

# Molecular Modeling and Dynamics of the Sodium Channel Inactivation Gate

Fernanda L. Sirota, Pedro G. Pascutti, and Celia Anteneodo

Instituto de Biofísica Carlos Chagas Filho UFRJ - Universidade Federal do Rio de Janeiro, Brazil

**ABSTRACT** The intracellular linker  $L_{III-IV}$  of voltage-gated sodium channels is known to be involved in their mechanism of inactivation. Its primary sequence is well conserved in sodium channels from different tissues and species. However, the role of charged residues in this region, first thought to play an important role in inactivation, has not been well identified, whereas the IFM triad (I1488-M1490) has been characterized as the crucial element for inactivation. In this work, we constructed theoretical models and performed molecular dynamics simulations, exploring the role of  $L_{III-IV}$ -charged residues in the presence of a polar/nonpolar planar interface represented by a dielectric discontinuity. From structural predictions, two  $\alpha$ -helical segments are proposed. Moreover, from dynamics simulations, a time-conserved motif is detected and shown to play a relevant role in guiding the inactivation particle toward its receptor site.

## INTRODUCTION

The  $\alpha$ -subunit of rat brain type IIA sodium channels has a 53 amino acid linker between domains III and IV ( $L_{III-IV}$ ) that has already been characterized as the inactivation gate (Catterall, 2000; West et al., 1992; Stühmer et al., 1989). This linker, rich in charged residues (12 positive and 5 negative), is highly conserved in different tissues and species. The positive charges were initially thought to play an important role in the inactivation phenomenon, similarly to the N-terminal positive charges in the potassium channel (Catterall, 2000). Consequently, several experiments to study the role of charged residues during inactivation have been performed (Patton et al., 1992; Moorman et al., 1990). In these studies, charges were shown to affect both inactivation and activation kinetics, although not in a crucial way. However, the critical element involved in inactivation was shown to be the hydrophobic IFM motif formed by residues 1488 through 1490 and termed inactivation particle (West et al., 1992). According to this model, the IFM would bind to its receptor site at the inner mouth of the ion channel pore, blocking it, and stabilizing the inactivated state. After the IFM triad identification, very few papers have addressed which role these highly conserved charges could be playing (Miller et al., 2000). Although none of the charges appear to be substantial for either inactivation or activation, they might play some other important function given their high

degree of conservation. The role of charges in the  $L_{III-IV}$  is not yet well understood mainly because of the lack of information on the detailed structure of the sodium channel. Although recently the three-dimensional structure of the sodium channel was investigated by cryo-electron microscopy (Sato et al., 2001), this very useful information is still not sufficient when looking at the atomic level. The apparent involvement of the  $L_{III-IV}$  during inactivation of the sodium channel led to the investigation of the  $L_{III-IV}$  peptide in solution using nuclear magnetic resonance (NMR) (Rohl et al., 1999). Certainly, a more complete and detailed structural model would better direct the studies on the role of the mainly positive charges present in the  $L_{III-IV}$ .

To this aim, we built theoretical models based on secondary structure prediction from the primary sequence. These models were used for molecular dynamics simulations where the polar/nonpolar vicinity of the  $L_{III-IV}$  was treated as a dielectric inhomogeneous medium. This environment is provided by the proximity of the  $L_{III-IV}$  to the biological membrane or to any local area where a dielectric discontinuity can be present, such as the hydrophobic inactivation particle receptor in the channel. Our model incorporates the fact that the loop inhabits an aqueous environment in the vicinity of a nonpolar region without considering electrostatic fields other than the intrinsic one due to the partial atomic charges and the polarization charges at the polar/nonpolar interface. Also the shielding effect of ions in solution is not taken into account. Explicit solvent representation would not be advantageous at the same time that many details of the sodium channel transmembrane portion and local electrostatic potential are not still well established, furthermore facing computational and time limitations. Although this representation of the  $L_{III-IV}$  environment is simplified, it constitutes a first approach to study the dynamics of the loop with special emphasis on the fact that the membranous and aqueous environments have different dielectric properties.

Submitted July 3, 2001, and accepted for publication October 12, 2001.

Address reprint requests to Celia Anteneodo, UFRJ - CCS - Instituto de Biofísica Carlos Chagas Filho, Bloco G - s. G<sub>0</sub>-028 - Ilha do Fundão - Rio de Janeiro - RJ, 21949-900 Brazil. Tel.: 55-21-2562-6575; Fax: 55-21-2280-8193; E-mail: [celia@cbpf.br](mailto:celia@cbpf.br).

Fernanda L. Sirota's present address is Service de Conformation de Macromolécules Biologiques et de Bioinformatique - Université Libre de Bruxelles, Bruxelles, Belgium. <http://www.bmm.icnet.uk/~3dpssm/>

Celia Anteneodo's present address is Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil. <http://www.cmpharm.ucsf.edu/~nomi/npredict.html>

© 2002 by the Biophysical Society

0006-3495/02/03/1207/09 \$2.00

## MATERIALS AND METHODS

### Structure prediction

For the construction of different structural models for the 53 amino acid sequence L<sub>III-IV</sub>, we made use of Prelude (Rooman et al., 1991), a software for protein structure prediction based on a protein library (Wintjens et al., 1996). Outputs are given in terms of seven backbone structure assignments: *A*, *B*, *C*, *G*, *E*, *O*, and *P*, in which each letter defines a ( $\phi$ ,  $\psi$ ,  $\omega$ ) domain for each residue (Rooman et al., 1991). Assignments *A* and *C* represent  $\alpha$  and  $3_{10}$ -helix conformations, respectively, *B* and *P* correspond to extended structures with *B* representing  $\beta$  strand conformations, domains *E* and *G* have positive  $\phi$  and occur essentially in glycine residues, and assignment *O* corresponds to *cis* peptide conformations occurring almost exclusively in proline residues.

Prelude allows one to consider distance constraints between residues. To test the prediction robustness, different intervals of distance between the C $_{\alpha}$  atoms of the N- and C-terminal residues were set as constraint allowing variations only within the fixed interval. As upper bound we considered the interval 90 to 95 Å. Other intervals were studied, but the range of 10 to 15 Å was used as constraint in most molecular dynamics (MD) simulation cases. This vinculum is a valid one, because the L<sub>III-IV</sub> is part of the sodium channel located in the cytosol before and after the putative transmembrane  $\alpha$ -helical segments IIS6 and IVS1, respectively. Within the arrangement of transmembrane helices proposed by Noda et al. (1986), segments IIS6 and IVS1 are almost adjacent, which is compatible with a distance constraint within the range of 10 to 15 Å. For the purpose of comparison, secondary structure predictions were also performed using the software 3D-PSSM\* (Kelley et al., 2000) and Nnpredict\*\* (Kneller et al., 1990; McClelland and Rumelhart, 1988), although these on-line programs do not allow the consideration of distance constraints.

