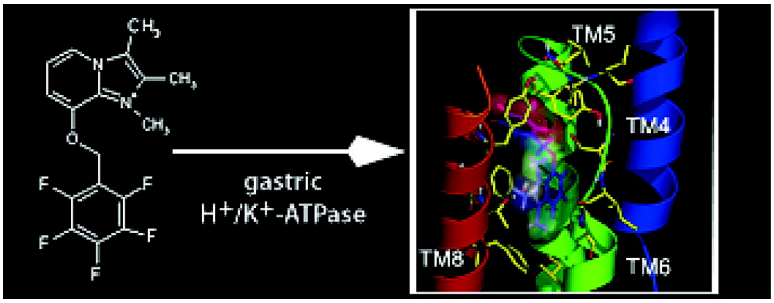
Journal of Medicinal Chemistry

Article

Ligand Docking in the Gastric H/K-ATPase: Homology Modeling of Reversible Inhibitor Binding Sites

Chang G. Kim, Jude A. Watts, and Anthony Watts

J. Med. Chem., **2005**, 48 (23), 7145-7152• DOI: 10.1021/jm050326o • Publication Date (Web): 20 October 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

Journal of Medicinal Chemistry is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Ligand Docking in the Gastric H^+/K^+-ATPase: Homology Modeling of Reversible Inhibitor Binding Sites

Chang G. Kim, Jude A. Watts, and Anthony Watts*

Biomembrane Structure Unit, Department of Biochemistry, Oxford University, South Parks Road, Oxford OX1 3QU, U.K.

Received April 8, 2005

Using the recent high-resolution X-ray structures determined for the Ca^{2+}-ATPase, we have generated two homology models of the gastric H^+/K^+-ATPase reflecting the E_1 and E_2 conformations adopted by P-type ATPases in their catalytic cycle. In regimes where the in situ solid-state NMR-determined structure for 1,2,3-trimethyl-8-(pentafluorophenylmethoxy)imidazo[1,2-*a*]pyridinium iodide (TMPFPIP), a reversible inhibitor of the gastric H^+/K^+-ATPase, was retained in its predefined conformation and was allowed full torsional flexibility in docking, the ligands localized to discrete binding volumes in the E_1 model and to a single central binding space, together with secondary peripheral locations, in the E_2 conformation. The results of these binding studies are in good agreement with current site-directed mutagenesis data and support the suggestion that the binding site is proximal to the loop between TM5 and TM6 and TM8, the transmembrane (TM) region considered important for cation translocation. Furthermore, the results of the simulation with the flexible ligand complement the solid-state NMR structural constraints of this inhibitor when bound in situ to the protein.

Introduction

Integral membrane proteins represent important targets for disease control and management, and they are triggered in an estimated 80% of all cellular communications. P-type ATPases, GTP-binding protein-coupled receptors (GPCRs), and ion channels are all targets for natural ligands as well as drugs.[1] Targeting of these proteins has been important in treating neurological diseases (schizophrenia,[2] depression and migraine,[3] and Alzheimer's and Parkinson's diseases[4]), as well as with physiological ailments such peptic ulceration.[5] Membrane proteins are frequently the end point of a stimulation cascade that results in a tissue-specific response. Where the acidification of the stomach is to be controlled as, for instance, in the management gastroesophagael reflux disease and gastric ulcer disease, it may be possible to interrupt the feeding response pathway at several stages. To minimize any deleterious effects resulting from regulating early steps in a sensory stimulus pathway, it may be more desirable to target the final stage, namely, the gastric H^+/K^+-ATPase, which is the point of acid secretion.

The gastric H^+/K^+-ATPase, together with Na^+/K^+-ATPase and Ca^{2+}-ATPase, is a member of the P-type ATPase family. These proteins engage in a common catalytic cycle with ion translocation coupled to phosphorylation and dephosphorylation of a conserved aspartate residue.[6] Upon phosphorylation, the inward-facing ion-binding site of high affinity [E_1] changes to an outward-facing ion-binding site with low affinity for K^+ [E_2].[7,8] Unlike the Na^+/K^+- and Ca^{2+}-ATPase, the H^+/K^+-ATPase engages in $2K^+/2H^+/1ATP$ electroneutral ion exchange,[9] generating a million-fold H^+-gradient across the mammalian canallicular membrane of the parietal cell and so acidifying the gastric lumen as part of the feeding response.[10]

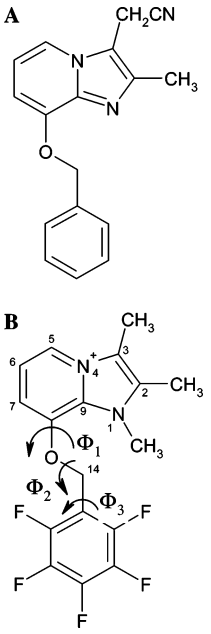


Figure 1. Chemical structure of SCH28080 (A) and its pentafluorinated analogue TMPFPIP (B), which shows the numbering system and definition of torsion angles of SCH28080 derivatives (TMPFPIP).

The gastric H^+/K^+-ATPase has been successfully targeted by substituted benzamidazoles such as omeprazole, lanzoprazole, and pantoprazole, which form a covalent complex with the protein at specific cysteine residues and dominate the pharmaceutical approach to control of gastric acid production. On the other hand, noncovalent inhibitors, such as 1,2,3-trimethyl-8-(pentafluorophenylmethoxy)imidazo[1,2-*a*]pyridinium iodide (TMPFPIP) (Figure 1), a derivative of the imidazo[1,2-

* To whom correspondence should be addressed. Phone: +44 (0) 1865 275268. Fax: +44 (0)1865 275234. E-mail: anthony.watts@bioch.ox.ac.uk.

10.1021/jm050326o CCC: $30.25
Published on Web 10/20/2005

a]pyridine SCH28080 series, are K^+-competitive inhibitors of the H^+/K^+-ATPase.[11] Preincubation of the protein with 2-methyl-3-cyanomethyl-8-(phenylmethoxy)imidazo-[1,2-*a*]pyridine (SCH28080) prevented modification of C813 by omeprazole,[12] and mutation of this residue to threonine (C813T) decreased the affinity for SCH28080 by about 6-fold.[13] Scanning mutagenesis has identified a number of residues in the transmembrane (TM) regions TM4−6 and -8, which display sensitivity to substituted imidazo[1,2-*a*]pyridines (SIPs). These residues are predicted to form a binding domain surface for SIPs to block the ion-translocating pathway.[14]

