A model of 3D-structure of H+, K+-ATPase catalytic subunit derived by homology modeling

Dong YAN, Yuan-dong HU[1], Song LI[1], Mao-sheng CHENG[2]

Shenyang Pharmaceutical University, Shenyang 110016; [1]Academy of Military Medical Sciences, Beijing 100850, China

ABSTRACT

AIM: To build a model of 3D-structure of H+, K+-ATPase catalytic subunit for theoretical study and anti-ulcer drug design. **METHODS:** The model was built on the basis of structural data from the Ca^{2+}-ATPase. Structurally conserved regions were defined by amino acid sequence comparisons, optimum interconnecting loops were selected from the protein databank, and amino (N)- and carboxyl (C)-terminal ends were generated as random coil structures. Applying molecular mechanics method then minimized the model energy. Molecular dynamics technique was used to do further structural optimization. **RESULTS:** The model of 3D-structure of H+, K+-ATPase was derived. The model is reasonable according to several validation criteria. There were ten transmembrane helices (TM1-TM10) in the model and inhibitor-binding site was identified on the TM5-8 riched negatively charged residues. **CONCLUSION:** The 3D-structure model from our study is informative to guide future molecular biology study about H+, K+-ATPase and drug design based on database searching.

INTRODUCTION

H+, K+-ATPase is a member of the P-type ATPase that responsible for acid secretion into the stomach and catalyzes electro neutral exchange of cytoplasmic H+ and external K+ coupled with ATP hydrolysis[1]. It is present in gastric parietal cells of various higher animals such as human, rabbits, pigs, dogs, rats, and frogs and is located in cytoplasmic vesicles or apical plasma membranes of the secretary canaliculus. H+, K+-ATPase is composed of two subunits, the alpha subunit, which contains the phosphorylation site, is the catalytic subunit, and the beta subunit, which has six or seven

N-linked glycosylation sites. It has been shown that the H+, K+-ATPase inhibitors bind to transmembrane helices of the alpha subunit, but the function of beta subunit is not so clear[2-4]. An accurate, high-resolution picture of the inhibitor-binding site will await determination of the 3-D structure of the alpha subunit of a H+, K+-ATPase.

The alpha subunit of human H+, K+-ATPase is comprised of about 1035 amino acid residues (M_r=114 047)[5], these residues are thought to be arranged in 10 transmembrane (TM) segment. The evidences have been obtained through biochemistry and molecule biology studies that covalent inhibitors, "razoles (such as omeprazole, pantoprazole, *etc*)", react with cys813 in the intercellular M5-M6 region of the H+, K+-ATPase. Trypsinolysis followed by tricine gradient gel separation of the SDS-solubilized membrane fraction and sequencing provided details of the sequences of the first eight membrane segments, as did labeling either with

thiophilic reagents such as omeprazole, lansoprazole or pantoprazole (M3/4, M5/6, and M7/8) or with imidazo-pyridine derivatives of SCH28080[6,7]. A combination of tryptic cleavage, peptide sequencing and *in vitro* translation tests demonstrated that covalent ligands bind to a common site, Cys 813 in external surfaces.

Knowledge of the three-dimensional structure can be great importance for the drug design. Structural information often greatly enhances our understanding of proteins function. X-ray crystallography, electron microscopy or diffraction and NMR spectroscopy are the three techniques to obtain actual 3D-structure of proteins. However, H^+, K^+-ATPase has not yet been crystallized, and the main disadvantage of NMR for solving H^+, K^+-ATPase is that high concentration of dissolved H^+, K^+-ATPase is needed, but H^+, K^+-ATPase refuse to dissolve in normal solvents. In the absence of experimental data, model building on the basis of the known three-dimensional structure of a homologous protein is at present the only reliable method to obtain structural information. In this article, we built the 3D-structural model of H^+, K^+-ATPase based on the known 3D-structure of Ca^{2+}-ATPase[8] with homology method.

COMPUTATIONAL METHODS

Model building by homology is a multistep process. The process proposed by the followed steps: template recognition, sequence alignment, backbone and canonical loops generation, side chain generation plus optimization, and overall model optimization.

