Journal of Biomolecular NMR, 2 (1992) 57–70
ESCOM

J-Bio NMR 026

Toxin III of the scorpion *Androctonus australis* Hector: Proton nuclear magnetic resonance assignments and secondary structure

Afaf Mikou[a,*], Steven R. LaPlante[a,**], Eric Guittet[a], Jean-Yves Lallemand[a], Marie-France Martin-Eau Claire[b] and Hervé Rochat[b]

[a]*Laboratoire de R.M.N. ICSN CNRS, 1 Av. de la terrasse, F-91190 Gif-sur-Yvette, France*
[b]*Laboratoire de Biochimie – URA 1455 CNRS, Université d'Aix-Marseille II, Faculté de Médecine Secteur Nord, Bd Pierre Darmard, F-13326 Marseille Cedex 15, France*

Received 18 April 1991
Accepted 8 July 1991

Keywords: Scorpion; Toxin; NMR; Assignment; Structure; Protein

SUMMARY

^{1}H NMR has been applied to a **3.5 mM,** pH 5.4, solution of toxin III (64 amino acids) from venom of the scorpion *Androctonus australis* Hector. The resonance assignment strategy began by applying a generalized main-chain directed method for rapid identification and resonance assignments of secondary structures. The remaining resonances were assigned by the sequential method. Major structural features include a helix of $2\frac{1}{2}$ turns (residues 20–28) which is linked by two disulfide bridges to the central strand of a triple-stranded anti-parallel β-sheet. Turns were identified at residues 15–17, 47–49 and also at residues 51–53. Numerous NOEs have been observed between hydrophobic residues which suggest the presence of a hydrophobic core; these include Leu^{37}, Leu^{23}, Val^{47}, Tyr^{14}, Trp^{45} and Tyr^{5}. The Trp^{45} and Tyr^{5} rings lie orthogonal to one another. No crystal structure has been solved for this AaH III toxin. Comparisons are made with other members of the scorpion toxin family.

INTRODUCTION

The venom of scorpion is well known to contain a complex mixture of toxins. These toxins can display various degrees of toxicity towards different animal classes (Miranda et al., 1970; Watt

* To whom correspondence should be addressed.
** Current address: Bio-Méga Inc., 2100 rue Cunard, Laval, Québec, Canada H7S 2G5.

Abbreviations: NOE, nuclear Overhauser effect; NOESY, 2D NOE spectroscopy; DQF-COSY, double-quantum filtered correlated spectroscopy; HOHAHA, homonuclear Hartmann-Hahn spectroscopy; MCD, main-chain directed; AMX, spin system of three different protons; AaH, *Androctonus australis* Hector; CsE, *Centruroides sculpturatus* Ewing.

The *nomenclature* used is similar to that described by Wüthrich , 1986.

and Simard, 1984; Zlotkin et al., 1972a,b). Those which are preferentially directed towards mammals have been classified into two families according to their different pharmacological activities and binding sites (Couraud et al., 1982). The α-toxins, like toxin II of *Androctonus australis* Hector, bind in a voltage-dependent manner to the sodium channels of excitable membranes and thus prolong inactivation of sodium current. The β-toxins, like toxin II of *Centroides suffusus suffusus*, bind in a voltage-independent manner and enhance activation of inward sodium current (Catterall, 1977; Couraud et al., 1982; Wang and Strichartz, 1982; Darbon et al., 1983; Wheeler et al., 1983; Meves et al., 1984). Sequence homologies between the two families along with secondary structure prediction studies (Fontecilla-Camps, 1980; Fontecilla-Camps et al., 1988; Mikou et al., unpublished) suggest that there are structural features in common. Indeed, the crystal structures solved for CsE-v3 (variant 3 from *Centruroides sculpturatus* Ewing, Fontecilla-Camps et al., 1980) of the β-family and for AaH II of the α-family show common structural themes. The overall global folds are similar. Both have an antiparallel sheet, helix and three of four disulfide bridges in common. Some structural differences exist which may be responsible for the evolutionary divergence in activity. As expected, these are located at regions where the sequences differ. This is most notably apparent for the orientation of the loop protruding from the dense core and the C-terminal stretch of AaH II as compared to the corresponding regions in CsE-v3 (Fontecilla-Camps et al., 1988).

In order to better understand the structure–function relationships of the toxins, it would be interesting to make further structural comparisons between and within the toxin families. We plan to determine, by NMR and molecular modelling, the 3D solution structure of the toxin III of *Androctonus australis* Hector, which belongs to the α-family. In this paper, we present the first part of the project where the ^{1}H NMR assignments of the AaH III toxin have been completed and secondary structures and some global fold features are discussed.

MATERIALS AND METHODS

Toxin AaH III was isolated and purified from the scorpion venom of *Androctonus australis* Hector as described by Miranda et al. (1970) with recent modifications (Martin and Rochat, 1986). Its amino acid sequence has been determined previously (Kopeyan et al., 1979; Bougis et al., 1989). The 3.5 mM protein solution was unbuffered and contained 15 mg in 0.6 ml of either 90% H_2O/10% D_2O or 99.996% D_2O at pH 5.4.

^{1}H NMR spectra were recorded on Bruker AM600 and AM400 spectrometers and were processed on an Aspect 3000 computer. Some spectra were processed on a μVAX using the GIFA software which was developed in our laboratory (Delsuc, Robin and Stoven, unpublished). NOESY spectra (Jeener et al., 1979; Macura et al., 1981) and DQF-COSY spectra (Shaka and Freeman, 1983; Rance et al., 1984) were recorded in D_2O at 25°C with 100, 150 and 300 ms mixing times for the NOESY experiments. Spectra taken in 90% H_2O/10% D_2O include DQF-COSY at 25°C, NOESY (125, 150 and 300 ms mixing times) and HOHAHA spectra (13, 38, 50 and 70 ms spinlock times) (Braunschweiler and Ernst, 1983; Davis and Bax, 1985) were taken at 15, 25, 35, and 45°C. The large water resonance was suppressed by presaturation during the relaxation delay; for NOESY experiments presaturation was also applied during the mixing time. Spectra were acquired with 2 K data points in t2 and 512 points in t1 with spectral widths of 8024 Hz at 600 MHz and 5000 Hz at 400 MHz. All 2D spectra were recorded in pure phase absorption mode using the

time-proportional phase incrementation method (Redfield and Kuntz, 1975; Marion and Wüthrich, 1983). HOHAHA and NOESY data were apodized using a shifted cosine function in both dimensions; a pure sine function was applied to DQF-COSY data. All data were zero-filled in t1 before Fourier transformation. DQF-COSY data were additionally zero-filled in t2.

RESULTS

Proton resonance assignments were determined by a strategy which includes three steps. The first consists of taking DQF-COSY, NOESY and HOHAHA spectra in D_2O. This is mainly to distinguish resonances of nonlabile protons from those of labile protons. The second step involves applying a generalized MCD (main-chain directed) method for rapid identification and resonance assignment of the antiparallel β-sheet and helix secondary structures. The remaining resonances are assigned in the third step by the classical sequential method.