To achieve high selectivity and efficient inhibition of a specific protein in a process of rational drug design, it is advantageous to possess high-resolution structural information about the potential ligand binding site(s), the effects of catalytic cycle conformational changes, and in particular, details about the in situ structure of the pharmacophore under investigation. Membrane proteins present particular difficulties in determining high-resolution structural information of their active sites or of bound ligands, since the various experimental techniques required for high-resolution data are frustrated by heterogeneity, low levels of expression in the membrane environment, and crystallization difficulties. In addition, large ($M_r \gg 30$ kDa) protein−membrane complexes are generally too large to tumble sufficiently rapidly to produce the averaging effects required for high-resolution nuclear magnetic resonance (NMR) structure determination of protein or ligand by solution-state NMR. The structure of the gastric H^+/K^+-ATPase is poorly defined, being currently limited to a resolution of 8 Å.[15] In silico homology structure modeling may yield viable models of an unknown system upon selection of an appropriate template protein of known structure and sequence alignment between the target and template protein sequences. Once an optimal alignment is obtained, a variety of methods may be employed to generate models of the protein, particularly "modeling by assembly of rigid bodies",[16,17] "modeling by segment matching of coordinate reconstruction"[18,19] and "modeling by satisfaction of spatial restraints".[20−22] The models may then be coupled with knowledge of structural constraints experienced by inhibitors to identify putative binding sites and proposed mechanisms of inhibition.

In this study, recent high-resolution structures of E_1·Ca^{2+}-ATPase (1EUL)[23] and E_2·(TG)-conformer (1IWO)[24] have provided templates with which to generate sequence-threaded models of the H^+/K^+-ATPase in both E_1 and E_2·K^+ conformations. The docking of TMPFPIP, for which structural constraints have been determined in situ,[11] has been simulated with these models of the gastric proton pump. The simulations suggest that there is a diffuse binding site for this SIP on the luminal face of the E_1 conformer of the protein. The binding locations are consistent with existing site-directed mutagenesis (SDM) analysis and occupy volumes proximal to C813 and within the TM1, -2, and -4−9 interhelical spaces. The modeled E_2 conformation proposes a single concerted ligand binding site for TMPFPIP occupying the vestibule created by TM5, -6, and -8, which is the location determined by SDM of several proposed cation-binding residues[14] and a volume that is unavailable to the ligand in the E_1 conformation. The possible roles in attracting and retaining this class of inhibitors at their high-affinity binding site are considered for the residues identified by mutagenesis.

Table 1. Comparison of *G*-factor and Percentage of Residues Located in the Ramachandran Plots That Are Generated by PROCHECK

	E_1 conformation		E_2 conformation	
	template (1EUL) %	model (R3−B6) %	template (1IWO) %	model (E2_08L2) %
core region[a]	78.5	86.8	77.3	87.8
allowed region[b]	18.9	0.5	22.4	11.0
generously allowed region	2.3	1.8	0.3	0.8
disallowed region[c]	0.3	0.9	0.0	0.4
G-factor	0.17	- 0.05	0.11	- 0.16

[a] The core region corresponds to conformations of the polypeptide backbone where there are no steric clashes. [b] The allowed region corresponds to conformations in which shorter van der Waals radii are used in the calculation, that is, the atoms are allowed steric clashes, permitting consideration of left-handed α-helices. [c] Disallowed regions involve steric hindrance between the side chain Cβ methylene group and main chain atoms. Hence it frequently occurs in turn regions of proteins where any other residue would be sterically hindered.

Results and Discussion

Sequence Alignment. The high-resolution structural information available for the E_1 and E_2 conformations of the Ca^{2+}-ATPase has been used as a template for homology modeling of the related gastric H^+/K^+-ATPase. Although overall the sequence identity between the Ca^{2+}-ATPase (P20647) and gastric H^+/K^+-ATPase (P19156) is 27%, the sequence identities in TM4, -5, -6, and -8, the membrane-spanning domains identified by SDM as important in ligand and ion binding, were 41%, 33%, 33%, and 32%, respectively (see Supporting Information, S2). In addition to the identical residues, these regions showed 30%, 25%, 44%, and 14% strong homology, where "strong homology" is determined by the proximity of aromatic residues, aliphaticity, charge, and hydrophilicity. The cytoplasmic regions containing actuator (A), phosphorylation (P), and nucleotide binding (N) domains had sequence identities of 28%, 42%, and 27%, respectively. Proteinase K cleavage sites K172 and E288 and the phosphorylation motif DKTGTLT beginning at D386, are well conserved. The PEGL motifs are known to play a central role in energy transduction and are aligned exactly. In the transmembrane region, amino acid sequences suggested to be cation binding sites (i.e., N793, E796, T824, D825, and E937 for cation binding site I and E309, T824, and D825 for cation binding site II) are also well aligned. Comparison of the boundaries of putative transmembrane helices reveals large changes in the transmembrane residue distribution during the transition between the E_1 and E_2 catalytic states.

Model Evaluation. PROCHECK[25] compares the orientation of residues in the model to those permitted by Ramachandran plots, $\chi-\chi^2$ plots, main- and side-chain torsions, and residue properties. The PROCHECK *G*-factor ranks values above −0.5 as positive candidates for homology models. Table 1 shows the summary of the *G*-factors and the percentages of the residues on the Ramachandran plot for the selected models R3−B6 (E_1 conformation) and E2_08L2 (E_2 conformation) with

Table 2. Comparison of ProsaII3.0 *Z*-Scores between Template and E_1 and E_2 Conformation Models

protein molecule	Z_p			Z_1		
	combined	pair	surface	combined	pair	surface
			E1			
template (1EUL)	−14.38	−12.30	−7.80	−3.79	−4.07	−3.85
model (R3−B6)	−9.50	−7.99	−5.04	−3.90	−3.94	−3.79
			E2			
template (1IWO)	−14.18	−12.40	−7.35	−3.79	−4.07	−3.85
model (E2_08L2)	−9.14	−7.36	−5.19	−3.90	−3.94	−3.79

those of the templates. Here, models generated by MODELLER,[22] based on the CLUSTALW[26] alignment, had values exceeding −0.5 and ranged from −0.34 to −0.05. The *G*-value of the selected E_1 conformer was −0.05 and E_2 conformer was −0.16. A low *G*-factor corresponds to a low probability of a given conformation. The *G*-values indicated here predict that the secondary structure of the models resides in the high-probability regime for $\chi-\chi^2$ values. As such, there are few prohibited conformations adopted by residues in the models, and the models may therefore be regarded as structurally realistic.

