



(51) International Patent Classification:

A61K 38/00 (2006.01) *G01N 33/566* (2006.01)
G01N 33/567 (2006.01)

(21) International Application Number:

PCT/US2014/035524

(22) International Filing Date:

25 April 2014 (25.04.2014)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/815,939 25 April 2013 (25.04.2013) US

(71) Applicant: THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA [US/US]; 3160 Chestnut Street, Suite 200, Philadelphia, PA 19104 (US).

(72) Inventor; and

(71) Applicant : XI, Jin [CN/US]; 805 Parkridge Drive, Media, PA 19063 (US).

(72) Inventors: LIU, Renyu; 805 Parkridge Drive, Media, PA 19063 (US). SAVEN, Jeffery, G.; 515 South 46th Street, Philadelphia, PA 19143 (US). PEREZ-AGUILAR, Jose,

Manuel; 1249 Park Avenue, Apt. 12D, New York, NY 10029 (US). MATSUNAGA, Felipe; 27108 Laurel Lane, Olmsted Township, OH 44138 (US).

(74) Agents: CALDWELL, John, W. et al.; Baker & Hostetler LLP, Circa Centre, 12th Floor, 2929 Arch Street, Philadelphia, PA 19104 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,

[Continued on next page]

(54) Title: WATER SOLUBLE G-PROTEIN COUPLED RECEPTOR

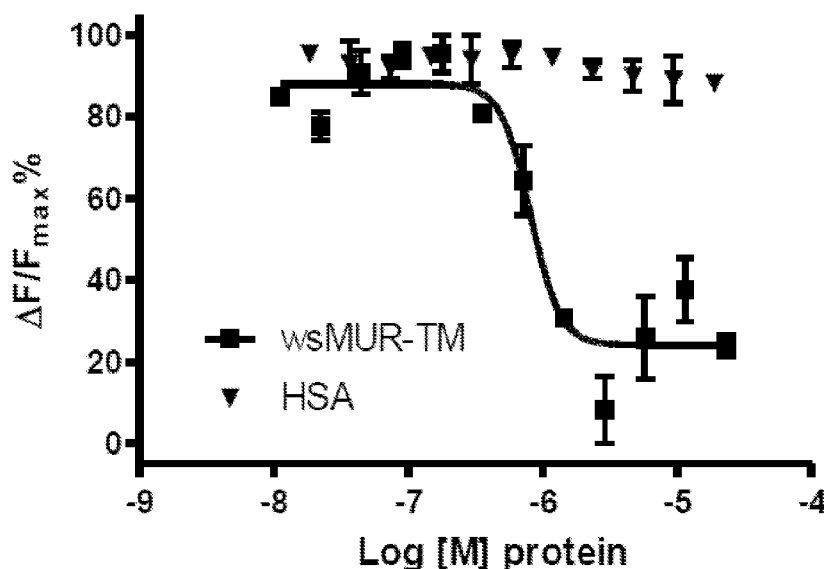


Figure 7

(57) Abstract: Described herein are recombinant integral membrane proteins having multiple transmembrane domains that have been engineered to be less hydrophobic, through alteration of the amino acid sequence of the native protein, but retain the ability to bind their natural ligand. The decreased hydrophobicity of the described proteins makes them more water soluble than the native protein, which allows the described proteins to be expressed in bacteria in large quantities and isolated in the absence of membranes, all while retaining the ability to interact with known ligands.



EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU,
LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK,
SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, KM, ML, MR, NE, SN, TD, TG).

— *before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments (Rule 48.2(h))*

Published:

— *with international search report (Art. 21(3))*

WATER SOLUBLE G-PROTEIN COUPLED RECEPTOR

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. Provisional Patent Application 61/815,939, filed on April 25, 2013, which is incorporated herein by reference in its entirety.

GOVERNMENT RIGHTS

[0002] The subject matter disclosed herein was made with government support under K08-GM-093115-01 awarded by the National Institutes of Health, and DMR-0425780, DMR08-32802, and DMR-1120901, which were awarded by the National Science Foundation. The Government has certain rights in the herein disclosed subject matter.

TECHNICAL FIELD

[0003] Described herein are recombinant, water soluble variants of membrane-spanning G-protein coupled receptors that may be expressed and isolated from bacteria in a manner that retains properties of the protein related to its functionality.

BACKGROUND

[0004] The G-protein-coupled receptor (GPCR) family of proteins have important roles in signal transduction and cellular response to extracellular stimuli. For this reason GPCRs are the target of many pharmaceuticals. The μ opioid receptor (MUR) is a GPCR that is the dominant target of opioids, many of which are potent analgesics widely used for the treatment of severe and chronic pain, e.g., morphine. Opioid use has soared in recent years and human MUR has been linked to abuse and many notorious side effects, including addiction and deadly respiratory depression.

[0005] The molecular mechanisms governing GPCR function remains obscure despite the profound insights obtained recently from multiple high-resolution crystal structures. Drug development and the study of the molecular mechanisms of GPCRs are impeded by limited solubility and difficulty in isolating sufficient quantities of functional receptors. These difficulties are caused in part by the large numbers of hydrophobic residues on the transmembrane, lipid-contacting protein exterior. Functional studies of MUR, and other GPCRs, could be carried out or greatly accelerated if forms of the protein existed that are water soluble, retain properties of native protein functionality, and are easily obtained in large quantity.

SUMMARY

[0006] Described herein are recombinant integral membrane proteins having multiple transmembrane domains computationally redesigned to increase their water solubility while retaining functionally related properties. The design involves several key steps: Comparative modeling using sequence alignment and known GPCR structures (the subsequently solved structure of murine MUR provided a means to assess the quality of the comparative model); Identification and computational redesign of transmembrane exterior residues; Overexpression in *E. coli* and purification; Characterization of structural and ligand-binding properties in aqueous buffer. The designed water-soluble human MUR has structurally and functionally related properties comparable to the native membrane-soluble human MUR.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] **Figure 1.** Model structure of the human μ opioid receptor transmembrane domain used during the computational design. (A) Comparative model structure of the transmembrane domain of the native human μ opioid receptor. (B) Model of the computationally designed transmembrane-only water-soluble variant (wsMUR-TM) of the human μ opioid receptor. Residues are colored by amino acid types: hydrophilic in gray (GNQSTY); hydrophobic in white (ACFILMPVW); basic in dark gray (HKR); and acidic in dark gray (DE).

[0008] **Figure 2.** Scheme of the computational design protocol. Homology modeling: Starting from the sequence alignment between known GPCR structures (bovine rhodopsin and β_2 adrenergic receptor) and MUR (A), 3D structures of MUR are generated (B). Identification of exposed sites in the transmembrane portion: A representative 3D model was selected from the generated models of MUR and the transmembrane lipid-exposed positions are identified (C; dark gray dots). Computational design of selected exterior positions to generate a water-soluble variant: The selected exterior positions are targeted of the computational design calculation with the intention to increase the protein's solubility in water. By maximizing an effective entropy function subject to different energy constraints, the computational approach generates site-specific probability profiles, that is, the probability of each amino acid to be present at each of the targeted sites. The amino acid identities of the sites where the probability of a particular amino acid is strongly favored (equal or larger than 0.8) was chosen to be that of this most probable amino acid (D; light gray dots). An iterative series of such calculations were performed until the probabilities of the different positions no longer fulfill these criteria (E; light gray dots).

At any remaining residues not yet specified with regard to amino acid identity, the most probable amino acid is selected (F; light gray dots).

[0009] Figure 3. (A) Sequences of the crystal structure of the mouse μ opioid receptor (PDB code 4DKL; Top) (1) and the human water-soluble variant wsMUR-TM (bottom). The murine sequence (top) corresponds to that whose structure is presented in the crystal structure of the mouse μ opioid receptor. The helical secondary structure is shown as rectangles. The gray residues in between TM5 and TM6 (MLSGSK) are absent in the crystal structure. The helical secondary structure of the wsMUR-TM model is indicated by lines under the sequence. (B) Superposition of the mouse μ opioid receptor (light gray) and the wsMUR-TM model (dark gray). (C) Rendering from the “extracellular” viewpoint of the crystal structure of mouse μ opioid receptor, where the side chain of the mutated positions in wsMUR are depicted as dark gray spheres. The majority of mutations (50 out of 55) are located at the exterior of the structure. Five remaining positions (see also residues in rectangular boxes in Figure 3A) are also rendered: Y130, T120, A306, N232, and K305. None of these positions are in direct contact with the irreversible antagonist β -FNA based on the crystal structure, where β -FNA was covalently attached to K235.

[0010] Figure 4. Overexpression and verification of wsMUR-TM. A SDS-PAGE gel for wsMUR-TM is shown where lane 1 correspond to the molecular weight standard, lane 2 to purified wsMUR-TM and lane 3 to expressed wsMUR-TM in the crude material. The band corresponding to the wsMUR-TM appears at approximately 36 kDa.

[0011] Figure 5. Mean residue ellipticity at 222 nm of wsMUR-TM in buffer solution (5 mM sodium phosphate, pH=7.0) as a function of temperature, from 10 to 90 °C. The spectrum of wsMUR-TM showed significant change near 62 °C and an almost complete loss in molar ellipticity at 90 °C.

[0012] Figure 6. Molar circular dichroism (CD) derived percentage of the original helical content (determined at 222 nm) of wsMUR-TM in the absence (inner-most dotted plot) and the presence (outer-most dotted plot) of cholesterol in buffer solution (5 mM sodium phosphate, 0.01% SDS, pH=7.0) as functions of the temperature. The addition of cholesterol stabilized the wsMUR-TM as indicated by the rightward shift of the thermostability curve.

[0013] Figure 7. Binding competition assay between the human μ opioid receptor expressed in HEK293 cells and the μ opioid water-soluble variants. Inhibition of the native μ opioid receptor constitutive signal in the presence of increasing concentrations of wsMUR-FL (dots, $IC_{50} = 8.4 \times 10^{-7}$ M, $R^2 = 0.9306$) or wsMUR-TM (squares, $IC_{50} = 8.6 \times 10^{-7}$ M, $R^2 = 0.9067$) in sodium

phosphate buffer. Data for the negative control is also included, HSA (inverted triangles). Data is used to calculate HTRF ratios, and represent the mean $\pm$ standard error of mean of quadruplicates. ΔF is used for the comparison of different runs of the same assay which reflects the signal to background of the assay. $\Delta F = [(Ratio_{sample} - Ratio_{background}) / Ratio_{background}] (\%)$.

[0014] **Figure 8.** Expression and purification of wsMUR-TM (SEQ ID NO: 2).

[0015] **Figure 9.** The secondary structure of wsMUR as indicated by CD spectra analysis.

[0016] **Figure 10.** The specific interaction of naltrexone with the wsMUR, similar to that indicated in figure 7.

[0017] **Figure 11.** Expression of 4 different versions of the wsMURs (SEQ ID NOs: 3-6). All 4 version of the receptors are expressed well in *E. Coli* and were purified successfully using affinity chromatography.

[0018] **Figure 12.** Four versions of the wsMUR demonstrate comparable affinities with naltrexone by using the methodology described for figure 7.

[0019] **Figure 13.** Expression and purification of a water-soluble variant of the beta-adrenergic receptor.

[0020] **Figure 14.** The secondary structure of the water soluble beta-adrenergic receptor as indicated by the CD spectra analysis.

DETAILED DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

[0021] Various terms relating to aspects of the description are used throughout the specification and claims. Such terms are to be given their ordinary meaning in the art unless otherwise indicated. Other specifically defined terms are to be construed in a manner consistent with the definitions provided herein.

[0022] As used in this specification and the appended claims, the singular forms “a,” “an,” and “the” include plural referents unless the content clearly dictates otherwise. Thus, for example, reference to “a cell” includes a combination of two or more cells, and the like.

