Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supplementary Information

From phage display to structure: interplay of Enthalpy and Entropy in binding of LDHSLHS polypeptide to silica.

Daniel Oliver,[a] Monica Michaelis,[a,b] Hendrik Heinz,[c] Victor V. Volkov[a] and Carole C. Perry[a,*]

[a]*Interdisciplinary Biomedical Research Centre, School of Science and Technology, Nottingham Trent University, Clifton Lane, Nottingham NG11 8NS, United Kingdom.*
[b]*Hybrid Materials Interfaces Group, University of Bremen, Am Fallturm 1, Bremen D-28359, Germany.*
[d]*Chemical and Biological Engineering, College of Engineering & Applied Science, University of Colorado Boulder, Boulder, Colorado 80309, United States.*

1. Molecular Dynamics: Positioning in respect to Interface.

To accomplish this, we take advantage that the Silicon Oxide component of the system is crystalline and stable. The slab of Silicon Oxide has two surfaces. Under Periodic Boundary Conditions the slab may slide about the Z axis of the simulation box. Then, according to the size of the box and the space left for water and for the polypeptide, there is a possibility that the polypeptide may become associated with the interface located underneath or above it. The initial placement of the polypeptide is such that it is closer to the surface which is below it. Furthermore, we confirmed that during the trajectory, the polypeptide did not reach the surface (initially designed to be above it) to establish an association (the distance is never shorter than 4 angstroms). Accordingly, for each MD snapshot (along the trajectory) we extract coordinates of H atoms which terminates Silicon Oxide bonds at the inorganic surface, which is located below the polypeptide centre of mass. However, again, under the Periodic Boundary Conditions the inorganic slab may slide about the Z axis of the simulation box. If the inorganic slab flips with its both surfaces to become above the polypeptide, we can recalculate the relative geometry about the Z-axis by proper subtraction of the box Z dimension from the Z coordinates of the H atoms which terminate Silicon Oxide surface which is assigned to be below. Next, we calculate the mean Z coordinate for these H atoms as a structural reference of the inorganic surface. In respect to this structural reference, we measure the heights (Z-components of the distances) of the selected structural moieties of the polypeptides.

2. Proximity of Na^+ cations and O^- defects.

According to the nature of classical simulations, the positions of the O^- defects are assigned to the specified sites and do not change, though, of course one should expect defect dynamics.[1-6] However, the translation motion of the polypeptide above the surface, and the dynamics of its side group partially constrained by the classical limit is compensated. At any moment, within the area of the surface underneath of the polypeptide, we may have a different number of O^- defects. Also, at any moment of time, next to the polypeptide, we would have a different number of Na^+ ions, according to the collisional diffusion and electrostatic interactions. The proximal to the polypeptide Na^+ ions may be, at the same time, next to the inorganic interface to compensate charges of O- defects. Here we wish to describe presence of three ionic species next to the polypeptide: i) number of O- defects within the surface area underneath of the polypeptide, ii) number of Na+ cations proximal to the polypeptide, and iii) the number of proximal Na+ cations but which are also next to O- defects within the surface area underneath of the polypeptide. To do this we develop the following geometric criteria.

First, since we know the coordinates of the atom of the polypeptide, we calculate the projection of the polypeptide geometric centre onto the XY plane of the simulation box $\{X_{mean}\ ,\ Y_{mean}\}_{polypeptide}$, as the mean for the two Cartesian components of its atoms. Then we collect all Na+ cations which happen to fall into the box $\{\{X_{min}$ - 1.9Å, X_{max} + 1.9Å $\},\{Y_{min}$ - 1.9Å, Y_{max} + 1.9Å $\},\{Z_{min} - 3.45Å, Z_{max} - 3.45Å\}\}$, where, for example, X_{min} is the smallest Cartesian component for the atoms of the polypeptide. This provides the statistics on the proximal to the polypeptide presence of cations. The indicated numerical extensions were attuned by practical trials to be able to collect reasonable statistics for the two types of trajectories. Second, once we accounted for the proximal Na+ cations, we recognize two aspects:

i) there negative Oxygen terminals, possibly, to interact with the polypeptide and Na+ are located at the inorganic interface, which is typically below the geometric centre of the polypeptide;

ii) ii) in respect to XY plane of the simulation box and of the $\{X_{mean}\ ,\ Y_{mean}\}$, the O^- defects can be positioned slightly more distant than proximal to $\{X_{mean}\ ,\ Y_{mean}\}_{polypeptide}$ Na+ cations.

To include such contributions, we develop a cut-off radius in the XY plane of the simulation box about $\{X_{mean}\ ,\ Y_{mean}\}_{polypeptide}$ to collect O^- defects with such. Using the coordinates of the atoms of the polypeptide, we know the geometry of the region $\{\{X_{min}\ ,\ X_{max}\},\{Y_{min}\ ,\ Y_{max}\}\}_{polypeptide}$. If there are proximal Na+ cations, we would extract analogous for $\{\{X_{min}\ ,\ X_{max}\},\{Y_{min}\ ,\ Y_{max}\}\}_{Na}$. We may calculate now all the distances between the $\{X_{mean}\ ,\ Y_{mean}\}_{polypeptide}$ and all the {X,Y} entries of both, polypeptide specific and Na+ specific set. Then we determine the cut-off radius for statistical sampling of O^- incrementing the longest distance by 1.5Å.

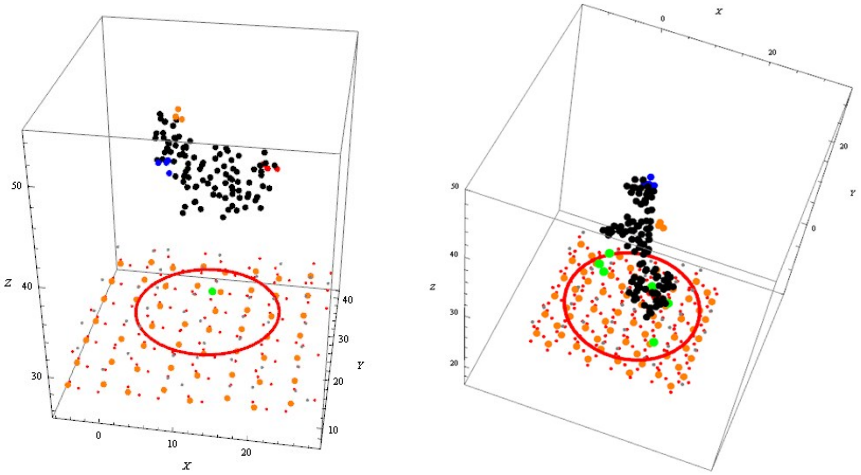

Fig. S1 *Graphical examples of the adopted sampling in the case of the LDHSLHS polypeptide.*

Additionally, as recommended by the Principle Investigator, there is one more diagnostics to be programmed and established to cover this topic. We will sort out and count such Na+ that are not only proximal to a polypeptide, but which are near (at the distance less than 2 angstroms) to the surface, as well.

3. Backbone and Side Groups Characterization

We characterize the backbone structure using Φ and Ψ backbone dihedral angles, consistent with such of the Ramachandran space. There are five pairs of the angle along the backbone of the hepta-peptide. The dihedral angle may be visualized as the angle between two planes formed by three vectors.

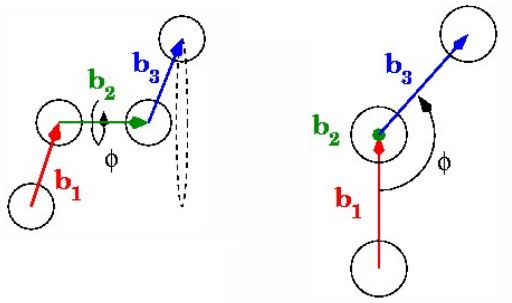

Fig. S2 *Vector Model used to calculate Φ and Ψ dihedral angles*

The extraction of dihedral angle may be conducted evaluating:

$$ArcTan[\vec{n}_1.\vec{n}_2, Cross[\vec{n}_1, Normalize[\vec{b}_2]].\vec{n}_2]]$$

where

$\vec{n}_1 = Normalize\left[Cross\left[\vec{b}_1,\vec{b}_2\right]\right], \vec{n}_2 = Normalize[Cross\left[\vec{b}_2,\vec{b}_3\right]]$

Let us consider the case of the first pair according to the first three Amino Acids LDH of LDHSLHS polypeptide, starting from the N-terminal. In the case of the Φ_1 angle, $\vec{b}_1,\vec{b}_2$ and $\vec{b}_3$ are by the coordinates of the C atom of the carbonyl of the Leucine, the C_α atom of the Aspartic Acid and the C atom of the carbonyl of the Aspartic Acid, respectively. After taking the arctangent, we subtract π from the numerical value to express the Φ_1 angle in [-2π, 2π] space. In the case of the Ψ_1 angle, $\vec{b}_1,\vec{b}_2$ and $\vec{b}_3$ are by the coordinates of the N atom of the Histidine, the C_α atom of the Aspartic Acid and the N atom of the Aspartic Acid, respectively. After taking the arctangent, we add π to the numerical value to express the Ψ_1 angle in [-2π, 2π] space.

To fit the conventional Ramachandran presentation of the numerical values of the two angles, we have to "fold" them into the [-π, π] space. This we do either by subtraction of 2π, if the angular value is more than π; or by addition of 2π, if the angular value is less than –π. However, for the correlation dynamics studies, we adopted the numerical values of the dihedral angles as they are, to avoid the 2π shifts as programmed to fold the values of the dihedral angles into the [-π, π] Ramachandran space.

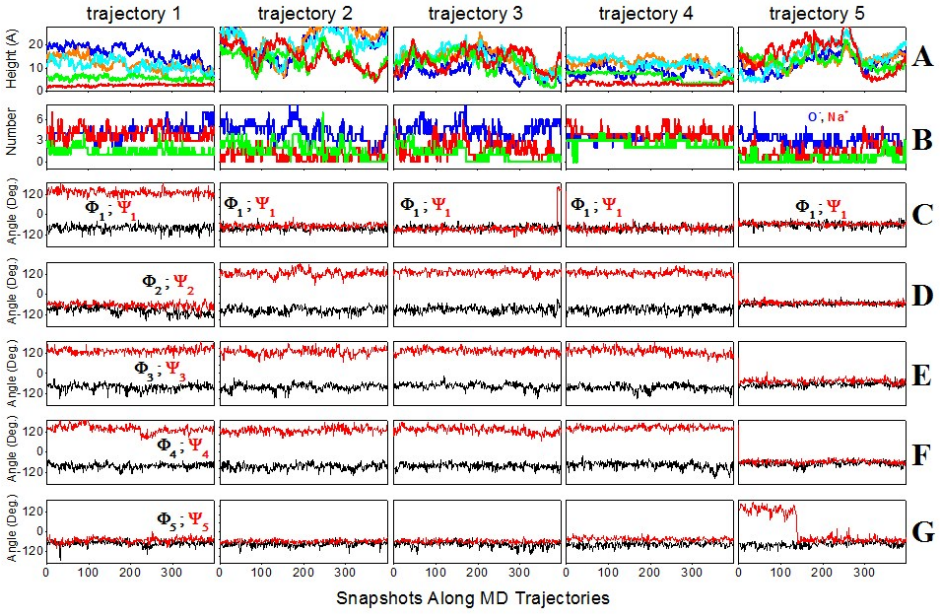

Fig. S3 *Structural Properties of LD82 columned according to five trajectories as indicated:* ***Set A:*** *Heights of N-terminal, COO^- of ASP2, centre of the side group of HSE3, centre of the side group of HSE6, C-terminal (by Blue, Orange, Cyan, Green and Red lines, respectively) above the mean (along the Z axis of the box) of H atom distribution at SiO_2 interface.* ***Set B:*** *number of Na+ and SiO^- within the cut-off radius around the polypeptide (Red and Blue lines, respectively); Green line shows number of Na+ within the cut-off radius around the polypeptide and proximal to the interface.* ***Sets C-G:*** *numeric values of {Φ; Ψ} angular pairs counted from N to C terminal.*

