2013 International Conference on Food and Agricultural Sciences
IPCBEE vol.55 (2013) © (2013) IACSIT Press, Singapore
DOI: 10.7763/IPCBEE. 2013. V55. 1

Homology Modeling of the Chimeric Human Sweet Taste Receptors Using Multi Templates

Ragheed Hussam Yousif [1+]

[1] Malaysia-Japan International Institute of Technology (MJIIT), University Teknologi Malaysia, Kuala Lumpur

Abstract. The sweet taste perception is mainly sensed by T1R2 and T1R3 human sweet taste receptors, which belong to the super family of G protein coupled receptors (GPCR). However, there is yet a clear study to describe the binding modes of T1R2 and T1R3. Therefore, further experimental and the computational data is needed to understand more about the GPCR, especially for the homology modeling as it is important to reduce the gap between the protein structures and sequences. In this research, 3MQ4 and 2E4U were selected as templates for the chimeric T1R2 and T1R3. MODELLER V9.10 was used to create the 3D structure of the target sequences, and finally the Ramachandran plot evaluations showed that 83% of the residues are located in the most favoured regions for the chimeric model.

Keywords: Homology modelling, human sweet taste receptors, MODELLER

1. Introduction

The sweet taste perception in human being is able to be sensed by T1R2 and T1R3 sweet taste receptors, they are heterodimeric belong to TR Family closely related to G protein coupled receptors (GPCR) [1], [2], which are super family of protein expressed on the eukaryotic cell membrane to function as sensor for several extracellular substances [3]. T1R2 and T1R3 are capable to recognize all different kinds of sweet substances, such as sugars, artificial sweeteners, amino acids, and sweet proteins [4], since they compose various ligand binding sites [5].

However, there is no clear study able to describe the binding ability of T1R2 and T1R3 with several ligands [5], which create a challenge in understanding the binding modes of the GPCR. Therefore, further experimental and computational data is required for discovering the GPCR. For the experimentally solved protein structure, it is necessary to provide a comparable template for unsolved protein structure, in order to perform the Homology modeling [3].

The homology modeling role is to reduce the gap between the proteins solved structure and protein primary sequences, in order to utilize the protein resources to understand the protein function [6], [7].

The protein structure prediction problem can be classified into three different dimensional levels, which are: (1D, 2D and 3D) dimensional levels. 1D dimensional level depicts the prediction of secondary structure and other protein structural topologies, and the prediction of spatial relationships between two amino acids belongs to the 2D dimensional level, and finally the prediction of three dimensional coordinates of each amino acid in the target protein belongs to the 3D dimensional level, which is the most important aspect of the protein structure prediction [6].

Although the multiple templates homology modeling is more complicated than the single template, the multiple templates method is beneficial to produce more reliable model, because of its capability to enhance the possibility of giving better template and it is qualified to cover more of the target sequence [8].

[+] Corresponding author: Tel.: + 603-2203-1200; fax: + 603-2203-1266.
E-mail address: hyragheed3@live.utm.my.

Moreover, the quality of the protein structure produced by Homology modeling depends on the similarity between the target sequence and the template. For instance, if the similarity is more than 50% it may produce high quality models, but if the similarity is less than 30% the produced model may probably contain significant errors [9]. Figure 1 shows the process of homology modelling, which includes template selection, sequence alignment, model building, and finally model evaluation.

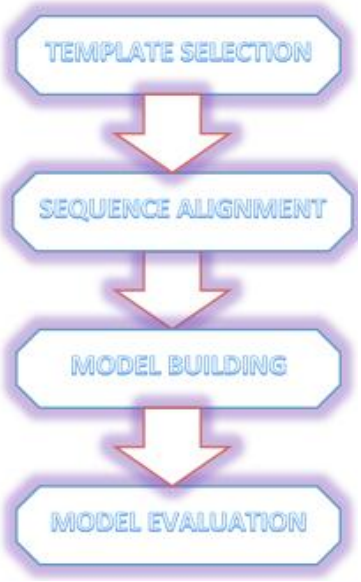

Fig. 1: Flow chart of the structure prediction process of chimeric human sweet taste receptors

2. Methodology

The chimeric or fusion of human sweet taste proteins T1R2 and T1R3 was prepared by overlapping each of T1R2 and T1R3 primary sequences, so the target sequence of chimeric human sweet taste receptors in fasta format was:

