Static Accessibility Model of Protein Antigenicity: The Case of Scorpion Neurotoxin

Jiří Novotný* and Edgar Haber

Molecular and Cellular Research Laboratory, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114

Received April 3, 1986; Revised Manuscript Received June 13, 1986

ABSTRACT: Scorpion neurotoxins are a family of homologous, 64 to 65 residue containing proteins with four invariant disulfide bridges. Previous experimental work established four antigenic epitopes in the *Androctonus australis* neurotoxin and localized them in the amino acid sequence. Using crystallographic coordinates of the *Centruroides sculpturatus* neurotoxin and computing its large sphere (radius 1 nm or 10 Å) accessibility profiles, we identified six antigenic sites clustered into four surface regions. Three of four computed sites coincided with the epitopes identified and localized experimentally in the *A. australis* neurotoxin, while two of the computed sites partially overlapped the fourth epitope. To investigate the relationship between antigenicity and segmental flexibility, 8-ps molecular dynamics simulations were performed on the *C. sculpturatus* structure, average backbone temperature factors computed from the simulation, and results compared with the X-ray-derived *B* values. Most of the neurotoxin structure and, in particular, three of the four antigenic sites were found inflexible, as judged by the computed and/or crystallographic temperature factors. The remaining epitope was associated with only marginal above-average maxima of backbone *B* values, corresponding to root mean square atomic displacements of 0.5 Å (50 pm). We conclude that neurotoxin antigenicity is determined by an exceptional surface exposure of relatively short loop segments and that segmental flexibility is not an essential component of antigenicity.

Proteins are capable of eliciting antibody response in foreign, i.e., nonself, organisms. The elicited antibody molecules specificaly recognize areas on protein surfaces approximately 3–4 nm^2 (300–400 Å^2) large (Kabat, 1978; Amit et al., 1986). These areas are referred to as antigenic sites or epitopes.

Some of the molecular characteristics of antigenic epitopes emerged from accumulated experimental data. Sela and coworkers (Sela, 1969) demonstrated that antiprotein antibodies can be elicited by synthetic peptides with amino acid sequences corresponding to segments of the parent protein. They also demonstrated that antibodies generally recognize a "single", well-defined conformation of a peptide; for example, antibodies against the intact, disulfide-bonded "loop" peptide of lysozyme (residues 64–82 of hen egg white lysozyme) did not react with the peptide in which the disulfide bond had been reduced (Arnon et al., 1971).

Using a panel of polyclonal antipeptide antibodies, Atassi (1975, 1978) described what seemed to be the complete collection of antigenic epitopes on myoglobin and lysozyme. Only a few discrete antigenic sites emerged from his studies, implying that certain regions of surface are more antigenic than others. A more recent interpretation (Benjamin et al., 1984; Berzofsky, 1985) of newer experimental evidence on the basis of data obtained from both polyclonal and monoclonal antibodies indicated, however, that many more mutually overlapping epitopes exist on protein surfaces than inferred by Atassi and that perhaps the whole surface can be antigenic. It should be noted that some antigenic sites correspond to a single continuous polypeptide chain segment, whereas others consist of amino acid residues that were not contiguous in the amino acid sequence (composite or discontinuous epitopes).

Westhof et al. (1984) noted that the continuous antigenic sites usually have higher than average Debye–Waller temperature factors. The (isotropic) crystallographic temperature factor, or the *B* value, is given by $B = (8/3)\pi^2\langle r^2\rangle$, where $\langle r^2\rangle$ is the mean square atomic displacement from the crystal equilibrium position (Ladd & Palmer, 1985). Thus, it has been hypothesized that flexibility is an essential component of antigenicity (Tainer et al., 1985), perhaps because structural adjustments are important for antibody complementarity. However, the *B* values are known to be significantly influenced by conditions such as crystal packing (Sheriff et al., 1985) and static crystal disorder (Finzel & Salemme, 1985) and do not necessarily reflect the dynamical aspects of the structure, such as segmental flexibility, in their pure form. Indeed, when Moudallal et al. (1985) reexamined some of the experimental data presented by Westhof et al. (1984), no clear-cut correlation of antigenicity with segmental flexibility was evident in the tobacco mosaic virus protein.

Padlan (1985) introduced a numerical measure of "immunogenic potential" based on evaluation of structural dissimilarity of amino acid residues near antigenic sites. In his model, then, the antigenic sites are those regions that are the most "dissimilar" from the rest of the structure. The predictive value of his method seems to be comparable to that of Fraga (1982) or Hopp and Woods (1981), which identifies solvent-exposed hydrophilic loops as the most probable locations of antigenic epitopes. A recent test of the Hopp and Woods predictive algorithm reported a success rate of only 56% (Tanaka et al., 1985).

Three research groups (Novotný et al., 1986; Fanning et al., 1986; Thornton et al., 1986) recently suggested that antigenicity of polypeptide chain segments is derived from their exceptional surface exposure, making them accessible to contacts with large antibody domains. In our previous paper (Novotný et al., 1986), surface areas on proteins accessible to contacts with a large spherical probe were evaluated by using the Lee and Richards (1971) surface algorithm; the probe size (1 nm or 10 Å in radius) was comparable to that of antibody domains that contain antigen-combining sites. All the reported antigenic sites corresponded to segments particularly accessible to a large sphere. In myoglobin and cyto-

0006-2960/86/0425-6748$01.50/0

chrome *c*, virtually all of the van der Waals surface was accessible to the large probe and therefore potentially antigenic, whereas in myohemerythrin (Tainer et al., 1984), distinct large probe inaccessible surface regions were apparent, which also were the regions that were not antigenic.

