



## Structural Determinants for Protein adsorption/non-adsorption to Silica Surface

Christelle Mathé, Stéphanie Devineau, Jean-Christophe Aude, Gilles Lagniel,  
Stéphane Chédin, Véronique Legros, Marie-Hélène Mathon, Jean-Philippe  
Renault, Serge Pin, Yves Boulard, et al.

### ► To cite this version:

Christelle Mathé, Stéphanie Devineau, Jean-Christophe Aude, Gilles Lagniel, Stéphane Chédin, et al..  
Structural Determinants for Protein adsorption/non-adsorption to Silica Surface. PLoS ONE, 2013,  
8 (11), pp.e81346. 10.1371/JOURNAL.PONE.0081346 . hal-02112602

**HAL Id: hal-02112602**

**<https://hal.science/hal-02112602>**

Submitted on 25 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Structural Determinants for Protein adsorption/non-adsorption to Silica Surface

Christelle Mathé<sup>1,2,3\*</sup>, Stéphanie Devineau<sup>1</sup>, Jean-Christophe Aude<sup>2</sup>, Gilles Lagniel<sup>2</sup>, Stéphane Chédin<sup>2</sup>, Véronique Legros<sup>3</sup>, Marie-Hélène Mathon<sup>4</sup>, Jean-Philippe Renault<sup>1</sup>, Serge Pin<sup>1</sup>, Yves Boulard<sup>2,5</sup>, Jean Labarre<sup>2\*</sup>

**1** Laboratoire de Radiolyse, SIS2M, IRAMIS and UMR3299 CEA-CNRS, Saclay, France, **2** Service de Biologie Intégrative et Génétique Moléculaire, iBITec-S, FRE3377 CEA-CNRS-Université Paris-Sud, Saclay, France, **3** Laboratoire Analyse et Modélisation pour la Biologie et l'Environnement, UMR 8587 CNRS-Université Evry Val d'Essonne, Evry, France, **4** Laboratoire Léon Brillouin, IRAMIS and UMR12 CEA-CNRS, Saclay, France, **5** Laboratoire Structure et Dynamique par Résonance Magnétique, SIS2M, IRAMIS and UMR3299 CEA-CNRS, Saclay, France

## Abstract

The understanding of the mechanisms involved in the interaction of proteins with inorganic surfaces is of major interest in both fundamental research and applications such as nanotechnology. However, despite intense research, the mechanisms and the structural determinants of protein/surface interactions are still unclear. We developed a strategy consisting in identifying, in a mixture of hundreds of soluble proteins, those proteins that are adsorbed on the surface and those that are not. If the two protein subsets are large enough, their statistical comparative analysis must reveal the physicochemical determinants relevant for adsorption versus non-adsorption. This methodology was tested with silica nanoparticles. We found that the adsorbed proteins contain a higher number of charged amino acids, particularly arginine, which is consistent with involvement of this basic amino acid in electrostatic interactions with silica. The analysis also identified a marked bias toward low aromatic amino acid content (phenylalanine, tryptophan, tyrosine and histidine) in adsorbed proteins. Structural analyses and molecular dynamics simulations of proteins from the two groups indicate that non-adsorbed proteins have twice as many  $\pi$ - $\pi$  interactions and higher structural rigidity. The data are consistent with the notion that adsorption is correlated with the flexibility of the protein and with its ability to spread on the surface. Our findings led us to propose a refined model of protein adsorption.

**Citation:** Mathé C, Devineau S, Aude J-C, Lagniel G, Chédin S, et al. (2013) Structural Determinants for Protein adsorption/non-adsorption to Silica Surface. PLoS ONE 8(11): e81346. doi:10.1371/journal.pone.0081346

**Received:** July 4, 2013; **Accepted:** October 11, 2013; **Published:** November 25, 2013

**Copyright:** © 2013 Mathé et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** The authors acknowledge support from the "programme de Toxicologie" of the CEA and from the "programme NanoSciences Ile-de-France". The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

\* E-mail: jean.labarre@cea.fr

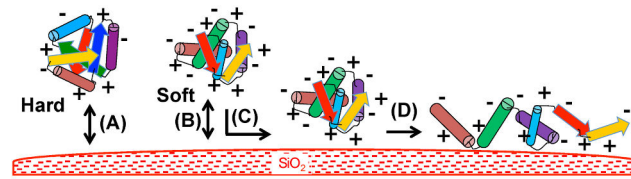
† Current address: Université de Lorraine, Laboratoire Ingénierie des Biomolécules (LIBio), 2 av. de la Forêt d'Haye, BP 172, 54505, Vandœuvre-lès-Nancy, France

## Introduction

The adsorption of proteins on surfaces is a quasi-universal phenomenon of major physiological and toxicological significance. However, the mechanisms and structural determinants of protein/surface interactions are still unclear [1]. A crude description would assume that the main determinants of protein adsorption are electrostatic interactions on charged surfaces, and hydrophobic interactions on hydrophobic surfaces. This scheme is perfectly functional for chromatographic techniques [2], but fails to explain the "nonspecific adsorption" of proteins that occurs, for example, on biosensors, implants, etc [3,4]. The only way to prevent such adsorption is to expose the surface to a passivating protein like BSA, which will saturate all sites [5] or to an antifouling compound such as poly(ethylene glycol) [6]. The

important question then is, rather than why a given protein is adsorbed, why should another protein not be adsorbed on a surface? In other words, is there a determinant of the relative sensitivity of proteins to nonspecific interactions? Obviously, the answer to this important question depends on the physical and chemical structure of the surfaces considered.

Owing to its omnipresence, silica ( $\text{SiO}_2$ ) is a reference material in the study of interactions of proteins with inorganic surfaces. The surface of crystalline silica is composed of silanol groups ( $\text{Si-OH}$ ) and siloxane bridges ( $\text{-Si-O-Si-}$ ). At pH higher than 3, the silanol groups tend to be deprotonated as  $\text{Si-O}^-$  leading to a negatively charged surface. Silica is used in a wide variety of applications. Nowadays, a growing utilization of silica is in the form of nanoparticles, mainly as an anti-agglomerant in drugs, creams, food, etc. Nanoparticles (NPs) are objects of nanometric size ( $< 100$  nm). Thanks to their extremely low size,



**Figure 1. Scheme for the interaction of proteins with the silica surface.** Interactions between the proteins and the surface are supposed to be initiated by electrostatic interactions between positively charged residues and the negative charges of the silica surface (A) and (B). This interaction remains transient and reversible in the case of “hard” proteins. In the case of “soft” proteins, the proteins can deform on the surface and establish other electrostatic contacts (C), which may lead to a spreading of the protein on the surface (D), a quasi-irreversible process.

doi: 10.1371/journal.pone.0081346.g001

they have two major properties: (i) the ability to penetrate cells and reach toxicological targets or medical targets and (ii) a high surface/mass ratio that considerably amplifies the material's surface available for interactions. These properties and the growing utilization of nanoparticles raise the question of their biological effects and health hazards. Upon contact with a biological fluid, e.g. when entering a cell or the blood stream, NPs are readily coated by proteins [7,8]. The nature of the proteins adsorbed on the NPs is a determining factor in the fate and biodistribution of the NPs in the organism (for review see [9]), a property that can be used in nanomedicine to design new carriers targeting specific organs or tissues [10,11]. In this context, a better understanding of adsorption mechanisms would be of great interest. Besides, some bio-surfaces, such as cell membranes or cell walls, are typically nanostructured [12,13], and probably present a higher degree of complexity [14] than standard NPs. Thus, studying protein-NPs interactions can be a relevant first approach to the more complex problem of understanding the interaction of proteins with extended nano-surfaces.

The consensual view is that protein adsorption to silica (see schematic model in Figure 1) and similar metal oxides results from both (i) the electrostatic properties of the protein and (ii) its ability to induce structural deformation on the surface [15–18]. Electrostatics (ionic interactions and hydrogen bonds) is considered of major importance, at least for the first contacts (Figure 1A and 1B). In particular, basic amino acids (Arg and Lys, which are positively charged at neutral pH) are supposed to be essential to establish electrostatic interactions with the electronegative silica surface. This first step is considered reversible. The second step of the adsorption process is supposed to depend on the degree of “hardness-softness” of the proteins [1]. Rigid and tightly structured proteins (“hard” proteins) do not deform on the surface and are not prone to adsorption (Figure 1A). In contrast, proteins with weak internal cohesion (“soft” proteins) are more able to deform and structurally rearrange on the surface leading to increase the number of interactions and to a spreading of the protein on the surface (Figure 1C and 1D). During this conformational change, interactions may occur between the silica hydrophobic sites (e.g. siloxane bridges) and hydrophobic residues of the protein exposed to the surface. This second step is often considered irreversible. The adsorption process is mainly

driven by an entropy gain arising from protein rearrangement and the release of water molecules in the bulk compartment.

