Scuola Internazionale Superiore di Studi Avanzati - Trieste

Structural predictions of HCN/CNG ion channels: Insights on channels' gating

Thesis submitted for the degree of

Doctor Philosophiæ

Candidate:

Alejandro Giorgetti

Supervisors:

Prof. Paolo Carloni

Prof. Vincent Torre

October 2004

SISSA – Via Beirut 2-4 – 34014 TRIESTE – ITALY

SISSA ISAS

SCUOLA INTERNAZIONALE SUPERIORE DI STUDI AVANZATI
INTERNATIONAL SCHOOL FOR ADVANCED STUDIES

Structural predictions of HCN/CNG ion channels: Insights on channels' gating

Thesis submitted for the degree of

Doctor Philosophiæ

Candidate:

Alejandro Giorgetti

Supervisors:

Prof. Paolo Carloni

Prof. Vincent Torre

October 2004

'...estas personas son las que hacen a la vida fantástica
y digna de ser bebida hasta la última gota.....'.
A ellos dedico mi vida entera y esta tesis,
A mis dos amores:

Ro y Santi

Table of Contents

1 Introduction and Motivation of this thesis.

Evolution built a membrane around the earliest forms of life in order to isolate them from the external environment. The cell membrane is constituted by two layers of phospholipids, which are molecules having a polar head and non-polar tails. Two films of these molecules are assembled together by hydrophobic forces building a very stable lipid bilayer.

Inserted in this amphiphilic environments are membrane proteins, which have both hydrophobic and hydrophilic regions on their surface. In highly evolved and specified cells this class of proteins carries out a variety of different activities essential for the cell and organism life, like the antibody recognition in lymphocytes and the nervous pulse transmission in neurons.

Although the presence of the membrane helps cells to retain vital ingredients, it prevents the access to necessary ionized substrates and ions, because the hydrophobic core is a high free energy barrier in the diffusion of charged molecules.

Membrane spanning pores are a common feature to ionic channels (Hille, 2001;Chang et al., 1998), and they are presents in different classes of other biological transporter proteins like bacterial porins and aquaporins.

Special membrane proteins, the ionic channels, form holes through the cell membrane, providing a feasible path for ion exchanges.

1.1 Ionic channels

Ion permeation is crucial for a variety of biological functions such as nervous signal transmission and osmotic regulation (Hille, 2001). Many human and animal diseases are also associated to defects in ionic channels function, the majority of them arising from mutations in the genes encoding the channel proteins. A lot of effort is still necessary to connect these mutations to the structural and functional changes causing the disorder.

Ionic channels are proteins inserted in the membrane lipid bilayer, as shown in Fig. 1.1, by forming aqueous pores through the cell membrane (Hille, 2001). They allow ions to cross the hydrophobic barrier of the core membrane, guarantying to the cell a controlled exchange of ionized particles.

In resting conditions the internal side of a cell has a lower electric potential than the external side (Hille, 2001). When ions flux through the channels, they produce an electric current accompanied by changes of the membrane potential. Different kinds of ionic channels are present in nature, differing in functional, gating and selectivity properties.

The open state of the channel pore may be regulated by different factors that open and close the channel mouth allowing and preventing the ionic flux. The gating mechanisms can be based on physical factors like the transmembrane potential difference such as happens in voltage gated channels(Catterall, 2000), or pressure-induced structural changes as the case of mechanosensitive channel(Chang et al., 1998). In other classes of channels, chemical factors are also determinant in triggering the channel gate, as the presence of ligand molecules (i.e. in cyclic nucleotide activated channels) or pH variations (i.e. in KcsA potassium channel (Cuello et al., 1998;Schrempf et al., 1995)).

An important property of the open pore is the selective permeability, which is the ability to allowing only to a restricted class of ions to flow trough the channel pore in large amount. Channels may require to permeating ions having some specific valence, selecting anions over cations like in chloride channels (Jentsch et al., 1995), or selecting divalent cations over monovalent cations like in calcium channels(Hille, 2001). In Ionic selectivity is commonly measured by patch clamp techniques using biionic conditions, namely bathing the outer (o) and inner (i) membrane sides with solutions containing different ionic species A,B. In these conditions the zero current reversal potential E_{rev} defines the permeability ratio P_A/P_B between the two species trough the equation:

$$E_{rev} = \frac{RT}{zF} \ln \frac{P_A [A]_o}{P_B [B]_i},$$

where z is the ionic valence, RT the Boltzmann factor, and F the Faraday constant.

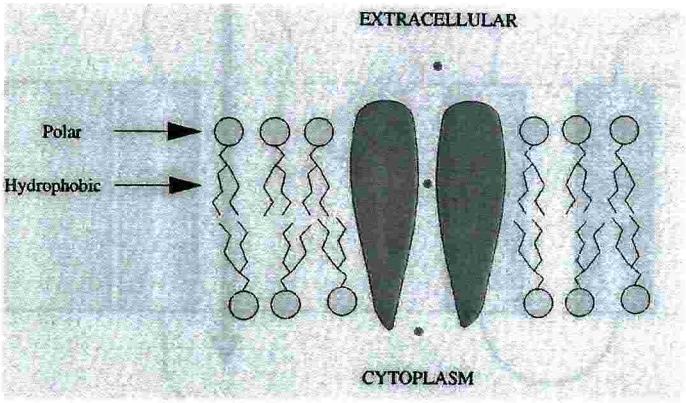


Fig. 1.1. Schematic representation of ions (green dots) flowing trough an ionic channel (green cartoon) embedded in the phospholipid bilayer of the cell membrane (gray sketches).

Thus, ionic channels differ in functional, gating and selectivity properties. Among the large variety of cells that accommodates channel proteins, a special subset of cell families is more than others linked to these biological electric devices: neurons. Cells of the human and animal nervous systems transduce and propagate electrical signals through membrane pores that are highly selective: the voltage-gated K^+ and Na^+ channels.

1.2 Voltage-gated K^+ channels

A mayor object of studies since the early years of biophysics has been the investigation of the mechanisms regulating the nervous signal propagation. A nervous pulse is an electrical signal produced by the ionic flux trough the plasmatic membrane of a neuronal cell.

In neurons, as in other cells, the intracellular K^+ concentration is much higher than that of Na^+. ATP-based active transporters maintain these gradient concentrations, pumping ions against their electrochemical gradient(Hille, 2001). In absence of signal the membrane potential, defined as the difference between the potential inside the cell and that outside the cell, is –60 mV and the membrane is mainly permeable to K^+ ions. During the signal transmission the membrane polarization changes, generating a 'potential wave', the action potential, that propagates along the axon. In these conditions the membrane assumes even positive potentials before restoring the equilibrium resting potential.

In 1952 quantitative measurements of membrane current of the squid giant axon demonstrated that the Na^+ and K^+ membrane permeabilities change during the

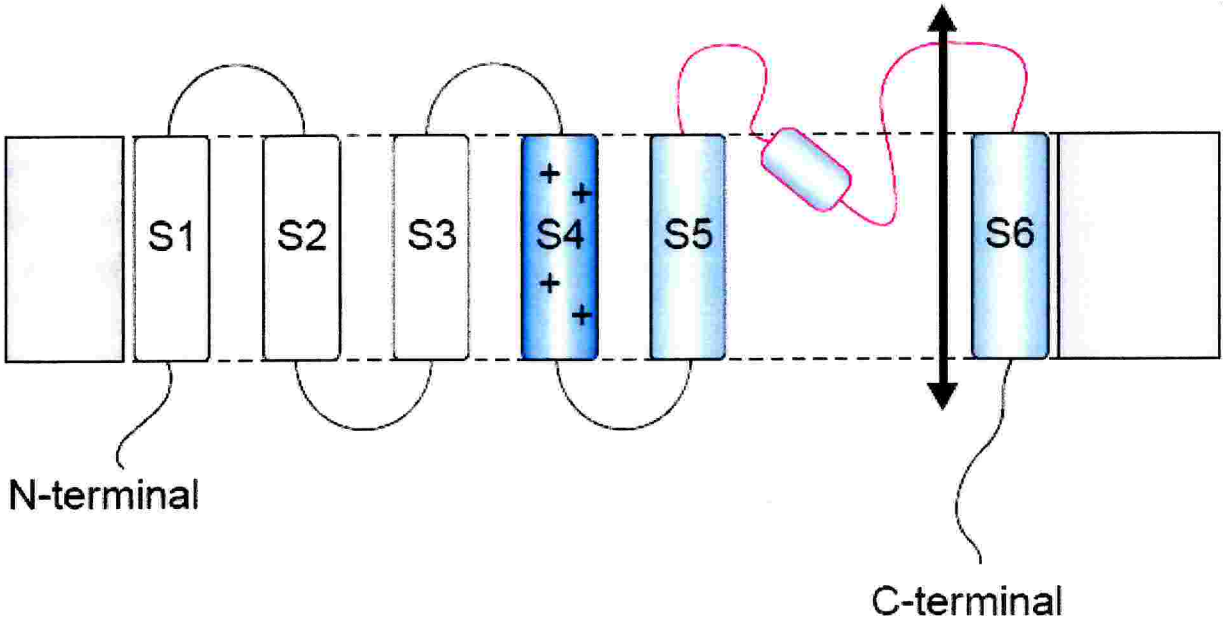


Fig. 1.2. *Schematic representations of the transmembrane domains of K^+ channels. These are made up of either six (S1-S6) or two (S5-S6) TMHs as well as of the so-called P-loop (in red). These domains are assembled in tetrameric structures.*

action potential (Hodgkin and Huxley, 1952). The Na^+ and K^+ ionic fluxes through the membrane may be separated in two different current contributions. Potassium and sodium follow different ways in crossing the membrane, using high selective channels, which open and close as a function of the membrane potential. The interplay between these two currents leads to the propagation of the nervous signal along the axon. During this process, although the membrane potential undergoes large variations (from –60mV to 30mV and again to –60mV), only a minor part of the alkali ions (~ 0.001 %) passes trough the plasma membrane.

Because their account for the action potentials in neuronal cells, Na^+ and K^+ channels of axons were the first channels to be recognized and described in details.

K^+ channels are highly selective proteins (in particular they are more selective than Ca^+ or Na^+ channels). Typical values for the P_{Na+}/P_{K+} permeability ratios are in the range 0.01-0.1 (Hille, 2001).

The first potassium channel gene was cloned in 1987 from fruit fly *Drosophila* (the *Shaker* gene)(Tempel et al., 1987). The *Shaker* channel became a paradigm of the voltage gated K^+ channels: not only opened the avenue to the discovery of other K^+ channel

genes, but also started the combined use of sequence manipulation with patch clamp techniques(Heginbotham et al., 1994).

The primary structure of the gene revealed the presence of one cluster of putative membrane-spanning segments. The overall structure of K^+ channels has been found to be tetrameric(MacKinnon, 1991), formed by four identical subunits assembled together around the channel pore. Hydrophobicity plots, gives us a picture of the transmembrane topology: six membrane spanning helices could be identified as shown in Fig. 1.2. Totally, a boundle of 24 helices is a common motif in voltage gated K^+ channels.

The S4 helix is responsible for the voltage dependence of the channel, and contains a highly conserved sequence of positively charged amino acids. This voltage sensor regulates the channel conductivity as a function of the membrane potential, and similar structures are also common features to Ca^+ and Na^+ channels.

Between S5 and S6 is located the so-called pore region, about 30 amino acids lining the channel pore. The selectivity filter is the fingerprint of potassium selectivity (see the sequence alignment in Fig. 1.3) consisting in a sequence of five highly conserved residues in all the potassium channel family located in the middle of the pore region. Modification of these conserved amino acids leads to strong alteration and often destruction of the potassium selectivity(Heginbotham et al., 1994). The selectivity filter was supposed to be the narrowest part of the channel pore, through which ions needs to flow during permeation. The simultaneous presence of more than one K^+ ion in the selectivity filter was postulated for the first time by Hodgkin and Keynes, observing that the ion flux, as a function of the potassium concentration, is better described assuming the ions acting like a multivalent unit than as single particles. In a narrow pore channel there is no space for two ions for exchanging position without exiting the pore. Therefore, the presence of many ions causes that the ionic flux needs to occur in a concerted, single-file motion(Hille, 2001).

The multi-ion nature of a channel is also reflected in the concentration dependent permeability ratios. In biionic conditions, indeed, a flux-coupling situation occurs providing strongly dependence of the two different ionic fluxes from each other.

1.3 The KcsA channel

As all membrane proteins, ionic channels are difficult to crystallize. In 1995 the cloning (Schrempf et al., 1995) of the first bacterial gene encoding for a potassium selective channels (the *KcsA* gene) opened the avenue to large production of purified K^+ channel, and led in 1998 to the *KcsA* channel crystallization and determination of the 3D structure(Doyle et al., 1998;Zhou et al., 2001).

In this section, after having summarized the general features of the KcsA bacterial channel, we describe its crystal structure in detail.

1.3.1 General Properties

Bacterial genomes shows sequences that are homologues to eukaryotic potassium channels (Schrempf et al., 1995;MacKinnon and Doyle, 1997).

After the KcsA gene from *Streptomyces lividans* was cloned(Schrempf et al., 1995), mM quantities of bacterial channels have been purified and successfully reconstructed in detergent micelles for functional studies(Schrempf et al., 1995).

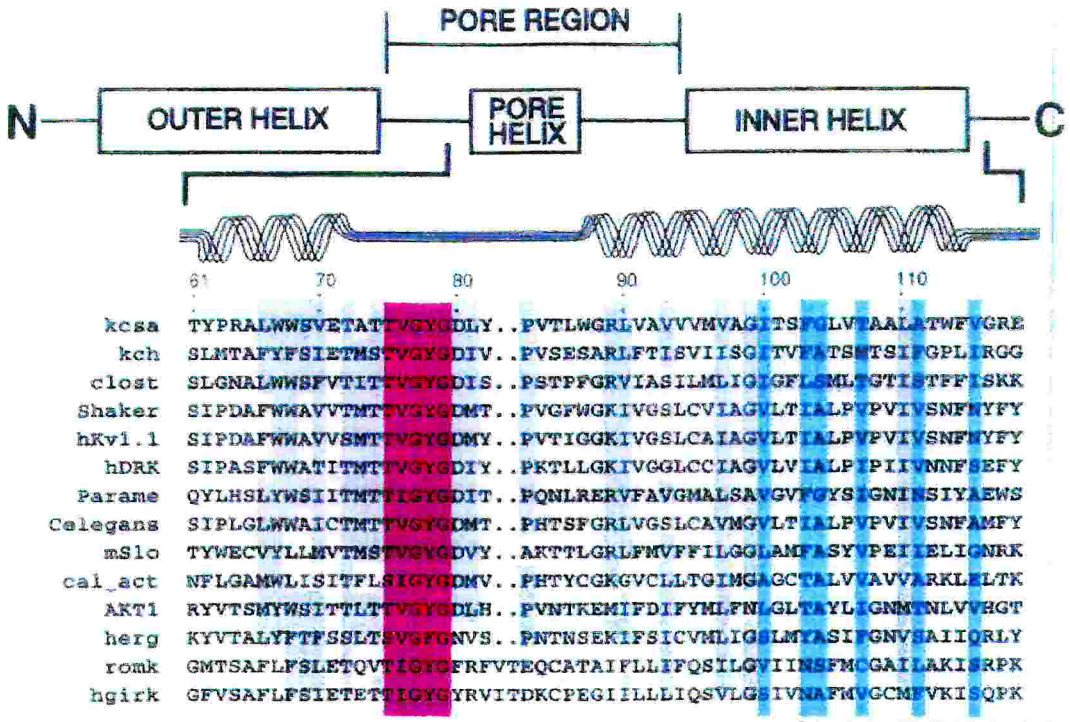


Fig. 1.3. Sequence alignment of the K^+ channel family. Figure from ref.(Doyle et al., 1998).

Single channel measurements on reconstructed proteins incorporated into lipid bilayers revealed a strong pH dependence of the channel opening probability(Cuello et al., 1998). High proton concentration (pH <5.0) promotes channel opening, while in neutral conditions no significant single channel activity was observed. Recent investigations suggested that the protonation site responsible of channel opening is located in the intracellular side(Heginbotham et al., 1999), despite previous works supposed an extracellular gating site(Heginbotham et al., 1992;Perozo et al., 1999;Perozo et al., 1998). The channel exhibits a K^+ selectivity over Na^+ and other cations. The measured permeability ratio P_{Na+}/P_{K+} does not show a strong selectivity: 0.35 (Schrempf et al., 1995), 0.1 (Heginbotham et al., 1998a), 0.09 ± 0.01(Meuser et al., 1999) (see also ref.(Cuello et al., 1998)). However, because the difficulties of such a measurements, it is not suitable attribute to them definitive quantitative values(Heginbotham et al., 1998b).
Other important issue about the similarity between KcsA an eukaryotic channels is the behavior of the different channels in presence of current blockers, like Ba^{2+}, tetraethylammonium (TEA) and toxins(Hille, 2001). The presence of intracellular barium blocks the KcsA channel as happens for all the other K^+ channel(Heginbotham et al.,

1998b;Meuser et al., 1999). Application of TEA to either side of the membrane shows that KcsA is also sensitive to the ion presence(Cuello et al., 1998;Meuser et al., 1999;Heginbotham et al., 1999).

Families of high specific K^+ channel inhibitors are peptide toxins present in scorpion venom. These powerful inhibitors tightly interact in specific sites with the extracellular mouth of the protein, occluding the pore and blocking ion transport. Using site directed mutagenesis on the bacterial channel has been possible to make a KcsA mutant sensible to a *Shaker*-acting toxin(Garcia et al., 1999), showing the structural conservation between prokaryotic and eukaryotic channels(MacKinnon et al., 1998).

The KcsA channel is much smaller than the eukaryotic analogues. It has only two transmembrane helices against six in voltage gated K^+ channels, namely the S5 and S6 helices of Fig. 1.2. The overall pore region sequence presents large similarity to the potassium channel family (Fig. 1.3), and, in particular, it has the same selectivity filter sequence (TVGYG) of *Shaker* channel. Despite the smaller dimensions the protein is still able to preserve pore properties like permeation, blocking and selectivity close to these of larger K^+ channels. On the other hand, its simplicity allowed it to be a perfect candidate for structural studies.

1.3.2 The structure

The KcsA potassium channel is formed by four identical subunits, each composed of 158 residues, among which 97 are detected in the crystallographic structure(Doyle et al., 1998;Zhou et al., 2001)(Fig 1.4). In panel A from Fig.1.4 the two opposite subunits are represented by ribbons. The tetrameric structure of the protein is emphasized in the Fig.1.4 A, by using different colors for the two subunits. Each monomer contains three α helices: two transmembrane (TM) helices that cross the cell membrane and one shorter helix, located at the extracellular side. The membrane-spanning overall structure is built by a TM helix bundle forming a sort of funnel with the helix converging point near the intracellular solution (Fig. 1.4). The outer helix TM1 is the analogue of S5 in voltage-gated channels (see Fig. 1.2), and forms the majority of the protein-membrane interface. The inner helix TM1 (analogue of voltage-gated S6) lines the central pore forming an hydrophobic cavity of ~ 8 Å diameter in the middle of the membrane region. The four

shorter helices (the Pore helices) lie near the selectivity filter, pointing their axes towards the cavity center.

The selectivity filter regions is a narrow tunnel with a diameter of ~ 3 Å, lined with backbone carbonylic oxygens (Doyle et al., 1998;Zhou et al., 2001). Protein oxygens replace the K^+ hydration shell in the bulk water, compensating for the loss in energy due to the ion dehydration.

Three binding sites for the potassium atoms were identified by Rb^+ and Cs^+ difference Fourier maps

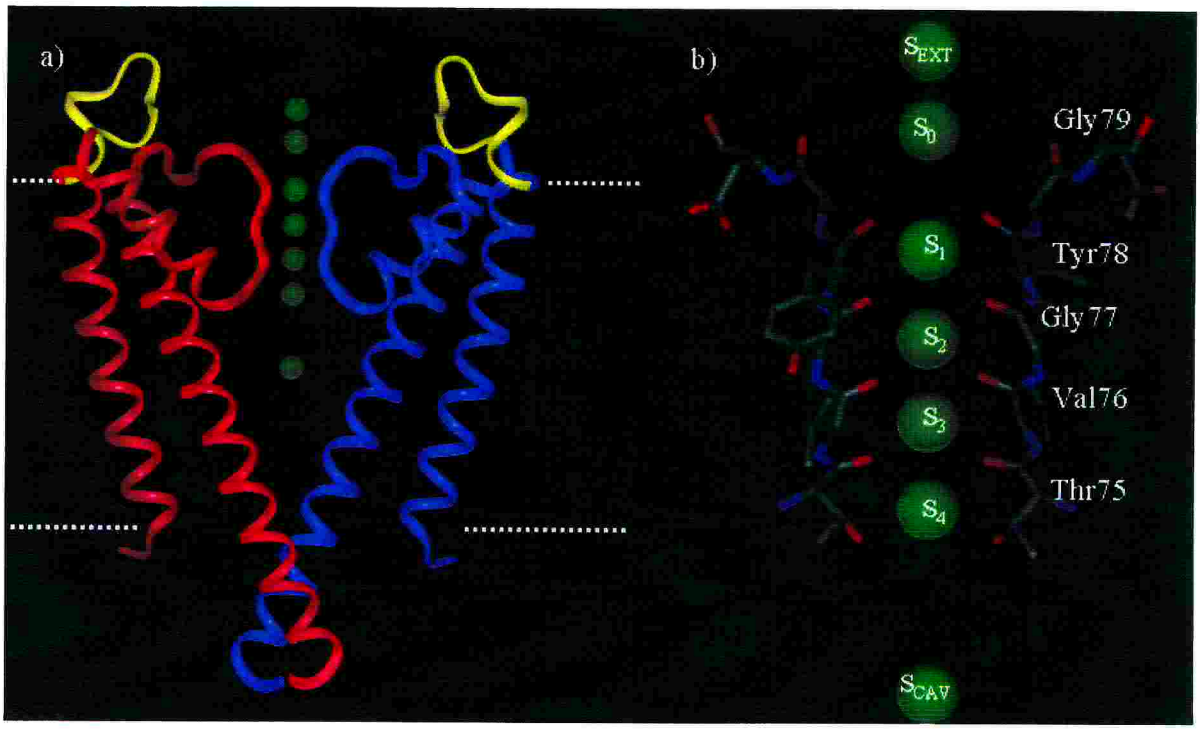


Fig. 1.4 *KcsA channel crystal structure (Zhou et al., 2001). For the sake of clarity only two opposite subunits are shown. View is from a direction parallel to the membrane plane. Putative K binding sites are shown.*

(Roux and MacKinnon, 1999;Doyle et al., 1998), in agreement with the traditional view that K^+ channels are multi-ion channels (green spheres in Fig1.4). Backbone carbonylic oxygens belonging to the all four chains bind the potassium ions inside the selectivity filter (Fig. 1.4B). Because the closeness between the two more intracellular binding sites detected in crystals ($K^+_{(2a)}$, $K^+_{(2b)}$) is unlikely that they are simultaneously occupied by two ions (Doyle et al., 1998). Crystallographers (Doyle et al., 1998;Roux and

MacKinnon, 1999) therefore proposed that the structure accommodate three ions: two into the selectivity filter binding sites ($K^+_{(1)}$, $K^+_{(2)}$) and one into the internal cavity ($K^+_{(3)}$). An additional K^+ ion is believed to be accommodated in the cavity center, as a diffuse electron density cloud revealed in the difference Fourier maps(Doyle et al., 1998;Roux and MacKinnon, 1999).

1.4 Recently solved crystal structures and Gating mechanism

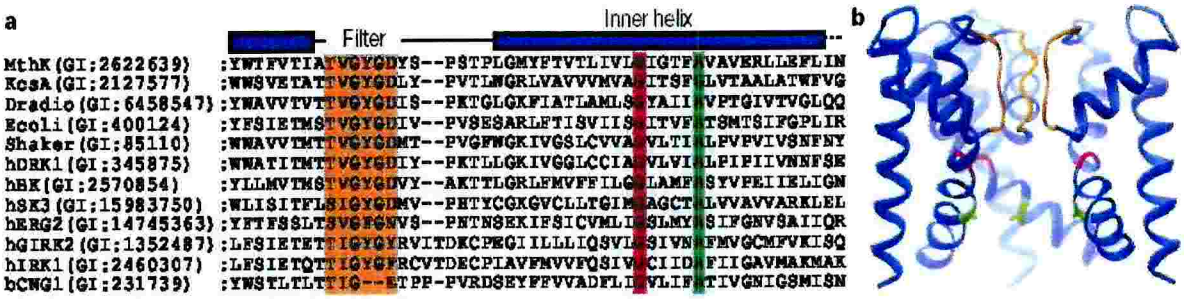


Fig. 1.5. A) Sequence alignment of the K channel family, together with bovine rod CNG channel. B) Crystal structure of the MthK K^+channel in the open configuration. A) and B) The conserved glycine (gating hinge) and the conserved Ala/Gly five residues downstream from the hinge are shown in red and green, respectively.

In the last three years, together with a refinement of the KcsA K^+ channel from *Streptommyces lividans* (Doyle et al., 1998;Zhou et al., 2001), other 3D structures of K^+ channels have been solved, namely: the complete sequence of the voltage gated K^+ channel from *Aeropyrum pernix* (KvAP)(Jiang et al., 2003) and the inward rectifier K^+ channel from *Burkholderia Pseudomallei* KirBac1.1 (Kuo et al., 2003) in the closed state, and the K^+ MthK channel from *M. Thermautotrophicum* in the open state (Jiang et al., 2002a)(Figs. 1.5 and 1.6). In fact, these structures on different K^+ channel families of ionic channels confirmed that potassium channels share several important features. Firstly, in both families the pore domain includes four identical subunits (Ludwig et al., 1998), each featuring a loop forming the inner pore region together with an additional small helix (P-helix), not spanning the lipid membrane. Each subunit of voltage gated K^+ channels features six transmembrane helices (referred to as S1-S6), while subunits of the

KcsA K^+ and MthK channels have only two transmembrane helices (referred to as TM1-TM2), structurally similar to S5 and S6 in voltage-gated K^+ channels. S6 and TM2 are involved in gating, whereas the P-helix-loop does not change conformation upon gating (Doyle et al., 1998;Zhou et al., 2001;Jiang et al., 2002a). Secondly, the aminoacid sequence of K^+ channels in the inner pore region is highly conserved and contains the GYG fingerprint, called the selectivity filter (Heginbotham et al., 1992), that do not change conformation upon channels gating(Heginbotham et al., 1994). Thirdly, based on the open MthK, the closed KcsA channel, and in the amino-acid sequence analysis, Mackinnon and collaborators, have proposed a structural basis for gating transitions in the transmembrane pore of K^+ channels. The four inner helices can exists in a straight (apparently relaxed) conformation in which case they form a bundle that closes the pore near its intracellular opening. Alternatively the inner helices can bend, causing the bundle to splay open to a diameter of about 12 Å. The principal structure feature of the gating conformational change is the presence of a gating hinge located deep within the membrane, conserved as a glycine residue in most K^+ channels. The gating conformational changes within the membrane are very large, but are mainly confined to the intracellular half of the channel. Based on amino-acid sequence conservation, Mackinnon and co-workers have proposed that different K^+ channels, ligand- and voltage-gated, as well as cyclic nucleotide-gated channels, undergo similar pore conformational changes.

