OPEN

Aggregated Gas Molecules: Toxic to Protein?

Meng Zhang[1,2], Guanghong Zuo[1], Jixiu Chen[3], Yi Gao[1] & Haiping Fang[1]

[1]Division of Interfacial Water and Laboratory of Physical Biology, Shanghai Institute of Applied Physics, Chinese Academy of Sciences, P.O. Box 800-204, Shanghai 201800, China, [2]University of the Chinese Academy of Sciences, Beijing 100049, China, [3]Department of Infectious Disease, Changhai Hospital, the Second Military Medical University - Shanghai 200433, China.

The biological toxicity of high levels of breathing gases has been known for centuries, but the mechanism remains elusive. Earlier work mainly focused on the influences of dispersed gas molecules dissolved in water on biomolecules. However, recent studies confirmed the existence of aggregated gas molecules at the water-solid interface. In this paper, we have investigated the binding preference of aggregated gas molecules on proteins with molecular dynamics simulations, using nitrogen (N_2) gas and the Src-homology 3 (SH3) domain as the model system. Aggregated N_2 molecules were strongly bound by the active sites of the SH3 domain, which could impair the activity of the protein. In contrast, dispersed N_2 molecules did not specifically interact with the SH3 domain. These observations extend our understanding of the possible toxicity of aggregates of gas molecules in the function of proteins.

An increased concentration of breathing gases (N_2, O_2, Ar, Kr, CO_2, Xe, etc.) in body tissues can cause severe biological effects, such as nitrogen narcosis[1–4]. Although some mechanisms have been proposed, such as the gas-induced alteration of ion permeability at the cell membrane or the binding of gas to proteins of neurotransmitter receptors[5–7], the mode of action remains unresolved. Earlier work mainly focused on the influences of dispersed gas molecules dissolved in water on biomolecules. The recent experimental and theoretical studies demonstrated that gas can be aggregated at the water-solid interface in the form of nanobubbles[8–18] and nanopancakes[19,20] after the great debate[21–23]. Thus, the interaction between aggregated gas molecules, as opposed to dispersed gas molecules, and biological macromolecules is an important area of investigation.

The interactions between nanoscale particles and biomolecules have attracted great attention. Previous studies have shown that nanoparticles can adsorb onto proteins[24–41]. If the protein-ligand binding sites are attacked or blocked by the nanoparticles, the biological functions of these proteins could be impaired. Considering that the binding sites and many of the aggregated gases are usually hydrophobic[42–45], we questioned if aggregated gas molecules showed similar toxicity toward proteins as observed with nanoparticles.

In this paper, we applied molecular dynamics simulations to study the adsorption capability of aggregated gas onto the Src-homology 3 (SH3) domain. The SH3 domain was chosen due to its widespread presence in protein families for regulating signaling pathways and for substrate specific-binding[46–52] (Figure 1). Figure 1 was generated using the VASCo PYMOL plug-in[53]. The functional loss of the SH3 domain would be lethal for the normal functions of living cells. As the gas, N_2 was employed due to its abundance in air. We found that aggregated N_2 molecules preferred to adhere to the SH3 domain at the ligand-binding sites of the SH3 domain. In contrast, the dispersed N_2 molecules showed no specific binding to the SH3 domain.

Results

In the initial configuration of the system, N_2 molecules were placed around the SH3 domain in the water box with box size of L_x = 7.5 nm, L_y = 7.5 nm, L_z = 7.5 nm (Figure 2). Different numbers of N_2 molecules were examined in six samples with 100 ns molecular dynamics simulations under the NPT (P = 1 atm, T = 300 K) ensemble.

