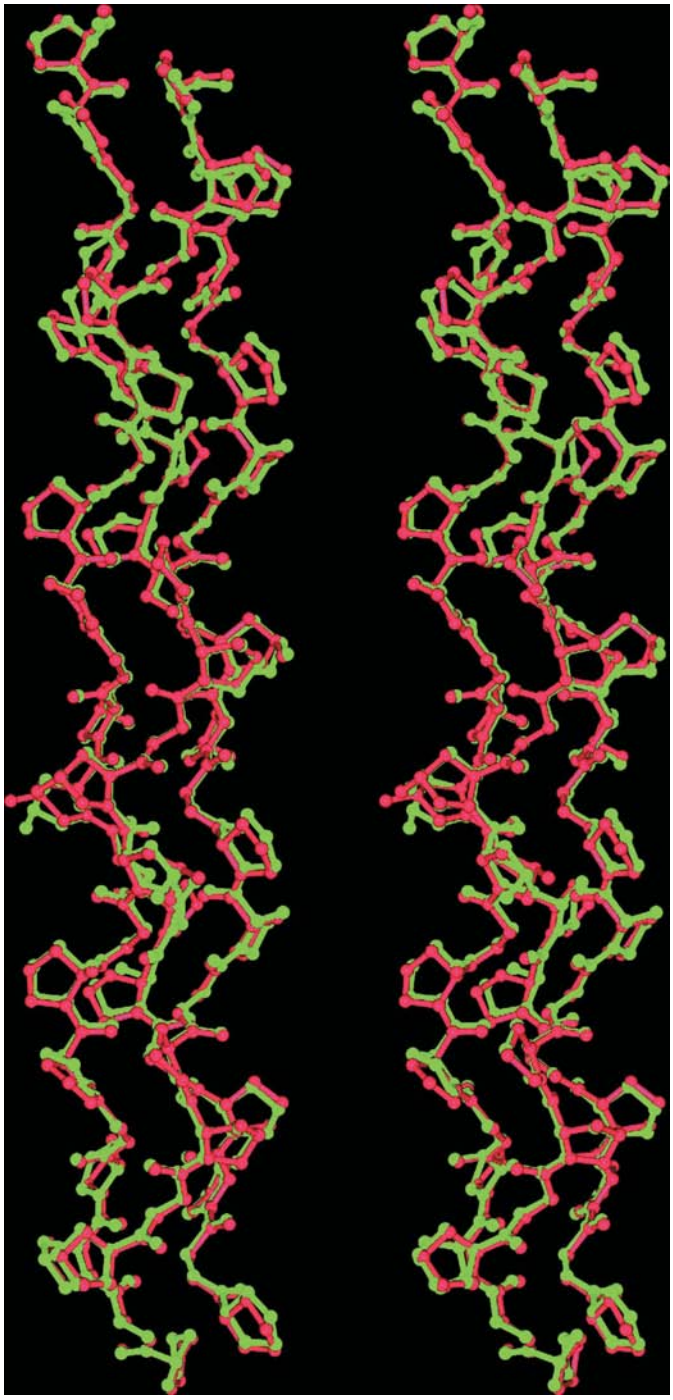

Figure 1
Stereoscopic superposed *ABC* molecular conformations of ppg9-OOG_H (red) and ppg9-OOG_T (green). Only the central 57 amino-acid residues from Pro5 to Pro23 of each strand are included in the figure. The r.m.s.d. of these two molecules was 0.32 Å^2.

3.3. Differences from the previously reported ppg9-OOG_T structure

After the publication of the crystal structure of ppg9-OOG_T (Okuyama *et al.*, 2009), we found that many of the ppg9-*XY*G crystals, including the ppg9-OOG_T and ppg9-OOG_H crystals, were pseudo-merohedrally twinned. In this study, the structure of ppg9-OOG_T was refined using the same intensity data set as in the previous analysis, but this time using the twin operator (h, $-k$, $-l$). Although we observed significant decreases in the R and $R_{\rm free}$ values, no significant improvement of the electron-density maps and no significant changes in the structure were observed. The r.m.s.d. between two independent molecules (*ABC* and *DEF* molecules) was 0.12 Å^2 in the previous analysis and was 0.16 Å^2 in this analysis, while the r.m.s.d. between the *ABC* molecules in the previous and the present structures was 0.07 Å^2 and that between the *DEF* molecules was also 0.07 Å^2; this demonstrates the high similarity of these two molecular structures. In this study, the puckering conformations of the Hyp residues in the *X* and *Y* positions are very important. All the puckering conformations observed in this study were essentially the same as previously reported conformations (Okuyama *et al.*, 2009). This is also true for the location of water molecules. In the previous analysis a total of 263 water molecules were found, whereas 239 water molecules could be located in the present analysis. The positional differences for 224 water molecules were less than 0.5 Å between these structures.