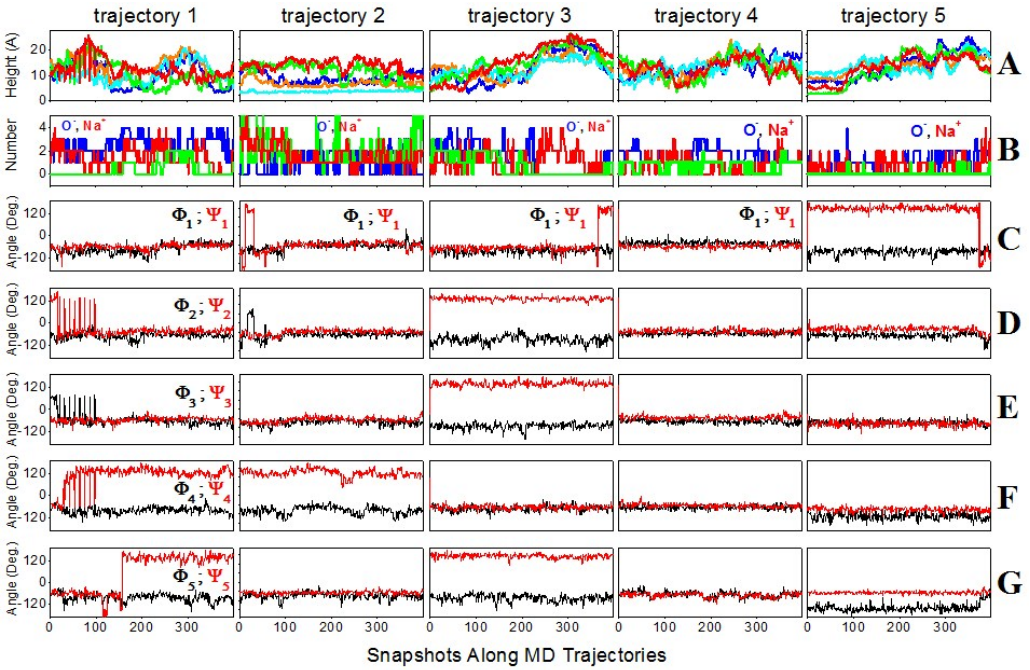

Fig. S4 *Structural Properties of LD91 columned according to five trajectories as indicated:* ***Set A:*** *Heights of N-terminal, COO^- of ASP2, centre of the side group of HSE3, centre of the side group of HSE6, C-terminal (by Blue, Orange, Cyan, Green and Red lines, respectively) above the mean (along the Z axis of the box) of H atom distribution at SiO_2 interface.* ***Set B:*** *number of Na+ and SiO^- within the cut-off radius around the polypeptide (Red and Blue lines, respectively); Green line shows number of Na+ within the cut-off radius around the polypeptide and proximal to the interface.* ***Sets C-G:*** *numeric values of {Φ; Ψ} angular pairs counted from N to C terminal.*

Additionally, in this study we extracted χ_1 and χ_2 dihedral angles to describe orientations of the side groups of Leucine, Aspartic Acid and Histidine, and χ_1 dihedral angles for Serine Amino Acids. In the case of the χ_1 angle, $\vec{b}_1$,$\vec{b}_2$ and $\vec{b}_3$ are by the coordinates of the N atom of the backbone, and the C_β and Z_γ atoms of the side group, were Z is either Carbon atom for all considered amino acids, save for Serine where it is Oxygen. In the case of the χ_2 angle, $\vec{b}_1$,$\vec{b}_2$ and $\vec{b}_3$ are by the coordinates of the C_α, C_γ and $Z_{\delta 1}$ atoms, where Z is Oxygen, Nitrogen and Carbon for Aspartic Acid, Histidine and Leucine, respectively.

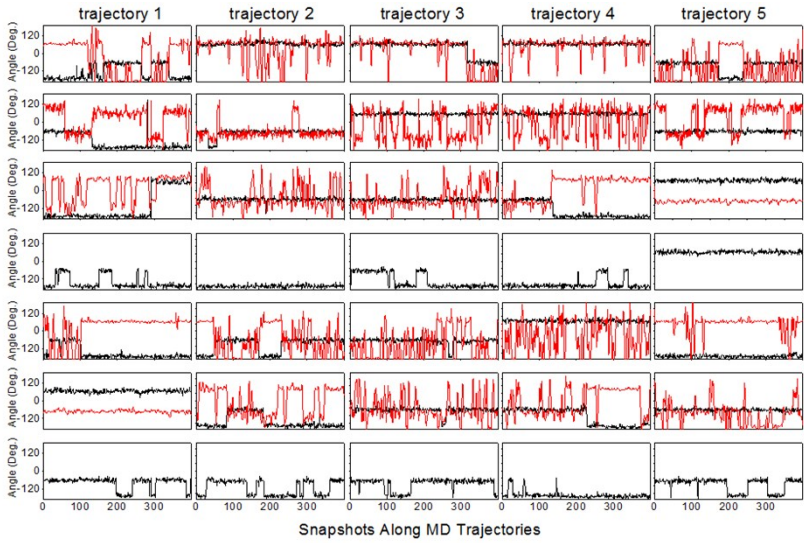

Fig. S5 *Dihedral angles χ_1 (black line) and χ_2 (red lines) for the residues from the N-terminal to the C-terminal listed in panels from top to bottom for the LD82 systems.*

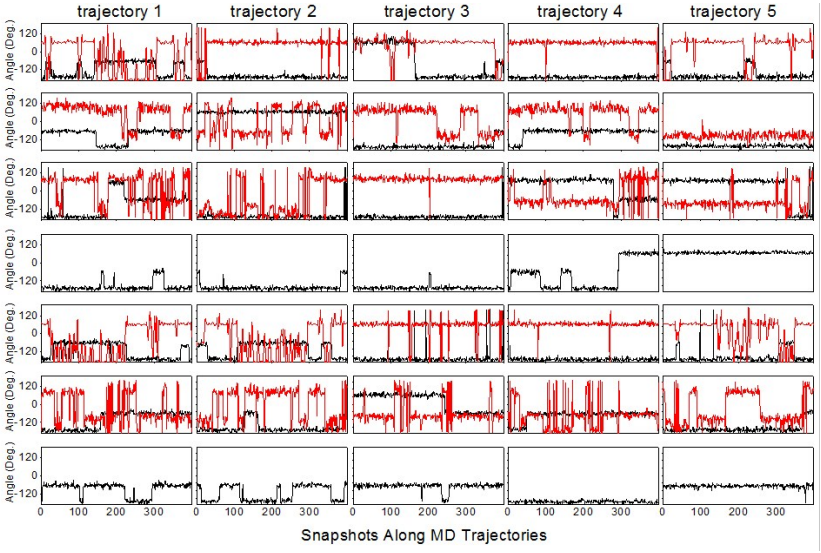

Fig. S6 *Dihedral angles χ_1 (black line) and χ_2 (red lines) for the residues from the N-terminal to the C-terminal listed in panels from top to bottom for the LD91 systems.*

4. MD studies: details on extraction of representative structural cases. Using the results of classical simulation, we extracted: positions of main referenced structural components of the polypeptides in respect to the SiO_2 surface; number of defects (Na^+ and $\equiv SiO^-$) proximal to polypeptides, and five pairs of Ramachandran dihedral angles for the five residues, see Fig. S3 and S4. The extractions are from five trajectories for the LDHSLHS polypeptide next to a moderately ionic surface, with 18% of $\equiv$Si-O$^-$ at the surface (the case LD82), and from five trajectories for a weakly ionic surface, with 9% of $\equiv$Si-O$^-$ at the surface (the case LD91) of silicon oxide: see the left and the right data sets. respectively. The

figures demonstrate the heights of: 1) N-terminal, by the mean of the Z coordinates for N atom and three Hydrogen atoms (blue line); 2) aspartic acid side group, by the mean of Z coordinates for carbon atom and two oxygen atoms (orange line); 3) histidine-3 and histidine-6 side groups, by the means of the Z coordinates of the two nitrogen atoms in each case (green and cyan lines, respectively); 4) C-terminal, by the mean of Z coordinates for C atom and two oxygen atoms connected to it (red line).

The main trends in peptide orientation in the case of LD82 systems is as follows. First, we see that trajectory 1 demonstrates that the C-terminal is very much attracted to the surface. This is contra-intuitive and even striking since it is negatively charged, and the surface is negative, as well. Initially, the N-terminal and other moieties are oriented away: the backbone is approximately 18^{o} to the normal at the surface. During the trajectory, the N-terminal slowly moves towards the surface nearly to level up with the C-terminal by the end of the simulation. At the end of the simulation, the tilt of the backbone in respect to the normal is about 73^{o}. It is interesting that the side group of histidine-3 shows a constant height of 4-5 Å in respect to the surface. Second, trajectories 2 and 3 demonstrate several rapid approaches of the N-terminal toward the surface: at the time-frames 90 and 160 in former case, at the timeframe 300 in the latter case. The C-terminal has a rapid contact with the surface at the end of the second trajectory at the time-frame 370. Third, in both trajectories 2 and 3, histidine-3 and the histidine-6 tend to follow the C-terminal and the N-terminal, respectively. At the timeframe 90, in the former trajectory histidine-3 dominates in the interaction with the surface. Both, the C-terminal and histidine-3 are next to silicon oxide at the end of the second trajectory. Fourth, trajectory 4 shows a similar (compared to trajectory 1) attraction of the C-terminal to the surface all the time. The system is close to being parallel to the surface: The N-terminal (and the moieties between the two terminals) slowly oscillate in respect to the surface. Here, again, in the beginning, the side group of histidine-3 shows a constant height of 4-5 Å in respect to the surface and, at the end of the trajectory, it moves even closer to the surface. Fifth, trajectory 5 starts when histidine-6 and the N-terminal are proximal to the surface. Yet, at the time-frame 80, the polypeptide starts drifting away from the surface for a while till the time-frame 320 when it has a brief attraction of the C-terminal with the surface.

Here, it is necessary to describe if and how the ionic neighboring species of the polypeptide in the LD82 systems affect orientation in respect to silicon oxide? The data set in Fig. S3(b) describes the proximity of O^{-} defects and Na^{+} cations to the polypeptides. It is interesting that the unexpected orientation, when the negative C-terminal is towards the surface, and the steady character of orientation dynamics in the first and in the fourth simulations correlate well with the high and equal presence of both charged species next to the polypeptide. During trajectories 2, 3 and 5, while the number of SiO^{-} defects at the surface proximal to the polypeptide is relatively high, the quantity of Na^{+} nearby is rather small. It would appear that cations diffuse to different compartments of the simulation box. This behavior correlates with a poor attraction of the polypeptide to the surface. One may notice though that either a slight increase of Na^{+} or a decrease of SiO^{-} help to enhance the association of the C-terminal with the surface: for example, the time-frame 380 in the second trajectory; the timeframe 100 in the

third trajectory; the time-frame 330 in the fifth trajectory. As a result, either the system shows an unexpected steady attraction of the C-terminal together with a moderate attraction of the histidine-3 to the surface, or the polypeptide demonstrates a relatively fast grazing dynamic moving towards and away from the surface, during which all the side groups (accept that of aspartic acid) participate in short-term attractions to Silicon oxide. The orientation dynamic is rather complex. We need more information to understand this.

What are the classic predictions in the case of the weakly ionic surface (the case LD91)? In this case, the character of the polypeptide orientation and dynamics in respect to the interface is somewhat different. In particular, the data set in Fig. S4(c) suggests that in contrast to the case LD82, most of the time the polypeptide tends to orient nearly parallel to the interface, with both terminals at about the same height above the interface. There are few moments when it gains an obvious tilt, such as at the time-frame 150 of the first trajectory. Specifically, we observe, first, that in the first trajectory, there are moments of associations of by the N-terminal and both histidines at time-frames 150-200 and 370. Second, during trajectory 2, the N-terminal, the side groups of the aspartic acid and histidine-6 are proximal to the surface. Third, trajectory 3 starts with the polypeptide wobbling next to the surface while its terminal switches in association with SiO_2. Fourth, during trajectory 4, the polypeptide is mostly in the aqueous environment, save, at time-frame 120, its C-terminal and the side group of histidine-3 get attracted to the surface for a short while. Fifth, trajectory 5 starts the structure when the histidine-3 is next to the surface while histidine-6 orients away from the surface. Consistently, the C-terminal is next to the third side group, at the time-frame 80 the polypeptide drifts away from the surface.