[0023] The term “about” as used herein when referring to a measurable value such as an amount, a temporal duration, and the like, is meant to encompass variations of up to $\pm 10\%$ from the specified value, as such variations are appropriate to perform the disclosed methods. Unless otherwise indicated, all numbers expressing quantities of ingredients, properties such as molecular weight, reaction conditions, and so forth used in the specification and claims are to be understood as being modified in all instances by the term “about.”

[0024] Notwithstanding that the numerical ranges and parameters setting forth the broad scope of the invention are approximations, the numerical values set forth in the specific examples are reported as precisely as possible. Any numerical value, however, inherently contain certain errors necessarily resulting from the standard deviation found in their respective testing measurements.

[0025] “Isolated” means altered “by the hand of man” from the natural state. If a molecule or composition occurs in nature, it has been “isolated” if it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living plant or animal is not “isolated,” but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is “isolated” as the term is employed herein.

[0026] “Polynucleotide,” synonymously referred to as “nucleic acid molecule” or “nucleic acids,” refers to any polyribonucleotide or polydeoxyribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. “Polynucleotides” include, without limitation single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, “polynucleotide” refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The term polynucleotide also includes DNAs or RNAs containing one or more modified bases and DNAs or RNAs with backbones modified for stability or for other reasons. “Modified” bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications may be made to DNA and RNA; thus, “polynucleotide” embraces chemically, enzymatically or metabolically modified forms of polynucleotides as typically found in nature, as well as the chemical forms of DNA and RNA characteristic of viruses and cells. “Polynucleotide” also embraces relatively short nucleic acid chains, often referred to as oligonucleotides.

[0027] A “vector” is a replicon, such as plasmid, phage, cosmid, or virus in which another nucleic acid segment may be operably inserted so as to bring about the replication or expression of the segment.

[0028] A cell has been “transformed” when exogenous or heterologous nucleic acids such as DNA have been introduced inside the cell. The transforming DNA may or may not be integrated (covalently linked) into the genome of the cell. In prokaryotes, yeast, and mammalian

cells for example, the transforming DNA may be maintained on an episomal element such as a plasmid. With respect to eukaryotic cells, a stably transformed cell, or “stable cell” is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population of daughter cells containing the transforming DNA. A “clone” is a population of cells derived from a single cell or common ancestor by mitosis. A “cell line” is a clone of a primary cell that is capable of stable growth in vitro for many generations. In some examples provided herein, cells are transformed by transfecting the cells with DNA.

[0029] The embodiments described herein are not limited to particular methods, reagents, compounds, compositions or biological systems, which can, of course, vary. Furthermore, the terminology used herein is for the purpose of describing particular antibodies or antigen-binding fragments only, and is not intended to be limiting.

[0030] Described herein are recombinant integral membrane proteins having multiple transmembrane domains that have been engineered to be less hydrophobic, through alteration of the amino acid sequence of the native protein, but retain the ability to bind their natural ligand. The decreased hydrophobicity of the described proteins makes them more water soluble than the native protein, which allows the described proteins to be expressed in bacteria in large quantities, and isolated in the absence of membranes, all while retaining the ability to interact with known ligands in the manner of the corresponding membrane protein.

[0031] In some embodiments the described recombinant integral membrane proteins have seven transmembrane domains, with 4 of these transmembrane domains each having at least 3 amino acid mutations that decrease the overall hydrophobicity of the recombinant integral membrane protein relative to the native protein. In another embodiment, the described recombinant integral membrane proteins having seven transmembrane domains, with at least 5 of these transmembrane domains each having at least 3 amino acid mutations that decrease the overall hydrophobicity of the recombinant integral membrane protein relative to the native protein. In some embodiments, the described recombinant integral membrane proteins are variants of a native protein characterized as a G-protein coupled receptor. For example, in some embodiments the described protein may be a recombinant form of a human mu opioid receptor. In another embodiment the described protein may be a recombinant form of a human β_2 adrenergic receptor.

[0032] In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 95% identical to that of SEQ ID NO: 1. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence

that is about 96% identical to that of SEQ ID NO: 1. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 97% identical to that of SEQ ID NO: 1. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 98% identical to that of SEQ ID NO: 1. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 99% identical to that of SEQ ID NO: 1. In one embodiment, the described recombinant integral membrane protein has the amino acid sequence of SEQ ID NO: 1.

[0033] In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 95% identical to that of SEQ ID NO: 2. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 96% identical to that of SEQ ID NO: 2. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 97% identical to that of SEQ ID NO: 2. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 98% identical to that of SEQ ID NO: 2. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 99% identical to that of SEQ ID NO: 2. In one embodiment, the described recombinant integral membrane protein has the amino acid sequence of SEQ ID NO: 2.

[0034] In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 95% identical to that of SEQ ID NO: 3. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 96% identical to that of SEQ ID NO: 3. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 97% identical to that of SEQ ID NO: 3. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 98% identical to that of SEQ ID NO: 3. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 99% identical to that of SEQ ID NO: 3. In one embodiment, the described recombinant integral membrane protein has the amino acid sequence of SEQ ID NO: 3.

[0035] In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 95% identical to that of SEQ ID NO: 4. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 96% identical to that of SEQ ID NO: 4. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 97% identical to that of SEQ ID NO: 4. In one embodiment, the described recombinant integral membrane

protein has an amino acid sequence that is about 98% identical to that of SEQ ID NO: 4. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 99% identical to that of SEQ ID NO: 4. In one embodiment, the described recombinant integral membrane protein has the amino acid sequence of SEQ ID NO: 4.

[0036] In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 95% identical to that of SEQ ID NO: 5. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 96% identical to that of SEQ ID NO: 5. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 97% identical to that of SEQ ID NO: 5. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 98% identical to that of SEQ ID NO: 5. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 99% identical to that of SEQ ID NO: 5. In one embodiment, the described recombinant integral membrane protein has the amino acid sequence of SEQ ID NO: 5.

[0037] In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 95% identical to that of SEQ ID NO: 6. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 96% identical to that of SEQ ID NO: 6. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 97% identical to that of SEQ ID NO: 6. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 98% identical to that of SEQ ID NO: 6. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 99% identical to that of SEQ ID NO: 6. In one embodiment, the described recombinant integral membrane protein has the amino acid sequence of SEQ ID NO: 6.

[0038] In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 95% identical to that of SEQ ID NO: 7. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 96% identical to that of SEQ ID NO: 7. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 97% identical to that of SEQ ID NO: 7. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 98% identical to that of SEQ ID NO: 7. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence

that is about 99% identical to that of SEQ ID NO: 7. In one embodiment, the described recombinant integral membrane protein has the amino acid sequence of SEQ ID NO: 7.

[0039] In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 95% identical to that of SEQ ID NO: 8. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 96% identical to that of SEQ ID NO: 8. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 97% identical to that of SEQ ID NO: 8. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 98% identical to that of SEQ ID NO: 8. In one embodiment, the described recombinant integral membrane protein has an amino acid sequence that is about 99% identical to that of SEQ ID NO: 8. In one embodiment, the described recombinant integral membrane protein has the amino acid sequence of SEQ ID NO: 8.

SEQ IN NO.	Construct Name	Amino Acid Sequence (excluding signal sequence)
1	wsMur-TM	SMITAIKIHEEYKKVCEEKGKGNKLVMEVIVRYTKMK TATNIYIFNLAKADALAESTLPFQSVNKLMTWPFGTI LCKKVISIDYYNMFTSIFTLCTMSVDRIYVCHPVKAL DFRTPRNAKEENEKNWKLSSSEIGKPVEKKATTKYRQG SIDCTLTFSHPTWYWEDKLKDEVFKKAFFEPVKKIKE CYGLMILRLKSVRMLSGSKEKDRNLRRITRMVLVVVE VFIKCWTEIHKYVKEGKLVITPETTFQTVSWHECIAKG YKNSCENPKLYEELDENFKRCFREFC
2	wsMUR-TM+7mut	SMITAIKILEEYKKVCEEGRKGNKLVMEVIVRYTKMK TATNIYIFNLAKADALAESTLPFQSVNKLMTWPFGTI LCKKVISIDYYNMFTSIFTLCTMSVDRIYVCHPVKAL DFRTPRNAKEHNEKNWKLSSSEIGKPVEKRATTKYRQG SIDCTLTFSHPTWYWEDKLKDTVFKKAFFEPVKKVIKE CYGLMILRLKSVRMLSGSKEKDRNLRRITRMVLVVVE VFIKCWTEIHKYVKEGKL
3	G-min	SEKKREKIFQEYKKVYEEGKEGNKLVVDVIERYTKM KTATNDYIRNLAEADMKATETLPYQSENYLKGTWPF GTEECKKVISQDYNNMFTSIETLKTMSEDRIYVCHPV KALDFRTPRDAQEKNKENWEKSKKIGEPVEKSATTKY RQGSIDCTLTFSHPTWYWENKQKQKVFEFAFKKPVEE IKKKHEEMQKRLKSVRMLSGSKEKDRNLRRITRMVM EVVQVFIIKHWDPIHKYVKDKAEKTIPETTFQTKKWHE SIIEGYKNSDHNPKLYDENDENFKRHFREFK

4	H-min	SEKKKEEIWKEYKEWIEKGKKGNKLVMEVIERYTKM KTATNDYIKNLAEADWKATETLPEQSKNYLEGTWPF GKEKCKEVISRDYYNMFTSIYTLKTMSKDRYIAVDHP VKALDFRTPREAKKENKKNWEESSKKIGEPVKKDATT KYRQGSIDCTLTFSHPTWYWENKQKEEVFKKA FEEPV KDIEEQKKKMDERLKSVRMLSGSKEKDRNLRRITRM VWEVVKKFFEKWPIHEEVKKKA EKTIPETTFQTEEW HKKIYEGYKNSEENPKLYDEKDENFKREFREFE
5	I-min	SEEKKKKIDEEYKKQIEEGKKGNKLVEDVIERYTKMK TATNIYIKNLAQADQGATKTLPEQSKNYLEGTWPF GK EKCKEVISKDYNNMFTSIWTLDTMSEDRYIAVEHPVK ALDFRTPRKAKEENKKNWEESSKKIGEPVKKEATTKYR QGSIDCTLTFSHPTWYWENKWKEEVFKKA FEEPVKKI EERKKKMEERLKSVRMLSGSKEKDRNLRRITRMVEN VVKRFEEHWKPIHERVKEKAKKTIPETTFQTEEWHKEI QKGYENSKENPKLYEKEDENFKREFREFK
6	D-min	SEETAEEIEKQYKEVIEKGKKGNKLVKEVIERYTKMK TATNIYIWNLAEADLKATETLPKQSQNYLEGTWPF GQ EDCKNVISIDYYNMFTSIWTLATMSEDRYIAVAHPVK ALDFRTPREA EKENKKNWEESSKKIGEPVKKDATT KYR QGSIDCTLTFSHPTWYWENDLKDDVFKA FEEPVKKI EEAYKKMQERLKSVRMLSGSKEKDRNLRRITRMVWK VVQIFIEAWDPIHKYVIEKAKETIPETTFQTEEWHKSIA EGYKNSAENPELYKKDDENFKRTFREFE