Template recognition and amino acids sequence alignment A BLASTP search using the Human Gastric H^+, K^+-ATPase alpha chain sequence (ATHA, swisspot ID: P20648) as the query sequence matched only one template, 1EUL, the crystal structure of Ca^{2+}-ATPase at 2.6 Å resolution. After finding the template, a pair wise sequence alignment between 1EUL and ATHA was obtained by ClustalW program. Then, removing the m-boxes and using the Modeler/Align2D command in Homology performed alignment correction.

3D model building Backbone of the protein's core regions were transferred directly from the corresponding coordinates of 1EUL. Side chains conformation for backbone residues was generated automatically by Homology. Coordinates for the remaining residues in the loops were obtained from PDB database using a fragment-searching approach. The 10 best candidate structures of the loops was extracted from PDB database according to the minimum RMS deviation of the

$C\alpha$ trace between the residues adjacent to the fragment considered (5 adjacent residues per sides). The lowest RMS loop was then selected on the basis of the analysis result of the Profile-3D program that measured the compatibility between the protein sequence profile and its 3D profile. The (C)- and (N)- terminals of the model were generated automatically through end-repair with Homology. The resulting model was subjected to an energy minimization by using the Refine routine in Homology. Using the parameters as a distance-dependent dielectric constant $\varepsilon=1.0$ and non-bonding cutoff of 15 Å, AMBER force field and Amber-all-atom charges, a 800-step steepest descent minimization was first performed until the maximum derivative is less than 50.0 $kcal \cdot mol^{-1} \cdot nm^{-1}$, and then a 1000-iteration conjugate gradient minimization was performed until the maximum derivative is less than 20.0 $kcal \cdot mol^{-1} \cdot nm^{-1}$. The above energy minimizations were started with the core main chain, then all the core side chains, the loop main chain, and finally all the loop side chain. The whole model was minimized to convergene.

In the next step, the model was subjected to molecular dynamics (MD) simulations using the same parameters and geometrical constrains as described above for 19 000 iterations. The model was kept 300 K during the entire simulation and wrote history file every 1000 steps. The averaged structure from the last 20 ps of the MD was then calculated and further optimized with steepest descent and conjugate gradient methods to generate the final refined structural model of H^+,K^+-ATPase.

RESULTS AND DISCUSSION

Amino acid sequence alignment and homology modeling Structural alignment of ATHA with the template ATA1_RABIT was shown in Fig 1. The sequences are highly conserved for the helices M4-M6 and M8 (shaded in Fig 1). This alignment was chosen not only according to the aligning score but also based on the natural sequence variability. This is because lipid-facing residues tend to be variable, whereas residues involved in helix-helix contacts tend to be conserved. Firstly, the structural checking was performed using the ProtStat in the Insight II/Homology. Some structural flaws produced by incorrectly packed side chains, distortions in incorrectly aligned regions, were found and corrected till all the bond lengths and dihedrals were in the normal range of geometrical criteria. And then, the Profiles-3D in Homology module was applied to

verify the residue environment according to the PDB classification in the smooth mode.

3-D Model of H⁺, K⁺ -ATPase On the basis of the sequence alignments, a model for the 3-D structure of H⁺,K⁺-ATPase has been derived (Fig 2). The cytoplasmic part was consisting of the three separated domains (named as A, B, and C). The phosphorylation site, Asp 351, is located in domain A. The transmembrane region (M) comprises ten α-helices (M1-M10, Fig 3). There is a clear segregation between M1-M6 and M7-M10 helices. The lengths of helices and the inclination to the membrane are variety. Some of the membranes (M2 and M5) are very long (60 Å) and very straight, but some are unwound (M4 and M6) or kinked (M10) at the middle of the membrane. K⁺-binding sites (M4-6, M8) show very different folding patterns. M5 is located in the center of the molecule and extends from the surface of the membrane to the center of the domain B. M2 and M3 are long helices that are rather isolated. The two long helices are across the membrane surface.

