ACADEMIA ROMÂNĂ
Revue Roumaine de Chimie
http://web.icf.ro/rrch/

Rev. Roum. Chim.,
2012, *57*(1), 39-44

3D HOMOLOGY MODEL OF THE HUMAN PROSTAGLANDIN E2 RECEPTOR EP4 SUBTYPE**

Daniela MARGAN,[a] Ana BOROTA,[b] Maria MRACEC[c] and Mircea MRACEC[b,c*]

[a] Technical CollegeTimişoara, 2 Huniade Iancu Sq., RO-300002, Timişoara, Roumania
[b] Chemistry Institute Timişoara of Roumanian Academy, 24, M.Viteazul Av., RO-300223, Timişoara, Roumania
[c] Molecular Forecast Research Center, 125, A 4, Paunescu-Podeanu str., RO-300569, Timişoara, Roumania

Received March 3, 2011

In order to achieve the 3D homology model of the human prostaglandin E2 receptor EP4 subtype (hPE2R4) we used as template the X-rays structure (at 2.6 Å resolution) of the human adenosine receptor A2a (hAA2AR), (PDB file 3EML). The amino acid sequence of the PE2R4 prostanoid receptor was taken from the SWISS-PROT database (code P35408). The obtained 3D homology model was verified with the PROCHECK program in view of establishing its quality. The model was used for docking dinoprostone (prostaglandin E2), the endogenous ligand of the four PE2R subtypes.

INTRODUCTION

G-protein-coupled receptors (GPCRs) are seven-helix transmembrane proteins that provide a molecular link between extracellular signals and intracellular reactions ranging from intercellular communication processes to physiological responses. They are the largest family of membrane receptors, being coded of approximately 3% of the genes from the human genome.[1] It is estimated that over 50% of the current medication targets are GPCR, which is in contrast with the small proportion of genes from the human genome which is predicted to codify GPCR, illustrating the importance of these proteins, from a medical as well as pharmaceutical point of view.[1]

Knowing the three-dimensional (3D) structure of GPCR is important for understanding the molecular mechanism of diseases and syndromes caused by mutations in these receptors, as well as for the design based on the structure of small molecules that act as therapeutic treatments.

For the first time the topology of a GPCR on an atomic level was revealed in the year 2000, when Palczewski and collaborators solved the X-rays structure of bovine rhodopsin.[2] Knowing the 3D structure of a protein is of major importance for the discovery of medicines, because it serves to identify new ligands through the computational (*in silico*) techniques such as *de novo* design and virtual screening.

Prostaglandin E2 (PGE2) exerts its actions by acting on four prostanoid receptor subtypes PE2R1, PE2R2, PE2R3, and PE2R4 (also known as EP1, EP2, EP3 and EP4 receptor subtypes). The PE2R subtypes exhibit differences in signal transduction, tissue localization, and regulation of expression. The most important physiological and pathophysiological actions of PE2R4 are the following: facilitates closure of ductus arteriosus,[3,4] induces bone formation,[5] protects against inflammatory bowel disease,[6] facilitates Langerhans cell migration and maturation,[7] mediates joint inflammation in collagen-induced arthritis.[8]

3D modeling of the hPE2R4 provides opportunities to design selective PE2R4 agonists and antagonists. The template we chose was the

* Corresponding author: mracec@acad-icht.tm.edu.ro
** Supplementary information on http://web.icf.ro/rrch/

human adenosine receptor A2a, based on the sequence and structural similarities between them. Both belong to the class A (rhodopsin-like) of the GPCR family. Sequential homology between hAA2AR and PE2R4 is low, under 30%, yet building a 3D model for this receptor is possible due to highly conserved amino acids and regions in the GPCR family. The crystal structure of the human adenosine receptor A2a was determined in complex with the subtype selective high affinity antagonist (4-(2-[7-amino-2-(2-furyl)-[1,2,4]triazolo-[2,3-a][1,3,5]triazin-5-ylamino]ethyl)-phenol (ZM241385).[9]

In this paper we report building, refinement and results of geometrical tests of a 3D model of PE2R4 based on the hAA2AR template.