In the first four samples, 330, 310, 290 and 270 nitrogen molecules were placed in the cubic water box, denoted by NS_1, NS_2, NS_3 and NS_4, respectively. The N_2 molecules were found to be aggregated and we denoted them as NS. In the other two samples, denoted by DS_1 and DS_2, with 216 and 205 N_2 molecules, respectively, the maximum aggregate size was no more than 6 N_2 molecules, and thus we denoted them as DS. The details of all samples appear in Table 1. The criterion to determine the existence of aggregated N_2 molecules followed the Wolde and Frenkel (TWF) theory[54]. Two N_2 molecules were considered as a neighboring pair if their distance was less than r_S (the Stillinger radius, in our systems, the r_S = $1.5\sigma_{NN}$ = 0.489 nm)[54,55]. If a N_2 molecule had more than
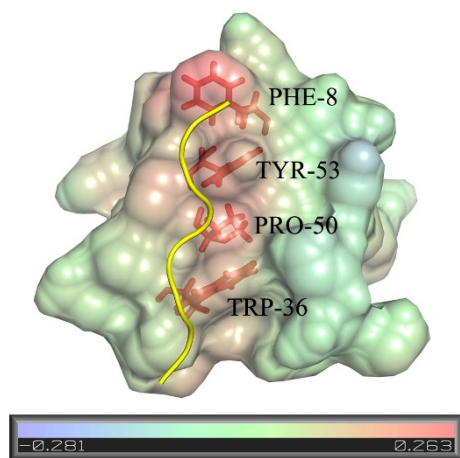
Figure 1 | A typical structure of the SH3 domain bound to its ligand. The yellow loop represents the corresponding ligand of the SH3 domain. The four key binding sites (8(PHE), 36(TRP), 50(PRO), 53(TYR)) of the SH3 domain are plotted as red sticks. The surface of the SH3 domain is shown in PYMOL-CGOs and colored according to lipophilic potential. The hydrophobic surface is indicated by red and the hydrophilic surface by blue.

two neighbors, this N_2 belonged to the aggregated N_2 molecules. More computational details are listed in the Methods section.

Representative snapshots of the trajectory for NS_1 and the change of the interface area of this sample are given in Figure 3. (For other samples, see Figure S3 in supplementary information). The interface area was defined as:

$$\text{Interface Area} = (\text{SASA}_{\text{SH3}} + \text{SASA}_{\text{NIT}} - \text{SASA}_{\text{SH3+NIT}})/2.$$

The SASA_{SH3}, SASA_{NIT} and $\text{SASA}_{\text{SH3+NIT}}$ were defined as the solvent accessible surface areas of the SH3 domain, the N_2 molecules, and the complex of the SH3 domain and the N_2 molecules, respectively[56]. Based on the snapshots and interface area, we could illustrate the whole simulation process. For NS situations, in the first ~5 ns, N_2 molecules gathered quickly, corresponding to a decline in the interface area to nearly zero. Then, the aggregated N_2 molecules were separated from the SH3 domain for about 25 ns. Next, the interface area rose quickly to about 3 nm², which was caused by a rapid adsorption of aggregated N_2 onto the SH3 domain. This increase in the interface area indicated a strong interaction between aggregated N_2 molecules and the hydrophobic surface of the SH3 domain. Then, the aggregated N_2 molecules attached the SH3 domain in the rest of the simulation (see the snapshots at t = 35 ns and t = 100 ns). There were some large fluctuations in the curve of the interface area, owing to the quite loose and flexible structure of the aggregated N_2 molecules.

Discussion

To further characterize the adsorption between the N_2 and the SH3 domain, we calculated the probability of the adsorption of N_2 molecules to each amino acid residue in the SH3 domain over the last 50 ns. The distances between the amino acids and N_2 molecules were evaluated in every frame (In simulations, we recorded one frame of trajectory every 1 ps). If the minimum distance of one amino acid to any N_2 molecule was less than r_S ($r_S = 1.5\sigma_{NN} = 0.489$ nm) in one frame, we defined the phenomenon as the adsorption of N_2 molecules by this particular amino acid in this specific time frame and the adsorption frequency of this amino acid added 1. The adsorption probability of a specific amino acid was the adsorption frequency of this amino acid divided by the total frame number (50,000 in our simulations).