ProsaII3.0[27] *Z*-scores indicate the compatibility between the model sequence and structure and those of the template. Where *Z*-scores are similar, the model adopts an electrochemical surface similar to that of the template protein. Z_p relates to the protein *in tota*, whereas Z_1 relates to a fragment of the protein with the lowest energy. The *Z*-scores of the models and their templates are shown in Table 2. Here, the Z_p-scores for the E_1 (−9.50) and E_2 (−9.14) models show reasonably good compatibility with their template comparisons (Z_p = −14.38 and −14.18 for E_1 and E_2 conformers, respectively). The differences in *Z*-scores between template and model are to be expected, however, since the overall identity between Ca^{2+}- and H^+/K^+-ATPase proteins extends to only 27% of the sequence. The resultant variation from the nonhomologous population of residues will necessarily produce peptides with varied electrochemical profiles. The Z_1-scores for template and model are rather closer, however, indicating that the lowest energy-state fragments of the proteins are sufficiently similar (Z_1 = −3.79 (models) and −3.90 (templates)) for the models to be retained. The lowest energy states of the proteins do not appear to change, and this suggests that there is a region common to both conformers that does not undergo a conformational change in the E_1 and E_2 forms. Prosa2003 energy graphs were determined as a function of residue with a window size of 3 residues to examine the quality of the model in the regions of particular interest. Figure 2 shows the combined surface and pairing energy scores determined for template and model in E_1 and E_2 conformations for the transmembrane regions defined by TM4−6 and -8, which are of particular interest in ligand binding. The quality of the model is generally no worse than the template in TM4 and -8 for both conformations of the protein. There is unfavorable combined energy associated with residue L345, but SDM studies have not reported this amino acid as being involved in ligand binding. In the TM5,6 region, residues I784 and N793

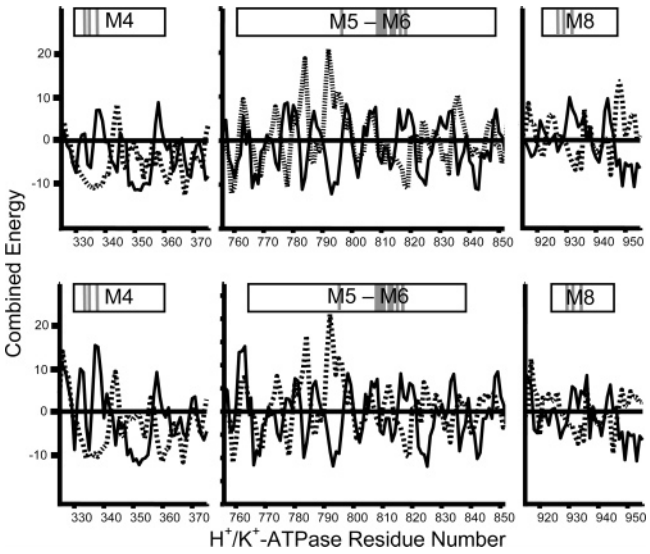


Figure 2. Comparison of Prosa2003 combined surface and pairing energy plots examined as a function of residue (window size of three residues) between template (solid line) and model (dashed line) in the E_1 (top) and E_2 (bottom) conformations, highlighting the proposed transmembrane regions M4, M5−6, and M8, which are indicated to be influential in binding of substituted imidazopyridines. In the model E_1 conformation, M4 is defined by residues 328−363, M5−6 by residues 762−850, and M8 by residues 923−948. In the E2 conformation, M4 is defined by residues 328−364, M5−6 by residues 766−835, and M8 by residues 924−945, as indicated above each plot. Positions of residues implicated in ligand binding by SDM are given in gray.

are calculated as being strained according to the Prosa2003 analysis. I784 is not reported as playing a role in ligand or cation binding, whereas N793 has been implicated in cation but not ligand binding. In M8, residue K948 is calculated as being unstable. This residue is at the terminal end of the suggested TM8 and is distal from the ligand binding site.

PROFIT[28] calculates the root-mean-square deviation (RMSD) between the backbone atoms of the template and those of the model. Here, the RMSDs between the transmembrane region of template and the E_1 and E_2 models were 4.2 and 4.3 Å, respectively. The RMSDs between TM4−6 and -8 of the template and of the E_1 and E_2 models were 2.5 and 2.4 Å, respectively. The transmembrane region defined by TM4−6 and -8 has identity ranging between 32% and 41%. These values are consistent with those for a template−model pair with this level of sequence identity[29] (http://www.expasy.org/swissmod/SM_LikelyPrecision.html).

Visual inspections were performed using PyMol. The directions of the residues, especially surrounding the cation biding sites, were examined to determine whether the side chains faced toward the cation binding sites. Thus, residues such as F332, L816, E820, and E936, which have been identified by SDM as important for cation[30] and SCH28080[31] binding, should face toward the interior of the region bounded by TM4−6 and -8. In the E_1 conformation, the −OH groups of E795, T823, and E936 face the proposed ion binding site I, located between TM5 and TM6, and the groups of E343 and D824 face toward the proposed ion binding site II, proximal to TM4. The locations of these side chains in the modeled system were thus consistent with the

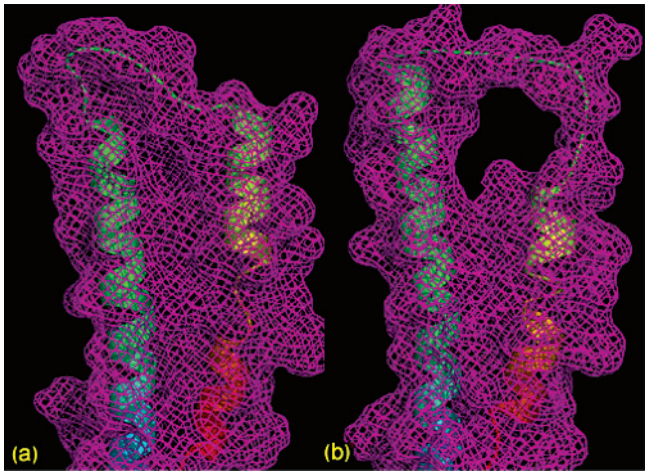


Figure 3. Effect of conformational change on the TM5−6 loop region of the gastric H^+/K^+-ATPase. The transition from E_1 to E_2 results in a rotation of the TM6 helix of the E_2 conformer such that there is an increase in the volume available beneath the TM5−6 loop. Panels a and b represent the side view of the E_1 and E_2 conformations, respectively, from the TM4 direction. It is proposed that this conformational change reveals a binding volume, which accommodates the K^+-competitive SCH28080 analogue, TMPFPIP, proximal to the cation translocation domain of the protein.

functions attributed to these groups by the SDM experiments, and reinforced the confidence in the in silico candidates as models of the protein in its catalytic conformations.