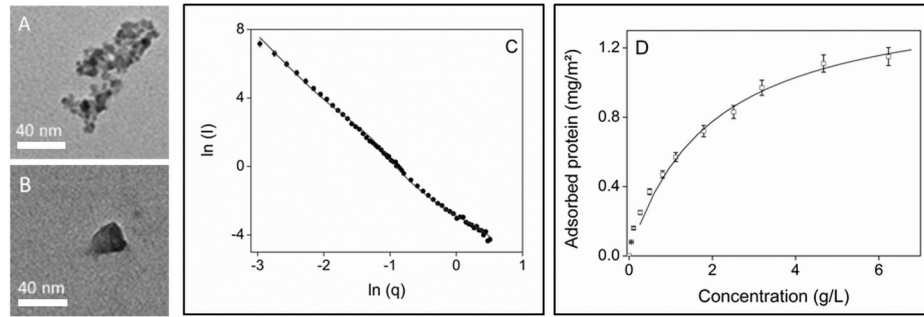
However, in spite of a wealth of studies on this question, the physicochemical parameters that could determine whether a protein is prone or resistant to adsorption have not been identified with precision. For example, it could be expected that a high content in positively charged amino acids or a low content of residues known to be important for protein “hardness” would favor adsorption. However, to our knowledge, no experimental approach has addressed the question of whether some specific amino acids promote or hinder the adsorption process. In addition, the protein adsorption studies have generally been performed on individual proteins, such as lysozyme and serum albumin [19,20]. Though this approach has greatly enhanced our understanding of the adsorption process for these model proteins, it does not allow easy comparisons or an overall vision of the phenomena involved.

To answer this fundamental question, we present in this work a proteomic strategy, which consists in identifying, in a large protein mixture, the proteins that bind to the surface and the proteins resistant to adsorption and in comparing the characteristics of the two protein subsets. This methodology has the advantage of throughput compared with most adsorption studies conducted on isolated, purified proteins. The strategy was tested with silica nanoparticles and our findings lead us to propose a refined model of protein adsorption.

## Results

### Preliminary characterizations

For this study, we used silica in the form of nanoparticles (NPs) as they provide a large specific surface for the subsequent analysis. The SiO<sub>2</sub> NPs were suspended in a phosphate buffer containing 150 mM NaCl (PBS for phosphate buffered saline). NP size and shape were characterized by transmission electron microscopy (TEM) images (Figure 2A and 2B) and small-angle neutron scattering (SANS) experiments (Figure 2C). Both techniques show two distinct populations of particles, a major one composed of large spherical particles with a mean diameter of  $26 \pm 2$  nm and a second one composed of smaller particles with a mean diameter of  $5 \pm 2$  nm, possibly coming from the fragmentation of the larger ones. In aqueous solution, the NPs tend to form



**Figure 2. Characterization of silica nanoparticles used in the study.** Transmission electron microscopy images of silica NPs: (A), small aggregated particles and (B), single large particle. (C) Small angle neutron scattering profiles of silica NPs; experimental data (■) and fitting by a particle size distribution (gray line). (D) adsorption isotherm of yeast proteins on silica NPs in PBS buffer; experimental data (□) and fitting by the Langmuir model (black line).

doi: 10.1371/journal.pone.0081346.g002

small aggregates of a mean size of 50 nm. NP surface area measured by gas adsorption was  $\text{SBET} = 168 \text{ m}^2/\text{g}$ . We determined a Zeta potential of  $-14.2 \text{ mV}$ , which indicates a negatively charged surface at pH 7.4 in PBS buffer, consistent with the isoelectric point of pH 2 [21].

The yeast protein mixtures were prepared and labeled with sulfur 35 (see Materials and Methods section) and were first used to determine the overall adsorption isotherm at an NP concentration of 5 g/L and a protein concentration ranging from 0.25 g/L to 6.5 g/L (Figure 2D). The isotherm could be fitted by the Langmuir model [22]. The maximum quantity of proteins adsorbed on the surface was  $1.58 \text{ mg/m}^2$  under these conditions, which is similar to values obtained for the adsorption of pure proteins, such as serum albumin and lysozyme [19].

### Comparative two-dimensional electrophoresis

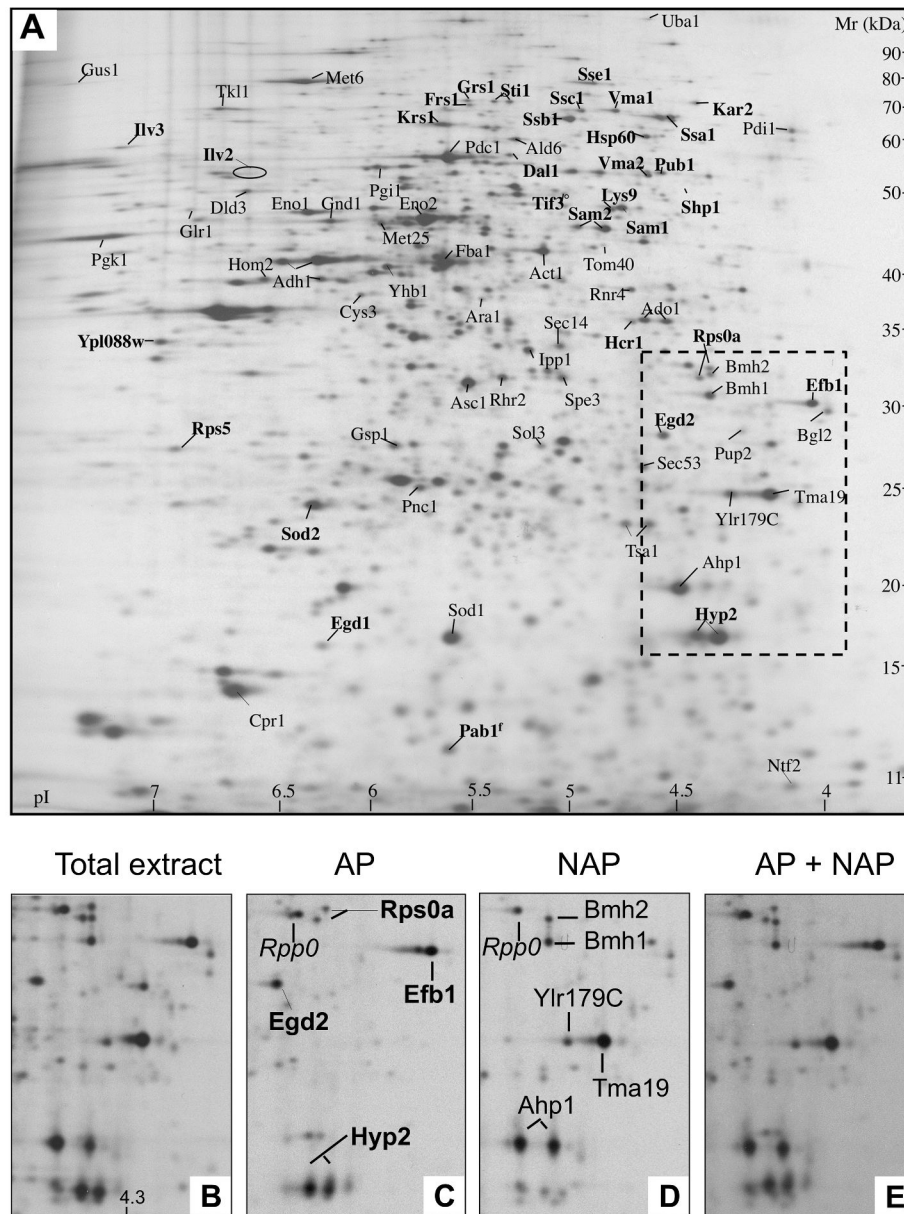
For the next experiments designed to separate proteins that bind NPs from proteins resisting adsorption, we chose a concentration of the protein mixture of 1 g/L. At this concentration, 37% of the proteins were bound to the NPs and the rest (63%) remained in solution. Under these conditions, the surface covered by proteins is far from being saturated ( $0.55 \text{ mg/m}^2$  of adsorbed proteins compared with a maximum capacity of  $1.58 \text{ mg/m}^2$ ), ensuring that unbound proteins are really resistant to adsorption. The  $^{35}\text{S}$ -labeled proteins bound to the NPs under these conditions were re-solubilized in the two-dimensional electrophoresis (2-DE) extraction buffer and run on 2-DE gels. In the same way, the labeled proteins remaining in the buffer (resistant to adsorption) were run on 2-DE gels (see Materials and Methods section). The autoradiographs of the gels revealed very different patterns of protein spots. This is illustrated in panels C (adsorbed proteins) and D (non-adsorbed proteins) of Figure 3, which show a selected part of the autoradiographs. A visual comparison of the 2 panels indicates a strong over-representation of some proteins (e.g. Efb1 and Egd2) in the adsorbed fraction and the presence of other proteins (e.g. Ahp1 and Tma19) exclusively in the non-adsorbed fraction. In this part of the gel, only one protein

(Rpp0) belongs to both fractions, with equivalent amounts in each gel. It indicates a marked partition of the protein mixture into the adsorbed fraction and the non-adsorbed fraction. Figure 3 also shows the 2-DE gel of the total extract prior to adsorption (panel A and B) and the superimposition of autoradiographs of the two fractions (panel E) that reconstitute almost perfectly the 2-DE map of the total extract, indicating that very few proteins are lost in the experimental procedure (washing, desorption step, electrophoresis).

Based on this visual examination, we proposed to classify the proteins into 3 groups: the adsorbed proteins (APs), the non-adsorbed proteins (NAPs) and the proteins showing an intermediate behavior. To define these groups, we quantified the radioactivity present in each protein spot of each gel. After standardization of the data, the calculation of the ratio (AP/NAP) of  $^{35}\text{S}$  intensity for each protein allowed us to classify the proteins as APs, NAPs and intermediate proteins (see Materials and Methods section). These groups were composed respectively of 31, 46 and 24 proteins. The 31 APs and the 46 NAPs are listed in Table S1 with the  $^{35}\text{S}$  ratios and are indicated on the 2-DE maps of the total protein extract (Figure 3A).