3.4. Distribution of helical twists

The helical twists of each triplet in the six strands of the *ABC* and *DEF* molecules were calculated using the in-house program *PHEL* (Okuyama, Wu *et al.*, 2006; Sugeta & Miyazawa, 1967). The distribution of the helical twists of ppg9-PPG showed a clear cyclic fluctuation with a 20 Å repeating period (Okuyama, 2008). The helical twists of the *ABC* and *DEF* molecules of the ppg9-OOG_T peptide showed a similar cyclic fluctuation with the same period (Fig. 2*a*), which was some-

what clearer than that in the previous structure (Okuyama *et al.*, 2009). However, these molecules were somewhat disordered compared with ppg9-PPG. For example, the *E* strand had no local minimum around residue 412. On the other hand, cyclic fluctuations were not observed in either molecule of the ppg9-OOG_H peptide (Fig. 2*b*), although the deviations of the helical twist from average values were in a similar range to those of ppg9-OOG_T (35–65°). The average helical twists of the *ABC* and *DEF* molecules in the ppg9-OOG_T crystal were 51.4 and 51.6°, respectively, while those in the ppg9-OOG_H crystal were 51.9 and 52.1°, respectively. Although the imino-acid content of the ppg9-OOG peptide was very high (67%), the values of the helical twist of each triplet were scattered over a wide range (35–65°). There are some reports of the sequence dependence of helical twist in which the local conformation of the imino-acid-rich region prefers 7/2-helical symmetry, while that of imino-acid-poor region prefers a relaxed conformation and is accordingly closer to 10/3-helical symmetry (Boudko *et al.*, 2008; Kramer *et al.*, 2001). In general, this is true as an average. However, even the Pro-Pro-Gly and Hyp-Hyp-Gly sequences could adopt helical twists close to that of the ideal 10/3-helix (36°; Fig. 2). The distribution of helical twists in the two crystal modifications of the ppg9-OOG peptide indicated the difficulty in estimating the local conformation from its short amino-acid sequence alone.

3.5. Proline-ring puckering

Understanding the mechanism of the stabilization and destabilization of a collagen helix induced by Hyp residues is one of important issues in structural studies of collagen since collagens are the only animal proteins with a high Hyp content. Recent explanations of this mechanism have been concerned with the puckering conformation of the proline ring (Fig. 3; Improta *et al.*, 2008; Vitagliano *et al.*, 2001). In a previous paper (Okuyama *et al.*, 2009), we proposed a hypothesis on the puckering propensities of the hydroxyproline in the *X* position based on model peptide structures. In this hypothesis, Hyp(*X*) residues involved in Hyp(*X*):Pro(*Y*) and Hyp(*X*):Hyp(*Y*) stacking pairs prefer the down-puckering and up-puckering conformations, respectively. The ppg9-OOG peptide contains both Hyp(*X*):Pro(*Y*) and Hyp(*X*):Hyp(*Y*) stacking pairs in the central parts of the triple helix. Since these stacking pairs were in different surroundings in the