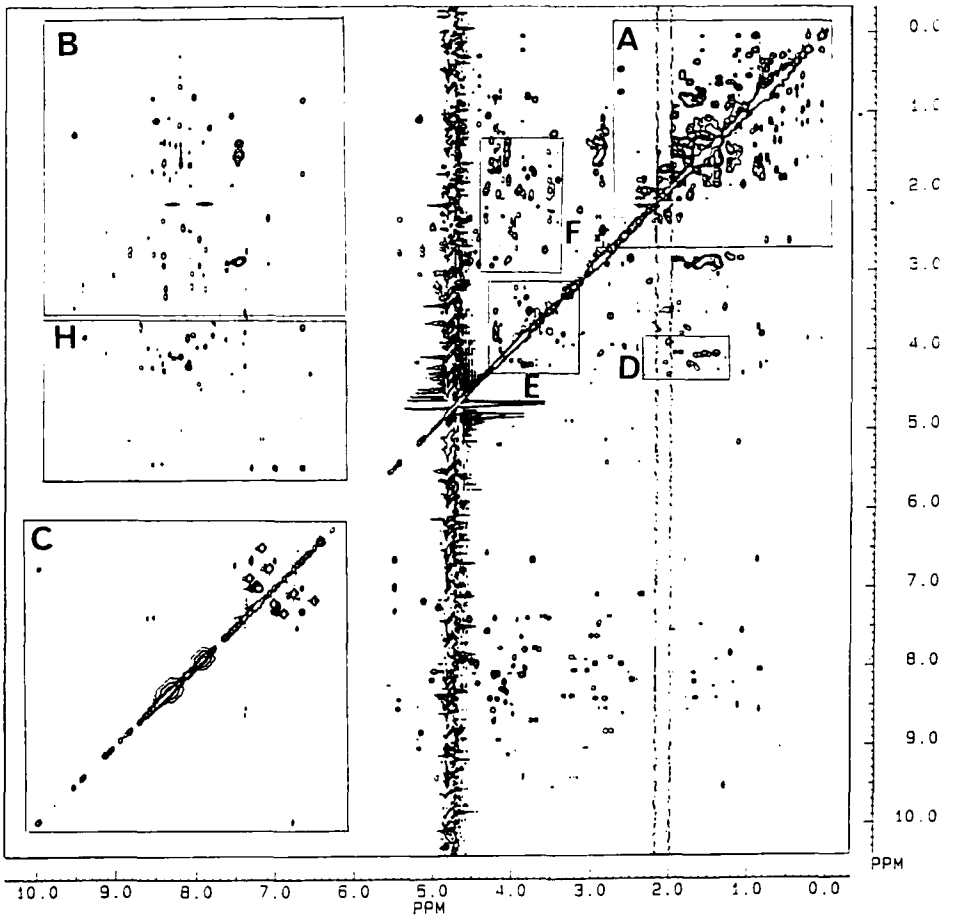

Fig. 1. The full HOHAHA spectrum at 600 MHz (70 ms and 25°C in H_2O) of AaH III scorpion toxin. Specific regions of interest are shown in boxes and labeled with letters.

Step I

^{1}H NMR spectra taken in H_2O contain the resonances of nonlabile and most labile protons of the protein. After lyophilization and subsequent addition of D_2O solvent, the labile protons exchanged completely with deuterium within a few hours. The remaining resonances in the low-field

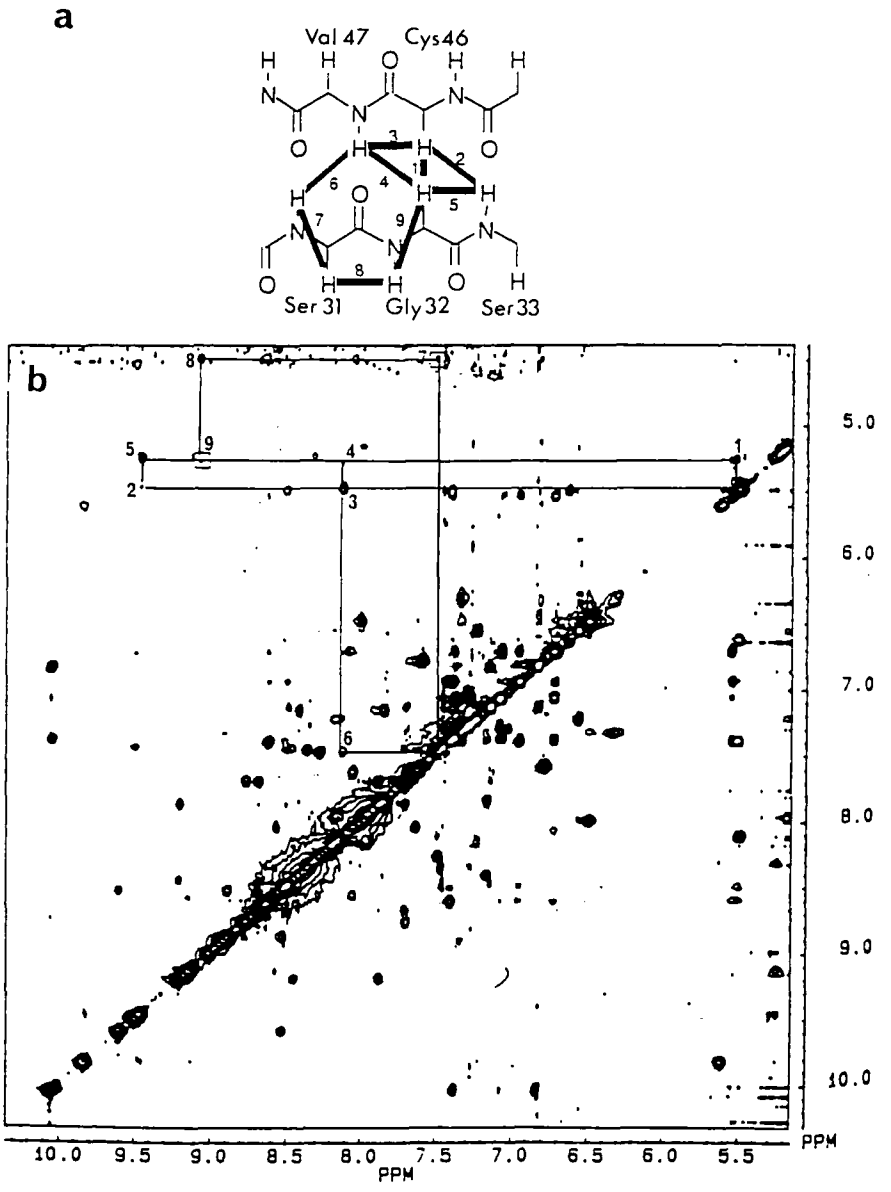

Fig. 2. Shown is a 400 MHz NOESY (150 ms) subspectrum of the αH, NH and aromatic resonances of AaH III in H_2O at 25°C. The characteristic signature of one inner loop begins at the αH-αH cross peak (peak 1). The characteristic MCD inner pattern for the antiparallel β-sheet shown as thick lines in A, corresponds to peaks 1, 2, 3, 4 and 5 shown in B. Cross peaks enclosed by a small square in B are also found as DQF-COSY cross peaks.