The three first lower energy Prelude output configurations, under the constraint 10 to 15 Å, were used for the backbone construction of theoretical models named TM1, TM2, and TM3, respectively. Side chains were adjusted with the aid of the software SCWRL (Bower et al., 1997) for side chain adjustments based on a rotamer library (Dunbrack and Cohen, 1997; Dunbrack and Karplus, 1993, 1994). Minor improvements were performed under optimization routines using the GROMOS 96 force field parameters (Scott et al., 1999), by means of the steepest descent and conjugated gradient algorithms (Press et al., 1992), implemented in the software THOR developed at our laboratory (Pascutti et al. 1999b; Moret et al., 1998; Arêas et al. 1995). To check the stereochemical quality of the theoretical models here proposed, we used the software Procheck (Rullmann, 1996; Laskowski et al., 1993).

A sequence of 21 residues, from Q1486 to S1506 in the 53 amino acid sequence that constitutes the inactivation gate of rat brain type IIA, had its structure elucidated by NMR in solution (Rohl et al., 1999), accessible under the Protein Data Bank (Berman et al., 2000) ID code 1byy. Therefore, we further compared the available NMR data with the corresponding residues in the theoretical models by means of root mean square deviation analysis. A hypothetical average structure based on the coordinates of all residues constituting the 10 available NMR structures was calculated. The deviation of each NMR and each theoretical model with respect to the NMR average structure was computed for two cases: one considering the 21 C $_{\alpha}$  atoms coordinates only and the other all the backbone atoms.

### Molecular dynamics simulations

Once we constructed the theoretical models, we performed MD simulations, by means of software THOR, starting from each one of the three different conformations (models TM1, TM2, and TM3). The GROMOS 96 force field was also considered for the MD simulations as it was for the optimization routines. All simulations were performed treating the environment either as a homogeneous medium characterized by a dielectric constant  $\epsilon$  or as an inhomogeneous one with a dielectric discontinuity ( $\epsilon =$

80 in the semispace  $x < 0$  and  $\epsilon = 2$  otherwise, hence, the interface is given by the  $y$ - $z$  plane). The value  $\epsilon = 80$  for the dielectric constant was used to represent a hydrophilic environment, whereas  $\epsilon = 2$  was used for a hydrophobic one. To take into account the presence of two media with different dielectric constants separated by a planar layer, which modifies the Coulomb's term of the interaction potential energy between charges (Pascutti et al., 1999a), the electrostatic images method was applied. The method basically consists in the substitution of the real configuration involving the point charges and the induced polarization charges at the surface by a simpler configuration formed by the point charges and their images (Griffiths, 1989). This representation has already been applied to the study of different biological systems (Cassiano and Arêas, 2001; Pascutti et al., 1999a,b; Arêas et al., 1995) providing results consistent with experimental evidence. Facing the simplicity of this description where the environment is not treated explicitly, additional vincula are required (Arêas et al., 1995). Here, hydrogen bonds in  $\alpha$ -helix structures were set as harmonic potentials. Therefore, the  $\alpha$ -helix secondary structures present in the L<sub>III-IV</sub> were stable during the simulations. In addition, backbone atoms of residues D1474 and D1526 had their atomic coordinates kept fixed along the MD simulations to mimic the fact that the L<sub>III-IV</sub> is attached to putative transmembrane helices.

The temperature of the system, proportional to the total kinetic energy, is controlled by appropriately scaling the velocity of each atom periodically. First, the temperature is raised at 5 K/ps up to 300 K in  $\epsilon = 80$ . After equilibration at 300 K, the interface of discontinuity is introduced, changing consequently the force field as described above. Different initial conditions were used. These include the distance and the orientation adopted by the different L<sub>III-IV</sub> structural models (TM1, TM2, and TM3) in  $\epsilon = 80$  facing the dielectric discontinuity interface in the  $y$ - $z$  plane at  $x = 0$ .

We define a spreading index  $S$  through the expression

$$S = \frac{1}{N} \sum_i |\vec{r}_i - \vec{r}_{cm}|^2,$$

in which  $\vec{r}_i$  are the atomic position vectors,  $\vec{r}_{cm}$  the position vector for the peptide center of mass, and  $N$  the number of atoms considered. Index  $S$  is related to the polar momentum of inertia and provides a measure of the degree in which a given structure is unfolded: the lower  $S$ , the higher the folding degree.  $S$  was used to investigate the global backbone deformation during MD simulations. For a local analysis of the deformation, both curvature and torsion angles were computed, considering only the backbone defined by C $_{\alpha}$ . The curvature angle, calculated over 3 consecutive C $_{\alpha}$  atoms, shows how much a curve deviates from the straight line. The torsion angle, like the dihedral in classical force fields, accounts for the angle between two planes formed by four consecutive C $_{\alpha}$ . The temporal deviation of curvature and torsion angles can be interpreted as a measure of flexibility.

Many results from MD simulations are illustrated through the outcomes from model TM2, as they are representative of the results obtained by starting from the other two theoretical models as well.

## RESULTS

### Structure prediction

The secondary structure prediction, proposed by Prelude, for the case where the distance between the C $_{\alpha}$  atoms of the N and C termini was constrained to the range of 10 to 15 Å, is plotted in Fig. 1 *a*. This figure shows a specific-position frequency distribution of the assignments occurrence in the first 500 backbone structures. The higher relative energy difference among the first 500 predicted structures is only

