Proceedings of the WELCOME Scientific Meeting on Hybrid Nanostructures, Toruń, Poland, August 28–31, 2011

De Novo Designed Proteins — Perspective Materials for Nanotechnology

J. Grzyb

Laboratory of Biological Physics, Institute of Physics, Polish Academy of Sciences
al. Lotników 32/46, 02-668 Warsaw, Poland

The paper explores the field of *de novo* protein design, as a source of material for effective hybrid nanostructures. Main design approaches, namely the intuitional and the computational strategy, are briefly overviewed. The achievements in the field are illustrated with several examples, starting from historical heme binding maquettes to novel non-natural enzymes. Separate paragraph covers the problem of designing peptides, which may act as anchor between biological and non-biological parts of nanostructures. The advantages of *de novo* designed proteins and still existing problems of the field are discussed.

PACS: 87.14.E–, 89.20.Bb

1. Introduction

Now-a-days science is challenged to become unified. What is more, today is not enough to ask "how can we understand the structure and behavior of the system" but the more important question is "how can we apply the knowledge to generate novel functions" [1]. Interdisciplinarity of research and application leads to creation of hybrid nanostructures — junction of biological and physical objects, till now studied by such distinct fields as biology and condense matter physics.

Hybrid nanostructures often contain proteins or protein complexes as their biological part [2]. Natural, native proteins have a lot of useful functionalities, but their long-term use is restricted by its other properties. Special attention has to be put on stability of structure and biological activity. Opposite to "physical" partner from hybrids, e.g. nanocrystals, which may be stable for months in room temperature, proteins may lose its properties during few minutes. This process, known as denaturation, is not only temperature-related, but may be caused by salts (especially heavy metals), chaotropic agents, high and low pH, etc.

This paper focuses on special type of proteins, which may become the answer for the problems with biological partner in hybrid structure. These proteins, called protein maquettes, *de novo* designed or artificial proteins have significant advantages: the increased structural stability, very well defined functionality and lack of several features responsible for denaturation and degradation. *De novo* designed proteins have been already successfully connected into hybrid structures with carbon nanotubes [3]. Variants of *de novo* designed proteins may participate in hybrid structure of different applications — from imaging-only to catalytical, specifically triggered reaction. Already known examples of these proteins are shortly described in following paragraphs.

Protein maquette is a term, invented by DeGrado in 1994 [4]. A bit earlier, concept of protein design was born. Today, the field is developing faster and faster, thanks to the involvement of advanced computer algorithms for designing and optimization of structures. In 1997, in their mini-review of protein design, Beasley and Hecht pointed out several problems, as β-sheet and mixed structured construction, as well as incorporation of thermodynamics, structural and functional properties of natural proteins, and finally getting enzymatic activity [5]. After only thirteen years, in 2010, the same group could conclude that the field of *de novo design* reached the stage when it is possible to design structures without natural analogues [6].

2. Strategy of experimental (intuitional) design

Historically first protein maquettes were heme binders. This example will be used here to illustrate design strategy known as empirical or intuitional. First sketch of protein designer is the overall structure. In natural (native) proteins the spatial organization (secondary and tertiary structure) is defined by hydrogen bonding pattern between following amino acids in peptide chain. Most popular secondary structure elements are α-helices and β-sheets. In α-helix, amino acid at 1st position creates hydrogen bond with amino acid in position 4th (compare Fig. 1a,c), giving approximately 3.6 residue for helix turn. In β-sheet, residues do not interact with such regularity, and for example, amino acid at 1st position may create a hydrogen bond with 13th position (compare Fig. 1b). In native proteins, one type of amino acid residue may be

present in different types of secondary structure elements; however every residue can be described by its tendency to create helix, sheet or coil (part of chain lacking stable structure). These probabilities are now well defined and included in the algorithms predicting protein structure (e.g. GOR [7] or Jpred [8]).