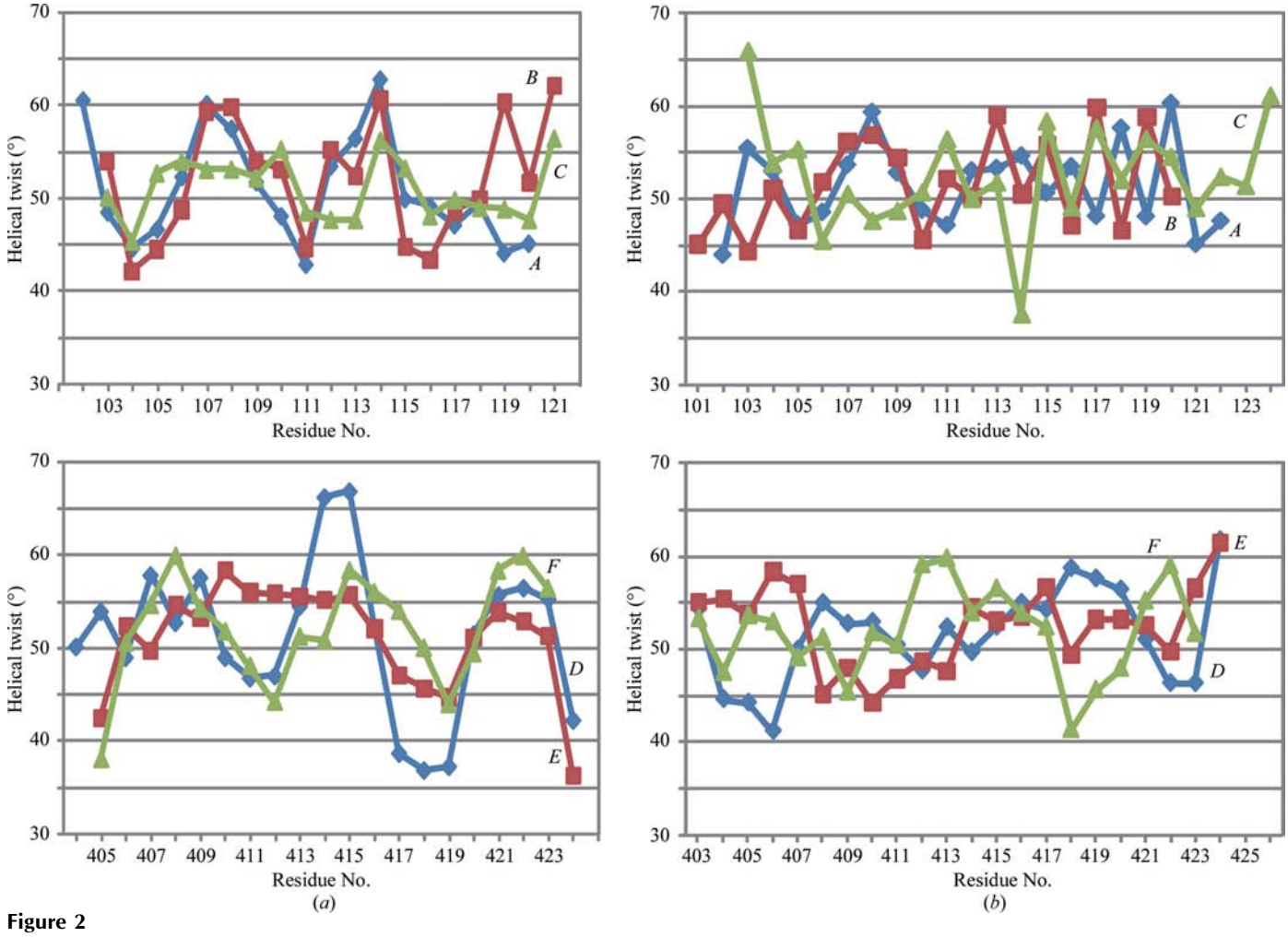

Figure 2
Helical twists of each triplet in (*a*) ppg9-OOG_T and (*b*) ppg9-OOG_H. The helical twists of ppg9-OOG_T showed a cyclic fluctuation with a seven-triplet period (20 Å), while those of ppg9-OOG_H did not.

Table 2
χ_1 (N—C^α—C^β—C^γ) values of Hyp and Pro residues in the Xaa:Yaa stacking pairs.

* denotes a symmetry-related residue.

Xaa	χ_1 (°)	Yaa	χ_1 (°)	Contact between Xaa and adjacent molecules
ppg9-OOG_H				
Hyp13*A*	−26.4	Hyp14*C*	−33.1	Hyp13*A* $O^\delta\cdots$Hyp14*E** C^δ, 3.22 Å
Hyp13*B*	−10.7	Hyp14*A*	−23.5	Hyp13*B* $O^\delta\cdots$Hyp14*C** C^δ, 3.59 Å†
Hyp13*C*	−11.6	Pro11*B*	−29.2	Hyp13*C* $O^\delta\cdots$Hyp13*B** C^γ, 3.47 Å†
Hyp13*D*	−25.2	Hyp14*F*	−22.0	Hyp13*D* $O^\delta\cdots$Pro20*F** C^δ, 3.53 Å
Hyp13*E*	−12.9	Hyp14*D*	−22.1	No contact less than 4 Å
Hyp13*F*	−21.8	Pro11*E*	28.2	Hyp13*F* $O^\delta\cdots$Pro17*C** C^δ, 3.23 Å†
Pro16*C*	22.7	Hyp14*B*	−26.3	
Pro16*F*	24.4	Hyp14*E*	−25.0	
ppg9-OOG_T				
Hyp13*A*	28.4	Hyp14*C*	−20.3	Hyp13*A* $O^\delta\cdots$Hyp14*E** C^δ, 3.29 Å‡
Hyp13*B*	−14.6	Hyp14*A*	−25.1	No contact less than 5 Å
Hyp13*C*	25.8	Pro11*B*	−26.2	Hyp13*C* $O^\delta\cdots$Pro16*F* C^β, 3.34 Å
Hyp13*D*	22.4	Hyp14*F*	−18.0	Hyp13*D* $O^\delta\cdots$Hyp14*B** C^δ, 3.44 Å‡
Hyp13*E*	−21.4	Hyp14*D*	−25.2	No contact less than 5 Å
Hyp13*F*	24.5	Pro11*E*	−25.5	Hyp13*F* $O^\delta\cdots$Pro7*B* C^β, 3.62 Å‡
Pro16*C*	27.0	Hyp14*B*	−21.8	
Pro16*F*	26.4	Hyp14*E*	−22.9	

† The distance would become shorter if the Xaa residue adopted the down-puckering conformation. ‡ The distance would become shorter if the Xaa residue adopted the up-puckering conformation.

ppg9-OOG_H and ppg9-OOG_T crystals, we examined the puckering conformations of Hyp and its interactions with the adjacent molecule in order to verify the above hypothesis.

