Supporting Information

# Physical Origin of Thermostabilization by a Quadruple Mutation for the Adenosine $A_{2a}$ Receptor in the Active State

Yuta Kajiwara,[1] Satoshi Yasuda,[2–4] Simon Hikiri,[2,3] Tomohiko Hayashi,[4] Mitsunori Ikeguchi,[5,6] Takeshi Murata,*[,2,3,7] and Masahiro Kinoshita*[,4]

[1]Graduate School of Energy Science, Kyoto University, Uji, Kyoto 611-0011, Japan
[2]Graduate School of Science, Chiba University, 1-33 Yayoi-cho, Inage, Chiba 263-8522, Japan
[3]Molecular Chirality Research Center, Chiba University, 1-33 Yayoi-cho, Inage, Chiba 263-8522, Japan
[4]Institute of Advanced Energy, Kyoto University, Uji, Kyoto 611-0011, Japan
[5]Graduate School of Medical Life Science, Yokohama City University, Tsurumi, Yokohama 230-0045, Japan
[6]RIKEN Medical Sciences Innovation Hub Program, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan
[7]JST, PRESTO, 1-33 Yayoi-cho, Inage, Chiba 263-8522, Japan

**1. Entropic Excluded-Volume Effect by Solvent**

First, we discuss a water-soluble protein. As illustrated in Fig. S1(A), the presence of a side chain generates an excluded volume (EV), a volume which the centers of water molecules cannot enter. When side chains are closely packed, the overlap of EVs occurs and the total EV reduces, leading to an increase in the total volume available to the translational displacement of water molecules in the system followed by a large gain of translational, configurational entropy of water. We have shown that this entropic effect plays a pivotal role in protein folding.[1–8]

For a membrane protein, the hydrocarbon ($CH_2$, $CH_3$, and CH) groups constituting nonpolar chains of lipid molecules should act as “solvent” just like water for a

water-soluble protein. Protein folding, especially the side-to-side association of helices illustrated in Fig. S1(B), is accompanied by a large gain of solvent entropy. The solvent possesses not only the translational entropy (TE) but also the orientational (rotational) and vibrational entropies. However, upon protein folding the increase in orientational and vibrational freedoms occurs only for the solvent molecules in the close vicinity of the protein whereas that in translational freedom reaches all of the solvent molecules in the system.[2,9] As a consequence, the TE contribution dominates irrespective of the solvent species.

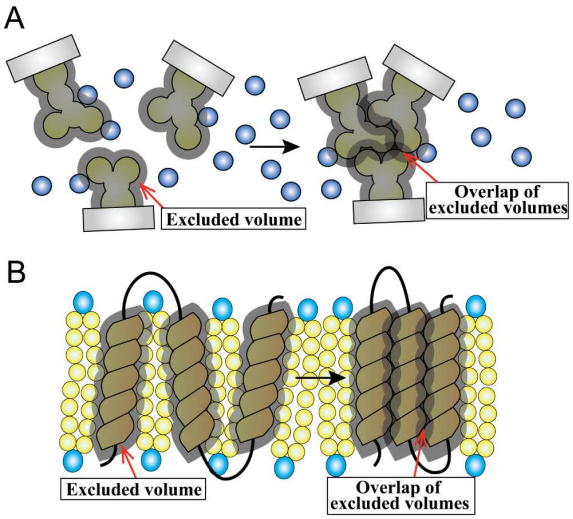


Figure S1. Close packing of side chains of a protein in solvent. (A) The solvent is water (the blue circles represent water molecules) for a water-soluble protein. (B) Side-to-side association of α-helices of a membrane protein in solvent. The association is characterized by close packing of side chains. The solvent is formed by hydrocarbon ($CH_2$, $CH_3$, and CH) groups represented by the yellow circles.

**2. Protein Intramolecular Hydrogen Bonding**

A gain of protein intramolecular van der Waals (vdW) attractive interaction upon protein folding is somewhat cancelled out by the loss of protein-solvent vdW attractive interaction accompanied. This cancellation of vdW energy is applicable to water-soluble and membrane proteins. For electrostatic energy, however, a membrane protein behaves quite differently from a water-soluble protein.[5] This is because the solvent is nonpolar for the former.

