Journal of Computer-Aided Molecular Design **18:** 119–133, 2004.

Comparison of a 3D-model of the classical α-scorpion toxin V from *Leiurus quinquestriatus quinquestriatus* with other scorpion toxins

Stefanie Bendels[a,b] & Hans-Dieter Höltje[a]
[a]*Institut für Pharmazeutische Chemie, Heinrich-Heine-Universität, Universitätsstr. 1, D-40225 Düsseldorf, Germany;* [b]*F. Hoffmann-La Roche Ltd., Pharma Research, CH-4070 Basel, Switzerland*

Received 22 May 2003; accepted in revised form 22 March 2004

Key words: α-toxins, electrostatic potential, LqqV, molecular dynamics, molecular modelling, scorpion toxins

Summary

In the present paper, a study of classical and insect α-scorpion toxins is described. A homology model of the classical α-toxin LqqV from *Leiurus quinquestriatus quinquestriatus* was developed. The model was compared to stable and energetically favourable conformations of AaHII from *Androctonus australis* Hector and LqhαIT from *Leiurus quinquestriatus hebraeus*, which are the most active α-toxins in mammals and insects. The conformations were retrieved from molecular dynamics simulations of known structures. The model of LqqV shows a C-terminal conformation similar to LqhαIT. This is mainly caused by electrostatic interactions between Lys10/Lys60 and Glu59, which are comparable to the cation-π interactions of Tyr10 and Arg64 in LqhαIT. During the simulations the structures of AaHII and LqqV were stabilised through electrostatic interactions between Glu32 and Lys50 and especially the loop adjacent to the α-helix is affected, which is in contrast to LqhαIT. When the molecular electrostatic potentials of the toxins were studied, a possibly important difference between the classical α-toxins and the insect α-toxin LqhαIT was found in the area around Lys30 and Arg56 of AaHII, where a positive potential is missing in LqhαIT. A large negative potential caused by Asp3, Glu15 and Asp19 in LqhαIT is also unique for this toxin. It is proposed that Arg18, which is important for activity of LqhαIT, restricts the negative potential in this area and is not essential for toxins where negatively charged residues in comparable positions are not present.

Abbreviations: AaHI-III, α-toxins I, II, III from *Androctonus australis* Hector; LqqV, α-toxin V from *Leiurus quinquestriatus quinquestriatus*; LqhαIT, insect α-toxin from *Leiurus quinquestriatus hebraeus*; BotIII from *Buthus occitanus tunetanus*; BomIII from *Buthus occitanus mardochei*; LqhII and LqhIII from *Leiurus quinquestriatus hebraeus*.

Introduction

In scorpion venoms a variety of biologically active components are found, of which the peptidic neurotoxins acting on different ion channels are the best studied groups (for reviews see [1, 2]). The sodium channel specific toxins are 60–70-residue polypeptides stabilised by four disulfide bridges. According to their pharmacological properties they are divided in α- and β-toxins. The α-toxins bind to receptor site 3 on the extracellular surface of the voltage-gated sodium channel and inhibit fast inactivation [3]. The β-toxins interact with site 4 and modify the activation process by shifting activation to a more negative membrane potential [4]. Due to differences in species specificity (mammal and/or insects) and in binding properties, the α- and β-toxin types can be divided into several subgroups. Classical α-toxins are highly active on mammals, including toxins such as AaHI-III from *Androctonus australis* Hector and LqqV from *Leiurus quinquestriatus quinquestriatus* (sequences shown in Figure 1). LqhαIT from *Leiurus quinquestriatus hebraeus* belongs to a second subgroup of α-toxins that are highly active on insects. A third group within the α-toxins class has been defined as α-like toxins, which are active on mice but do not compete for

```
                  βββ               αα  αααααααα    βββββββ        βββββββ
AaHII    (1)  VKDGYIVDDVNCTYFCGRNA-YCNEECTKLKGESGYCQWASPYGNACYCYK  (50)
LqqV     (1)  LKDGYIVDDKNCTFFCGRNA-YCNDECKKKGGESGYCQWASPYGNACWCYK  (50)
AaHI     (1)  KRDGYIVYPNNCVYHCVPP---CDGLCKKNGGSSGSCSFLVPSGLACWCKD  (48)
AaHIII   (1)  VRDGYIVNSKNCVYHCVPP---CDGLCKKNGAKSGSCGFLIPSGLACWCVA  (48)
BotIII   (1)  VKDGYIVDDRNCTYFCGRNA-YCNEECTKLKGESGYCQWASPYGNACYCYK  (50)

LqhII    (1)  IKDGYIVDDVNCTYFCGRNA-YCNEECTKLKGESGYCQWASPYGNACYCYK  (50)

LqhaIT   (1)  VRDAYIAKNYNCVYECFRDA-YCNELCTKNGASSGYCQWAGKYGNACWCYA  (50)
LqqIII   (1)  VRDAYIAKNYNCVYECFRDS-YCNDLCTKNGASSGYCQWAGKYGNACWCYA  (50)
BotIT1   (1)  VRDAYIAQNYNCVYFCMKDD-YCNDLCTKNGASSGYCQWAGKYGNACWCYA  (50)

BomIII   (1)  GRDGYIAQPENCVYHCFPGSSGCDTLCKEKGATSGHCGFLPGSGVACWCDN  (51)
LqhIII   (1)  VRDGYIAQPENCVYHCFPGSSGCDTLCKEKGGTSGHCGFKVGHGLACWCNA  (51)


                  β
AaHII    (51) LPDHVRTKGPG-RCH*   (64)
LqqV     (51) LPDRVSIKEKG-RCN*   (64)
AaHI     (49) LPDNVPIKDTSRKCT    (63)
AaHIII   (49) LPDNVPIKDPSYKCHS   (64)
BotIII   (51) VPDHVRTKGPG-RCN*   (64)

LqhII    (51) LPDHVRTKGPG-RCR    (64)

LqhaIT   (51) LPDNVPIRVPG-KCR    (64)
LqqIII   (51) LPDNVPIRVPG-KCH    (64)
BotIT1   (51) LPDNVPIRIPG-KCHS   (65)

BomIII   (52) LPNKVPIVVGGEKCH    (66)
LqhIII   (52) LPDNVGIIVEGEKCHS*  (67)
```

Figure 1. Sequence alignment (CLUSTALW [34]) of polypeptides isolated from scorpions. An asterisk (*) indicates amidation at the C-terminus. Classical α-scorpion toxins (highly active on mammals): AaHI (SWISS-PROT Database [35]: SCX1_ANDAU), AaHII (SCX2_ANDAU), AaHIII (SCX3_ANDAU, some references adduce D8 instead of N8) from *Androctonus australis* Hector; LqqV (SCX5_LEIQU) from *Leiurus quinquestriatus quinquestriatus*; BotIII (SCX3_BUTOC) from *Buthus occitanus tunetanus*; α-toxin (active on mammals and insects): LqhII [6] from *Leiurus quinquestriatus hebraeus*; insect α-toxins (highly active on insects): LqhαIT (SCXA_LEIQH) from *Leiurus quinquestriatus hebraeus;* LqqIII (SCX3_LEIQU) from *Leiurus quinquestriatus quinquestriatus*; BotIT1 (SIX1_BUTOC) from Buthus occitanus tunetanus; α-like toxins: BomIII (SCX3_BUTOM) from *Buthus occitanus mardochei*; LqhIII (SCL3_LEIQH) from *Leiurus quinquestriatus hebraeus;* PDB accession codes: AaHII (1aho [18], 1ptx [29]); LqhαIT (1lqh, 1lqi [19]); LqqIII (1lqq [36]); LqhIII (1bmr [37]). For AaHII the regions of secondary structure are indicated above the sequence: α-helix: 19–28, β-strands: 2–4, 32–37, 45–51.

AaHII binding on rat brain synaptosomes [2]. BomIII from *Buthus occitanus mardochei* [5] and LqhIII from *Leiurus quinquestriatus hebraeus* [6] belong to this group.

