

Genetic Mapping of the Human C5a Receptor

IDENTIFICATION OF TRANSMEMBRANE AMINO ACIDS CRITICAL FOR RECEPTOR FUNCTION*

Received for publication, June 26, 2000

Published, JBC Papers in Press, August 21, 2000, DOI 10.1074/jbc.M005602200

Adi Geva‡, Tracey B. Lassere‡, Olivier Lichtarge§, Sonia K. Pollitt¶, Thomas J. Baranski‡||

From the ‡Departments of Medicine and Molecular Biology and Pharmacology, Washington University School of Medicine, St. Louis, Missouri 63110, the §Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas 77030, and the ¶Department of Cellular and Molecular Pharmacology, University of California, San Francisco, California 94143

Many hormones and sensory stimuli signal through a superfamily of seven transmembrane-spanning receptors to activate heterotrimeric G proteins. How the seven transmembrane segments of the receptors (a molecular architecture of bundled α -helices conserved from yeast to man) work as “on/off” switches remains unknown. Previously, we used random saturation mutagenesis coupled with a genetic selection in yeast to determine the relative importance of amino acids in four of the seven transmembrane segments of the human C5a receptor (Baranski, T. J., Herzmark, P., Lichtarge, O., Gerber, B. O., Trueheart, J., Meng, E. C., Iiri, T., Sheikh, S. P., and Bourne, H. R. (1999) *J. Biol. Chem.* 274, 15757–15765). In this study, we evaluate helices I, II, and IV, thereby furnishing a complete mutational map of the seven transmembrane helices of the human C5a receptor. Our analysis identified 19 amino acid positions resistant to non-conservative substitutions. When combined with the 25 essential residues previously identified in helices III and V–VII, they delineate two distinct components of the receptor switch: a ligand-binding surface at or near the extracellular surface of the helix bundle and a core cluster in the cytoplasmic half of the bundle. In addition, we found critical amino acids in the first and second helices that are predicted to face the lipid membrane. These residues form an extended surface that might mediate interactions with lipids and other membrane proteins or function as an oligomerization domain with other receptors.

G protein-coupled receptors, a superfamily of seven transmembrane proteins, act as molecular switches that, upon activation by extracellular stimuli, transmit signals to heterotrimeric G proteins on the cytoplasmic face of the plasma membrane. These receptors then catalyze ligand-dependent exchange of GTP for GDP on the α -subunit of the heterotrimer, causing dissociation of α -GTP from the $\beta\gamma$ -dimer; α -GTP and free $\beta\gamma$ subsequently activate effector enzymes and ion chan-

nels (1, 2). More than 1000 G protein-coupled receptors of mammals share with their counterparts in yeast and plants a conserved three-dimensional architecture, comprising seven α -helices in a transmembrane bundle (3–6). The switch mechanism is also conserved, as indicated by the abilities of mammalian receptors to activate G protein trimers in yeast (7–9). The switch clearly resides in the seven-helix bundle: swapping of extra- or intracellular loops preserves the ability of ligands to activate G proteins while transferring specificity of ligand binding or G protein activation, respectively, from one receptor to another (10–13).

Our understanding of receptor mechanisms remains limited by a lack of precise structural information and the general difficulties of characterizing integral membrane proteins. A low-resolution (5 Å) electron cryomicroscopic structure of rhodopsin (14), the retinal light receptor, reveals relative positions and tilts of seven transmembrane helices in the plane of the membrane. Based on many mutations, the rhodopsin structure, and an analysis of the primary structures of >500 rhodopsin-like G protein-coupled receptors, Baldwin *et al.* (15) constructed an α -carbon template of the helix bundle, hereafter termed the Baldwin model. In this model, transmembrane helices I–VII bundle together in clockwise order as viewed from the cytoplasm. The probable arrangement of helices and the positions of specific amino acids in the model are inferred from patterns of conserved hydrophobic and hydrophilic residues in many receptors. The Baldwin model specifies which helix corresponds to which density in the electron projection map, approximate orientations of cognate amino acids around the helical axes, and the cytoplasmic and extracellular limits of each transmembrane sequence (16).

Random saturation mutagenesis provides a comprehensive method to explore the structural characteristics of protein segments and has recently been used to elucidate the helical orientations of an inwardly rectifying potassium channel (17). This comprehensive approach determines the relative importance of side chains in each α -helix by identifying those that cannot be altered in mutant receptors selected for maintenance of function. We undertook a systematic genetic analysis of a single G protein-coupled receptor, the human C5a receptor, with the goal of identifying functionally important residues and sites of helix/helix interactions that relay the ligand signal to G protein activation. The human chemoattractant C5a receptor, a member of the rhodopsin family, couples to $G\alpha_i$ and $G\alpha_{16}$. The C5a ligand, a 74-amino acid glycoprotein produced as a cleavage product in the complement cascade, mediates leukocyte chemotaxis, vascular permeability, and smooth muscle contraction (18).

Previously, we selected functional C5a receptors after random saturation mutagenesis of four transmembrane helices

* This work was supported in part by a grant from the Edward J. Mallinckrodt, Jr. Foundation, a Young Investigator Award and American Cancer Society Grant IRG-58–010-43 (to T. J. B.), and a Career Development Award from the American Heart Association (to O. L.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

|| To whom correspondence should be addressed: Depts. of Medicine and Molecular Biology and Pharmacology, Washington University School of Medicine, P. O. Box 8103, 660 S. Euclid Ave., St. Louis, MO 63110. Tel.: 314-747-3997; Fax: 314-362-7058; E-mail: baranski@pcg.wustl.edu.

