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

Rong-Jin Guan1†, Ye Xiang1†, Xiao-Lin He1, Chun-Guang Wang1
Miao Wang1, Ying Zhang1, Eric J. Sundberg2 and Da-Cheng Wang1*

1Center for Structural and Molecular Biology, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101 People’s Republic of China
2Center for Advanced Research in Biotechnology, W.M. Keck Laboratory for Structural Biology, University of Maryland Biotechnology Institute, Rockville, MD 20850 USA

Non-proline *cis* peptide bonds have been observed in numerous protein crystal structures even though the energetic barrier to this conformation is significant and no non-prolyl-*cis/trans*-isomerase has been identified to date. While some external factors, such as metal binding or co-factor interaction, have been identified that appear to induce *cis/trans* isomerization of non-proline peptide bonds, the intrinsic structural basis for their existence and the mechanism governing *cis/trans* isomerization in proteins remains poorly understood. Here, we report the crystal structure of a newly isolated neurotoxin, the scorpion *α*-like toxin *Buthus martensii* Karsch (BmK) M7, at 1.4 Å resolution. BmK M7 crystallizes as a dimer in which the identical non-proline *cis* peptide bond between residues 9 and 10 exists either in the *cis* conformation or as a mixture of *cis* and *trans* conformations in either monomer. We also determined the crystal structures of several mutants of BmK M1, a representative scorpion *α*-like toxin that contains an identical non-proline *cis* peptide bond as that observed in BmK M7, in which residues within or neighboring the *cis* peptide bond were altered. Substitution of an aspartic acid residue for lysine at residue 8 in the BmK M1 (K8D) mutant converted the *cis* form of the non-proline peptide bond 9–10 into the *trans* form, revealing an intramolecular switch for *cis*-to-*trans* isomerization. *Cis/trans* interconversion of the switch residue at position 8 appears to be sequence-dependent as the peptide bond between residues 9 and 10 retains its wild-type *cis* conformation in the BmK M1 (K8Q) mutant structure. The structural interconversion of the isomeric states of the BmK M1 non-proline *cis* peptide bond may relate to the conversion of the scorpion *α*-toxins subgroups.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: *cis/trans* isomerization; non-proline peptide bond; structural mechanism; intramolecular switch; crystal structure

*Corresponding author

**Introduction**

The partial double-bond character of the peptide bond in proteins restricts its rotation such that consecutive Cα atoms exhibit coplanarity, and consequently, only *trans* and *cis* conformations are possible. The majority of peptide bonds adopt the *trans* conformation, as it is intrinsically favored energetically,1,2 with the primary exception being peptide bonds formed between any amino acid and proline (Xaa–Pro), which have been observed in the *cis* configuration with greater frequency.3 The observation of a *cis* peptide bond, especially of the non-proline *cis* variety, has been generally noted with great emphasis in the literature. Surveys of several databases3−5 of protein structures, however,

† R.-J.G. & Y.X. contributed equally to this work.
Abbreviations used: Aah, *Androctomus australis* Hector; BmK, *Buthus martensii* Karsch; CT, C-terminal segment, residues 63–66; rBmK, recombinant BmK; RT, reverse turn residues, 8–12.
E-mail address of the corresponding author: dchwang@sun5.ibp.ac.cn

0022-2836/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
have revealed that the frequency of non-proline cis peptide bonds approaches one in every two thousand instances. Additionally, non-proline cis peptide bonds unequivocally identified in protein structures occur at or near functionally important sites more often than if their spatial distribution was random.\textsuperscript{5,6} Renewed awareness of non-proline cis peptide bonds has led to a series of detailed analyses of their structural characteristics\textsuperscript{7–7} and functional roles.\textsuperscript{8–12} The structural mechanisms by which peptide bonds are converted from the energetically preferred trans conformation to that of the less favorable cis conformation, however, have received less attention.

Structural analyses of several proteins have shown that transformation of non-proline peptide bonds from trans to cis conformations can be induced by intermolecular binding events involving specific metals or co-factors that result in isomer-specific functional states.\textsuperscript{8–11} The cis–trans isomerization of non-proline peptide bonds has also been observed in protein unfolding and refolding events in the absence of binding partners,\textsuperscript{7,13–15} revealing that structural properties intrinsic to the protein can control the peptide bond conformations. The questions of how non-proline cis peptide bonds attain their energetically unfavorable conformation and which specific intramolecular structural elements are responsible for cis and trans conformations and their isomerization, however, have yet to be addressed rigorously. Here, we describe structural factors that govern the occurrences of cis and trans conformations of a non-proline peptide bond, as well as their co-existence, intrinsic to a class of scorpion neurotoxins. We also identify the intramolecular switch for cis-to-trans isomerization of this non-proline peptide bond.

Scorpion neurotoxins are known to interact specifically with voltage-dependent sodium channels.\textsuperscript{6,17} Three phylogenetically distinct groups have been identified, classical \(\alpha\)-, \(\alpha\)-like and \(\alpha\)-insect toxins, which are highly toxic to mammals, insects and both insects and mammals, respectively.\textsuperscript{18} The scorpion \textit{Buthus martensii} Karsch (BmK) produces numerous neurotoxins, including BmK M1, M7 (both \(\alpha\)-like) and M8 (classical \(\alpha\)). The crystal structures of BmK M8\textsuperscript{19} and M1\textsuperscript{20} have revealed that the non-proline peptide bond between residues Pro9 and His10 exists in the cis conformation in BmK M1 but in the trans conformation in BmK M8, similar to other representative classical \(\alpha\)-toxins such as Aah2 (\textit{Androctonum australis} Hector toxin 2).\textsuperscript{21} This peptide bond is located within a five-residue reverse turn motif that contacts the C-terminal residue when it adopts the cis conformation in BmK M1, but makes no such intramolecular contacts when in the trans conformation in BmK M8.