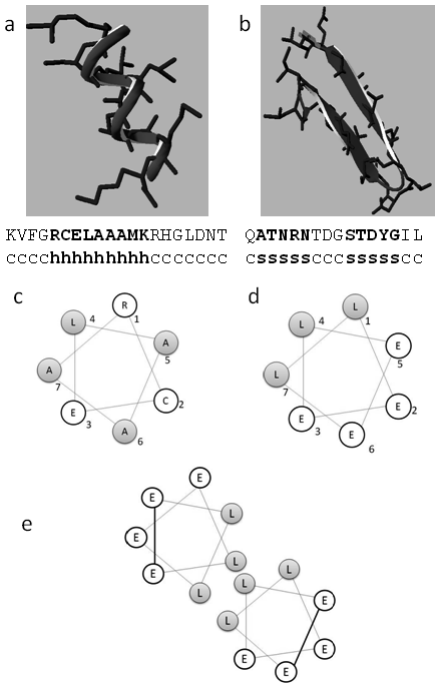

Fig. 1. The main type secondary structure elements, illustrated with the lysozyme [pdb:1lyz] fragments. Secondary structure properties (h: α-helix, s: β-sheet, c: coil) is denoted below amino acid code, cartoon representation is given for (a) α-helix and (b) β-sheet. (c) The helix (a) projection on heptad repeat, color code represents hydrophobic (white) and hydrophilic (gray) residues — note lack of discrete hydrophobic and hydrophilic parts, opposite to ideal amphipathic helix (d). Dimerization of amphipathic helix occur in water environment (e) (more details in the text).

The same knowledge is used to build protein. As helices are much more stable and well defined, it is easier to build them from scratch. Classic α-helix can be approximated by heptad repeat, representing spatial structure projected in two dimensions (Fig. 1c,d). Figure 1d presents how to simply choose amino acid sequence in order to create helix, polarized by hydrophobic properties. Due to rule of 3.6 residues per turn, helices longer than two heptades would be distorted if only this method was used. However, computer modeling enables more sophisticated designs, including e.g. non-natural right-handed coiled-coil of helices with undecarepeat [9].

Tertiary structure should be also defined by designer. Simple and smart way of controlling takes advantage of hydrophobic interaction. The helix, designed as in Fig. 1d, when placed in water environment, is creating bundles to bury hydrophobic residues (Fig. 1e). The number of helices involved in the bundle depends on the relative size of hydrophobic region, and may give even 6-helix bundle [10]. However, 4-helix bundle is very convenient for various purposes, and its design is well established [11].

Useful protein maquette has well defined functionality. In the historical approach, heme binding was assured by substitution of histidine into the place of one hydrophobic residue. Bis-histidine heme place was achieved by dimerization of helices. In advanced generation of these proteins (neuroglobin maquette [12]), the binding was intentionally distorted to make place for oxygen in the 6th coordination place. The heme-binding peptides were also the basis to creation of (bacterio)chlorophylls and (bacterio)chlorophyllides binders [13, 14].

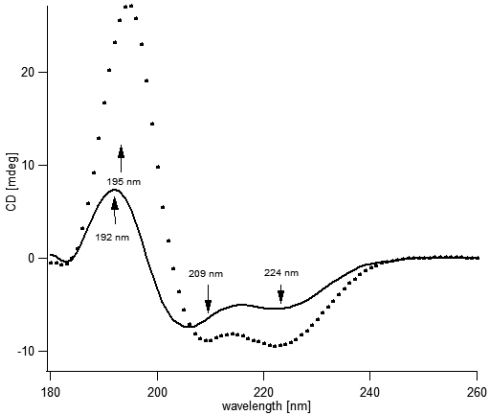

Fig. 2. Circular dichroism spectra, representing changes in the secondary structure of heme binding maquette, during heme titration into protein solution (unpublished data of author's laboratory). The structure changed from partially helical (solid line) to strictly helical (dotted). Arrows point to main spectral features of α-helix.

Protein structuralization may depend on presence of ligand (Fig. 2). The effect is usually unwanted as structural rearrangement has high energetic cost and would decrease affinity. However, the cost may get compensated by more favorable ligand burial in the hydrophobic interior. If protein folds about the ligand, the complex may be very dilution-resistant. Ligand-dependent structuralization was shown for several proteins, among others for heme binding maquettes (Fig. 2) and for coiled-coil iron-sulfur protein (CCIS1, [15]).