Can we distinguish if and how the ionic neighbor(s) of the polypeptide in the LD91 systems affects orientation in respect to silicon oxide? Here, it is instructive to see if there are any effects of the proximity of the ionic species on the orientational dynamics of the polypeptide in respect to the interface. The data set in Fig. S4(d) shows that in these simulations with a less ionized surface the effects of ionic species on the orientational dynamics are much less obvious (if present) than for the case of LD82. Along the first trajectory, when the polypeptide weakly bounces towards and away the interface, there is a moderate proximity of ionic species and the N-terminal leans towards the interface when proximal O^- are not compensated as Na^+ diffuses away. During the second simulation, both ionic species are not proximal, and the interaction with the interface has a stable character via the side groups of histidine-6, aspartic acid and the N terminal. In the beginning of the third trajectory, when SiO^- seems to be uncompensated, again there is a slight tendency for the N-terminal to drift next to the surface. However, once sodium ion concentration builds up around the polypeptide it drifts away from the surface. Finally, here, apparently, the weak ionic presence during the fourth and the fifth trajectories provides the weak associative tendencies – the polypeptide prefers to be in the aqueous phase.

In summary, the classical simulations predict that overall the polypeptide is not a strong and effective binder to the SiO_2 interface. However, when simulated, it demonstrates rather diverse orientational

dynamics in respect to a neighboring SiO_2 interface in dependence on local ionic conditions. Much to our surprise, its positively charged N-terminal does not demonstrate the initially expected propensity to associate with the mineral surface. Possibly, this is due to a steric interference with the side group of histidine-6. However, it is even more surprising that simulations indicate that the negatively charged C-terminal and the side group of histidine-3 may play a leading role in association with the surface when under a more ionic presence (LD82), when Na^+ equilibrates with SiO^- defects at the surface. Using variances of dihedral angles and proximity of selected structural moieties to the SiO_2 interface, we extract six representative structural cases for the polypeptide attached to the Silicon oxide surface, see Table 1

Table 1. *Dihedral angular pairs for the five residues (from the N-terminal to the C-terminal) to characterize secondary structures of LDHSLHS polypeptide when in bio-inorganic composites as detected along trajectories in the system LD82 and LD91.*

Residues Φ and Ψ dihedral angle (standard deviation)					trajectory: case(s)
D	H	S	L	H	
-81°(18°) 129°(30°)	-95°(23°) -64°(17°)	-82°(19°) 132°(31°)	-83°(16°) 123°(30°)	-69°(15o) -52°(15°)	LD82_1: 350
-81°(16°) -68°(15°)	-92°(19°) 128°(20°)	-82°(19°) 132°(30°)	-83°(20°) 135°(26°)	-76°(16°) -55°(15°)	LD82_4: 158
-63°(27°) -58°(16°)	-68° (18°) -45° (13°)	-65° (16°) -52° (12°)	-79° (16°) 128° (19°)	-85° (21°) 134° (17°)	LD91_1: 215, 250 and 360
-88°(18°) 148°(12°)	-62°(12°) -33°(12°)	-71°(13°) -78°(16°)	-112°(20°) -77°(13°)	-143°(15°) -57°(7°)	LD91_5: 64

5. MD studies: entropy and energy.

Having possible structural extracts, we may address the thermodynamics of binding and/or dissociation and of structural realizations of the polypeptide. Here, we evaluate configuration entropies for the polypeptide according to the approach developed by Schlitter.[7] For this task we used polypeptide coordinates of consecutive time-frames along the MD trajectories. Prior to expressing covariance matrices for eigenvalue analysis, first, each structure was positioned according to its center of mass. Next, its coordinates were projected into axes of inertia tensor specific to the structure. In Fig. S7 we present changes of energies) and configurational entropies along the MD trajectories.

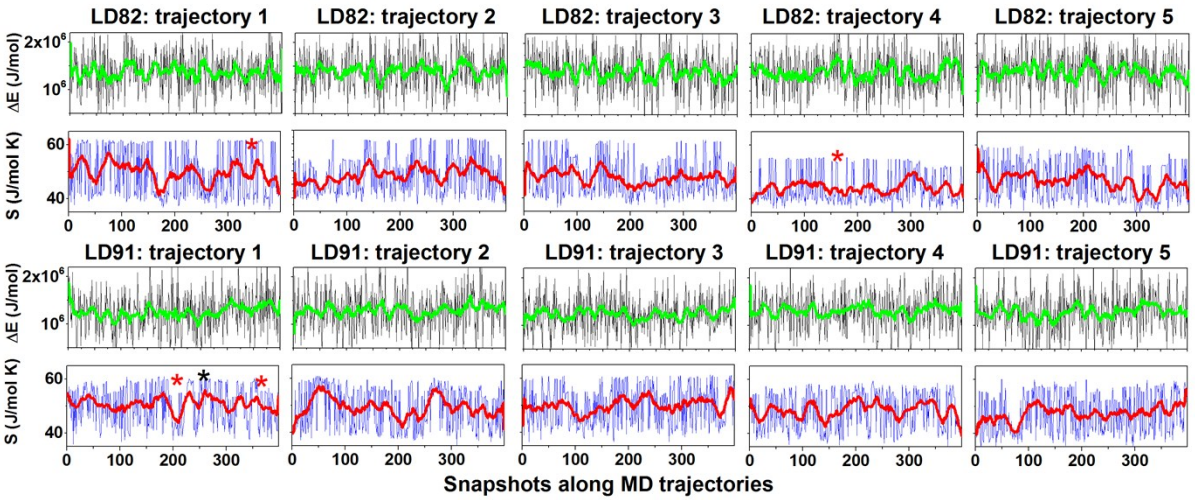

Fig. S7 *Energy changes (in respect to minimum reference) and configuration Entropy of the backbone along MD trajectories for the LD82 and LD91 structural cases, as indicated. Green and red lines are the smoothed visual guide-lens. Stars mark the snapshots where the structures of the polypeptide at interface are found to be relevant by modelling Infrared and Raman spectral responses.*

6. Volume and Surface of Convex Hull

Trying to explore the stearic role of the side groups in compacting the molecule and in water screening, we develop calculations of the volume and the surface of Convex Hull constructed using the positions of the atoms of the polypeptide as vertexes. Convex Hull of a set X of points in the Euclidean plane or in a Euclidean space is the smallest convex set that contains X.[8,9] A convex set is a region such that, for every pair of points within the region, every point on the straight-line segment that joins the pair of points is also within the region.

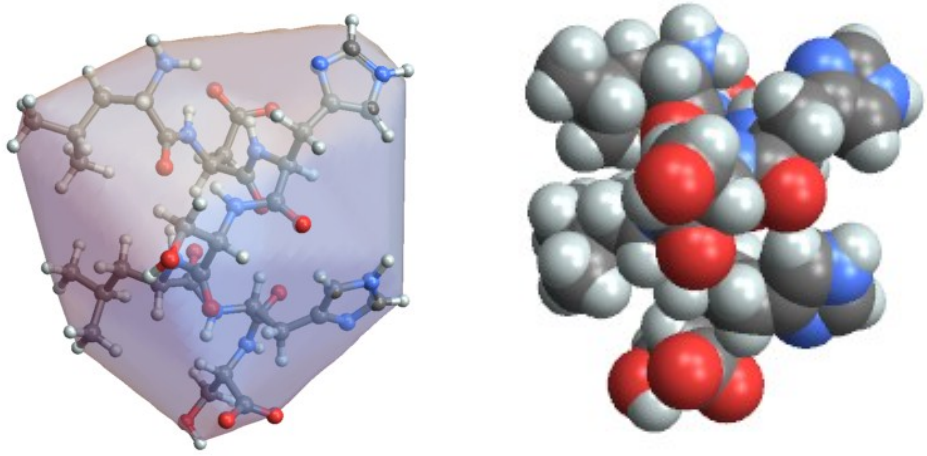

Fig. S8 *Graphical example: at the left side we show the Convex Hull of the polypeptide formed by the vertexes of the atoms. At the right side, we present a graphic presentation for the polypeptide where each atom takes its volume, which is specified by Van der Waals radius of the atom.*

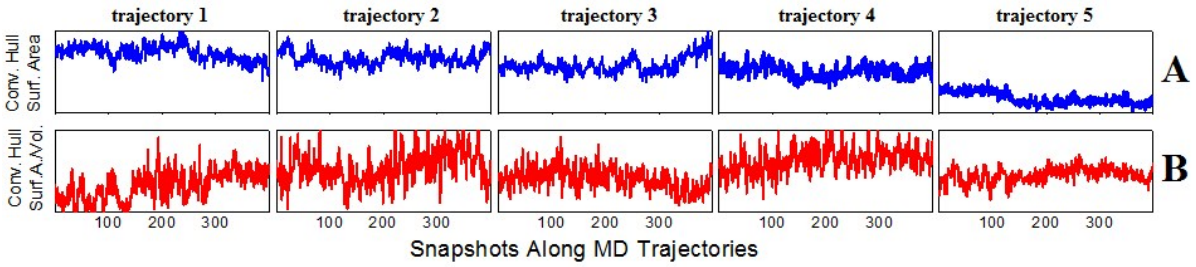

Fig. S9 *Structural properties of LD82 system.* ***Set A:*** *Surface Area of Convex Hulls.* ***Set B:*** *Ratio of Surface Area to Volume of Convex Hulls.*

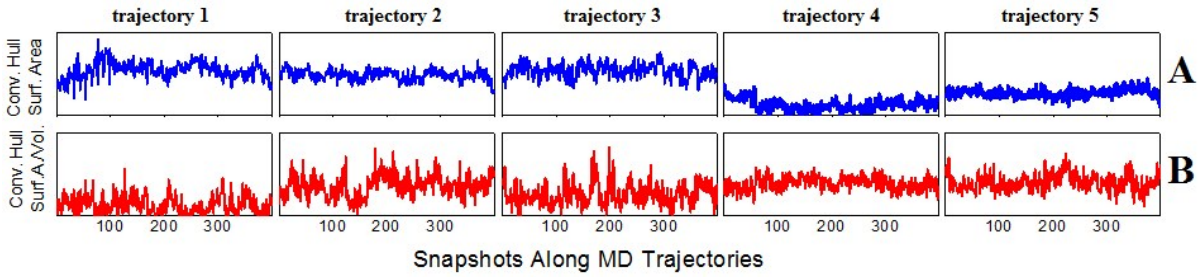

Fig. S10 *Structural properties of LD91 system.* ***Set A:*** *Surface Area of Convex Hulls.* ***Set B:*** *Ratio of Surface Area to Volume of Convex Hulls.*

Trying to understand better the stearic role of the side groups, we develop calculations of the volume and the surface of Convex Hull constructed using the positions of the atoms of the polypeptide as vertexes. Convex Hull of a set X of points in the Euclidean plane or in a Euclidean space is the smallest convex set that contains X.[8,9] A convex set is a region such that, for every pair of points within the region, every point on the straight-line segment that joins the pair of points is also within the region.

Using the vertexes of atoms of a polypeptide we may anticipate the maximal convex 3D space it occupies (at a given moment of time) according to spatial distribution of its all atoms, not the atoms of the backbone only. The volume of the Convex Hull can be compared to the volume, which is specified by Van der Waals radius of the atom. While the latter presentation would not show a time dependence, both, the surface and the volume of the Convex Hull would change according to the relative arrangements of the atoms in space.

Dynamics of the Convex Hull (by polypeptide atoms) may help to monitor the degree of the compactness of the polypeptide packing in space. Together with the information on accessibility to solvent and to polar moieties at inorganic interface, properties of the Convex Hull may instruct on the roles of the side groups and the first neighbour interactions in hierarchy of dynamical events.

Trajectory LD82_2. The secondary structure is more folded initially: this is due to the Ψ1 and Ψ2. The polypeptide has a more “globular” look and it stays stable like this. The distance between the terminals decreases by 3 angstroms, approximately.

Towards the end of trajectory, the polypeptide is driven by its Hys-3 and the C-terminal attractions towards the surface. This leads to compacting the hydrophobic side groups together. However, the side group of the Hystidin-6 stretches out and hence increasing slightly the contribution of the Convex Hull surface area in respect to the volume.

In terms of backbone structure, trajectories 3 and 4 identical to the trajectory 2. However, the properties of Convex Hulls and the orientations in respect to the surface indicate that there is an obvious variant contribution into dynamics due to the side groups. Towards the end of the third trajectory the volume of Convex Hulls become larger and the contribution of surface decreases.

Indeed, along the trajectories there is a slight tendency for the Convex Hull to become more compact. This means that the polypeptide has a propensity to form a dome-like structure where its side group can have dominant structural roles over traditional amide backbone. However, more careful examination of the positioning of the side groups above the interface indicate that the side group play does not demonstrate the expected "structural coherence". There are just short-term tendencies at the time frame 400 (of the first and third trajectories), and at the time frame 50 of the fifth trajectories. However, these are not surviving stable structures for several nanoseconds.