Here, we describe crystal structures of BmK M7, which is unique in both its dimeric form and its extended C terminus, as well as of several mutants of BmK M1. In these structures, both the cis and trans conformations are observed for the non-proline peptide bond 9–10, as well as their co-existence within a monomer in the asymmetric unit. We show that mutation at a site proximal to the non-proline cis peptide bond can act as an intramolecular switch for cis-to-trans isomerization. Furthermore, receptor specificity of these wild-type and variant scorpion toxins appears to be controlled by the backbone geometry of this particular non-proline peptide bond.

**Results**

**Crystallographic sequencing of BmK M7**

BmK M7 is an \(\alpha\)-like toxin from the scorpion \textit{Buthus martensii} Karsch (BmK), which resides in eastern Asia. It is comprised of 66 amino acid residues cross-linked by four disulfide bridges. Compared with most scorpion \(\alpha\)-toxins containing 64 residues, which have only one residue after the terminal disulfide bridge Cys63–Cys12, BmK M7 is unique in having three C-terminal residues (Figure 1). The entire 66-residue sequence of BmK M7 was determined by crystallographic analysis and is shown in Figure 1. At 1.4\(\text{\AA}\) resolution, the electron density defining most residues is unambiguous. Loop residues 39–43, which exhibited poor electron density, were defined with the aid of homologous sequences from other BmK toxins (BmK M1, M2 and M4; Figure 1), as well as the experimentally determined molecular mass of 7237.4 from MALDI-TOF mass spectroscopic analysis. Glu/Gln and Asp/Asn residues were differentiated according to both the chemical environment as well as the homologous BmK toxin sequences.

**Overall structure of BmK M7 dimer**

The refined BmK M7 model contains two molecules in the asymmetric unit, related to one another by non-crystallographic 2-fold symmetry\textsuperscript{22} (~179.8°), that form a homodimer through nine hydrogen bonds and numerous van der Waals contacts (Table 1). The structure of BmK M7 is the first scorpion toxin observed as a dimer. The protein regions associated with the non-proline cis peptide bond between residues Pro9 and His10, including the five residue reverse turn comprised of residues 8–12 and the neighboring C-terminal residues 63–66, are not involved in dimer formation. The dimer interface in BmK M7 displays a high degree of shape complementary, exhibiting a shape correlation coefficient (\(S_S\)) of 0.795 \((S_S = 1.0\) for interfaces with geometrically perfect fits\textsuperscript{23}) and a large buried surface area of 911 \(\text{\AA}^2\), relative to the overall surface area of a 66-residue protein.

The two monomers in the BmK M7 dimer both adopt a general fold similar to the BmK M1, a representative BmK \(\alpha\)-like scorpion toxin. All are composed of a dense core of secondary structural elements, including an \(\alpha\)-helix formed by residues 19–28 and a three-stranded antiparallel \(\beta\)-sheet formed by residues 2–5, 32–37 and 45–51.
Figure 1. Sequence of BmK M7 and structure-based alignment of the amino acid sequences of different subgroups of scorpion alpha toxins. The positions are numbered in terms of BmK M1. Secondary structure elements are denoted on top of the sequences. The highly conserved cysteine residues are marked with dark in background. The residues of the unique site RT-CT formed by a five-residue reverse turn, in which a cis peptide bond occurs between residues 9 and 10, and the C-terminal segment are boxed. BmK M7 conspicuously has two additional residues at the C terminus. The sequences are from Possani et al. and Hutchinson et al. Sequence alignment was performed with CLUSTAL W and the Figure was prepared with ALSCRIPT.
Table 1. Interaction on the monomer–monomer interface in BmK M7

<table>
<thead>
<tr>
<th>Hydrogen bonds</th>
<th>N19A O¹</th>
<th>I38B O</th>
<th>3.12 Å</th>
</tr>
</thead>
<tbody>
<tr>
<td>C36A O</td>
<td>I38B N</td>
<td>3.00 Å</td>
<td></td>
</tr>
<tr>
<td>N37A O¹</td>
<td>N19B O¹</td>
<td>2.87 Å</td>
<td></td>
</tr>
<tr>
<td>N37A O¹</td>
<td>N37B O¹</td>
<td>2.94 Å</td>
<td></td>
</tr>
<tr>
<td>N37A N²</td>
<td>N19B N²</td>
<td>3.12 Å</td>
<td></td>
</tr>
<tr>
<td>I38A N</td>
<td>C36B O</td>
<td>2.98 Å</td>
<td></td>
</tr>
<tr>
<td>N44A O¹</td>
<td>N44B N²</td>
<td>2.75 Å</td>
<td></td>
</tr>
<tr>
<td>W47A N¹</td>
<td>Y35B OH</td>
<td>3.05 Å</td>
<td></td>
</tr>
</tbody>
</table>

Hydrogen bonds were calculated using a cut-off of 2.5–3.5 Å.