(III and V–VII) (19). In our current study, we examine the remaining three transmembrane segments (I, II, and IV) of the human C5a receptor. We anticipated that these helices would demonstrate the same characteristics as the previously studied four helices: essential amino acids would be located on surfaces of helices that face other helices, and these critical residues would cluster at either the extracellular side of the receptor or at the core of the helix bundle. The results of scanning helices I and IV conform to these expectations; however, helix II possesses essential amino acids that map to a surface that is predicted to be in contact with the surrounding lipid. In addition, the functional receptors mutated in helix I all share a hydrophobic patch that faces the lipid membrane and, with helix II, forms an extended hydrophobic surface. These helix I and II residues facing the lipid bilayer might function as an interaction domain with other membrane molecules or as an oligomerization domain with other receptors. The results of this study, combined with our previous work (19), provide the first complete mutational map of the transmembrane segments of a G protein-coupled receptor and will serve as a useful model for further structure/function studies.

EXPERIMENTAL PROCEDURES

Library Construction—To construct mutagenized C5a receptors, we engineered restriction endonuclease sites (which did not alter the amino acid sequence) into the C5a receptor gene at the approximate boundaries of the cDNA encoding helices I (*Mlu*I and *Bst*EII, 5'- and 3'-ends, respectively), II (*Nde*I, 5'-end), and IV (*Bsr*GI, 3'-end). Endogenous restriction endonuclease sites bound the cDNA encoding helices II (*Bsr*GI, 3'-end) and IV (*Sfi*I, 5'-end); however, for the helix IV library, the *Bsr*GI site at the 3'-end of the second encoded helix and a second *Sfi*I site in the 3'-untranslated region of the C5a receptor gene were removed. A *Nco*I site at the start methionine introduced aspartic acid in place of asparagine at position 2. To ensure efficient library construction and to eliminate the possibility of contamination by wild-type receptor DNA, for each helix library, we created a subcloning vector that contained non-receptor DNA substituted between the flanking restriction sites. The following oligonucleotides were used to encode the transmembrane segments (bases doped with 20% non-wild-type nucleotides are underlined): helix I, 5'-AAAATACGCGTTCCAGATATTC-TTGCTCTTGTATCTTTGCTGTGTCTTTCTTGTGGTGTCTTGG-TAATGCTCTTGGTTGGTTGGGTGACCCA-3'; helix II, 5'-AACTTACAC-ATATGGTTTCTTAATCTTGCTGTGGCCGATTTCTCTCTGTCTTG-CTCTTCATCCTTTTCACTTCTATTGTACAAT-3'; and helix IV, 5'-CTAATATTGGCCGGCTTGGCCTGGATTGCTTGTGCTGTCGCTTGG-GGTCTTGTCTTCTTCTTACAATTCCATCTTTCTGTACAGA-3'.

Oligonucleotides (Protein and Nucleic Acid Chemistry Laboratory, Washington University, St. Louis, MO) encoding helix I, II, or IV were mutually primed via extended palindromic sequences at their 3' termini, and the complementary strands were synthesized using Klenow DNA polymerase. The products were digested with the appropriate restriction enzymes and subcloned into the C5a receptor gene in the pBS-SK Bluescript vector. After determining the complexity and quality of the library (by DNA sequencing and restriction mapping), we subcloned each mutated helix library into the yeast shuttle vector p1303ADE2 (derived from p1303; Cadus Pharmaceuticals). The number of recombinants obtained equaled or exceeded the initial size of the mutated helix library.

Yeast Strains—Strain BY1142, engineered by standard yeast strategies, has the genotype *far1*Δ1442 *tbt1-1* *FUS1-HIS3* *can1* *ste14::trp1::LYS2* *ste3*Δ1156 *GPA1(41)-Gα₂* *lys2* *ura3* *leu2* *trp1* *his3* *ade2*. BY1142 expresses a yeast/human Gα chimera in which the N-terminal 41 residues of yeast GPA1 (a region predicted to interact with Gβγ based on the crystal structure of the G_i heterotrimer) (20) replaces the first 33 residues of human Gα₂. A *Sac*I restriction site introduces a point mutation that substitutes leucine for valine at position 34 of Gα₂. The chimera is designed for optimal ability to sequester yeast Gβγ and to release it upon C5a receptor activation, thereby triggering the mating response pathway. Expression of the FUS1/HIS3 reporter enzyme, stimulated by the mating response pathway, allows BY1142 cells (*his3*) to grow in histidine-deficient medium. The *far1* deletion blocks cell cycle arrest induced by pheromone. The *tbt1-1* mutation increases electroporation transformation efficiency; and *ste14*, a carboxymethylase deletion, reduces basal signaling of the pheromone response pathway. In-

clusion of an *ADE2* gene in the C5a receptor plasmid p1303 (p1303ADE2, *PGK-hC5aRADE2 REP3* 2μm-ori AmpR flori) allowed us to infer from the color of a colony whether its growth on histidine-deficient medium depends on the C5a receptor (red colonies lack *ADE2* and, by implication, the receptor plasmid). A separate plasmid, p1297 (*ADH1-mfα1-hC5aURA3 REP3* 2μm-ori AmpR flori; Cadus Pharmaceuticals), allowed autocrine expression of the C5a ligand as an prepro-α-factor/C5a ligand fusion protein; autocrine expression was necessary because C5a cannot traverse the yeast cell wall.

Yeast Transformation and Receptor Selection—We cotransformed BY1142 yeast cells by electroporation with plasmid libraries of mutant C5a receptors and the C5a ligand plasmid p1297. Colonies bearing functional receptors were selected by replica plating on histidine-deficient medium containing 1 mM 3-amino-1,2,4-triazole (AT¹; Sigma). We recovered the plasmids encoding functional C5a receptors and confirmed the phenotypes by retransforming each candidate C5a receptor plasmid and the C5a ligand plasmid (or without C5a ligand, to assess constitutive signaling) into the parental yeast strain, BY1142. Growth after 3 days at 30 °C on plates containing different concentrations of AT (0, 1, 2, 5, 10, 20, 50, and 100 mM) served to quantitate the relative intensities of signals mediated by mutant receptors. For each mutant receptor, phenotypes were scored from the average growth of at least 20 independent transformants. In each case, the wild-type C5a receptor expressed in parallel yeast cells served as a positive control. The results of 10 independent transformations of the wild-type C5a receptor were growth at 5, 5, 5, 5, 5, 5, 10, 5, 10, and 5 mM AT. The mutant receptor sequences were obtained by DNA sequencing with *Taq* polymerase using ABI Prism Big Dye Terminator Cycle sequencing (Ready Reaction kit, PerkinElmer Life Sciences).

