

## Arginine- and Lysine-Specific Polymers for Protein Recognition and Immobilization

Christian Renner,<sup>†</sup> Jacob Piehler,<sup>‡</sup> and Thomas Schrader<sup>\*†</sup>

Contribution from the Philipps-Universität Marburg, Fachbereich Chemie, Hans-Meerwein-Strasse, 35032 Marburg, Germany, and Johann Wolfgang Goethe-Universität Frankfurt, Institut für Biochemie, Marie Curie Strasse 9, 60439 Frankfurt, Germany

Received September 15, 2005; E-mail: schradet@staff.uni-marburg.de

**Abstract:** Free radical polymerization of methacrylamide-based bisphosphonates turns weak arginine binders into powerful polymeric protein receptors. Dansyl-labeled homo- and copolymers with excellent water solubility are accessible through a simple copolymerization protocol. Modeling studies point to a striking structural difference between the stiff rodlike densely packed homopolymer **1** and the flexible copolymer **2** with spatially separated bisphosphonate units. Fluorescence titrations in buffered aqueous solution (pH = 7.0) confirm the superior affinity of the homopolymer toward oligoarginine peptides reaching nanomolar  $K_D$  values for the Tat peptide. Basic proteins are bound almost equally well by **1** and **2** with micromolar affinities, with the latter producing much more soluble complexes. The Arg selectivity of the monomer is transferred to the polymer, which binds Arg-rich proteins 1 order of magnitude tighter than lysine-rich pendants of comparable pI, size, and (Arg/Lys vs Glu/Asp) ratio. Noncovalent deposition of both polymers on glass substrates via polyethyleneimine layers results in new materials suitable for peptide and protein immobilization. RfS measurements allow calculation of association constants  $K_a$  as well as dissociation kinetics  $k_D$ . They generally confirm the trends already found in free solution. Close inspection of electrostatic potential surfaces suggest that basic domains favor protein binding on the flat surface. The high specificity of the bisphosphonate polymers toward basic proteins is demonstrated by comparison with polyvinyl sulfate, which has almost no effect in RfS experiments. Thus, copolymerization of few different comonomer units without cross-linking enables surface recognition of basic proteins in free solution as well as their effective immobilization on surfaces.

### Introduction

Protein–protein interactions play a key role in numerous biological processes, characterized, inter alia, by hormone-receptor, protease-inhibitor, or antibody–antigen complexes.<sup>1</sup> Many proteins exist in dimeric or oligomeric states, with extended surface contact areas. The morphology and composition of such interfaces have been intensely studied in recent years, and systematic surveys in structural databases have disclosed some general trends and basic rules.<sup>2</sup> Earlier reports emphasize the factors hydrophobicity and complementarity; large hydrophobic patches on one protein's surface, surrounded by charged residues, are matched by its complexation partner.<sup>3</sup> Argos and many others found that the amino acid composition in protein interfaces does not resemble the nonpolar protein interior but is rather intermediate between interior and exterior. Notably, an unusually high number of arginines has repeatedly been reported in protein–protein complexes.<sup>4</sup> Especially, if these aggregates are nonobligatory and involve proteins that must exist

also independently (many enzyme–inhibitor or antibody–protein complexes), charged residues play a major role, since they are engaged in dynamic reversible interactions. On the other hand, permanently homodimeric proteins, which are only functional in the oligomeric state (homodimer cytochrome *c*), recombine via highly hydrophobic interfaces, which are also more closely packed.<sup>5</sup> A database of >2000 Ala mutants has been compiled; its systematic analysis revealed hot spots on protein interfaces, enriched in aromatic amino acids and in arginine. These are often surrounded by energetically less important residues, that most likely serve to occlude bulk solvent from the hot spot to lower the local dielectric constant.<sup>6</sup> Recently, Sheinerman et al. pointed out that electrostatic aspects in protein–protein interactions do not only largely determine the specificity of association but also contribute substantially to complex stabilization. Structural and mutational analyses suggest that clusters of charged and polar residues compensate desolvation penalties by forming networks of ion pairs and hydrogen bonds. Importantly, these attractive long-range interactions drive protein association kinetically.<sup>7</sup> As a relevant illustration it was now found that the transcription factor P53,

<sup>†</sup> Philipps-Universität Marburg.

<sup>‡</sup> Johann Wolfgang Goethe-Universität Frankfurt.

(1) Review: Stites, W. E. *Chem. Rev.* **1997**, *97*, 1233–1250.

(2) Larsen, T. A.; Olsen, A. J.; Goodsell, D. S. *Structure* **1998**, *6*, 421–427.

(3) Chothia, C.; Janin, J. *Nature* **1975**, *256*, 705–708.

(4) Argos, P. *Protein Eng.* **1988**, *2*, 101–113. Janin, J.; Miller, S.; Chothia, C. *J. Mol. Biol.* **1988**, *204*, 155–164.

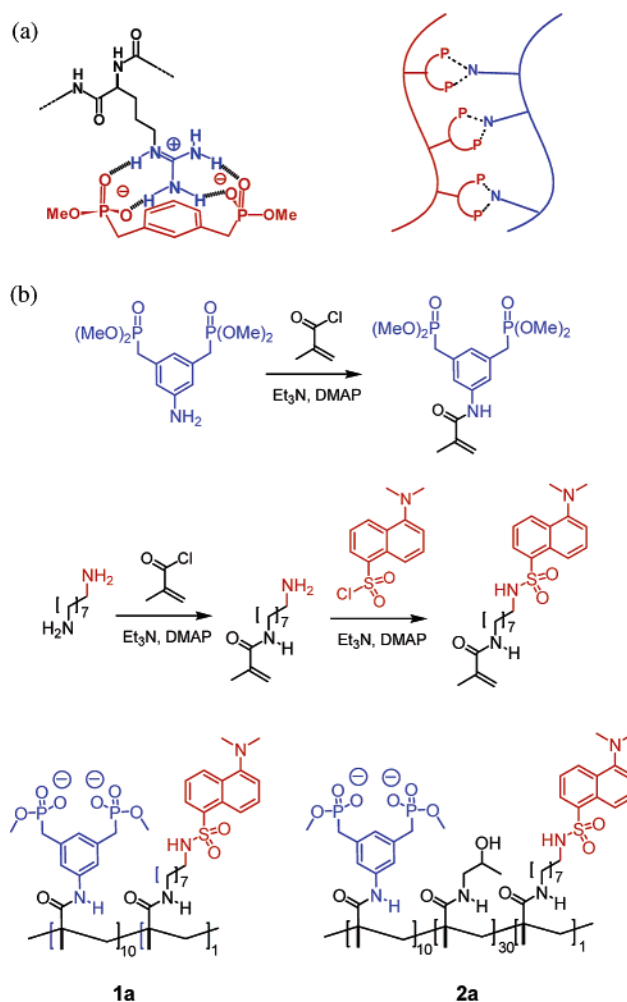
(5) Jones, S.; Thornton, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13–20.

(6) Bogan, A. A.; Thorn, K. S. *J. Mol. Biol.* **1998**, *280*, 1–9.

