Molecular Dynamics Simulation Analysis of Structural Dynamic Cross Correlation Induced by Odorant Hydrogen-Bonding in Mouse Eugenol Olfactory Receptor

Chisato Okamoto and Koji Ando*
Department of Information and Sciences, Tokyo Woman's Christian University,
2-6-1 Zenpukuji, Suginami-ku, Tokyo 167-8585, Japan

Structural fluctuations and dynamic cross-correlations in the mouse eugenol olfactory receptor (Olfr73) were studied by molecular dynamics (MD) simulation. We focused on the hydrogen (H) bond of the odorant eugenol to Ser113, Asn207, and Tyr260 of the receptor protein, the importance of which has been suggested by previous experimental studies. The H-bond was not observed in docking simulations, but in subsequent MD simulations the H-bond to Ser113 was formed in 2–4 ns. The lifetime of the H-bond was in the range of 1–20 ns. On the trajectory with the most stable (20 ns) H-bond, the structural fluctuation of the α-carbon atoms of the receptor main chain was studied by calculating the root-mean-square fluctuations, the dynamic cross-correlation map, and the time-dependent dynamic cross-correlation. The analysis suggested a correlation transfer pathway Ser113 → Phe182 → (Leu259 or Tyr260) → Tyr291 induced by the ligand binding with a time scale of 4–6 ns.

I. INTRODUCTION

The mammalian olfactory systems are capable of detecting and discriminating between a wide range of chemically diverse molecules [1–9]. Olfactory receptors (ORs) are located on the cell membrane of olfactory nerves in the nasal cavity. Binding of odorant molecules to ORs leads to the dissociation of heterotrimeric G-proteins from the intracellular domain of the OR protein, which transduces chemical signals into neuronal electrical responses. The OR proteins belong to class A of G-protein coupled receptors (GPCRs) [10], which are characterized by seven transmembrane (TM) helices. Due to the current lack of experimental 3D structural data of OR proteins, molecular modeling studies of the OR mechanisms are limited [11–19]. Recent studies have mainly used homology modeling techniques with the available 3D structures of other class A proteins such as bovine rhodopsin and human β_2-adrenergic receptor as templates to construct 3D structural models of OR proteins.

The mouse eugenol olfactory receptor Olfr73 (mOR-EG) has been studied as a prototype of mammalian ORs. By combining computational molecular modeling with experimental techniques such as site-directed mutagenesis, eleven critical amino acid residues (Cys106, Ser113, Phe182, Phe203, Asn207, Glu208, Leu212, Phe252, Ile256, Leu259, Tyr260, see Figure 1) located in TM3, TM5 and TM6 and the second extracellular loop were identified as essential for the ligand binding [16, 17]. Molecular dynamics (MD) simulations of the apo-form OR protein suggested an interesting mechanism for the transfer of structural change information to a residue (Tyr291) linked to the G-protein [18]. Another MD simulation study with a fingerprint interaction analysis showed that the binding pocket of Olfr73 is smaller but more flexible than those of non-olfactory GPCRs [19].

In this work, we further investigate the dynamic response of Olfr73 to ligand binding. The artificial intelligence tool AlphaFold2 [20–22] is used to construct a 3D structural model of Olfr73. The dynamics of both the apo and holo forms are studied by calculating the root-mean-square fluctuations (RMSF), the dynamic cross-correlation map (DCCM) [23–27], and the time-dependent dynamic cross-correlation (TDDCC) from MD trajectories. The dynamic structural response induced by the ligand binding suggests pathways of correlation transfer and their corresponding time scales.

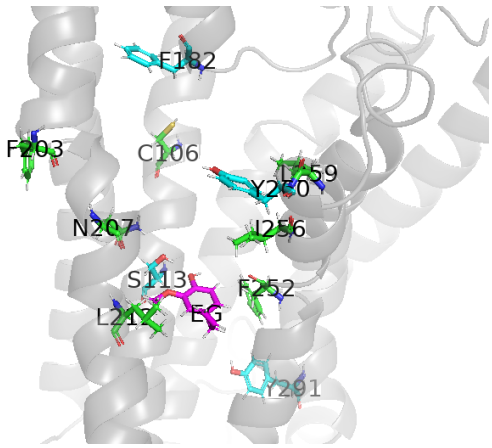

FIG. 1: A snapshot of the MD simulation showing the positions of the twelve key residues, Cys106, Ser113, Phe182, Phe203, Asn207, Glu208, Leu212, Phe252, Ile256, Leu259, Tyr260, and Tyr291, relative to the odorant ligand eugenol (EG).

