生物化学与生物物理进展 *Progress in Biochemistry and Biophysics*, 2006, 33(3): 231~240
www.pibb.ac.cn

PIBB

研究报告

A Dimeric Structure of The Scorpion Toxin BmK M1 at 1.4 Å Resolution: Non-proline *cis* Peptide Bond and Its Inntrinsic Structural Elements*

HE Xiao-Lin**, GUAN Rong-Jin**, ZHANG Ying, WANG Da-Cheng***
(*Center for Structural and Molecular Biology, Institute of Biophysics, The Chinese Academy of Sciences, Beijing* 100101, China)

Abstract Non-proline *cis* peptide bond is rarely found in proteins. The recent surveys revealed that this unusual peptide configuration is by no means a curiosity, but overwhelmingly occurs at functionally important sites. However one still doubts whether it is related to crystal packing interactions, since all non-proline *cis* peptide bonds identified so far are from crystal structures. A toxin BmK M1 from the scorpion *Buthus martensii* Karsch have been crystallized as a dimer in space group $P2_12_12_1$ with unit-cell dimensions a = 76.39 Å, b=52.77 Å, c=27.12 Å. This dimeric structure was solved by molecular replacement and refined to R=0.109 for all reflections at a resolution of 1.4 Å. The extensively refined structure definitely shows that the *cis* peptide bond Pro9-His10 equally occurs in both molecule A and molecule B in the dimer. The observation manifested that this striking non-proline *cis* peptide is not related to crystal packing, but caused by certain intrinsic factors. The detailed analyses and comparison with the structure of BmK M8, which is homologous to M1 but has *trans* peptide bond 9-10, showed that the five-residue reverse (8～12) with a consensus sequence (-KPXNC-) may be the intrinsic structural element for the *cis* form of this peptide bond. A pair of well organized main-chain hydrogen bond between residues 10 in *cis* unit and 64 at C-terminus forms main tertiary interactions to stabilize this energetically unfavorable peptide bond.

Key words Non-proline *cis* peptide bond, crystal structure, α-like toxin, scorpion *Buthus martensii* Karsch

1 Introduction

It is a general principle in protein science for many years that the peptide bonds in protein structures occur predominantly in the *trans* configuration due to its double-bond character[1]. Especially, the non-proline *cis* peptide (Xaa-non-Pro) in which the amide nitrogen is provided by an amino acid rather than proline is rarely found in protein structure. The occasional occurrence of such *cis* peptide is usually considered as a curiosity. Statistic analyses based on respective 32 539[2] and 153 209[3] peptide bonds showed that only 0.05% or 0.03% of all Xaa-nPro peptide bonds occurred in the *cis* conformation. Correspondingly, in X-ray crystal structure analysis most refinement programs widely used for many years (such as X-PLOR, Brunger *et al*., 1992[4]) allow only to build a *trans* conformation of a Xaa-nPro bond into the structural model. However, this dogma for the peptide bond is gradually changed in recent years. A series of surveys and analyses[3, 5～7] based on more three-dimensional structures at rather higher resolution revealed that the non-Pro *cis* peptide bonds are by no means a curiosity and overwhelmingly occur at functional important site in protein structures. Recently several reports from some proteins, such as concanavalin A[8], phosphoribosyltransferase[9] and an antibody CDR region[10], in which non-Pro *cis* peptide

*This work was supported by the grants of The National Natural Science Foundation of China (30370320) and the Chinese Academy of Sciences (KSCX1-SW-17, KSCX2-SW-322)), X-ray Data collection was supported by the Photon Factory, KEK (2005G253).
**These two authors made equal contribution to this paper.
***Corresponding author .
Tel: 86-10-64888547, Fax: 86-10-64888560
E-mail: dcwang@sun5.ibp.ac.cn
Received: October 9, 2005 Accepted: December 30, 2005

bonds were unequivocally identified, have showed that the non-Pro *cis* peptide should be crucial structural factors governing the certain functional performances. However, what does one make of these peptide bonds to be *cis* form? Investigations still doubt whether these structures were forced into the high-energy *cis* conformation by crystal packing interactions [11], because all non-Pro *cis* peptide bonds identified so far are found in crystallographic conditions. Here we report a dimeric structure of BmK M1, in which a *cis* non-Pro peptide occurs in both monomers in an asymmetric unit.

BmK M1 is an α-like neurotoxin acting on the sodium channel [12,13] from the scorpion *Buthus martensii* Karsch, which is composed of 64 residues cross-linked by four disulfide bridges. Previously the crystal structure of BmK M1 was determined as a monomer in an asymmetric unit at 1.7 Å resolution, in which a non-Pro *cis* peptide bond was found between residues Pro9 and His10 [13]. Meanwhile several other structures of BmK toxins, such as BmK M2, M4 and M8, were also determined as monomers, in which either *cis* or *trans* form of 9-10 peptide bond was observed in the respective structures [13~16]. In order to investigate the possible influence of the crystal packing on this unusual *cis* peptide bond a new crystal form with two molecules of BmK M1 in the asymmetric unit has been crystallized and the structure has been determined at 1.4 Å resolution. In this paper the dimeric structure of BmK M1 and the *cis* peptide bonds occurred in this structure as well as a possible sequence motif related to the occurrence of this *cis* peptide bond will be described.