(7) Sheinerman, F. B.; Norel, R.; Honig, B. *Curr. Opin. Struct. Biol.* **2000**, *10*, 153–159.

an important tumor suppressor, binds cooperatively to DNA only as a dimer, stabilized by two intermolecular arginine-glutamate salt bridges.<sup>8</sup>

High affinity molecular recognition of protein surfaces combined with high specificity also remains a premier challenge for artificial receptor systems, especially in light of their solvent-exposed nature. Pioneering work by Hamilton et al. focused on the development of aspartate-rich cyclopeptides on calixarene scaffolds or glutamate-rich peptides on porphyrins for cationic protein surface recognition.<sup>9</sup> Often multiple copies of single weak binding motifs were used to increase affinities toward proteins: thus, a tetraguanidinium ligand was introduced as a helical protein surface binder for the tetramerization domain of the above-mentioned protein P53.<sup>10</sup> Similarly, linear anionic oligomers were reported to adopt heparin-like properties<sup>11</sup> or efficiently inhibit human leukocyte elastase ( $K_i$  values of  $\leq 0.2 \mu\text{M}$ ).<sup>12</sup> Merritt et al. demonstrated how a high-affinity inhibitor for cholera toxin evolves, if five copies of an  $\alpha$ -D-galactoside (MNPG) are coupled to a pentacyclen core unit ( $\text{IC}_{50} \sim 1 \mu\text{M}$ ).<sup>13</sup> Specific recognition of phosphorylated peptides and proteins was achieved by a fluorescent chemosensor carrying two Zn(II)-dipicolylamine units.<sup>14</sup> Strong and selective binding of carbonic anhydrase was also achieved with multivalent transition metal complexes, matching the protein's histidine surface pattern.<sup>15</sup> By contrast, anionically functionalized amphiphilic nanoparticles (i.e., monolayer-protected gold clusters - MMPCs) use nonspecific interactions to efficiently inhibit chymotrypsin through electrostatic binding followed by protein denaturation.<sup>16</sup> Similarly, Kiessling et al. developed postsynthetically modified (PSM) polymers in the form of multivalent mannose displays which nonspecifically inhibited hemagglutination.<sup>17</sup> In a molecular imprinting approach on the protein surface, the shape of lanthanide ion-carrying liposomes is reconstructed and used for protein sensing.<sup>18</sup> Finally, protein-protein interactions may be probed with designed protein binders (generated, e.g., by combinatorial library screening of "monobodies")<sup>19</sup> or disrupted with synthetic  $\beta$ -turn mimetics (e.g., the interaction of the nerve



**Figure 1.** (a) Left: structural key element - arginine residue embraced by bisphosphonate dianion. Right: schematic depicting the multiplication of bisphosphonate units for oligoarginine recognition. (b) Synthesis of monomer building blocks as well as schematic structure of homo- and copolymer **1a** and **2a**.

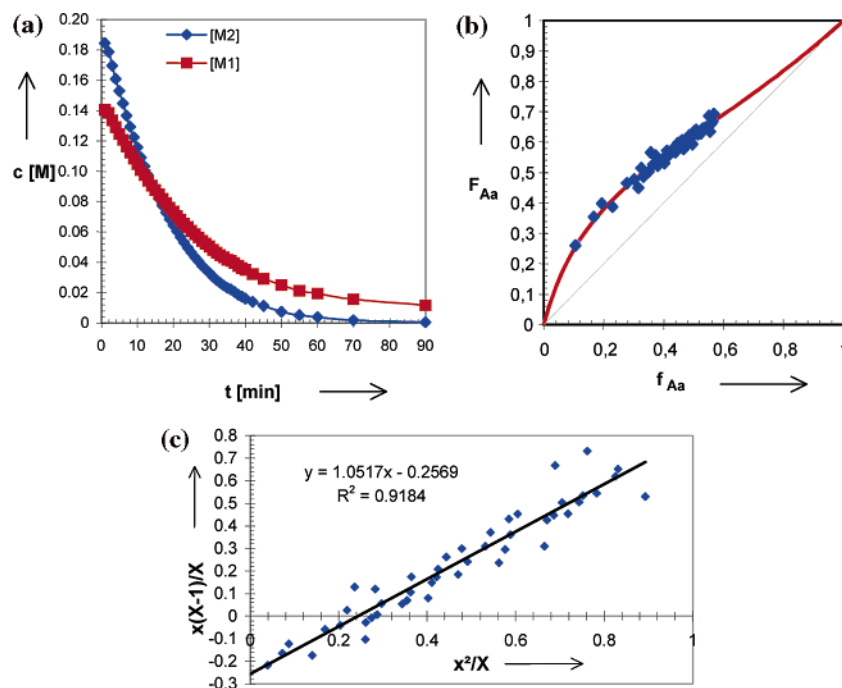
growth factor with its transmembrane tyrosine kinase receptor TrkA);<sup>20</sup> however, this area is still in its infancy.

## Concept

Some years ago, we discovered that small bisphosphonate dianions bind to arginine ( $K_a = 86\,000 \text{ M}^{-1}$  in  $d_6$ -DMSO) and lysine residues ( $K_a = 4000 \text{ M}^{-1}$  in  $d_6$ -DMSO) in a peptidic environment with remarkable affinity, while almost other amino acid side chains are completely rejected (Figure 1a).<sup>21</sup> However, transition to water causes a drastic drop in free binding energy ( $K_a \leq 100 \text{ M}^{-1}$ ), because the recognition process relies mainly on electrostatic attraction, enforced by  $\pi$ -cation attraction. The original high affinity for basic amino acids could now be restored and markedly enhanced by applying the concept of multivalency often found in natural surface recognition processes:<sup>22</sup> the weak arginine binder was therefore polymerized and thus transformed into an efficient receptor site for basic proteins with  $K_D$  values in buffered aqueous solution reaching the submicromolar regime. To this end, we performed a simple free radical

- (8) Dehner, A.; Klein, C.; Hansen, S.; Müller, L.; Buchner, J.; Schwaiger, M.; Kessler, H. *Angew. Chem.* **2005**, *117*, 5381–5386.
- (9) (a) Lin, Q.; Park, H. S.; Hamuro, Y.; Lee, C. S.; Hamilton, A. D. *Biopolymers* **1998**, *47*, 285–297. (b) Park, H. S.; Lin, Q.; Hamilton, A. D. *J. Am. Chem. Soc.* **1999**, *121*, 8–13. (c) Jain, R. K.; Hamilton, A. D. *Org. Lett.* **2000**, *2*, 1721–1723. (d) Baldini, L.; Wilson, A. J.; Hong, J.; Hamilton, A. D. *J. Am. Chem. Soc.* **2004**, *126*, 5656–5657.
- (10) Salvatella, X.; Martinell, M.; Gairi, M.; Mateu, M. G.; Feliz, M.; Hamilton, A. D.; de Medoza, J.; Giralt, E. *Angew. Chem., Int. Ed.* **2004**, *43*, 196–198.
- (11) Benezra, M.; Vlodaysky, I.; Yayon, A.; Bar-Shavit, R.; Regan, J.; Chang, M.; Ben-Sasson, S. *Cancer Res.* **1992**, *52*, 5656–5662.
- (12) Regan, J.; McGarry, D.; Bruno, J.; Green, D.; Newman, J.; Hsu, C.-Y.; Kline, J.; Barton, J.; Travis, J.; Choi, Y. M.; Volz, F.; Pauls, H.; Harrison, R.; Zilberstein, A.; Ben-Sasson, A. A.; Chang, M. *J. Med. Chem.* **1997**, *40*, 3408–3422.
- (13) Merritt, E. A.; Zhang, Z.; Pickens, J. C.; Ahn, M.; Hol, W. G. J.; Fan, E. *J. Am. Chem. Soc.* **2002**, *124*, 8818–8824.
- (14) Ojida, A.; Mito-oka, Y.; Inoue, M.; Hamachi, I. *J. Am. Chem. Soc.* **2002**, *124*, 6256–6258. Ojida, A.; Inoue, M.; Mito-oka, Y.; Hamachi, I. *J. Am. Chem. Soc.* **2003**, *125*, 10184–10185. Ojida, A.; Kohira, T.; Hamachi, I. *Chem. Lett.* **2004**, *33*, 1024–1025. Review: Ojida, A.; Miyahara, Y.; Kohira, T.; Hamachi, I. *Biopolymers* **2004**, *76*, 177–184.
- (15) Fazal, Md. A.; Roy, B. C.; Sun, S.; Mallik, S.; Rodgers, K. R. *J. Am. Chem. Soc.* **2001**, *123*, 6283–6290.
- (16) Fischer, N. O.; McIntosh, C. M.; Simard, J. M.; Rotello, V. M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5018–5023. Fischer, N. O.; Verma, A.; Goodman, C. M.; Simard, J. M.; Rotello, V. M. *J. Am. Chem. Soc.* **2003**, *125*, 13387–13391.
- (17) Strong, L. E.; Kiessling, L. L. *J. Am. Chem. Soc.* **1999**, *121*, 6193–6196.
- (18) Santos, M.; Roy, B. C.; Goicoechea, H.; Campiglia, A. D.; Mallik, S. J. *J. Am. Chem. Soc.* **2004**, *126*, 10738–10745.
- (19) Koide, A.; Abbateillo, S.; Rothgery, L.; Koide, S. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1253–1258.