region were then easily identified as the nonlabile aromatic protons. DQF-COSY and HOHAHA spectra in D_2O allowed identification of the individual aromatic spin systems (see Wüthrich, 1986). Those resonances which had disappeared in D_2O solvent could be classified as main-chain amide protons and labile protons of aromatic and side-chain asparagine, lysine and arginine residues. The furthest downfield resonance at 10.05 ppm in H_2O solvent (Fig. 1) could be assigned as the labile $N^{\varepsilon}H$ proton of the single tryptophan residue due to its unique chemical shift (Wüthrich, 1986).

Step II

The second step in the assignment strategy entails identification and assignment of secondary structures. The generalized MCD method described elsewhere (Saudek et al., 1989; LaPlante et al., 1990) begins by locating MCD patterns (Di Stefano and Wand, 1987; Englander and Wand, 1987) which represent the characteristic NMR cross-peak signature of antiparallel β-sheets and helices. One advantage of this approach is that pattern searches focus directly on spin systems of residues involved in the secondary structures. This differs from the classical MCD strategy (Englander and Wand, 1987) where the NH-αH-βH spin systems of all residues need to be identified before conducting searches for MCD patterns.

Antiparallel β-sheet

The generalized MCD approach for antiparallel β-sheets begins by locating the readily apparent αH-αH NOESY cross peaks. These key cross peaks, which often are located downfield of the solvent resonance, allow targeting of spin systems of residues which are involved in the sheet. For example, Fig. 2B shows an αH-αH cross peak (peak 1) with its corresponding rectangle of alternating strong (peaks 3 and 5) and weak cross peaks (peaks 2 and 4) in the NH-αH region. This is

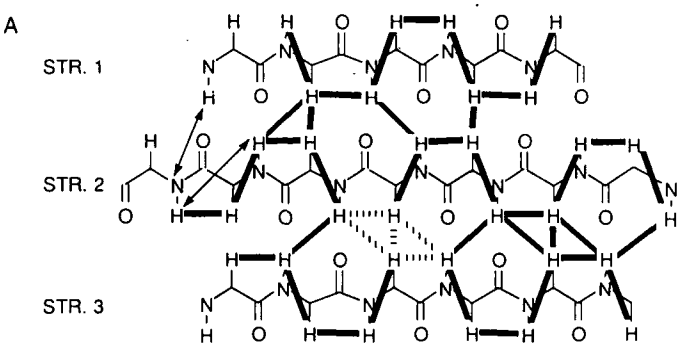

Fig. 3. Schematic representation of the triple-stranded antiparallel β-sheet of AaH III. The three strands (denoted as STR.) are indicated in A. The dashed lines show the inner loop MCD patterns as discovered from the NOESY spectrum of Fig. 2. The thick lines represent discovered NOEs and/or DQF-COSY cross peaks. Other important interstrand NOE connectivities are indicated by double-headed arrows. Amino acid assignments of the sheet are given in B.

the inner loop MCD pattern within the antiparallel β-sheet shown in Fig. 2A. In all, an extensive combination of DQF-COSY and NOESY cross peaks was found from spectra taken in H_2O and D_2O which could be represented as interconnecting full loop MCD patterns as shown in Fig. 3.

Once the loop patterns were discovered, they were then assigned to specific residues of the primary structure (Fig. 3B). The task was much simpler than imagined. Rather than having to define the complete spin systems of all the residues within the sheet, only the readily determinable spin systems such as valine, isoleucine, leucine, alanine and glycine had to be unequivocally identified. The remaining residues were classified as AMX (denoted as '*') or long chain (denoted as ' + ') spin systems. As a result, the three strands of the sheet were found to match peptide segments of the primary structure of AaH III scorpion toxin. For example, strand 1 (STR.1 of Fig. 3) was shown to consist of a peptide sequence of [(+)-(*)-(Gly)-(*)]. This was found to uniquely match [Arg^2-Asp^3-Gly^4-Tyr^5] in the toxin sequence. In a similar manner for strand 2 (STR. 2 of Fig. 3), a sequence of [(Ala)-(*)-(*)-(*)-(Val)-(Ala)-(Leu)] was found to uniquely correspond to [Ala^{43}-

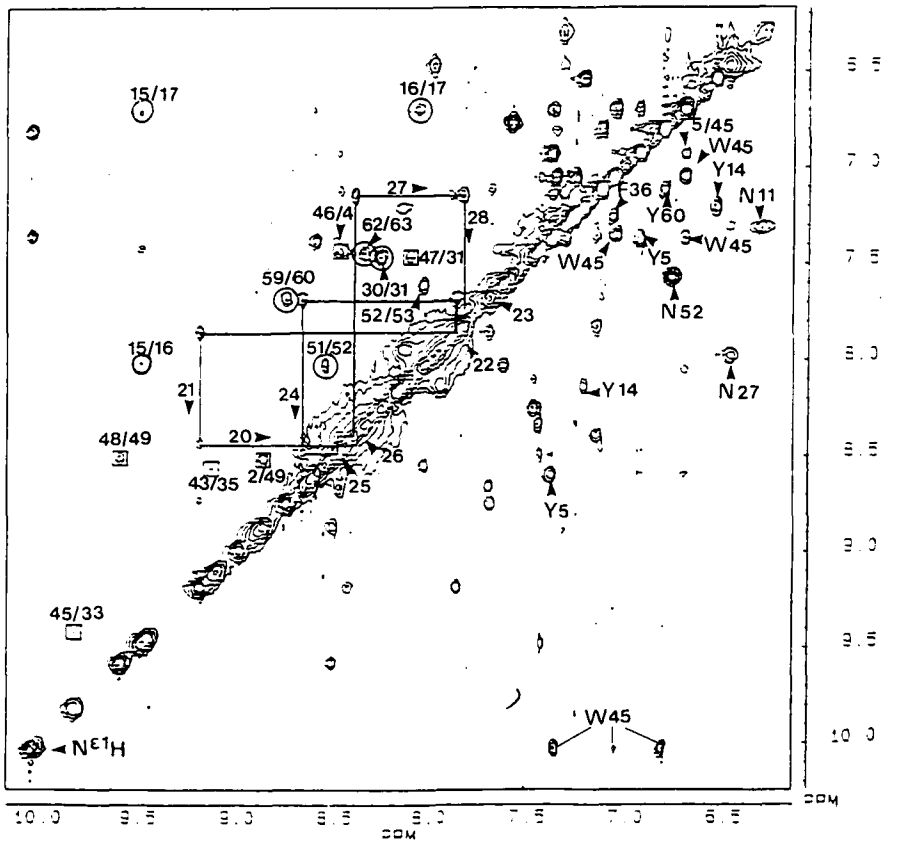

Fig. 4. Shown is the 400 MHz NOESY spectrum (150 ms) in H_2O at 25°C of the NH and aromatic region. The $N^{\varepsilon 1}H$ resonance of Trp^{45} at 10.05 ppm is indicated at the bottom left. The right side of the diagonal gives assignments of intraresidue connectivities of asparagines ($N^{\delta}H$) and aromatic side-chain resonances. An interresidue NOE is shown which involves the aromatic rings of Y5 and W45. On the left side of the diagonal, NOEs resulting from the helix are connected by lines and define the 'NOESY walk.' The squares correspond to NOEs observed in the β-sheet. Lastly, the circles show NOEs of possible turns.

