Small-Molecule Binders

DOI: 10.1002/ange.201005059

Powerful Protein Binders from Designed Polypeptides and Small Organic Molecules—A General Concept for Protein Recognition**

*Lotta T. Tegler, Guillaume Nonglaton, Frank Büttner, Karin Caldwell, Tony Christopeit, U. Helena Danielson, Karin Fromell, Thomas Gossas, Anders Larsson, Paola Longati, Thomas Norberg, Ramesh Ramapanicker, Johan Rydberg, and Lars Baltzer**

Drug development, in vitro and in vivo diagnostics, industrial protein purification, innumerable bioanalytical research applications, and many other areas of biotechnology are critically dependent on access to molecules that recognize and bind proteins specifically and with high affinity in complex biological media. While antibody technologies and organic synthesis remain invaluable tools for biotechnology and biomedicine, the quest continues for robust and efficient binder molecules that expand chemical diversity and are less costly and time consuming to develop. Currently, protein binders are almost exclusively of a purely biological origin, that is, they are antibodies, engineered proteins, and aptamers, or purely synthetic, for example, small organic pharmaceuticals. Here we report on a new concept for protein recognition based on a set of designed polypeptides conjugated to small organic molecules. The resulting hybrid molecules bind proteins with specificities and affinities that compare well to antibodies, while being by comparison easy to prepare and more than an order of magnitude smaller.

At the core of the technology are polypeptides not developed specifically for a given protein but selected from a designed set of 42-residue sequences of general applicability. An analogy may be found in nature, where a small subset of amino acids at a protein–protein interface, the so called hot spots, dominate the interaction between proteins, whereas the remaining amino acid residues at the protein–protein interface can be mutated extensively without greatly reducing the affinity or specificity.[1–3] The small organic molecules, or "hot-spot mimics", need only to bind with micromolar dissociation constants for the polypeptide conjugates to have K_D values in the nanomolar to picomolar range. The technology enables chemists with access to small-molecule binders that would normally be considered "failed", "fragments", or "at an early stage of development" to prepare powerful binders with relative ease and use them for target validation, biomarker identification, protein purification, diagnostics, as pharmaceuticals, etc. The concept is demonstrated here for the C-reactive protein (CRP), a protein of interest as a diagnostic biomarker.[4] CRP plays a key role in inflammation and the so called high-sensitivity CRP test is used as a cardiovascular risk marker.[5]

A 16-membered set of 42-residue polypeptides was designed to serve as scaffolds to which small molecules could be attached to form binder candidates (Figure 1). The polypeptides have no natural origin and were designed de novo to have some propensity for folding into helix-loop-helix motifs.[6] The underlying principle was that hydrophobic interactions would provide the binding energy and charged residues the selectivity in protein binding. Helices were designed to be amphiphilic with charged residues introduced in four different combinations to give total charges of −7, −4, −1, and +2, with the hydrophobic residues unaltered throughout the set. The site of incorporation of the small-molecule ligand was also varied and positioned at the early part of helix I (position 8), the latter part of helix I (position 17), the loop region (position 22), and the middle of helix II (position 34). Only modest amounts of extra binding energy are required for enhanced affinities on the order of 3–4 orders of magnitude over that of the small molecule, as the entropic cost for binding has been reduced as a consequence of conjugation of the polypeptide to the small molecule. Energies of 20–30 kJ mol^{-1} correspond to hydrophobic inter-

[*] Dr. L. T. Tegler, Dr. G. Nonglaton,[+] Dr. F. Büttner,[#] T. Christopeit, Prof. U. H. Danielson, Dr. T. Gossas, Dr. P. Longati,[$] Prof. T. Norberg, Prof. R. Ramapanicker,[++] Dr. J. Rydberg, Prof. L. Baltzer
Department of Biochemistry and Organic Chemistry
Uppsala University
P.O. Box 576, 75123 Uppsala (Sweden)
Fax: (+46) 18-471-3818
E-mail: lars.baltzer@biorg.uu.se
Homepage: http://www.biorg.uu.se
Prof. K. Caldwell, Dr. K. Fromell
Department of Physical and Analytical Chemistry
Uppsala University (Sweden)
Prof. A. Larsson
Department of Medical Sciences, Akademiska sjukhuset
75185 Uppsala (Sweden)
[+] Current address: Département micro-Technologie pour la Biologie et la Santé CEA-LETI, Minatec, Grenoble (France)
[++] Current address: Department of Chemistry, Indian Institute of Technology, Kanpur (India)
[$] Current address: Department of Clinical Science, Intervention and Technology, Novum Centrum, Karolinska University Hospital, Huddinge (Sweden)
[#] Current address: Leipziger Arzneimittelwerk, Leipzig (Germany)

