ABSTRACT

Several recent studies employing differential epitope tagging, selective immunoisolation of receptor complexes, and fluorescence or bioluminescence resonance energy transfer techniques have provided direct evidence for heterodimerization between closely related members of the G-protein coupled receptor (GPCR) family. Since heterodimerization appears to play a role in modulating agonist affinity, efficacy, and/or trafficking properties, molecular models of interacting GPCRs would be required to understand receptor function. To advance knowledge in this field, we present here a computational approach based on **correlated mutation analysis**. The new subtractive correlated mutation (SCM) method is designed to predict pairs of residues preserved in evolution at the contact interface between transmembrane (TM) regions of GPCR heterodimers. The interpretation of results with the use of molecular models of GPCRs based on the rhodopsin crystal structure reveals likely intermolecular contacts amongst the 49 alternatives that are possible for all 7 TM domains. The algorithm filters out likely intramolecular pairs of interacting residues. Among the critical aspects of the SCM approach that will be discussed in the presentation are the number of sequences considered in the multiple sequence alignments, and the criteria to be used for eliminating the significant number of false positives.

Using Correlated Mutation Analysis to Predict the Heterodimerization Interface of GPCRs.

Marta Filizola,* Osvaldo Olmea, and Harel Weinstein

Department of Physiology and Biophysics, Mount Sinai School of Medicine, New York, NY
** e-mail address: marta.filizola@physbio.mssm.edu*

INTRODUCTION

Recent biophysical methods based on luminescence and fluorescence energy transfer are supporting the idea that GPCRs exist as dimers or even higher-order oligomers [see (1) and (2) for recent reviews]. In particular, these complexes can either involve identical proteins (homodimers) or be the result of the association of non-identical proteins (heterodimers).

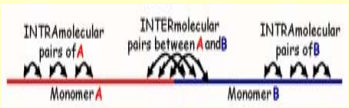
Current reports on heterodimerization of closely and distantly related members of the GPCR family suggest potential roles for this phenomenon in modulating agonist affinity, efficacy, and/or trafficking properties. Heterodimerization seems to be selective, so that GPCRs will heterodimerize with one type of receptors and not another. Heterodimerization between closely related members of the GPCR family has been observed for $GABA_BR1$-$GABA_BR2$ [3-5], M_2-M_4 muscarinic [6, 7], κ-δ opioid [8], μ-δ opioid [9, 10], 5HT1B-5HT1D serotonin [11], SSTR1-SSTR5 somatostatin [12], and CCR2-CCR5 chemokine [13] receptors. Recent examples of heterodimerization between distantly related members of the GPCR family are adenosine A1-D1 dopamine [14], angiotensin AT1-bradykinin B2 [15], somatostatin SSTR5-D2 dopamine [16], β2-adrenergic-δ-opioid [17], β2-adrenergic-κ-opioid [17], and metabotropic glutamate 1alpha-adenosine A1 [18] receptors. Finally, examples of GPCR subtypes that cannot heterodimerize are μ opioid with κ opioid receptors [8], somatostatin SSTR5 with SSTR4 [12], and chemokine CCR2 with CXCR4 [13] receptors.

Since the effect that GPCR heterodimerization *in vivo* has in the modulation of receptor function is not known yet, molecular models of interacting GPCRs should be used to advance knowledge in this field. Although a controversial mode of receptor interaction involving the swapping of domains has been proposed for homodimers and symmetric chimeric heterodimers [19], converging evidence suggests that receptor heterodimers are likely to contain only "contact dimers". There are 49 different configurations in which two tightly packed bundles of 7 transmembrane domains (TM) can be positioned next to each other. In order to reduce this number of possible configurations to a limited number of the most likely interfaces for specific GPCR heterodimerization, we have designed a computational approach based on ***correlated mutation analysis (CMA)*** and the structural information contained in three-dimensional (3D) molecular models of GPCRs built using the rhodopsin crystal structure [20] as a template.

SUBTRACTIVE CORRELATED MUTATION METHOD (SCM)

It has been recently demonstrated that oligomer interfaces are significantly conserved with respect to the protein surface [21], and correlated mutations have been shown to contain information about inter-domain contacts [22]. The correlation has been interpreted as a result of the tendency of positions in proteins to mutate coordinately: sequence changes occurring during evolution at the interface of dimerization of a given monomer A must be compensated by changes in the interacting monomer B in order to preserve the interaction interface.