RESULTS AND DISCUSSION

Selection of Functional Receptors—To select functional receptors from populations of mutant receptors, we used a genetic selection that couples functional receptors with the G protein-mediated response to mating pheromones of *Saccharomyces cerevisiae* (19). In yeast strain BY1142, activation of the C5a receptor induces expression of the *HIS3* gene, thereby allowing cells with functional receptors to survive and proliferate in medium lacking histidine. In this selection procedure, the relative intensities of signals mediated by mutant receptors can be quantitated by assessing growth in the presence of different concentrations of AT, a competitive inhibitor of the *HIS3* gene product (21). Expression of both C5a and the C5a receptor allowed BY1142 cells to grow in selective medium lacking histidine and containing up to 5 mM AT (Fig. 1, *first column*). Similar yeast cells harboring chimeras of yeast/human G proteins have been utilized to expression clone mammalian receptors (22) and to generate novel ligands for orphan G protein-coupled receptors (23).

Previously, we described how we used this genetic selection system to analyze transmembrane helices III and V–VII of the C5a receptor (19). Here, using the same mutagenesis and genetic selection, we have studied the remaining three transmembrane helices of the C5a receptor. Helix by helix, we subjected stretches of amino acid sequence to random saturation mutagenesis for a total of 64 positions in helices I, II, and IV. The quality of the libraries was assessed in two ways. First, DNA sequencing of 15–20 independent recombinants from each library, prior to selection, revealed oligonucleotide mutagenesis rates between 19 and 23%, resulting in amino substitution rates of 31–43% (Table I). Second, we determined the substitution rate at the third nucleotide position of degenerate codons for amino acid positions at which few substitutions were observed in functional receptors (see below). The observed substitution rate of 22% closely matches the predicted rate of 20%, demonstrating that little or no bias occurred in the randomization of the libraries.

We then selected recombinants from three separate libraries of mutant receptors, each library containing mutations in a

¹ The abbreviation used is: AT, 3-amino-1,2,4-triazole.

helix III-mutated receptors and 16 helix VI-mutated receptors demonstrated ligand-independent signaling. We did not specifically attempt to select for constitutively active receptors; nonetheless, the locations of the mutations in the 22 activated C5a receptors provide a comprehensive map of the "hot spots" for activation within the seven transmembrane helices of a single G protein-coupled receptor (Table I). These results from the genetic screen of the C5a receptor reflect the sensitivity of many other G protein-coupled receptors to activating mutations: naturally occurring mutations and artificially introduced substitutions that result in activated receptors predominantly map to helices III and VI (24). The abundance of activating mutants in helices III and VI, when compared with the rarity of such mutants in the other five helices, strengthens the possible role of helices III and VI as essential elements of the turn-on switch.

Likewise, the helices can be classified as either more or less resistant to amino acid substitutions, an indicator of the relative importance of that helix for the overall folding and stability of the molecule. In the Baldwin model, helices I, II, and IV reside on the periphery of the seven-helix bundle (see Fig. 4 and below); as such, they might be expected to tolerate more mutations than the centrally located helix III (which, in the model, makes contacts with five other helices) or helix VII (which contacts four other helices). Helices I, II, and IV displayed intermediate levels of resistance to substitution, tolerating 6.1, 5.3, and 4.2 amino acid substitutions per mutant receptor, respectively (Table I). The most resistant transmembrane segment, helix VII, tolerated only an average of 3.2 amino acid substitutions per receptor (15% mutation rate *versus* a 36% average amino substitution rate prior to selection) (Table I, as taken from Ref. 19). Conversely, helix VI tolerated more than twice the number of amino acid substitutions (8.1 or 31%) (Table I) (19). Despite its central location in the model, helix III tolerated a similar rate of mutations as helices I, II, and IV. However, the propensity for mutations in helix III and VI to activate the receptor likely increases the observed mutation rates in these two helices.

Patterns of Preserved Residues Confirm the Structural Model—We analyzed the mutant C5a receptors, amino acid position by position, in an attempt to rank their relative importance in that helix for the function of the receptor. We expected that, in functional mutant receptors, the side chains that face other helices would tolerate fewer mutations than the side chains that face the surrounding lipid since association of transmembrane helices with one another depends upon preservation of complementary shapes that maximize van der Waals interactions (25). We defined side chain character as "preserved" if we observed at that position no more than 1 amino acid substitution or if all substitutions at that position involved closely related amino acids (log odds score of 1.0 or greater in the PAM 250 scoring matrix) (26). We assume that most amino acid substitutions occur at less important parts of the receptor switch; some mutations, however, may compensate for functional defects caused by other mutations in the same helix. Single amino acid substitutions were included to allow for the rare mutation that might have occurred at a critical position, but only in the setting of other compensatory substitutions. For example, the single P90I substitution obtained in R05 occurs in the setting of substitutions in the two neighboring residues, I89L and L91I (Fig. 2). In addition, some combinations of mutations might increase the function of the C5a receptor: all of the mutant receptors obtained from selections of the three libraries functioned at levels equal to or greater than the wild-type C5a receptor (Figs. 1–3). Although these criteria may be subjective, they nonetheless accurately