[**] We are indebted to the European Union for financial support through the IP CARE-MAN (contract no. NMP4-CT-2006-017333), to the Knut and Alice Wallenberg Foundation, and to UppsalaBio through its Bio-X project. We also thank the Swedish Research Council for financial support through a grant to L.B. We are indebted to Greta Hulting for technical assistance and Dr. Jennifer Neff of Allvivo LLC for the generous gift of Pluronic-F108 PDF.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201005059.

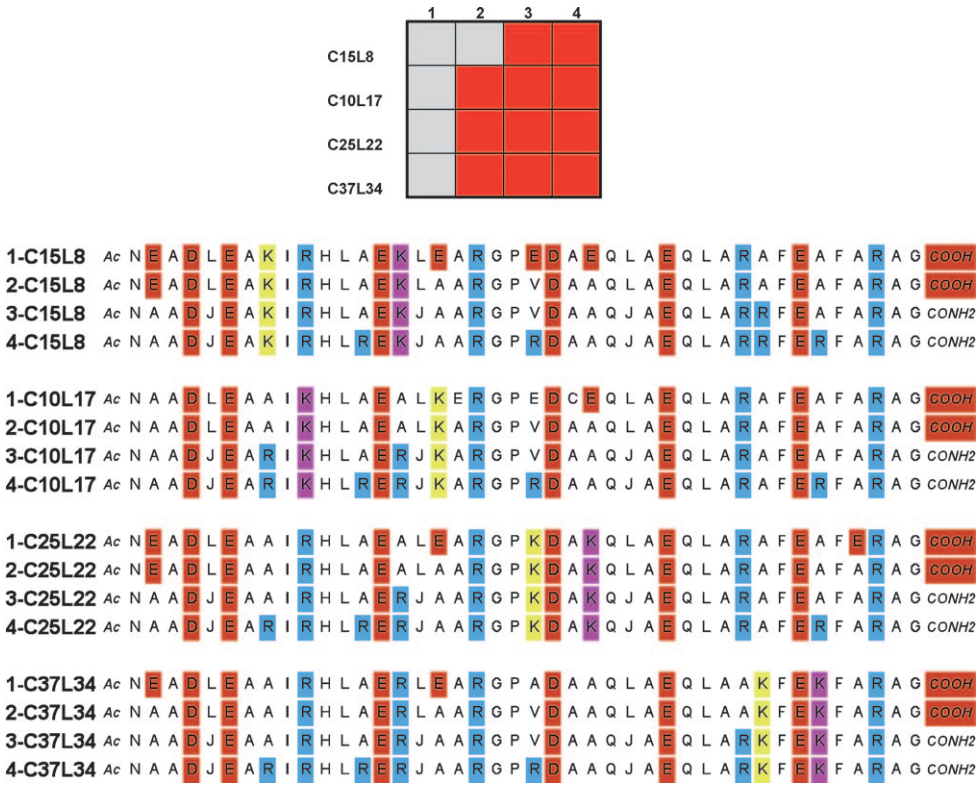

Figure 1. Polypeptide sequence library and graphical representation of the hit profile (red squares). The 16-membered set of 42-residue sequences varies in terms of charge and the site of incorporation of PC-6. To illustrate the variation pattern, Asp and Glu residues are color coded red, Arg residues are blue, linkage sites (L) for PC-6 are yellow, and linkage sites for the chromophore (C) are violet. All N terminals are acetylated.

actions involving, for example, 2–4 leucin side chains in the polypeptide. A principle behind the design is that if a number of residues are in position to bind, some of them are likely to find productive binding sites, although it is not possible to predict which ones. Charge–charge interactions are individually weak, but may contribute through cooperativity. Helices are expected to present binding residues better than unordered polypeptides for reasons of preorganization.