Based on these observations and a computational method for identifying the correlated mutations [23], we have developed a new subtractive correlated mutation (SCM) method aiming at the identification of the most likely heterodimerization interfaces between interacting proteins that are structurally similar to each other (such as GPCRs in subfamilies). Given two structurally similar interacting proteins A and B with different amino acid sequences, a list of both intra- and intermolecular pairs of correlated residues is predicted from the multiple sequence alignment of the corresponding species of proteins A and B treated as if they were a single protein. The algorithm then filters out from that list the intramolecular pairs of correlated residues within A and within B. A schematic representation of the SCM method is shown below.



The process requires analysis of four different multiple sequence alignments: 1) the sequence alignment of A+B, obtained by appending the sequences of protein A to the corresponding sequences of protein B, which are then treated as if each were from a single protein, 2) the multiple alignment of all known sequences of A from different organisms, 3) the multiple alignment of all known sequences of B from different organisms, and 4) the multiple alignment of all known sequences of A together with all known sequences of B.

Using the multiple sequence alignment of A+B as an input to calculate correlated mutations, a list of all intra- and intermolecular pairs of residues (CM(A+B)) is expected as output. In contrast, correlated mutations based on the multiple sequence alignments of A and B will provide a list of likely intramolecular pairs of residues (CM(A) and CM(B), respectively). Since the two monomers A and B are structurally similar to each other, the correlated mutations (CM(A,B)) calculated using the multiple sequence alignment of all known sequences of A together with all known sequences of B will provide an additional filter to eliminate **likely intramolecular** pairs of residues. **Likely intermolecular** pairs of residues (I) will then be the result of the following equation:

$$\mathbf{I = CM(A + B) – CM(A) – CM(B) – CM(A,B)}$$

In order to better identify the residues that are at the heterodimerization interface of A and B, the results of the subtractive correlated mutation method are further pruned based on solvent accessibility values calculated for each residue of A and B from the atomic coordinates of their 3D structures. Specifically, the intermolecular pairs where either one or both residues are completely or partially inaccessible to the solvent are eliminated from the list. The remaining residues of each monomer are then considered to be candidates for the interface of heterodimerization between the two proteins.

FINDINGS

1. Testing the method

Crystallographic structures of dimeric complexes were retrieved from the Protein Quaternary Structure File Server (PQS; http://pqs.ebi.ac.uk) and considered for analysis if they fulfilled the following criteria. **i)** The two proteins in the complex must have less than 80% amino acid sequence identity. This will ensure elimination of homodimers from the test set; **ii)** The mean loss of accessible surface area per chain upon assembly formation compared to the isolated chains must be more than 400 Å. In addition, sequences must contain more than 50 amino acids. These requirements exclude both fragments and peptides from the test set; **iii)** The two monomers must have similar 3D structures (rmsd ≤ 3.0 Å). This condition is required since meaningful 3D models of GPCRs [24] are currently built using the same rhodopsin crystal structure as a template; **iv)** At least 5 corresponding species of the two proteins in the heterodimeric complex must be available.

Among the initial 883 heterodimeric complexes retrieved from the PQS server on December 11, 2001, only 4 structures satisfied all criteria listed above. Application of the SCM method to these 4 structures demonstrated the ability of the method to predict residues at the interface between structurally related proteins. This predictive ability is shown below for one (PDB code: 15C8) of these 4 dimeric complexes.

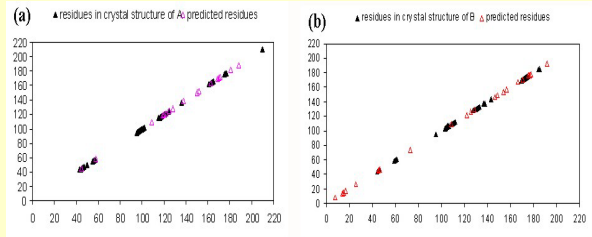


Figure 1. Residues at the heterodimerization interface of A (a) and B (b) in the dimeric complex corresponding to the 15C8 PDB code.

The heterodimer corresponding to the 15C8 PDB code consists of two proteins (A and B) that share a 23% sequence identity and a 2.6 Å structural similarity. Twenty-seven corresponding species of A and B were appended to each other and treated as if they were one protein in order to identify the intra- and intermolecular pairs of correlated residues derived by their multiple sequence alignment. Application of the SCM method identified likely intermolecular residues for both A (magenta triangles) and B (red triangles).

As shown in Figures 1a and b, most of the residues predicted to be at the interface between A and B are either corresponding or very close (< i+7) to residues at the heterodimerization interface of the 15C8 crystallographic structure. Thus, the method predicts 36 % of residues of A and 44% of residues of B at the heterodimerization interface. The remaining predicted residues of A and B (11% and 23%, respectively) that are distant more than 7 residues from the heterodimerization interface are considered to be false positives.

