WATER SOLUBLE MEMBRANE PROTEINS AND METHODS FOR THE PREPARATION AND USE THEREOF

RELATED APPLICATION(S)

This application claims the benefit of U.S. Provisional Application No. 61/445,740, filed on February 23, 2011. The entire teachings of the above application(s) are incorporated herein by reference.

BACKGROUND OF THE INVENTION

Membrane proteins play vital roles in all living systems. Approximately ~30% of all genes in almost all sequenced genomes, code for membrane proteins. However, our detailed understanding of their structure and function lags far behind that of soluble proteins. As of February 2012, there are over 79,500 structures in the Protein Data Bank (http://www.rcsb.org/pdb/home/home.do), however, there are 952 membrane protein structures with 320 unique structures including 8 G-protein coupled receptors. Although there are about 400 functional olfactory receptors in human, not a single olfactory receptor has been determined.

There are several bottlenecks in elucidating the structure and function of olfactory receptors and their recognition and odorant-binding properties although they are of great interest. The most critical and challenging task is that it is extremely difficult to produce milligrams quantities of soluble and stable receptors. Inexpensive large-scale production methods are desperately needed, and have thus been the focus of extensive research. It is only possible to conduct detailed structural studies once these preliminary obstacles have been surmounted. Therefore, there is a need in the art for improved methods of studying G-protein coupled receptors, including olfactory receptors.

SUMMARY OF THE INVENTION

The present invention is directed to water-soluble membrane peptides, compositions comprising said peptides, methods for the preparation thereof and methods of use thereof.

The invention encompasses a water-soluble polypeptide comprising a modified α-helical domain, wherein the modified α-helical domain comprises an amino acid sequence in

which one or more hydrophobic amino acid residues within a α-helical domain of a native membrane protein is replaced with one or more hydrophilic amino acid residues. The invention also encompasses a method of preparing a water-soluble polypeptide comprising replacing one or more hydrophobic amino acid residues within the α-helical domain of a native membrane protein with one or more hydrophilic amino acid residues. The invention additionally encompasses a polypeptide prepared by replacing one or more hydrophobic amino acid residues within the α-helical domain of a native membrane protein with one or more hydrophilic amino acid residues.

The invention further encompasses a method of treatment for a disorder or disease that is mediated by the activity a membrane protein in a subject in need thereof, comprising administering to said subject an effective amount of a water-soluble polypeptide comprising a modified α-helical domain, wherein the modified α-helical domain comprises an amino acid sequence in which one or more hydrophobic amino acid residues within a α-helical domain of the membrane protein is replaced with one or more hydrophilic amino acid residues. In certain aspects, the water-soluble polypeptide retains the ligand-binding activity of the membrane protein. Examples of disorders and diseases that can be treated by administering a water-soluble peptide of the invention include, but are not limited to, cancer (such as, small cell lung cancer, melanoma, triple negative breast cancer), Parkinson's disease, cardiovascular disease, hypertension, and bronchial asthma.

The invention also encompasses a pharmaceutical composition comprising a water-soluble polypeptide of the invention and pharmaceutically acceptable carrier or diluent.

In some aspects, the α-helical domain is a 7-transmembrane α-helical domain. In an additional embodiment, the native membrane protein is a G-protein coupled receptor (GPCR). In some aspects of this embodiment, the GPCR is selected from the group comprising purinergic receptors ($P2Y_1$, $P2Y_2$, $P2Y_4$, $P2Y_6$), M_1 and M_3 muscarinic acetylcholine receptors, receptors for thrombin [protease-activated receptor (PAR)-1, PAR-2], thromboxane (TXA_2), sphingosine 1-phosphate ($S1P_2$, $S1P_3$, $S1P_4$ and $S1P_5$), lysophosphatidic acid (LPA_1, LPA_2, LPA_3), angiotensin II (AT_1), serotonin ($5\text{-}HT_{2c}$ and $5\text{-}HT_4$), somatostatin (sst_5), endothelin (ET_A and ET_B), cholecystokinin (CCK_1), V_{1a} vasopressin receptors, D_5 dopamine receptors, fMLP formyl peptide receptors, GAL_2 galanin receptors, EP_3 prostanoid receptors, A_1 adenosine receptors, α_1 adrenergic receptors, BB_2 bombesin receptors, B_2 bradykinin receptors, calcium-sensing receptors, chemokine

receptors, KSHV-ORF74 chemokine receptors, NK_1 tachykinin receptors, thyroid-stimulating hormone (TSH) receptors, protease-activated receptors, neuropeptide receptors, adenosine A2B receptors, P2Y purinoceptors, metabolic glutamate receptors, GRK5, GPCR-30, and CXCR4. In yet an additional embodiment, the native membrane protein or membrane protein is an integral membrane protein. In a further aspect, the native membrane protein is a mammalian protein. In yet a further aspect, the native membrane protein is an olfactory receptor. In additional embodiments, the olfactory receptor is mOR103-15.

In some aspects, the hydrophilic residues (which replace one or more hydrophobic residues in the α-helical domain of a native membrane protein) are selected from the group consisting of glutamine (Q), threonine (T), tyrosine (Y) and any combination thereof. In additional aspects, one or more hydrophobic residues selected from leucine (L), isoleucine (I), valine (V) and phenylalanine (F) are replaced.

In certain embodiments, one or more phenylalanine residues of the α-helical domain of the protein are replaced with tyrosine. In certain additional embodiments, one or more isoleucine and/or valine residues of the α-helical domain of the protein are replaced with threonine. In yet additional aspects, one or more phenylalanine residues of the α-helical domain of the protein are replaced with tyrosine and one or more isoleucine and/or valine residues of the α-helical domain of the protein are replaced with threonine. In additional embodiments, one or more leucine residues of the α-helical domain of the protein are replaced with glutamine. In yet additional embodiments, one or more leucine residues of the α-helical domain of the protein are replaced with glutamine and one or more isoleucine and/or valine residues of the protein are replaced with threonine. In further embodiments, one or more leucine residues of the α-helical domain of the protein are replaced with glutamine and one or more phenylalanine residues of the α-helical domain of the protein are replaced with tyrosine. In yet additional aspects, one or more leucine residues of the α-helical domain of the protein are replaced with glutamine, one or more phenylalanine residues of the α-helical domain of the protein are replaced with tyrosine, and one or more isoleucine and/or valine residues of the α-helical domain of the protein are replaced with threonine.

In additional embodiments, the water-soluble polypeptide retains at least some of the biological activity of the native membrane protein. In an aspect of this embodiment, the water-soluble polypeptide retains the ability to bind the ligand which normally binds to the

native membrane protein. In another embodiment, one or more amino acids within potential ligand binding sites of the native membrane protein are not replaced. In an aspect of this embodiment, examples of native membrane proteins with one or more amino acids not replaced within potential ligand-binding sites are purinergic receptors ($P2Y_1$, $P2Y_2$, $P2Y_4$, $P2Y_6$), M_1 and M_3 muscarinic acetylcholine receptors, receptors for thrombin [protease-activated receptor (PAR)-1, PAR-2], thromboxane (TXA_2), sphingosine 1-phosphate ($S1P_2$, $S1P_3$, $S1P_4$ and $S1P_5$), lysophosphatidic acid (LPA_1, LPA_2, LPA_3), angiotensin II (AT_1), serotonin ($5\text{-}HT_{2c}$ and $5\text{-}HT_4$), somatostatin (sst_5), endothelin (ET_A and ET_B), cholecystokinin (CCK_1), V_{1a} vasopressin receptors, D_5 dopamine receptors, fMLP formyl peptide receptors, GAL_2 galanin receptors, EP_3 prostanoid receptors, A_1 adenosine receptors, α_1 adrenergic receptors, BB_2 bombesin receptors, B_2 bradykinin receptors, calcium-sensing receptors, chemokine receptors, KSHV-ORF74 chemokine receptors, NK_1 tachykinin receptors, thyroid-stimulating hormone (TSH) receptors, protease-activated receptors, neuropeptide receptors, adenosine A2B receptors, P2Y purinoceptors, metabolic glutamate receptors, GRK5, GPCR-30, and CXCR4.

In another embodiment, one or more amino acids within potential odorant binding sites of the native membrane protein are not replaced.

In one embodiment, water-soluble polypeptide comprising a modified α-helical domain comprises the amino acid sequence of MERRNHTGRV SEFVLLGFPA PAPQRALQFF QSLQAYVQTL TENIQTITAI RNHPTLHKPM YYFLANMSFYL ETWYTTVTTP KMQAGYIGSE ENHGQLISFE ACMTQLYFFQ GLGCTECTLL AVMAYDRYVA TCHPLHYPVI VSSRQCVQMA AGSWAGGFGT SMTVKVYQISR LSYCGPNTIN HFFCDVSPLL NLSCTDMSTA ELTDFILAIF ILLGPLSVTG ASYMAITGAV MRIPSAAGRH KAFSTCASHL TTVITYYAAS IYTYARPKAL SAFDTNKLVS VLYAVIVPLL NPIIYCLRNQ EVKKALRRTL HLAQGDANT KKSSRDGGSS GTETSQVAPA (SEQ ID NO: 2). In yet an additional embodiment, the water-soluble polypeptide comprising a modified 7-transmembrane α-helical domain comprises one or more of the following amino acid sequences:

a. PQRALQFFQSLQAYVQTLTENIQTITAI R (SEQ ID NO: 3)
b. M YYFLANMSFYLETWYTTVTTPKMQAGYI (SEQ ID NO: 4)
c. CMTQLYFFQGLGCTECTLLAVMAYDRYVA TC (SEQ ID NO: 5)
d. RQCVQMAAGSWAGGFGTSMTVKVYQ (SEQ ID NO: 6)

e. LTDFILAIFILLGPLSVTGASYMAITGAV (SEQ ID NO: 7)

f. HKAFSTCASHLTTVITYYAAS IYTY (SEQ ID NO: 8)

g. TNKLVSVLYAVIVPLLNPIIYCLRN (SEQ ID NO: 9)

In certain aspects of the invention, the secondary structure of the water-soluble peptide is determined. In some embodiments, the secondary structure is determined using circular dichroism.