During refinement it was impossible to restrain the ω value of the peptide bond Pro9–His10 in molecule A near ±180°, whereas the ω angle went favorably towards 0°. Both 2F_o−F_c and F_o−F_c electron density maps (Figure 3(A) and (B)) definitively show that this peptide bond adopts a cis conformation in which the amide nitrogen is provided by an amino acid (His10) other than proline, classifying this as a non-proline cis peptide bond. Moreover, the C-terminal segment in BmK M7, which extends beyond other α-like toxins, exhibits an ordered conformation with clearly defined electron density (Figure 3(A)). Three main-chain hydrogen bonds between this segment and residues 9 and 10 were observed (Figure 5(A)). These local tertiary interactions likely play a crucial role in stabilization of the energetically unfavorable cis conformation of the peptide bond 9–10. An intermolecular contact between the main-chain O of the terminal residue Ala66 and the N^\text{z} atom of residue Arg58 from a symmetry-related molecule (Figures 2 and 5(A)), may provide energetic stabilization for the C terminus conformation.
Co-existence of cis and trans isomers in molecule B of BmK M7

In the dimer structure of BmK M7, the peptide bond Pro9–His10 in molecule B exhibits a cis/trans dual-configuration (Figure 4), rather than the cis conformation observed in molecule A. Throughout the refinement, electron densities around the Pro9–His10 peptide bond in molecule B appeared relatively broad, and could not be fitted exclusively by either cis or trans conformations. During the earlier stages of the refinement, no special restriction was imposed on this peptide bond. After several cycles of model rebuilding and refinement, the quality of the model and the density maps were greatly improved (with $R_{\text{free}}$ below 0.20). An unbiased difference ($F_o - F_c$) electron density map was calculated prior to inclusion of either cis or trans conformations for the peptide bond 9–10, which showed a co-existence of the two conformations.

In order to identify whether the cis and trans forms existed coincidentally at the peptide 9–10, two $\sigma$-weighted $F_o - F_c$ electron density maps were further calculated. The first calculation included the cis peptide bond prior to inclusion of the trans configuration, into which a trans conformation model can be built into positive electron density while a cis peptide bond fits well into the negative density (Figure 4(B)). Conversely, when an $F_o - F_c$ map is calculated using a model for the trans conformation, the cis peptide bond 9–10 can be built into the positive density, while a trans conformation can be built into negative density (Figure 4(C)).

The final $2F_o - F_c$ and $F_o - F_c$ electron density maps reveal that the C-terminal segment comprised of residues 64–66 is disordered (Figure 4(D)). All of the intramolecular contacts between the C-terminal residues and the peptide bond 9–10 observed in
Figure 4. Electron densities around residues 9 and 10 and 63–66 in molecule B of BmK M7 showing the co-existence of
molecule A are absent in molecule B, providing a rationale for the structural basis for the existence of strict cis conformation of this peptide bond versus a mixed population.

In the final stage of refinement, the dual cis and trans conformations for residues Pro9 and His10 (with atomic occupancy 0.5) were built into the model of molecule B, resulting in increases in both $R_{	ext{cryst}}$ and $R_{	ext{free}}$ from 0.1679 to 0.1669 and from 0.1864 to 0.1840, respectively. This implies that the inclusion of the dual conformation of the peptide 9–10 in molecule B is closer to the actual BmK M7 structure and co-existence of cis and trans conformations in the peptide bond 9–10 is objective.

Expression, purification and bioassay of BmK M1 mutants

To explore the key structural factors responsible for the cis/trans conversion of the peptide bond between positions 9 and 10, all residues in the vicinity of the cis-peptide bond are taken into consideration during the mutagenesis design, including residues 8, 9, 10 and 11. However, structural and sequence comparisons clearly show that the residue type of position 10 varies irregularly, which can be the same residue (Glu) in both cis- and trans-containing toxins, and also the different types in the same cis-containing toxins (His or Glu) or the same trans-containing toxins (Val or Glu or Lys). For residue Asn11, it is invariant in either cis or trans form toxins (Figure 1). Therefore, residues 10 and 11 should not be related to the cis or trans form of the peptide bond 9–10. At the same time residue 8 is conserved as Asp in all trans form toxins and residue 9 is conserved as Pro8 in all cis form toxins (Figure 1). Based on these analyses, residues 10 and 11 are set aside from the mutagenesis considerations and the emphasis is focused on residues 8 and 9.

Five mutants of BmK M1 were constructed, including mutations of the flanking Lys residue at position 8 with Asp (K8D), Asn (K8N), Glu (K8E) and Gln (K8Q), as well as replacement of Pro at position 9 with Ser (P9S). Of these, the K8N and K8E mutants did not express in our recombinant system. The molecular masses of the three other variants, including mutations of the flanking Lys residue at position 9 with Ser (P9S), were measured by mass spectrometry and found to agree well with their respective theoretical values.

The LD$_{50}$ of wild-type BmK M1 determined by the method described by Meier & Theakston was ~0.53 mg/kg, which is consistent with that of native BmK M1. Excluding K8N and K8E, which were not expressed, the other three mutants (K8D, K8Q and P9S) were used for bioassays. The toxicity of mutant K8D was lost in mice, while mutant K8Q was almost unchanged in comparison with unmodified rBmK M1 (Table 2). Mutant P9S exhibited the toxicity in the same order (55%).

Cis–trans conversion of the peptide bond 9–10 in the BmK M1 K8D mutant

The structure of mutant K8D at 1.5 Å resolution clearly revealed that the non-proline cis peptide bond between residues 9 and 10 in the native BmK M1 was converted to the trans conformation in the K8D mutant. Correspondingly, both $2F_o-F_c$ and $F_o-F_c$ electron density maps show the C$\alpha$ atoms of residues 9 and 10 in K8D located on opposite sides of the 9–10 peptide bond. Accompanying the cis–trans conversion, both the N main-chain atom of His10 and the side-chain of Asp8 are rotated into the reverse turn relative to wild-type BmK M1. These changes force the five-residue reverse turn (residues 8–12) to adopt a new conformational state, in which the residues at positions 8 and 10 form an extensive hydrogen-bonding network (Figure 6), characteristic of the trans peptide bond-containing reverse turn in classical α scorpion toxins.