The structural differences between the E_1 and E_2 models of the gastric H^+/K^+-ATPase are consistent with the major structural rearrangements noted between the E_1 and E_2 states of the Ca^{2+}-ATPase.[23,24] Large structural changes are observed in both template and model systems in the cytoplasmic domains, especially in the N domain. There are also conformational changes that occur in the transmembrane domain, where SCH28080 and its reversible inhibitor analogues are proposed to bind to the proton pump.[14] Notably, TM6 and -8 rotate about their helical axes to lengthen the TM5−6 interhelical loop in the E_2 conformation, which results in an increase in the space beneath the loop region and the putative ion-binding sites. Figure 3 shows an intramembranous volume, which results upon conformational change. The directions of residues E343, N792, E795, T823, D824, and E936 also change so as to generate additional volume in the interhelical space, as well as creating a hydrophobic environment around a putative inhibitor binding pocket. These residues correspond to those in the E_2 conformer of the Ca^{2+}-ATPase that interact with Ca^{2+} ions and are directed to the proposed cation binding site. By analogy with the Ca^{2+}-ATPase, cations bound to the E_1 conformer could be released to the extracellular regions during the transition of the E_1 to E_2 catalytic states, with an accompanying directional change of the residues.

Access to the Central Binding Site upon Conformational Change. In shifting from the E_1 to E_2 conformations, it is apparent in the models that TM6 and -8 rotate and the TM5−6 loop elongates (Figure 3). In SDM studies, mutation of L809F and L811F, which reside on the TM5−6 loop and face toward the core of the transmembrane domain, resulted in a very large increase of inhibition constants to 6.15 and 0.625 mM, respectively, whereas L809V and L811V resulted in a relatively small increase of inhibition constants to 288 and 95 nM (WT 60−100 nM).[14] Thus, the volume afforded by the E_1−E_2 transition may permit the reversible inhibitor access to the high-affinity cation-competitive location in the membrane interior from its low-affinity binding positions on the extramembranous face of the protein.

Protein−Ligand Docking. Docking Simulations with the E_1 Conformation of the Gastric H^+/K^+-ATPase. A docking simulation was performed using AUTODOCK3.0 on the E_1 model R3−B6 with a predefined conformation (bowed ligand (Φ_1 (165°), Φ_2 (180°), and Φ_3 (0°)) of TMPFPIP (Figure 1), which is one of the SCH28080 derivatives whose bound structure had been determined by solid-state NMR.[11] The ligand binds at a position in the extracellular domain with an estimated inhibition constant (K_i) of 490 nM at 298 K, which corresponds to a final docking energy from simulation of −8.6 kcal/mol. The inhibition constant is high when compared to the experimentally determined value of 60−100 nM.[33,34] This value is acceptable however, when one considers that the high-affinity binding site for SCH28080 exists in the E_2−P conformation.[6,33] The simulated ligand binds at the TM5−6 loop proximal to C813, a residue previously demonstrated to react with the substituted benzamidazole, omeprazole.[35] This localization is also consistent with the observation that preincubation of the protein with substituted imidazopyridines prevents omeprazole binding at this position.[12] Furthermore, the ligand appears to occlude the K^+ ion pathway formed by TM4, -5, -6, and -8.

When the ligand was permitted flexibility about its three variable torsion angles (Figure 1), the TMPFPIP ultimately clustered to four locations (Figure 4). The lowest-energy ligand in each cluster was selected as representing the local population. The binding conformations varied in docking energy from −9.09 to −8.66 kcal/mol, which corresponded to K_i for the ligand ranging from 207 to 456 nM. The simulated torsion angles (Φ_1, Φ_2, and Φ_3) of the ligands (L_1−L_4) were L_1 (103°, 94°, 75°), L_2 (93°, 134°, 72°), L_3 (−102°, 137°, 65°), and L4 (101°, 106°, −2°), respectively. When compared to the torsions angles determined by the NMR experiments, none of the ligand conformations simulated to dock with the E_1 form of the ATPase is within the range of the experimental data (Φ_1 (165° ± 15°), Φ_2 (180° ± 30°), and Φ_3 (0° ± 60°)).[11] The inhibitor is K^+-competitive and binds with high affinity to the E_2 conformation, which has a low-affinity binding site for K^+ [6,7]. As such, the experimental data therefore refers to the conformation of the ligand bound to the phosphorylated E_2 conformation of the protein, and a variation in the simulation that gives torsion angles dissimilar to the experimental torsion angles is to be expected. Although the simulated bound ligands have several conformations, their binding clusters are consistently proximal to the TM5−6 loop and to the cation binding sites that become accessible from the extracellular face of the protein upon the transition from E_1 to E_2 of the protein. The fixed-ligand and free-ligand simulations gave similar results for the low-affinity ligand binding sites and

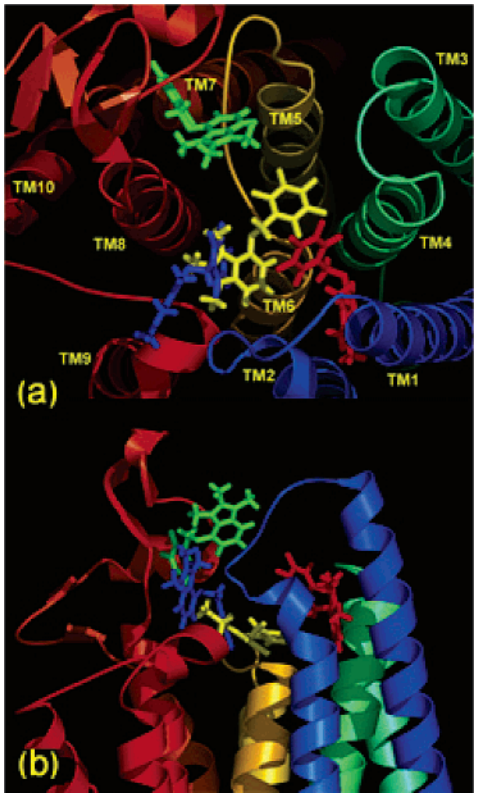


Figure 4. Representations of the top (a) and side view (b) of ligand (TMPFPIP) binding site in the gastric H^+/K^+-ATPase E_1 conformer. The ligand does not occupy a single concerted binding location but rather resides in a diffuse cluster at the interface between the transmembrane domain and the extracytoplasmic face of the protein at the luminal ends of transmembrane helices TM1, -2, -5, and -6−9. The ligand positions represent the lowest energy conformations of a series of clusters. The transmembrane domain color grades from blue to red from TM1−TM10.