It has been known that surface exposure and flexibility are strongly correlated (Petsko & Ringe, 1984; Bennett & Huber, 1984; Karplus & McCammon, 1983); thus, if one of these properties is essential to antigenicity, the other one will show a strong correlation, too, without even being requisite to antigenicity. In order to distinguish between the two possibilitites, namely, that (1) flexibility, but not accessibility, is essential or that (2) accessibility, but not flexibility, is essential, we analyzed antigenicity, large probe accessibility, and segmental flexibility of neurotoxins from scorpion species *Centruroides sculpturatus* Ewig and *Androctonus australis* Hector. The neurotoxins are ideally suited for this type of analysis: they are smaller than other protein models used to date, the antibody response against the *A. australis* toxin and its peptide fragments has been extensively studied (El Ayeb et al., 1983, 1984; Bahraoui et al., 1986), and the three-dimensional structure of the *C. sculpturatus* has been solved (Almassy et al., 1983) to a high resolution.

Methods of Computation

Cartesian atomic coordinates of *Centruroides sculpturatus* neurotoxin, refined at 1.8-Å (180 pm) resolution (Almassy et al., 1983), were obtained from the Brookhaven Protein Data Bank (Bernstein et al., 1977). The crystallographic residual (*R* factor) for this structure is 16%, and the data contain the isotropic *B* factor values for all the (nonhydrogen) atoms of the structure. The structure was manipulated for structural analysis, energy minimization, dynamics simulation, drawing, etc. with the program CHARMM version 16 (Brooks et al., 1983). An explicit hydrogen atom representation was used, with all the polar hydrogen atoms explicitly constructed by the program; nonpolar hydrogens were combined with the carbon atoms to which they were attached into "extended" atoms. Stereo drawings were prepared with the use of the Unified Graphics software package (Robert C. Beach, Stanford Linear Accelerator Center) as described before (Novotný et al., 1984). Alternatively, the ARPLOT (Lesk & Hardman, (1982) and PLT2 (D. States, personal communication) programs were used to generate pictures. Large-sphere accessibility was computed by using the Lee and Richards (1971) algorithm as originally implemented in CHARMM and subsequently modified to include large probe radii by Dr. R. E. Bruccoleri (Novotný et al., 1986).

The crystallographic coordinates were energy minimized by the adopted basis Newton–Raphson method (Brooks et al., 1983) using a variant of the constrained minimization protocol of Bruccoleri and Karplus (1986). After the initial 100 cycles of hydrogen atom adjustment with the positions of the carbon, nitrogen, and sulfur atoms kept fixed, 400 cycles of ABNR minimization followed. Harmonic constraints of 82 kJ (20 kcal) were assigned to all the atoms and were periodically reassigned to the current atomic positions every 50 cycles. The lists of hydrogen-bonded and nonbonded interactions were recompiled every 25 cycles.

The molecular dynamics simulation protocol employed was essentially that of Northrup et al. (1980). The simulation was done at room temperature (300 K) with the program CHARMM version 16. Altogether, it consisted of 50.856 ps of dynamics with a time step size 0.978 fs. The Verlet (1967) algorithm was used to integrate the equations of motion for all the carbon, oxygen, nitrogen, and polar hydrogen atoms in the energy-minimized neurotoxin structure. Lengths of bonds that involved hydrogen atoms were kept fixed by use of the "SHAKE" algorithm (Ryckaert et al., 1977). A cut-off distance of 8 Å (800 pm) was imposed on evaluations of all the nonbonded interactions, and the lists of hydrogen-bonded and nonbonded interactions were recompiled every 10 cycles. The heating phase of the simulation involved 11.736 ps of 2.5° temperature increments every 97.8 fs, spanning the temperature range from 0 to 300 K. During each temperature increase, the velocities of atoms were reassigned according to a Gaussian distribution. When the temperature of the molecule reached 300 K, the equilibration phase was started. The velocities of the atoms were reassigned every 0.195 ps, for the total time interval of 7.824 ps, to the Gaussian distribution of the final temperature. Following an additional 11.736 ps of an unperturbed equilibration, 19.56 ps of the dynamics simulation was accumulated, i.e., velocities and positions of the atoms were periodically recorded every 0.195 ps. Root mean square averages, $\langle r^2 \rangle^{1/2}$, of all the atomic positions were computed from the last 7.824 ps, and the average temperature factors of the backbone atoms N, Cα, C, and O were then obtained as $B = (8/3)\pi^2\langle r^2 \rangle$.

Results and Discussion

Neurotoxin from C. sculpturatus and Energy Minimization of Its Structure. Most of the scorpion neurotoxins are proteins of 64 or 65 residues, with 4 disulfide bridges. The crystal structure solved by Almassy et al. (1983) is that of *C. sculpturatus* variant 3 consisting of 65 amino acid residues (492 nonhydrogen atoms), with disulfide bonds between residues 12–65, 16–41, 25–46, and 29–48. The structure, as schematically depicted in Figures 1 and 2, consists of a dense, hydrogen-bonded core of three β-strands and an α-helix; two shorter segments of polypeptide chain in an extended conformation append to the sides of the β-sheet by hydrogen bonding.

The in vacuo potential energy of the original crystallographic structure, as evaluated by the CHARMM program, was 236 kJ (56.4 kcal). After 400 cycles of energy minimization the potential energy of the structure improved to −2571 kJ (−614.4 kcal), corresponding to an average stabilization energy of −4.3 kJ/atom (there are 492 heavy atoms and 109 hdyrogen atoms in the CHARMM-constructed structure). The small root mean square shift in atomic positions upon minimization, namely, 0.33 Å (33 pm), indicated that the original crystallographic structure was very good and that acceptable values of potential energy can be achieved by small adjustments of coordinates. The main purpose of energy minimization was to eliminate the possibility of small areas of energetic strain (e.g., partial atomic overlaps in the crudely determined parts of the structure) that would lead to undesirably large forces in the initial steps of dynamics simulation.