2. Application of the SCM method to the δ-μ opioid receptor heterodimer.

The five available sequences of δ opioid receptor from different organisms were appended to the corresponding sequences of μ opioid receptor and arranged in a multiple sequence alignment. Application of the SCM method to the δ-μ opioid receptor heterodimer identified more than one dimerization interface:

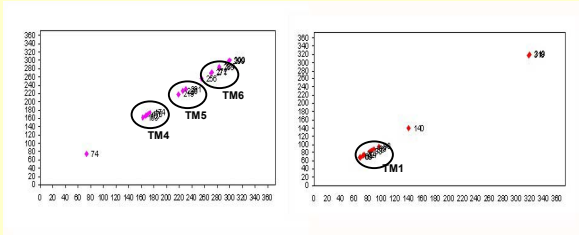


Figure 2. Residues of δ (magenta) and μ (red) opioid receptors predicted to be at the most likely heterodimerization interfaces of the δ-μ complex by the SCM method.

Based on these results, the number of 49 different configurations in which the bundles of 7 TM domains of δ and μ opioid receptors can be positioned next to each other is reduced to a limited number of possibilities. Specifically, the most likely heterodimerization interfaces of the δ-μ heterodimer involve helices TM4, TM5, and TM6 of the δ opioid receptor with helix TM1 of the μ opioid receptor.

Interestingly, application of the SCM method to the μ-κ opioid heterodimer, which is a known pair of receptors that cannot heterodimerize [8], correctly predicts that no residues are likely to be at the heterodimerization interface.

LIMITATIONS OF THE METHOD

The ability of the SCM method to identify heterodimerization interfaces can be influenced by many factors:

1. The analysis requires multiple sequence alignments of **the same GPCR cloned from different organisms**. The sequence alignment has to be limited strictly to the specific receptor for which dimerization is considered.
2. Only a **few sequences from different organisms** are known for each GPCR. As a result, the number of sequences in the multiple sequence alignments is often inadequate for a statistical analysis of the data.
3. **Predictions are limited to the TM regions of the GPCRs under study**, due to the low sequence identity of extracellular and intracellular loops among GPCRs. Therefore, limited reliability is expected for their corresponding multiple sequence alignments and their resulting 3D models based on the rhodopsin crystal structure.
4. **The lack of a statistical validation of the method** due to the presence in the literature of only a few known structures of heterodimeric complexes of structurally similar proteins.