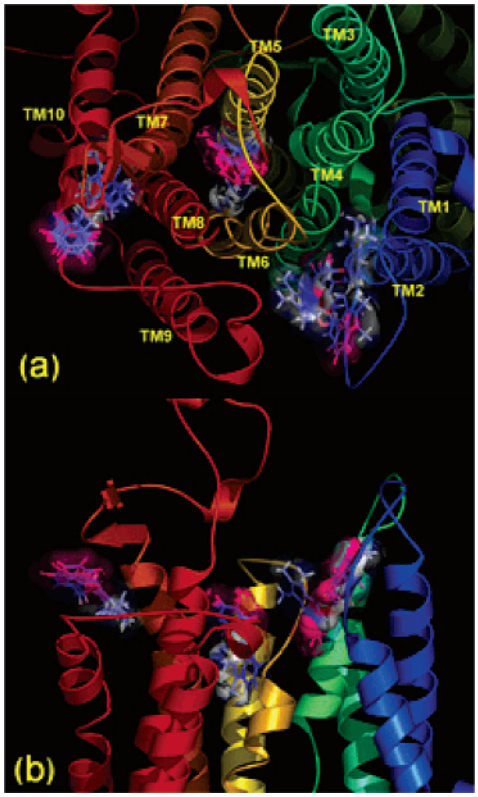


Figure 5. Top view (a) and side view (b) of ligand (TMPFPIP) binding site in the gastric H^+/K^+-ATPase E_2 conformer. The ligand occupies three binding locations: a central binding site defined by transmembrane helices TM4−8, a peripheral diffuse site at the membrane interface bounded by the termini of TM1, -2, -4, -6 and -9, and a low-population peripheral site bounded by the termini of TM7−10.

support the contention that the model of the E_1 conformer is realistic.

Docking Simulations with the E_2 Conformation of the Gastric H^+/K^+-ATPase−A Central Binding Site. Docking simulations between TMPFPIP and the E_2 conformation of the gastric H^+/K^+-ATPase were performed in the same way as those of the E_1 conformer. The results are shown in Figure 5. The bowed fixed-angled ligand binds with an estimated K_i of 99 nM at 298 K and final docking energy of −9.6 kcal/mol. In the case where the ligand was permitted flexibility in its variable bonds, K_i was 22 nM at 298 K and the final docking energy was −10.44 kcal/mol. These values of K_i are in good agreement with the experimental data (60−100 nM). The final torsion angles for the flexible ligand were 111° ± 31°, 122° ± 38°, and 65° ± 25° and are generally consistent with experimentally determined data (165° ± 15°, 180° ± 30°, 0° ± 60°),[11] with overlap in the uncertainties of measurement. Minor deviations in ligand structure may account for the difference in inhibition constant between that of SCH28080, that of TMPFPIP modeled here, and that expected for TMPFPIP in vitro. Since binding affinity is frequently intimately related to conformation,[33] the high-affinity binding conformation (111° ± 31°, 122° ± 38°, and 65° ± 25°) of the fluorinated ligand in this study may reflect the modest deviation in structure from that determined previously (165° ± 15°, 180° ± 30°, 0° ± 60°) by solid-state NMR methods.[11]

Here, the flexible ligand ultimately adopts a bowed conformation in the high-affinity binding site. The central binding location is the same under both docking regimes and is shown in Figure 6. The SCH28080 analogue TMPFPIP appears to occupy a binding pocket consisting of L809, P810, L811, and C813, which are on the TM5−6 loop, I814, I816, and F818 in TM6, and Y925, T929, and F932 on TM8, all of which were identified as sensitive to SCH28080 binding either by engaging with the ligand or in maintaining the binding volume.[31,33] The phenyl moiety of the substituted imidazopyridine is permitted some space to move and is directed toward the extracellular space. As suggested by SDM studies, the imidazopyridine moiety is proximal to F332, M334, V337, and A335 in TM4.[34,36] In this model, the phenyl moiety of F332 points away from the ligand into the TM4−5 interhelical space. Similarly, M334 and V337 face the TM1−2 interhelical space.

In TM4, M334 and V337 face the TM1−2 interhelical space, however, and it is likely that changes in chain length of the side chains result in indirect effects on the binding pocket by distorting of the interhelical space in this region, probably forcing TM4 into the binding volume that would otherwise be readily occupied by the ligand. Similarly, large changes were seen in SDM studies with other proximal residues such as E795 of TM5[30] and L809, L811,[14] and C813 of the TM5−6 loop.[13] Any negative charge carried by E795, the carboxyl group

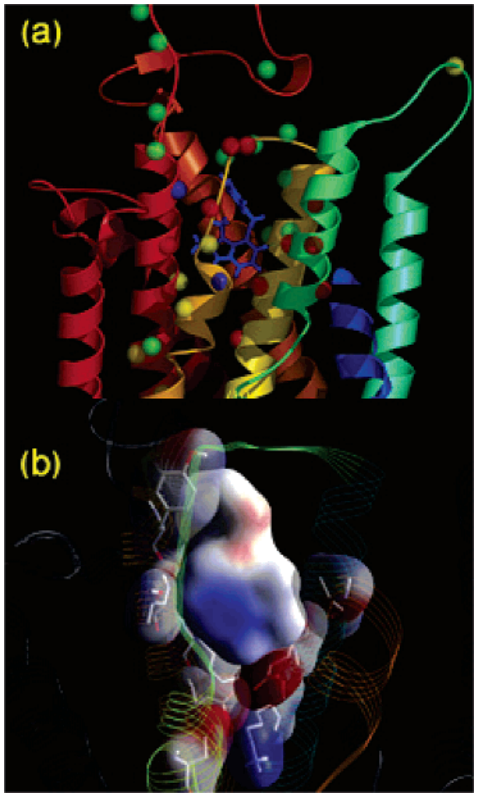


Figure 6. Side view of the central binding site of TMPFPIP modeled in the gastric H^+/K^+-ATPase E_2 conformer (a). The red, blue, yellow, and green colors of the spheres indicate where the mutant of the indicated residue results in an inhibition constant K_i over 500, 300, 200, and 100 nM, respectively. Close-up view of the TMPFPIP bound to its central inhibitory site with electropotential surface overlay (b). Residues which, upon mutation, experienced a change in inhibition constant K_i to over 500 nM are indicated in stick form. Electron-dense areas are indicated by the red electropotential surface. The schematic representation shows that while the imidazopyridine part of TMPFPIP is tightly caught by Y928, T929, and F932 of TM8, I820 and I816 of TM6, and Y802 of TM5, its phenyl ring part has some space to move. Furthermore, the electron-poor imidazopyridine moiety appears to be stabilized by proximal electron-rich residues.

of which models here as lying within 5 Å of the electron-poor N4, may stabilize the active protonated SCH28080 or the imidazo N4 anion, which arises upon methylation of the N1 to the quaternary ammonium form.

Within the TM5−6 loop, only the side chains of P810 and C813 are directed toward this ligand. Accordingly, mutagenesis of L809 and L811 may result in a steric deformation of the loop gate that is revealed upon transition to the E_2 state, rather than affecting electronic interaction with the inhibitor.