METHODS

Homology modeling of the human prostaglandin E2 receptor EP4 subtype, hPE2R4, was performed using as template the X-ray structure at 2.6 Å resolution of the human adenosine receptor A2a, PDB code file 3EML. The sequence of the human PE2R4 in FASTA format was extracted from the SwissProt database[10] and the one of hAA2AR from PDB.[11]

The two sequences were aligned using the SWISS-MODEL software and the resulted alignment was manually refined and submitted further for automatic model building to the Swiss-PdbViewer server.[12,13] The resulted model was then evaluated using the PROCHECK program.[14,15] The refinement of the model was performed using the Protein Preparation Wizard[16] program from Schrödinger software package.

The structure of the endogenous ligand, PGE2, was prepared for docking experiments with the LigPrep 2.3. application implemented in Schrödinger suite and accessible from the Maestro interface.[17,18] The LigPrep 2.3. program used the next criteria: Ionizer, Generate tautomers, Generate possible conformers at pH 7, Force field: OPLS_2005.[17]

Ligand docking was performed with the Schrödinger's Induced Fit[19] Docking (IFD) protocol, taking in account the flexibility of the ligand and protein.

RESULTS

In order to obtain a good 3D homology model we took into account that in the alignment of the two sequences the presence of empty spaces in transmembrane regions is not allowed because the transmembrane structure is conserved in all GPCRs even better than the sequence similarity. Initially, we tried to align the amino acid sequence of hPE2R4 (P35408 ID) with the sequence of the human adenosine receptor A2a (3EML) by using the automatic alignment program T-Coffee.[20] The alignment resulted through running the T-Coffee[20] program is given as Supplementary material. This alignment could not be used by the SWISS-MODEL server[12] for generating the 3D homology model, since empty spaces occurred in the transmembrane regions of both hPE2R4 and 3EML_A2a sequences. Therefore, we tried another automatic alignment by using the SWISS-MODEL server. The result of alignment is a little better, but not good enough for the building of a 3D model. A disadvantage of this method of automatic alignment is that the segments that received a high score, which prove consistency within the data set, have the score improved through extension so that they become insensitive to penalties due to empty spaces in the alignment. It must again be underlined that, due to structural homology, no empty space is allowed in transmembrane regions. For this reason we tried a manual refinement of the alignment obtained with the SWISS-MODEL server, by taking into consideration that the conserved amino acids must correspond in the alignment of the two sequences, and the empty spaces appeared in the alignment should not be situated in the regions of major importance, *i.e.* the transmembrane regions. The empty spaces were moved towards the loops. Following the manual refinement we obtained a qualitatively superior alignment of the two sequences, alignment which was taken over by the SWISS-MODEL server, which generated the raw 3D homology model of hPE2R4. The refined manual alignment is presented in Fig. 1.

This alignment presents a sequential identity of 16.1% and a sequential similarity of 37.7%, satisfactory percentages for obtaining a 3D model of the hPE2R4 prostanoid receptor.

The alignment of the hPE2R4 and hAA2AR (template) sequences showed the presence of highly conserved residues. In order to identify the conserved amino acids, the notation system proposed by Balerestos and Weinstein[21,22] is used. According to this convention, each amino acid residue is identified through the number of the transmembrane to which it pertains (from 1 to 7), and the most conserved amino acid of each transmembrane has the number 50 assigned to it. The other residues are numbered taking into consideration the most conserved amino acid position in each transmembrane.[22,23]