7. Molecular Dynamics: Accessibility to Water

There are several approaches to characterize accessibility of structural moieties of proteins, polypeptides to solvent. Water may associate with several moieties specific to a polypeptide: Amide carbonyls, Amide NH groups, OH moieties, COO^- side groups and charged terminals. It would be most instructive to understand what role the side group and the backbone play in accessibility of water to the central structural moieties of polypeptides. According to the geometry, as shown in the

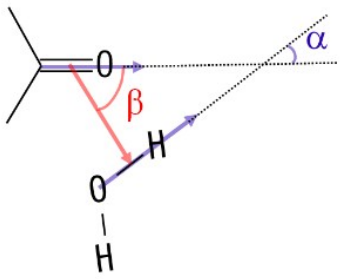

Fig. S11 *Geometric definitions used for Hydrogen Bonding analysis*

Fig., we suggest the following criteria for an optimal Hydrogen bonding (next to a carbonyl): the red arrow mark to connect the two bonds should not be longer than 5 angstroms, and $\alpha \in [90^0,160^0]$, $\beta \in [10^0,50^0]$.[9]

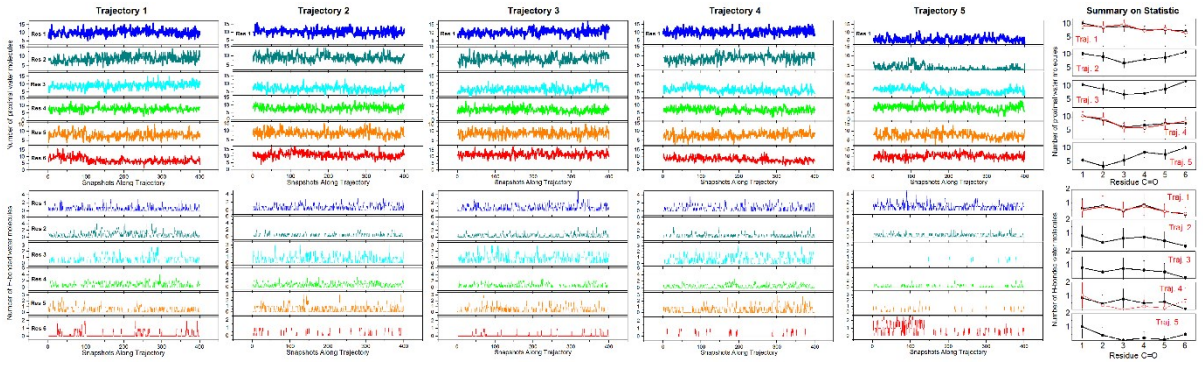

Fig. S12 *Hydration of residues for LD82. Upper set: Number of all water molecules within 5 Angstroms next to C=O moiety of the six Amino acids along the trajectories as indicated and the statistics for such at the right side. Upper set: Number of optimally Hydrogen bonded molecules (by the geometry as shown in Fig. S11) within 5 Angstroms next to C=O moiety of the six Amino acids along the trajectories as indicated and the statistics for such at the right side.*

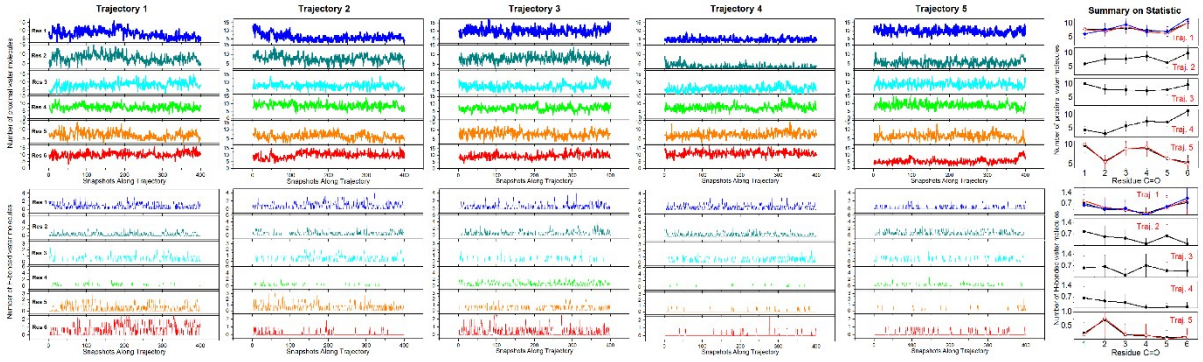

Fig. S13 *Hydration of residues for LD91. Upper set: Number of all water molecules within 5 Angstroms next to C=O moiety of the six Amino acids along the trajectories as indicated and the statistics for such at the right side. Upper set: Number of optimally Hydrogen bonded molecules (by the geometry as shown in Fig. S11) within 5 Angstroms next to C=O moiety of the six Amino acids along the trajectories as indicated and the statistics for such at the right side.*

8. Calculations of the frequencies of C=C stretching modes specific to neutral and positively charges Histidines.

The classic simulations for the project were conducted for Histidine side group where only one nitrogen, the one which is more distant to the linker, was protonated. At the same time, practically, under neutral

pH (pD) we may have this side group to have second nitrogen to be protonated as well. To anticipate the effect of deuteration of both Nitrogens on calculated Raman responses in the frequency range from 1500 – 1650 cm^{-1}, we conducted a series of DFT calculation on various conformers of methylketon-methylamine-Histidine dipeptide, as shown in Fig. S13 under, where the side group was either single or double protonated (or deuterated).

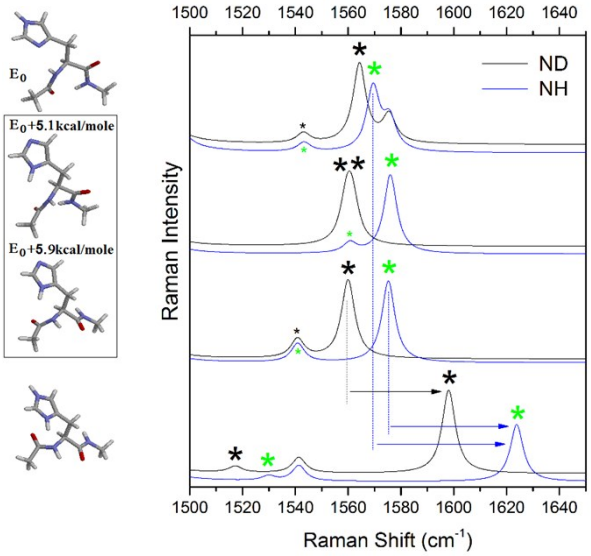

Fig. S14 *Left: structure at the top represent the lowest energy conformer methylketon-methylamine-Histidine dipeptide in vacuum where the side group is single protonated (or deuterated); two boxed structures represent higher energies single protonated (or deuterated) polypeptide, where there side groups are similarly twisted (comparing to the top structure) but the backbone is either folded (the upper structure in the pair) or unfolded (the lower structure in the pair); structure at the bottom shows the dipeptide where the side group is double protonated and the backbone is unfolded. Right: Raman spectra arranged (from top to bottom) to be parallel to the indicated structures at the left; blue and black lines shows responses for the molecules where the Nitrogens of the side group were either protonated or deuterated, respectively; large and small start indicate the normal mode which involve either strong or weak participations of the atoms of the side group in a vibration. The scaling factor is 0.97.*

The results of DFT studies indicate that under various conditions, upon addition of a second proton (deuterium), there will be one dominant mode to experience a higher frequency shift. In the case of double protonation, the shift is of 50-55 cm^{-1}, while in the case of double deuteration the shift is 35-38 cm^{-1}.

In result, in our studies we adopt 38 cm^{-1} shift to higher frequency for the imidazole mode specific to the ring C=C stretching upon addition of the second proton to the Nitrogen which is next to the linker. Assuming that half of Histidines in our cases may be completely protonated, we calculate

spectra for polypeptide (either next to silica or alone in aqueous environment) by averaging spectral responses of polypeptide where Histidine side groups are single and double protonated. The spectral response for the latter case is anticipated by giving 38 cm^{-1} shift to higher frequency for the imidazole mode specific to the ring C=C stretching. In our discussion, for the purpose of clarity, when we describe the normal modes specific to Histidines, we provide the non-scaled frequencies of the normal modes of interests and, in brackets, we provide both, scaled and shifted frequencies.

9. DFT studies: nature of the Normal Modes of in polypeptide-silica composites

We develop our structural analysis comparing spectral dispersions of predicted Infrared and Raman activities with such detected in experiments in a frequency range from 1350 to 1750 cm^{-1}. This spectral range is too wide to use one scaling factor for the calculated normal modes. To address this systematically, we explored and developed a linear scaling function $1.0909^{\omega_{nm}}$ - 230 to map the frequencies of the calculated normal modes, $^{\omega_{nm}}$, in the indicated spectral range. Consistently, in our discussion, whenever we provide a calculated frequency, we show its value as calculated and, next to it, in brackets, we provide the scaled frequency according to the scaling function. The numbers in brackets correspond to resonance in calculated spectral as we present in figures in this article.

We will start our DFT based analysis reviewing theoretical predictions for the structural case LD82_1:350. Angular mapping in the corresponding Ramachandran space (see Fig. 3 in the main text) indicates that the polypeptide residues (one after another) switch their locations in two different angular regions. We may consider this as a random and rather extended (not far from being linear) secondary structure. All polar moieties of the structure are relatively well hydrated, save C-terminal: at average: there are 8-10 proximal water molecules within 5Å next to each Amide unit but 6 waters next to C-terminal. The reason for the slightly limited water access to this moiety is that it orients towards the inorganic interface.

Let us describe the nature of the normal modes and their optical activities in Amide spectral range for this structure, case LD82_1:350, when it is next to SiO_2, see Fig. S15. First, theory predicts that at the higher frequency side of the Amide I manifold we should expect two infrared active states at 1749 (1678) and 1740 (1668) cm^{-1}, which are localized on Amide carbonyls of LEU_1 and SER_4, respectively. The next, lower frequency, mode at 1721 (1648) cm^{-1} should demonstrate most intense both, Infrared and Raman. We anticipate it to be delocalized over fourth, fifth and sixth units. The following calculated resonance at 1701.2 (1626) cm^{-1} involves the same but antisymmetric activities of the fourth and fifth units and shows a larger delocalization to involve the antisymmetric stretching of the C-terminal, as well. At the lower frequency side of Amide I region, theory predicts Infrared active resonances at 1695 and 1685.05 (1619 and 1608) cm^{-1} which are specific to the Aspartic acid Amide I and the antisymmetric stretching of the C-terminal delocalised on the previous two Amide 1 groups,

respectively. The next resonance at 1681 (1603) cm^{-1} is characteristic to the Amide I of the HIS_3 residue. In the spectral region 1550-1580 (1450-1500) cm^{-1} theory anticipates normal modes where a decrease of CNH angle is admixed with a stretching out C-N bond and a decrease of N-C_α bond length – characteristic signatures of Amide III. While the predictions for the Amide I region fit well to the results of FTIR, the predictions in the spectral range specific to Amide III are inconsistently blue shifted comparing to the experimental results. Finally, here DFT suggests that the C=C stretching modes of both Histidine side groups (residues 3 and 6) should demonstrate resonances at 1599 (1515.1+38 = 1553.1) and 1602 (1517.9 + 38 = 1555.9) cm^{-1}, respectively. For this structure, theory fails to predict the observed Raman intensity at 1610 cm^{-1}.