The conformational angles χ_1 (N—C^α—C^β—C^γ) of all the stacking pairs with Hyp residues in the *X* and/or *Y* positions are listed in Table 2 together with information on the contacts between Xaa and adjacent molecules. All six hydroxyproline residues in the *Y* position of the Xaa-Hyp-Gly sequence in both crystal modifications adopted the up-puckering conformation (Table 2) as observed in other collagen-like peptides (Kawahara *et al.*, 2005; Okuyama *et al.*, 2004; Schumacher *et al.*, 2005). To date, no down-puckering conformation of a Hyp(*Y*) residue has been reported. The reason for this may be attributed to the agreement between the intrinsic residue preference of Hyp and the positional preference of the *Y* position (Vitagliano *et al.*, 2001). In contrast, four of the six hydroxyproline residues in the *X* position of ppg9-OOG_T adopted down-puckering conformations, against the residue preference, while all six Hyp residues in the *X* position of ppg9-OOG_H adopted up-puckering conformations (Table 2). Since two of the Hyp residues of the former peptide, Hyp13*C* and Hyp13*F*, were involved in Hyp(*X*):Pro(*Y*) stacking pairs, their down-puckering conformations agreed with the hypothesis (Okuyama *et al.*, 2009). The other two residues, Hyp13*A* and Hyp13*D*, were involved in Hyp(*X*):Hyp(*Y*) stacking pairs and their down-puckering conformations were in conflict with the hypothesis. In these cases, if Hyp(*X*) adopted the up-puckering conformation the distances between the O^δ atom of Hyp(*X*) and some of the atoms of Hyp in the adjacent molecule would become smaller than the corresponding van der Waals distances. For example, as shown in Fig. 4, if the down-puckering conformation of Hyp13*A* (or Hyp13*D*) were replaced by the up-puckering conformation, the distance between the O^δ atom of Hyp13*A* (Hyp13*D*) and the C^γ atom of Hyp14*E** (Hyp14*B**), where * denotes that the residue was generated by crystallographic symmetry, would become 2.37 Å (2.43 Å). Therefore, because of the intermolecular interaction, Hyp(*X*) is forced to adopt a down-puckering conformation in these cases. Similarly, three conflicting down-puckering conformations of Hyp(*X*) residues in the (Gly-Hyp-Hyp)$_9$ structure (Schumacher *et al.*, 2005) were explained by steric hindrance between Hyp(*X*) and adjacent molecules. In contrast, in the (Hyp-Hyp-Gly)$_{10}$ structure (Kawahara *et al.*, 2005) three Hyp(*X*) residues adopted down-puckering conformations to form hydrogen bonds to hydroxyl groups of Hyp(*X*) in adjacent molecules. However, in the solution state there is no such interaction between triple helices. Therefore, these Hyp(*X*) residues should adopt down-puckering conformations according to the hypothesis.

Since all the Hyp residues in the *X* position of the ppg9-OOG_H peptide adopted the up-puckering conformation, Hyp13*C* and Hyp13*F* involved in Hyp(*X*):Pro(*Y*) stacking

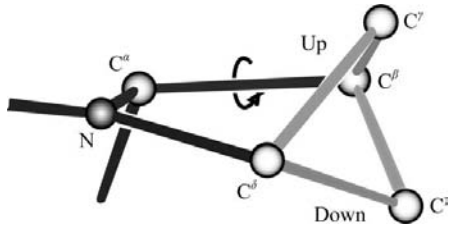

Figure 3
Up-puckering and down-puckering conformations of the proline ring. Up and down conformations are defined by negative and positive values of the dihedral angle χ_1 (N—C^α—C^β—C^γ), respectively.

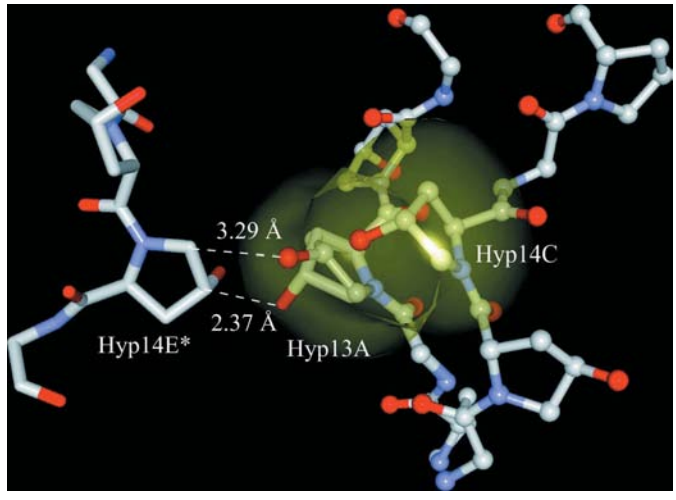

Figure 4
Intermolecular contacts between Hyp13*A* and Hyp14*E**, showing the shortest distance (3.29 Å) between Hyp13*A* O^δ and Hyp14*E** C^δ. The *A* and *C* strands are shown as ball-and-stick models, while the *E* strand and Hyp13*A* with the up-puckering conformation are shown as stick models. Hyp13*A* with the up-puckering conformation was generated by superimposing the main-chain atoms of Gly12*E*-Hyp13*E* (up) of ppg9-OOG_T on those of Gly12*A*-Hyp13*A* (down). In this case, the shortest distance would be 2.37 Å between Hyp13*A* O^δ (up) and Hyp14*E** C^γ. The van der Waals surface of the Hyp13*A*:Hyp14*C* stacking pair is shown in yellow. Hyp14*E** denotes the residue generated from Hyp14*E* by the symmetry operation.