Structurally, the polypeptides show the hallmarks of molten globules. The ^{1}H NMR spectrum of the sequence 4-C10L17PC6 is characterized by broadened resonances and poor shift dispersion. The CD signature shows a high content of secondary structure, with a mean residue ellipticity $Q_{222} = -20000\ \mathrm{deg\,cm^2\,dmol^{-1}}$ (see the Supporting Information), but there is no structural evidence yet to show that the polypeptide adopts a helical conformation in a complex with CRP. The polypeptide scaffolds are not preorganized to fit a unique surface epitope of a protein, but adapt to the protein surface in the vicinity of the small-molecule binding site. The binder for CRP was designed based on the crystal structure of the protein complexed to its natural ligand, phosphocholine (PC, Figure 2).[7] PC binds CRP with an affinity of approximately 6 μM[8,9] in the presence of two Ca^{2+} ions, which intercalate between the phosphate group of PC and Asp60, Asn61, Glu 138, Gln139, Asp140, Glu147, and Gln150 in the PC binding site. A spacer was introduced by replacing one of the methyl groups of the trimethylammonium group by a functionalized spacer to form PC6 (Figure 2). The strategy to link PC through the trimethylammonium group was based on the expectation that important interactions of the phosphate group would be weakened if it was chemically modified.[10] The 4-nitrophenyl ester in PC6 was chosen to enable incorporation of the spacer-modified phosphocholine into the polypeptides by forming an amide linkage at the side chain of a lysine residue in a one-step reaction in DMSO solution. Only one free lysine residue was accessible in each sequence. For a detailed description of the synthesis of PC6 and the conjugation reaction see the Supporting Information.

PC6 was treated with each member of the set of 16 polypeptides to form 16 binder candidates and each one was evaluated with regards to affinity and selectivity. The affinity for CRP was estimated in a screening experiment based on fluorescence intensity using a microtiter plate reader. A 500 nM solution of polypeptide conjugate in 50 mM tris(hydroxymethyl)aminomethane (Tris) buffer at pH 7 containing 5 mM $CaCl_2$ and 150 mM NaCl was titrated in three steps with CRP. All polypeptides carried 7-methoxycoumarin chromophore probes and the intensity of the fluorescence emission was found to be affected by binding to CRP. Purified human CRP was added from a stock solution at a concentration of 2.3 $\mathrm{mg\,mL^{-1}}$, to give total CRP concentrations of 500 nM, 1000 nM, and 1500 nM. Each measurement was carried out in triplicate and compared to a negative control, where the binder alone was titrated with buffer. The purpose of titrating the binder in three steps was to determine within experimental accuracy whether binding to the polypeptide conjugate was saturated at a 500 nM concentration of CRP, in which case an apparent K_D value of 10 nM or lower could be estimated assuming an experimental error of 10% or less (see the Supporting Information). Binders for which changes in the fluorescence intensity at 410 nm were observed after the addition of one equivalent of CRP, but no significant further changes after two and three additions were considered hits. There were 11 hits in the set of 16 binder candidates (Figure 1), although the evaluation is somewhat subjective, since the overall change in fluorescence intensity is often small and difficult to evaluate. Polypeptides with a total charge of +2 or −1 gave the most hits, which is not surprising in view of the fact that CRP is an acidic protein with a protein