In certain embodiments, ligand binding to the water-soluble polypeptide is measured. In some aspects, ligand binding affinity of the water-soluble polypeptide is compared to that of the native protein. In additional aspects, ligand binding is measured using microscale thermophoresis, calcium influx assay or any combination thereof.

In yet an additional embodiment, the invention encompasses a cell transfected with a water-soluble peptide comprising a modified α-helical domain. In certain embodiments, the cell is a mammalian cell. One example of a mammalian cell that can be transfected is a HEK293 cell.

BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing and other objects, features and advantages of the invention will be apparent from the following more particular description of preferred embodiments of the invention, as illustrated in the accompanying drawings in which like reference characters refer to the same parts throughout the different views. The drawings are not necessarily to scale, emphasis instead being placed upon illustrating the principles of the invention.

FIG. 1 shows the amino acid sequences of native mOR103-15 and mutated mOR103-15 using glutamine, threonine and tyrosine (QTY) replacements. Use of QTY replacements to systematically mutate key residues on the 7-transmembrane α-helices to convert a water-insoluble olfactory receptor into a water-soluble one. We only change the positions of *b, c, f* with the more water-soluble residues Q, T, Y. These positions are on the hydrophilic face of the helices. We maintain the positions *a, d, e, g* that are on the hydrophobic face. It is likely that these changes will maintain the individual α -helices. The mutations are labeled in capital blue letters on top of the receptor sequence. The small letters, *abcdefg*, are helical wheel positions. The underlines are the locations of 7-transmembrane α -helices. The numbers (8, 7, 3, 5, 4, 5, 4) are mutations in each α -helix. There are 36-residue changes, ~10.5% of the total 340 residues.

FIG. 2 shows molecular models of a QTY Replacement olfactory receptor mOR103-15. A total 36 mutations have been made (~10.5%) in the 7-transmembrane helical segments. These mutations do not change the charged residues, so the variant receptor mass and pI remain largely unchanged. The molecular shapes and sizes of amino acids, Q, T, and Y are very similar to L, V/I and Y, so there are minimal overall local shape changes. A segment of 20 amino acids at the C-terminus are not modeled for clarity.

FIGs. 3A-3C A) Top view of the QTY replacements and B) side view of the QTY replacements. Note the mutations are only on one side of the helices. The native receptor without mutations has a folded structure similar to $\alpha_1$adrenergic receptor, whereas after mutation, the structure is similar to the $\beta_2$adrenergic receptor. C) Simulated structures of superimposed native mOR103-15 (red) and designed QTY mutation of mOR103-15 (blue). The overall structural difference is ~0.8Å average.

FIG. 4 Circular dichroism spectrum of CXCR4 and designed QTY mutation of CXCR4-QTY.

FIG. 5 SDS Gel showing comparison of molecular weight between native CXCR4 and CXCR4 with QTY mutations (SEQID NO:10: CXCR4 QTY).

FIG. 6 Use of QTY replacements to systematically mutate key residues on the 7-transmembrane alpha-helices and few other hydrophobic residues to convert the water-insoluble membrane form CXCR4 into a water-soluble form. **A**) We have changed positions *b, c, f* with the more water-soluble residues Q, T, Y. We do not change the positions *a, d, e, g*. These positions are believed to maintain the specific clustering of individual alpha-helices. **B**). The superimpositions of membrane form CXCR4 (red) and QTY water-soluble CRCR4 (blue). **C**) The native residues are labeled in red letters and **D**) mutations are labeled in blue letters in the sequence. A total of 29 QTY mutations among 352 residues have been made (about8.2%) in the seven transmembrane helical segments. These mutations do not change the charged residues, so the variant receptor mass and pI remain largely unchanged. The molecular shapes and sizes of amino acids, Q, T and Y are very similar to L, V/I and Y, so there are minimal overall, local shape changes.

DETAILED DESCRIPTION OF THE INVENTION

A description of preferred embodiments of the invention follows.

The words "a" or "an" are meant to encompass one or more, unless otherwise specified.

In some aspects, the invention is directed to the use of the QTY (Glutamine, threonine and tyrosine) replacement method to systematically change the 7-transmembrane α-helix hydrophobic residues leucine (L), isoleucine (I), valine (V), and phenylalanine (F) of a native protein to the hydrophilic residues glutamine (Q), threonine (T) and tyrosine (Y). This invention will convert the native membrane protein from a water-insoluble one to a water-soluble counterpart.

Another innovation of the invention is to convert the water-insoluble olfactory receptor mOR103-15 into a water-soluble one with about 10.5% specific residues changes (36aa/340aa). This will be accomplished by systematically and selectively changing key residues at the α-helical positions *b, c, f* that usually face the hydrophilic surface, while maintaining the hydrophobic residues at α-helical positions *a, d, e, g*. The synthetic biology design method is general and broadly applicable to the study of other olfactory receptors and G-protein coupled receptors. This strategy has the potential to overcome the bottleneck of crystallizing olfactory receptors, as well as additional GPCRs and other membrane proteins.

We used synthetic biology methods to convert a water-insoluble olfactory receptor into a water-soluble one with ~10.5% of the residues changes (36aa/340aa) (FIG. 1 and 2). We have systematically and selectively changed key residues at the α-helical positions *b, c, f* (which usually form the hydrophilic surface), but maintained the hydrophobic residues at α-helical positions *a, d, e, g* (FIG. 1). Our synthetic biology design method is general in nature, thus it is broadly applicable to the study other olfactory receptors, chemokine CXCR4 as well as other G-protein coupled receptors (GPCRs) and other membrane proteins. This simple strategy may partly overcome the bottleneck of structural studies of olfactory receptors, GPCRs, and other membrane proteins if the converted water-soluble membrane proteins remain biologically functional.

In order to facilitate the study of the structural aspects of olfactory receptors and their binding properties, we will use the QTY replacement method to design a water-soluble 7-bundle helical olfactory receptor mOR103-15 (FIGs. 1-3). It is known that seven amino acids have α-helical forming tendencies (32): leucine (L) (1.30), glutamine (Q) (1.27),

phenylalanine (F) (1.07), tyrosine (Y) (0.72), isoleucine (I) (0.97), valine (V) (0.91) and threonine (T) (0.82). We also know that side chains of Q, Y and T can all form hydrogen bonds with water: Q can form 4 H-bonds (2 H-donors from $-NH_2$, 2 H-acceptors from C=O), and T and Y can form 3 H-bonds each (-OH, 1-H donor from –H and 2 acceptors from -O). The Q, T, Y residues are more water-soluble than L, F, I, or V, which cannot form any hydrogen bonds with their side chains. The proposed substitutions will not have any positive- or negative-charges changes. Furthermore, the molecular shapes and sizes are very similar for the pairs: leucine/glutamine, phenylalanine/tyrosine, valine/threonine, and isoleucine/threonine (33-34). The proposed changes should thus increase the solubility of 7-transmembrane α--helices while maintaining the overall helical structure (FIG. 3C).

In this soluble olfactory receptor design, we have performed the following substitutions: phenylalanine to tyrosine (F->Y), isoleucine/valine to threonine (I/V->T), and leucine to glutamine (L->Q). The secondary structure of the water-soluble olfactory receptor, as well as measure its odorant-binding capabilities can be examined. If odorant-binding is detected with the QTY replacements, then it is likely that we have preserved important components of the original structure. The secondary structure and binding of the designed water-soluble olfactory receptor with the native olfactory receptor can be prepared. Milligram quantities of the water-soluble receptor can be produced and crystal screens can be set up with and without odorants.

In one embodiment, the native membrane protein is a G-protein coupled receptor (GPCR). In yet another embodiment, the native membrane protein is an olfactory receptor. In some embodiments, the olfactory receptor is a mammalian receptor. In yet another embodiment, the olfactory receptor is mOR103-15. In certain aspects, the water-soluble polypeptide retains at least some of the biological activity of the native membrane protein. In yet another aspect, the membrane protein is a membrane receptor that mediates a disease or condition.

