Modeling Human Neurokinin-1 Receptor Structure using the Crystal Structure of Bovine Rhodopsin

Santosh A Khedkar, Alpeshkumar K Malde and Evans C Coutinho*

Department of Pharmaceutical Chemistry, Bombay College of Pharmacy, Kalina, Santacruz (E), Mumbai-400 098. INDIA.

Received xxx; Preprint published xxx; Accepted xxx ; Published xxx

Internet Electron. J. Mol. Des. **2004**, *3*, 000–000

Abstract

G protein-coupled receptors (GPCRs) regulate a wide range of physiological processes by transmitting signals to cells in response to stimuli such as light, Ca^{2+}, odorants, amino acids, nucleotides, peptides or proteins. GPCRs are by far the most successful drug targets as evidenced by the fact that 50% of the marketed drugs treat diseases by targeting nearly 20 GPCRs. The lack of high-resolution structures of GPCRs limits the application of structure-based drug design on these targets. However, the recent publication of the crystal structure of bovine rhodopsin has changed the scenario in GPCR structure modeling. Neurokinin-1 receptor (NK1R) is a member of the family A of GPCR, which on modulation by substance P (SP), produces a variety of physiological and pathophysiological conditions. A high-resolution structure of NK1R is not yet available and hence alternative approaches must be used for building a model 3D-structure of the NK1 receptor, which can then be used for structure-based drug design.

A 3D-structure of NK1R has been built using the recently published high-resolution crystal structure of bovine rhodopsin (PDB code: 1L9H) with *INSIGHT II* (*Homology, Discover and Analysis* modules). Due to the low sequence identity, between the 7 TM regions, a segmented approach for model building was used. The loop and end regions were modeled using simulated annealing and stringent energy minimization protocols. The model retains the global arrangement of the GPCRs and is energetically and geometrically consistent. The loops in the NK1R model are longer than those in rhodopsin and their orientation in the model, in particular the extracellular loops, would be of use in structure based drug design studies. The lipophilic potential surface of the final NK1R model has been calculated and reflects the characteristics of this membrane protein. This work provides a first complete model of human NK1R with which ligand - GPCR interactions can now be investigated at the atomic level.

Keywords. G-Protein Coupled Receptor (GPCR), Neurokinin-1 Receptor (NK1R), Homology Modeling, Lipophilic Potential.

*Correspondence Evans C. Coutinho; phone: +91-22-26670871; fax: +91-22-26670816; E-mail: evans-im@eth.net

1 INTRODUCTION

G Protein-Coupled Receptors (GPCRs) are membrane-embedded proteins responsible for signal transduction by mediating transmission of environmental stimuli such as light, odor, taste, as well as hormonal, neurotransmittal, and other types of communications across cellular membranes [1]. Nearly 50% of all recently launched drugs are targeted against GPCRs with annual worldwide sales

exceeding $30 billion in 2001. Among the 100 top-selling drugs, 25% are targeted at members of this protein family. The GPCR family represents nearly 30% of the portfolio of the targets of many pharmaceutical companies [2]. GPCRs share many structural features, the most important being a bundle of seven transmembrane α–helices connected by six loops of varying lengths. Although GPCRs share a common membrane topology, they are remarkably diverse in sequence and vary especially in the size of the extracellular amino terminus, the cytoplasmic loops, and the carboxy terminus. Based on these differences, mammalian GPCRs have been classified as family A, family B, and family C [3-6]. Family A (rhodopsin-like or adrenergic-receptor-like) is by far the largest, displaying a short amino terminal tail and highly conserved amino acid residues within each transmembrane helix. Family B (glucagon-receptor-like or secretin-receptor-like) receptors display a longer amino terminal tail with a set of six conserved cysteine residues. Lastly, members in family C (metabotropic glutamate receptor) generally have a longer amino tail (500-600 residues), which is folded and forms the ligand-binding domain [7].

The neurokinin-1 receptor (NK1R) is a member of the family A of the GPCR superfamily, and is an integral membrane protein activated by the tachykinin peptide hormone substance P (SP) [8]. The three mammalian tachykinin receptor subtypes (*i.e.*, NK1R, NK2R and NK3R) exhibit a preference for binding a particular endogenous tachykinin peptide: NK1R for SP, NK2R for neurokinin A and NK3R for neurokinin B. NK1R preferentially binds SP to elicit various biological responses, one of which is transmission of pain. NK1R activation triggers the activity of phospholipase C that increases the level of intracellular calcium. Due to the amphiphilic nature of SP, it has been proposed that it penetrates into the lipid phase adopting a partial α̃-helical conformation, and then diffuses within the membrane to the receptor binding site. In addition to SP, substance K and neurokinins A and B also bind to NK1R however with lower affinity [9].