3. The examples of *de novo* designed proteins

3.1. De novo designing of metal sites

The binding of the porphyrine ring is, in general, coordination of its central atom, Fe. The heme binding maquettes and their offsprings, assembling other porphyrines and chlorines, were described in previous paragraph. Introduction of non-heme iron and other metal sites is also possible and well documented in the scientific

literature. The basic example was modification of native protein, thioredoxin, by only few amino acid substitution in various regions of the sequence, bringing iron binding sites and catalytical properties [16]. Similarly, small sequence substitution changed maltose binding protein into zinc-sensing protein. Studies on binding of four-iron four-sulfur clusters (4Fe4S), common cofactors in redox biochemistry were started with short peptides [17]. The same short peptide, but as a loop in helix-loop-helix motif, also supported 4Fe4S assembly [17]. The rubredoxin--mimics has simpler structure of the cluster (1Fe0S), but it is worth mentioning due to its structure, with still rare use of β-sheets, and high stability during the redox cycling [18]. Already mentioned CCIS1 is four helix bundle, but with 4Fe4S coordinating sphere within hydrophobic core [15]. This motif has no natural analogue.

3.2. Mixed cofactor maquettes

Very few examples are known for maquettes acquiring more than one type of cofactor for single working unit. Heme binders can bind also other porphyrines, as Ru-porphyrin, coproporphyrine and (bacterio)chlorophyll(ides) [19, 20]. Rabanal et al. [19] used smart approach with porphyrine/chlorine dimers in bis-histidine place and covalently attached quinone, in mimic of photosynthetic reaction centre. Gibney et al. [17] proposed ferredoxin-heme maquette, however did not present any data for cofactors common presence in protein scaffold. Recently, we designed, expressed and reconstituted redox-chain maquette (RCM), showing that it may bind both heme and 4Fe4S cluster within monomerics protein [21].

3.3. Hydrophobic and hydrophobic–hydrophylic maquettes

Membrane proteins are of high interest due to their important role in biological processes. Because of hydrophobicity, these proteins demand much more effort in purification and characterization, and are much less understood. However, the examples of *de novo* designed transmembranal helices are already known [22]. Short hydrophobic peptides were designed with viral-like activity of cell transformation induction [23]. Of special interest are proteins with strictly hydrophobic and hydrophilic region. Here belongs the combination of HP1 maquette and transmembranal protein of influenza virus [24], binding heme and hydrophobic cofactors [20]. Separate design algorithms are developed for hydrophobic membrane proteins [25].

3.4. Artificial enzymes and directed evolution

De novo design leads to obtaining of biotechnologically important enzymes, catalyzing various reaction [26]. Simply, native activity or affinity to one over another substrate may be improved [27]. More complicated, however possible, is introduction of catalytical properties into non-catalytical protein. Dwyer et al. [28] proved that few (18–22) point mutations (not more than 8 percent of protein) may introduce triose phosphate isomerase activity into ribose-binding protein, a receptor that normally lacks enzyme activity. Mentioned already, introduction of Fe sites into thioredoxin, belongs also to this category of manipulation. Notable examples are O_2-dependent phenol oxidase [29] and the AlleyCat, allosterically controlled eliminase, created on the basis of calmodulin [30].

De novo design has let recently to creation of enzymes, catalyzing reaction unknown to protein word. These are *Rosetta* designed catalyst of retro-aldol reaction [31] and Kemp elimination [32]. Optimization of activity may demand few rounds of design and testing, so-called directed evolution in laboratory tube [33].

4. Strategy of computational design

Computer enables designing of much more complicated structures that were before. Computational design is derivative of structure prediction and includes several steps optimizing structure for new protein folds, improving catalysis or increasing binding affinity [34]. There are known several programs and modeling platform (to be explored, e.g. by www.expasy.org). Just to give few examples, a DEZYMER was used for thioredoxin modification [35], ProtCAD for CCIS1 design [15], or PROBE used for improving termostability of subtilisin [36]. The best known modeling platform, Rosetta, has now a lot of different subdomains, including RosettaDesign [37]. Recent release, RosettaRemodel [38], deals with flexible proteins. However, the same rule which is written into heptad repeat is of high importance for every computer designed helix, and physicochemical intuition towards computational design is not in the opposition; design based only on experiments and intuition provide proofs for understanding of rules, which may become codified into computer algorithms [11].

One of approaches, using computational design, is known as evolution *in vitro*. It starts from library of sequences, expressed and tested. Several sequences with desired functionality are selected and subjected second round of evolution [39]. This approach may save time, but increase costs. Generally, it is much easier when the functionality is unnatural and may be tested even in bacterial cell lysate.

5. Nanomaterial specific recognition

In nanohybrids, very important is stable junction. It can be achieved by chemical modification or by surface or electrostatic absorption. However, this system works well only in small laboratory scale. Being time- and work--consuming, it is not the system of choice for biotechnology. Long process or harsh chemistry will lead to loss of activity even for improved novel proteins. Parallel problem is stability of nanomaterials, reacting with in water or with so common biological agents, as phosphates (compare Fig. 3). For toxicity of released metals see [40].