Substituted pyridyl-2-methylsulfinyl-benzimidazoles binding site Substituted pyridyl methysulfinyl benzimidazols are a class of covalently inhibitors of the enzyme. They are weak bases and undergo an acid-catalyzed rearrangement to form tetracyclic sulfenamides forming a disulfide bond with cysteine side located on extracellular surface of the enzyme after accumulate in acidic spaces[9]. Some works have been shown that at least 4 cysteines, Cys321, 813, 822, and 892, respectively, are labeled by one or more of benzimidazole[6,10]. To have insight into the chemical environments of binding sites, a 3D-model of the tetracyclic sulfenamide was constructed by SYBYL, and disulfide bonds were formed between the 3D-model and the above cysteines manually. Quantum mechanical analyses were performed for the disulfide bonds complexes. The bond energies of the disulfide bonds were obtained by the analysis. Comparing the energies of the complexes, a interesting result was found, the disulfide bond energy was rising followed the razoles' activity, the most effectual compound, rabeprazole, has the highest disul-
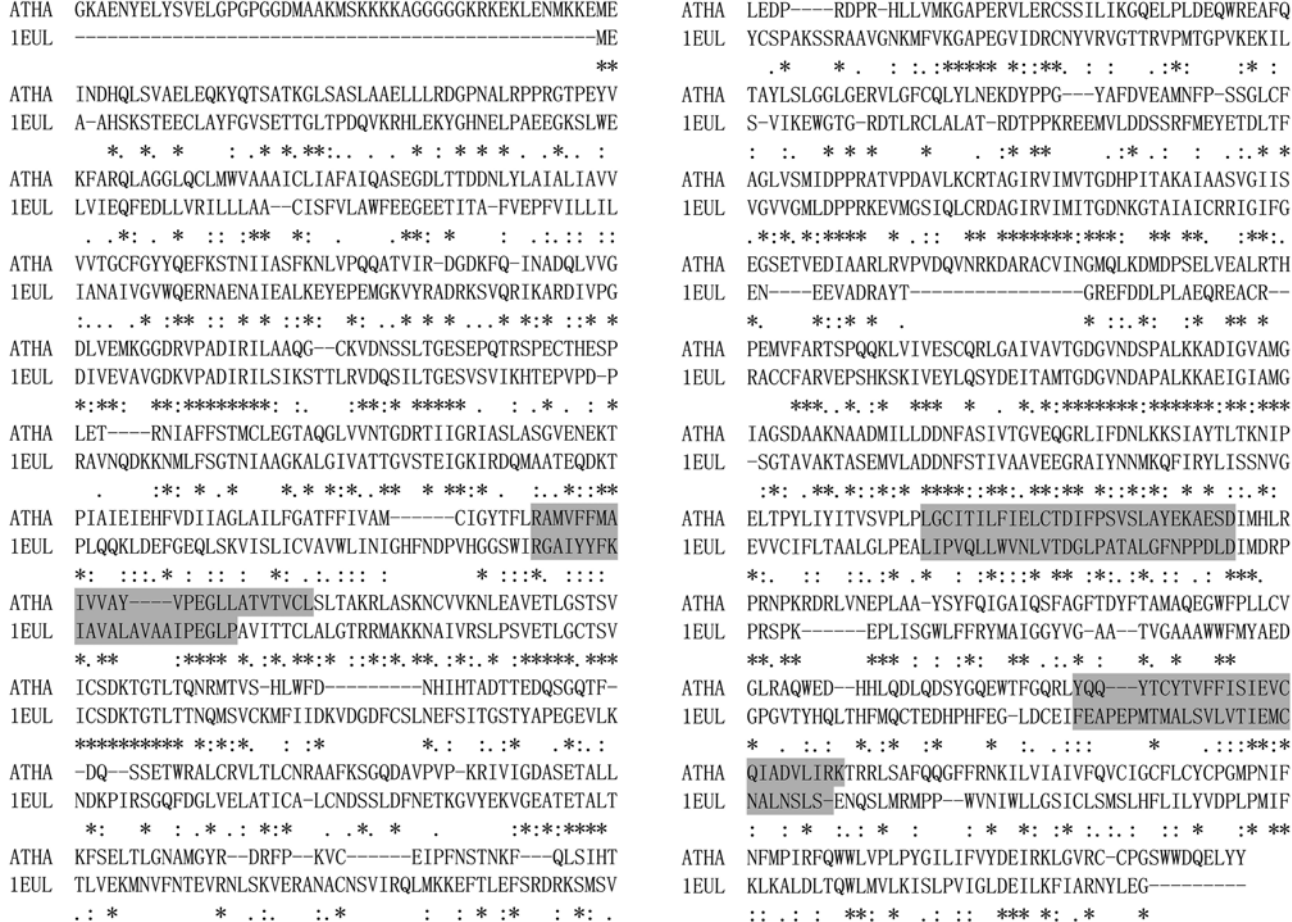
```
ATHA  GKAENYELYSVELGPGPGGDMAAKMSKKKKAGGGGGKRKEKLENMKKEME
1EUL  -------------------------------------------------ME
                                                       **
ATHA  INDHQLSVAELEQKYQTSATKGLSASLAAELLLRDGPNALRPPRGTPEYV
1EUL  A-AHSKSTEECLAYFGVSETTGLTPDQVKRHLEKYGHNELPAEEGKSLWE
         *. *. *      : .* *.**:.. . . *: * ** . ..*,. :
ATHA  KFARQLAGGLQCLMWVAAAICLIAFAIQASEGDLTTDDNLYLAIALIAVV
1EUL  LVIEQFEDLLVRILLLAA--CISFVLAWFEEGEEITITA-FVEPFVILLIL
       . .*:. *  ::  :** *:  .    .**: *    : .:.:: ::
ATHA  VVTGCFGYYQEFKSTNIIASFKNLVPQQATVIR-DGDKFQ-INADQLVVG
1EUL  IANAIVGVWQERNAENAIEALKEYEPEMGKVYRADRKSVQRIKARDIVPG
       :.... .* :** :: * * ::*:  *:  ..* * * ...*  *:* :** *
ATHA  DLVEMKGGDRVPADIRILAAQG--CKVDNSSLTGESEPQTRSPECTHESP
1EUL  DIVEVAVGDKVPADIRILSIKSTTLRVDQSILTGESVSVIKHTEPVPD-P
       *:.** **:********: :.    :**:* ***** .   : .*. : *
ATHA  LET----RNIAFFSTMCLEGTAQGLVVNTGDRTIIGRIASLASGVENEKT
1EUL  RAVNQDKKNMLFSGTNIAAGKALGIVATTGVSTEIGKIRDQMAATEQDKT
       .    :*: *.*   *. * *:*..** * **:* .  :..*::*:**
ATHA  PIAIEIEHFVDIIAGLAILFGATFFIVAM------CIGYTFLRAMVFFMA
1EUL  PLQQKLDEFGEQLSKVISLICVAWVLINIGHFNDPVHGGSWIRGAIYYFK
       *:  :::.* : :: :  *:  .:.::: : :      * :::*. ::::
ATHA  IVVAY----VPEGLLATVTVCLSLTAKRLASKNCVVKNLEAVETLGSTSV
