Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics

SUPPORTING INFORMATION

The importance of shape of protein-water interface of kinesin motor domain for dynamics of the surface atoms of the protein

Anna Kuffel and Jan Zielkiewicz

Gdańsk University of Technology, Department of Chemistry

Narutowicza 11/12, 80–952 Gdańsk, Poland.

I. DESCRIPTION OF COMPUTATIONAL METHODS

Definition of solvation layer. On the kinesin surface, 30 regions were selected and hydration shell of each one of them was separately analyzed. Generally, the main criterion to define a separate region was its secondary structure. These amino acids, which were buried inside the protein, were not taken into account. List of all of the regions along with numbers of amino acids belonging to them is presented in Table S1. Figure S1 illustrates the way, in which solvation shells have been defined.

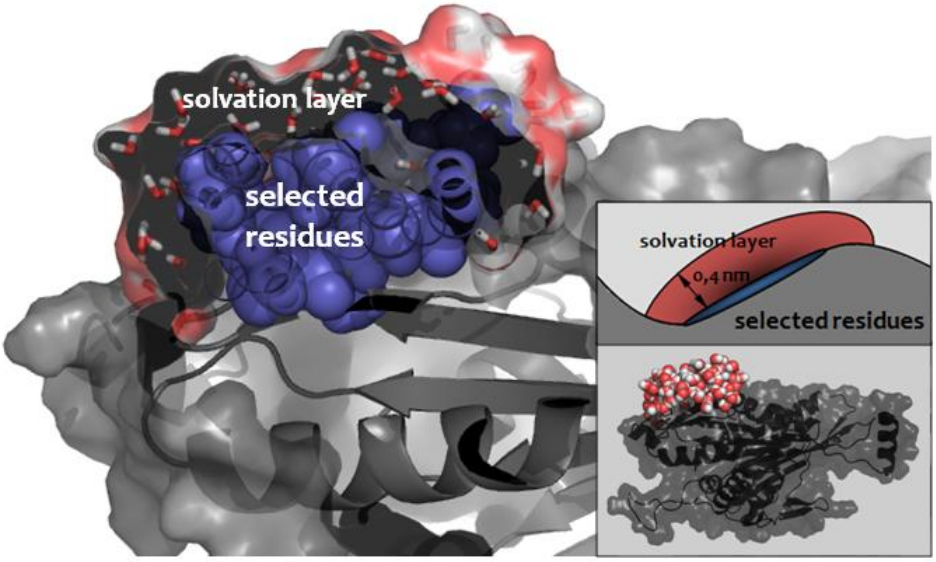

Figure S1. An illustration of solvation shell next to one selected region on kinesin surface. Water molecules, whose centers of mass are within the distance *R*=0.4 nm from selected protein residues are taken into account. On the right lower corner, a whole kinesin motor domain and water molecules surrounding one selected region are shown. Image was created using PyMOL.[1]

It should be noted that the regions α7 and L11were omitted on Figures 5 and 6 in the main text. These regions protrudes very substantially and can move to a high degree independently of the rest of the molecule. Therefore, we have decided to omit them it the discussion concerning protein dynamics. They are however included in all of the rest figures.

Modification of water model. We modified SPCE model of water following the conception described by Sorin *et al.*[2] According to this idea, both ε and σ Lennard-Jones parameters have

been modified, keeping the density of water unchanged. The parameters, along with selected physical properties of modified models, are collected in the table below.