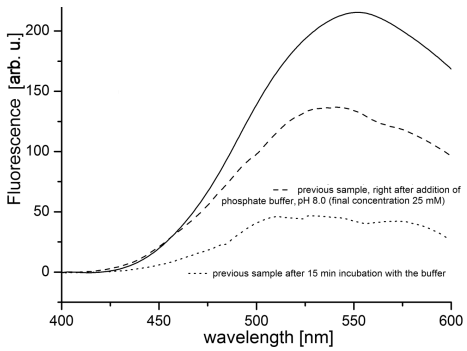

Fig. 3. Fluorescence emission spectra ($\lambda_{exc} = 325$ nm) of ZnO/MgO (100 μg/ml) nanopowders, prepared by method of [41] ilustrating fast decomposition after addition of phosphate buffer.

The perfect solution should be a protein tail (N- or C-terminal, part of loop, etc.) working as a tag specifically binding to e.g. semiconductor material of quantum dot (QD) and as nanoparticle stabilizers. Protein tail connection was used already to join quantum dots and redox elements, modifying surface charge of nanocrystal and its luminescent properties [42].

In nature, sequences with such properties are not consciously present — at least no one knows the place with natural selection, promoting organisms being able to recognize e.g. CdSe. However, it is possible to grow organisms under this specific selection criteria in laboratory. The *in vitro* evolution has brought already peptide sequences binding to crystals of, for example, CdSe, CdTe, ZnS, ZnO, ZnSe. Yeast display [43], *E.coli* display [44] and phage display [45] experiments showed high affinity of histidine to mentioned semiconductors, modulated by presence of tryptophane, cysteine and methionine. Modulation is dependent on material electronegativity and peptide total charge, or electrostatic attraction with lattice spacing elements [43, 46]. Computer modeling of platinum–peptide interaction suggests strongly that peptide organisation on surface is dependent on lattice [47]. Selective binding between Cu_2O and ZnO was achieved by presence of diaminoacid clusters: arginine–arginine and arginine–lysine [44]. Cited experiments were done on macroscopic materials — flat surfaces or single crystals. The obtained peptides are usually short (6–12 amino acid residues). Broader studies, describing longer peptide–metallic surface interactions are still missing. Inclusion of conformational problem in is also necessary, which was visible even for short, Pt-binding peptide [48].

Peptide–metal interaction is very important not only for stabilization or specific recognition in single nanohybrid, but also for organization of proteins in dense arrays on metallic surfaces. Molecular dynamics simulation has already started to face this problem [44].

Smart approach is the use of peptide as a building template for creation of nanocrystals in nature, there are several bio-composites, which became inspiration [49]. Engineered peptides for inorganic (GEPI) [50] were designed and are known to help in nucleation of hydroxyapatite [51] and ZnO powders [52, 53]. Biologically programmed core/shell QD were created in presence of two-domain peptide [54]. However, the methods of synthesis and assembly demands temperatures which are too high for stability of almost every known protein and there is still a lot to do in this field.

6. Conclusions

When biology meets physics there is a great possibility to create structures with completely new properties. This is, with no doubt, the future of medicine, both diagnostics and therapy. Nanohybrid structures are promising material also for light harvesting, light-dependent and light-triggered processes. Nonetheless, to become useful in the industrial scale, nanohybrids have to be less expensive and more stable. Stabilization of the biological part in nanohybrid junction may be achieved with application of designed proteins, with very specific functionality and possibility to be produced in low-cost bacterial expression systems. The protein part may be enriched in a specific anchor to a nanocrystal, may stabilize the nanoparticle in water environment, and finally, may become a building scaffold for it. However, these are only few known examples, still imperfect. Today a demand is co-development of biological and physical part of nanohybrid, as they are forced to exist and cooperate in the same process.

Acknowledgments

Support of Foundation for Polish Science, Homing PLUS Programme, co-financed by the European Union within the European Regional Development Fund is acknowledged. Partial support was also received from Polish Ministry of Science and Higher Education grant No. N N202 130039.

References

[1] P. Marguet, F. Balagadde, C. Tan, L. You, *J. Royal Society Interface* **4**, 607 (2007).

[2] S. Mackowski, *J. Phys., Condens. Matter* **22**, 193102 (2010).