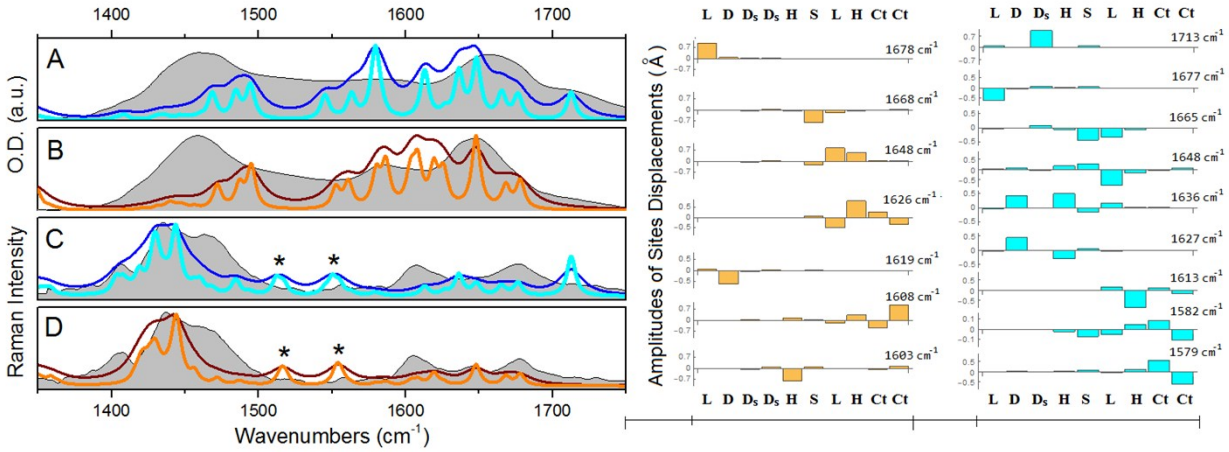

Fig. S15 *Panels A and C: calculated Infrared and Raman spectra for the structural case LD82_1:350, respectively, when in aqueous environment, assuming 7 and 20 cm^{-1} line-width (cyan and blue lines) and compared with the experimental results given as grey backgrounds. Panels B and D: calculated Infrared and Raman spectra for the structural case LD82_1:350, respectively, when at SiO_2 interface in aqueous environment, assuming 7 and 20 cm^{-1} line-width (orange and brown lines) and compared with the experimental results given as grey backgrounds. The stars marks indicate Raman responses of C=C Histidine rings stretching modes as discussed previously. At the right, Amplitudes of C···O displacements of the Amide I carbonyls and carboxyl moieties of the Amino Acids as indicated by the letter labels for the selected normal modes (see the scaled eigen-frequencies): Ds and Ct labels indicate the C···O displacements of the side group of the Aspartic Acid and of the C-terminal.*

When in aqueous environment the IR active vibrations states at 1780.9 (1712.8) and 1747 (1676.5) cm^{-1} are specific mainly to carbonyl stretching of the COOH side group of Aspartic acid and to Amide I of LEU_1, respectively. Here, we see that the anticipated state at 1780.9 (1712.8) cm^{-1} demonstrate too intense Raman activity comparing to the experiment. The pair of resonances at 1737.5 (1665) and 1721.8 (1648) cm^{-1} belong to the Amide I modes delocalized from ASP_2 till HIS_6. However, regardless the anticipated delocalizations, the symmetric and antisymmetric vibration mixings of the

contributions of SER_4 and LEU_5 dominate in these two modes, respectively. In general, these two transitions resemble in character the predicted vibrations at 1721 (1648) and 1701.2 (1625.8) cm^{-1} for this structure when next to SiO_2, though when in water the Amide I modes are more delocalized. Next, the set of resonances at 1711 (1636) and 1701.2 (1627) cm^{-1} belong to the Amide I modes delocalized over the residues 2-4, where symmetric and antisymmetric mixed contributions of Amide I carbonyls of ASP_2 and HIS_3 dominate, respectively. The lower frequency subset in the Amide I spectral range accounts three resonances at 1689.5 (1613), 1661 (1582) and 1658.5 (1579) cm^{-1} which are specific to: a) Amide I stretching of HIS_6 strongly admixed with COO^- asymmetric vibration of C-terminal; b) vibration activity delocalized over residues 3, 4, 5 and 6; c) COO^- asymmetric stretching of C-terminal, respectively. The first two states resemble such at 1701 (1626) and 1685.05 (1608) cm^{-1} for such structure but when polypeptide is next to silica. Further below, the calculated Raman spectrum accounts 1596 (1512 + 38 = 1550) and 1601 (1516 + 38 = 1555) cm^{-1} resonances of the C=C stretchings of the imidazole rings of HIS_3 and HIS_6, respectively. Overall, if we compare to experimental results, we can state that DFT calculation for this structure in aqueous environment yields inconsistently intense Raman activity at 1781 (1713) cm^{-1}, fails to explain Raman and Infrared activities at 1610 cm^{-1} and does not provide frequencies to explain optical density in FTIR in the spectral region 1430-1500 cm^{-1}. It is important to notice, however, that DFT calculations of the normal modes for this structure indicate that removal of silica and protonation of the Aspartic Acid side group make noticeable impact on the nature of the normal mode – overall, they are more delocalized in the aqueous case.

The next structural case is LD91_1:250. Exploring its structural expression in the angular space of Ramachandran plot, see Fig. 3 in the main text, we see that more than half of its residues, starting from N-terminal, participate in formation of a complete turn, where dihedral angles of residues are proximal to helical. Nonetheless, the other amino acids (LEU_5 and HIS_6) depart from the helical architecture to give C-terminal an extended unstructured (nearly linear) character. The side group of HIS_6, which belongs to this unstructured segment, makes the contact with the inorganic interface – Hydrogen of a proximal OH moiety of a silanol group points toward Nitrogen of the NH moiety of the imidazole ring. Regardless the contact, overall the polypeptide sticks away from the surface. The helical part is oriented such that Amid 1 carbonyls of LEU_1 and of LEU_5 are oriented inside or away from the aqueous environment. In result, while around any other amide unit there are 6-10 water molecules within 5 Å cut-off radius, next to LEU_1 and LEU_5 units there are 5 water molecules – these residues are slightly screened. Overall, this structural case does not look like it is in a binding "mood". On the other side, since it is next to the interface it may help us to understand better on structural scenarios which may lead to engagement or disengagement.

Let us review the nature of the normal modes for this structural case LD91_1:250, when next to SiO_2 interface, see Fig. S16. First, the anticipated IR activities account intense absorption at 1770 (1701) cm^{-1}, which is specific to the Amide I unit of HIS_6. Next, there is a quartet of the following states: a weak infrared resonance at 1750 (1679) cm^{-1} and intense transitions at 1747.5 (1676) cm^{-1},

1741 (1669) cm^{-1}, and 1731.4 (1659) cm^{-1}. These Amide I modes are very delocalized over LEU_1, ASP_2, HIS_3 and SER_4 residues. Apparently, the extend of delocalization is not only because of through bond coupling, but, also, through space due to a favourable relative orientation of polar moieties according to the helical structure at the side of N-terminal, see Fig. 3 in the main text. In this spectral region, beside strong infrared absorption, theory anticipates very strong Raman activity at 1747.5 (1676.4) cm^{-1}, which does not agree well with the experimental data.

The set of calculated vibrations in the lower frequency spectral region accounts localized contribution due to the LEU_5 carbonyl, strong IR asymmetric stretching of the C-terminal and the strong IR asymmetric stretching of the Aspartic Acid at 1712.5 (1638.2), 1696 (1620.5) and 1692 (1616.0) cm^{-1}, respectively. For this structure, C=C stretchings of the Histidine rings of the third and the sixth residue contribute into very strong Raman resonances at 1605.5 (1521.5 + 38 = 1559.5) and 1610.8 (1527.2 + 38 = 1565.2) cm^{-1}, respectively. Additionally, theory anticipates intense Raman activities due to the CH bending modes in the spectral range centred at 1530 (1440) cm^{-1}. The lower frequency transition of this subset fits well the experimentally observed Raman intensity at 1405 cm^{-1}. Overall, both calculated IR and Raman responses of the structure demonstrate some resemblance with the experimental results. We see, however, that the observed Raman scattering at 1610 cm^{-1} is under accounted and the calculated absorption does not match the observed experimental results in the spectral region 1430-1470 cm^{-1}.

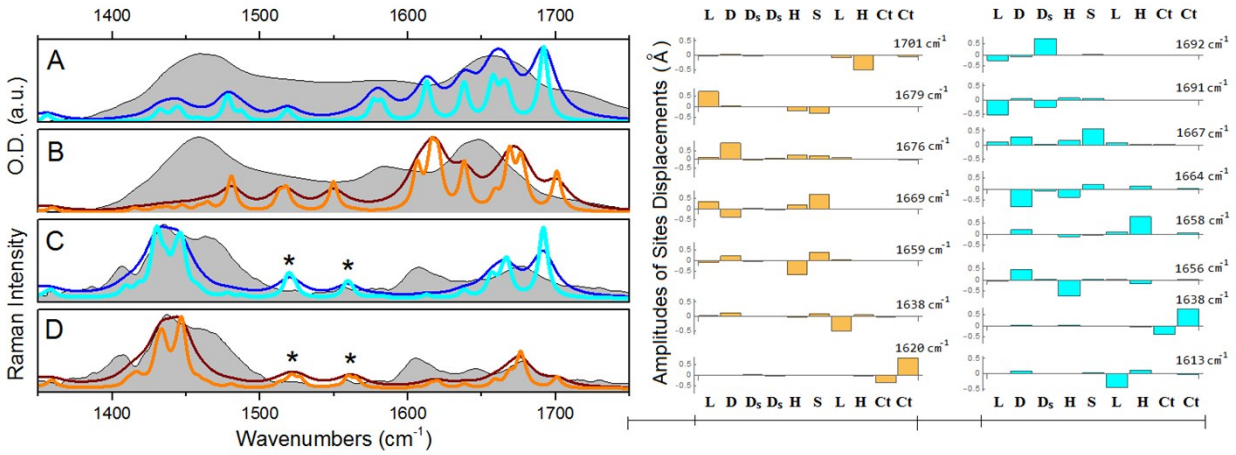

Fig. S16 *Calculated Infrared and Raman spectra in comparison to the experimental results for the structural case LD91_1:250. The definitions are as in Fig. S15.*

When in aqueous environment, for the structure theory anticipates doublet of resonances at 1762 (1692) and 1761 (1691) cm^{-1} due to COOH of the Aspartic acid admixed with Amide I carbonyl of LEU_1 with antisymmetric and symmetric characters, respectively. These differ from the higher frequency modes of the Amide I subset of the structure next to silica. This is due to the adopted protonation of the side group of the Aspartic acid which provides the COOH moiety to couple with the

carbonyl of LEU_1 Further down the energy ladder, we have resonances at 1739 (1667), 1736 (1664), 1730.64 (1658.0) and 1728.72 (1656) cm^{-1} due to mixings of the contributions of the carbonyls of the residues 2-6. These are comparable to the states 1747.5 (1676), 1740.6 (1669), 1731.4 (1659) and 1712.5 (1638) cm^{-1}, when the polypeptide of the selected structure is next to silica. At the lower frequency side of the Amide subset, for the selected structure in aqueous environment theory anticipates resonances at 1712 (1638) and 1689 (1613) cm^{-1} specific to COO asymmetric of the C-term and Amide carbonyl of LEU_3. These correspond well (save the order at the energy ladder) to two resonances in the same spectral range calculated for this structure when next to silica. The calculated Raman spectra demonstrate the characteristic to C=C stretchings of the rings of the HIS_3 and HIS_6 resides at 1604 (1520 + 38 = 1558) and 1607 (1523+ 38 = 1561) cm^{-1} respectively. Overall the calculated spectra for the structure in aqueous environment resemble those when this structure is next to silica interface. Consequently, the aqueous case lacks Raman intensity at 1610 cm^{-1}, demonstrates a mismatch in 1650-1670 cm^{-1} spectral range comparing to the results of FTIR studies, and a complete disagreement for infrared absorption in the spectral region 1430-1470 cm^{-1}. Here, also, we wish to notice that, as in the previous case, DFT calculations of the normal modes indicate that upon removal of silica and protonation of the Aspartic Acid side group the Amide I normal mode become slightly more delocalized.

Next considered structural case is LD91_1:360. According to MD, it is evolutionary related to the previous case. Exploring the representation of its backbone in Ramachandran plot, Fig. 3 in the main text, we see that, comparing to the previous case, the helical motive on the side of N-terminal is more folded about Ψ angles, while the backbone on the side of C-terminal is more unfolded in respect to Φ dihedral angles. Also, in contrary to the previous case, here, both Histidines stick away from the surface, the main axis of the polypeptide is parallel to the inorganic surface, the positively charged N-terminal is next to SiO^- defect, and Oxygen of the carboxyl of C-terminal is favourably hydrated with an aqueous cluster associated with Silanol moieties. On the other side, such association still does not look determined – oriented toward the surface the side groups of both Leucines cannot help an effective and lasting association, see Fig. S4. Comparing to the previous structural case, here, MD predicts that water screening for the first residue would increase slightly: see Fig. S13.