1EUL  IAVALAVAAIPEGLPAVITTCLALGTRRMAKKNAIVRSLPSVETLGCTSV
       *.**    :**** *.:*.**:* ::*:*.**.:*.:.* :***** ***
ATHA  ICSDKTGTLTQNRMTVS-HLWFD---------NHIHTADTTEDQSGQTF-
1EUL  ICSDKTGTLTTNQMSVCKMFIIDKVDGDFCSLNEFSITGSTYAPEGEVLK
       ********* *:*:*.   : :*        *.: :.:*   .*:.:
ATHA  -DQ--SSETWRALCRVLTLCNRAAFKSGQDAVPVP-KRIVIGDASETALL
1EUL  NDKPIRSGQFDGLVELATICA-LCNDSSLDFNETKGVYEKVGEATETALT
        *:   *   : .*.:*:*   ..*. *  .      :*:*:*:****
ATHA  KFSELTLGNAMGYR--DRFP--KVC------EIPFNSTNKF---QLSIHT
1EUL  TLVEKMNVFNTEVRNLSKVERANACNSVIRQLMKKEFTLEFSRDRKSMSV
       .: :      *   *. .* :.* : * * :**:.
```

```
ATHA  LEDP----RDPR-HLLVMKGAPERVLERCSSILIKGQELPLDEQWREAFQ
1EUL  YCSPAKSSRAAVGNKMFVKGAPEGVIDRCNYVRVGTTRVPMTGPVKEKIL
        .*   *.  : :.:***** *::**.  : :   .:*:    :* :
ATHA  TAYLSLGGLGERVLGFCQLYLNEKDYPPG---YAFDVEAMNFP-SSGLCF
1EUL  S-VIKEWGTG-RDTLRCLALAT-RDTPPKREEMVLDDSSRFMEYETDLTF
       : :.  *** *   . :* **    .:*.:  : .:.* * :.:.* *
ATHA  AGLVSMIDPPRATVPDAVLKCRTAGIRVIMVTGDHPITAKAIAASVGIIS
1EUL  VGVVGMLDPPRKEVMGSIQLCRDAGIRVIMITGDNKGTAIAICRRIGIFG
       .*:*.*:****  * .:: ** *******:***: ** **.  :**:.
ATHA  EGSETVEDIAARLRVPVDDQVNRKDARACVINGMQLKDMDPSELVEALRTH
1EUL  EN----EEVADRAYT---------------GREFDDLPLAEQREACR--
       *.   *::* * .                * ::.*:  :* ** *
ATHA  PEMVFARTSPQQKLVIVESCQRLGAIVAVTGDGVNDSPALKKADIGVAMG
1EUL  RACCFARVEPSHKSKIVEYLQSYDEITAMTGDGVNDAPALKKAEIGIAMG
       ***.*.:* ***  * .  *.*:*******:******:**:***
ATHA  IAGSDAAKNAADMILLDDNFASIVTGVEQGRLIFDNLKKSIAYTLTKNIP
1EUL  -SGTAVAKTASEMVLADDNFSTIVAAVEEGRAIYNNMKFIRYLISSNVG
       :*: .** *.*:*:* ****:.:**:. **:** *:*:*: * * ::.*:
ATHA  ELTPYLIYITVSVPLPLGCITILFIELCTDIFPSVSLAYEKAESDIMHLR
1EUL  EVVCIFLTAALGLPEALIPVQLLWVNLVTDGLPATALGFNPPDLDIMDRP
       *:. ::  :: :.:* .* : :*:::* ** :*:.:*.:: .: ***.
ATHA  PRNPKRDRLVNEPLAA-YSYFQIGAIQSFAGFTDYFTAMAQEGWFPLLCV
1EUL  PRSPK------EPLISGWLFFRYMAIGGYVG-AA--TVGAAAWWFMYAED
       **.**    ***  : : **:   ** .:*:   * . *    **
ATHA  GLRAQWED--HHLQDLQDSYGQEWTFGQRLYQQ---YTCYTVFFISIEVC
1EUL  GPGVTYHQLTHFMQCTEDHPHFEG-LDCEIFEAPEPMTMALSVLVTIEMC
       *  . .:  *.:*  :*   * *  :. .:::     *    .:::**:*
ATHA  QIADVLIRKTRRLSAFQQGFFRNKILVIAIVFQVCIGCFLCYCPGMPNIF
1EUL  NALNSLS-ENQSLMRMPP--WVNIWLLGSICLSMSLHFLILYVDPLPMIF
       : : *  :.: *    : * * :**::.: ::   : * * **
ATHA  NFMPIRFQWWLVPLPYGILIFVYDEIRKLGVRC-CPGSWWDQELYY
1EUL  KLKALDLTQWLMVLKISLPVIGLDEILKFIARNYLEG---------
       ::.: : : **: *  .:  : *** *: .* *
```