Model	ε [kJ·mol^{-1}] ([kcal·mol^{-1}])	σ [nm]	Density, ρ [g·cm^{-3}]	Translational diffusion, D_T [10^{-9}m^2·s^{-1}]	n_{HB}
mod10	0.41840 (0.10)	0.32277	0.999	1.52	3.48
mod20	0.83680 (0.20)	0.31431	0.998	3.71	3.09
mod25	1.04600 (0.25)	0.31313	0.999	4.54	2.87

The hydrogen bond analysis. During the analysis of the results of computer simulations, it is important to choose such definition of hydrogen bond that will be physically reasonable. This criterion is fulfilled by, for example, the most popular and commonly used, so called "geometric" definition. It resorts to results of neutron scattering data, which suggest the limitation of HB angles to the interval (0° ÷ 30°).[3] Another definition is the "cone" one, put forward by Wernet *et al.* (see reference 32 in the main text). This one is well founded on the results of *X*-ray scattering measurements. It should be noted here that the "cone" definition seems to be more physically reasonable than the geometric one – for illustration see Figure 5 in reference 4. This is why we applied it in our investigations.

According to this definition, the hydrogen bond between two water molecules is formed if the oxygen-oxygen distance fulfills the relation:

$$R_{OO} \leq -0.000044\beta_{OOH}^2 + 0.33[\text{nm}]$$

where β_{OOH} is the O-O-H angle (in degrees). We additionally demanded the total interaction energy between two hydrogen-bonded molecules to be negative – it is roughly equivalent to demanding the force acting between these molecules to be attractive. This total interaction energy was computed as the sum of both Lennard-Jones and electrostatic terms between two molecules, for which the relation above was fulfilled.

To describe the hydrogen bond network structure, we slightly modified Chau and Hardwick's[5] idea; the adopted procedure has been described in details in our previous paper (ref.30 in the main text). The main idea of this approach is the analysis of all angles, θ, between the vectors linking the oxygen atom of the central water molecule with the oxygen atoms of all its partners which form the hydrogen bonds with this molecule (it should be noted here that the number of these partners, of course, is not fixed). Next, we build histogram describing the probability distribution of the angles. This histogram has been normalized, assuming a total area bounded by the graph to be equal to 100. By adopting this procedure to water molecules within solvation layer, probability distribution is obtained, which reflects overall (geometrical) structure of the hydrogen bond network within this space region.

The Voronoi tessellation(η and γ parameters). Voronoi polyhedra were developed by Voronoi in 1908. Bernal and Finney[6] found them to be useful in describing the structure of liquids in the 1960s. Richards[7] applied this method to proteins. Since Procacci and Scateni[8], who used this method to analyze the results of simulation of protein in water, the interest in it is recently increasing.[9] In our paper, we have slightly modified this procedure comparing with its most common use. Usually all water molecules are treated as points and represented by oxygen atoms (ref. 36 in the main text) or by centers of mass (ref. 33 and 34 in the main text). Herein, we treat a tagged water molecule as a point represented by its center of mass, whereas surrounding particles are treated as a set of individual atoms. This approach assures consistency of results of calculations if we adapt the tessellation procedure to water molecules in the neighborhood of the protein surface. In this case, instead of some atoms of water molecules, atoms of protein molecule are present in the vicinity. The asphericity factor, η, used in this work was introduced by Ruocco *et al.*(ref. 33 and 34 in the main text). This dimensionless parameter was designed as a rational measure of the cell shape and it is calculated from the relation: $\eta=A^3/(36\pi V^2)$,where A symbolizes the cell area, while V symbolizes its volume. This method has also been used for calculations of γ parameter (the volume of the first solvation shell, V_{solv}, and the surface area of contact of every selected patch with water, S_{solv}).

Velocity autocorrelation functions and diffusion coefficients. Let $\mathbf{v}_K(t)$ denote a vector of translational ($K=T$) or rotational velocity ($K=R$) of water molecule. The "short time" diffusion coefficients: translational (D_T), and rotational (D_R), were calculated from the appropriate velocity auto-correlation function using the well-known Green-Kubo relation:[10]

$$D_K = \frac{1}{3}\lim_{t\to\infty}\left(\int_0^t C_K(t)dt\right)$$

where $C_K(t)$ symbolizes the translational (or angular) velocity auto-correlation function:

$$C_K(t) = \lim_{T\to\infty}\left(\frac{1}{T}\int_0^T \mathbf{v}_K(x)\mathbf{v}_K(x+t)dx\right) = \left\langle \mathbf{v}_K(x)\mathbf{v}_K(x+t)\right\rangle$$

Determination of autocorrelation function requires, of course, monitoring of the state of each molecule over the time; in this work we used 1.5 ps as a correlation time (it is assumed to be "infinity" in the equation above). During this period each molecule moves around, changing its position in space. The problem: how to ascribe particular molecules to the solvation shells, has been described in detail in our previous paper(ref. 37 in the main text). The adopted procedure allows us to determine diffusion coefficients within relatively thin solvation shell.