7	BAD3	MAHHHHHHVMSGQPGNGSAFLLAPNGSHAPDHDVTQ QRDEEWVKGQGGKKMSEIVKKIVEGNKLVITAIKKFER LQTVTNYFITS LAEADLKMGEAVVPYGA AHILKKMW TYGNKWCEYWTSIDVLTVTAS IETLDVIAEDRYKAITS PFKYQSELTKNKAREEIKKVWERSGKTSFDPIQKHKY RATHQEAINCYANETCCDFFTNQDYAKKSSKESFYEP LKKMKEVYSRVEQEAKRQLQKIDKSEGRFHVQNLSQ VEQDGRTGHGLRRSSKESLKEHKALKTLGEIMGTFTK QWEPFFKVNEEHVKQDNKIRKEEYIKLNWEGYKNSG ENPKIYERSPDFRIAFQELKSLRRSSLKAYGNGYSSNG NTGEQSGYHVEQEKENKLLAEDLPGTEDFVGHQGTV PSDNIDSPGRNASTNDSLL
8	BAD4	MAHHHHHHVMSGQPGNGSAFLLAPNGSHAPDHDVTQ QRDEEWVKG TGRQMSEIVKKIVEGNKLVITAIQKFER LQTVTNYFITS LAEADLKMGEAVVPYGA AHILKKMW TYGNRWCEYWTSIDVLTVTAS IETLDVIAEDRYKAITS PFKYQSELTKNKAREEIKKVWERSGKTSFDPIQKHKY RATHQEAINCYANETCCDFFTNQDYAKKSSKQSFYEP LQKMKDVYSRVEQEAKRQLQKIDKSEGRFHVQNLSQ VEQDGRTGHGLRRSSKESLKEHKALKTLGEIMGTFTR QWDPFFKVNEEHVKQDNKIRKEEYIKLNWEGYKNSG ENPKIYERSPDFRIAFQELRSLRRSSLKAYGNGYSSNG NTGEQSGYHVEQEKENKLLAEDLPGTEDFVGHQGTV PSDNIDSPGRNASTNDSLL
9	Native MUR-TM	SMITAITIMALYSIVCVVGLFGNFLVMYVIVRYTKMKT ATNIYIFNLALADALATSTLPFQSVNYLMGTWPFGTIL CKIVISIDYYNMFTSIFTLCTMSVDRYIAVCHPVKALDF RTPRNAKIINVCNWILSSAIGLPVMFMATTKYRQGSID CTLTFSHPTWYWENLLKICVFIFAFIMPVLIITVCYGLM ILRLKSVRMLSGSKEKDRNLRRITRMVLVVAVFIVC WTPIHIVIIKALVTIPETTFQTVSWHFCIALGYTNSCL NPVLYAFLDENFKRCFREFC

[0040] In some aspects the described recombinant integral membrane proteins can be further modified to have additional sequences present such as a signal sequence or an epitope tag to allow for selective binding or purification of the protein without the need to contact structural epitopes of the variant protein. As discussed herein, the epitope tag may be a polyhistidine tag or an HA epitope tag. In some embodiments the polyhistidine tag will include at least 5 consecutive histidine amino acid residues.

[0041] Polynucleotides encoding the described nucleotide sequences are also within the scope of the subject matter described herein. A polynucleotide encoding any one of the amino acid sequences for the described recombinant integral membrane proteins is provided. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 95% identical to that of SEQ ID NO: 1. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 96% identical to that of SEQ ID NO: 1. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 97% identical to that of SEQ ID NO: 1. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 98% identical to that of SEQ ID NO: 1. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 99% identical to that of SEQ ID NO: 1. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having the amino acid sequence of SEQ ID NO: 1.

[0042] In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 95% identical to that of SEQ ID NO: 2. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 96% identical to that of SEQ ID NO: 2. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 97% identical to that of SEQ ID NO: 2. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 98% identical to that of SEQ ID NO: 2. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 99% identical to that of SEQ ID NO: 2. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having the amino acid sequence of SEQ ID NO: 2.

[0043] In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 95% identical to that of SEQ ID NO: 3. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 96% identical to that of SEQ ID NO: 3. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 97% identical to that of SEQ ID NO: 3. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 98% identical to that of SEQ ID NO: 3. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 99% identical to that of SEQ ID NO: 3. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having the amino acid sequence of SEQ ID NO: 3.

[0044] In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 95% identical to that of SEQ ID NO: 4. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 96% identical to that of SEQ ID NO: 4. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 97% identical to that of SEQ ID NO: 4. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 98% identical to that of SEQ ID NO: 4. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 99% identical to that of SEQ ID NO: 4. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having the amino acid sequence of SEQ ID NO: 4.

[0045] In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 95% identical to that of SEQ ID NO: 5. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 96% identical to that of SEQ ID NO: 5. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 97% identical to that of SEQ ID NO: 5. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 98% identical to that of SEQ ID NO: 5. In one embodiment the described polynucleotide encodes a recombinant integral

membrane protein having an amino acid sequence that is about 99% identical to that of SEQ ID NO: 5. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having the amino acid sequence of SEQ ID NO: 5.

[0046] In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 95% identical to that of SEQ ID NO: 6. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 96% identical to that of SEQ ID NO: 6. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 97% identical to that of SEQ ID NO: 6. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 98% identical to that of SEQ ID NO: 6. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 99% identical to that of SEQ ID NO: 6. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having the amino acid sequence of SEQ ID NO: 6.

[0047] In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 95% identical to that of SEQ ID NO: 7. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 96% identical to that of SEQ ID NO: 7. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 97% identical to that of SEQ ID NO: 7. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 98% identical to that of SEQ ID NO: 7. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 99% identical to that of SEQ ID NO: 7. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having the amino acid sequence of SEQ ID NO: 7.

[0048] In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 95% identical to that of SEQ ID NO: 8. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 96% identical to that of SEQ ID NO: 8. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 97% identical to that of SEQ ID

NO: 8. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 98% identical to that of SEQ ID NO: 8. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having an amino acid sequence that is about 99% identical to that of SEQ ID NO: 8. In one embodiment the described polynucleotide encodes a recombinant integral membrane protein having the amino acid sequence of SEQ ID NO: 8.

[0049] In some embodiments the described polynucleotides may be a segment of a plasmid, vector, phage genome, YAC, or other gene expression system. The polynucleotides described herein may be used to transform bacteria, yeast, or mammalian cells to allow for expression of the protein that the polynucleotide encodes. Accordingly, described herein are bacteria transformed with a polynucleotide encoding any one of the recombinant integral membrane proteins described herein. In some embodiments the bacterium transformed with a polynucleotide encoding any one of the recombinant integral membrane proteins described herein may be *E. coli*.

[0050] Methods of use for the described proteins are also provided herein. In one embodiment the described recombinant integral membrane proteins may be used in a method of obtaining a recombinant, soluble integral membrane protein having seven transmembrane domains in bacteria by: expressing in bacteria a polynucleotide encoding the recombinant integral membrane protein described herein, lysing the bacteria, and collecting a recombinant, soluble integral membrane protein having seven transmembrane domains. The expressed recombinant protein may be collected from the bacterial culture supernatant, the lysed bacterial pellet, or both. Additionally, the recombinant integral membrane protein may be collected by any number of known methodologies, such as centrifugation, affinity chromatography, size exclusion chromatography, molecular weight filtration (such as dialysis or size exclusion centrifugation).

[0051] Also provided herein are methods of identifying a ligand for any one of the recombinant integral membrane proteins described herein by contacting the recombinant integral membrane protein of interest with a compound and determining whether the two have a specific interaction. In some embodiments a specific interaction between a compound and a recombinant integral membrane protein may be identified by determining a binding affinity between the two. Alternatively, the affinity of one of the recombinant integral membrane proteins described herein for a ligand could be determined by contacting the ligand with one or more such recombinant integral membrane proteins to determine the binding affinity between the two. The affinity of

the interaction may be determined by any number of mechanisms, such as calorimetry, spectral absorption, time-resolved fluorescence resonance energy transfer, or surface plasmon resonance. In some embodiments, the recombinant integral membrane protein may be attached to a surface, for example by conjugation to an antibody specific for a protein tag added to the recombinant protein, to allow one or more compounds to be tested for interaction with the protein. Similar methods could also be used to assess the structural changes the described recombinant integral membrane proteins undergo upon ligand binding. For example, in one embodiment the structure of the recombinant integral membrane protein could be assessed before and after ligand binding occurs.

[0052] The following examples are provided to describe the embodiments described herein with greater detail. They are intended to illustrate, not to limit, the embodiments.

Example 1 - Design of a Water-Soluble Variant of the Human MUR

[0053] Studies were initiated to produce a comparative model of the human MUR transmembrane domain (288 residues, 66-353) using known GPCR structures (Figure 1A). To identify the site-specific amino acid probabilities of the target positions, a statistical entropy-based formalism was used. Energy functions to quantify sequence-structure compatibility are derived from a molecular mechanics force field. To account for solvation effects and for the tendency of different amino acids to be exposed to or sequestered from water (hydrophobicity), an energy term (herein environmental energy) based on the local density of C_{β} atoms of each residue and parameterized using a large database of globular proteins was used. In this case the environmental energy term was constrained to a value expected for soluble proteins of 288 residues, the size of the segment of the human MUR encompassing the TM domain. The conformational variability of the amino acid residues is addressed using a rotamer library of side chain conformations. The site-specific probabilities of the amino acids at each of the target positions are determined by maximizing an effective entropy function subject to constraints on the two energies. These probabilities were used to identify specific sequences. Residues suitable for mutation were identified as exposed, hydrophobic amino acids. Exposure is determined via inspection of model and crystallographic structures, hydrophobic scoring of the amino acids based upon empirical energy scales, and the solvent accessible area calculated for each amino acid. This resulted in identifying 55 exterior amino acids suitable for mutation. After the residues suitable for mutations were identified, the remaining residues were fixed at their wild type identities, and their side chain conformations were allowed to vary to accommodate

designed mutations. All amino acids but proline and cysteine were permitted at each of the identified variable positions. A hydrophobicity scoring function (environmental energy) was applied and selected to have a value consistent with that of a globular water-soluble protein having 288 amino acids. Identification of sequence proceeded iteratively (Figure 2). In all, 55 exterior transmembrane residues were selected for the computational redesign. A first calculation using the method described above to calculate the site-specific probabilities of the amino acids at each of 55 variable residues identified 31 positions where the probability of the selected amino acid exceeded 0.8; each such residue was mutated to this most probable amino acid, yielding the following mutations: A75E^{1.37}, S78K^{1.40}, I79K^{1.41}, V83E^{1.45}, F89K^{1.51}, Y93E^{1.55}, T120E^{2.54}, K187K^{4.43}, I188E^{4.44}, V191E^{4.47}, C192K^{4.48}, A199E^{4.55}, L202K^{4.58}, M205E^{4.61}, N232D^{5.36}, L233K^{5.37}, I240K^{5.44}, F241K^{5.45}, I244E^{5.48}, M245E^{5.49}, L248K^{5.52}, V252E^{5.56}, A289E^{6.42}, V293K^{6.46}, P297E^{6.50}, I300K^{6.53}, I303K^{6.56}, I304E^{6.57}, A306K^{6.59}, L326K^{7.41}, and V336K^{7.51}. The superscript notation is consistent with the Ballesteros and Weinstein indexing system: (number of the transmembrane helix).(residue number relative to most conserved residue in transmembrane helix, which is assigned position 50). These residue identities were fixed in subsequent calculations. Similarly, second and third calculations specified one (V82E^{1.44}) and two (T72K^{1.34} and L333E^{7.48}), respectively, additional positions with the same probability threshold. Using the results of a fourth calculation, the most probable amino acid was selected at the remaining 21 positions, yielding a sequence and model structure for wsMUR-TM as presented in figure 1B. The designed sequence is presented in figure 3A. The recent structure of the closely related murine MUR provides an opportunity to evaluate the structure and the location of the mutated positions in wsMUR-TM. The human and mouse receptors have 94% sequence identity. The model of the human MUR and the murine crystal structure superimpose well (Figure 3B), particularly with regard to the transmembrane helices. Only five positions in wsMUR-TM were not located in the exterior of the murine structure (T120E, Y130K, N232D, K305G, and A306K) and could in principle affect ligand binding (Figure 3C). In the murine structure, however, these five positions residues were not among the residues that directly contact beta-Funaltrexamine (β -FNA), an irreversible antagonist of the receptor.