The interaction of neurokinin (also termed as tachykinin) with NK1R plays a pivotal role in the induction and progression of inflammatory diseases. Neurokinin is also involved in a variety of physiological and pathophysiological conditions such as pain, inflammation, smooth muscle contraction, vasodilation, and activation of the immune system [9]. Thus, NK receptor antagonists have emerged as interesting agents for the treatment of primary pain, emesis, asthma, and in other disorders such as anxiety, arthritis, migraine, cancer and schizophrenia [10-13].

Until recently, the X-ray structure of bacteriorhodopsin [14,15] was being used as a three dimensional template for modeling GPCRs. It is a seven transmembrane helical protein, but not a true GPCR. The first structure solved of a GPCR was that of bovine (*Bos taurus*) rhodopsin (PDB code: 1F88) by X-ray crystallography at 2.8 Å resolution [16]. Thereafter, three other bovine rhodopsin structures [17-19] have been determined at different resolutions (PDB codes: 1HZX, 1L9H and 1GZM at 2.8 Å, 2.6 Å and 2.65 Å resolutions, respectively). Two solution structures [20,21] of bovine rhodopsin have also been reported in the literature (PDB codes: 1JFP and 1LN6).

The crystal structure of bovine rhodopsin provides the first three-dimensional information on GPCRs to support homology modeling studies and structure-based drug design approaches for other members of the GPCR family.

2 MATERIALS AND METHODS

All computations and molecular modeling of human NK1R were carried out on a Silicon Graphics O2 workstation (R5000 MIPS processor) using the *INSIGHT II* molecular modeling package (Accelrys Inc., USA) [22]. Comparative protein modeling was carried out with the *Homology* module in *InsightII*. The amino acid sequence of NK1R was obtained from the NCBI protein database (GI: 128359, accession no.: P25103). The PSI-BLAST [23] algorithm was used to identify homologous structures for NK1R by searching the structural database of protein sequences in the Protein Data Bank (PDB) [24]. The X-ray structure of bovine rhodopsin with the highest resolution (PDB code: 1L9H [18]) was used as a template for NK1R 3D structure modeling.

2.1 Sequence Alignment

The start and end regions of the transmembrane segments (TM1 to TM7) of NK1R were assigned based on information in the literature [8]. The segmented approach was used for the alignment of NK1R, wherein the N-terminus and the seven transmembrane helices (TM1 to TM7) were individually aligned to the respective segments of bovine rhodopsin using ClustalW [25] and the PAM250 scoring matrix [26]. The intracellular (IC1 to IC3), extracellular (EC1 to EC3) loops and the C-terminus were finally aligned manually with the respective regions in the crystal structure wherever a good alignment could be obtained. The alignment of NK1R with 1L9H is shown in Figure 1.

2.2 Assignment of Coordinates and Side-chain Rotamer Search

The coordinates for the residues in N-terminus and transmembrane helices (TM1 to TM7) were assigned from the N-terminus and corresponding TMs of the bovine rhodopsin crystal structure. The co-ordinates for the ECs, ICs and the C-terminus were extracted either from the crystal structure or from loops identified [27] in the PDB database, or were constructed with the "loop generate" algorithm [28] as implemented in the *Homology* module. The percentage identities (and similarity – identical plus conservative substitutions) between 1L9H and NK1R for each of the seven transmembrane regions are TM1 13% (48%); TM2 41% (55%); TM3 18% (59%); TM4 33% (48%); TM5 28% (64%); TM6 41% (68%); TM7 28% (48%). The N-terminal sequence of NK1R is almost the same length as that of bovine rhodopsin (31 amino acids) and corresponds to 13% identity and 29% similarity based on conservative substitution. The co-ordinates for this region were read from the crystal structure and then refined further. Sequences similar to the intra- and extracellular loops (IC1 to IC3 and EC1 to EC3) were searched in the PDB database and the best hits were used. The long C-terminal region (98 amino acids) is shorter than the C-terminal of