>chimera:|SEQUENCE
SDFYLPGDYLLGGLFSLHANMKGIVHLNFLQVPMCKEYEVKVIGYNLMQAMRFAVEEINNDSSLLP
GVLLGYEIVDVCYISNNVQPVLYFLAHEDNLLPIQEDYSNYISRVVAVIGPDNSESVMTVANFLSLFL
LPQITYSAISDELRDKVRFPALLRTTPSADHHIEAMVQLMLHFRWNWIIVLVSSDTYGRDNGQLLGE
RVARRDICIAFQETLPTLQPNQNMTSEERQRLVTIVDKLQQSTARVVVVFSPDLTLYHFFNEVLRQN
FTGAVWIASESWAIDPVLHNLTELRHLGTFLGITIQSVPIPGFSEFREWGPQAGPPPLSRTSQSYTCNQ
ECDNCLNATLSFNTILRLSGERVVYSVYSAVYAVAHALHSLLGCDKSTCTKRVVYPWQLLEEIWKV
NFTLLDHQIFFDPQGDVALHLEIVQWQWDRSQNPFQSVASYYPLQRQLKNIQDISWHTINNTIPMSM
CSKRCQSGQKKKPVGIHVCCFECIDCLPGTFLNHTEDEYECQACPNNEWSYQSETSCFKRQLVFLE
WHEAPTIAVALLAALGFLSTLAILVIFWRHMLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYV
LGGLFPLGEAEEAGLRSRTRPSSPVCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTC
SEPVVAMKPSLMFLAKAGSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGA
SMELLSARETFPSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGIC
IAHEGLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEA
WLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQGLEEDVV
GQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDPVKPWQLLENM
YNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRTERLKIRWHTSDNQKP
VSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPERSTRCFRRRSRF
LA

The templates were searched by using Basic Local Alignment Search Tools (BLAST), to find out the similar templates to the target sequences [10]. Then the uncovered residues by the template were removed, and phylogenetic analysis was done, using MEGA5 [11] as shown in Figure 2, to locate the neighboring template in PDB according to their Phylogenetic evaluation together with the highest score template in BLAST search. The uncovered residues were from 1to 24, 549 to 844, and 1410 to 1691. ClustalW program was used to align between the target and the templates sequences [12]

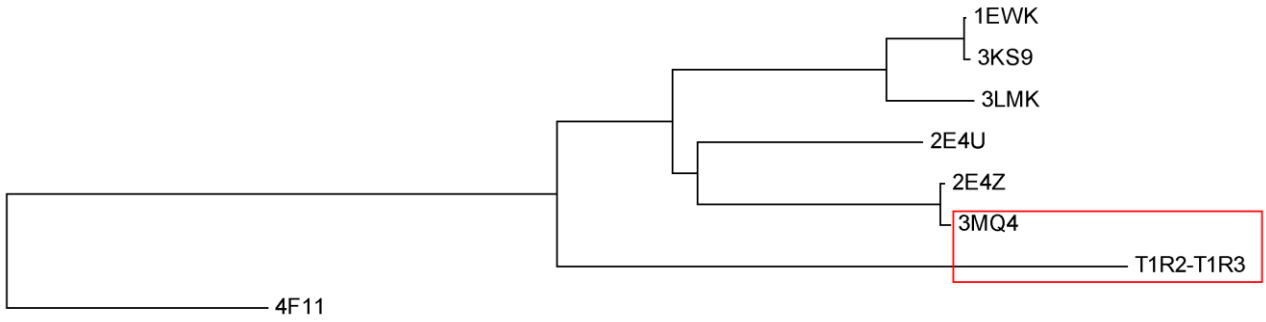

Fig. 2: The phylogenetic analysis for the chimeric T1R2 and T1R3

The three dimensional model was generated using MODELLER v9.10 [13] for the chimeric sequences, and the Ramachandran plot has been chosen to evaluate the model with lowest energy [14].

3. Results

The results show that the neighbors' templates were Metabotropic glutamate receptor mGluR7 complexed with LY341495 antagonist (3MQ4), and its fasta format follows:

>3MQ4:A|PDBID|CHAIN|SEQUENCEGAMDMYAPHSIRIEGDVTLGGLFPVHAKGPSGVPCGDIKREN
GIHRLEAMLYALDQINSDPNLLPNVTLGARILDTCSRDTYALEQSLTFVQALIQKDTSDVRCTNGEPP
VFVKPEKVVGVIGASGSSVSIMVANILRLFQIPQISYASTAPELSDDRRYDFFSRVVPPDSFQAQAMV
DIVKALGWNYVSTLASEGSYGEKGVESFTQISKEAGGLSIAQSVRIPQERKDRTIDFDRIIKQLLDTPN
SRAVVIFANDEDIKQILAAAKRADQVGHFLWVGSDSWGSKINPLHQHEDIAEGAITIQPKRATVEGF
DAYFTSRTLENNRRNVWFAEYWEENFNCKLTISGSKKEDTDRKCTGQERIGKDSNYEQEGKVQFVI
DAVYAMAHALHHMNKDLCADYRGVCPEMEQAGGKKLLKYIRNVNFNGSAGTPVMFNKNGDAPG
RYDIFQYQTTNTSNPGYRLIGQWTDELQLNIEDMQWGK