Antigenicity of A. australis Neurotoxin Compared to Large Probe Accessibility of the C. sculpturatus Structure. The first step of our analysis consisted of computing the large probe (radius 1 nm or 10 Å) accessibility profile of the *C. sculpturatus* X-ray structure, and its comparison with the experimentally determined antigenic sites of the *A. australis* neurotoxin. The comparison requires an assumption that antigenicity of the two closely homologous structres is virtually identical, and it is worthwhile to review briefly the arguments on which the assumption was based. First, the assumption is implicit (albeit tacitly) in all the previous abundant experimental work that reported on inhibition of protein antigen–antibody reactions by closely homologous protein antigens [e.g., the ternary reaction system of hen egg white lysozyme, turkey

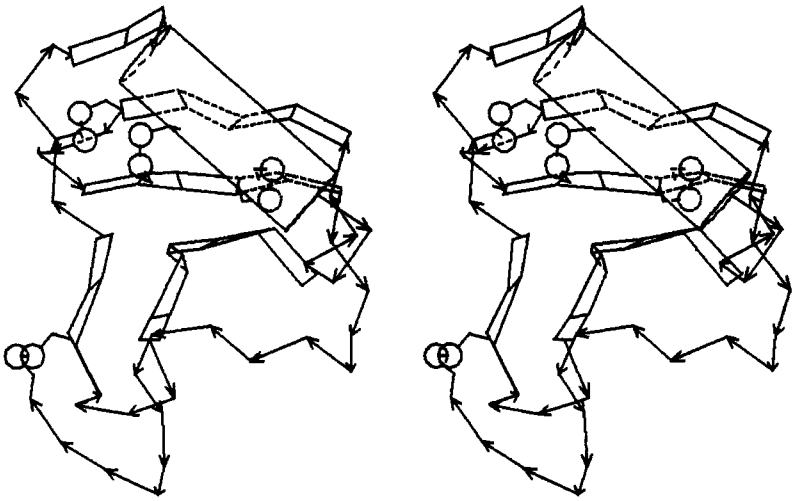

FIGURE 1: Schematic stereo diagram of three-dimensional structure of scorpion neurotoxins. The N → C polarity of polypeptide bonds is indicated by arrows, ribbons indicate β-strands, and cylinders indicate α-helix. The four cystine residues are schematically represented as balls and sticks. The diagram was prepared from the crystallographic coordinates of the *C. sculpturatus* neurotoxin (Almassy et al., 1983) by using the computer program of Lesk and Hardman (1982).

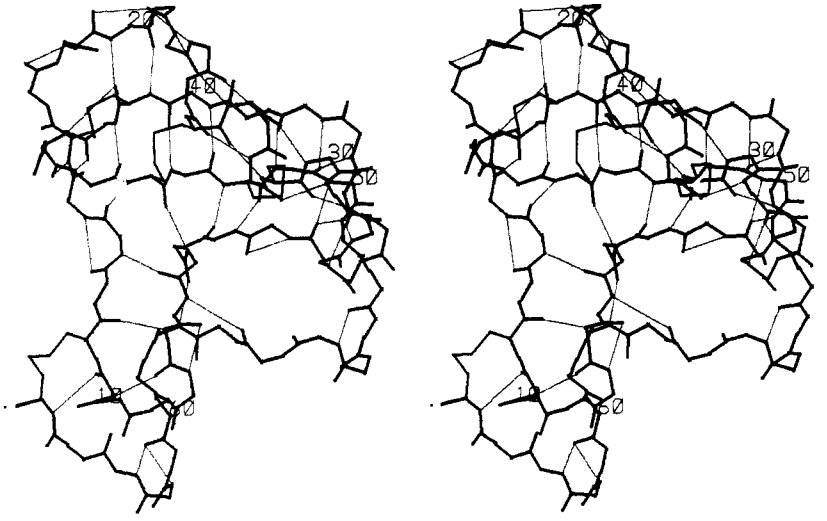

FIGURE 2: Stereo drawing of the *C. sculpturatus* neurotoxin polypeptide backbone. Backbone-to-backbone hydrogen bonds, as generated by the program CHARMM (Brooks et al., 1983), are drawn in light lines. The Cα atoms of every tenth residue are numbered.

lysozyme, and anti-(hen egg white lysozyme) antibody]. Second, Novotný et al. (1986) have previously shown that large sphere accessibility profiles of bonito and tuna fish cytochrome *c* are virtually identical, so that the profile obtained on one molecule can be directly extrapolated to a homologous molecule. Third, the amino acid sequence homology between the *C. sculpturatus* and *A. australis* neurotoxins is 44% identical residues. In families of homologous proteins, such a degree of sequence identity usually guarantees identical tertiary structures (Dayhoff et al., 1978). For example, the β-sheets of β_2 microglobulin and immunoglobulin CH3[1] domain can be superimposed with an average rms distance 0.9 Å (90 pm, i.e., less than the length of a carbon–carbon bond; Becker & Reeke, 1985), although their amino acid sequences have only 22% identical residues. Since the scorpion neurotoxins are relatively small and the fold of their chain is maintained by the four invariant disulfide bridges, it is safe to conclude that their overall molecular shapes, as well as their gross antigenic properties (although perhaps not the details), are virtually identical.