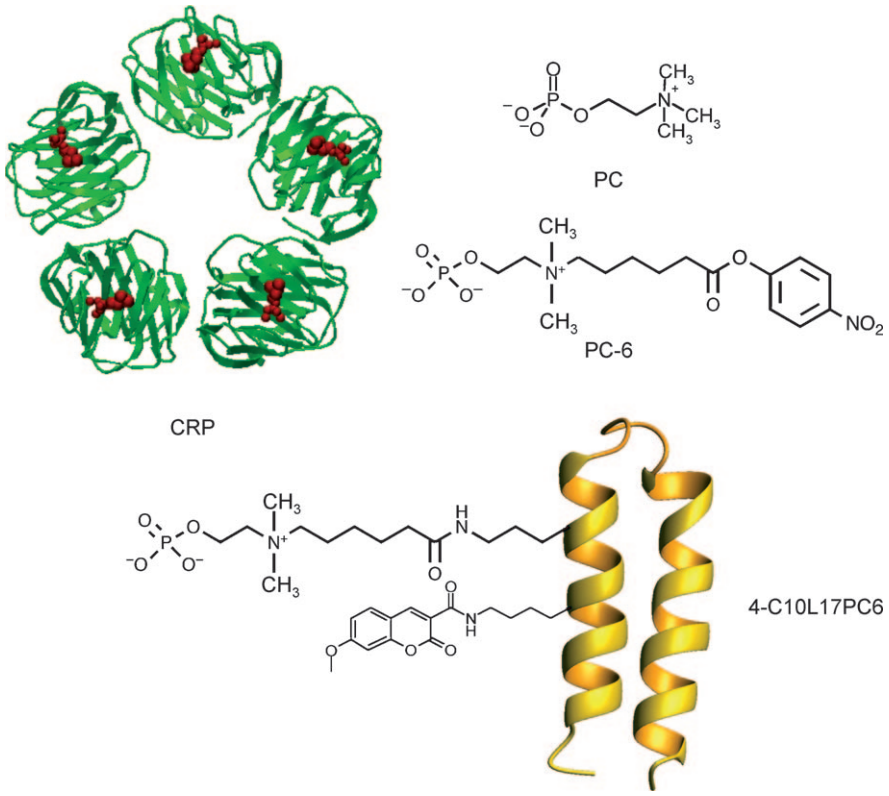

Figure 2. The crystal structure of CRP in a complex with PC and Ca^{2+} ions shows that PC binds to each subunit of CRP in the presence of the Ca^{2+} ions. Modification of PC by introduction of a spacer to form PC-6, followed by reaction with a polypeptide gives the binder molecule. A chromophore is routinely introduced to enable determinations of the concentration and for fluorescence titrations, but probably contributes little or nothing to affinity. C denotes coumarin, but dansyl (D) is also used.

isoelectric point (PI) of 5.3. It was not possible to evaluate which amino acids were critical for binding. We synthesized an N-terminal 20-mer and an N-terminal 32-mer of 4-C10L17PC6, and analysis showed that the full-length sequence binds better than the truncated ones, and could not in this case have been replaced by shorter ones (see the Supporting Information). Multivalency is not the reason for the high affinity towards CRP, as demonstrated by surface plasmon resonance (SPR) and fluorescence titrations (see the Supporting Information).

All binder candidates were also characterized by SPR biosensor analysis (Biacore) to obtain a more detailed insight into their interactions with CRP, as described previously for PC.[9] The protein was immobilized on the chip under standard conditions and the interactions of each binder with CRP was studied in the presence of 10 mM $CaCl_2$, with the binder concentrations varied in the range from 0.5 to 1200 nM (Figure 3). The results from the SPR analysis were in broad agreement with those from the fluorescence-based screening, as the strongest interactions with CRP were found among the binder molecules with total charges of +2 and −1. All binders interacted reversibly with CRP, and several tight binders were identified. The dissociation of 4-C10L17PC6 was extremely slow and essentially irreversible. The exceptional stability of the complex shows that the affinity of 4-C10L17PC6 for CRP is very high, and higher than that of the phosphocholine residue by probably three to four orders of magnitude. The affinity of the polypeptide 4-C15L8 without PC attached could not be measured under the conditions used—no change in the fluorescence was observed upon addition of 2 mM CRP to a 500 nM solution of the scaffold—and one can only estimate that it has a dissociation constant that is higher than 100 mM. PC was shown to inhibit the binding of 4-C15L8 coupled to PC6, thus demonstrating that the PC6 residue of the conjugate binder targets the PC binding site of CRP as designed (data not shown). The overall affinity is thus due to cooperativity between the PC binding group, the polypeptide, and possibly also the aliphatic spacer. A suitable mechanistic model could not be identified by fitting equations representing a variety of interaction models to the SPR data, and thus kinetic parameters and dissociation constants could not be determined. The binders were ranked by calculating from the data set their concentration that provides half saturation (B_{50} values; see the Supporting information). The seven highest ranking binders identified by SPR were all in the group of ten identified by fluorescence and only one of the ten highest ranked binders according to SPR was not identified on the basis of fluorescence.