- (20) Burgess, K. *Acc. Chem. Res.* **2001**, *34*, 826–835.
- (21) Rensing, S.; Springer, A.; Grawe, T.; Schrader, T. *J. Org. Chem.* **2001**, *66*, 5814–5821.
- (22) Review: Mammen, M.; Choi, S. K.; Whitesides, G. M. *Angew. Chem., Int. Ed.* **1998**, *37*, 2755–2794.



**Figure 2.** Determination of copolymerization parameters: (a) Decrease of comonomer concentrations; (b) corresponding copolymerization diagram; (c) Fineman–Ross diagram.

polymerization of the weakly binding monomeric unit, preferably in its anionic state. We describe in this manuscript its facile synthesis and the remarkable binding properties of this new class of bisphosphonate polymers toward basic peptides and proteins, in free solution and on surfaces. It appears that Nature itself uses the closely related phosphate–guanidinium interaction extensively for molecular recognition events between DNA/RNA and regulatory proteins; the chelate type arrangement Phosphate<sup>−</sup>...Arginine<sup>+</sup>...Phosphate<sup>−</sup> has even been coined the “arginine fork” and is largely responsible for the sequence-selective Tat/TAR recognition at an early critical stage of the HIV life-cycle.<sup>23</sup> A recent investigation about receptor heteromerizations between adenosine A<sub>2A</sub>/dopamine D<sub>2</sub> or glutamate NMDA/dopamine D<sub>2</sub> revealed the paramount importance of electrostatic interactions between Arg-rich epitopes and phosphorylated serines, reaching “amazing covalent-like” stabilities.<sup>24</sup>

With our strategy of developing highly efficient lysine and arginine binders<sup>25</sup> we are currently pursuing a 2-fold goal, i.e., targeting arginine- and lysine-rich protein domains as well as immobilization of Arg-tagged proteins. On one hand this would ultimately allow specific interference with gene regulation processes such as histone/DNA packing or RNA complexation by regulatory proteins. Numerous proteins (inter alia, cytochrome *c*, chymotrypsin, lysozyme) contain active sites surrounded by basic amino acids which might be reversibly blocked by receptor caps. On the other hand, any Arg-tagged protein might be immobilized onto surfaces in a functional state by means of these Arg-tag recognition units.

## Synthesis

The monomeric arginine binding building block was obtained by amide coupling between a *m*-amino-substituted xylylene

bisphosphonate<sup>26</sup> and methacryloyl chloride (Figure 1b). This could be homopolymerized in DMF or copolymerized in the same solvent with an amino alcohol-based methacrylamide,<sup>27</sup> in order to increase its water solubility. In a final polymer-analogous reaction, all phosphonic acid methyl esters were quantitatively cleaved by direct nucleophilic attack in a dipolar aprotic solvent.<sup>28</sup> Prior to dealkylation, the polymers were well soluble in chloroform (>100 mM); afterward, they switched to perfect water solubility (>100 mM). Alternatively, the bisphosphonate unit itself was mildly dealkylated with LiBr and subsequently copolymerized with a water-soluble initiator in aqueous solution. The neutral as well as the corresponding anionic state of the polymers could be well characterized by <sup>1</sup>H and <sup>31</sup>P NMR spectroscopy in CDCl<sub>3</sub> and D<sub>2</sub>O. Experimental evidence was thus produced for the complete polymerization of all monomers, and also for the quantitative monodealkylation of all dimethyl phosphonate ester groups. For an even distribution of binding motifs along the polymer chain, it was very important to guarantee a statistical copolymerization in all cases; to this end, we determined the respective copolymerization parameters in DMF and water (Figure 2). Monomer concentrations were calculated from their respective NMR integrations, calibrated with pyridine as an internal standard. Both, the integrated form of the copolymerization equation<sup>29</sup> as well the Fineman–Ross evaluation method<sup>30</sup> gave comparable results. With *r*<sub>1</sub> and *r*<sub>2</sub> values between 0.3 and 2.0, an almost ideal statistical copolymerization occurs in all cases, irrespective of the solvent. Interestingly, in water, the smaller *r*<sub>2</sub> values document a certain reluctance of the dianionic bisphosphonate

(26) Rensing, S.; Schrader, T. *Org. Lett.* **2002**, *4*, 2161–2164.

(27) Gros, L.; Ringsdorf, H.; Schupp, H. *Angew. Chem., Int. Ed. Engl.* **1981**, *20*, 305–325.

(28) Karaman, R.; Goldblum, A.; Breuer, E.; Leader, H. J. *J. Chem. Soc., Perkin Trans. 1* **1989**, 765–774. Krawczyk, H. *Synth. Commun.* **1997**, *27*, 3151–3161.

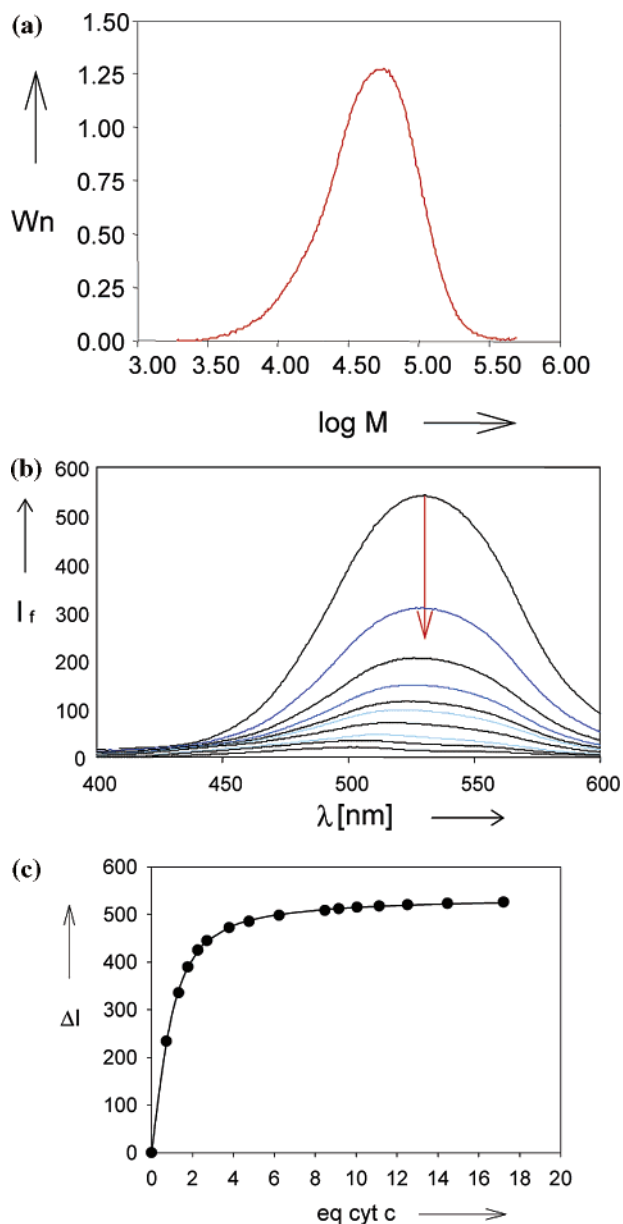
(29) Aguilar, M. R.; Gallardo, A.; del Mar Fernández, M.; San Román, J. *Macromolecules* **2002**, *35*, 2036–2041. Taden, A.; Tait, A. H.; Kraft, A. *J. Polym. Sci., Part A: Polym. Chem.* **2002**, *40*, 4333–4343.