2 Materials and methods

2.1 Crystallization and data collection

The purification of BmK M1 was described earlier by Li *et al*. [12]. The crystallization was performed by hanging-drop vapor-diffusion method. The protein sample was dissolved in 1.0 mmol/L acetic acid at 20 g/L, and then mixed with equal volume of reservoir solution containing 2.8 mol/L NaH_2PO_4 at pH 4.5. The drops were equilibrated against the reservoir solution for about 15 days. Crystals with dimensions of 1.0 mm × 0.3 mm × 0.1 mm were finally obtained. The preliminary X-ray diffraction analysis showed the crystal form belonged to the space group $P2_12_12_1$. There are two molecules in an asymmetric unit with a Matthews coefficient value of 1.88 Dalton·Å^{-3}.

The crystals could well diffract to 1.8 Å resolution on a Mar-Research Image Plate area detector with rotating anode Cu $K\alpha$ radiation. The data set to 1.4 Å resolution, which was used for the structure determination, were collected with the synchrotron radiation at Photon Factory in KEK of Tsukuba, Japan (Beamline BL6A2, wavelength 1.0 Å). A total of 78 828 observations were recorded, which were scaled and merged by the CCP4 package[17] into a set of 25 220 unique reflections with R_{merge} of 6.2%. The completeness was 91.4% for the whole data set (86.3% for the highest resolution shell 1.5～1.4 Å).

2.2 Structure determination

The structure of BmK M1 in the new crystal form was solved using the molecular replacement method with the program AMoRe [18], using the monomer structure in space group $P2_12_12$ of BmK M1 as a probe (1SN1 in Protein Data Bank) [16]. All the water molecules were discarded from 1SN1 and the temperature factors of the remaining atoms were set to 15.0 Å^2. Using a Patterson cutoff radius of 18 Å and data in the resolution range of 10～3.5 Å, a list of 99 rotation function peaks were obtained, with the top peak having an AMoRe Correlation Coefficient (*CC*) of 18.3, and the second peak having a *CC* of only 11.8. The results clearly suggested that the top peak is a correct solution. A translation function, using the data in the same resolution range, proved that the top peak of the rotation search gave a *CC* value of 40.3, and an *R* factor of 50.2%, whereas the second peak gave only a *CC* of 22.8 with an *R* factor of 54.6%. This solution (α = 46.92, β = 41.13, γ = 164.52, x = 0.0852, y = 0.0562, z = 0.4480) was then fixed and the translation function calculation was further performed to find the second molecule. This process seemed straightforward and only one solution made the *CC* rise. After rigid-body refinement the final solutions for two molecules in the asymmetric unit are: $\alpha 1$ = 47.32, $\beta 1$ = 42.14, $\gamma 1$ = 164.15, $x1$ = 0.0867, $y1$ = 0.0566, $z1$ = 0.4480; $\alpha 2$ = 70.49, $\beta 2$ = 76.28, $\gamma 2$ = 233.75, $x2$ = 0.4396, $y2$ = 0.7517, $z2$ = 0.0629. They together gave a *CC* of 64.1 and an *R* factor of 39.6%.

2.3 Crystallographic refinement

The structure was refined on a Silicon Graphics O_2 workstation and modeled with the program O [19]. The refinement was carried out against 95% of the

data. The remaining 5% randomly selected from the full data were used for cross-validation in which the free R factor (R_{free}) was calculated to monitor the progress of refinement[20].

Program XPLOR 3.851 [4] was used in initial refinement. The model was first subjected to rigid-body refinement for the data between 10 Å and 3 Å. Then simulated annealing with the slow-cooling protocol provided in XPLOR was applied to the data between 10 Å and 1.8 Å, which allowed the R value to be dropped to 38.7% and R_{free} to 32.3%. This process was followed by minimization and B factor refinements. Finally using all the data between 20～1.2 Å, several cycles of model rebuilding and minimization gave an R of 15.9 and R_{free} of 20.2, with 120 water molecules incorporated.

The structural model was then further refined by SHELXL 97 [21]. With all atoms having individual isotropic temperature factors, the refinement preceded to R=16.1%. When all atoms adopted anisotropic B factors, the R value was lowered to 13.4%. Then adding more waters and giving half occupancy to those waters with high temperature factors reduced the R value to 12.0%. Adding riding hydrogen atoms gave an R factor value of 11.1% and further refinement converged to an R value of 10.8% and R_{free} of 15.9%. Finally, a full-matrix least square minimization was carried upon all the data including the 5% set-aside data, giving a final R value of 10.9%.

2.4 Homologous sequence search

The non-Pro *cis* peptide bonds of BmK toxins occur in a distinct five-residue reverse turn (8～12) which has a conservative sequence in the same group of toxins. For surveying whether this consensus sequence occur also in other proteins, the program BLASTp[22] was used for the homologous search with sequence data from all sequence database. The search was performed on a remote server run by NIH (www.ncbi.nlm.nih.gov/blast/). The fragment sequence (residues 6～14) of BmK M1, which contains the consensus sequence (residues 8～12), was used as a template to compare. The sequences found were aligned based on the BLAST agrithrym. The sequence comparison and the postscript figure of alignment were prepared with ALSCRIPT[23].

3 Results and discussion

3.1 Quality of the model

The structure of BmK M1 in $P2_12_12_1$ space group was refined to 1.4 Å resolution with a final R factor of 0.109 for all data between 20 Å and 1.4 Å. The result of refinement was summarized by the statistics in Table 1. Electron density maps with high quality fit to all residues very well. Figure 1 shows 2Fo-Fc maps around the five-residue reverse turn (8～12), in which a *cis* peptide bond appeared between residues 9 and 10. Ramachandran plot [1] shows that 91.7% of the non-glycine and non-proline residues fall in the core region and the rest 8.3% fall in the allowed region. No residues were found in the generously allowed region or disallowed region. The upper limit of the error on the atomic positions is estimated to be less than 0.10 Å by means of Luzzati plot [24]. A total of 3 alternative conformations for the side-chains were modeled, including Arg2 in molecule A and Val39 and Lys41 in molecule B. The other residues all have definite electron densities.