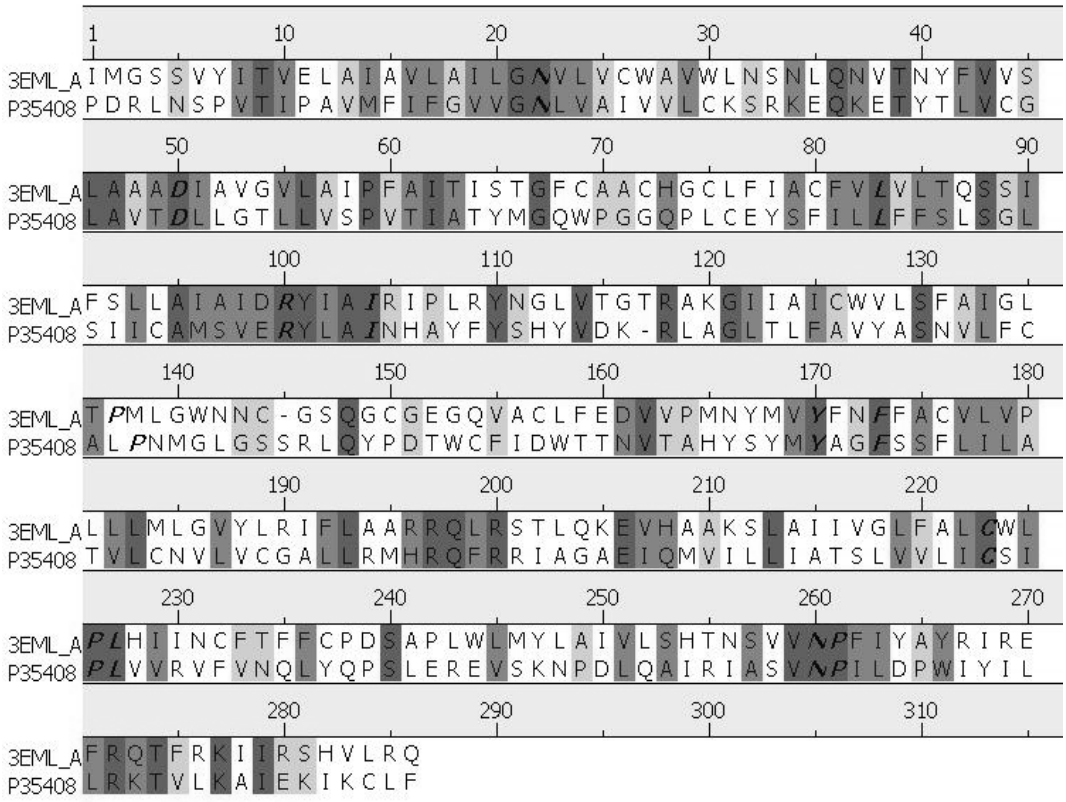

Fig. 1 – Manual alignment of the two sequences of the human adenosine receptor A2a (3EML) and hPE2R4 (P35408) achieved with the Discovery studio program; dark blue bold italic – highly conserved amino acids in the GPCR family; blue – highly similar amino acids; light blue – similar amino acids.

Thus, in transmembrane 1 (TM1) of the hPE2R4 prostanoid receptor the most conserved amino acid is a residue of asparagine, noted as N1.50. It is preceded by the glycine G1.49 and followed by the valine V1.52 (leucine L1.52 for 3EML_A2a). A residue of aspartic acid, D2.50, preceded by the leucine L2.46 and alanine A2.47 and followed after five positions by leucine L2.56 represent the residues conserved from the TM2. On TM3, the most conserved residue is arginine R3.50 situated at the cytoplasmic end of the TM3, being preceded by a residue of aspartic or glutamic acid, D/E3.49 (D3.49 at hAA2AR and E3.49 at hPE2R4) and followed by tyrosine Y3.51 and a branched amino acid (isoleucine or valine) I/V3.54. This structural motif, (D/E)RY--(V/I), is known in literature under the name of "the arginine cage" (contains the DRY region) and has the role of maintaining the receptor in inactive conformation.[24] Also on TM3 there is another leucine residue L3.43 (only at hAA2AR) and a cysteine in the 3.25 position which intersects with another cysteine residue situated on the extracellular loop between TM4 and TM5. TM4 contains only a highly conserved residue (serine S4.53), while TM5 contains two: the phenylalanine F5.43 and the leucine L5.53. TM6 is one of the most conserved transmembrane segments, containing cysteine C6.47, proline P6.50, and leucine L6.51. The last transmembrane (TM7) is characterized by the following set of conserved amino acids: asparagine N7.49, proline P7.50 and arginine 7.61.

The validation step was performed using the PROCHECK software which verifies the normality of torsion angles, bond angles, bond lengths and distances between unbounded neighbor atoms. The tests of geometric parameters of main and side chains are given as Supplementary data.