### Statistical comparison of physicochemical parameters of APs versus NAPs

We undertook a statistical analysis based on amino acid sequences of the proteins of both subsets. We compared 46 features such as amino acid composition, hydrophobicity, polarity, size, isoelectric point, etc... that can be calculated based on amino acid sequences (listed in Table S2). It is clear that many of these parameters are partially linked, but the analysis was designed to identify the most significant ones, if any. For each feature we assessed the significance of the difference in between-group means using the Brunner-Munzel non-parametric test (see Materials and Methods section). Relevant features showing the greatest differences between the two groups are listed in Table 1, ordered by increasing p-values. For some selected parameters, the data are also shown as empirical cumulative distributions representative of



**Figure 3. Autoradiograms of two-dimensional electrophoresis gels of <sup>35</sup>S-labeled yeast proteins.** (A) Reference map of total soluble yeast proteins; names of adsorbed and non-adsorbed proteins are indicated in bold and standard characters, respectively. The area of the gel focused in B, C, D and E (Mr 35-15 kDa, pI 4-4.6) is framed in the dotted line. (B) Total soluble proteins of the focused region. (C) Adsorbed proteins. (D) Non-adsorbed proteins. (E) Superposition of autoradiograms (C) and (D). Is also shown in italics a protein spot (Rpp0), which belongs to the intermediate group. Proteomic experiments were repeated 3 times with similar results.

doi: 10.1371/journal.pone.0081346.g003

the 2 sets of proteins (Figure 4). This representation allows a visual comparison of the two distributions for the parameter considered.

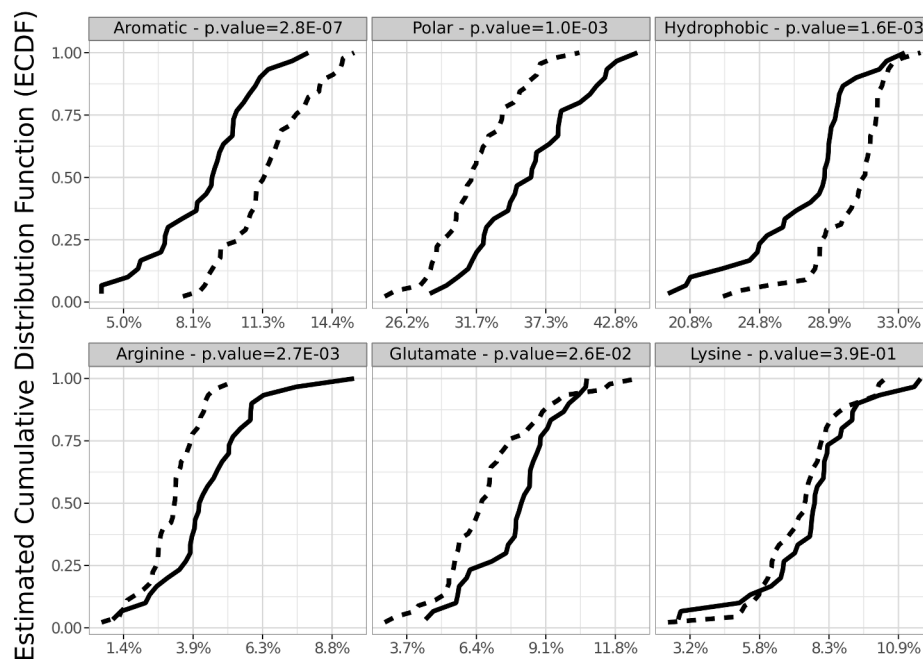
A striking result is the low abundance of aromatic residues (Phe + Trp + Tyr + His) in the APs compared with the NAPs (p-value = 2.8E-7). Remarkably, this bias involves each of the four aromatic amino acids (Table 1), including His, which is also

considered as aromatic according to the Hückel rule (see 23). The second important result concerns the charged amino acids Arg and Glu, which are found in excess in the AP subset (Table 1 and Figure 4). This result is consistent with a major importance of electrostatic interactions in the adsorption process. Though no significant difference is observed for the two other charged amino acids, Lys (Figure 4, p-value = 0.41)

**Table 1.** Physicochemical parameters showing significant differences between the AP and NAP groups<sup>a</sup>.

Parameter tested	In AP/NAP	p-value
Aromatic amino acids (F, W, Y, H)	less	2.8E-07
Polar amino acids (R, K, E, D, Q, N)	more	1.0E-03
$\beta$ -strand amino acids (V, I, Y, C, W, F, T)	less	1.4E-03
Hydrophobic amino acids (C, L, V, I, M, F, W)	less	1.6E-03
$\alpha$ -helix amino acids (E, A, L, M, Q, K, R, H)	more	2.2E-03
Basic amino acids (R, K)	more	2.7E-03
Arginine (R)	more	2.7E-03
Neutral amino acids (all except K, R, E, D)	less	3.1E-03
Tyrosine (Y)	less	3.7E-03
Histidine (H)	less	2.1E-02
Phenylalanine (F)	less	2.1E-02
Acidic amino acids (E, D)	more	2.6E-02
Glutamate (E)	more	2.6E-02
Tryptophan (W)	less	2.7E-02

**a.** In this Table are reported the most significant parameters (with p-value < 0.05; see Table S2) ordered by increasing p-values, i.e beginning with the most significant ones.  
doi: 10.1371/journal.pone.0081346.t001

**Figure 4. Cumulative distribution for selected parameters.** The AP group is represented by a solid line and the NAP group by a dotted line. The parameters presented are aromatic amino acids (Phe + Tyr + Trp + His), polar amino acids, hydrophobic amino acids, Arginine, Glutamate and Lysine.

doi: 10.1371/journal.pone.0081346.g004

and Asp (p-value = 0.44), the APs contain an excess of polar residues charged positively as well as negatively (Table 1). Conversely, the APs contain a reduced number of hydrophobic and neutral amino acids. Finally, residues involved in  $\alpha$ -helix are enriched in APs and residues involved in  $\beta$ -strands are enriched in NAPs.

Most of the remaining features analyzed were either not significant (p-value > 0.05). For instance, protein size (p-value

= 0.08) and isoelectric point (p-value = 0.53) were not relevant in this study. Though this latter result was somewhat surprising, we noted that it is consistent with previous data [24]. Hence, though the total number of positively and negatively charged residues seems important for protein adsorption, the net charge of the proteins does not seem relevant in the process.

**Table 2.** Interaction type per amino acid measured in 3D structures in both groups of proteins<sup>a</sup>.

Interaction type	Ionic/AA <sup>b</sup> (6Å)	Cation- $\pi$ /AA (6Å)	$\pi$ - $\pi$ /AA (4.5-7Å)	H bond /AA SC <sup>c</sup> -SC	Hydrophobic /AA (5Å)	Total attractive interactions /AA
AP	0.100	0.02	0.026	0.278	0.785	1.204
NAP	0.102	0.026	0.046	0.325	0.858	1.359
% of increase	2%	30%	77%	17%	9.5%	13%

a. The mean values were calculated for 12 and 18 protein structures of the AP and NAP groups respectively (see list in Table S1).

b. AA: amino acid.

c. SC: side chain.

doi: 10.1371/journal.pone.0081346.t002

## Structural analyses

Whereas the importance of charged amino acids for binding could be interpreted in terms of electrostatic interactions (see Introduction), the relevance of aromatic residues in limiting protein binding to NPs was more surprising. Aromatic amino acids have different functions in proteins including folding, protein-protein recognition, ligand binding, catalytic activity and structural stabilization. In order to determine the consequences of the amino acid bias in the AP versus NAP subsets in terms of structures and interactions, we searched, among the two sets, the proteins with known structure. We found respectively 12 and 18 structures from the AP and NAP subsets (the structures are listed in Table S1), which represented 39% of the proteins in each group (12/31 for AP subset and 18/46 for NAP subset). Based on the monomer structures of each protein, we calculated the accessibility of charged residues for both groups. No significant difference was observed between the two groups (data not shown). The ability to bind NPs could not be explained by a difference in positive charge (or negative charge) accessibility at the surface of the proteins. In a second step, we quantified the number of different types of intramolecular interactions in the two subsets (see Table 2). These noncovalent attractive interactions are crucial for maintaining the 3D structure of proteins. Our results clearly show that, except for ionic type (salt bridges), there is a significant increase of the interactions in the NAP group. Proteins of the NAP subset thus appear to be more structured with an increased internal cohesion. Moreover, two of the strongest interactions, namely  $\pi$ - $\pi$  and cation- $\pi$  interactions, represent by far the highest augmentation, respectively 77% and 30%, and are well correlated with a higher content of aromatic residues in the NAP subset. This analysis thus shows a clear correlation between the content in aromatic residues, the amount of  $\pi$ -mediated interactions and protein structural stability. We also noted that the 30% augmentation of cation- $\pi$  interactions occurs despite a shortage of basic amino acids in the NAP group. A consequence of this increased implication of basic residues in cation- $\pi$  interactions in the NAPs is that the number of positively charged residues remaining free to interact with negatively charged NPs is thus reduced in the NAPs. In the same line, the unchanged amount of ionic interactions in the NAPs in spite of a lower content of basic residues decreases the available positive charges in the NAP proteins. We conclude that the positively charged residues

available for interactions with silica tend to be markedly higher in the APs than in the NAPs.