Table 1　Refinement statistics of the dimeric structure of BmK M1

Resolution range	20.0～1.4 Å
I/I(σ) cutoff	None
Number of reflections	25 220
Completeness	91.0% (80.4%)
R_{merge}	7.2% (25%)
R_{cryst}	0.109
R_{free}	0.159
R.M.S. deviation	
Bond length	0.012 Å
Bond angle	2.3°
Angle distance	0.030 Å
Zero chiral volumes	0.128 Å^3
Non-zero chiral volumes	0.125 Å^3
Anti-bumping distance restraints	0.046 Å
Rigid-body ADP components	0.004 Å^2
Similar ADP components	0.058 Å^2
Approximately isotropic ADPs	0.200 Å^2
Ramachadran plot	
Core region	91.3 %
Allowed region	8.7 %
Protein model	
Protein atoms	1021
Solvent atoms	192

Values in parentheses refer to the outer resolution shell.

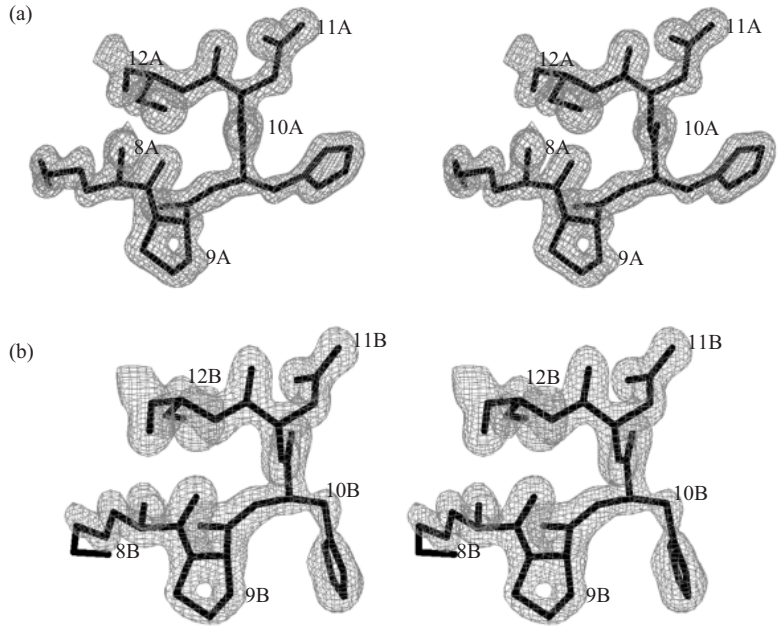

Fig. 1 Electron density maps (2Fo-Fc, contoured at 2σ) around the five-residue (8~12) reverse turn in molecule A (a) and molecule B (b), respectively

3.2 Dimeric structure of BmK M1

There are two M1 molecules, denoted as molecule A and B respectively, in the asymmetric units. These two molecules are related by a pseudo 2-fold axis of 168° to form a crystallographic dimer (Figure 2a). In each monomer, the secondary structure is dominated by one α-helix and a three-stranded β-sheet that form the dense core of the molecule, which is cross-linked by 4 disulfide bridges (Cys12-Cys63, Cys16-Cys36, Cys22-Cys46, Cys26-Cys48). The rest of the protein is made up of loops and turns joining the helix and β-strands. The general fold of the molecule is similar to the monomer structure determined in space group $P2_12_12$ [14]. The only major difference occurs in the loop 38～44, as indicated by the overlapping of the main chains. In fact, this loop is the most flexible part of the molecule as shown by the relatively high temperature factors. This loop has been proposed to be the characteristic of the Old World toxins and may play a role in the functional performance for the α- and α-like toxins[25].

Till now, most of scorpion toxin structures, either by crystallographic or NMR analysis, are in monomeric forms. During the purification and characterization no aggregation was detected in solution. Thus these dimer in the asymmetric unit should be formed mainly due to the crystallization. It was found that organic agents were needed in crystallizing BmK M1. However, when less than 0.5% alcohol added into the reservoir solution, the precipitation of the protein became very difficult even the concentration of precipitant was raised by 3 times. This observation suggests that there is dimer-monomer balance in solution, but packing interactions in crystallization promoted a bias towards the dimeric state. Actually the solved structure showed that the interaction between the two monomers mainly encompasses a 4 aromatic residue cluster (Figure 2b). In addition, Glu15 is also essential for dimer formation, because there are several hydrogen bonds related with it (E15A OE1---A17B N, 2.88 Å; E15B OE1---A17A N, 2.86 Å; E15A OE1---R18B N, 3.19 Å; E15B OE1---R18A N, 3.13 Å). Furthermore, a pair of hydrogen bonds between C-terminal residues and the side chain of K28 (K28A NZ---H64B OXT, 2.91 Å; K28B NZ---H64A OXT, 2.86 Å) also contribute to the dimer forming interactions. Obviously, the stereochemical environment described above in the present dimeric BmK M1 is completely different from that in the monomeric structure [14]. Therefore, the present dimeric structure provided a chance to inspect the possible influence of crystal packing on the *cis* peptide bond.