To make our results more reliable, many separate calculations have been performed and results in Table S2 represent obtained averages.

The use of the s_{conf} parameter instead of the s_{ort} one. The local ordering parameter has been defined previously (see ref. 30 in the main text) using the two-particle correlation function for water and there are two versions of it. The first one is called $s_{ort}(R_c)$ and it is sensible to the mutual orientations of a central molecule and molecules surrounding it. The second one is called$s_{conf}(R_c)$;it is a simplified form of the $s_{ort}(R_c)$ parameter and its value is calculated assuming that surrounding molecules are represented by points. In the cited paper we have demonstrated that parameter $s_{conf}(R_c)$ can be successfully used instead of $s_{ort}(R_c)$, which is more difficult to calculate. Now we want to answer the question, whether there still is an interdependence of both of these parameters for water near different regions of kinesin surface.

Because of difficulties in determination of the $s_{ort}(R_c)$ quantity (for more details see our previous papers, references 30 and 31 in the main text), we confirm the relationship between $s_{conf}(R_c)$ and $s_{ort}(R_c)$ parameters only within the solvation layer of double thickness (equal to 0.8 nm); it is shown in Figure S2. We assume, however, that it is also true within the first solvation layer (0.4 nm thick). The validity of this assumption can be supported by the results presented on Figure S3, described below. Therefore, parameter $s_{conf}(R_c)$ will be used as a measure of orientational effects of water molecules within solvation layers.

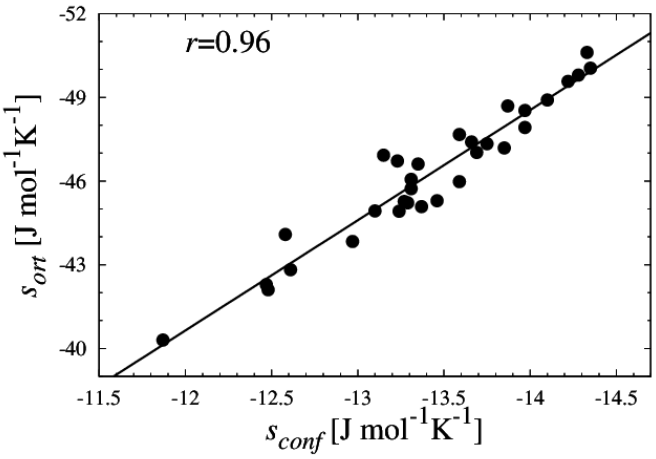

Figure S2. Correlation between the parameters $s_{ort}\,(0.58)$ and $s_{conf}\,(0.58)$, calculated around different patches on the kinesin surface and using double thickness of solvation layer. Straight line on this graph was plotted using the least-square fitting method. The value of correlation coefficient *r* is also included.

Previously[11] for a system consisting solely of water, a correlation has been found between a mean number of hydrogen bonds in water and the two-particle contribution to the entropy, $s_{conf}^{(2)}$ (if water is the only constituent, we have $s_{conf} = s_{conf}^{(2)}$, what is a simple consequence of a definition of s_{conf} parameter). Therefore we can say that a mean number of hydrogen bonds formed by a water molecule reflects in a way the degree of structural ordering of water. However, the results discussed in that paper relate to the situation, when thermal motions are responsible for breaking of hydrogen bonds and for changing the values of $s_{conf}^{(2)}$. Because of that, this kind of correlation does not have to occur for water in solvation layer. Nevertheless, as can be seen on Figure S3, the values of s_{conf} parameter are undoubtedly correlated with the mean number of hydrogen bonds, in which a single molecule participate; note that Figure S3 demonstrates the existence of the correlation within solvation layer of both – single and double – thickness. However, in our opinion, in this case the degree of local corrugation of the surface is responsible for the observed correlation. It is discussed in the main text.