El Ayeb et al. (1983) reported that the stoichiometry of neurotoxin–antineurotoxin Fab complexes was 1:4 when rabbit polyclonal antibodies toward three different toxin molecules, including the variant 2 toxin of *A. australis*, were studied. Thus, either only four antibody species from a polyclonal serum can bind at a time because of steric factors or no more than four antigenic epitopes are present on the neurotoxin molecules. El Ayeb et al. (1984) presented the following supportive evidence for the second of the above possibilities. Two peptides representing the segments 50–59 and 5–14–SS–60–64 (disulfide bonded) of the *A. australis* sequence were synthetized and used to isolate, by affinity chromatography, subpopulations of rabbit antineurotoxin antibodies specific for the peptides. Each of the two isolated site-specific IgG populations had association constants for their respective peptides in the nanomolar range (≈3 nM), their indices of heterogeneity being 0.9. In addition, the two populations were shown to be specific to one antigenic site only by the multideterminant effect on the slopes of binding curves [method of Berzofsky et al., (1976)], and their functional independence was further demonstrated by the lack of hindrance or cooperativity between them. Thus, the two regions of neurotoxin structure around the disulfide bridge 12–SS–63 and spanning residues 50–59 were unequivocally identified as independent antigenic epi-

[1] Abbreviations: CH3, the third immunoglobulin heavy chain constant region; Fab, antigen-binding fragment of antibody; rms, root mean square; IgG, immunoglobulin G.

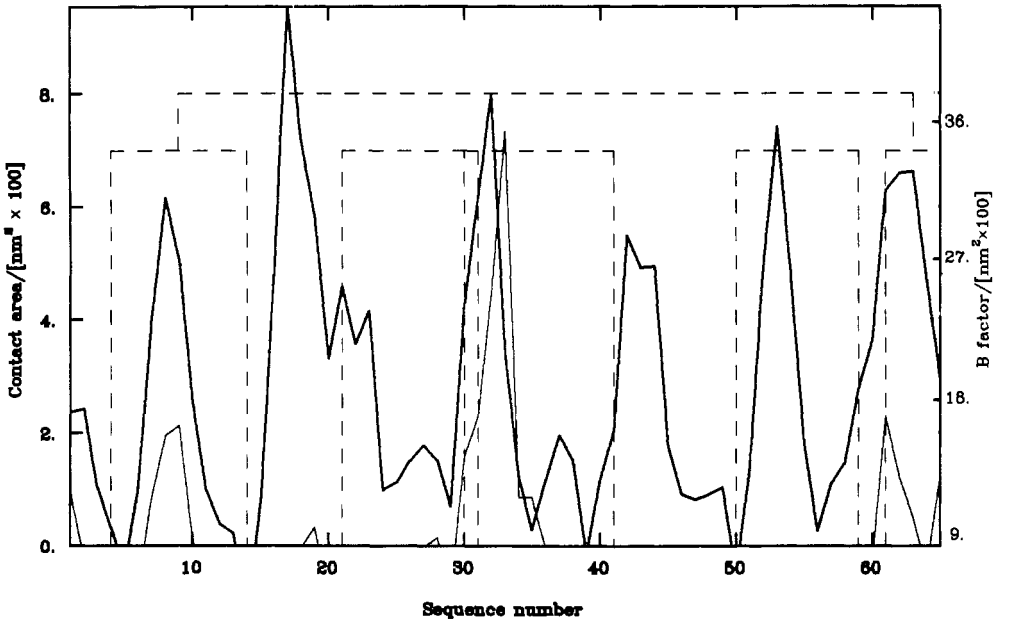

FIGURE 3: Large probe accessibility (antigenicity) profile and average backbone temperature factors of the X-ray-derived 3D structure of the *C. sculpturatus* neurotoxin. Heavy line, the smoothed contact area profile computed with the 1-nm radius probe (Novotný et al., 1986); light line, average crystallographic *B* values of the backbone atoms N, C, Cα, and O. The average backbone value is 8.6 Å^2. The dashed boxes indicate positions of four experimentally determined antigenic epitopes of the *A. australis* neurotoxin structure, i.e., residues 22–31, 30–41, 50–59, and 5–14–S–6–64 in the *C. sculpturaturs* neurotoxin sequence.

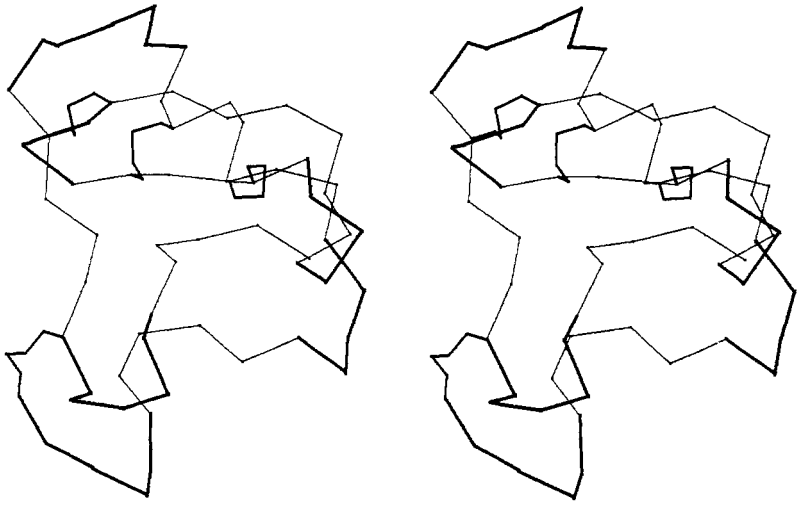

FIGURE 4: Stereo diagram showing locations of the six large probe accessibility maxima, as indicated in Figure 3, in the three-dimensional structure of the *C. sculpturatus* neurotoxin. Cα backbone atoms are connected by virtual bonds (light lines), disulfides are indicated by medium lines, and the computed antigenic sites by heavy lines. It can be seen that the six large probe accessibility maxima cluster into four spacial regions (epitopes) located approximately at the four apices of the tetrahedrally shaped neurotoxin fold.

topes. In a subsequent communication, Bahraoui et al. (1986) used a similar methodology to demonstrate that two additional segments of the *A. australis* sequence, namely, residues 19–29 and 28–39, also include major antigenic epitopes of this toxin.