Hydrogen bonding is an essential component of the electrostatic attractive interaction. In aqueous environment, the formation of protein intramolecular hydrogen bonds (IHBs) is accompanied by the break of protein-water hydrogen bonds. However,

this is not the case in nonpolar environment: There is no energetic penalty arising from the break of protein-solvent hydrogen bonds. Therefore, the contribution to structural stability from the formation of IHBs is much stronger in nonpolar environment.[5,8]

### 3. Solvent Model

Recent experimental data[10] have shown that many membrane proteins fold and oligomerize quite efficiently in nonpolar environments which bear little similarity to a membrane (e.g., those provided by amphipols). This strongly suggests that folding of a membrane protein occurs only if the surrounding solvent molecules thermally move (i.e., the solvent-entropy effect comes into play) and the intramolecular hydrogen bonding is a sufficiently powerful contributor to the structure formation: The details of specific characteristics of nonpolar chains of lipid molecules are not very important. Further, we note that a membrane is immersed in water. When a membrane protein takes a structure with a larger EV, the membrane also generates a larger EV for water molecules. Thus, water indirectly acts as the solvent. Consequently, we should take the view that a membrane protein is immersed in bulk solvent. We have shown that water can be modeled as "simple fluids" in any theoretical method focused on the entropic EV effect at ambient temperature and pressure.[3,6,7]

The above discussions have motivated us to employ a simplified model for the solvent: an ensemble of neutral hard spheres whose diameter and packing fraction are set at those of water at 298 K and 1 atm. This parameter setting for nonpolar chains of lipid molecules can be justified as follows. The solvent-entropy effect becomes larger as the solvent diameter decreases or the packing fraction increases.[3] The diameters of $CH_2$, $CH_3$, and CH groups are larger than the molecular diameter of water but their packing fraction is higher than the water value. These two properties are rather compensating, and the parameter setting mentioned above becomes reasonable.

The simplified solvent model thus obtained allows us to quantify the solvent-entropy gain upon protein folding within the membrane on the basis of statistical mechanics. Its validity was shown to provide a good result for membrane proteins in our earlier works.[5,11–14]

### 4. Application to GPCR folding

The stability of a G protein-coupled receptor (GPCR) should be governed by that of its transmembrane (TM) region.[11–14] Therefore, only the TM region is considered. For GPCR folding, we account for only the two physical factors which should be the most essential: a gain of solvent entropy and lowering of intramolecular electrostatic energy whose principal component is hydrogen bonding.

**5. Calculation of Solvent-Entropy Gain upon GPCR Folding**

The solvent-entropy gain arising from the close packing of side chains in the side-to-side association of the seven helices is considered as the dominant contributor to the net gain. The solvent-entropy gain $\Delta S$ is calculated using a hybrid of an integral equation theory (IET)[15] and our morphometric approach (MA).[16,17] The former is a statistical-mechanical theory of solvation and the latter is necessitated to treat a large protein with complex polyatomic structure with moderate computational burden. Using this hybrid, we can finish the calculation of $\Delta S$ in less than 0.5 sec for a folded structure given.

The idea of the MA is to express $\Delta S$ by the linear combination of changes in only four geometric measures of a GPCR:

$$\Delta S/k_B=C_1\Delta V_{ex}+C_2\Delta A+C_3\Delta X+C_4\Delta Y. \qquad \text{(S1)}$$

Here, $V_{ex}$ is the EV generated by the GPCR, $A$ is the solvent-accessible surface area (SASA), $X$ and $Y$ are the integrated mean and Gaussian curvatures of the solvent-accessible surface, respectively, $\Delta X$ is the change in quantity $X$ upon GPCR folding, and $k_B$ is the Boltzmann constant. The solvent-accessible surface is the surface that is accessible to the centers of solvent particles. In the MA, the GPCR structure enters $\Delta S$ only via the four geometric measures. Therefore, the four coefficients ($C_1$–$C_4$), which are dependent only on the solvent species and its thermodynamic state, can be determined in simple geometries: They are calculated beforehand from the solvation entropies of isolated, spherical solutes with various diameters. Once $C_1$–$C_4$ are determined, $\Delta S$ for a folded structure given is obtained from Eq. S1 only if $\Delta V_{ex}$, $\Delta A$, $\Delta X$, and $\Delta Y$ are calculated from the ($x$, $y$, $z$) coordinates of centers and the diameters of all the constituent atoms.