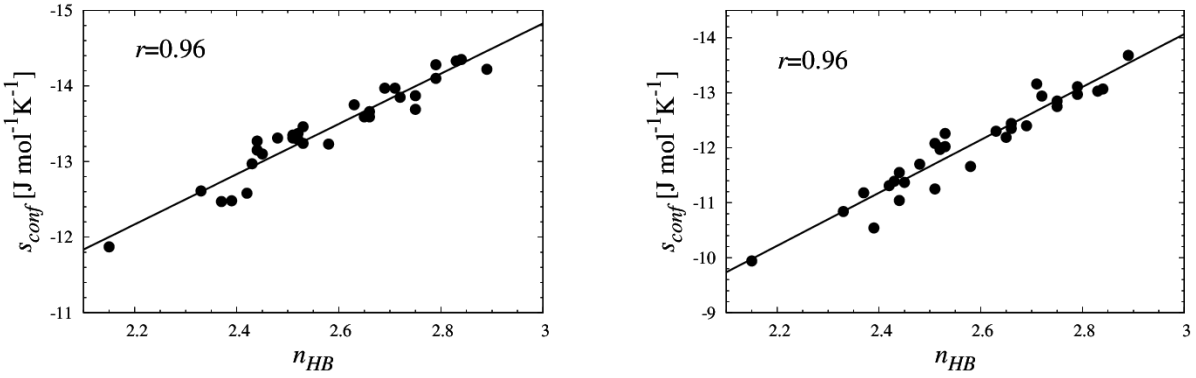

Figure S3. Correlation between mean number of the hydrogen bonds per water molecule n_{HB} and the parameter s_{conf} for solvation layers of single (0.4 nm, left) and double (0.8 nm, right) thickness. The values of correlation coefficients, *r*, are also included.

Modification of the local ordering parameter idea– comparison of results of calculations obtained using various d_0 values.

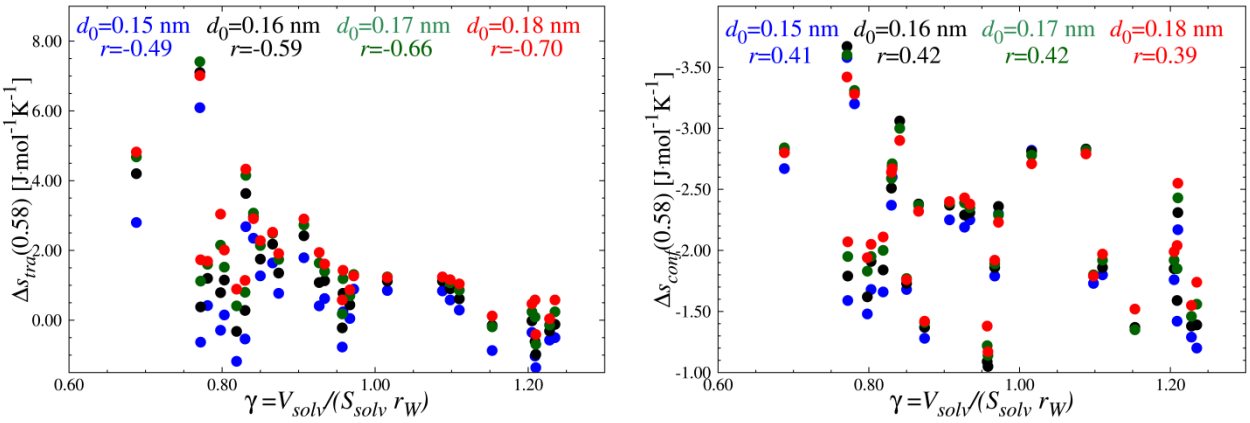

Figure S4.Left – Δs_{conf} quantity versus the parameter *γ*, calculated using various d_0 distances. Right –Δs_{tra} quantity versus the parameter *γ*, calculated using various d_0 distances. The values of correlation coefficient, *r*, are included.