Figure 3 gives the large probe accessibility profile computed from the *C. sculpturatus* three-dimensional structure (solid heavy line), together with the location of the four epitopes identified by El Ayeb et al. (1984) and Bahraoui et al. (1986) on the surface of the *A. australis* toxin (light broken line). Correspondence between four of the maxima present in the profile and three experimentally identified epitopes is striking and lends further support to our suggestion (Novotný et al., 1986) that antigenicity and large probe accessibility are strongly correlated. The fourth antigenic peptide, spanning residues 19–29, overlaps two of the computed maxima, particularly the one located on residue 18. Altogether there are six distinct maxima in the profile (Figure 3), and their spatial arrangement on the surface of the *C. sculpturatus* is indicated in Figure 4. Disulfide bonds occurring between cysteine residues 15–65 and 16–41 bring two pairs of the computed antigenic segments close together in space, thus reducing the number of spacially independent antigenic regions to four. The shape of the neurotoxin molecule is roughly tetrahedral, and the six computed antigenic sites occupy the four apices of its molecular contour.

Our computations predict that two segments, residues 16–23 and 42–45 of the *C. sculpturatus* sequence, respectively, belong among those of the highest antigenic potential. Antigenicity of these regions is yet to be studied experimentally, although the segment 16–23 partially overlaps the peptide 19–29 of the *A. australis* sequence studied by Bahraoui et al. (1986).

Small contact area peaks at positions 1–4, 25–28, and 36–37 (cf. the profile in Figure 3) might be indicative of additional epitopes located around these positions. These peaks are

considerably smaller than the major maxima discussed above and may represent epitopes of weak antigenicity. In the absence of more experimental data, their significance remains unclear. For example, the antigenic potential of the peptide 19–29 described by Bahraoui et al. (1986) can be explained either by its coincidence with the small computed maximum centered at residue 27 or by its overlap with the large contact area peak centered on residue 18.

Crystallographic B Factors of Computed Antigenic Epitopes. Next we discuss segmental flexibility of the known and predicted antigenic sites, as it emerges from the X-ray-derived *B* values. In the following paragraph, we then compare the crystallographic temperature factors with the *B* values obtained from dynamical simulations on the crystallographic structure.

The solid light line in Figure 3 traces the *B* values of backbone atoms obtained by averaging the X-ray-derived temperature factors of the N, Cα, C, and O atoms in each residue. Almassy et al. (1983) have presented a similar graph in Figure 6a of their original paper and some of the following discussion merely repeats points made by them earlier. First, it is to be noted that the *C. sculpturatus* neurotoxin crystals are characterized by lower *B* values than, e.g., crystals of lysozyme, myoglobin, or myohemerythrin. The majority of average backbone *B* values in neurotoxin fluctuate between 6–7 Å^2 (0.06–0.07 nm^2), compared to all-backbone averages of sperm whale myoglobin and human lysozyme 10 Å^2 (0.10 nm^2) and 17 Å^2 (0.17 nm^2), respectively. Either the neurotoxin molecule is more rigid (presumably due to its high disulfide content) or its crystals are more regularly packed than those of the two other structures, or both.

Figure 3 shows the epitopes 19–29 and 50–59 as not being associated with any significant above-average *B* maximum. These epitopes could thus be considered as rigid, although the segment 50–59 is marginally involved in contacts with surrounding molecules in the crystal (Almassy et al., 1983) and may become more flexible in solution. The epitope 12–SS–64 coincides with a small above-average maximum of *B* values, the highest *B* value of the epitope being 15 Å^2 (0.15 nm^2) at residue 61. This value corresponds to a root mean square atomic displacement of about 0.5 Å (50 pm). Such a displacement is rather small, on the order of one-third of the length of a carbon–carbon bond (1.53 Å or 153 pm), and hardly indicative of concerted atomic motions (i.e., segmental flexibility) of any significance. The most significant *B* factor maximum of Figure 3 is the large peak centered on residue 33 (B = 29 Å^2 or 0.29 nm^2). This exceptionally high temperature factor coincides with the antigenic site 28–39 (cf. Figure 3), and its possible significance has been discussed by Almassy et al. (1983) who concluded "...that the loop defined by residues 30 to 34 either displays large thermal motion or is disordered". Subsequent unpublished refinement of the *C. sculpturatus* structure to 1.3 Å (130 pm; Dr. Charles Bugg, personal communication) has identified two alternative conformations adopted by this loop. Originally, Almassy et al. (1983) observed that one molecule of the solvent 2-methyl-2,4-pentanediol cocrystallized with each molecule of neurotoxin, "tightly bound...as evidenced by atomic temperature factors that are of the same average magnitude as those of large side-chain atoms of the protein.... The two hydroxyl groups (of methylpentanediol) are hydrogen-bonded to side chains of Lys-8 and Asn-33 from a symmetry related protein molecule.... It appears that only one of the enantiomers is bound at this site.... The alternate enantiomer might be included in the model; ...the electron density maps...are not sufficiently sensitive to eliminate this possibility, at the present resolution.... However...the contact to Asn-33 would be unreasonably short if the 5-methyl group were interchanged with the (enantiomeric) 4-hydroxyl group (Almassy et al., 1983)."