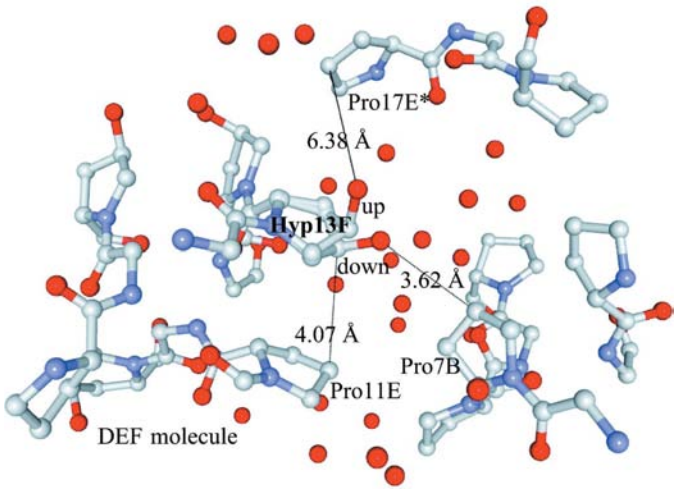

Figure 5
Intermolecular contacts between Hyp13*F* and surrounding molecules in the ppg9-OOG_T crystal. For both the up-puckering and down-puckering conformations of Hyp13*F* no short contacts were observed with the surrounding molecules. The up-puckering conformation was generated by superimposing the main-chain atoms of the Gly12*E*-Hyp13*E* (up) dipeptide onto the Gly12*F*-Hyp13*F* (down) dipeptide. The distance (4.07 Å) between the C^{γ} atoms of Hyp13*F* (*X*) and its stacking mate, Pro11*E* (*Y*), is also shown. Pro17*E** denotes the residue generated from Pro17*E* by the symmetry operation.

pairs were in conflict with the hypothesis. The reason for this was also attributed to the intermolecular interactions between Hyp(*X*) and adjacent molecules. In the case of Hyp13*C*, the distance between the O^{δ} atom of this residue and Hyp13*B** C^{γ} was 3.47 Å (Table 2). This distance would be 2.50 Å if Hyp13*C* adopted the down-puckering conformation. Similarly, the shortest distance would be 3.0 Å if Hyp13*F* adopted the down-puckering conformation. These distances were shorter than the van der Waals distance between the corresponding atoms. Therefore, two conflicting up-puckering conformations in the ppg9-OOG_H peptide were explained by steric hindrance with adjacent molecules.

The conflicting puckering conformations of Hyp(*X*) observed in this study, the up-puckering conformation of Hyp(*X*) in the Hyp(*X*):Pro(*Y*) stacking pair and the down-puckering conformation of Hyp(*X*) in the Hyp(*X*):Hyp(*Y*) stacking pair, could be explained by the steric hindrance that would occur if they adopted the opposite puckering conformation. Fig. 5 shows that the distance between Hyp13*F* O^{δ} of the ppg9-OOG_T peptide and Pro7*B* C^{γ} is 3.62 Å, while this distance would be longer (5.24 Å) if Hyp13*F* adopted the up-puckering conformation. Although the Hyp13*F* residue has sufficient space to adopt both the up-puckering and the down-puckering conformations without any steric clash with surrounding molecules, it adopts the down-puckering confor-