Correlation of s_{tra}(0.58)and Δs_{tra}(0.58) with translational diffusion coefficient, D_T.

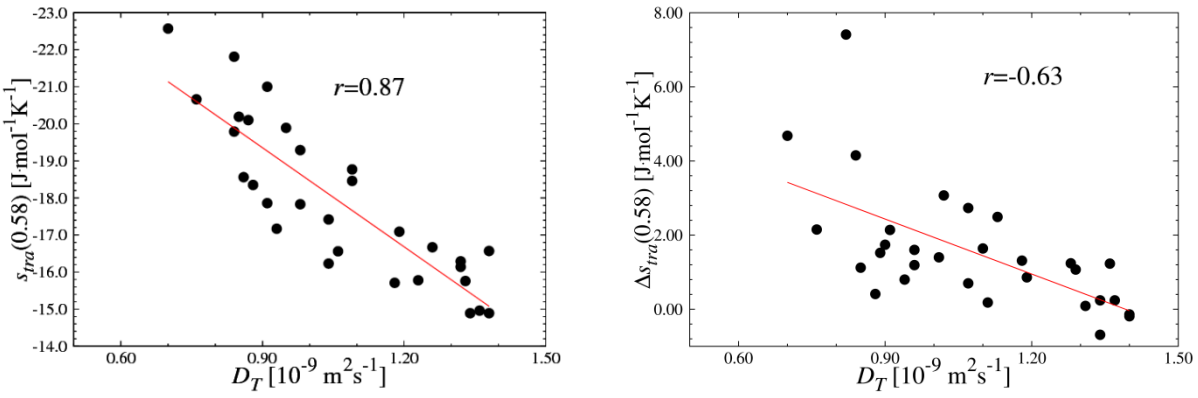

Figure S5. Left – s_{tra} quantity versus the translational diffusion coefficient, D_T. Right – Δs_{tra} quantity versus the translational diffusion coefficient, D_T. The values of correlation coefficient, *r*, are included.

The analysis of protein inner motions was carried out using a covariance matrix of internal atomic displacements, and the procedure is available in the literature(see reference 35in the main text).The first step was to determine mean position of each one of all heavy atoms of the protein (that is: besides hydrogen atoms). These mean coordinates have been calculated over the total time of the simulation (20 ns) and over shorter periods of time (1 ns) after removing translational and rotational motion of the molecule as a whole. Next, for each heavy atom, standard deviation from the average position has been calculated. It measures the mobility of atoms in analyzed time scale. Moreover, covariance matrix of atomic displacements has been calculated:

$$C_{ij} = \langle (x_i - \langle x_i \rangle) \cdot (x_j - \langle x_j \rangle) \rangle$$

to be used in PCA analysis of fluctuations in the system.

Results of analysis of higher PCA modes.

