J Biosci (2019) 44:138
DOI: 10.1007/s12038-019-9944-9

Presence of carbohydrate binding modules in extracellular region of class C G-protein coupled receptors (C GPCR): An *in silico* investigation on sweet taste receptor

Elaheh Kashani-Amin[1], Amirhossein Sakhteman[2,3], Bagher Larijani[4] and Azadeh Ebrahim-Habibi[1]*

[1]*Biosensor Research Center, Endocrinology and Metabolism Molecular-Cellular Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran*

[2]*Department of Medicinal Chemistry, School of Pharmacy, Shiraz University of Medical Sciences, Shiraz, Iran*

[3]*Medicinal Chemistry and Natural Products Research Center, Shiraz University of Medical Sciences, Shiraz, Iran*

[4]*Endocrinology and Metabolism Research Center, Endocrinology and Metabolism Clinical Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran*

**Corresponding author (Email, aehabibi@sina.tums.ac.ir, azadehabibi@yahoo.fr)*

MS received 16 January 2019; accepted 8 July 2019; published online 25 October 2019

Sweet taste receptor (STR) is a C GPCR family member and a suggested drug target for metabolic disorders such as diabetes. Detailed characteristics of the molecule as well as its ligand interactions mode are yet considerably unclear due to experimental study limitations of transmembrane proteins. An *in silico* study was designed to find the putative carbohydrate binding sites on STR. To this end, α-D-glucose and its α-1,4-oligomers (degree of polymerization up to 14) were chosen as probes and docked into an ensemble of different conformations of the extracellular region of STR monomers (T1R2 and T1R3), using AutoDock Vina. Ensembles had been sampled from an MD simulation experiment. Best poses were further energy-minimized in the presence of water molecules with Amber14 forcefield. For each monomer, four distinct binding regions consisting of one or two binding pockets could be distinguished. These regions were further investigated with regard to hydrophobicity and hydrophilicity of the residues, as well as residue compositions and non-covalent interactions with ligands. Popular binding regions showed similar characteristics to carbohydrate binding modules (CBM). Observation of several conserved or semi-conserved residues in these binding regions suggests a possibility to extrapolate the results to other C GPCR family members. In conclusion, presence of CBM in STR and, by extrapolation, in other C GPCR family members is suggested, similar to previously proposed sites in gut fungal C GPCRs, through transcriptome analyses. STR modes of interaction with carbohydrates are also discussed and characteristics of non-covalent interactions in C GPCR family are highlighted.

Keywords. α-D-glucose; carbohydrate binding modules; CBM; oligomer; sweet taste receptor

Nomenclature and abbreviations used: 3Glc to 14Glc: Ligands, Alpha-D-glucose will be called glucose, disaccharide composed of two 1,4-α-D-glucoses units maltose and oligomers of 3 to 14 units of glucose will be called 3Glc to 14Glc. To clarify characteristics of atomic interactions of oligomers, each pyranose unit is numbered as follows, the first α-D-glucose with its C1 binded to C4 of the next glucose unit is numbered as 1 and others are numbered serially; for example, 3Glc-2 means the second gluco-pyranose unit of the tri-saccharide ligand. Other nomenclature adopted by YASARA (software used for docking, please see details in Methods section), Atoms of each molecule are shown as the name of the molecule or residue followed by the name of the atom; for example, 3Glc-2-O3 means atom O3 of 3Glc-2 and GLY 502-O means oxygen atom of G502. Core residues, residues in common between binding pockets located in each distinct region; EC, extracellular part of STR; H-Bond, hydrogen bond; HP-interaction, hydrophobic interaction; MDS, molecular dynamic simulations; STR, sweet taste receptor; TMD, trans-membrane domain

Electronic supplementary material: The online version of this article (https://doi.org/10.1007/s12038-019-9944-9) contains supplementary material, which is available to authorized users.

1. Introduction

Human sweet taste receptor (STR) is responsible for sensing the sweet taste of a wide range of chemical structures (Li *et al.* 2002). Components of STR heterodimer, T1R2 and T1R3 monomers, are members of class C G-protein coupled receptors (C GPCR) (Nelson *et al.* 2001). C GPCRs consist of three domains: the Venus flytrap domain or module (VFTM), which is the large extracellular (EC) part and responsible for ligand binding and activation of the receptor and consists of two lobes (LB1 and 2), the cysteine-rich domain (CR) responsible for transmitting conformational changes from extracellular domain to the transmembrane domain (TMD) (Muto *et al.* 2007; Kniazeff *et al.* 2011).

STR is expressed in gastrointestinal cells and functions as a chemosensor in this organ (Sternini *et al.* 2008), leading to glucose-stimulated glucagon-like peptide-1 (GLP-1) secretion (Rozengurt *et al.* 2006; Jang *et al*. 2007; Margolskee *et al.* 2007; Gerspach *et al.* 2011; Steinert *et al*. 2011; Xu *et al*. 2016; Wang *et al.* 2017). The relation between expression of mammalian STR components and glucose metabolism has been established by several experimental studies (Sclafani and Mann 1987; Nie *et al*. 2005; Sternini *et al.* 2008). The receptor appears as a promising therapeutic target (Laffitte *et al.* 2014) for the control of metabolic disorders such as diabetes and obesity due to suggested roles in controlling metabolic processes in various organs (e.g. glucose homeostasis and adipogenesis) (Behrens and Meyerhof 2011; Laffitte *et al*. 2014). Recently, a transcriptome sequence analyses and homology study on three strains of anaerobic gut fungi (Seppälä *et al.* 2016) has identified one hundred C GPCRs with non-canonical putative sugar binding domains.

Carbohydrate binding modules (CBM) are non-catalytic components of carbohydrate active enzymes (CAZymes) and fold into discrete functional units (Ficko-Blean and Boraston 2012; Abbott and van Bueren 2014). Several roles have been suggested for CBMs in various organisms including cell wall polysaccharides recognition by plant CBMs (Gilbert *et al.* 2013), human glycome recognition by microbial organisms (Ficko-Blean and Boraston 2012), or facilitation of glycoside hydrolases substrate association (Boraston *et al*. 2004).

Existence of carbohydrate binding sites on STR may bring forward new aspects of this molecule and be used in controlling metabolic disorders (e.g. diabetes and obesity).

STR is a single molecule that senses quite diverse chemical structures (Cui *et al.* 2006). Several studies have been performed to understand architecture as well as ligand binding modes of the STR and its key residues in ligand binding with the use of experimental methods such as point-directed mutations analysis, cell-based assays or saturation transfer difference NMR spectroscopy (Xu *et al.* 2004; Assadi-Porter *et al*. Assadi-Porter et al. 2010a; Assadi-Porter *et al*. Assadi-Porter et al. 2010b; Servant *et al*. 2010; Masuda *et al.* 2012a, b). However, experimental methods for GPCR studies are faced with severe limitations, due to instability of the protein when removed from the membrane (Congreve *et al.* 2015). As a substitute, computational methods could provide valuable insights in the study of GPCR characteristics (Hillisch 2004).

In this study, we have investigated putative carbohydrate binding sites on STR with molecular modeling techniques, using an improved version of the most recently introduced STR model by Chéron *et al.* (2017) along with α-D-glucose and its oligomers (up to 14 degrees of polymerization). While simple sugars (e.g. glucose and sucrose) have a sweet taste that is recognized by STR (Sclafani and Mann 1987; Nie *et al*. 2005), detection of longer oligomers appear to be STR-independent (Kochem 2017; Pullicin *et al*. 2017; Spector and Schier 2018). However, longer oligomers of α-D-glucose (up to 14 degrees of polymerization) were used as carbohydrate probes to explore possible large binding sites on the receptor. Actually, humans can taste glucose oligomers up to 14 degrees (Pullicin *et al*. 2017), and the receptor is also present in the gut. The hypothesis was that the binding sites located in the digestive tract might not activate signals to perceive sweet taste but may undergo other functions.

In an attempt to present the results as clearly as possible: first, an overall assessment of residue composition in detected binding pockets will be considered. Then, characteristics of each binding pocket (such as residue arrangement, binding energy characteristics and hydrophilicity/hydrophobicity of the binding regions) will be reported in details, considering the best poses for each ligand.

2. Materials and methods

2.1 *Receptor preparation*

The docking experiments were carried out on the extracellular (EC) region of T1R2 and T1R3 monomers. In order to obtain accurate conformations, the EC regions were sampled from 50 ns simulations of complete structures (EC+TMD) of T1R2 and T1R3. Then, extracellular regions were separated from TMD region for further use.

Homology modeling: T1R2 and T1R3 monomers were prepared as an improved model based on the one that has been previously introduced by Chéron *et al*. (2017). To improve the models, VFTM and CR domains were built separately and, further, assembled based on available literature data about the STR structure and the template structures (Muto *et al.* 2007). Model development and quality assessment of each constructed domain are discussed in Supplementary Information-Note 1.