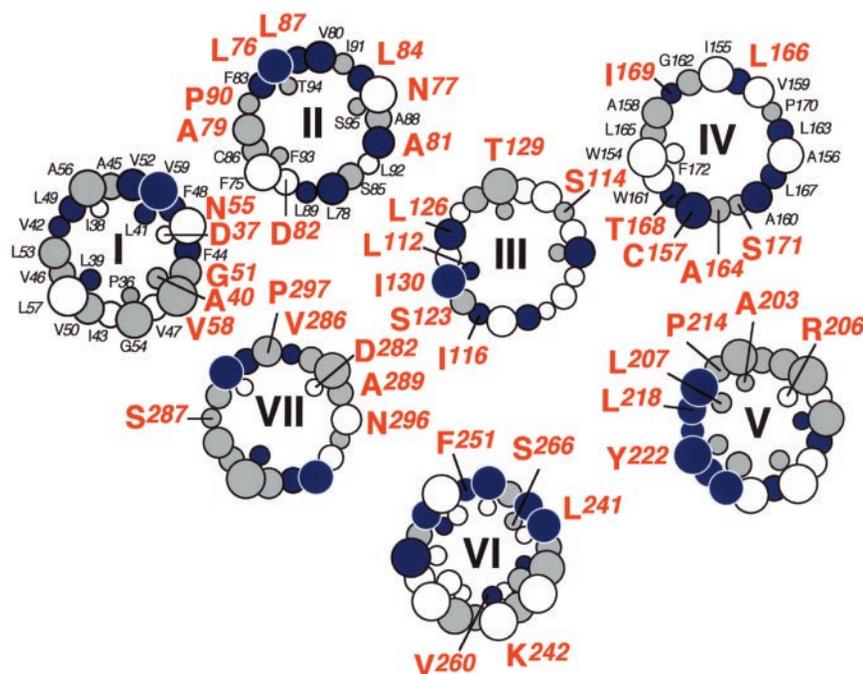
identified essential amino acids when applied to the results obtained with helices III and V–VII (19); as reported previously, single amino acid substitutions at 14 of 17 preserved positions in helices III and V–VII significantly impaired receptor function (19).

The functional mutant receptors preserved side chain character at 19 of the 64 positions mutated (Figs. 1–3). When localized on helical wheel diagrams, the helix I preserved residues (Asp³⁷, Ala⁴⁰, Gly⁵¹, Asn⁵⁵, and Val⁵⁸) (Fig. 4, *red letters*) map to a single face of the α -helix that, in the Baldwin model, points toward helices II and VII. Likewise, for helix IV, 4 of the 6 preserved residues (Cys¹⁵⁷, Ala¹⁶⁴, Thr¹⁶⁸, and Ser¹⁷¹) map to a single face of helix IV that points toward helices III and V. For helix II, the 8 preserved residues (Leu⁷⁶, Asn⁷⁷, Ala⁷⁹, Ala⁸¹, Asp⁸², Leu⁸⁴, Leu⁸⁷, and Pro⁹⁰) are distributed on multiple faces of the helix and thus provide little information to help orient the helix relative to the other transmembrane segments. Our results (Fig. 4) fit equally well with an orientation that rotates the helix 180° relative to its current position, thereby pointing the majority of the preserved residues inward. However, helix II, repositioned in that way, orients 2 highly conserved polar residues, Asp⁸² and Asn⁷⁷, into the lipid bilayer. Moreover, biochemical data obtained from studies of three different members of the G protein-coupled receptor family support the Baldwin helix II orientation as evidenced by studies of the following: (i) accessibility of substituted cysteines at each position in helix II of the D₂-dopamine receptor to a water-soluble chemical probe (27), (ii) metal ion-binding sites formed by substituted histidines in extracellular ends of helices II and III of the NK-1 receptor (28), and (iii) site-directed spin labeling studies of the intracellular end of rhodopsin helix II (29). For these reasons, we favor the helix orientation in the Baldwin model. In our genetic mapping of functional residues in the C5a receptor, the preservation of residues in helix II that face the lipid bilayer suggests that these residues serve as interaction domains with membrane molecules (see below).

Our analysis also identified positions in the mutant receptors that either retained their hydrophobic character or, conversely, allowed substitutions of polar amino acids. Based on the established role of hydrophobic interactions in stabilizing the tertiary structure of soluble proteins (30, 31), we reasoned that surfaces retaining hydrophobicity might be important for helix packing or for interactions with other hydrophobic molecules. For our classification system, hydrophobicity is considered preserved at positions where only hydrophobic amino acids with a transfer_{octanol} free energy transfer of >0.1 kcal/mol (32) are observed. Twenty-three of the 64 mutated positions preserved their hydrophobic character in functional receptors; this group includes 8 of the 19 positions at which side chain character was preserved (Figs. 1–3). For helix IV, 4 of the preserved hydrophobic residues point toward helix V, whereas the remaining three hydrophobic residues encircle the helix. For helices I and II, the majority of the preserved hydrophobic residues (Fig. 4, *blue circles*) face the lipid bilayer, providing evidence, in addition to the preserved residues in helix II, for a potential interaction domain with another membrane molecule.

We identified amino acid positions in helices I, II, and IV that allowed polar substitutions, reasoning that polar substitutions are more likely to occur on surfaces exposed to the aqueous cleft in the helix bundle or at the aqueous borders of the helices. For example, in helix I, Leu⁵⁷ might be located at the intracellular border of the lipid bilayer since this position tolerates arginine substitutions and is unlikely to be pointing inward to a water-accessible cleft. Likewise, in helix IV, polar substitutions were tolerated in three consecutive residues (Trp¹⁵⁴, Ile¹⁵⁵, and Ala¹⁵⁶) (Fig. 3). An aqueous/lipid border on the cytoplasmic side

FIG. 4. **Complete functional map of the seven transmembrane helices of the human C5a receptor.** Helical wheel representations are based on the Baldwin model (15). Helices are presented as viewed from the cytoplasm and in positions that correspond to their relative positions at the middle of the bilayer (15). *Larger circles* indicate residues closer to the cytoplasm. *Red letters and numbers* indicate residues and positions at which side chain character is preserved in functional mutant receptors (see "Results and Discussion"). *Blue circles* indicate positions at which hydrophobicity is preserved; *white circles* indicate positions that tolerate polar substitutions (as defined in the legend to Fig. 1); and *gray circles* indicate other positions. Results of functional scanning of helices I, II, and IV are combined with previous results for helices III and V–VII (19).



positions 64 and 153 (C5a receptors numbers, by alignment). The differences between these two results might reflect a detergent effect on solubilized rhodopsin, the length of the nitroxide spin label, limits to the resolution of either approach, or a fundamental difference between the structures of the two receptors. On the extracellular side of the helices, polar substitutions that point outward from the receptor core were obtained only in helix I (Ile³⁸), making it difficult to make any assignments regarding the aqueous/lipid borders.