The typical adsorption probabilities of each residue are illustrated by a histogram (Figure 4a) showing data for NS and DS (Figures of
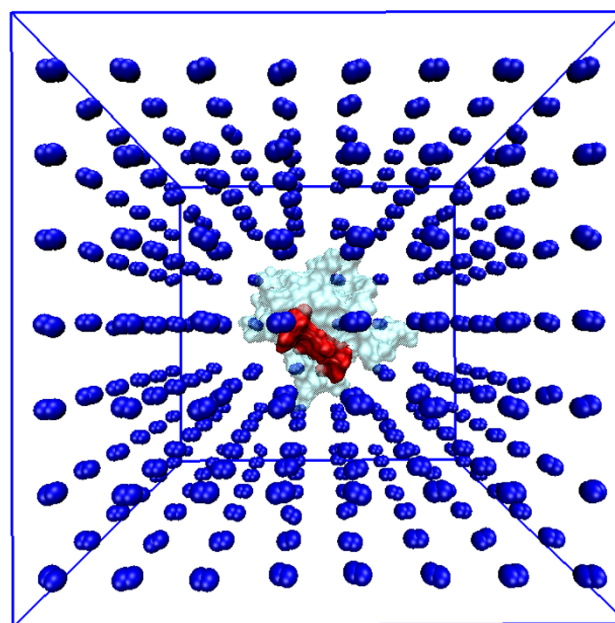

Figure 2 | The initial configuration of the simulation. N_2 molecules were placed around SH3 in the water box with box size of $L_x = 7.5$ nm, $L_y = 7.5$ nm, $L_z = 7.5$ nm. The key residues of the protein domain are noted by the red surfaces. The rest of residues of protein are shown as the cyan semitransparent surface. N_2 molecules are shown as blue balls.

more samples can be found in Figure S2 in supplementary information). A high probability means strong adsorption. Additional details appear in Table 1, which contains the favorable N_2-binding residues, defined as those residues with a probability of greater than 0.7 of all samples. In the first four samples (NS_1–NS_4 in Table 1), the SH3 domain had the same favorable binding sites with the aggregated N_2 molecules. The amino acid residues 7(LEU), 8(PHE), 35(GLN), 36(TRP), 50(PRO), 52(PRO) and 53(TYR) presented more favorable N_2 binding in NS cases. It is known that 8(PHE), 36(TRP), 50(PRO) and 53(TYR) (Figure 1 and Table 1) are the key binding sites of the SH3 domain and its corresponding ligand[51]. Thus, we can conclude that the binding of the aggregated N_2 molecules to the SH3 domain had specificity for the active sites. Figure 1 shows a typical structure of the ligand binding to the SH3 domain. The hydrophobicity of the surface was plotted by lipophilic potential and the surface of the SH3 domain in contact with the ligand was hydrophobic[57]. Since the aggregated N_2 molecules completely covered the ligand-binding sites of the SH3 domain, aggregates of nitrogen might block the key binding sites and hinder the normal process of ligand binding at the SH3 domain. In another words, binding of aggregated N_2 molecules might lead to the loss of protein function. Because proteins were the functional units essential for life, our simulation indicated that aggregated gas might be toxic.

In the last two samples (DS_1 and DS_2) in Table 1, N_2 molecules did not aggregate, a completely different outcome that with the NS samples. Another difference was an absence of probabilities greater than 0.7 for the most active binding sites (8, 36, 50 and 53) for DS, which means the adsorption abilities of DS were much weaker than the adsorption abilities of NS at these sites. More importantly, a greater number of N_2 molecules did not correspond to the high adsorption probabilities when N_2 molecules were dispersed, which could be illustrated by DS_1 and DS_2 with 216 N_2 and 205 N_2, respectively (shown in Figure S2 in supplementary information). The differences in the probabilities of binding could only be from the existence of the aggregated N_2, as shown in Table 1. To further illustrate the difference between DS and NS, we calculated the