Classical α-toxins like AaHII and LqqV, which show a high binding affinity in rat brain synaptosomes, are able to inhibit LqhαIT binding on cockroach membranes as illustrated in Table 1 [5]. However, concentrations of the toxins 3–4 orders of magnitude higher than LqhαIT are needed. In contrast, the insect α-toxin LqhαIT was found to have a very low ability to inhibit the binding of AaHII in rat brain synaptosomes, as observed by Gordon et al. [5], who did not detect a significant inhibition at 1 μM toxin.

Information about important residues of the toxins is derived by chemical modification and mutational analysis. For the classical α-toxin AaHII chemical modifications of Trp38 [7], Arg56 [8] and Lys58 [9] showed an influence on the activity of the toxin. In addition, a positively charged residue in position 2 of AaHI and AaHIII [10] seems to be important. On the other hand, no significant effect on the interaction with the mammalian receptor site was measured after modification of Arg18 and Arg62 in AaHII.

So far, data from mutational analysis are only available for LqhαIT, an α-toxin highly active on insects. Substitutions of the positively charged residues Lys8, Arg18, Lys62 and Arg64 as well as the aromatic residues Tyr10, Phe17 [11] and Tyr21 [12] decreased the binding affinity for the receptor site on cockroach neuronal sodium channels. Lys41 was also considered in this study, although the increase of the IC_{50} for the double mutant G40S/K41P was only 26-fold [11]. Charged residues usually contrib-

Table 1. Binding affinity of scorpion toxins.

	Rat brain IC_{50}[a] (nM)	Cockroach central nervous system IC^{b}_{50} (nM)
AaHII[c]	0.2–0.3	58.8 ± 12.6
LqqV[c]	1.0	120.4 ± 10.1
AaHI[c]	4.5	160.5 ± 59.6
AaHIII[c]	3.0	325.6 ± 120.2
LqhII[d]	0.4 ± 0.1	5.72 ± 0.63
LqhαIT[c]	>>1000	0.02 ± 0.01
LqqIII[d]	700	0.03 ± 0.01
BomIII[d]	>>1000	29.3 ± 8.1
LqhIII[d]	>1000	0.43 ± 0.2

[a]Competition for ^{125}I-AaHII binding in rat brain synaptosomes.
[b]Competition for ^{125}I-LqhαIT binding in cockroach nervous system membranes.
[c]See Gordon et al. [5] for references.
[d]See Sautiere et al. [6] for references.

ute smaller amounts to the binding free energy in a protein–protein complex, if the desolvation penalty is large [13–16]. Lys41, which is part of a tight β-turn (Gly40–Gly43), is solvent exposed, so that the energy penalty for desolvation should be large.

Usually the function of important residues could be (i) a direct interaction with the receptor site 3 on the sodium channel; (ii) a stabilisation of the molecule or important residues; or (iii) a contribution to a specific electrostatic potential.

In the present study, we developed a homology model for the classical α-scorpion toxin LqqV based on the crystal structure of AaHII (Protein Data Bank PDB [17]: 1aho [18]). The 3D model was compared with AaHII and LqhαIT (NMR structures: 1lqh, 1lqi [19]), the most active α-toxins on mammals and insects, to get more information about the 'functional sites' of the toxins. The specific functional site describes both the nature and spatial organisation of the elements by which the toxin interacts with its receptor site and exerts its function [2, 20].

We used molecular dynamics (MD) simulations to retrieve conformations of the toxins which are stable and energetically favourable under identical conditions (method, forcefield, temperature, pH, dielectric constant). This was important, because although comparisons of X-ray and NMR structures have yielded close similarities between the two types of structures, differences are usually localised in mobile loop regions and surface residues, features least well characterised in NMR structures due to the sparsity of experimental restraints, and in X-ray structures due to crystal contacts [21]. Although MD simulations are limited in their precision and accuracy due to the application of empirical forcefields and to trajectory lengths, it was shown that MD simulations of proteins starting from high-resolution crystal structures with explicit solvent molecules could be used as an alternative method to NMR for studying protein structure in solution (for references see [21]).

The molecular electrostatic potential (MEP) of the α-scorpion toxins, which inhibit inactivation of sodium channels, was studied. The inactivation process is voltage-dependent, mostly due to its coupling to the activation process driven by transmembrane movements of the voltage sensors (for reviews see [22, 23]). Recently a new mechanism for the gating charge movement in a voltage-dependent K^+ channel was published based on X-ray studies [24, 25]. The S4 segment in domain IV seems to be important in this process and α-scorpion toxins as well as sea anemone toxins, which uncouple activation from inactivation, bind to receptor site 3 at the extracellular site of the channel. Part of this binding site is the loop between segments 3 and 4 in domain IV [3]. A strong MEP of the toxins probably influences the movement of the voltage sensor. A possible contribution of a positive potential to the inhibition of inactivation arose from investigations on sea anemone toxins [26, 27] and scorpion toxins [11].

Methodology

A 3D model of LqqV

Scorpion toxins targeting sodium channels show analogous 3D structures with a common secondary structure motif, which comprises an α-helix and a three-stranded antiparallel β-sheet [2]. Comparing the sequence of LqqV with published 3D structures in the PDB, LqqV has the highest amino acid sequence identities to the classical α-toxin AaHII (78%) and to the acidic toxin Bmk M8 from *Buthus martensii* Karsch (59%; PDB: 1snb [28]), which shows only a weak toxicity in mice and is so far not considered a classical α-toxin. Considering the structurally conserved regions within the scorpion toxins, the high percentage of identical residues and the same pharmacological class (classical α-toxins) AaHII was used as template for LqqV.

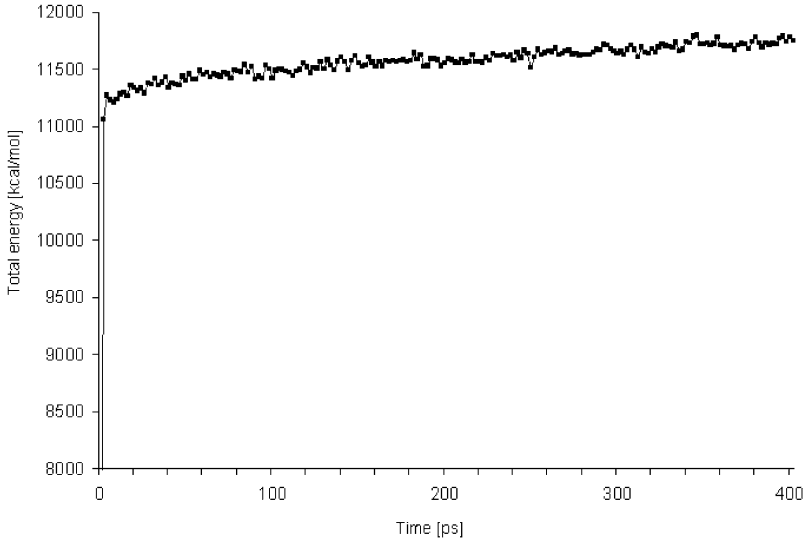

Figure 2. Total energy record for the simulation of LqqV. The relaxation of the water molecules (20 ps) is not shown. During the first 80 ps a decreasing force constant (200, 100, 50, 10 kcal/mol·Å^2) was used to keep the atoms in their original positions. No strong variations in the energy were observed during the sampling phase with free atoms (80–400 ps).

In the Protein Data Bank two X-ray structures of AaHII are available: 1aho [18] (resolution 0.96 Å) and 1ptx [29] (resolution 1.3 Å). These structures have a small root mean square deviation of 0.155 Å for backbone atoms. In this study coordinates for AaHII were taken from 1aho after amidation at the C-terminus (SWISS-PROT Database: SCX2_ANDAU). These coordinates were assigned to LqqV using the HOMOLOGY module of the INSIGHTII [30] software package. An additional conformational search with rotamer libraries of the HOMOLOGY module did not lead to further improvement of the LqqV model.

