Appendix S1

Helix 8 in chemotactic receptors of the complement system

Szymon Wisniewski, Paulina Dragan, Anna Makal, Dorota Latek

Faculty of Chemistry, University of Warsaw, Poland

Figure A. Global changes of C5aR1 induced by Gi subunits. (I) - A superposition of inactive C5aR1 (blue-to-red, the C5aR1 crystal structure - 5O9H) and active C5aR1 (grey, the homology model based on 6OMM, the cryo-EM structure of FPR2 with G\textsubscript{i}). Both conformations of C5aR1, active and inactive, were subjected to microsecond MD. The structures extracted from the last frames of MD simulations starting from inactive and active C5aR1 were superposed in (II). A global movement of TM6 in inactive C5aR1 towards the active conformation induced by interactions with the G\textsubscript{a} subunit was observed. The G\textsubscript{a} subunits (orange) also overlapped regardless of the starting model of C5aR1 used for building the simulation system (inactive vs. active).

Figure B. A homology model vs. a microsecond MD-refined model of active C5aR1 – the location of G protein subunits. (I) – An RMSD plot representing the differences in the positions of C\textalpha in the G\textalpha subunit. Here, the TM cores in all of the frames were superposed to show only the variability of the G\textalpha position with respect to the starting homology model. (II) – The same RMSD plot for G\textbeta. Here, the G\textalpha subunits of each frame were superposed to observe only the
variability in the Gβ position with respect to Gα. (III) – An RMSD plot for Gγ. Here, the Gβ subunits of each frame were superposed to observe the variability in the Gγ position with respect to Gβ. Interfaces Gα – Gβ and Gβ – Gγ were of the least conformational variability. (IV) A superposition of the homology model of C5aR1 based on FPR2 with Gi (dark blue) and the C5aR1 conformation extracted from the last microsecond MD simulation frame (blue-to-red and orange). The N-terminus of Gα was slightly rotated inwards (a green arrow) in comparison to the FPR2 template, but its C-terminus interacting with the receptor remained in the place.
Figure C. First steps of the C5aR1 activation observed in microsecond MD simulations. Contact maps generated for a starting homology model of inactive C5aR1 with Gα (I, III, V) and for a microsecond MD-refined model of C5aR1 with Gα (II, IV, VI). (I, II) - contact maps for the receptor only. The most visible changes in the receptor conformation were marked with circles (magenta). They included: breaking the TM3-TM6 lock and TM6-TM7 interactions during activation and changes in the TM7-ICL4-H8 region. The most visible was the loss of TM6-TM2 interactions through the TM core of the receptor. (III, IV) - contact maps for the receptor and the Gα subunit. Interacting sequence regions were marked with the same circles as in Fig. 2. In comparison to Fig. 2, a very similar pattern of interactions was obtained at the end of simulations regardless of the starting model - inactive from the C5aR1 crystal structure or active based on FPR2. (V, VI) - contact maps for the whole simulation system. A gradient contact map for the receptor TM core only was included as S3 Appendix.
Figure D. Results of microsecond MD simulations performed for C5aR2. Here, RMSD was computed for the TM core only, without the flexible C-terminus (I-II), and for the bottom half of TM7 and H8 (III-IV). (I) and (III) correspond to the results of MD performed with the inactive receptor structure based on the crystal structure of C5aR1. (II) and (IV) correspond to the results of MD performed with the active receptor structure based on the cryo-EM structure of FPR2.
**Figure E. Amino acid composition of ICL4 loops in C5aR receptors.** In both cases, C5aR1 (I) and C5aR2 (II), polar amino acids are the most populated except for one Leu residue, but ICL4 (green) in C5aR2 is shorter than in C5aR1. Flexible ICL4 loops in both receptors did not form any regular secondary structure during MD simulations.