bovine rhodopsin and is the least conserved. Therefore, it was modeled using the "loop generate" algorithm. The coordinates of the side chains and the backbone atoms were copied to the target sequence only if, identical amino acids were found at corresponding positions in the sequence; for 'similar' but not 'identical' amino acids, only the common side chain atoms were copied, while for the rest, amino acid conformations from the library were used. The side chains of all residues in the model were explored for their optimal conformations, and those with minimum steric clashes (bumps) were assigned to the model. The splice points (the amide groups) were refined by minimization with the consistent valence force field (CVFF) [29] as implemented in the *Discover* program (v 98, Accelrys Inc., USA)

```
L9HM 001: MNGTEGPNFYVPFSNKTGVVR-SPFEAPQYYLAEPWQFSMLAAYMFLLIM
NK1R 001: MDNVLPVDSDLSPNISTNTSEPNQFVQPAWQ-------IVLWAAAYTVIV
          *.     .  ..    *     . *  *           * *  . .*.

L9HM 050: LGFPINFLTLY-VTVQHKKLRTPLNYILLNLAVADLFMVFGGFTTTLYTS
NK1R 044: VTSVVGNVVVMWIILAHKRMRTVTNYFLVNLAFAEASMAAFNTVVNFTYA
          .   .  . .      **.**  **.*.*** *.  *

L9HM 099: LHGYFVFGPTGCNLEGFFATLGGEIALWSLVVLAIERYVVVCKPMSNFRF
NK1R 094: VHNEWYYGLFYCKFHNFFPIAAVFASIYSMTAVAFDRYMAIIHPLQPRLS
                *   * .. **.  .   .. *.  .*..**. . .*.

L9HM 149: GENHAIMGVAFTWVMALACAAPPLVGWSRYIPEGMQCSCGIDYYTPHEET
NK1R 144: ATATK-VVICVIWVLALLLAFPQGY-YSTTETMPSRVVCMIEWPEHPNKI
                . .   **.**  * *     *       .  * *.

L9HM 199: NNESFVIYMFVVHFIIPLIVIFFCYGQLVFTVKEAAA|ATTQKAEKEVTR
NK1R 192: YEKVYHICVTVLIYFLPLLVIGYAYTVVGITLWASEIPGDSSDRYHEQVS
              . * . *. ...**.** . *  . .*              *

L9HM 253: -------MVIIMVIAFLICWLPYAGVAFYIFTHQGSDFG---PIFMTIPA
NK1R 242: AKRKVVKMMIVVVCTFAICWLPFHIFFLLPYINPDLYLKKFIQQVYLAIM
                *.*..* .* *****.    .

L9HM 293: FFAKTSAVYNPVIYIMMNKQFRNCMVTTLCCGKNPLGD|----------
NK1R 292: WLAMSSTMYNPIIYCCL-------------NDRFRLGFKHAFRCCPFISA
           .* .*..***.**                 ..  **

L9HM 334: ---------STTVSKTETSQVAPA
NK1R 329: GDYEGLEMKSTRYLQTQGSVYKVSRLETTISTVVGAHEEEPEDGPKATPS
                  **    *. *    .

NK1R 379: SLDLTSNCSSRSDSKTMTESFSFSSNVLS
```

Figure 1. Alignment of NK1R and bovine rhodopsin (1L9H) sequences using the segmented approach. The 7 TM helices (TM1 32-54, TM2 65-86, TM3 107-128, TM4 149-169, TM5 195-219, TM6 249-270, TM7 284-308) are highlighted in magenta, the intracellular loops (IC1 55-64, IC2 129-148, IC3 220-248) in green, the extracellular loops (EC1 87-106, EC2 170-194, EC3 271-283) in blue and the N-terminus and C-terminus are shown in black. Identical amino acids are indicated with an "*" whereas the 'conserved' amino acid residues, which meet the criteria of either highly conservative or semi-conservative substitutions as defined by the PAM 250 matrix are indicated with "·".

2.3 Refinement and Validation

The disulfide bond between the extracellular loops EC1 and EC2, which is conserved across all GPCR families [30], was manually created between Cys105 in EC1 and Cys180 in EC2. All the six loops (ECs and ICs), the N-terminal and the C-terminal segments were refined by an initial minimization with steepest descents followed by conjugate gradient methods to a gradient convergence of 0.001 kcal/mole/Å. A simulated annealing (SA) procedure was then carried out wherein all degrees of freedom for these regions were allowed to relax, but the heavy atoms of all other residues (TM1 to TM7) were held rigid. The protocol used for SA involved a slow heating to 600 K in steps of 100 K, followed by slow cooling to 300 K, for a period of 25 picoseconds at every temperature step. The lowest energy structure from the 300 K trajectory was then subjected to a final round of minimization, with all heavy atoms tethered by a force constant of 100 kcal/mole/Å^2. The minimization was done using 10,000 steps each of steepest descents and conjugate gradients.

The bond lengths, bond angles, torsions (*omega*, *phi* and *psi*) and chirality of the Cα atoms in the model structure were analyzed with the *ProStat* submodule of *Homology* (Accelrys Inc., USA). The *phi-psi* map (*i.e.*, Ramachandran plot) of the final model was constructed.