Cys^{44}-Trp^{45}-Cys^{46}-Val^{47}-Ala^{48}-Leu^{49}]. Likewise, the sequence [(+)-(*)-(Gly)-(*)-(*)-(Gly)] of strand 3 matched [Lys^{30}-Ser^{31}-Gly^{32}-Ser^{33}-Cys^{34}-Gly^{35}].

Helix and turns

The generalized MCD approach for helices (Saudek et al., 1989; LaPlante et al., 1990) begins by locating the characteristic series of relatively strong, aligned NH-NH NOESY cross peaks. This key feature which is shown in Fig. 4 allows targeting of spin systems of residues which are involved in the helix. For instance, the NH-NH 'NOESY walk' in Fig. 4 was used as a starting point for an analogous 'NOESY walk' in the NH-βH region (not shown). This is possible due to the inherently short intraresidue and interresidue distances between NHs and βHs of helical structures. Medium-range NOEs were additionally found as a result of other short distances $d_{NN}(i,i+2)$, $d_{\alpha N}(i,i+3)$ and $d_{\alpha\beta}(i,i+3)$ (see Fig. 5). This further confirmed the presence of an α-helix of $2\frac{1}{2}$ turns. The residues of the helix (Fig. 4) could be assigned since the order of spin systems [(*)-(*)-(Gly)-(Leu)-(*)-(+)-(+)-(*)-(Gly)] uniquely matched the sequence [Cys^{20}-Asp^{21}-Gly^{22}-Leu^{23}-Cys^{24}-Lys^{25}-Lys^{26}-Asn^{27}-Gly^{28}].

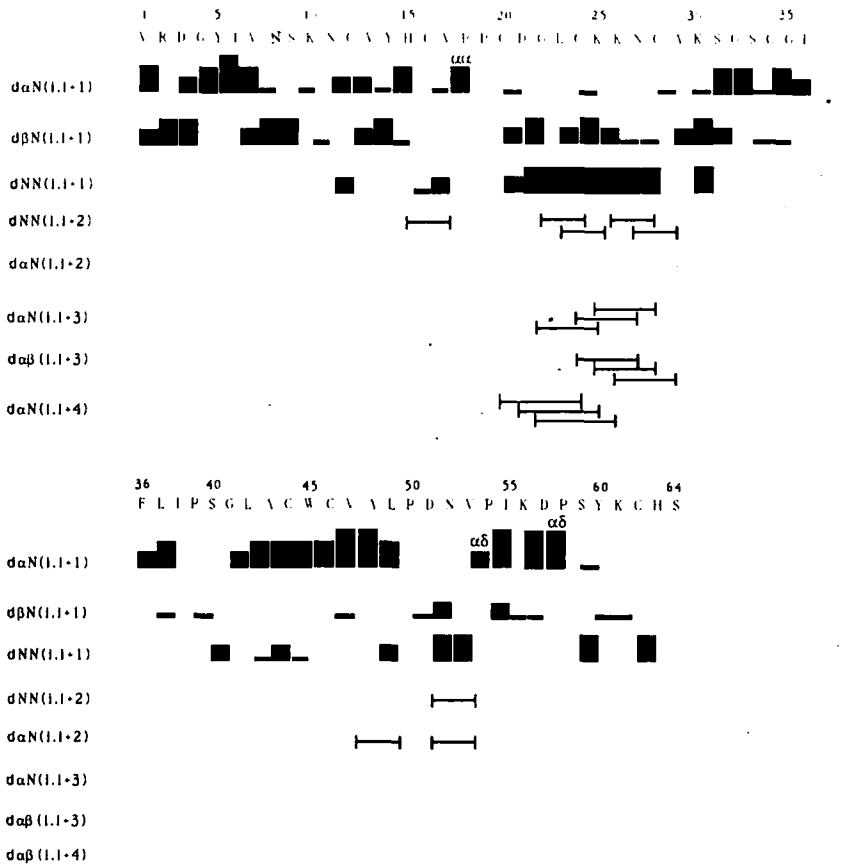

Fig. 5. Sequence of AaH III scorpion toxin together with a summary of the short-range interresidue NOEs involving the NH, αH and βH protons mostly from spectra taken at 25°C. The NOEs are reported as very strong, strong, medium or weak as shown by the thickness of the lines. Note that $d_{\alpha\delta}$, $d_{\alpha\alpha}$ connectivities for prolines are shown as $d_{\alpha N}$.

Turns also have characteristically short NH-NH distances and should give rise to a strong cross peak in the NOESY spectrum of Fig. 4 (Wagner et al., 1986). A preliminary analysis for turns is conducted where the remaining unattributed NOESY cross peaks in the NH-NH region (Fig. 4) are considered as possible indications of turns. These relatively strong NOEs resulting from $d_{NN}(i,i+1)$ type distances are shown as circles on the left side of the diagonal in Fig. 4. Medium-range distances such as $d_{NN}(i,i+2)$ and $d_{\alpha N}(i,i+2)$ are also pointed out. These characteristic NOEs allowed identification of turns at residues 15-17, 47-49 and 51-53. A more in-depth analysis was conducted upon completion of Step III.