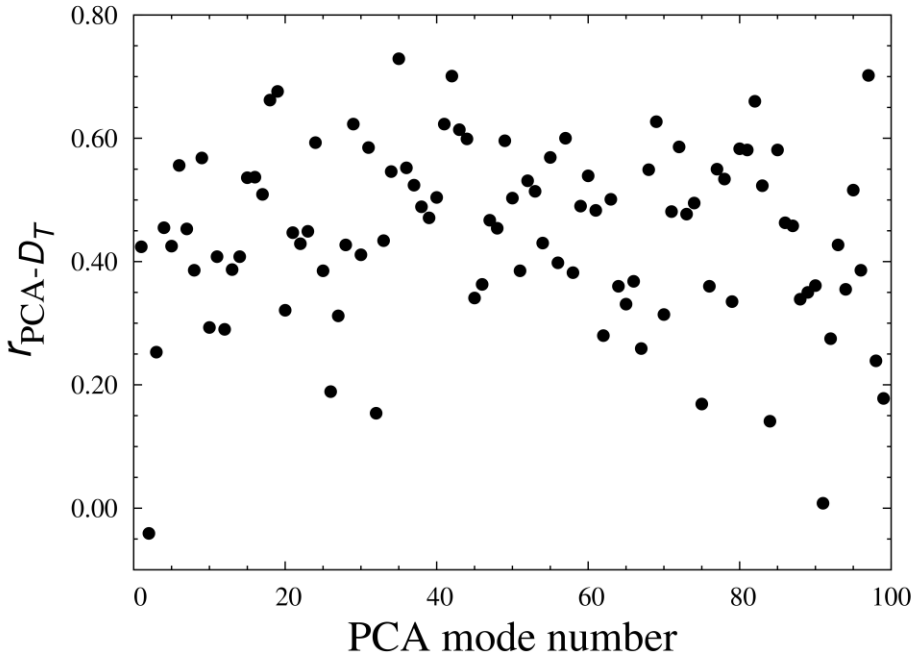

Figure S6. Correlation coefficients, $r_{PCA\text{-}DT}$, between the translational diffusion coefficient of water, D_T, and the mean mobility of the surface atoms of the kinesin belonging to various patches on the surface, for the first 99 modes obtained from PCA analysis.

Cited references

1. *The PyMOL Molecular Graphics System*, Version 1.1, Schrödinger, LLC.

2. Reference 61 in the main text

3. Teixeira, J.; Bellissent-Funel, M.-C.; Chen, S.-H. *J. Phys. Chem. Condens. Matter* 1990, **2**, SA105-SA108.

4. Kumar, R.; Schmidt, J. R.; Skinner, J. L. *J. Chem. Phys*. 2007,**126**, 204107.

5. Chau, P.-L.; Hardwick, A. J. *Mol. Phys.* 1998, **93**, 511.

6. Bernal, J. D.; Finney, J. L. *Discuss. Faraday Soc.* 1967, **43**, 62.

7. Richards, F. M. J. Mol. Biol. 1974, **82**, 1.

8. Procacci, P.; Scateni, R. *Int. J. Quant. Chem.*1992, **42**, 1515.

9. Poupon, A. *Curr. Op. Struct. Biol.*2004, **14**, 233.

10. D. Chandler, *Introduction to Modern Statistical Mechanics*, Oxford University Press, 1987, Chapter 8.

11. Reference 31 in the main text.

II. SELECTION OF PATCHES ON THE KINESIN SURFACE.

Table S1. Definition of selected regions on the protein surface.

No.	Name of surface area	Residue numbers
1	α0	20-25
2	α1a	58-60,62-64
3	α1b	66-68,70-72,74-75
4	α2b	108-110,112-114,116-121
5	α3	176-178,180-182,184-185,187-190
6	α3a	197-204
7	α4	256-257,259-260,262-263,265-270
8	α5L12	271-279
9	α5L13	281-282,284-285,287-293
10	α61	306-311,313-314
11	α62	315-318,320-321,323-324
12	α7	336-349 (for 1MKJ only)
13	β0/β0β9	4-8,326-329 (4-8 for 1BG2)
14	β1aβ1b	32-40
15	β1c	41-47
16	β1β8	9-10,12,48-50,295-297
17	β5aβ5b	154-157,163-167
18	β5L8	158-162
19	β10	330-335 (for 1MKJ only)
20	β4β5	136-143
21	β4567_1	123-129,214-216,222-223
22	β4567_2	131,168-172,209,211,213,224,226,228
23	Lβ1α0β2α1a	16-19,53-57
24	Lα0β1a	26-31
25	L5	97-105
26	L8	145-153
27	L9	191-196
28	L10	217-221
29	L11_1	233-238,250-255
30	L11_2	239-249

III. NUMERICAL RESULTS

Table S2. Values of parameters describing properties of solvation water next to selected regions on kinesin surface (for conformation 1MKJ and also for pure SPC/E water in the last line). All values are forsolvation layer 0.4 nm thick(maximum distance $R$from the atoms on kinesin surface), except from values of local ordering parameters, which were calculated for two distances (R=0.4 nm and R=0.8 nm).