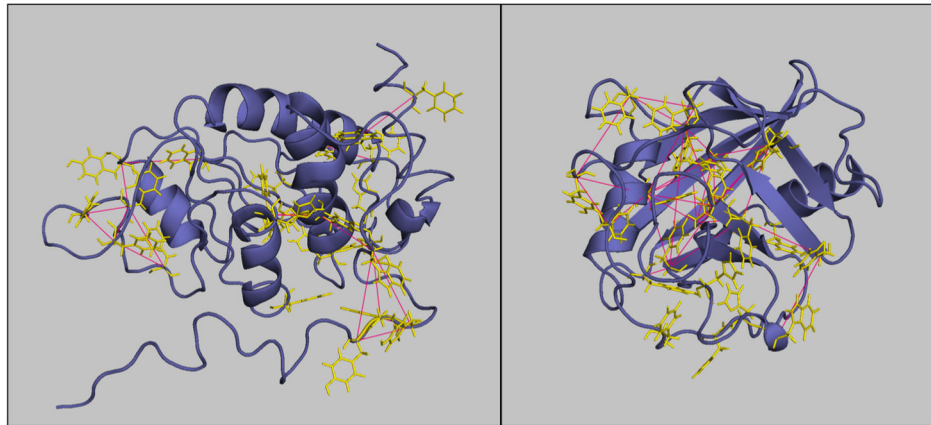
The two protein structures presented in Figure 5 are typical of each group. The protein Cpr1 representative of the NAP subset is more structured and enriched in aromatic clusters. The protein Rps0A belonging to the AP subset is more disordered and the aromatic amino acids are less numerous and clustered.

Among these proteins with known structures, we selected in each group the five proteins with the best structural resolution, to analyze their structural stability by molecular dynamics simulations. The idea was to test on a small protein subset whether protein adsorption and protein flexibility were indeed correlated. Figure 6 shows the change in root mean square deviation (RMSD) over the 5 ns production period of the simulation time for the proteins of the two subsets. Three proteins (Sse1, Rps0A and Hyp2) of the AP subset present strong deviations, rapidly reaching an RMSD value of ca. 4 Å while the deviations for the NAPs remain significantly lower (average RMSD value of 1.35 Å). We carefully checked that these large RMSD values were not the result of artifacts such as the energy terms and the temperature. The largest RMSD values could be explained due to specific structural features: the structure of Sse1 is extended and the C terminal part of Rps0A is unstructured, which confer larger flexibility properties for these molecules. For the 3 other proteins (Hyp2, Ilv2 and Lys9) the mean RMSD value remains always significantly higher (1.95 Å) than those of the NAP subset. This is confirmed with the analysis of B-factor per residue, which reflects the local flexibility of the structure. The mean value per residue is 95 for the AP subset whereas it is 33 for the NAP subset (data not shown). From these simulations, we conclude that proteins binding to silica NPs are more flexible than the NAPs.

## Discussion

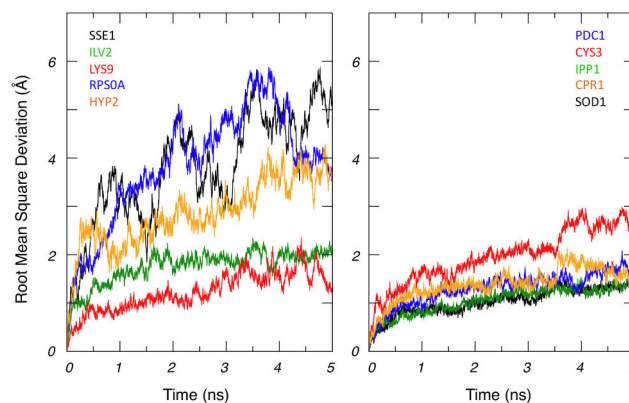
### The methodological approach

Omics methods have gained great popularity in cellular biology and biochemistry. However, despite their high accessibility nowadays, they have seldom been used to resolve physicochemical issues. We demonstrate here that proteomics, associated with molecular modeling and rigorous statistical analysis, can help to shed light on a longstanding issue: the physicochemical determinants of protein adsorption/non-



**Figure 5. Representative structures of proteins belonging to AP and NAP groups.** Left: Rps0a for the AP group. Right: Cpr1 of the NAP group. The secondary structures of the protein are displayed in cartoon mode. The aromatic residues are shown in yellow in line mode and the interactions between backbone carbon atoms of the aromatic residues are displayed in red to show the aromatic clusters in both groups.

doi: 10.1371/journal.pone.0081346.g005



**Figure 6. Root Mean Square Deviation calculated from molecular dynamics trajectories.** AP group (left) and NAP group (right).

doi: 10.1371/journal.pone.0081346.g006

adsorption. The method developed and validated in this work consists in sorting, among a pool of proteins, those proteins prone to adsorb silica NPs from those that resist adsorption. If the proteins of the two subsets are sufficiently numerous and varied, the statistical comparison of these two subsets allows the identification of the physicochemical determinants responsible for adsorption and non-adsorption. This strategy has the advantage of throughput compared with most adsorption studies conducted on isolated, purified proteins. It allows to directly compare the adsorption properties of many different proteins under the same physicochemical conditions.

We used an extract of total soluble proteins produced in yeast. The proteome that can be analyzed by 2-DE is composed of more than 100 soluble proteins with isoelectric points ranging from 3.5 to 7.5 and molecular weights from 10 to 120 kDa. Most of these proteins are globular, highly soluble and have metabolic functions in yeast [25]. Their abundance

ranges from a few percent of the total soluble proteins (e.g. Tdh3, Eno2) to less than 0.01% (e.g. Bgl2, Kar2) [26]. Among the yeast proteins, a high number (31 for the APs and 46 for the NAPs) clearly behaved differently with respect to silica. The size of these groups permitted a statistical comparison and the determination of the physicochemical parameters essential for adsorption and non-adsorption. Though many studies have analyzed and listed the proteins adsorbed on NPs in blood plasma (for review see 8), the approach developed here has, to our knowledge, never been used. We noticed, however, the study of Tenzer et al. [24], which showed that, in accordance with our data, protein charge and protein size do not correlate with protein adsorption/non-adsorption. Unfortunately, this study did not analyze other protein parameters such as amino acid composition. The problem of plasma protein pools is their high dynamic range in terms of protein abundance, i.e. the 10 proteins the most abundant represent more than 99% of the

total bulk mass of proteins present in the plasma [27]. In the best case, even with an efficient LC/MS/MS analysis, the number of proteins analyzed are too low to constitute groups of sufficient size and thus to determine relevant criteria for protein adsorption. This difficulty emphasizes the advantage of using an extract of yeast proteins, which presents a lower dynamic range of protein abundance. The analysis can also be deepened by the use of a labeled pool of proteins. Using a pool of <sup>35</sup>S-labeled proteins facilitates the visual and quantitative analysis of the data even with low amounts of protein on the gels. In such a way, we could routinely analyze about 100 proteins. Note that the method proposed in this paper can be applied to any other type of surface. A possible drawback of the method is that we cannot exclude that some of the proteins identified as adsorbed on the NPs are in fact proteins with a strong affinity for a true AP protein. However, this occurrence should be rare since our search in the yeast databases retrieved only 3 physical interactions between members of the AP subset.

### Importance of charged amino acids

An important result of this analysis is that the net charge of the proteins is not a distinctive criterion among the two groups. However, the total number of negatively and positively charged amino acids is markedly increased in the AP subset (*p*-value = 2.6E-2 and 2.7E-3, respectively), meaning enhanced charge density, polarity and hydrophilicity. A direct consequence is the presence of an increased number of positive charge clusters in the AP subset (Figure S1), which may help to initiate locally a series of electrostatic interactions with the SiO<sub>2</sub> groups. This interpretation is consistent with molecular dynamics simulations of the initial steps of lysozyme interaction with the silica surface [28], suggesting a major importance of several positive charges (mainly Arg residues) in a small area of the protein. Interestingly, the analysis of the 3D structures (Table 2) showed that the number of basic residues (Lys and Arg) not involved in intra-molecular interactions (salt bridges or cation- $\pi$ ) and thus available for electrostatic interactions with the silica surface is higher in the APs than in the NAPs. This effect concerns mainly Arg, which is preferred to Lys in salt bridges [29,30] and in cation- $\pi$  interactions [29,31] probably contributes to increase the differences between APs and NAPs in the probability of electrostatic interaction with the silica surface.

Among basic amino acids, only Arg residues are over-represented in APs, with no bias in favor of Lys and His. This significant difference between Arg and Lys is quite unexpected as one generally assumes that these two basic amino acids have a similar capacity for electrostatic interactions. The difference may arise from the specific interaction of Arg with silica. When we consider the free amino acids, Arg adsorbs better to nanosized silica than His, Lys, and ornithine [32]. One can suppose that the guanidinium group (specific to the Arg residue), with its delocalized positive charge on three nitrogen atoms, may have a higher probability of establishing an electrostatic interaction with the silica surface. Consistently, molecular dynamics simulations of lysozyme adsorption on a silica surface show that Arg residues play a greater role than Lys residues [28]. One hypothesis, already proposed by

Hoefling et al. [33], is that Arg, more than Lys, facilitates the initial contact(s) as it is more prone to penetrate the water layer, and is able after a first contact to change the conformation of the charged guanidinium in order to enhance the interaction surface. In addition, the charged guanidinium group can potentially establish more hydrogen bonds than lysine side chain with the OH groups of the silica surface [34,35]. Arg would thus, more than Lys, be able to deal with the spatial constraints to establish stronger interactions with silica.