The values of the four coefficients are as follows: $C_1$=−0.26966 Å$^{-3}$, $C_2$=0.21418 Å$^{-2}$, $C_3$=−0.18719 Å$^{-1}$, and $C_4$=0.05103. The contributions from $C_1\Delta V$ and $C_2\Delta A$ are much larger than those from the other two terms. It is straightforward that $C_1$ is negative and a decrease in the EV leads to higher solvent entropy. Positive $C_2$ may be counterintuitive but can be interpreted as follows.[7] First, we explain "solvent crowding". The presence of a solvent particle also generates an EV for the *other* solvent particles. Thus, all the solvent particles are entropically correlated. This entropic correlation, which is referred to as "solvent crowding", is serious within the TM region. Significantly many solvent particles come in contact with the protein, leading to the formation of a denser layer of solvent particles in the immediate vicinity of the protein. By this formation, the EVs generated by the protein and by the solvent particles in contact with the protein overlap. The total volume available to the *other* solvent particles then increases by the overlapped volume, which leads to the

mitigation of the solvent crowding. The entropic gain originating from this reduction is larger than the entropic loss occurring for the solvent particles in contact with the protein. Larger SASA allows a larger number of solvent particles to come in contact with the protein surface, reducing the solvent crowding to a larger extent.

## 6. Calculation of lowering of intramolecular electrostatic energy upon GPCR Folding

Hydrogen bonding is the principal component of the lowering of intramolecular electrostatic energy. We examine all the donors and acceptors for backbone-backbone, backbone-side chain, and side chain-side chain IHBs within the TM region of a folded structure given. The examination is made using the criteria proposed by McDonald and Thornton[18] with the modification that the maximum distances between H and the acceptor and between the donor and the acceptor are 3.0 Å and 4.4 Å, respectively. (H is the hydrogen atom covalently bound to the donor.) When an IHB is formed, an energy decrease of $D$ is considered. $D$ is made dependent on the distance between centers of H and the acceptor, $d$: This is why the modification of the criteria mentioned above is necessitated. For $d \leq 1.5$Å, $D$ is set at $E^*$. As d increases from 1.5Å, $|D|$ decreases lineally and becomes zero for $d \geq 3.0$Å. As a result, $\Delta A$ calculated can be expressed as $xN_{\mathrm{IHB}}E^*$ where $N_{\mathrm{IHB}}$ denotes the number of IHBs and $0<x\leq 1$.

The value of $E^*$ calculated using quantum chemistry for the formation of a hydrogen bond in gas phase is $-10k_{\mathrm{B}}T_0$.[19] If an H-acceptor pair was in vacuum, $E^*$ could be set at $-10k_{\mathrm{B}}T_0$. However, it is in the environment where atoms with positive and negative partial charges are present. $E^*$ can be regarded as the potential of mean force between the pair in such environment. In this respect, $|E^*|$ should be significantly smaller than $10k_{\mathrm{B}}T_0$ (factor 1). For example, it has been suggested that $E^*$ be $\sim -3.4k_{\mathrm{B}}T_0$.[20] On the other hand, the net lowering of electrostatic energy upon GPCR folding arises from not only the intramolecular hydrogen bonding but also the other electrostatic interaction. When the latter is incorporated in $E^*$, the resulting $|E^*|$ should become larger (factor 2). In our earlier work,[14] $E^*$ was optimized to $-5k_{\mathrm{B}}T_0$ by accounting for factors 1 and 2.

## REFERENCES

(1) Harano, Y.; Kinoshita, M. Translational-entropy gain of solvent upon protein folding. *Biophys. J.* **2005**, *89*, 2701−2710.