After refinement of the initial model, water molecules from the AaHII structure were transferred to LqqV and the new assembly was soaked in an equilibrated waterbox. Using the consistent valence forcefield (CVFF) of DISCOVER, MD simulations and energy minimisations according to the following protocol were performed to retrieve an optimised, hydrated polypeptide. A temperature of 310 K, periodic boundary conditions, a cutoff of 15 Å and a dielectric constant $\varepsilon = 4$ were chosen. In addition $\varepsilon = 1$ was tested, but unfavourable geometric properties were detected using PROCHECK [31], probably due to strong intramolecular interactions of the numerous charged residues. For the ionisation of the acidic and basic groups a physiological pH = 7.4 was considered.

After relaxation of the water molecules (20 ps) with fixed heavy atoms of LqqV and a short minimisation, the subsequent simulation was performed for 20 ps using a tethering constant of 200 kcal/mol·Å^2 for the entire polypeptide, restraining atoms to their original position. This was important for avoiding high kinetic energies due to unfavourable contacts of atoms at the beginning of the calculation. The constant was decreased to 100 kcal/mol·Å^2 for the next 20 ps. Another 20 ps were performed using a constant of 50 kcal/mol·Å^2 only on backbone atoms, and the force was decreased to 10 kcal/mol·Å^2 in the next step. In the following 320 ps the entire system was flexible and one conformation was archived every 2 ps (total energy record shown in Figure 2). During the MD simulation the *trans* configured omega torsions were restrained to be *trans* using an energy constant of 20 kcal/mol.

MD simulations of AaHII and LqhαIT

For the classical α-scorpion toxin AaHII the crystal structure deposited in the Protein Data Bank as 1aho was used for simulation after amidation at the C-terminus. Conformation A was considered for the amino acid residues, which are modelled in two different conformations (Asp9, Glu24, Cys12, and Cys63).

Two MD simulations with different starting molecules were performed for the insect α-toxin LqhαIT, whose solution structure was determined by NMR

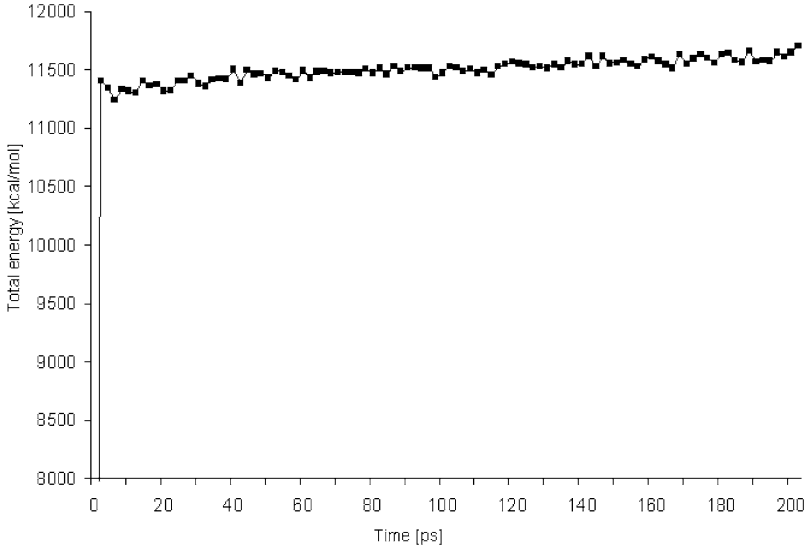

Figure 3. Total energy record for the simulation of AaHII. The relaxation of the water molecules (20 ps) is not shown. During the first 80 ps a decreasing force constant (200, 100, 50, 10 kcal/mol·Å^2) was used to keep the atoms in their original positions. No strong variations in the energy were observed during the sampling phase with free atoms (80–200 ps).

spectroscopy [19]. For the first trajectory conformation 1 of 1lqi (29 frames) was used, including the additional Met0 (N-terminus) of the original NMR structure. In a second calculation the minimised average structure of LqhαIT (1lqh) was investigated without Met0. Using frames obtained in the first simulation more favourable superpositions with other toxins were possible and therefore the second simulation will not be mentioned in detail.

The same protocol (minimisation, MD simulation), used for modelling LqqV was taken for the other α-scorpion toxins AaHII and LqhαIT. Only the sampling phase, where the polypeptides were completely flexible, was reduced to 120 ps, so that 60 frames of each toxin were retrieved for further investigations.

In summary, the time periods of the simulations were 220 ps and during this time the systems reached equilibrium conditions (see Figure 3). Although this is perhaps not sufficient for the description of the configuration space, it was possible to retrieve relaxed conformations of the toxins, which was the major task. Therefore longer simulations, which would cause high calculation times due to the application of explicit water molecules, were not performed.

Representative MD conformations of the toxins

Cluster analyses were performed for each trajectory ensemble to group related conformers into one common cluster. Therefore the program NMRCORE [32] was used, which determines the local structural domains for the examined polypeptide. For clustering NMRCLUST [33] was taken. The frames were superimposed on the largest local structural domain and the root mean square deviations were calculated for all heavy atoms. For each individual cluster NMRCLUST also provides one representative frame, that is the structure closest to the centroid of the cluster. These structurally diverse frames representing one cluster were geometry-optimised using comparable conditions to the previous MD simulation (aqueous environment, periodic boundary conditions, cutoff = 15 Å, dielectric constant $\varepsilon = 4$, 20 kcal/mol for trans configured omega torsions). At the beginning the steepest descent algorithm was used (2000 iterations) and the entire polypeptide was tethered with 500 kcal/mol·Å^2 to restrain atoms to their original positions. This relatively high force constant was used at the start to avoid extensive movements due to unfavourable interactions. Subsequently conjugate gradient was employed and the force constant was decreased from 500 to 100, 50 and 10 kcal/mol·Å^2 (each step 2000 iterations) until all atoms were flexible (5000 iterations).

For each toxin one minimised frame was chosen considering the size of the represented cluster, because a large population is an indication for a stable conformation. In addition the total energy and the geometrical properties (PROCHECK) were investigated

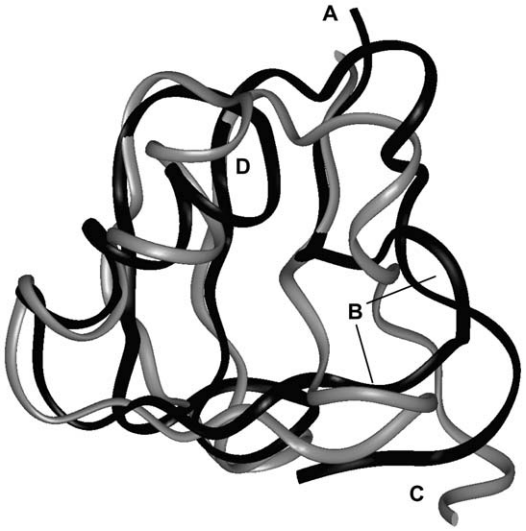

Figure 4. Backbone displayed in ribbon of AaHII (black) and LqhαIT (grey) superimposed on backbone atoms. The N-terminus (A), the 5-residue turn (B), the C-terminus (C) and the loop following the helix (D) are indicated. The toxins are superimposed on backbone atoms of the secondary structures (AaHII: α-helix: 19–28, β-strands: 2–4, 32–37, 45–51).

as characteristics for one individual conformation and taken into account for the selection.

Molecular electrostatic potential

The molecular electrostatic potential (MEP) was calculated using DelPhi [30]. The program solves the Poisson–Boltzmann equation by the finite differences method. A focussing method was employed so that the percent of the cubic box edge occupied by the polypeptide's greatest dimension was 40% in the first calculation and 70% in the second. Simultaneously the grid resolution was increased from 0.8 Å to 0.5 Å per grid point. For the interior of the protein a dielectric constant $\varepsilon = 4$ was used; for the surrounding solvent medium, which is defined as the region outside the solvent-accessible surface of the molecule, $\varepsilon = 80$ was taken with a physiological ionic strength of 0.145 mol/L.

All computations were performed on Silicon Graphics O2 (R5000, R12000) and Origin 2000 (R10000) computers.