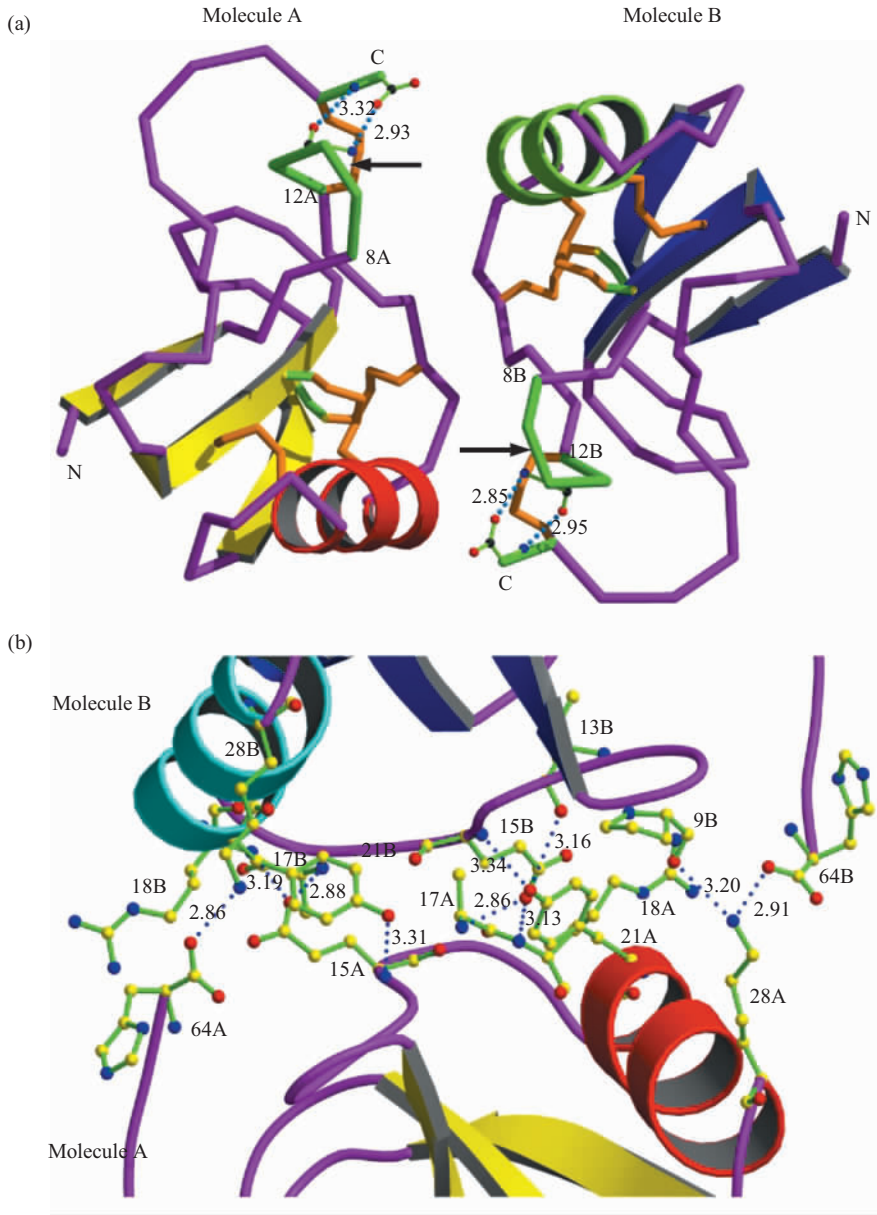

Fig. 2　Dimeric structure of BmK M1 (a) and intermolecular contacts on the dimer-forming surface (b)

The five-residue reverse turn with the *cis* peptide bond is highlighted in green color. Arrows indicate the location of the *cis* peptide bond 9-10 in respective monomer A and B. The figures were prepared with MOLSCRIPT[26] and rendered by RASTER3D[27].

3.3　Non-proline *cis* peptide bond in the dimer

The non-proline *cis* peptide bond between residues Pro9 and His10 occurs equally in both molecule A and B of the dimer. During refinement it was impossible to restrain the ω values of these peptide bonds near ±180°, whereas the ω angles went favorably towards 0° under rebuilding (Table 2). The high-resolution 2Fo-Fc and Fo-Fc maps clearly showed the *cis* configuration for these peptides in the respective monomer A and B (Figure 3). In these *cis* peptide bonds, the amide nitrogen is provided by an amino acid (His10) rather than proline. Therefore they belong to the non-proline *cis* peptide bond.

Compared with the *cis* peptide Pro9-His10 observed in the $P2_12_12$ monomeric structure, the ω values in the dimer structure have a small amount of deviation ($\omega_A = -11.6°$, $\omega_B = -4.1°$) from that in the monomeric structure ($\omega = 0°$) (Table 2, Figure 4). The present result showed that the geometry of the *cis* peptide Pro9-His10 in the dimeric structure is basically

the same as in the monomeric structure. The small variation of the geometry may be due to the different resolutions and refinement extent between the dimeric structure (1.4 Å) and monomeric structure (1.7 Å). Now there are three molecular structures from the same BmK M1 but in different crystal packing, including two as monomers of the dimer in $P2_12_12_1$ cells and one as an individual molecule in $P2_12_12$ lattice. The extensively refined structures at high resolutions showed that the non-proline *cis* peptide bond Pro9-His10 can occur in these structures with the same geometric status. We therefore conclude that this unusual non-proline *cis* peptide bond is not related to crystal packing interactions, but caused by certain intrinsic structural reasons.