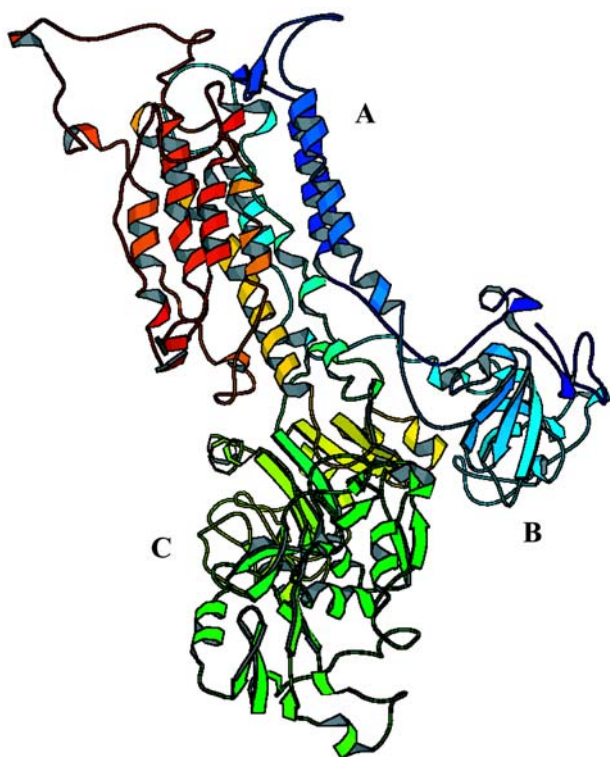**Fig 1. Structural alignment of ATHA with the template ATA1_RABIT (1EUL).**