II. METHODS

A. Structure Model from AlphaFold2

We used the ColabFold web server [28] to construct a 3D structural model of Olfr73. The default options of ColabFold (AlphaFold2 with MMseqs2) were used. The amino-acid sequence was taken from UniProt entry Q920P2.

*E-mail: ando_k@lab.twcu.ac.jp

B. MD Simulation of Apo Form

The structure model of Olfr73 from AlphaFold2 is embedded in a palmitoyl-oleolyl-glycelo-phosphatidyl-ethanolamine (POPE) membrane in a saline solution, using the membgene3 package of the myPresto5 software suite [29]. The number of Na^+ and Cl^- ions was adjusted to neutralize the system. The resulting numbers of molecules and ions in the simulation box were 229 POPE, 30 Na^+, 41 Cl^-, and 10815 H_2O. After structural optimization by minimizing the total energy with positional constraints on the main chain atoms by harmonic potentials, constant pressure and temperature (NPT) simulations were performed to heat and equilibrate the system at 310 K and 1 atm. The simulation time step was 0.5 fs. The Berendsen method [30] with a relaxation time of 100 fs was used to control the pressure and temperature. Equilibration of the box lengths and the total potential energy was achieved in a 1.1 ns simulation. The box lengths were then averaged over an additional 100 ps simulation. The resulting box size was $105.9 \times 78.9 \times 77.7$ Å^3 (Figure 2). Next, a 20 ns NVT (constant volume and temperature) MD simulation was performed. The Gaussian constraint method [31, 32] was used to control the temperature. The simulation time step was 2 fs with the SHAKE method [33, 34] to constrain the bond lengths and angles involving hydrogen atoms. The zero-dipole summation method [35, 36] was used to calculate the long-range electrostatic interaction. We used the software psygene-G [36] for the MD simulation. The Amber99 [37] and Lipid14 [38] potential force-fields were used.

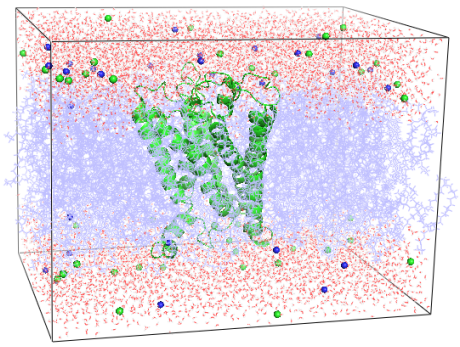

FIG. 2: A snapshot of the MD simulation box.

C. Structure Sampling and Odorant Docking

The first 1 ns trajectory of the apo form was divided into twenty windows of 50 ps each. One configuration per each window of the Olfr73 protein was randomly sampled. An eugenol molecule was docked to target points of the sampled configurations near Ser113 and Asn207, covering the spatial region surrounded by TM3, TM5 and TM6 helices. The software sievgene [39] was used for the docking simulation. The atomic charge parameters for eugenol were determined by the electrostatic potential fitting with the RHF/6-31G(d) wave function. the software GAMESS [40] was used for the electronic structure calculations.