2.4 Lipophilic Potential

The lipophilic potential (LP) of the final model was calculated and visualized using the *MOLCAD* program as implemented in the *SYBYL* molecular modeling package (Tripos Inc., USA) [31]. Charges to atoms were assigned by the Gasteiger-Hückel method. The electron density isosurface of NK1R model was calculated with a cut value of 0.003 and step width of 0.5 Å, over which the LP surface was generated and visualized. The LP was calculated on the basis of Crippen's atomic partial lipophilicities [32], which are corrected for proteins.

3 RESULTS AND DISCUSSION

The overall percentage sequence identity between bovine rhodopsin and other members of the family A GPCRs is as low as 20%. Normally, when the sequence identity between the model and the template is below 30%, the sequence alignment becomes the main bottleneck in the homology modeling procedure. However, the family A GPCRs are the exception to this rule, as the sequence homology in the transmembrane regions of the GPCRs with bovine rhodopsin is significantly higher. Therefore, a strategy of segmented alignment was used to overcome the problems arising due to low overall sequence identity. The transmembrane residues were modeled using the crystal structure of bovine rhodopsin, while the loop and end regions were modeled separately.

The Brookhaven Protein Data Bank currently contains four structures of bovine rhodopsin solved by X-ray crystallography (PDB codes: 1F88, 1HZX, 1L9H and 1GZM) at different resolutions (2.8 Å, 2.8 Å, 2.6 Å and 2.65 Å respectively) and two solution structures (PDB codes: 1JFP and 1LMX) solved using NMR. In addition to these, the PDB also contains two theoretical

models [33] of bovine rhodopsin (PDB codes: 1BOJ and 1BOK), which were previously [34] used for GPCR modeling. The NMR structures are not ideal for homology modeling, as they are an *ensemble* of structures instead of a single representative structure and have several chirality errors. Therefore, the PDB structure 1L9H with the highest resolution was chosen as the reference protein for the NK1R modeling.

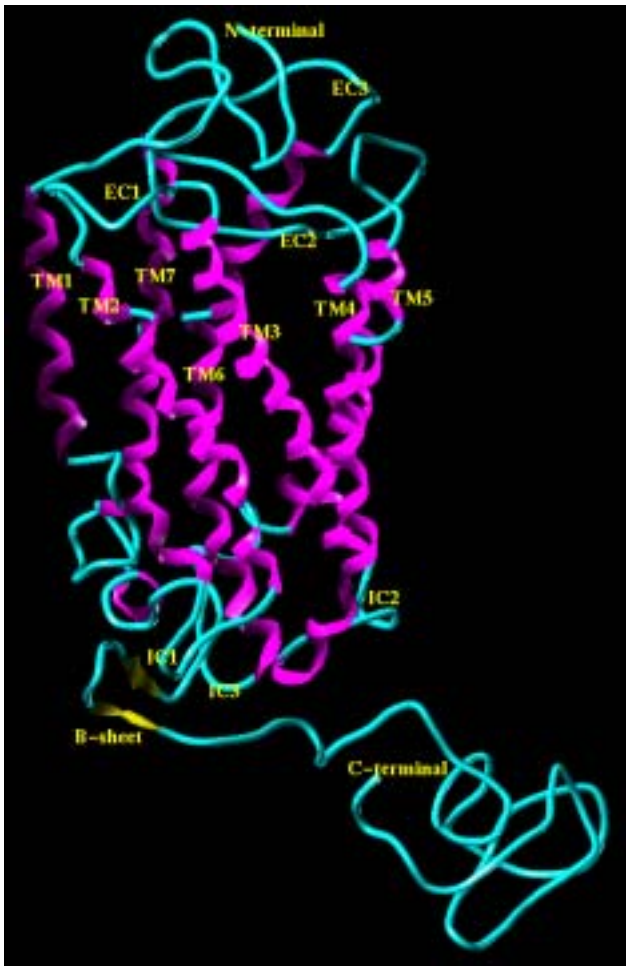

Figure 2. The topology of human NK1R model generated by homology modeling using segmented approach. The transmembrane helices are shown in magenta and the rest of structure (ICs, ECs, N- and C- terminals) is shown in cyan. The long C-terminus folds to form a globular structure. A β-sheet is within the start of the C-terminus tail.

Transmembrane Helices: The complete topology of the NK1R receptor is shown in Figure 2. A comparison with the bovine rhodopsin structure (Figure 3) shows that the global arrangement of the TMs is roughly maintained. However, differences between the primary sequences of the NK1 receptor and bovine rhodopsin lead to some differences between their 3D structures. The seven TM helices are arranged in an anticlockwise fashion. A small kink is observed in TM1 at residues Val48 through Val51 just before the start of IC1. TM1, TM2 and TM3 run almost parallel to each other but diverge near their IC ends, and TM4 angles out with TM3. TM6 and TM7 deviate from the rhodopsin structure. TM7 has a kink running from residues Ile259 to Trp261, which is also seen

in rhodopsin; but after this kink, the helix axis deviates in NK1R than in the template structure. TM6 and TM7 progress parallel to TM1, and the seven TMs together form a circular bundle of seven helices within the membrane.