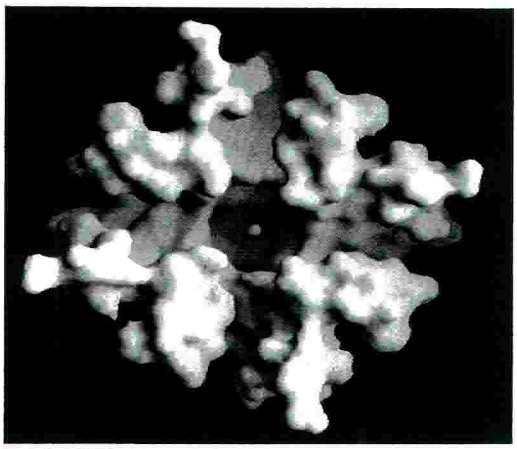


Fig 1.6. *The open pore of the MthK channel allows entry of large molecules from the intracellular solution. Molecular surface of the MthK pore viewed from the intracellular side. A K^+ ion in the selectivity filter is shown as a green sphere*(Jiang et al., 2002a).

1.5 Cyclic Nucleotide gated channels.

As previously said, similarities in structure and mechanism may be expected also for cyclic nucleotide gated channels. In this case, we will study two nonselective cation channels that are important effectors of cAMP. Both channels have the 6-TM architecture of the voltage-gated K^+ channel superfamily and a classical cyclic nucleotide binding domain in their cytoplasmic C-terminal tail. They illustrate nicely the evolutionary innovation of new protein functions by combining functional domains from several unrelated proteins (Hille, 2001).

In the following two sections we give a brief introduction to their cellular function as well as to their already known features, as a starting point for the modelling process that we will describe in the following chapters.

1.5.1 HCN Channels

Hyperpolarization-activated cation currents (I_h) have puzzled physiologists since their initial discovery over 25 years ago(Noma and Irisawa, 1976). Unlike most voltage-gated channels, Ih are activated by hyperpolarizing voltage steps to potentials negative to –60 mV, near the resting potential of most cells. This property earned them the designation of If for 'funny'(Brown et al., 1979) or Iq for 'queer'(Halliwell and Adams, 1982). Although these channels clearly contribute to determining the resting potential and resting membrane properties of cells in which they are expressed, their role in dynamic signalling events is less well agreed upon. Ih is perhaps the most widely known for its proposed role in the generation of spontaneous pacemaker activity, both in the heart(DiFrancesco, 1993) and central nervous system(Pape, 1996). As a result, Ih is often referred to as the pacemaker current. However, it is clear that the spontaneous firing of certain cells, such as Purkinje neurons in the cerebellum (Raman and Bean, 1999) and respiratory neurons in the brainstem(Thoby-Brisson et al., 2000), do not require the participation of Ih to generate automaticity. Even in the heart, the importance of Ih as the prime generator of depolarizing pacemaker current has been questioned(Irisawa et al., 1993;Noma et al., 1983).

Cellular function

The recent identification of a family of four mammalian genes that encode the hyperpolarization-activated, cation nonselective channels, HCN1-4(Ludwig et al., 1998), presents the opportunity of addressing the physiological function of Ih at the molecular level.

On the basis of amino acid sequence, these four mammalian genes were shown to be members of the voltage-gated K channel superfamiliy, with a single gene corresponding to a single K^+ channel subunit. Related HCN genes have been cloned from a sea urchin sperm cDNA library (spHCN) and from insect antenna (hvHCN). Indeed, these channels are activated by membrane Hyperpolarization and are regulated by Cyclic Nucleotides, such as cyclic adenosine-mono-phosphate (cAMP). The I_h current that flows through HCN channels (Clapham, 1998)) is carried by a mixture of Na^+ and K^+ ions (Wollmuth and Hille, 1992;Ho et al., 1994). The I_h current is blocked by Cs^+ ions (Fain et al., 1978) in a voltage-dependent way. Furthermore, it was shown that HCN channels control fundamental biological events such as heart beat and neuronal rhythmic activity, providing the biophysical mechanism for the pacemaker ((DiFrancesco, 1993); (Brown and Ho, 1996);(Halliwell and Adams, 1982); (Pape, 1996)) in these cells.

Architecture of the functional domains

HCN isoforms are highly conserved in their core transmembrane regions and cyclic nucleotide binding domains (80-90% identical)(Santoro and Tibs, 1999). However, the genes products diverge in their amino- and carboxy-terminal extreme cytoplasmic regions. Each HCN subunit is composed of six transmembrane segments, with a positively charged S4 voltage sensor, similar to the voltage sensor of depolarization-activated channels (Fig. 1.2).

All four isoforms have been expressed in heterologous cells and shown to form homomeric, hyperpolarization-activated, non selective cation currents that are directly modulated by cAMP. However, the four HCN isoforms give rise to channels that differ in their kinetics, steady-state voltage dependence and potency of modulation by cAMP. Given the diversity of native Ih in different cell types, a key goal is to relate the cellular heterogeneity to the molecular composition of the channel. This is complicated by the

fact that native pacemaker currents exhibit greater diversity, for example with regard to voltage dependence, than the expressed isoforms.

A striking feature of the four HCN isoforms is the presence of a 120-amino acid cyclic nucleotide binding domain (CNBD) in their cytoplasmic carboxy terminus, which mediates the response to cAMP. The CNBD is homologous to similar regions in other cyclic nucleotide binding proteins, including the cAMP-and cGMP-dependent protein kinases, a bacterial cAMP binding protein, the catabolite-activating protein (or CAP), and the cyclic nucleotide-gated ion channels of olfactory neurons and photoreceptors.

Deletion of the CNBD mimics the effect of cAMP by shifting the voltage dependence of HCN gating to more positive voltages by an amount similar to the maximal shift seen with saturating concentrations of cAMP. This indicates that cAMP enhances gating by relieving a tonic inhibitory action of the CNBD, which shifts gating in the absence of cAMP to more negative potentials. Differences in the voltage dependence of gating and potency of cAMP modulation between HCN1 and HCN2 were shown to result from differences in the extent of inhibition produced by the CNBD, with HCN2 having a much greater inhibition than HCN1.

The finding that HCN channels contain a positively charged S4 voltage sensor, similar in structure to the S4 region of depolarization-activated Shaker K^+ channels, raises the question as to the mechanism of hyperpolarization gating. Is S4 the HCN voltage sensor and does it move in response to changes in the membrane potential similar to the movement of S4 in depolarization-activated channels? The fact that Ih activation is very steeply regulated by voltage, comparable to depolarization-activated channels, suggests that the S4 domain must act as the voltage sensor because it is the only transmembrane segment that contains the requisite charge necessary to produce such a high voltage sensitivity.

The pore region

As already said, the aminoacid sequence of HCN channels is composed of a cytoplasmic binding domain for cAMP and a transmembrane domain. Whilst the X-ray structure of the cytoplasmatic domain for cAMP of mouse HCN channel (mHCN2: pdb accession number: 1Q3E) has been recently determined (Zagotta et al., 2003), structural information on the pore domain is still lacking. Studies carried on the different HCN

isoforms have suggested a structural similarity between the pore regions of HCN and of K^+ channels for which the 3D structure has been solved, namely the KcsA K^+ channel (Doyle et al., 1998;Zhou et al., 2001), the voltage gated K^+ channel from *Aeropyrum pernix* (KvAP)(Jiang et al., 2003) and the inward rectifier K^+ channel from *Burkholderia Pseudomallei* KirBac1.1(Kuo et al., 2003) in the closed state, and the K^+ MthK channel in the open state (Jiang et al., 2002a). In fact, several important features are shared by both families of ionic channels:

i) the pore domain includes four identical subunits (Ludwig et al 1998). Each subunit contributes to the pore architecture by providing two transmembrane helices (called S5 and S6 in HCN as in voltage-gated K^+ channels), and the P-helix.. Subunits of voltage gated and HCN channels have six transmembrane helices.

ii) The aminoacid sequence of K^+ and HCN channels in the filter region is highly conserved and contains the GYG fingerprint(Heginbotham et al., 1992;Heginbotham et al., 1994).

iii) The closed and open states in the two channel families are believed to be similar, at least to some extent. In K^+ channels, the only significant structural difference during channel opening is the bending by 30 deg of the S6/TM2 helix towards the lipid phase (Jiang et al., 2002b;Jiang et al., 2002a). A similar conformational change is believed to occur also in HCN channels, where S6 bends in the open state (Rothberg et al., 2002;Shin et al., 2001). This hypothesis was first inferred by the conservation of a glycine residue providing the gating hinge, and an alanine (or a glycine) residue, five residues downstream, in HCN and K^+ channels. This view was substantiated by electrophysiological experiments with wild type (wt) and mutant spHCN channels with the bradycardic agent ZD7288, and by analyzing Cd^{2+} blockage in cysteine mutants of spHCN channels(Rothberg et al., 2002).

In spite of these similarities, the main biophysical properties of HCN channel pore domains differ markedly from those of K^+ channels: The permeability ratio (P_{Na}/P_K) between Na^+ and K^+ ions in HCN channels ranges between 0.24 and 0.31 (Ludwig et al., 1998), whereas in K^+ channels ranges between 0.18 and 0.03 (Hille, 2001;Laio A. and

Torre, 1999), indicating that HCN are less selective than K^+ channels. (Rauer et al., 2000;Wrisch and Grissmer, 2000)

1.5.2 CNG channels

Ion channels that are directly activated by cyclic nucleotides were first discovered by Fesenko et al. in 1985 (Fesenko et al., 1985), studying light-dependent channels of rods. Subsequently, within a relatively short time, similar channels were identified in cone photoreceptors (Bonigk et al., 1993), chemosensitive cilia of olfactory sensory neurons (Nakamura and Gold, 1987), and the pineal gland (Dryer and Henderson, 1991). It is now known that CNG channels are not unique to photoreceptors and olfactory sensory neurons, but are also expressed in other neurons and non-neuronal tissues. CNG channels belong to a heterogeneous gene superfamily of ion channels that share a common transmembrane topology and pore structure and that harbor in their C-terminal region a binding domain for nucleoside 3'-5'-cyclic monophosphates. Other members of this superfamily are the previously described HCN channels (Kaupp and Seifert, 2002), the ether-a-gogo (EAG) and human eag-related gene family of voltage activated K^+ channels (Morais Cabral et al., 1998) (Ganetzky et al., 1999), and several plant K^+ channels commonly refereed to as KAT, AKT, and KST channels (Schachtman, 2000).

In the following sections we briefly review some of the most important known functional and structural facts on CNG channels with a particular focus on rod photoreceptors channels, which are one of the objects of this thesis. Extensive reviews are those by Zagotta and Siegelbaum (Zagotta and Siegelbaum, 1996) and by Kaupp and Seifert (Kaupp and Seifert, 2002).

Cellular function

The function of CNG channels has been firmly established in rod and cone photoreceptors, in extraretinal photoreceptors, and in sensory neurons of the olfactory epithelium. Electrophysiological studies of CNG channels in these sensory neurons provided a wealth of information regarding their ligand sensitivity, mechanism(s) of activation, modulation, and ion selectivity. Although CNG channels also exist in other neurons and non-neuronal tissues, their specific functions are yet to be determined

rigorously. In this section we briefly describe the cellular function of CNG channels from rod photoreceptors.

Rods are photoreceptor cells of the vertebrate retina (about 125 million in a human eye) involved with black and white or night vision. They are constituted by three distinct regions: i) A region, synapsed with the neural layer of the retina, that contains the nucleus and other cytoplasmic organelles. ii) The inner segment, rich in mitochondria and connected through a thin 'neck' to the outer segment. iii) The outer segment, that consists of a stack of disks rich in rhodopsin, the pigment that absorbs light.

Rods respond to a light stimulus with a brief hyperpolarization by closing CNG channels in the surface membrane of the outer segment (Yau and Baylor, 1989).

In the dark, channels are activated by the binding of cGMP, allowing a steady cation current (Ohyama et al., 2000) to flow into the outer segment. Light triggers a sequence of enzymatic reactions that leads to the hydrolysis of cGMP. When CNG channels close, the inward current ceases and the cell hyperpolarizes. The CNG channel is crucially important for the control of cytoplasmic Ca^{2+} concentration ($[Ca^{2+}]_i$), because it provides the only source for Ca^{2+} influx into the outer segment. In rods, between 10 and 18% of the dark current (30 pA) is carried by Ca^{2+} (Gray-Keller and Detwiler, 1994).

The cascade of events associated to rod CNG channel activity is summarized in Fig. 3.1. Photons are adsorbed by the photopigment rhodopsin. Light activated rhodopsin (rhodopsin*) complexes with the G protein transducin which, in turn, activates another enzyme, phosphodiesterase (PDE). The latter catalyzes the hydrolysis of cyclic GMP, causing the closure of CNG channels and hence the decrease in $[Ca^{2+}]_i$. Then, the rod cell initiates the recovery from the light response by enhancing the synthesis of new cGMP molecules and adjusts the sensitivity of the transduction machinery, a process known as light adaptation. $[Ca^{2+}]_i$ decline controls at least three biochemical processes. First, the activity of the guanylyl cyclase (GC) that synthesizes cGMP is stimulated as Ca^{2+} levels decrease. Second, number of activated PDE is reduced through the phosphorylation of light-activated rhodopsin (rhodopsin*) by rhodopsin kinase. Finally, the ligand sensitivity of the CNG channel increases as $[Ca^{2+}]_i$ decreases. The regulation of ligand sensitivity by Ca^{2+} is mediated by a third Ca^{2+}-dependent protein, calcium calmodulin (Hsu and Molday, 1993)(Molday, 1996). All three reactions help to restore the dark state and to adjust the light sensitivity of the cell (Leskov et al., 2000).

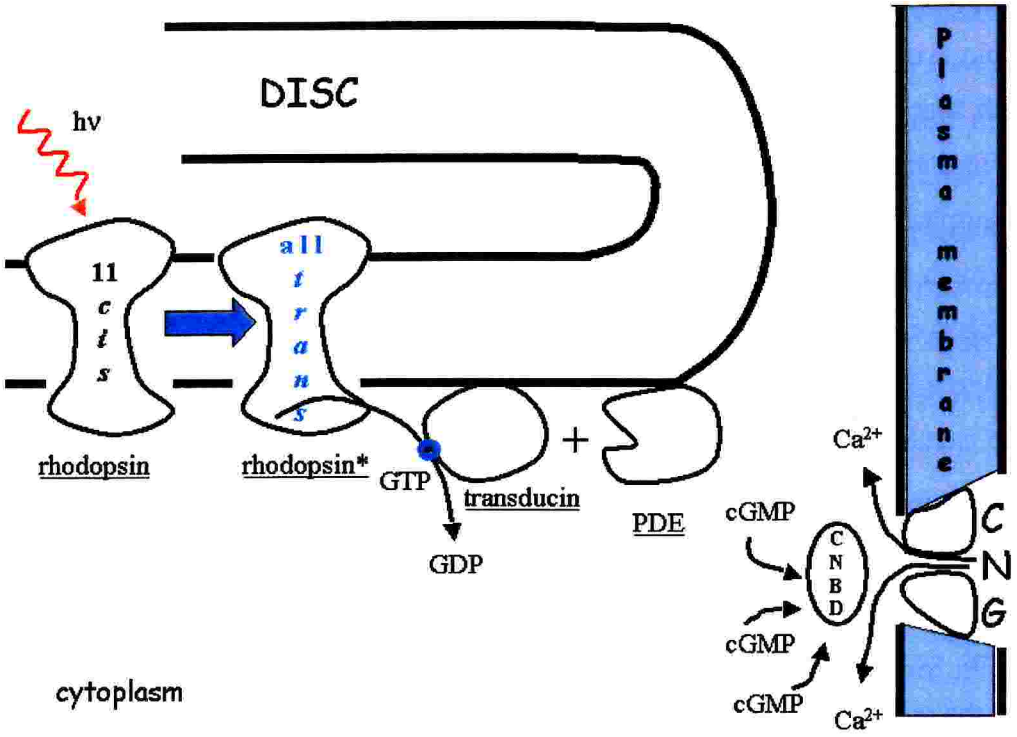


Fig. 3.1 The "vision cascade" (see text).

Architecture of the functional domains

All CNG channels form heterotetramers (Shammat and Gordon, 1999; He et al., 2000). While some of the constitutive subunits were identified to form functional channels on their own (A subunits), other subunits were not and can be considered as modulatory subunits (B subunits). All along the present work we will follow the nomenclature introduced by Bradley et al. (Bradley et al., 2001), which is reported in Table 1.1.

The focus of the present study is the *bovine* rod CNG channel A subunit (CNGA1) (Kaupp et al., 1989). The CNGA1 subunit was the first CNG channel sequence to be cloned (Kaupp et al., 1989) and, for this reason, a great amount of experimental data concerning the homomeric CNGA1 channel have accumulated over the years. Overall, CNGA1 channels have properties similar to those of the native channels (with small differences in the ion permeation and selectivity properties, Kaupp and Seifert, 2002). CNGA1, like other CNG channel subunits, comprises several functional domains, which are described in the following sections.

ADOPTED NOMENCLATURE FOR CNG CHANNEL SUBUNITS	
Adopted nomenclature	**Channel**
CNGA1	Rod channel
CNGA2	Olfactory neurons channel
CNGA3	Cone channel
CNGA4	Modulating subunit of olfactory neurons channel
CNGB1	Modulating subunit of rod channel
CNGB3	Modulating subunit of cone channel

Table 1.1 *Adopted nomenclature for CNG channel subunits (adapted from Bradley et al. (Bradley et al., 2001)).*

Transmembrane region

Overall Structure

No high-resolution 3D structure of the transmembrane moiety is at present available. Therefore, up to now, the channel transmembrane topology has been built based on low-resolution experimental data. The most important experimental facts have been: (i) immunogold labeling for electron microscopy of the rod CNG channel that localized both the N- and the C-terminus to the cytoplasmic side of the plasma membrane. This also implies that CNG channels have an even number of TMHs; (ii) a gene fusion approach consistent with the presence of six TMHs; (iii) the recent determination by electron microscopy (Higgins et al., 2002) of the structure of the native rod channel at low resolution (35 Å), also consistent with the six TMHs model, on the basis of the comparison with electron microscopy structures of other channels (Sokolova et al., 2001; Sato et al., 2001); (iv) several studies that identified, in the region between the putative S5 and S6 TMHs, the determinant for most of the ion selectivity and ion permeation properties of the channel (P region) (Kaupp and Seifert, 2002).

The model for the transmembrane topology of A and B subunits of CNG channels, based on these data, is shown in Fig. 1.2. It comprises six TMHs, designated S1-S6, and the so called P region comprising ~25 residues and located between S5 and S6.
CNG channels, although exhibiting a very low voltage dependence, possess on the S4 transmembrane segment a charged sequence motif, which is reminescent of the "voltage-sensing" S4 segment of voltage-gated K^+ channels. Experiments aiming at understanding why the S4 segment does not confer voltage sensitivity to CNG channels were not conclusive, although evidences suggest that the loop connecting S3 and S4 plays a role in the S4 inhibition (probably by locking the sensor in a fixed conformation) (Tang and Papazian, 1997).

The Pore region

The P region (from Y347 to V368) is part of the ion permeation pathway and is involved in ion selectivity (Kaupp and Seifert, 2002). As previously mentioned, the sequence of the P region has a striking resemblance to that one of voltage-gated K^+ channels. Here, we briefly recall the experimental data concerning ion selectivity, permeation, pore blockage and channel conductance in CNG channels.

Selectivity and permeation

- CNG channels strongly select cations over anions and discriminate poorly among monovalent cations (Kaupp et al., 1989; Cervetto et al., 1988). In particular, the rod native channel has the following sequence of permeability ratios $Na^+ \sim K^+ > Li^+ > Rb^+ > Cs^+$ (Luhring et al., 1990; Furman and Tanaka, 1990).

- By measuring the channel permeability with large organic cations (such as NH_4^+, $CH_3NH_3^+$, $(CH_3)_2NH_2^+$, $(CH_3)_3NH^+$, $CH_2CH_3NH_3^+$), it was possible to estimate the diameter of the narrowest part of the CNG channel pores, which for the bovine CNGA1 is between 4 and 6 Å (Goulding et al., 1993; Bucossi et al., 1996; Laio and Torre, 1999).

- Conductance features depend much on the CNG channel that is considered. In particular, homomeric CNGA1 channels exhibit a single conductance of about 28 pS. In contrast, multiple pH dependent conductance levels have been observed in several CNGA1 channel mutants (Bucossi et al., 1997).
- CNG channels are also permeable to divalent cations such as Ca^{2+}, Mg^{2+}, Sr^{2+}, Ba^{2+} (Park and MacKinnon, 1995). Among these, Ca^{2+} is the divalent cation with the highest affinity for the channel (Park and MacKinnon, 1995). Permeability to this divalent cation is fundamental to the cellular function. In fact, changes in channel activity are associated with changes in Ca^{2+} concentration, which in turn is responsible for initiation of various excitatory or adapting cellular processes. Ca^{2+} acts as a permeant blocker of the channel in that the current carried by monovalent cations is progressively suppressed by increasing concentrations of external Ca^{2+}. Voltage dependence of the Ca^{2+} blockade measurements indicates that a single Ca^{2+} ion, binding to an intrapore binding site, is able to completely block the channel (Seifert et al., 1999a), although at high external concentration pore Ca^{2+} occupancy of the pore is also possible (Dzeja et al., 1999). Glutamate E363 has been identified as a crucial factor in Ca^{2+} binding. When E363 is replaced by a neutral amino acid (e.g. glutamine), Ca^{2+} blockade is almost entirely abolished (Eismann et al., 1994). Substitution with an aspartate residue, *vice versa*, enhances the Ca^{2+} affinity (Park and MacKinnon, 1995). Furthermore, Ca^{2+} affinity is weakened by protonation of E363 (Rho and Park, 1998).

Pore topology

Substituted Cysteines Accessibility Method (SCAM) (Akabas et al., 1992) (see Methods) has been widely used to determine the accessibility of residues in the channel pore. SCAM is particularly effective in identifying residues lining the channel pore region, since residue modification in this region is likely to cause significant current alterations, either through the occlusion of the pore or by impairement of the gating mechanism.

- Overall, SCAM studies (Becchetti et al., 1999; Becchetti and Roncaglia, 2000; Roncaglia and Becchetti, 2001; Liu and Siegelbaum, 2000) have suggested that the topology of the CNG channel P region (Fig. 1.2) is rather similar to that

observed in K^+ channels (Lu and Miller, 1995), for which a crystal structure is also available (Zhou et al., 2001). Indeed, these experiments have pointed to: (i) the internal accessibility of W353C, T359C, T360C and I361C (Becchetti et al., 1999); (ii) the external accessibility of V348, T364C, P365C, and P366C (Becchetti et al., 1999) and Val348, Leu351, Thr355 and Leu358 (Liu & Siegelbaum 2000); (iii) the wider diameter of the extracellular vestibule lumen, lined by residues T364, P365, and P366, with respect to the diameter of the pore lumen at the level of the residues G362–E363; this is supported by two facts: first, the large thiol reagent MTSET can readily reach T364, P365, and P366; second, in mutant channels T364C, P365C, and P366C, no pore occlusion via formation of disulfide bridges between cysteines of neighbouring subunits is observed; (iv) the helical periodicity of the accessibility pattern of residues spanning from Y347 to T360, suggesting that this region folds into an α helix (Liu and Siegelbaum, 2000) like the corresponding KcsA region (pore helix); (v) the location of I361 at the narrowest part of the pore; this, in particular, was deduced from the formation of disulfide bridges in mutant channels I361C (Becchetti et al., 1999); (vi) the helical periodicity of the accessibility pattern of residues located in the C terminal part of S6 (T389 to S399) (Flynn and Zagotta, 2001).

- The residue accessibility is in general state dependent, suggesting the occurrence of structural rearrangements in the pore during channel gating (Becchetti et al., 1999; Liu and Siegelbaum, 2000; Flynn and Zagotta, 2001). This hypothesis is also supported by data on CNGA1 block by intracellular tetracaine. Indeed, Fodor et al. (Fodor et al., 1997a) reported that the tetracaine block, mediated by E363 (Fodor et al., 1997b), is strongly state dependent. Tetracaine affinity for the closed state is three orders of magnitude higher that that one for the open state.

Conformational changes in the pore region during gating

- The only residue in the putative filter region, for which there are experimental evidences indicating a conformational rearrangement is Thr360. Wether a Cd^{2+} blockage of mutant channel T360C is observed only when Cd^{2+} is added in the presence of 1 mM cGMP (open channel), no blockage was observed when Cd^{2+}

was added in the absence of cGMP. These experimental results and the MTSES potentiation in the open and closed state of mutant T360C indicate that residues in position 360 are accessible both in the open and closed state but change their relative conformation. This conformational change so to allow Cd^{2+} ions to coordinate to at least two S atoms of mutant T360C in the open but not in the closed state.