Table 1 | Summary of simulations. The left-most column shows the source of data. N_{water} is the total number of water molecules. N_T, N_N and N_D mean the total number of N_2 molecules, the final number of N_2 molecules in the aggregated state and the final number of N_2 molecules in the dispersed state, respectively. The numbers in the middle of the table are the index numbers of residues whose probabilities to adsorb N_2 molecules are greater than 0.7 (defined in Figure 4a). The bold numbers (8, 36, 50, 53) in table represent the key binding sites of the SH3 domain. σ is the standard deviation of each residue's adsorption probability to N_2. L represents the final length of the cubic box. C is the final concentration of N_2 molecules. ''Ref'' indicates the active binding sites in reference 51

Sample	N_{water}	N_T	N_N	N_D	Index number of residues									σ	L (nm)	C (kg·m^{-3})
NS_1	12834	330	220	110	7	**8**	10	35	**36**	**50**	52	**53**		0.298	7.64	34.6
NS_2	12834	310	204	106	7	**8**	10	35	**36**	**50**	52	**53**		0.301	7.61	32.9
NS_3	12834	290	190	100	7	**8**		35	**36**	**50**	52	**53**		0.255	7.60	30.9
NS_4	12834	270	174	96	7	**8**	10	35	**36**	**50**	52	**53**		0.298	7.57	29.1
DS_1	13106	216	6	210									56	0.175	7.52	23.8
DS_2	13106	205	5	200								**53**		0.189	7.52	22.6
Ref						**8**			**36**	**50**		**53**				

standard deviation of each residue's adsorption probability to N_2 (Table 1). All standard deviations of NS were greater than 0.25, whereas none of those of DS was greater than 0.19. Furthermore, the distribution of probability of DS tended to be well-distributed compared with NS, which indicated the adsorption of dispersed N_2 molecules to the SH3 domain had no specific binding sites. This fact implied that the characteristic binding sites were related to the state of N_2. Only when the N_2 molecules were aggregated could they adsorb well onto the SH3 domain at some specific sites.

Finally, it should be noted that the SH3 domain is a small domain with 57 residues. Each residue in SH3 could be in contact with water

and N_2 molecules. The calculated exposed area of each residue in water is shown in Figure 4b. Compared with Figure 4a, the residues with the largest exposed areas (15(GLU), 43(GLU) and 57(TYR)) did not correspond to large adsorption probabilities. On the other hand, the exposed area of the key residue 50(PRO) was very small even though its adsorption probability was very large. Considering that the four key residues (8(PHE), 36(TRP), 50(PRO) and 53(TYR)) are all located at the hydrophobic groove (red surface in Figure 1) of the SH3 domain, the ability of each residue to adsorb nitrogen was related to the hydrophobicity of the exposed residues rather than to the exposed areas of these residues.
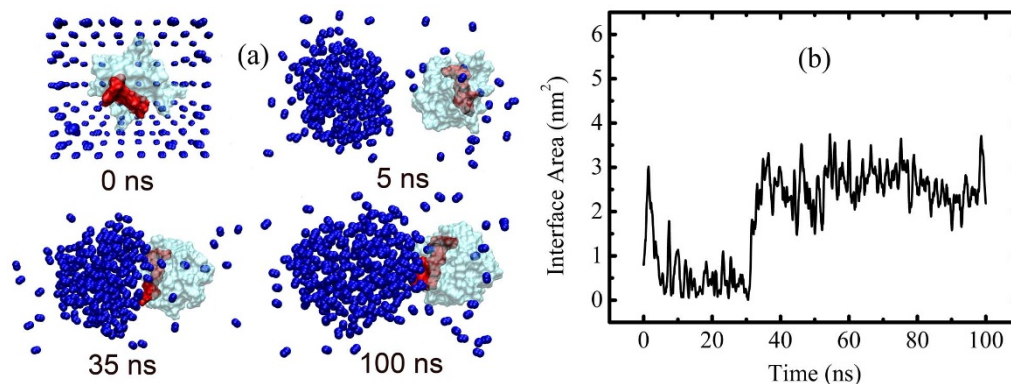