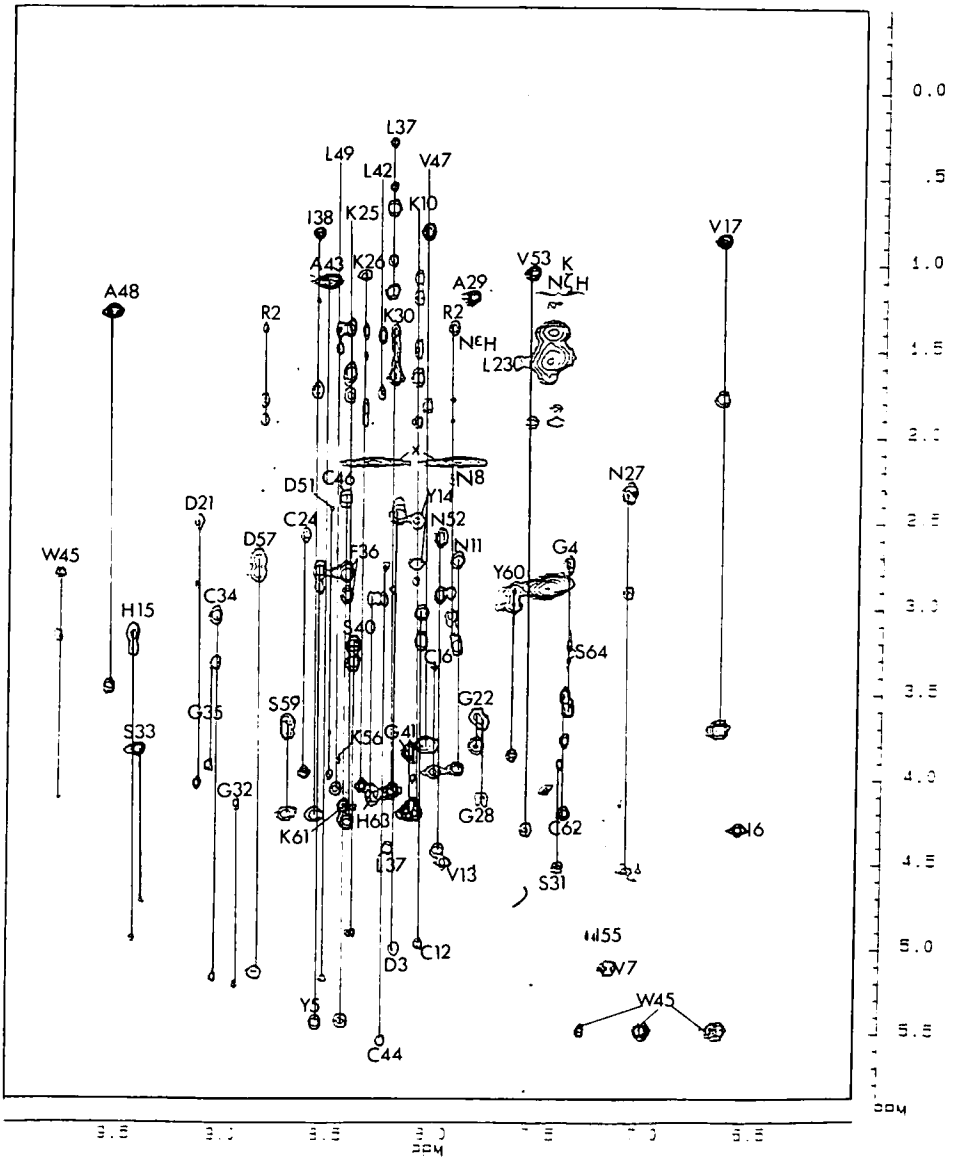

Fig. 6. NH (horizontal axis) to aliphatic (vertical axis) region of the HOHAHA spectrum (70 ms) of AaH III scorpion toxin in H_2O at 25°C. Each line represents a spin system. 'x' indicates an impurity.

Step III

A sequential method (Wüthrich, 1986) was applied to complete the resonance assignments. All the remaining spin systems of the HOHAHA spectrum such as the region shown in Fig. 6 are first classified as belonging to general residues types and then the spin systems of sequentially neighboring residues are identified using the NOESY spectrum.

Spin system classification

Given that the amino acids (except proline) have a backbone amide proton, regions should exhibit cross-peak correlations between the amide resonances with their respective side-chain resonances (regions B and H of Fig. 1). Hence, this relatively less crowded region can potentially contain a résumé of the information of the full spectrum. This is illustrated in Fig. 6 where the lines represent most of the correlations between the side-chain protons with their respective amide proton. Therefore, since toxin AaH III has 64 amino acids, Fig. 6 should contain a determinable number of lines (64 residues minus 6 prolines plus side-chain exchangeable protons).

It should be noted that not all the expected cross peaks were found in Fig. 6. This is due to factors such as incomplete magnetization transfer, overlap of the resonances and cross-peak suppression from irradiation of the H_2O resonance. Most of these problems were overcome by taking spectra at different temperatures. Varying the spinlock times in the HOHAHA experiments was also helpful. The lines in Fig. 6 could then be classified according to spin system types.

Like the AMX spin systems, the aromatic residues had three cross peaks along a line in Fig. 6 (NH to αH,βH and β'H), but in addition, relatively strong NOESY cross peaks were observed between the βH and some aromatic protons. This was not the case for His^{63} where no NOEs could be found between the main-chain and aromatic resonances. On the other hand, His^{15} had a weak NOE between δ2 and the amide resonance.

The favorable chemical shift dispersion within the methyl region (region A of Fig. 1) allowed identification of partial spin systems of valine, isoleucine and leucine. The most notable features were pairs of equally intense HOHAHA cross peaks between the βH proton and both methyls of valine residues and between the γH proton and both methyls of leucine residues. The isoleucine spin systems did not have this feature. In all cases, DQF-COSY spectra helped identify the spin systems.

The alanine residues gave rise to unusually sharp and intense HOHAHA cross peaks for the αH-βCH_3 (region D of Fig. 1) and NH-βCH_3 (region B of Fig. 1) correlations. All the alanines are involved in the secondary structures and were therefore assigned during Step II.

The glycines were identified by the large active coupling constants of the αH-α'H cross peaks in the DQF-COSY spectrum (not shown).

The asparagines were classified as AMX spin systems and additionally had NOEs between the βH proton and $N^{\delta}H$ and $N^{\delta'}H$ protons. The $N^{\delta}H$ and $N^{\delta'}H$ resonances had been independently identified since they gave rise to very strong NOE and HOHAHA cross peaks when in H_2O solvent, but not in D_2O solvent. The $N^{\delta}H$ and $N^{\delta'}H$ resonances were found in the aromatic region (Fig. 4).

Lysine and arginine are the only residues which belong to the long-chain class since there are no methionine, glutamic acid or glutamine residues in AaH III. The presence of numerous cross peaks along a line in Fig. 6 is a good indication of a long-chain spin system. In fact, we observed

two lines in Fig. 6 for each arginine spin system; one is from the amide proton (e.g. R2 of Fig. 6) and the other is from the $N^{\varepsilon}H$ side-chain proton (e.g. R2 $N^{\varepsilon}H$ of Fig. 6). Similar observations were apparent for the lysine residues, however, broad cross peaks were found along the $N^{\zeta}H$ line (e.g. K26 $N^{\varepsilon}H$ and K30 $N^{\varepsilon}H$ in Fig. 6).

Of all the amino acids, prolines spin systems were the most difficult to identify since they have no backbone amide proton and therefore are not found in Fig. 6. However, the δH to γH and βH cross peaks in the HOHAHA spectrum fall within a unique region (region F of Fig. 1). The αH protons often resonate close to the solvent peak (e.g. Pro^{39}) and could be observed by changing the temperature.