MD simulation: models of each constructed MD simulation was performed for the complete structure (EC+TMD) of each monomer using YASARA2 force field (Krieger *et al.* 2009); a built-in force field incorporated in YASARA suite (Krieger and Vriend 2015). This forcefield was chosen based

on the fact that it has been optimized for structure prediction, refinement and energy minimization in addition to knowledge based-interaction potentials for homology models (Krieger *et al.* 2009). Protein side-chains pKa was predicted (Krieger *et al.* 2006) and protonation state was predicted at physiological pH 7.4. Sizes of the simulation cells were set as 20 Å larger than the protein, and filled with water and 0.9% NaCl (physiological solution) (Krieger *et al*. 2004). The main simulation was then run for 50 ns (2 × 25 ns) with Particle Mesh Ewald (PME) (Essmann *et al.* 1995) and 8.0 Å cutoff for non-bonded real space forces, a 4 fs time-step, constrained hydrogen atoms, and at constant pressure and temperature (NPT ensemble) (Krieger and Vriend 2015). Initial equilibration time was 250 ps which further was omitted from analysis. Snapshots were saved each 100 ps.

Cutoff and time step values were applied by YASARA forcefield based on a set of algorithms to boost MD simulation; explanation was provided in Supplementary Note-2 (Krieger and Vriend 2015).

Ensembles of sampled receptor conformations: Carbon-alpha RMSDs (CA-RMSDs) were gathered from the whole simulation time and classified in six clusters based on minimum, median and maximum values of CA-RMSDs. Each cluster was further sub-classified, and the receptors used for the study were sampled from each sub-cluster. For both T1R2 and T1R3 monomers, a total of 10 structures were selected and superposed to prepare ensemble of receptor conformations; one out of ten receptors was the initial conformation which underwent MDS experiment. The others were nine different conformations taken from the pre-mentioned clusters. Pre-mentioned structures were superposed and ensembles of T1R2 and T1R3 EC regions were created and centered in a box with dimensions as following: T1R2 89 × 179.58 × 89 (Å ^{3}) and T1R3 94 × 171.61 × 76 (Å ^{3}).

2.2 *Ligand preparation*

Ligands were built and optimized in YASARA software (Krieger and Vriend 2014), using 'Build' command, by 1,4-alpha-glycosylation of α-D-glucose units. Dimer to 14-units oligomers (2Glc to 14Glc) were built from α-1, 4-glycosylation of α-D-glucose (1Glc).

For parameter assignments, the AutoSMILES method (automatic force filed parameter assignment for organic molecules) is employed by all force fields available in YASARA. SMILES strings are used to uniquely identify residues based on their chemical connectivity only. In case of oligosaccharides simulation, if it contains a matching residue, YASARA will assign parameters from the GLYCAM06j (Kirchner *et al.* 2007), otherwise, it derives GAFF/AM1BCC parameters (Jakalian *et al.* 2002; Wang *et al.* 2004).

Docking procedures: YASARA molecular modeling program (Krieger and Vriend 2014) was used for visualization, computing H-bonds and HP-interactions and setting up and running docking procedures using AutoDock Vina (Trott and Olson 2010). Vina has been previously used for carbohydrate docking experiments (Lima *et al.* 2013; Aronsson *et al*. 2018). Energy minimization was also performed for the best docking poses in each binding region to include water molecules. For each receptor in an ensemble, 50 runs were performed and 200 poses were clustered according to 5.0 (Å) heavy atom RMSDs of the ligands. Docking results were reported according to YASARA report parameters as binding energies (the energy required to disassemble a whole into separate parts, usually positive); where more positive binding energy means more favorable interaction (Krieger and Vriend 2014).

Strength of HP-interaction was calculated by YASARA analyze interaction command; the calculation is based on an algorithm basically designed from a knowledge-based potential extracted from high-resolution X-ray structure from PDB for six types of hydrophobic atom interactions (Krieger and Vriend 2014).

To analyze the composition of binding sites, residues with similar physicochemical properties were considered identical. The classification was as follows: Residues N, C, Q, M and S as neutral polar, residues A, V, I, L as hydrophobic-aliphatic, residues F, W and Y as hydrophobic aromatic, residues R, H and K as basic, D and E as acidic and, finally, G and P as unique amino acids.

Energy minimization: Energy minimization in presence of water was performed to consider water solvent effects on carbohydrate conformation as well as ligand-receptor interactions. This procedure was performed by a macro (em_runclean) of MD simulations using a built-in force field incorporated in YASARA suite (Krieger and Vriend 2015). Best poses in each binding regions were refined by Amber14 force field (Hornak *et al*. 2006). This force field combines ff14SB (Piana *et al.* 2011) for proteins, and GLYCAM06 (Kirschner *et al*. 2008) for carbohydrates, some molecules from the AMBER parameter database, ion parameters with optimized ion-oxygen distance and GAFF (Wang *et al.* 2004) for the rest. A force cut off of 8 Å and PME (Essmann *et al.* 1995) to treat long range electrostatic interaction were applied. A cubic cell was created around each ligand-receptor complex, and a steepest descent minimization followed by simulated annealing (time step 2 fs, atom velocities scaled down by 0.9 every 10th step) was done until reach the convergence (Krieger and Vriend 2015).

2.3 *BLAST search*

Sequence alignment was performed for each retrieved binding region, to find identical or similar residues between T1R2 and T1R3. BLAST (BLASTP), using position-specific iterated BLAST (PSI-BLAST) (Altschul *et al.* 1997) was applied and iterations continued to reach E-value of 0.00 to obtain the best match.

3. Results

3.1 *Overall residue composition in detected binding pockets*

Neutral polar and aliphatic residues compose almost halves of residues involved in the interactions between α-D-glucose and its oligomers with STR monomers. Overall analyzing residues of binding pockets showed that, regardless of the number of glucose units, T1R2 and T1R3 binding pockets consisted majorly of neutral polar amino acids in the first place and hydrophobic-aliphatic amino acids in the second place (figure 1A and B). However, this consistency disappeared in the next binding pockets ranks; the third place was occupied by hydrophobic-aromatic and basic residues, respectively, for T1R2 and T1R3. Number of glucose units had no significant effect on residue composition (supplementary figure 1A and B).

3.2 *Individual binding pockets characteristics*

Poses obtained from docking experiments were clustered based on side chain heavy atoms (Methods section). Clusters were observed in distinct binding pockets (BP) distributed in four main regions (R) in both monomers: CR domain region (figure 2), region involving parts of CR and VFTM domains (figure 3), a region located in VFTM LB-1 (figure 4) and regions involving (around) orthosteric binding site (figures 5 and 6). Some binding regions consisted either of one or of more binding pockets with some residues in common (called as core residues). Further, each binding pocket was distinguished by related additional residues. These binding sites will be next discussed from different aspects for both monomers in details, such as residue composition, binding energy characteristics and hydrophilicity/hydrophobicity of the binding regions, and best poses in each region. Tables 1 and 2 show that most clusters gathered in one of the binding pockets belonging to regions involving orthosteric binding pocket as well as the region involving CR-VFTM in both monomers.

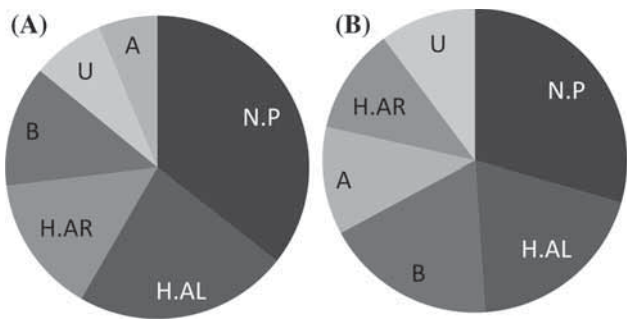

Figure 1. Overall composition of residues in binding pockets found in T1R2 and T1R3. (**A**) T1R2, (**B**) T1R3. N.P.: neutral polar, H.AL: hydrophobic aliphatic, H.AR: hydrophobic aromatic, A: acidic, B: basic and U (unique): P and G.

3.2.1 *Hydrophilicity/hydrophobicity of the binding regions:* Residue composition of each individual binding pocket of T1R2 and T1R3 have been included in supplementary tables 1 and 2, respectively. Figure 7 depicts the hydrophilicity/hydrophobicity of individual binding pocket based on the portion of each amino acid category as a percentage of whole numbers of amino acids involved in each binding pocket. Neutral polar amino acids were the most frequent type observed for both T1R2 and T1R3 binding pockets. However, the least frequently positioned residues could not be clearly attributed to an individual category; for T1R2, acidic and unique residues were the least frequent or even absent and hydrophobic-aromatic residues took this place in T1R3 (details in figure 7).