Figure 3 | (a) Some representative snapshots of one trajectory (NS_1). The key residues of the protein domain are noted by red surfaces. The rest of residues of protein are shown in cyan semi-transparent surface. The N_2 are shown as blue balls. (b) The change of interface area of N_2 molecules and the SH3 domain corresponding to the trajectory shown in Figure 3a.
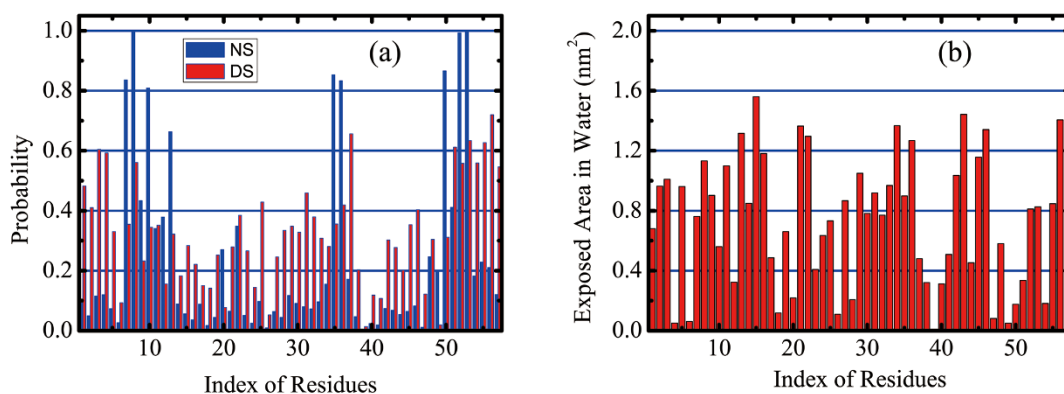

Figure 4 | (a) A typical adsorption probability plot of each residue. The X-axis represents the index of residues in the SH3 domain while the Y-axis represents the probability that the minimum distance of the amino acid to any one nitrogen molecule will be less than 0.489 nm in the last 50 ns. The blue rectangles stand for the probabilities of the aggregated N_2 system (NS_1), and the red rectangles stand for the probabilities of the dispersed N_2 system (DS_1). (b) The exposed area in water of each residue of the SH3 domain.

3

In general, N_2 molecules aggregated and were adsorbed onto the surface of the SH3 domain at several specific binding sites. Considering non-specific adsorptions of dispersed N_2 molecules in the system, this binding could be concluded as the hydrophobic interactions between aggregated N_2 molecules and the hydrophobic residues in the key binding sites of the SH3 domain. Furthermore, the occupation of the binding sites of the SH3 domain by aggregated N_2 molecules would interfere with ligand binding at these same positions, which would poison the protein function. Our study might provide insights into the potentially common phenomenon of the toxicity of aggregated gases in the biological systems and could provide an alternative way to understand the gas effects in anesthesia.

Methods

The molecular dynamics simulations were carried out with the Gromacs 4.5.1 program package[58]. In simulations, Amber03 force field was applied in protein and water, the nitrogen atoms were modeled as uncharged Lennard-Jones particles with a radius of cross-section of $\sigma_{NN} = 3.260 \times 10^{-1}$ nm, and a depth of the potential well of $\varepsilon_{NN} = 2.888 \times 10^{-1}$ kJ·mol^{-1} [59]. The N-N bond lengths of 1.400×10^{-1} nm were maintained by harmonic potentials with spring constants of 1.280×10^6 kJ/mol·nm^2 [59]. The pdb file of the SH3 domain in the simulations was downloaded from RCSB Protein Data Bank[50]. Simulations were performed under the isothermal-isobaric (NPT, 1 atm and 310 K to mimic in vivo condition) ensemble with periodic boundary conditions applied in three dimensions. A time step of 1 fs was used, and data were collected every 1 ps. The particle mesh Ewald method was used to treat the long-range electrostatic interactions with cut-off of 1 nm[60].