In TM6, I814 is directed toward the TM1−6 interhelical space, but away from the central binding location, suggesting that its mutation may result in a steric distortion of the core-region interhelical volume. However, the δ-carbon of I816 may lie as close as 3.6 Å to N1 of the imidazopyridine and thus engage the conjugate ring system of the ligand more directly. In TM8, the aromatic ring of the SIP appears to be stabilized by Y925, whereas T929 and F932 engage the ligand at its central binding site via the imidazopyridine delocalized ring system proximal to the electron-poor N4. This central inhibitor binding site is marginally above a proposed cation binding site defined by E796, D825, and E939. Indeed, the C5 hydrogen is 3.7 Å from the delocalized carboxyl of E795. The planar imidazopyridine moiety is almost immobilized in the vestibule created by these helices. These observations are consistent with ^{2}H-observed solid-state NMR data, which indicated that the imidazopyridine moiety is motionally restricted on the NMR time scale in its binding site.[11]

Mutants I816L in TM6 and Y925F in TM8 experienced an increase in K_i for SCH28080 to over 300 nM.[33] Thus, steric rearrangement of the nonpolar side chain at I816 appears to have a volumetric influence on the binding space available to the delocalized imidazo ring system of the inhibitor. Residues I814 and F818 in TM6 and F932 in TM8, when mutated, showed a modest decrease in sensitivity to SCH28080 to exceed 200 nM. In the model presented here, I814 and F818 face away from the central binding site and F932 is not in van der Waals contact with the ligand, possibly accounting for the diminished effects of their mutation. Nonetheless, the preponderance of SIP-sensitive residues in this region favors the suggestion that the binding model of this class of reversible inhibitors is realistic and that the high-affinity binding site is revealed to the ligand upon structural rearrangement of the protein to its E_2 conformation. In particular, the central binding site appears to be defined by hydrophobic residues and a bank of aromatic groups provided by tyrosine (Y802, Y925, Y928) and phenylalanine (F932) residues in TM8 (Figure 6b). These groups may provide opportunities for $\pi-\pi$ stacking arrangements or $\pi-$cation interactions between the protein and the SIP, which has been observed by solid-state NMR as changes in the chemical shift of C14.[37] The ultimate electrostatic draw for the inhibitor is the negative charge provided by glutamate residues (E343 and E820) residing at the cytoplasm-terminal ends of TM4 and -6 (Figure 6).

Docking Simulations with the E_2 Conformation of the Gastric H^+/K^+-ATPase−Two Peripheral Binding Sites. Autodock simulations indicated two clustered binding locations for the E_2 ATPase conformation for TMPFPIP other than the one that is proximal to the proposed cation binding region (Figure 5). These clusters exist at the luminal interfacial region at the TM1−2 loop, TM4, and the TM7−10 interhelical space. The first cluster is dispersed with no concerted binding location or structure predicted for the ligand. Similar residue types are suggested to be present as those that constitute the central core binding pocket. Thus, F124, D137, L142, Y170, and L811 appear to define what may be a lower-affinity binding region similar to that found with the E_1 conformer of the protein (Figure 5). The omeprazole binding site, which is close to the central binding site and includes C813,[35] is also proximal to one of the outlying proposed positions for the ligand modeled in this region. As such, the binding site may represent a region of preferential association between the ligand and protein, namely, a region that localizes the ligand close to the cation-translocation pathway, which becomes available to the ligand upon conformational change of the protein to the E_2 conformer. The second smaller cluster is defined by TM7−10 residues A883, W887, L891, C892, C927, Q996, and W997. The effects of these mutations have, to date, been little reported in

the literature. However, neighboring Y925, T929, F932, M937, and I940 were found, upon mutation, to increase K_i for SCH28080 to between 200 and 500 nM,[33] as discussed above, but the side chains of these residues all engage with the central binding site in this model, rather than directing toward this alternative peripheral site. The existence of peripheral binding sites in addition to a K^+-competitive inhibitory site is in good agreement with the estimation of 2.2 ligand binding sites per phosphorylation site.[38] It is possible therefore that a centrally bound inhibitor blockades the cation-binding region of the H^+/K^+-ATPase and prevents a catalytic rearrangement of the protein to its E_1 cycle conformation. This blockade is assisted by peripherally bound ligands, which are either retained at the membrane interface to replace ligand departing its central site or engaged in further steric challenges to protein rearrangement.

Conclusions

In silico models have been produced for the gastric H^+/K^+-ATPase based on templates provided by its related P-type ATPase, the Ca^{2+}-ATPase. Using MODELLER software, we generated candidate models of the H^+/K^+-ATPase in the E_1 and E_2 conformations of its catalytic cycle. Evaluation of these models using quality parameters assessed by PROCHECK, ProsaII3.0, and PROFIT and correlation between the results of ligand docking and existing mutagenesis information for the protein showed that the models are realistic and could reveal an insight into the binding mechanism for a class of site-specific reversible inhibitors of the gastric H^+/K^+-ATPase. It is proposed that these inhibitors are first localized to low-affinity binding sites on the extracellular face at the TM5, -6, and -8 interface with the protein in the E_1 conformation. Upon conformational change to the phosphorylated E_2 conformer, there is a dramatic change in the intramembranous volume available to the ligand, which results from the rotation of TM6 and -8 and the extension of the TM5−6 interhelical loop. A central binding pocket is revealed that stabilizes the imidazopyridine derivative, whereby the planar conjugate ring system is tightly bound within the interior of transmembrane domain. The phenyl derivative is permitted a degree of motional freedom beneath the TM5−6 loop. The proposed binding conformation of a bowed ligand is supported by existing high-resolution solid-state NMR data,[11,37] and the location of the proposed binding site is supported by existing site-directed mutagenesis studies. The site is sufficiently proximal to the putative sites of cation transport to explain the K^+-competitive behavior of this class of inhibitor. It is suggested that these reversible inhibitors bind this important target in the treatment of gastric ulcer and gastroesophagael reflux diseases by preventing the gastric H^+/K^+-ATPase from returning to the E_1 conformation and completing its catalytic cycle and maintaining the protein in its E_2 conformation until the inhibitor is released by K^+-competition or entropic drive.