[0054] Other attempts to produce a water-soluble MUR protein were not successful. While some of these constructs did not express in bacteria, as was the case with the native protein, those that did express were not functional, including the native MUR protein. In all, wsMUR-TM was only one of 11 recombinant MUR constructs to have increased water solubility that could be expressed in bacteria and also bind to a native MUR ligand with comparable

affinity to the native protein. Following the production of wsMUR-TM; however, several other variants were produced that could be expressed in *E. coli* and also retained functionality (described in Example 5).

Example 2 – Expression and Purification of wsMUR-TM

[0055] A synthetic cDNA encoding the transmembrane-only water-soluble MUR variant (wsMUR-TM) was produced by DNA2.0 Inc. (Menlo Park, CA). The sequences were subcloned between the NdeI and XhoI restriction sites of the expression plasmid pET-28b(+) (EMD/Novagen®). *E. coli* BL21(DE3) cells (EMD/Novagen®) were used for expression. Cells were grown in shake flasks with Lysogeny broth medium with 30 µg/mL kanamycin to an optical density (OD) of 1.0, induced with 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) for 3 hours at 37 °C, then pelleted by centrifugation. Cell pellets were stored at 20 °C until purification. For solubility testing, 1 OD aliquots of cells were pelleted in microcentrifuge tubes, suspended in 150 µL of TE (50 mM Tris-HCl, 1 mM EDTA, pH=8.0), then shaken with 0.3 g of glass beads (0.1 mm diameter) for 5 min. Aliquots of the resulting lysates were spun in a microcentrifuge for 1 min. Aliquots of total lysate, or the supernatant and pellet fractions after centrifugation, were analyzed on reducing sodium dodecyl sulfate (SDS) gels.

[0056] Frozen cells from 250 mL of fermentation (500-550 ODs) were thawed, and then suspended in 33.5 mL of 50 mM Tris-HCl, 1 M urea, pH=8.0. Once the pellet was fully resuspended, EDTA was added to 1 mM, Triton™ X-100 to 1%, and hen egg lysozyme to 1 µg per OD of cells, in a total volume of 37 mL. After the slurry was incubated for 20 min at room temperature (RT), MgCl₂ was added to 3 mM, followed by 100 units of benzonase. The suspension was swirled, incubated another 5 min at RT, and then spun in an Oak Ridge tube at 10,000 rpm for 20 min at 20 °C in an SS-34 rotor (ravg = 6.98 cm, rmax = 10.70 cm).

[0057] The resulting pellet was resuspended into 35 mL of 50 mM Tris-HCl, 1 M urea, pH=8.0. Triton™ X-100 (1.5 mL of a 25% solution) and 2-mercaptoethanol (2-ME) was added to 40 mM. The tube was inverted several times, and then spun as above.

[0058] The following steps were designed to resemble those that had been used to dissolve and purify recombinant forms of native µ opioid receptor. The pellet from the above washes was resuspended into 5 mL of buffer phosphate Tris buffer (100 mM phosphate, 10 mM Tris, adjusted to pH=8.0 with NaOH) and dispersed by drawing through a pipet followed by a 25 gauge needle. The volume was then raised to 37 mL by addition of phosphate Tris buffer, and 2-ME was then added to 40 mM. The tube was inverted to mix, then spun as above.

[0059] The resulting pellet was dispersed into 36 mL of PT as described above. The suspension was then mixed with an equal volume of phosphate Tris buffer containing 0.2% SDS and 10 mM 2-ME. The suspension was rocked until it became almost clear (60-90 min). The suspension was then poured into two 38 mL Oak Ridge tubes. These were spun tube at 12,000 rpm for 20 minutes at 20 °C in an SS-34 rotor.

[0060] Since an initial exposure to 0.1% SDS was required during purification, the purified wsMUR-TM in solution may still contain small amounts of SDS due to the difficulty of removing SDS from proteins. In order to avoid protein aggregation, 0.01% of SDS was utilized in the final buffer solutions for functional assays. Using binding and crystallographic studies, it has been shown that such small amounts of SDS do not disrupt the tertiary structure and/or the ligand binding capabilities of some proteins. Conversely, a much higher concentration of SDS (0.1%) and other anionic detergents are required for the “solubilization” of the native human MUR.

[0061] Attempts to express the native full-length human MUR in *E. coli* were unsuccessful presumably due to the protein's toxicity. In contrast, wsMUR-TM expressed well and was isolated with high purity using affinity chromatography (Figure 4). The yield was ~20 mg/L of shake flask culture. An initial exposure to ~0.1% sodium dodecyl sulfate (SDS) was required to purify the receptors. After dialysis to remove non-bound SDS, the purified variant were soluble at 6 mg/mL in buffer solution (130 mM NaCl, 20 mM NaHPO₄, pH=7.0).

Example 3 - Protein Structure Characterization and Thermostability

[0062] The secondary structure of the water-soluble variant was determined through circular dichroism (CD). Circular dichroism (CD) spectra were recorded by using CD Spectrometer (Chirascan, AppliedPhotophysics Limited, Leatherhead, United Kingdom) with a scan speed of 1 nm/s and 1mm path length. Corresponding blanks were used for calibration for each assay and subtracted from raw data. Two data sets were recorded and averaged to increase the signal-to-noise ratio. The CDNN CD spectra deconvolution software was utilized to determine the secondary structure content of the proteins. CD spectroscopy for wsMUR-TM at different temperatures were recorded with 6 µM of the receptor in buffer (5 mM sodium phosphate, pH=7.0) from 10 °C to 90 °C in increments of 2 °C per min. Absorbance was maintained lower than 1.0 to ensure sufficient light transmission. The temperature-dependence curve was plotted using GraphPad Prism® (version 5, GraphPad Software, Inc. La Jolla).

[0063] The CD spectra indicated predominantly helical structures with a helical secondary structure content of ~48% (estimations based on the molar ellipticity over the range 205 to 260 nm). The comparison of the helical content with that of the native human MUR expressed in yeast system in the presence of high concentration of detergent (0.1% SDS) is presented in Table 1.

Table 1. Helical content comparison for the native and engineered receptors

205-260 nm	wsMUR-TM (pH 7.0 in NaHPO₄)	Native MUR (pH 7.0 + 0.1% SDS)
Helix	48.0%	40.6%
Turn	14.6%	18.9%
Others	37.4%	40.5%

wsMUR-TM: transmembrane-only water-soluble human mu receptor variant;

MUR: human μ receptor

[0064] As monitored by CD, wsMUR-TM started to lose ellipticity significantly near 62 °C and was almost fully unfolded at 90 °C (Figure 5). The stability of wsMUR-TM was also investigated upon addition of cholesterol, which has been found to modulate the stability of several GPCRs. The inclusion of cholesterol caused a shift of the melting point from 82.9 °C to 89.3 °C, suggesting that it may stabilize the helical structure of wsMUR-TM (Figure 6).

[0065] CD and intrinsic tryptophan fluorescence were used to probe disulfide bond formation in the water-soluble variant. The structure of wsMUR-TM was monitored with increasing concentrations of urea and the reducing agent 2-mercaptoethanol (2-ME). After addition of urea, the molar ellipticity at 222 nm and the intensity of the intrinsic tryptophan fluorescence of wsMUR-TM decreased. Even in 8 M urea, the protein retains some helical structure (Table 2). Upon addition of 2-ME, both the molar ellipticity and fluorescence further decreased, becoming more pronounced at the higher concentration of the reducing agent (200 mM). Thus the presence of an intramolecular disulfide bond is corroborated in the case of wsMUR-TM.

Table 2. Effects of denaturant and reducing agent on the wsMUR-TM

	None	Urea (8 M)	Urea (8 M) 2-ME (25 mM)	Urea (8 M) 2-ME (200 mM)
Molar Ellipticity (%; 222 nm)	100.0	40.0	25.1	0.0
Fluorescence Peak Intensity (%; 300-350 nm)	100.0	28.4	23.9	4.5

wsMUR-TM: transmembrane-only water-soluble human μ receptor variant;

Values are normalized to the condition without denaturant or reducing agent (None). 2-ME: 2-mercaptoethanol.

[0066] Intrinsic tryptophan fluorescence was used to provide qualitative information of the conformations adopted by the water-soluble receptors; wsMUR-TM contains just six tryptophan residues (W135^{2.69}, W194^{4.50}, W228^{EC2}, W230^{EC2}, W295^{6.48}, and W320^{7.35}). Of particular interest are the tryptophan residues located in the partially buried transmembrane locations of the model structure (positions 194, 295, and 320). The fluorescence associated with these residues is expected to be sensitive to the local hydrophobic environment and overall folding of the protein. The observed decrease in the tryptophan fluorescence and the red shift in the emission with increasing denaturant (urea) concentration suggest that at least some of these tryptophan residues are located in the interior of the protein.

[0067] The decrease of the tryptophan fluorescence under denaturing conditions and in the presence of 2-ME is consistent with the changes in CD spectra observed under similar conditions. The requirement of the reducing agent to fully denature and unfold the protein indicates the relevance of an intramolecular disulfide bond in stabilizing the receptor structure. Although these observations suggest the presence of a disulfide bond, they do not specify which bond is formed given the existence of 11 cysteine residues in wsMUR-TM. However, the CD and ligand-binding studies are consistent with the adoption of the proper protein tertiary structure and by extension with the formation of the correct disulfide bond.

Example 4 - Ligand Binding Properties of the wsMUR-TM

[0068] A recently developed methodology which uses a fluorescently labeled ligand and the native MUR was used to investigate the ligand-binding capabilities of the water-soluble receptors. Naltrexone binding was monitored using a competitive TR-FRET (time-resolved fluorescence resonance energy transfer) based assay with fluorescently labeled wild type MUR and a naltrexone-derived antagonist. The ratio of fluorescence emission at 665 nm and 620 nm decreased in a dose-dependent manner with increasing concentrations of wsMUR-TM. The determined K_d values for naltrexone were 65 ± 1.8 nM (wsMUR-TM) (Figure 7). As a negative control, human serum albumin (HSA, a soluble helical protein), rather than a water-soluble variant, was introduced with no significant change in the fluorescence ratio upon HSA addition.

[0069] This binding assay has been applied to study several GPCRs and particularly MUR, where the K_i values for the morphinan opioids naloxone and naltrindole were estimated (5.1 nM and 8.1 nM for naloxone and naltrindole, respectively) and found to be in agreement with values obtained using other techniques, wsMUR-TM competes with native MUR expressed in HEK293 cells for the potent opioid antagonist naltrexone. This study demonstrates that the wsMUR-TM can compete with the native MUR for the fluorescent antagonist with binding affinities in nM range. The HSA (negative control) results indicate that the interaction of the water-soluble variant with naltrexone is selective and specific.

Example 5 – Additional MUR constructs

[0070] Constructs having unique sequences, but similar properties to the wsMUR-TM construct were also produced and analysed as described above. One such construct is a second wsMUR recombinant protein (wsMUR-TM+7mut – SEQ ID NO: 2). Studies performed to characterize wsMUR-TM+7mut demonstrate its production and isolation using bacterial expression (Figure 8), its alpha-helical nature as measure by CD (Figure 9), and binding activity was also observed for related MUR constructs wsMUR-TM+7mut (Figure 10). Similar characteristics were observed for the MUR constructs G-min (SEQ ID NO: 3), H-min (SEQ ID NO: 4), I-min (SEQ ID NO: 5), and D-min (SEQ ID NO: 7) (see figures Figure 11 and 12).

**Example 6 – Production and Isolation of a Water-Soluble
Human Beta₂ Adrenergic Receptor**

[0071] Studies were conducted to engineer and generate a more water soluble human β_2 adrenergic receptor (BAD). After analyzing the native protein sequence, as described above for MUR, amino acid sequence changes were made to cause the engineered BAR to be less hydrophobic. Two recombinant BAR sequences were designed (SEQ ID NOs: 7 and 8). To assess expression and isolation from bacteria, *E. coli* were transformed with a construct encoding SEQ ID NO: 8 (BAD4), cultured and then lysed. BAD4 was identified on a western blot following purification from the bacterial cell lysate (Figure 13). The isolated protein was also assessed for helical structural content by CD spectroscopy and was shown to have a spectral profile consistent with high alpha-helical content (Figure 14).