In a further embodiment, the native membrane protein is a GPCR selected from the group comprising purinergic receptors ($P2Y_1$, $P2Y_2$, $P2Y_4$, $P2Y_6$), M_1 and M_3 muscarinic acetylcholine receptors, receptors for thrombin [protease-activated receptor (PAR)-1, PAR-2], thromboxane (TXA_2), sphingosine 1-phosphate ($S1P_2$, $S1P_3$, $S1P_4$ and $S1P_5$), lysophosphatidic acid (LPA_1, LPA_2, LPA_3), angiotensin II (AT_1), serotonin ($5\text{-}HT_{2c}$ and $5\text{-}HT_4$), somatostatin (sst_5), endothelin (ET_A and ET_B), cholecystokinin (CCK_1), V_{1a}

vasopressin receptors, D_5 dopamine receptors, fMLP formyl peptide receptors, GAL_2 galanin receptors, EP_3 prostanoid receptors, A_1 adenosine receptors, α_1 adrenergic receptors, BB_2 bombesin receptors, B_2 bradykinin receptors, calcium-sensing receptors, chemokine receptors, KSHV-ORF74 chemokine receptors, NK_1 tachykinin receptors, thyroid-stimulating hormone (TSH) receptors, protease-activated receptors, neuropeptide receptors, adenosine A2B receptors, P2Y purinoceptors, metabolic glutamate receptors, GRK5, GPCR-30, and CXCR4. In a further embodiment, the invention is directed to a pharmaceutical composition or method of treatment described herein wherein the native membrane protein is a GPCR selected from the group comprising purinergic receptors ($P2Y_1$, $P2Y_2$, $P2Y_4$, $P2Y_6$), M_1 and M_3 muscarinic acetylcholine receptors, receptors for thrombin [protease-activated receptor (PAR)-1, PAR-2], thromboxane (TXA_2), sphingosine 1-phosphate ($S1P_2$, $S1P_3$, $S1P_4$ and $S1P_5$), lysophosphatidic acid (LPA_1, LPA_2, LPA_3), angiotensin II (AT_1), serotonin ($5\text{-}HT_{2c}$ and $5\text{-}HT_4$), somatostatin (sst_5), endothelin (ET_A and ET_B), cholecystokinin (CCK_1), V_{1a} vasopressin receptors, D_5 dopamine receptors, fMLP formyl peptide receptors, GAL_2 galanin receptors, EP_3 prostanoid receptors, A_1 adenosine receptors, α_1 adrenergic receptors, BB_2 bombesin receptors, B_2 bradykinin receptors, calcium-sensing receptors, chemokine receptors, KSHV-ORF74 chemokine receptors, NK_1 tachykinin receptors, thyroid-stimulating hormone (TSH) receptors, protease-activated receptors, neuropeptide receptors, adenosine A2B receptors, P2Y purinoceptors, metabolic glutamate receptors, GRK5, GPCR-30, and CXCR4

In another embodiment, the water-soluble polypeptide retains the at least some of the ligand-binding activity of the membrane protein. In some embodiments, the GPCRs are mammalian receptors.

In a further embodiment, one or more amino acids within potential ligand binding sites of the native membrane protein are not replaced. In an aspect of this embodiment, examples of native membrane proteins with potential ligand-binding sites having one or more amino acids not replaced include purinergic receptors ($P2Y_1$, $P2Y_2$, $P2Y_4$, $P2Y_6$), M_1 and M_3 muscarinic acetylcholine receptors, receptors for thrombin [protease-activated receptor (PAR)-1, PAR-2], thromboxane (TXA_2), sphingosine 1-phosphate ($S1P_2$, $S1P_3$, $S1P_4$ and $S1P_5$), lysophosphatidic acid (LPA_1, LPA_2, LPA_3), angiotensin II (AT_1), serotonin ($5\text{-}HT_{2c}$ and $5\text{-}HT_4$), somatostatin (sst_5), endothelin (ET_A and ET_B), cholecystokinin (CCK_1), V_{1a} vasopressin receptors, D_5 dopamine receptors, fMLP formyl peptide receptors, GAL_2 galanin

receptors, EP_3 prostanoid receptors, A_1 adenosine receptors, α_1 adrenergic receptors, BB_2 bombesin receptors, B_2 bradykinin receptors, calcium-sensing receptors, chemokine receptors, KSHV-ORF74 chemokine receptors, NK_1 tachykinin receptors, thyroid-stimulating hormone (TSH) receptors, protease-activated receptors, neuropeptide receptors, adenosine A2B receptors, P2Y purinoceptors, metabolic glutamate receptors, GRK5, GPCR-30, and CXCR4.

The invention further encompasses a method of treatment for a disorder or disease that is mediated by the activity of a membrane protein, comprising the use of a water-soluble polypeptide to treat said disorders and diseases, wherein said water-soluble polypeptide comprises a modified α-helical domain, and wherein said water-soluble polypeptide retains the ligand-binding activity of the native membrane protein. Examples of such disorders and diseases include, but are not limited to, cancer, small cell lung cancer, melanoma, breast cancer, Parkinson's disease, cardiovascular disease, hypertension, and asthma.

As described herein, the water-soluble peptides described herein can be used for the treatment of conditions or diseases mediated by the activity of a membrane protein. In certain aspects, the water-soluble peptides can act as "decoys" for the membrane receptor and bind to the ligand that activates the membrane receptor. As such, the water-soluble peptides described herein can be used to reduce the activity of a membrane protein. These water-soluble peptides can remain in the circulation and bind to specific ligands, thereby reducing the activity of membrane bound receptors. For example, the GPCR CXCR4 is over-expressed in small cell lung cancer and facilitates metastasis of tumor cells. Binding of this ligand by a water-soluble peptide such as that described herein may significantly reduce metastasis.

The chemokine receptor, CXCR4, is known in viral research as a major coreceptor for the entry of T cell line-tropic HIV (Feng, et al. (1996) Science 272: 872-877; Davis, et al. (1997) *J Exp Med 186*: 1793-1798; Zaitseva, et al. (1997) Nat Med 3: 1369-1375; Sanchez, et al. (1997) *J Biol Chem 272*: 27529-27531). T Stromal cell derived factor 1 (SDF-1) is a chemokine that interacts specifically with CXCR4. When SDF-1 binds to CXCR4, CXCR4 activates Gαi protein-mediated signaling (pertussis toxin-sensitive) (Chen, et al. (1998) *Mol Pharmacol 53*: 177-181), including downstream kinase pathways such as Ras/MAP Kinases and phosphatidylinositol 3-kinase (PI3K)/Akt in lymphocyte, megakaryocytes, and hematopoietic stem cells (Bleul, et al. (1996) *Nature 382*: 829-833; Deng, et al. (1997) *Nature 388*: 296-300; Kijowski, et al. (2001) *Stem Cells 19*: 453-466; Majka, et al. (2001)

Folia. Histochem. Cytobiol. 39: 235-244; Sotsios, et al. (1999) *J. Immunol. 163*: 5954-5963; Vlahakis, et al. (2002*) J. Immunol. 169:* 5546-5554). In mice transplanted with human lymph nodes, SDF-1 induces CXCR4-positive cell migration into the transplanted lymph node (Blades, et al. (2002) *J. Immunol. 168*: 4308-4317).

Recently, studies have shown that CXCR4 interactions may regulate the migration of metastatic cells. Hypoxia, a reduction in partial oxygen pressure, is a microenvironmental change that occurs in most solid tumors and is a major inducer of tumor angiogenesis and therapeutic resistance. Hypoxia increases CXCR4 levels (Staller, et al. (2003) *Nature 425*: 307-311). Microarray analysis on a sub-population of cells from a bone metastatic model with elevated metastatic activity showed that one of the genes increased in the metastatic phenotype was CXCR4. Furthermore, overexpression CXCR4 in isolated cells significantly increased the metastatic activity (Kang, et al. (2003) *Cancer Cell 3*: 537-549). In samples collected from various breast cancer patients, Muller et al. (Muller, et al. (2001) *Nature 410*: 50-56) found that CXCR4 expression level is higher in primary tumors relative to normal mammary gland or epithelial cells. Moreover, CXCR4 antibody treatment has been shown to inhibit metastasis to regional lymph nodes when compared to control isotypes that all metastasized to lymph nodes and lungs (Muller, et al. (2001)). As such a decoy therapy model is suitable for treating CXCR4 mediated diseases and disorders.

In another embodiment of the invention relates to the treatment of a disease or disorder involving CXCR4-dependent chemotaxis, wherein the disease is associated with aberrant leukocyte recruitment or activation. The disease is selected from the group consisting of arthritis, psoriasis, multiple sclerosis, ulcerative colitis, Crohn's disease, allergy, asthma, AIDS associated encephalitis, AIDS related maculopapular skin eruption, AIDS related interstitial pneumonia, AIDS related enteropathy, AIDS related periportal hepatic inflammation and AIDS related glomerulo nephritis.

In another aspect, the invention relates to the treatment of a disease or disorder selected from arthritis, lymphoma, non-small lung cancer, lung cancer, breast cancer, prostate cancer, multiple sclerosis, central nervous system developmental disease, dementia, Parkinson's disease, Alzheimer's disease, tumor, fibroma, astrocytoma, myeloma, glioblastoma, an inflammatory disease, an organ transplantation rejection, AIDS, HIV-infection or angiogenesis.

The invention also encompasses a pharmaceutical composition comprising said water-soluble polypeptide and a pharmaceutically acceptable carrier or diluent.