Sequential assignments

The sequential method is based on observation of NOEs which arise from short distances between neighboring residues such as $d_{\alpha N}(i,i+1)$ and $d_{NN}(i,i+1)$ (sometimes $d_{\beta N}(i,i+1)$). The strategy basically functions by deciphering which spin systems in the NMR maps belong to neighboring residues in the primary structure. Confident assignments are attained when a discovered series of neighboring spin systems is sufficiently long to uniquely match a peptide segment in the sequence (Wüthrich, 1986).

For this study, the sequential assignment process began by overlaying a plot of regions B and H of the HOHAHA spectrum in Fig. 1 (plotted on transparent paper) over a NOESY plot of the corresponding region. This allowed distinction between intraresidue and interresidue NOEs. The intraresidue NOEs were overlapped by HOHAHA cross peaks whereas the interresidue NOEs ($d_{\alpha N}(i,i+1)$, $d_{NN}(i,i+1)$ and $d_{\beta N}(i,i+1)$) were not superimposed by HOHAHA connectivities. Where possible, the assigned residues from Step II were used as confident starting points for attributing neighboring residues. Otherwise, peptide segments which corresponded uniquely to the primary structure were discovered as in the classical sequential assignment method (Wüthrich, 1986; Basus et al., 1988; Basus, 1989; Breg et al., 1989; Dyson et al., 1989; Kördel et al., 1989). These segments included Ile^{6} up to Tyr^{14}, Phe^{36} up to Leu^{42} and Pro^{54} to Pro^{58}. As no NOEs were found between His^{63} and Ser^{64}, this last residue was assigned by a process of elimination. Figure 5 summarizes the short range interresidue NOEs observed for AaH III.

In general, it was difficult to identify a continuous assignment pathway across the proline residues since they have no amide proton. Sequential NOE contacts such as $d_{\alpha\delta}(i,i+1)$ were observed between Val^{53} and Pro^{54} and also between Asp^{57} and Pro^{58}. This suggests that both Pro^{54} and Pro^{58} have trans conformations. Furthermore, a $d_{\alpha\alpha}(i,i+1)$ NOE was observed between Val^{17} and Pro^{18} which is consistent with Pro^{18} having a cis conformation. No interresidue NOEs were found between Pro^{19} and its neighboring amino acid and this residue was assigned by default. In summary, termination of the assignment process as described in these three steps was reassured by the fact that all spin systems were accounted for.

DISCUSSION

Structural aspects of AaH III

Upon completion of the resonance assignments of AaH III, numerous medium and long-range NOEs were identified. Although an intensive analysis will have to await calculation of a 3D structure (Mikou et al., unpublished), a qualitative description can be given here. As mentioned in the

Results section, NOEs have been found which clearly describe the presence of a triple-stranded sheet. It folds in such a way that strand 2 is disrupted by a turn at Ala^{48} (see Fig. 3). This turn forces strand 2 to approach strand 1 thus giving rise to a relatively strong NOE from Leu^{49} to both Ala^{2} and Ala^{48} as indicated by arrows in Fig. 3.

Strand 3 is disrupted after Lys^{30} by a folding back of the peptidyl chain and becomes an α-helix starting from Gly^{28} down to Cys^{20}. The helix lies parallel to the sheet as indicated by numerous NOEs with both strands 1 and 2. The disulfide bridge between Cys^{20} of the helix and Cys^{44} of strand 2 most likely dictates this fold. On the same face of the helix (about one turn away), there is another disulfide bridge between Cys^{24} of the helix and Cys^{46} of strand 2. The bridging of the secondary structures by the disulfide bonds certainly contributes to the stability of the molecular framework of AaH III. Another contributing factor is the presence of a hydrophobic core as evidenced by numerous NOEs between hydrophobic residues. This core is found roughly on the opposite side of the sheet as the helix and is discussed in more detail in the next section.

Structural comparisons with other scorpion toxins

Even though the scorpion toxins should largely have the same general structure given the extensive sequence homology, it is important to know whether or not the differences in the observed activity are due to specific conformational variations. Unfortunately, too few atomic-resolution studies have been completed to adequately respond to this question. The crystal structure of CsE-v3 has been solved (Fontecilla-Camps et al., 1980; Almassy et al., 1983). For this same toxin, ^{1}H NMR assignments along with some structural features in solution have been reported (Krishna et al., 1989; Nettesheim et al., 1989) and recently similar studies have been done for AaHiT, anti-insect toxin (Darbon et al., 1991). For the α-family, the crystal structure of AaH II has been solved. Thus, with additional structural information from this report concerning AaH III, some interesting comparisons can still be made.

The sheet and helix provide a common structural foundation since both are approximately the same for all the structures and are located in homologous sequential regions. The hydrophobic cluster most likely also plays a significant role in this stable framework. Interestingly, the crystal structures of CsE-v3 and AaH II along with the NMR study of CsE-v3 report an orthogonal arrangement of aromatic rings within the hydrophobic cluster. In our study of AaH III, we observe an orthogonal type of arrangement between the Trp^{45} and Tyr^{5} rings. This orthogonality is suggested by strong NOEs observed between protons ε3 and ζ3 of Trp^{45} with aromatic protons of Tyr^{5} while no NOEs were observed between Tyr^{5} with η2 and ζ2 of Trp^{45}. This also explains the unusual upfield shift observed for ε3 of Trp^{45} which would be induced by a ring-current effect from Tyr^{5}. Although the orthogonal rings were observed in all the toxin structures (termed the 'herringbone' arrangement), the number of rings involved differs. In CsE-v3, the Tyr^{4}, Tyr^{47}, Tyr^{42}, Tyr^{40}, Tyr^{38} and Tyr^{58} rings were orthogonal with respect to one another. For AaH II, the same distribution was observed for Tyr^{5}, Tyr^{47}, Tyr^{42} and Tyr^{49} aromatic rings. In this study, we report an orthogonal arrangement for Tyr^{5} and Trp^{45}. It is apparent that the orthogonal arrangements in AaH II and AaH III are smaller than in CsE-v3; this would be due to the fact that both AaH toxins do not have aromatic residues at homologous positions of the sequence as compared to CsE-v3. Interestingly, it is possible that the hydrophobic arrangement of AaH II may be somewhat conserved in AaH III since NOEs were observed which suggest a Tyr^{5}, Trp^{45}, Val^{47} arrangement. The hydrophobic valine residue seemed to have substituted the aromatic ring found in AaH

TABLE I
^{1}H NMR CHEMICAL SHIFTS OF THE AaH III SCORPION TOXIN AT 25°C AND pH 5.4[a]