[3] G. Grigoryan, Y.H. Kim, R. Acharya, K. Axelrod, R.M. Jain, L. Willis, M. Drndic, J.M. Kikkawa, W.F. DeGrado, *Science* **332**, 1071 (2011).

[4] D.E. Robertson, R.S. Farid, C.C. Moser, J.L. Urbauer, S.E. Mulholland, R. Pidikiti, J.D. Lear, A.J. Wand, W.F. DeGrado, P.L. Dutton, *Nature* **368**, 425 (1994).

[5] J.R. Beasley, M.H. Hecht, *J. Biol. Chem.* **272**, 2031 (1997).

[6] B.A. Smith, M.H. Hecht, *Curr. Opinion Chem. Biol.* **15**, 421 (2011).

[7] J. Garnier, J.F. Gibrat, B. Robson, *Methods Enzymol.* **266**, 540 (1996).

[8] C. Cole, J.D. Barber, G.J. Barton, *Nucl. Acids Res.* **36**, W197 (2008).

[9] P.B. Harbury, J.J. Plecs, B. Tidor, T. Alber, P.S. Kim, *Science* **282**, 1462 (1998).

[10] M.J.I. Andrews, A.B. Tabor, *Tetrahedron* **55**, 11711 (1999).

[11] V. Nanda, R.L. Koder, *Nature Chem.* **2**, 15 (2010).

[12] R.L. Koder, J.L. Anderson, L.A. Solomon, K.S. Reddy, C.C. Moser, P.L. Dutton, *Nature* **458**, 305 (2009).

[13] D. Noy, P.L. Dutton, *Biochemistry* **45**, 2103 (2006).

[14] I. Cohen-Ofri, M. van Gastel, J. Grzyb, A. Brandis, I. Pinkas, W. Lubitz, D. Noy, *J. Am. Chem. Soc.* **133**, 9526 (2011).

[15] J. Grzyb, F. Xu, L. Weiner, E.J. Reijerse, W. Lubitz, V. Nanda, D. Noy, *Biochim. Biophys. Acta* **1797**, 406 (2010).

[16] C.M. Kraemer-Pecore, A.M. Wollacott, J.R. Desjarlais, *Curr. Opinion Chem. Biol.* **5**, 690 (2001).

[17] B.R. Gibney, S.E. Mulholland, F. Rabanal, P.L. Dutton, *Proc. Natl. Acad. Sci.* **93**, 15041 (1996).

[18] V. Nanda, M.M. Rosenblatt, A. Osyczka, H. Kono, Z. Getahun, P.L. Dutton, J.G. Saven, W.F. DeGrado, *J. Am. Chem. Soc.* **127**, 5804 (2005).

[19] F. Rabanal, B.R. Gibney, W.F. DeGrado, C.C. Moser, P. Leslie Dutton, *Inorg. Chim. Acta* **243**, 213 (1996).

[20] D. Noy, B.M. Discher, I.V. Rubtsov, R.M. Hochstrasser, P.L. Dutton, *Biochemistry* **44**, 12344 (2005).

[21] J. Grzyb, F. Xu, R. Łuczkowska, E.J. Reijerse, W. Lubitz, V. Nanda, D. Noy, *Biochim Biophys Acta - Bioenergetics* **1817**, 1256 (2012).

[22] A. Zouzoulas, A.G. Therien, R. Scanzano, C.M. Deber, R. Blostein, *J. Biol. Chem.* **278**, 40437 (2003).

[23] L.L. Freeman-Cook, D. DiMaio, *Oncogene* **24**, 7756 (2005).

[24] B.M. Discher, D. Noy, J. Strzalka, S. Ye, C.C. Moser, J.D. Lear, J.K. Blasie, P.L. Dutton, *Biochemistry* **44**, 12329 (2005).

[25] P. Barth, B. Wallner, D. Baker, *PNAS* **106**, 1409 (2009).

[26] R. Sterner, F.X. Schmid, *Science* **304**, 1916 (2004).

[27] B. Schreier, C. Stumpp, S. Wiesner, B. Höcker, *Proc. Natl. Acad. Sci.* **106**, 18491 (2009).

[28] M.A. Dwyer, L.L. Looger, H.W. Hellinga, *Science* **304**, 1967 (2004).

[29] J. Kaplan, W.F. DeGrado, *PNAS* **101**, 11566 (2004).