The coupling between the P helix and S6

- Single mutant channel L356C inactivates in the presence of a steady cGMP concentration. However, the double mutant channel L356C & F380C does not inactivate and has a normal gating. These results suggest that when the interaction between residues in position 356 and 380 is disrupted there is an abnormal gating, as in mutant L356C, but when the hydrophobic interaction between side chain of Leu356 and Phe380 is substituted by an S-S bond as in mutant channel L356C&F380C an almost normal gating is restored.

S6 and N-terminal region of C-Linker.

Cd^{2+} blockage

- The spatial profile of Cd^{2+} blockage can be resumed mapping all these results into an ideal alpha helix, as shown in Fig. 1.8. In panels A and B are shown respectively Cd^{2+}. blockage in open and closed state respectively. green indicates residues for which the cGMP activated current in the corresponding cysteine mutant was significantly affected by 100 μM Cd^{2+}. Yellow indicates residues for which the corresponding cysteine mutant was reversibly blocked by Cd^{2+}. Red and orange denotes residues where in the corresponding cysteine mutant there was a complete (larger than 70 %) and significant (between 40 and 70 %) irreversible blockage of the cGMP activated current.

Inspection of Fig.1.8 provides an immediate observation on the spatial pattern of Cd^{2+} action: for residues upstream of Gly395 blockage is reversible- or almost absent - and poorly state dependent. On the contrary Cd^{2+} blockage is irreversible and strongly state dependent for residues downstream Asn400.

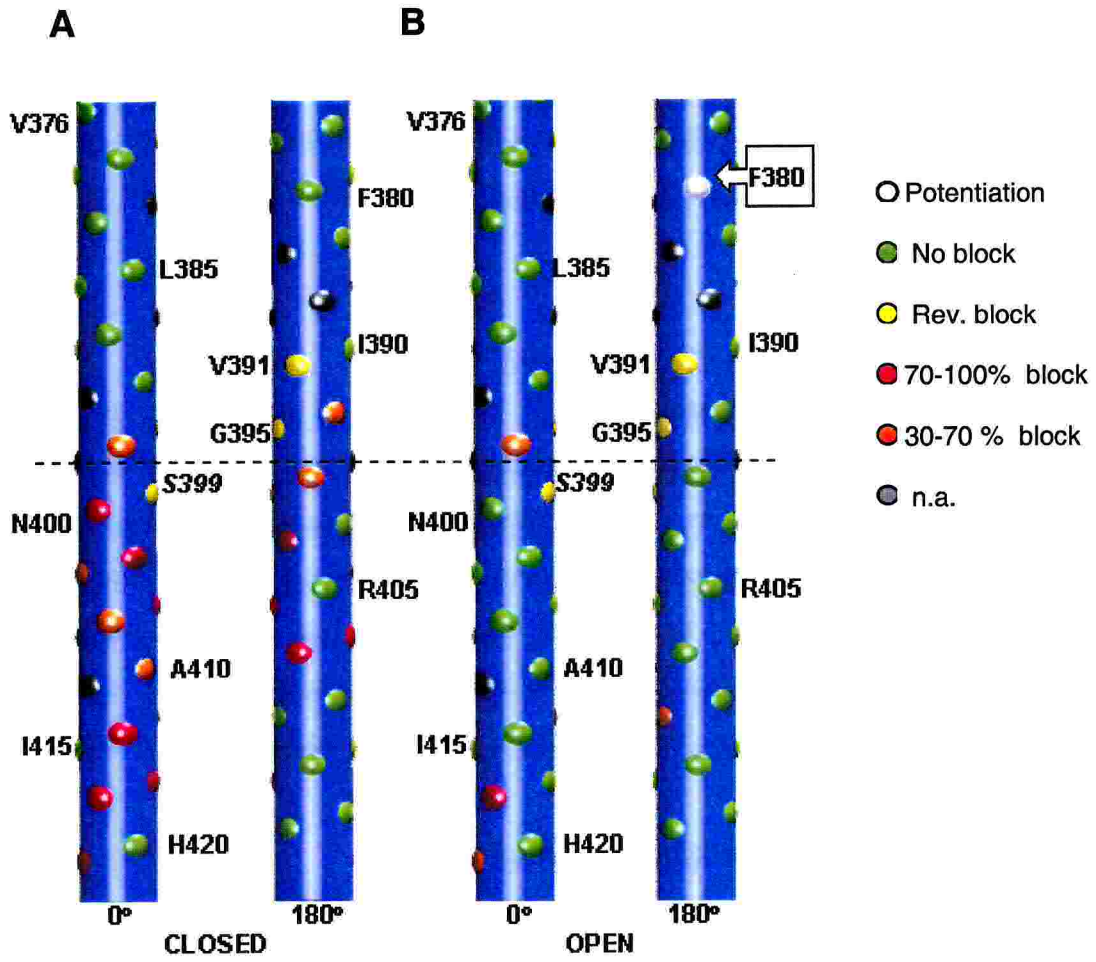


Fig. 1.8: *Summary of Cd^{2+} blockage in S6 from F375C to A424C illustrated as a map plotted on an ideal alpha helix. In each pannel blockage is shown in a color-coded map (Red indicates blockage higher than 70 %, orange between 30 and 70%, green between 0 and 30%, white indicates potentiation and gray are the residues where the corresponding cysteine mutant did not produce functional channels with a cGMP gated current) with residues located on an ideal alpha helix. Two views of the alpha helix, rotated by 180 degrees around its symmetry axis, are shown. First two helices refer to blockage in the presence of 1 mM cGMP and second two helices refer to blockage in the absence of cGMP.*

1.6 Motivation of the present work

The structure the KcsA K+ channel (Doyle et al., 1998) (Zhou et al., 2001; Fig. 1.4) rationalized years of structural studies on ion channels, opening the way to a more profound comprehension of fundamental issues such as ion selectivity and permeation. The impact of the KcsA structure was further amplified by its evolutionary relationship with a large number of other ion channels, comprising voltage gated K^+ channels, voltage gated Ca^{2+} channels and cation selective cyclic nucleotide gated (CNG) channels (Heginbotham et al., 1992). These channels, although much different in many of their physiological characterisations (e.g. selectivity, permeation, gating, modulation), conserve several sequence motives in correspondence to the ion permeation pathway and share a similar 3D folding. This raises the possibility of exploiting structure-based strategies to predict their structures and to design specific inhibitors targeting pharmacologically relevant channels.

In this thesis, we have predicted the structures of two ion channels belonging to different families, namely the HCN and CNG channels, both of them extensively studied in Prof. Torre's lab. Because the homology between these channels and the crystallized K^+ channels is low (around 18%), structural models were refined with the inclusion of experimental data, 'translated' as structural restraints. These two classes of ion channels have allowed us to use rather diverse approaches.

In the HCN channels, for which it is possible to identify few experimentally derived structural constraints, we have constructed a large family of models generated from different initial conditions. Conclusions regarding the rigidity/flexibility properties of the filter and of the gating can be drawn, within the highly approximate models.

In the CNG channels for which a very large number of constraints could be obtained, we have been able to present a rather well-defined structure of the open and closed states.

Thus, we have been able therefore, to present a novel model of the gating, which provides a rational to years of experimental research in this field.

1.7 Outline

In chapter 2 we will summarize the state of the art of ionic channel structural prediction, as an introduction to our work. The main results of the thesis will be presented in Chapters 3-4, where we focus on the structures of two ion channels currently investigated in Prof. Torre's lab. The gating mechanisms emerging from the present study for HCN and CNG channels will be compared between them, and with the one proposed by Mackinnon and coworkers (section 1.4) for K^+ channels. The methodological tools used along the thesis can be found in Appendix A and B.

2 Ion channel modelling: State of the art

Recent advances in membrane protein crystallography have greatly increased our knowledge of ion channel structures. Based on this structural information, structural bioinformatics techniques and molecular dynamics calculations are providing structural information of ion channels for which the three-dimensional structure is not known. Most of the structure prediction studies are based on the KcsA K^+ channel structure. After a brief introduction to the methods used, this chapter presents a survey of the applications appeared so far, which focus on K^+ and Na^+ channels.

2.1 Techniques Used for Structure Prediction

Essentially two approaches have been followed to construct structural models of ion channels. The first is based on *molecular dynamics (MD) simulations,* which explores the conformational space of biomolecules given an appropriate, empirically derived interatomic potential energy function. The protein is often inserted into a lipid bilayer or a membrane-mimic such as a slab of an organic liquid such as n-octane(Zhong et al., 2000) (water/n-octane/water) to take into account the presence of the cell membrane. The explicit membrane model is more realistic than that based on an organic liquid such as n-octane, although at times difficulties might be encountered in equilibrating the system because of the slow dynamics of lipid molecules in the bilayer (the diffusion coefficient of lipids being of the order $10^{-3} - 10^{-4}$ $Å^2$ ps^{-1}) (Faraldo-Gomez et al., 2002).

The second approach is based on *structural bioinformatics tools,* including: 1- secondary structure predictions (SSP) for each residue of the primary sequence, which are based onto statistical methods, physico-chemical methods, sequence pattern matching and evolutionary conservation (Rost, 2001). 2- Identification of transmembrane helices (TMH) based on their hydrophobicity and their known minimum length (at least 15 residues)(Chen and Rost, 2002). 3-Homology modeling (HM), which constructs structural models (targets) that are homologous to other protein(s) whose 3D structure, is known (templates). HM is based on the larger conservation of protein folds than primary sequences during evolution (Richardson, 1977). All of HM-based models of K^+ channels

so far reported use the KcsA as template (Table 2.1). Comparative modeling techniques will be explained in detail in the Appendix A: Methods.

The MD and bioinformatics approaches are highly complementary and they are often used together. For the sake of simplicity, here we schematically distinguish two protocols that are often used: **(i)** SSP followed by TMH identification along with prediction of loops connecting the helices. MD or simulated annealing simulations may finally used to refine the model; **(ii)** HM techniques followed by MD simulations (Tab. 2.1). Occasionally, biological data may be taken into account.

2.4 Structural Predictions

Voltage-gated K+ channels. These channels open and close in response to changes in the potential difference across the membrane. As KcsA, they are selective for K^+ over Na^+ ions (MacKinnon et al., 1998;LeMasurier et al., 2001). Their predicted topology (Fig. 1.2) consists of four TMHs (S1-S4), and the pore region (S5, P-loop, S6). The ion conduction pore is believed to be structurally similar to that of KcsA(Lu et al., 2001).

Ranatunga et al.(Ranatunga et al., 2001) built a model of the pore region of the *Shaker* potassium channel from fruit-fly *D. melanogaster.* This is the most extensively experimentally investigated channel of this family(Tempel et al., 1987). The model was built based on the significant sequence identity with KcsA (Tab. 2.1 and Fig. 2.1). Brief MD simulations of the channel *in vacuo*, in which a K^+ ion was placed in different locations along the pore, were used to calculate the potential energy profiles for ion/protein and ion/water interactions.

The binding of the κ-conotoxin-PVIIA toxin to *Shaker* channel was also investigated(Moran, 2001). A structural model of *Shaker* pore region, similar to that of Ranatunga et al.(Ranatunga et al., 2001), was first built (Tab. 2.1 and Fig. 2.1). The toxin was then docked on the channel structural model using the Gramm geometric recognition algorithm(Katchalski-Katzir et al., 1992). Two basic domains of the toxin turned out to interact with residues at the open mouth of the selectivity filter, consistently with experimental data(Garcia et al., 1999;Terlau et al., 1999;Jacobsen et al., 2000).

The binding of the charybdotoxin (ChTX) toxin to a *human* voltage-gated K^+ (Kv1.3) channel was recently investigated by Rauer et al.(Rauer et al., 2000). This channel is regarded as therapeutic target, as it regulates the membrane potential of resting and

activated T-cells(Cahalan and Chandy, 1997). A structural model of the pore region was constructed based on the relatively high sequence identity with KcsA (Tab. 2.1 and Fig. 2.1). ChTX was docked onto the protein outer mouth guided by mutant cycle analysis, which is an experimental procedure that measures the strength of coupling between interactive pairs of residues belonging to the toxin and channel.

In a similar approach, Wrisch and Grissmer(Wrisch and Grissmer, 2000), docked a peptide toxin, kaliotoxin (KTX) onto a structural model of the pore region of a *mouse* voltage-gated K^+ channel (Kv1.1), which also share a fairly good degree of structural identity with KcsA(Table 2.1 and Fig. 2.1). Results point to significant differences between turrets of both channels (KcsA and Kv1.1).

Based on these toxin/protein complex structures, novel-blocking agents of voltage gated-channels and site-directed mutagenesis experiments have been designed (Rauer et al., 2000).

Fig. 2.1. *Sequence of K^+ channels aligned to KcsA (See Text). The residues forming the selectivity filter region are marked as a red square. Initial and final residues aligned are specified.*

Table 2.1. *Methods used for structure prediction of K^+ channels.*

SSP methods: TopPred2, TMAP(Persson and Argos, 1997), DAS, PHD , Prelude.

Ref.	Channel	Protocol	Sequence Identity	Secondary Structure Predictor	Force Field in MD
Ranatunga	K^+ Shaker	I + II	KcsA/ 34%	DAS, TMAP, PHD, TopPred2	CHARMM
Moran	K^+ Shaker	II	KcsA/ 34%	~	~
Capener	K^+ Kir6.2	I+II	KcsA/ 15%	DAS, TMAP, PHD, TopPred2	GROMOS
Thompson	K^+ Kir2.1	II	KcsA/ 15 %	~	~
Rauer	K^+ Kv1.3	II	KcsA/ 50%	~	~
Wrisch	K^+ Kv1.1	II	KcsA/ 50 %	~	~
Rauer	K^+ IKCa	II	KcsA/ 50%	~	~
Giorgetti	HCN	I + II	KcsA/ 18 %	PHD, TopPred2	~
Giorgetti	CNG	I+II	KcsA/ 18 %	PHD, TopPred2	~
Lipkind	Na^+ P loop	I	KcsA	~	cvff-Insight*
Lipkind	Na^+ S5-S6	II	KcsA	~	cvff-Insight*
Sirota	Na^+ L_{III-IV}	I	~	Prelude	GROMOS

* Consistent Valence force field- Insight (Biosym Technologies, Inc., San Diego)

Inward Rectifier K^+ Channels (Kir). These channels conduct K^+ ions from the extracellular side to the cytoplasmatic side more efficiently than in the opposite direction. Kir channels are involved in processes controlling excitability in many mammalian cell types(Doupnik et al., 1995;Reimann and Ashcroft, 1999). They are gated by binding to G-proteins or ATP. Their TM topology is the same as that of KcsA (Fig. 1.2).

Thompson et al.(Thompson et al., 2000) built HM models of the *human* Kir channel Kir2.1 based on KcsA structure. Because of the low-sequence identity between KcsA and Kir2.1 (Tab. 2.1 and Fig. 2.1), the alignment is non-trivial. Two alternative alignments were used, leading to two different models. This structural information, along with sited directed mutagenesis experiments, provided insights on the structural basis of the voltage dependent blockade of Rb^+ (Standen and Stanfield, 1980) and Cs^+ (Hagiwara et al., 1976) ions. Specifically, it was shown that residues S165 and T141 located in the S6 helix and in the P-loop respectively provide the binding sites for both cations.

A model of inward rectifier K^+ channel (Kir6.2) was also built based on the KcsA structure (Capener et al., 2000) (Table 2.1 and Fig. 2.1). MD simulations in a phospholipid bilayer were carried out. Analysis of the interactions of the Kir6.2 channels model with K^+ ions suggest that single-file motion of K^+ ions and water molecules

through the selectivity filter takes place, in a fashion similar to that observed in simulations of KcsA(Shrivastava and Sansom, 2000). The calculations also suggested that filter flexibility plays a role for ion selectivity.

Calcium activated K^+ channels. Calcium-activated potassium channels of intermediate conductance (IKCa) play a prominent role in a large variety of cellular events(Yamamoto et al., 1999;Devor et al., 1996;Singh et al., 2001). These channels are highly selective for K^+ ions. As for voltage-gated channels, the structure is composed of six TMHs and the P-loop (Fig. 1.2). The only region with a relatively high sequence identity with KcsA is the pore region (P-loop and S5/S6).

Using the same approach as for Kv1.3 K^+ channel, Rauer et al.(Rauer et al., 2000) constructed a model of ChTX/ IKCa complex (Table 2.1 and Fig. 2.1). Structural differences between the two complexes were exploited to design three novel charybdotoxin analogs highly selective for IKCa1. These results demonstrate the feasibility of using structure-based strategies to design selective inhibitors for mammalian voltage-gated K^+ channels.

Voltage-Gated Sodium Channels. Experimental structural information at atomic resolution is still lacking. So far, the only channels for which structural models have been constructed are the voltage-gated (VG) sodium channels. These receptors are highly selective towards Na^+ ions, and provide the current source for depolarization and conduction of the action potential in most excitable cells of muscle and nerves(Hodgkin and Huxley, 1952). VG channels are of paramount importance for pharmacological and anesthetic applications as they are targets for most commonly used anti-arrhythmic, anticonvulsant, and local anesthetic drugs.

VG channels share the common transmembrane topological motif of four homologous domains or subunits (I, II, III, IV), each of which is composed of six transmembrane segments arranged in a similar way as voltage-gated K^+ channels(MacKinnon, 1991;Catterall, 1992). The linker between domains III and IV (L_{III-IV}), which includes the IFM triad, is known to be involved in channel inactivation(West et al., 1992;Stuhmer et al., 1989;Catterall, 2000). As for K+ channels, the pore region is formed by an assembly

of four domains each consisting of S5 and S6 TMHs as well as the extracellular loop connecting them (P-loop). The latter forms the selectivity filter(Guy and Conti, 1990).
So far, the structures of the pore and L_{III-IV} have been modeled.
The pore region from *rat brain* (Lipkind and Fozzard, 2000) was constructed in two steps(Tab. 2.1). First, the P-loop was constructed based on the KcsA structure as an assembly of four (Favre et al., 1996;Sun et al., 1997) α-helix-turn-β-strands. This structure was refined by docking with tetrodotoxin (TTX) and saxitoxin (STX). These are naturally occurring high affinity marine toxins that selectively block the VG Na^+ channel. The docking was done, exploiting the experimentally-established high affinity of these toxins for the P-loop region (Favre et al., 1996;Sun et al., 1997;Terlau et al., 1991;Lipkind and Fozzard, 1994;Penzotti et al., 1998;Penzotti et al., 2001;Favre et al., 1996;Sun et al., 1997;Lipkind and Fozzard, 1994;Penzotti et al., 2001;Penzotti et al., 1998). The resulting model was symmetric, with the β-strands of the αβ-structures located toward the central axis and the α-helices located outside the pore. The P-loop structure was then docked onto S5 and S6 helices, which initially were oriented as in KcsA (Doyle et al., 1998). The resulting structure was geometry-optimized, causing small rearrangements of the helices. The structural properties of the selectivity filter turned out to be consistent with Na^+ channel permeation properties(Lipkind and Fozzard, 1994). However, as ion selectivity in Na+ channel is governed by side chain interactions as opposed to main-chain carbonyls in K+ channels(Lipkind and Fozzard, 1994), the predictive power of the model should be taken therefore with some caution at the present stage.
The model of L_{III-IV} was constructed using SSP tools Sirota et al.(Sirota et al., 2002) (Tab. 2.1). It consisted of two α-helical segments (F1476-K1482 and T1491-K1502). The second helix, which had already been detected by NMR(Rohl et al., 1999), formed a hairpin motif with the 1505-1512 loop. MD simulations L_{III-IV} suggested that this structure is stable both in a polar or non-polar implicit solvent.

2.5 Conclusions

Structural bioinformatics techniques have already shown to provide reliable structural predictions of the transmembrane domain of a variety of K^+ channels, and, in some cases, also of other domains of these channels. Models of VG Na^+ channels have also started to

appear. This information in turn may be very useful to designing experiments directed towards the understanding of function and blocking mechanisms of the channels. It also raises the possibility of exploiting structure-based strategies to design specific inhibitors that target pharmacological relevant ionic channels.

3 HCN channels

In this chapter, structural models of the pore region of HCN channels are presented. These models combine homology modeling with constraints derived from electrophysiological experiments. The modeling is based on two major assumptions, justified by electrophysiological observations: the first is that the topology of the inner pore of HCN and K^+ channels is similar (Roncaglia, Mistrik & Torre 2002). We further assume that(Zhou et al., 2001;Jiang et al., 2002a;Jiang et al., 2003) the orientation of the S5 and S6 helices of HCN channels in the closed state is very similar to that of the corresponding helices of the K^+ KcsA and K^+ KirBac1.1 channels (Zhou et al., 2001). Thus, the templates of the closed state are the X-ray structure of these K^+ channels(Kuo et al., 2003). The second assumption is that S6 helix in the open state bends upon gating (Rothberg et al., 2002). For this reason, the template of the open conformation is the X-ray structure of the MthK channel (Jiang et al., 2002b)(Rothberg et al., 2002). The structural models of the closed state are consistent with the available electrophysiological data, whereas those of the open state require the introduction of an appropriate, experimentally derived constraint. Our modeling provides a structural framework for understanding several functional properties of HCN channels: i - the cysteine ring at the inner mouth of the pore is a sensor of the intracellular oxidizing/reducing conditions; ii - the bending of the S6 helix upon gating is significantly smaller than that occurring in MthK channels; iii - the reduced ionic selectivity of HCN channels compared to that of K^+ channels could be caused to the larger flexibility of the inner pore of HCN channels.

3.1 Computational Details

3.1.1 Bioinformatics

Sequence Alignment. The S5/S6 regions in the aminoacid sequence of mHCN2 and spHCN channels were identified on the basis of the prediction of transmembrane helices (TMHs), using the programs: DAS [http://www.biokemi.su.se/~server/DAS (Cserzo et al., 1997)], tmap, [http://www.embl-heidelberg.de/tmap/tmap_info.html (Persson and Argos, 1994;Persson and Argos, 1997)], PHDhtm [http://www.embl-heidelberg.de/predictprotein (Rost et al., 1995;Rost, 2001) see **Appendix A** for a detailed description] and TopPred2 [http://www.biokemi.su.se/~server/toppred2 (von Heijne, 1992)]. The sequences including S5 and S6 turned out to be 368-477 in spHCN and 340-449 in mHCN2. The sequences of spHCN and mHCN2 channels were aligned with residues 26-124 of KcsA, 62-151 of KirBac1.1 and 19-98 of MthK, which correspond to the truncated sequence visible in the X-ray structures (Zhou et al., 2001;Jiang et al., 2002a). The alignment was performed using the program ClustalW [(Thompson et al., 1994) (http://www2.ebi.ac.uk/clustalw) see also **Appendix A**]. In the alignment, the following criteria were used: (i) similarity of the P region (P-helix-loop) (Roncaglia et al., 2002); (ii) Fully conservation of the G-Y-G motif in the filters (Jiang et al., 2002b;Ludwig et al., 1998). (iii) Alignment of S5 (spHCN) residues with TM1 in KcsA (Zhou et al., 2001), KirBac1.1(Kuo et al., 2003) and MthK (Jiang et al., 2002a)). (iv) Alignment of S6 with TM2 in KcsA (Zhou et al., 2001), KirBac1.1 (Kuo et al., 2003) and MthK (Jiang et al., 2002a). Notice that a long gap (> 11 residues) was inserted between the S5 and the pore-helix in the potassium channels in order to align the predicted TMHs of the HCN channels with those of the templates.

3.1.1.1 Building of 3D models.

3D models of the closed spHCN and mHCN2 and open spHCN were produced using Modeller 6v2 (Sanchez and Sali, 1997)(**Appendix A**). The S5, S6 and the P-helix were restrained to adopt a helical conformation, and the non-crystallographic fourfold symmetry present in the KcsA, KirBac1.1 and MthK structures was enforced. The large extracellular loops, present in the two channels, have a very high variability and cannot be compared to existing regions in K^+ channels of known 3D structure. The best-scoring models were selected for each channel. These models provided also the best stereochemistry, as monitored by ProcheckV3.5 (Laskowski et al., 1996) and WhatifCheck(Vriend, 1990) programs. Models of the channel in the closed state were

consistent with all the available experimental data (See Results Section). In contrast, the model of spHCN in the open state was not in agreement with the estimates of distances of opposing C_α along the S6 helix (Shin et al., 2001). Therefore the MthK channel was used as a template but a distance constraint between opposing C_α's of Thr-464. The experimental data on residues in positions 461, 465 and 468 were used for model validation (Tab. 3.1).

	Closed spHCN C_α distance (Å)	Open spHCN C_α distance (Å)	KcsA C_α distance (Å)	MthK C_α distance (Å)
Gly-461	9	13.3	9.8 (A108)	23.7 (E92)
Thr-464	9	10.7	9.5 (A111)	30.7 (L95)
Asn-465	11	15.1	11.5 (T112)	29.3 (E96)
Gln-468	10	15.3	10.9 (V115)	No data

Table 3.1. *Distances of selected (opposite) C_α 's along the S6 helix from spHCN structural models in the open and closed conformations. The distances for opposite Thr-464 in the open conformations were restrained in order to get agreement with experimental data. Distances corresponding to the homologous residues in the templates are also shown.*

3.1.1.2 Number of H-bonds.

In order to estimate the rigidity/flexibility of HCN channels, relative to the KcsA channel, the number of H-bonds was counted. Counting H-bonds in KcsA is straight forward, but some care must be used when homology models are investigated. Therefore several models for the spHCN, mHCN2 and mHCN3 channels were generated using Modeller6v2(Sanchez and Sali, 1997) with different initial conditions, so to explore a variety of possible side chains configurations. The differences in the initial conditions were inserted by changing the initial random number generator and the rms deviation from the initial structure. Once models were generated, an H-bond counting was carried out on the P-helix-loop from Tyr-415 to Pro-436 in spHCN, Tyr-387 to Pro-408 in mHCN2 and Tyr-297 to Pro-318 in mHCN3. All N-O and O-O atoms separated by a distance between 2.195 and 3.350 Å (following Swiss-Pdb Viewer protocols -

http://www.expasy.org/spdbv) were considered as putative H donors and acceptors, respectively. The configuration space of possible models was sampled in the following way: for each channel 10 new models were generated using modeller6v2 and H-bonds were counted and listed. The procedure of model generation was stopped when for a new set of 10 different models no new H-bond distribution appeared. The procedure stopped after the generation of 6 groups of 10 different models.