Results and discussion

Initially the classical α-scorpion toxin AaHII and the insect α-toxin LqhαIT were superimposed. Important residues of LqhαIT should be at least partially present in AaHII, which can inhibit LqhαIT binding on cockroach nervous system membranes (see Table 1). Apart from common features also distinctive differences are expected, because LqhαIT has a very low ability to inhibit binding of AaHII in rat brain synaptosomes. Common and opposing properties should also be present in other classical α-scorpion toxins, e.g. LqqV. Therefore, the sequences of classical and insect α-toxins, as well as α-like toxins, were analysed. In addition the 3D model of LqqV was included.

Although mainly hydrophobic interactions contribute to the free energy of binding in protein–protein complexes, the long-range electrostatic interactions strongly affect the rate of association and the specificity of binding [13–16]. For superposition of the toxins, hydrophobic and especially charged residues had to be considered.

Superposition of AaHII and LqhαIT

When comparing the NMR structure of LqhαIT with the crystal structure of AaHII Tugarinov et al. found differences at the 5-residue turn comprising residues Asp8–Cys12 (AaHII) and the C-terminal segment. Superimposition of the same toxins using conformations retrieved from molecular dynamics simulations shows that these observed differences became more distinctive (see Figure 4: B,C). The most poorly defined regions within the NMR structure of LqhαIT are the Lys8–Cys12 5-residue turn, the Gly40–Gly43 turn, and four C-terminal residues [19]. Therefore the conformational changes in the regions B and C are not unexpected.

In addition, a different conformation of the loop (D) following the α-helix is also observed. During the simulation this loop is stabilised through electrostatic interactions between Lys50 and Glu32 in AaHII (see Figure 5). Both residues are missing in LqhαIT, leading to a conformational change in the simulation, but nevertheless the superposition of the MD conformation and the minimised average NMR structure of LqhαIT shows that only slight changes occurred (see Figure 6a). The differences in the structures of AaHII (MD conformation and original crystal structure superimposed in Figure 6b) were also relatively small except for the 5-residue turn (B). This change is caused by electrostatic interaction between Arg56 and Asp8/9.

It was possible to overlay residues Lys8, Arg18, Lys41 and Phe17 of LqhαIT, which are important

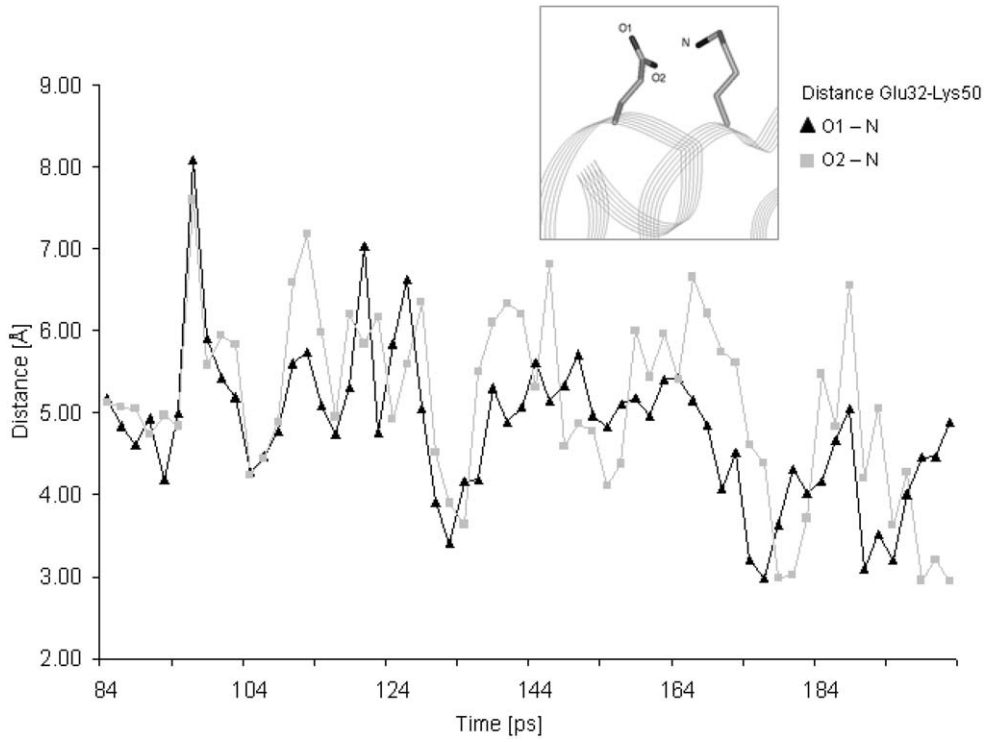

Figure 5. Distance between Glu32 and Lys50 of the toxin AaHII during the simulation. All atoms are completely flexible. The salt bridge (distance between the negatively and the positively charged atoms $\leq$ 4.1 Å) recurs and in several frames the criteria for hydrogen bonds are additionally fulfilled (13%).

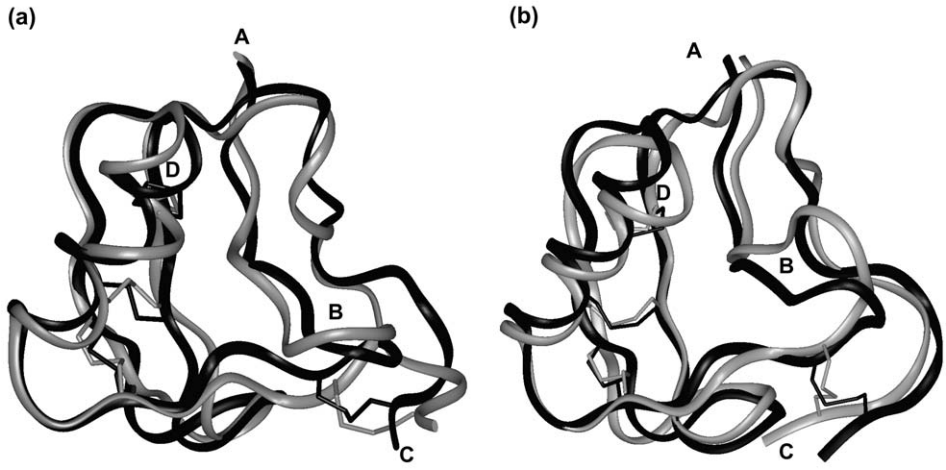

Figure 6. Backbone displayed in ribbon of different LqhαIT and AaHII conformations after superimposition on backbone atoms. Disulfide bridges are displayed and the N-terminus (A), the 5-residue turn (B), the C-terminus (C) and the loop following the helix (D) are indicated. (a) LqhαIT conformations: minimised average structure [1lqh]: black; MD conformation: grey; root mean square deviation (RMSD) = 1.9 Å. (b) AaHII conformations: crystal structure [1aho]: black; MD conformation: grey, RMSD = 2.0 Å.

Table 2. Superimposition of AaHII and LqhαIT.

AaHII	α-NH_2[a]	Lys2	Lys28	Arg18	Arg62	Lys58	Asp9	Tyr21	Phe15
LqhαIT	α-NH_2[a]	Arg2	Lys8	Arg18	Lys41	Arg58	COO^- (Arg64)[b]	Phe17	Trp38

[a]α-NH_2: N-terminal α-amino group.
[b]C-terminal carboxyl group.

for binding in cockroach nervous system membranes, with corresponding residues of AaHII (see Table 2 and Figure 7). The conserved hydrophobic surfaces (AaHII: Tyr5, Tyr35, Tyr42, Tyr47 and Tyr49) were found in comparable orientation (not displayed).

The important residues Tyr10 and Tyr21 and also the charged residues Lys62 and Arg64 could not be superimposed with equivalent parts in AaHII. Since the major function of Tyr10 is probably the stabilisation of the 5-residue turn relative to the C-terminus in LqhαIT through cation-π interaction with Arg64, it is not directly interacting with the channel. AaHII shows a different C-terminal conformation, so it was expected that Tyr10 is not overlaid. In addition to Tyr10, the absence of Tyr21 in AaHII is supposed to be insignificant, because hydrophobic interactions are not as specific as electrostatic interactions, therefore residues in comparable orientation should be able to compensate missing residues.