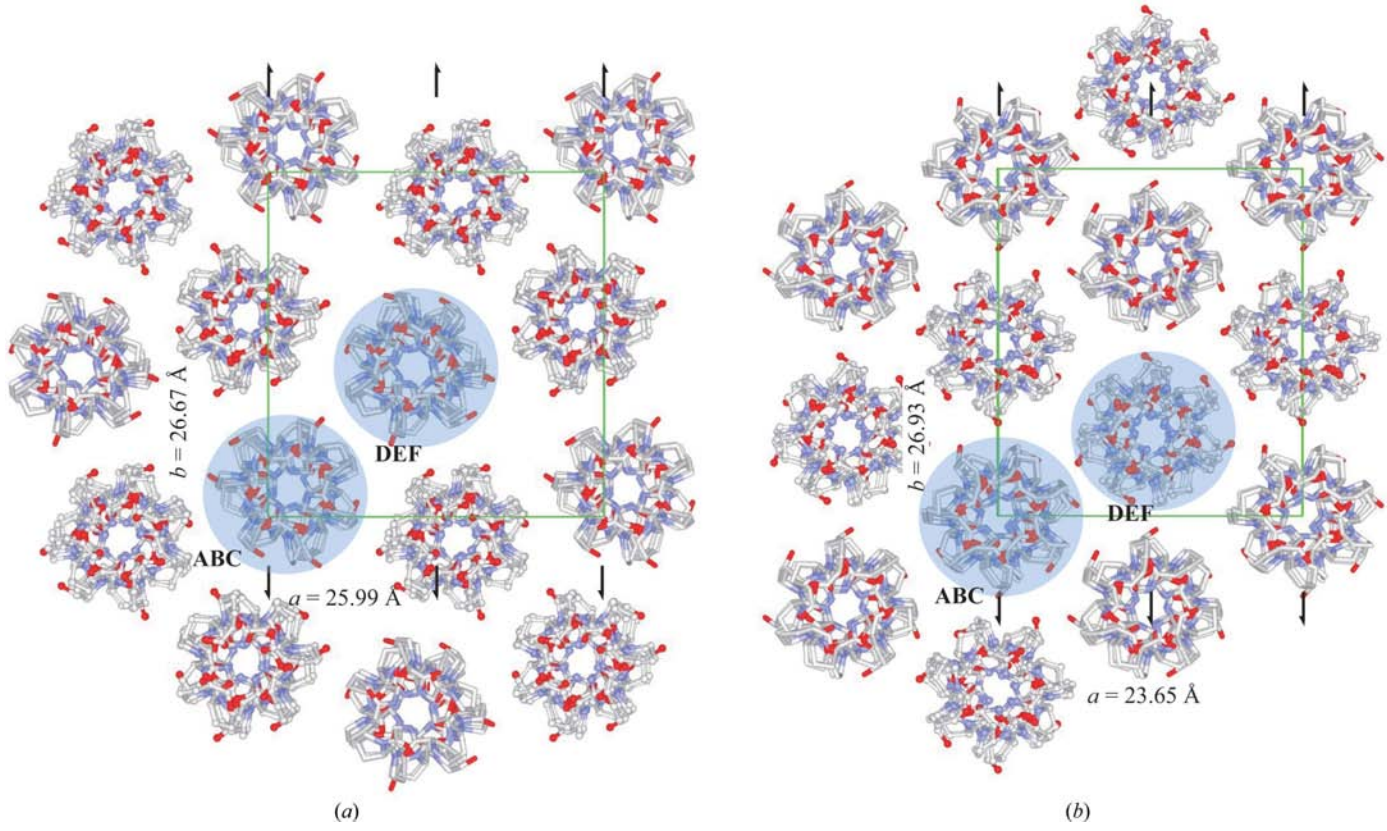

Figure 6
Packing arrangements of ppg9-OOG. (*a*) Pseudo-tetragonal packing of ppg9-OOG_T and (*b*) pseudo-hexagonal packing of ppg9-OOG_H. In both cases the asymmetric unit consists of *ABC* and *DEF* molecules (marked by blue filled circles). Molecules shown as stick models are directed upwards from the paper, whereas those shown as ball-and-stick models are directed downwards from the paper. Molecules in a line along the *b* axis at $u \simeq 0$ and $u \simeq 0.5$ consist of *ABC* and *DEF* molecules, respectively, and are related by the crystallographic 2_1 symmetry along the *b* axis at $u = 0$ and $u = 0.5$, respectively.

mation, which agrees with the hypothesis. A similar situation was also observed for the Hyp13*C* residue of the ppg9-OOG_T peptide. These experimental observations further confirmed that Hyp(*X*) in Hyp(*X*):Pro(*Y*) stacking pairs prefers the down-puckering conformation.

Apart from Hyp in the *X* position, in the collagen helix imino-acid residues in the *X* position prefer a down-puckering conformation and those in the *Y* position prefer an up-puckering conformation; this was deduced from high-resolution analysis of the (Pro-Pro-Gly)$_{10}$ peptide (Vitagliano *et al.*, 2001). In general, these positional preferences are true for both Pro-Pro-Gly and Pro-Hyp-Gly sequences. Specifically, there is no exception for the Pro in the *X* position of the Pro-Pro-Gly sequence or for the Hyp in the *Y* position of the Pro-Hyp-Gly sequence. On the other hand, in some cases the Pro in the *Y* position of the Pro-Pro-Gly sequence and the Pro in the *X* position of the Pro-Hyp-Gly sequence adopt down-puckering and up-puckering, respectively; *e.g.* those in (Pro-Pro-Gly)$_9$ (Hongo *et al.*, 2001, 2005), (Pro-Pro-Gly)$_{10}$ (Berisio *et al.*, 2002; Vitagliano *et al.*, 2001), (Pro-Hyp-Gly)$_{10}$ and (Pro-Hyp-Gly)$_{11}$ (Okuyama *et al.*, 2004). In these cases, hydrophobic interactions between adjacent molecules seem to be the cause of the opposing puckering. In this study, the Pro11*E* residue in the Pro10*E*-Pro11*E*-Gly12*E* sequence of the ppg9-OOG_H peptide adopted a down-puckering conformation (Table 2) in order to make a hydrophobic interaction between Pro11*E* and Pro17*D** (the distance between C$^{\gamma}$ atoms is 3.91 Å).

In conclusion, the puckering conformations of imino-acid residues in the collagen helix can be summarized as follows. (i) Except for Hyp residues in the *X* position, imino acids in the *X* and *Y* positions prefer the down-puckering and up-puckering conformations, respectively. (ii) Hyp(*X*) residues that are involved in Hyp(*X*):Hyp(*Y*) and Hyp(*X*):Pro(*Y*) stacking pairs prefer the up-puckering and down-puckering conformations, respectively.

3.6. Two types of packing arrangements of the ppg9-OOG peptide

Since the first single-crystal structure analysis of a collagen model peptide in 1981 (Okuyama *et al.*, 1981), a considerable number of crystal structures of model peptides have been reported. Although there have been some exceptional cases (Kramer *et al.*, 2001), the rod-like triple-helical molecules of model peptides are usually packed in a parallel fashion, similar to those in the fibrillar structures of native collagen (Fraser *et al.*, 1987). These have been classified into two types of packing arrangements: one is the pseudo-tetragonal arrangement, which is mainly observed for peptides with Pro-Pro-Gly-rich sequences, and the other is the pseudo-hexagonal arrangement, which is mainly observed for peptides with Pro-Hyp-Gly-rich sequences. However, two peptides with similar amino-acid sequences, (Gly-Hyp-Hyp)$_9$ (Schumacher *et al.*, 2005) and (Hyp-Hyp-Gly)$_{10}$ (Kawahara *et al.*, 2005), showed different packing arrangements: the pseudo-hexagonal and the pseudo-tetragonal arrangements, respectively. In the present study, we observed that the same peptide molecule could be crystallized in both packing arrangements under the same crystallization conditions.