D. MD Simulation of Holo Form

The sampled and docked configurations of Olfr73 and eugenol were returned to the simulation box of the membrane solution. The atomic coordinates of the membrane and the saline solution were optimized by energy minimization with the harmonic constraints to the positions of Olfr73 and eugenol. After heating the temperature to 310 K in 100 ps, the constant NVT MD simulation was performed for 10 ns. In two out of twenty trajectories, stable H-bonding between Ser113 and eugenol was observed for 2–3 ns. The 3 ns H-bond trajectory was divided into ten 300 ps windows, and one configuration per each window was randomly sampled. For each configuration, the atomic velocities were randomized in the Maxwell-Boltzmann distribution to restart the constant NVT MD simulation. Six trajectories maintained the H-bond between eugenol and Ser113 for 1–20 ns. From the 20 ns H-bond trajectory (Supplementary Figure S2), the root-mean-square fluctuation (RMSF), the dynamic cross-correlation map (DCCM) [23–27], and the time-dependent dynamic cross-correlation (TDDCC) were calculated.

III. RESULTS AND DISCUSSION

A. Structure from AlphaFold2

The resulting 3D structure from AlphaFold2 contained a disulfide bond between Cys98 and Cys180. The amino-acid sequences corresponding to the seven TM helices that were determined from the main-chain H-bond structures are; TM1 = Leu24–Asn53, TM2 = Ile58–Thr88, TM3 = Ile94–Pro130, TM4 = Ser138–Leu165, TM5 = Thr193–Met229, TM6 = Ser231–Val262, and TM7 = Ser268–Leu293.

B. Sampling of Hydrogen-Bonded Trajectories

Two representative trajectories of the distance between the O_H atom (the oxygen atom of the OH group) of eugenol and the O_γ atom of Ser113, the O_γ atom of Asp207, and the O_η atom of Tyr260 starting from the docking configurations are shown in Figure 3. Since the O-O distance shorter than 3.5 Å can be considered to form a H-bond, both panels in Figure 3 indicate that a stable H-bond is formed between eugenol and Ser113 for 2–3 ns. For Asp207, Glu208, and Tyr260, stable H-bonds with eugenol were not observed in Figure 3 or in the other eighteen sampled trajectories shown in the Supplementary Figure S1.

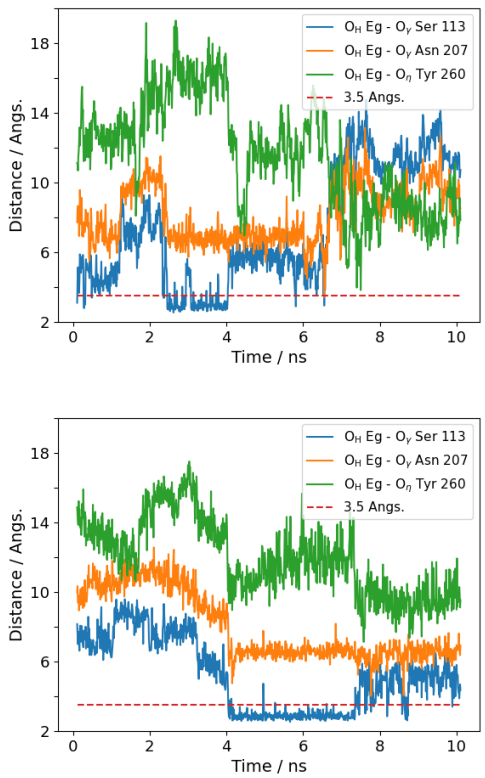

FIG. 3: Representative trajectories of the distance between the O_H atom of eugenol and the O atoms of Ser113, Asn207, and Tyr260, starting from the docking configurations.

Figure 4 shows a snapshot from the MD trajectory of the H-bond structure between eugenol and Ser113, indicating that eugenol is also H-bonded to Glu112. The double H-bond would be the main reason for the long lifetime (> 2 ns) of the ligand binding.

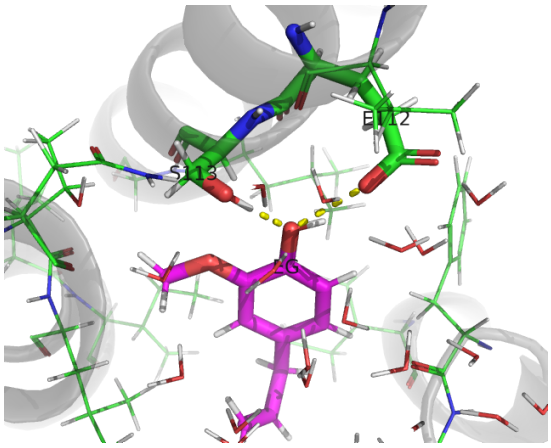

FIG. 4: A snapshot of the hydrogen-bond structure between eugenol (EG) and Ser113 and Glu112.