Expression and toxicity of the five BmK M1 mutants

<table>
<thead>
<tr>
<th>Toxins</th>
<th>Expression (mg/l)</th>
<th>Toxicity (LD$_{50}$) (mg/kg)</th>
<th>Relative toxicity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild-type</td>
<td>~5</td>
<td>0.53</td>
<td>100</td>
</tr>
<tr>
<td>rBmK M1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K8D</td>
<td>10–11</td>
<td>&gt;50</td>
<td>&lt;1</td>
</tr>
<tr>
<td>K8N</td>
<td>None</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>K8E</td>
<td>None</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>K8Q</td>
<td>5–6</td>
<td>0.50</td>
<td>106</td>
</tr>
<tr>
<td>P9S</td>
<td>4–5</td>
<td>0.95</td>
<td>55</td>
</tr>
</tbody>
</table>

Table 2. Expression and toxicity of the five BmK M1 mutants
Cis form of the peptide bond 9–10 in mutants K8Q and P9S

Contrary to the K8D mutant, replacement of residue Lys8 with Gln results in retention of the cis conformation of the peptide bond 9–10 found in wild-type BmK M1. Consequently, the intramolecular contacts between residues from the reverse turn and the C terminal are similar to those in wild-type BmK M1 (Figure 6). The structure of mutant P9S also revealed a cis conformation for the peptide bond 9–10. Sequence alignment indicates that Pro9 is conserved in cis-containing toxins (Figure 1). The structure of the P9S mutant indicates that the Pro residue at position 9 is not required for the peptide bond 9–10 to adopt the cis conformation. Lack of expression of the K8E and K8N mutants may indicate the importance of residue 8 not only in governing the conformational state of the five-residue reverse turn in the folded molecule, but also in subsequent correct global folding of the toxin molecule.

Discussion

Structural properties of differential peptide bond 9–10 isomeric states

Surveys of structural databases3–5 have revealed the occurrence and biological importance of non-proline cis peptide bonds in protein structures. Detailed analyses of the conformational characteristics and functional roles of these unusual peptide bonds have been reported.8–12 External structural factors, such as substrate binding, metal binding, or co-factor interaction, have been identified as inducers of the cis/trans isomerization. How non-proline peptide bonds attain their distinct cis/trans isomeric states and which structural factors specifically govern the occurrence of different isomers, however, have yet to be addressed rigorously. The observations here reveal one way by which the specific intramolecular structural factors govern a non-proline peptide bond to attain cis or trans conformations, or cis/trans co-existence.

In the crystal structures of dimeric BmK M7 and three BmK M1 mutants (K8D, K8Q and P9S) three distinct isomeric states of the non-proline peptide bond 9–10 are observed, including cis (BmK M7 molecule A and BmK M1 mutants K8Q and P9S), trans (BmK M1 mutant K8D), and the co-existence of cis and trans (BmK M7 molecule B) conformations. Detailed inspection and comparison of these structures reveal that these isomeric states are dependent on the distinct tertiary arrangement of the five-residue reverse turn (residues 8–12) and the C-terminal segment (residues 63–66). In the cis conformation, both the main-chain O and N atoms of residues 9 and 10, respectively, extend away from the reverse turn and interact with the C-terminal residue 64 via a hydrogen bond between the main-chain N and O atoms of residues 10 and 64, respectively. This results in a protrusion of the side-chain of residue 8 from the reverse turn. In the trans conformation, however, the main-chain N atom of residue 10 turns into the core of the reverse turn and, correspondingly, interacts with the side-chain of residue 8, giving rise to the extensive hydrogen-bonding network of the reverse-turn structure (Figure 6). This results in the abrogation of the interaction between the peptide bond 9–10 and C terminus. When cis and trans conformations co-exist within a single molecule, the main-chain N atom of residue 10 contacts neither a C-terminal residue nor the side-chain of residue 8 (Figure 5), resulting in a novel reverse turn structure that cannot be classified as one of the canonical β-turn defined by Hutchinson & Thornton.23

Regardless of the peptide bond 9–10 heterogeneity exhibited in these structures, the reverse turn motifs in all BmK toxins share a number of common structural features. The final two turn residues, Asn11 and Cys12, are strictly conserved (Figure 1) and form multiple contacts with the C-terminal segment in all scorpion α-toxin structures determined so far,20 including a disulfide bond. The highly constrained reverse turn-C terminus (RT-CT) structural motif accounts for the tendency of the cis conformation, isomerization to the trans conformation and the co-existence of both isomers for the peptide bond 9–10 in BmK toxins. The structures reported here indicate that the cis conformation of the peptide bond 9–10 is energetically favorable to the structural organization of the RT-CT motif. In this case, the myriad interactions between the reverse turn in which the peptide bond 9–10 resides and the C-terminal segment appear capable of overcoming the normally energetically favored trans peptide bond conformation. Conversely, altered interactions within the core of the reverse turn arising from mutation of Lys at position 8 to Asp in BmK M1 invert the energetic parameters as to allow the K8D mutant to adopt the trans conformation.