The highest score template in BLAST search was the Crystal structure of the extracellular region of the group II metabotropic glutamate receptor complexed with L-glutamate (2E4U) and its fasta format follows:

>2E4U:A|PDBID|CHAIN|SEQUENCE
DHNFMRREIKIEGDLVLGGLFPINEKGTGTEECGRINEDRGIQRLEAMLFAIDEINKDNYLLPGVKLG
VHILDTCSRDTYALEQSLEFVRASLTKVDEAEYMCPDGSYAIQENIPLLIAGVIGGSYSSVSIQVANL
LRLFQIPQISYASTSAKLSDKSRYDYFARTVPPDFYQAKAMAEILRFFNWTYVSTVASEGDYGETGIE
AFEQEARLRNICIATAEKVGRSNIRKSYDSVIRELLQKPNARVVVLFMRSDDSRELIAAANRVNASFT
WVASDGWGAQESIVKGSEHVAYGAITLELASHPVRQFDRYFQSLNPYNNHRNPWFRDFWEQKFQC
SLQNKRNHRQVCDKHLAIDSSNYEQESKIMFVVNAVYAMAHALHKMQRTLCPQTTKLCDAMKI
LDGKKLYKEYLLKIQFTAPFNPNKGADSIVKFDTFGDGMGRYNVFNLQQTGGKYSYLKVGHWAET
LSLDVDSIHWSRNSVPTSQCSDPCAPNEMKNMQPGDVCCWICIPCEPYEYLVDEFTCMDCGPGQWP
TADLSGCYNLPEDYIKWEDALVPR

The target and the templates alignment results to ClustalW program shown in Figure 3. The 3D model of chimeric Sweet Taste Receptors is shown in Figure 4, and The Ramachandran plot analysis, which is 83.9% of the residues located in the most favoured regions as shown in Figure 5.

4. Conclusion

The main purpose of this research is to achieve a multi template homology modeling for a chimeric T1R2 and T1R3 human sweet taste receptors, by overlapping their own primary sequences, selecting the closest templates, building the 3D model, and finally performing the model evaluation.