C. Root-Mean-Squares Fluctuation

Next, we examine the fluctuation of the protein structure induced by the ligand binding. The root-mean-square fluctuation (RMSF) of the Cartesian coordinate $\boldsymbol{r}_i$ of the i-th atom is defined by

$$\mathrm{RMSF}_i = \sqrt{\langle \Delta \boldsymbol{r}_i^2 \rangle} \tag{1}$$

where $\Delta \boldsymbol{r}_i = \boldsymbol{r}_i - \langle \boldsymbol{r}_i \rangle$ is the deviation from the average $\langle \boldsymbol{r}_i \rangle$. In this work, the fluctuations of the amino-acid residues are represented by those of the α-carbon atoms.

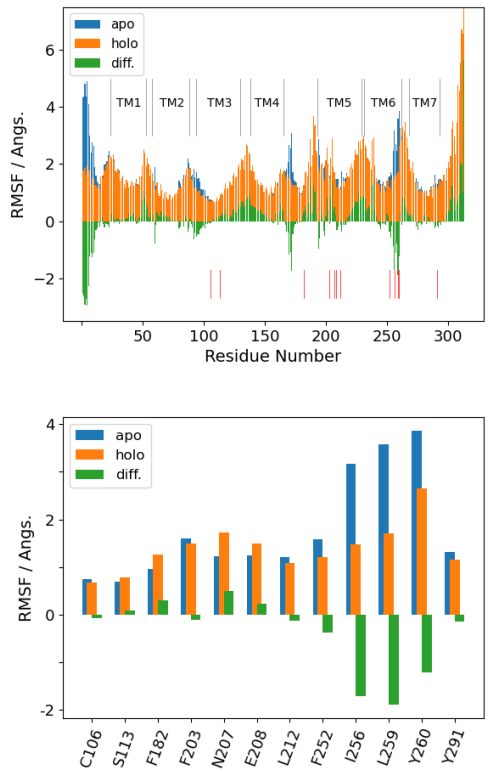

FIG. 5: Root-mean-square fluctuation (RMSF) of the C_α atoms of Olfr73 in the apo and holo forms and their difference (holo minus apo). The left panel shows all residues. TM1–TM7 denote the transmembrane α-helices. The vertical lines below the bar plot indicate the twelve key residues shown in the right panel.

The calculated RMSFs in the apo and holo forms and their difference (holo minus apo) are shown in Figure 5. In both apo and holo forms, the fluctuation is comparatively smaller in the central regions of the TM helices. The RMSF is suppressed in the holo form for small residue numbers (1–12), in the beginning of TM3, after TM4, and in the latst region of TM6. On the other hand, the RMSF is enhanced in the holo-form in the regions between TM1 and TM2, between TM3 and TM4, before TM5, between TM5 and TM6, and after TM7. T welve key residues are taken from the left panel of Figure 5 and displayed in the right panel. Eleven of them (Cys106, Ser113, Phe182, Phe203, Asn207, Glu208, Leu212, Phe252, Ile256,