The enrichment in Glu residues in the AP subset may be considered as counter-intuitive: unlike Arg, Glu is negatively charged at pH 7.4 and the binding of this amino acid to the electronegative silica surface is therefore unexpected [36]. We thus suppose that Glu does not bind directly to the silica surface, though this hypothesis cannot be totally excluded [37]. An interpretation is that positive cations such as Na<sup>+</sup> cations present at high concentration (150 mM) in the PBS buffer used in our study would be attracted by the silica surface [35] and would screen at least partially the electronegative surface, favoring the attraction of glutamate residues in some places [1]. Another explanation is to consider that the Glu bias is linked to the overall enrichment in both negatively and positively charged residues in the AP subset. In other words, the excess in Glu residues would participate in the electrostatic compensation for the excess of Arg in the AP subset.

### “Hard” proteins are more resistant to adsorption

Our work shows a close correlation between 4 properties of proteins: (i) their aromatic residue content, (ii) their amount of  $\pi$ - $\pi$  interactions, (iii) their level of structural stability (hardness) and (iv) their capacity to resist adsorption. These correlations have been suggested by both statistical comparison of the primary sequences and structural analysis of APs and NAPs. They were further confirmed by analyzing the known 3D-structures of proteins from the 2 groups and by molecular dynamics simulations.

The notion that protein plasticity is an important determinant in the adsorption process is not new, but strong evidence demonstrating this concept is still lacking, though a body of arguments and data supports this notion [15-18]. An interesting example is the  $\alpha$ -lactalbumin, for which the less stable form (Ca<sup>2+</sup>-depleted) has a higher affinity for silica surfaces than the stable form (Ca<sup>2+</sup>-replete) [38]. In this context, our data showing that APs are significantly less structured and more flexible than NAPs help strengthen the above hypothesis.

### Importance of $\pi$ - $\pi$ interactions in protein “hardness”

The origin of protein structural stability is quite controversial. Usually, hydrophobic cores (along with hydrogen bonds and salt bridges) are thought to be a major determinant of structural stability [39-42]. The amino acids classically considered to be mainly involved in the hydrophobic cores are Leu, Ile and Val. Aromatic residues such as Phe and Trp are also considered, but are of lesser importance because of their lower abundance in proteins. In addition, aromatic residues (in the IUPAC definition of the word) like His and Tyr are usually not considered in the hydrophobic core. Alternatively, other studies suggest a major importance of aromatic residues in the

hydrophobic core during the first steps of protein folding [43-45]. It has also been shown that protein structural stability is strengthened by “pure”  $\pi$ - $\pi$  interactions [46,47] and by multiple interactions constituting spatial clusters of aromatic amino acids [48-50]. In our study, the impact of classical hydrophobic amino acids seems limited. In the proteins of known structure, the occurrence of hydrophobic interactions was increased by 9.5% in NAPs compared with APs (Table 2), whereas the increase in  $\pi$ - $\pi$  interactions (involving Phe/Trp/Tyr/His) reached 77%. Moreover, another argument supporting the preeminence of aromatic interactions over hydrophobic interactions is that, among the aromatic residues, the strongest bias was not for the hydrophobic Phe and Trp but for His and Tyr, which do not have hydrophobic properties. Thus, beyond the question of “to bind or not to bind” to silica NPs, our work strongly suggests that aromatic clusters are of major importance in the cohesive core of the proteins and the control of their “softness/hardness”.

## Conclusions

Our study highlights two major determinants of adsorption/non-adsorption on silica surface: (i) electrostatics, with an important role of arginine residues and (ii) protein flexibility, partly controlled by aromatic residues. These two characteristics are key determinants in the two steps of the rough adsorption model depicted in Figure 1. Our data strengthen the notion that aromatic clusters ( $\pi$ - $\pi$  interactions) are of major importance for maintaining the rigid structure of “hard” proteins, which prevents their spreading on the surface. The “soft” proteins, less rich in aromatic residues, are more prone to deform and to establish an increased number of electrostatic interactions and hydrogen bonds between the protein and the surface. Alternatively, during the spreading process, some apolar amino acids normally buried inside the protein core become exposed outside, favoring hydrophobic interactions with siloxane sites (-Si-O-Si-) [17]. In terms of thermodynamics, the driving force would be mainly enthalpy in the first step and entropy in the second step, with the removal of surface-bound water molecules and salt ions and with the structural rearrangements of the protein.

In this work, we proposed a strategy that we validated with silica NPs. The method can be used to study protein interactions with other different types of surfaces (e.g. with other chemical compositions or other sizes). We must keep in mind that the data shown in this work were obtained with NPs for which we expect quasi uniform surfaces properties. For micro-, meso- and macro-materials, the surface charge pattern can be less homogeneous, which may lead to a variable importance of the role of the charged amino-acids from one surface to another. On the other hand, the role of protein softness in enabling deformation and maximizing interactions (of any type) with surfaces is expected to be a general phenomenon, irrespective of the surface electrostatics or of its geometrical characteristics.

Overall, this study validates the proof of concept of the proposed strategy, deepens our understanding of protein adsorption on SiO<sub>2</sub> surfaces and opens the way of using these

parameters (arginine and aromatic content in proteins) to develop models predicting adsorption/non adsorption.

## Materials and Methods

### Silica nanoparticle characterizations

Silica nanopowder (SiO<sub>2</sub>) of 99.5% purity was from Sigma-Aldrich (637238) (CAS Number 7631-86-9). According to the manufacturer, the NPs are 10-20 nm in size and have a specific surface area of 140-180 m<sup>2</sup>/g. NPs were prepared in phosphate buffer saline (PBS) pH 7.4 at the final concentration of 5 g/L and characterized by TEM, SANS, zetametry and gas adsorption isotherms.

TEM images were recorded at IMAGIF (Centre de Recherche de Gif, CNRS) on a Jeol JEM-1400 transmission electron microscope operating at 60 kV and equipped with an Orius SC1000 camera. 3  $\mu$ L of 0.5 mg/mL NP solution was deposited on an ionized 400 mesh carbon-coated Cu grid and left to evaporate for 5 min. Particle size distribution was obtained by image analysis with ImageJ software. Surface area was measured using nitrogen adsorption-desorption isotherms at 77K recorded on a Micromeritics apparatus (ASAP 2010 Instrument). Samples were prepared by drying 1 g of NPs under vacuum at 90°C for 1 h and at 105°C for 2 h to ensure complete moisture desorption before analysis. The specific surface area, noted S<sub>BET</sub>, was calculated as between 0.06 to 0.2 P<sub>00</sub> using the Brunauer-Emmet-Teller method [51]. Zeta potentials of NP solutions (0.1 mg/mL) were measured in buffered solutions on a Malvern Zeta Nanosizer. Zeta potentials were calculated by fitting the electrophoretic mobility to the Smoluchowski model, which is valid for aqueous solutions with medium electrolyte concentrations [52].

SANS experiments were performed on a PAXY small-angle spectrometer of the ORPHEE facility at Laboratoire Léon Brillouin. Small-angle neutron scattering spectra were recorded at room temperature using dry NPs in a 1 mm quartz cell according to two different setups defined respectively by a sample-detector length of 5 m and 2 m and a selected wavelength of 9 Å and 6 Å. Overlapping of wavenumber transfer  $q$  ensures correct absolute values for the scattered intensities measured and lack of multiple scattering. Data were processed according to [53] to determine the size distribution of the particles.

### Production of the protein mixture

The protein mixture used in this study was produced by the yeast (*Saccharomyces cerevisiae*) strain S288C (*Mat $\alpha$  SUC2 mal mel gal2 CUP1*) [54]. Cells were grown with shaking at 30°C in 1 L of a synthetic defined (SD) yeast medium (6.7 g/L yeast nitrogen base (with (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>) and 20 g/L glucose) to obtain a cellular concentration of 2 x 10<sup>7</sup> cells/mL. Cells were collected by centrifugation, resuspended in phosphate buffered saline (PBS) buffer containing 5% glycerol and a cocktail of protease inhibitors (complete, EDTA-free from Roche and 1 mM PMSF) and broken using a French press. The extract was centrifuged (14,000 rpm, 30 min, 4°C) and the supernatant containing soluble proteins was recovered. The protein concentration was determined by the method of Bradford [55].

Typically, 1 L of yeast cells yielded 4 mL of soluble proteins at about 20 µg/µL.

To produce <sup>35</sup>S-labeled proteins, 50 mL of the unlabeled yeast cell culture (see above) was centrifuged, resuspended in a medium devoid of sulfate for 30 min and the cells were labeled with 1 to 2 mCi of <sup>35</sup>S-methionine for 30 min. Under these conditions, nearly all the <sup>35</sup>S-methionine is used for protein synthesis and the amount of <sup>35</sup>S accumulated in <sup>35</sup>S-metabolites is less than 1% (data not shown). The proteins were prepared by the same protocol as for the unlabeled protein mixture except that the volumes were ten-fold lower and that the cells were broken with beads instead of the French press. In typical experiments, we obtained about 0.5 mL of soluble <sup>35</sup>S-protein extract at a concentration of 2 µg/µL containing about 2 µCi/µL.