Preserved Residues Identify Working Parts of the Switch—Similar to our previous analysis of helices III and V–VII (19), we mapped the preserved residues from helices I, II, and IV onto the α -carbon template for the rhodopsin family of G protein-coupled receptors (15). Positions preserved in helices I, II, and IV of the mutant C5a receptors map to and extend the two distinct clusters formed by the preserved residues from helices III and V–VII, as indicated by the *red* and *yellow balls* in Fig. 5. To identify the positions in helices I, II, and IV that are conserved throughout the evolution of G protein-coupled receptors, we applied the evolutionary trace analysis method (34) to 62 G protein-coupled receptors closely related to the human C5a receptor. Figs. 1–3 indicate the top ranked positions (labeled *Evol. Conserved*). As expected, the evolutionary trace approach identified many of the same residues found to be conserved by Baldwin's analysis of 199 unique G protein-coupled receptor sequences (16). *Red balls* in Fig. 5 indicate positions at which residues are both preserved in mutant receptors and conserved evolutionarily; *yellow balls* indicate positions that are important for the function of mutant C5a receptors, but that are not highly conserved in other G protein-coupled receptors throughout evolution. (*Blue balls* indicate positions that are deemed evolutionarily important, but that are not preserved in our genetic screen (see below).) Most of the *yellow* positions are located at or near the receptor's extracellular face, whereas *red* positions cluster tightly in the cytoplasmic half of the transmembrane helices (Fig. 5).

In the two-site binding model proposed for C5a binding (35–37), the carboxyl-terminal tail of C5a probably inserts into the interhelical crevice of the receptor, whereas other portions of the C5a ligand interact with the amino terminus of the C5a receptor. We propose that the 22 preserved residues, 9 residues

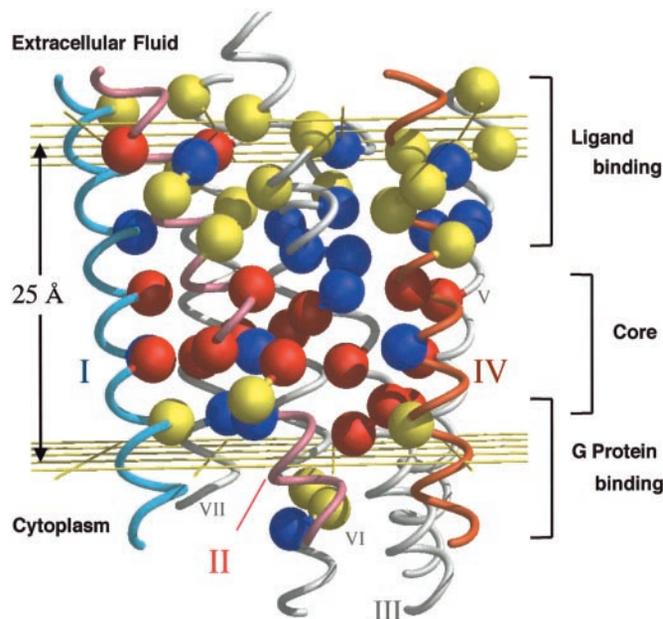


FIG. 5. **Comparison of the locations of preserved amino acids in the functional mutant C5a receptors versus highly conserved residues in evolution.** Colored tubes represent the α -carbon backbones of the helices in the Baldwin model. Portions of helices I (*cyan tubes*), II (*pink tubes*), and IV (*orange tubes*) were subjected to random saturation mutagenesis; *gray tubes* indicate residues or helices not tested. Results of functional scanning of helices I, II, and IV are combined with previous results for helices III and V–VII (19). *Yellow* and *red balls* indicate α -carbon positions at which side chain character is preserved in functional C5a receptors; *blue* and *red balls* indicate positions that are conserved as assessed by the evolutionary trace method (34). *Red balls* indicate positions that are both preserved in the genetic screen and also conserved in evolution of G protein-coupled receptors. The *yellow grids* (lines separated by 10 Å) indicate proposed borders of the lipid core of the membrane (15).

from helices I, II, and IV (Asp³⁷, Ala⁴⁰, Leu⁸⁴, Leu⁸⁷, Pro⁹⁰, Leu¹⁶⁶, Thr¹⁶⁸, Ile¹⁶⁹, and Ser¹⁷¹) and 13 residues from helices III and V–VII (19), cluster near the extracellular fluid and, as identified by their resistance to substitutions, define an interhelical receptor surface that interacts with the carboxyl-termi-

nal tail of C5a (Fig. 5, labeled *Ligand binding*). Not all these residues necessarily contact the C5a ligand; some residues, *e.g.* Pro⁹⁰, may dictate the overall shape of the binding pocket or help mediate a conformational change that occurs upon ligand binding. A P90A mutation in the C5a receptor caused nearly a 10-fold decrease in ligand binding affinity, but also resulted in receptors that were misfolded as assessed by aberrant localization when expressed in mammalian cells (38).

The cluster of preserved residues, proposed as the ligand-binding crevice in the C5a receptor, likely represents a common binding site in other G protein-coupled receptors. For the preserved Ser¹⁷¹ position in helix IV of the C5a receptor, mutations at cognate positions of the vasopressin, NK-1, and thyrotropin receptors (Gln¹⁸⁵, Gln¹⁶⁵, and Pro⁵⁵⁶, respectively) diminish ligand binding without altering processing of the receptors (39–41). In helix II, a Q104A mutation in the vasopressin receptor (preserved Leu⁸⁷ in the C5a receptor) also decreases ligand binding affinity (39). As noted previously (19), small amines, the retinal chromophore, and the 74-amino acid C5a ligand interact with amino acids at many of the same positions in helices III and V–VII (albeit with different side chains) to activate G protein-coupled receptors. This suggests that G protein-coupled receptors that bind very different agonists share the first element of a common activation mechanism: a conserved location for the ligand-binding domain.