In the columns from the left:

γ – local corrugation coefficient;

V – mean volume of Voronoi cell;

D_T – translational diffusion coefficient;

D_R – rotational diffusion coefficient;

η – mean value of asphericity parameter;

n_{HB} – mean number of hydrogen bonds per water molecule (formed with other water molecules);

Δn_{HB} – mean number of hydrogen bonds per water molecule (formed with other water molecules) after extracting direct influence of the shape of the surface.

s_{tra}, s_{conf} and s_{ort} – local ordering parameters;

Δs_{tra}, Δs_{conf} – local ordering parameters after extracting direct influence of the shape of the surface.

Surface area name	γ	V [nm^3]	D_T [10^{-9}m^{2}s^{-1}]	D_R [10^{11}rad^{2}s^{-1}]	η	n_{HB}	Δn_{HB}	s_{tra}	Δs_{tra}	s_{conf}	Δs_{conf}	s_{conf}	s_{ort}
									[J mol^{-1}K^{-1}]			[J mol^{-1}K^{-1}]	
	(R=0.4 nm)											(R=0.8nm)	
α0	1.228	0.0180	1.40	1.71	1.680	2.84	0.38	-14.89	-0.14	-13.07	-1.46	-14.35	-50.04
α1a	0.850	0.0178	0.91	1.47	1.646	2.45	0.46	-17.86	2.14	-11.37	-1.77	-13.10	-44.93
α1b	0.907	0.0179	1.07	1.50	1.643	2.52	0.47	-17.42	2.73	-12.00	-2.39	-13.37	-45.08
α2b	0.866	0.0184	1.13	1.74	1.658	2.58	0.55	-18.77	2.49	-11.66	-2.37	-13.23	-46.72
α3	0.874	0.0179	0.90	1.44	1.648	2.44	0.35	-18.35	1.74	-11.04	-1.41	-13.15	-46.93
α3a	0.819	0.0180	0.88	1.44	1.647	2.44	0.35	-20.10	0.41	-11.55	-2.00	-13.27	-45.27
α4	0.934	0.0181	1.01	1.49	1.656	2.63	0.59	-17.83	1.40	-12.30	-2.35	-13.75	-47.33
α5L12	1.205	0.0183	1.37	1.73	1.674	2.79	0.43	-15.76	0.24	-12.97	-1.92	-14.10	-48.91
α5L13	0.688	0.0179	0.70	1.35	1.641	2.33	0.59	-22.57	4.68	-10.84	-2.84	-12.61	-42.82
α61	0.803	0.0178	0.89	1.37	1.640	2.48	0.31	-18.56	1.52	-11.70	-1.95	-13.31	-45.72
α62	0.841	0.0182	1.02	1.55	1.655	2.52	0.69	-19.29	3.07	-11.97	-3.00	-13.29	-45.22
α7	1.235	0.0180	1.34	1.61	1.667	2.79	0.27	-14.89	0.24	-13.11	-1.56	-14.28	-49.80
β0β9	0.967	0.0180	1.07	1.58	1.662	2.66	0.53	-16.56	0.70	-12.44	-1.89	-13.59	-47.67
β10	0.798	0.0178	0.76	1.35	1.642	2.39	0.48	-21.81	2.15	-10.54	-1.83	-12.48	-42.10
β1aβ1b	1.098	0.0181	1.29	1.61	1.669	2.69	0.39	-15.78	1.07	-12.40	-1.80	-13.97	-48.53
β1c	1.088	0.0182	1.36	1.71	1.666	2.71	0.51	-16.14	1.23	-13.16	-2.82	-13.97	-47.92
β1β8	0.771	0.0179	0.82	1.39	1.640	2.37	0.63	-20.66	7.41	-11.18	-3.60	-12.47	-42.29
β5aβ5b	0.927	0.0182	1.10	1.62	1.656	2.53	0.45	-18.46	1.64	-12.02	-2.39	-13.46	-45.29
β5L8	1.209	0.0182	1.31	1.65	1.670	2.75	0.50	-16.57	0.09	-12.85	-1.85	-13.69	-47.02
β4β5	0.781	0.0180	0.96	1.52	1.644	2.53	0.68	-21.00	1.60	-12.26	-3.31	-13.24	-44.92
β4567_1	0.958	0.0179	0.96	1.41	1.651	2.51	0.33	-17.17	1.19	-11.25	-1.14	-13.35	-46.61
β4567_2	0.831	0.0181	0.84	1.40	1.643	2.42	0.69	-19.79	4.15	-11.31	-2.71	-12.58	-44.09
Lβ1α0β2α1a	0.830	0.0182	0.94	1.48	1.645	2.51	0.47	-19.89	0.80	-12.08	-2.59	-13.31	-46.06
Lα0β1a	1.016	0.0182	1.28	1.65	1.665	2.72	0.51	-16.67	1.24	-12.94	-2.78	-13.85	-47.18
L5	1.110	0.0180	1.19	1.60	1.664	2.75	0.44	-15.71	0.86	-12.75	-1.92	-13.87	-48.69
L8	0.972	0.0180	1.18	1.66	1.662	2.66	0.50	-17.09	1.31	-12.35	-2.29	-13.66	-47.40
L9	1.210	0.0182	1.34	1.81	1.676	2.89	0.51	-16.29	-0.69	-13.68	-2.43	-14.22	-49.57
L10	0.957	0.0179	1.11	1.44	1.665	2.65	0.25	-16.23	0.18	-12.19	-1.22	-13.59	-45.98
L11_1	0.772	0.0177	0.85	1.35	1.637	2.43	0.42	-20.19	1.12	-11.39	-1.95	-12.97	-43.83
L11_2	1.153	0.0180	1.40	1.66	1.671	2.83	0.30	-14.96	-0.19	-13.03	-1.35	-14.33	-50.61
SPC/E water	-	0.0187	2.75	2.38	1.725	3.27	-	-12.88	-	-15.34	-	-15.34	-54.27