In pseudo-tetragonal packing (Fig. 6*a*), the up-pointing *ABC* molecule close to the origin ($u = v = 0$) is surrounded by two down-pointing *ABC* molecules on the *b* axis and three *DEF* molecules, of which two are down-pointing molecules on the *a* axis and one is an up-pointing molecule located close to the centre of the unit cell. The *DEF* molecule is in a similar situation to that of the *ABC* molecule because of the pseudo-2_1 symmetry parallel to the *a* axis. As a result, both *ABC* and *DEF* molecules directly contact five surrounding molecules, mainly by hydrophobic interactions. In the pseudo-hexagonal packing (Fig. 6*b*), the up-pointing *ABC* molecule at the origin is surrounded by two down-pointing *ABC* molecules on the *b* axis and four *DEF* molecules: two up-pointing and two down-pointing molecules. The situation for the *DEF* molecule is very similar to that of the *ABC* molecule, for the same reason as in ppg9-OOG_T. In this case, both the *ABC* and *DEF* molecules directly contact six surrounding molecules mainly by hydrophobic interactions, which leads to the pseudo-hexagonal

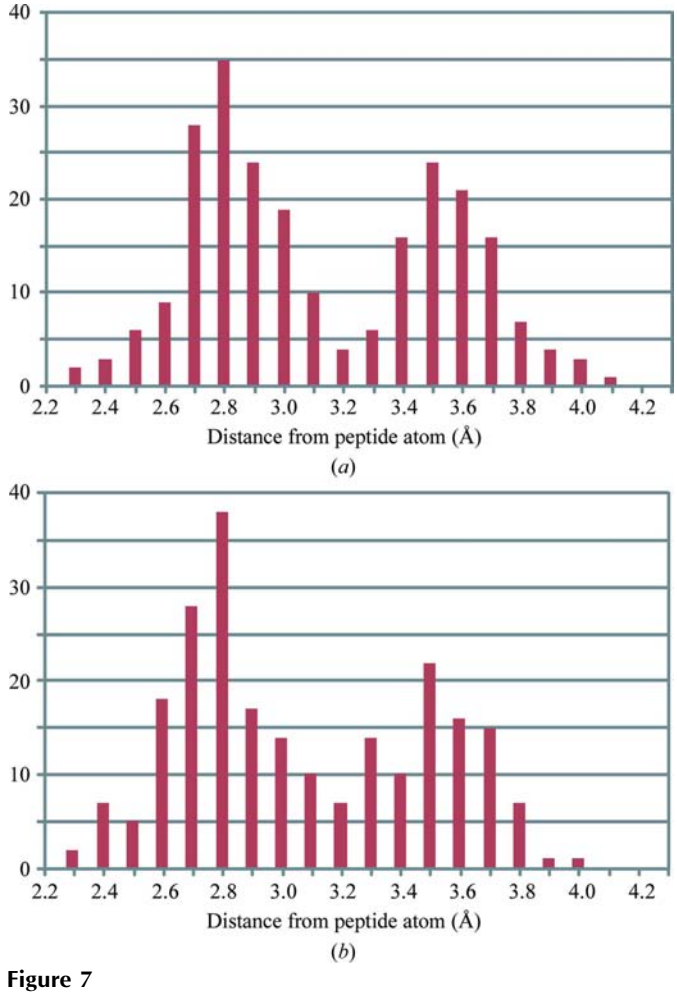

Figure 7
Distribution diagrams of water molecules in (*a*) ppg9-OOG_T and (*b*) ppg9-OOG_H crystals. In both cases, the first peak around 2.8 Å corresponds to the first hydration shell and the second peak around 3.5 Å corresponds to the second hydration shell.

packing arrangement. In the pseudo-hexagonal packing of $(\text{Pro-Hyp-Gly})_n$ (n = 10, 11; Berisio *et al.*, 2001; Okuyama *et al.*, 2004) there is a direct hydrogen bond between Hyp residues of adjacent molecules located along the direction of the *a* axis. However, no such O—H···O hydrogen bond is observed in both crystal modifications.

3.7. Hydration state

In a previous paper (Okuyama *et al.*, 2009), we reported the hydration states of ppg9-*XYG* together with peptides containing repetitions of the Hyp-Hyp-Gly or Pro-Hyp-Gly sequence and concluded that the number of water molecules in the first hydration shell depends on the number of hydrophilic (Hyp) residues in the peptides, while that in the second hydration shell depends on the number of hydrophobic (Pro) residues. Furthermore, we reported that despite the different packing arrangement the distributions of water molecules around the peptide helices of $(\text{Gly-Hyp-Hyp})_9$ and $(\text{Hyp-Hyp-Gly})_{10}$ were very similar to each other.