In the Baldwin model, the α -carbons of all 17 preserved positions in helices I–VII cluster at the core of the receptor (Fig. 5, *red balls*, labeled *Core*) and are located within 10 Å of the α -carbon of at least one other highly conserved residue (“conserved” is defined as identical in at least 60% of 199 G protein-coupled receptors assessed by Baldwin *et al.* (15)). This cluster, highly preserved both in evolution and in our genetic selection for functional C5a receptors, is positioned to form an interacting network that is crucial for the structure and function of G protein-coupled receptors. The precise roles of the residues in the conserved core cannot be determined without careful biochemical characterization of mutations at each position. However, this proves to be a difficult task since many of these critical residues also govern the overall structure of the receptor; mutations at these positions often result in misfolded proteins (42). Based on the overlap between the results of our genetic analysis and our evolutionary analysis, we propose that these jointly identified positions form a structural core that allows conformational changes induced by ligand binding to be transmitted to the G protein interaction surface.

Some positions (Fig. 5, *blue balls*) that are highly important in many G protein-coupled receptors throughout evolution are nonetheless sites of frequent substitution in functional mutant C5a receptors. They may mediate a specific function subject to a selective pressure in evolution that was absent in our yeast screen. One possibility is that these *blue* residues, located primarily in the extracellular half of the helix bundle, participate in interhelical interactions responsible for maintaining the receptor switch in the off state in the absence of ligand. We have not tested whether single mutations at the *blue* positions systematically produce constitutive activation of the receptor.

On the cytoplasmic side of helices I, II, and IV, the genetic screen identified three positions that did not tolerate substitutions, yet were not highly conserved in other G protein-coupled receptors (Val⁵⁸, Leu⁷⁶, and Cys¹⁵⁷) (Fig. 5, *yellow balls*). When combined with Leu²⁴¹ and Lys²⁴² from the scan of helix VI, the positions of these 5 residues begin to identify regions of the receptor that might mediate specific interactions with G proteins. Supporting this role, a C157S mutation in helix IV of the C5a receptor completely abrogated G protein activation while leaving binding activity intact (38). Likewise, rhodopsin bear-

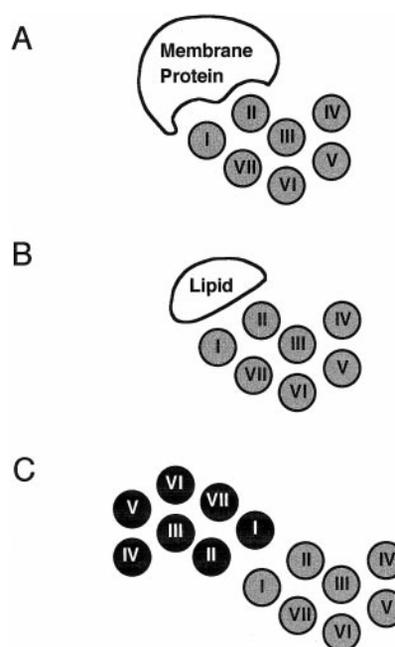


FIG. 6. Models for potential interactions of the helix I/II surface that faces the lipid bilayer. A, the receptor could interact with other membrane proteins such as molecular chaperones or scaffolds. B, the helix 1,2-surface might mediate specific contacts with membrane sterols, phospholipids, or fatty acid modifications on the heterotrimeric G proteins. C, G protein-coupled receptors might dimerize via a 1,2-contact interface.

ing a T58R mutation (aligns with Val⁵⁸ in the C5a receptor) forms *meta*-rhodopsin II properly upon illumination, but cannot activate transducin (43). The regions of the receptor that contact G protein most likely extend farther into the cytoplasm than was scanned in our analysis: ample evidence implicates the intracellular loops and carboxyl-terminal tail of the receptors in binding G proteins (5, 6).

Helix I/II Interaction Domain—Our results suggest that the lipid-oriented faces of helices I and II might play an important role in G protein-coupled receptor structure and function. From the results of our functional mapping of the seven transmembrane helices of the C5a receptor, helices I and II display patches of preserved residues (Leu⁴¹, Val⁴², Phe⁴⁸, Leu⁴⁹, Val⁵², and Val⁵⁹, and Leu⁷⁶, Ala⁷⁹, Val⁸⁰, Phe⁸³, Leu⁸⁴, Leu⁸⁷, and Pro⁹⁰, respectively) that face the lipid membrane (Fig. 4). Other helices might also specifically interact with membrane proteins; for example, helix IV displays 3 preserved hydrophobic residues (Leu¹⁶³, Leu¹⁶⁶, and Ile¹⁶⁹) to the lipid membrane. However, the helix I/II surface represents the largest group of preserved residues from the genetic scanning of the C5a receptor that face the lipid bilayer.

Fig. 6 displays three simple models for possible interactions of the lipid-facing residues in helices I and II with accessory membrane proteins, other lipids, or other G protein-coupled receptors. First, as an example of a membrane protein/receptor interaction (Fig. 6A), a recent study identified a new class of proteins called RAMP (receptor activity-modifying protein) that interact with receptors to alter their intracellular processing and ligand binding specificity (44). When coexpressed with RAMP1, the calcitonin receptor-like receptor binds the calcitonin gene-related peptide; however, if coexpressed with RAMP2 or RAMP3, the same receptor no longer binds the calcitonin gene-related peptide, but is activated by an entirely different ligand, adrenomedullin (44). Other accessory proteins for G protein-coupled receptors include ODR-4, a type II membrane protein required for the proper membrane localization of odor-

ant receptors in *Caenorhabditis elegans* (45), and NinaA, a cyclophilin that acts as a molecular chaperone for opsins (46). No accessory protein has been identified for the pheromone receptors in yeast.