The initial 3D model of the PE2R4 was refined in many steps. To avoid close contacts and to correct some distorted bond lengths and angles or planarity of some aromatic rings signaled by PROCHECK the geometrical parameters of the residues have been optimized by using the OPLS_2005 force field. The refined 3D model contained 226 residues in the most favored regions (87.9 % in A, B, L regions of the Ramachandran plot), 28 in the additional allowed regions (10.9% in a, b, l, p regions), 2 in generously allowed regions (~a, ~b, ~p, ~l regions) and one in disallowed regions (white area) one end-residuu,

14 glycine residues, 13 proline residues. Asp167 from the disallowed region is situated on the third extracellular loop, and it does not influence the binding site, which is placed aproximately at the middle of the lipidic bilayer between the transmembranes TM3, TM5, TM6, and TM7. The Ramachandran map for the refined 3D model of hPE2R4 based on the hAA2AR template is displayed in Fig. 2.

The 3D model, represented as a solid ribbon is displayed in Fig. 3

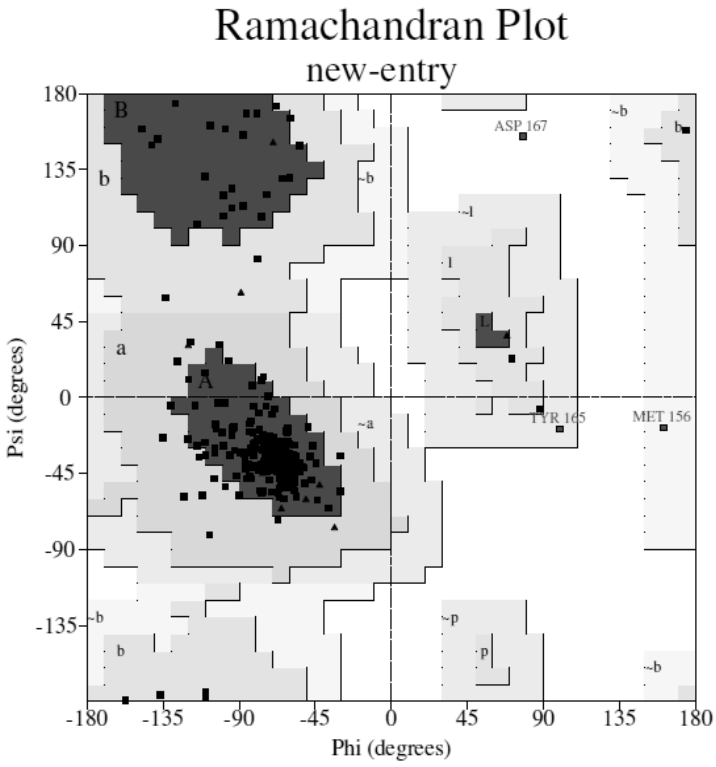

Fig. 2 – Ramachandran map for hPE2R4 after refinement steps.

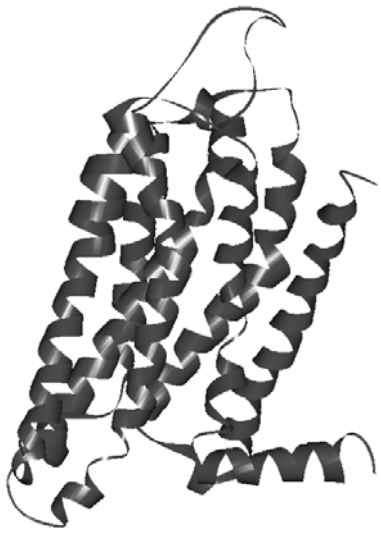

Fig. 3 – 3D-structure of hPE2R4 obtained by homology modeling using the hAA2AR structure as a template.

From site-directed mutagenesis experiments performed by some research groups the following key amino-acids involved in ligand binding have been identified: threonine T199 (T5.51), serine S278 (S6.41), serine S285 (S6.48), and arginine R316 (R7.44).[25] This information was used to verify the validity of our 3D model. In order to validate the model we used the docking software Induced Fit to dock PGE2, the endogenous ligand of all four PGE2 receptor subtypes. The structure of PGE2 was initially prepared with LigPrep program and then optimized with the OPLS_2005 force field. Through docking experiments we obtained 10 poses for the ligand-receptor complex, from which we selected as the best the one that fulfiled the following two criteria: 1) the best binding affinity reported to GlideScore; 2) the best interactions between key amino-acids from protein's binding site and certain atoms from the ligand. The pose selected as the best is presented in Fig. 4.