1. Fenn, W. O. Inert gas narcosis. *Ann. NY Acad.Sci.* **117**, 760–767 (1965).
2. Miller, K. W. Inert-gas narcosis, high-pressure neurological syndrome, and critical volume hypothesis. *Science* **185**, 867–869 (1974).
3. Fowler, B., Ackles, K. N. & Porlier, G. Effects of inert-gas narcosis on behavior - a critical-review. *Undersea Biomedical Research* **12**, 369–402 (1985).
4. Bennett, P. B. & Rostain, J. C. Inert gas narcosis. *Bennett and Elliott's Physiology and Medicine of Diving, edited by Brubakk AO and Neumann TS (5th ed.)* 300–322 (2003).
5. Franks, N. P. & Lieb, W. R. *Molecular and cellular mechanisms of general anaesthesia* (2000).
6. Trudell, J. R., Koblin, D. D. & Eger, E. I. A molecular description of how noble gases and nitrogen bind to a model site of anesthetic action. *Anesthesia and Analgesia* **87**, 411–418 (1998).
7. Rostain, J. C. & Balon, N. Recent neurochemical basis of inert gas narcosis and pressure effects. *Undersea & Hyperbaric Medicine* **33**, 197–204 (2006).
8. Lou, S. T. *et al.* Nanobubbles on solid surface imaged by atomic force microscopy. *Journal of Vacuum Science & Technology B* **18**, 2573–2575 (2000).
9. Ishida, N., Inoue, T., Miyahara, M. & Higashitani, K. Nano bubbles on a hydrophobic surface in water observed by tapping-mode atomic force microscopy. *Langmuir* **16**, 6377–6380 (2000).
10. Koishi, T. *et al.* Nanoscale hydrophobic interaction and nanobubble nucleation. *Phys. Rev. Lett.* **93**, 4 (2004).
11. Agrawal, A. *et al.* Controlling the location and spatial extent of nanobubbles using hydrophobically nanopatterned surfaces. *Nano Lett.* **5**, 1751–1756 (2005).
12. Zhang, X. H., Khan, A. & Ducker, W. A. A nanoscale gas state. *Phys. Rev. Lett.* **98**, 136101–136104 (2007).
13. Li, Z. X. *et al.* Analysis of the gas states at a liquid/solid interface based on interactions at the microscopic level. *J. Phys. Chem. B* **111**, 9325–9329 (2007).
14. Zhang, L., Chen, H., Li, Z., Fang, H. & Hu, J. Long lifetime of nanobubbles due to high inner density. *Science in China Series G-Physics Mechanics & Astronomy* **51**, 219–224 (2008).
15. Ducker, W. A. Contact angle and stability of interfacial nanobubbles. *Langmuir* **25**, 8907–8910 (2009).
16. Craig, V. S. J. Very small bubbles at surfaces-the nanobubble puzzle. *Soft Matter* **7**, 40–48 (2011).
17. Karpitschka, S. *et al.* Nonintrusive optical visualization of surface nanobubbles. *Phys. Rev. Lett.* **109**, 066102 (2012).
18. Ball, P. Nanobubbles are not a superficial matter. *ChemPhysChem* **13**, 2173–2177 (2012).
19. Zhang, X. H. *et al.* Detection of novel gaseous states at the highly oriented pyrolytic graphite-water interface. *Langmuir* **23**, 1778–1783 (2007).
20. Hampton, M. A. & Nguyen, A. V. Accumulation of dissolved gases at hydrophobic surfaces in water and sodium chloride solutions: Implications for coal flotation. *Minerals Engineering* **22**, 786–792 (2009).
21. Parker, J. L., Claesson, P. M. & Attard, P. Bubbles, cavities, and the long-ranged attraction between hydrophobic surfaces. *J. Phys. Chem.* **98**, 8468–8480 (1994).
22. Doshi, D. A., Watkins, E. B., Israelachvili, J. N. & Majewski, J. Reduced water density at hydrophobic surfaces: Effect of dissolved gases. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 9458–9462 (2005).