Apparently, the above mentioned crystal contacts caused slightly different positions of the Asn-33 loop in different molecules, giving rise to a local crystalline disorder. It thus seems that the conspicuously high *B* values of this loop are of static, rather than of dynamic, origin.

C. sculpturatus Temperature Factors from Molecular Dynamics Simulations. In order to obtain independent evidence about the nature and magnitude of segmental flexibility in the *C. sculpturatus* neurotoxin, we computed temperature factors from room temperature dynamical simulations of the isolated molecule in vacuo and compared them with those obtained by Almassy et al. (1983) from the crystallographic data.

In molecular dynamics simulations, equations of motion for all the atoms are being simultaneously solved, on the basis of all the forces acting on the atoms and being specified by the potential empirical energy function (McCammon et al., 1977; Brooks et al, 1983). Instantaneous atomic positions are recorded in regular time intervals, and a series of computer files containing the positions at various times constitute the simulated dynamical trajectory. Average atomic positions and their root mean square deviations can be readily obtained from the trajectories. Rms fluctuations, $\langle r^2 \rangle^{1/2}$ are then related to the crystallographic *B* factors as $B = (8/3)\pi^2\langle r^2 \rangle$. Previous work has shown that the rms atomic fluctuations obtained from dynamical simulations agreed well with the experimental temperature factors in all the proteins studied, namely, the cytochrome *c* (Northrup et al., 1980), myoglobin (Levy et al., 1985), pancreatic trypsin inhibitor, crambin, ribonuclease, lysozyme (Levitt et al., 1985; Post et al., 1986), and L7/L12 ribosomal protein (Åquist et al., 1985). Levitt et al. (1985) used the method of normal-mode dynamics, rather than the dynamical trajectories, to compute the *B* factor values.

The *B* factor values of the *C. sculpturatus* neurotoxin backbone were computed from the last 7.824 ps of dynamical simulations. As described under Methods of Computation, no energy adjustments were applied during this part of the simulation and, although the kinetic energy of the molecule fluctuated in a typical manner around a stable average value of 315 K, the total energy of the system remained constant indicating that the system has reached thermal equilibrium. The computed average backbone temperature factors are displayed in Figure 5. The crystallographic average backbone *B* values are also included for comparison.

It can be readily seen that positions and amplitudes of the computed maxima in positions 9 and 61 of the amino acid sequence coincide rather well with the crystallographic values. However, four small maxima appear in the computed temperature factors around positions 14, 15, 18, and 43–44 of the sequence that are absent from the crystallographic *B* value profiles. Conversely, the major maximum in position 33, present in the crystallographic data, does not appear in the computed profile. We explain these differences as follows. Positions 18 and 42 and their immediate vicinity (residues 14 and 15 included) were implicated by Almassy et al. (1983) in crystal contacts. They are likely to be parts of segments that are somewhat flexible in solution but become constrained in the crystal. The exceptionally high maximum centered on residue 33 occurs in a region affected by crystalline disorder (cf. discussion above). As it does not exist in the computed profiles, it is reasonable to assume that it is, indeed, of static origin.

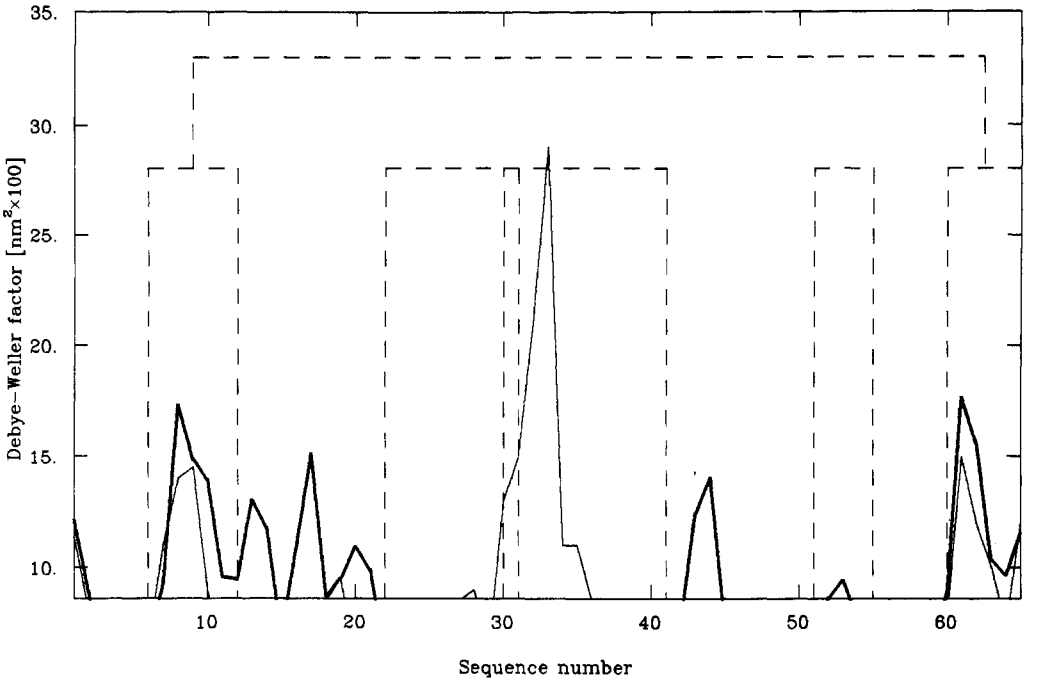

FIGURE 5: Comparison of temperature factors (Debye–Weller or B values given in Å^2) computed from dynamics simulations with those derived from X-ray crystallography (Almassy et al., 1983). Heavy line, the average backbone B values computed from dynamics simulations; light line, values obtained from crystallographic data. Dotted lines indicate locations of the four antigenic epitopes depicted in Figure 3.