[30] I.V. Korendovych, D.W. Kulp, Y. Wu, H. Cheng, H. Roder, W.F. DeGrado, *Proc. Natl. Acad. Sci.* **108**, 6823 (2011).

[31] L. Jiang, E.A. Althoff, F.R. Clemente, L. Doyle, D. Rothlisberger, A. Zanghellini, J.L. Gallaher, J.L. Betker, F. Tanaka, C.F. Barbas, 3rd, D. Hilvert, K.N. Houk, B.L. Stoddard, D. Baker, *Science* **319**, 1387 (2008).

[32] D. Rothlisberger, O. Khersonsky, A.M. Wollacott, L. Jiang, J. DeChancie, J. Betker, J.L. Gallaher, E.A. Althoff, A. Zanghellini, O. Dym, S. Albeck, K.N. Houk, D.S. Tawfik, D. Baker, *Nature* **453**, 190 (2008).

[33] L. Yuan, I. Kurek, J. English, R. Keenan, *Microbiol. Mol. Biol. Rev.* **69**, 373 (2005).

[34] S.M. Lippow, B. Tidor, *Curr. Opin. Biotechnol.* **18**, 305 (2007).

[35] A.L. Pinto, H.W. Hellinga, J.P. Caradonna, *Proc. Natl. Acad. Sci.* **94**, 5562 (1997).

[36] M.K. DiTursi, S.-J. Kwon, P.J. Reeder, J.S. Dordick, *Protein Eng. Design Select.* **19**, 517 (2006).

[37] Y. Liu, B. Kuhlman, *Nucl. Acids Res.* **34**, W235 (2006).

[38] P.-S. Huang, Y.-E.A. Ban, F. Richter, I. Andre, R. Vernon, W.R. Schief, D. Baker, *PLoS ONE* **6**, e24109 (2011).

[39] A.D. Keefe, J.W. Szostak, *Nature* **410**, 715 (2001).

[40] K. Jomova, M. Valko, *Toxicology* **283**, 65 (2011).

[41] B. Debasis, D. Bern, L. Qian, P.H. Holloway, *J. Phys. D* **41**, 182002 (2008).

[42] I.L. Medintz, T. Pons, S.A. Trammell, A.F. Grimes, D.S. English, J.B. Blanco-Canosa, P.E. Dawson, H. Mattoussi, *J. Am. Chem. Soc.* **130**, 16745 (2008).

[43] B.R. Peelle, E.M. Krauland, K.D. Wittrup, A.M. Belcher, *Langmuir* **21**, 6929 (2005).

[44] C.K. Thai, H. Dai, M.S. Sastry, M. Sarikaya, D.T. Schwartz, F. Baneyx, *Biotech. Bioeng.* **87**, 129 (2004).

[45] S.R. Whaley, D.S. English, E.L. Hu, P.F. Barbara, A.M. Belcher, *Nature* **405**, 665 (2000).

[46] T. Hayashi, K. Sano, K. Shiba, Y. Kumashiro, K. Iwahori, I. Yamashita, M. Hara, *Nano Lett.* **6**, 515 (2006).

[47] E.E. Oren, C. Tamerler, M. Sarikaya, *Nano Lett.* **5**, 415 (2005).

[48] U.O. Seker, B. Wilson, S. Dincer, I.W. Kim, E.E. Oren, J.S. Evans, C. Tamerler, M. Sarikaya, *Langmuir* **23**, 7895 (2007).

[49] M. Sarikaya, C. Tamerler, A.K. Jen, K. Schulten, F. Baneyx, *Nature Mater.* **2**, 577 (2003).

[50] C. Tamerler, D. Khatayevich, M. Gungormus, T. Kacar, E.E. Oren, M. Hnilova, M. Sarikaya, *Biopolymers* **94**, 78 (2010).

[51] M. Gungormus, M. Branco, H. Fong, J.P. Schneider, C. Tamerler, M. Sarikaya, *Biomaterials* **31**, 7266 (2010).

[52] L. Song, Y. Liu, Z. Zhang, X. Wang, J. Chen, *Prot. J.* **29**, 516 (2010).

[53] M.-K. Liang, O. Deschaume, S.V. Patwardhan, C.C. Perry, *J. Mater. Chem.* **21**, 80 (2011).

[54] S. Singh, K. Bozhilov, A. Mulchandani, N. Myung, W. Chen, *Chem. Commun.* **46**, 1473 (2010).