23. Kikuchi, K. *et al.* Concentration determination of oxygen nanobubbles in electrolyzed water. *J. Colloid Interface Sci.* **329**, 306–309 (2009).
24. Park, K. H., Chhowalla, M., Iqbal, Z. & Sesti, F. Single-walled carbon nanotubes are a new class of ion channel blockers. *J. Biol. Chem.* **278**, 50212–50216 (2003).
25. Karajanagi, S. S., Vertegel, A. A., Kane, R. S. & Dordick, J. S. Structure and function of enzymes adsorbed onto single-walled carbon nanotubes. *Langmuir* **20**, 11594–11599 (2004).
26. Cedervall, T. *et al.* Detailed identification of plasma proteins adsorbed on copolymer nanoparticles. *Angewandte Chemie-International Edition* **46**, 5754–5756 (2007).
27. Shen, J.-W., Wu, T., Wang, Q. & Kang, Y. Induced stepwise conformational change of human serum albumin on carbon nanotube surfaces. *Biomaterials* **29**, 3847–3855 (2008).
28. Zuo, G. H., Huang, Q., Wei, G. H., Zhou, R. H. & Fang, H. P. Plugging into proteins: Poisoning protein function by a hydrophobic nanoparticle. *ACS Nano* **4**, 7508–7514 (2010).
29. Zuo, G. H., Gu, W., Fang, H. P. & Zhou, R. H. Carbon nanotube wins the competitive binding over proline-rich motif ligand on sh3 domain. *J. Phys. Chem. C* **115**, 12322–12328 (2011).
30. Calzolai, L., Franchini, F., Gilliland, D. & Rossi, F. Protein-nanoparticle interaction: Identification of the ubiquitin-gold nanoparticle interaction site. *Nano Lett.* **10**, 3101–3105 (2010).
31. Brewer, S. H., Glomm, W. R., Johnson, M. C., Knag, M. K. & Franzen, S. Probing bsa binding to citrate-coated gold nanoparticles and surfaces. *Langmuir* **21**, 9303–9307 (2005).
32. Casanova, D. *et al.* Counting the number of proteins coupled to single nanoparticles. *J. Am. Chem. Soc.* **129**, 12592–12593 (2007).
33. De, M., You, C. C., Srivastava, S. & Rotello, V. M. Biomimetic interactions of proteins with functionalized nanoparticles: A thermodynamic study. *J. Am. Chem. Soc.* **129**, 10747–10753 (2007).
34. Dutta, D. *et al.* Adsorbed proteins influence the biological activity and molecular targeting of nanomaterials. *Toxicol. Sci.* **100**, 303–315 (2007).
35. Gessner, A., Lieske, A., Paulke, B. R. & Muller, R. H. Functional groups on polystyrene model nanoparticles: Influence on protein adsorption. *J. Biomed. Mater. Res. Part A* **65A**, 319–326 (2003).
36. Gref, R. *et al.* 'stealth' corona-core nanoparticles surface modified by polyethylene glycol (peg): Influences of the corona (peg chain length and surface density) and of the core composition on phagocytic uptake and plasma protein adsorption. *Colloid Surf. B-Biointerfaces* **18**, 301–313 (2000).
37. Hong, R. *et al.* Control of protein structure and function through surface recognition by tailored nanoparticle scaffolds. *J. Am. Chem. Soc.* **126**, 739–743 (2004).
38. Kaufman, E. D. *et al.* Probing protein adsorption onto mercaptoundecanoic acid stabilized gold nanoparticles and surfaces by quartz crystal microbalance and zeta-potential measurements. *Langmuir* **23**, 6053–6062 (2007).
39. Linse, S. *et al.* Nucleation of protein fibrillation by nanoparticles. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 8691–8696 (2007).
40. Lundqvist, M., Sethson, I. & Jonsson, B. H. Protein adsorption onto silica nanoparticles: Conformational changes depend on the particles' curvature and the protein stability. *Langmuir* **20**, 10639–10647 (2004).