Let us account quantum theory predicted normal modes for this structural when at silica interface in the spectral range specific to Amide, see Fig. S17. At the higher frequency side, theory anticipates resonance at 1752 (1682) cm^{-1} which is due to the Amide carbonyls' stretching of HIS_6 and LEU_5. It is to show the strongest Amide I Raman scattering in Amide I spectral region. At the same time its IR should be rather weak. Further down (in energy), theory predicts for this structure a triplet of resonances at 1738 (1666), 1735.2 (1663) and 1731 (1659) cm^{-1}, which involve the same Amide I carbonyls of LEU_1, ASP_2, HIS_3, SER_4, LEU5 and HIS_6. Even though the same residues are involved, the normal modes are anticipated to combine stretchings of their Amide I carbonyls under different phase relations. The anticipated strong delocalizations in this spectral region resemble such in the

previously described structural case - LD91_1:250 next to silica. Both structures demonstrate complete turns of helical character at their N-terminals, see Fig. 3 in the main text. Consistently, their spectral responses in Amide I resemble such that would be expected from helices. Among the listed above transitions, the vibration at 1738 (1666) cm^{-1} is predicted to show the strongest IR absorption in the manifold of Amide 1, and a relatively strong Raman scattering. The resonance at 1735.2 (1662.9) is calculated to show moderately strong both, Infrared and Raman responses.

At the lower frequency side of the Amide I, here, theory predicts vibration states at 1722 (1649), 1713 (1639), which reflect SER_4 contribution (with small participation of all previous three Amide I units at the side of N-terminal) and antisymmetric stretching of Amide I carbonyls of LEU_1 and ASP_2 residues, respectively. DFT anticipates moderately strong Infrared absorption and Raman scattering for the former one. Finally, here, the calculated transitions at 1687.59 (1611) cm^{-1} and 1674 (1596.7) cm^{-1} are specific to the antisymmetric stretchings of the side groups of Aspartic acid and of C-terminal, respectively. Additionally, the calculated Raman spectra demonstrate characteristic to C=C stretchings of the rings of the HIS_3 and HIS_6 resides at 1614 (1530.7 + 38 = 1568.7) and 1622.90 (1540.4 + 38 = 1578.4) cm^{-1}, respectively. Overall, the calculated spectra show some resemblance with the experimental results, save the lack of Raman scattering at 1610, slight mismatch in the spectral region of 1630-1670 cm^{-1}, and a strong disagreement in the spectral region from 1430 – 1500 cm^{-1} – frequencies of the calculated Amide III normal modes are blue shifted.

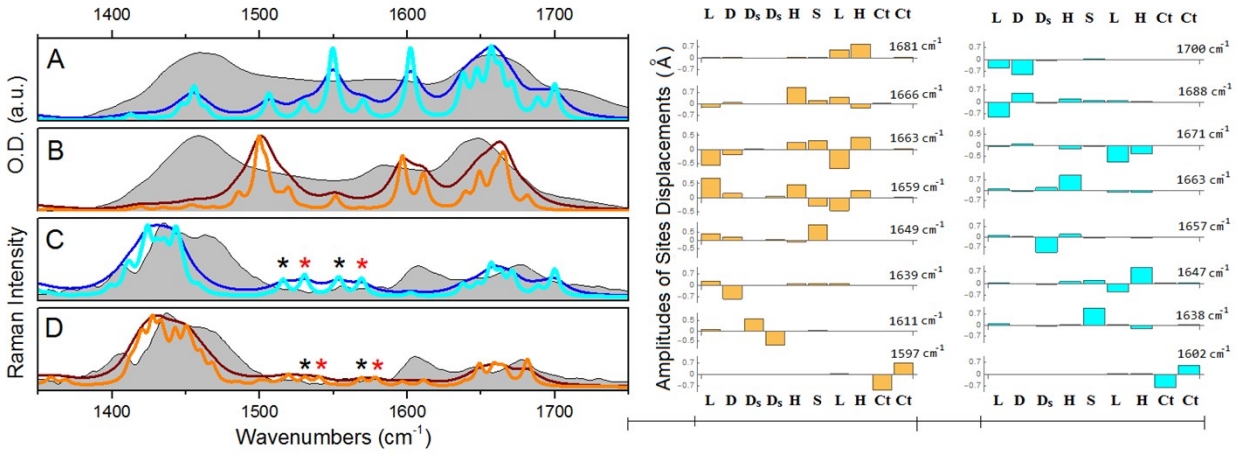

Fig. S17 *Calculated Infrared and Raman spectra in comparison to the experimental results for the structural case LD91_1:360. The definitions are as in Fig. S15.*

In aqueous environment, for the structural case LD91_1:360, theory predicts at the higher frequency side of the Amide I manifold resonances at 1769 (1700) and 1758 (1688) cm^{-1}. These transitions represent symmetric and antisymmetric carbonyl stretchings of LEU_1 and ASP_2, respectively. Both calculated activities demonstrate intermediate intensity Infrared absorption, and the higher energy resonance shows intense Raman scattering. Next, in the lower frequency range, theory calculates strongly separated doublet at 1742 (1671) and 1720 (1647) cm^{-1} which is due to symmetric

and antisymmetric stretchings of the fifth (LEU_5) and sixth (HIS_6) residue with a slight admixture of the first three residues (from N-terminal) in both activities. The first transition is anticipated to show intense both, Infrared and Raman, while the second resonance would be expected to show moderate Infrared absorption. The following pair of calculated transitions at 1735 (1663) and 1729 (1657) cm^{-1} represent HIS_3 stretching with slight participation of COOH side group of ASP_3, and COOH stretching of the side group of ASP_3 with slight admixing of HIS_3 Amide I carbonyl, respectively. Both transitions demonstrate strong Infrared and Raman, and this is consistent with the experimental observations. At the lower frequency side of Amide I region, there are two transitions at 1712 (1637.8) and 1680 (1602.2) cm^{-1} which are due to Amide I carbonyl of SER_4 and C-terminal COO antisymmetric stretching. These resonances are calculated to show intense Infrared absorption. This agrees with the experimental results. On the other side, rather weak Raman activity of the two vibrations does not correspond well to the results of Raman experimental studies. Finally, here, the calculated Raman spectra demonstrate the characteristic to C=C stretchings of the rings of the HIS_3 and HIS_6 resides at 1614 (1530.7 + 38 = 1568.7) and 1600.3 (1515.8 + 38 = 1553.8) cm^{-1} respectively, It is important to add here, that anticipated Raman activities for this structure in water cannot explain the experimentally detected scattering at 1610 cm^{-1}. Overall, the predicted spectral responses of the structural case LD91_1:360 when in aqueous environment resemble the experimental results: there is a reasonable match in Infrared absorption for Amide I and there are two Amide III modes specific to HIS_3 admixed with the COH bending of SER_4 side group to contribute favourable in the spectral range 1430-1500cm^{-1}. On the other side, theory for this structure does not explain observed Raman activity at 1610 cm^{-1}. Finally, here, it is interesting to notice that in contrary to two previous cases, DFT calculations of the normal modes indicate that upon removal of silica and protonation of the Aspartic Acid side group the Amide I normal mode become somewhat more local.

Let us describe theory predictions for the case LD82_4:158 at SiO_2 interface. In this structure, the central three residues (HIS_3, SER_4 and LEU_5) occupy the Ramachandran region which would correspond to a β-turn like structure but slightly more folded. At the same time, the terminal residues (ASP_2 and HIS_6) "twist out" in a helical fashion, see the corresponding Ramachandran plot in Fig. 3 in the main text. MD studies indicate that N-terminal and C-terminal residues should be relatively well hydrated (10-7 water molecules within 5Å next to them), while HIS_3 and SER_5 would be partially screened from water (3-5 water molecules within 5Å from the carbonyls): see Fig. S12.

At the higher frequency side of the Amide I spectral range, DFT predicts for this structure resonances at 1767 (1698), 1744 (1673), and 1730 (1657) cm^{-1}, see Fig. S18. These should show relatively strong Infrared and Raman intensities. The first resonance is the signature of the carbonyl of SER_4 with small participations of the carbonyls of the third and fifth residues. The second vibration is mainly due to the symmetric stretching of the Amide I carbonyls of LEU_1 and ASP_2. The third vibration is localized on the carbonyl of HIS_6. Further, down the energy ladder, quantum theory anticipates a doublet at 1729 (1656) and 1726 (1653) cm^{-1}, which would involve delocalised over LEU_1, ASP_2, HIS_3

and SER_4 residues Amide I carbonyls vibrations. The higher frequency mode is calculated to show intense both Raman and Infrared, while the lower frequency one is anticipated to show weak contributions in both signals. Next are the resonances at 1689 (1612), 1669 (1591) and 1639 (1558) cm^{-1}. The first vibration is strongly localized on the Amide 1 carbonyl of LEU_5. It demonstrates an intense Infrared absorption and a very strong Raman scattering. Weak Raman activity in this spectral range would fit well the results of Raman experiment. The other two vibrations are specific to the side group carboxyl of ASP_2 and to the carboxyl of C-terminal, respectively. Both carboxyls are predicted to show intermediate intensity Infrared absorptions, which would fit to the experimental observation. Next, DFT predicts strong Raman activities for C=C stretchings of the Histidine imidazole rings at 1601 (1516 + 38 = 1554) and 1600 (1515 + 38 = 1553) cm^{-1} for HIS_6 and HIS_3 residues, respectively. Here, it is interesting to note that LD82_4:158 at SiO_2 interface is predicted to show the best agreement, so far, with the FTIR and Raman experimental results in the spectral region 1430-1500 cm^{-1}, which is specific to Amide III. In the current case, the better results are due to Amide III activities of LEU_1 and ASP_2 residues, where both NH groups orient toward well-hydrated COO^- moiety of the second residue. Interesting to notice that the structural case LD82_4:158 is the only, where the main central set of residues structurally belong to β-turn like region in Ramachandran space, see Fig. 3 in the main text. Also, only in this structural case, theory suggests that the distinct shoulder observed in Raman at 1400 cm^{-1} may be assigned to two CH_2 scissor modes at 1484 (1389) and 1501 (1408) cm^{-1} of the side groups of ASP_2 and HIS_3, respectively. Overall, the anticipated spectral properties of this structure resemble the most the experimental results.

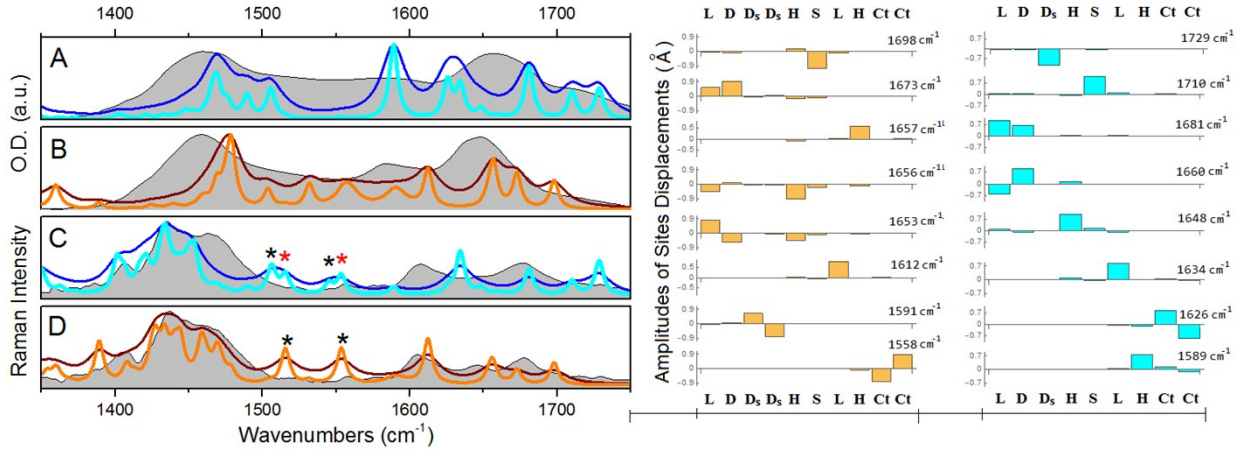

Fig. S18 *Calculated Infrared and Raman spectra in comparison to the experimental results for the structural case LD82_4:158. The definitions are as in Fig. S15.*

Let us describe the anticipated optical responses of LD82_1:158 when in aqueous environment. At the higher frequency region of the carbonyl manifold we predict resonances at 1795 (1729) and 1779 (1710) cm^{-1}. The former vibration localizes mainly on the COOH moiety of the side group of ASP_2

Residue. The latter one is mainly specific to the carbonyl of SER_4 residue. Both resonances are anticipated to show moderate infrared absorption, what would agree with experiment. At the same time, the higher frequency resonance is predicted to show Raman intensity to be larger than that observed in experiment. Next, in the lower frequency range, at 1752 (1681) and 1733 (1660) cm^{-1}, we observe symmetric and antisymmetric doublet of carbonyls stretchings of LEU_1 and ASP_2 with admixture of the contribution from the Amide I of HIS_3. The two resonance are analogous to the vibrations at 1744 (1673) and 1726 (1653) cm^{-1} for this structure when at silica interface. In the aqueous environment, DFT theory anticipates the higher frequency resonance to show Infrared and Raman of strong intensities, while the lower frequency resonance to have weak optical activities. The decrease of infrared absorption for the second resonance may play a significant role in comparison of theoretical predictions and experimental results. In the lower frequency range of the Amide I transitions we observe a resonance at 1721.9 (1648) cm^{-1} and a transition at 1709 (1634) cm^{-1}. These are specific to the carbonyl of HIS_3 and to the vibration localised on LEU_5, respectively. Theory predicts the second resonance to demonstrate intense both, Infrared and in Raman. This resonance corresponds to that at 1689 (1612) cm^{-1} when the structure is next to silica. When in aqueous environment, the blue shift of this vibration makes its Raman contribution to be less consistent with the experimental results.