3.2.2 *Ligand binding energy characteristics:* Tables 3 and 4 show that highest ligand binding energy values belonged to CR-VFTM regions and binding pockets involving orthosteric binding sites in both monomers. The best energies in T1R2 belonged to CR-VFTM binding pocket (table 3), while in T1R3, the highest energy values were observed for one of the regions involving orthosteric binding site (table 4). Comparison of highest values of ligand binding energies in both monomers showed that overall, T1R3 binding sites had stronger interactions with ligands. Considering lowest and highest energy values in each monomer, fewer diversities were found in T1R2 interactions.

3.2.3 *Binding pocket selection: Role of ligand type and receptor conformation:* Most of the binding pockets were picked up by specific ligands and were found in specific conformations of the receptor. Residue composition of binding regions are available from supplementary tables 1 and 2.

3.2.3.1 CR domain binding regions CR domain could be regarded as a region for interacting with smaller ligands (supplementary figure 2A and B).

T1R2: Ten clusters were located in T1R2 CR domain. CR domain could be regarded as a region for interacting with smaller ligands while interacting with α-D-glucose to α-D-triose and with minimum CA-RMSD conformations (supplementary figure 2A).

T1R3: Forty-four clusters were fitted in this region. On the contrary to T1R2, this binding region was observed for almost all ligands (supplementary figure 2B). Ligands with less than three units of α-D-glucose could be retained as the main ligands interacting with these residues; these ligands interacted with at least one of conformations in minimum to maximum CA-RMSDs clusters (supplementary figure 2B).

Identical residues between T1R2 and T1R3 CR-binding regions: Alignment of sequences revealed that some residues in CR region binding pockets of T1R2 and T1R3 could be aligned which were either conserved or showed

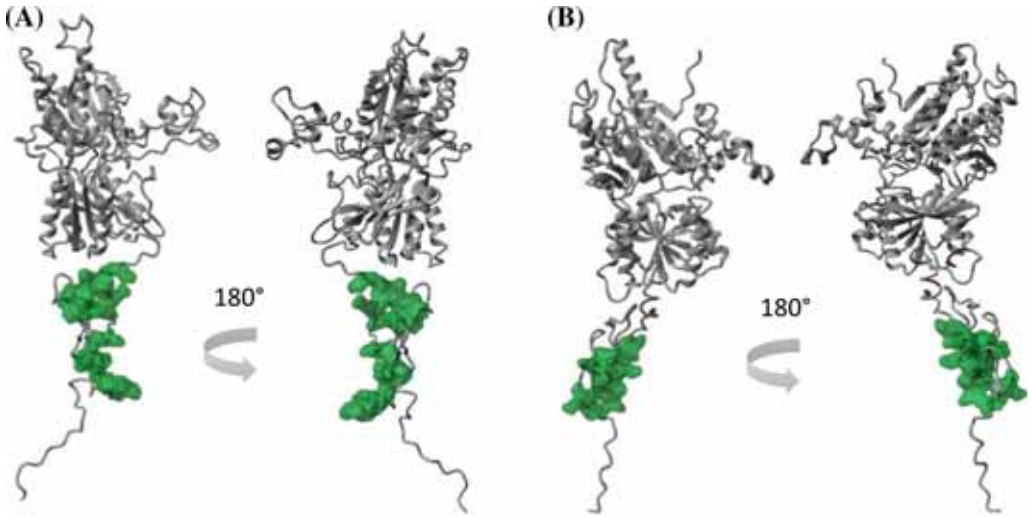

Figure 2. CR domain binding regions. (**A**) T1R2, (**B**) T1R3. Green color shows the molecular surface of residues in each region. Pictures have been rotated 180 degrees to cover back and front views of the pockets.

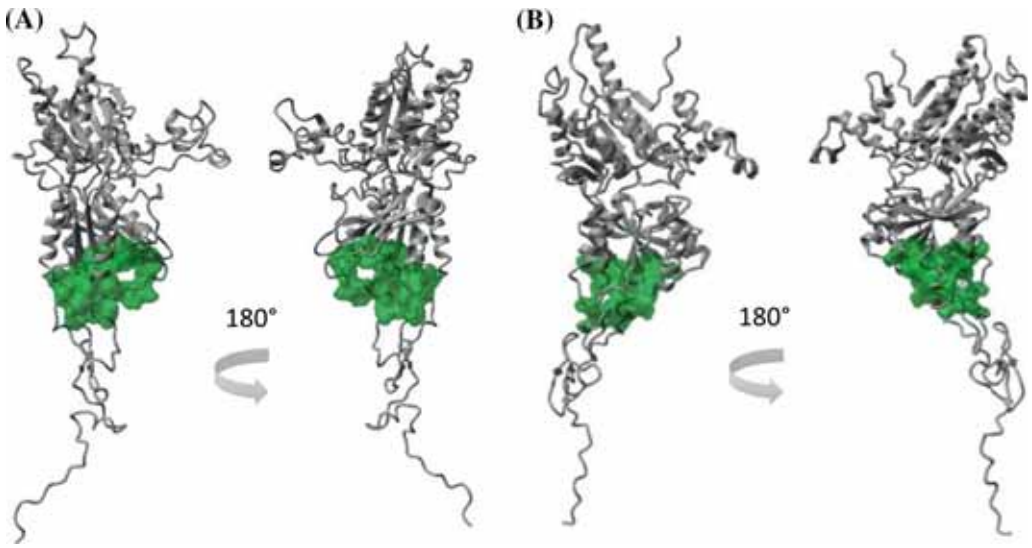

Figure 3. CR-VFTM involving binding pockets. (**A**) T1R2, (**B**) T1R3. Green color shows the molecular surface of residues in each region. Pictures have been rotated 180 degrees to cover back and front views of the pockets.

similar side chain characteristics to conserved residues (Chéron *et al.* 2017) in C GPCR family members. These were S493/S497, C495/C499, S496/S500, C499/C503, C517/C521, D531/D534, G502/G506, F515/Y519, F525/Y529, W543/W546, from T1R2/T1R3 respectively.

3.2.3.2 CR and VFTM involved binding region This binding region was the most popular one (tables 1 and 2) for both monomers, and the residues, regardless of receptor conformation, interacted with all ligands (supplementary figure 3A and B).

Identical residues between T1R2 and T1R3 CR-VFTM involving regions: Aligned residues of T1R2/T1R3 monomers which were either conserved (Chéron *et al.* 2017) in C GPCR or showed similar side chain characteristics with conserved residues are: W203/W206, S493/S497, W205/W208, V271/V270, W483/W487, S496/S500, C513/C517, F515/Y519 from T1R2/T1R3 respectively.

3.2.3.3 VFTM binding pockets These binding pockets were subcategorized to regions belonging to lobe 1 of VFTM (VFTM LB-1) and regions involving VFTM orthosteric binding site.

LB-1: In both monomers, no clear recognition pattern could be defined for this region (figure 4A and B). Two binding pockets in LB-1 region of T1R2 showed different profiles of interaction (figure 4A); while BP-1 was found by glucose oligomers with less than seven α-D-glucose units in almost every conformation, BP-2 was found by almost all ligands only in conformations with maximum CA-RMSDs