The compositions can also include, depending on the formulation desired, pharmaceutically-acceptable, non-toxic carriers or diluents, which are defined as vehicles commonly used to formulate pharmaceutical compositions for animal or human administration. The diluent is selected so as not to affect the biological activity of the pharmacologic agent or composition. Examples of such diluents are distilled water, physiological phosphate-buffered saline, Ringer's solutions, dextrose solution, and Hank's solution. In addition, the pharmaceutical composition or formulation may also include other carriers, adjuvants, or nontoxic, nontherapeutic, nonimmunogenic stabilizers and the like. Pharmaceutical compositions can also include large, slowly metabolized macromolecules such as proteins, polysaccharides such as chitosan, polylactic acids, polyglycolic acids and copolymers (such as latex functionalized SEPHAROSE™, agarose, cellulose, and the like), polymeric amino acids, amino acid copolymers, and lipid aggregates (such as oil droplets or liposomes).

The compositions can be administered parenterally such as, for example, by intravenous, intramuscular, intrathecal or subcutaneous injection. Parenteral administration can be accomplished by incorporating a composition into a solution or suspension. Such solutions or suspensions may also include sterile diluents such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents. Parenteral formulations may also include antibacterial agents such as, for example, benzyl alcohol or methyl parabens, antioxidants such as, for example, ascorbic acid or sodium bisulfite and chelating agents such as EDTA. Buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose may also be added. The parenteral preparation can be enclosed in ampules, disposable syringes or multiple dose vials made of glass or plastic.

Additionally, auxiliary substances, such as wetting or emulsifying agents, surfactants, pH buffering substances and the like can be present in compositions. Other components of pharmaceutical compositions are those of petroleum, animal, vegetable, or synthetic origin, for example, peanut oil, soybean oil, and mineral oil. In general, glycols such as propylene glycol or polyethylene glycol are preferred liquid carriers, particularly for injectable solutions.

Injectable formulations can be prepared either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection can also be prepared. The preparation also can also be emulsified or encapsulated in liposomes or micro particles such as polylactide, polyglycolide, or copolymer for enhanced adjuvant effect, as discussed above. Langer, *Science* 249: 1527, 1990 and Hanes, *Advanced Drug Delivery Reviews 28*: 97-119, 1997. The compositions and pharmacologic agents described herein can be administered in the form of a depot injection or implant preparation which can be formulated in such a manner as to permit a sustained or pulsatile release of the active ingredient.

Transdermal administration includes percutaneous absorption of the composition through the skin. Transdermal formulations include patches, ointments, creams, gels, salves and the like. Transdermal delivery can be achieved using a skin patch or using transferosomes. [Paul et al., *Eur. J. Immunol. 25*: 3521-24, 1995; Cevc et al., *Biochem. Biophys. Acta* 1368: 201-15, 1998].

"Treating" or "treatment" includes preventing or delaying the onset of the symptoms, complications, or biochemical indicia of a disease, alleviating or ameliorating the symptoms or arresting or inhibiting further development of the disease, condition, or disorder. A "patient" is a human subject in need of treatment.

An "effective amount" refers to that amount of the therapeutic agent that is sufficient to ameliorate of one or more symptoms of a disorder and/or prevent advancement of a disorder, cause regression of the disorder and/or to achieve a desired effect.

The invention will be better understood in connection with the following example, which is intended as an illustration only and not limiting of the scope of the invention. Various changes and modifications to the disclosed embodiments will be apparent to those skilled in the art and such changes and may be made without departing from the spirit of the invention and the scope of the appended claims.

EXAMPLES

Example 1: Systematic analyses of the ligand-binding properties of olfactory receptors

The Q (Glutamine) T (Threonine) Y (Tyrosine) QTY replacement are used to convert a water-insoluble olfactory receptor to a water-soluble one for biochemical, biophysical and structural analyses. Our specific aims are to:

1) Use the QTY (Glutamine, threonine and tyrosine) replacement method to systematically change the 7-transmembrane α-helix hydrophobic residues leucine (L), isoleucine (I), valine (V), and phenylalanine (F) to the hydrophilic residues glutamine (Q), threonine (T) and tyrosine (Y). This method converts the protein from a water-insoluble olfactory receptor to a water-soluble one.

2) Produce and purify milligram quantities of native and bioengineered olfactory receptors using commercial cell-free *in vitro* translation systems (Invitrogen and Qiagen).

3) Determine the secondary structure of the purified olfactory receptors using circular dichroism (CD).

4) Determine the binding affinity of the native and bioengineered olfactory receptor variants using microscale thermophoresis.

5) Transfect the native and variant OR genes into HEK293 cells, and use calcium influx assays to measure odorant activation of the native and mutant olfactory receptors. These measurements will correlate the microscale thermophoresis binding data to functional responses within cells.

6) Systematically screen the native and bioengineered olfactory receptors for crystallizing conditions in the presence and absence of odorants and the presence and absence of detergent.

RESEARCH STRATEGY

Use QTY replacement to design a soluble 7-helical bundle olfactory receptor mOR103-15. An innovation of our study is to convert the water-insoluble olfactory receptor mORI03-15 into a water soluble one with about 10.5% specific residues changes (36aa/340aa). We have systematically and selectively changed key residues at the α-helical positions *b, c, f* that usually face the hydrophilic surface, while maintaining the hydrophobic residues at a-helical positions *a, d, e,* g. Our synthetic biology design method is general and broadly applicable to the study of other olfactory receptors and G-protein coupled receptors. This strategy has the potential to overcome the bottleneck of crystallizing olfactory receptors, as well as additional GPCRs and other membrane proteins. While our design to change the solubility of the sequence is focused on the *b,c,f* positions of the helical wheel, some further changes to other parts of the sequence can be made without significantly affecting the function or structure of the peptide, polypeptide or protein. For example conservative

mutations can be made.

The experimental approach

1) Use of QTY replacements to design a water-soluble 7-helical bundle olfactory receptor mOR103-15. We used synthetic biology methods to convert a water-insoluble olfactory receptor into a water-soluble one with ~10.5% of the residues changes (36aa/340aa) (FIGs. 1-3). We have systematically and selectively changed key residues at the α-helical positions *b,* c, *f* (which usually form the hydrophilic surface), but maintained the hydrophobic residues at α-helical positions *a, d, e, g* (FIG. 1). Our synthetic biology design method is general in nature, thus it is broadly applicable to the study other olfactory receptors as well as other G-protein coupled receptors (GPCRs). This simple strategy may partly overcome the bottleneck of structural studies of olfactory receptors, GPCRs, and other membrane proteins if the converted water-soluble membrane proteins remain biologically functional.

In order to facilitate the study of the structural aspects of olfactory receptors and their binding properties, we can use the QTY replacement method to design a water-soluble 7-bundle helical olfactory receptor mORI03-15 (FIGs. 1-3). It is known that seven amino acids have α-helical forming tendencies (32): leucine (L) (1.30), glutamine (Q) (1.27), phenylalanine (F) (1.07), tyrosine (Y) (0.72), isoleucine (I) (0.97), valine (V) (0.91) and threonine (T) (0.82). We also know that side chains of Q, Y and T can all form hydrogen bonds with water: Q can form 4 H-bonds (2 H-donors from $-NH_2$, 2 H-acceptors from C=O), and T and Y can form 3 H-bonds each (-OH, I-H donor from -H and 2 acceptors from -0). The Q, T, Y residues are more water-soluble than L, F, I, or V, which cannot form any hydrogen bonds with their side chains. The substitutions will not have any positive- or negative-charges changes. Furthermore, the molecular shapes and sizes are very similar for the pairs: leucine/glutamine, phenylalaine/tyrosine, valine/threonine, and isoleucine/threonine. The changes increase the solubility of 7-transmembrane a-helices while maintaining the overall helical structure.

In this soluble olfactory receptor design, we have performed the following substitutions: leucine -> glutamine (L->Q), isoleucine/valine -> threonine (IN->T) and phenylalanine -> tyrosine (F->Y). In the study, we can examine the secondary structure of the water-soluble olfactory receptor, as well as measure its odorant-binding capabilities. If odorant-binding is measured with the QTY replacements, then it is likely that we have

preserved important components of the original structure. We can compare the secondary structure and binding of native olfactory receptor with the designed water-soluble olfactory receptor. We can also produce milligram quantities of the water-soluble receptor, and set up crystal screens with and without odorants.

```
MERRNHTGRV SEFVLLGFPA PAPQRALQFF QSLQAYVQTL TENIQTITAI RNHPTLHKPM YYFLANMSYL ETWYTTVTTP
                       abcdefga bcdefgabcd efgabcdefg a        a bcdefgabcd efgabcdefg
KMQAGYIGSE ENHGQLISFE ACMTQLYFFQ GLGCTECTLL AVMAYDRYVA TCHPLHYPVI VSSRQCVQMA AGSWAGGFGT
abcdefg                abcdefgab cdefgabcde fgabcdefga bc            abcdefg abcdefgabc
SMTKVYQISR LSYCGPNTIN HFFCDVSPLL NLSCTDMSTA ELTDFIQAIY TLLGPLSTTG ASYMAITGAV MRIPSAAGRH
defgabcd                                   abcdefgab cdefgabcde fgabcdefga          a
KAFSTCASHL TTVITYYAAS IYTYARPKAL SAFDTNKLVS VLYAVITPLQ NPITYCQRNQ EVKKALRRTL HLAQGQDANT
bcdefgabcd efgabcdefg abcd           abcdef gabcdefgab cdefgabc
KKSSRDGGSS GTETSQVAPA. (36aa mutations/340aa, ~10.5% mutations)
```

2) Produce and purify milligram quantities of native and bioengineered variants of olfactory receptors. We can use commercial cell-free systems to produce milligrams of native and water-soluble mORI03-15. We can use the optimized protocols we have developed in our lab: this is the key advancement and innovation we have accomplished in the last few years. We can produce and purify the native and variant olfactory receptors in one day using immunoaffinity purification. Gel filtration can then be used to separate the monomeric and dimeric receptor forms.