(2) Yoshidome, T.; Kinoshita, M.; Hirota, S.; Baden, N.; Terazima, M. Thermodynamics of apoplastocyanin folding: comparison between experimental and theoretical results. *J. Chem. Phys.* **2008**, *128*, 225104.

(3) Kinoshita, M. Roles of translational motion of water molecules in sustaining life. *Front. Biosci.* **2009**, *14*, 3419−3454.

(4) Kinoshita, M. Importance of translational entropy of water in biological self-assembly processes like protein folding. *Int. J. Mol. Sci.* **2009**, *10*, 1064−1080.

(5) Yasuda, S.; Oshima, H.; Kinoshita, M. Structural stability of proteins in aqueous and nonpolar environments. *J. Chem. Phys.* **2012**, *137*, 135103.

(6) Kinoshita, M. A new theoretical approach to biological self-assembly. *Biophys. Rev.* **2013**, *5*, 283−293.

(7) Oshima, H.; Kinoshita M. Essential roles of protein-solvent many-body correlation in solvent-entropy effect on protein folding and denaturation: comparison between hard-sphere solvent and water. *J. Chem. Phys.* **2015**, *142*, 145103.

(8) Hayashi, T.; Yasuda, S.; Škrbić, T.; Giacometti, A.; Kinoshita, M. Unraveling protein folding mechanism by analyzing the hierarchy of models with increasing level of detail. *J. Chem. Phys.* **2017**, *147*, 125102.

(9) Kinoshita, M. Molecular origin of the hydrophobic effect: analysis using the angle-dependent integral equation theory. *J. Chem. Phys.* **2008**, *128*, 024507.

(10) Popot, J. L.; Engelman D. M. Membranes do not tell proteins how to fold. *Biochemistry* **2016**, 55, 5−18.

(11) Yasuda, S.; Kajiwara, Y.; Takamuku, Y.; Suzuki, N.; Murata, T.; Kinoshita, M. Identification of thermostabilizing mutations for membrane proteins: Rapid method based on statistical thermodynamics. *J. Phys. Chem. B* **2016**, *120*, 3833−3843.

(12) Kajiwara, Y.; Ogino, T.; Yasuda, S.; Takamuku, Y.; Murata, T.; Kinoshita, M. Physical origins of remarkable thermostabilization by an octuple mutation for the adenosine $A_{2a}$ receptor. *Chem. Phys. Lett.* **2016**, *657*, 119−123.

(13) Kajiwara, Y.; Yasuda, S.; Takamuku, Y.; Murata, T.; Kinoshita, M. Identification of thermostabilizing mutations for a membrane protein whose three-dimensional structure is unknown. *J. Comput. Chem.* **2017**, *38*, 211−223.

(14) Yasuda, S.; Kajiwara, Y.; Toyoda, Y.; Morimoto, K.; Suno, R.; Iwata, S.; Kobayashi, T.; Murata, T.; Kinoshita, M. Hot-spot residues to be mutated common in G protein-coupled receptors of class A: Identification of thermostabilizing mutations followed by determination of three-dimensional structures for two example receptors. *J. Phys. Chem. B* **2017**, *121*, 6341−6350.

(15) Hansen, J.-P.; McDonald, I. R. Theory of Simple Liquids, 3rd ed.; Academic Press: London, U.K., **2006**.

(16) König, P. M.; Roth, R.; Mecke, K. R. Morphological thermodynamics of fluids: shape dependence of free energies. *Phys. Rev. Lett.* **2004**, *93*, 160601.

(17) Roth, R.; Harano, Y.; Kinoshita, M. Morphometric approach to the solvation free energy of complex molecules. *Phys. Rev. Lett.* **2006**, *97*, 078101.

(18) McDonald, I. K.; Thornton, J. M. Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol*. **1994**, *238*, 777−793.

(19) Mitchell, J. B. O.; Price, S. L. *Chem. Phys. Lett.* On the relative strengths of amide…amide and amide…water hydrogen bonds. **1991**, *180*, 517−523.

(20) Bowie, J. U. Membrane protein folding: how important are hydrogen bonds? *Curr. Opin. Struct. Biol.* **2011**, *21*, 42−49.