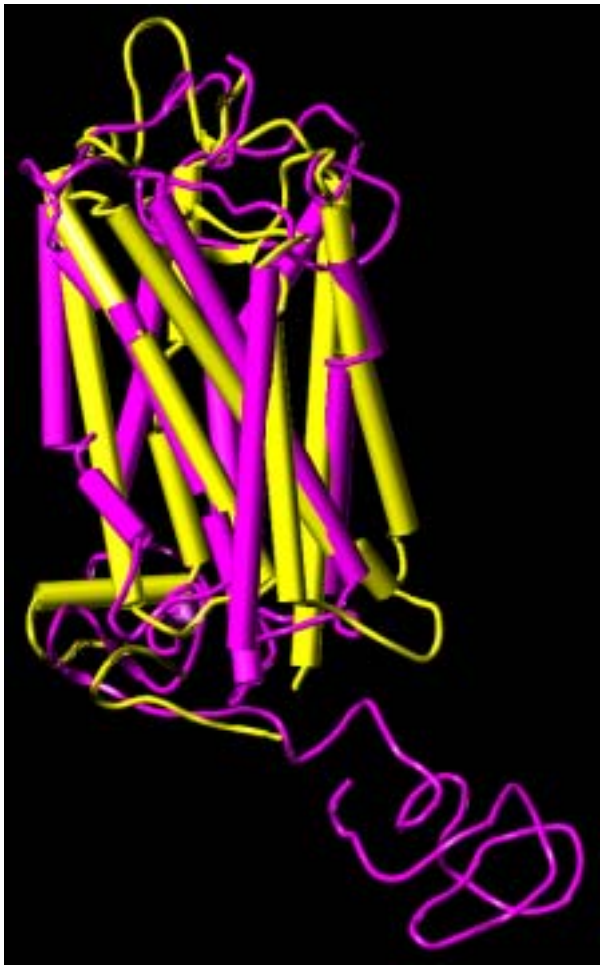

Figure 3. A Superposition of the human NK1R model (magenta) and the bovine rhodopsin (yellow) crystal structure (PDB 1L9H).

The N-terminal region in NK1R model (although extracted from the template structure) does not adopt a β-sheet configuration like 1L9H after refinement with simulated annealing. The C-terminal region after optimization by SA forms a small antiparallel β-sheet, spanning segments Thr339 to Arg340 and Gly346 to Ser347. The rest of the C-terminal residues folds into a globular structure. The ICs embedded in the cytoplasmic region of the cell have crucial structures as they are linked to the G-protein to whom the message must be conveyed for elucidation of the final response.

Extracellular Region: The antiparallel β-sheet found between segments Tyr178 to Glu181 and Ser186 to Ile189 in EC2 in 1L9H, does not exist in EC2 of the NK1R model, even though the loops are of equal length. Except for EC2, all other ECs in the NK1R model are longer than the corresponding ones in bovine rhodopsin. The other major topological feature of the extracellular region of the NK1R is the presence of a disulfide bond connecting Cys105 in the extracellular end

of TM3 (*i.e*. EC1) with Cys180 in the EC2 loop. This linkage is observed in many receptors with the seven-TM motif and, in some cases, is the most conserved feature in distantly related receptors. In NK1R, the disulfide bond is seen to be essential for binding to SP with high affinity [35]. Interestingly, engineering additional cysteines into NK1R does not disrupt the wild type disulfide bond [36]. The presence of the disulfide bond creates two pseudo-loops out of EC2, the first connects TM3 to TM4 and the second connects TM3 to TM5. This ultimately affects the relative orientation of TM4 and TM5 with respect to TM1, TM2 and TM3 as discussed above. Also, importantly, many of the residues implicated in SP binding by mutational or photoaffinity labeling studies are in close proximity of this disulfide bond [35].