Fig 2. Three-dimensional model of the H⁺, K⁺-ATPase catalytic subunit.

fide bond energy with the cysteine (Tab 1). At the same time, the long acyclic substitutes in pyridine of lansoprazole and rabeprazole push the side chains of cysteine to make the sulfinyl more exposed (Fig 4). It can be presumed that the energy of the sulfide bond and the substitutes in phenyl were related with the activity.

SCH28080 binding site SCH28080, a K⁺-competitive gastric H⁺,K⁺-ATPase inhibitor, has been found interacted with Phe[126] in M1-M2 extracellular loop and prevented substituted benzimidazoles such as omeprazole from interacting with Cys[813][11]. These information suggest that SCH28080 and omeprazole binding sites overlap and the SCH28080 binding pocket on the protein should be formed by a limited set of amino acid in several trans-membranes, the pocket's size should correspond to the size of SCH28080. Because the negatively charged residues located in or around the M5-M6 domain are important for K⁺ binding and conserved in other P-type ATPase[12], several negatively charged residues such as Glu[822], Lys791, Glu[795], in above regions, Phe[126] and Cys[813] were defined as binding pocket in FlexiDock of SYBYL. After cycles of dock procedures, the amount of residues were reduced and

Tab 1. Disulfide bonds energy of razoles with Cysteine.

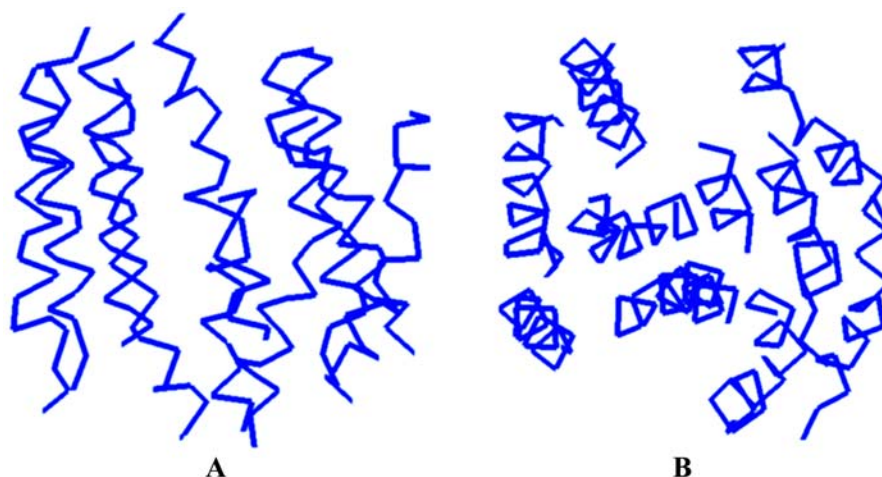| Razoles	Omeprazole		Pantoprazole		Lansoprazole		Rabeprazole	
Formed disulfide bond with	Cys[813]	Cys[822]	Cys[813]	Cys[822]	Cys[813]	Cys[822]	Cys[813]	Cys[822]
Bond energy /kcal·mol⁻¹	0.122	0.010	0.344	0.004	0.570	0.075	0.281	0.197