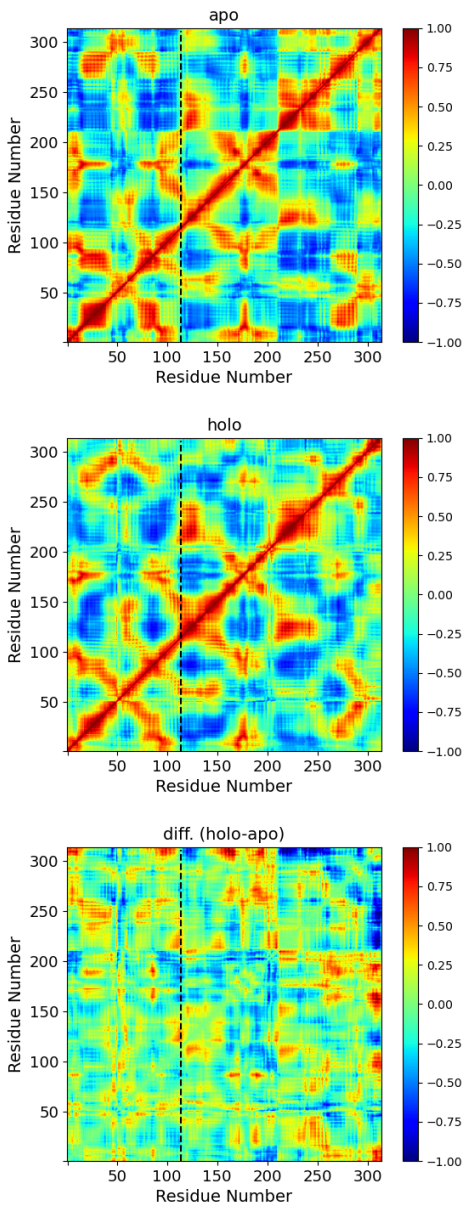

FIG. 6: Dynamic cross-correlation map (DCCM) of the C_α atoms of Olfr73 in the apo and holo forms and their difference (holo minus apo).

Leu259, Tyr260) are those reported in ref [17] as essential for the olfactory function, and one (Tyr291) is bonded to the G-protein [18]. A snapshot of their positions is shown in Figure 1. It can be seen in Figure 5 that the suppression of RMSF in the holo-form is remarkable for Ile256, Leu259, and Tyr260, which are located near the end or TM6.

D. Dynamic Cross Correlation Map

The dynamic cross-correlation map (DCCM) [23–27] is a map of the cross-correlation or covariance between the fluc-

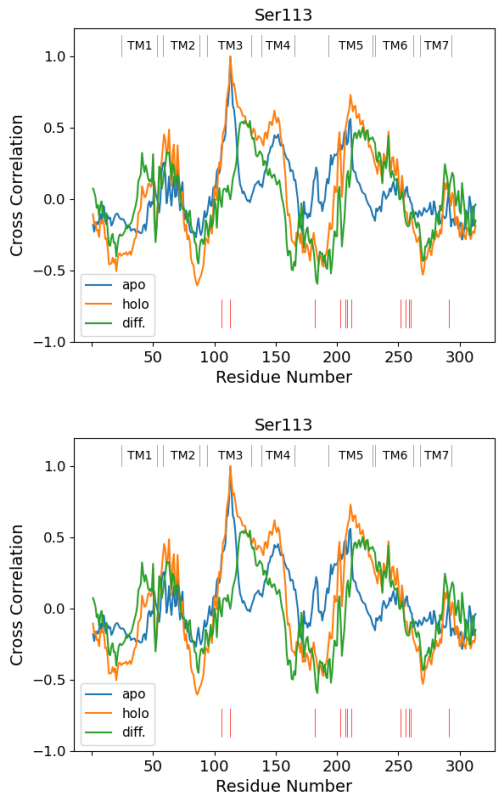

FIG. 7: Sections of the DCCMs in Figure 6 at Ser113 and difference of the DCCMs (holo minus apo) for the key residues.

tuations of the atomic Cartesian coordinates $\boldsymbol{r}_i$ and $\boldsymbol{r}_j$, defined by

$$C_{ij} = \frac{\langle \Delta \boldsymbol{r}_i \cdot \Delta \boldsymbol{r}_j \rangle}{\sqrt{\langle \Delta \boldsymbol{r}_i^2 \rangle}\sqrt{\langle \Delta \boldsymbol{r}_j^2 \rangle}}. \tag{2}$$

The numerator is the average of the inner product between the coordinate shift vectors. The denominator is the normalization factor, which limits the value between −1 and +1. The DCCM is close to +1 when the atomic coordinate fluctuations are in the same direction, close to −1 when they are in the opposite direction, and close to 0 when there is no directional correlation.