3.1.1.3 Electrostatic potential along the channel axis.

In order to estimate the electrostatic profile in the absence of a reliable 3D structure, a variety (see later) of 3D models of mHCN1, mHCN2, mHCN3 and spHCN channels were generated and the electrostatic potential profile was computed by solving the coulomb equation along the channels' z-axis. In order to take into account the presence of water molecules (different from 'bulk water') in the cavity and in the inner pore region, a dielectric constant of 4 was used. Because far from the filter, water become 'bulk' – i. e. the dielectric constant increases - , electrostatic profiles only in the neighborhoods of the inner pore will be considered. Used partial charges correspond to the amber force field (Cornell et al., 1995). The different 3D models for each of the HCN channels were generated by considering all possible rotamers of Arg-405 (for mHCN2), of Lys-433 (for spHCN), of Ala–352 (for mHCN1) and Gln-315 (for mHCN3). Rotamers were modeled by using Swiss-Pdb Viewer tools (http://www.expasy.org/spdbv) and the distribution was as follow: 18, 16, 6 and 2 different rotamer configurations for Arg, Lys, Gln and Ala respectively.

3.2 Results

3.2.1 Overall structural determinants of HCN channels

The aminoacid sequence in the pore region (TM1/S5, P-helix-loop and TM2/S6) of the KcsA, KirBac1.1, MthK, spHCN and mHCN2 channels are aligned in Fig.3.1A. The overall sequence identity (SI) between targets and templates is low (≈18% for all the templates). In the P-helix-loop (residues Tyr-415 to Pro-436 in spHCN and Tyr-387 to Pro-408 in mHCN2), however, SI increases significantly, and becomes about 30%

between targets and KcsA channel. The sequence alignment between K^+ and HCN channels requires the inclusion of a gap of more than eleven residues in KcsA, KirBac1.1 and MthK channel, between residues Gly-53 and Ala-54 (KcsA). This region, in HCN channels, corresponds to a long extracellular loop (highlighted in yellow in Fig. 3.1 A) absent in K^+ channels.

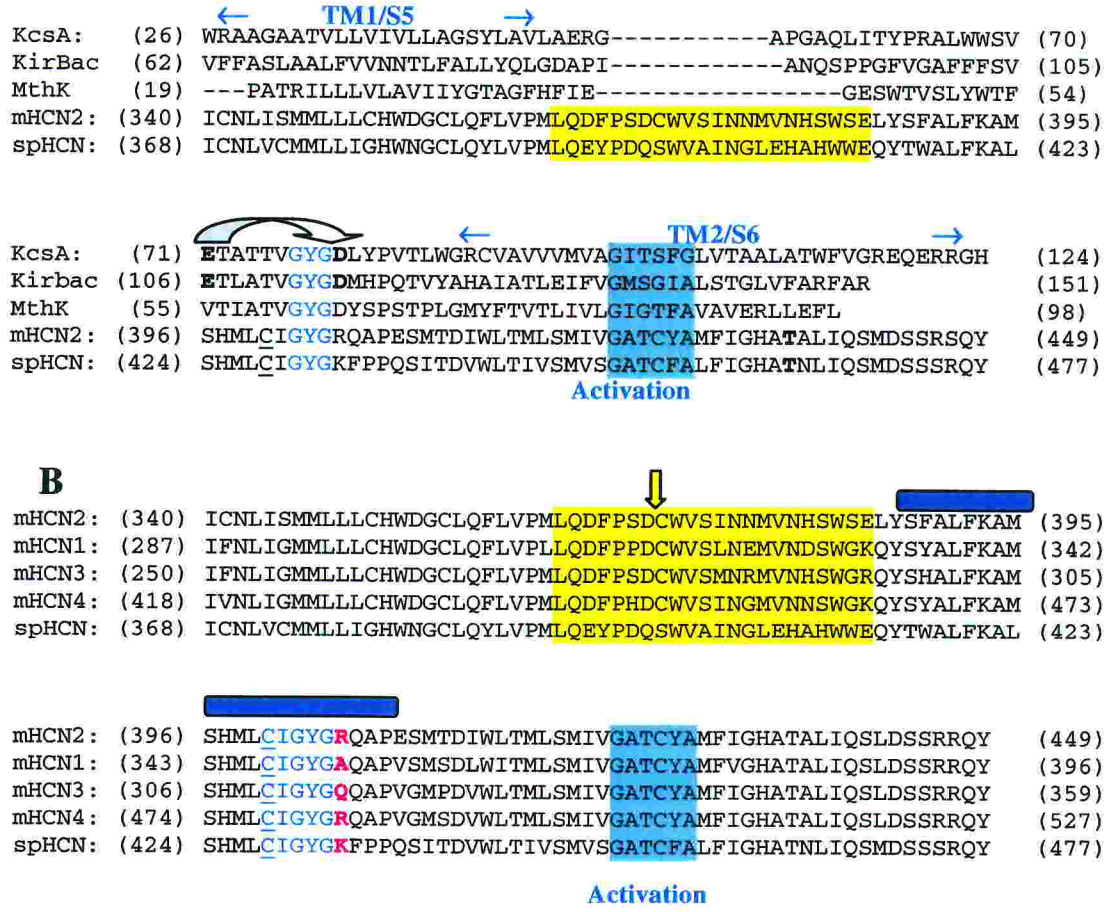


***Fig 3.1**: (A) Sequence alignment of KcsA, KirBac1.1, MthK, mHCN2 and spHCN channels. The filter GYG fully conserved signature (Ludwig et al., 1998) is shown in blue. The activation gate (Jiang et al., 2002b), highlighted in cyan, begins with the fully conserved glycine residue acting as gating hinge and the Gly/Ala five residues downstream (Jiang et al., 2002b). In the KcsA channel, E71 and D80 (light blue) interaction is represented here with an arrow. The long extracellular loop present in spHCN and mHCN2 is highlighted in yellow. Double-headed arrows indicate location of TM1/S5 and TM2/S6. (B): Sequence alignment of pore regions of spHCN, mHCN1, mHCN2, mHCN3 and mHCN4 channels. The conserved residues in the inner pore region are shown in blue, while the residues present in the extracellular end of the inner pore are shown in red. The activation gate in the S6 transmembrane helix and the large extracellular loop were highlighted in cyan and yellow, respectively. The cysteine in the extracellular loop conserved in HCN1-4 channels, responsible for MTSES and MTSEA channel block (Xue and Li, 2002) is marked with a yellow arrow.*

Fig. 3.1B illustrates the sequence alignment in the same region of the mHCN1, mHCN2, mHCN3, mHCN4 and spHCN channels. The overall SI is very high (of the order of 90 %), but some significant differences are evident. In the inner pore region there is a cysteine conserved in all HCN channels (blue underlined cysteine residues in Fig.3.1 A and B). According to the sequence alignment in Fig. 3.1A and the 3D structure of the KcsA channel, these cysteines may form a ring at the intracellular mouth of the channel inner pore. The presence at the inner pore of a ring of cysteines able to oxidize and form disulfide bonds is a remarkable feature and some properties of these cysteines will be studied in spHCN and mHCN2 channels in the following sections. The inner pore region in HCN channels is composed of the residues C-I-G-Y-G-X, (in Fig. 3.1 B, CIGYG underlined residues, and X residues shown in red) where X can be positively charged, polar or neutral residue. X is a lysine in spHCN, an arginine in mHCN2 and mHCN4, but an alanine in mHCN1 and a glutamine in mHCN3. The S5 and S6 transmembrane helices are well conserved among the HCN channels family. In the activation gate region the motif GATCYA (highlighted in cyan in Fig. 3.1 B) is fully conserved in HCN1-4 channels, but in the spHCN channel a tyrosine residue is replaced by a phenylalanine.

The large extracellular loop (highlighted in yellow in Fig. 3.1) has a similar aminoacidic composition among HCN1-4 channels, but there are some differences between those and spHCN channel. For instance in the middle of the loop there is a cysteine (indicated by a yellow arrow in Fig. 3.1 B) conserved in all HCN channels, which is replaced by a serine in spHCN channels. This cysteine mediates the blockage of the I_h current by extracellular MTSEA and MTSES (Xue & Li, 2002), indicating that the extracellular loop could approach the pore at a distance comparable to MTSEA length, i.e. about 6 Å. As Expected from a hydrophilic domain, residues forming the extracellular loops are polar and have a high probability for H-bond formation.

Models were constructed using homology modeling (**Appendix A**) based on the sequence alignment shown in Fig. 3.1A (models are available at: http://www.sissa.it/~giorget/pdb_files). As precise atomic details cannot be derived from comparative models, more than 180 models were generated so to obtain an exhaustive sampling of all possible conformations of the channel. Therefore, our conclusions are based on physical quantities obtained by sampling a large number of putative alternatives, minimizing bias intrinsic to specific models. As expected, the 3D

architecture of the K^+ channels is maintained. The long loop regions are not shown because of the high variability and uncertainty of their location after using our modeling approach. The stereochemistry and quality of the final models, investigated with ProcheckV3.5 and WhatCheck programs, validated the proposed models (Tab. 3.2).

Table 3.2. Results of WathCheck and Procheck analysis of the final models. As can be appreciated all the values are well inside the limits of a correctly folded structure

	spHCN_Closed	*spHCN_Open*	*mHCN2*
Procheck-Average	-0.3000	-0.0300	-0.0800
Z-score Bond angles	1.525	1.311	1.318
Z-score bond lengths	0,957	0,945	0,947
Bumps	No	No	No
Backbone Oxygen Eval.	Ok	Ok	Ok
Improper dihedral RMS Z-score	1,154	1.023	0,836
Pro Puckering amplitude	Ok	Ok	Ok
Side chain planarity	Ok	Ok	Ok
Atoms connected to aromatic rings	Ok	Ok	Ok
Chi-1/Chi-2 angle correlation Z-score	-2,012	-1,411	-1,81
Ramachandran Z-score	2,428	0,617	2,086

3.2.2 The inner pore

Structural properties. Homology models of the inner pore region of mHCN2 and spHCN channels are illustrated in Fig. 3.2A and B respectively. The C_α of residues (shown as balls) from His-397 to Arg-405 for the mHCN2 and from His-425 to Lys-433 for the spHCN channels are shown. A remarkable feature -almost a marker - of all known HCN channels is the presence of a cysteine residue two residues upstream of the filter motif GYG (Cys-428 for spHCN and Cys-400 for mHCN2 - see blue underlined cysteine residues in Fig 3.1A and B). Given the tetrameric structure of these channels, these

cysteines form an intracellular ring (Fig. 3.2 A and B). In mHCN2 and spHCN channels at the outer mouth of the pore there is an arginine and a lysine, respectively, forming an extracellular ring of positive charges. The intracellular ring of cysteines and the extracellular ring of positive residues are key features of the inner pore of mHCN2 and spHCN channels.

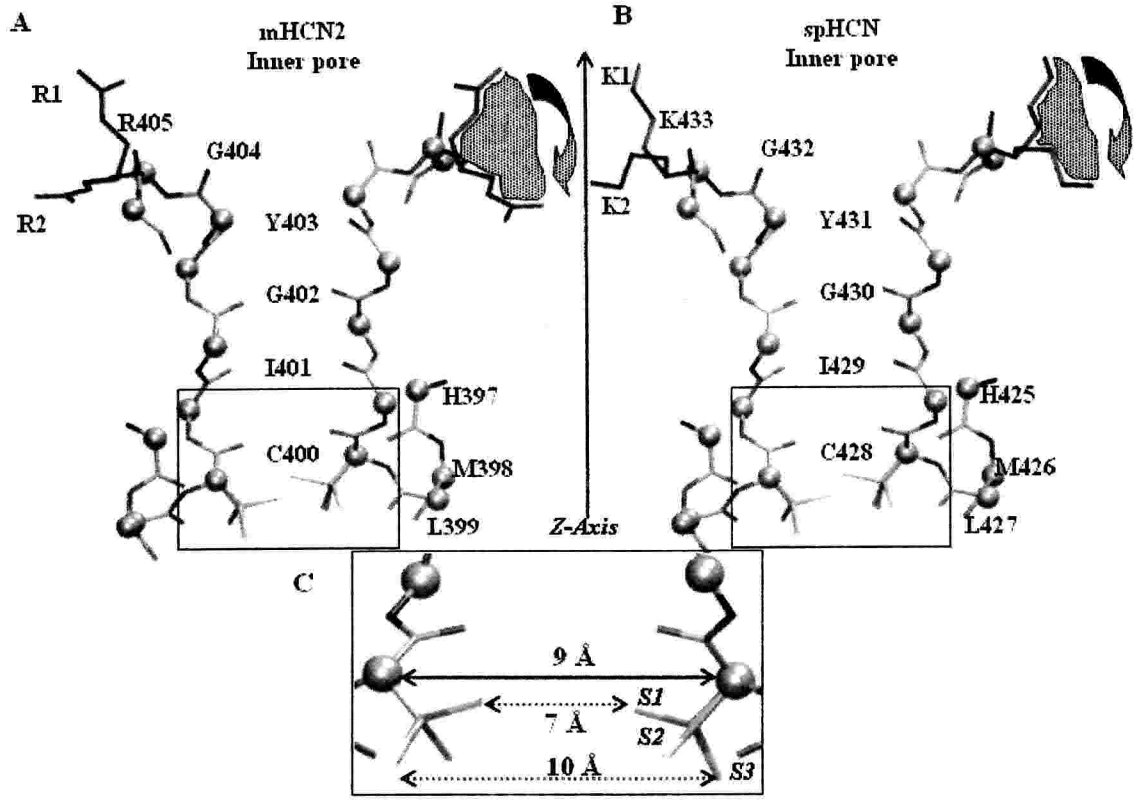


Fig 3.2: *(A) and (B) homology model of the inner pore region of mHCN2 and spHCN channels respectively. The model for the mHCN2 channel shows residues from His-397 to Arg-405, while the model for spHCN from His-425 to Lys-433. Side chains are not shown, but the rotamer space of Arg-405 and Lys-433 is shown as a shaded area. R1, R2 and K1, K2 indicates the extreme rotamers for Arg-405 and Lys-433 respectively. (C): Cysteine rotamer configurations for Cys-400 and Cys-428: S1, S2 and S3 indicate three different rotamer configurations. The homology model of the intracellular vestibule for mHCN2 and spHCN channels is the same. Solid line indicates the distance between C_α of opposing cysteines and the dotted lines indicate examples of distances between S atoms of opposing cysteines. Just two opposite subunits are shown.*

Given the limitation of homology modeling it is difficult to determine the precise location of side chains and therefore we have investigated the configuration space of all possible rotamers of Arg-405 and Lys-433. These rotamers span wide angles centered on the corresponding C_α indicated by the shaded arrows in Fig 3.2 A and B respectively. The side chain of Arg-405 can point in the two extreme directions indicated by R1 and R2. Similarly the side chain of Lys-433 can point almost upward (see K1) and downward (see K2). These side chains do not seem to be anchored as the side chain of Asp-80 in the homologous position in the KcsA channel.

We now compare our structural models of the pore with electrophysiological experiments of Roncaglia and coworkers (Roncaglia et al., 2002) and Giorgetti et al.(submitted). These experiments are summarized in the next Section.

3.2.2.1 Comparison with experimental data

We now compare our models with the structural properties of the cysteine ring inferred by the electrophysiological experiments done by Giorgetti et al (submitted) and in Roncaglia et al 2002. An analysis of Cd^{2+} and Cys containing protein structures performed by Krovetz et al. (Krovetz et al., 1997), suggests that the distance between the C_α for a pair of cysteines binding to Cd^{2+} to be ≈4- 9 Å (Ermler et al., 1998)(Maroney, 1999) and to be ≈ 4-6.5 Å for pairs of cysteines forming S-S bridges (Srinivasan et al., 1989). Because the approximate diameter of Cd^{2+} is 1.8 Å and considering thermal fluctuations (Glusker, 1991) the maximum distance d(C_α@Cys-C_α@Cys) is expected to be ≤ 11 Å (Johnson and Zagotta, 2001;Careaga and Falke, 1992;Krovetz et al., 1997;Ermler et al., 1998;Maroney, 1999). In our models, d(C_α@Cys-C_α@Cys) is ≈9 Å and ≈6.5 Å for opposite and adjacent cysteines, respectively, showing fully consistency with the experiments (Fig. 3.2 A). We have also investigated the configuration space of all possible rotamers of Cys-400 and Cys-428, three limiting rotamers were illustrated in Fig 3.2 C. The S atom can point in several different directions: for some rotamers the distance of opposite S is 7 Å and for other rotamers 10 Å. In our models the expected distance for a Cd^{2+}-S interaction is approximately 2.5 Å. This distance is in agreement with that determined by X-ray diffraction between S atoms of cysteines forming the binding pocket for divalent cations (Furey et al 1986; Changela et al. 2003).

3.2.2.2 Flexibility of the inner pore region

The 3D structure of the P-helix-loop of the proposed models of HCN channels is - rather obviously - very similar to that of the KcsA K^+ channel. However, the resulting molecular interactions between neighboring domains are different: of particular interest is the absence in the spHCN channels of the anchoring interactions present in the KcsA structure, presumably causing the stability of the KcsA inner pore (Zhou et al. 2001). In the KcsA channel, Glu-71 in the P-helix forms a strong hydrogen bond with Asp-80 in the outer pore (Zhou et al., 2001) (Fig 3.3A), resulting in a powerful stabilization of the channel inner pore. In HCN channels this interaction is lost due to the presence of a serine residue in the corresponding position of Glu-71 (Fig.3.3B): the serine side chain is shorter than that of glutamate and therefore there is a significant reduction of strong H-bonds with residues in the outer mouth of the channel. As a consequence, the inner pore of spHCN and mHCN2 channels is expected to be less rigid than that of K^+ channels. The H-bond present in KcsA channels, is not conserved in all K^+ channels: in the Inward Rectifier K channels (Kir) a salt bridge between residues E138 and R148 is formed. These residues correspond in the to KcsA channel to V70 and L81, respectively. When this interaction is disrupted by appropriate point mutations, the mutant channel is not selective for K^+ ions (Yang et al. 1997).

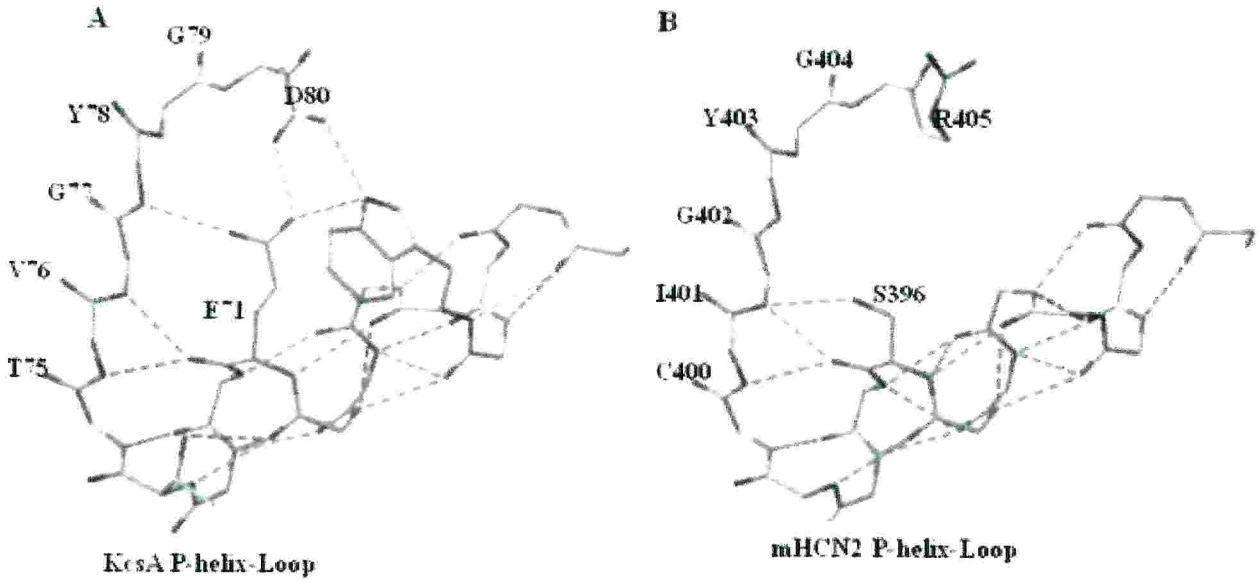


Fig 3.3: *H-bond distribution in KcsA and mHCN2 channels in the P-helix-loop region (inner pore and P-helix). (A): 3D crystallographic structure of 1 subunit of the KcsA channel (Zhou et al. 2001); (B) 3D model of one configuration of the mHCN2 channel. In (A) and (B) only side chains capable of H-bonds formation are shown. H-bonds are indicated by dotted lines. The stabilizing H-bonds network formed by Glu-71 and Asp-80 in the KcsA channel is absent in HCN channels. The average number of H-bonds was 21.3±1.5, 21.1±1.4 and 21.1±1.6 in the P-helix-loop region for spHCN, mHCN2 and mHCN3 channels respectively. The configurations were obtained as described in the Methods section.*

In order to obtain a more quantitative validation for this hypothesis we have compared the H-bonds pattern in the P-helix-loop region in the KcsA and HCN channels. This is expected to affect the filter stability and rigidity. Of course, the same criterion for defining H-bond has been used (see Methods section). In the KcsA channel X-ray structure, 26 H-bonds are present in the P-helix-loop (from Tyr-62 to Pro-83).

We have generated 60 models for each of the pores of spHCN, mHCN2 and mHCN3 channels so as to include a wide range of side-chain configurations of the P-helix-loop. Residues from Tyr-415 to Pro-436 for spHCN and from Tyr-387 to Pro-408 in mHCN2 (indicated with a blue thick line in Fig. 3.1 B) were considered. 21.3±1.5, 21.1±1.4 and 21.1±1.6 H-bonds are in the P-helix-loop region for spHCN, mHCN2 and mHCN3 channels respectively. Thus, the number of H-bonds is consistently lower than in the

KcsA channel. The variation of H-bonds depends primarily on the orientation of a serine side-chain (Ser-396 in mHCN2, Ser-306 in mHCN3 and Ser-424 in spHCN). As shown in figure 3.3 A and B the KcsA channel has more H-bonds (dotted lines) in the upper part of the inner pore. These considerations suggest that the inner pore of HCN channels is less rigid than that of K^+ channels.

3.2.2.3 The electrostatic potential profile within the pore

As shown in Fig.3.1B the inner pore of HCN channels is conserved (underlined residues) with the exception (bold residues) of the Lys-433 in spHCN, which becomes an arginine in mHCN2 and mHCN4, an alanine in mHCN1 and a glutamine in mHCN3. The outer ring of positively charged residues in spHCN, mHCN2 and mHCN4 is expected to affect electrostatics in the pore and in the outer vestibule.

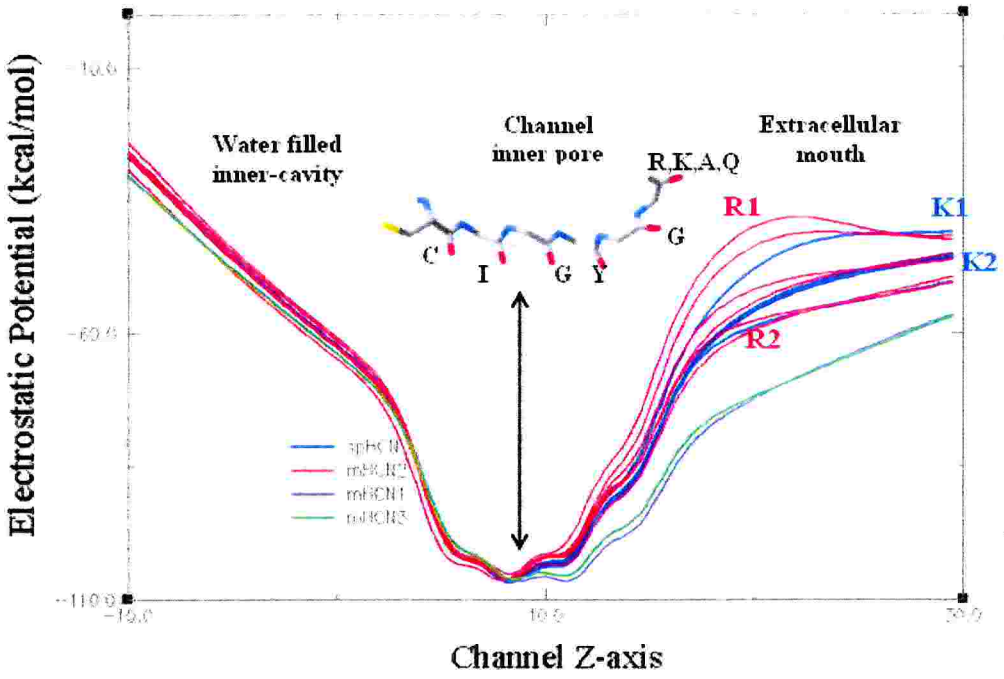


Fig 3.4: *Electrostatic potential profile along the channel z-axis for different rotamer configurations of mHCN2 (red), mHCN1 (violet), mHCN3 (green) and spHCN (blue) channels. The plot indicates the position of the different carbonyl oxygen from inner pore. The electrostatic profiles corresponding to the limiting rotamers -shown in Fig. 2 - are labeled as R1, R2 and K1, K2.*

In order to estimate the electrostatic potential profile in the absence of a reliable 3D structure, a variety of 3D models of mHCN1, mHCN2, mHCN3 and spHCN channels were generated and for each of them the electrostatic profile along the pore axis was computed. All models had the same backbone, illustrated in Fig. 3.2, but differed in the direction of side chains of the outer ring of arginine, lysine, alanine and glutamine. In Fig. 3.2 A and B the allowed rotamers for Arg-405 and for Lys-433 are shown as shaded regions. The two extreme rotamers are marked *R1*, *R2* and *K1*, *K2* respectively. As discussed in the Methods section for each HCN channel a family of electrostatic potential profiles was obtained.