In summary, the computed B values confirm our previous conclusions concerning the relative rigidity of the epitopes 19–29, 28–39, 50–59, and 12–SS–64. That is, the epitopes 19–29, 28–39, and 50–59 are completely rigid, while the 12–SS–64 epitope is characterized by a marginal segmental flexibility, on the order of 0.5 Å (50 pm). It is very likely that amplitudes of the motions computed in vacuo will be significantly damped in solution by frictional forces (van Gunsteren & Karplus, 1981), resulting in average displacements even smaller than 50 pm.

Conclusions

Previous experiments have established that scorpion neurotoxins contain four antigenic epitopes (El Ayeb et al., 1983, 1984; Bahraoui et al., 1986). Using the available crystallographic coordinates of *C. sculpturatus* neurotoxin (Almassy et al., 1983), we have found six antigenic regions with use of our large probe accessibility method. Four of the epitopic regions computed by us coincide with the three experimental sites, while the fourth experimental site overlaps two of the epitopes computed by us.

Cumulative evidence from X-ray crystallography and room temperature dynamical simulation in vacuo suggest that polypeptide chain backbones of all the antigenic sites are fairly rigid. That is, two of them, segments 19–29 and 50–59 (which contains two proline residues), appear inflexible by all the available criteria; the epitope 28–39 appears rigid by dynamical simulation, although it is associated with an exceptionally high B value of static origin in the crystallographic structure; and the remaining epitope displays only marginal flexibility, as judged both by crystallographic and computed B values. Thus, it is difficult to escape the conclusion that protein antigenicity is synonymous with an exceptional surface exposure of protein segments and that segmental flexibility is not an essential component of antigenicity.

Theoretical considerations suggest that this conclusion is likely to be valid generally. The measured binding constant of an antigen–antibody reaction, K_{as}, relates to the Gibbs free energy of binding, ΔG, as $-RT \log K_{as} = \Delta G = \Delta H - T\Delta S$ (T, temperature; R, gas constant; H, enthalpy; and S, entropy of the reaction). Entropy S involves number of degrees of freedom available to the system, and since its change enters into the free energy balance with a negative sign, an amount of Gibbs energy on the order of $T\Delta S_{conf}$ would always be required to compensate for the S_{conf} entropy units of conformational freedom lost (immobilized) in the process of binding, whenever the K_{as} is to remain favorable. Consequently, an excessive conformational freedom of antigens and antibodies is always directed against, and not in favor of, the strength and specificity of binding [see, e.g., the article by Sachs et al. (1972) for a detailed theoretical analysis of the conformational equilibria in antigen–antibody reactions]. Indeed, the highest binding constants have been observed in antibodies directed toward small, rigid organic molecules such as fluorescein (K_{as} in picomolar range; Kranz et al., 1982) or digoxin (K_{as} of 1 nM to 10 pM, Mudgett-Hunter et al., 1985), while antibodies against molecules endowed with many degrees of freedom (oligosaccharides, peptides) typically have binding constants that are significantly lower [in the micromolar range, cf. Kabat (1966), Arnon et al. (1971), Haber et al. (1976), and Dyson et al. (1985)].

Acknowledgments

We are indebted to John Newell, Head of the Cardiac Computer Center at the Massachusetts General Hospital, for his generous support. We thank Dr. Charles Bugg (University of Alabama) and Drs. Hervé Rochat and C. Granier (CNRS Laboratory of Biochemistry, Marseille) for communicating to us their results prior to publication. We are indebted to Dr. Robert E. Bruccoleri (Massachusetts General Hospital) for his comments on the manuscript.

References

Almassy, R. J., Fontecilla-Camps, J. C., Suddath, F. L., & Bugg, C. E. (1983) *J. Mol. Biol. 170*, 497–527.

Amit, A. G., Mariuzza, R. A., Phillips, S. E. V., & Poljak, R. J. (1986) *Science (Washington, D.C.) 233*, 747–753.