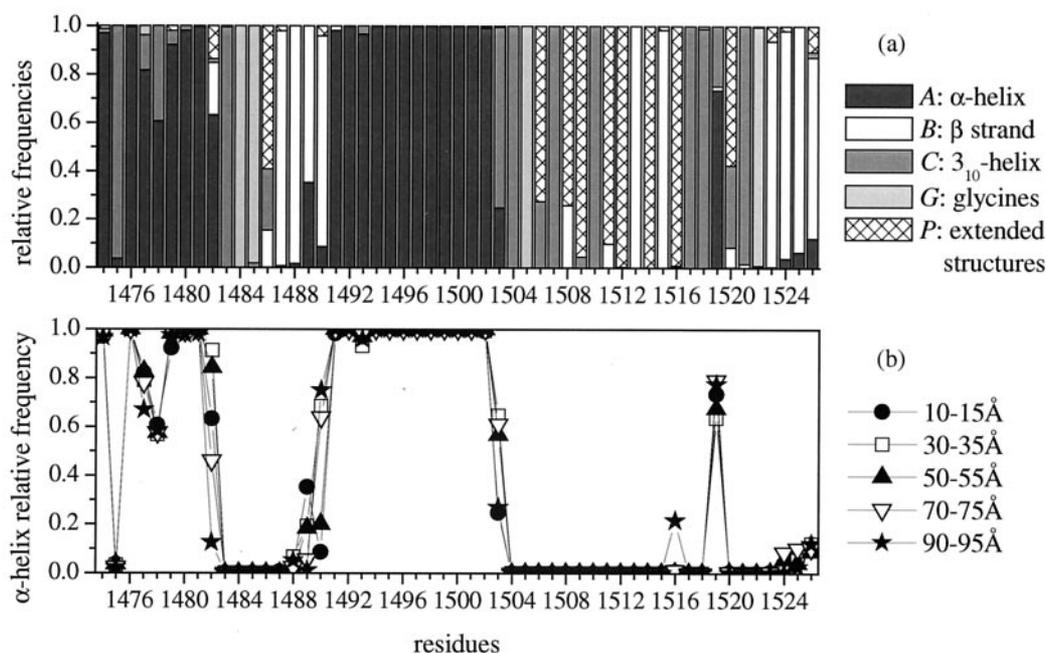


FIGURE 1 Specific-position relative frequency distribution for the Prelude (Rooman et al., 1991) assignments occurrence. (a) Relative frequency of occurrence of Prelude assignments A, B, C, G, E, O, and P (see methodology) in the first 500 backbone structures proposed as represented by a stacked columns graph. The distance between the  $C_{\alpha}$  atoms of the N- and C-terminal residues was constrained to the range 10 to 15 Å. There was only one hit for E at position 1522 and no hit for O, which corresponds to *cis* peptide conformations occurring almost exclusively in proline residues. (b) Relative frequency of occurrence of Prelude assignment A ( $\alpha$ -helix) in the first 500 backbone structures proposed using the distance constraints indicated in the figure.

2.6%, therefore frequencies were computed by direct counting without the need of weights. This statistical analysis predicts the very stable  $\alpha$ -helix between residues T1491 and K1502 also observed by NMR in solution. In addition, a shorter  $\alpha$ -helix near the N terminus involving residues F1476 through K1482 is further proposed. The statistics shows a slightly lower frequency for the  $\alpha$ -helix near the N terminus than for the  $\alpha$ -helix elucidated by NMR (Rohl et al., 1999), indicating the possibility that the small  $\alpha$ -helix formation is more dynamic and therefore more difficult to

be experimentally measured. The protein database used by Prelude is a very general one covering proteins of different sizes and with high-resolution structures (better than 2 Å). Although these structures were obtained in solution, some proteins may contribute with portions that provide a hydrophobic environment, similar to the one in the vicinity of the L<sub>III-IV</sub>. Like in synthetic antimicrobial peptides (Oh et al., 2000), this short  $\alpha$ -helix might require a hydrophobic neighborhood for its stability, being unstable in aqueous environment. In solution, this very small  $\alpha$ -helix may be either

TABLE 1 Predicted sequences

|           | 1474 |            |                     |            |            |            | 1526 |                |
|-----------|------|------------|---------------------|------------|------------|------------|------|----------------|
| Sequence: | D    | NFNQKKKFG  | <b>GQDIFM</b> TTEEQ | KKYYNAMKKL | GSKKPQKPIP | RPANKFQGMV | FD   | $d_{1-53}$ (Å) |
| TM1       | A    | CAAAAAACG  | GPBBBBAAAA          | AAAAAAACC  | GPCPPCPPBP | BPCCAPCGBB | BB   | 10.4           |
| TM2       | A    | CAAAAAACG  | GPBBBBAAAA          | AAAAAAACC  | GPCBPCPPBP | BPCCAPCGBB | BB   | 13.0           |
| TM3       | A    | CAAAAAACG  | GCBABAAAA           | AAAAAAACC  | GPCPPCPPBP | BPCCAPCGBB | BB   | 10.5           |
| SP1       | -    | -AAAAAAA-  | --BBBB-AAA          | AAAAAAA-   | -----      | ---AAABBB- | --   |                |
| SP2       | -    | ---AA----- | -----AAA-           | AAAAAAA-   | -----      | -----A     | ---  |                |
| SP3       |      |            | -----AAA            | AAAAAAA-   | ---        |            |      |                |

TM1, TM2, and TM3 are the first three lower energy sequences provided by the Prelude output file. Sequence SP1 was obtained using 3D-PSSM (Kelley et al., 2000), SP2 using Nnpredict (Kneller et al., 1990; McClelland and Rumelhart, 1988), and SP3 is the fold recognition given by 3D-PSSM to the PDB ID 1byy (Berman et al., 2000; Rohl et al., 1999).

Letter code for secondary structure assignments is the same as in Fig. 1, expect for "--", which represents nonpredicted segments. The IFM triad is shown in bold.

The distance between  $C_{\alpha}$  of the terminal residues 1 and 53 is shown in the right hand column.

**TABLE 2** Ramachandran plot statistics

|  | TM1 |         | TM2 |         | TM3 |         |
|--|-----|---------|-----|---------|-----|---------|
| Residues in most favorable regions           | 35  | (81.4%) | 39  | (90.7%) | 38  | (88.4%) |
| Residues in additional allowed regions       | 8   | (18.6%) | 4   | (9.3%)  | 5   | (11.6%) |
| Residues in generously allowed regions       | 0   | (0%)    | 0   | (0%)    | 0   | (0%)    |
| Residues in disallowed regions               | 0   | (0%)    | 0   | (0%)    | 0   | (0%)    |
| Number of nonglycine and nonproline residues | 43  | (100%)  | 43  | (100%)  | 43  | (100%)  |

The statistics were provided by Procheck (Laskowski et al., 1993; Rullmann, 1996) for the three theoretical models TM1, TM2, and TM3. There is a total of 4 glycine and 4 proline residues per model in a total of 53 amino acid residues.

unstable or fold and unfold so quickly in time that the NMR technique cannot detect. Hence, new NMR experiments either in the presence of micelles or in an organic solvent could clarify the existence of this small  $\alpha$ -helix, which has been theoretically predicted.