What is Claimed:

1. A recombinant integral membrane protein having seven transmembrane domains, comprising 4 transmembrane domains each having at least 3 amino acid mutations that decrease the overall hydrophobicity of the recombinant integral membrane protein relative to that of the native protein.
2. The recombinant integral membrane protein of claim 1, comprising at least 5 transmembrane domains each having at least 3 amino acid mutations that decrease the overall hydrophobicity of the recombinant integral membrane protein relative to that of the native protein.
3. The recombinant integral membrane protein of claims 1 or 2, wherein said protein assumes an active conformation when bound to a native ligand.
4. The recombinant integral membrane protein of any one previous claim, wherein said protein is characterized as being a G-protein-coupled receptor (GPCR).
5. The recombinant integral membrane protein of claim 4, wherein the GPCR is a human mu opioid receptor or a human β_2 adrenergic receptor.
6. The recombinant integral membrane protein of any one previous claim, wherein the protein has an amino acid sequence that is at least 95% identical to SEQ ID NO: 1.
7. The recombinant integral membrane protein of any one previous claim, wherein the protein has an amino acid sequence of SEQ ID NO: 1.
8. The recombinant integral membrane protein of any one previous claim, wherein the protein is water soluble.
9. The recombinant integral membrane protein of any one previous claim, wherein the protein further comprises an epitope tag.

10. The recombinant integral membrane protein of claim 9, wherein the epitope tag comprises 5 consecutive histidine amino acids.
11. A polynucleotide encoding the recombinant integral membrane protein of any one previous claim.
12. A bacterium transformed with a polynucleotide encoding the recombinant integral membrane protein of any one previous claim.
13. The bacterium of claim 12, wherein said bacterium is *E. coli*.
14. A method of obtaining a recombinant, soluble integral membrane protein having seven transmembrane domains in bacteria, comprising: expressing in bacteria a polynucleotide encoding the recombinant integral membrane protein of claim 9, lysing said bacteria, and collecting a recombinant, soluble integral membrane protein having seven transmembrane domains.
15. A method of identifying a ligand for the recombinant integral membrane protein of any one of claims 1 to 10, comprising contacting said protein with a compound and determining the affinity of said protein for said compound.
16. The method of claim 15, wherein the recombinant integral membrane protein is attached to a surface.
17. The method of claim 15 or 16, wherein the affinity of the compound for the recombinant integral membrane protein is measured by calorimetry, spectral absorption, time-resolved fluorescence resonance energy transfer, or surface plasmon resonance.
18. A method of obtaining a recombinant, soluble integral membrane protein having seven transmembrane domains in bacteria comprising: expressing in bacteria a polynucleotide encoding the recombinant integral membrane protein of claim 1, lysing said bacteria, and collecting a recombinant, soluble integral membrane protein having seven transmembrane domains.

19. A method of identifying the relative affinities of a set of ligands for the recombinant integral membrane protein of any one of claims 1 to 10, comprising contacting said protein with said ligands and assessing the affinities of said protein for said ligands.
20. A method of identifying the relative affinities of a set of recombinant integral membrane proteins for a given ligand, comprising contacting said ligand with one or more of the proteins of any one of claims 1 to 10 and assessing the affinities of said proteins for said ligand.
21. A method of assessing structural changes in a recombinant integral membrane protein of any one of claims 1 to 10, comprising contacting said recombinant integral membrane protein with a compound and assessing the structure of the compound.
22. A method of identifying a ligand for the recombinant integral membrane protein of any one of claims 1 to 10, comprising contacting said protein with a compound.
23. A method of identifying recombinant integral membrane protein(s) of any one of claims 1 to 10 that bind a given ligand, comprising contacting said protein(s) with a compound.
24. A method of identifying recombinant integral membrane protein(s) of any one of claims 1 to 10 that bind a given ligand, comprising contacting said protein(s) with a compound and determining the affinities of said protein(s) for said compound.