(30) Fineman, M.; Ross, S. D. *J. Polym. Sci.* **1950**, *5*, 259–262.

(23) Calnan, B. J.; Tidor, B.; Biancalana, S.; Hudson, D.; Frankel, A. D. *Science* **1991**, *252*, 1167–1171.

(24) Woods, A. S.; Ferre, S. J. *Proteome Res.* **2005**, *4*, 1397–1402.

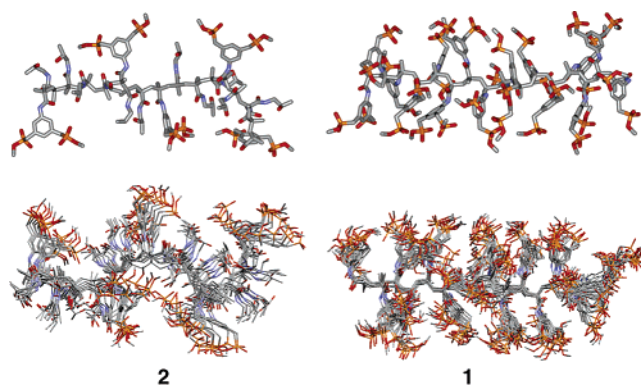
(25) Arginine hosts: Schrader, T. *Chem.–Eur. J.* **1997**, *3*, 1537–1541. Lysine hosts: Fokkens, M.; Schrader, T.; Klärner, F.-G. *J. Am. Chem. Soc.* **2005**, *127*, 14415–14421.



**Figure 3.** (a) Molecular weight and polydispersity determination for homopolymer **1** by GPC; (b) fluorescence spectra from titration of **1a** (1  $\mu$ M) with cytochrome *c*; (c) corresponding titration curve (inverted intensity decrease).

unit to homopolymerize, probably due to electrostatic repulsion. In DMF,  $r_2$  becomes 1 and restores perfect statistical behavior.

For the second generation of polymers a fluorescence label was incorporated by addition of 10% of a methacrylamide-based dansyl monomer to the copolymerization mixture to ensure a general sensitive spectroscopic characterization of the complexation event on protein surfaces (Figure 1b). All polymers were characterized by GPC, either in DMF (with LiCl additive) or in water. Molecular weights between 20 000 and 80 000 were detected, accompanied with the expected monomodal molecular weight distributions according to Schultz–Flory:  $M_w/M_n$  quotients of  $>2$  were observed in all cases, characteristic of the conventional free radical polymerization mechanism (Figure 3a). However, the molecular weight of the copolymers was always a little higher than that of the homopolymers, a fact that has to be kept in mind for direct comparisons. Thus, the homopolymers with a molecular weight of  $\sim 40$  000 contain



**Figure 4.** Top: side view from molecular mechanics calculations of homo- and copolymer **1** and **2** in water (3000 steps). Bottom: respective MD calculations (300 K, water, 10 ps).

approximately 100 bisphosphonate units, whereas the copolymers with twice the molecular weight contain only  $\sim 95$  bisphosphonates. The first generation of polymers was used in UV/vis titrations with proteins carrying a natural chromophore, such as cytochrome *c*. To be able to generalize these experiments, several basic proteins were subsequently fluorescence-labeled selectively at their *N*-terminus with OregonGreen.<sup>31</sup> However, in all cases, the effects remained very small and unreliable, probably due to the fact that the label was too far away from the binding region of the polymer. The statistical distribution of 10% of dansyl monomers within the polymer chain of the second receptor generation places them in the immediate vicinity of the complexation region; thus, strong effects can be expected on binding to a peptide or protein. Indeed, approach of such a polycationic guest usually leads to a marked increase in emission intensity of the polymer. As an exception, the interaction with iron-containing proteins such as cyt *c* and ferritin effects a complete fluorescence emission quenching, as illustrated in Figure 3b.<sup>32</sup> These binding events could be quantitatively monitored by fluorescence titrations and were evaluated by nonlinear regression (Figure 3c).<sup>33</sup> All the experiments described in this paper were conducted in 30 mM  $\text{NaH}_2\text{PO}_4$  buffer (pH 7.0).

## Polymer Structures

Hexadecamer fragments of homo- and copolymer were subjected to molecular mechanics calculations, followed by MD simulations (MacroModel7.0, Amber\*, water, 3000 steps; MD: 300 K, 10 ps). The resulting structures and their conformational flexibility, depicted in Figure 4, are strikingly different: Homopolymer **1** is characterized by a remarkable steric congestion with simultaneous electrostatic phosphonate anion repulsion, resulting in a stiff rodlike overall shape, with multiple interior amide groups surrounding the backbone and a high exterior negative charge density. In addition to this structural reminiscence of DNA, it also reaches a similar

(31) Molecular Probes, <http://www.molecularprobes.com/media/pis/mp06153.pdf>, 2001.

(32) Petrat, F.; de Groot, H.; Sustmann, R.; Rauen, U. *Biol. Chem.* **2002**, *383*, 489–502. Fluorescence quenching by redoxactive iron as a sensitive tool for the determination of chelatable iron: Petrat, F.; Weisheit, D.; Lensen, M.; de Groot, H.; Sustmann, R.; Rauen, U. *Biochem. J.* **2002**, *362*, 137–147.

(33) (a) Schneider, H. J.; Kramer, R.; Simova, S.; Schneider, U. *J. Am. Chem. Soc.* **1988**, *110*, 6442–6448. (b) Wilcox, C. S. In *Frontiers in Supramolecular Chemistry*; Schneider, H. J., Ed.; Verlag Chemie: Weinheim, 1991; p 123.

**Table 1.** Overview of All Binding Experiments Involving Homo- and Copolymers with Peptides and Proteins in Free Solution, According to Fluorescence Data<sup>a</sup>

peptide	pI <sub>theor</sub>	pI <sub>exptl</sub>	MW [kDa]	fluorescence titration <sup>b</sup> [M <sup>-1</sup> ]			
				homopolymer		copolymer	
				BP:guest	K <sub>a</sub> <sup>c</sup>	BP:guest	K <sub>a</sub> <sup>c</sup>
Arg	12	12	0.2	1:1	6e+1		
Arg <sub>2</sub>	12	12	0.3	2:1	9e+3		
Arg <sub>3</sub>	12	12	0.5	4:1	4e+4	2:1	4e+3
Arg <sub>4</sub>	12	12	0.6	8:1	6e+5 <sup>d</sup>	2:1	1e+4
Arg <sub>6</sub>	12	12	1.0	insoluble	>e+7	4:1	1e+6
Tat <sup>e</sup>	10	10	1.5		~e+9		5e+8