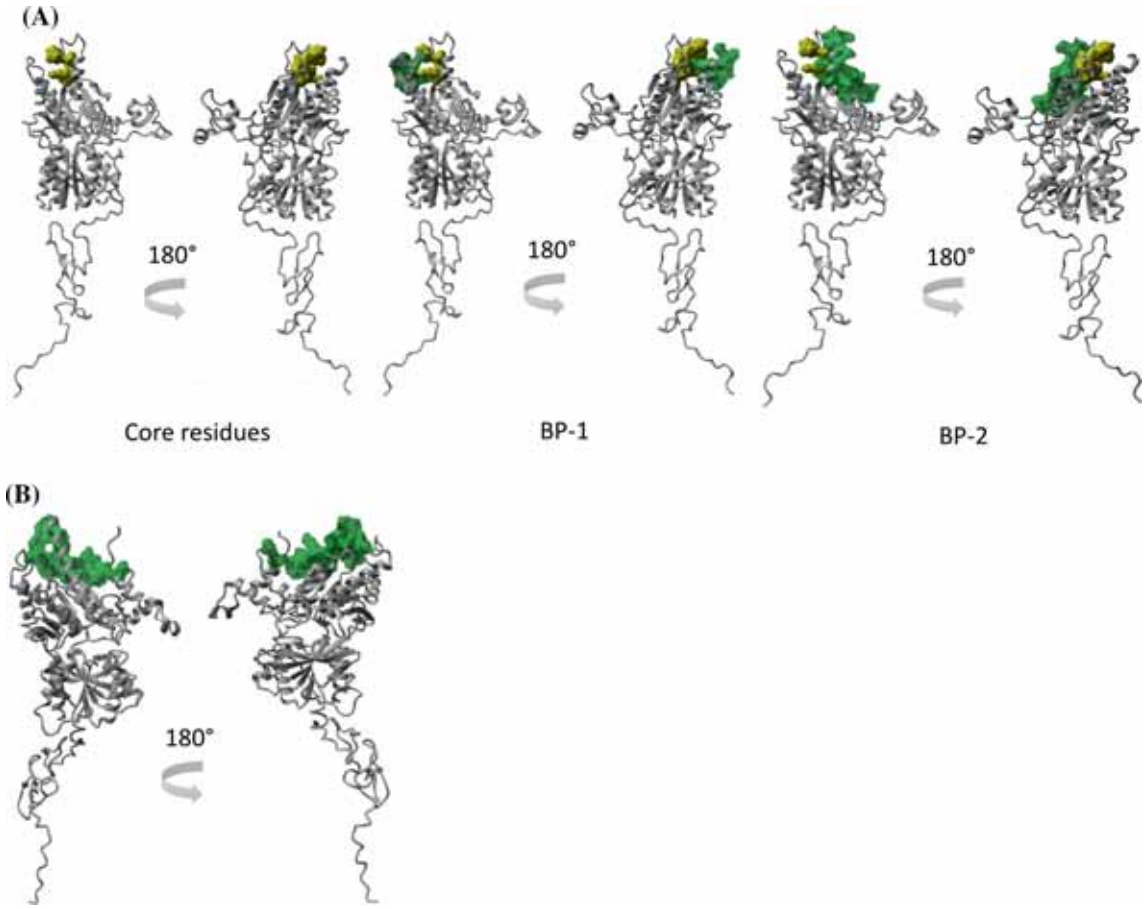

Figure 4. Binding pockets involving VFTM Lobe-1. (**A**) T1R2; yellow and green colors respectively show the molecular surface of core residues and additional residues of each binding pocket, (**B**) T1R3; green color shows the molecular surface of residues in each region. Pictures have been rotated 180 degrees to cover back and front views of the pockets. BP: binding pocket, numbers are defined in the text.

(supplementary figure 4A1 and A2). The LB-1 region of T1R3 showed also no clear interaction pattern (supplementary figure 4B).

Identical residues between T1R2 and T1R3 LB-1 binding regions: T1R2/T1R3 sequence alignment revealed aligned pair of residues G31/G33, D32/D34, Y33/Y35, V92/L93, L112/L114 which were found conserved (Chéron *et al.* 2017) with regard to side chain characteristics among C GPCR family members.

Pockets involving orthosteric binding site: These binding pockets were subcategorized in four and three distinct regions for T1R2 and T1R3, respectively. After CR-VFTM involving regions, the most popular and stable binding sites could be found here (table 1 and 2; supplementary figures 5 and 6).

T1R2

Region 1: This binding site was occupied by 183 clusters distributed in two distinct binding pockets with a few residues in common (figure 5A; supplementary table 1) and seemed stable during MD simulation for almost all ligands (supplementary figure 5A2). Both binding pockets were recognized by almost all α-D-glucose oligomers though diversities were observed in receptor conformations. BP-1 was observed in initial conformation (supplementary figure 5A1). supplementary figure 5A2 shows that BP-2 (the popular binding pocket with 119 clusters) was recognizable in at least one conformation of each CA-RMSD by all α-D-glucose oligomers (exceptions were α-D-glucose and α-D-maltose).

Region 2: This region was located in LB-1 with core residues located around orthosteric binding region (figure 5B). Supplementary table 1 and supplementary figure 5B show the related residue composition and ligand and receptor characteristics, respectively. Overall, no clear pattern could be extracted, neither by ligand type nor by receptor conformation (supplementary figure 5B1 and B2).

Region 3: This binding pocket was another popular binding region (table 1) with no limitations in ligand type or receptor conformation (supplementary figure 5C).

Region 4: This binding pocket (figure 5D) could be considered as a binding region for receptors in minimum CA-RMSD conformation (supplementary figure 5D).

T1R3

Region 1: This region was a popular and stable binding region (table 2 and supplementary figure 6A) which involved the orthosteric binding site (figure 6A).

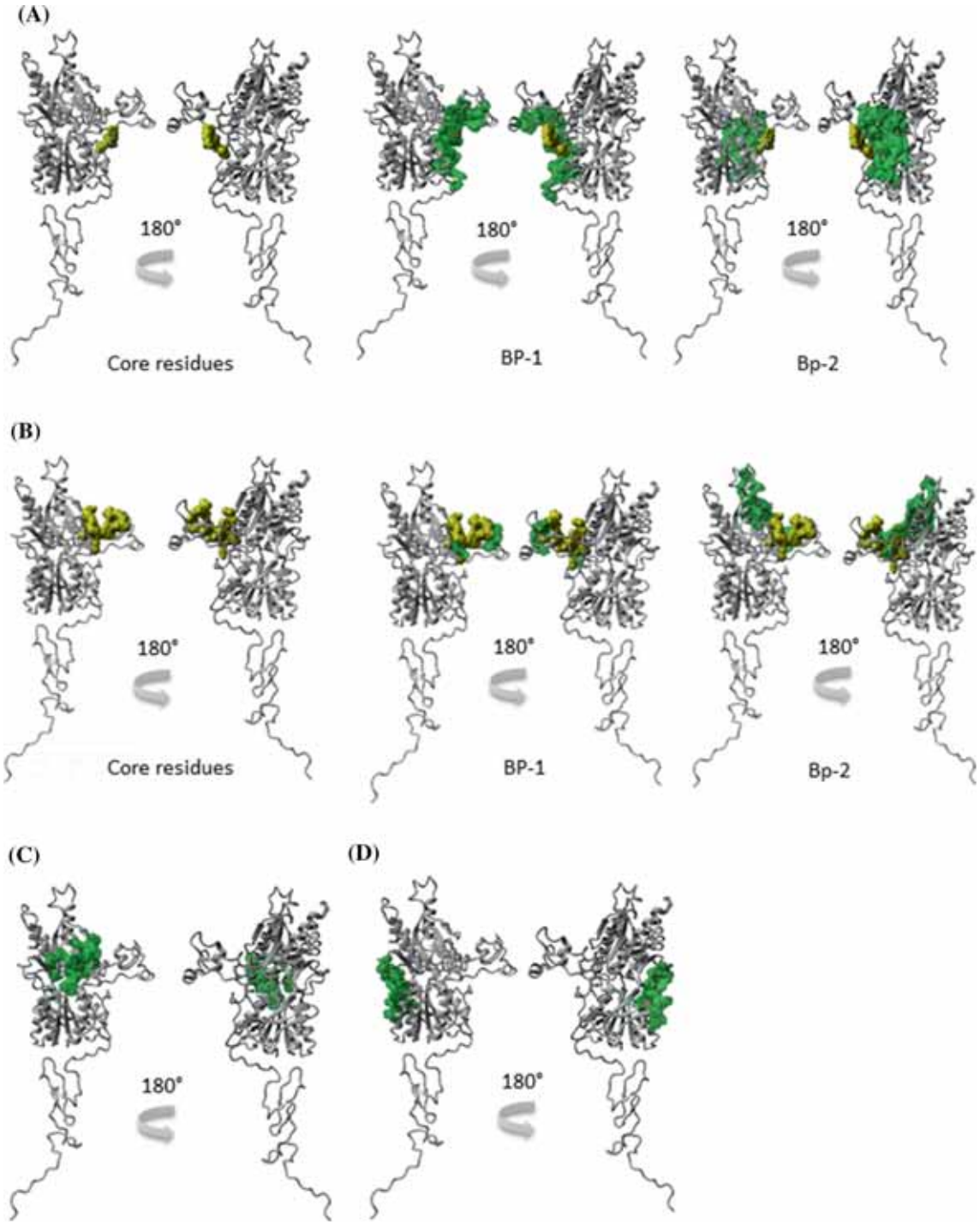

Figure 5. T1R2 VFTM binding pockets involving orthosteric binding site. (**A**) Region 1, (**B**) Region 2, (**C**) Region 3 and (**D**) Region 4. Yellow and green colors respectively show the molecular surface of core residues and additional residues of each binding pocket. Pictures have been rotated 180 degrees to cover back and front views of the pockets. BP: binding pocket, numbers are defined in the text as a descriptor for the related BP.

Region 2: Though not as crowded as region1, this region was also popular to some extent (table 2) and found to be stable in different conformations (supplementary figure 6B).