3) Determine secondary structure using circular dichroism. We can use circular dichroism (CD) spectral analysis to measure the secondary structures of the purified receptors. CD is a very sensitive technique that is be able to detect any small structural changes between the native and mutant receptors. Specifically, CD analysis can be used to calculate the percentage of α-helices and β-sheets in a protein. If a proteins' structure is altered, it can be revealed in the CD analysis. In addition to determining whether specific mutations alter receptor structure, CD can also be used to measure any odorant-induced structural changes. See FIG. 4

4) Assay ligand-binding of olfactory receptors. Microscale thermophoresis are used to measure the binding affinity of the native and bioengineered proteins and their odorant ligands. The key advantages of this technique over SPR or other ligand binding technologies are that they are totally surface-free and label free. Thus, the receptors do not need to be modified. The measurements can be performed in solution using native tryptophan as a signal source. Additionally, small ligands (MW ~200 Daltons) can be reliably measured. Furthermore, each measurement needs 0.5J.t1 (lJlg/J.t1) of sample thus, save the precious receptor samples. These results show whether the mutant olfactory receptors are capable of

binding odorants as efficiently as the native protein.

5) Use calcium influx activation assay to measure olfactory receptor activation. We can use calcium influx assays to examine odorant-induced activation of the native and variant olfactory receptors in HEK293 cells. This data is be correlated to the microscale thermophoresis measurements. Microscale thermophoresis directly measures ligand binding, while calcium influx assays measure activation. Combined, these assays can verify whether specific mutations affect binding, activation, or both. Additionally, we can distinguish between agonist and antagonist ligands.

6) Systematic screen for crystallization conditions. We can systematically screen the native and bioengineered variant olfactory receptors for crystallizing conditions in the absence and presence of odorants. The technology for crystallization screening of water-soluble proteins is well developed. Commercial screens are available which supply a variety of precipitants, salts, buffers with fine tuned pH gradients, and a range of cationic and anionic substances. All of these variables are well known and will be used in crystallizing membrane proteins. An additional unique ingredient of membrane protein screens is the presence of one of more detergent molecules. However, precipitation techniques involving slow water removal from the hanging drop may continue to be effective. Although it is useful to form large crystals, the results of a crystal screen may yield smaller crystals.

Surface Plasmon Resonance analysis of CXCR4 QTY

Human CXCR4 and our CXCR4 QTY proteins obtained from cell-free production and purified with affinity beads were captured in different flow cells on a Biacore CM5 chip with immobilized 1D4 Antibody (Ab) in a Biacore 2000 instrument. Different concentrations of SDF1α, the native ligand for hCXCR4 receptor, were injected over the surface to allow interaction with the receptors.

HUMAN CXCR4 QTY (SEQ ID NO:10)

MEGISIYTSDNYTEEMGSGDYDSMKEPCFREENANYNKTFLPTIYSIIYQTGTVGNGL
VI
LVMGYQKKLRSMTDKYRLHLSTADLQFVTTLPYWATDATANWYFGNFLCKAVHVI
YTVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWTPAQLLTTPDY
TFANVSEADDRYICDRFYPNDLWVVVFQFQHIMVGLILPGIVILSCYCIIISKLSHSKG
HQKRKALKTTTTLIQAFFACWQPYYTGISIDSYILLEIIKQGCEFENTVHKWISITEAQ

AFYHCCTNPTQYAYLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVSTESESSS SFHS

Immobilization of 1D4 Antibody

Biacore CM5 chips were activated with 1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride and N-hydroxysuccinimide according to the manufacturer's protocol prior to a 7 minute injection at 5µl/min of 1D4 Ab to flow cells 2-4 at 70µg/ml followed by deactivating of the surfaces in all the 4 flow cells with a short Ethanolamine pulse. The immobilization level of 1D4 Ab range from 8000-25000 Response units (RU).

Capture of GPCRs

CXCR4 and CXCR4 QTY mutant are captured by the 1D4 Ab on the CM5 chip by injecting a 0.1mg/ml sample of the protein to a single flow cell at 5µl/min during 15min with both sample and running buffer containing 0.2% Fos-Choline-14 detergent. The receptors were captured to a level of 800-3000 RU.

Interaction analyzis

SDF1α were injected over all flow cells to allow interaction with both the receptors and flow cell one is used as a reference cell without any immobilized protein. Injections were made at 0, 7.8nM, 15.6nM, 31.25nM, 62.5nM, 125nM, 250nM, 500nM, 1µM in triplicates, at 20ul/min for 2 minutes with 15 min waiting time to allow dissociation. HBST (50mM Hepes, pH 7.4, 150mM NaCl, 0.005% Tween-20) with the addition of 0.2% BSA and 0.2% Fos-Choline-14 was used as both running buffer and for dilution of the SDF1α samples.

Conclusion: The above described study shows ligand binding by CXCR4 QTY.

References

1. Choma C, Gratkowski H, Lear JD & DeGrado WF. (2000) Asparagine-mediated self-association of a model transmembrane helix. *Nat Struct Biol* **7,** 161-6.

2. Slovic AM, Kono H, Lear JD, Saven JG & DeGrado WF. (2004) Computational design of water-soluble analogues of the potassium channel KcsA. *Proc Natl Acad Sci U SA* **101,** 1828-33.

3. Walters RF & DeGrado WF. (2006) Helix-packing motifs in membrane proteins. *Proc Natl Acad Sci U S A* **103,** 13658-63.

4. Zhang Y, Kulp DW, Lear JD & DeGrado WF. (2009) Experimental and computational evaluation of forces directing the association of transmembrane helices. *J Am Chem Soc* **131,** 11341-11343.

5

While this invention has been particularly shown and described with references to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the scope of the invention encompassed by the appended claims.

CLAIMS

What is claimed is:

1. A water-soluble polypeptide comprising a modified α-helical domain, wherein the modified α-helical domain comprises an amino acid sequence in which one or more hydrophobic amino acid residues within a α-helical domain of a native membrane protein is replaced with one or more hydrophilic amino acid residues.

2. The water-soluble polypeptide of claim 1, wherein the α-helical domain is 7-transmembrane α-helical domain.

3. The water-soluble polypeptide of claim 1, wherein the native membrane protein is a G-protein coupled receptor (GPCR).

4. The water-soluble polypeptide of claim 1, wherein the native membrane protein is an integral membrane protein.

5. The water-soluble polypeptide of claim 1, wherein the hydrophilic residues are selected from the group consisting of glutamine (Q), threonine (T), tyrosine (Y) and any combination thereof.

6. The water-soluble polypeptide of claim 1, wherein one or more of the hydrophobic residues leucine (L), isoleucine (I), valine (V) and phenylalanine (F) are replaced.

7. The water-soluble polypeptide of claim 1, wherein one or more phenylalanine residues of the α-helical domain of the native membrane protein are replaced with tyrosine.

8. The water-soluble polypeptide of any one of claims 1 and 7, wherein one or more isoleucine and/or valine residues of the α-helical domain of the protein are replaced with threonine.

9. The water-soluble polypeptide of any one of claims 1, 7 and 8, wherein one or more leucine residues of the α-helical domain of the native membrane protein are replaced with glutamine.

10. The water-soluble polypeptide of claim 1, wherein the native membrane protein is an olfactory receptor or chemokine receptor CXCR4.

11. The water-soluble polypeptide of claim 1, wherein the native membrane protein is a mammalian protein

12. The water-soluble polypeptide of claim 10, wherein the native membrane protein is mOR103-15 or or chemokine receptor CXCR4.

13. The water-soluble polypeptide of claim 12, wherein the hydrophobic amino acid residues at α-helical positions b, c and f are replaced.

14. The water-soluble polypeptide of claim 12, wherein the hydrophobic amino acid residues at α-helical positions a, d, e and g are not replaced.