Lipophilic Potential: The lipophilic potential (LP) is an indicator of lipophilic and hydrophilic centers. The GPCRs are well known transmembrane proteins, with their TM regions embedded in the cell membrane built of phospholipids arranged in a bilayer. The interactions of the TMs with the lipid bilayer are of hydrophobic origin, responsible for stabilizing the TM regions in the membrane. The extracellular and intracellular regions of the GPCRs are less lipophilic (rather, more hydrophilic) in nature, being exposed to the extracellular and the intracellular aqueous fluid, respectively. Figure 4 shows the LP plotted on the electron density surface of the NK1R model. The potential varies from + 0.16 (lipophilic, brown in color) to – 0.14 (hydrophilic, blue in color). The green regions indicate a neutral potential. A large fraction of the region shaded in brown (lipophilic) is located in the TM segments, whereas most of the regions colored blue (hydrophilic) are located outside the TM regions. The most hydrophilic surface is that of the intracellular C-terminal, which is folded into a globular structure.

4 CONCLUSIONS

Using the recently published high-resolution crystal structure of bovine rhodopsin, we have modeled the N-terminal and seven TM regions of NK1R. The loops and the C-terminal of the receptor protein were built with fragments from the PDB library or with the "loop generate" algorithm and refined using simulated annealing coupled with stringent minimization protocols. The complete model was assessed for geometry and energetics. The *phi* and *psi* dihedrals of all residues in the NK1R are present within the allowed region of the Ramachandran plot. The model preserves the overall topology of the family A GPCR. The lipophilic potential surface of the NK1R shows that hydrophobic residues are present within the TM region, consistent with the fact that these residues are present in the hydrophobic environment of the lipid bilayer. The other regions *i.e.* loops and end regions are relatively hydrophilic, being surrounded by an aqueous environment the extra or intracellular fluid. This work provides the first complete model of human NK1 receptor making possible the study of ligand – GPCR as well as G-protein – GPCR interactions at the atomic level.

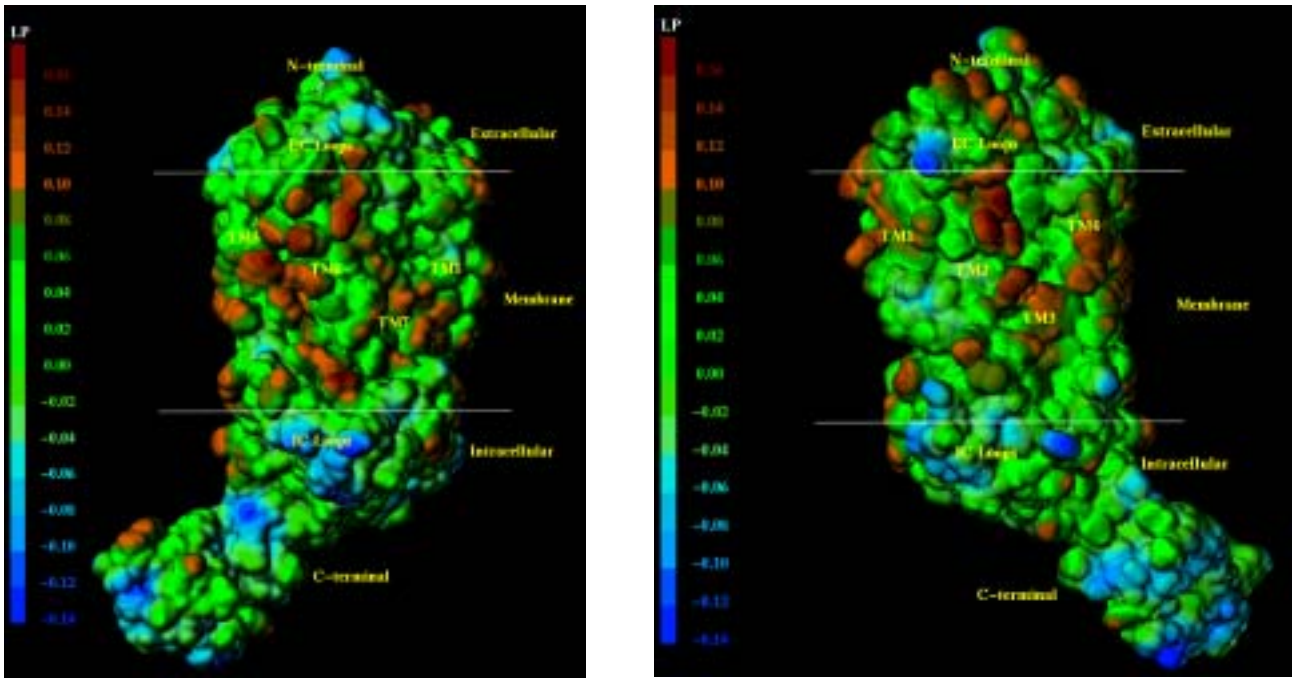

Figure 4: The Lipophilic Potentials calculated for the NK1R model. The figure to the left shows the view of TM6, TM7 and TM1 while the figure to the right displays the TM1, TM2, TM3 and TM4 segments. A deep brown color identifies lipophilic regions, while a deep blue color indicates hydrophilic regions. The inset strip is the color spectrum along with the potential values showing gradation of the LP values.