Fig. 3.4 illustrates the electrostatic potential profile along the pore axis for different rotamers of the mHCN1, mHCN2, mHCN3 and spHCN channels. For HCN channels with an alanine or a glutamine the electrostatic potential profiles for the different rotamers are very similar, as expected from residues not bearing a net charge. As a consequence, the electrostatic potential profile for mHCN1 and mHCN3 channels is more negative near the extracellular vestibule than that of HCN channels with a positively charged ring, such as spHCN and mHCN2. The electrostatic potential profile for the mHCN2 and spHCN channels varies from about -60 to -40 Kcal/mol, according to the orientation of the side chain of the arginine or lysine. When side chains point toward the lipid phase (configurations R2 and K2 of Fig 2 and corresponding potential profile in Fig 3.4) the electrostatic potential profile is more negative than when side chains point upward and are almost parallel to the pore axis (configurations R1 and K1).

3.2.3 Closed and open states

The spHCN channel is blocked by the compound ZD7288 (Shin et al., 2001) only in the open state. Shin and coworkers (Shin et al., 2001) have shown that three residues in S6 region - Phe-456, Leu-458, and Ile-460 - mediate the irreversible blockage of ZD7288. These residues are presumably located inside the cavity of the channel pore and the channel needs to open so that ZD7288 can bind to these residues. Upon channel closure the blocker is trapped irreversibly inside the channel cavity. As a consequence an intracellular gate (Rothberg et al., 2002;Shin et al., 2001) involving conformational rearrangements of the S6 transmembrane helix as in K^+ channels (Jiang et al., 2002a) has

also been proposed for HCN channels. In order to verify the existence of this gate and to identify the involved residues, cysteine scanning mutagenesis of the S6 transmembrane helix of spHCN channels was used (Rothberg et al., 2002). When Thr-464 was mutated into a cysteine, small amounts of intracellular Cd^{2+} produced an irreversible blockage of the mutant channel, likely to be caused by the coordination of three or more S atoms with a Cd^{2+} ion. As a consequence the distance between opposing C_α of these exogenous cysteines is expected to be about or less than 11 Å. When Glu-461 and Asn-465 were mutated into a cysteine, Cd^{2+} ions powerfully blocked mutant channels G461C and N465C in the open state. Cd^{2+} blockage of mutant channels G461C, T464C, N465C and Q468C in the open state are not in agreement with a model of the spHCN channel using the MthK as the template (see Fig. 3.5 A and Table 3.1). Therefore, the extent of bending of the S6 helix, upon gating, in K^+ and HCN channels is different. As discussed in the Methods section, in order to obtain a realistic model of the spHCN channel in the open state, the MthK channel was used as a template but with the additional constraint that the distance between opposite Thr-464 to be of ≈11 Å.

Fig. 3.5 A reproduces the templates used for our homology modeling for the open (MthK channel in red) and closed (KcsA channel in blue) channel. Fig. 3.5 B illustrates the final models of the spHCN channels in the open (red) and closed state (blue). The bending of the S6 transmembrane helix associated to gating in K^+ channels is ≈30° whereas in our model of spHCN channels, this angle decreases to about 12°. Rothberg et al (2003) have shown that the mutant channel Q468C of the spHCN channel is blocked in the closed state very powerfully by Cd^{2+} ions with an affinity of 72 μM suggesting the simultaneous binding of four Cd^{2+} ions. The proposed model of the pore of the spHCN channel in the closed state is compatible with the proposed Cd^{2+} blockage. In our model the distance between opposite C_α of exogenous cysteines in the mutant channel is ≈14 Å so that it may be possible to accommodate among neighboring S a Cd^{2+} ion and obtain the proposed configuration S-Cd-S-Cd-S-Cd-S-Cd (Furey et al 1986; Changela et al. 2003).

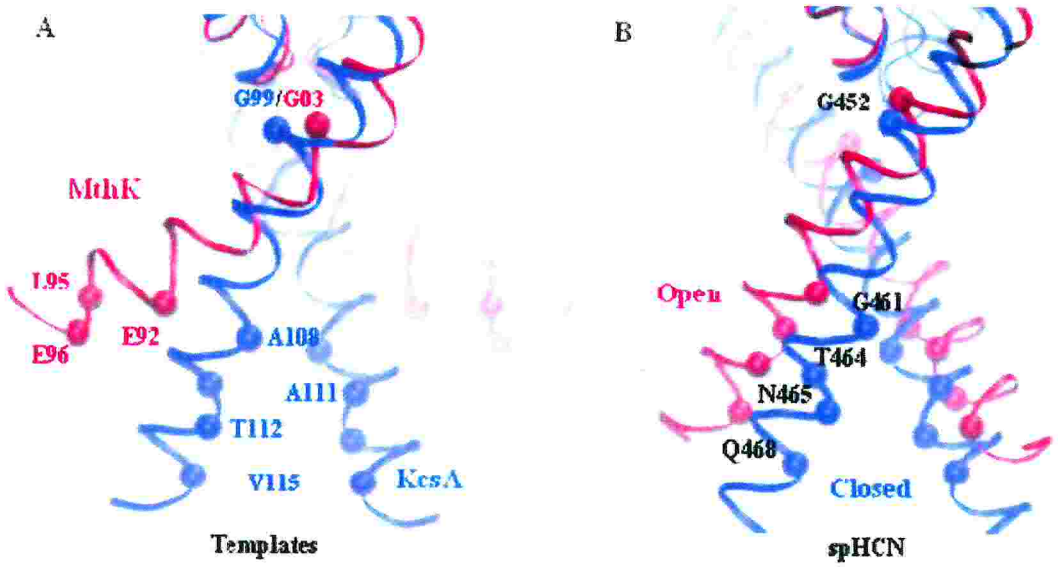


Fig 3.5: *The open (red) and close (blue) channel conformations. (A) Templates used for homology modeling for the open (MthK channel in red) and closed (KcsA channel in blue) channel. (B) Final models of the spHCN channels in the open (red) and closed state (blue). The bending of the S6 transmembrane helix associated to gating in K^+ channels is $\approx 30^o$ whereas in our model of spHCN channels the bending is about $\approx 12^o$. Only two subunits are shown. Selected residues are shown as spheres. Distances between opposite C_α–C_α atoms obtained from the modeling are compared with estimates from experimental data in Table 3.1.*

Very recently, Zagotta and coworkers (Zagotta et al., 2003) have solved the crystal structure of the cytoplasmic domain of mHCN2 channel. As the solved structure was obtained in the presence of cAMP bound, it represents the structure of the HCN channel in the open configuration. The solved structure encompasses residues very near the C-terminal end of the S6 transmembrane helix, such as Asp-443, thus providing a direct experimental validation of the proposed model of the open channel configuration. In the crystallographic 3D structure of the mHCN2 channel, the distance from Asp-443 C_α in the S6 helix to the channel axis is ≈ 16 Å, a value intermediate (Zagotta et al., 2003) between that obtained from the MthK open structure and that of the KcsA closed structure. In our model of the open configuration of the spHCN channel, S6 helix was constrained to adopt an intermediate conformation between KcsA and MthK crystal TM2 structures, in agreement with Zagotta and coworkers' (Zagotta et al., 2003) result.

3.3 Discussion

The present chapter provides molecular models of the pore of spHCN and mHCN2 channels. These models were obtained by homology modeling using the crystal structure of the KcsA channel (Zhou et al., 2001) and KirBac1.1 (Kuo et al., 2003) as templates for the closed conformation and the MhtK channel (Jiang et al., 2002a) for the open conformation. In order to obtain models consistent with electrophysiological experiments performed previously (Roncaglia et al., 2002) (Rothberg et al., 2002;Shin et al., 2001), experimental data were translated into structural constraints of the homology modeling. The validity of these assumptions was discussed in the Introduction and in the Computational Details section. Let us now discuss possible insights on structure/function relationships in HCN channels based on the proposed models.

3.3.1 The inner pore in HCN and KcsA channels.

The great majority of HCN and K^+ channels share the GYG motif in the pore, as shown in Fig.3.1A. The ionic selectivity of these channels, however, is different; K^+ channels are primarily permeable for K^+ and Rb^+ ions and the permeability ratio P_{Na}/P_K is below 0.1. On the contrary a distinct feature of all HCN channels is a significant permeability also to Na^+ ions, with P_{Na}/P_K around 0.3. A recent theory of selectivity of ionic channels (Laio & Torre 1999) suggests the inner pore radius and its flexibility are the primary determinants of ionic selectivity. In this view K^+ channels have a radius of about 1.5 Å, large enough for the permeation of alternating, single-file of K^+ ions and water molecules, as confirmed by high resolution X-ray crystallography (Zhou et al. 2001). If the inner pore is rigid, Na^+ ions with three water molecules attached will not be able to pass and the channel will be highly selective for K^+ ions. Occasionally, if the inner pore becomes less rigid Na^+ ions coordinated to three water molecules can permeate through it, accounting for increased Na^+ permeability. As a consequence, the less pronounced selectivity of HCN channels compared to K^+ channels could be ascribed to a more flexible inner pore in HCN channels. As shown in Fig. 3.3, HCN channels lack a major structural anchor present in the KcsA channel, where Glu-71 forms a strong hydrogen bond with Asp-80. This H-bond anchors rather tightly the inner pore to the P-helix.

Indeed, in the high-resolution structure of the KcsA channel (Zhou et al., 2001) there are approximately 26 H-bonds in the P-helix and the inner pore region from Tyr-62 to Pro-83. In the models of the pore region of spHCN, mHCN2 and mHCN3 channels, shown in Fig. 3.3, there are on average approximately 21.2 H-bonds, significantly less than in the KcsA channel. This structural difference between HCN and K^+ channels maybe the reason for their different ionic selectivity.

3.3.2 The Cysteine Ring.

All HCN channels bear a cysteine in the inner pore region, as shown in Fig. 3.1B (blue underlined Cys). Cd^{2+} ions from the cytoplasmic side of the membrane block spHCN and mHCN2 channels. These results show that these cysteines are located at the intracellular side of the inner pore and form the cytoplasmic vestibule (Roncaglia et al 2002).

Under normal conditions HCN channels are open and therefore these cysteines do not form S-S bonds. The distance between opposite C_αs of these cysteines is $\approx$9 Å, suggesting that S-S bridges can be readily formed in the presence of an oxidizing medium. As a consequence, the inner pore of HCN channels is likely to be a sensor of the intracellular oxidizing/reducing conditions. Given the limitation of homology modeling it is very difficult and almost impossible to provide an explanation why Cd^{2+} ions bind with different affinities the ring of the native cysteines in position 428 and of exogenous cysteines introduced in position 464 or 468

3.3.3 The pore electrostatics

As shown in Fig. 3.4 the electrostatic potential profile obtained from our modeling is slightly different for the various HCN channels. In agreement with the K^+ pore structure the carbonyl groups of the signature GYG point towards the pore axis and provide an attracting binding site for cations. The computed electrostatic potential profile in the outer part of the pore is more negative for mHCN1 and mHCN3 channels. Given the positive charge of Arg and Lys present in the outer mouth of mHCN2, mHCN4 and spHCN channels the electrostatic potential profile for all considered rotamers was consistently less negative than for mHCN1 and mHCN3. Therefore differences of single

channel properties conductance among HCN channels could be expected. Given the small single channel conductance of HCN channels (Roncaglia et al., 2003; Shin et al., 2001; Rothberg et al., 2002) it may be difficult to experimentally establish these differences.

3.3.4 Closed and open states.

The transition between the open and closed state in bacterial K^+ channels is associated to a significant bending of the S6 transmembrane helix around a hinge located around a conserved glycine (see Fig. 3.1A see also chapter 1). This bending in bacterial K^+ channels is associated to an angular rotation of about 30° around the hinge (Jiang et al., 2002a and b). Gating in HCN channels seems to occur by a similar mechanisms (Rothberg et al., 2002), but with a smaller rotation of about 12°. Gating in the large superfamily of voltage gated ionic channels is likely to be caused by a movement of the S6 transmembrane helix, but its extent and properties are likely to depend on the specific ionic channel.

4 CNG Channels

In this chapter we present structural models of the CNG channel tetrameric pore structure from bovine rod in both open and closed states, as obtained by combining homology modeling based techniques, experimentally derived spatial constrains and structural patterns present in the PDB database. Gating is initiated by a clockwise rotation of the N-terminal region of the C-linker, which is then transmitted trough S6 transmembrane helices to the P-helix, and in turn from this to the pore lumen, which opens up by about 3 Å (from 2 to 5 Å) thus allowing for ion permeation. The approach here presented is expected to present a general methodology for model ion channels and their gating when reliable templates are available and an extensive electrophysiological analysis has been performed.

4.1 Computational Details

4.1.1 Bioinformatics

Modeling the S6 domains of CNG channels.

Molecular modeling of the 3D structure of the S6 domains of CNGA1 channels in the open and closed configuration is based on comparative modeling techniques (**Appendix A**). Structural models of the S6 domain of CNGA1 channels were built following 5 steps: 1 – Template identification; 2 – Sequence alignment; 3 - Molecular modeling; 4 - Insertion of experimental constrains; 5 - Model structural analysis. These steps were as follows:

Template identification

The 3D structure of three channels were considered as possible templates for the S5 and S6 domain of the CNGA1 channel: KcsA K^+ channel from *Streptommyces lividans* (Doyle et al., 1998;Zhou et al., 2001), the K^+ MthK channel from *M. Thermautotrophicum* in the open state (Jiang et al., 2002a) as a template for the open configuration and the crystal structure of the mHCN2 channel (Zagotta et al., 2003) as a template for the final portion of the S6 domain in the open configuration.

Sequence Alignment

The S5/S6 domains in the aminoacid sequence of CNGA1 channels were identified on the basis of the prediction of transmembrane helices, using the programs: DAS [http://www.biokemi.su.se/~server/DAS (Cserzo et al., 1997)], http:// www.embl-heidelberg.de/tmap/tmap_info.html (Persson and Argos, 1994;Persson and Argos, 1997)], PHDhtm [http://www.embl-heidelberg.de/predictprotein (Rost et al., 1995;Rost, 2001) and **Appendix A**] and TopPred2 [http://www.biokemi.su.se/~server/toppred2 (von Heijne, 1992)]. The aminoacid sequence of the P helix and the S6 domain of CNGA1 channels was aligned with residues from 39 to 124 of KcsA and with residues from 67 to 98 of the MthK channel. The N-terminal C-linker region of CNGA1 channels (from residues 402 to 422) was aligned with residues from 443 to 460 of the HCN2's C-linker(Flynn & Zagotta 2003). The alignment was performed using the program ClustalW [(Thompson et al., 1994) (http://www2.ebi.ac.uk/clustalw)].

Molecular modelling

3D- structural models of the closed and open CNG filter region were produced using Modeller 6v2 (Sanchez and Sali, 1997) (**Appendix A**). This region was restrained to satisfy several distance constraints between filter residues (Tab. 4.1), and the non-crystallographic fourfold symmetry present in the KcsA structure was enforced. The three best-scoring models were selected for each channel configuration. These models provided also the best stereochemistry, as monitored by Procheck (Laskowski et al.,

1996) and WhatCheck (Vriend, 1990) programs. Furthermore, they were structurally similar to other top-scoring models.

Comparison with experimental data and Model generation

Models were constructed by imposing several restraints either based on direct experimental data or assumed from the data: these restraints are imposed in the generation of the models as single gaussian harmonic potentials, in which the center and the standard deviations are defined. The helices are restrained by Phi/Psi binormal restraints, N-O hydrogen bonds, Ca-Ca distances for i-j E (2-9), Ca-O distances for i-j E (2-9), and O-O distances for i-j E (2-6) (Modeller6v2 (Sanchez and Sali, 1997)).The spatial conformation of loops connecting the S5, S6 and the P helix were modeled by an extensive scanning of a loop database, and an afterward minimization of bond lengths, bond angles and dihedral angles. These operations were carried out by using the Swiss-PdbViewer utilities (http://www.expasy.org/spdbv) (Guex and Peitsch, 1997). Experimental data on Cd^{2+} blockage were used as experimental constraints on the C_α of opposing and/or neighboring exogenous cysteines. A reversible Cd^{2+} blockage is usually caused by its coordination to two cysteines and an irreversible blockage to the coordination to three or more cysteines (Rothberg et al., 2002). More specifically we will assume that in an irreversible Cd^{2+} the distance between the C_α of 'opposite' exogenous cysteines to be slightly less than 11 Å allowing a Cd^{2+} ion bind to three or four cysteines (Tab. 4.2). We will assume also that a reversible blockage report a distance slightly less than 11 Å between the C_α of 'adjacent' exogenous cysteines, allowing one Cd^{2+} ion to bind only to two out of four cysteines. The model of the S6 domains of CNGA1 channels in the closed state was obtained by using as template the 3D structure of the KcsA channel and constraining the distance between C_α of opposite residues, as specified in Table 4.1. The same for the open configuration, with the additional restraint that S6 helix rotates in the anticlockwise sense around the helix axis with respect to the closed conformation to permit residues from 402 to 419 to connect to the 3D structure of the C-linker of the mHCN2 channel. This rotation was about 45°

for residues between 400 and 402 and progressively decreased to approximately 30° for residues in the upper portion of S6 between 379 and 381.

Model structural analysis

The overall average G factor for the protein by PROCHECK was -0.49 and -0.50 for the closed and open finale structures, respectively. Both of the models show a very good Ramachandran plot that is about 92 % residues in the most favored regions, and no residues in the forbidden regions. WhatCheck results, show values in the normal variability ranges, as well as side chain planarity of aromatic aminoacids. The structural average quality control value is within normal ranges (Average value= -0.910). No residue with unusual rotamers was found.

4.2 Results and Discussion

CNG channels transmembrane domain is a tetramer, in which each subunit consists of six transmembrane helices (S1-S6) and a pore helix (P-helix), with the same topology of voltage-gated K^+ channels (Becchetti et al., 1999;Sesti et al., 1995) (Fig. 1 A). The pore, unselective for Na^+ and K^+, is believed to gate via a conformational change of S6 transmembrane helix (TMH) 'induced' by a coupling with the cytoplasmic domain which is transduced to the pore via the four P-helices (Fig. 1 A) (Johnson and Zagotta, 2001;Matulef et al., 1999).

Here we provide a molecular basis to this proposal by constructing models of the transmembrane region of channel from bovine rod, which includes S6, P-helix and the pore wall, along with the C-linker N-term, for which there exist a great amount of experimental data. Models of P-helix-loop and S6 are based on the KcsA X-ray structure (Jiang et al., 2002b;Zhou et al., 2001), whose topology has been suggested to be similar to that of CNG channels (Becchetti et al., 1999; Becchetti and Roncaglia, 2000; Roncaglia and Becchetti, 2001; Liu and Siegelbaum, 2000). The C-linker domain was modeled using the C-linker of the HCN channel in the opened state from *mouse*, for

which the X-ray structure has been recently solved (Zagotta et al., 2003). This template shares a high sequence identity (>30%) with CNG channels in this particular region.

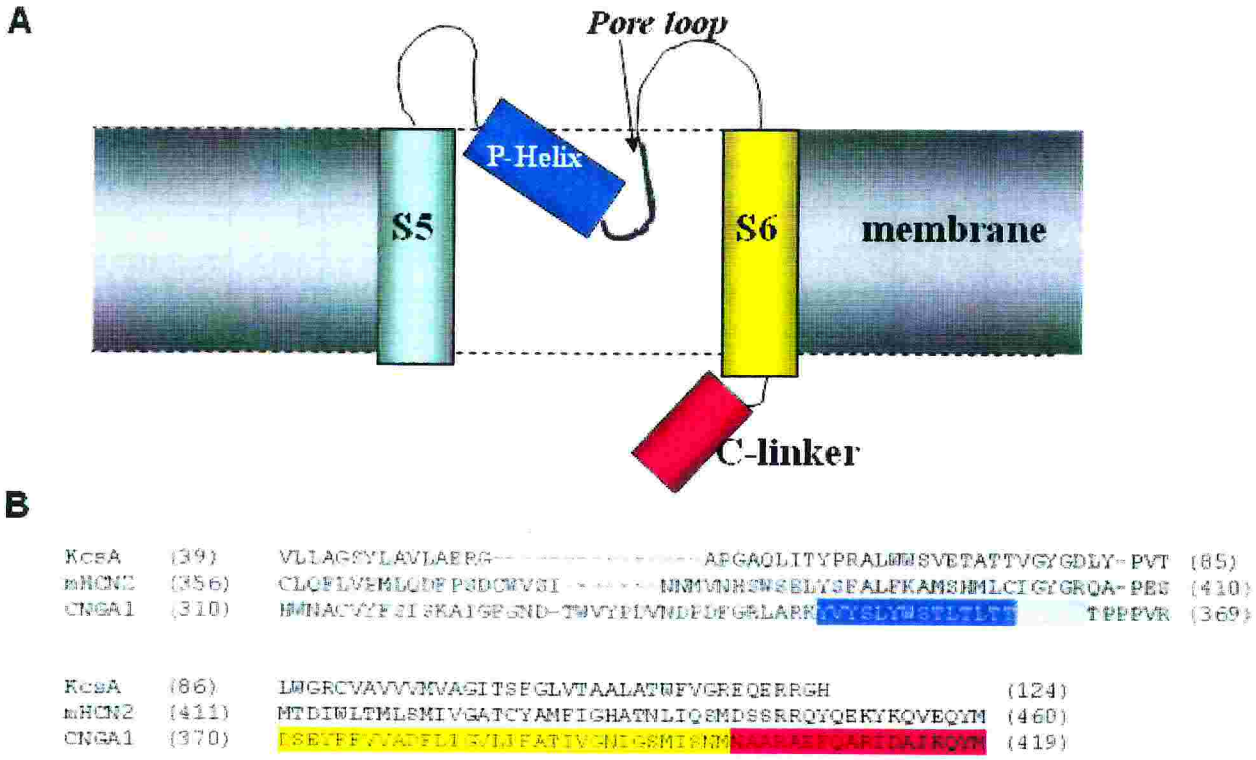


Fig. 4.1 *(A) Cartoon representing one of the four identical subunits of the the transmembrane domain and the N-Term@C-linker. A rearrangement of the C-linker, induced by a conformational change of the nucleotide binding domain (not shown) induces in turn a rearrangement of S6 helix, which in turns affects the conformation of the P Helix and that, in turn, of the pore loop (wall).* ***(B)*** *Sequence alignment of the P-helix-loop, S6 and N-Term@C-linker of the CNGA1 channel from bovine rod used for building up the structure by homology modeling. The templates are: KcsA K^+ channel from Streptommyces lividans (Doyle et al., 1998;Zhou et al., 2001) for P-helix-loop and TM2/S6, and HCN from mouse for the C-linker (Zagotta et al., 2003).* ***Color coding****: Gray: Pore-loop. Blue: P-helix. Yellow: S6 transmembrane helix. Red: C-linker N-term.*

Table 4.1. *Spatial constraints involving P-helix-loop, S6 and N-term@C-linker of CNGA1 channels from bovine rod, based on experimental data on cysteine mutants (see* ***section 1.5.2****). Distance restraints always refer to opposite Cα in the tetramer, unless specified. Accessibilities patterns were used to constraint the P-helix orientation in agreement with (Liu and Siegelbaum, 2000). References corresponds to: [1] (Liu and Siegelbaum, 2000) [2]Nair et al.,2004 (in preparation) [3] (Becchetti et al., 1999) [4]Mazzolini et al., 2004 (in preparation)*

Mutant	Cd(II)		MTS		CuP		Ref	Restraint
	Open	Closed	Open	Closed	Open	Closed		
Val348	-	-	B	B	-	-	1 & 2	Accessible extracellular
Ser350	-	-	N. E.	N.E.	-	-	2	-
Leu351	-	-	B	B	-	-	1 & 3	Accessible extracellular
Thr355	-	-	I	B & I	-	-	1 & 3	"
Leu356	-	-	-	-	-	-	-	Inactivation
Leu358	-	-	N.E.	B	-	-	1	Accessible extracellular
Thr360	B	N.E.	P MTSET MTSES	P MTSET MTSES	P	P	2	D(Cα- Cα) ≈ 11 Å (Open) D(Cα- Cα) > 14 Å (Closed) Accessible Intracellular
Glu363	-	-	B	B	-	-	1	Accessible Extracellular
Thr364	-	-	B	B	-	-	1	"
Pro365	-	-	B	B	-	-	1	"
Phe375	N.E.	N.E.	-	-	-	-	4	D(Cα Cα) > 14 Å
Val376	"	"	-	-	-	-	4	"
Val377	"	"	-	-	-	-	"	"
Ala378	"	"	-	-	-	-	"	"
Asp379	"	"	-	-	-	-	"	"
Phe380	P	-	-	-	Locked	Locked	4	D(F380Cα- C314Cα) < 8 Å Rotation of 30° anticlockwise in Open state.
F380-L356	N.E.	N.E	-	-	-	-	4	D(F380Cα- L356Cα) ≈ 6 Å
I382 to 383	N.E	N.E	-	-	-	-	4	D(Cα Cα) > 14 Å
V384	N.E	N.E	B MTSEA	B MTSEA			4 5	D(Cα Cα) > 14 Å Face Central pore
L385	N.E	N.E	N.E	N.E			4 & 5	D(Cα Cα) > 14 Å
I386	N.E	N.E	N.E	N.E			4 & 5	D(Cα Cα) > 14 Å
F387	N.E	N.E	B MTSEA	B MTSEA			4 & 5	D(Cα Cα) > 14 Å Face Central pore
A388	N.E	N.E	B MTSEA	B MTSEA			4 & 5	D(Cα Cα) > 14 Å Face Central pore
Val391	R.B	R.B	-	-	B	B	4 &5	D(Cα Cα) ≈ 14 Å