Further, below, there are vibrations at 1701 (1626), and 1667 (1589) cm^{-1} which represent antisymmetric stretching activity of the C-terminal alone and of the carbonyl of HIS_6 residue. These two transitions demonstrate relatively strong Infrared absorptions and rather weak Raman activities. The Histidine rings C=C stretching modes of the residues six and three dominate in anticipated Raman at 1600 (1516 + 38 = 1554) and 1592 (1507 + 38 = 1545) wavenumbers. In result we may clearly see that when in the aqueous environment the structure does not demonstrate Infrared and Raman responses to match well the experimentally observed in mid-IR spectral range. The disagreement is more obvious for the Raman response. On the other side, the same as when next to silica, this structure demonstrate very good agreement with the experiment in the spectral range 1390 – 1500 cm^{-1}. In the considered structural case either when the polypeptide is next to silica or in aqueous environment, theory does not predict any significant change in delocalization character of the normal modes in the two cases.

Let us review the case of LD91_1:215 when next to SiO_2 interface. Starting from the N-terminal, we observe a rather unfolded but a helical like structure. This involves the first three residues. The other three amino acids, towards the C-terminal form a rather linear backbone. Interestingly, overall, the polypeptide is relatively well hydrated, though the number of water (within 5Å next to a residue) decreases from 12 next to LEU_1 down to 6 next to HIS_6. Interesting to notice that on the N-terminal side, the initial helical tern stands nearly perpendicular towards the silica interface. With the side groups of LEU_1 and SER_4 stretching towards the surface, the helical tern forms an ark structure, which may accommodate 2 water molecules within: the detailed data on hydration of each residue is provided in Fig. S13. With the rest of the polypeptide backbone aligned nearly parallel to surface and with the side groups of LEU_5 and HIS_4 oriented towards silica, as well, overall polypeptide forms a

dome-like structure with relatively large space underneath to be filled with water. This explains the anticipated well hydration of the residues of the considered structure of the polypeptide next to the interface. In the considered case, the role of the side groups in formation of a hydrophobic dome like structure is indispensable and may be even competing. In case of stabilization and locking of a considerable aqueous cluster underneath and next to silica surface, such dome-like structures may assist in formation of dynamically "frozen" and vibrationally and thermodynamically uncoupled subsets of water next to interface.

Let us review the vibration properties in the mid-IR spectral range, see Fig. S19. At the higher frequency side of the carbonyl manifold, we calculate prominent Infrared and Raman signatures local vibration states specific to Amide I carbonyls of LEU_5, ASP_2, HIS_3, and LEU_1 at 1760.1 (1690), 1747 (1675), 1737 (1665) and 1730 (1658) cm^{-1}, respectively. While the anticipated Raman spectral signatures of this modes are consistent with the experimental results, their infrared absorptions are relatively blue shifted comparing to the main Amide I optical density in FTIR. In the lower frequency range of the Amide I, the calculated dispersion demonstrates a triplet of vibrations specific to: asymmetric stretchings of Amide I and COO^- of the side group of ASP_2 at 1712 (1637) cm^{-1}; Amide I of HIS_6 at 1703 (1627) cm^{-1}; Amide I of SER_4 at 1692 (1616) cm^{-1}. These vibrations demonstrate very strong IR absorption and moderate Raman scattering. Raman activity of the normal modes specific to HIS_6 at 1703 (1627) cm^{-1} may correspond to the observed Raman scattering at 1610 cm^{-1}. Next, the calculated Infrared resonance at 1658 (1579) cm^{-1} is specific to asymmetric COO^- vibration of the C-terminal. Its Raman contribution is rather weak. The calculated Raman response show strong Raman scatterings specific to C=C Ram of the Histidine rings of HIS_6 and HIS_3 at 1609.5 (1526 + 38 = 1564) and 1593.8 (1509 + 38 = 1547) cm^{-1}, respectively. Finally, here, there is no agreement between FTIR absorption in the spectral range 1430-1500 cm^{-1} and the frequencies specific to Amide III vibrations. Overall, the calculated responses for the considered structure show rather distant resemblance with the experimental data.

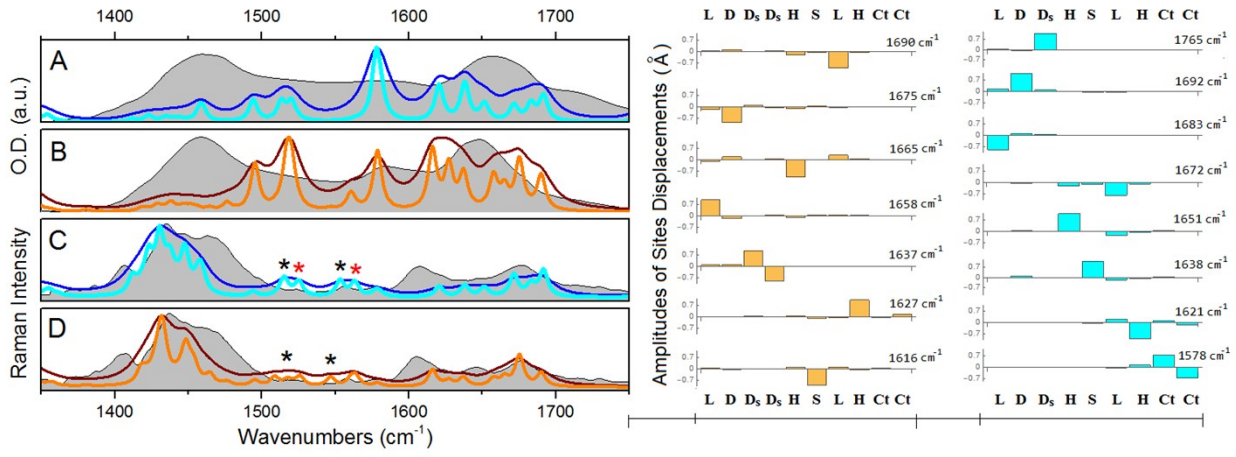

Fig. S19 *Calculated Infrared and Raman spectra in comparison to the experimental results for the structural case LD91_1:215. The definitions are as in Fig. S15. The C-terminal specific resonance at 1765 cm^{-1} is outside of the provided spectral region.*

When in aqueous environment, in the spectral region specific to carbonyl moiety, the theoretical spectra for the structural case LD91_1:215 demonstrate a very blue shifted (outside of the spectral region in Fig. S19) stretching localized on the COOH side group of the Aspartic Acid at 1829 (1765) cm^{-1}. Further down the frequency ladder, we observe a doublet of state at 1761 (1692) and 1753 (1683) cm^{-1}, which are localized on the Amide I of ASP_2 and LEU_1, respectively. These two vibrations correspond to those at 1747 (1675) and 1730 (1658) cm^{-1} for this structure when next to silica, respectively. Next, theory predicts a doublet of vibrations at 1743 (1672) and 1724.7 (1651) cm^{-1}. The first vibration involves Amide I stretching of LEU_5 with a small symmetric admixing of Amide I of HIS_3. The second one is due to Amide I stretching of HIS_3 with a small symmetric admixing of Amide I of LEU_5. At the bottom of the carbonyl specific spectral manifold, we observe resonances at 1712 (1638) and 1697 (1621) cm^{-1}. The former one is mainly due to the residues four and five with a participation of the second and sixth Amino acid. The latter is localized on Amide I unit of HIS_6. The calculated responses when in aqueous environment show no resemblance to the experimentally detected spectra. It is important to notice here that either in the case when next to silica or when in aqueous environment theory predicts for this structure normal modes (specific to carbonyls) to be of very local character, to demonstrate eigen-frequencies to be rather separated from each other and to demonstrate rather good Raman and Infrared activities (this indicates low symmetry at the local sites).

Here, let us explore the theoretical predictions for the structural case LD91_5:64. With the structural switch from β-like ASP_2 towards a helical-like realization along the His_3-Ser_4-Leu_5-His_6 sequence, and with a present intramolecular Hydrogen bond pattern (NH of Asp_2 and C=O of His_6, C=O of Asp_2 and NH groups of Leu_5 and His_6) the polypeptide demonstrates a typical hairpin architecture.[11-13] With strong intramolecular Hydrogen bonding pattern, the accessibility of internal residues for water is expected to be reduced. Indeed, MD simulation shows that only five or less water molecules (in average) are present within 5 Å distance range from the second and the fifth residues as the carbonyls of these amino acids orient towards NH groups of the backbone within the helical turn: the detailed data on hydration of each residue is provided in Fig. S13.

Let us describe the normal modes and their properties for this structural case when next to silica in the spectral region of Amide vibrations, see Fig. S20. First, quantum theory anticipates that at the high frequency side of the carbonyl manifold the structure would demonstrate Infrared and Raman moderately active resonances at 1770 (1706) and 1763 (1698) cm^{-1} due to local Amide I vibrations of HIS_3 and a symmetric stretching which involves Amide I unites of the first and second residues, respectively. In the middle of the Amide I spectral range, the structure is anticipated to show a nearly degenerate doublet of Amide I symmetric and antisymmetric stretchings delocalised mainly over the

first, fourth and fifth residues at 1736 (1669) and 1736 (1668) cm^{-1}, respectively. The higher frequency state dominates over the lower in both, Infrared absorption and Raman response. Very next to the doublet, at the lower frequency side, there is a transition at 1730 (1663) cm^{-1}. This is due to a delocalized vibration, which involves all the residues from the first till the fifth one. However, the main contributions into this vibration comes from the displacements specific to Amide I groups of LEU_1 and LEU_5 mixed in the antisymmetric fashion. This transition has a very strong IR absorption and a weak Raman activity. At the lower frequency side of the carbonyl/carboxyl manifold, the system is predicted to show a spectral doublet at 1703.4 (1633.2) and 1694.9 (1624) cm^{-1} due to the C-terminal delocalizations to involve Amide I of HIS_6 and COO antisymmetric stretching, respectively. The former resonance should demonstrate a stronger Raman intensity, the latter demonstrate intense IR absorption. Further, down the energy ladder, at 1656 (1582) cm^{-1} we should expect a contribution of antisymmetric stretching of COO^- of ASP_2. The transition is predicted to demonstrate a strong Infrared absorption and a weak Raman. The strong Raman scattering of HIS_6 and HIS_5 C=C stretchings of the rings are anticipated at 1611.2 (1533 +38 = 1571) and 1592.2 (1512 + 38 = 1550) cm^{-1}, respectively. Overall, the calculated spectra for this structure do not correspond well to the spectra observed in experiments. There is a poor agreement in Amide spectral I region for both IR and Raman responses. There is an obvious lack of Raman scattering at 1610 cm^{-1}. There is in a poor agreement with the spectral range between 1430 and 1500 cm^{-1}. However, it is interesting to notice that, possibly, due to the hairpin architecture the considered structure demonstrates delocalized vibrations which involve residues located distantly or even next to the two terminals.