The calculated DCCM for the apo and holo forms and their difference (holo minus apo) are shown in Figure 6. Their cross sections at Ser113 are shown in the left panel of Figure 7. (The corresponding plots at the other eleven key residues are displayed in Supplementary Figure S3.) The difference of the correlations is remarkable; between TM2 and TM3 (negative), in the last part of TM3 (positive), between TM4 and TM5 (negative), in the last part of TM5 (positive), between TM6 and TM7 (negative). This means that negative or positive correlations in the structural fluctuation are induced by the H-bond of the ligand eugenol to Ser113 and Glu112.

The difference map for the 12 key residues is shown in the right panel of Figure 7. The map suggests several possible pathways of the correlation transfer. In particular, we note the

correlations between Ser113 and Phe182 (negative), Phe182 and Leu259 (positive), and Leu259 and Tyr291 (negative). Leu259 and Tyr260 show similar behavior. Therefore, one of the simplest correlation transfer pathways would be Ser113 → Phe182 → (Leu259 or Tyr260) → Tyr291.

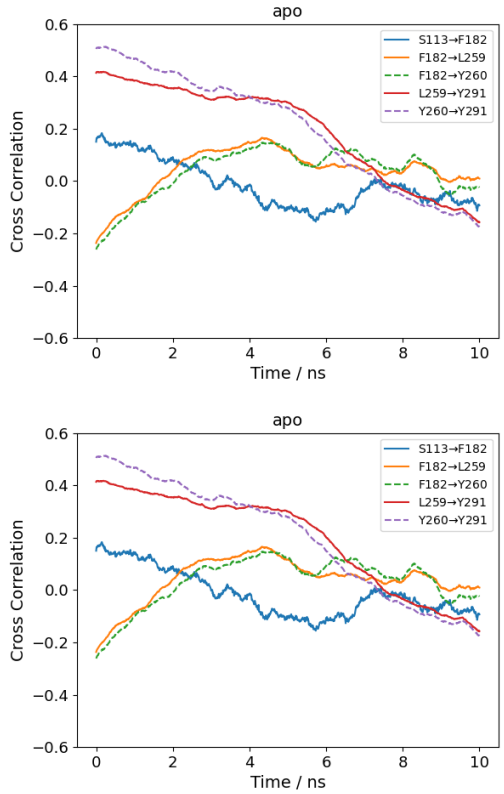

FIG. 8: Time-dependent dynamic cross-correlation (TDDCC) of the apo and holo forms of Olfr73 between the key residues identified as being in the major correlation transfer pathway.

E. Time-Dependent Dynamic Cross Correlation

For the correlation transfer pathway identified in the previous section, we calculated the time-dependent dynamic cross-correlation (TDDCC) defined by

$$C_{ij}(t) = \frac{\langle \Delta \boldsymbol{r}_i(0) \cdot \Delta \boldsymbol{r}_j(t) \rangle}{\sqrt{\langle \Delta \boldsymbol{r}_i^2 \rangle} \sqrt{\langle \Delta \boldsymbol{r}_j^2 \rangle}} \tag{3}$$

This is the cross-correlation with the time delay t of the fluctuations of the atomic coordinates $\boldsymbol{r}_i$ and $\boldsymbol{r}_j$. The DCCM of Eq. (2) is the simultaneous ($t = 0$) part of the TDDCC. Since we compute the TDDCC from the equilibrium simulation, we assume that the essential dynamics of the non-equilibrium response induced by the ligand binding is embedded in the time correlation in the equilibrium (fluctuation-dissipation theorem).

The calculated TDDCC along the path Ser113 → Phe182 → (Leu259 or Tyr260) → Tyr291 is shown in Figure 8. The results with Leu259 and Tyr260 are almost identical. All TDDCCs change their qualitative behavior after the ligand binding. The TDDCC between Ser113 and Phe182 is overall a decaying function of time in the apo form, while it is basically an increasing function of time in the holo form. Interestingly, they have minimum or maximum at $t \simeq 6$ ns. The TDDCC between Phe182 and Leu259 (or Tyr260) decays from positive to negative in 10 ns in the apo form, while in the holo form it oscillates near zero on an ns timescale. The TDDCC between Leu259 (or Tyr260) and Tyr291 increases from negative to positive in ~ 4 ns and stays near zero (or slowly decays to zero) in the apo form, while it decreases from positive to negative overall with oscillation in ns timescale.