Åquist, J., van Gunsteren, W. F., Leijonmarck, M., & Tapia, O. (1985) *J. Mol. Biol. 183*, 461–477.
Arnon, R., Maron, E., Sela, M., & Anfinsen, C. B. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 1450–1455.
Atassi, M. Z. (1975) *Immunochemistry 12*, 423–438.
Atassi, M. Z. (1978) *Immunochemistry 15*, 909–936.
Bahraoui, E., El Ayeb, M., Van Rietschoten, J., Rochat, H., & Granier, C. (1986) *Mol. Immunol. 23*, 357–366.
Becker, J. W., & Reeke, G. N., Jr. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 4225–4229.
Benjamin, D. C., Berzofsky, J. A., East, I. J., Gurd, F. R. N., Hannum, C., Leach, S. J., Margoliash, E., Michael, J. G., Miller, A., Prager, E. M., Reichlin, M., Sercarz, E. E., Smith-Gill, S. J., Todd, P. E., & Wilson, A. C. (1984) *Annu. Rev. Immunol. 2*, 67–101.
Bennett, W. S., & Huber, R. (1984) *CRC Crit. Rev. Biochem. 15*, 291–384.
Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Price, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. J. (1977) *J. Mol. Biol. 112*, 535–542.
Berzofsky, J. A. (1985) *Science (Washington, D.C.) 229*, 932–940.
Berzofsky, J. A., Curd, J. G., & Schechter, A. N. (1976) *Biochemistry 15*, 2113–2121.
Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187–217.
Bruccoleri, R. E., & Karplus, M. (1986) *J. Comput. Chem. 7*, 165–175.
Dayhoff, M. O., Barker, W. C., Hunt, L. T., & Schwartz, R. M. (1978) in *Atlas of Protein Sequence and Structure*, Suppl. 3, pp 9–24, National Biomedical Research Foundation, Georgetown University Medical Center, Washington, D.C.
Dyson, H. J., Cross, K. J., Houghten, R. A., Wilson, I. A., Wright, P. E., & Lerner, R. A. (1985) *Nature (London) 318*, 480–483.
El Ayeb, M., Martin, F., Delori, P., Bechis, G., & Rochat, H. (1983) *Mol. Immunol. 20*, 697–708.
El Ayeb, M., Bahraqui E. M., Granier, C., Delori, P., Van Rietschoten, J.,& Rochat, H. (1984) *Mol. Immunol. 21*, 223–232.
Fanning, D. W., Smith, J. A., & Rose, G. D. (1986) *Biopolymers 25*, 863–883.
Finzel, B. C., & Salemme, F. R. (1985) *Nature (London) 315*, 686–688.
Fraga, S. (1982) *Can. J. Chem. 60*, 2606–2610.
Haber, E., Margolies, M. N., & Cannon, L. E. (1976) *Cold Spring Harbor Symp. Quant. Biol. 41*, 647–659.
Hopp, T. P., & Woods, K. R. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 3824–3828.
Kabat, E. A. (1966) *J. Immunol. 97*, 1–11.
Kabat, E. A. (1978) *Adv. Protein Chem. 32*, 1–75.
Karplus, M., & McCammon, A. D. (1983) *Annu. Rev. Biochem. 52*, 263–300.
Kranz, D. M., Herron, J. N., & Voss, E. W. (1982) *J. Biol. Chem. 257*, 6987–6995.
Ladd, M. F. C., & Palmer, R. A. (1985) in *Structure Determination by X-ray Crystallography*, pp 165–166, Plenum, New York.
Leed, B. K., & Richards, F. M. (1971) *J. Mol. Biol. 55*, 379–400.
Lesk, A. M., & Hardman, K. D. (1982) *Science (Washington, D.C.) 216*, 539–540.
Levitt, M., Sander, C., & Stern, P. S. (1985) *J. Mol. Biol. 181*, 423–447.
Levy, R. M., Sheridan, R. P., Keepers, J. W., Dubey, G. S., Swaminathan, S., & Karplus, M. (1985) *Biophys. J. 48*, 509–518.
McCammon, J. A., Gelin, B. R., & Karplus, M. (1977) *Nature (London) 267*, 585–590.
Moudallal, Z. A., Briand, J. P., & Van Regenmortel, M. H. V. (1985) *EMBO J. 4*, 1231–1235.
Mudgett-Hunter, M., Anderson, W., Haber, E., & Margolies, M. N. (1985) *Mol. Immunol. 22*, 477–488.
Northrup, S. H., Pear, M. R., McCammon, J. A., Karplus, M. & Takano, T. (1980) *Nature (London) 287*, 659–660.
Novotný, J., Bruccoleri, R. E., & Karplus, M. (1984) *J. Mol. Biol. 177*, 787–818.
Novotný, J., Handschumacher, M., Haber, E., Bruccoleri, R. E., Carlson, W. E., Fanning, D. W., Smith, J. A., & Rose, G. D. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 226–230.
Padlan, E. A. (1985) *Mol. Immunol. 22*, 1243–1254.
Petsko, G. A., & Ringe, D. (1984) *Annu. Rev. Biophys. Bioeng. 13*, 331–372.
Post, C. B., Brooks, B. R., Karplus, M., Dobson, C. M., Artymiuk, P. J., Cheetham, J. C., & Phillips, D. C. (1986) *J. Mol. Biol. 190*, 455–479.
Ryckaert, J. P., Ciccotti, G., & Berendsen, H. J. (1977) *J. Comput. Phys. 23*, 327–341.
Sachs, D. H., Schechter, A. N., Eastlake, A., & Anfinsen, C. B. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 3790–3794.
Sela, M. (1969) *Science (Washington, D.C.) 166*, 1365–1374.
Sheriff, S., Hendrickson, W. A. Stenkamp, R. E., Sieker, L. C. & Jensen, L. H. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1104–1107.
Tainer, J. A., Getzoff, E. D., Alexander, H., Houghten, R. A., Olson, A. J., Lerner, R. E., & Hendrickson, W. A. (1984) *Nature (London) 312*, 127–134.
Tainer, J. A., Getzoff, E. D., Paterson, A., Olson, A. J., & lerner, R. A. (1985) *Annu. Rev. Immunol. 3*, 423–438.
Tanaka, T., Slamon, D. J., & Cline, M. J. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 3400–3404.
Thornton, J. M., Edwards, M. S., Taylor, W. R., & Barlow, D. J. (1986) *EMBO J. 5*, 409–413.
van Gunsteren, W. F., & Karplus, M. (1981) *Nature (London) 293*, 677–678.
Verlet, L. (1967) *Phys. Rev. 159*, 98–103.
Westhof, E., Altschuh, D., Moras, D., Bloomer, A. C., Mondragon, A., Klug, A., & Van Regenmortel, M. H. V. (1984) *Nature (London) 311*, 123–126.