Materials and Methods

Model Generation. Multiple sequence alignment between the gastric H^+/K^+-ATPase (gHKA, P19156) and Ca^{2+}-ATPase (1EUL, P20647) was carried out using CLUSTALW, v1.81.[26] The parameters were chosen with the default values provided by the CLUSTALW server, but the gap extension penalty was reduced to avoid long consecutive gaps in the alignment (Gap open penalty was set at 10, Gap extension penalty was 0.05, and BLOSUM62 position specific weight matrix). Models were generated as pdb files using MODELLER6v2 using distance and dihedral angle restraints on the target sequence derived from its alignment with the template structure. Initially, 30 models were generated for both E_1 and E_2 protein conformer using MODELLER6v2. Five loop-optimized models were generated for each pregenerated model using a loop optimization routine in MODELLER6v2. Candidates were evaluated using PROCHECK and ProsaII3.0. Side chain directions were inspected visually using PyMol to determine whether there were any violations of the experimentally determined data. PROCHECK produced a *G*-factor, which assessed the probability of the conformation having permitted stereochemistry by reference to a Ramachandran plot, Gly and Pro Ramachandran plot, $\chi_1-\chi_2$ plots, main-chain parameters, and residue properties. Models producing high *G*-factors were selected. ProsaII3.0 assessed the compatibility between the sequence and structure of a model, and the sequence and structure of the template to produce a *Z*-score. Structures were selected by closeness of model and template *Z*-scores. PROFIT determined the RMSD between the model and the template. Candidates that had a low (<5 Å) RMSD were selected. The model that ranked the highest under these evaluation criteria was used in subsequent ligand-docking simulations.

Ligand Docking. The K^+-competitive inhibitor SCH28080 analogue TMPFPIP was generated as a pdb file and imported into AUTODOCK3.0. Ligand docking to both E_1 and E_2 gastric H^+/K^+-ATPase models was performed using rigid bowed ligands, which represented the conformations of this ligand as determined previously by REDOR NMR ($\Phi_1 = 165°$, $\Phi_2 = 180°$, and $\Phi_3 = 0°$)(Φ_1, Φ_2 and Φ_3, Figure 1),[11] or using TMPFPIP, which was permitted full torsional flexibility about its three variable bonds. Docking between ligand and protein was simulated for free ligand starting from distal "lumenal" positional grid. Docking behavior was predicted by a linear regression analysis in an AMBER force field and a search method using a Lamarckian genetic algorithm.[32] For the fixed-conformation ligand, a population size of 50 with 250 000 energy evaluations, 27 000 generations, a gene mutation rate of 0.02, and a crossover of 0.8 were used. Each docking simulation contained 200 iterative steps to generate clusters of ultimate docking positions with a positional RMSD tolerance of 3 Å. For the flexible ligand, a population size of 50 with 90 million energy evaluations, 90 000 generations, a gene mutation rate of 0.02, and crossover of 0.8 were used. Each docking simulation contained 3000 iterative steps to generate clusters of ultimate docking positions with a positional RMSD tolerance of 2 Å. The partial charge of ligands was calculated by WebLab ViewerPro. Ligands were initially placed in arbitrary positions restricted to the volume above the extracellular domain of the gastric H^+/K^+-ATPase.

Docking energy and inhibition constants were calculated in AutoDock according to the relations:

$$\Delta G_{\text{obs}} = \Delta G_{\text{vdW}} \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) + \Delta G_{\text{hbond}} \sum_{i,j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{6}} \right) + \Delta G_{\text{elec}} \sum_{i,j} \frac{q_i q_j}{\epsilon(r_{ij}) r_{ij}} + \Delta G_{\text{tor}} N_{\text{tor}} + \Delta G_{\text{sol}} \sum_{i,j} (S_i V_j + S_j V_i)\, e^{(-r_{ij}^2/(2\sigma^2))} \quad (1)$$

and

$$\Delta G_{\text{obs}} = RT \ln K_i \quad (2)$$

where ΔG terms on the right-hand side of eq 1 are empirically calculated using linear regression analysis. The final ligand−protein complexes were visualized using Accelerys DS ViewerPro 5.0.

Acknowledgment. This project has been funded in part by the MRC, BBSRC, Magnex Scientific Ltd, Varian Inc., and CJ Corp. (Korea).

Supporting Information Available: Details for molecular modeling coordinates in pdb format and sequence alignment between gastric H^+/K^+-ATPase and Ca^{2+}-ATPase. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Watts, A.; Burnett, I. J.; Glaubitz, C.; Gröbner, G.; Middleton, D. A.; Spooner, P. J. R.; Watts, J. A.; Williamson, P. T. F. Membrane Protein Structure Determination by Solid State NMR. *Nat. Prod. Rep*. **1999**, *16*, 419−423.

(2) Dean, B.; Opeskin, K.; Pavey, G.; Hill, C.; Keks, N. Changes in Protein Kinase C and Adenylate Cyclase in the Temporal Lobe From Subjects With Schizophrenia. *J. Neural Transm.* **1997**, *104* (11−12), 1371−1381.

(3) Brys, R.; Josson, K.; Castelli, M. P.; Jurzak, M.; Lijnen, P.; Gommeren, W.; Leysen, J. Reconstitution of the Human 5-HT-(1D) Receptor-G-Protein Coupling: Evidence For Constitutive Activity and Multiple Receptor Conformations. *Mol. Pharmacol*. **2000**, *57* (6), 1132−1141.

(4) Pronin, A. N.; Morris, A. J.; Surguchov, A.; Benovic, J. L. Synucleins Are a Novel Class of Substrates For G-Protein-Coupled Receptor Kinases. *J. Biol. Chem*. **2000**, *275* (34), 26515−26522.

(5) Sachs, G.; Carlsson, E.; Lindberg, P.; Wallmark, B. Gastric H,K-ATPase as Therapeutic Target. *Annu. Rev. Pharmacol. Toxicol.* **1988**, *28*, 269−284.

(6) Walderhaug, M. O.; Post, R. L.; Saccomani, G.; Leonard, R. T.; Briskin, D. P. Structural Relatedness of Three Ion-Transport Adenosine Triphosphatases Around Their Active Sites of Phosphorylation. *J. Biol. Chem*. **1985**, *260* (6), 3852−3859.

(7) Sachs, G., Shin, J. M.; Briving, C.; Wallmark, B.; Hersey, S. The Pharmacology of the Gastric Acid Pump: the H^+, K^+ ATPase. *Annu. Rev. Pharmacol. Toxicol.* **1995**, *35*, 277−305.

(8) Munson, K.; Lambrecht N.; Sachs G. Analysis of the Membrane Domain of the Gastric H^+/K^+-ATPase. *J. Exp. Biol.* **2000**, *203* (1), 161−170.

(9) Lee, J.; Simpson, G.; Scholes, P. An ATPase From Dog Gastric Mucosa: Changes of Outer pH in Suspensions of Membrane Vesicles Accompanying ATP Hydrolysis. *Biochem. Biophys. Res. Commun*. **1974**, *60* (2), 825−832.

(10) Wolosin, J. M. Ion Transport Studies With H+-K+-ATPase-Rich Vesicles: Implications for HCl Secretion and Parietal Cell Physiology. *Am. J. Physiol*. **1985**, *248* (6, Part 1), G595−607.