### Mixing proteins and SiO<sub>2</sub> nanoparticles

The stock solution of dispersed SiO<sub>2</sub> NPs was prepared in PBS buffer pH 7.4 at the concentration of 25 g/L. The NPs were diluted in the protein solution to give a final concentration of 5 g/L, representing a surface of 0.168 m<sup>2</sup> for adsorption per experiment (in a final volume of 0.2 mL). The protein solution was a mixture of unlabeled and <sup>35</sup>S-labeled yeast proteins at the relevant protein concentrations. The <sup>35</sup>S-labeled proteins were used as a tracer in the isotherm experiments and two-dimensional electrophoresis (2-DE).

For the isotherm experiments, the final protein concentration ranged from 0.15 g/L to 6.5 g/L. The mixtures of proteins and NPs were incubated at room temperature for 3 hours with gentle shaking. NPs (with adsorbed proteins) and non-adsorbed proteins were separated by centrifugation (4000 rpm, 2 min, 20°C). The NPs (with adsorbed proteins) were washed in PBS buffer and counted by liquid scintillation (liquid scintillation counter Wallac 1409 DSA). Preliminary tests showed that no radioactive quenching was induced by the presence of SiO<sub>2</sub> NPs in the counted samples.

For the other experiments, the final concentration of unlabeled + labeled proteins was 1 g/L. At this concentration, 37% of the proteins are adsorbed and 63% remain in solution. The adsorbed protein fraction and the non-adsorbed protein fraction were prepared as previously and processed for 2-DE.

### Two-dimensional gel electrophoresis

Proteins of the supernatant were concentrated 5-fold by evaporation and 4 µL of the concentrated solution was added to 16 µL of 2-DE buffer [DNase 1 (100 U/mL), Rnase A (10 U/mL), CaCl<sub>2</sub> (20 mM), MgCl<sub>2</sub> (50 mM), TrisHCl pH 7 (100 mM), SDS 0.02%, ampholine (1%), CHAPS (0.8%), β-mercaptoethanol (1%) and urea (8 mM)]. After incubation for 15 min at 30°C, 0.5 µL of bromophenol blue was added and the sample was processed on 2-DE as described in [25].

NP pellets containing the adsorbed proteins, were washed twice with 200 µL PBS buffer and then incubated for 15 min at 30°C in 20 µL of 2-DE buffer. The mixture was centrifuged (13,000 rpm, 3 min) to separate the NPs (in the pellets) and the desorbed proteins (in the supernatant). Scintillation counting of the NPs confirmed that all the proteins bound to the NPs were solubilized in the 2-DE buffer. Bromophenol blue (0.5 µL) was

added to the supernatant and the sample was run on 2-DE gels.

The typical quantity of protein used for the 2-DE gels was about 50 µg for adsorbed proteins and about 10 µg for non-adsorbed proteins. This value (10 µg) was limited because the high salt concentration (PBS buffer) in the non-adsorbed protein extract perturbs the isofocusing electrophoresis. After electrophoresis, gels were stained with Coomassie Blue R-250, dried, and processed for autoradiography by standard procedures. As expected, exposure times necessary to obtain comparable image intensities were 5-fold higher for non-adsorbed proteins than for the adsorbed proteins. The <sup>35</sup>S images of the gels were also recorded on PhosphorImager (Molecular Dynamics STORM). The experiments were repeated 3 times independently with identical results.

### Protein identification and mass spectrometry

A 2-DE gel reference map of total soluble yeast proteins is available in the laboratory based on previous identifications by different methods [25] including mass spectrometry [56,57]. Adsorbed and non-adsorbed protein fractions on 2-DE gels were identified essentially by comparing the two corresponding gels with the gel of the total protein extract (see for example Figure 3B, 3C and 3D). When an unknown protein spot appeared on the gel(s) or for some confirmations of previously known spots, identifications were performed by mass spectrometry. Spots of interest were extracted from non-radioactive gels and the proteins were digested in-gel by trypsin using standard protocols [58]. Spot extracts were resuspended in 5% acetonitrile, 0.1% formic acid (buffer A). Peptide extracts were submitted to nanoLC-MS/MS using an Ultimate 3000 nanoLC system (Dionex, Thermo Fisher Scientific) connected to an ESI-LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific) equipped with a nanoelectrospray ion source. The ion spray voltage was set at 1.4 kV with a transfer capillary temperature of 200°C. Five microliters of each sample was loaded on a C18 precolumn (300 µm inner diameter x 5 mm, Dionex), and desalted for 5 min at a flow rate of 20 µL/min with 100% buffer A. Then, the precolumn was switched connected to an analytical capillary C18 column (75 µm inner diameter x 15 cm; AcclaimPepMap 100, Dionex, Thermo Fisher Scientific) equilibrated in 100% solvent A. Peptides were eluted using a 0-70% gradient of solvent B (80% acetonitrile, 0.1% formic acid) over 50 min at a flow rate of 300 nL/min. The mass spectrometer was operated in the data-dependent mode to automatically switch between Orbitrap MS and LTQ-MS/MS acquisition. Survey full scan MS spectra were acquired in the Orbitrap in the 400-1600 *m/z* range with the resolution set to a value of 60,000 at *m/z* 400. Sequential isolation of the 5 most intense precursor ions was carried out, followed by their fragmentation by collision-induced dissociation in the linear ion trap.

Data were analyzed using the locally installed Mascot search engine (version 2.2.1, Matrix Science). Peak lists were searched against *S. cerevisiae* sequences in the Swiss-Prot database. Mass accuracy was set to 10 ppm for MS mode and 0.6 Da for MS/MS mode and one missed tryptic cleavage site was allowed in the search. Oxidation of methionine was

searched for as a variable modification and carbamidomethylation was set as a fixed modification.

### Quantitative protein analysis

Radioactive images of 2-DE gels of total protein extract, adsorbed proteins and non-adsorbed proteins, were analyzed using the software Progenesis SameSpots v3.0. The automatic matching of the 3 gels by the software was checked, spot by spot, and corrected if necessary. The intensity (volume value) of each protein spot was normalized to the total intensity of the gel. This normalization allows a direct comparison of the relative intensities of each protein spot in the different gels. Using the quantitative data of volume intensity provided by the software, we calculated for each protein the ratio of relative intensity in the gel of the adsorbed proteins to the relative intensity in the gel of the non-adsorbed proteins. High ratios correspond to adsorbed proteins (APs) and low ratios to non-adsorbed proteins (NAPs). To classify the proteins into 3 groups (adsorbed, non-adsorbed and intermediate), we arbitrarily chose the values  $10^{(0.2)} = 1.58$  and  $10^{(-0.2)} = 0.63$  as thresholds. If the ratio was higher than 1.58, the protein was classified as adsorbed; if the ratio was lower than 0.63, the protein was classified as non-adsorbed; and if  $0.63 < \text{ratio} < 1.58$ , the protein was put in the intermediate group.

### Protein primary structure analysis

Protein coding sequences were extracted from the *S. cerevisiae* full genome version sacCer2 released by UCSC. Statistical analysis was conducted using the R software [59].

Structural and physiochemical features of proteins were calculated using ProFeat [60] or in specific cases using our own implementations (length, molecular weight, isoelectric point, aromatic amino acid composition). The statistical differences among features between the group of adsorbed proteins and the group of non-adsorbed proteins were exhibited using the Brunner-Munzel test [61] as implemented in the R lawstat package [62]. The non-parametric Brunner-Munzel test is used to assess the stochastic equality of two samples ( $H_0: P(X < Y) = P(X > Y)$  against  $H_1: P(X < Y) \neq P(X > Y)$ ). This test is commonly known as a "generalization of the Wilcoxon-Mann-Whitney test" for heteroscedastic cases (i.e. unequal variance or shape distribution). Indeed, we cannot use a student t-test because 40% of the features distributions, such as hydrophobic or neutral amino acids composition, do not fulfill the Gaussian distribution assumption (assess using the robust Jarque-Bera test of normality [63]). We cannot either use the Wilcoxon-Mann-Whitney test because 11% of the features distributions, such as the arginine residue composition, do not fulfill the homoscedasticity assumption (assess using the robust Brown-Forsythe test of equality of variance [64]). Nevertheless, using the Wilcoxon-Mann-Whitney test leads to similar conclusions as shown in the Table S2. Resulted p-values were adjusted for multiple comparisons, by controlling the false discovery rate, using the Benjamini-Hochberg method [65].

We also calculated positive (respectively negative) charge profiles, where positively (respectively negatively) charged amino acids are counted over a sliding window of twelve amino

acids. The results were filtered by only keeping counts above a threshold of three positively (respectively negatively) charged amino acids per sliding window. To compare profiles between the AP and NAP groups, for each protein we calculated the ratio of the filtered charged window over the total number of sliding windows (denoted charge clusters hereafter). Then charge profile distributions were compared using the same Brunner-Munzel statistical test as previously described.