Region 3: Two distinct binding sites with some core residues formed this region (figure 6C) which was recognized by 138 clusters. BP-1 and BP-2 seemed stable for some ligands (supplementary figure 6C1 and C2).

Identical residues between T1R2 and T1R3 regions involving orthosteric binding site: PSI-BLAST alignment indicated the following residues from binding pockets

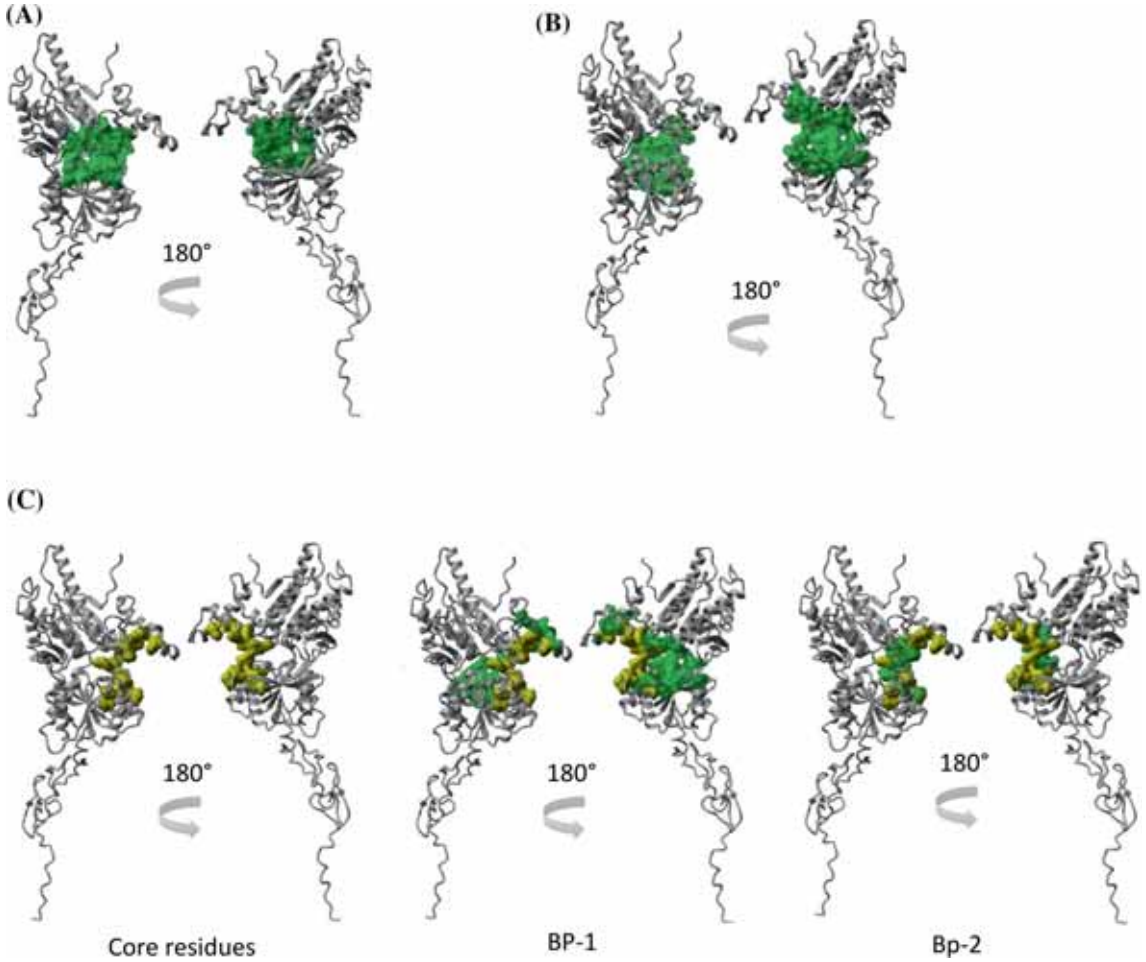

Figure 6. T1R3 VFTM binding pockets located in regions involving orthosteric binding site. **(A)** Region 1, (**B**) Region 2 and (**C**) Region 3. Yellow and green colors show the molecular surface of core residues and additional residues of each binding pocket, respectively, if more than one binding pocket existed in each region. Otherwise, green color shows the molecular surface of residues in each region. Pictures have been rotated 180 degrees to cover back and front views of the pockets; BP: binding pocket, numbers are defined in the text as a descriptor for the related BP.

involving orthosteric binding site of T1R2/T1R3: H42/E45, K65/N68, Y103/S104, E145/E148, Y282/H281, D307/S306, V309/L308, I332/L331, L379/L385, S380/N386, G381/H387, E382/H388, R383/Q389, Y386/F391. Furthermore, C59/C62 and C363/C370 were totally conserved among C GPCR family members (Chéron *et al.* 2017).

Binding pockets solvent accessibility and ligand replacement in T1R2 and T1R3: Supplementary figure 7 depicts the direct relationship between solvent accessibility of the binding pockets and replacement of the bulky ligands. The overall trend is that an increase in the volume of solvent accessible surface led the receptor to receive bulkier ligands, those having 8 units or more of α-D-glucose.

3.2.4 *Best poses in each region:*

H-bond characteristics: In order to better understand receptor-ligand interaction, for each region, the poses with highest binding energies were selected and further discussed.

Results showed that, for both monomers, the best poses occurred in interaction with oligomers larger than 8-unit-oligomers except for T1R2 VFTM-LB1 which occurred with α-D-tetrasaccharide (supplementary tables 3 to 6). Results represented in supplementary tables 3 to 6 suggest that most of the residues could be considered as either H-bond acceptors or donors in interaction of sweet taste receptor with α-D-glucose-1,4 oligomers. However, there were some residues that played both H-bond acceptor or donor roles; those included I104, S105, Q221, D231, Q244, G381 for T12R and R30, K32, R123, Q202, H278, T305, Q330, N386, Q389, E545 for T1R3. Comparing the energy of H-bonds in both monomers showed that regardless of the size of the ligand, the best pose in each binding region can bond as strong as 25 $kJ.mol^{-1}$ (supplementary tables 3 to 6).

Identical residues between T1R2 and T1R3 H-bond donors or acceptors: Alignment of T1R2 and T1R3 sequences suggested that some aligned residues played

Table 1. Clusters, binding pockets and binding regions characteristics in T1R2

T1R2 binding region	Binding pockets observed in each region	Clusters observed in each binding pocket	Volume of solvent accessible surface for each binding pocket (Å^3)
CR	BP	10	7542
CR-VFTM	BP	802	8942
VFTM-LB1	BP-1	44	5651
	BP-2	46	7626
Orthosteric-R1	BP-1	64	9983
	BP-2	119	7562
Orthosteric-R2	BP-1	14	11066
	BP-2	17	5544
Orthosteric-R3	BP	86	4791
Orthosteric-R4	BP	24	6608

BP: Binding pocket; CR: CR domain region; CR-VFTM: region involving parts of CR and VFTM domains; VFTM-LB1: region located in VFTM LB-1; Orthosteric-R1/R2/R3/R4: regions involving orthosteric binding site; Volume of Solvent Accessible Surface (Å3): the volume enclosed by the solvent accessible surface of the selected atoms in Å^3; Å^3: cubic angstrom.

Table 2. Clusters, binding pockets and binding regions characteristics in T1R3

T1R3 binding region	Binding pockets observed in each region	Clusters observed in each binding pocket	Volume of solvent accessible surface for each binding pocket (Å^3)
CR	BP	44	2685
CR-VFTM	BP	658	3528
VFTM-LB1	BP	66	3251
Orthosteric-R1	BP	362	5264
Orthosteric-R2	BP	62	5037
Orthosteric-R3	BP-1	160	3624
	BP-2	76	3423

BP: Binding pocket; CR: CR domain region; CR-VFTM: region involving parts of CR and VFTM domains; VFTM-LB1: region located in VFTM LB-1; Orthosteric-R1/R2/R3: regions involving orthosteric binding site; Volume of Solvent Accessible Surface (Å3): the volume enclosed by the solvent accessible surface of the selected atoms in Å^3; Å^3: cubic angstrom.

H-bond donors: H50/S53, R134/R137, R230/R233, R383/Q389 from T1R2/T1R3. Also, aligned pair G381/H387 from T1R2/T1R3 played acceptor role (supplementary tables 3 to 6).

HP-interaction characteristics: Information in this regard is available from supplementary tables 7 and 8, for T1R2 and T1R3 respectively. Irrespective of the size of ligands, it was observed that these interactions were week (<5). The most frequently observed residues in HP-interactions (supplementary tables 7 and 8) were neutral polar residues (33% for T1R2 and 24% for T1R3 when considering all residues involved in HP-interactions). Furthermore, supplementary figures 8 to 17 were provided as a visual aid for an insight into binding modes of ligands in detected best poses.