REFERENCES

1. Angers, S., Salahpour, A., Bouvier, M. "DIMERIZATION: An Emerging Concept for G-Protein Coupled Receptor Ontogeny and Function" *Annu. Rev. Pharmacol. Toxicol.* 2002, 42: 409-435
2. Devi, L.A. "Heterodimerization of G-Protein Coupled Receptors: Pharmacology, Signaling and Trafficking" *TRENDS in Pharmacol. Sci.* 2001, 22: 532-537
3. Jones KA, Borowsky B, Tamm JA, Craig DA, Durkin MM, Dai M, Yao WJ, Johnson M, Gunwaldsen C, Huang LY, Tang C, Shen Q, Salon JA, Morse K, Laz T, Smith KE, Nagarathnam D, Noble SA, Branchek TA, Gerald C. "GABA(B) receptors function as a heteromeric assembly of the subunits GABA(B)R1 and GABA(B)R2" *Nature.* 1998, 396: 674-679.
4. Kaupmann K, Malitschek B, Schuler V, Heid J, Froestl W, Beck P, Mosbacher J, Bischoff S, Kulik A, Shigemoto R, Karschin A, Bettler B. "GABA(B)-receptor subtypes assemble into functional heteromeric complexes" *Nature.* 1998, 396: 683-687.
5. White JH, Wise A, Main MJ, Green A, Fraser NJ, Disney GH, Barnes AA, Emson P, Foord SM, Marshall FH. "Heterodimerization is required for the formation of a functional GABA(B) receptor", *Nature.* 1998; 396: 679-682.
6. Maggio R, Barbier P, Colelli A, Salvadori F, Demontis G, Corsini GU. "G protein-linked receptors: pharmacological evidence for the formation of heterodimers." *J Pharmacol Exp Ther.* 1999 291: 251-257.
7. Sawyer GW, Ehlert FJ. "Muscarinic M3 receptor inactivation reveals a pertussis toxin-sensitive contractile response in the guinea pig colon: evidence for M2/M3 receptor interactions." *J Pharmacol Exp. Ther.* 1999, 289: 464-476.
8. Jordan, B.A., Devi, L.A. "G-protein coupled receptor heterodimerization modulates receptor function". *Nature*, 1999, 399: 697-700.
9. George, S.R., Fan, T., Xie, Z., Tse, R., Tam, V., Varghese, G., O'Dowd, B.F. "Oligomerization of mu and delta opioid receptors". *J. Biol. Chem.* , 2000, 275: 26128-26135.
10. Gomes, I., Jordan, B.A., Gupta, A., Trapaidze, N., Nagy, V., Devi, L.A. "Heterodimerization of mu and delta opioid receptors: a role in opioid synergy". *J. Neurosci.* 2001, 20: RC110(1-5).
11. Xie, Z., Lee, S.P., O'Dowd, B.F., George, S.R., "Serotonin 5-HT1B and 5-HT1D receptors form homodimers when expressed alone and heterodimers when co-expressed", *FEBS Lett.* 1999, 456: 63-67.
12. Rocheville, M., Lange, D.C., Kumar, U., Sasi, R., Patel, R.C., Patel, Y.C. "Subtypes of the somatostatin receptor assemble as functional homo and heterodimers". *J. Biol. Chem.* 2000, 275: 7862-7869.
13. Mellado M, Rodriguez-Frade JM, Vila-Coro AJ, Fernandez S, Martin de Ana A, Jones DR, Toran JL, Martinez-A C. "Chemokine receptor homo- or heterodimerization activates distinct signaling pathways". *EMBO J.* 2001 , 20: 2497-507.
14. Gines, S., Hillion, J., Torvinen, M., Crom, S., Casado, V., Canela, E.I., Rondin, S., Lew, J.Y., Watson, S., Zoli, M., Agnati, L.F., Vernier, P., Lluis, C., Ferre, S., Fuxe, K., Franco, R. "Dopamine D1 and adenosine A1 receptors from functionally interacting heteromeric complexes". *Proc. Natl. Acad. Sci. USA* 2000, 97: 8606-8611.
15. AbdAlla, S., Lother, H., Quitterer, U. "AT1-receptor heterodimers show enhanced G-protein activation and altered receptor sequestration". *Nature.* 2000, 407: 94-98.
16. Rocheville, M., Lange, D.C., Kumar, U., Patel, S.C., Patel, R.C., Patel, Y.C. "Receptors for dopamine and somatostatin: formation of hetero-oligomers with enhanced functional activity". *Science* 2000, 288: 154-157.
17. Jordan, B.A., Trapaidze, N., Gomes, I., Nivarthi, R., Devi, L.A. "Oligomerization of opioid receptors with β2-adrenergic receptors: a role in trafficking and mitogen-activated protein kinase activation". *Proc. Natl. Acad. Sci. USA* 2001, 98: 343-348.
18. Ciruela F, Escriche M, Burgueno J, Angulo E, Casado V, Soloviev MM, Canela EI, Mallol J, Chan WY, Lluis C, McIlhinney RA, Franco R. "Metabotropic glutamate 1alpha and adenosine A1 receptors assemble into functionally interacting complexes". *J Biol Chem.* 2001, 276(21):18345-51.
19. Gouldson, P.R., Dean, M.K., Snell, C.R., Bywater, R.P., Gkoutos, G., Reynolds, C.A. "Lipid-facing correlated mutations and dimerization in G-protein coupled receptors". *Prot. Eng.* 2001, 14: 759-767.
20. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Le Trong I, Teller DC, Okada T, Stenkamp RE, Yamamoto M, Miyano M. "Crystal structure of rhodopsin: A G protein-coupled receptor". *Science* 2000, 289: 739-745
21. Valdar, W.S. & Thornton, J.M. "Protein-protein interfaces: Analysis of amino acid conservation in homodimers" *PROTEINS: Struct., Func., and Gen.* 2001, 42: 108-124.
22. Pazos, F., Helmer-Citterich, M., Ausiello, G., Valencia, A. "Correlated mutations contain information about protein-protein interaction" *J. Mol. Biol.* 1997, 271: 511-523.
23. Olmea, O. & Valencia, A. "Improving contact predictions by the combination of correlated mutations and other sources of sequence information" Folding & Design 1997, 2 (Suppl.1) S25-S32.
24. Visiers, I., Ballesteros, J.A. and Weinstein, H. "Three-dimensional representations of G protein-coupled receptor structures and mechanisms" *Meth. Enzymol.* 2002, 343:329-371.