The conformational state of the main-chain N atom of residue 10 is the structural signature of the distinct isomeric states of the peptide bond 9–10. When this atom protrudes into the reverse turn and interacts with the Asp side-chain in the K8D mutant, the peptide bond adopts the trans conformation. When it instead protrudes out of the reverse turn and interacts with the main-chain carbonyl of the C-terminal residue 64 it is found in the cis conformation (Figure 6). In the absence of intramolecular atomic interactions for the residue 10 N atom, the peptide bond 9–10 co-exists in both the cis and trans isomers (Figure 5(B)). Although additional interactions, such as between main-chain atoms of residues 9 and 10 with those from the C-terminal segment, are present in this region of the molecule and may play a role in stabilization of the local RT-CT structural motif, they appear insufficient to control the isomeric state of the peptide bond 9–10.
An intramolecular switch governs peptide bond isomerization

The molecular switch for cis/trans peptide bond isomerization at position 8 of BmK M1 is sequence-dependent. As described above, mutation of the wild-type Lys residue at this position to Asp (K8D) results in a cis-to-trans isomerization of the peptide bond, while mutation to Gln (K8Q) maintains the cis conformation of the wild-type protein. Our attempts to express BmK M1 mutants with similar amino acid changes at position 8, including Glu (K8E) and Asn (K8N), were unsuccessful (Table 2). This suggests the intriguing possibility that these mutants represent a third outcome of the molecular switch for peptide bond isomerization at position 8, which results in a local structure that is incompatible with the global folding of the protein. As the five-residue reverse turn in which this molecular switch resides is, respectively, conserved in different groups of scorpion α-toxin (Figure 1), its sequence-dependent nature may also extend to other members of this toxin family. Thus, it can be predicted that replacing residue 8 in other scorpion

Figure 5. Structure of the five-residue reverse turn and C-terminal segment in M7-A (A) and M7-B (B). In M7-A the peptide bond 9–10 adopts cis form and both NH10 and CO9 groups of the peptide are out of the reverse turn to interact with the C-terminal residues. In M7-B the group NH10 contacts neither residue 8 nor C-terminal residues in flexible state and, thus, the peptide bond 9–10 takes cis/trans co-existence. For clarity the side-chains of residues 8–12 are only shown corresponding to the cis form of the peptide 9–10, with deletion of His10 not shown in the Figures.
Figure 6. Distinct structures of the five-residue reverse turn and C-terminal residues of the *trans*-containing site RT-CT in BmK M1 K8D mutant (A) and native BmK M8 (B), and the *cis*-containing RT-CT in BmK M1 K8Q mutant (C) and native BmK M1 (D). In the *trans*-containing form (A) and (B), the peptide group NH10 is situated inside the turn and the residue 8 must be Asp; these two interact with each other via hydrogen bond N10...O\text{d}18, but there is no contact between the NH10 group and the C-terminal residue. In *cis*-containing form (C) and (D) group NH10 is out of the reverse turn and interacts with the C-terminal residue via hydrogen bond N10...O64. Meanwhile the residue 8 is non-aspartic and protrudes from the turn. In this case the reverse turn is well stabilized by two main-chain hydrogen bonds between residues 8 and 10. For clarity the side-chain of residue 10 is not shown in all Figures. The Figures were prepared with MOLSCRIPT.42
toxins with Asp, Lys/Gln, or Asn/Glu will result in a peptide bond 9–10 adopting a trans, cis, or misfolded protein, respectively. It is possible to experimentally validate this prediction.

Possible biological significance of variable RT-CT structural motif conformational states

Variable conformations of the peptide bond 9–10 result in three unique structural organizations of the RT-CT motif. In the first motif, termed transRT-farCT and exemplified by BmK M8 and Aah2 (Figure 7(A)), the trans peptide bond 9–10 mediates an extensive hydrogen bonding network internal to the RT and a paucity of intramolecular interactions with the CT residue 64. The second motif, cisRT-closeCT (Figure 7(B)), is characterized by the cis-containing RT structures of BmK M7 molecule A, BmK M1 and associated mutants K8Q and P9S, in which extension of the main-chain amine and carbonyl groups of the peptide bond 9–10 to the periphery of the RT facilitates intimate contact with the CT segment. The third motif, transRT-closeCT (Figure 7(C)), exhibited by the BmK M1 K8D mutant, is a fusion of an RT structure common to the trans peptide bond 9–10 and the interacting CT segment similar to toxins exhibiting cis-containing RT structures. The unique structures of the RT-CT motif are stabilized by the disulfide bridge Cys12–Cys63, which exhibits structural variability amongst different toxins, such as BmK M8, M1 and Aah2.

These three RT-CT conformational motifs provide a possible structural basis for the binding preference of three α-toxin subgroups for phylogenetically distinct, yet closely related, target sites on the sodium channel. The pathogenicity of scorpion α-toxins is caused by binding to the fourth domain of the α-subunit of the target sodium channel, resulting in a prolonged action potential due to retardation of the inactivation event. Scorpion α-toxins can be classified into pharmacological subgroups according to their preferential toxicity to mammals or insect. The evidence supports their binding to a homologous cluster of partially overlapping sites located on the sodium channel surfaces in both mammals and insects.

The functional residue Arg58 in these structures is solvent-accessible and able to engage ligand. Mutation of Lys8 to Asp functions not only as a molecular switch in isomerizing the peptide bond 9–10 from the cis to trans conformation and consequent adoption of the transRT-closeCT motif (Figure 7(C)), but also as a switch for biological function. The bioassay revealed that the toxic effects of this mutant have been lost in mice (Table 2), while electrophysiological characterization using the cloned Para/tipE insect sodium channel showed that the BmK M1 K8D mutant retained its functionality responsible for insect toxicity (J. Tytget, personal communication). Accordingly, the conformational state of the transRT-closeCT motif in α-like toxins may also be related to sodium channel binding site selectivity. These observations show one way in which high levels of molecular specificity and biological function may be achieved primarily through cis/trans isomerization of a critical peptide bond.

Comparison of the conversion mechanism for cis/trans isomerization of BmK toxins with that of other proteins

Structural analyses of several proteins have identified that conversion of non-proline peptide bonds from trans to cis conformations can be induced by external structural factors, such as metal binding, substrate binding or cofactor interaction that result in isomer-specific functional states.