### Structural analyses

All simulations were performed using the AMBER 9 program suite [66] with the Parm99SB force field. From the 30 structures of proteins that were analyzed, 28 of them were imported from the Protein Data Bank [67] and 2 from the Protein Model Portal [68]. These two models are considered as reliable thanks to the high sequence identity between the template structure and the protein of interest (62% for Krs1 and 69% for Sam1). All the calculations were performed on monomer forms. Hydrogen atoms were added with the LEaP program to complete the X-ray structures. Proteins were then neutralized with Na<sup>+</sup> cations and immersed in a water box with at least 10 Å-deep solvation shell using the TIP3P water molecules. The structures were then minimized and used to initiate molecular dynamics. All simulations were performed in the isothermal isobaric ensemble ( $P = 1 \text{ atm}$ ,  $T = 300 \text{ °K}$ ), regulated with the Berendsen barostat and thermostat [69], using periodic boundary conditions and Ewald sums for treating long range electrostatic interactions [70]. The hydrogen atoms were constrained to the equilibrium bond length using the SHAKE algorithm [71]. A 2 fs time step for the integration of Newton's equations was used. The nonbonded cutoff radius of 10 Å was used. All simulations were run with the PMEMD module of the AMBER package.

Molecular dynamics simulations of 5 ns for the production period, without constraints, were used to provide quantitative information about the stability of the models (root mean square deviation; RMSD) and to determine the flexible regions of the structures (B factor). The structures were analyzed with CMTview program version 1.1 [72]. The amino acid accessibilities were calculated with Naccess program version 2.1.1 [73] that use the Lee and Richards method [74].

### Supporting Information

**Figure S1. Cumulative distribution for positive charge clusters (left panel) and negative charge clusters (right panel).** The AP group is represented by a solid line and the NAP group by a dotted line. The electrostatic charge distribution over the protein sequences was evaluated by enumerating the number of positive (respectively negative) charges within a sliding window of 12 amino acids along the polypeptide sequences. A threshold of three charged amino acids per sliding window was used to exhibit positively (respectively negatively) charged clusters for each protein. (EPS)

**Table S1. List of adsorbed and non-adsorbed proteins.**

(XLSX)

**Table S2. List of features analysed and statistically compared between the AP and the NAP groups.**  
(XLSX)

## Acknowledgements

This work has benefited from the facilities and expertise of the Platform for transmission electron microscopy of IMAGiF (Centre de Recherche de Gif - [www.imagif.cnrs.fr](http://www.imagif.cnrs.fr)). We also

thank Claire Boulogne for her help with transmission electron microscopy imaging and Romain Dagnelie for his help with surface area measurement

## Author Contributions

Conceived and designed the experiments: SP JPR SC CM JL. Performed the experiments: CM SD JCA GL VL MHM JPR SP YB JL. Analyzed the data: CM SD JCA GL VL MHM JPR SP YB JL. Wrote the manuscript: CM SD JCA SP YB JPR JL.

## References

- Norde W (2008) My voyage of discovery to proteins in flatland ...and beyond. *Colloids Surf B Biointerfaces* 61: 1–9. doi:10.1016/j.colsurfb.2007.09.029. PubMed: 18023976.
- Janson JC, Jönsson JA (1989) Introduction to chromatography. In: JC Janson, L Ryden. *Protein purification, principles, high resolution methods and applications*. New York. pp. 35–62.
- Gifford R, Kehoe JJ, Barnes SL, Kornilayev BA, Alterman MA et al. (2006) Protein interactions with subcutaneously implanted biosensors. *Biomaterials* 27: 2587–2598. doi:10.1016/j.biomaterials.2005.11.033. PubMed: 16364432.
- Wang K, Zhou C, Hong Y, Zhang X (2012) A review of protein adsorption on bioceramics. *Interface Focus* 2: 259–277. doi:10.1098/rsfs.2012.0012. PubMed: 23741605.
- Peng Q, Zhang S, Yang Q, Zhang T, Wei X-Q et al. (2013) Preformed albumin corona, a protective coating for nanoparticles based drug delivery system. *Biomaterials* 34: 8521–8530. doi:10.1016/j.biomaterials.2013.07.102. PubMed: 23932500.
- Banerjee I, Pangule RC, Kane RS (2011) Antifouling Coatings: Recent Developments in the Design of Surfaces That Prevent Fouling by Proteins, Bacteria, and Marine Organisms. *Adv Mater* 23: 690–718. PubMed: 20886559.
- Nel AE, Mädler L, Velegol D, Xia T, Hoek EMV et al. (2009) Understanding biophysicochemical interactions at the nano-bio interface. *Nat Mater* 8: 543–557. doi:10.1038/nmat2442. PubMed: 19525947.
- Walkey CD, Chan WCW (2012) Understanding and controlling the interaction of nanomaterials with proteins in a physiological environment. *Chem Soc Rev* 41: 2780–2799. doi:10.1039/c1cs15233e. PubMed: 22086677.
- Aggarwal P, Hall JB, Mcleland CB, Dobrovolskaia MA, Mcneil SE (2009) Nanoparticle interaction with plasma proteins as it relates to particle biodistribution, biocompatibility and therapeutic efficacy. *Adv Drug Deliv Rev* 61: 428–437. doi:10.1016/j.addr.2009.03.009. PubMed: 19376175.
- Göppert TM, Müller RH (2005) Polysorbate-stabilized solid lipid nanoparticles as colloidal carriers for intravenous targeting of drugs to the brain: comparison of plasma protein adsorption patterns. *J Drug Target* 13: 179–187. doi:10.1080/10611860500071292. PubMed: 16036306.
- Foot M (2007) Using nanotechnology to improve the characteristics of antineoplastic drugs: improved characteristics of nab-paclitaxel compared with solvent-based paclitaxel. *Biotechnol Annu Rev* 13: 345–357. doi:10.1016/S1387-2656(07)13012-X. PubMed: 17875482.
- Palaniyar N, Ridsdale RA, Hearn SA, Heng YM, Ottensmeyer FP et al. (1999) Filaments of surfactant protein A specifically interact with corrugated surfaces of phospholipid membranes. *Am J Physiol* 276: L631–L641. PubMed: 10198361.
- Müller DJ, Dufrène YF (2011) Atomic force microscopy: a nanoscopic window on the cell surface. *Trends Cell Biol* 21: 461–469. doi:10.1016/j.tcb.2011.04.008. PubMed: 21664134.
- Scopelliti PE, Borronovo A, Indrieri M, Giorgetti L, Bongiorno G et al. (2010) The Effect of Surface Nanometre-Scale Morphology on Protein Adsorption - PLOS ONE 5: e11862. doi:10.1371/journal.pone.0011862.
- Andrade J, Hlady V (1986) Protein adsorption and materials biocompatibility: a tutorial review and suggested hypotheses. Springer. Berlin, Heidelberg. 63 pp.
- Fenoglio I, Fubini B, Ghibaudi EM, Turci F (2011) Multiple aspects of the interaction of biomacromolecules with inorganic surfaces. *Adv Drug Deliv Rev* 63: 1186–1209. doi:10.1016/j.addr.2011.08.001. PubMed: 21871508.
- Ikeda T, Kuroda A (2011) Why does the silica-binding protein “Si-tag” bind strongly to silica surfaces? Implications of conformational adaptation of the intrinsically disordered polypeptide to solid surfaces. *Colloids Surf B Biointerfaces* 86: 359–363. doi:10.1016/j.colsurfb.2011.04.020. PubMed: 21592750.
- Shemetov AA, Nabiev I, Sukhanova A (2012) Molecular interaction of proteins and peptides with nanoparticles. *ACS Nano* 6: 4585–4602. doi:10.1021/nn300415x. PubMed: 22621430.
- Norde W, Favier JP (1992) Structure of adsorbed and desorbed proteins. *Colloids and Surfaces* 64: 87–93. doi:10.1016/0166-6622(92)80164-W.
- Nakanishi K, Sakiyama T, Imamura K (2001) On the adsorption of proteins on solid surfaces, a common but very complicated phenomenon. *J Biosci Bioeng* 91: 233–244. doi:10.1263/jbb.91.233. PubMed: 16232982.
- Orts-Gil G, Natte K, Drescher D, Bresch H, Mantion A et al. (2010) Characterisation of silica nanoparticles prior to in vitro studies: from primary particles to agglomerates. *J Nanopart Res* 13: 1593–1604.
- Langmuir I (1918) The adsorption of gases on plane surfaces of glass, mica and platinum. *J Am Chem Soc* 40: 1361–1403. doi:10.1021/ja02242a004.
- McNaught AD, Wilkinson AR (1997) *Compendium of Chemical Terminology*. Wiley-Blackwell. p. 1.
- Tenzen S, Docter D, Rosfa S, Wlodarski A, Kuharev J et al. (2011) Nanoparticle Size Is a Critical Physicochemical Determinant of the Human Blood Plasma Corona: A Comprehensive Quantitative Proteomic Analysis - ACS Nano 5: 7155–7167. doi:10.1021/nn201950e.
- Maillet I, Lagniel G, Perrot M, Boucherie H, Labarre J (1996) Rapid identification of yeast proteins on two-dimensional gels. *J Biol Chem* 271: 10263–10270. doi:10.1074/jbc.271.17.10263. PubMed: 8626593.
- Baudouin-Cornu P, Lagniel G, Chédin S, Labarre J (2009) Development of a new method for absolute protein quantification on 2-D gels. *Proteomics* 9: 4606–4615. doi:10.1002/pmic.200800975. PubMed: 19750513.
- Anderson NL (2002) The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics* 1: 845–867. PubMed: 12488461.
- Kubiak K, Mulheran PA (2009) Molecular Dynamics Simulations of Hen Egg White Lysozyme Adsorption at a Charged Solid Surface. *J Phys Chem B* 113: 12189–12200. doi:10.1021/jp901521x. PubMed: 19518099.
- Karlin S, Zuker M, Brocchieri L (1994) Measuring residue associations in protein structures. Possible implications for protein folding. *J Mol Biol* 239: 227–248. doi:10.1006/jmbi.1994.1365. PubMed: 8196056.
- Musafia B, Buchner V, Arad D (1995) Complex salt bridges in proteins: statistical analysis of structure and function. *J Mol Biol* 254: 761–770. doi:10.1006/jmbi.1995.0653. PubMed: 7500348.
- Gallivan JP, Dougherty DA (1999) Cation- $\pi$  interactions in structural biology. *Proc Natl Acad Sci U S A* 96: 9459–9464. doi:10.1073/pnas.96.17.9459. PubMed: 10449714.
- Vlasova NN, Golovkova LP (2004) The adsorption of amino acids on the surface of highly dispersed silica. *Colloid Journal* 66: 657–662. doi:10.1007/s10595-005-0042-3.
- Hoefling M, Monti S, Corni S, Gottschalk KE (2011) Interaction of  $\beta$ -sheet folds with a gold surface. *PLOS ONE* 6: e20925. doi:10.1371/journal.pone.0020925. PubMed: 21687744.
- Borders CL, Broadwater JA, Bekeney PA, Salmon JE, Lee AS et al. (1994) A structural role for arginine in proteins: multiple hydrogen