The differences in the charged residues are especially interesting, because in a protein–protein complex they are often important for specificity. Substitution of Arg64 by His revealed a three-fold increased insect and anti-mammalian toxicity of LqhαIT [11] and the side chain of His64 in the R64H mutant is oriented in the opposite direction relative to that of Arg64 in the unmodified toxin [19]. Therefore neither the positive charge nor the conformation in this part should be essential for toxicity of LqhαIT. It is conceivable that a steric effect caused by the exposed side chain of Arg64 contributes to the lower toxicity of LqhαIT in comparison to R64H.

Relatively minor changes in the sequence of scorpion toxins may change their relative selectivity, which was shown by Sautiere et al. [6]. LqhII only differs in two residues (LqhII: Ile1, Arg64) compared to AaHII, although it reveals significantly higher activity to insects. This is consistent with our results, because it was not possible to superimpose the important C-terminal Lys62 from LqhαIT with equivalent residues of AaHII, although other important residues could be considered, and therefore an additional Arg at the C-terminus, present in LqhII, should compensate for this restriction.

Looking at AaHII, the important residue Lys58 could be overlaid with Arg58 of LqhαIT, but it was not possible to consider Arg56. In addition, the positively charged Lys30, which is present at the same surface as Arg56, has no equivalent in LqhαIT. Lys50, found in electrostatic interaction with Glu32, is also not present in LqhαIT. If these residues are important features for activity in mammals, they should also be present in other classical α-scorpion toxins like LqqV.

Superposition of AaHII and LqqV

Although the 3D model of LqqV is based on the crystal structure of AaHII, differences in the conformation of the C-terminus and the 5-residue turn were found. The changes were caused by electrostatic interactions between Glu59 and Lys10/Lys60 in LqqV (see Figure 7). It is proposed that a similar conformation could also be present in AaHI (Asn10, Asp57, Arg60) and AaHIII (Lys10, Asp57, Tyr60) (for alignment see Figure 1). This pattern is comparable to LqhαIT, where interactions between Tyr10 and Arg64 are proposed to stabilise the 5-residue turn relative to the C-terminus.

The residues Lys30 and Arg56 of AaHII are of special interest while superpositioning AaHII and LqqV, because they are probably important for the binding to the mammalian sodium channel, although they are only present in a few classical α-scorpion toxins like AaHII and BotIII. Nevertheless, in the sequences of LqqV, AaHI and AaHIII at least two positively charged residues are found near Lys30, a feature which is different to insect α-toxins like LqhαIT. In the α-like toxins BomIII and LqhIII, which do not inhibit binding of AaHII in rat brain synaptosomes (see Table 1), Lys28 and Lys30 are also found, but in addition the negatively charged Glu29 is present, which should disrupt a positive potential in this area.

In Table 3, the superposition of AaHII and LqqV is presented. It was possible to superimpose residues Arg18, Lys28, Asp8 and Asp9 of both toxins and also the conserved hydrophobic surfaces were in a similar orientation. The important residues Arg56 (LqqV: Arg54) and Lys58 as well as Lys2 and Lys30 (Lys29) in AaHII were found in a comparable spatial ar-

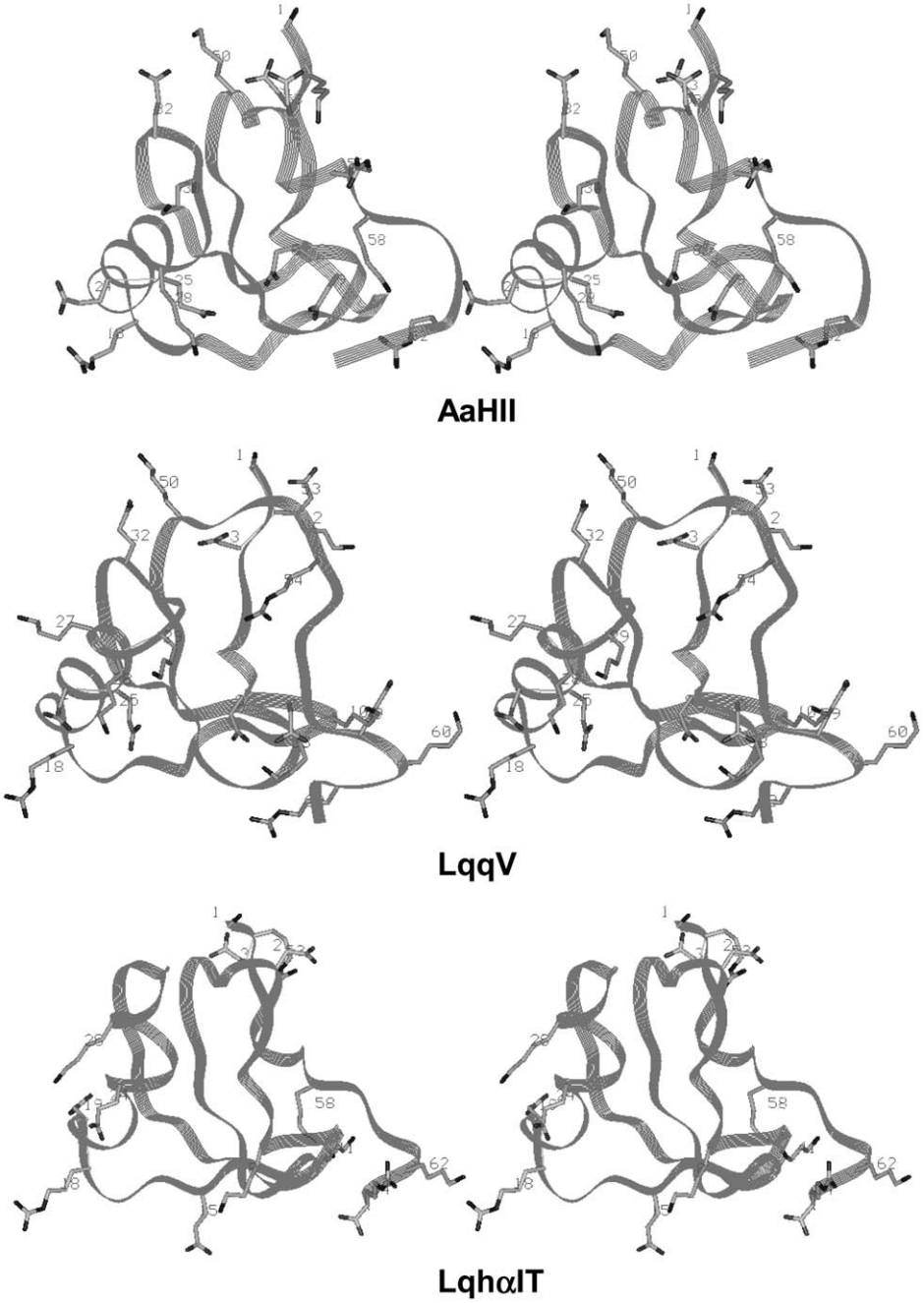

Figure 7. Stereoview of the scorpion toxins AaHII, LqqV and LqhαIT in front view. The backbone (in ribbon) and the charged residues are displayed.

Table 3. Superimposition of AaHII and LqqV.

AaHII	α-NH_2[a]	Lys2	Arg18	Lys28	Lys30	Arg56	Lys58	Asp8/9
LqqV	α-NH_2[a]	Lys2	Arg18	Lys28	Lys29	Arg54	Lys58	Asp8/9

[a]α-NH_2: N-terminal α-amino group.

Table 4. Superimposition of LqqV and LqhαIT.