Acknowledgment

This work was possible by a grant (SR/FST/LSI-083/2003) from the Department of Science and Technology (DST), through their FIST program to ECC. SAK thanks the Lady Tata Memorial Trust, Mumbai, and AKM the University Grants Commission (UGC), New Delhi, for financial support.

5 REFERENCES

[1] O. M. Becker, S. Sachman, Y. Marantz and S. Noiman, Modeling the 3D structure of GPCRs: Advances and applications to drug discovery, *Curr. Opin Drug Discov. Develop.* **2003**, *6*, 353-361.

[2] T. Klabunde and G. Hessler, Drug design strategies for targeting G-protein-coupled receptors, *ChemBioChem* **2002**, *3*, 928-944.

[3] H. Bourne, How receptors talk to trimeric G protein, *Curr. Opin. Cell. Biol.* **1997**, *9*, 134-142.

[4] J. Wess, G-protein coupled receptors: molecular mechanisms involved in activation and selectivity of G-protein recognition, *FASEB J.* **1997**, *11*, 346-354.

[5] H. Hamm, The many faces of G protein signaling, *J. Biol. Chem.* ***1998, 273,*** 669-672.

[6] J. Bockaert and J. P. Pin, Molecular tinkering of G protein-coupled receptors: an evolutionary success, *EMBO J.* **1999**, *18*, 1723-1729.

[7] T. K. Attwood, A compendium of specific motifs for diagnosing GPCR subtypes, *Trends Pharmacol. Sci.* **2001**, *22*, 162-165.

[8] Y. Takeda, K. B. Chou, J. Takeda, B. S. Sachais and J. E. Krause, Molecular cloning, structural characterization and functional expression of the human substance P receptor, *Biochem. Biophys. Res. Commun.* **1991,** *179*, 1232-1240.

[9] C. A. Maggi and T. W. Schwartz, The dual nature of the tachykinin NK1 receptor, *Trends Pharmacol. Sci.* **1997**, *18*, 351-355.

[10] P. Datar, S. Srivastava, E. Coutinho and G. Govil, Substance P: structure, function and therapeutics, *Curr. Top. Med. Chem.* **2004**, *4*, 75-103.

[11] S. McLean, Nonpeptide antagonists of the NK1 tacykinin receptor, *Med. Res. Rev.* **1996**, *16*, 297-317.

[12] P. L. R. Andrews and P. Bhandari, Resinferatoxin, an ultrapotent capsaicin analogue, has anti-emetic properties in ferret, *Neuropharmacology* **1993**, *32*, 799-806.

[13] F. D. Tattersall, W. Rycroft, R. J. Hargteaves and R. G. Hill, The tachykinin NK1 receptor antagonist CP-99,994 attenuates cisplatin induced emesis in ferret, *Eur. J. Pharmacol.* **1993**, *250*, R5-R6.

[14] D. Roper, E. Jacoby, P. Cruger, M. Engles, J. Grotzinger, A. Wollmer and W. Strassburger, Modeling G-protein coupled receptors with bacteriorhodopsin as a template. A novel approach based on interaction energy differences, *J. Recept. Res.* **1994**, *14*, 167-186.

[15] H. Luecke, B. Schobert, H. T. Richter, J. P. Cartailler and J. K. Lanyi, Structure of bacteriorhodopsin at 1.55 Å resolution, *J. Mol. Biol.* **1999**, *291*, 899-911.

[16] K. Palczewski, T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, I. Le Trong, D. C. Teller, T. Okada, R. E. Stenkamp, M. Yamamoto and M. Miyano, Crystal structure of rhodopsin: A G protein-coupled receptor, *Science* **2000**, *289*, 739-745.

[17] D. C. Teller, T. Okada, C. A. Behnke, K. Palczewski and R. E. Stenkamp, Advances in determination of a high-resolution three-dimensional structure of rhodopsin, a model of G protein coupled receptor, *Biochemistry* **2001**, *40*, 7761-7772.

[18] T. Okada, Y. Fugiyoshi, N. Cilow, J. Navarro, E. M. Landau, Y. Shichida, Functional role of internal water molecules in rhodospin revealed by X-ray crystallography, *Proc. Natl. Acad. Sci USA.* **2002**, *99*, 5982-5987.

[19] J. Li, P. Edwards, M. Burghammer, C. Villa, G. F .X. Schertler, Structure of rhodopsin in a trigonal crystal form, *J. Mol. Biol.* **2004**, *343*, 1409-1438.