								Face central Pore
Gly392	N.E.	N.E.	B MTSET	B MTSET	-	-	4 &5	D(Cα Cα) > 14 Å Face Central pore
Asn393	"	"	N.E	N.E	-	-	4 & 5	"
Ile394	"	"	-	-	-	-	4	"
Gly395	R.B.	R.B.	B MTSEA	B MTSEA	-	-	4 & 5	D(Cα Cα) ≈ 14 Å Face Central Pore
Ser396	B	B	-	-	-	-	4	D(Cα Cα) < 14 Å
Ile398	N.E.	B	-	-	-	-	4	D(Cα Cα) > 14 Å (Open) D(Cα Cα) ≈ 11 Å (Closed)
Ser399	R.B	R.B	B MTSEA	B MTSEA	-	-	4 &5	D(Cα Cα) ≈ 14 Å Face Central pore
Asn400	N.E.	B	-	-	-	-	4	D(Cα Cα) > 14 Å (Open) D(Cα Cα) ≈ 11 Å (Closed)
Met401	N.E.	N.E.	-	-	-	-	4	D(Cα Cα) > 14 Å
Asn402	N.E.	B	-	-	-	-	4	D(Cα Cα) > 14 Å (Open) D(Cα Cα) ≈ 11 Å (Closed)
Ala403	N.E.	B	-	-	-	-	4	D(Cα Cα) > 14 Å (Open) D(Cα Cα) ≈ 11 Å (Closed)
Ala404	N.E.	B	-	-	-	-	4	D(Cα Cα) > 14 Å (Open) D(Cα Cα) ≈ 11 Å (Closed)
Arg405	N.E.	N.E.	-	-	-	-	4	D(Cα Cα) > 14 Å
Ala406	N.E.	B	N.E MTSET	N.E MTSET	-	-	4	D(Cα Cα) > 14 Å (Open) D(Cα Cα) ≈ 11 Å (Closed)
Asp407	N.E.	B	-	-	-	-	4	D(Cα Cα) > 14 Å (Open) D(Cα Cα) ≈ 11 Å (Closed)
Phe408	N.E.	B	-	-	-	-	4	D(Cα Cα) > 14 Å (Open) D(Cα Cα) ≈ 11 Å (Closed)
Gln409	N.E.	B	-	-	-	-	4	D(Cα Cα) > 14 Å (Open) D(Cα Cα) ≈ 11 Å (Closed)
Ala410	N.E.	N.E.	N.E MTSET	N.E MTSET	-	-	4	D(Cα Cα) > 14 Å
Ile412	N.E.	N.E.	-	-	-	-	4	D(Cα Cα) > 14 Å
Asp413	N.E.	N.E.	B	B	-	-	4	D(Cα Cα) > 14 Å

Ala414	N.E.	B	-	-	-	-	4	D(Cα Cα) > 14 Å (Open) D(Cα Cα) ≈ 11 Å (Closed)
Ile415	N.E.	N.E.	-	-	-	-	4	D(Cα Cα) > 14 Å
Lys416	N.E.	N.E.	-	-	-	-	4	D(Cα Cα) > 14 Å
Gln417	N.E.	B	N.E	N.E	-	-	4	D(Cα Cα) > 14 Å (Open) D(Cα Cα) ≈ 11 Å (Closed)
Tyr418	N.E.	B	-	-	-	-	4	D(Cα Cα) > 14 Å (Open) D(Cα Cα) ≈ 11 Å (Closed)
Met419	N.E.	N.E.	-	-	-	-	4	D(Cα Cα) > 14 Å
His420	N.E.	N.E.	N.E	N.E	-	-	4	D(Cα Cα) > 14 Å
Phe421	N.E.	N.E.	-	-	-	-	4	D(Cα Cα) > 14 Å
Arg422	N.E.	B	-	-	-	-	4	D(Cα Cα) > 14 Å (Open) D(Cα Cα) ≈ 11 Å (Closed)

The models, obtained based on the alignments shown in Fig.4.1 B, were refined by the inclusion (Appendix B) of an extensive dataset (more than 70) of spatial constraints inferred from ion current measurements of cysteine mutants in the presence of a variety of chemicals(Becchetti et al., 1999;Liu and Siegelbaum, 2000;Flynn et al., 2001;Flynn and Zagotta, 2001;Flynn and Zagotta, 2003)(Nair et al. in preparation) (Mazzolini et al. in preparation) (Tab. 4.1): (i) Cd^{2+}, which can block the channel *reversibly or irreversibly.* In the first case, the ion binds to two adjacent cysteines (see Tab. 4.2). Estimates obtained by a calculation of residue-residue distribution function based on the RCSB Protein Data Bank(Berman et al., 2000) suggest that the Cα of adjacent cysteine are located at less than 11 Å (Tab. 4.2). In the second case, the Cd^{2+} ion binds to three or four cysteines (Rothberg et al., 2002), suggesting that the Cα of opposite cysteines are located atless than 11 Å. ii) Mild-oxidizing agent copper phenanthroline (CuP), favors disulfide bridge formation between two cysteines separated by a distance going from 3.6 to 7.5 Å (see Tab 4.2).

In addition, information on solvent accessibilities is obtained by using solutions rich in the differently sized and charged sulfhydryl-specific reagents such as MTSET+,

MTSEA+ and MTSES-, those can produce irreversible changes in current flow upon reaction with solvent accessible cysteines (Akabas et al., 1992)(Tab. 4.1) [1].

In S6 TMH and the N-Term@C-Linker, Cd^{2+} blockage is reversible (or almost absent) for residues upstream Gly395 in both open and close states whilst it is irreversible and strongly state dependent for residues downstream Asn400 (Fig 1.8 and Tab 4.1). As a result, in the open state, the C-linker@N-term is bent around a hinge located approximately between Val391 and Gly395 and it is also rotated in clockwise sense by about 45° related to the closed conformation, assuming a configuration similar to that of the template (Fig 4.2). These conformational changes resemble somehow that proposed for the Shaker K^+ channel (Webster et al., 2004), in which the S6 TMH bends around a valine residue (V374). This residue is in correspondence to CNG's Val391, suggesting a common mechanism in the gating of a variety of ion channels.

In the P-helix-loop region, three structural features can be established. First, L356@P helix forms hydrophobic interactions with F380@S6. In fact, L356C and F380C single mutants, show inactivation and locking effects, respectively; whilst L356C & F380C (Tab. 4.1) behaves as wt, suggesting that in the latter case the hydrophobic interactions are substituted by a S-S bridge (Fig. 4.2). In addition, the P helix changes its location in space on passing from closed to open form. Indeed, Thr355 and Leu358 of the P-helix are accessible to extracellular solvent only in the closed state, whilst Val348 and Leu351 are accessible in both states, as described above (Tab. 4.1). Finally, in the intraceullar side, Thr360 is solvent accessible and it exhibits a different conformation in the two states, as shown by the differences in response upon Cd^{2+} addition (Fig 4.2). Because Thr360 is directly linked to the pore wall through Ile361 (already belongs to the wall), also the latter is expected to rearrange upon gating.

[1] Because the topology of MthK is similar to that of KcsA, we have attempted using the former as a template for the open configuration. The MthK based constructs from Cd^{2+} adducts (Tab 4.1) resulted in

Table 4.2. *Residue-residue distribution functions. For cysteines making disulphide bond, these were obtained by extensive scanning of the non redundant RCSB protein data bank (PDB)(Berman et al., 2000). Instead, for cysteines bound to* Cd^{2+} *a direct inspection of the 3D structure of just metallothioneins deposited in the PDB(Berman et al., 2000) was performed. Thus,* Cd^{2+} *and CuP can be used to have a rough evaluation on the residue relative distances.*

	Mean (Å)	**S.D. (Å)**	**Range (Å)**	**Allowed distance(Å)[2]**
Cα@Cys- Cα@Cys	5.5	0.7	3.6 -7	9
Cd – Cα@Cys	4.4	0.4	3 – 5	6
Cα@Cys - Cd -Cα@Cys	7	1.3	5 – 9.2	11

Finally, the model pore lumen diameter increases from about 3 Å to about 6 Å, in agreement with measurements based on the permeability of the channel to large organic cations (such as NH_4^+, $CH_3NH_3^+$, $(CH_3)_2NH_2^+$, $(CH_3)_3NH^+$, $CH_2CH_3NH_3^+$). Using these techniques it was possible to estimate the diameter of the narrowest part of the CNG channel pores in the open configuration, which for the bovine CNGA1 is between 4 and 6 Å (Goulding et al., 1993; Bucossi et al., 1996; Laio and Torre, 1999).

On the basis of these finding, we propose that gating occurs by bending and a clockwise rotation by about 45° (around channel axis) of N-terminal section of C-linker (Fig. 4.2 and 4.3). This rotation is transmitted upwards, making the upper part of S6 to rotate anticlockwise by about 30° (around the helix axis). Because of the S6 interacts directly with the P-helix, this motion is transmitted to the pore helix, which rearranges so as its terminal Thr360 residues and, therefore the lower part of the wall pore, leading to the opening of the pore wall region (Fig 4.3).

In conclusion, the initial event of cyclic nucleotide binding is transmitted to the pore walls by a remarkable and sophisticated coupling of conformational changes spanning

---

extremely large rearrangements (more than 10 Å), while causing relatively small rearrangements (overall RMSD of 3 Å) (Fig. 4.2) in the KcsA-based model. Thus, MthK was not considered any further.

[2] Maximum allowed distance considering the thermal fluctuations of the protein(Careaga and Falke, 1992).

throughout the entire cytoplasmic domain of the channel. The novel computational approach can be extended to uncover mechanisms of action in other ion channels.

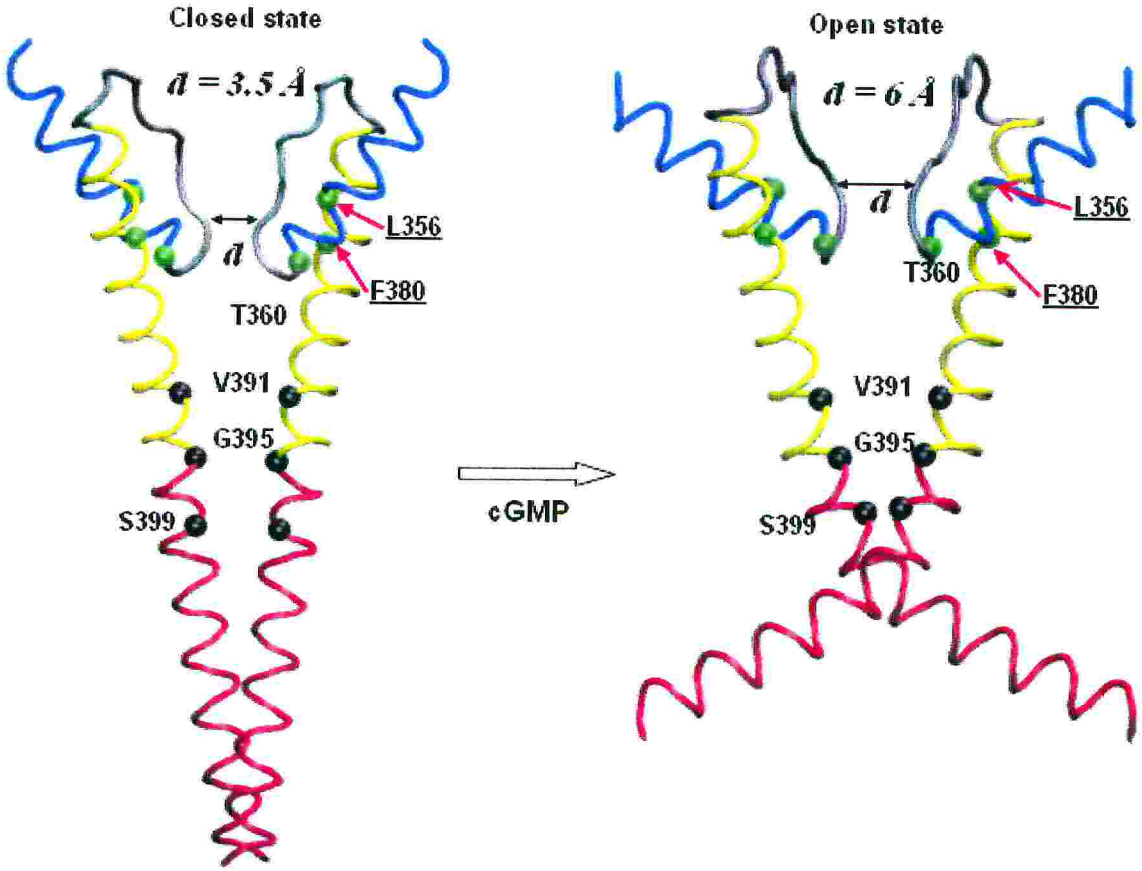


Fig 4.2. *Structural models of S6, the P loop of the transmembrane domain along with the N-term@C-linker the closed and open state. Only two adjacent subunits are shown for the sake of clarity. The structures were obtained by homology modelling (See Fig. 1B) and inclusion of the spatial constraints in Tab. 1SI. The following residues are shown: (i) F380 and L356, which are underlined, interact via hydrophobic contacts. (ii) T360, which rearranges upon gating (See text). (iii) Residues V391, G395 and S399, whose mutation into cysteines, along with* Cd^{2+} *addition, causes a reversible block of ion current in both states. Only two adjacent subunits are shown for the sake of clarity.*

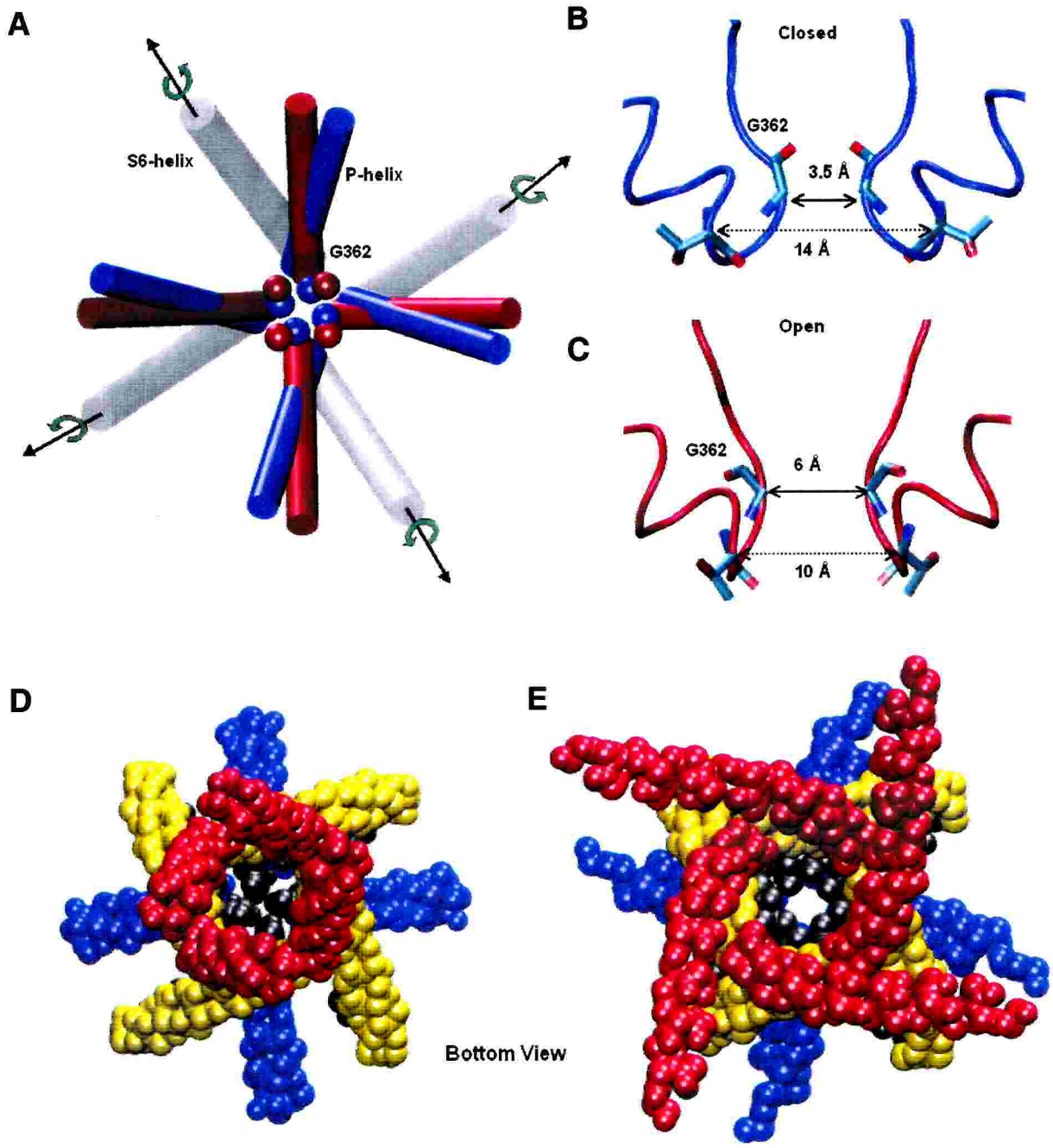


Fig 4.3: *Model of the closed and open pore region.* ***A:*** *top view of the hypothesized configuration of the pore in the closed (in blue) and open (in red) state. The S6 alpha helices are shown in grey and the hypothesized anticlockwise rotation initiating gating is indicated by the green arrows around the axis through the S6 helices.* ***B*** *and* ***C:*** *the hypothesized conformation of two opposing subunits in the closed and open configuration respectively* ***D and E*** *Bottom view of the full model for the closed and open configuration, respectively (the color coding is the same as for figure 4.1 and 4.2). The model in the closed state configuration was obtained (see Methods) by homology modeling using as template the KcsA structure and the constraints suggested by the experiments here described.*

5 Conclusions

The structures of two different ionic channels, HCN and CNG channels, both sharing low sequence similarity with their templates, were built by use of comparative modelling techniques. In the absence of high sequence identity with the template, state-of-the-art comparative modelling techniques cannot provide a unique structure for the target proteins. For this reason, available experimental data and multiple models generation were used in order to guide the comparative modelling at all levels of the procedure. At the basis of the whole study was the idea to combine into an integrated approach theoretical structural methods and experimental techniques. In this approach, theoretical modelling played a fundamental role by providing a working hypothesis for the interpretation of experimental data, whereas experiments were crucial for validation and refinement of the theoretical models.

All the available experimental data were used in the construction of models in both open and closed states. These models were used to rationalize structural features and mechanisms of the cyclic nucleotide gated channels.

HCN channels. Our model provides a structural framework for understanding several structural properties of HCN channels: i - the cysteine ring at the inner mouth of the pore is a sensor of the intracellular oxidizing/reducing conditions; ii - the bending of the S6 helix upon gating is significantly smaller than that occurring in MthK channels; iii - the reduced ionic selectivity of HCN channels compared to that of K^+ channels could be caused to the larger flexibility of the inner pore of HCN channels.

CNG channels. The large number of experimental data has allowed us to provide a structural basis for the gating. The latter initiated by a clockwise rotation of the N-terminal region of the C-linker, which is then transmitted trough S6 transmembrane helices to the P-helix, and in turn from this to the pore lumen, which opens up by about 3 Å (from 2 to 5 Å) thus allowing for ion permeation.

We conclude that the studied channels exhibit slightly different gating mechanisms, although a similar initial conformational change is transmitted to S6, generating an opening of the channel by helix bending. In fact, in CNG channels this conformational

change is transmitted to the P-helix-loop region, whilst in HCN the coupling between S6 and P-helix seems to be less strong, not allowing in this way a conformational change to be transmitted to the filter region.

The fundamental differences in gating might be the cause of differences in rigidity/flexibility of the channel pore and so, directly related with the highly divergent selectivity properties of both channels(Laio A. and Torre, 1999). Thus, HCN channels exhibit intermediate properties between pure voltage-gated K^+ channels and pure Cyclic-nucleotide gated channels.

In conclusion, because of the constantly improving bioinformatics techniques and of the rapidly increasing number of high-resolution protein structures, we anticipate that our combined experimental/computational approach to play an increasingly important role in membrane structure predictions in the next future.

Appendix A: Methods

A.1 Biological Basis of Sequence Alignment techniques.

Sequence alignment. Proteins are a product of evolution. Their sequences are encoded histories of mutation and selection over millions of years. Closely related protein sequences, if aligned, exhibit several residues at the same at many positions but differ at others. Furthermore, alignment of many such sequences will show rich patterns of variation and consistency. The alignment among biological sequences is the cornerstone of structural bioinformatics, because protein evolution has led to families of structures that have the same fold at the level of secondary structures, motifs, domains or entire globular proteins. Since the number of stable folds is limited, tertiary structures turn out to be more conserved in evolution than the amino acid sequences (Richardson, 1977).

Comparative modelling is a structural prediction method that exploits the observation that proteins evolving from a common ancestor maintained similar core 3D structures. In other words, the core structures of two homologous (i.e. evolutionary related) proteins are expected to be similar. Comparative modelling is then used to construct structural models of proteins (targets) that are homologous to other proteins whose 3D structure is known (templates). It is therefore necessary to define mathematically the similarity between aminoacids, which provides the plausibility of a particular substitution. This leads to a 20x20 matrix, where the value of each element represents the similarity of the row and column amino acids.

The simplest is the identity matrix, which gives identical amino acids a similarity of 1 and 0 for non identical ones. The scheme based on this scoring is referred to as the percentage identity. Percentage is a very widely used measure of sequence similarity because it is easily understood. However, there is no obvious way to directly relate percentage identity with evolutionary distance.

More sophisticated substitution matrices, instead, are based on observed substitutions on highly similar sequences (Barker et al., 1978).

In our work, we have used two common substitution matrices for the initial alignments: PAM250 and BLOSUM62. PAM250 are appropriate for aligning sequences that are about 20 % identical, and BLOSUM62 is the best for detecting the majority of weak

protein similarities (Henikoff and Henikoff, 1992). In our case we have used both of them to get two different initial alignments. In the following paragraphs we will describe how these two matrices are obtained.

PAM250 Matrix. This is based on The Percent Accepted Mutations (PAM) matrix, which weights how often amino acids replace each other during evolution(Barker et al., 1978). The model of evolution that Dayhoff(Barker et al., 1978) used, assumed that proteins diverged as a result of accumulated, uncorrelated mutations[3]. Based on this model, the PAM matrix was calculated over a data base of 1,572 changes in 71 groups of closely related proteins.

A contribution of the relative mutability of different amino acids is added to the PAM matrix in such a way that each matrix element gives the probability that the amino acid in one column will be replaced by the amino acid in some row after a given evolutionary interval. For example, a matrix with an evolutionary distance of 0 PAMs would have ones on the main diagonal and zeros elsewhere. A matrix with an evolutionary distance of **1** PAM would have numbers close to one on the main diagonal and small numbers off the main diagonal. One PAM would correspond to roughly 1% divergence in a protein (one amino acid replacement per hundred). To derive a mutational probability matrix for a protein sequence that has undergone N % accepted mutations **(PAM-N)** matrix, the **PAM-1** matrix is multiplied by itself **N** times. This results in an extrapolated family of scoring matrices.

By trial and error Dayhoff *et al.* (Barker et al., 1978) found that for weighting purposes a PAM250 matrix works well. At this evolutionary distance, only one amino acid in five remains unchanged and the divergence has increased to roughly 80%. The amino acids vary greatly in their mutability. According to Dayhoff *et al.*, roughly 55% of the tryptophans, 52% of the cysteines and 27% of the glycines would still be unchanged, but

---

[3] This is the called ***Neutralist Model***. In this model most of the observed substitutions in a multiple sequence alignments are neutral (i.e. have negligible effects on the function): rather than representing improvements in a protein, indicate how tolerant the protein is to change at this position. In an already optimized protein, regions that are more constrained by function will show slower rates of substitution and vice versa. There exist another model, less used: ***Selectionist Model***. According to this model, the majority of accepted mutations confer a selective advantage whereas neutral mutations are rare. Most of the mutations observed in a multiple alignment therefore represent evolutionary adaptations to different environments and have arisen through natural selections rather than random neutral drift.

only 6% of the highly mutable asparagines would remain. Several other amino acids particularly alanine, aspartic acid, glutamic acid, glycine, lysine and serine are more likely to occur in place of an original asparagine than asparagine itself at this evolutionary distance.

What is actually tabulated is the odds matrix, which takes the elements of the previous matrix (target frequencies or $P(i,j)_{match}$) and divides each term by the frequency of the replacement residue (background or random frequencies[4]): thus, each term represents the probability of replacement, **j** to **i** per occurrence of residue **j**.

$$R(i,j) = \frac{P(i,j)_{match}}{P(i,j)_{random}} \qquad (1)$$

where, if f_i and f_j are the fractional frequencies of i and j respectively, then $P(i,j)_{random} = f_i \times f_j$

Due to historical reasons, the relatedness odds matrix is usually converted to the log-odds matrix, thus:

$$M(i,j) = \log_{10} R(i,j) \qquad (2)$$

These log odds PAM250 matrix is shown in Table 1: larger is the value, more often a particular mutation is found.

Some of the properties that are visible from this matrix and go into its makeup are - size, shape, local concentrations of electric charge, conformation of van der Waals surface, ability to form salt bonds, hydrophobic bonds, and hydrogen bonds. Interestingly, these patterns are imposed principally by natural selection and only secondarily by the constraints of the genetic code. Limitations of the method are: i) it assumes that all sites are equally mutable; ii) The ensemble of proteins used to calculate the matrix exhibited only few differences; iii) the matrix is biased because it is based on small globular proteins.

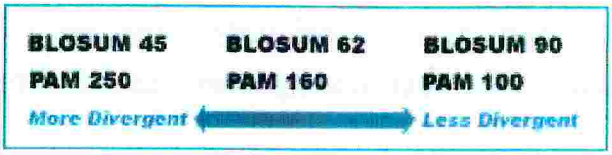
[4] That is, if there were no selection for fitness, the frequencies of each possible substitution would be primarily influenced by the frequencies of the different amino acids.