IV. CONCLUSION

Structural fluctuations and dynamic cross-correlations in the mouse eugenol olfactory receptor Olfr73 were studied by molecular dynamics (MD) simulation. Among the candidate residues for H-bonding with the odorant ligand eugenol, Ser113, Asn207, and Tyr260, we find that only Ser113 forms a stable H-bond in the sampled trajectories. The lifetime of the H-bond was in the range of 1–20 ns. The structural fluctuations of the C_α atoms of the receptor main chain were investigated by calculating the root-mean-square fluctuations (RMSF), the dynamic cross-correlation map (DCCM), and the time-dependent dynamic cross-correlation (TDDCC). The analysis suggested a correlation transfer pathway Ser113 → Phe182 → (Leu259 or Tyr260) → Tyr291 induced by the ligand binding. The TDDCC indicated that the time scale of the correlation transfer was 4–6 ns. However, to our knowledge, there is no experimental evidence to verify these results. Time-dependent spectroscopies such as fluorescence resonance excitation transfer (FRET) [41], resonance Raman [42], and transient-grating [43] could be useful to clarify the time scale of the structural correlation transfer.

In this work, we focused on the twelve key residues identified in the previous experimental studies. The correlation transfer pathways highlighted in this work appear to be the most straightforward ones that directly involve the key residues. More extensive analysis would be required to clarify the role of other residues that may be directly or indirectly involved in the correlation transfer. In addition, the precise mechanism by which the dissociation of the G-protein is induced is still unclear. Work on these issues is ongoing.

Acknowledgments

We thank Dr. S. Katada and Prof. K. Touhara for useful discussions. This work has been supported by JSPS KAKENHI No. 19K22173.