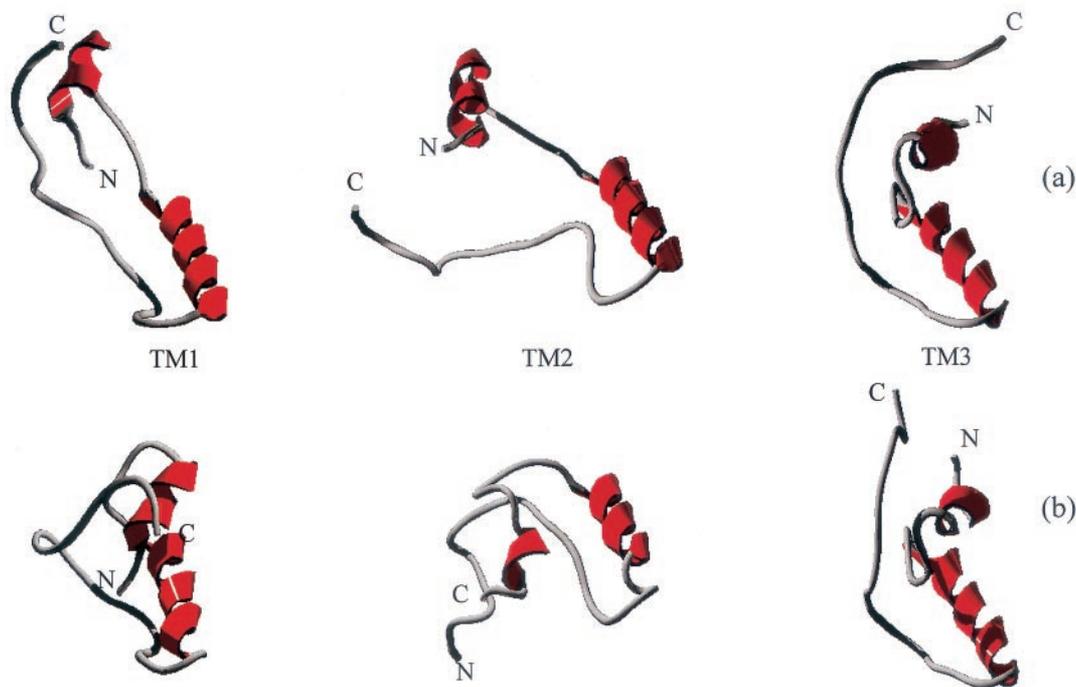
Because the precise distance between N and C termini is unknown, it is necessary to cover a broad range of options for the distance hindrance. Therefore, we repeated the statistics of Fig. 1 *a*, obtained with the constraint 10 to 15 Å, using different restriction ranges. We observed that, up to a constraint range of 90 to 95 Å, results are almost insensitive to this hindrance, yielding histograms essentially similar to those exhibited in Fig. 1 *a*. Constraints as large as 110 to 115 Å do not allow computation, probably due to geometric incompatibilities, as the peptide is only 53 residue long. In particular, we exhibit in Fig. 1 *b* the  $\alpha$ -helix relative frequency analysis for different constraint intervals. This figure illustrates the robustness of the secondary structure

prediction under conditions where the distance between termini varies from 10 to 95 Å.

The predictions performed using 3D-PSSM and Nnpredict were consistent with the Prelude prediction as shown in Table 1, evidencing also the possibility of a shorter  $\alpha$ -helix formation near the N terminus. Nevertheless, for models construction, we used the Prelude predictions, because they satisfy constraints on the distance between the N and C termini.

The stereochemical quality of the theoretical models is presented through the Ramachandran plot statistics for all three models in Table 2. After adjusting side chains and optimizing as described in the methodology section, we obtained the models shown in Fig. 2 *a*.

The comparison between the experimental NMR (Rohl et al., 1999) and our theoretical models through root mean square deviation analysis is presented in Table 3. Inspection of this table shows that the deviation of the theoretical



**FIGURE 2** Theoretical models TM1, TM2, and TM3. (*a*) Optimized Prelude outputs and (*b*) their corresponding conformations after MD heating up to and equilibration at 300 K in  $\epsilon = 80$  before the insertion of the dielectric discontinuity interface.

**TABLE 3 RMS deviation analysis**

| Model | RMS deviation<br>(Å) C $\alpha$ | RMS deviation<br>(Å) backbone |
|-------|---------------------------------|-------------------------------|
| NMR   |                                 |                               |
| 1     | 2.82                            | 2.93                          |
| 2     | 1.86                            | 1.82                          |
| 3     | 1.88                            | 1.87                          |
| 4     | 1.73                            | 1.72                          |
| 5     | 1.87                            | 1.90                          |
| 6     | 1.70                            | 1.71                          |
| 7     | 3.64                            | 3.73                          |
| 8     | 2.79                            | 2.93                          |
| 9     | 2.44                            | 2.44                          |
| 10    | 2.54                            | 2.70                          |
| TM1   | 3.15                            | 2.95                          |
| TM2   | 2.88                            | 2.68                          |
| TM3   | 2.87                            | 2.67                          |

RMS deviation of NMR (Rohl et al., 1999) and theoretical models with respect to the average model calculated over the 10 NMR structures coded 1byy (Rohl et al., 1999). All calculations were performed over the 21 residues from Q1486 to S1506 in the 53-amino-acid linker, because this is the region confined in the 1byy. Computations over C $\alpha$  atoms only and over all backbone atoms are shown in the second and last columns, respectively.

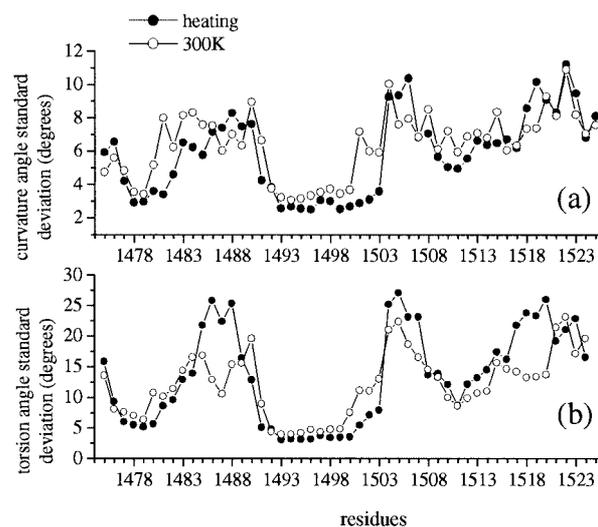
structures is within the NMR deviation range for backbone atoms.

### Molecular dynamics simulations

After the optimized theoretical structures were heated up to and equilibrated at 300 K in  $\epsilon = 80$ , all of them packed to structures of spreading index  $S \approx 0.12$ , having started within the range 0.19 to 0.23. However, all structures were very flexible along heating and equilibration, showing no specific folding (Fig. 2 b), except for the region involving residues T1491 through P1512, which is very well conserved in time, adopting the shape similar to a hairpin in all theoretical models studied (Fig. 3). In Fig. 4 we show the deviation of the curvature and torsion angles computed along time and averaged over the three models for both



**FIGURE 3** The time conserved hairpin motif (T1491-P1512). Superposition of the motif in the three models equilibrated at 300 K in  $\epsilon = 80$  with its pattern of charged residues exhibited.