(11) Watts, J. A.; Watts, A.; Middleton, D. A. A Model of Reversible Inhibitors in the Gastric H^+/K^+-ATPase Binding Site Determined by REDOR NMR. *J. Biol. Chem.* **2001**, *276* (46), 43197−43204.

(12) Hersey, S. J.; Steiner, L.; Mendlein, J.; Rabon, E.; Sachs, G. SCH28080 Prevents Omeprazole Inhibition of the Gastric H+/K+-ATPase. *Biochim. Biophys. Acta* **1988**, *956* (1), 49−57.

(13) Lambrecht, N.; Munson K.; Vagin O.; Sachs G. Comparison of Covalent With Reversible Inhibitor Binding Sites of the Gastric H,K-ATPase By Site-Directed Mutagenesis. *J. Biol. Chem*. **2000**, *275* (6), 4041−4048.

(14) Swarts, H. G.; Klaasse, C. H.; de Boer, M.; Fransen, A. M.; De Pont, J. J. Role of Negatively Charged Residues in the Fifth and Sixth Transmembrane Domains of the Catalytic Subunit of Gastric H+, K+-ATPase. *J. Biol. Chem*. **1996**, *271* (47), 29764−29772.

(15) Xian, Y.; Hebert, H. Three-Dimensional Structure of the Porcine Gastric H,K-ATPase From Negatively Stained Crystals. *J. Struct. Biol.* **1997**, *118* (3), 169−177.

(16) Greer, J. Comparative Modeling Methods: Application to the Family of the Mammalian Serine Proteases. *Proteins* **1999**, *7*, 317−334.

(17) Blundell, T. L.; Sibanda B. L.; Sternberg, M. J. E.; Thornton, J. M. Knowledge-Based Prediction of Protein Structures and the Design of Novel Molecules. *Nature* **1987**, *326*, 347−352.

(18) Levitt, M. Accurate Modeling of Protein Conformation by Automatic Segment Matching. *J. Mol. Biol.* **1992**, *226*, 507−533.

(19) Unger, R.; Harel, D.; Wherland, S.; Sussman, J. L. A 3-D Building Blocks Approach to Analyzing and Predicting Structure of Proteins. *Proteins* **1989**, *5*, 355−373.

(20) Aszodi, A.; Taylor, W. R. Homology Modeling by Distance Geometry. *Folding Des.* **1996**, *1*, 325−334.

(21) Brocklehurst, S. M.; Perham, R. N. Prediction of the Three-Dimensional Structures of the Biotinylated Domain from Yeast Pyruvate Carboxylase and of the Lipolyated H Protein from the Pea Leaf Glycine Cleavage System: a New Automated Method for the Prediction of Protein Tertiary Structure. *Protein Sci.* **1993**, *2*, 629−639.

(22) Sali, A.; Blundell, T. L. Comparative Protein Modelling by Satisfaction of Spatial Sestraints. *J. Mol. Biol*. **1993**, *234* (3), 779−815.

(23) Toyoshima, C.; Nakasako M.; Nomura H.; Ogawa H. Crystal Structure of the Calcium Pump of Sarcoplasmic Reticulum at 2.6 Å Resolution. *Nature* **2000**, *405* (8), 647−655.

(24) Toyoshima, C.; Nomura, H. Structural Changes in the Calcium Pump Accompanying the Dissociation of Calcium. *Nature* **2002**, *418* (8), 605−611.

(25) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK: a Program to Check the Stereochemical Quality of Protein Structures. *J. Appl. Crystallogr*. **1993**, *26*, 283−291.

(26) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment Through Sequence Weighting, Positions-Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Res.* **1994**, *22*, 4673−4680.

(27) Sippl, M. J. Recognition of Errors in Three-Dimensional Structures of Proteins. *Proteins* **1993**, *17* (4), 355−362.

(28) McLachlan, A. D. Rapid Comparison of Protein Structures. *Acta Crystallogr.* **1982**, *A38*, 871−873.

(29) Martí-Renom, M. A.; Stuart, A. C.; Fiser, A.; Sanchez, R.; Melo, F.; Sali, A. Comparative Protein Structure Modeling of Genes and Genomes. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291−325.

(30) Vagin, O.; Munson, K.; Lambrecht, N.; Karlish, S. J. D.; Sachs, G. Mutational Analysis of the K^+−Competitive Inhibitor Site of Gastric H, K.-ATPase. *Biochemistry* **2001**, *40*, 7480−7490.

(31) Asano, S.; Matsuda, S.; Tega Y.; Shimizu, K.; Sakamoto, S.; Takeguchi N. Mutational Analysis of Putative SCH 28080 Binding Sites of the Gastric H+, K+-ATPase. *J. Biol. Chem*. **1997**, *272* (28), 17668−17674.

(32) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated Docking Using a Lamarckian Genetic Algorithm and an Emperical Binding Free Energy Function. *J. Comput. Chem.* **1998**, *19* (14), 1639−1662.

(33) Vagin, O.; Denevich, S.; Munson, K.; Sachs, G. SCH28080, A K^+-Competitive Inhibitor of the Gastric H,K-ATPase, Binds Near the 5−6 Luminal Loop and Prevents K^+ Access to the Ion Site. *Biochemistry* **2002**, *41* (42), 12755−12762.

(34) Munson, K. B.; Lambrecht, N.; Sachs, G. Effects of mutations in M4 of the gastric H+, K+-ATPase on inhibition kinetics of SCH28080. *Biochemistry* **2000**, *39* (11), 2997−3004.

(35) Besancon, M.; Simon, A.; Sachs, G.; Shin, J. M. Sites of Reaction of the Gastric H,K-ATPase with Extracytoplasmic Thiol Reagents. *J. Biol. Chem*. **1997**, *272* (36), 22438−22446.

(36) Munson, K.; Vagin O.; Sachs G. The K^+-Competitive Inhibitor SCH28080 Provides Tertiary Structural Information on the K Transport Domain of the H,K ATPase. *Biophys. J.* **2002**, *82* (1), 1263.

(37) Middleton, D. A.; Robins, R.; Feng, X.; Levitt, M. H.; Spiers, I. D.; Schwalbe, C.; Reid, D. G.; Watts, A. The Conformation of an Inhibitor Bound to Gastric Proton Pump. *FEBS Letts*. **1997**, *410*, 269−274.

(38) Keeling, D. J.; Taylor, A. G.; Schudt, C. The Binding of a K^+ Competitive Ligand, 2-methyl,8-(phenylmethoxy)imidazo(1,2-a)-pyridine 3-acetonitrile, to the Gastric $(H^+ + K^+)$-ATPase. *J. Biol. Chem*. **1989**, *264* (10), 5545−5551.

JM050326O