		Chemical shifts				
Residue	NH	$C^{\alpha}H$	$C^{\beta}H$	$C^{\gamma}H$	$C^{\delta}H$	others
Val[1]		4.20	1.62	0.92, 0.83		
Arg[2]	8.89	4.96[+]	2.00, 1.88	1.47, 1.47	3.14, 3.01	8.00 ($N^{\varepsilon}H$)
Asp[3]	8.26	5.08	2.48, 2.52			
Gly[4]	7.46	3.70, 2.84				
Tyr[5]	8.63	5.51	2.97, 2.78		7.42, 7.42	6.97, 6.97($C^{\varepsilon}H$)
Ile[6]	6.63	4.38	1.82	1.32*, 1.32*	0.91*	0.72 ($C^{\gamma}H3$)
Val[7]	7.24	5.21	1.92	0.67, 0.61		
Asn[8]	8.00	4.92[+]	3.24, 2.34			
Ser[9]	8.17		4.20, 4.07			
Lys[10]	8.15	4.27	1.73, 2.01	1.27, 1.16	1.73, 1.58	2.91, 2.85 ($C^{\varepsilon}H$), 7.54 ($N^{\zeta}H$)
Asn[11]	7.97	4.03	3.32, 2.82		7.35, 6.33 ($N^{\delta}H$)	
Cys[12]	8.14	5.05	3.28, 3.12			
Val[13]	8.03	4.58	2.69	0.84, 0.56		
Tyr[14]	8.18	4.53	2.85, 2.59		7.24, 7.24	6.58, 6.58 ($C^{\varepsilon}H$)
His[15]	9.49	5.02	3.34, 3.25			8.61 ($C^{\varepsilon 1}H$), 7.46 ($C^{\delta 2}H$)
Cys[16]	8.08	4.02	3.85, 3.42			
Val[17]	6.72	3.80	1.87	0.95, 0.95		
Pro[18]		4.22	3.60, 3.60	2.12, 1.91	2.40, 2.19	
Pro[19]						
Cys[20]	8.46	4.22	2.78, 2.39			
Asp[21]	9.20	4.11	2.94, 2.59			
Gly[22]	7.88	3.90, 3.72				
Leu[23]	7.70	3.95	1.67, 1.21	0.99	0.32, 0.14	
Cys[24]	8.68	4.04	2.67, 2.63			
Lys[25]	8.49	4.34	1.88, 1.88	1.48, 1.48	1.71, 1.71	3.01, 3.01 ($C^{\varepsilon}H$), 7.53 ($N^{\zeta}H$)
Lys[26]	8.43	4.13	1.99, 1.91	1.49, 1.49	1.70, 1.70	2.98, 2.98 ($C^{\varepsilon}H$), 7.54 ($N^{\zeta}H$)
Asn[27]	7.17	4.61	3.00, 2.42	8.01, 6.50		
Gly[28]	7.86	4.23, 3.78				
Ala[29]	7.91	4.70	1.28			
Lys[30]	8.28	4.15	1.75, 1.75	1.49, 1.49	1.58, 1.58	2.98, 2.98 ($C^{\varepsilon}H$), 7.54 ($N^{\zeta}H$)
Ser[31]	7.49	4.58	3.98, 3.98			
Gly[32]	9.01	5.28, 4.24				
Ser[33]	9.47	5.01[+]	3.93, 3.93			
Cys[34]	9.11	5.24	3.41, 3.13			
Gly[35]	9.16	4.02, 3.45				
Phe[36]	8.49	5.11[+]	3.11, 2.90		7.29, 7.29	7.09, 7.09 ($C^{\varepsilon}H$)
Leu[37]	8.29	4.48	1.25, 0.75	1.08	0.64, 0.39	
Ile[38]	8.63	4.29	1.81	1.32, 0.87	0.75	0.91 ($C^{\gamma}H3$)
Pro[39]		4.45[+]	3.51, 3.49	2.20, 1.96	1.62, 1.62	
Ser[40]	8.46	4.99	3.41, 3.32			
Gly[41]	8.19	4.29, 3.92				
Leu[42]	8.35	4.99[+]	1.85, 1.50	1.85	1.00, 0.99	
Ala[43]	8.59	5.26	1.20			
Cys[44]	8.31	5.63	3.02, 2.84			
Trp[45]	9.84	4.62	3.26, 2.88			10.05 ($N^{\varepsilon 1}H$), 7.38 ($C^{\zeta 2}H$), 6.83 ($C^{\delta 1}H$), 7.08 ($C^{\eta 2}H$), 6.73 ($C^{\zeta 3}H$), 5.57 ($C^{\varepsilon 3}H$)

TABLE I (continued)

Residue	NH	Chemical shifts				
		$C^{\alpha}H$	$C^{\beta}H$	$C^{\gamma}H$	$C^{\delta}H$	others
Cys^{46}	8.51	5.52	2.88, 2.45			
Val^{47}	8.12	3.89	1.90	0.88, 0.88		
Ala^{48}	9.61	3.54	1.38			
Leu^{49}	8.54	4.13	1.65, 1.65 ×	1.46	0.74, 0.67	
Pro^{50}		4.38	1.99, 1.99	2.01, 2.01	3.98, 3.98	
Asp^{51}	8.57	4.08	2.53, 2.48			
Asn^{52}	8.05	4.48	2.98, 2.52		7.60, 6.80 ($N^{\delta}H$)	
Val^{53}	7.63	4.38	2.02	1.12, 1.12		
Pro^{54}		4.26	2.17*, 2.10*	1.80*, 1.80*	3.94, 3.83	
Ile^{55}	7.34	5.01	1.73	1.18*, 0.89*	0.80	0.47 ($C^{\gamma}H3$)
Lys^{56}	8.52	3.97				
Asp^{57}	8.80	5.19	2.88, 2.80			
Pro^{58}		4.41	2.47, 2.37*	2.27*, 2.06	4.00, 3.76	
Ser^{59}	8.77	4.37	3.81, 3.76			
Tyr^{60}	7.70	5.01 +	3.08, 3.00		7.15, 7.15	6.84, 6.84 ($C^{\varepsilon}H$)
Lys^{61}	8.50	4.22				
Cys^{62}	7.47	4.27	3.87, 3.52			
His^{63}	8.38	4.12	3.02, 3.02			8.70 ($C^{\varepsilon 1}H$), 7.47 ($C^{\delta 2}H$)
Ser^{64}	7.45		3.38, 3.30			

[a] Chemical shifts are relative to the H_2O resonance which is 4.78 ppm at 25°C relative to DSS. A '+' indicates that the chemical shift was taken from spectra at 45°C. A '*' indicates that the assignment can be interchanged with another assignment within the same residue which also has a '*'. A '×' means that the chemical shift could be 1.65 or 1.46.

II. On one side of the ring of Trp^{45}, NOEs were observed between ε1 and ζ1 of Trp^{45} with a methyl of Val^{47} whereas on the other side of the ring NOEs were observed from η2 and ζ2 of Trp^{45} with Tyr^{5}.

Others have pointed out that differences in conformation exist between AaH II and CsE-v3 which may be the source of the variance in activity (Fontecilla-Camps et al., 1988). This was the orientation of the loop protruding from the dense core and C-terminal stretch of AaH II as compared to the corresponding region in CsE-v3. At this point in our analysis of AaH III, no definitive comparisons can be made concerning these regions.