**FIGURE 4** Standard deviation of curvature (a) and torsion (b) angles computed along both heating and equilibration at 300 K in  $\epsilon = 80$ . Symbols correspond to the average over the three models. Lines are guides to the eyes.

heating and equilibration at 300 K in  $\epsilon = 80$ . Although the residues preceding the long  $\alpha$ -helix showed a coil-coiled conformation, they were all in a flexible structure kept parallel to the  $\alpha$ -helix (Fig. 3). On the other hand, local flexibility did not occur within the  $\alpha$ -helix, because of the H-bond constraints described in methodology. The local flexibility of the hairpin structure observed in the simulations could be the reason why the NMR (Rohl et al., 1999) could not measure all parameters needed for the hairpin motif determination, albeit sufficient to determine the stable  $\alpha$ -helix portion. The presence of the hairpin motif was observed also for the case where a distance of 30 Å was used as constraint along simulations. To mimic a more realistic articulation of the termini, in one simulation, 10 residues were added to each terminus, reproducing part of the putative IIIS6 and IVSI helices, respectively. They were kept fixed along MD simulations and their atomic charges were taken as zero. Even in this case, the motif was stable as well (data not shown). It is noteworthy that this motif presents a particular pattern of charges as exhibited in Fig. 3.

Residues G1484, G1485, P1512, P1514, and P1516 were described as involved in the hinge mechanism necessary for normal rapid inactivation (Kellenberger et al., 1997). Interestingly, none of those residues belong to the region of the motif here proposed (T1491-P1512), except for P1512, which is located precisely at its end. In Fig. 4 it can be observed that, besides all glycine residues, a few residues preceding F1489 also present high flexibility, which may aid fitting the F1489 into its receptor site.

The dielectric discontinuity interface is introduced at  $x = 0$ , i.e., at the  $y$ - $z$  plane, with  $\epsilon = 80$  for  $x < 0$  and  $\epsilon =$

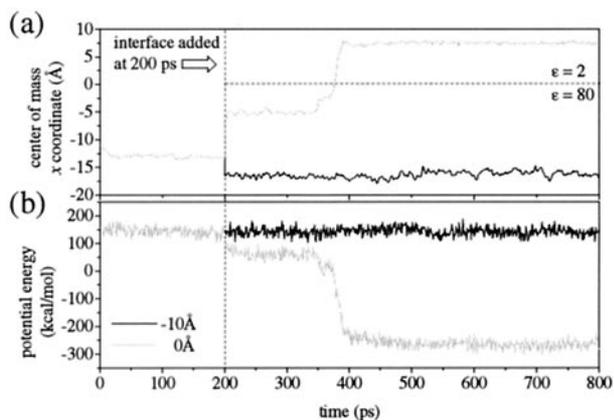


FIGURE 5 Effect of the position of the loop ends with respect to the dielectric interface. (a) Trajectory of the center of mass  $x$  coordinate along MD simulations of TM2. Up to 200 ps the medium was treated as a dielectric continuum with  $\epsilon = 80$ . At time 200 ps, the planar interface of dielectric discontinuity ( $y$ - $z$  plane) was added at  $x = 0$ , such that  $\epsilon = 80$  for  $x < 0$  and  $\epsilon = 2$  otherwise. Before the insertion of the interface, the equilibrated structure was positioned in the region  $x < 0$ , such that its termini were located at  $x = \delta$ , in which  $\delta$  was taken as  $-10$  or  $0 \text{ \AA}$ . The jump in the curve at 200 ps is due to the translation of the loop to set  $\delta$ . (b) Potential energy as a function of time for the different values of  $\delta$ . When the ends are at  $-10 \text{ \AA}$ , the loop does not fold to the lower dielectric medium, whereas for the ends at  $0 \text{ \AA}$ , the linker quickly visits the  $\epsilon = 2$  environment.

2 otherwise, whereas the linker is positioned at the semispace  $x < 0$  (hence  $\epsilon = 80$ ). Because the precise value for the  $x$  coordinate, let us call it  $\delta$ , of the loop ends (taken the same for both  $C_\alpha$ ) with respect to the interface is not known, we considered different values as initial condition when adding the interface. The choice of  $\delta$  is crucial to the dynamics of the loop: for  $\delta$  below a threshold, the loop core does not approach the interface, whereas for  $\delta$  above the threshold, it moves toward the interface, as can be seen in Fig. 5 by following the evolution of the center of mass  $x$  coordinate for  $\delta = 0 \text{ \AA}$  and  $\delta = -10 \text{ \AA}$ . This effect is mainly due to the presence of negative charged residues at the termini and can be understood through the potential energy profile that results when the structure is rigidly translated along the  $x$  axis (Fig. 6). The potential energy in Fig. 6 is calculated over a static and rigid structure, (i.e., all internal coordinates are kept fixed) at a given position, along the  $x$  axis, in front of the interface. When the structure has its termini in  $\epsilon = 80$  far from the interface, the force on the center of mass is in the opposite direction to the interface, whereas for the ends sufficiently close to it, the center of mass feels a force toward the low dielectric moiety. For intermediate values, the potential barrier can be overcome through a conformational change of the loop. Furthermore, as the loop enters a region of low dielectric constant, the formation of salt bridges, involving mainly the residue pairs E1492 through K1495 and E1493 through K1496, yields

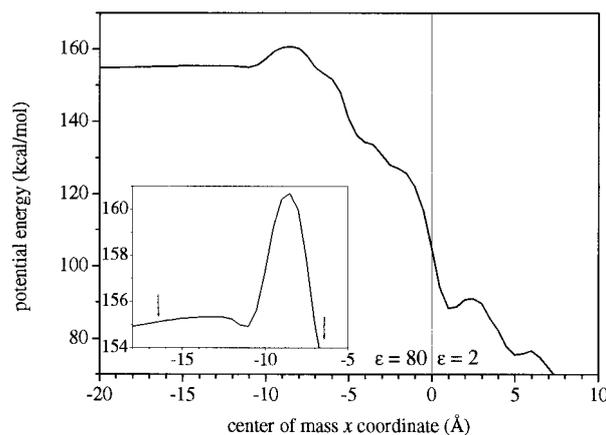


FIGURE 6 Potential energy profile of a rigid linker (all internal coordinates are kept fixed) whose center of mass is moved along the  $x$  axis in the presence of a dielectric discontinuity at  $x = 0$ . The linker conformation corresponds to time 200 ps of Fig. 5, that is, before the interface insertion. Arrows indicate the position of the center of mass  $x$  coordinate corresponding to each value of  $\delta$  considered in Fig. 5. Notice that the force on the center of mass along the  $x$  direction is given by  $F_x = -\partial U/\partial x$ , in which  $U$  is the potential energy. The inset is a magnification of the main plot in the vertical axis.

the observed reduction of the potential energy, therefore stabilizing the loop in such environment.

Our result suggests that the conformational change that the channel suffers when passes from one state to another (particularly from open to inactivated) may yield immersion, at the angstrom scale, of the loop termini through the membrane interface with consequent movement of the loop toward a polar/nonpolar environment.