Table S3. Standard deviations from mean positions of heavy atoms after removing translational and rotational movement of the whole protein (averaged over the all surface atoms in every selected region separately) for 25 ns run (*std 25*) and 1 ns run (*std 1*). The last column contains results from PCA analysis (*PCA max*).

Surface area name	*std 25* [nm]	*std 1* [nm]	*PCA max* [nm]
α0	0.211	0.138	0.122
α1a	0.144	0.104	0.038
α1b	0.158	0.109	0.039
α2b	0.138	0.102	0.030
α3	0.148	0.098	0.052
α3a	0.162	0.118	0.038
α4	0.156	0.110	0.043
α5L12	0.182	0.135	0.034
α5L13	0.108	0.078	0.027
α61	0.160	0.095	0.086
α62	0.147	0.098	0.064
α7	0.483	0.259	0.227
β0β9	0.141	0.103	0.028
β10	0.132	0.093	0.036
β1aβ1b	0.194	0.135	0.050
β1c	0.185	0.129	0.066
β1β8	0.152	0.098	0.031
β5aβ5b	0.195	0.115	0.090
β5L8	0.321	0.167	0.201
β4β5	0.137	0.093	0.046
β4567_1	0.160	0.115	0.047
β4567_2	0.125	0.090	0.026
Lβ1α0β2α1a	0.123	0.093	0.032
Lα0β1a	0.198	0.131	0.099
L5	0.181	0.134	0.017
L8	0.135	0.099	0.034
L9	0.165	0.129	0.038
L10	0.180	0.125	0.054
L11_1	0.242	0.126	0.130
L11_2	0.581	0.244	0.382