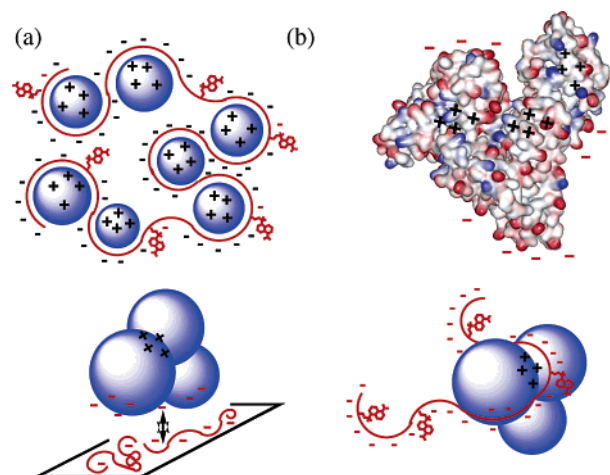
protein	pI <sub>theor</sub>	pI <sub>exptl</sub>	MW [kDa]	fluorescence titration <sup>b</sup> [M <sup>-1</sup> ]			
				homopolymer		copolymer	
				stoich	K <sub>a</sub> <sup>c</sup>	stoich	K <sub>a</sub> <sup>c</sup>
Hist H1	10.4		22.0		>e+6		>e+5
Cyt c	9.2		12.3	2:1	6e+5	3:1	2e+5
lysozyme	9.1	11.0	14.3		>e+5	4:1	4e+6
ADH	8.0	5.4	141	no shift		no saturation	
Trypsin	8.3	10.5	22.0		>e+5	4:1	4e+4
Prot K	7.7		28.9		>e+4	4:1	2e+4
BSA	5.8	5.8	66.0	5:2	2e+5	3:2	2e+5
ferritin	5.4	6.0	455.3	1:9	1e+6	1:9	1e+6

<sup>a</sup> Dansyl-labeled homo- and copolymers (0.1–1 μM) were titrated with arg<sub>1</sub>–arg<sub>6</sub> peptides and proteins of varying pI (5–10) and size (8–450 kDa). <sup>b</sup> Each measurement in 30 mM NaH<sub>2</sub>PO<sub>4</sub>. <sup>c</sup> Averaged virtual 1:1 binding constant, calculated for each complexation step, assuming no cooperativity between peptide/protein and polymer. <sup>d</sup> In 150 mM NaCl. <sup>e</sup> Alexa 488-labeled YRRKKQRRRC.

diameter of ~20 Å. By contrast, copolymer **2** with a statistical number of 2–4 amido alcohol comonomers between each bisphosphonate pair remains highly flexible and has spatially well separated receptor units, ideally suited for induced fit processes on a protein surface.

**Peptide Recognition in Solution.** Association constants were initially calculated for a whole series of oligoarginines, after establishing their individual stoichiometries on complex formation with the homo- and copolymer.<sup>34</sup> Surprisingly, complex formation with the homopolymer is highly economic, and on average each bisphosphonate unit binds exactly one arginine residue (Job plots, Table 1). A strong synergistic increase in binding affinity is observed, reaching nanomolar K<sub>D</sub> values for the Tat peptide (seven basic residues).<sup>35</sup> As a certain drawback the solubility of these extremely strong complexes drops below the micromolar regime and, hence, hampers the K<sub>a</sub> determination. In this respect the copolymers proved to be much easier to handle and produced completely soluble peptide complexes in all cases. However, the respective binding free energies remained about one kcal/mol (~1 order of magnitude in K<sub>a</sub>) behind those of their homopolymer counterparts. Again, the stoichiometric ratios reflect the statistical distance between two bisphosphonate receptor units, separated on average by 3 amidopropanol comonomer units.

**Protein Recognition in Solution.** Encouraged by these preliminary experiments we then focused our attention on protein surfaces. Proteins with pI values from slightly acidic (~5) to strongly basic (~10) and a broad spectrum of biological functions were chosen for comparison. When the homopolymer



**Figure 5.** (a) Schematic depicting the postulated binding mode between **1a** and cyt *c*; (b) Protein topology as determination factor for surface binding events - BSA's concave lysine domain is accessible for **1a** and **2a** only in free solution.<sup>36</sup>

was used as the host, the solubility problem became much more obvious, because even for the naked eye a considerable number of protein complexes precipitated from solution ( $c \approx \mu\text{M}$ ). Nevertheless, strong binding could be inferred from the remaining soluble cases, whose K<sub>a</sub> values were in the range of 10<sup>5</sup> M<sup>-1</sup> to 10<sup>6</sup> M<sup>-1</sup>, with stoichiometries between 3:2 and 4:1 indicating the efficient complexation of multiple protein molecules by one single polymer strand (Table 1, Figure 5a). Two proteins deserve a special comment: ADH failed to give any change in fluorescence emission, although its pI of 8.4 is clearly basic; on the contrary, Ferritin, the least basic protein with a pI of 5.5, binds very tightly to the polymer with a capacity to hold 9 copies at a time. The latter case is easily explained: in order to draw hundreds of Fe<sup>2+</sup> ions inside the giant molecular capsule, its interior is covered with numerous aspartate and glutamate carboxylates, leaving an excess of basic amino acid residues on the outside.<sup>37</sup> Transition to the copolymer restores the solubility of most complexes, and this time it is accompanied only with a marginal decrease in K<sub>a</sub>, corresponding to much less than 1 order of magnitude (Table 1). It should be emphasized that the molecular weights of most examined proteins decrease in reciprocal relation to their basicity increase.

**Arginine Selectivity.** To verify the assumed arginine preference of our polymeric hosts, the total number of solvent-exposed lysine vs arginine residues in each examined protein was drawn from the crystal structures and visualized in a synoptic view as gray and dark blue columns (Figure 6a).<sup>38</sup> Three pairs can be identified: Cyt *c* and trypsin contain almost exclusively lysine, whereas lysozyme and proteinase K carry predominantly arginine residues. HSA and ADH are mixed in their Arg/Lys composition. To take into consideration the counterbalancing acidic residues, another diagram was created which shows the total number of aspartates and glutamates on each protein surface, accompanied by a second column with the total number of basic residues (Figure 6b).<sup>38</sup> Proteins, which differ only in

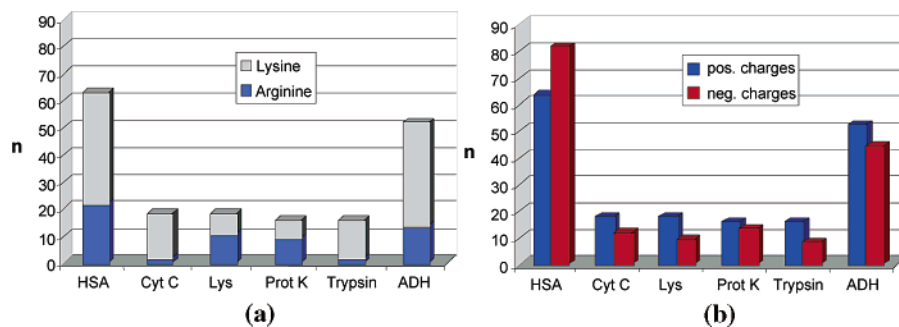
(36) Instead of BSA's crystal structure, which has not yet been solved, the known structure of the related HSA is shown (PDB-code 1BJ5).

(37) Reviews: Carrondo, M. A. *EMBO J.* **2003**, *22*, 1959–1968. Theil, E. C.; Ferritin. In *Handbook of Metalloproteins*; Messerschmidt, A., Poulos, H. R., Weighardt, K., Eds.; Wiley: Chichester, U.K., 2001; Vol. 2, pp 771–781.

(38) PROVE program (PROtein Volume Evaluation): Pontius, J.; Richelle, J.; Wodak, S. J. *J. Mol. Biol.* **1996**, *264*, 121–136.

(34) (a) Job, P. *Compt. Rend.* **1925**, *180*, 928–930. (b) Blanda, M. T.; Horner, J. H.; Newcomb, M. J. *Org. Chem.* **1989**, *54*, 4626–4636.

(35) Cooperation with Göbel, M. Frankfurt University. The K<sub>D</sub> value was determined by Fluorescence Correlation Spectroscopy (FCS).