1/14

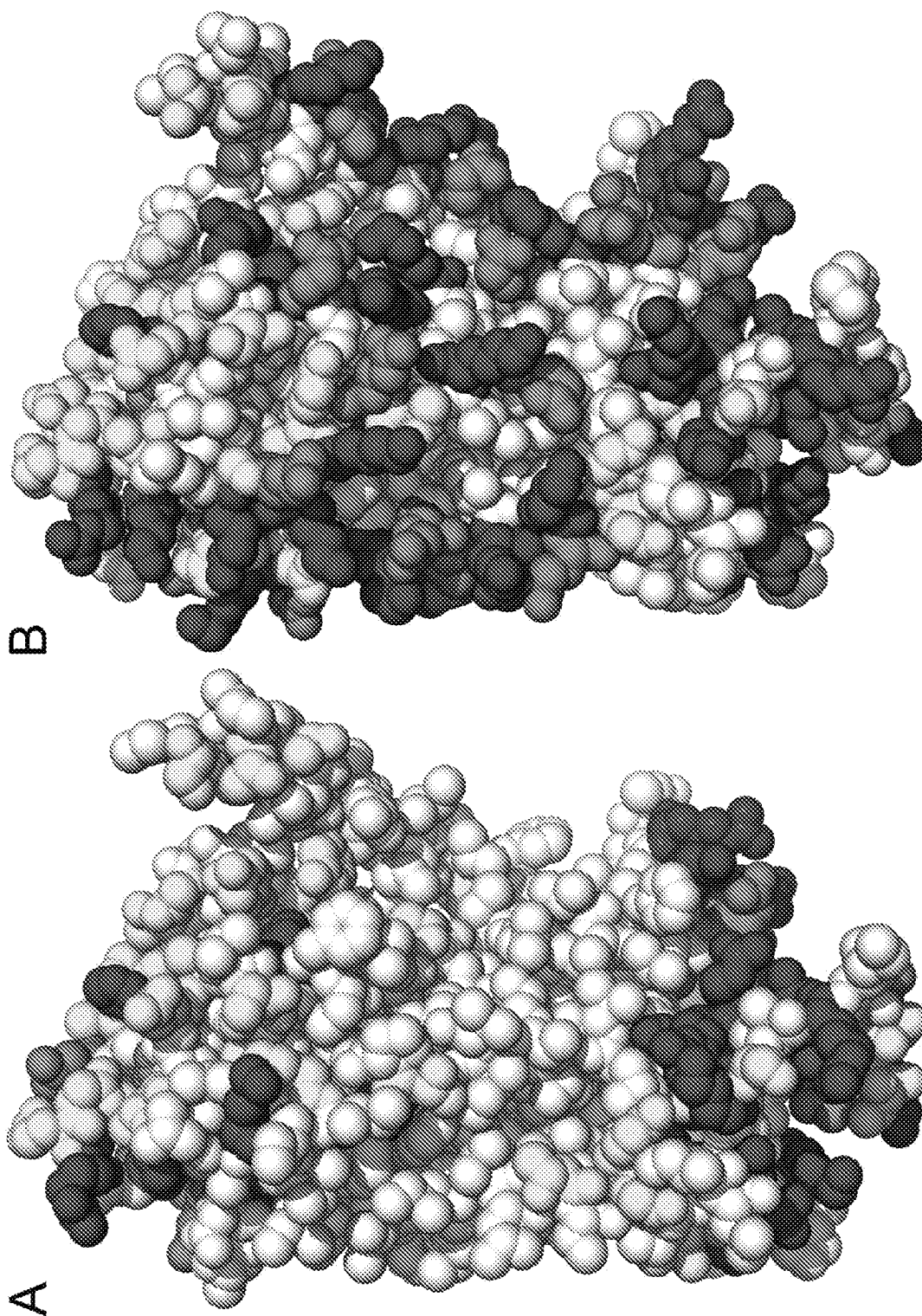


Figure 1

2/14

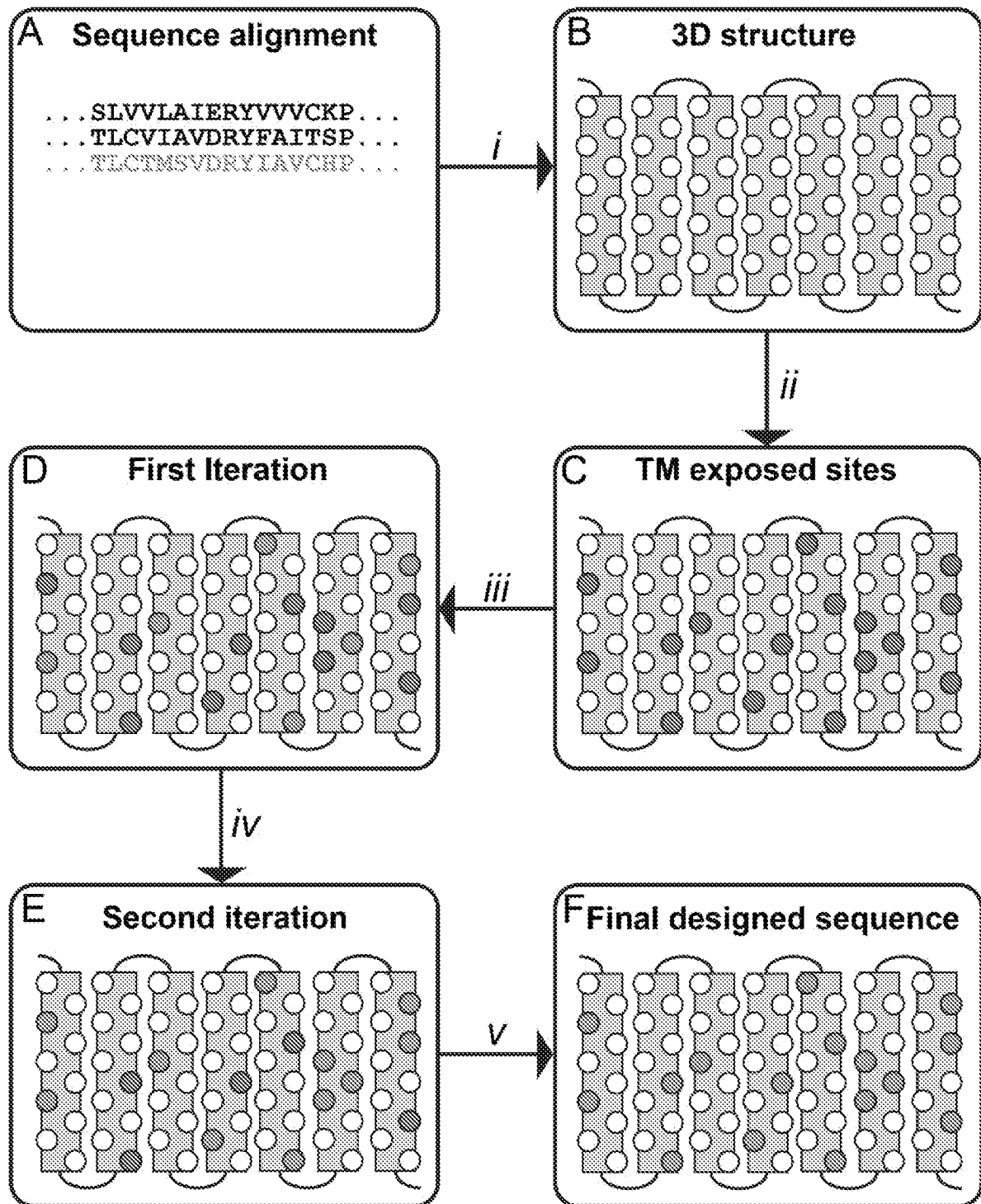


Figure 2

3/14

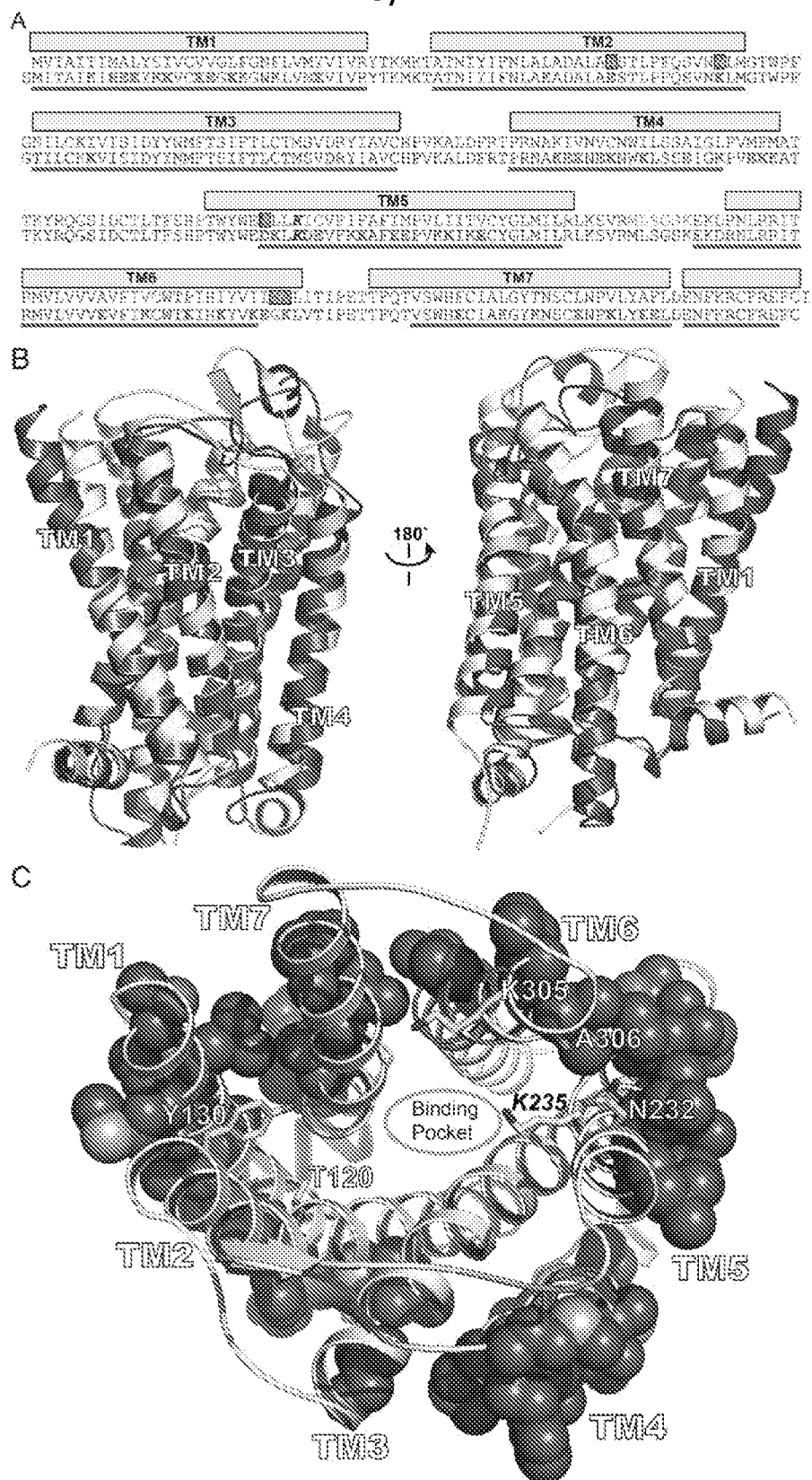


Figure 3

4/14

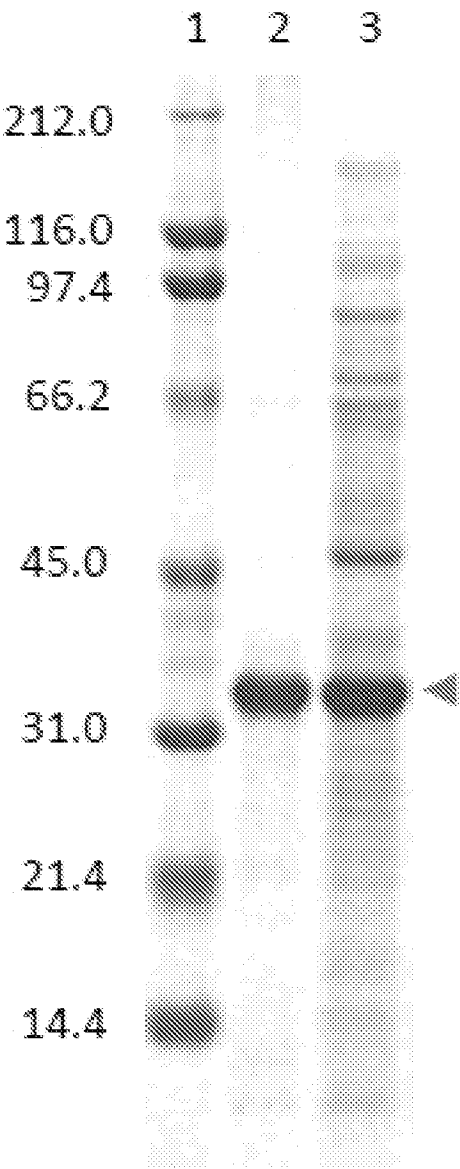


Figure 4

5/14

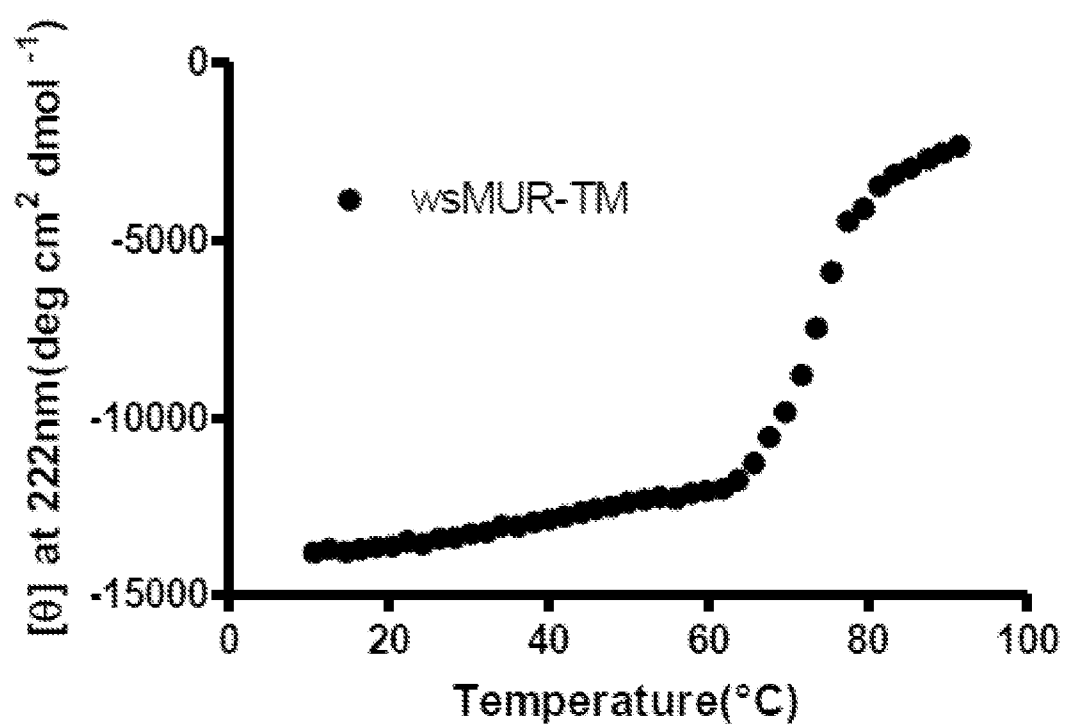


Figure 5

6/14

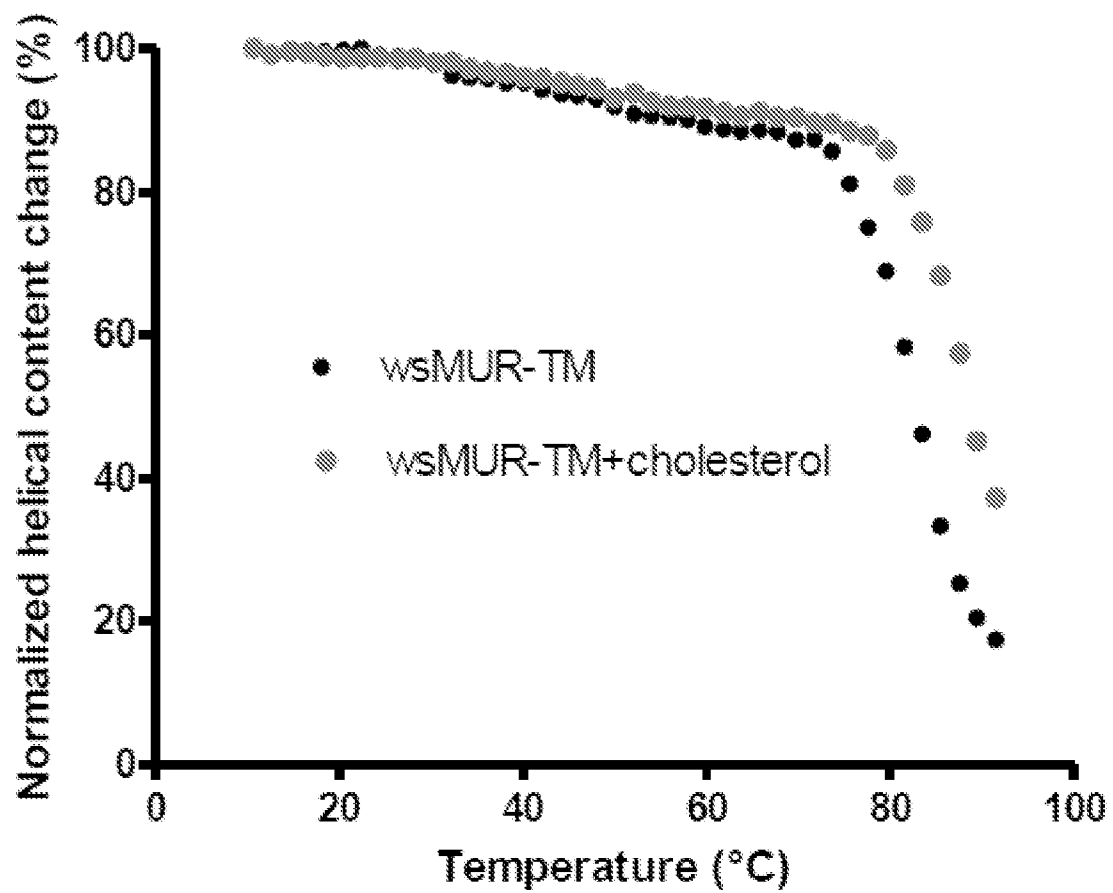


Figure 6

7/14

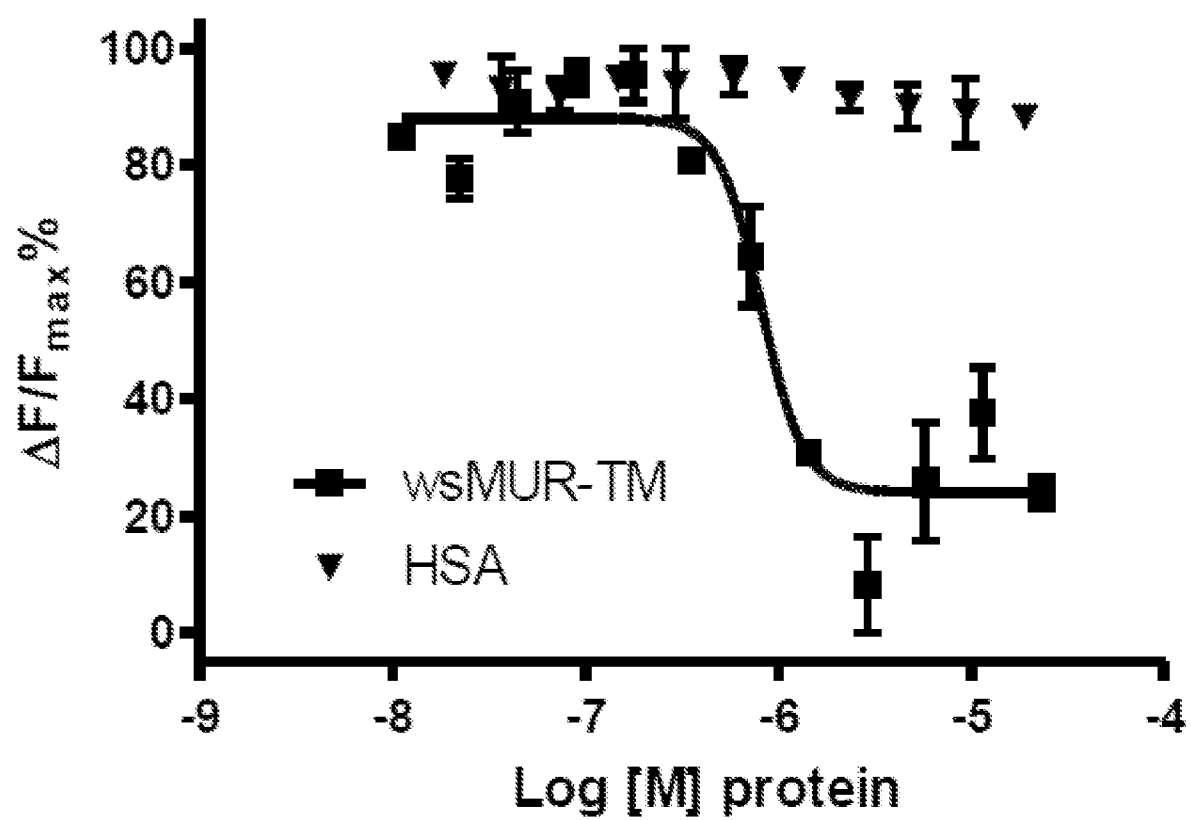


Figure 7

8/14

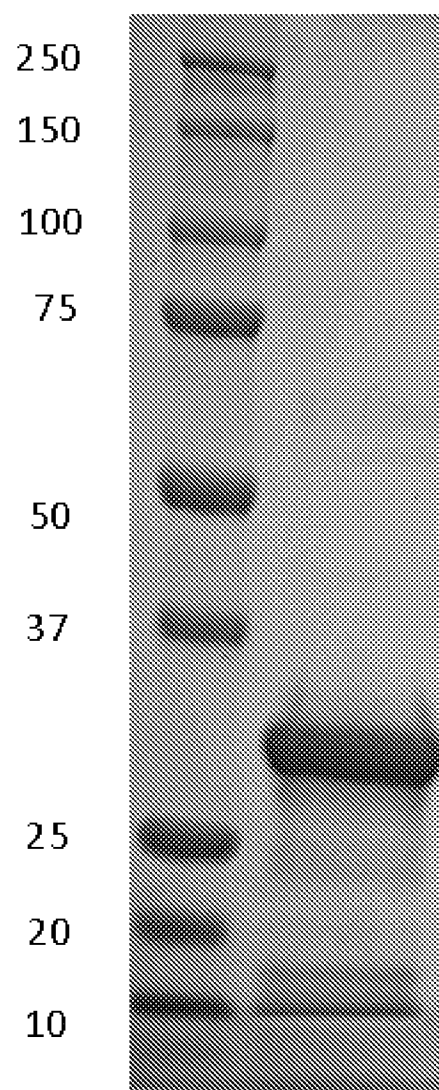


Figure 8

9/14

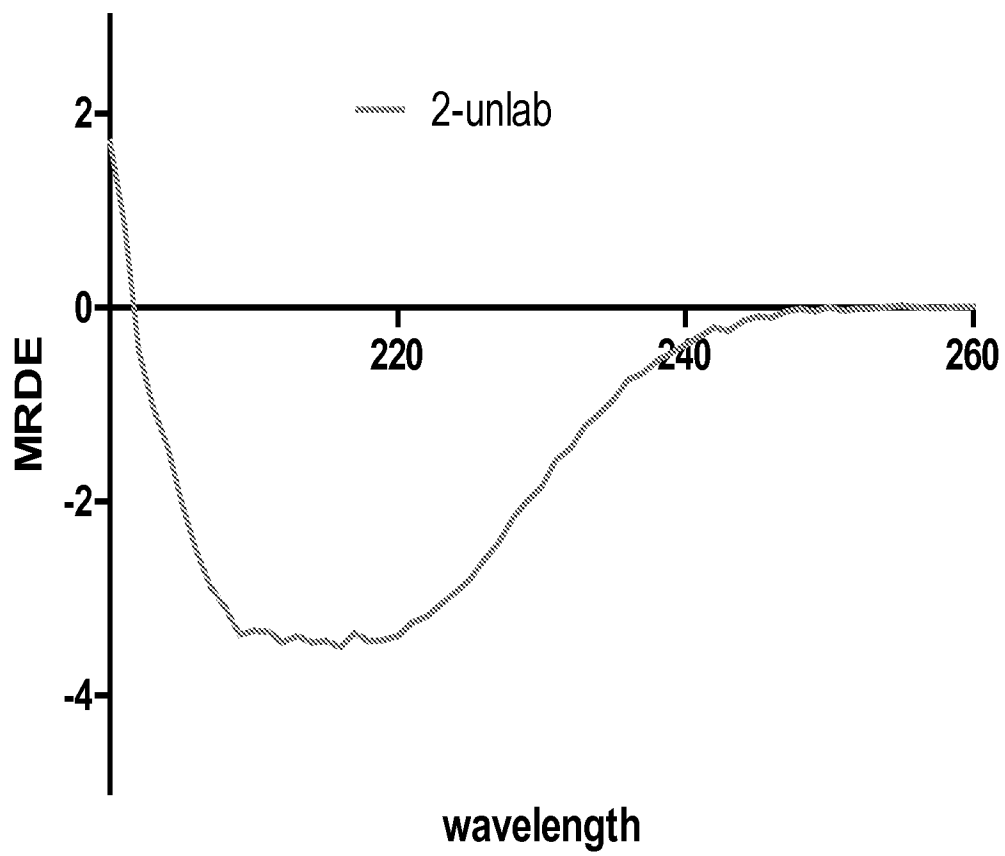


Figure 9

10/14

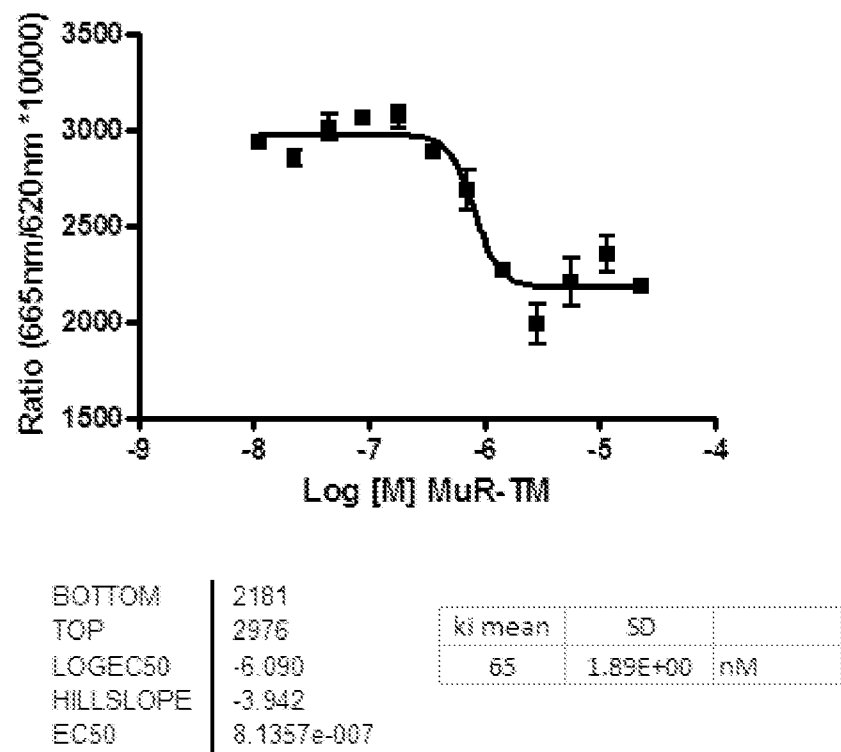


Figure 10

11/14

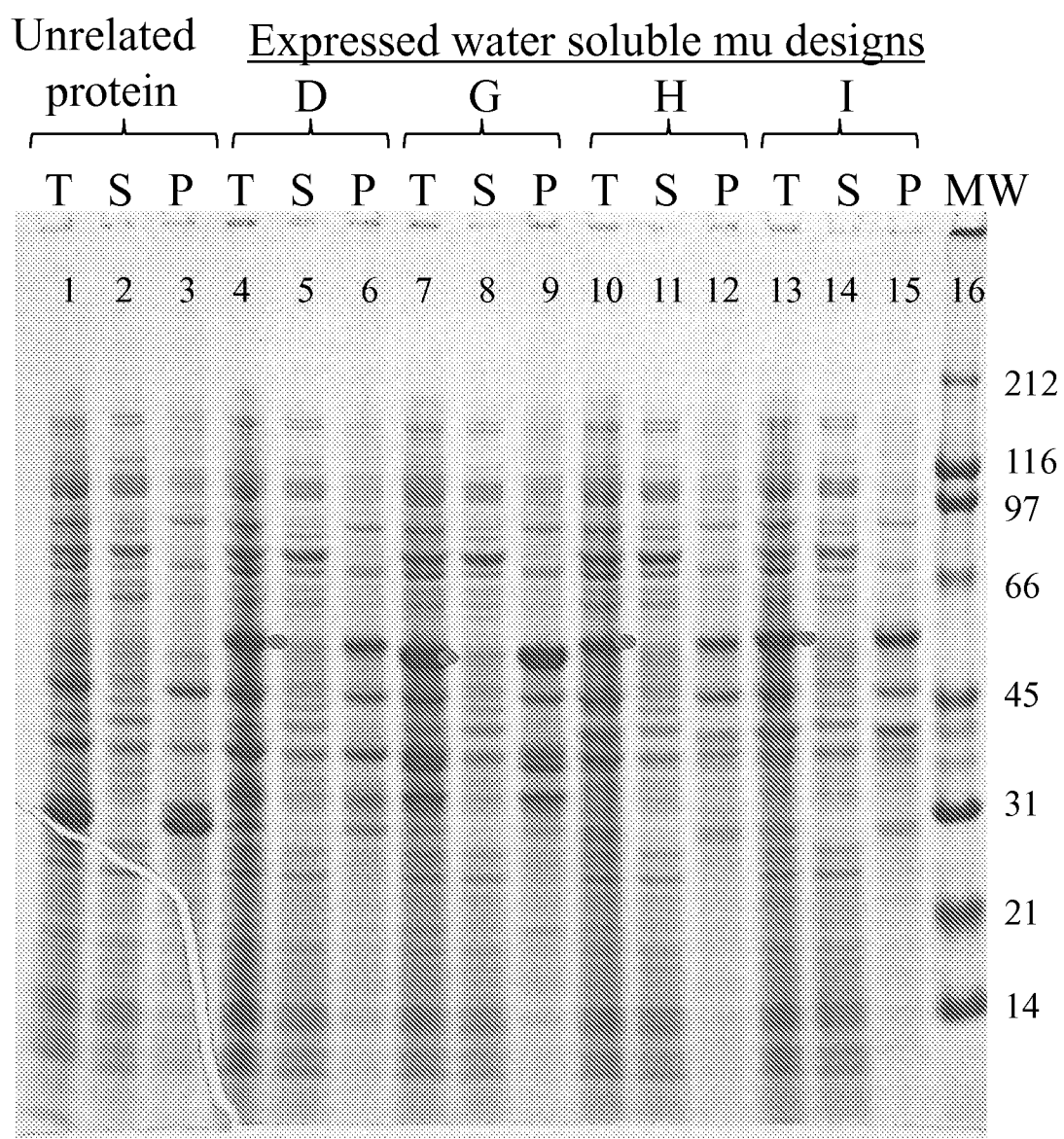


Figure 11

12/14

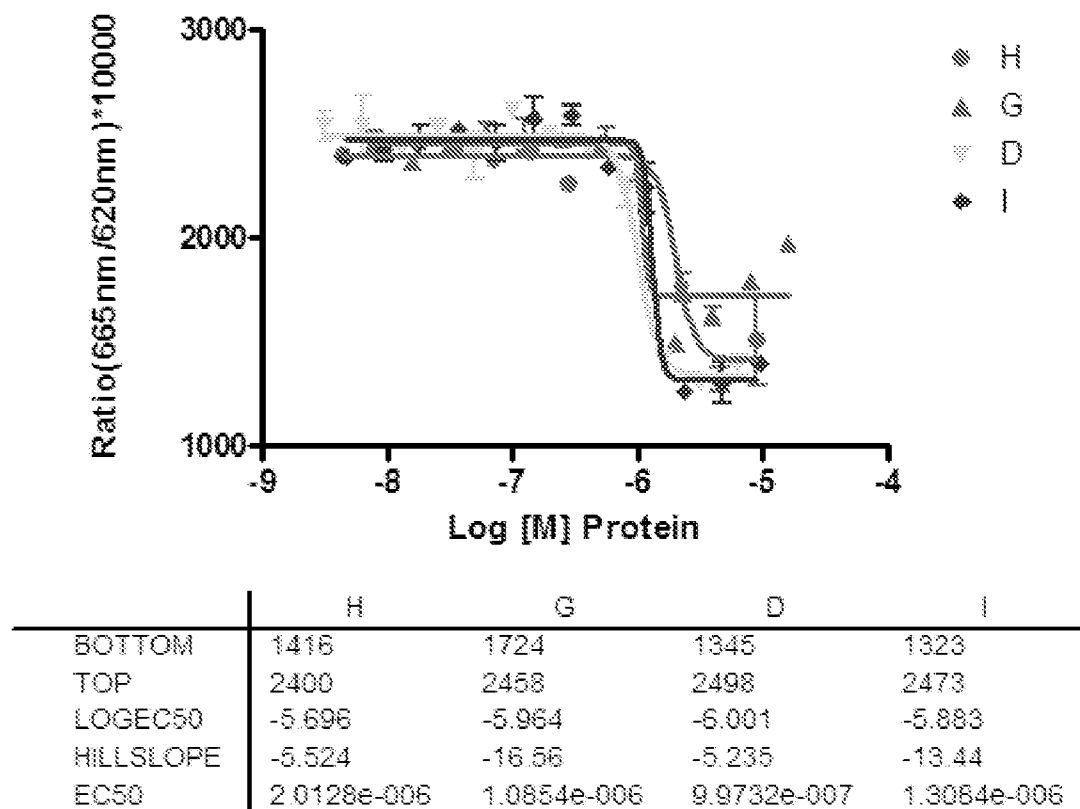


Figure 12

13/14

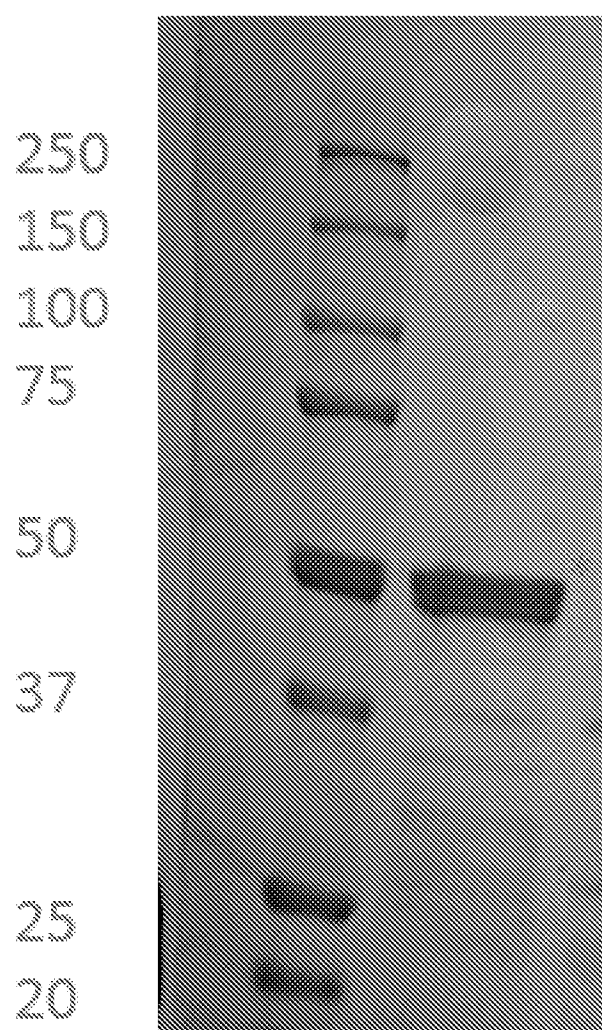


Figure 13

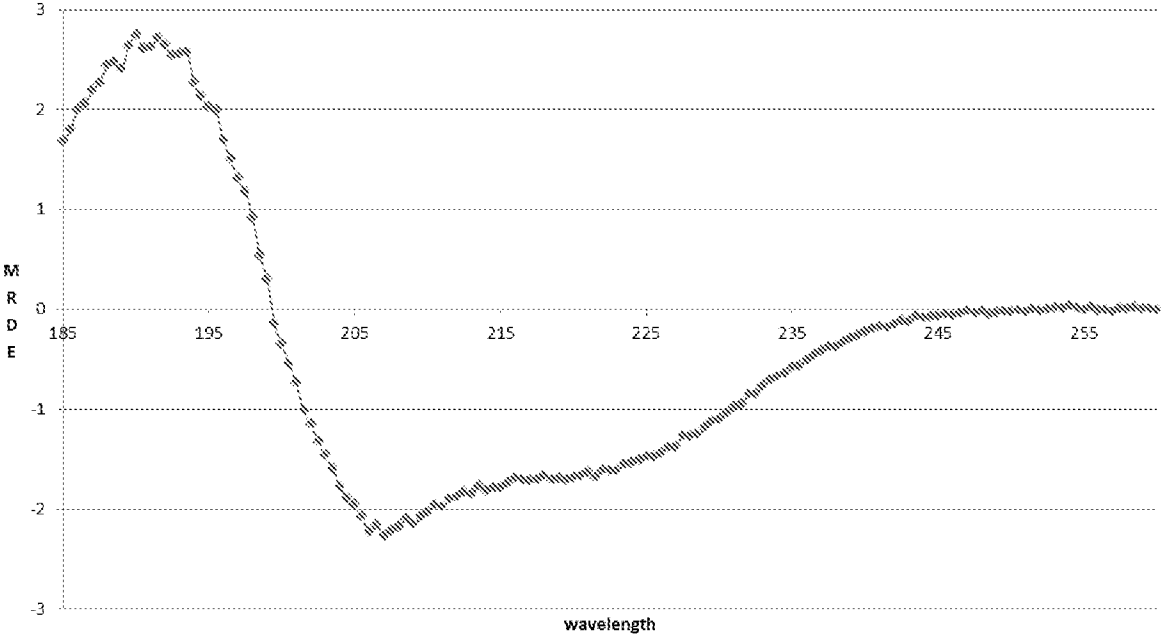


Figure 14

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 14/35524

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - A61K 38/00; G01N 33/567; G01N 33/566 (2014.01)

CPC - A61K 38/00; C07K 14/705

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC(8) - A61K 38/00; G01N 33/567; G01N 33/566; C07K 14/00; C12P 21/06 (2014.01)

USPC - 530/402; 435/7.2; 436/501; 530/350; 435/69.1; CPC - A61K 38/00; C07K 14/705; G01N 33/582; C07K 14/47

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