4. Discussion

This study was designed to find out putative carbohydrate binding modules on extracellular region of human STR to better understand the architecture of this molecule which has not been experimentally observed so far. The probes α-D-glucose and its α-1,4-oligomers with oligomerization degree up to 14, were docked into various conformations of STR monomers sampled from an MD simulation experiment. Considering the 'carbohydrate' characteristic of ligands, further analysis on binding sites were carried out after they were energy minimized in presence of water molecules with AMBER14 forcefield (Bryce *et al.* 2001; Hornak *et al.* 2006). Resulted poses could be classified in four major binding regions in CR domain, binding regions involving CR-VFTM, VFTM LB1 binding region and binding regions involving orthosteric binding site in T1R2 and T1R3 extracellular segments.

4.1 *Ligand binding crucial residues were found in regions involving orthosteric binding site*

Some residues in regions involving orthosteric binding site of the both monomers (supplementary tables 1 and 2), which are shown as a pair of aligned T1R2/T1R3 residues here, have been mentioned as key residues in ligand binding of STR by experimental studies (Chéron *et al*. 2017). Actually, K65 (Zhang *et al*. 2010)/N68, Y103 (Zhang *et al.* 2010; Masuda *et al.* 2012a, b; Maillet *et al.* 2015) /S104, E145/E148, Y282/H281, D307 (Zhang *et al.* 2010; Maillet *et al*. 2015)/S306, V309/L308, E382/H388, R383 (Zhang *et al.* 2010)/Q389. H42 (Waksmonski and Koppel 2016)/E45 and Y386/F391 have been reported as key ligand binding residues (Chéron *et al*. 2017). Residues such as E382/H388 and R383, that are important for small molecule binding (Masuda *et al*. 2012a, b), were also found to interact with glucose rings in orthosteric region (supplementary table 7).

4.2 *Reported interactions are in accordance with existing knowledge*

T1R2 K65, V309 and Y103 (supplementary table 7) have been suggested to form HP-interactions as 'hydrophobic pincers' (Zhang *et al.* 2010). Another study has also mentioned the role of Y103 (supplementary table 7) aromatic ring in HP-interaction (Masuda *et al.* 2012a, b). This residue

T1R2 binding regions	Binding pockets of each region	N.P.	H.AL.	H.AR.	A	B	U
CR		32	14	14	11	11	18
CR-VFTM		44	21	15	0	15	6
VFTM-LB1	Bp-1	33	24	14	10	19	0
	Bp-2	33	24	14	10	19	0
Orthosteric-R1	Bp-1	30	24	12	12	21	0
	Bp-2	27	27	10	10	10	17
Orthosteric-R2	Bp-1	54	4	4	18	14	7
	Bp-2	30	18	15	18	12	6
Orthosteric-R3		53	24	12	6	0	6
Orthosteric-R4		17	17	4	38	21	4

T1R3 binding regions	Binding pockets of each region	N.P.	H.AL.	H.AR.	A	B	U
CR		8	1	2	5	3	4
CR-VFTM		11	3	5	1	5	2
VFTM-LB1		9	5	1	2	6	4
Orthosteric-R1		10	8	2	11	8	6
Orthosteric-R2		6	6	2	3	8	4
Orthosteric-R3	BP-1	9	10	1	5	10	6
	BP-2	8	8	1	5	4	3

Figure 7. Hydrophilicity/hydrophobicity of individual binding pockets in T1R2 and T1R3 monomers. Numbers represent portion of each amino acid category as a percentage of the whole residues recognized in each binding pocket. N.P.: neutral polar; H.AL: hydrophobic aliphatic; H.AR: hydrophobic aromatic; A: acidic; B: basic and U (unique): P and G. Green and red colors are representative for the highest and the lowest portion of amino acids, respectively; BP: binding pocket; CR: CR domain region; CR-VFTM: region involving parts of CR and VFTM domains; VFTM-LB1: region located in VFTM lobe-1; Orthosteric-R1/R2/R3/R4: regions involving orthosteric binding site.

Table 3. Ligand-binding energies for each individual binding pocket of T1R2

T1R2 binding region	Binding pockets	Binding energy (kcal.mol^{-1})		
		Lowest	Highest	Average
CR		4.7	6.1	5.4
CR-VFTM		4.4	8.7	**6.9**
VFTM-LB1	BP-1	4.6	7.8	6.4
	BP-2	4.7	8.4	6.5
Orthosteric-R1	BP-1	4.4	7.9	6.2
	BP-2	4.5	8.2	6.7
Orthosteric-R2	BP-1	5.1	7.3	6.1
	BP-2	5.6	6.8	6.4
Orthosteric-R3		4.7	8.0	6.5
Orthosteric-R4		4.5	7.2	6.1

BP: Binding pocket; CR: CR domain region; CR-VFTM: region involving parts of CR and VFTM domains; VFTM-LB1: region located in VFTM LB-1; Orthosteric-R1/R2/R3/R4: regions involving orthosteric binding site. Highest values are in boldface.

Table 4. Ligand-binding energies for each individual binding pocket of T1R3

T1R3 binding region	Binding pockets	Binding energy (kcal.mol^{-1})		
		Lowest	Highest	Average
CR		5.3	8.1	6.7
CR-VFTM		4.5	9.2	7.1
VFTM-LB1		5.5	8.3	6.8
Orthosteric-R1		5.8	**9.7**	**7.9**
Orthosteric-R2		4.7	8.4	6.8
Orthosteric-R3	BP-1	4.7	8.5	7.1
	BP-2	4.6	8.8	7.3

BP: Binding pocket; CR: CR domain region; CR-VFTM: region involving parts of CR and VFTM domains; VFTM-LB1: region located in VFTM LB-1; Orthosteric-R1/R2/R3: regions involving orthosteric binding site. Highest values are in boldface.

is crucial in sucrose (Zhang *et al*. 2010), sucralose (Zhang *et al.* 2010), aspartame (Masuda *et al*. 2012a, b) and D-tryptophan (Masuda *et al.* 2012a, b) reception by STR. T1R2 V309, L310, R383 (supplementary table 7) were mentioned to form HP-interaction (Liu *et al.* 2011). T1R2 K65 was found to form H-bond with 10Glc while previously mentioned to form H-bond with sweet taste enhancers (Zhang *et al.* 2010).

4.3 *Results could be extrapolated to other C GPCR family members*

It is possible to extrapolate the results of this study to C GPCR family members. Each binding region in T1R2 and T1R3 was found to have similar or identical residues in both monomers. Some of the residues are totally conserved in C GPCR family (Chéron *et al.* 2017), such as residues S493/S497, C499/C503, C517/C521 and D531/D534 (CR region), residues W203/W206 and S493/S497 (CR-VFTM involving regions). Some others are either conserved in most of C GPCR family members or have similar side chain characteristics according to Chéron *et al.* results (Chéron *et al.* 2017); these include residues S496/S500, G502/G506, F515/Y519, F525/Y529 and W543/W546 (CR region); residues of CR-VFTM involving region F201/F204 (conserved as N.P side chain residues), W205/W208 (conserved in C GPCR family unless in GABA receptors), R270/Q269 (conserved also in mGluRs), V271/V270 (conserved as H.AL. side chain residues), W483/W487 (conserved as N.P side chain residues), S496/S500, F515/Y519 (conserved as N.P side chain residues) (Chéron *et al.* 2017). VFTM LB1 region also possess such residues: G31/G33 (conserved in C GPCR family unless in GABA receptors)(Chéron *et al*. 2017), D32/D34 (conserved in C GPCR family unless in mGluR6 and GABA receptors)(Chéron *et al*. 2017), V92/L93 and L112/L114 (conserved as H.AL. side chain residues save for GABA receptors) (Chéron *et al*. 2017). So, existence of conserved or semi-conserved residues prompts further

investigation about the existence of such binding regions in other C GPCR family members.