[1] L. Buck and R. Axel, Cell **65**, 175 (1991),
[2] P. Kraft, J. A. Bajgrowicz, C. Denis, and G. Fráter, Angew. Chem. Int. Ed. Engl. **39**, 2980 (2000),
[3] S. Firestein, Nature **413**, 211 (2001),
[4] S. Zozulya, F. Echeverri, and T. Nguyen, Genome Biology **2**, research0018.1 (2001),
[5] X. Zhang and S. Firestein, Nat. Neurosci **5**, 124 (2002),
[6] S. Serizawa, K. Miyamichi, H. Nakatani, M. Suzuki, M. Saito, Y. Yoshihara, and H. Sakano, Science **302**, 2088 (2003),
[7] B. Malnic, P. A. Godfrey, and L. B. Buck, Proceedings of the National Academy of Sciences **101**, 2584 (2004),
[8] R. R. Reed, Cell **116**, 329 (2004),
[9] C. S. Sell, *Chemistry and the Sense of Smell* (Wiley, 2014),
[10] D. Yang, Q. Zhou, V. Labroska, S. Qin, S. Darbalaei, Y. Wu, E. Yuliantie, L. Xie, H. Tao, J. Cheng, et al., Signal Transduction and Targeted Therapy **6** (2021),
[11] A. Di Pizio and M. Y. Niv, Israel Journal of Chemistry **54**, 1205 (2014),
[12] L. Doszczak, P. Kraft, H.-P. Weber, R. Bertermann, A. Triller, H. Hatt, and R. Tacke, Angewandte Chemie International Edition **46**, 3367 (2007),
[13] L. Gelis, S. Wolf, H. Hatt, E. M. Neuhaus, and K. Gerwert, Angewandte Chemie International Edition **51**, 1274 (2011),
[14] W. B. Floriano, N. Vaidehi, W. A. Goddard, M. S. Singer, and G. M. Shepherd, Proceedings of the National Academy of Sciences **97**, 10712 (2000),
[15] W. B. Floriano, Chemical Senses **29**, 269 (2004),
[16] S. Katada, T. Hirokawa, Y. Oka, M. Suwa, and K. Touhara, J. Neurosci. **25**, 1806 (2005),
[17] O. Baud, S. Etter, M. Spreafico, L. Bordoli, T. Schwede, H. Vogel, and H. Pick, Biochemistry **50**, 843 (2010),
[18] O. Baud, S. Yuan, L. Veya, S. Filipek, H. Vogel, and H. Pick, Sci. Rep. **5**
[19] S. Yuan, T. Dahoun, M. Brugarolas, H. Pick, S. Filipek, and H. Vogel, Commun. Biol. **2** (2019),
[20] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, et al., Nature **596**, 583 (2021), ISSN 1476-4687,
[21] M. Akdel, D. E. V. Pires, E. P. Pardo, J. Jänes, A. O. Zalevsky, B. Mészáros, P. Bryant, L. L. Good, R. A. Laskowski, G. Pozzati, et al., Nat. Struct. Mol. Biol. **29**, 1056 (2022),
[22] Z. Yang, X. Zeng, Y. Zhao, and R. Chen, Signal Transduct. Target. Ther. **8** (2023),
[23] J. A. McCammon, Rep. Prog. Phys. **47**, 1 (1984),
[24] R. P. da Penha Valente, R. C. de Souza, G. de Medeiros Muniz, J. E. V. Ferreira, R. M. de Miranda, A. H. L. e Lima, and J. L. da Silva Gonçalves Vianez Junior, Sci. Rep. **10** (2020),
[25] P. K. Parida, D. Paul, and D. Chakravorty, Phytother. Res. **34**, 3420 (2020),
[26] Z. Wang, B. Hu, Y. An, and J. Wang, Oxid. Med. Cell. Longev. **2022**, 1 (2022),
[27] R. Dash, Y. A. Munni, S. Mitra, H. J. Choi, S. I. Jahan, A. Chowdhury, T. J. Jang, and I. S. Moon, Sci. Rep. **12** (2022),
[28] M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, and M. Steinegger, Nat. Methods **19**, 679 (2022), ISSN 1548-7105,
[29] Y. Fukunishi, Y. Mikami, and H. Nakamura, J. Phys. Chem. B **107**, 13201 (2003),
[30] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, J. Chem. Phys. **81**, 3684 (1984),
[31] D. J. Evans, W. G. Hoover, B. H. Failor, B. Moran, and A. J. C. Ladd, Phys. Rev. A **28**, 1016 (1983),
[32] S. Nosé, Prog. Theor. Phys. Suppl. **103**, 1 (1991),
[33] J.-P. Ryckaert, G. Ciccotti, and H. J. Berendsen, J. Comput. Phys. **23**, 327 (1977),
[34] K. Kasahara, B. Ma, K. Goto, B. Dasgupta, J. Higo, I. Fukuda, T. Mashimo, Y. Akiyama, and H. Nakamura, Biophys. Physicobiol. **13**, 209 (2016),
[35] I. Fukuda, Y. Yonezawa, and H. Nakamura, J. Chem. Phys. **134**
[36] T. Mashimo, Y. Fukunishi, N. Kamiya, Y. Takano, I. Fukuda, and H. Nakamura, J. Chem. Theory Comput. **9**, 5599 (2013),
[37] J. Wang, P. Cieplak, and P. A. Kollman, J. Comput. Chem. **21**, 1049 (2000),
[38] C. J. Dickson, B. D. Madej, Å. A. Skjevik, R. M. Betz, K. Teigen, I. R. Gould, and R. C. Walker, J. Chem. Theory Comput. **10**, 865 (2014),
[39] Y. Fukunishi, Y. Mikami, and H. Nakamura, J. Mol. Graph. Model **24**, 34 (2005),
[40] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, et al., J. Comput. Chem. **14**, 1347 (1993),
[41] T. A. Nguyen, H. L. Puhl, K. Hines, D. J. Liput, and S. S. Vogel, Nat. Commun. **13** (2022),
[42] S. Yamashita, M. Mizuno, and Y. Mizutani, J. Chem. Phys. **156**
[43] M. Terazima, Acc. Chem. Res. **54**, 2238 (2021),