IPC(8) - A61K 38/00; G01N 33/567; G01N 33/566; C07K 14/00; C12P 21/06 (2014.01) - see keyword below

USPC - 530/402; 435/7.2; 436/501; 530/350; 435/69.1; CPC - A61K 38/00; C07K 14/705; G01N 33/582; C07K 14/47- see keyword below

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PubWEST(USPT,PGPB,EPAB,JPAB); PatBase; Medline, Google: G-protein-coupled receptor, GPCR, mutation, mutant, substitute, transmembrane domain, TM, helix, seven, ligand, binding, conformation, hydrophobic, hydrophobicity, charge, soluble, solubility, express, recombinant, purify, isolate, bacteria, E.coli, lysing, lyse, concentrate, collect,

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CUI et al. NMR Structure and Dynamics of a Designed Water-soluble Transmembrane Domain of Nicotinic Acetylcholine Receptor. Biochim Biophys Acta. 2012, Vol. 1818(3), p. 617-626. Abstract; pg 617, col 1; pg 618, col 1, para 2 and last para; pg 619, Fig 1; pg 620, col 1, para 2-3, and col 2, last para; pg 621, col 1, top para; pg 624, col 1, up para, and col 2, last para; and pg 625, Fig 5	1-3, 18
Y	US 2012/0270230 A1 (HENDERSON et al.) 25 October 2012 (25.10.2012), Abstract, para [0003], [0009], [0012]-[0015], [0017]-[0021], [0025], [0033], [0073], [0083], [0185], [0188], [0204], [0263], pg 25, col 1, Table (vi); and Figs: 9-11, 15, and 17-25	1-3, 18