Table 2 The main-chain conformation analyses of the reverse turn K8-RT in dimer and monomer of Bmk M1 and D8-RT in BmK M8

BmK M1	Mol A in dimer			Mol B in dimer			Monomer			BmK M8			
	ϕ_i	φ_i	ω_{i-1}	ϕ_i	φ_i	ω_{i-1}	ϕ_i	φ_i	ω_{i-1}		ϕ_i	φ_i	ω_{i-1}
Lys8	−117.7	156.7	166.0	−94.2	141.3	166.6	−116.3	161.0	172.8	Asp8	−81.73	−162.8	179.9
Pro9	−58.4	137.5	−172.4	−57.7	136.0	178.4	−60.2	136.6	−178.9	Ser9	−75.11	−11.13	−177.3
His10	−109.7	114.0	**−11.6**	−111.8	107.1	**−4.1**	−120.0	106.8	**−0.6**	Glu10	−104.9	30.0	**179.8**
Asn11	58.0	48.1	175.7	61.0	45.1	−179.0	63.1	50.4	179.9	Asn11	49.6	42.1	−174.5
Cys12	−123.3	154.5	174.6	−117.8	147.9	179.4	−124.5	143.0	−179.5	Cys12	−111.2	148.3	178.5

Data for M1 monomer and M8 were taken from 1SN1 and 1SBB in the Protein Data Bank.

The ω values related to *cis* or *trans* peptide bond (9-10) are highlighted with boldface.

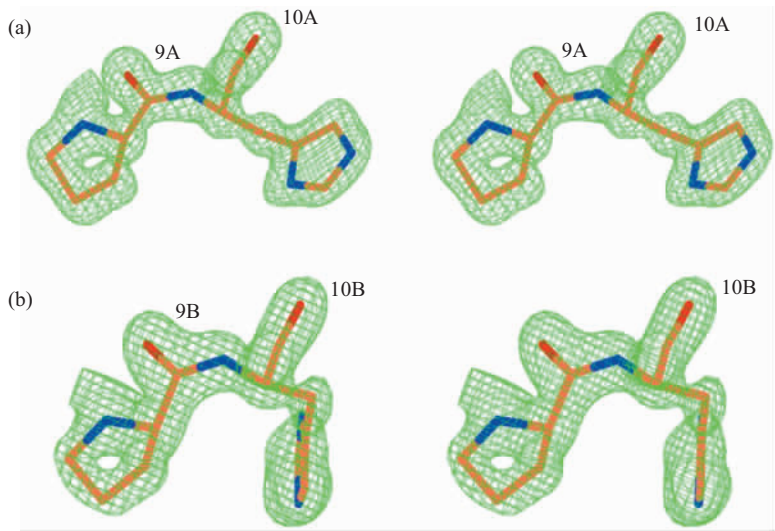

Fig. 3 The stereo view of the 2Fo-Fc maps (contoured at 2σ) for the residues 9 and 10 in molecule A (a) and B (b), respectively

These maps clearly show the *cis* configuration of the peptide bond Pro9-His10.

3.4 Intrinsic structures elements for the *cis* peptide bond

Generally the *cis* geometric configuration is determined by certain local structural elements. The *cis* peptide bond Pro9-His10 occurs in a five-residue reverse turn (RT) in BmK M1 (Figure 4). There are several peculiarities in the constitution of this turn. Firstly, the last residue, Cys12, is identical in all scorpion toxins including α and β- types. It covalently links another identical residue Cys63 at the C-terminus through a disulfide bridge. Secondly, the *i*+3 residue Asn 11, is identical in all α and α-like toxins (Figure 5). Interestingly, the well determined structures[13~16] showed that the position 8 is distinct,

where residue Lys 8 is conservatively occurred in *cis* peptide-containing reverse turn (K8-RT), but Asp8 in *trans* peptide-containing reverse turn (D8-RT) (Figure 4 and 5). Moreover, the local structures of K8-RT of D8-RT are also distinctive. The K8-RT (BmK M1) is well stabilized by a couple of main-chain hydrogen bonds between residues 8 and 12 (Figure 4). However, a more complicated hydrogen bond network, including a pair of H-bonds from the Glu10 side-chain and the main-chain of residue 10 and 12, occurs in D8-RT (BmK M8) (Figure 4). The geometry of K8-RT and D8-RT is also different. The striking distinctions of residue 8 in both sequence and structure suggest that the K8- or D8-containing five-residue reverse turn should be the intrinsic structural elements for the *cis* or *trans* form of peptide bond 9-10. Most recently, mutagenesis experiments and structural analyses have identified that K8/D8 is really a structural switch for *cis*/*trans* isomerization of the peptide bond 9-10, which is related to the phylogenetic selectivity of BmK M1 for the sodium channels of mammal or insect[28, 29].

It is also notable that a pair of well organized main-chain H-bonds between residues 10 at the *cis* bond and 64 at the C-terminus (Figure 4) are found in all structures containing the *cis* peptide bond, but not in that containing the *trans* peptide bond. Evidently these tertiary interactions are the structural elements for stabilizing the *cis* configuration, since it is energetically unfavorable. In fact, in the structure of a new toxin BmK M7 with longer C-terminal segment the loss of these H-bonds induces the *cis*/*trans* isomerization of this peptide bond[28]. Moreover, the conformation of the functionally important C-terminal segment[28 ~30] is mediated by the *cis* peptide bond, which provides a special way to reach a structural state necessary for the functional performance through the strained backbone geometry.