4.4 *Probable existence of structures resembling carbohydrate binding modules*

Most clusters were fitted in the regions which shows similarities with carbohydrate binding modules (CBM) in carbohydrate-active enzymes (Armenta *et al*. 2017); these are CR-VFTM and regions around orthosteric binding site for both T1R2 and T1R3 (tables 1 and 2). CBMs have a broadly conserved conformation to accept a wide variety of carbohydrates in different shape and conformations. From ligand-receptor topology point of view, CBMs are classified in three classes of A, B and C (Armenta *et al*. 2017). In a previous experimental study, possibility of CBMs presence in gut fungal C GPCRs has been suggested (Seppälä *et al.* 2016). Two bundles mimicking beta-sandwich folds, the major fold in CBMs, can be observed around orthosteric binding site of both monomers in LB1 and 2 (figure 8). CR-VFTM involved binding regions were occupied by 65% and 46% of clusters, respectively in T1R2 and T1R3 (tables 1 and 2). These regions engaged parts of the beta-sandwich folds in LB2. A considerable amount of clusters were positioned in binding regions involving orthosteric binding site such as region 1-BP2 (10%) and region 3 (7%) in T1R2 as well as region 1 (25%) and region 3-BP1 (11%) in T1R3. These regions mimic CBM type B (Armenta *et al*. 2017). They consist of loops that form a cavity and are able to interact with larger oligomers. Further, these are in close proximity of the beta sandwich folds in LB 1and LB2 of the STR. Aromatic residues play major role in all types of CBMs and polar residues stabilize H-bonds (Armenta *et al.* 2017). STR CBM mimicking binding regions consisted of polar amino-acids and aromatic residues (figure 7). Week HP-interaction as well as strong H-bonds, regardless of ligand size, were found to be a common interaction pattern among binding regions in both monomers (supplementary tables 2–8). However, such H-bond dominance was majorly observed in CBM mimicking regions, as in CBM type B (Pell *et al*. 2003). supplementary figure 7 implies the ligand promiscuity observed in CBMs (Charnock *et al*. 2002; Gilbert 2010), accompanied by flatness in binding regions, as a cause for interacting with various ligand sizes and conformations (Samaei-Daryan *et al.* 2017).

In brief, CBMs characteristics may not be observed in STR, most popular binding regions recognized by α-D-glucose oligomers strongly mimic CBMs patterns and could be suggested as carbohydrate binding modules specific to C GPCR family.

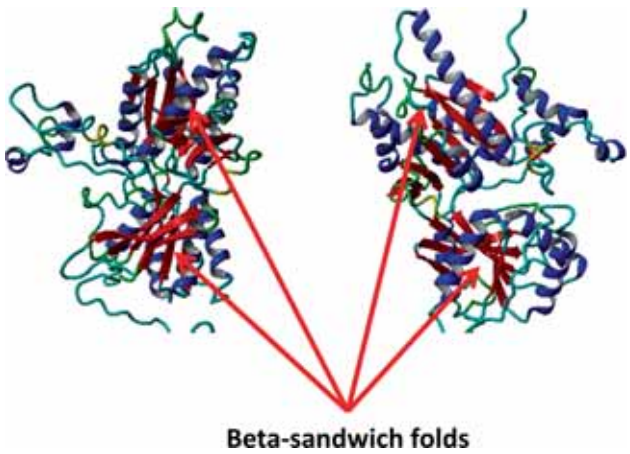

Figure 8. Beta-sandwich folds in T1R2 and T1R3 VFTM domains. VFTM segments of T1R2 and T1R3 are shown. Alpha-helices are represented in blue and beta-sheets in red; beta-sandwiches as connected anti-parallel beta-sheets are presented in the picture.

5. Conclusions

The current study shades light onto structural properties of C GPCR family by investigating carbohydrate binding characteristics on one of its members, namely STR. Four putative carbohydrate binding regions were found for the sweet taste receptor heterodimer which could be structurally aligned in both monomers: CR region, CR-VFTM involving region, VFTM LB1 and regions involving orthosteric binding site. In accordance with CBMs existence in gut fungal C GPCRs (Seppälä *et al.* 2016), CBM-like characteristics can be found in STR and maybe other C GPCR family members, especially for CR-VFTM involving binding regions and some regions involving orthosteric binding sites in both monomers.

T1R2/T1R3 residues are suggested as either H-Bond donors (H50/S53, R134/R137, R230/R233, R383/Q389) or H-Bond acceptors (G381/H387). Furthermore, some residues from T1R2 (I104, S105, Q221, D231, Q244, G381) and T1R3 (R30, K32, R123, Q202, H278, T305, Q330, N386, Q389, E545) are suggested to play both H-Bond donors or acceptors in the interaction of sweet taste receptor with α-D-glucose-1,4 oligomers. Existence of conserved or semi-conserved residues prompts further investigation about the existence of such binding regions in other C GPCR family members.

Acknowledgements

This study was performed as a Ph.D. thesis granted by Endocrinology and Metabolism Research Institute, Tehran University of Medical Sciences, Tehran, Iran.

References

Abbott DW and AL van Bueren 2014 Using structure to inform carbohydrate binding module function. *Curr. Opinion Struct. Biol.* **28** 32–40

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.***25** 3389–3402

Armenta S, Moreno-Mendieta S, Sánchez-Cuapio Z, Sánchez S and Rodríguez-Sanoja R 2017 Advances in molecular engineering of carbohydrate-binding modules. *Proteins Struct. Funct. Bioinform.* **85** 1602–1617

Aronsson A, Guler F, Petoukhov MV, Crennell SJ, Svergun DI, Linares-Pasten JA and Nordberg Karlsson E 2018 Structural insights of RmXyn10A - A prebiotic-producing GH10 xylanase with a non-conserved aglycone binding region. *Biochim. Biophys. Acta* **1866** 292–306

Assadi-Porter FM, Maillet EL, Radek JT, Quijada J, Markley JL and Max M 2010 Key amino acid residues involved in multi-point binding interactions between brazzein, a sweet protein, and the T1R2-T1R3 human sweet receptor. *J. Mol. Biol.* **398** 584–599

Assadi-Porter FM, Tonelli M, Maillet EL, Markley JL and Max M 2010 Interactions between the human sweet-sensing T1R2-T1R3 receptor and sweeteners detected by saturation transfer difference NMR spectroscopy. *Biochim. Biophys. Acta Biomembranes* **1798** 82–86

Behrens M and Meyerhof W 2011 Gustatory and extragustatory functions of mammalian taste receptors. *Physiol. Behav.* **105** 4–13

Boraston AB, Bolam D N, Gilbert H J and Davies GJ 2004 Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem. J.* **382** 769–781

Bryce R, Hillier I and Naismith J 2001 Carbohydrate-protein recognition: molecular dynamics simulations and free energy analysis of oligosaccharide binding to concanavalin A. *Biophys. J.* **81** 1373–1388

Charnock SJ, Bolam DN, Nurizzo D, Szabó L, McKie VA, Gilbert HJ and Davies GJ 2002 Promiscuity in ligand-binding: the three-dimensional structure of a Piromyces carbohydrate-binding module, CBM29-2, in complex with cello-and mannohexaose. *Proc. Nat. Acad. Sci.* **99** 14077–14082

Chéron J-B, Golebiowski J, Antonczak S and Fiorucci S 2017 The anatomy of mammalian sweet taste receptors. *Proteins Struct. Funct. Bioinform.* **85** 332–341

Congreve M, Doré AS, Jazayeri A and Nonoo R 2015 Engineering G protein-coupled receptors for drug design; in *Multifaceted Roles of Crystallography in Modern Drug Discovery* (Springer) pp1–18

Cui M, Jiang P, Maillet E, Max M, Margolskee RF and Osman R 2006 The heterodimeric sweet taste receptor has multiple potential ligand binding sites. *Curr. Pharmaceut. Design* **12** 4591–4600

Essmann U, Perera L, Berkowitz ML, Darden T, Lee H and Pedersen LG 1995 A smooth particle mesh Ewald method. *J. Chem. Phys.***103** 8577–8593

Ficko-Blean E and Boraston AB 2012 Insights into the recognition of the human glycome by microbial carbohydrate-binding modules. *Curr. Opinion Struct. Biol.* **22** 570–577

Gerspach AC, Steinert RE, Schönenberger L, Graber-Maier A and Beglinger C 2011 The role of the gut sweet taste receptor in regulating GLP-1, PYY, and CCK release in humans. *Am. J. Physiol. Endochrinol. Metab.* **301** E317–E325

Gilbert HJ 2010 The biochemistry and structural biology of plant cell wall deconstruction. *Plant Physiol.* **153** 444–455

Gilbert HJ, Knox J P and Boraston AB 2013 Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. *Curr. Opinion Struct. Biol.* **23** 669–677

Hillisch A 2004 Utility of homology models in the drug discovery process. *Drug Discov. Today* **9** 659–669

Hornak V, Abel R, Okur A, Strockbine B, Roitberg A and Simmerling C 2006 Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **65** 712–725

Jakalian A, Jack D B and Bayly CI 2002 Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **23** 1623–1641

Jang HJ, Kokrashvili Z, Theodorakis MJ, Carlson OD, Kim BJ, Zhou J, Hyeon HK, Xu X, Chan SL, Juhaszova M, Bernier M, Mosinger B, Margolskee RF and Egan JM 2007 Gut-expressed gustducin and taste receptors regulate secretion of glucagon-like peptide-1. *Proc. Nat. Acad. Sci. USA* **104** 15069–15074