The selectivity of the tightest binder according to Biacore ranking, 4-C10L17PC6, was investigated by attaching it to a polystyrene (PS) nanoparticle through a disulfide bridge to the Cys residue at position 24, and incubating it with patient serum having a CRP concentration of 60 mg L^{-1} that was obtained from the Uppsala University hospital. After incubation of the nanoparticles with serum, the mixture was centrifuged and the supernatant removed. The particles were washed three times with buffer, treated with DTT to reduce the disulfide bonds, and the resulting supernatant analyzed by sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE; Figure 4). As a control, an avian antibody (IgY) directed against CRP was attached to nanoparticles and treated the same way. Several bands were observed, most likely because CRP binds not only the binder and the antibody, but also several proteins in the complement activation system, which are captured from the serum as well. The critical observation is that both the 4-C10L17PC6 binder and the anti-CRP antibody extract the same pattern of proteins, thereby giving rise to the same bands on the gel. The synthetic binder is as selective as the antibody in human serum.

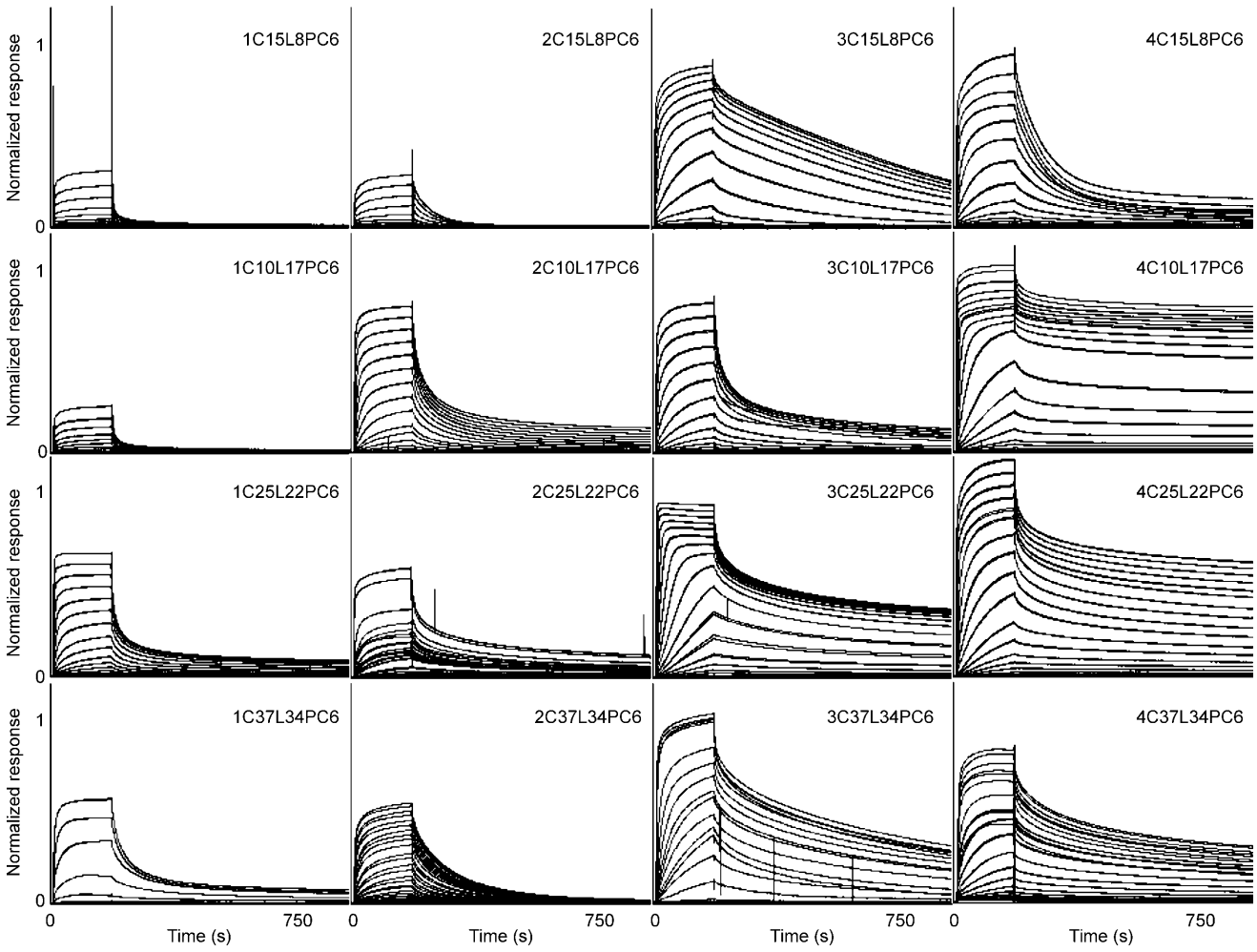

Figure 3. Sensorgrams from SPR biosensor analysis of interactions between the 16 binder candidates and immobilized CRP. The binder 4-C10L17PC6 dissociates extremely slowly. The concentrations of binders used were 1200, 800, 533, 355, 237, 158, 105, 70.2, 46.8, 31.2, 20.8, 13.8, 9.24, 6.61, 4.11, 2.74, 1.82, 1.21, 0.81, and 0.54 nM.

Figure 4. SDS-PAGE of extracts by 4-D10L17PC6 and an avian anti-CRP antibody from human serum. Lanes from left to right: 1) anti-CRP in human serum, 2) anti-CRP in CRP-free serum, 3) 4D10L17PC6 in human serum, 4) 4D10L17 in CRP-free serum, 5) blank, PS beads in human serum, 6) blank, PS beads in CRP-free serum, 7) HSA, 8) CRP, 9), 10) standard proteins (97, 66, 40, 30, 20.1, and 14.4 kDa), 11) prestained standard.

To further demonstrate the selectivity in biological fluids we introduced tight binders in well-known bioanalytical test formats (see the Supporting Information). Several of the polypeptide conjugate binders provided excellent results in measuring CRP in human serum in an enzyme-linked immunosorbent assay (ELISA). A nitrocellulose-based assay showed that 4-C10L17PC6 compared well with the anti-CRP antibody in a commercially available point of care assay. A binder for CRP based on the polypeptide sequence 3-D37L34 was recently shown to perform extremely well on chip in a point of care assay.[11]

The small set of polypeptides required to develop excellent binding performance contrasts sharply with the huge libraries that need to be accessed to find specific high-affinity antibodies, aptamers, or phage-display-generated proteins. The specific binding of specific proteins is associated with unique sequences in biomolecular recognition. The introduction of a small organic residue with several functional groups by chemical synthesis dramatically relaxes the need for structural complexity in the polypeptide conjugates. The molten globule-like character of the polypeptide conjugates