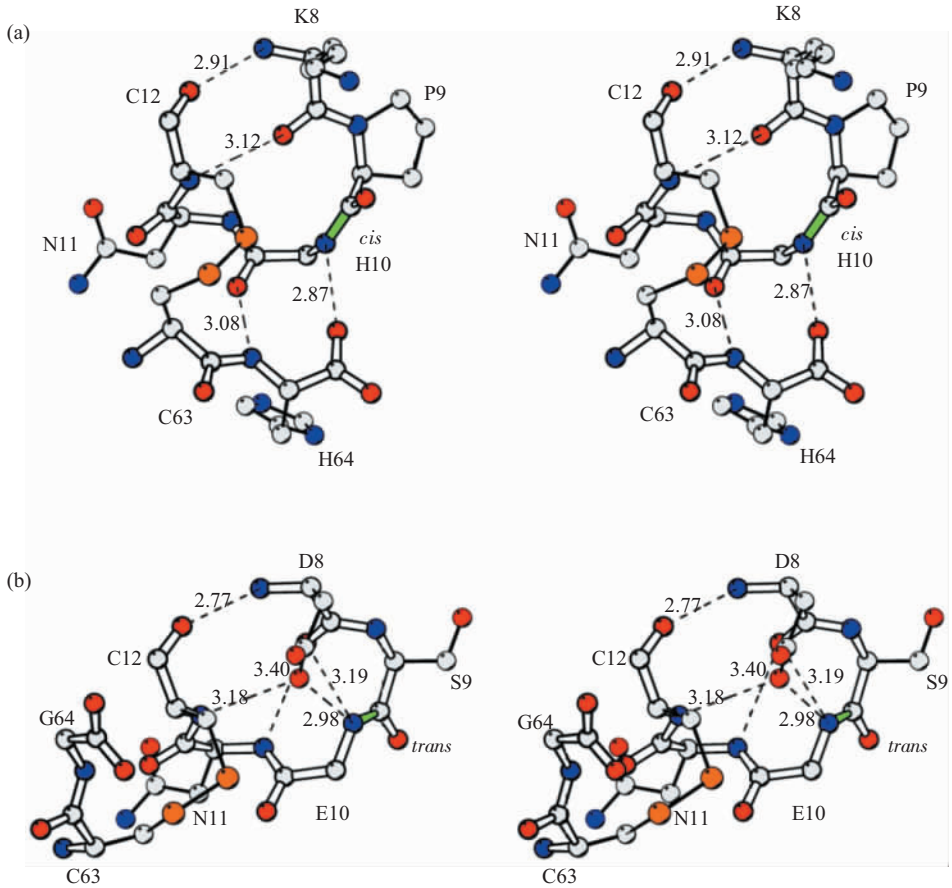

Fig. 4　The local structure of the five-residue (8~12) reverse turns and C-terminal residues in BmK M1-A (K8-RT) with *cis* peptide bond 9-10 (a) and in BmK M8 with *trans* peptide bond 9-10 (D8-RT) (b)

It is noticeable that K8-RT and D8-RT are distinctive in their geometry and contacts with C-terminal residues. The figures were prepared with MOLSCRIPT[26].

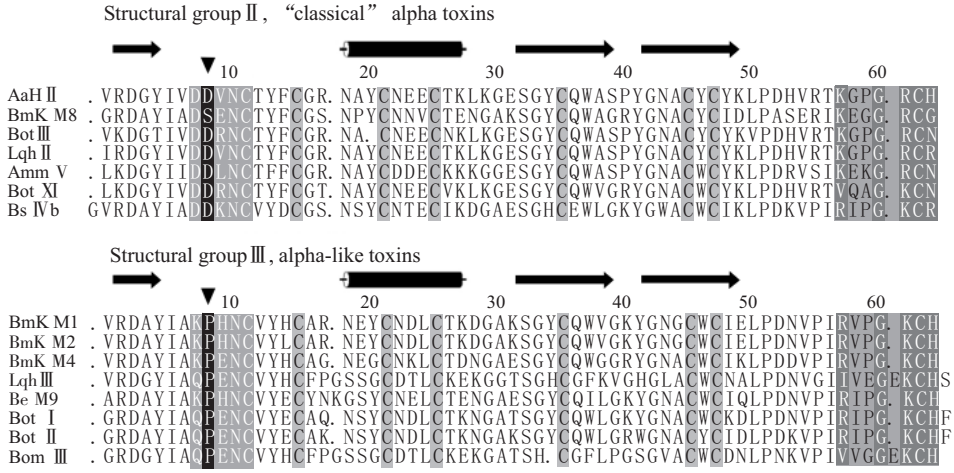

Fig. 5 Amino acid sequences of some classical α toxins and α-like toxins

Those residues corresponding to the reverse turn with the *trans* or *cis* peptide bond 9-10 in classical α toxins and α-like toxins,respectively, are highlighted by box.

3.5 Possible consensus sequences for the reverse turn with the non-proline *cis* peptide bond

The above analyses and comparisons showed that the sequence motif (-KPXNC-) (X could be any residues) might be a consensus sequence for a *cis*-containing five-residue reverse turn. In order to know whether this sequence could be appeared in other proteins, a search for identical sequence based on the database in the SWISSPORT have been performed. More than 18 local sequences identical to the sequence motif (-KPXNC-) were found in different kinds of proteins (Figure 6). One can assume that the local structure determined by this unique sequence motif in the proteins may take the same as the K8-RT reverse turn with a *cis* peptide bond like in the BmK M1 structure. Unfortunately all these proteins have no structures available at the moment. Therefore, the question whether the assumption described above is correct is still open.