- bonds to backbone carbonyl oxygens. *Protein Sci* 3: 541–548. PubMed: 8003972.
35. Patwardhan SV, Emami FS, Berry RJ, Jones SE, Naik RR et al. (2012) Chemistry of aqueous silica nanoparticle surfaces and the mechanism of selective peptide adsorption. *J Am Chem Soc* 134: 6244–6256. doi: 10.1021/ja211307u. PubMed: 22435500.
  36. Gao Q, Xu W, Xu Y, Wu D, Sun Y et al. (2008) Amino acid adsorption on mesoporous materials: influence of types of amino acids, modification of mesoporous materials, and solution conditions. *J Phys Chem B* 112: 2261–2267. doi:10.1021/jp0763580. PubMed: 18217746.
  37. Willett RL, Baldwin KW, West KW, Pfeiffer LN (2005) Differential adhesion of amino acids to inorganic surfaces. *Proc Natl Acad Sci U S A* 102: 7817–7822. doi:10.1073/pnas.0408565102. PubMed: 15901900.
  38. Norde W, Anusiem ACI (1992) Adsorption, desorption and re-adsorption of proteins on solid surfaces. *Colloids and Surfaces* 66: 73–80. doi:10.1016/0166-6622(92)80122-1.
  39. Tanford C (1962) Contribution of hydrophobic interactions to the stability of the globular conformation of proteins. *J Am Chem Soc* 84: 4240–4247. doi:10.1021/ja00881a009.
  40. Kellis JT, Nyberg K, Fersht AR (1989) Energetics of complementary side-chain packing in a protein hydrophobic core. *Biochemistry* 28: 4914–4922. doi:10.1021/bi00437a058. PubMed: 2669964.
  41. Richards FM, Lim WA (1993) An analysis of packing in the protein folding problem. *Q Rev Biophys* 26: 423–498. doi:10.1017/S0033583500002845. PubMed: 8058892.
  42. Pace CN, Fu H, Fryar KL, Landua J, Trevino SR et al. (2011) Contribution of hydrophobic interactions to protein stability. *J Mol Biol* 408: 514–528. doi:10.1016/j.jmb.2011.02.053. PubMed: 21377472.
  43. Brocchieri L, Karlin S (1995) How are close residues of protein structures distributed in primary sequence? *Proc Natl Acad Sci U S A* 92: 12136–12140. doi:10.1073/pnas.92.26.12136. PubMed: 8618859.
  44. Baud F, Karlin S (1999) Measures of residue density in protein structures. *Proc Natl Acad Sci U S A* 96: 12494–12499. doi:10.1073/pnas.96.22.12494. PubMed: 10535950.
  45. Vondrášek J, Bendová L, Klusák V, Hobza P (2005) Unexpectedly strong energy stabilization inside the hydrophobic core of small protein rubredoxin mediated by aromatic residues: correlated ab initio quantum chemical calculations. *J Am Chem Soc* 127: 2615–2619. doi: 10.1021/ja044607h. PubMed: 15725017.
  46. Burley SK, Petsko GA (1985) Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science* 229: 23–28. doi: 10.1126/science.3892686. PubMed: 3892686.
  47. Anderson DE, Hurley JH, Nicholson H, Baase WA, Matthews BW (1993) Hydrophobic core repacking and aromatic-aromatic interaction in the thermostable mutant of T4 lysozyme Ser 117→Phe. *Protein Sci* 2: 1285–1290.
  48. Hecht MH, Sturtevant JM, Sauer RT (1984) Effect of single amino acid replacements on the thermal stability of the NH<sub>2</sub>-terminal domain of phage lambda repressor. *Proc Natl Acad Sci U S A* 81: 5685–5689. doi:10.1073/pnas.81.18.5685. PubMed: 6237363.
  49. Kannan N, Vishveshwara S (2000) Aromatic clusters: a determinant of thermal stability of thermophilic proteins. *Protein Eng* 13: 753–761. doi: 10.1093/protein/13.11.753. PubMed: 11161106.
  50. Lanzarotti E, Biekofsky RR, Estrin DA, Marti MA, Turjanski AG (2011) Aromatic-aromatic interactions in proteins: beyond the dimer. *J Chem Inf Model* 51: 1623–1633. doi:10.1021/ci200062e. PubMed: 21662246.
  51. Brunauer S, Emmett PH, Teller E (1938) Adsorption of gases in multimolecular layers. *J Am Chem Soc* 60: 309–319. doi:10.1021/ja01269a023.
  52. Hunter RJ (1981) Zeta potential in colloid science: principles and applications. Academic Press. London, New York. 386 pp.
  53. Mathon MH, De Novion CH (1999) De l'intensité à la structure des matériaux. *Le Journal de Physique IV* 09: Pr1–127–Pr1–146.
  54. Mortimer RK, Johnston JR (1986) Genealogy of principal strains of the yeast genetic stock center. *Genetics* 113: 35–43. PubMed: 3519363.
  55. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72: 248–254. doi: 10.1016/0003-2697(76)90527-3. PubMed: 942051.
  56. Boucherie H, Sagliocco F, Joubert R, Maillet I, Labarre J et al. (1996) Two-dimensional gel protein database of *Saccharomyces cerevisiae*. *Electrophoresis* 17: 1683–1699. doi:10.1002/elps.1150171106. PubMed: 8982601.
  57. Shevchenko A, Jensen ON, Podtelejnikov AV, Sagliocco F, Wilm M et al. (1996) Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. *Proc Natl Acad Sci U S A* 93: 14440–14445. doi:10.1073/pnas.93.25.14440. PubMed: 8962070.
  58. Shevchenko A, Tomas H, Havli sbreve J, Olsen JV, Mann M (2007) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* 1: 285–2860.
  59. R Development Core Team (2011) R: A Language and Environment for Statistical Computing. Vienna, Austria. Available: <http://www.r-project.org/>. Accessed 2013 October 24.
  60. Li ZR, Lin HH, Han LY, Jiang L, Chen X et al. (2006) PROFEAT: a web server for computing structural and physicochemical features of proteins and peptides from amino acid sequence. *Nucleic Acids Res* 34: W32–W37. doi:10.1093/nar/gkj014. PubMed: 16845018.
  61. Brunner E, Munzel U (2000) The nonparametric Behrens-Fisher problem: asymptotic theory and a small-sample approximation. *Biometrical Journal* 42: 17–25. doi:10.1002/(SICI)1521-4036(200001)42:1.
  62. Noguchi K, Noguchi K, Hui WLW, Hui WLW, Gel YR et al. (2009) lawstat: An R package for biostatistics, public policy, and law.
  63. Jarque CM, Bera AK (1980) Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Econ Lett* 5: 177–82. PubMed: 12312014.
  64. Brown MB, Forsythe AB (1974) Robust Tests for the Equality of Variances. *Journal of the American Statistical Association* 69: 364–367. doi:10.1080/01621459.1974.10482955.
  65. Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *J Roy Stat Soc B Met* 57: 289–300.
  66. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R et al. (2005) The Amber biomolecular simulation programs. *J Comput Chem* 26: 1668–1688. doi:10.1002/jcc.20290. PubMed: 16200636.
  67. Protein Data Bank. Available: <http://www.pdb.org/>. Accessed 2013 October 24.
  68. Protein Portal Model. Available: <http://www.proteinmodelportal.org/>. Accessed 2013 October 24.
  69. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. *J Chem Phys* 81: 3684–3690. doi:10.1063/1.448118.
  70. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J Chem Phys* 98: 10089–10092. doi:10.1063/1.464397.
  71. Ryckaert J-P, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics* 23: 327–341. doi:10.1016/0021-9991(77)90098-5.
  72. Protein contact map visualization and analysis (CMview) Available: <http://www.bioinformatics.org/cmview/>. Accessed 2013 October 24.
  73. Naccess. Available: <http://www.bioinf.manchester.ac.uk/naccess/>. Accessed 2013 October 24.
  74. Lee B, Richards FM (1971) The interpretation of protein structures: estimation of static accessibility. *J Mol Biol* 55: 379–400. doi: 10.1016/0022-2836(71)90324-X. PubMed: 5551392.