Fig 3. Ten transmembrane region of the model. A) View from the membrane; B) View from the cytoplasm.
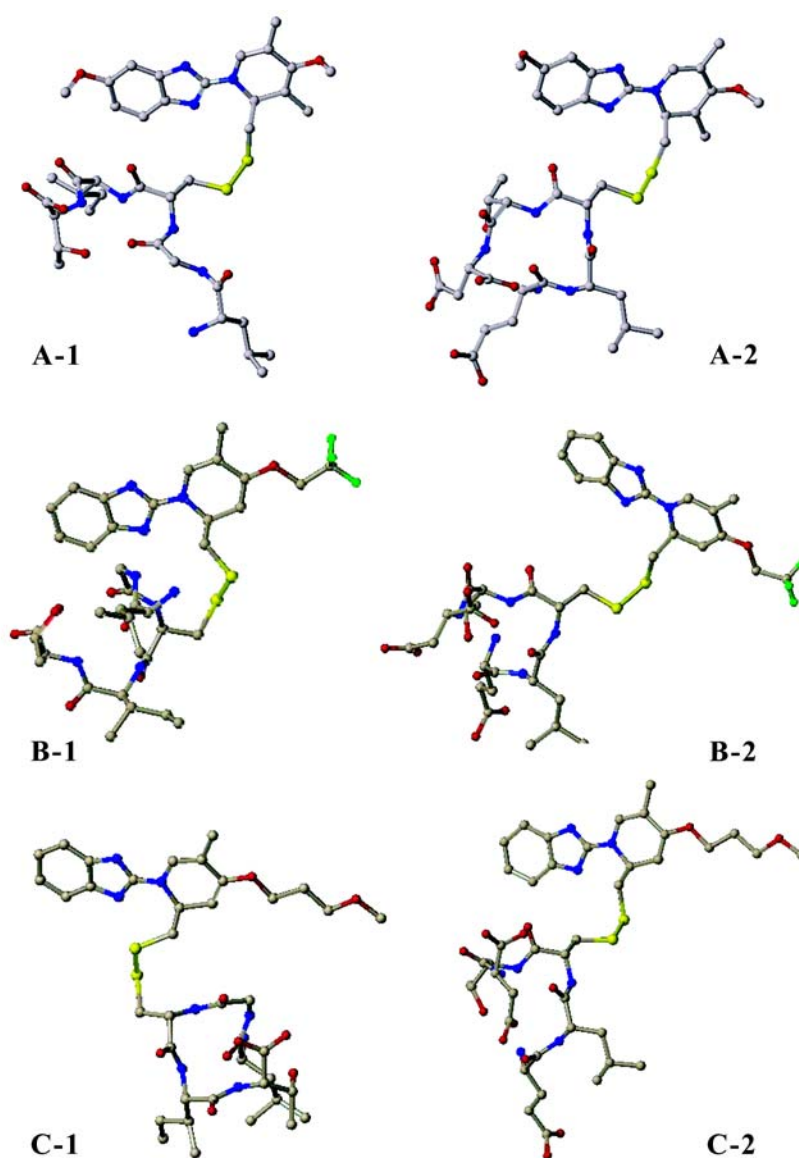
Fig 4. The razoles-cysteine complexes (hydrogen were not present). A-1: omeprazole with Cys813; A-2: omeprazole with Cys822; B-1: lansoprazole with Cys813; B-2: lansoprazole with Cys822; C-1: rabeprazole with Cys813; C-2: rabeprazole with Cys822.

a favorable energy of SCH28080 and enzyme complex was obtained. In this complex, phenyl ring of the Phe[126] toward to the extracellular surface of the enzyme, the conformation of the SCH28080 like a bow integrates in the pocket surround by Glu[795], Glu[822], Glu[936], Phe[126] and Cys[813]. The phenyl of SCH28080 interacted with the phenyl of Phe[126], and the polar part interacted with negatively charged residues (Fig 5).