Table 1 - The log odds matrix for 250 PAMs (multiplied by 10)

	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y
A	2	-2	0	0	-4	1	-1	-1	-1	-2	-1	0	1	0	-2	1	1	0	-6	-3
C		12	-5	-5	-4	-3	-3	-2	-5	-6	-5	-4	-3	-5	-4	0	-2	-2	-8	0
D			4	3	-6	1	1	-2	0	-4	-3	2	-1	2	-1	0	0	-2	-7	-4
E				4	-5	0	1	-2	0	-3	-2	1	-1	2	-1	0	0	-2	-7	-4
F					9	-5	-2	1	-5	2	0	-4	-5	-5	-4	-3	-3	-1	0	7
G						5	-2	-3	-2	-4	-3	0	-1	-1	-3	1	0	-1	-7	-5
H							6	-2	0	-2	-2	2	0	3	2	-1	-1	-2	-3	0
I								5	-2	2	2	-2	-2	-2	-2	-1	0	4	-5	-1
K									5	-3	0	1	-1	1	3	0	0	-2	-3	-4
L										6	4	-3	-3	-2	-3	-3	-2	2	-2	-1
M											6	-2	-2	-1	0	-2	-1	2	-4	-2
N												2	-1	1	0	1	0	-2	-4	-2
P													6	0	0	1	0	-1	-6	-5
Q														4	1	-1	-1	-2	-5	-4
R															6	0	-1	-2	2	-4
S																2	1	-1	-2	-3
T																	3	0	-5	-3
V																		4	-6	-2
W																			17	0
Y																				10

BLOSUM matrix. The BLOSUM matrices have been constructed in a similar fashion than PAM matrices, but make use of a different strategy for estimating the target frequencies (Henikoff and Henikoff, 1992). The underlined data are derived from the BLOCKS database (Henikoff and Henikoff, 1992) which contains local multiple sequence alignments (blocks) involving distantly related sequences (as opposed to the closely related sequences of PAM). Although there is no evolutionary model in this case, it is advantageous to have data generated by direct observation, rather than extrapolation. Different levels of the BLOSUM matrix can be created by differentially weighting the degree of similarity between sequences. For example, a BLOSUM62 matrix is calculated from protein blocks such that if two sequences are more than 62% identical, then the contribution of these sequences is weighted to sum to one. In this way the contributions of multiple entries of closely related sequences is reduced.

The BLOSUM62 matrix turns out to be roughly equivalent to a PAM 150 matrix. Conversely, the well known PAM250 is equivalent to BLOSUM45 matrix.



Indeed, visual inspection shows that BLOSUM matrices are less tolerant of substitutions to or from hydrophilic amino acids, while more tolerant of hydrophobic changes and of cysteine and tryptophan mismatches.

Table 2 - The log odds matrix for BLOSUM 62

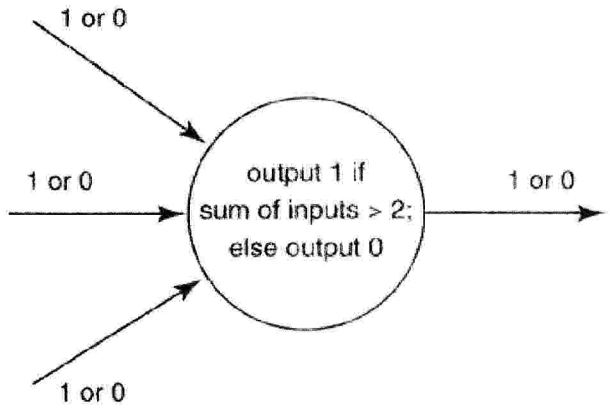
	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y
A	4	0	-2	-1	-2	0	-2	-1	-1	-1	-1	-2	-1	-1	-1	1	0	0	-3	-2
C		9	-3	-4	-2	-3	-3	-1	-3	-1	-1	-3	-3	-3	-3	-1	-1	-1	-2	-2
D			6	2	-3	-1	-1	-3	-1	-4	-3	1	-1	0	-2	0	-1	-3	-4	-3
E				5	-3	-2	0	-3	1	-3	-2	0	-1	2	0	0	-1	-2	-3	-2
F					6	-3	-1	0	-3	0	0	-3	-4	-3	-3	-2	-2	-1	1	3
G						6	-2	-4	-2	-4	-3	0	-2	-2	-2	0	-2	-3	-2	-3
H							8	-3	-1	-3	-2	1	-2	0	0	-1	-2	-3	-2	2
I								4	-3	2	1	-3	-3	-3	-3	-2	-1	3	-3	-1
K									5	-2	-1	0	-1	1	2	0	-1	-2	-3	-2
L										4	2	-3	-3	-2	-2	-2	-1	1	-2	-1
M											5	-2	-2	0	-1	-1	-1	1	-1	-1
N												6	-2	0	0	1	0	-3	-4	-2
P													7	-1	-2	-1	-1	-2	-4	-3
Q														5	1	0	-1	-2	-2	-1
R															5	-1	-1	-3	-3	-2
S																4	1	-2	-3	-2
T																	5	0	-2	-2
V																		4	-3	-1
W																			11	2
Y																				7

Insertions and deletions must be often introduced in the alignments. This procedure decreases the alignment score for each gap introduced. A number of strategies have been proposed for penalizing gaps. The most common formulation, known as the affine gap penalties, involves a fixed deduction for introducing gaps plus an additional deduction proportional to the length of the gap.

A.2 Secondary Structure prediction

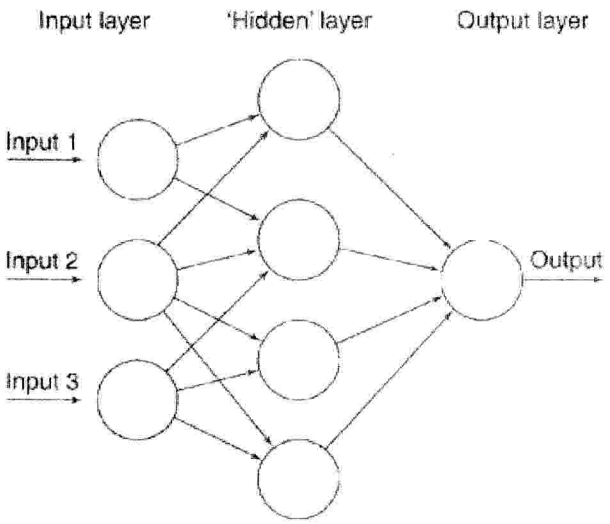
To align the sequences, it is very useful to first identify secondary structure elements and, in membrane proteins, putative transmembrane regions. The most powerful methods for secondary structure prediction are based in neural networks. Neural networks are a class

of general computational structures based loosely on the anatomy and physiology of biological nervous systems. They have been applied successfully to a wide variety of pattern recognition, classification, and decision problems. A single neurone, in the computational scheme, is a node in a directed graph, with one or more entering connections designated as input, and a single leaving connection called the output:



In physiology, one says that the neurone 'fired' if the output is 1, and that the neurone 'did not fire' if the output is 0. Simulated neurons can differ in the number of input and output connections, and in the formula for deciding whether to fire.

Networks, assemble several neurons and connect the outputs of some to the inputs of others. Some nodes contain connections that provide input to the entire network; some deliver output information from the network to the outside world; and others, that do not interact directly with the outside, are called *hidden layers:*



An unlimited degree of complexity is available by assembling and connecting neurons, and by varying the strengths of the connections. That is, instead of taking a simple sum of inputs $i1 + i2 + i3$, take a weighted sum - for instance, $10i1 + 5i2 + i3$, which would make the neurone most sensitive to input 1 and least sensitive to input 3. Biologically, this may correspond to changing the strengths of synapses.

Neural networks applicable to secondary structure prediction contain three basic layers.

1) The input layer is a sliding 15-13-residue window that scans the sequence. That is, it treats a 15-13- residue region, predicts the secondary structure of the central residue and then moves the window one residue along the amino acid sequence and repeats the process. To each of the 15-13 residues in the current window there corresponds 20 nodes in the input layer of the network, one of which will be triggered according to the amino acid in that position.

2) A hidden layer of ~ 100 units connects the input with the output. Each node of the hidden layer is connected to *all* input and output units; not all the connections are shown.

3) The output layer consists of only three nodes that signify prediction that the central residue in the window be in a helix, strand, or other conformation.

Most methods feed the input layer with the identities of the amino acid at successive positions, together with a profile derived from a multiple sequence alignment.

To have an idea of how these neural networks work, in the following section we will briefly describe the method used in this thesis for the prediction of secondary structure elements (PHDsec) and transmembrane helix prediction (PHDhtm): PHD.

A.2.1 Generating the multiple sequence alignment

PHD first generates a multiple sequence alignment, which is feed into a neural network system. The correctness of the multiple sequence alignment is as crucial for prediction accuracy as that the alignment contains a broad spectrum of homologous sequences. By default, PHD uses the program MaxHom (Fig. A.1) that generates a pairwise profile-based multiple sequence alignment. A key feature of MaxHom is the compilation of a length-dependent cut-off for significant pairwise sequence identity.

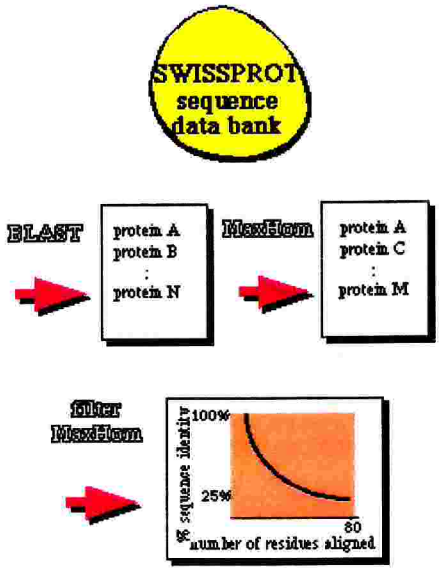


Fig. A.1 *First, for each protein, the SWISSPROT data base is searched for sequence homologues with a fast alignment method (BLAST , (Altschul et al., 1990)). Second, the list of putative homologues found is re-examined with a more sensitive profile-based multiple alignment method (MaxHom,(Sander and Schneider, 1991)). Third, a length dependent cut-off for significant pairwise sequence identity is applied (typically 30 %).*

A.2.2 Multiple levels of computations

The PHD methods process the input information on multiple levels (Fig.A.2). The first level is a feed-forward neural network with three layers of units (input, hidden, and output). Input to this first level sequence-to-structure network consists of two contributions: one from the local sequence, i.e., taken from a window of 13 adjacent residues, and another from the global sequence (Fig.A.2). Output of the first level network is the 1D structural state of the residue at the centre of the input window. For PHDsec and PHDhtm the second level is a structure-to-structure network (see below). The next level consists of an arithmetic average over independently trained networks (jury decision). The final level is a simple filter.

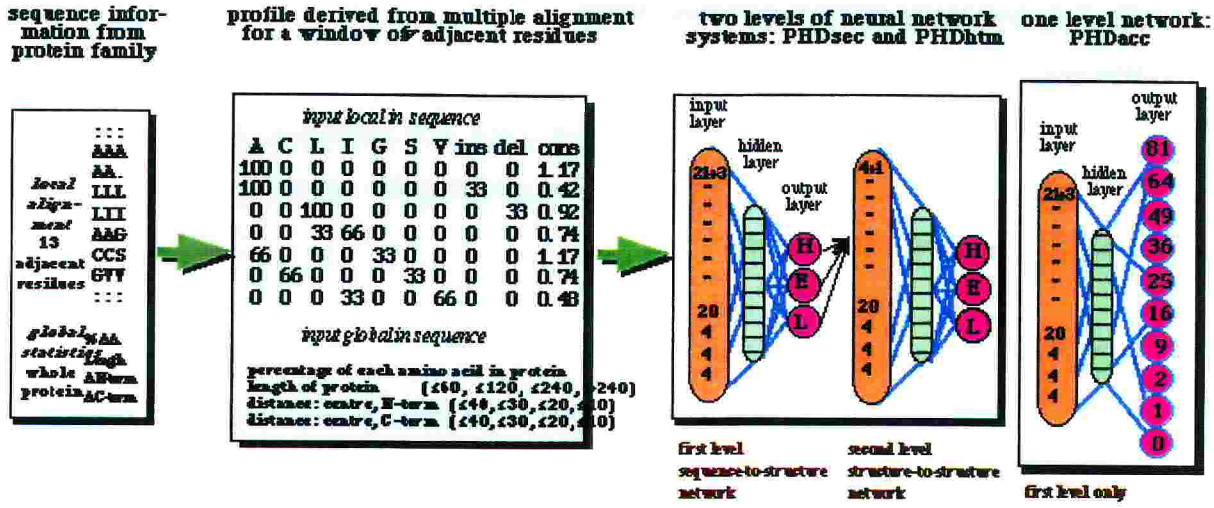


Fig A.2. *First, a window of w =13 adjacent residues is chosen from the alignment. Second, for each residue the profile and global information is compiled from the protein. Third, the local and global information is fed into neural network systems. PHDsec and PHDhtm consist of two network levels. First level, sequence-to-structure networks, for each residue position 24 units are used, 20 for the amino acid types, one for a 'spacer' allowing the window to extent over the protein ends, two for the numbers of insertions (ins) and deletions (del) in the alignment at that position, and one for the conservation weight (cons); the global information is coded by 20 units for the amino acid composition, four for the protein length, and eight for the distances of the window with respect to the protein ends. The output units code for the 1D structural state of the central residue. For PHDsec, three output units code for helix, strand, and restand for PHDhtm, two units code for transmembrane or not transmembrane helix. Second level, structure-to-structure networks: the output of the first level is fed into a second level of structure-to-structure network, which additionally uses global information and the conservation weight as input, e.g., for PHDsec: first level output = 3 units -> local input to second level = 3 + 1 (spacer) + 1 (cons). The output of the second level is the same as that of the first level.*

A.2.3 Number of output units determined by task

Secondary structure is coded by three units: helix, H; strand, E; and none of the above, denoted loop, L. Transmembrane locations are coded by two units, one for residues being in a transmembrane helix, the other for non-membrane bound residues (assignments from SWISSPROT (Bairoch and Boeckmann, 1994)).

A.3 Comparative modelling

We now describe how in practice we can construct the 3D structure of proteins whose structure is not known based on sequence alignment. The modelling procedure consists of four principal steps: template recognition, target/template sequence alignment, model building, and model evaluation.

A.3.1 Template recognition

The first step is the identification of all protein structures related to the target sequence, i.e. proteins that may be regarded as possible templates. The templates are identified from protein sequence databases (Barker et al., 1998), GenBank (Benson et al., 1998), or TrEMBL/SWISS-PROT (Bairoch and Apweiler, 1999) and/or from protein structure databases (Berman et al., 2000), SCOP (Hubbard et al., 1999), DALI (Holm and Sander, 1999), and CATH (Orengo et al., 1999)).

In our case, the template recognition was straightforward due to the existence of just a few crystal structures of ion channels (see Introduction).

A.3.2 Target/template sequence alignment

The target sequence is aligned to all the template structures (Holm and Sander, 1996; Taylor, 1996; Baxevanis, 1998). The quality of the alignment is the single most important factor determining the accuracy of the 3D model. If SI is about 30%, the fraction of residues that can be correctly aligned is about 80% (Johnson and Overington, 1993). If SI is lower, one should use not only sequence information but also: (i) structural information as predicted by algorithms that give a reasonable guess of the target topology (e.g. secondary structure predictions); (ii) available experimental information, which can provide suggestions on the alignment of single residues.

In all circumstances, one might have to edit manually the positions of insertions and deletions to ensure that they occur in a reasonable structural context. For example, gaps are favoured outside secondary structure segments and in exposed regions. More than one plausible alignment may be possible, and the decision on the most appropriate one is often taken by directly evaluating the corresponding model on the basis of its structural properties and of the available experimental data.

A.3.3 Model building by satisfaction of spatial restraints

To build our models we used Modeller6v2 (Sanchez and Sali, 1997), a program that implements modelling by satisfying stereo chemical, homology-derived spatial restraints and external restraints. This is achieved by introduction of molecular restraints, obtained by maximization of probability density functions (pdfs)[5].

Molecular pdf

The molecular pdf gives a probability for the simultaneous occurrence of combinations of the feature pdfs. The molecular pdf can be written as a product of features pdfs:

$$P = \prod_i p^{f_i} \tag{3}$$

where p^{f_i} is defined as any quantity associated with a particular set of atoms as, for example, the distance between two atoms, or the angle among three atoms, and is calculated as:

$$p^f = \prod_i p^{b_i} \tag{4}$$

Every p^{f_i} can be restrained by several basis pdfs (p^{b_i}). A feature pdf is hence a pdf that combines all the information about the possible values that a particular feature can assume. As all restraints have to be satisfied in all structures, the feature pdf is obtained by multiplication of the p^{b_i}

We now briefly review the functional form of the single pdf's used in MODELLER.

Stereochemical basis pdfs are easily obtained from the amino acid sequence of a protein. The geometric restraints that are, bond lengths, bond angles and dihedral angles are expanded as Gaussian functions. The mean values and standard deviations are obtained from the CHARMM parameter set. The Coulomb and van der Waals energy terms are the same as the CHARMM force field (MacKerell et al., 1998).

i) Geometric Restraints: Bond lengths, bond angles and dihedral angles.

For example, the classical harmonic model for the bond length b (b_0 is a reference distance) between two atoms gives the vibrational potential of a bond as:

[5] The pdfs must be non-negative and integrate to 1 over the values of r.

$$E(b) = \frac{1}{2} c(b - b_0)^2 \tag{5}$$

So, the basis pdf for the bond length is then found, from classical statistical, to be a Gaussian pdf (Hill, 1960):

$$p^b(b) = \frac{1}{\sigma_b \sqrt{2\pi}} \exp\left[-\frac{1}{2} \left(\frac{b - \bar{b}}{\sigma_b} \right)^2 \right] \tag{6}$$

Where $\sigma_b = \sqrt{kT/c}$ and $\bar{b} = b_0$. Only two parameters, the mean $\bar{b}$ and the standard deviation σ_b (c is the force constant), are needed to describe this distribution. As said before, these parameters are taken from CHARMM (MacKerell et al., 1998) parameter set.

The derivation of the pdf for the bond angle is equivalent to that for the bond length. The final result is again a Gaussian pdf $p^\alpha(\alpha) = N(\alpha_0, \sigma_\alpha)$, were α_0 and σ_α are the average and standard deviation angles between two specified atoms, respectively. Similarly a monomodal pdf for a torsional or an improper dihedral angle ζ is $N(\zeta_0, \sigma_\alpha)$, where ζ_0 and σ_α are the average and standard deviation for dihedral angles, respectively. This monomodal pdf is used to restraint several different features of protein structure. They include peptide and ring planarities, and chiralities of Cα atoms, Thr and Ile side-chains.

ii) Coulomb Restraint.

$$p(r) = \frac{q_i q_j}{r} s(r, r_1, r_2) \tag{7}$$

$$s(r, r_1, r_2) = \begin{cases} 1; & r \le r_1 \\ \dfrac{(r_2 - r)^2 (r_2 + 2r - 3r_1)}{(r_2 - r_1)^3}; & r_1 < r \le r_2 \\ 0; & r > r_2 \end{cases} \tag{8}$$

Where q_i and q_j are the atomic charges of atoms i and j and r_1 and r_2 are cutoff distances, obtained from the CHARMM topology file(MacKerell et al., 1998). These charges are at a distance r. The function $s(r, r_1, r_2)$ is a switching function that smoothes the potential

down to zero in the interval from r_1 to r_2 ($r_1 < r_2$). The total Coulomb energy of a molecule is a sum over all pairs of atoms that are not in the same bonds or bond angles.

iii) Lennard-Jones restraint

Usually used for non-bonded distances (r):

$$p(r) = \left[\left(\frac{A}{r}\right)^{12} - \left(\frac{B}{r}\right)^{6}\right] s(r, r_1, r_2) \tag{9}$$

The parameters r_1 and r_2 of the switching function can be different from those in Eq. 8, and together with the parameters A and B, are obtained from the CHARMM parameter file (NONBOND section).

The total Lennard-Jones energy should be evaluated over all pairs of atoms that are not in the same bonds or bond angles.

Homology-derived basis pdfs are defined based on a database of 105 family alignments. In order to clarify how they were obtained, a representative example of the C_α-C_α atoms distance restraint is given.

In this case, the property of interest is defined as the difference (d-d') between two equivalent C_α-C_α distances: d' from a known structure (template), d from an unknown structure (target).

The difference (d-d') between two equivalent C_α-C_α is considered as a function of four independent variables: (a) the corresponding C_α-C_α distance in the template (i.e. d'); (b) the fractional sequence identity between two aligned sequences; (c) the average solvent accessibilities of the two residues spanning the distance d' in the known structure and (d) the average distance from a gap of the residues spanning the distance d'. By considering all possible homologous protein pairs (order 10^3) in the database of protein family alignments, it is then possible to extract the conditional pdf, $p(D_{C\alpha\text{-}C\alpha} \mid a,b,c,d)$:

$$p^{d}(D_{C\alpha-C\alpha} \mid a,b,c,d) = \frac{1}{\sigma(a,b,c,d)\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{d-d'}{\sigma(a,b,c,d)}\right)^{2}\right] \tag{10}$$

where $\sigma(a,b,c,d)$ is a combination of properties (a) to (d) with coefficients extracted from the database analysis.

Thus, a C_α-C_α distance restraint is defined on a target sequence once (a), (b), (c) and (d) are specified.

'Alpha' restraint makes restraints enforcing an α-helix for the residue segment specified. The helices are restrained by Phi/Psi binormal restraints, N-O hydrogen bonds, Ca-Ca distances for i-j ∈ (2-9), Ca-O distances for i-j ∈ (2-9), and O-O distances for i-j ∈ (2-6).

External distance restraints are included as single gaussian functions (harmonic potentials) (eq. 2), in which the expected distance between two atoms (not bonded between them) and the standard deviation have to be provided. In the case of our models these expected distances come from experimental data. The inclusion of these restraints is performed in the refinement procedure of the program.

The final models are obtained by optimizing the objective function in Cartesian space employing methods of conjugate gradients and molecular dynamics with simulated annealing. Several slightly different models can be calculated by varying the initial structure (Fig. A.3).

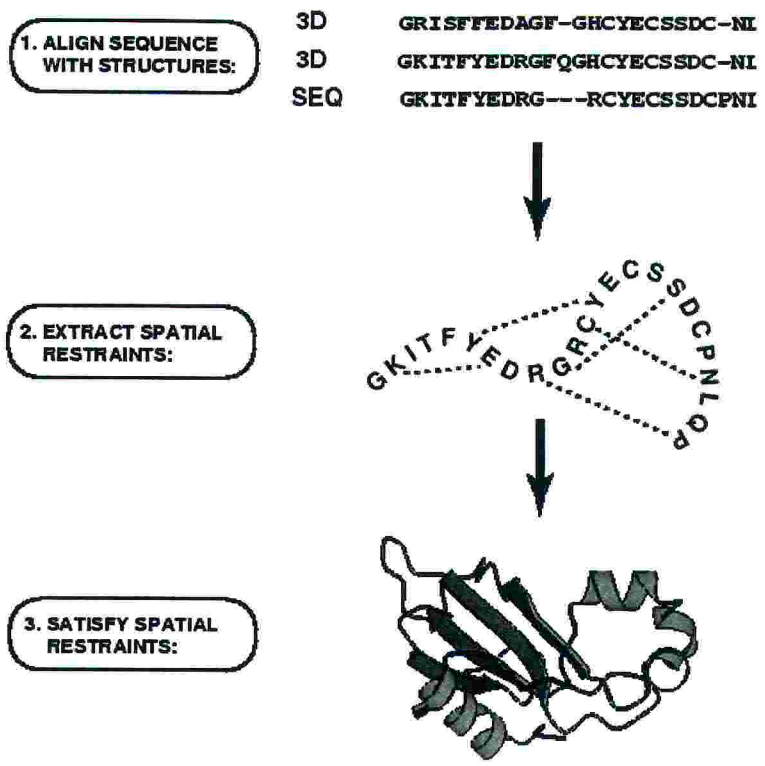


Fig. A.3 .First, the known, template 3D structures are aligned with the target sequence to be modelled. Second, spatial features are transferred from the template to the target. Thus, a number of spatial restraints on its structure are obtained. Third, the 3D model is obtained by satisfying the restraints(Sanchez and Sali, 1997).

A.3.4 Model evaluation

The main criteria for assessing the accuracy of comparative models are: (i) the stereochemistry. This can be investigated by using specific programs (such as PROCHECK (Laskowski et al., 1996), and WHATIF (Vriend, 1990)), which check the structure of chemical bonds, bond angles, aromatic ring planarity, torsion angles, clashes between non-bonded pairs of atoms and the chirality of the C_α atoms; (ii) the comparison with the available experimental data.

Appendix B: Experimental methods.

Site directed mutagenesis experiments were performed in the SISSA laboratories by Monica Mazzolini, Anil Nair and Paolo Codega under the supervision of Prof. Vincent Torre. Although the author did not directly participate in the lab work, experiments and theoretical modeling were part of a unitary project, with comparative modeling and MD simulations giving the gross structural framework for interpretation of the experimental data, while experiments provided tests and refinements of the theoretical hypotheses.
In the Introductory material for HCN and CNG channels we have outlined the basic assumptions underlying the experiments and their structural implications.

B.1 Methods based on cysteine substitution

The study of cysteine mutants by application of thiol specific reagents (substitute cysteine scanning accessibility method or SCAM) has been largely used to infer the residue solvent accessibility in ion channels (Akabas et al., 1992; Lu and Miller, 1995; Becchetti et al., 1999; Wu et al., 2000).
In this approach, the protein residues are systematically replaced by a cysteine and the accessibility of the cysteine reactive thiol group is tested from either side of the membrane with thiol specific reagents like MTSES, MTSET, MTSEA[6], or heavy ions such as Ag^{+} or Cd^{2+}. There are two basic assumptions in SCAM. The first is that only cysteine-substituted residues that are found at a water accessible surface of the protein, will react with the MTS reagents. The second is that cysteine-substituted residues, which react with MTS compounds, produce an irreversible change in channel function, i.e. a change that can be detected by channel activity (current) measurements.
A variant of SCAM consists in the use of divalent cations to assess spatial contiguity between residues (Johnson, Jr. and Zagotta, 2001). In this case it is assumed that, in order to exert a permanent effect on channel function, divalent cations must bind to more than one cysteine-mutated residue at a time (divalent cations are usually tetracoordinated). In the same way, oxidizing reagent such as CuP, which is known to enhance disulfide bond

[6] MTSES, sulfonatoethyl methanethiosulfonate; MTSET, trimethylammonium methanethiosulfonate; MTSEA, ethylammonium methanethiosulfonate

formation, may be used to promote binding of two contiguous cysteine-mutated residues (Mazzolini et al., 2002 in press). In both cases, the application of these reagents to cysteine mutants results in an indirect estimate of the relative residue distance.

B.2 Reagents

In this section we summarize the characteristics of the reagents that were used to probe the pore regions structure: namely MTSET, MTSES, MTSEA, Cd^{2+} and CuP.

MTSET is a bulky thiol specific reagent (approximating to a cylinder with a diameter of 6 Å and a height of 10 Å (Akabas et al., 1992)) and can be used to probe the accessibility to residues.