5. Acknowledgements

The authors would like to thank Malaysia - Japan International Institute of Technology (MJIIT)-Universiti Teknologi Malaysia Kuala Lumpur for supporting this research.

```
gi|116242831|sp|Q8TE23.2|Chime     DIACTFCGQDEWSPERSTRCFRRRSRFLAGAMDMYAPHSIRIEGDVTLGG 1150
3MQ4_A|PDBID|CHAIN|SEQUENCE        ------------------------------GAMDMYAPHSIRIEGDVTLGG 21
2E4U_A|PDBID|CHAIN|SEQUENCE        -----------------------------------DHNFMRREIKIEGDLVLGG 19
                                                                      . :  :.*:****:.***

gi|116242831|sp|Q8TE23.2|Chime     LFPVHAKGPSGVPCGDIKRENGIHRLEAMLYALDQINSDPNLLPNVTLGA 1200
3MQ4_A|PDBID|CHAIN|SEQUENCE        LFPVHAKGPSGVPCGDIKRENGIHRLEAMLYALDQINSDPNLLPNVTLGA 71
2E4U_A|PDBID|CHAIN|SEQUENCE        LFPINEKGTGTEECGRINEDRGIQRLEAMLFAIDEINKDNYLLPGVKLGV 69
                                   ***:: **..   ** *:.:.**:******:*:*:**.*  ***.*.**.

gi|116242831|sp|Q8TE23.2|Chime     RILDTCSRDTYALEQSLTFVQA-LIQKDTSDVRCTNGEPPVFVK-PEKVV 1248
3MQ4_A|PDBID|CHAIN|SEQUENCE        RILDTCSRDTYALEQSLTFVQA-LIQKDTSDVRCTNGEPPVFVK-PEKVV 119
2E4U_A|PDBID|CHAIN|SEQUENCE        HILDTCSRDTYALEQSLEFVRASLTKVDEAEYMCPDGSYAIQENIPLLIA 119
                                   :***************** **:* * : * ::  *.:*. .:  : *  :.

gi|116242831|sp|Q8TE23.2|Chime     GVIGASGSSVSIMVANILRLFQIPQISYASTAPELSDDRRYDFFSRVVPP 1298
3MQ4_A|PDBID|CHAIN|SEQUENCE        GVIGASGSSVSIMVANILRLFQIPQISYASTAPELSDDRRYDFFSRVVPP 169
2E4U_A|PDBID|CHAIN|SEQUENCE        GVIGGSYSSVSIQVANLLRLFQIPQISYASTSAKLSDKSRYDYFARTVPP 169
                                   ****.* ***** ***:****************:.:***. ***:*:*.***

gi|116242831|sp|Q8TE23.2|Chime     DSFQAQAMVDIVKALGWNYVSTLASEGSYGEKGVESFTQISKEAGGLSIA 1348
3MQ4_A|PDBID|CHAIN|SEQUENCE        DSFQAQAMVDIVKALGWNYVSTLASEGSYGEKGVESFTQISKEAGGLSIA 219
2E4U_A|PDBID|CHAIN|SEQUENCE        DFYQAKAMAEILRFFNWTYVSTVASEGDYGETGIEAFEQEARLRN-ICIA 218
                                   * :***:**.:*:: :.*.****:****.***.*:*:* * ::  . :.**

gi|116242831|sp|Q8TE23.2|Chime     QSVRIPQERKDRTIDFDRIIKQLLDTPNSRAVVIFANDEDIKQILAAAKR 1398
3MQ4_A|PDBID|CHAIN|SEQUENCE        QSVRIPQERKDRTIDFDRIIKQLLDTPNSRAVVIFANDEDIKQILAAAKR 269
2E4U_A|PDBID|CHAIN|SEQUENCE        TAEKVG--RSNIRKSYDSVIRELLQKPNARVVVLFMRSDDSRELIAAANR 266
                                    : ::   *.:   .:* :*::**:.**:*.**:* ..:* ::::***:*

gi|116242831|sp|Q8TE23.2|Chime     ADQVGHFLWVGSDSWGSKINPLHQHEDIAEGAITIQPKRATVEGFDAYFT 1448
3MQ4_A|PDBID|CHAIN|SEQUENCE        ADQVGHFLWVGSDSWGSKINPLHQHEDIAEGAITIQPKRATVEGFDAYFT 319
2E4U_A|PDBID|CHAIN|SEQUENCE        VNAS--FTWVASDGWGAQESIVKGSEHVAYGAITLELASHPVRQFDRYFQ 314
                                   .:    * **.**.**:: . ::  *.:* *****::    .*. ** **

gi|116242831|sp|Q8TE23.2|Chime     SRTLENNRRNVWFAEYWEENFNCKLTISGSKKEDTDRKCTGQERIGKDSN 1498
3MQ4_A|PDBID|CHAIN|SEQUENCE        SRTLENNRRNVWFAEYWEENFNCKLTISGSKKEDTDRKCTGQERIGKDSN 369
2E4U_A|PDBID|CHAIN|SEQUENCE        SLNPYNNHRNPWFRDFWEQKFQCSLQN----KRNHRQVCDKHLAID-SSN 359
                                   * .  **:** ** ::**::*:*.*       *.:  : *  :  *. .**

gi|116242831|sp|Q8TE23.2|Chime     YEQEGKVQFVIDAVYAMAHALHHMNKDLCADYRGVCPEMEQAGGKKLLK- 1547
3MQ4_A|PDBID|CHAIN|SEQUENCE        YEQEGKVQFVIDAVYAMAHALHHMNKDLCADYRGVCPEMEQAGGKKLLK- 418
2E4U_A|PDBID|CHAIN|SEQUENCE        YEQESKIMFVVNAVYAMAHALHKMQRTLCPQTTKLCDAMKILDGKKLYKE 409
                                   ****.*: **::***********:*::: **.:   :*  *:  .**** *

gi|116242831|sp|Q8TE23.2|Chime     YIRNVNFNG------SAGTPVMFNKNGDAPGRYDIFQYQTTNTSNPGYRL 1591
3MQ4_A|PDBID|CHAIN|SEQUENCE        YIRNVNFNG------SAGTPVMFNKNGDAPGRYDIFQYQTTNTSNPGYRL 462
2E4U_A|PDBID|CHAIN|SEQUENCE        YLLKIQFTAPFNPNKGADSIVKFDTFGDGMGRYNVFNLQQTGG-KYSYLK 458
                                   *: :::*..       .*.: * *:. **. ***::*: * *.  : .*

gi|116242831|sp|Q8TE23.2|Chime     IGQWTDELQLNIEDMQWGK---------------------------------- 1610
3MQ4_A|PDBID|CHAIN|SEQUENCE        IGQWTDELQLNIEDMQWGK---------------------------------- 481
2E4U_A|PDBID|CHAIN|SEQUENCE        VGHWAETLSLDVDSIHWSRNSVPTSQCSDPCAPNEMKNMQPGDVCCWICI 508
                                   :*:*:: *.*:::.::*.:

gi|116242831|sp|Q8TE23.2|Chime     ------------------------------------------------
3MQ4_A|PDBID|CHAIN|SEQUENCE        ------------------------------------------------
2E4U_A|PDBID|CHAIN|SEQUENCE        PCEPYEYLVDEFTCMDCGPGQWPTADLSGCYNLPEDYIKWEDALVPR 555
```