This work offered a computed model of gastric H[+], K[+]-ATPase catalytic subunit, it provided a valuable insight into the structural of the enzyme and the structural environments of the inhibitors binding sites. Struc-

tural information of the H[+], K[+]-ATPase is lacking, but the homology modeling method provided a model suggesting a mechanism for inhibitors interacting with the enzyme through comparing theoretical 3D-structural data with the biochemical information of P-type ATPase. The model proposes that the electrostatic and hydrophobic factors around the binding sites influence the inhibitors interacting with the enzyme, and the reversible inhibitor SCH28080 partially spans the extracellular surface of the enzyme to interact with residues in the M1 to M6 transmembrane domains. These information could provided some useful information for new
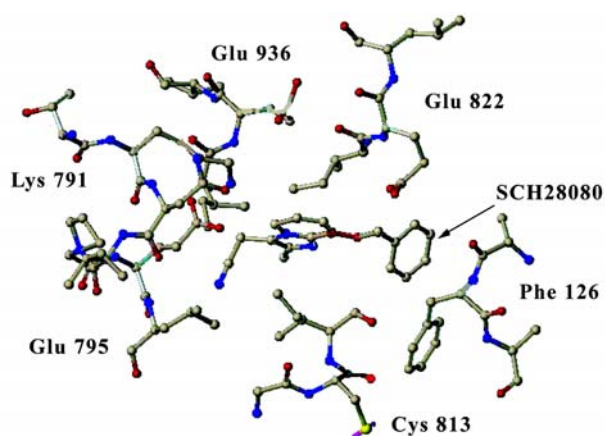
Fig 5. SCH28080 binding with the active pocket (hydrogen atoms were not present).

class of inhibitors, but the more detail information was waiting for the farther structural information becoming available.

REFERENCES

1   Slayman CL, Zuckier GR. Differential function properties of a P-type ATPase/proton pump. Ann N Y Acad Sci 1989; 574: 233-45.

2   Glynn IM, Karlish SJ. Occluded cations in active transport. Annu Rev Biochem 1990; 59: 171-205.

3   Toyoshima C, Sasabe H, Stokes DL. Three-dimensional cryo-electron microscopy of the calcium ion pump in the sarco-plasmic reticulum membrane. Nature 1993; 362: 467-71.

4   Carafoli E. The Ca^{2+} pump of the plasma membrane. J Biol Chem 1992; 267: 2115-8.

5   Jaunin P, Horisberger JD, Richter K, Good PJ, Rossier BC, Geering K. Processing, intracellular transport, and functional expression of endogenous and exogenous alpha-beta 3 Na^+, K^+-ATPase complexes in *Xenopus* oocytes. J Biol Chem 1992; 267: 577-85.

6   Besancon M, Simon A, Sachs G, Shin JM. Sites of reaction of the gastric H^+,K^+-ATPase with extracytoplasmic thiol reagents. J Biol Chem 1997; 272: 22438-46.

7   Munson K, Lambrecht N, Shin JM, Sachs G. Analysis of the membrane domain of the gastric H^+/K^+-ATPase. J Exp Biol 2000; 203 Pt 1: 161-70.

8   Toyoshima C, Nakasako M, Nomura H, Ogawa H. Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 A resolution. Nature 2000; 405: 647-55.

9   Lindberg P, Nordberg P, Alminger T, Brandstrom A, Wallmark B. The mechanism of action of the gastric acid secretion inhibitor, omeprazole. J Med Chem 1986; 29: 1327-9.

10 Besancon M, Shin JM, Mercier F, Munson K, Miller M, Hersey S, *et al.* Membrane topology and omeprazole label-ing of the gastric H^+, K^+-ATPase. Biochemistry 1993; 32: 2345-55.

11 Hersey SJ, Steiner L, Mendlein J, Rabon E, Sachs G. SCH28080 prevents omeprazole inhibition of the gastric H^+/K^+-ATPase. Biochim Biophys Acta 1988; 956: 49-57.

12 Swarts HG, Klaassen CH, de Boer M, Fransen JA, De Pont JJ. 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: 29764-72.