In concanavalin A (ConA), binding of a metal ion in the binding site S2 is the driving force to initiate the trans-to-cis isomerization of the Ala207–Asp208 peptide bond, resulting in a locked state of ConA. While metal ions are released from the binding site the cis peptide bond is destabilized and the structure tends to refold with a trans (Ala207–Asp208) peptide. In hypoxanthine-guanine phosphoribosyltransferase (HGPR) the Leu78–Lys79 peptide bond changes its geometric status during the course of catalysis, in which cis/trans isomerization of this peptide bond is related to the substrate binding (trans to cis) and product release (cis to trans), respectively. The authors proposed that a portion of the energy released upon substrate binding to the apoenzyme is used to drive the Leu78–Lys79 peptide bond into the cis configuration, and that the energy released upon isomerization of the cis peptide bond back to the trans ground state helps to propel the substrate out of the active site. The similar mechanism is reported in oxidation–reduction procedures of flavodoxin. In these reports cis/trans conversion of non-proline peptide bond is governed by intermolecular binding events involving substrates, metal ions or cofactors. However, the present study revealed that the conversion of cis and trans isomeric states of a non-proline peptide bond in scorpion BmK toxins is governed by intramolecular structural factors as described above. A local unique structural motif (RT-CT) accounts for the tendency of the...
Figure 7. Three conformational states of site RT-CT formed by the five-residue reverse turn (in dark green) in association with C-terminal residues (in red) observed in scorpion α-toxins. (A) transRT-farCT observed in BmK M8 (thick lines) and Aah2 (thin lines), in which the NH group of peptide bond 9–10 is situated inside the turn in compiling with the side-chain of the specific Asp8 and the C-terminal segment is far away from the peptide without any contact in between. (B) cisRT-closeCT observed in BmK M1 (thick lines) and M7 (thin lines), in which both NH and CO groups of the cis peptide bond 9–10 are out of the reverse turn and the C-terminal segment is connect to the peptide via hydrogen bond N10–O64, in turn, close to the cis peptide bond 9–10. (C) transRT-closeCT observed in BmK M1 K8D mutant, in which the reverse turn adopts trans-containing organization but the orientation of the C-terminal residue is still close to the peptide bond 9–10 though there is no contact in between.
cis conformation, isomerization to the trans conformation and the co-existence of both isomers for the peptide bond 9–10. The cis/trans conversion is controlled by specific residues at position 8, which would result in the peptide 9–10 adopting a trans, cis or misfolded protein when it is replaced with Asp, Lys/Gln or Asn/Glu, respectively. All these properties found in BmK toxins showed a novel structural mechanism for cis/trans isomerization of a non-proline peptide bond, which is distinct from that observed in some other proteins, such as ConA, HGPRT and flavodoxin.

It was reported that many non-proline cis peptide bonds contain an aromatic residue and the presumed reason for this was the occurrence of an aliphatic–aromatic interaction from C–H to the Pi-system. The aliphatic–aromatic interaction is an assumed reason for this was the occurrence of an aliphatic–aromatic interaction from C–H to the Pi-system. The aliphatic–aromatic interaction is an important factor for the stabilization of non-proline cis peptide bonds and may influence the cis/trans isomerization of these special peptide bonds. In principle, His10 of BmK M1 may accept such an interaction. However, careful examination of the contacts between residues 9 and 10 in Bmk M1 and other cis-containing BmK toxins showed that there is no such an aliphatic–aromatic interaction as observed at non-proline cis peptide bonds in some other proteins. In fact, a homology toxin of BmK M1, BmK M4, in which residue 10 is Glu other than His, also showed a cis form of peptide bond 9–10. It seems that the stabilization of the cis form in BmK toxins is not dependent on the close aliphatic–aromatic contacts, but the tertiary interaction between the cis peptide bond and the C-terminal segment.

A series of experimental data suggest that the cis/trans isomerization found in BmK toxins is pH-independent. Actually the crystal structures of BmK M1, M2, M4 and mutant (K8Q) obtained in a wide range of pH values (from pH 4.6–8.5) can present the same cis form of peptide bond 9–10. The toxins with either His (BmK M1) or Glu (BmK M4) at position 10 can adopt the cis peptide bond, which seems not relevant to the protonating status of His10. The mutant K8D crystallized in a range of pH 5–8 can present the same conformation as trans. Besides, it is definitely known that the co-existence of cis and trans conformations is due to the abrogation of the interaction between the peptide bond 9–10 and C terminus, but not related to the side-chain interactions. All these data indicate that the cis/trans isomerization observed in BmK toxins is pH-independent. Actually the distinct isomeric states of the peptide bond 9–10 in BmK toxins are mainly governed by the specific intramolecular interactions and the molecular switch at position 8 is sequence-dependent rather than pH-dependent.

Experimental Procedures

Protein production

BmK M7 was purified from BmK venom by gel-filtration and ion-exchange chromatography as described. The molecular mass, isoelectric point and toxicity of isolated BmK M7 were confirmed prior to crystallization. Five mutants of BmK M1 (Lys8Asp (K8D), Lys8Asn (K8N), Lys8Glu (K8E), Lys8Gln (K8Q) and Pro9Ser (P9S)) were produced by one-step PCR using synthetic primers: primer1 (5’-CGCTATAGAAAA GAATTTCTGTCGGGATGCTTATATTGCC-3’ (K8N), 5’-CGCTAAGATAAAAGAAATTC TGTTTCGGGATGCTTATA TTGCCCAATCCCATAAC-3’ (K8Q), 5’-CGCTAGATAAAAGAAATTC TGTTTCGGGATGCTTATA TTGCCCAATCCCATAAC-3’ (K8E), 5’-CGCTAGATAAAAGAAATTC TGTTTCGGGATGCTTATA TTGCCCAATCCCATAAC-3’ (K8D), BmK M4, in which residue 10 is Glu other than His, also showed a cis form of peptide bond 9–10. It seems that the stabilization of the cis form in BmK toxins is not dependent on the close aliphatic–aromatic contacts, but the tertiary interaction between the cis peptide bond and the C-terminal segment.