Fig. 3: The ClustalW alignment between the target and the templates sequences

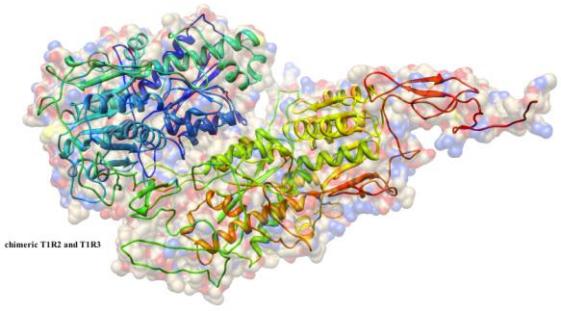

Fig. 4: The 3D structure for chimeric T1R2 and T1R3 using multi templates

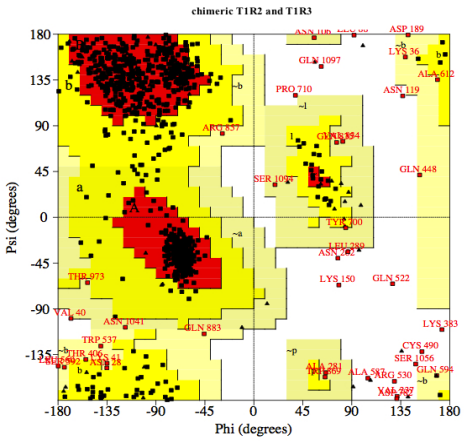

Fig. 5: Ramachandran Plot of the chimeric T1R2 and T1R3 using multi templates

6. References

[1] K. Ohta, et al., "Introduction of a negative charge at Arg82 in thaumatin abolished responses to human T1R2–T1R3 sweet receptors," *Biochemical and Biophysical Research Communications,* 2011, **413** (1): 41-45.

[2] F. Zhang, et al., "Molecular mechanism of the sweet taste enhancers," vol. 107, Washington, DC, ETATS-UNIS*: National Academy of Sciences*, 2010.

[3] S. Costanzi, Modeling G protein-coupled receptors and their interactions with ligands. Current opinion in structural biology, 2013.

[4] J. Chandrashekar, et al., "The receptors and cells for mammalian taste," *Nature*, 2006, **444** (7117): 288-294.

[5] K. Masuda, et al., "Characterization of the modes of binding between human sweet taste receptor and low-molecular-weight sweet compounds," *PLoS One*, 2012, **7**(4): e35380.

[6] J. Cheng, et al., "The MULTICOM toolbox for protein structure prediction," *BMC Bioinformatics,* 2012, **13** (1): 65.

[7] M. J. Sippl, "Recognition of errors in three-dimensional structures of proteins," *Proteins: Structure, Function, and Bioinformatics*, 1993, **17** (4): 355-362.

[8] J. Ko, H. Park, and C. Seok, "GalaxyTBM: template-based modeling by building a reliable core and refining unreliable local regions," *BMC bioinformatics*, 2012, **13** (1): 198.

[9] C. A. Floudas, et al., "Advances in protein structure prediction and de novo protein design: A review," *Chemical Engineering Science*, 2006, **61** (3): 966-988.

[10] S. F. Altschul, et al., "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs," *Nucleic Acids Research*, 1997, **25** (17): 3389-3402.

[11] K. Tamura, et al., "MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods," *Molecular Biology and Evolution*, 2011, **28** (10): 2731-2739.

[12] K. B. Li, "Clustalw-MPI: clustalw analysis using distributed and parallel computing," *Bioinformatics,* 2003, **19** (12): 1585-1586.

[13] A. Sali, et al., "Evaluation of comparative protein modeling by MODELLER," *Proteins: Structure, Function, and Bioinformatics,* 1995, **23** (3): 318-326.

[14] B. K. Ho and R. Brasseur, "The ramachandran plots of glycine and pre-proline,". *BMC structural biology*, 2005, **5**.