From the best pose one can see the formation of multiple hidrogen bonds between ligand and the amino acids from the binding site. One important interaction is between the hydrogen atom of the OH group of S285 (S6.48), which was confirmed also by mutagenesis studies as key amino-acid, and the $O_{18}H$ group of the ligand. The same OH group of the ligand is involved in another H-bond (through the O_{18} atom) with S103 (S3.36). Ligand stability in the binding site is provided also by four H-bonds formed between oxygen atoms and OH groups of the ligand and the following amino acids: tyrosine Y186 (Y5.36), phenylalanine F191 (F5.43), aspartic acid D311 (D7.39), and leucine L195 (L5.47).

The analysis of stereochemical parameters as much as the docking of the endogenous ligand validate the refined 3D homology model of the PE2R4 prostanoid receptor.

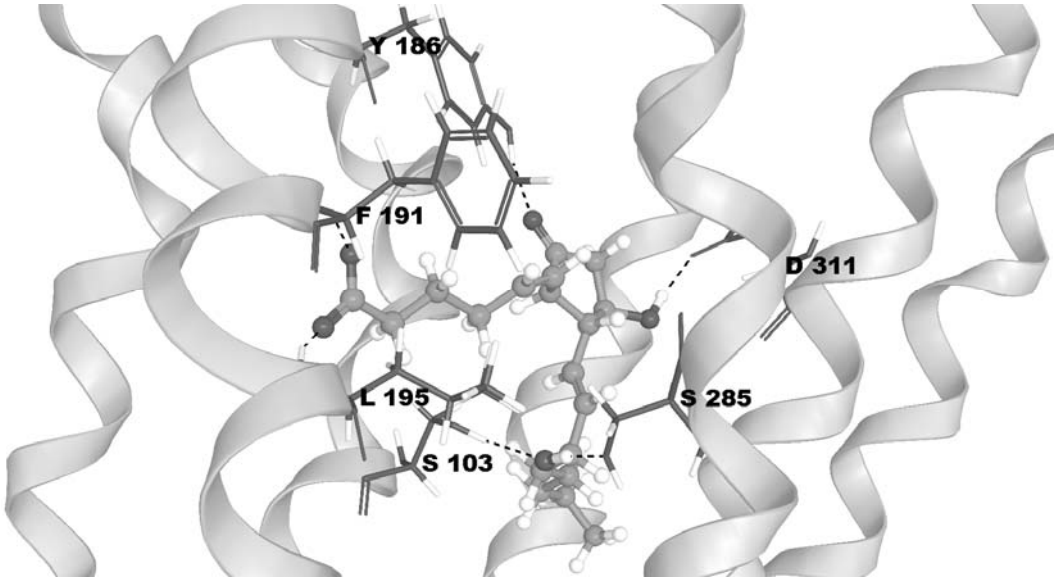

Fig. 4 – PGE2 docked in the 3D model of PE2R4.

CONCLUSIONS

The 3D model of hPE2R4 was obtained as an attempt to have a better view of the ligand behavior in the receptor's binding site. The homology modeling of the hPE2R4 was possible due to the structural similarities (seven transmembrane helices) with hAA2AR and conserved residues in the GPCR family.

The performed docking experiment highlighted the amino acids involved in endogenous ligand's interactions and responsible for its binding with the protein. A key role, confirmed by docking study and also by known experimental data, is played by serine S285.

The obtained 3D model can be a useful tool to design compounds with better selectivity for hPE2R4.

Acknowledgments: The authors thank to the National Councilor for Research and High Education (CNCSIS) for allowing the financial backgrounds use for purchasing the HyperChem 7.52 package by CNCSIS grant no.776/2005/Agreement 27658/2005/GR177/2006/1973/2006; and for Schrödinger package by Grant PN-II-PCE-ID no. 1268/Agreement 248/2007; additional agreement 2/2009, additional agreement 3/2010.