15. The water-soluble polypeptide of claim 1, comprising the amino acid sequence of MERRNHTGRV SEFVLLGFPA PAPQRALQFF QSLQAYVQTL TENIQTITAIRNHPTLHKPM YYFLANMSFYL ETWYTTVTTP KMQAGYIGSE ENHGQLISFE ACMTQLYFFQ GLGCTECTLL AVMAYDRYVA TCHPLHYPVI VSSRQCVQMA AGSWAGGFGT SMTVKVYQISR LSYCGPNTIN HFFCDVSPLL NLSCTDMSTA ELTDFILAIF ILLGPLSVTG ASYMAITGAV MRIPSAAGRH KAFSTCASHL TTVITYYAAS IYTYARPKAL SAFDTNKLVS VLYAVIVPLL NPIIYCLRNQ EVKKALRRTL HLAQGDANT KKSSRDGGSS GTETSQVAPA (SEQ ID NO: 2);

or

MEGISIYTSDNYTEEMGSGDYDSMKEPCFREENANYNKTFLPTIYSIIYQTGTV GNGLVILVMGYQKKLRSMTDKYRLHLSTADLQFVTTLPYWATDATANWYF GNFLCKAVHVIYTVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKVVY VGVWTPAQLLTTPDYTFANVSEADDRYICDRFYPNDLWVVVFQFQHIMVGLI LPGIVILSCYCIIISKLSHSKGHQKRKALKTTTTLIQAFFACWQPYYTGISIDSYI LLEIIKQGCEFENTVHKWISTTEAQAFYHCCTNPTQYAYLGAKFKTSAQHALT SVSRGSSLKILSKGKRGGHSSVSTESESSSSFHS (SEQ ID NO:10).

16. The water-soluble polypeptide of claim 2, wherein the modified 7-transmembrane α-helical domain comprises one or more of the following amino acid sequences:

    a. PQRALQFFQSLQAYVQTLTENIQTITAI R (SEQ ID NO: 3)
    b. M YYFLANMSFYLETWYTTVTTPKMQAGYI (SEQ ID NO: 4)
    c. CMTQLYFFQGLGCTECTLLAVMAYDRYVA TC (SEQ ID NO: 5)
    d. RQCVQMAAGSWAGGFGTSMTVKVYQ (SEQ ID NO: 6)
    e. LTDFILAIFILLGPLSVTGASYMAITGAV (SEQ ID NO: 7)
    f. HKAFSTCASHLTTVITYYAAS IYTY (SEQ ID NO: 8)
    g. TNKLVSVLYAVIVPLLNPIIYCLRN (SEQ ID NO: 9).

17. The water-soluble polypeptide of any one of claims 1 to 16, wherein the polypeptide retains the biological activity of the native membrane protein.

18. The water-soluble polypeptide of any one of claims 1 to 16, wherein one or more amino acids within potential odorant binding sites are not replaced.

19. A method of preparing a water-soluble polypeptide comprising replacing one or more hydrophobic amino acid residues within the α-helical domain of a native membrane protein with one or more hydrophilic amino acid residues.

20. The method of claim 19, wherein the α-helical domain is a 7-transmembrane α-helical domain.

21. The method of claim 19, wherein the native membrane protein is a G-protein coupled receptor (GPCR).

22. The method of claim 19, wherein the hydrophilic residues are selected from the group consisting of glutamine (Q), threonine (T), tyrosine (Y) and any combination thereof.

23. The method of claim 19, wherein the hydrophobic residues leucine (L), isoleucine (I), valine (V) and phenylalanine (F) are replaced.

24. The method of claim 20, wherein one or more phenylalanine residues of the 7-transmembrane α-helical domain of the native membrane protein are replaced with tyrosine.

25. The method of any one of claims 19 and 24, wherein one or more isoleucine and/or valine residues of the α-helical domain of the native membrane protein are replaced with threonine.

26. The method of any one of claims 19, 24 and 25, wherein one or more leucine residues of the α-helical domain of the native membrane protein are replaced with glutamine.

27. The method of claim 19, wherein the native protein is an olfactory receptor.

28. The method of claim 19, wherein the native protein is a mammalian protein.

29. The method of claim 27, wherein the native protein is mOR103-15, or chemokine receptor CXCR4.

30. The method of claim 19, wherein the polypeptide comprises the amino acid sequence
MERRNHTGRV SEFVLLGFPA PAPQRALQFF QSLQAYVQTL TENIQTITAI
RNHPTLHKPM YYFLANMSFYL ETWYTTVTTP KMQAGYIGSE
ENHGQLISFE ACMTQLYFFQ GLGCTECTLL AVMAYDRYVA TCHPLHYPVI
VSSRQCVQMA AGSWAGGFGT SMTVKVYQISR LSYCGPNTIN
HFFCDVSPLL NLSCTDMSTA ELTDFILAIF ILLGPLSVTG ASYMAITGAV
MRIPSAAGRH KAFSTCASHL TTVITYYAAS IYTYARPKAL SAFDTNKLVS
VLYAVIVPLL NPIIYCLRNQ EVKKALRRTL HLAQGDANT KKSSRDGGSS
GTETSQVAPA (SEQ ID NO: 2);
or
MEGISIYTSDNYTEEMGSGDYDSMKEPCFREENANYNKTFLPTIYSIIYQTGTV
GNGLVILVMGYQKKLRSMTDKYRLHLSTADLQFVTTLPYWATDATANWYF
GNFLCKAVHVIYTVNLYSSVLILAFISLDRYLAIVHATNSQRPRKLLAEKVVY
VGVWTPAQLLTTPDYTFANVSEADDRYICDRFYPNDLWVVVFQFQHIMVGLI
LPGIVILSCYCIIISKLSHSKGHQKRKALKTTTTLIQAFFACWQPYYTGISIDSYI
LLEIIKQGCEFENTVHKWISTTEAQAFYHCCTNPTQYAYLGAKFKTSAQHALT
SVSRGSSLKILSKGKRGGHSSVSTESESSSSFHS (SEQ ID NO:10).

31. The method of claim 24, wherein the modified 7-transmembrane α-helical domain comprises one or more of the following amino acid sequences:

    a. PQRALQFFQSLQAYVQTLTENIQTITAI R (SEQ ID NO: 3)
    b. M YYFLANMSFYLETWYTTVTTPKMQAGYI (SEQ ID NO: 4)
    c. CMTQLYFFQGLGCTECTLLAVMAYDRYVA TC (SEQ ID NO: 5)
    d. RQCVQMAAGSWAGGFGTSMTVKVYQ (SEQ ID NO: 6)
    e. LTDFILAIFILLGPLSVTGASYMAITGAV (SEQ ID NO: 7)
    f. HKAFSTCASHLTTVITYYAAS IYTY (SEQ ID NO: 8)
    g. TNKLVSVLYAVIVPLLNPIIYCLRN (SEQ ID NO: 9) SEQ ID NO: 2.

32. The method of any one of claim 19 to 31, wherein the water-soluble polypeptide retains the biological activity of the native membrane protein.

33. The method of any one of claim 19 to 31, further comprising determining the secondary structure of the polypeptide.

34. The method of claim 33, wherein the secondary structure is determined using circular dichroism.

35. The method of any one of claims 19 to 31, further comprising measuring ligand binding.

36. The method of any one of claims 19 to 31, wherein the polypeptide is prepared using a cell-free system.

37. The method of any one of claim 19 to 31, wherein the protein is an olfactory receptor and further comprising measuring odorant binding to the olfactory receptor.

38. The method of claim 35, wherein the ligand binding affinity of the water-soluble polypeptide is compared to the native membrane protein.

39. The method of claim 35, wherein ligand binding is measured using microscale thermophoresis.

40. The method of claim 35, wherein ligand binding is measured using a calcium influx assay.

41. A polypeptide produced according to the method of any one of claims 19 to 31.

42. A cell transfected with the polypeptide of any one of claims 1 to 18 and 41.

43. The method of claim 42, wherein the cell is a mammalian cell.

44. The cell of claim 43, wherein the cell is a HEK293 cell.

45. A method for treating a mammal suffering from a disorder or disease that is mediated by the activity of a native membrane protein, comprising administering to said mammal an effective amount of a water-soluble polypeptide comprising a modified α-helical domain, wherein the modified α-helical domain comprises an amino acid sequence in which one or more hydrophobic amino acid residues within a α-helical domain of the native membrane protein is replaced with one or more hydrophilic amino acid residues.

46. The method of claim 45, wherein the α-helical domain is 7-transmembrane α-helical domain.

47. The method of claim 45, wherein the native membrane protein is a G-protein coupled receptor (GPCR).

48. The method of claim 45, wherein the native membrane protein is an integral membrane protein.

49. The method of claim 45, wherein the hydrophilic residues are selected from the group consisting of glutamine (Q), threonine (T), tyrosine (Y) and any combination thereof.

50. The method of claim 45, wherein one or more of the hydrophobic residues leucine (L), isoleucine (I), valine (V) and phenylalanine (F) are replaced.

51. The method of claim 45, wherein one or more phenylalanine residues of the α-helical domain of the native membrane protein are replaced with tyrosine.

52. The method of any one of claims 45 and 50, wherein one or more isoleucine and/or valine residues of the α-helical domain of the protein are replaced with threonine.

53. The method of any one of claims 45 and 50, wherein one or more leucine residues of the α-helical domain of the native membrane protein are replaced with glutamine.

54. The method of claim 45, wherein said native membrane protein is abnormally activated, up-regulated, or over-expressed in a disorder or disease.