Kirchner K, Yongye A, Tschampel S, Daniels C, Foley B and Woods R 2007 GLYCAM06: a generalizable biomolecular force field. *J. Comput. Chem.* **29** 622–655

Kirschner KN, Yongye AB, Tschampel SM, Gonzalez-Outeirino J, Daniels CR, Foley BL and Woods RJ 2008 GLYCAM06: a generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem.***29** 622–655

Kniazeff J, Prézeau L, Rondard P, Pin J-P and Goudet C 2011 Dimers and beyond: the functional puzzles of class C GPCRs. *Pharmacol. Therapeut.* **130** 9–25

Kochem M 2017 Type 1 Taste receptors in taste and metabolism. *Ann. Nutr. Metab.* **70** 27–36

Krieger E, Darden T, Nabuurs SB, Finkelstein A and Vriend G 2004 Making optimal use of empirical energy functions: force-field parameterization in crystal space. *Proteins Struct. Funct. Bioinform.* **57** 678–683

Krieger E, Joo K, Lee J, Lee J, Raman S, Thompson J, Tyka M, Baker D and Karplus K 2009 Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: four approaches that performed well in CASP8. *Proteins* **77** 114–122

Krieger E, Nielsen JE, Spronk CA and Vriend G 2006 Fast empirical pKa prediction by Ewald summation. *J. Mol. Graph. Model* **25** 481–486

Krieger E and Vriend G 2014 YASARA View—molecular graphics for all devices—from smartphones to workstations. *Bioinformatics* **30** 2981–2982

Krieger E and Vriend G 2015 New ways to boost molecular dynamics simulations. *J. Comput. Chem.* **36** 996–1007

Laffitte A, Neiers F and Briand L 2014 Functional roles of the sweet taste receptor in oral and extraoral tissues. *Curr. Opin. Clin. Nutr. Metab. Care* **17** 379–385

Li X, Staszewski L, Xu H, Durick K, Zoller M and Adler E 2002 Human receptors for sweet and umami taste. *Proc. Nat. Acad. Sci.* **99** 4692–4696

Lima MA, Oliveira-Neto M, Kadowaki MA, Rosseto FR, Prates ET, Squina FM, Leme AF, Skaf MS and Polikarpov I 2013 Aspergillus niger beta-glucosidase has a cellulase-like tadpole molecular shape: insights into glycoside hydrolase family 3

(GH3) beta-glucosidase structure and function. *J. Biol. Chem.* **288** 32991–33005

Liu B, Ha M, Meng X-Y, Kaur T, Khaleduzzaman M, Zhang Z, Jiang P, Li X and Cui M 2011 Molecular mechanism of species-dependent sweet taste toward artificial sweeteners. *J. Neurosci.* **31** 11070–11076

Maillet EL, Cui M, Jiang P, Mezei M, Hecht E, Quijada J, Margolskee RF, Osman R and Max M 2015 Characterization of the binding site of aspartame in the human sweet taste receptor. *Chem. Senses* **40** 577–586

Margolskee RF, Dyer J, Kokrashvili Z, Salmon KSH, Ilegems E, Daly K, Maillet EL, Ninomiya Y, Mosinger B and Shirazi-Beechey SP 2007 T1R3 and gustducin in gut sense sugars to regulate expression of Na+-glucose cotransporter 1. *Proc. Nat. Acad. Sci. USA* **104** 15075–15080

Masuda K, Koizumi A, Nakajima K-I, Tanaka T, Abe K, Misaka T and Ishiguro M 2012 Characterization of the modes of binding between human sweet taste receptor and low-molecular-weight sweet compounds. *PLoS One* **7** e35380

Masuda K, Koizumi A, Nakajima K-I,, Tanaka T, Abe K, Misaka T and Ishiguro M 2012 Characterization of the models of binding between human sweet taste receptor and low-molecular-weight sweet compounds. *PLoS One* **7** e35380

Muto T, Tsuchiya D, Morikawa K and Jingami H 2007 Structures of the extracellular regions of the group II/III metabotropic glutamate receptors. *Proc. Nat. Acad. Sci.* **104** 3759–3764

Nelson G, Hoon MA, Chandrashekar J, Zhang Y, Ryba NJ and Zuker CS 2001 Mammalian sweet taste receptors. *Cell* **106** 381–390

Nie Y, Vigues S, Hobbs JR, Conn GL and Munger SD 2005 Distinct contributions of T1R2 and T1R3 taste receptor subunits to the detection of sweet stimuli. *Curr. Biol.* **15** 1948–1952

Pell G, Williamson MP, Walters C, Du H, Gilbert HJ and Bolam DN 2003 Importance of hydrophobic and polar residues in ligand binding in the family 15 carbohydrate-binding module from Cellvibrio japonicus Xyn10C. *Biochemistry* **42** 9316–9323

Piana S, Lindorff-Larsen K and Shaw DE 2011 How robust are protein folding simulations with respect to force field parameterization? *Biophys. J.* **100** L47–49

Pullicin AJ, Penner MH and Lim J 2017 Human taste detection of glucose oligomers with low degree of polymerization. *PLoS One* **12** e0183008

Rozengurt N, Wu SV, Chen MC, Huang C, Sternini C and Rozengurt E 2006 Colocalization of the α-subunit of gustducin with PYY and GLP-1 in L cells of human colon. *Am. J. Physiol. Gastrointestinal Liver Physiol.* **291** G792-G802

Samaei-Daryan S, Goliaei B and A. Ebrahim-Habibi 2017 Characterization of surface binding sites in glycoside hydrolases: a computational study. *J. Mol. Recog.* **30** e2624

Sclafani A and Mann S 1987 Carbohydrate taste preferences in rats: glucose, sucrose, maltose, fructose and polycose compared. *Physiol. Behav.* **40** 563–568

Seppälä S, Solomon KV, Gilmore SP, Henske JK and O'Malley MA 2016 Mapping the membrane proteome of anaerobic gut fungi identifies a wealth of carbohydrate binding proteins and transporters. *Microb. Cell Fact.* **15** 212

Servant G, Tachdjian C, Tang X-Q, Werner S, Zhang F, Li X, Kamdar P, Petrovic G, Ditschun T, Java A, Brust P, Brune N, DuBois GE, Zoller M and Karanewsky DS 2010 Positive allosteric modulators of the human sweet taste receptor enhance sweet taste. *Proc. Nat. Acad. Sci.USA* **107** 4746–4751

Spector AC and Schier LA 2018 Behavioral evidence that select carbohydrate stimuli activate T1R-independent receptor mechanisms. *Appetite* **122**: 26–31

Steinert R, Gerspach A, Gutmann H, Asarian L, Drewe J and Beglinger C 2011 The functional involvement of gut-expressed sweet taste receptors in glucose-stimulated secretion of glucagon-like peptide-1 (GLP-1) and peptide YY (PYY). *Clin. Nutr.* **30** 524–532

Sternini C, Anselmi L and Rozengurt E 2008 Enteroendocrine cells: a site of 'taste' in gastrointestinal chemosensing. *Curr. Opinion Endocrinol. Diabetes Obesity* **15** 73–78

Trott O and Olson AJ 2010 AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31** 455–461

Waksmonski JC and Koppel K 2016 Variation in human sweet taste receptor may result in different levels of sweet intensity variability between sweet stimuli. *Int. J. Food Sci. Technol.* **51** 1958–1966

Wang F, Song X, Zhou L, Liang G, Huang F, Jiang G and Zhang L 2017 The downregulation of sweet taste receptor signaling in enteroendocrine L-cells mediates 3-deoxyglucosone-induced attenuation of high glucose-stimulated GLP-1 secretion. *Arch. Physiol. Biochem.* **124** 430–435

Wang J, Wolf RM, Caldwell JW, Kollman PA and Case DA 2004 Development and testing of a general amber force field. *J. Comput. Chem.***25** 1157–1174

Xu H, Staszewski L, Tang H, Adler E, Zoller M and Li X 2004 Different functional roles of T1R subunits in the heteromeric taste receptors. *Proc. Nat. Acad. Sci. USA* **101** 14258–14263

Xu Z, Wang W, Nian X, Song G, Zhang X, Xiao H and Zhu X 2016 Dose-dependent effect of glucose on GLP-1 secretion involves sweet taste receptor in isolated perfused rat ileum. *Endocrine J.* **63** 1141–1147

Zhang F, Klebansky B, Fine RM, Liu H, Xu H, Servant G, Zoller M, Tachdjian C and Li X 2010 Molecular mechanism of the sweet taste enhancers. *Proc. Nat. Acad. Sci.* **107** 4752–4757

Corresponding editor: Ravindra Venkatramani