Second, helices I and II of the C5a receptor might form a specific and critical protein/lipid interaction (Fig. 6B). Lipids play ill defined but important roles in the function of G protein-coupled receptors. Removal of membrane sterol from cells dramatically decreases the binding affinity and activity of the oxytocin receptor (47). Potential mechanisms for this effect include an allosteric interaction between the cholesterol and the receptor helix bundle; a more general effect on the biophysical properties of the membrane or the sterol, by virtue of their ability to form discrete microdomains, might localize the receptors with other signaling molecules on the plasma membrane. Caveolae, cholesterol-enriched organelles proposed as signaling centers on the plasma membrane (48), might serve as sites for sterol/receptor interactions; however, these structures have not been demonstrated in yeast. Alternatively, lipid modifications of the heterotrimeric G proteins (farnesylation of β -subunits and palmitoylation and myristoylation of α -subunits) may increase the specificity and efficiency of G protein activation by directly interacting with the transmembrane segments of the receptors.

Third, numerous recent studies demonstrate that G protein-coupled receptors might form oligomeric structures (49–52). Our genetic data support a role for helices I and II in the putative dimerization domain for G protein-coupled receptors (Fig. 6C). In favor of a protein/protein interaction, three of the mutant receptors contain an L41C mutation in helix I (a rare mutation since it requires 2 nucleotide substitutions in the leucine codon) at a position (Leu⁴¹) that otherwise does not tolerate other mutations (Fig. 1). Although we have no direct evidence that the substituted cysteine in the mutant receptors is involved in a disulfide bond, the crystal structure of rhodopsin demonstrates receptor monomers contacting each other at the first transmembrane segments where the cysteine substitutions occur in the mutant C5a receptors (53). Also, recent studies of bradykinin receptor B2 demonstrate receptor dimers and implicate the junction of the N terminus and first transmembrane helix as part of the dimerization domain (54).

In summary, we have performed a comprehensive genetic analysis of transmembrane helices I, II, and IV of the C5a receptor, which, combined with our previous work (19), provides a complete mutational map of the transmembrane helices of a single G protein-coupled receptor. The results identify 44 residues located primarily in two clusters that are essential for C5a receptor function. One cluster includes residues at or near the extracellular face of the receptor that probably constitute a binding pocket for interaction with the C5a ligand; strikingly, this cluster shares a very similar "footprint" with residues in distantly related receptors that interact with retinal, peptide ligands, and biogenic amines. The second cluster, at the core of the helix bundle, consists of residues that are conserved in most G protein-coupled receptors. Both clusters thus argue strongly for an activation mechanism that is conserved in all or most G protein-coupled receptors. We propose that proper orientation of ligands (C5a, biogenic amines, and probably others) in the binding pocket induces a conformational change that is transmitted through the conserved core of the helix bundle to G proteins. In addition, we found critical amino acids in the first and second helices of the receptor that face the lipid membrane and that might mediate interactions with lipids or other membrane proteins or function as oligomerization domains with other receptors. At present, we can only infer that the locations of these critical residues indicate that they perform specific

functions. To define the roles of these residues more precisely will require further biochemical and structural information.

Acknowledgments—We thank Nathan Durnell for technical assistance; Josh Trueheart (formerly of Cadus Pharmaceuticals), Henry Bourne, Basil Gerber, and Elaine Meng for helpful discussions; and Joyce Baldwin for the α -carbon template for the rhodopsin family of G protein-coupled receptors.