The coordinates and the structural factors have been deposited into the PDB databank, with the accession number of 1DJT and RCSB010142.

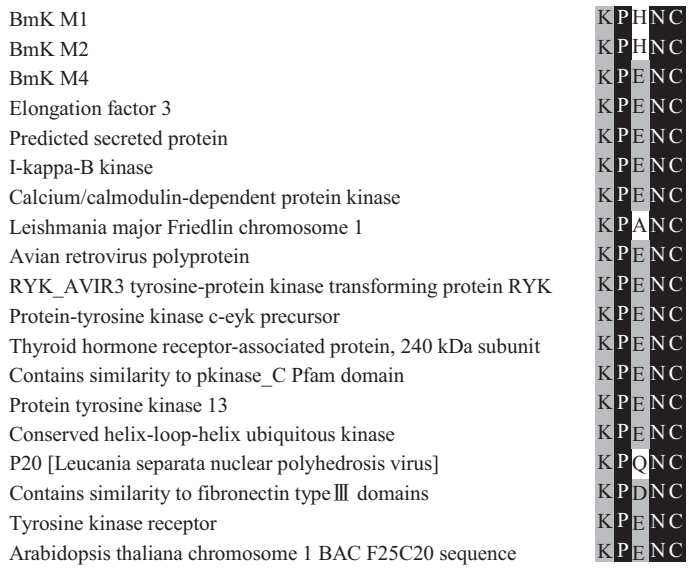

Fig. 6 Some protein sequence segments which show a local sequence (-KPXNC-) identical to the consensus sequence for the *cis* peptide bond 9-10 in the toxin BmK M1

References

1 Ramachandran G N, Sasisekharan V. Conformation of polypeptides and proteins. Adv Protein Chem, 1968, **23**: 283～438

2 Stewart D E, Sarkar A, Wampler J E. Occurrence and role of *cis* peptide bonds in protein structures. J Mol Biol, 1990, **214**: 253～260

3 Weiss M S, Jabs A, Hilgenfeld R. Peptide bonds revisited. Nat Struct Biol, 1998, **5**: 676

4 Brunger A T. XPLOR: A System for X-ray Crystallography and NMR. New Haven, CT: Yale University, 1992

5 Jabs A, Weiss M S, Hilgenfeld R. Non-proline *cis* peptide bonds in proteins. J Mol Biol, 1999, **286**: 291～304

6 Pal D, Chakrabarti P. *cis* Peptide bonds in proteins: residues involved, their conformations, interactions and locations. J Mol Biol, 1999, **294**: 271～288

7 Pappenberger G, Aygun H, Engels J W, *et al*. Nonprolyl *cis* peptide bonds in unfolded proteins cause complex folding kinetics. Nat Struct Biol, 2001, **8**: 452～458

8 Bouckaert J, Dewallef Y, Poortmans F, *et al*. The structural features of concanavalin A governing non-proline peptide isomerization. J Biol Chem, 2000, **275**: 19778～19787

9 Heroux A, White E L, Ross L J, *et al*. Crystal structures of the Toxoplasma gondii hypoxanthine-guanine phosphoribosyltransferase-GMP and -IMP complexes: comparison of purine binding interactions with the XMP complex. Biochemistry, 1999, **38**: 14485～14494

10 Bates P A, Dokurno P, Freemont P S, *et al*. Conformational analysis of the first observed non-proline *cis*-peptide bond occurring within the complementarity determining region (CDR) of an antibody. J Mol Biol, 1998, **284**: 549～555

11 Heroux A, White E L, Ross L J, *et al*. Crystal structure of Toxoplasma gondii hypoxanthine-guanine phosphoribosyltransferase with XMP, pyrophosphate, and two Mg (2+) ions bound: insights into the catalytic mechanism. Biochemistry, 1999, **38**: 14495～14506

12 Li H M, Zhao T, Wang D C, *et al*. A series of bioactivity-variant neurotoxins from scorpion *Buthus martensii* Karsch: purification, crystallization and crystallographic analysis. Biol Crystallography (Acta Cryst. D), 1999, **55**: 341～344

13 Wang C G, Nicolas G, Wang D C, *et al*. Exploration of the functional site of a scorpion alpha-like toxin by site-directed mutagenesis. Biochemistry, 2003, **42**: 4699～4708

14 He X L, Li H M, Wang D C, *et al*. Crystal structures of two scorpion alpha-like toxins: non-proline *cis* peptide bonds and implications for new binding site selectivity on sodium channel. J Mol Biol, 1999, **292**: 125～135

15 Li H M, Wang D C, Zeng Z H, *et al*. Crystal structure of an acidic neurotoxin from scorpion *Buthus martensii* Karsch at 1.85 Å resolution. J Mol Biol, 1996, **261**: 415～431

16 He X L, Deng J P, Wang D C, *et al*. Crystal structure of a new neurotoxin from the scorpion *Buthus martensii* Karsch at 1.76 Å. Biol Crystallography (Acta Cryst. D), 2000, **D56**: 25～33

17 Collaboratory Computational Project, Number 4. The CCP4 suit. Acta Cryst, 1994, D50, 760～763

18 Navaza J. AMoRe - an automated package for molecular replacement. Acta Cryst, 1994, **A50**: 157～163

19 Jones T A, Zou J Y, Cowan S W, *et al*. M., Improved methods for binding protein models in electron density maps and the location of errors in these models. Acta Cryst, 1991, **A47**: 110～119