55. The method of claim 45, wherein the disease is cancer.

56. The method of claim 55, wherein the cancer is small cell lung cancer.

57. The method of claim 55, wherein the cancer is melanoma.

58. The method of claim 55, wherein the cancer is breast cancer.

59. The method of claim 45, wherein the disease is Parkinson's disease.

60. The method of claim 45, wherein the disease is cardiovascular disease.

61. The method of claim 45, wherein the disease is hypertension.

62. The method of claim 45, wherein the disease is asthma.

63. The method of claim 45, wherein the native membrane protein is selected from the group comprising purinergic receptors ($P2Y_1$, $P2Y_2$, $P2Y_4$, $P2Y_6$), M_1 and M_3 muscarinic acetylcholine receptors, receptors for thrombin [protease-activated receptor (PAR)-1, PAR-2], thromboxane (TXA_2), sphingosine 1-phosphate ($S1P_2$, $S1P_3$, $S1P_4$ and $S1P_5$), lysophosphatidic acid (LPA_1, LPA_2, LPA_3), angiotensin II (AT_1), serotonin ($5\text{-}HT_{2c}$ and $5\text{-}HT_4$), somatostatin (sst_5), endothelin (ET_A and ET_B), cholecystokinin (CCK_1), V_{1a} vasopressin receptors, D_5 dopamine receptors, fMLP formyl peptide receptors, GAL_2 galanin receptors, EP_3 prostanoid receptors, A_1 adenosine receptors, α_1 adrenergic receptors, BB_2 bombesin receptors, B_2 bradykinin receptors, calcium-sensing receptors, chemokine receptors, KSHV-ORF74 chemokine receptors, NK_1 tachykinin receptors, thyroid-stimulating hormone (TSH) receptors,

protease-activated receptors, neuropeptide receptors, adenosine A2B receptors, P2Y purinoceptors, metabolic glutamate receptors, GRK5, GPCR-30, and CXCR4.

64. The method of any one of claims 45 to 63, wherein said water-soluble polypeptide retains the activity of binding the ligand of the native membrane protein.

65. The method of any one of claims 45 to 64, wherein said water-soluble polypeptide retains the ability to bind the ligand which normally binds to the native membrane protein.

66. The method of any one of claims 45 to 65, wherein one or more amino acids within potential ligand binding sites are not replaced.

67. The method of claim 45 wherein the mammal is a human.

68. The method according to any of claims 45-67, wherein said hydrophobic amino acid is present in the hydrophilic face of the α-helical domain.

69. The method according to any of claims 45-67, wherein said hydrophobic amino acid is present in the b, c or f position on a helical diagram.

70. The method according to any of claims 19-29, wherein about 10% to about 100% of the hydrophobic amino acid residues present in the hydrophilic face of α-helical domain is replaced.

71. The method according to claim 70, wherein at least about 15%, or about 25%, or about 50, or about 70% of the hydrophobic amino acid residues of the α-helical domain is replaced.

72. The method according to claim 70 or 71, wherein the helical content of the modified helical domain is at least about 30% or about 50% or about 75% or about 95% of the helical domain of the native protein.

73. A method for producing a native ligand binding modified polypeptide of a transmembrane protein comprising the step of replacing or mutating at least some of the hydrophobic amino acid residues by hydrophilic amino acid residues wherein said

modified polypeptide retains at least some of the ligand binding properties of said transmembrane protein.

74. A method for treating a disease or disorder mediated by ligand binding of a transmembrane protein comprising the step of administering a modified polypeptide wherein said polypeptide corresponds to a mutated amino acid sequence of said protein wherein transmembrane region of the sequence is modified so that at least some of the hydrophobic amino acid residues are replaced by hydrophilic amino acid residues.

75. The method according to claim 73 or 74, wherein said transmembrane region is α-helical.

76. The method according to claim 73 or 74, wherein hydrophobic amino acid residues corresponding to at least one of b, c or f positions of the helical wheel is replaced by hydrophilic amino acid residues.

77. The method according to claim 76, wherein said hydrophilic residues are selected from glutamine (Q), threonine (T), tyrosine (Y) and any combination thereof.

78. The method according to claims 75 or 76, wherein said hydrophobic residues are selected from leucine (L), isoleucine (I), valine (V) and phenylalanine (F).

79. The method according to any of claims 73-78, wherein said modified polypeptide retains ligand binding affinity of said transmembrane protein.

80. The method according to claim 79, wherein said ligand binding affinity of said polypeptide is substantially similar or better than the ligand binding affinity of said transmembrane protein.

81. The method according to claim 79, wherein said ligand binding affinity of said polypeptide is about 10%, or about 20% or about 30% or about 40% or about 50% or about 60% or about 70% or about 80% or about 90% of the ligand binding affinity of said transmembrane protein.

82. The method according to any of claims 73-81, wherein said modified polypeptide is water soluble.

83. The method according to any of claims 73-82, wherein said transmembrane protein is a G-protein coupled receptor (GPCR).

84. The method according to any of claims 73-82, wherein said transmembrane protein is selected from purinergic receptors ($P2Y_1$, $P2Y_2$, $P2Y_4$, $P2Y_6$), M_1 and M_3 muscarinic acetylcholine receptors, receptors for thrombin [protease-activated receptor (PAR)-1, PAR-2], thromboxane (TXA_2), sphingosine 1-phosphate ($S1P_2$, $S1P_3$, $S1P_4$ and $S1P_5$), lysophosphatidic acid (LPA_1, LPA_2, LPA_3), angiotensin II (AT_1), serotonin ($5\text{-}HT_{2c}$ and $5\text{-}HT_4$), somatostatin (sst_5), endothelin (ET_A and ET_B), cholecystokinin (CCK_1), V_{1a} vasopressin receptors, D_5 dopamine receptors, fMLP formyl peptide receptors, GAL_2 galanin receptors, EP_3 prostanoid receptors, A_1 adenosine receptors, α_1 adrenergic receptors, BB_2 bombesin receptors, B_2 bradykinin receptors, calcium-sensing receptors, chemokine receptors, KSHV-ORF74 chemokine receptors, NK_1 tachykinin receptors, thyroid-stimulating hormone (TSH) receptors, protease-activated receptors, neuropeptide receptors, adenosine A2B receptors, P2Y purinoceptors, metabolic glutamate receptors, GRK5, GPCR-30, and CXCR4.

85. The method according to claim 84, wherein said modified polypeptide binds to the corresponding native ligand.

86. The method according to claim 84 or 85, wherein said modified polypeptide is CXCR4 QTY.

87. The method according to any of claims 73-86, wherein said modified polypeptide is at least about 10% or about 20% or about 30% or about 40% or about 50% or about 60% or about 70% or about 80% or about 90% or about 100% or about 125% or about 150% or about 175% or about 200% or about 225% or about 250% or about 300% more water soluble than said transmembrane protein.

88. The method according to any of claims 45-54 or 63-87, wherein said disease or disorder is selected from arthritis, lymphoma, non-small lung cancer, lung cancer, breast cancer, prostate cancer, multiple sclerosis, central nervous system

developmental disease, dementia, Parkinson's disease, Alzheimer's disease, tumor, fibroma, astrocytoma, myeloma, glioblastoma, an inflammatory disease, an organ transplantation rejection, or angiogenesis.

89. A modified polypeptide produced by the method according to any of claims 83-87.

90. A pharmaceutical composition comprising an effective amount of a water-soluble polypeptide of any one of claims 1 to 9, 11 or 88, and a pharmaceutically acceptable diluent or carrier.

SEQUENCE LISTING

<110> Zhang, Shuguang
Rich, Alexander
Corin, Karolina
Tegler, Lotta T.

<120> Water Soluble Membrane Proteins and
Methods for the Preparation and Use Thereof