**Figure 6.** (a) Total number  $n$  of Arg and Lys residues displayed on the examined protein surfaces; (b) counterbalance of acidic residues on the same proteins (total numbers  $n$ ; net charges  $\approx$  experimental pI values).

one single parameter, can be compared with respect to their binding strength. Cyt *c* and lysozyme, e.g., have almost the same molecular weight and a similar ratio of acidic and basic residues on their surfaces. However, lysozyme, which is rich in arginines (Arg: 61% of all basic residues), binds to the polybisphosphonate  $\sim 1$  order of magnitude more strongly than Cyt *c* (89% Lys; Table 1,  $K_{\text{Copol}}$ ). Similarly, proteinase K (60% Arg) and trypsin (88% Lys) have a comparable size, but trypsin is definitely more basic than proteinase K. Nevertheless, both bind to the polymeric tweezer equally well. In both cases, the arginine-rich protein is superior to the lysine-rich competitor, establishing a distinct arginine selectivity of the polymeric receptor. This important feature had hitherto only been proven for the monomer in organic solution (N/C-protected Arg vs Lys:  $K_a \sim 14:1$ ).

**Charge Density.** Another important factor governing the efficiency of protein binding is the sheer size of the target molecule. This is really an entropy argument: two proteins with the same absolute number of basic residues and a comparable pI may differ in their relative sizes. Hence, one protein will offer a much higher surface charge density which makes it statistically easier for each bisphosphonate moiety to find an Arg or Lys counterpart on the protein surface. In our series, Cyt *c* and trypsin both carry 14 or 16 lysine residues, respectively, on their surfaces and have a resembling surface charge ratio. However, Cyt *c* is much smaller than trypsin (12 vs 22 kD) and is hence bound much more tightly by the polybisphosphonate (Table 1,  $K_{\text{Copol}}$ ). This entropy effect is also reflected in the stoichiometry factors. Transition to larger proteins such as albumine (66 kD) and ferritin (455 kD) illustrates this effect even more impressively: although both proteins are slightly acidic (pI = 5.5–6.0), they are bound extremely well in solution, probably due to the high number of accessible basic residues.

### Peptide Recognition on Surfaces

Noncovalent immobilization of Arg-tags and proteins was accomplished by way of irreversible electrostatic attraction of a monolayer of the polymeric binders onto a SiO<sub>2</sub> surface densely coated with a polyethyleneimine (PEI) layer. The consecutive deposition of carrier polymer (PEI) and host polymer and the formation of stably adsorbed layers were monitored in a time-resolved manner by RfS (Reflectometric Interference Spectroscopy) in a flow-through system (Figure 7a).<sup>39</sup> This technique detects interactions at surfaces as an increase in optical thickness of a thin interference layer.<sup>40</sup> Subsequent equilibration of the immobilized bisphosphonate

polymer with (Arg)<sub>*n*</sub> and protein solutions produced, in most cases, a marked increase in optical thickness, proportional to the offered concentration. For most of the proteins investigated, the major part of the binding amplitude was caused by reversible interaction, and only a minor part of the protein was stably attached to the surface or dissociated only slowly. From the reversible binding signal titration curves could be derived whose nonlinear regression analysis furnished association constants for the surface binding process (Figure 7).<sup>41</sup> These are summarized in Table 2; in general, they remained  $\sim 0.5$  order of magnitude behind those measured in free solution. This makes sense, because a considerable number of negative charges have already been consumed for the electrostatic attraction of the host to the PEI surface. For Arg-tags, a similar picture evolves as already found in free solution: a steady increase in affinity was observed for higher oligomers, with the homopolymer clearly showing a performance superior to that of the related copolymer.

### Protein Recognition on Surfaces

The same trend is also true for protein binding to the polymer-coated surface. In principle, the amount of protein deposition followed the pI scale, with highly basic histone and lysozyme taking the lead, followed by moderately basic cyt *c* and trypsin (a survey is shown in Figure 8a). Several noteworthy exceptions remain:

**Topology.** BSA which was bound tightly in free solution did not show any affinity toward the immobilized bisphosphonate polymers, even with the copolymer carrying the hydrophobic dansyl moieties. This is the first indication of a different binding mode sensitive toward the protein topology:<sup>42</sup> while in free solution, the thin polymer thread can be smoothly wound around the BSA molecule, even deep inside the cleft, where most of the basic residues are concentrated, and this place becomes inaccessible to the flat anionic surface; even worse, several acidic domains are located on the edges of the heart-shaped molecule, leading to significant electrostatic repulsion (Figure 5b). Similar reasons may account for the fact, that ADH (Figure 6: 52+/44–) and proteinase K (16+/14–) completely refuse to interact with the anionic host in its immobilized state.

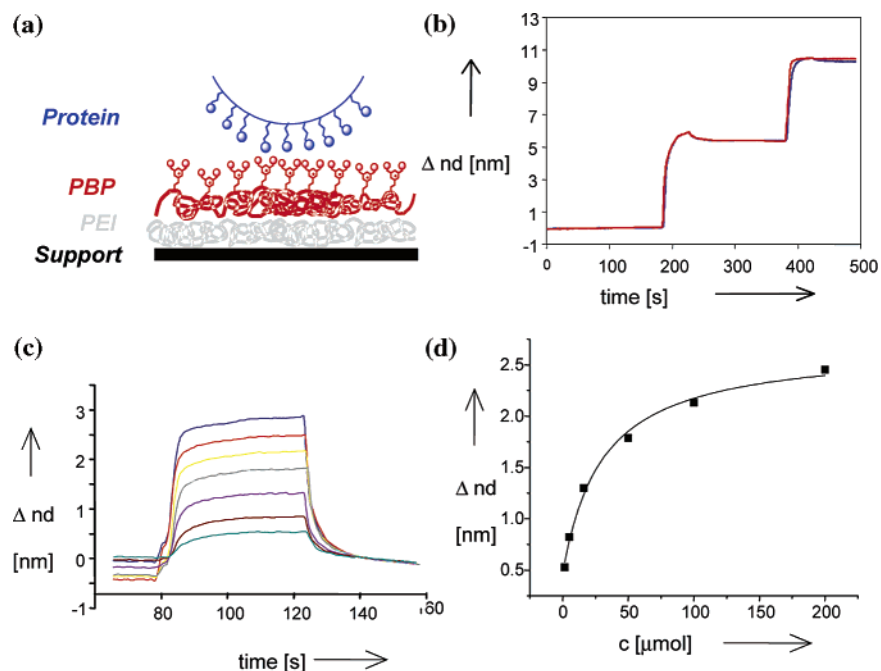
**Crystallization.** Ferritin, although *in toto* acidic, binds most strongly to the bisphosphonate surface, in an almost entirely

(39) Brecht, A.; Gauglitz, G.; Polster, J. *Biosens. Bioelectron.* **1993**, *8*, 387–392; Gauglitz, G.; Brecht, A.; Kraus, G.; Nahm, W. *Sens. Actuators, B* **1993**, *11*, 21–27.

(40) Piehler, J.; Brecht, A.; Gauglitz, G.; Zerlin, M.; Maul, C.; Thiericke, R.; Grabley, S. *Anal. Biochem.* **1997**, *249*, 94–102.

(41) Connors, K. A. *Binding constants*; Wiley: New York, 1987.

(42) BSA is the only protein investigated in the above-discussed series which substantially deviates from a globular topology (compare Figure 7).



**Figure 7.** (a) Schematic showing the attachment of molecular PEI and PBP layers onto the signal transducer; (b) Corresponding stepwise increase in RfS signal. (c) RfS signal upon injection of trypsin at different concentrations, binding on **1** with reversible and irreversible contributions; (d) corresponding titration curve furnishing  $K_a$ .