contrast sharply with the high level of structural preorganization of biomacromolecules.

High-resolution structural information is not yet available, but the competitive binding by PC demonstrates that the PC binding site of CRP is targeted, and the variation in binder affinities over many orders of magnitude is clear evidence that the polypeptide sequence interacts with CRP. The lack of structural definition in the molten globule-like conjugate molecules suggests an "adapted fit" binding mechanism. From the perspective of application, the binder concept offers robust binders that do not denature and can be stored lyophilized or in solution for years. They are chromatographically pure, with little or no batch to batch variation. The chemical nature and absence of a biological history of designed binder molecules minimize the risk for undesirable interactions with endogenous biomolecules that will interfere with, for example, diagnostic tests and other applications.[12] The set of polypeptides is not limited to forming binders for CRP. We reported previously that the conjugation of the 42-residue polypeptide KE2 to benzenesulfonamide, a known inhibitor of human carbonic anhydrase II (HCAII) with a K_D value of 1.5 μM, afforded a hybrid molecule that binds HCAII with a K_D value of 20 nM in buffer.[13] A binder molecule for HCAII, based on benzensulfonamide but developed using the 16 polypeptides reported here, discriminated between HCAI and HCAII—two isoforms of human carbonic anhydrase with 60% homology—by a factor of 30, although the affinity for benzenesulfonamide is the same for both proteins.[14] Several binders are under development using the same set of polypeptides. A concept for enhancing affinity was reported previously by King and Burgen, where 4-substituted benzenesulfonamides were shown to bind more strongly to human carbonic anhydrase II the longer the aliphatic chain in the 4-position. The concept presented here is very different. A wide range of affinities is obtained even though the hydrophobic side chains are the same in all 16 polypeptides. Aliphatic side chains enhance affinity non-specifically in a hydrophobic binding pocket, but the combination of hydrophobic substituents with charged amino acid residues introduce enhanced affinity as well as selectivity.[15]

While the theory behind cooperativity in biomolecular recognition is now well understood,[16] the quantification of the level of cooperativity obtained for these binders requires that the affinity of the polypeptide can be measured precisely for the protein surface area in proximity to the small-molecule binding site, which is unrealistic.