We also investigated the consequences, on the dynamics, of changing the orientation of the loop with respect to the plane of the interface for fixed  $\delta$ . From several simulations, we observed that the angle  $\theta$  that the hairpin structure forms with the interface plane, as defined in Fig. 7 *b*, is the relevant variable concerning the orientation of the linker as a whole. In Fig. 7 *a*, we show the electrostatic potential energy profile that results when the hairpin, by separate, is rigidly moved along the  $x$  axis (fixed internal coordinates), considering different orientations with respect to the interface. The most favorable orientations, when close to the interface, are those corresponding to  $\theta$  in the range of  $90^\circ$  to  $180^\circ$ , whereas for orientations of the hairpin with its turn pointing to the interface, as for  $\theta = -90^\circ$ , a potential barrier forbids the approximation of the motif. In fact, in dynamics simulations, when the initial structure has its hairpin oriented as in Fig. 8 *a* ( $\theta \approx -30^\circ$ ) with  $\delta = 0$ , the structure does not approach the interface immediately, whereas with an initial orientation as in Fig. 8 *b* ( $\theta \approx 0^\circ$ ) it does. Plots of the potential energy and the center of mass  $x$  coordinate as a function of time, for dynamics simulations starting from configurations as in Fig. 8, *a* and *b*, are shown in

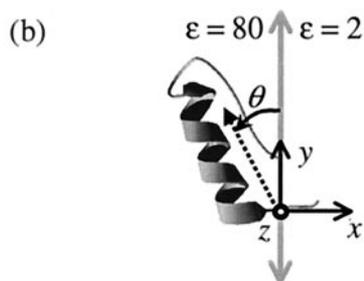
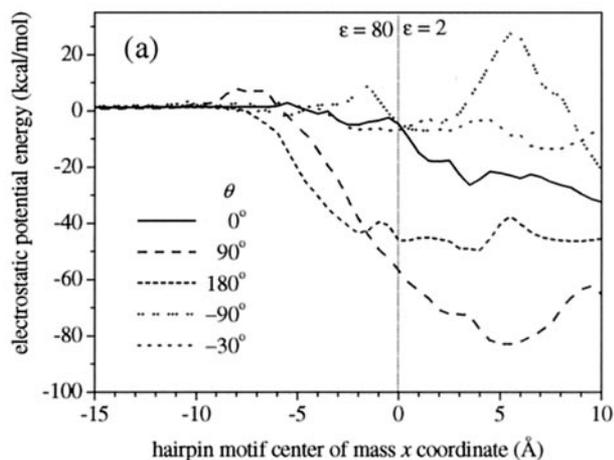


FIGURE 7 Electrostatic potential energy profile for a rigid hairpin whose center of mass is moved along the  $x$  axis in the presence of a dielectric discontinuity at  $x = 0$  ( $y$ - $z$  plane, normal to the sheet of paper) for different orientations (a) as defined in b. For each given position of the hairpin center of mass, the motif presents a preferential orientation. The change of this lowest energy orientation with the position of the center of mass can also be noticed from the dynamics simulation results shown in Figs. 8 and 9.

Fig. 9. As illustrated in Fig. 8, b through d, following the case  $\theta \approx 0^\circ$ , for initial orientations of the hairpin with  $\theta$  approximately in the range  $0^\circ$  to  $180^\circ$ , the structure soon comes closer to the low dielectric medium changing orientation, toward the most favorable one, in such a way

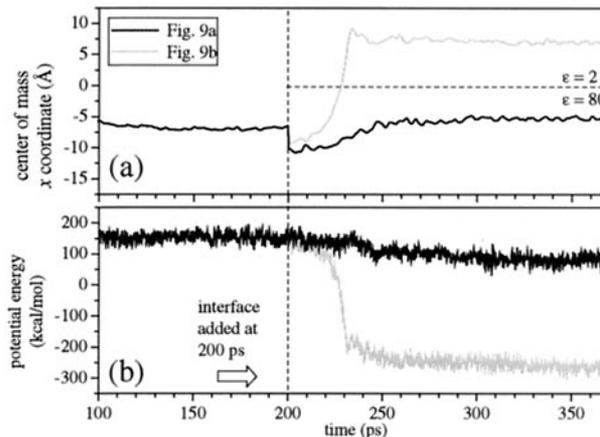


FIGURE 9 The orientation of the loop with respect to the dielectric interface. (a) Trajectory of the center of mass  $x$  coordinate along MD simulations. Up to 200 ps the medium was treated as a dielectric continuum with  $\epsilon = 80$ . At time 200 ps, the planar interface of dielectric discontinuity was added as in Fig. 5. Before the insertion of the interface, the equilibrated structure was oriented as in Fig. 8, a and b. (b) Potential energy as a function of time for the two initial orientations. For the orientation in Fig. 8 a, the loop does not fold to the lower dielectric medium, whereas for the orientation as in Fig. 8 b the loop tends to visit the low dielectric environment. Notice that, following Fig. 7 a,  $\theta \approx 0^\circ$  is not the most favorable initial orientation, and that is why the approximation does not occur immediately.

that the hairpin drives the IFM motif toward the low dielectric environment. This movement can be interpreted as the approximation of the IFM to its hydrophobic receptor site in the channel.

It has been already suggested in the literature (Catterall, 2000) that the rigid  $\alpha$ -helix serves as a scaffold to present the IFM motif and T1491 to a receptor in the mouth of the pore as the gate closes. Here, we propose that the T1491 to P1512 hairpin motif serves to orient the F1489 to its site. The presence of image charges due to the dielectric interface, without considering additional electrostatic potentials, may be, at least partially, responsible for furnishing the long range interaction that guides the IFM toward its hydrophobic receptor site.

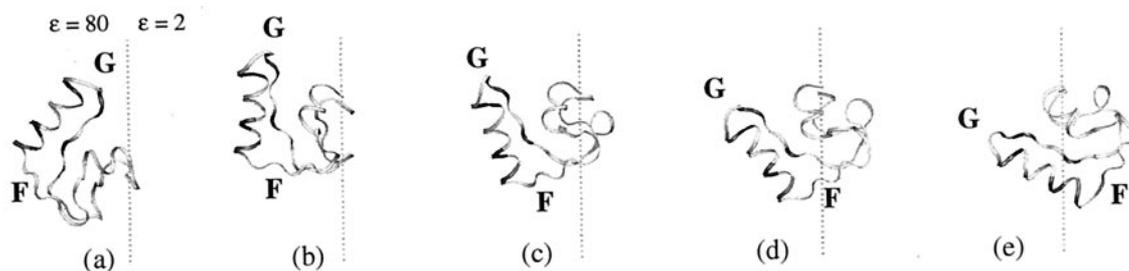


FIGURE 8 Effect of the orientation of the linker facing the dielectric discontinuity interface. TM2 in a and b differs in the orientation facing the plane of the interface (normal to the sheet) with  $\delta = 0$ . Orientations a and b were used as starting conditions for MD simulations in the presence of the interface (added at 200 ps). Starting from a, the structure remains stable, whereas starting from b, the loop folds toward  $\epsilon = 2$ , as can be observed in successive instants of the dynamics schematized in c through e, corresponding to times 218, 225, and 229 ps, respectively. Residues F1489 and G1505 are indicated. Line separates  $\epsilon = 80$  (left side) from  $\epsilon = 2$  (right side).