**Table 2.** Overview of All Binding Experiments Involving Homo- and Copolymers with Peptides and Proteins on the PBP-Coated PEI-Surface, According to RfS Data<sup>a</sup>

peptide	$pI_{theor}$	$pI_{exptl}$	MW [kDa]	RfS titration (PEI surface) <sup>b</sup> [M <sup>-1</sup> ]			
				$K_a$ reversible		$k_D$ irreversible	
				homopolymer	copolymer	homopolymer	copolymer
Arg	12	12	0.2				
Arg <sub>2</sub>	12	12	0.3				
Arg <sub>3</sub>	12	12	0.5	1e + 4	2e + 3		
Arg <sub>4</sub>	12	12	0.6	4e + 4	4e + 3		
Arg <sub>6</sub>	12	12	1.0	3e + 5	3e + 5		
Tat <sup>c</sup>	10	10	1.5				

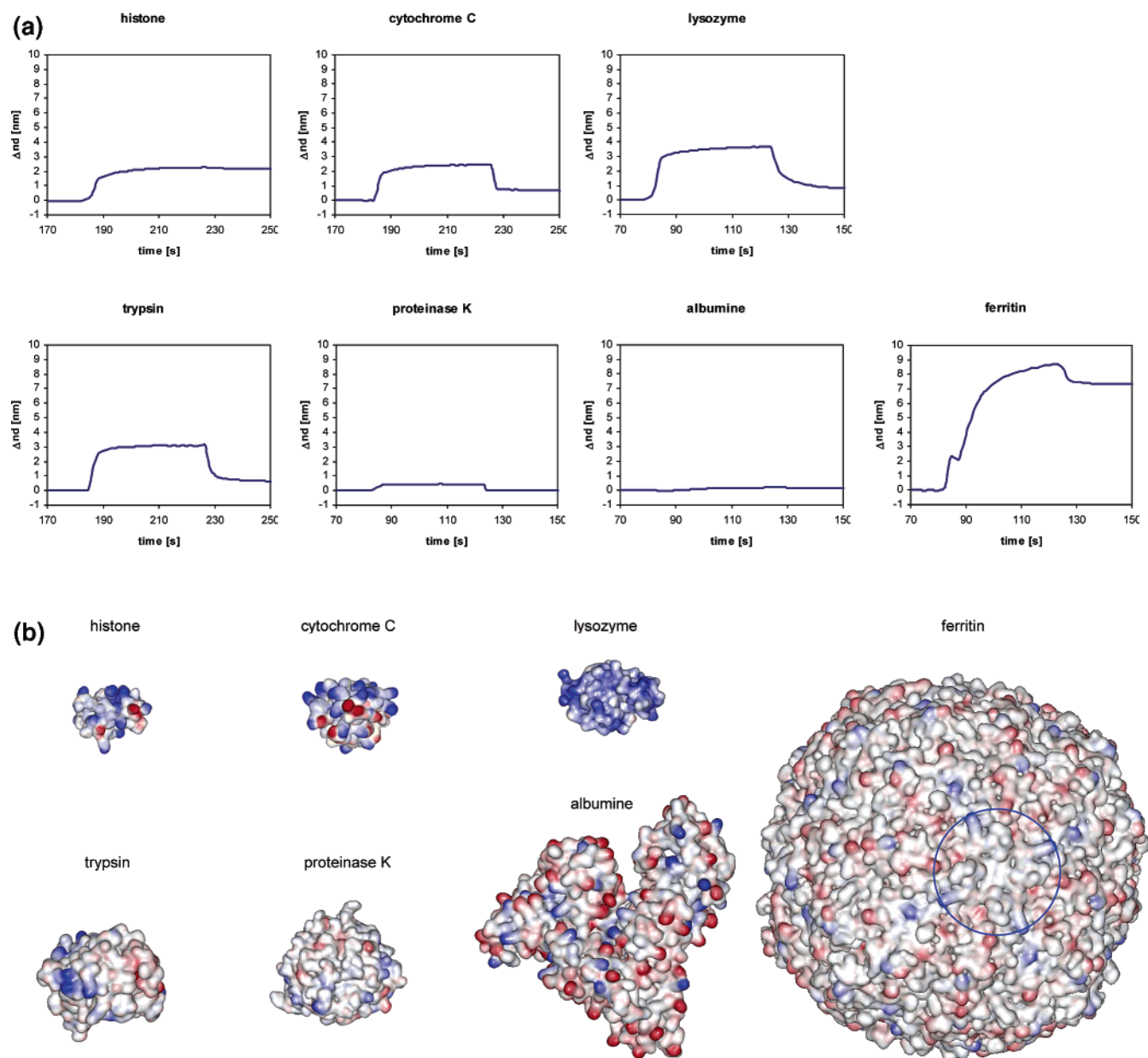
protein	$pI_{theor}$	$pI_{exptl}$	MW [kDa]	RfS titration (PEI surface) <sup>b</sup> [M <sup>-1</sup> ]			
				$K_a$ reversible		$k_D$ irreversible	
				homopolymer	copolymer	homopolymer	copolymer
histone H1	10.4		22.0	irreversible	strong	3e - 3	2e - 4
Cyt c	9.2		12.3		2e + 4	3e - 3	3e - 4
lysozyme	9.1	11.0	14.3	8e + 4	3e + 4	2e - 3	5e - 3
ADH	8.0	5.4	141	no effect		no effect	
trypsin	8.3	10.5	22.0	1e + 5	4e + 4	6e - 3	2e - 1
Prot K	7.7		28.9	no effect		no effect	irreversible
BSA	5.8	5.8	66.0	no effect		no effect	
ferritin	5.4	6.0	455.3	strong		crystallization	

<sup>a</sup> Unlabeled homo- and copolymers ( $\sim 2 \mu\text{M}$ ) were immobilized on the cationic PEI polymer and titrated with arg<sub>1</sub>-arg<sub>6</sub> peptides and proteins of varying  $pI$  (5–10) and size (8–450 kDa). <sup>b</sup> Each measurement in 30 mM NaH<sub>2</sub>PO<sub>4</sub>. <sup>c</sup> Alexa 488-labeled YRRKKQRRRC.

irreversible fashion. Due to its self-assembled dodecamer structure, crystallization is obviously greatly facilitated on the well-ordered two-dimensional arrangement on the PEI/PBP layer.

Additional information about the binding mode was drawn from an analysis of the interaction kinetics. In most cases, diffusion controlled association kinetics was observed as expected for electrostatic interaction. An exception was Ferritin, which associated much slower, in good agreement with the more complex layer formation on the PBP surface proposed for this protein. Typically, multiphasic dissociation of the adsorbed proteins was observed: the major part dissociated very fast,

followed by slower dissociation kinetics, which was characteristic for the different proteins, and for homo- and copolymer. To this purpose the dissociation rate constants were determined by fitting the decay dissociation curve with a mono- or biexponential decay. Indeed, the homopolymer shows again the slowest dissociation processes or even binds completely irreversibly (histone H1, Ferritin). To draw as much systematic information about the immobilization process as possible from the accumulated data, a graphical survey of all time-dependent RfS curves was produced, showing the association and dissociation kinetics of each individual protein immobilization event (Figure 8a). The same scale was used for the y-axis, indi-



**Figure 8.** (a) Protein recognition by homopolymer-coated surface **1** according to RfS (top series): (histone, ferritin) completely irreversible binding; (cytochrome *c*, lysozyme, trypsin) partially irreversible binding; (proteinase K, HSA) almost no binding. (b) Bottom series: the Connolly surface of these proteins, depicted in correct relative sizes, is patterned with the electrostatic surface potential (ESP), showing basic (blue) and acidic domains (red) on the protein surfaces. Histone, cytochrome *c*, and lysozyme with  $pI > 9$  are uniformly basic, and trypsin has a basic domain at one end and an opposing acidic domain at the other; proteinase K, human serum albumine (HSA), and ferritin show a statistic distribution of acidic and basic amino acids.