In conclusion, the problem of creating high-affinity binders for proteins has been reduced to that of finding small-molecule binders that bind at a micromolar level, a problem that is of considerably less complexity. We believe that synthetic binders for proteins that are robust and easy to modify, and that furthermore can be generated from a comparatively simple set of polypeptide scaffolds will have a large impact on clinical diagnostics and drug development in the near future.

Experimental Section

Polypeptides were synthesized by solid-phase peptide synthesis and purified by HPLC. Affinities were determined by fluorescence titration and surface plasmon resonance (Biacore). Selectivity was evaluated by ELISA and by pull-down experiments followed by SDS-PAGE. For full experimental detail see the Supporting Information.

Received: August 12, 2010
Revised: December 12, 2010
Published online: January 14, 2011

Keywords: bioorganic chemistry · C-reactive protein · phosphocholine · polypeptides · proteins

[1] W. L. DeLano, *Curr. Opin. Struct. Biol.* **2002**, *12*, 14–20.
[2] A. A. Kossiakoff, S. Koide, *Curr. Opin. Struct. Biol.* **2008**, *18*, 499–506.
[3] A. A. Bogan, K. S. Thorn, *J. Mol. Biol.* **1998**, *280*, 1–9.
[4] J. Westhuyzen, H. Healy, *Ann. Clin. Lab. Sci.* **2000**, *30*, 133–143.
[5] S. Verma, P. E. Szmitko, P. M. Ridker, *Nat. Clin. Pract. Cardiovasc. Med.* **2005**, *2*, 29–36.
[6] W. F. DeGrado, C. M. Summa, V. Pavone, F. Nastri, A. Lombardi, *Annu. Rev. Biochem.* **1999**, *68*, 779–819.
[7] D. Thompson, M. B. Pepys, S. P. Wood, *Structure* **1999**, *7*, 169–177. PDB ID: 1B09.
[8] J. K. Anderson, R. M. Stroud, J. E. Volanakis, *Fed. Proc.* **1978**, *37*, 1495–1499.
[9] T. Christopeit, T. Gossas, U. H. Danielson, *Anal. Biochem.* **2009**, *391*, 39–44.
[10] M. B. Pepys, G. M. Hirschfield, G. A. Tennant, J. R. Gallimore, M. C. Kahan, V. Belotti, P. N. Hawkins, R. M. Myers, M. D. Smith, A. Polara, A. J. Cobb, S. V. Ley, J. A. Aquilina, C. V. Robinson, I. Sharif, G. A. Gray, C. A. Sabin, M. C. Jenvey, S. E. Kolstoe, D. Thompson, S. P. Wood, *Nature* **2006**, *440*, 1217–1221.
[11] C. Albrecht, P. Fechner, D. Honcharenko, L. Baltzer, G. Gauglitz, *Biosens. Bioelectron.* **2010**, *25*, 2302–2308.
[12] L. J. Kricka, *Clin. Chem.* **1999**, *45*, 942–956.
[13] T. Andersson, M. Lundquist, G. T. Dolphin, K. Enander, B.-H. Jonsson, J. W. Nilsson, L. Baltzer, *Chem. Biol.* **2005**, *12*, 1245–1252.
[14] L. T. Tegler, K. Fromell, B.-H. Jonsson, J. Viljanen, C. Winander, J. Carlsson, L. Baltzer, *ChemBioChem* **2010**, DOI: 10.1002/cbic.201000556.
[15] R. W. King, A. S. V. Burgen, *Proc. R. Soc. London Ser. B* **1976**, *193*, 107–125.
[16] A. Fersht, *A Guide to Enzyme Catalysis and Protein Folding*, W. H. Freeman and Company, New York, **1999**, chap. 11.