## DISCUSSION AND CONCLUSIONS

We performed a theoretical prediction of the topology of the intracellular loop L<sub>III-IV</sub> of rat brain type II sodium channels. The results shown in Fig. 1 and Table 1 indicate the presence of two helical segments (F1476–K1482 and T1491–K1502). The latter helix had already been observed by NMR measurements of the linker in aqueous solution (Rohl et al., 1999), whereas the former was not detected by the same technique. The software Prelude (Rooman et al., 1991) used for predictions yielded results consistent with those of other secondary structures predictors and partially consistent with the NMR measurements of the L<sub>III-IV</sub> in solution. This prediction tool was very useful when studying the region that the NMR could not determine as a stable structure. The statistics presented in Fig. 1 indicates that the small  $\alpha$ -helix near the N terminus is a bit less frequent than the one elucidated by NMR. Although the Prelude databank is composed of proteins whose structures were obtained in solution, they may include hydrophobic portions, which can be representative of the L<sub>III-IV</sub> neighborhood. That is why as a result of Prelude analysis, what comes out is the structure of L<sub>III-IV</sub> contemplating different environments, instead of resulting exactly the same structure experimentally observed when the loop is in solution. Here we propose that the formation of the helix close to the N terminus is influenced by the proximity of a hydrophobic environment, such as the biological membrane, having low stability in solution, reason why it was not detected in the above-mentioned NMR experiments. This proposal could be elucidated by carrying out NMR experiments of the linker, either in an organic solvent or in the presence of micelles.

MD simulations were developed by treating the environment as a continuous and linear dielectric, either homogeneous or with a jump at a planar interface. In this continuum model, the heterogeneous properties of the solution and mainly of the protein are neglected. However, bearing in mind the complexity of the medium inhabited by the L<sub>III-IV</sub> and its surroundings (e.g., transmembrane protein segments, intracellular loops, phospholipid bilayer, ionic diffuse double layers, etc.), our effort constitutes a first approach toward understanding the dynamics of the linker.

Our results from MD indicate the formation of a hairpin motif (T1491–1512). This motif can be responsible for providing the long-range interactions, which optimize the approximation of the hydrophobic triad IFM toward its docking site, probably, of hydrophobic nature. This function would explain, at least partially, the presence of evolutionary well-conserved charges in this region of the channel. Our results show that the orientation of the hairpin motif with respect to the polar/nonpolar interface is important for determining the approximation of the IFM to the plane of the interface. At the same time, the precise localization of the loop termini with respect to that plane is also shown to influence the folding of the loop toward it. It is worth

recalling that electrostatic potentials may be overestimated because the shielding due to the density distribution of small ions is not being considered in the present model.

The G1505 is the only glycine residue that was not classified in the list of residues participating in the hinge mechanism as proposed by Kellenberger et al. (1997). Coincidentally, G1505 is found to be the element responsible for the hairpin motif turn following the  $\alpha$ -helix. By substituting G1505 by alanine, valine, or proline (data not shown), a similar hairpin motif is formed after heating and equilibration MD, although starting conformations are extended. This indicates once more the tendency that the segment T1491 through P1512 has to form the hairpin motif with its characteristic charge pattern. Ionic pairs between atoms of the hairpin side chains and the  $\alpha$ -helix backbone allow the hairpin formation, despite the excess of positive charges. On the other hand, it is probable that the motif also interacts with the voltage-sensing device similarly as discussed in the work of Patton et al. (1992). A change in the electric field across the membrane may produce conformational changes in regions of the channel, such as the voltage sensor. These changes, in turn, alter the electric field felt by the intracellular loop. A different electrostatic field, as such present during activation, could induce unfolding of the hairpin motif, a movement that can be required at that stage of channel operation. In any case, flexibility at position 1505 is important for proper inactivation. In fact, “cut” and “cut/addition” mutants at position 1505 (Stühmer et al., 1989) yield dramatic decrease in the rate of inactivation. Also mutants G1505A and G1505V had their fast inactivation unaltered and slightly diminished, respectively, as reported in the literature (Kellenberger et al., 1997), whereas G1505P presented drastically diminished inactivation, besides macroscopic activation being affected in all these mutants (Kellenberger et al., 1997). Because proline confers rigidity to structures, as its side chain is attached to the protein backbone, this might be the reason why the major effects are observed for mutant G1505P.

Concerning the pattern of charges in the hairpin motif, let us discuss some results on deletions and point mutations involving charged residues. All deletions in the work by Patton et al. (1992) comprise a region of 10 consecutive residues each, where deletion 1 starts at N1475, deletion 2 at G1485 and so on, until deletion 5, whose last residue is R1515. Deletions 2 through 4 involve the hairpin motif. Deletions 2 and 3 were reported to fail in expressing functional channels (Patton et al., 1992). In deletion 4, activation was similar to the wild type, but inactivation was very slow. Deletion 4, which did not affect activation, could have had the hairpin motif positive charges replaced by the charges present in the region corresponding to deletion 5. The effect of diminished inactivation could have occurred due to the presence of P1516 almost replacing G1505, similarly as Kellenberger et al. (1997) observed in mutant G1505P, where inactivation was drastically diminished. Moreover, mutant B<sub>Q</sub> (Patton et al., 1992), which precisely neutralizes

all the charges in the hairpin motif, presented a reminiscent current indicating that not all channels inactivated properly, which could be due to the inadequate IFM guidance.

Additionally, a phosphorylation site at S1506 (West et al., 1991), important for regulating the electric activity in excitable cells, is located close to the motif turn. In the case of the delayed rectifier  $K^+$  channel, phosphorylation affects voltage gating by electrostatic interactions (Perozo and Bezanilla, 1990). Notice that phosphorylation of serine implies the presence of a negative charge interrupting the pattern of positive charges.

The hairpin motif, besides optimizing the approximation of the IFM to its docking site in the channel may interact with the voltage sensor as suggested in the literature (Patton et al., 1992). This point could be elucidated by simulating the electrostatic field produced by the voltage sensor at different stages of channel operation. Still, our simplified model proved to be suitable for discussing molecular aspects of the  $L_{III-IV}$  dynamics, whereas the  $Na^+$  channel three-dimensional structure is yet to be clarified at the angstrom scale.

We thank Prof. Walter Stühmer for revising and discussing this manuscript and Prof. Jean-Marie Ruysschaert and Prof. Paulo M. Bisch for the fruitful discussions. This work was financially supported by FAPERJ and CNPq.