A series of experimental data suggest that the cis/trans isomerization found in BmK toxins is pH-independent. Actually the crystal structures of BmK M1, M2, M4 and mutant (K8Q) obtained in a wide range of pH values (from pH 4.6–8.5) can present the same cis form of peptide bond 9–10. The toxins with either His (BmK M1) or Glu (BmK M4) at position 10 can adopt the cis peptide bond, which seems not relevant to the protonating status of His10. The mutant K8D crystallized in a range of pH 5–8 can present the same conformation as trans. Besides, it is definitely known that the co-existence of cis and trans conformations is due to the abrogation of the interaction between the peptide bond 9–10 and C terminus, but not related to the side-chain interactions. All these data indicate that the cis/trans isomerization observed in BmK toxins is pH-independent. Actually the distinct isomeric states of the peptide bond 9–10 in BmK toxins are mainly governed by the specific intramolecular interactions and the molecular switch at position 8 is sequence-dependent rather than pH-dependent.

Biossay of BmK M1 mutants

Using 0.9% (w/v) NaCl as a negative control and rBmK M1 as a positive control, the toxicity of the mutants was determined in mice (male, specified pathogen-free level, 18–20 g of body weight). Each group consisted of ten mice. Various doses of toxin mutants were dissolved in 0.9% NaCl and injected into the mice through the tail vein. Survival times (times between injection and death), reaction, and doses were recorded. Evaluation of toxicity was based on the determination of LD50 (the dose capable of statistically killing 50% of the mice) as defined in the method described by Meier & Theakston. The final purified samples were examined by ESI mass spectrometry using an API 3000 triple quadrupole mass spectrometer (Perkin-Elmer Sciex Instruments, Canada).

Crystallization and data collection

BmK M7 was crystallized at room temperature by the hanging-drop vapor-diffusion method using equal volumes of BmK M7 (10 mg/ml) and 0.65 M ammonium sulfate, 100 mM Tris–HCl (pH 8.5) using 1% (v/v) ethanol as an additive. The largest crystals grew to 0.5 mm within one month. The K8D, K8Q and P9S mutants of BmK M1 were crystallized similarly to BmK M7, with the exception of the crystallization buffer and additives used. These were 1.5 M sodium phosphates at pH 6.0 for K8D, 3.2 M sodium phosphate at pH 4.6 and 5% (v/v) PEG400 as an additive for P9S and 30% (v/v) PEG8000 and 0.1 M sodium cacodylate at pH 6.5 with 0.2 M ammonium sulfate as an additive for K8Q.

Diffraction data of BmK M7 and the K8D and P9S BmK M1 mutants were collected at room temperature using synchrotron radiation (λ = 1.0 Å) on an ADSC Quantum 4 CCD detector at the beam line BL18B of the Photon Factory in Tsukuba, Japan. For BmK M7, the crystal could diffract to at least 1.4 Å. Diffraction data of mutant K8Q
were collected on a Mar345 image-plate detector using CuKα radiation (λ = 1.5418 Å). The diffraction data were processed and analyzed by using DPCS/MOSFLM/CCP434–38 and the HKL suite of programs. The results showed that there are two monomers in the asymmetric unit of BmK M7, but only one in those of each of the three BmK M1 mutants. The crystal parameters and data collection statistics are listed in Table 3.

**Structure determination and refinement**

The structures of BmK M7 and three mutants of BmK M1 were solved by molecular replacement using the program AMoRe.35 For BmK M7, a homologous model of BmK M8 with PDB code 1snb19 was used as a model. With data from 8 Å to 3 Å, in the first run of AMoRe, one solution was found with correlation coefficient (cc) of 0.34 and R factor of 0.50, indicating a potential location of the first molecule in the asymmetric unit. Fixing the position of this first molecule and extending the molecular replacement search for a second unique solution resulted in improved cc and R factor values of 0.57 and 0.42, respectively, after rigid-body refinement. The structures of the three mutants of BmK M1 were also solved by molecular replacement using the structure of the native BmK M1 (PDB code 1sn1) as a model. The unique solutions for the rotation and translation searches were obvious, exhibiting initial cc and R factor values of about 0.60 and 0.40, respectively, after rigid body refinement. The initial structures of BmK M7 and the three mutants of BmK M1 were refined with CNS36 and the models were rebuilt with TURBO-FRODO44 and O37 with 10% of the data retained for cross-validation purposes. For BmK M7, no non-crystallographic symmetry restriction was applied during the refinement. Iterative rounds of model building and minimization lowered the Rcryst to 0.253 and Rfree to 0.280. After individual B factor refinement and adding water molecules, the Rcryst was 0.1679 and Rfree was 0.1864. When applying the cis and trans dual configurations to the 9–10 peptide bond in molecule B, the Rcryst and Rfree were reduced to 0.1669 and 0.1840, respectively. Finally, anisotropic B factor refinement using REFMAC538 lowered the Rcryst and Rfree values to 0.142 and 0.164, respectively. Structural refinements of the three BmK M1 mutant structures were performed using iterative rounds of positional and individual B factor refinement, as well as the addition of solvent molecules resulting in a final refinement of the K8D, K8Q and P9S BmK M1 mutant structures at resolutions of 1.50 Å, 1.85 Å and 1.40 Å resulted in Rcryst values of 0.165, 0.171 and 0.192 and corresponding Rfree values of 0.183, 0.197 and 0.216, respectively (Table 3).