LqqV	α-NH_2[a]	Lys2	Lys10/Lys60	Arg18	Lys27	Lys58	Asp9
LqhαIT	α-NH_2[a]	Arg2	Arg64/Lys62	Arg18	Lys28	Arg58	COO^- (Arg64)[b]

[a]α-NH_2: N-terminal α-amino group.
[b]C-terminal carboxyl group.

rangement in LqqV. The conformational change in the C-terminus had a beneficial influence on the superposition of Arg56 and Arg54. Nevertheless, due to this change the C-terminal Arg62 could not be overlaid in the toxins. This is consistent with experiments, which showed no important function of Arg62 for the activity of AaHII.

In the MD simulation the position of the positively charged side chains of the residues Lys28, Arg56 (LqqV: Arg54) and Lys30 (Lys29) were influenced by the negatively charged residues Asp8/Asp9 and Asp24/Asp25 (LqqV: Glu24/Asp25). The interaction of Arg56 (Arg54) with the negatively charged residues 8 and 9 influenced the relative position of the 5-residue turn.

Lys28 and Lys30 are found in BotIII, another classical α-toxin; however, Lys30 does not occur in the sequence alignment of AaHI and AaHIII (shown in Figure 1). In addition, the negatively charged residues in positions 8/9 and 24/25 are also missing and two gaps at the beginning of the helix (AaHII: 19–28) are present in AaHI and AaHIII compared to AaHII. It is perhaps possible that Lys25 and Lys26 of AaHI and AaHIII could adopt a similar orientation as Lys28 and Lys30 in AaHII.

The salt bridge between Glu32 and Lys50, which is found in AaHII, LqqV and BotIII, is probably important for the specificity, although this bridge is not present in some classical α-toxins like AaHI and AaHIII (for review see [2]). The structure of AaHII is stabilised by four disulfide bridges (Cys12-Cys63, Cys16-36, Cys22-Cys46, Cys26-Cys48), a β-sheet comprising three strands (residues 2–4, 32–37, 45–51) and one α-helix (residues 19–28). The salt bridge between the residues 32 and 50 enhances the stability of AaHII and LqqV and especially the conformation of the loop adjacent to the α-helix is affected. Comparable to Glu32/Lys50 a stabilisation of the loop in

Figure 8. Backbone displayed in ribbon of LqqV (black) and LqhαIT (grey) superimposed on backbone atoms. The N-terminus (A), the 5-residue turn (B) and the C-terminus (C) are indicated.

the toxin AaHI could be possible through interactions between Ser30/31 and Lys47 or Asp48, whereas in AaHIII Lys30 could take part in this function (for alignment see Figure 1). It is possible that the function of the salt bridge between Glu32 and Lys50 for mammalian specificity is the stabilisation of special parts in the toxins.

Superposition of LqqV and LqhαIT

The classical α-toxin LqqV is able to inhibit the binding of LqhαIT in cockroach nervous system membranes, but in comparison to AaHII, the IC_{50} value for LqqV is increased (see Table 1). Superimposition of the 3D model of LqqV and insect α-toxin LqhαIT (see Figure 8) shows close similarities in the N-terminus (A), the 5-residue turn (B) and the C-terminus (C). The conformation of the 5-residue turn relative to

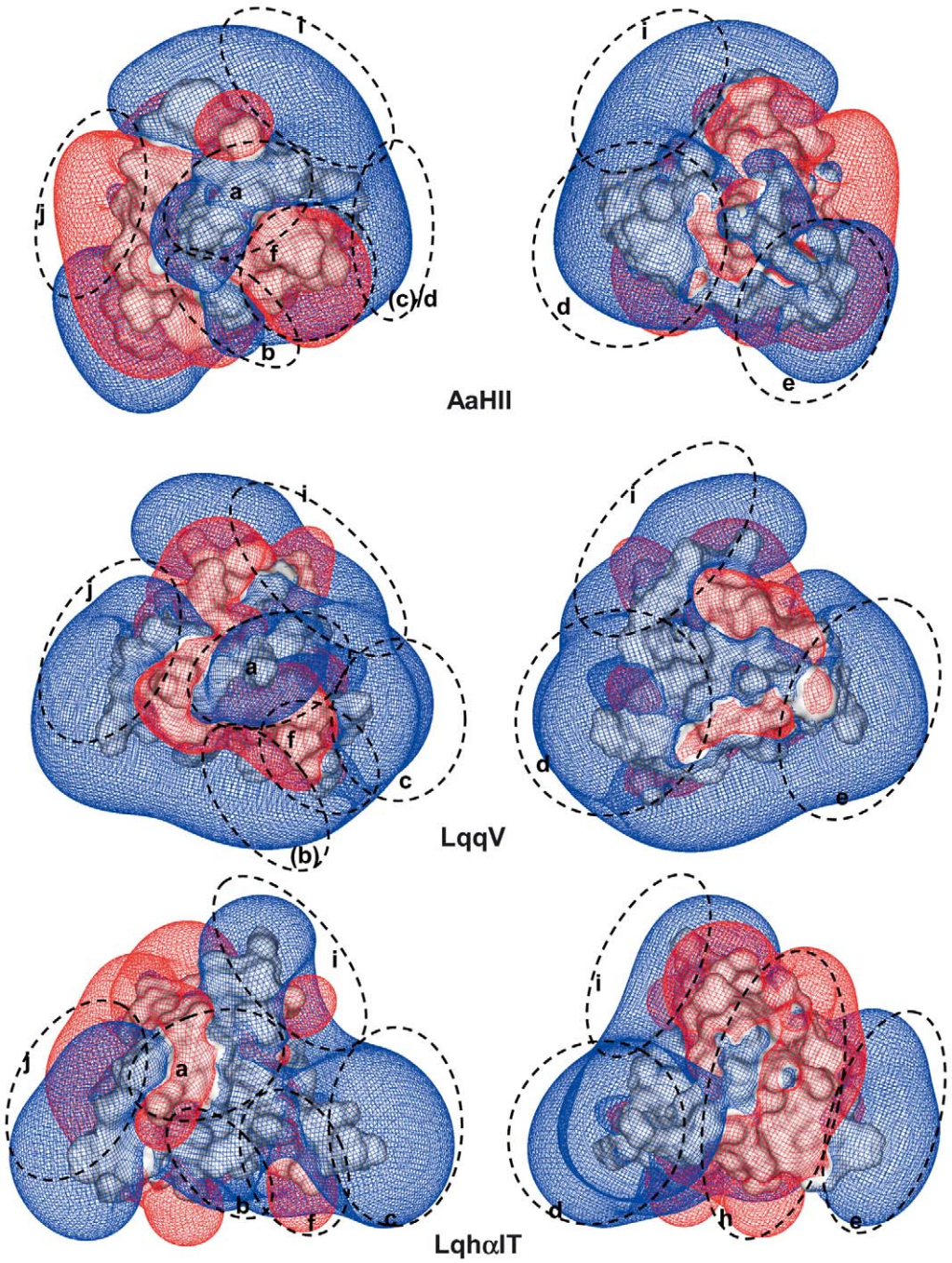

Figure 9. Molecular electrostatic potentials (MEP) of AaHII, LqqV and LqhαIT in the front view (left) and rotated at 180° (right) contoured at +0.2 kT/e (blue) and −0.2 kT/e (red).

the C-terminus in LqqV is caused by electrostatic interactions between Glu59 and Lys10/Lys60, which is comparable to the interaction between Tyr10 and Arg64 in LqhαIT [2].

In the superposition of LqqV and LqhαIT the conserved hydrophobic surfaces were found in a comparable orientation. The charged residues Arg2, Arg64/Lys62, Arg18, Lys28, Arg58 and the C-terminal carboxyl group of LqhαIT could be overlaid with Lys2, Lys10/Lys60, Arg18, Lys27, Lys58 and Asp9 from LqqV (see Table 4). Nevertheless, it was not possible to find equivalent residues in LqqV for the important residues Lys8 and Lys41 of LqhαIT. The absence of these residues should have an important impact on the affinity of LqqV to cockroach sodium channels. Comparing AaHII with LqhαIT, Lys8 and Lys41 of LqhαIT were both superimposed with Lys28 and Arg62, and AaHII shows a higher affinity than LqqV for cockroach sodium channels.