REFERENCES

- Clapham, D. E., and Neer, E. J. (1997) *Annu. Rev. Pharmacol. Toxicol.* **37**, 167–203
- Hamm, H. E. (1998) *J. Biol. Chem.* **273**, 669–672
- Gether, U., and Kobilka, B. K. (1998) *J. Biol. Chem.* **273**, 17979–17982
- Bockaert, J., and Pin, J. P. (1999) *EMBO J.* **18**, 1723–1729
- Wess, J. (1997) *FASEB J.* **11**, 346–354
- Bourne, H. R. (1997) *Curr. Opin. Cell Biol.* **9**, 134–142
- Erickson, J. R., Wu, J. J., Goddard, J. G., Tigyi, G., Kawanishi, K., Tomei, L. D., and Kiefer, M. C. (1998) *J. Biol. Chem.* **273**, 1506–1510
- Price, L. A., Kajkowski, E. M., Haddock, J. R., Ozenberger, B. A., and Pausch, M. H. (1995) *Mol. Cell Biol.* **15**, 6188–6195
- Price, L. A., Strnad, J., Pausch, M. H., and Haddock, J. R. (1996) *Mol. Pharmacol.* **50**, 829–837
- Cotecchia, S., Ostrowski, J., Kjelsberg, M. A., Caron, M. G., and Lefkowitz, R. J. (1992) *J. Biol. Chem.* **267**, 1633–1639
- Fong, T. M., Yu, H., and Strader, C. D. (1992) *J. Biol. Chem.* **267**, 25668–25671
- Gether, U., Johansen, T. E., Snider, R. M., Lowe, J. D., Nakanishi, S., and Schwartz, T. W. (1993) *Nature* **362**, 345–348
- Wong, S. K.-F., Parker, E. M., and Ross, E. M. (1990) *J. Biol. Chem.* **265**, 6219–6224
- Krebs, A., Villa, C., Edwards, P. C., and Schertler, G. F. (1998) *J. Mol. Biol.* **282**, 991–1003
- Baldwin, J. M., Schertler, G. F., and Unger, V. M. (1997) *J. Mol. Biol.* **272**, 144–164
- Baldwin, J. M. (1994) *Curr. Opin. Cell Biol.* **6**, 180–190
- Minor, D. L., Jr., Masseling, S. J., Jan, Y. N., and Jan, L. Y. (1999) *Cell* **96**, 879–891
- Gerard, C., and Gerard, N. P. (1994) *Annu. Rev. Immunol.* **12**, 775–808
- Baranski, T. J., Herzmark, P., Lichtarge, O., Gerber, B. O., Trueheart, J., Meng, E. C., Iiri, T., Sheikh, S. P., and Bourne, H. R. (1999) *J. Biol. Chem.* **274**, 15757–15765
- Wall, M. A., Coleman, D. E., Lee, E., Iniguez-Lluhi, J. A., Posner, B. A., Gilman, A. G., and Sprang, S. R. (1995) *Cell* **83**, 1047–1058
- Stevenson, B. J., Rhodes, N., Errede, B., and Sprague, G. F., Jr. (1992) *Genes Dev.* **6**, 1293–1304
- Chambers, J. K., Macdonald, L. E., Sarau, H. M., Ames, R. S., Freeman, K., Foley, J. J., Zhu, Y., McLaughlin, M. M., Murdock, P., McMillan, L., Trill, J., Swift, A., Aiyar, N., Taylor, P., Vawter, L., Naheed, S., Szekeres, P., Hervieu, G., Scott, C., Watson, J. M., Murphy, A. J., Duzic, E., Klein, C., Bergsma, D. J., Wilson, S., and Livi, G. P. (2000) *J. Biol. Chem.* **275**, 10767–10771
- Klein, C., Paul, J. I., Sauve, K., Schmidt, M. M., Arcangeli, L., Ransom, J., Trueheart, J., Manfredi, J. P., Broach, J. R., and Murphy, A. J. (1998) *Nat. Biotechnol.* **16**, 1334–1337
- Scheer, A., and Cotecchia, S. (1997) *J. Recept. Signal Transduct. Res.* **17**, 57–73
- MacKenzie, K. R., Prestegard, J. H., and Engelman, D. M. (1997) *Science* **276**, 131–133
- Dayhoff, J. U., and Schwartz, R. M. (1979) in *Atlas of Protein Sequence and Structure* (Dayhoff, J. U., ed) Vol. 5, pp. 353, National Biomedical Research Foundation, Washington, D. C.
- Javitch, J. A., Ballesteros, J. A., Chen, J., Chiappa, V., and Simpson, M. M. (1999) *Biochemistry* **38**, 7961–7968
- Elling, C. E., and Schwartz, T. W. (1996) *EMBO J.* **15**, 6213–6219
- Altenbach, C., Klein-Seetharaman, J., Hwa, J., Khorana, H. G., and Hubbell, W. L. (1999) *Biochemistry* **38**, 7945–7949
- Chothia, C. (1984) *Annu. Rev. Biochem.* **53**, 537–572
- Dill, K. A. (1990) *Biochemistry* **29**, 7133–7155
- Fauchere, J.-L., and Pliska, V. (1983) *Eur. J. Med. Chem.* **18**, 369–375
- Farahbakhsh, Z. T., Ridge, K. D., Khorana, H. G., and Hubbell, W. L. (1995) *Biochemistry* **34**, 8812–8819
- Lichtarge, O., Bourne, H. R., and Cohen, F. E. (1996) *J. Mol. Biol.* **257**, 342–358
- Gerard, C., Showell, H. J., Hoeprich, P. D. Jr., Hugli, T. E., and Stimler, N. P. (1985) *J. Biol. Chem.* **260**, 2613–2616
- Mery, L., and Boulay, F. (1994) *J. Biol. Chem.* **269**, 3457–3463
- Siciliano, S. J., Rollins, T. E., DeMartino, J., Konteatis, Z., Malkowitz, L., Van, R. G., Bondy, S., Rosen, H., and Springer, M. S. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 1214–1218
- Kolakowski, L. F., Jr., Lu, B., Gerard, C., and Gerard, N. P. (1995) *J. Biol. Chem.* **270**, 18077–18082
- Mouillac, B., Chini, B., Balestre, M. N., Elands, J., Trumpp-Kallmeyer, S., Hoflack, J., Hibert, M., Jard, S., and Barberis, C. (1995) *J. Biol. Chem.* **270**, 25771–25777
- Fong, T. M., Yu, H., Cascieri, M. A., Underwood, D., Swain, C. J., and Strader, C. D. (1994) *J. Biol. Chem.* **269**, 14957–14961
- Gu, W. X., Du, G. G., Kopp, P., Rentoumis, A., Albanese, C., Kohn, L. D., Madison, L. D., and Jameson, J. L. (1995) *Endocrinology* **136**, 3146–3153
- van Rhee, A. M., and Jacobson, K. A. (1996) *Drug Dev. Res.* **37**, 1
- Min, K. C., Zvyaga, T. A., Cypess, A. M., and Sakmar, T. P. (1993) *J. Biol. Chem.* **268**, 9400–9404

44. McLatchie, L. M., Fraser, N. J., Main, M. J., Wise, A., Brown, J., Thompson, N., Solari, R., Lee, M. G., and Foord, S. M. (1998) *Nature* **393**, 333–339
45. Dwyer, N. D., Troemel, E. R., Sengupta, P., and Bargmann, C. I. (1998) *Cell* **93**, 455–466
46. Baker, E. K., Colley, N. J., and Zuker, C. S. (1994) *EMBO J.* **13**, 4886–4895
47. Gimpl, G., Burger, K., and Fahrenholz, F. (1997) *Biochemistry* **36**, 10959–10974
48. Okamoto, T., Schlegel, A., Scherer, P. E., and Lisanti, M. P. (1998) *J. Biol. Chem.* **273**, 5419–5422
49. Rodriguez-Frade, J. M., Vila-Coro, A. J., de Ana, A. M., Albar, J. P., Martinez, A. C., and Mellado, M. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 3628–3633
50. Overton, M. C., and Blumer, K. J. (2000) *Curr. Biol.* **10**, 341–344
51. Jordan, B. A., and Devi, L. A. (1999) *Nature* **399**, 697–700
52. Angers, S., Salahpour, A., Joly, E., Hilal-Dandan, D., Dennis, M., and Bouvier, M. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97**, 3684–3689
53. Unger, V. M., Hargrave, P. A., Baldwin, J. M., and Schertler, G. F. (1997) *Nature* **389**, 203–206
54. AbdAlla, S., Zaki, E., Lothar, H., and Qwitterer, U. (1999) *J. Biol. Chem.* **274**, 26079–26084