**Protein Data Bank accession code**

The coordinates and structure factors of BmK M7 and the three mutants of BmK M1 (K8D, K8Q and P9S) have been deposited to the RCSB PDB with accession numbers 1KVO, 1T7A, 1T7B and 1T7E, respectively.

### Table 3. X-ray data collection and refinement statistics of BmK M7 and three mutants of BmK M1

<table>
<thead>
<tr>
<th>Species</th>
<th>BmK M7</th>
<th>K8D</th>
<th>K8Q</th>
<th>P9S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a (Å)</td>
<td>32.76</td>
<td>47.45</td>
<td>47.59</td>
<td>46.91</td>
</tr>
<tr>
<td>b (Å)</td>
<td>32.76</td>
<td>44.34</td>
<td>44.06</td>
<td>46.91</td>
</tr>
<tr>
<td>c (Å)</td>
<td>176.82</td>
<td>25.49</td>
<td>25.51</td>
<td>53.75</td>
</tr>
<tr>
<td>α, β, γ (deg.)</td>
<td>90, 90, 120</td>
<td>90, 90, 90</td>
<td>90, 90, 90</td>
<td>90, 90, 120</td>
</tr>
<tr>
<td>Space group</td>
<td>F2221</td>
<td>F2221</td>
<td>F2221</td>
<td>F2221</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>29.3–1.40</td>
<td>30.0–1.50</td>
<td>30.0–1.85</td>
<td>30.0–1.40</td>
</tr>
<tr>
<td>Highest resolution shell</td>
<td>1.48–1.40</td>
<td>1.55–1.50</td>
<td>1.92–1.85</td>
<td>1.48–1.40</td>
</tr>
<tr>
<td>Number of observations</td>
<td>57,131</td>
<td>41,017</td>
<td>28,173</td>
<td>97,281</td>
</tr>
<tr>
<td>Number of unique reflections</td>
<td>19,260</td>
<td>8946</td>
<td>4831</td>
<td>13,889</td>
</tr>
<tr>
<td>Completeness* (%)</td>
<td>85.8 (74.6)</td>
<td>98.1 (93.3)</td>
<td>97.6 (93.8)</td>
<td>100.0 (100.0)</td>
</tr>
<tr>
<td>Rmerge* (%)</td>
<td>4.0 (20.1)</td>
<td>5.2 (17.5)</td>
<td>8.3 (28.3)</td>
<td>5.3 (35.5)</td>
</tr>
<tr>
<td>l of 〈l〉(Å)</td>
<td>10.0 (3.6)</td>
<td>28.3 (6.7)</td>
<td>22.0 (6.0)</td>
<td>7.9 (2.0)</td>
</tr>
<tr>
<td>B factor from Wilson plot (Å²)</td>
<td>13.0</td>
<td>17.5</td>
<td>19.8</td>
<td>20.4</td>
</tr>
<tr>
<td>Rcryst,A</td>
<td>0.142</td>
<td>0.165</td>
<td>0.171</td>
<td>0.192</td>
</tr>
<tr>
<td>Rcryst,B</td>
<td>0.164</td>
<td>0.183</td>
<td>0.197</td>
<td>0.216</td>
</tr>
<tr>
<td>Rfree (%)</td>
<td>0.171</td>
<td>0.197</td>
<td>0.216</td>
<td>0.236</td>
</tr>
<tr>
<td>Rfree (10% data)</td>
<td>0.164</td>
<td>0.183</td>
<td>0.197</td>
<td>0.216</td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond length (Å)</td>
<td>0.017</td>
<td>0.004</td>
<td>0.008</td>
<td>0.013</td>
</tr>
<tr>
<td>Bond angles (deg.)</td>
<td>1.983</td>
<td>1.300</td>
<td>1.550</td>
<td>1.670</td>
</tr>
<tr>
<td>Chiral (deg.)</td>
<td>0.192</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ramachandran plot</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Core region (%)</td>
<td>91.3</td>
<td>92.6</td>
<td>90.7</td>
<td>87.3</td>
</tr>
<tr>
<td>Additional allowed region (%)</td>
<td>8.7</td>
<td>7.4</td>
<td>9.3</td>
<td>12.7</td>
</tr>
<tr>
<td>Protein model</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein non-H atoms</td>
<td>1008</td>
<td>516</td>
<td>517</td>
<td>520</td>
</tr>
<tr>
<td>Number of phosphate groups</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>3</td>
</tr>
<tr>
<td>Water molecules</td>
<td>118</td>
<td>84</td>
<td>50</td>
<td>51</td>
</tr>
<tr>
<td>Average temperature factors (Å²)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>14.5, 19.1</td>
<td>10.4</td>
<td>17.5</td>
<td>16.6</td>
</tr>
<tr>
<td>Phosphate groups</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>22.0</td>
</tr>
<tr>
<td>Water molecules</td>
<td>31.4</td>
<td>27.5</td>
<td>30.3</td>
<td>32.3</td>
</tr>
</tbody>
</table>

* Values in parentheses refer to the outer resolution shell.

\[ R_{\text{cryst}} = \sum |F_o| - |F_c| / \sum |F_o|, \quad R_{\text{free}} = R_{\text{cryst}} \text{ as for } R_{\text{cryst}} \text{ but calculated for a randomly selected 10% of reflections not included in the refinement.} \]
Acknowledgements

This work was supported by Grants from the NNSF (30370320), the MOST (G19990756, 2002BA711A13) and the CAS (KSCX1-SW-17). Data collection was supported by KEK (00G290). We thank Professor N. Sakabe for his help during data collection in Photon Factory in Japan.

References


30. Nielsen, C., Arvai, A., Szebenyi, D. M. E., Deacon, A.,


Edited by R. Huber

(Received 8 January 2004; received in revised form 15 June 2004; accepted 22 June 2004)