The model of LqqV shows that the C-terminal conformation of LqhαIT itself should not be an important feature for the higher affinity to cockroach sodium channels, nor for the very low affinity to mammalian channels. Furthermore, it is interesting that the solvent-exposed side chain of Arg64 in LqhαIT is close to the 5-residue turn, where Lys8 is found, which is probably in direct interaction with the sodium channel [11]. The comparison of LqhαIT with the 3D model of LqqV suggests that the increased affinity of the LqhαIT mutant R64H is due to an unfavourable steric effect caused by the volume of the exposed side chain of Arg64 in the original toxin.

Comparable to the superposition of AaHII and LqhαIT, where Lys30 and Arg56 of AaHII could not be considered, equivalent residues for the positively charged residues Lys28, Lys29 and Arg54 of LqqV were not found in LqhαIT.

Molecular electrostatic potentials

AaHII and LqqV. Molecular electrostatic potentials (MEPs) of the classical α-toxin AaHII (see Figure 9) show a large positive area **i** caused by the N-terminal α-amino group, Lys2 and Arg56. This potential extends over areas **a** (Arg56/Lys30) and **b** (Lys28). The positive MEPs in area **i** (α-amino group, Lys2 and Arg54) and area **a** are also found in LqqV, but the potential in region **a** is not only caused by Arg54 and Lys29, but Lys28 contributes to this MEP as well. Although potential **b** in AaHII and LqqV could therefore not be directly superimposed, in LqqV a large positive potential mainly due to the influence of Arg62 and Arg18 is found in close proximity. Additionally, the positive MEPs in area **d**, which are attributed to Lys58/Lys62 in AaHII and to Lys58/Lys60 in LqqV, and also in area **e**, which in both toxins is due to Arg18, were found to be in good agreement. The potential in part **c** is smaller in AaHII than in LqqV, because in the first toxin only the slightly buried Lys58 contributes to this MEP, in contrast to LqqV, where the additional Lys10 and also Lys58 and Arg62 play a role. The negative potential in part **f**, which is caused by Asp8 and Asp9, is larger in the toxin AaHII. But this MEP should not be important for the interaction with the sodium channel, because these two residues are not conserved in classical α-toxins. Instead it is proposed that they are important for the stabilisation of positively charged residues. An opposite MEP was found in area **j** due to Asp24 in AaHII and Lys27 in LqqV.

The investigation shows that the positive potentials of the toxin LqqV are more distinctive in comparison to AaHII, especially in the areas **c** and **j**. Nevertheless, AaHII is the most active toxin in mammals, so obviously not a large global positive potential is important, but instead MEPs in particular parts of the molecule probably contribute to the interaction with the channel and the inhibition of inactivation. Taking into consideration the clear difference in the large potential in area **j**, it is proposed that this part of the classical α-toxins (residues 24–27 of the helix) is also not important for the direct interaction. This is in agreement with investigations of El Ayeb et al., whose group was able to show that the α-helix region (residues 23–32) is still accessible to its respective antibodies when the toxin is bound to the channel [10].

AaHII and LqhαIT. Comparing the MEP of AaHII and LqhαIT (see Figure 9) after superposition, according to Table 2, common positively charged potentials were found in area **i** (AaHII: α-amino group/Lys2/Arg56; LqhαIT: α-amino group/Arg2), in area **b** (AaHII: Lys28; LqhαIT: Lys8), in area **d** (AaHII: Lys58/Arg62; LqhαIT: Lys41/Lys58) and in area **e** (both: Arg18). The positive potential in **c** is not so distinctive in toxin AaHII (Lys58), because in LqhαIT Lys58, Lys62 and Arg64 are present in this part of the molecule.

The negative MEP in area **f** is much smaller in LqhαIT, because only the C-terminal carboxyl group could be superimposed on Asp9 in AaHII. Nevertheless, the difference in this part of the toxins should not

be important for activity, as Asp8 and Asp9 are not conserved in classical α-toxins. The opposite potential in **j** is comparable to the difference between AaHII and LqqV, so it should not be an important feature in the comparison of AaHII and LqhαIT.

More interesting is the difference of the potential in region **a**. Both residues contributing to the large positive MEP in AaHII, Arg56 and Lys30, are not present in LqhαIT, therefore the positive area **a** is very small. In addition a negative potential is found in close proximity due to the influence of Glu24 and this negative MEP overlays the position of Lys30 in AaHII.

Furthermore, a distinctive negative potential **h**, extending over Asp3 and Asp19, was found in LqhαIT, which also covers an area between Lys8 and Lys41 caused by Glu15. This negative potential **h** is confined by Arg18. This residue is not important for the activity of classical α-scorpion toxins like AaHII, and in the sequences of the classical α-toxins the negatively charged residues Glu15 and Asp19, which are found in α-insect toxins like LqhαIT and LqqIII, are not present. In addition, the sequences of the α-like toxins BomIII and LqhIII, which are able to inhibit binding of LqhαIT to cockroach sodium channels, do not contain Arg18, but also lack the negatively charged residues in positions 15 and 19. So the importance of Arg18 for the α-insect toxins is probably the restriction of the negative potential in the area **h**. Therefore, the positively charged residue in position 18 is not important for toxins, where a negative potential like region **h** is not present.

LqqV and LqhαIT. Comparing the molecular electrostatic potentials of LqqV and LqhαIT after superposition (Table 4), similarities were found in the positive MEPs in area **i**, caused by the α-amino group and Arg2 in LqhαIT (LqqV: α-amino group/Lys2/Arg54), in area **c** due to Lys58, Lys62 and Arg64 (LqqV: Lys10/Lys58/Arg62) and in area **d** through Lys41 and Lys58 (LqqV: Lys58/Lys60). Further, agreement in the positive potentials is found for region **e**, in both toxins caused by Arg18, and in region **j**, where Lys27 in LqqV and Lys28 in LqhαIT are involved. The negative MEP in area **f** is slightly shifted and also the positive potential **b** caused by Lys8 in LqhαIT could not be directly overlaid. Comparable to the superposition of AaHII and LqhαIT, the large positive potential **a** in toxin LqqV (Lys28/Lys29) is partially found in the negative potential in LqhαIT caused by Glu24.

Comparison with LqhIII and BomIII

The absence of equivalent residues for Lys8 and Lys41 in LqqV in comparison to LqhαIT is proposed to have an important impact on the reduced binding affinity of LqqV to cockroach sodium channel. This point is interesting when analysing the α-like toxins BomIII and LqhIII (for alignment see Figure 1). In comparison to AaHII and LqhαIT, where eight positively charged residues are present, BomIII and LqhIII only have five positively charged residues but both are able to inhibit binding of LqhαIT in cockroach nerve system membranes (see Table 1). Important positively charged residues in LqhαIT are Lys8, Arg18, Lys41, Lys62 and Arg64. The positive charge in position 64 itself is not essential for activity, as was shown through the mutant R64H [11] and, from the data presented in this study, it is proposed that Arg18 is only important in toxins in which additional negatively charged residues in positions 15 and 19 are present. This is supported by the modification of Arg18 in AaHII, where no significant effect on the interaction with the mammalian receptor site was measured [8]. Arg18 is found in the sequences of LqqV and AaHII, but both toxins do not have negatively charged residues in positions 15 and 19, therefore Arg18 is not an important residue for the interaction with the receptor.

The sequence alignment demonstrates that Arg2 and Lys62 of LqhαIT are also found in BomIII (Arg2, Lys64) and LqhIII (Arg2, Lys64). In position 8 Gln is present within the α-like toxins instead of the important Lys. But the superposition of LqhαIT and AaHII showed that Lys8 is equivalent to Lys28 in AaHII. In BomIII and LqhIII Lys28 and Lys30 are found and, therefore, it is proposed that one of these positively charged residues can substitute Lys8 in LqhαIT, comparable to the situation in toxin AaHII.