In this study, we examined the distributions of water molecules in two different packing arrangements of ppg9-OOG peptide crystals. The distributions of water molecules in these peptide crystals are shown in Fig. 7, which indicates a quite similar distribution despite the different packing arrangements. The number of water molecules in the first hydration shell is 136 for ppg9-OOG_T and 139 for ppg9-OOG_H, while the number in the second hydration shell is 102 for ppg9-OOG_T and 93 for ppg9-OOG_H. These data confirm our previous finding that the hydration state of triple-helical molecules only depends on the ratio of the hydrophobic amino-acid residues.

We gratefully acknowledge the BL6A staff at the Photon Factory for beam time and support. We thank Ms Maki Hyakutake for the crystallization and diffraction experiments of ppg9-OOG_H in the initial stages. This work was supported in part by a Grant-in-Aid for Scientific Research (B) from Japan Society for the Promotion of Science.

References

Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1948–1954.
Bann, J. G. & Bächinger, H. P. (2000). *J. Biol. Chem.* **275**, 24466–24469.
Berisio, R., Granata, V., Vitagliano, L. & Zagari, A. (2004). *J. Am. Chem. Soc.* **126**, 11402–11403.
Berisio, R., Vitagliano, L., Mazzarella, L. & Zagari, A. (2001). *Biopolymers*, **56**, 8–13.
Berisio, R., Vitagliano, L., Mazzarella, L. & Zagari, A. (2002). *Protein Sci.* **11**, 262–270.
Bhatnagar, R. S., Pattabiraman, N., Sorensen, K. R., Langridge, R., MacElroy, R. D. & Renugopalakrishnan, V. (1988). *J. Biomol. Struct. Dyn.* **6**, 223–233.
Boudko, S. P., Engel, J., Okuyama, K., Mizuno, K., Bächinger, H. P. & Schumacher, M. A. (2008). *J. Biol. Chem.* **283**, 32580–32589.
Burjanadze, T. V. (1979). *Biopolymers*, **18**, 931–938.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Fraser, R. D., MacRae, T. P. & Miller, A. (1987). *J. Mol. Biol.* **193**, 115–125.
Hongo, C., Nagarajan, V., Noguchi, K., Kamitori, S., Okuyama, K., Tanaka, Y. & Nishino, N. (2001). *Polymer J.* **33**, 812–818.
Hongo, C., Noguchi, K., Okuyama, K., Tanaka, Y. & Nishino, N. (2005). *J. Biochem.* **138**, 135–144.
Improta, R., Berisio, R. & Vitagliano, L. (2008). *Protein Sci.* **17**, 955–961.
Inouye, K., Kobayashi, Y., Kyogoku, Y., Kishida, Y., Sakakibara, S. & Prockop, D. J. (1982). *Arch. Biochem. Biophys.* **219**, 198–203.
Kawahara, K., Nishi, Y., Nakamura, S., Uchiyama, S., Nishiuchi, Y., Nakazawa, T., Ohkubo, T. & Kobayashi, Y. (2005). *Biochemistry*, **44**, 15812–15822.
Kramer, R. Z., Bella, J., Brodsky, B. & Berman, H. M. (2001). *J. Mol. Biol.* **311**, 131–147.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Mizuno, K., Hayashi, T., Peyton, D. H. & Bächinger, H. P. (2004). *J. Biol. Chem.* **279**, 38072–38078.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Okuyama, K. (2008). *Connect. Tissue Res.* **49**, 1–12.
Okuyama, K., Hongo, C., Fukushima, R., Wu, G., Narita, H., Noguchi, K., Tanaka, Y. & Nishino, N. (2004). *Biopolymers*, **76**, 367–377.
Okuyama, K., Hongo, C., Wu, G., Mizuno, K., Noguchi, K., Ebisuzaki, S., Tanaka, Y., Nishino, N. & Bächinger, H. P. (2009). *Biopolymers*, **91**, 361–372.
Okuyama, K., Okuyama, K., Arnott, S., Takayanagi, M. & Kakudo, M. (1981). *J. Mol. Biol.* **152**, 427–443.
Okuyama, K., Takayanagi, M., Ashida, T., Kakudo, M., Sakakibara, S. & Kishida, Y. (1977). *Polym. J.* **9**, 341–343.
Okuyama, K., Wu, G., Jiravanichanun, N., Hongo, C. & Noguchi, K. (2006). *Biopolymers*, **84**, 421–432.
Okuyama, K., Xu, X., Iguchi, M. & Noguchi, K. (2006). *Biopolymers*, **84**, 181–191.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Persikov, A. V., Ramshaw, J. A., Kirkpatrick, A. & Brodsky, B. (2003). *J. Am. Chem. Soc.* **125**, 11500–11501.
Ramachandran, G. N. & Kartha, G. (1955). *Nature (London)*, **176**, 593–595.
Rich, A. & Crick, F. H. (1955). *Nature (London)*, **176**, 915–916.
Schumacher, M., Mizuno, K. & Bächinger, H. P. (2005). *J. Biol. Chem.* **280**, 20397–20403.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Sugeta, H. & Miyazawa, T. (1967). *Biopolymers*, **5**, 673–679.
Vitagliano, L., Berisio, R., Mazzarella, L. & Zagari, A. (2001). *Biopolymers*, **58**, 459–464.
Yeates, T. O. (1997). *Methods Enzymol.* **276**, 344–358.