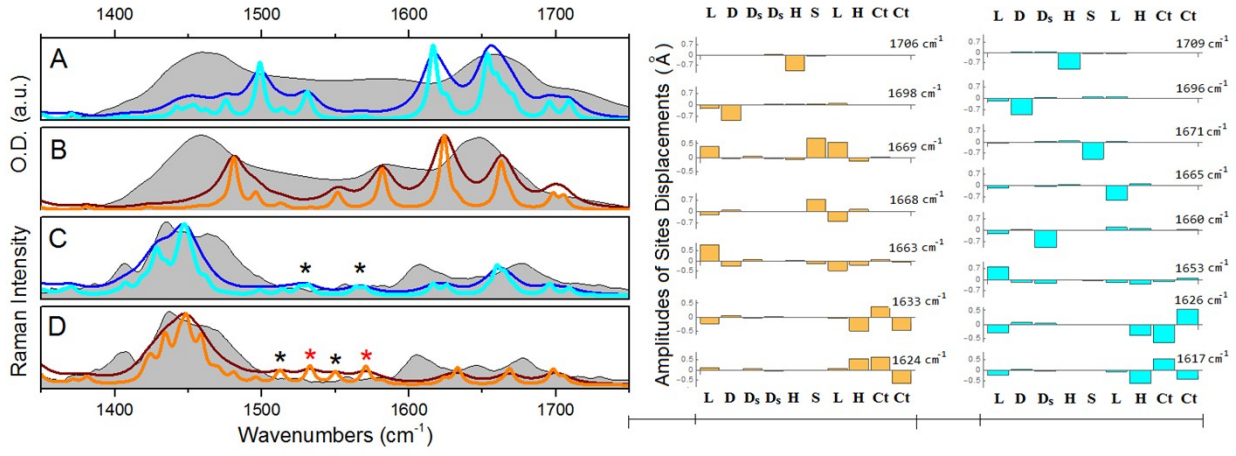

Fig. S20 *Calculated Infrared and Raman spectra in comparison to the experimental results for the structural case LD91_5:64. The definitions are as in Fig. S15.*

Let us explore normal modes prediction for the structure in aqueous environment. At the high frequency side LD91_5:64 is expected to show two resonances at 1778 (1709) and 1765 (1696) cm^{-1}. These are very analogous to the two calculated vibrations at 1770 (1705) and 1763 (1698) cm^{-1} for this

structure but when polypeptide is next to silica. Consistently, both resonances demonstrate moderate Infrared absorption and Raman signals, which correspond relatively well to the experimental observations when in deuterium oxide. Further below, there is a pair of states at 1743 (1671) and 1738 (1665) cm^{-1} which represent vibrational activities of Amide I more or less localized on Amide I carbonyls of SER_4 and LEU_5 residues, respectively. Both resonance should demonstrate strong Infrared absorption, and the lower frequency mode would be expected to show a moderate Raman activity. In the very middle of the Amide I spectral range, the anticipated IR spectrum of LD91_5:1664 contains intense pair of resonances at 1732 (1660) and 1726 (1653) cm^{-1}. The vibrations are due to the side COOH moiety of ASP_2 residue and due to the vibration, which is completely analogous to that at 1730 (1663) cm^{-1} for this structure but when polypeptide is next to silica. The higher frequency resonance demonstrates a strong Raman activity. At the lower frequency side of the Amide I manifold, theory anticipates for this structure two resonances: one is at 1701 (1626) cm^{-1} and another at 1693 (1616) cm^{-1}. Both transitions would be specific to the antisymmetric stretching of the COO group of the C-terminal coupled through bond to Amide I of HIS_6 and through space (since the hairpin like turn) to Amide I of LEU_1. The difference is in the phase relations of relative displacements of the carbonyl of HIS_6. The Raman activities of the Histidines C=C ring modes are at 1615 (1532 + 38 = 1570) and 1609 (1526 + 38 = 1564) cm^{-1} for the third and the sixth residues, respectively. Overall, the calculated spectra for the case LD91_5:64 demonstrate some resemblance with the experimentally observed FTIR and Raman spectral dispersions, save infrared absorption in the spectral region specific to Amide III. Here, it is interesting to notice that DFT calculations of the normal modes indicate that upon removal of silica and protonation of the Aspartic Acid side group, at the higher frequency limit of the Amide I region, the normal mode become somewhat more local.

10. DFT studies: details on fitting.

Fig. S21 compares theoretically calculated Infrared and Raman spectra when the polypeptide is next to a silica surface and in aqueous environment for the selected structural cases with the corresponding optical responses detected by experiment. First, the calculations demonstrate that for the structural case LD82_1:350 when next to silica, while the predictions for the Amide I region fit well to the results of FTIR, the theoretically anticipated Amide III resonances are inconsistently blue shifted comparing to experiment. Also, for this structure, theory fails to predict the observed Raman intensity at 1610 cm^{-1}. For this structure in water, the anticipated state at 1780.9 (1712.8) cm^{-1} demonstrates too intense Raman activity comparing to the experiment

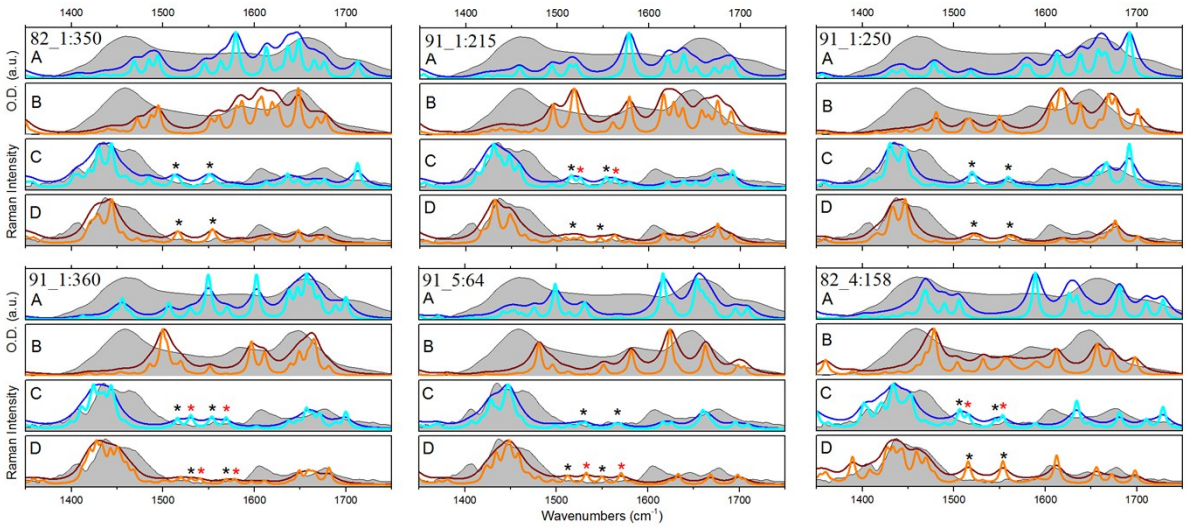

Fig. S21 *Comparisons of calculated optical properties for the considered structural cases with experimentally detected optical responses. (a) and (c): calculated Infrared and Raman spectra for the indicated structural case, respectively, when in aqueous environment, assuming 7 and 20 cm^{-1} line-width (cyan and blue lines) and compared with the experimental results given as grey backgrounds. (b) and (d): calculated Infrared and Raman spectra for the indicated structural case, respectively, when at SiO_2 interface in aqueous environment, assuming 7 and 20 cm^{-1} line-width (orange and brown lines) and compared with the experimental results given as grey backgrounds. The stars marks indicate Raman responses of C=C histidine rings stretching modes.*

In the case LD91_1:250, when next to silica, both calculated IR and Raman responses of the structure demonstrate some resemblance to the experimental results. However, at the higher frequency side, beside strong infrared absorption, theory anticipates very strong Raman activity at 1747.5 (1676.4) cm^{-1}, which does not agree with the experimental data. Also, the observed Raman scattering at 1610 cm^{-1} is under accounted for by DFT and the calculated absorption does not match the observed experimental results in the spectral region 1430-1470 cm^{-1}

Quantum chemistry predicts similar optical responses for LD91_1:250 when in water: the aqueous case lacks Raman intensity at 1610 cm^{-1}, demonstrates a mismatch in 1650-1670 cm^{-1} spectral range comparing to the results of FTIR studies, and a complete disagreement for infrared absorption in the spectral region 1430-1470 cm^{-1}. The structural case LD91_1:360 is evolutionarily related to the case LD91_1:250. Therefore, analogously, the spectra calculated for polypeptide next to silica show some resemblance to the experimental results, save the lack of Raman scattering at 1610 cm^{--1}, slight mismatch in the spectral region of 1630-1670 cm^{-1}, and a strong disagreement in the spectral region from 1430 to 1500 cm^{-1} as frequencies of the calculated Amide III normal modes are blue shifted. In the aqueous environment, for the structural case LD91_1:360, the predicted spectral responses resemble the experimental results: there is a reasonable match in Infrared absorption for Amide I and there are two Amide III modes specific to histidine-3 admixed with the COH bending of the serine-4 side group that contribute favorably in the spectral range 1430-1500cm^{-1}. However, theory for this structure does not explain the observed Raman activity at 1610 cm^{-1}.

For the structural case LD91_1:215, both when next to silica or in an aqueous environment, DFT anticipates spectral properties of Raman scattering and infrared absorption, which show no resemblance to the experimentally detected spectra. Next, reviewing results for LD91_5:64, when at the silica interface, we observe poor agreement between calculated and experimental Raman and infrared in the Amide I region and the spectral range between 1430 and 1500 cm^{-1}. Also, there is an obvious lack of Raman scattering at 1610 cm^{-1}. However, for the LD91_5:64 structure when in the aqueous environment: the calculated spectra demonstrate some resemblance with the experimentally observed FTIR and Raman spectral dispersions, save a mismatch for infrared absorption in the spectral region specific to Amide III. Finally, here, for the case LD82_4:158 at SiO_2: the anticipated spectral properties resemble the experimental results very well. However, when in aqueous environment, the disagreements between the theoretical and the experimental results are obvious.

11. Isothermal titration calorimetry data.

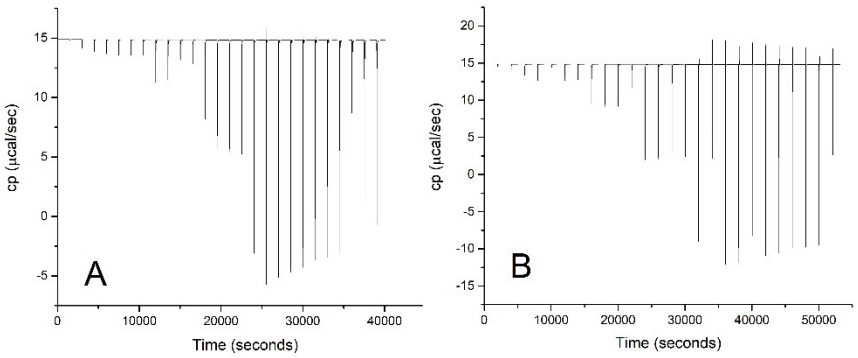

Fig. S22 Isothermal titration calorimetry data. (A) without salt, (B) with salt.

REFERENCES

1. K. R. Farmer and R. A. Buhrman, *Semiconductor Science and Technology*, 1989, **4**, 1084-1105.

2. J. Horbach and W. Kob, *Phys. Rev. B*, 1999, **60**, 3169.

3. A. Saksaengwijit and A. Heuer, *Phys. Rev. E*, 2006, **73**, 061503.

4. G. Biroli and J. P. Garrahan, *J. Chem. Phys.*, 2013, **138**, 12A301.

5. K. Vollmayr-Lee and A. Zippelius, *Phys. Rev. E*, 2013, **88**, 052145.

6. Y. Wimmer, A. El-Sayed, W. Gös, T. Grasser and A. Shluger, *Proc. Math. Phys. Eng. Sci.*, 2016, **472**, 2190.

7. J. Schlitter, *Chem. Phys. Lett*. 1993, **215**, 617-621

8. M. de Berg, O. Cheong and M. van Kreveld, M. Overmars. *Computational Geometry: Algorithms and Applications*, Springer, 2008, pp. 2–8.

9. D. E. Knuth, *Axioms and hulls, Lecture Notes in Computer Science*, Springer-Verlag, Berlin, Heidelberg, 1992, 606.

10. V. Volkov, F. Nuti, R. Takaoka, Y., Chelli, A. M. Papini and R. Righini, *JACS*, 2006, **128**, 9466-9471.

11. K.-C. Chou, *Anal. Bioch.*, 2000, **286**, 1-16.

12. X. Shi, X. Hu, S. Li and X. Liu. *J. Theor. Biol., 2011*, **286**, 24-30.

13. S. Aravinda, U. S. Raghavender, R. Rai, V. V. Harini, N. Shamala and P. Balaram, *Org. Biomol. Chem.*, 2013, **11**, 4220-4231.