41. Ge, C. C. *et al.* Binding of blood proteins to carbon nanotubes reduces cytotoxicity. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 16968–16973 (2011).
42. Privalov, P. L. & Gill, S. J. Stability of protein-structure and hydrophobic interaction. *Adv. Protein Chem.* **39**, 191–234 (1988).
43. Dill, K. A. Dominant forces in protein folding. *Biochemistry* **29**, 7133–7155 (1990).
44. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association - insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins-Structure Function and Genetics* **11**, 281–296 (1991).
45. Zhou, R. H., Huang, X. H., Margulis, C. J. & Berne, B. J. Hydrophobic collapse in multidomain protein folding. *Science* **305**, 1605–1609 (2004).
46. Koch, C. A., Anderson, D., Moran, M. F., Ellis, C. & Pawson, T. Sh2 and sh3 domains - elements that control interactions of cytoplasmic signaling proteins. *Science* **252**, 668–674 (1991).
47. Pawson, T. & Gish, G. D. Sh2 and sh3 domains - from structure to function. *Cell* **71**, 359–362 (1992).
48. Kuriyan, J. & Cowburn, D. Structures of sh2 and sh3 domains. *Curr. Opin. Struct. Biol.* **3**, 828–837 (1993).
49. Feller, S. M., Ren, R. B., Hanafusa, H. & Baltimore, D. Sh2 and sh3 domains as molecular adhesives - the interactions of crk and abl. *Trends in Biochemical Sciences* **19**, 453–458 (1994).
50. Wu, X. D. *et al.* Structural basis for the specific interaction of lysine-containing proline-rich peptides with the n-terminal sh3 domain of c-crk. *Structure* **3**, 215–226 (1995).
51. Ball, L. J., Kuhne, R., Schneider-Mergener, J. & Oschkinat, H. Recognition of proline-rich motifs by protein-protein-interaction domains. *Angew. Chem.-Int. Edit. Engl.* **44**, 2852–2869 (2005).
52. Ahmad, M., Gu, W. & Helms, V. Mechanism of fast peptide recognition by sh3 domains. *Angew. Chem.-Int. Edit. Engl.* **47**, 7626–7630 (2008).
53. Steinkellner, G., Rader, R., Thallinger, G. G., Kratky, C. & Gruber, K. Vasco: Computation and visualization of annotated protein surface contacts. *BMC Bioinformatics* **10**, 11 (2009).
54. ten Wolde, P. R. & Frenkel, D. Computer simulation study of gas-liquid nucleation in a lennard-jones system. *J. Chem. Phys.* **109**, 9901–9918 (1998).

55. Stillinger, F. H. Rigorous basis of frenkel-band theory of association equilibrium. *J. Chem. Phys.* **38**, 1486–1494 (1963).
56. Nadassy, K., Wodak, S. J. & Janin, J. Structural features of protein-nucleic acid recognition sites. *Biochemistry* **38**, 1999–2017 (1999).
57. Heiden, W., Moeckel, G. & Brickmann, J. A new approach to analysis and display of local lipophilicity hydrophilicity mapped on molecular-surfaces. *Journal of Computer-Aided Molecular Design* **7**, 503–514 (1993).
58. Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. Gromacs 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **4**, 435–447 (2008).
59. Rappe, A. K., Casewit, C. J., Colwell, K. S., Goddard, W. A. & Skiff, W. M. Uff, a full periodic-table force-field for molecular mechanics and molecular-dynamics simulations. *J. Am. Chem. Soc.* **114**, 10024–10035 (1992).
60. Essmann, U. *et al*. A smooth particle mesh ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).

Acknowledgements

Author contributions

Additional information