A difference between BomIII and LqhIII, which inhibits binding of LqhαIT in much lower concentrations than BomIII, is the positive charge of Lys40, which is part of a flexible β-turn. We propose that this difference is important for the higher binding affinity of LqhIII in comparison to BomIII. Looking at the alignment of LqhIII and LqhαIT, it is supposed that the α-like toxin shows positive potentials in the important areas **i** (α-amino group/Arg2), **b** (dependent on the structure Lys28 or alternatively Lys30), **c** (Lys64) and **d** (Lys40). This is in agreement with the potentials supposed to be important in classical and insect α-toxins for interaction with the voltage-gated sodium channel as well as inhibition of inactivation.

Conclusions

In this study differences and common features of classical and insect α-scorpion toxins were investigated. The charged residues and the molecular electrostatic potentials were of special interest, because of their importance for specificity and association of protein–protein complexes.

The insect α-toxin LqhαIT, which is not able to inhibit binding of AaHII in rat brain synaptosomes, shows very interesting differences in the molecular electrostatic potential in comparison to the classical α-toxins AaHII and LqqV. The large positive potential in area **a** is not present in LqhαIT. Instead it is partly overlaid by a negative potential mainly due to Glu24. The lack of equivalent residues for Lys30 and Arg56 not only influences the potential, but also prevents direct interactions with the channel. A second large negative potential was found in area **h** in LqhαIT due to residues Asp3, Asp19 and Glu15. Within the classical α-toxins only the conserved Asp3 forms a small, not extended negative potential. Both negative MEPs found in LqhαIT are proposed to lower the affinity of the toxin to mammalian sodium channels.

Large positive potentials were found for all α-toxins in areas **i**, **b**, **c** and **d**, and in most cases, functionally important residues contribute to these MEPs (classical toxins: Lys2, Arg56, Lys58; insect toxins: Lys8, Lys41, Lys62). It is conceivable that a large positive potential of a toxin bound at the loop, which connects the segments 3 and 4 in domain IV of sodium channels and which is therefore in close proximity to the voltage sensor in this domain, could influence the gating charge movement. For the classical α-toxins an additional large positive potential in region **a** should also be important. The importance of the residues in this part of the toxin structure is probably their contribution to the potential as well as a direct interaction with the channel.

The classical α-toxin AaHII is more active in mammals than LqqV. From the data presented here, several features could be important for this difference: (i) AaHII could have a conformational advantage, which possibly is a more beneficial conformation in the C-terminus and in addition in the 5-residue turn and the loop following the helix; (ii) the positive potential in area b is shifted in LqqV; (iii) the residues Lys30 and Arg56 of AaHII are in a more optimal position for direct interaction with the channel than Lys29 and Arg54 in LqqV.

References

1. Possani, L.D., Becerril, B., Delepierre, M. and Tytgat, J., Eur. J. Biochem., 264 (1999) 287.
2. Gordon, D., Savarin, P., Gurevitz, M. and Zinn-Justin, S., Toxicol. Toxin. Rev., 17 (1998) 131.
3. Rogers, J.C., Qu, Y., Tanada, T.N., Scheuer, T. and Catterall, W.A., J. Biol. Chem., 271 (1996) 15950.
4. Cestele, S., Qu, Y., Rogers, J.C., Rochat, H., Scheuer, T. and Catterall, W.A., Neuron, 21 (1998) 919.
5. Gordon, D., Martin-Eauclaire, M.F., Cestele, S., Kopeyan, C., Carlier, E., Khalifa, R.B., Pelhate, M. and Rochat, H., J. Biol. Chem., 271 (1996) 8034.
6. Sautiere, P., Cestele, S., Kopeyan, C., Martinage, A., Drobecq, H., Doljansky, Y. and Gordon, D., Toxicon, 36 (1998) 1141.
7. Kharrat, R., Darbon, H., Rochat, H. and Granier, C., Eur. J. Biochem., 181 (1989) 381.
8. Kharrat, R., Darbon, H., Granier, C. and Rochat, H., Toxicon, 28 (1990) 509.
9. Darbon, H., Jover, E., Couraud, F. and Rochat, H., Int. J. Pept. Protein Res., 22 (1983) 179.
10. El Ayeb, M., Darbon, H., Bahraoui, E.M., Vargas, O. and Rochat, H., Eur. J. Biochem., 155 (1986) 289.
11. Zilberberg, N., Froy, O., Loret, E., Cestèle, S., Arad, D., Gordon, D. and Gurevitz, M., J. Biol. Chem., 272 (1997) 4810.
12. Gilles, N., Krimm, I., Bouet, F., Froy, O., Gurevitz, M., Lancelin, J.M. and Gordon, D., J. Neurochem., 75 (2000) 1735.
13. Clackson, T. and Wells, J.A., Science, 267 (1995) 383.
14. Bogan, A.A. and Thorn, K.S., J. Mol. Biol., 280 (1998) 1.
15. Cunningham, B.C. and Wells, J.A., J. Mol. Biol., 234 (1993) 554.
16. Sheinerman, F.B., Norel, R. and Honig, B., Curr. Opin. Struct. Biol., 10 (2000) 153.
17. Bermann, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. and Bourne, P.E., Nucleic Acids Res., 28 (2000) 235. (http://www.rcsb.org/pdb/)
18. Smith, G.D., Blessing, R.H., Ealick, S.E., Fontecilla-Camps, J.C., Hauptman, H.A., Housset, D. and Langs, D.A., Acta Crystallogr., D53 (1997) 551.
19. Tugarinov, V., Kustanovich, I., Zilberberg, N., Gurevitz, M. and Anglister, J., Biochemistry, 36 (1997) 2414.
20. Menez, A., Toxicon, 36 (1998) 1557.
21. Philippopoulos, M. and Lim, C., Proteins, 36 (1999) 87.
22. Catterall, W.A., Neuron, 26 (2000) 13.
23. Bezanilla, F., Physiol. Rev., 80 (2000) 555.
24. Jiang, Y., Lee, A., Chen, J., Ruta, V., Cadene, M., Chait, B.T. and MacKinnon, R., Nature, 423 (2003) 33.
25. Jiang, Y., Ruta, V., Chen, J., Lee, A. and MacKinnon, R., Nature, 423 (2003) 42.
26. Schweitz, H., Vincent, J.P., Barhanin, J., Frelin, C., Linden, G., Hugues, M. and Lazdunski, M., Biochemistry, 20 (1981) 5245.
27. Khera, P.K., Benzinger, G.R., Lipkind, G., Drum, C.L., Hanck, D.A. and Blumenthal, K.M., Biochemistry, 34 (1995) 8533.
28. Li, H.M., Wang, D.C., Zeng, Z.H., Jin, J. and Hu, R.Q., J. Mol. Biol., 261 (1996) 415.
29. Housset, D., Habersetzer-Rochat, C., Astier, J.P. and Fontecilla-Camps, J.C., J. Mol. Biol., 238 (1994) 88.
30. INSIGHTII/DISCOVER/HOMOLOGY/DELPHI, Biosym Technologies, Inc./Molecular Simulations (MSI), San Diego, CA (Version 98.0 and 2000).

31. Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M., J. Appl. Crystallogr., 26 (1993) 283.
32. Kelley, L.A., Gardner, S.P. and Sutcliffe, M.J., Protein Eng., 10 (1997) 737.
33. Kelley, L.A., Gardner, S.P. and Sutcliffe, M.J., Protein Eng., 9 (1996) 1063.
34. Thompson, J.D., Higgins, D.G. and Gibson, T.J., Nucleic Acids Res., 22 (1994) 4673. (http://www2.ebi.ac.uk/clustalw)
35. Bairoch, A. and Boeckmann, B., Nucleic Acids Res., 19 Suppl (1991) 2247.
(http://www.expasy.ch/sprot/sptot-top.html)
36. Landon, C., Cornet, B., Bonmatin, J.M., Kopeyan, C., Rochat, H., Vovelle, F. and Ptak, M., Eur. J. Biochem., 236 (1996) 395.
37. Krimm, I., Gilles, N., Sautiere, P., Stankiewicz, M., Pelhate, M., Gordon, D. and Lancelin, J.M., J. Mol. Biol., 285 (1999) 1749.