## REFERENCES

- Arêas, E. P. G., P. G. Pascutti, S. Schreier, K. C. Mundim, and P. M. Bisch. 1995. Molecular dynamics simulations of signal sequences at a membrane/water interface. *J. Phys. Chem.* 99:14885–14892.
- Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. 2000. The protein data bank. *Nucleic Acids Res.* 28:235–242.
- Bower, M. J., F. E. Cohen, and R. L. Dunbrack Jr. 1997. Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: a new homology modeling tool. *J. Mol. Biol.* 267:1268–1282.
- Cassiano, M. M., and J. A. G. Arêas. 2001. Study of bovine  $\beta$ -casein at water/lipid interface by molecular modeling. *J. Mol. Struct. Theochem.* 539:279–288.
- Catterall, W. A. 2000. From ionic currents to molecular mechanisms: the structure and function of voltage-gated sodium channels. *Neuron.* 26:13–25.
- Dunbrack, R. L., Jr., and F. E. Cohen. 1997. Bayesian statistical analysis of protein sidechain rotamer preferences. *Protein Sci.* 6:1661–1681.
- Dunbrack, R. L., Jr., and M. Karplus. 1993. Backbone-dependent rotamer library for proteins: application to side-chain prediction. *J. Mol. Biol.* 230:543–574.
- Dunbrack, R. L., Jr., and M. Karplus. 1994. Conformational analysis of the backbone-dependent rotamer preferences of protein sidechains. *Nat. Struct. Biol.* 1:334–340.
- Griffiths, D. J. 1989. Introduction to Electrodynamics. Prentice-Hall, Inc., Englewood Cliffs, NJ.
- Kellenberger, S., J. W. West, W. A. Catterall, and T. Scheuer. 1997. Molecular analysis of potential hinge residues in the inactivation gate of brain type IIA  $Na^+$  channels. *J. Gen. Physiol.* 109:607–617.
- Kelley, L. A., R. M. MacCallum, and M. J. E. Sternberg. 2000. Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J. Mol. Biol.* 299:499–520.
- Kneller, D. G., F. E. Cohen, and R. Langridge. 1990. Improvements in protein secondary structure prediction by an enhanced neural network. *J. Mol. Biol.* 214:171–182.
- Laskowski, R. A., M. W. MacArthur, D. S. Moss, and J. M. Thornton. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26:283–291.
- McClelland, J. L., and D. E. Rumelhart. 1988. Explorations in Parallel Distributed Processing, Vol 3. MIT Press, Cambridge, MA. 318–362.
- Miller, J. R., M. K. Patel, J. E. John, J. P. Mounsey, and J. R. Moorman. 2000. Contributions of charged residues in a cytoplasmic linking region to  $Na$  channel gating. *Biochim. Biophys. Acta.* 1509:275–291.
- Moorman, J. R., G. E. Kirsch, A. M. Brown, and R. H. Joho. 1990. Changes in sodium channel gating produced by point mutations in a cytoplasmic linker. *Science.* 250:688–691.
- Moret, M. A., P. G. Pascutti, P. M. Bisch, and K. C. Mundim. 1998. Stochastic molecular optimization using generalized simulated annealing. *J. Comput. Chem.* 19:647–653.
- Noda, M., T. Ikeda, T. Kayano, H. Suzuki, H. Takeshima, M. Kurasaki, H. Takahashi, and S. Numa. 1986. Existence of distinct sodium channel messenger RNAs in rat brain. *Nature.* 320:188–192.
- Oh, D., S. Y. Shin, S. Lee, J. H. Kang, S. D. Kim, P. D. Ryu, K. S. Hahn, and Y. Kim. 2000. Role of hinge region and the tryptophan residue in the synthetic antimicrobial peptides, cecropin A(1-8)-magainin 2(1-12) and its analogues, on their antibiotic activities and structures. *Biochemistry.* 39:11855–11864.
- Pascutti, P. G., L. J. El-Jaik, P. M. Bisch, K. C. Mundim, and A. S. Ito. 1999b. Molecular dynamics simulation of  $\alpha$ -melanocyte stimulating hormone in a water-membrane model interface. *Eur. Biophys. J.* 28:499–509.
- Pascutti, P. G., K. C. Mundim, A. S. Ito, and P. M. Bisch. 1999a. Polarization effects on peptide conformations at water-membrane interface by molecular dynamics simulations. *J. Comput. Chem.* 20:971–982.
- Patton, D. E., J. W. West, W. A. Catterall, and A. L. Goldin. 1992. Amino acid residues required for fast  $Na^+$ -channel inactivation: charge neutralizations and deletions in the III-IV linker. *Proc. Natl. Acad. Sci. U. S. A.* 89:10905–10909.
- Perozo, E., and F. Bezanilla. 1990. Phosphorylation affects voltage gating of the delayed rectifier  $K^+$  channel by electrostatic interactions. *Neuron.* 5:685–690.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and R. T. Flannery. 1992. Numerical Recipes in Fortran: The Art of Scientific Computing. Cambridge, University Press.
- Rohl, C. A., F. A. Boeckman, C. Baker, T. Scheuer, W. A. Catterall, and R. E. Klevit. 1999. Solution structure of the sodium channel inactivation gate. *Biochemistry.* 38:855–861.
- Rooman, M. J., J. P. Kocher, and S. J. Wodak. 1991. Prediction of protein backbone conformation based on seven structure assignments: influence of local interactions. *J. Mol. Biol.* 221:961–979.
- Rullmann, J. A. C. 1996. AQUA, Computer Program. Department of NMR Spectroscopy, Bijvoet Center for Biomolecular Research, Utrecht University, The Netherlands.
- Sato, C., Y. Ueno, K. Asai, K. Takahashi, M. Sato, A. Engel, and Y. Fujiyoshi. 2001. The voltage-sensitive sodium channel is a bell-shaped molecule with several cavities. *Nature.* 409:1047–1051.
- Scott, W. R. P., P. H. Hunenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Kruger, and W. F. van Gunsteren. 1999. The GROMOS biomolecular simulation program package. *J. Phys. Chem. A.* 103:3596–3607.
- Stühmer, W., F. Conti, H. Suzuki, X. Wang, M. Noda, N. Yahagi, H. Kubo, and S. Numa. 1989. Structural parts involved in activation and inactivation of the sodium channel. *Nature.* 339:597–603.
- West, J. W., R. Numann, B. J. Murphy, T. Scheuer, and W. A. Catterall. 1991. A phosphorylation site in the  $Na^+$  channel required for modulation by protein kinase C. *Science.* 254:866–868.
- West, J. W., D. E. Patton, T. Scheuer, Y. Wang, A. L. Goldin, and W. A. Catterall. 1992. A cluster of hydrophobic amino acid residues required for fast  $Na^+$ -channel inactivation. *Proc. Natl. Acad. Sci. U. S. A.* 89:10910–10914.
- Wintjens, R. T., M. J. Rooman, and S. J. Wodak. 1996. Automatic classification and analysis of  $\alpha$ -turn motifs in proteins. *J. Mol. Biol.* 255:235–253.