REFERENCES

1. C. L. Worth, G. Kleinau and G. Krause, *Public Library of Science one (PLoS1)*, **2009**, *4*, e7011.
2. C. Vasseur, P. Rodien, I. Beau, A.Desroches and C. Gérard, *N. Engl. J. Med.*, **2003**, *34*, 753-759.
3. E. Segi, Y. Sugimoto, A. Yamasaki, Y. Aze, H. Oida, T. Nishimura, T. Murata, T. Matsuoka, F. Ushikubi, M. Hirose, T. Tanaka, N. Yoshida, S. Narumiya and A. Ichikawa, *Biochem. Biophys. Res. Commun.,* **1998**, *246*, 7-12.
4. U. Yokoyama, S. Minamisawa, Q. Hong, S. Ghatak, T. Akaike, E. Segi-Nishida, S. Iwasaki, M. Iwamoto, S. Misra, K. Tamura, H. Hori, S. Yokota, B. P. Toole, Y. Sugimoto and Y. Ishikawa, *J. Clin. Invest.,* **2006,** *116*, 3026-3034.
5. K Yoshida, H. Oida, T. Kobayashi, T. Maruyama, M.Tanaka, T. Katayama, K. Yamaguchi, E. Segi, T. Tsuboyama, M. Matsushita, K. Ito, Y. Ito, Y. Sugimoto, F. Ushikubi, S. Ohuchida, K. Kondo, T. Nakamura and S. Narumiya, *Proc. Natl. Acad. Sci. U. S. A.,* **2002**, *99*, 4580-4585.
6. K. Kabashima, T. Saji, T. Murata, M. Nagamachi, T. Matsuoka, E. Segi, K. Tsuboi, Y. Sugimoto, T. Kobayashi, Y. Miyachi, A. Ichikawa and S. Narumiya, *J. Clin. Invest*., **2002**, *109*, 883-893.
7. K. Kabashima, D. Sakata, M. Nagamachi, Y. Miyachi, K. Inaba, and S. Narumiya, *Nat. Med..,* **2003**, *9*, 744-749.
8. T. Honda, E. Segi-Nishida, Y. Miyachi and S. Narumiya, *J. Exp. Med.,* **2006**, *203*, 325-335.
9. V.P. Jaakola1, M. T. Griffith1, M. A. Hanson1, V. Cherezov1, E. Y.T.Chien1, J.R.Lane, A. P. Ijzerman and R. C. Stevens1, *Science,* **2008**, *322*, 1211-1217.
10. http://www.expasy.org/sprot/sprot-top.html
11. http://www.rcsb.org/pdb/home/home.do
12. www.expasy.ch/swissmod/SWISS-MODEL.html
13. T. Schwede, J. Kopp, N. Guex and M.C. Peitsch, *Nucleic Acids Res.*, **2003**, *31*, 3381-3385.
14. www.biochem.ucl.ac.uk/~roman/procheck/procheck.html
15. L.A. Laskowski, M.W. MacArthur, D.S. Moss and J.M. Thornton, *J. Appl. Cryst.*, **1993**, *26*, 283-291.
16. Schrödinger Suite 2009 Protein Preparation Wizard; Epik version 2.0, Schrödinger, LLC, New York, NY, 2009; Impact version 5.5, Schrödinger, LLC, New York, NY, 2009; Prime version 2.1, Schrödinger, LLC, New York, NY, 2009.
17. LigPrep, version 2.3, Schrödinger, LLC, New York, NY, 2009.
18. Maestro, version 9.1, Schrödinger, LLC, New York, NY, 2010.

19. Schrödinger Suite 2009 Induced Fit Docking protocol; Glide version 5.5, Schrödinger, LLC, New York, NY, 2009; Prime version 2.1, Schrödinger, LLC, New York, NY, 2009.
20. www.ch.embnet.org/software/TCoffee.html
21. J.A. Ballesteros and H. Weinstein, *Methods Neurosci.,* **1995**, *25*, 366-428.
22. J.A. Ballesteros and K. Palczewski, *Curr. Opin. Drug Discov. Devel.,* **2001**, *4*, 561-574.
23. J. Ballesteros, S. Kitanovic, F. Guarnieri, P. Davies, B.L. Fromme, K. Konvicka, L. Chi, R.P. Millar, J.S. Davidson, H. Weinstein and S.C. Sealfon, *J. Biol.. Chem.,* **1998**, *273*, 10445-10453.
24. G. Muller, *Curr. Med. Chem.,* **2000**, *7*, 861-888.
25. V. K. Honn and L.J. Marnett, "Eicosanoids and Other Bioactive Lipids in Cancer, Inflammation, and Radiation Injury 3", Kluwer Academic Publishers, 2002, p. 638.