The divalent cation Cd^{2+} is usually tetracoordinated and binds preferentially to cysteines, histidines and negatively charged residues. Direct inspection of the 3D structure of metallothioneins deposited in the Protein Data Bank (Berman et al., 2000) indicates that the distance between the C_α atoms of two cysteines binding to the same Cd^{2+} ion ranges between 4 and 9 Å (Krovetz et al., 1997; Ermler et al., 1998; Maroney, 1999).

CuP is an oxidizing agent that favours disulfide bridge formation. The distance between the C_α of two cysteines forming a disulfide bond ranges between 4 and 6.5 Å (Srinivasan et al., 1990).

Given the thermal fluctuations of the protein, the maximum distance between the C_α of two cysteines able to form a disulfide bond or to coordinate one Cd^{2+} ion (establishing bonds at the previously reported distances) can be estimated as being around 10 Å (Johnson, Jr. and Zagotta, 2001; Careaga and Falke, 1992; Krovetz et al., 1997; Ermler et al., 1998; Maroney, 1999). Thus, Cd^{2+} and CuP can be used to have a rough evaluation on the residue relative distances.

Note

Part of the work reported in the present thesis has been (or will be) published in the articles listed below.

- Giorgetti, A. and P. Carloni. 2003. Molecular Modeling of Ion Channels: Structural Predictions. Current Opinion in Chemical Biology. 7: 150-156.
- Giorgetti, A., P. Carloni, P. Mistrik and V. Torre. A Homology Model of the Pore Region of HCN channels. Submitted.
- Giorgetti, A., A.V. Nair, P. Codega, P. Carloni and V. Torre. Structural Basis of Gating of CNG channels. In preparation
- Nair, A.V., M. Mazzolini, K. Gamel, P. Codega, A. Giorgetti and V. Torre. Conformational changes in the pore of CNGA1 channels during gating revealed by electrophysiological experiments with mutant channels. In preparation.
- Mazzolini, M., P. Codega, A.V. Nair, A. Giorgetti and V. Torre. Conformational changes in the S6 domain of CNG channels probed by Cd^{2+} blockage. In preparation.

And finally, a paper which is not directly related to ion channels, but in which the same techniques and methodologies were applied in the generation of the models presented:

- Giorgetti, A. and P. Carloni. HIV-1 negative factor binding to human thioesterase II. 2004. Actualités de Chimie Thérapeutique-30^{e} serie. 89-112.

Acknowledgements

I would like to acknowledge GSK for financial support all these years, and also for very useful discussions.

Of course, I wish to thank Paolo and Vincent, who showed me how to work in this fascinating field, in which collaboration between theoreticians and experimentalists is fundamental.

I wish also to thank the 'Zii' (uncles) Michele, Katrin, Lorenzo, Ciras, Ruben and Valentina, Pedro, Andrea, Alessandra and Angelo, because they made us feel like home, and principally, because in these years they were our 'local family'. Hey...they were there when Santiago was born...this was great and very important for Roxana and me.

All the great people from our group: Simone, Claudio, Marco (Berrera and Punta), Pietro and Giovanna, Matteo and Lucia, Kamil, Andrea, Giacomo, Francoise and Juraj. Among them, I wish to say 'gracias' to Sergio, Claudia and Alejandro.

Also 'gracias' to our 'Argentinean' group: Marco, Dani and Marcelo; Agustin, Caro and Marcelo, and last but not least: Eugenio

Reference List

1. Akabas,M.H., D.A.Stauffer, M.Xu, and A.Karlin. 1992. Acetylcholine receptor channel structure probed in cysteine-substitution mutants. *Science* 258:307-310.

2. Altschul,S.F., W.Gish, W.Miller, E.W.Myers, and D.J.Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.

3. Bairoch,A. and B.Boeckmann. 1994. The SWISS-PROT protein sequence data bank: current status. *Nucleic Acids Res.* 22:3578-3580.

4. Barker,W.C., L.K.Ketcham, and M.O.Dayhoff. 1978. A comprehensive examination of protein sequences for evidence of internal gene duplication. *J. Mol. Evol.* 10:265-281.

5. Becchetti,A., K.Gamel, and V.Torre. 1999. Cyclic nucleotide-gated channels. Pore topology studied through the accessibility of reporter cysteines. *J. Gen. Physiol* 114:377-392.

6. 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.

7. Brown,H.F., D.DiFrancesco, and S.J.Noble. 1979. How does adrenaline accelerate the heart? *Nature* 280:235-236.

8. Brown,H. and W.Ho. 1996. The hyperpolarization-activated inward channel and cardiac pacemaker activity. *In* Molecular Physiology and Pharmacology of

Cardiac Ion Channels and Transporters. M.Morad, S.Ebashi, W.Trautwein, and Y.Kurachi, editors. Kluwer, Dordrecht, The Netherlands. 17-30.

9. Cahalan,M.D. and K.G.Chandy. 1997. Ion channels in the immune system as targets for immunosuppression. *Curr. Opin. Biotechnol.* 8:749-756.

10. Capener,C.E., I.H.Shrivastava, K.M.Ranatunga, L.R.Forrest, G.R.Smith, and M.S.Sansom. 2000. Homology modeling and molecular dynamics simulation studies of an inward rectifier potassium channel. *Biophys. J* 78:2929-2942.

11. Careaga,C. and J.Falke. 1992. Thermal motions of surface alpha-helices in the D-galactose chemosensory receptor. *J. Mol. Biol.* 226:1219-1235.

12. Catterall,W.A. 1992. Cellular and molecular biology of voltage-gated sodium channels. *Physiol Rev.* 72:S15-S48.

13. Catterall,W.A. 2000. From ionic currents to molecular mechanisms: the structure and function of voltage-gated sodium channels. *Neuron* 26:13-25.

14. Chang,G., R.H.Spencer, A.T.Lee, M.T.Barclay, and D.C.Rees. 1998. Structure of the MscL homolog from Mycobacterium tuberculosis: a gated mechanosensitive ion channel. *Science* 282:2220-2226.

15. Chen,C.T. and B.Rost. 2002. State-of-the-art in membrane protein prediction. *Applied Bioinformatics* 1:21-35.

16. Clapham,D. 1998. Not so funny anymore: pacing channels are cloned. *Neuron* 21:5.

17. Cserzo,M., E.Wallin, I.Simon, G.von Heijne, and A.Elofsson. 1997. Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. *Protein Eng* 10:673-676.

18. Cuello,L.G., J.G.Romero, D.M.Cortes, and E.Perozo. 1998. pH-dependent gating in the Streptomyces lividans K+ channel. *Biochemistry* 37:3229-3236.

19. Devor,D.C., A.K.Singh, R.A.Frizzell, and R.J.Bridges. 1996. Modulation of Cl- secretion by benzimidazolones. I. Direct activation of a Ca(2+)-dependent K+ channel. *Am. J. Physiol* 271:L775-L784.

20. DiFrancesco,D. 1993. Pacemaker mechanisms in cardiac tissues. *Annu. Rev. Physiol.* 55:455-472.

21. Doupnik,C.A., N.Davidson, and H.A.Lester. 1995. The inward rectifier potassium channel family. *Curr. Opin. Neurobiol.* 5:268-277.

22. Doyle,D.A., C.J.Morais, R.A.Pfuetzner, A.Kuo, J.M.Gulbis, S.L.Cohen, B.T.Chait, and R.MacKinnon. 1998. The structure of the potassium channel: molecular basis of K+ conduction and selectivity. *Science* 280:69-77.

23. Ermler,U., W.Grabarse, S.Shima, M.Goubeaud, and R.Thauer. 1998. Active sites of transition metals enzymes with a focus on nickel. *Curr. Op. Struct. Biol.* 8:749-758.

24. Fain,G., F.Quandt, B.Bastian, and GerschenfeldHM. 1978. Contribution of a cesium-sensitive conductance increase to the rod photoresponse. *Nature* 272:466-469.

25. Faraldo-Gomez,J.D., G.R.Smith, and M.S.Sansom. 2002. Setting up and optimization of membrane protein simulations. *Eur. Biophys. J* 31:217-227.

26. Favre,I., E.Moczydlowski, and L.Schild. 1996. On the structural basis for ionic selectivity among Na+, K+, and Ca2+ in the voltage-gated sodium channel. *Biophys. J.* 71:3110-3125.

27. Flynn,G.E., J.P.Johnson, Jr., and W.N.Zagotta. 2001. Cyclic nucleotide-gated channels: shedding light on the opening of a channel pore. *Nat. Rev. Neurosci.* 2:643-651.

28. Flynn,G.E. and W.N.Zagotta. 2001. Conformational changes in S6 coupled to the opening of cyclic nucleotide-gated channels. *Neuron* 30:689-698.

29. Flynn,G.E. and W.N.Zagotta. 2003. A cysteine scan of the inner vestibule of cyclic nucleotide-gated channels reveals architecture and rearrangement of the pore. *J. Gen. Physiol* 121:563-582.

30. Garcia,E., M.Scanlon, and D.Naranjo. 1999. A marine snail neurotoxin shares with scorpion toxins a convergent mechanism of blockade on the pore of voltage-gated K channels. *J. Gen. Physiol* 114:141-157.

31. Glusker,J. 1991. Structural aspects of metal liganding to functional groups in proteins. *Adv. Protein Chem.* 42:1-76.

32. Guy,H.R. and F.Conti. 1990. Pursuing the structure and function of voltage-gated channels. *Trends Neurosci.* 13:201-206.

33. Hagiwara,S., S.Miyazaki, and N.P.Rosenthal. 1976. Potassium current and the effect of cesium on this current during anomalous rectification of the egg cell membrane of a starfish. *J. Gen. Physiol* 67:621-638.

34. Halliwell,J. and P.Adams. 1982. Voltage-clamp analysis of muscarinic excitation in hippocampal neurons. *Brain. Res.* 250:71-92.

35. Heginbotham,L., T.Abramson, and R.MacKinnon. 1992. A functional connection between the pores of distantly related ion channels as revealed by mutant K+ channels. *Science* 258:1152-1155.

36. Heginbotham,L., L.Kolmakova-Partensky, and C.Miller. 1998b. Functional reconstitution of a prokaryotic K+ channel. *J. Gen. Physiol* 111:741-749.

37. Heginbotham,L., L.Kolmakova-Partensky, and C.Miller. 1998a. Functional reconstitution of a prokaryotic K+ channel. *J. Gen. Physiol* 111:741-749.

38. Heginbotham,L., M.LeMasurier, L.Kolmakova-Partensky, and C.Miller. 1999. Single streptomyces lividans K(+) channels: functional asymmetries and sidedness of proton activation. *J. Gen. Physiol* 114:551-560.

39. Heginbotham,L., Z.Lu, T.Abramson, and R.MacKinnon. 1994. Mutations in the K+ channel signature sequence. *Biophys. J* 66:1061-1067.

40. Henikoff,S. and J.G.Henikoff. 1992. Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. U. S. A* 89:10915-10919.

41. Hille,B. 2001. Ionic Channels of Excitable Membranes. Sunderland,MA.

42. Ho,W., H.Brown, and D.Noble. 1994. High selectivity of the If channels to Na+ and K+ in rabbit isolated sinoatrial node cells. *Pflugers Arch.* 426:68-74.

43. Hodgkin, A. L. and Huxley, A. F. A quantitative description of membrane current and its application to conduction and excitation in nerves. J.Physiol. 117, 500-544. 1952.

Ref Type: Generic

44. Irisawa,H., H.F.Brown, and W.Giles. 1993. Cardiac pacemaking in the sinoatrial node. *Physiol Rev.* 73:197-227.

45. Jacobsen,R.B., E.D.Koch, B.Lange-Malecki, M.Stocker, J.Verhey, R.M.Van Wagoner, A.Vyazovkina, B.M.Olivera, and H.Terlau. 2000. Single amino acid substitutions in kappa-conotoxin PVIIA disrupt interaction with the shaker K+ channel. *J. Biol. Chem.* 275:24639-24644.

46. Jentsch,T.J., W.Gunther, M.Pusch, and B.Schwappach. 1995. Properties of voltage-gated chloride channels of the ClC gene family. *J. Physiol* 482:19S-25S.

47. Jiang,Y., A.Lee, J.Chen, M.Cadene, B.T.Chait, and R.MacKinnon. 2002a. Crystal structure and mechanism of a calcium-gated potassium channel. *Nature* 417:515-522.

48. Jiang,Y., A.Lee, J.Chen, M.Cadene, B.T.Chait, and R.MacKinnon. 2002b. The open pore conformation of potassium channels. *Nature* 417:523-526.

49. Jiang,Y., A.Lee, J.Chen, V.Ruta, M.Cadene, B.T.Chait, and R.MacKinnon. 2003. X-ray structure of a voltage-dependent K+ channel. *Nature* 423:33-41.

50. Johnson,J. and W.N.Zagotta. 2001. Rotation movement during cyclic nucleotide-gated channel opening. *Nature* 412:917-921.

51. Katchalski-Katzir,E., I.Shariv, M.Eisenstein, A.A.Friesem, C.Aflalo, and I.A.Vakser. 1992. Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. *Proc. Natl. Acad. Sci. U. S. A* 89:2195-2199.

52. Krovetz,H., H.VanDongen, and A.VanDongen. 1997. Atomic distances estimates from disulfides and high-affinity metal-binding sites in a K channel poreKrovetz, H. S., H. M. A. VanDongen, and A. M. J. VanDongen. *Biophys. J* 72:117-126.

53. Kuo,A., J.M.Gulbis, J.Antcliff, T.Rahman, E.Lowe, J.Zimmer, J.Cuthbertson, F.M.Ashcroft, T.Ezaki, and D.A.Doyle. 2003. Crystal structure of the potassium Channel KirBac1.1 in the closed state. *Science* 3000:1922-1926.

54. Laio A. and V.Torre. 1999. Physical origin of selectivity in ionic channels of biological membranes. *Biophys. J* 76:129-148.

55. Laskowski,R.A., J.A.Rullmannn, M.W.MacArthur, R.Kaptein, and J.M.Thornton. 1996. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol. NMR* 8:477-486.

56. LeMasurier,M., L.Heginbotham, and C.Miller. 2001. KcsA: it's a potassium channel. *J. Gen. Physiol* 118:303-314.

57. Lipkind,G.M. and H.A.Fozzard. 1994. A structural model of the tetrodotoxin and saxitoxin binding site of the Na+ channel. *Biophys. J.* 66:1-13.

58. Lipkind,G.M. and H.A.Fozzard. 2000. KcsA crystal structure as framework for a molecular model of the Na(+) channel pore. *Biochemistry* 39:8161-8170.

59. Liu,J. and S.A.Siegelbaum. 2000. Change of pore helix conformational state upon opening of cyclic nucleotide-gated channels. *Neuron* 28:899-909.

60. Lu,Z., A.M.Klem, and Y.Ramu. 2001. Ion conduction pore is conserved among potassium channels. *Nature* 413:809-813.

61. Ludwig,A., X.Zong, M.Jeglitsch, F.Hofmann, and M.Biel. 1998. A family of hyperpolarization-activated mammalian cation channels. *Nature* 393:587-591.

62. MacKerell,A.D., D.Bashford, M.Bellot, R.Dunbrack, J.Evanseck, M.Field, S.Fischer, J.Gao, H.Guo, S.Ha, D.Joseph-McCarthy, L.Kuchnir, K.Kuczera, F.Lau, C.Mattos, S.Michnick, T.Ngo, D.Nguyen, B.Pro'hom, I.Reiher, B.Roux, M.Schlenkrich, J.Smith, R.Stote, J.Straub, M.Watanabe, J.Wiorkiewicz-Kuczera, D.Yin, and M.Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* 102:3586-3616.

63. MacKinnon,R. 1991. Determination of the subunit stoichiometry of a voltage-activated potassium channel. *Nature* 350:232-235.

64. MacKinnon,R., S.L.Cohen, A.Kuo, A.Lee, and B.T.Chait. 1998. Structural conservation in prokaryotic and eukaryotic potassium channels. *Science* 280:106-109.

65. MacKinnon,R. and D.A.Doyle. 1997. Prokaryotes offer hope for potassium channel structural studies. *Nat. Struct. Biol.* 4:877-879.

66. Maroney,M. 1999. Structure/function relationships in nickel metallobiochemistry. *Curr. Opin. Chem Biol* 3:188-199.

67. Matulef,K., G.E.Flynn, and W.N.Zagotta. 1999. Molecular rearrangements in the ligand-binding domain of cyclic nucleotide-gated channels. *Neuron* 24:443-452.

68. Meuser,D., H.Splitt, R.Wagner, and H.Schrempf. 1999. Exploring the open pore of the potassium channel from Streptomyces lividans. *FEBS Lett.* 462:447-452.

69. Moran,O. 2001. Molecular simulation of the interaction of kappa-conotoxin-PVIIA with the Shaker potassium channel pore. *Eur. Biophys. J.* 30:528-536.

70. Noma,A. and H.Irisawa. 1976. Membrane currents in the rabbit sinoatrial node cell as studied by the double microelectrode method. *Pflugers Arch.* 364:45-52.

71. Noma,A., M.Morad, and H.Irisawa. 1983. Does the "pacemaker current" generate the diastolic depolarization in the rabbit SA node cells? *Pflugers Arch.* 397:190-194.

72. Pape,H. 1996. Queer current and pacemaker: the hyperpolarization-activated cation currents in neurons. *Annu. Rev. Physiol.* 58:299-327.

73. Penzotti,J.L., H.A.Fozzard, G.M.Lipkind, and S.C.Dudley, Jr. 1998. Differences in saxitoxin and tetrodotoxin binding revealed by mutagenesis of the Na+ channel outer vestibule. *Biophys. J.* 75:2647-2657.

74. Penzotti,J.L., G.Lipkind, H.A.Fozzard, and S.C.Dudley, Jr. 2001. Specific neosaxitoxin interactions with the Na+ channel outer vestibule determined by mutant cycle analysis. *Biophys. J.* 80:698-706.

75. Perozo,E., D.M.Cortes, and L.G.Cuello. 1998. Three-dimensional architecture and gating mechanism of a K+ channel studied by EPR spectroscopy. *Nat. Struct. Biol.* 5:459-469.

76. Perozo,E., D.M.Cortes, and L.G.Cuello. 1999. Structural rearrangements underlying K+-channel activation gating. *Science* 285:73-78.

77. Persson,B. and P.Argos. 1994. Prediction of transmembrane segments in proteins utilising multiple sequence alignments. *J. Mol. Biol.* 237:182-192.

78. Persson,B. and P.Argos. 1997. Prediction of membrane protein topology utilizing multiple sequence alignments. *J. Protein Chem.* 16:453-457.

79. Raman,I.M. and B.P.Bean. 1999. Ionic currents underlying spontaneous action potentials in isolated cerebellar Purkinje neurons. *J. Neurosci.* 19:1663-1674.

80. Ranatunga,K.M., R.J.Law, G.R.Smith, and M.S.Sansom. 2001. Electrostatics studies and molecular dynamics simulations of a homology model of the Shaker K+ channel pore. *Eur. Biophys. J* 30:295-303.

81. Rauer,H., M.D.Lanigan, M.W.Pennington, J.Aiyar, S.Ghanshani, M.D.Cahalan, R.S.Norton, and K.G.Chandy. 2000. Structure-guided transformation of charybdotoxin yields an analog that selectively targets Ca(2+)-activated over voltage-gated K(+) channels. *J. Biol. Chem* 275:1201-1208.

82. Reimann,F. and F.M.Ashcroft. 1999. Inwardly rectifying potassium channels. *Curr. Opin. Cell Biol.* 11:503-508.

83. Richardson,J.S. 1977. beta-Sheet topology and the relatedness of proteins. *Nature* 268:495-500.

84. 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.

85. Roncaglia,P., P.Mistrik, and V.Torre. 2002. Pore topology of the hyperpolarization-activated cyclic nucleotide- gated channel from sea urchin sperm. *Biophys. J.* 83:1953-1964.

86. Rost,B. 2001. Review: protein secondary structure prediction continues to rise. *J. Struct. Biol.* 134:204-218.

87. Rost,B., R.Casadio, P.Fariselli, and C.Sander. 1995. Transmembrane helices predicted at 95% accuracy. *Protein Sci.* 4:521-533.

88. Rothberg,B., K.Shin, P.Phale, and G.Yellen. 2002. Voltage-controlled gating at the intracellular entrance to a hyperpolarization-activated cation channel. *J. Gen. Physiol* 119:83-91.

89. Roux,B. and R.MacKinnon. 1999. The cavity and pore helices in the KcsA K+ channel: electrostatic stabilization of monovalent cations. *Science* 285:100-102.

90. Sanchez,R. and A.Sali. 1997. Advances in comparative protein-structure modelling. *Curr. Opin. Struct. Biol.* 7:206-214.

91. Sander,C. and R.Schneider. 1991. Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins* 9:56-68.

92. Santoro,B. and G.Tibs. 1999. **The HCN gene family: molecular basis of the hyperpolarization-activated pacemaker channels.** *Ann NY Acad Sci* 868:714-764.

93. Schrempf,H., O.Schmidt, R.Kummerlen, S.Hinnah, D.Muller, M.Betzler, T.Steinkamp, and R.Wagner. 1995. A prokaryotic potassium ion channel with two predicted transmembrane segments from Streptomyces lividans. *EMBO J.* 14:5170-5178.

94. Sesti,F., E.Eismann, U.B.Kaupp, M.Nizzari, and V.Torre. 1995. The multi-ion nature of the cGMP-gated channel from vertebrate rods. *J. Physiol* 487 (Pt 1):17-36.

95. Shin,K., B.Rothberg, and G.Yellen. 2001. Blocker state dependence and trapping in hyperpolarization-activated cation channels: evidence for an intracellular activation gate. *J. Gen. Physiol* 117:91-101.

96. Shrivastava,I.H. and M.S.Sansom. 2000. Simulations of ion permeation through a potassium channel: molecular dynamics of KcsA in a phospholipid bilayer. *Biophys. J.* 78:557-570.

97. Singh,S., C.A.Syme, A.K.Singh, D.C.Devor, and R.J.Bridges. 2001. Benzimidazolone activators of chloride secretion: potential therapeutics for cystic fibrosis and chronic obstructive pulmonary disease. *J. Pharmacol. Exp. Ther.* 296:600-611.

98. Sirota,F.L., P.G.Pascutti, and C.Anteneodo. 2002. Molecular modeling and dynamics of the sodium channel inactivation gate. *Biophys. J* 82:1207-1215.

99. Srinivasan,N., R.Sowdhamini, C.Ramakrishnan, and P.Balaram. 1989. Conformations of disulfide bridges in proteins. *Int. J. Peptide Res* 36:147-155.

100. Standen,N.B. and P.R.Stanfield. 1980. Rubidium block and rubidium permeability of the inward rectifier of frog skeletal muscle fibres. *J. Physiol* 304:415-435.

101. Stuhmer,W., F.Conti, H.Suzuki, X.D.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.

102. Sun,Y.M., I.Favre, L.Schild, and E.Moczydlowski. 1997. On the structural basis for size-selective permeation of organic cations through the voltage-gated sodium channel. Effect of alanine mutations at the DEKA locus on selectivity, inhibition by Ca2+ and H+, and molecular sieving. *J. Gen. Physiol* 110:693-715.

103. Tempel,B.L., D.M.Papazian, T.L.Schwarz, Y.N.Jan, and L.Y.Jan. 1987. Sequence of a probable potassium channel component encoded at Shaker locus of Drosophila. *Science* 237:770-775.

104. Terlau,H., A.Boccaccio, B.M.Olivera, and F.Conti. 1999. The block of Shaker K+ channels by kappa-conotoxin PVIIA is state dependent. *J. Gen. Physiol* 114:125-140.

105. Terlau,H., S.H.Heinemann, W.Stuhmer, M.Pusch, F.Conti, K.Imoto, and S.Numa. 1991. Mapping the site of block by tetrodotoxin and saxitoxin of sodium channel II. *FEBS Lett.* 293:93-96.

106. Thoby-Brisson,M., P.Telgkamp, and J.M.Ramirez. 2000. The role of the hyperpolarization-activated current in modulating rhythmic activity in the isolated respiratory network of mice. *J. Neurosci.* 20:2994-3005.

107. Thompson,G.A., M.L.Leyland, I.Ashmole, M.J.Sutcliffe, and P.R.Stanfield. 2000. Residues beyond the selectivity filter of the K+ channel kir2.1 regulate permeation and block by external Rb+ and Cs+. *J. Physiol* 526 Pt 2:231-240.

108. Thompson,J.D., D.G.Higgins, and T.J.Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673-4680.

109. von Heijne,G. 1992. Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *J. Mol. Biol.* 225:487-494.

110. Vriend,G. 1990. *WHAT IF: A molecular modeling and drug design program. J. Mol. Graph* 8:52-56.

111. Webster,S.M., D.del Camino, J.P.Dekker, and G.Yellen. 2004. Intracellular gate opening in Shaker K+ channels defined by high-affinity metal bridges. *Nature* 428:864-868.

112. 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.

113. Wollmuth,L. and B.Hille. 1992. Ionic selectivity of Ih channels of rod photoreceptors in tiger salamanders. *J. Gen. Physiol* 100:749-765.

114. Wrisch,A. and S.Grissmer. 2000. Structural differences of bacterial and mammalian K+ channels. *J. Biol. Chem* 275:39345-39353.

115. Yamamoto,Y., K.Imaeda, and H.Suzuki. 1999. Endothelium-dependent hyperpolarization and intercellular electrical coupling in guinea-pig mesenteric arterioles. *J. Physiol* 514 (Pt 2):505-513.

116. Zagotta,W.N., N.Olivier, K.Black, E.Young, R.Olson, and E.Gouaux. 2003. Stuctural basis for modulation and agonist specificity of HCN pacemakers channels. *Nature* 425:200-205.

117. Zhong,Q., D.M.Newns, P.Pattnaik, J.D.Lear, and M.L.Klein. 2000. Two possible conducting states of the influenza A virus M2 ion channel. *FEBS Lett.* 473:195-198.

118. Zhou,Y., J.H.Morais-Cabral, A.Kaufman, and R.MacKinnon. 2001. Chemistry of ion coordination and hydration revealed by a K+ channel- Fab complex at 2.0 A resolution. *Nature* 414:43-48.