[20] P. L. Yeagle, G. Choi, A. D. Albert, Studies on the structure of the G protein-coupled receptor rhodopsin including the putative G-protein binding site in unactivated and activated forms, *Biochemistry* **2001**, *40*, 11932-11937.

[21] G. Choi, J. Landin, J. F. Galan, R. R. Birge, A. D. Albert, and P. L. Yeagle, Structural studies of metarhodopsin-II, the activated form of the G-protein coupled receptor, rhodospsin, *Biochemistry* **2002**, *41*, 7318-7324.

[22] InsightII, version 98, Accelrys Inc., San Diego, CA, USA

[23] S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389-3402.

[24] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, The protein data bank. *Nucleic Acids Res.* **2000**, *28*, 235-242.

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

[26] M. O. Dayhoff, W. C. Barker and L. T. Hunt, Establishing homologies in protein sequences, *Methods Enzymol.* **1983**, *91*, 524-545.

[27] U. Hobohm, M. Scharf, R. Schneider and C. Sander, Selection of representative protein data sets, *Protein Science* **1992**, *1*, 409-17.

[28] P. S. Shenkin, D. L. Yarmush, R. M. Fine, H. Wang and C. Levinthal, Predicting antibody hypervariable loop conformation. I. Ensembles of random conformations for ringlike structures, *Biopolymers* **1987**, *26*, 2053-85.

[29] P. Dauber-Osguthorpe, V. A. Roberts, D. J. Osguthorpe, J. Wolff, M. Genest and A. T. Hagler, Structure and energetics of ligand binding to proteins: *Escherichia coli* dihydrofolate reductase-trimethoprim, a drug-receptor system, *Proteins* **1988**, *4*, 31-47.

[30] U. Gether, Uncovering molecular mechanisms involved in activation of G protein-coupled receptors, *Endocrine Rev.* **2000**, *21*, 90-113.

[31] Sybyl, version 6.7, Tripos Inc., St. Louis, MO, USA.

[32] A. Ghose and G. Crippen, Atomic physicochemical parameters for three-dimensional structure-directed quantitative structure-activity relationships i. partition coefficients as a measure of hydrophobicity, *J Comput. Chem.* **1986**,*7*, 565-577.

[33] I. D. Pogozheva, A. L.Lomize and H. I. Mosberg, Transmembrane 7 alpha-bundle of rhodospin: distance geometry calculation with hydrogen bonding constraints, *Biophys. J.* **1997**, *72*, 1963-1985.

[34] A. J. W. Orry, B. A. Wallace, Modeling and docking the endothelin G-protein-coupled receptor, *Biophys. J.***2000**, *79*, 3083-3094.

[35] N. D. Boyd, R. Kage, J. J. Dumas, J. E. Krause and S. E. Leeman, The peptide binding site of substance P (NK-1) receptor localized by a photoreactive analogue of substance P: presence of a disulfide bond, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 433-437.

[36] C. E. Elling, U. Raffetesder, S. M. Nielsen and T. W. Schwartz, Disulfide bridge engineering in the tachykinin NK1 receptor, *Biochemistry* **2000**, *39*, 667-675.

Biographies

Santosh A Khedkar is a Ph.D. researcher at the Bombay College of Pharmacy, under the guidance of Prof. Evans Coutinho. He has pursued an M. Pharm. Sci. degree in Pharmaceutical Chemistry at the University Institute of Chemical Technology (UICT), Mumbai. His research interests include homology modeling, docking studies, new QSAR methods and *ab initio* studies in peptide design. He has presented papers in several conferences and published in reputed international journals.

Alpeshkumar K Malde is a Ph.D. researcher at the Bombay College of Pharmacy, under the guidance of Prof. Evans Coutinho. His research interests include applications of *ab initio* calculations, protein modeling, new QSAR methods, solution structures of peptides by 2D-NMR and design of novel boron peptides. He has presented / published papers in reputed conferences and journals. He is a recipient of the Indian Pharmaceutical Association (IPA) and National Institute of Pharmaceutical Education and Research (NIPER) Gold Medals

Evans C Coutinho is Professor of Pharmaceutical Chemistry at the Bombay College of Pharmacy. He received his PhD degree from the University of Mumbai and Masters and Bachelors degree in Pharmaceutical Technology from the University Institute of Chemical Technology (UICT). His research interests are molecular modeling and computational chemistry in the areas of QSAR, docking, structure-based drug design, quantum mechanical calculations, and multinuclear and multidimensional NMR of proteins and peptides. He has received the Cipla Distinguished Fellowship Award, the UGC Research Award, and the Career Award for Young Teachers from the All India Council of Technical Education.