Other studies have shown that biotinylation of lysine residues of AaH II resulted in various degrees of loss of pharmacological activity and binding to synaptosomes (Darbon et al., 1983; Fontecilla-Camps et al., 1988). It was pointed out that Lys^{58} is the most important lysine with regard to toxicity of AaH II. The same was observed when Lys^{56} of AaH I was carboxymethylated (Sampieri and Habersetzer-Rochat, 1978). The crystal structure of AaH II shows that Lys^{58} is relatively buried within the protein and anchored down by hydrogen bonds between $N^{\varepsilon}H$ of Lys^{58} and the carbonyls of Asn^{11} and Gly^{61} (Fontecilla-Camps et al., 1988). The modification of Lys^{56} of AaH III has also been shown to provoke a drastic loss of pharmacological activity and binding to synaptosomes (Darbon, personal communication). Although it is not absolutely definitive, we have observed an NOE between $N^{\varepsilon}H$ of Lys^{56} and $N^{\delta}H$ of Asn^{11} which would be consistent with Lys^{56} being buried and hydrogen bonded as found for AaH II.

ACKNOWLEDGEMENTS

We thank Dr. D. Guedin and ROUSSEL UCLAF for providing the protein. We also thank Dr. M.-A. Delsuc for valuable discussions.

REFERENCES

Almassy, R.J., Fontecilla-Camps, J.C., Suddath, F.L. and Bugg, C.E. (1983) *J. Mol. Biol.*, **170,** 497–527.
Basus, V.J. (1989) In *Methods in Enzymology, Vol.* **177,** *part B,* (Eds, Oppenheimer, N.J. and James, T.L.), p. 137.
Basus, V.J., Billeter, M., Love, R.A., Stroud, R.M. and Kuntz, I.D. (1988) *Biochemistry* **27,** 2763–2771.
Bougis, P.E., Rochat, H. and Smith, L.A. (1989) *J. Biol. Chem.*, 19259–19265.
Braunschweiler, L. and Ernst, R.R. (1983) J. Magn. Reson., **53,** 521–528.
Breg, J.N., Boelens, R., George, V.E. and Kaptein, R. (1989) *Biochemistry,* **28,** 9826–9833.
Catterall, W.A. (1977) *J. Biol. Chem.*, **252,** 8660–8668.
Couraud, F., Jover, E., Dubois, J.M. and Rochat, H. (1982) *Toxicon,* **20,** 9–16.
Darbon, H., Jover, E., Couraud, F. and Rochat, H. (1983) *Int. J. Pept. Prot. Res.*, **22,** 179–186.
Darbon, H., Weber, C. and Braun, W. (1991) *Biochemistry,* **30,** 1836.
Davis, D.G. and Bax, A. (1985) J. Am. Chem. Soc., **107,** 2821–2822.
Di Stefano, D.L. and Wand, A.J. (1987) *Biochemistry,* **26,** 7272–7281.
Dyson, H.J., Holmgren, A. and Wright, P.E. (1989) *Biochemistry,* **28,** 7074–7087.
Englander, S.W. and Wand, A.J. (1987) *Biochemistry,* **26,** 5953–5958.
Fontecilla-Camps, J.C. (1980) Ph. D. Thesis, University of Alabama, Birmingham.
Fontecilla-Camps, J.C., Almassy, R.J., Suddath, F.L., Watt, D.D. and Bugg, C.E. (1980) *Proc. Natl. Acad. Sci. USA,* **77,** 6496–6500.
Fontecilla-Camps, J.C., Habersetzer-Rochat, C. and Rochat, H. (1988) *Proc. Natl. Acad. Sci. USA,* **85,** 7443–7447.
Jeener, J., Meier, B.H., Bachmann, P. and Ernst, R.R. (1979) *J. Chem. Phys.*, **71,** 4546–4553.
Kopeyan, C., Martinez, G. and Rochat, H. (1979) *Eur. J. Biochem.*, **94,** 609–615.
Kördel, J., Forsén, S. and Chazin, W.J. (1989) *Biochemistry,* **28,** 7065–7074.
Krishna, N.R., Nettesheim, D.G., Klevit, R.E., Drobny, G., Watt, D.D. and Brugg, C.E. (1989) *Biochemistry,* **28,** 1556–1562.
LaPlante, S.R., Mikou, A., Robin, M., Guittet, E., Delsuc, M., Charpentier, I. and Lallemand, J.-Y. (1990) *Int. J. Pept. Prot. Res.*, **36,** 227–230.
Macura, F., Huang, Y., Suter, D. and Ernst, R.R. (1981) *J. Magn. Reson.*, **43,** 259–281.
Marion, D. and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.*, **113,** 967–974.
Martin, M.F. and Rochat, H. (1986) *Toxicon,* **24,** 1131–1139.
Meves, H., Rubly, N. and Watt, D.D. (1984) *Pflüger's Arch.*, **402,** 24.
Miranda, F., Kupeyan, C., Rochat, H., Rochat, C. and Lissitzky, S. (1970) *Eur. J. Biochem.*, **16,** 514.
Nettesheim, D.G., Klevit, R.E., Drobny, G., Watt, D.D. and Krishna, N.R. (1989) *Biochemistry,* **28,** 1548–1555.
Rance, M., Sorensen, O.W., Bodenhausen, O., Wagner, G., Ernst, R.R. and Wüthrich, K. (1984) *Biochem. Biophys. Res. Commun.*, **117,** 479–485.
Redfield, A.G. and Kuntz, S.D. (1975) *J. Magn. Reson.*, **19,** 250–254.
Sampieri, F., and Habersetzer-Rochat, C., (1978) *Biochim. Biophys. Acta,* **535,** 100–109.
Saudek, V., Atkinson, R.A., Williams, R.J.P. and Ramponi, G. (1989) *J. Mol. Biol.*, **205,** 229–239.
Shaka, A.J. and Freeman, R. (1983) *J. Magn. Reson.*, **51,** 169–173.
Wagner, G., Neuhaus, D., Worgotter, E., Vasaak, M., Kagi, J.H.R. and Wüthrich, K. (1986) *J. Mol. Biol.*, **187,** 131–135.
Wang, G.K. and Strichartz, G.R. (1982) *Biophys. J.*, **40,** 175–179.
Watt, D.D. and Simard, J.M. (1984) *J. Toxicol. Toxin Rev.*, **3,** 181–221.
Wheeler, K.P., Watt, D.D. and Lazdunski, M. (1983) *Pflüger's Arch.*, **397,** 164–165.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids,* Wiley, New York.
Zlotkin, E., Miranda, F. and Lissitsky, S. (1972a) *Toxicon,* **10,** 207.
Zlotkin, E., Miranda, F. and Lissitsky, S. (1972b) *Toxicon,* **10,** 211.