<130> 4041.3017 WO

<140> 13/403,725
<141> 2012-02-23

<150> US 61/445,740
<151> 2011-02-23

<160> 10

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 340
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 1
Met Glu Arg Arg Asn His Thr Gly Arg Val Ser Glu Phe Val Leu Leu
1 5 10 15
Gly Phe Pro Ala Pro Ala Pro Leu Arg Ala Leu Leu Phe Phe Leu Ser
20 25 30
Leu Leu Ala Tyr Val Leu Val Leu Thr Glu Asn Ile Leu Ile Thr Ala
35 40 45
Ile Arg Asn His Pro Thr Leu His Lys Pro Met Tyr Phe Phe Leu Ala
50 55 60
Asn Met Ser Glu Phe Leu Glu Ile Trp Tyr Val Thr Val Thr Ile Pro
65 70 75 80
Lys Met Leu Ala Gly Phe Ile Gly Ser Glu Glu Asn His Gly Gln Leu
85 90 95
Ile Ser Phe Glu Ala Cys Met Thr Gln Leu Tyr Phe Phe Leu Gly Leu
100 105 110
Gly Cys Thr Glu Cys Val Leu Leu Ala Val Met Ala Tyr Asp Arg Tyr
115 120 125
Val Ala Ile Cys His Pro Leu His Tyr Pro Val Ile Val Ser Ser Arg
130 135 140
Leu Cys Val Gln Met Ala Ala Gly Ser Trp Ala Gly Gly Phe Gly Ile
145 150 155 160
Ser Met Val Lys Val Phe Leu Ile Ser Arg Leu Tyr Thr Cys Gly Pro
165 170 175
Asn Thr Ile Asn His Phe Phe Cys Asp Val Ser Pro Leu Leu Asn Leu
180 185 190
Ser Cys Thr Asp Met Ser Thr Ala Glu Leu Thr Asp Phe Ile Leu Ala
195 200 205
Ile Phe Ile Leu Leu Gly Pro Leu Ser Val Thr Gly Ala Ser Tyr Met
210 215 220
Ala Ile Thr Gly Ala Val Met Arg Ile Pro Ser Ala Ala Gly Arg His
225 230 235 240
Lys Ala Phe Ser Thr Cys Ala Ser His Leu Thr Val Val Ile Ile Phe
245 250 255
Tyr Ala Ala Ser Ile Phe Ile Tyr Ala Arg Pro Lys Ala Leu Ser Ala
260 265 270

```
Phe Asp Thr Asn Lys Leu Val Ser Val Leu Tyr Ala Val Ile Val Pro
        275                 280                 285
Leu Leu Asn Pro Ile Ile Tyr Cys Leu Arg Asn Gln Glu Val Lys Lys
    290                 295                 300
Ala Leu Arg Arg Thr Leu His Leu Ala Gln Gly Gln Asp Ala Asn Thr
305                 310                 315                 320
Lys Lys Ser Ser Arg Asp Gly Gly Ser Ser Gly Thr Glu Thr Ser Gln
                325                 330                 335
Val Ala Pro Ala
            340


<210> 2
<211> 341
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 2
Met Glu Arg Arg Asn His Thr Gly Arg Val Ser Glu Phe Val Leu Leu
 1               5                  10                  15
Gly Phe Pro Ala Pro Ala Pro Gln Arg Ala Leu Gln Phe Phe Gln Ser
            20                  25                  30
Leu Gln Ala Tyr Val Gln Thr Leu Thr Glu Asn Ile Gln Thr Ile Thr
        35                  40                  45
Ala Ile Arg Asn His Pro Thr Leu His Lys Pro Met Tyr Tyr Phe Leu
    50                  55                  60
Ala Asn Met Ser Phe Tyr Leu Glu Thr Trp Tyr Thr Thr Val Thr Thr
65                  70                  75                  80
Pro Lys Met Gln Ala Gly Tyr Ile Gly Ser Glu Glu Asn His Gly Gln
                85                  90                  95
Leu Ile Ser Phe Glu Ala Cys Met Thr Gln Leu Tyr Phe Phe Gln Gly
            100                 105                 110
Leu Gly Cys Thr Glu Cys Thr Leu Leu Ala Val Met Ala Tyr Asp Arg
        115                 120                 125
Tyr Val Ala Thr Cys His Pro Leu His Tyr Pro Val Ile Val Ser Ser
    130                 135                 140
Arg Gln Cys Val Gln Met Ala Ala Gly Ser Trp Ala Gly Gly Phe Gly
145                 150                 155                 160
Thr Ser Met Thr Val Lys Val Tyr Gln Ile Ser Arg Leu Ser Tyr Cys
                165                 170                 175
Gly Pro Asn Thr Ile Asn His Phe Phe Cys Asp Val Ser Pro Leu Leu
            180                 185                 190
Asn Leu Ser Cys Thr Asp Met Ser Thr Ala Glu Leu Thr Asp Phe Ile
        195                 200                 205
Leu Ala Ile Phe Ile Leu Leu Gly Pro Leu Ser Val Thr Gly Ala Ser
    210                 215                 220
Tyr Met Ala Ile Thr Gly Ala Val Met Arg Ile Pro Ser Ala Ala Gly
225                 230                 235                 240
Arg His Lys Ala Phe Ser Thr Cys Ala Ser His Leu Thr Thr Val Ile
                245                 250                 255
Thr Tyr Tyr Ala Ala Ser Ile Tyr Thr Tyr Ala Arg Pro Lys Ala Leu
            260                 265                 270
Ser Ala Phe Asp Thr Asn Lys Leu Val Ser Val Leu Tyr Ala Val Ile
        275                 280                 285
Val Pro Leu Leu Asn Pro Ile Ile Tyr Cys Leu Arg Asn Gln Glu Val
    290                 295                 300
Lys Lys Ala Leu Arg Arg Thr Leu His Leu Ala Gln Gly Asp Ala Asn
305                 310                 315                 320
Thr Lys Lys Ser Ser Arg Asp Gly Gly Ser Ser Gly Thr Glu Thr Ser
                325                 330                 335
Gln Val Ala Pro Ala
            340


<210> 3
```

<211> 29
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 3
Pro Gln Arg Ala Leu Gln Phe Phe Gln Ser Leu Gln Ala Tyr Val Gln
 1   5   10   15
Thr Leu Thr Glu Asn Ile Gln Thr Ile Thr Ala Ile Arg
            20   25

<210> 4
<211> 29
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 4
Met Tyr Tyr Phe Leu Ala Asn Met Ser Phe Tyr Leu Glu Thr Trp Tyr
 1   5   10   15
Thr Thr Val Thr Thr Pro Lys Met Gln Ala Gly Tyr Ile
            20   25

<210> 5
<211> 31
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 5
Cys Met Thr Gln Leu Tyr Phe Phe Gln Gly Leu Gly Cys Thr Glu Cys
 1   5   10   15
Thr Leu Leu Ala Val Met Ala Tyr Asp Arg Tyr Val Ala Thr Cys
            20   25   30

<210> 6
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 6
Arg Gln Cys Val Gln Met Ala Ala Gly Ser Trp Ala Gly Gly Phe Gly
 1   5   10   15
Thr Ser Met Thr Val Lys Val Tyr Gln
            20   25

<210> 7
<211> 29
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 7

Leu Thr Asp Phe Ile Leu Ala Ile Phe Ile Leu Leu Gly Pro Leu Ser
1 5 10 15
Val Thr Gly Ala Ser Tyr Met Ala Ile Thr Gly Ala Val
20 25

<210> 8
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 8
His Lys Ala Phe Ser Thr Cys Ala Ser His Leu Thr Thr Val Ile Thr
1 5 10 15
Tyr Tyr Ala Ala Ser Ile Tyr Thr Tyr
20 25

<210> 9
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 9
Thr Asn Lys Leu Val Ser Val Leu Tyr Ala Val Ile Val Pro Leu Leu
1 5 10 15
Asn Pro Ile Ile Tyr Cys Leu Arg Asn
20 25

<210> 10
<211> 352
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 10

Met Glu Gly Ile Ser Ile Tyr Thr Ser Asp Asn Tyr Thr Glu Glu Met
1 5 10 15

Gly Ser Gly Asp Tyr Asp Ser Met Lys Glu Pro Cys Phe Arg Glu Glu
20 25 30

Asn Ala Asn Tyr Asn Lys Thr Phe Leu Pro Thr Ile Tyr Ser Ile Ile
35 40 45

Tyr Gln Thr Gly Thr Val Gly Asn Gly Leu Val Ile Leu Val Met Gly
50 55 60

Tyr Gln Lys Lys Leu Arg Ser Met Thr Asp Lys Tyr Arg Leu His Leu
65 70 75 80

Ser Thr Ala Asp Leu Gln Phe Val Thr Thr Leu Pro Tyr Trp Ala Thr
85 90 95

Asp Ala Thr Ala Asn Trp Tyr Phe Gly Asn Phe Leu Cys Lys Ala Val
100 105 110

His Val Ile Tyr Thr Val Asn Leu Tyr Ser Ser Val Leu Ile Leu Ala
115 120 125

Phe Ile Ser Leu Asp Arg Tyr Leu Ala Ile Val His Ala Thr Asn Ser
130 135 140

Gln Arg Pro Arg Lys Leu Leu Ala Glu Lys Val Val Tyr Val Gly Val
145 150 155 160

Trp Thr Pro Ala Gln Leu Leu Thr Thr Pro Asp Tyr Thr Phe Ala Asn
165 170 175

Val Ser Glu Ala Asp Asp Arg Tyr Ile Cys Asp Arg Phe Tyr Pro Asn
180 185 190

Asp Leu Trp Val Val Val Phe Gln Phe Gln His Ile Met Val Gly Leu
195 200 205

Ile Leu Pro Gly Ile Val Ile Leu Ser Cys Tyr Cys Ile Ile Ile Ser
210 215 220

Lys Leu Ser His Ser Lys Gly His Gln Lys Arg Lys Ala Leu Lys Thr
225 230 235 240

Thr Thr Thr Leu Ile Gln Ala Phe Phe Ala Cys Trp Gln Pro Tyr Tyr
245 250 255

Thr Gly Ile Ser Ile Asp Ser Tyr Ile Leu Leu Glu Ile Ile Lys Gln
260 265 270

Gly Cys Glu Phe Glu Asn Thr Val His Lys Trp Ile Ser Thr Thr Glu
275 280 285

Ala Gln Ala Phe Tyr His Cys Cys Thr Asn Pro Thr Gln Tyr Ala Tyr
290 295 300

Leu Gly Ala Lys Phe Lys Thr Ser Ala Gln His Ala Leu Thr Ser Val
305 310 315 320

Ser Arg Gly Ser Ser Leu Lys Ile Leu Ser Lys Gly Lys Arg Gly Gly
325 330 335

His Ser Ser Val Ser Thr Glu Ser Glu Ser Ser Ser Ser Phe His Ser
340 345 350