A	SLOVIC et al. Computational design of a water-soluble analog of phospholamban. Protein Sci. 2003, Vol. 12(2), p. 337-48. Entire documentation, especially Abstract	1-3, 18
A	MA et al. NMR studies of a channel protein without membranes: structure and dynamics of water-solubilized KcsA, Proc Natl Acad Sci U S A. 2008, Vol. 105(43), p. 16537-42. Entire documentation, especially Abstract	1-3, 18
A	Crasto, Hydrophobicity profiles in G protein-coupled receptor transmembrane helical domains. J Receptor Ligand Channel Res. 2010, Vol. 2010(3), p. 123-133. Entire documentation, especially Abstract; pg 125, Fig 1; and pg 129, Fig 4 and Fig 5	1-3, 18
P,X	PEREZ-AGUILAR et al. A Computationally Designed Water-Soluble Variant of a G-Protein-Coupled Receptor: The Human Mu Opioid Receptor. PLoS One. 2013 Jun 14, Vol. 8(6):e66009. Print 2013. PDF file: pg 1-10. Entire documentation, especially Abstract; pg 2, col 1, para 1; pg 3, col 2, middle para; pg 4, Fig 1; pg5, Fig 2; and pg9, col 1, top para	1-3, 18

☐ Further documents are listed in the continuation of Box C.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 19 August 2014 (19.08.2014)	Date of mailing of the international search report 12 SEP 2014
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 14/35524

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 4-17, 19-24
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.