**Figure F. Loss of crucial interactions between C5aR2 and Gi subunits.** (I, II) – RMSD plots showing a loss of interactions between C5aR2 and the Ga and Gb subunits. In both cases, the distances exceeded 8 Å, the maximal distance between the centers of mass of two residues to describe them as still being in contact. (I) – The RMSD plot for the Ca - Ca distance in the C5aR2 complex corresponding to the 325-349 distance in the C5aR1 complex (see Fig. 3D). This shows the loss of interactions between C5aR2 and Ga. (II) - The RMSD plot for the distance in the C5aR2 complex corresponding to the 331-57 distance in the C5aR1 complex (see Fig. 4D). This shows the loss of interactions between C5aR2 and Gb. (III, IV) - A comparison of the last simulation frames of C5aR1 (grey and dark blue) and C5aR2 (blue-to-red and orange) showed completely different relative positions of H8 and the C-terminus of Ga in these two receptors despite their similar starting conformation based on active FPR2 with Gi. It was caused by shorter ICL3 and ICL4 loops in C5aR2 in comparison to C5aR1. (III) – a side view, (IV) – an intracellular view. Similar loss of contacts between the receptor and Gb can be observed in contact maps (V-VIII). (V, VI) – contact maps generated for the simulation started from active C5aR2 based on FPR2, (VII-VIII) – contact maps generated for the simulation started from inactive C5aR2 based on the crystal structure of C5aR1. Left panels (V and VIII) correspond to the beginning of simulations while right panels correspond to last frames of simulations. Regardless the starting conformation of C5aR2 used for the simulation, loss of several interactions was observed confirming the decreased stability of such simulation system. For example, intra-receptor interactions (ICL2 - C-terminus incl. H8, marked with right black circles) were lost. Also, C-terminal regions of both, the receptor and Ga (dark blue circles) did not interact by the end of simulations so close like in case of C5aR1 (see Fig. 2). Glu325 – Lys349 contact on the C5aR1 – Ga interface (see Fig. 3) was lost in the case of C5aR2. Moreover, there were discrepancies between simulations started from active C5aR2 and inactive one in this region (compare VI and VIII). Finally, regions involved in three-body interactions in C5aR1 (see Fig. 4, yellow circles), in C5aR2 were involved in far fewer contacts (yellow circles) suggesting that such interactions were not present at all.
Figure G. A comparison of loop modeling algorithms in Rosetta. KIC (I), CCD (II), 50 top-scoring loop models were selected from 5000 generated models. Here, an inactive structure of C5aR1 (PDB id: 5O9H) was used for loop modeling. CCD allowed for a more extensive conformational search also including TM helices but was slightly biased towards regular secondary structures.
Figure H. Beta-arrestin binding site - C5aR receptors vs. β₁AR. (I) β₁AR in complex with arrestin-2 (PDB id: 6TKO). Polar contacts were marked with dashed yellow lines. (II) a superposition of β₁AR (green), C5aR1 (orange), and C5aR2 (grey). Here, a crystal structure of inactive C5aR1 and a homology model of inactive C5aR2 based on it were used. For both C5aR receptors, alanine (ICL1 - blue, ICL2 - green) substitutions of active site residues were observed in comparison to β₁AR. Only for C5aR2 Arg/Leu (TM3 – marine green) and Gln/Gly (TM2 – light blue) substitutions were observed. H8 (red) was in the same place in all three receptors (II). Arrestin binding sites in details: β₁AR (III), C5aR1 (based on β₁AR) (IV), and C5aR2 (based on β₁AR) (V). In (II-V) receptors were shown in blue-to-red color schemes.
Figure I. Residues important for C5aR1 function. Here (I - II), the residues described in Figures 3 and 4, involved in the G protein coupling and signal transduction, were shown in orange, with the ‘DRF’ motif shown in red. The residues described in Table 1 as involved in the receptor activation were shown in yellow, the residues involved in the β-arrestin binding were shown in green, and the residues involved in the formation of H8 were shown in magenta. The latter were shown in detail in (II) – the side view. Both, (I) and (II), represent the refined C5aR1 structure extracted from the last frame of microsecond simulations, started from active C5aR1 based on FPR2.
|          | Value   | Value   | Value   | Value   | Value   | Value   | Value   | Value   |
|----------|---------|---------|---------|---------|---------|---------|---------|---------|
|          | 0.997   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   |
|          | 0.997   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   |
|          | 0.997   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   |
|          | 0.997   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   |
|          | 0.997   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   |
|          | 0.997   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   |
|          | 0.997   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   |
|          | 0.997   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   | 0.000   |

Figure J. Prediction of the secondary structure of C5aRI and C5aR2.