cating absolute optical thicknesses which generally correlate with immobilized protein material. Histone demonstrates superior affinity (most likely  $> 10^6 M^{-1}$ ) toward the immobilized poly-bisphosphonate by a completely irreversible RfS profile. Lysozyme, trypsin, and cyt *c* are also strongly bound ( $\sim 10^5 M^{-1}$ ) but in an almost fully reversible manner, although with relatively slow dissociation kinetics ( $k_D \sim 10^{-3}$ ). Ferritin not only binds irreversibly but also shows again a peculiar stepwise binding behavior: during the first 5 s, a protein monolayer seems to be deposited very rapidly on the bisphosphonate tweezer polymer, followed by slow crystallization of multilayers on top.<sup>43</sup>

**Basic Domains.** A closer inspection of the arginine and lysine distribution on the respective protein surfaces gives further insight into the factors controlling the binding process; to this end, crystal structures were drawn from the PDB and depicted

with Connolly surfaces,<sup>44</sup> colored in blue and red according to their electrostatic surface potential (ESP).<sup>45</sup> Those protein faces showing the highest density of basic residues are depicted in Figure 8b. At first glance, it becomes obvious, that especially lysozyme and, to a lesser extent, also histone H1 and cyt *c* are almost evenly covered with basic residues, accompanied with a high charge density because of the overall small surface area. Trypsin, however, has a basic and an acidic face, divided by an essentially neutral gap. In solution, this is a clear disadvantage

(43) Granier, T.; Langlois d'Estaintot, B.; Gallois, B.; Chevalier, J.-M.; Précigoux, G.; Santambrogio, P.; Arosio, P. *J. Biol. Inorg. Chem.* **2003**, *8*, 105–111. Zeth, K.; Offermann, S.; Essen, L.-O.; Oesterheld, D. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 13780–13785.

(44) (a) Connolly, M. L. *Science* **1983**, *221*, 709–713. (b) Connolly, M. L. *J. Appl. Crystallogr.* **1983**, *16*, 548–558.

(45) Sanner, M. F.; Spohner, J. C.; Olson, A. J. *Biopolymers* **1996**, *38*, 305–320.



for effective polymer attraction (cyt *c*,  $10^5 \text{ M}^{-1}$ ; trypsin,  $10^4 \text{ M}^{-1}$ ), but on a two-dimensional surface, the one basic face oriented toward the host suffices and eliminates any discrimination ability from related basic proteins such as cyt *c* (both  $10^4 \text{ M}^{-1}$ ). Proteinase K's and ferritin's EPS look pale, due to multiple mutual neutralization of neighboring lysine and aspartates; this may account for the extraordinarily low affinity of the former toward the immobilized bisphosphonate polymer. In the case of giant ferritin, hydrophobic interactions between the large contact areas will contribute to strong binding, in addition to formation of surrounding basic domains of at least four lysine residues (Figure 8b, blue circle). These domains mark a path for the anionic polymer chain along the lysines and arginines across the large ferritin surface, avoiding almost any contact with acidic residues; thus strong binding can also be expected in solution.

### Specificity

An important issue with multiple interactions relying mainly on Coulomb attraction is the proof of specificity. One could argue that the monomeric tweezer has already been extensively examined in this respect and evolved as an arginine- and lysine-selective receptor. However, this might have been overridden by the multiplication effect in the polymer. It was therefore an important additional experimental evidence to find out that those proteins, which bound strongly to the immobilized bisphosphonate polymer, exhibited only very weak effects with a polyvinyl sulfate layer, which was attached to the surface in the same manner. This was individually checked with Arg<sub>6</sub>, histone H1, and Cyt *c*.

In an exploratory experiment we also checked the intended selective binding of Arg-tagged proteins to the polybisphosphonate surface: A protein of essentially neutral pI was chosen (maltose binding protein, MBP), equipped with a doubly orthogonal His<sub>4</sub>-Arg<sub>4</sub>-His<sub>4</sub>-tag at its *N*-terminus. It stimulated an intense RfS signal, with a fully reversible profile. On the contrary, exchange of the Arg<sub>4</sub> for the Ala<sub>4</sub> sequence completely eliminated the protein's affinity to the anionic surface decorated with bisphosphonate host polymers, confirming a highly specific interaction of the protein with PBP through the Arg<sub>4</sub> motif.

### Thermodynamics

The multivalency of our new hosts leads to stepwise addition of binding enthalpies and also includes an entropy gain, if the subsequent binding events take advantage from a favorable orientation through the previous step.<sup>22</sup> More or less one-dimensional recognition is realized, if (Arg)<sub>*n*</sub> oligomers are bound by homopolymers in solution. Their superiority over copolymers is easily explained by their perfect complementarity, enabling both complexation partners to line up along one another. In the RfS measurements this effect is greatly attenuated, because the two-dimensional chip surface per se produces a substantial degree of preorganization for both immobilized polymers. However, protein surface recognition by electrostatic interactions imposes more difficult requirements: since arginines and lysines are scattered across the surface, interspersed with acidic residues, a homopolymer will win statistically because it may establish attractive contacts to the maximum number of accessible residues, especially in the case of small proteins with a high charge density.<sup>1</sup> This is indeed

the case with histone and lysozyme (Table 1). A copolymer, however, can minimize repulsive interactions with acidic residues and by induced fit select just positively charged amino acids for ion pairing. BSA and Ferritin appear to be good examples for the latter case. Finally, the combination of a globular protein surface with the rough but essentially flat surface of an immobilized polymer on a chip leads to an enthalpic disadvantage, because the mutual contact areas are minimized. On the other hand, proteins with large basic domains find a preorganized even host arrangement, rendering docking enthalpically and also entropically favorable. Even small hydrophobic patches on the receiving surface will add to this attraction (exclusion of solvent molecules, hydrophobic effect) and most likely lead to irreversible complexation. In our investigation, trypsin with its distinct basic domain proved to be even superior to lysozyme with its uniformly high density of surface arginines.

### Conclusion and Outlook

We conclude that polymeric bisphosphonate hosts are accessible via highly economic synthetic pathways and that they display high affinities toward Arg oligomers and most basic proteins in buffered aqueous solution. The unusually effective combination of homopolymers and Arg-tags suggests a promising application for the controlled assembly of functional Arg-tagged proteins on surfaces. The factors governing protein recognition by the new polybisphosphonate hosts do not follow the simple pI scale but are much more delicate: among others, the Arg/Lys ratio, the surface charge density, the occurrence of basic domains on the outer surface, the overall protein topology, and the sheer size of the protein guest as well as its related crystallization tendency play a decisive role in determining its affinity toward the artificial host. Basic as well as Arg-tagged proteins can likewise be immobilized on cationic surfaces and follow similar rules, complemented by retarded dissociation kinetics for the irreversible contribution to their free binding energy. In the future, we will investigate whether the tight complex formation with our polymers even affects the activity or function of the bound proteins. We also intend to combine these binding monomers with others tailored for the remaining classes of amino acid residues, to achieve a highly specific recognition of protein surfaces, i.e., ultimately create artificial antibodies. To this end, additional specific monomers with hydrophilic, hydrophobic, and aromatic binding sites will be incorporated.

**Acknowledgment.** We thank the group of Prof. M. Göbel, Frankfurt University, for the determination of  $K_D$  values for the complex between homo- and copolymer and the Tat peptide by Fluorescence Correlation Spectroscopy (FCS). We are also indebted to the Deutsche Forschungsgemeinschaft for funding in the framework of a researcher group entitled: "Synthesis of functional chemical-biological hybrid compounds".

**Supporting Information Available:** Synthetic procedures, polymerization procedures and characterization, determination of copolymerization parameters, fluorescence titrations, RfS measurements. This material is available free of charge via the Internet at <http://pubs.acs.org>.

JA0560229