20 Kleywegt G J, Brunger A T. Checking your imagination: application of the free R value. Structure, 1996, **4**: 897～904

21 Sheldrick G M, Schneider T R. Methods Enzymol, 1997, **277**: 319～343

22 Altschul S F, Madden T L, Schaffer A A, *et al*. Gapped BLAST and PSI-BLAST: a new generation of protein data base search programs. Nucleic Acids Res, 1997, **25**: 3389～3402

23 Barton G J. ALSCRIPT: a tool to format multiple sequence alignments. Protein Eng, 1993, **6**: 37～40

24 Luzzati V. Traitment statistique des erreuis dans la determination des structure cristallines. Acta Cryst, 1952, **5**: 802～810

25 Possani L D, Becerril B, Delepierre M, *et al*. Scorpion toxins specific for Na^+-channels. Eur J Biochem, 1999, **264**: 287～300

26 Lawrence M C, Colman P M. Shape complementarily at protein/protein interfaces. J Mol Biol, 1993, **234**: 946～950

27 Merritt E, Murphy M E P. Raster3D version 2.0: a program for potorealistic molecular graphics. Acta Cryst, 1994, **D50**: 869～873

28 Guan R J, Xiang Y, He X L, *et al*. Structural mechanism governing *cis* and *trans* isomeric states and an intramolecular switch for *cis*/*trans* isomerization of a non-proline peptide bond observed in crystal structures of scorpion toxins. J Mol Biol, 2004, **341**: 1189～1204

29 Liu L H, Bosmans F, Maertens C, *et al*. Molecular basis of the mammalian potency of the scorpion a-like toxin, BmK M1. FASEB J, 2005, **19**: 594～596

30 Xiang Y, Bosmans F, Li C, *et al*. Structural basis for the voltage-gated Na^+ channel selectivity of the scorpion a-like toxin BmK M1. J Mol Biol, 2005, **353**: 788～803

二聚体蝎钠通道毒素的 1.4 Å 分辨率晶体结构：反常非脯氨酸顺式肽键及其内在结构因素 *

何小林 ** 管荣津 ** 张 英 王大成 ***
(中国科学院生物物理研究所，北京 100101)

摘要 报道了以二聚体存在的 dimo-BmK M1 的 1.4 Å 分辨率晶体结构. 蛋白质中的肽键是局部双键，不可旋转，因此具有顺式 (*cis*) 和反式 (*trans*) 两种构型，它们不能通过旋转操作相互转换. 非脯氨酸顺式肽键是指形成该肽键的氨基是由脯氨酸以外的氨基酸提供的 (Xaa-nonPro)，这类肽键的顺式构型的自由能远比反式高，因此极少出现在天然蛋白质结构中. 事实上，在长时间中，多肽链的“反式肽键连接”被视为蛋白质结构的一条基本规则，把顺式肽键视为不可能. 随着高分辨率精确蛋白质结构数量的增加，近年来有详细的统计分析揭示，非脯氨酸顺式肽键 (Xaa-nPro) 在蛋白质结构中出现的几率为 0.03%～0.05%，而且大多存在于功能敏感的结构区域，可能具有重要意义. 但由于所用的基本结构数据都来自晶体结构，对这种反常肽键是否由结晶环境影响而形成，存在疑问. 此前曾在以单体形式存在的蝎神经毒素 mono-BmK M1 的高分辨率结构中发现其中肽键 Pro9-His10 是非脯氨酸顺式肽键，并详细分析了其结构 - 功能意义. 以二聚体存在的 dimo-BmK M1 的 1.4 Å 分辨率晶体结构表明，它与 mono-BmK M1 有不同的空间群、不同的分子堆积方式，不同的晶体环境. 结构模型被高度精化，R_{cryst} 达到 0.109. dimo-BmK M1 结构显示，在不对称单位中的两个 M1 分子在同一位置 (残基 9-10 之间) 都清晰地存在顺式肽键. 立体化学分析显示，这一肽键的几何参数和局部结构与 mono-BmK M1 中的 (9-10) 顺式肽键基本相同. 这一结果表明，非脯氨酸顺式肽键 9-10 的存在与结晶环境无关，是 BmK M1 分子的固有结构特征. 在此基础上，综合分析了与顺式、反式肽键相关的结构元素，发现与残基 (8-19) 序列模体 -KPXNC- (X 为任意氨基酸) 所决定的特征回折结构可能是分子内在的主要结构因素，其中第 8 位残基是 Lys 或 Asp 对决定肽键是顺式还是反式有关键作用. 近来的突变实验及其晶体结构测定已证实，Lys8/Asp8 是 (9-10) 肽键顺式 / 反式异构的结构开关，它们对该类分子与不同种属钠通道作用的专一选择性具有重要作用. 通过 BLAST 搜索，发现在其他 18 个蛋白质中也存在相同的序列模体 -KPXNC-，推测在这些蛋白质的相应肽键位置也可能存在反常的脯氨酸顺式肽键.

关键词 非脯氨酸顺式肽键，晶体结构，神经毒素，蝎 *Buthus martensii* Karsch
学科分类号 Q617

---

*国家自然科学基金(30370320)和中国科学院知识创新工程 (KSCX1-SW-17, KSCX2-SW-322) 资助项目.
并列第一作者. * 通讯联系人. Tel: 010-64888547, Fax: 010-64888560, E-mail: dcwang@ibp.ac.cn
收稿日期：2005-10-09，接受日期：2005-12-30