Construction of an Artificial Receptor Protein ("Anticalin") Based on the Human Apolipoprotein D

Martin Vogt and Arne Skerra*[a]

Human apolipoprotein D (ApoD) is a prominent member of the lipocalin family of proteins and transports arachidonic acid and progesterone in various body fluids. Lipocalins share a structurally conserved β barrel as their central folding unit, which supports a set of four hypervariable loops that form the entrance to the ligand pocket. Based on this structural pattern ApoD was employed as a scaffold for the combinatorial design of artificial receptor proteins termed anticalins. After randomization of 24 amino acids located within the loop region, several ApoD variants were selected against hemoglobin, a biochemically well-characterized model target, by using bacterial phagemid display and colony screening. One variant, dubbed HbgA, was further investigated by surface plasmon resonance interaction analysis and found to complex hemoglobin specifically and with a dissociation constant of about 2 μM. While our previous work on the structurally related insect bilin-binding protein was focused on the generation of binding activity towards low-molecular-weight ligands, this study demonstrates for the first time that a lipocalin can also be tailored to recognize a protein target. The fact that even a human member of this protein family has now been successfully recruited for anticalin construction opens the possibility for future application of such engineered lipocalins as target-recognition vehicles in medical therapy.

Introduction

The concept of "anticalins" has recently been introduced for the generation of small artificial ligand-binding proteins as an alternative to antibody fragments.[1, 2] A prototypic member of the lipocalin family of proteins,[3] the bilin-binding protein (BBP) from *Pieris brassicae*, has been employed as a rigid and stable scaffold for creating cognate receptor proteins toward hapten-like ligands with high affinities and pronounced specifities by using the methodology of combinatorial protein design.[1, 4, 5]

Natural members of the lipocalin family are compact proteins that serve for the transport or storage of biomolecular compounds, like vitamins, hormones, and secondary metabolites, in many organisms. The tertiary structure of these proteins comprises a circularly closed eight-stranded antiparallel β sheet. This β barrel supports four loops at one of its ends which form the entrance to the ligand pocket. Despite low mutual sequence homology, the β barrel is structurally well conserved.[2] However, the four loops exhibit large conformational differences and even size variation between individual lipocalins, which gives rise to diverse natural ligand specificities. Hence, the lipocalin architecture resembles that of immunoglobulins, where altogether six hypervariable loops are supported on a rigid framework. In contrast to immunoglobulins, however, the lipocalins are composed of a single polypeptide chain, have a much smaller size, and their set of four loops can be more easily manipulated.[5]

To exploit the potential of the lipocalin fold for the preparation of artificial receptor proteins—termed "anticalins"[1, 2]—we first set out to extensively reshape the ligand pocket of BBP[6–8] by combinatorial protein design. For the construction of lipocalin variants that recognize prescribed small ligands, 16 amino acid positions in the four loops that essentially form the center of the pocket for the natural ligand biliverdin IX_γ were subjected to simultaneous random mutagenesis by using a PCR assembly procedure.[1, 5] A phagemid display library with 3.7×10^8 mutants was prepared and employed for panning with several immobilized compounds. In the case of fluorescein, which was chosen as a first model hapten, mutants with dissociation constants (K_D) in the low nanomolar range were selected.[1]

BBP variants with remarkable specificities for other organic molecules, including the widely applied digoxigenin group, were generated as well.[4] The crystal structures of a fluorescein-binding and of a digoxigenin-binding anticalin were elucidated recently.[9, 10] Both artificial ligands were found to be complexed in the mutated BBP structure at a very similar position to that of the natural ligand biliverdin in the wild-type protein. While the β barrel architecture was perfectly retained in both anticalins, the loops at the entrance to the ligand pocket exhibited large conformational differences among each other and BBP, most likely as a direct result of the side chain replacements. This

[a] *Dr. M. Vogt, Prof. Dr. A. Skerra*
Lehrstuhl für Biologische Chemie
Technische Universität München
85350 Freising-Weihenstephan (Germany)
Fax: (+49) 8161-714352
E-mail: skerra@wzw.tum.de

 DOI: 10.1002/cbic.200300703

observation reveals remarkable structural plasticity of the lipocalin loop region.

Anticalins can be produced at high levels in *Escherichia coli* and they are highly stable, with typical denaturation temperatures above 60 °C.[11] They are also amenable to the production of functional fusion proteins, for example with enzymes like alkaline phosphatase, which may serve as a convenient reporter enzyme for diagnostic applications.[4] Furthermore, the linking of two anticalins that recognize different ligands results in useful bispecific binding proteins, termed "duocalins".[12]

So far, only BBP has been utilized as a scaffold for the preparation of anticalins. However, it seems likely that other members of the lipocalin family, especially those belonging to the prototypic subgroup,[2] should be similarly adaptable to this approach. Human lipocalins are of particular interest in this respect because anticalins derived from an autologous protein scaffold should possess lower immunogenicity. Such anticalins might be valuable for therapeutic applications, for example, by providing the recognition module for cell surface receptors in recombinant immunotoxins, thus replacing the single-chain antibody fragments currently in use.[13] Among the possible lipocalin candidates, human apolipoprotein D (ApoD) has been postulated to be structurally most closely related to BBP.[14] Therefore, ApoD provides an interesting model for the generation of a first "human" anticalin.

Natural ApoD is an abundant plasma protein that is peripherally associated with high-density lipoprotein (HDL) particles through a disulfide bond with ApoAII.[15] In contrast to conventional apolipoproteins, the primary structure of ApoD contains no amphipathic α helices but instead appears to belong to the lipocalin fold.[14] Human ApoD also occurs as a monomeric protein in certain tissues and seems to play a lesser role in lipid transport in other mammals such as mice and rabbits, where the characteristic Cys residue needed for the covalent cross-link to HDL is missing (for further discussion, see ref. [16]).

Human ApoD is a glycoprotein and has been reported to complex a variety of small ligands, depending on the tissue from which it was isolated. Nevertheless, it can also be produced as a soluble protein without glycosylation. For example, secretion into the periplasm of *E. coli* yields functional protein whereby the two intrachain disulfide bonds are efficiently formed.[16] Thorough binding studies with this recombinant protein revealed measurable affinities (in the low micromolar range) only for progesterone and arachidonic acid. Thus, although previous reports have indicated a more promiscuous binding behavior, ApoD clearly possesses rather specific recognition properties for these two low-molecular-weight compounds.

In the present study, we utilized recombinant human ApoD as a scaffold for the generation of anticalins. In extension to our previous work, which was mainly directed at small prescribed ligands, we chose to use hemoglobin as a larger, antigen-like protein target that is readily available and whose three-dimensional structure is known.[17]

Results

Construction of a combinatorial library for ApoD

Based on the published structural model of human ApoD,[14] a set of 24 amino acids assumed to be located within the four structurally variable loops of the generic lipocalin ligand-binding site[2] were chosen for site-directed random mutagenesis (Figure 1): residues Thr34, Thr35, Phe36, Glu37, and Asn38 in peptide

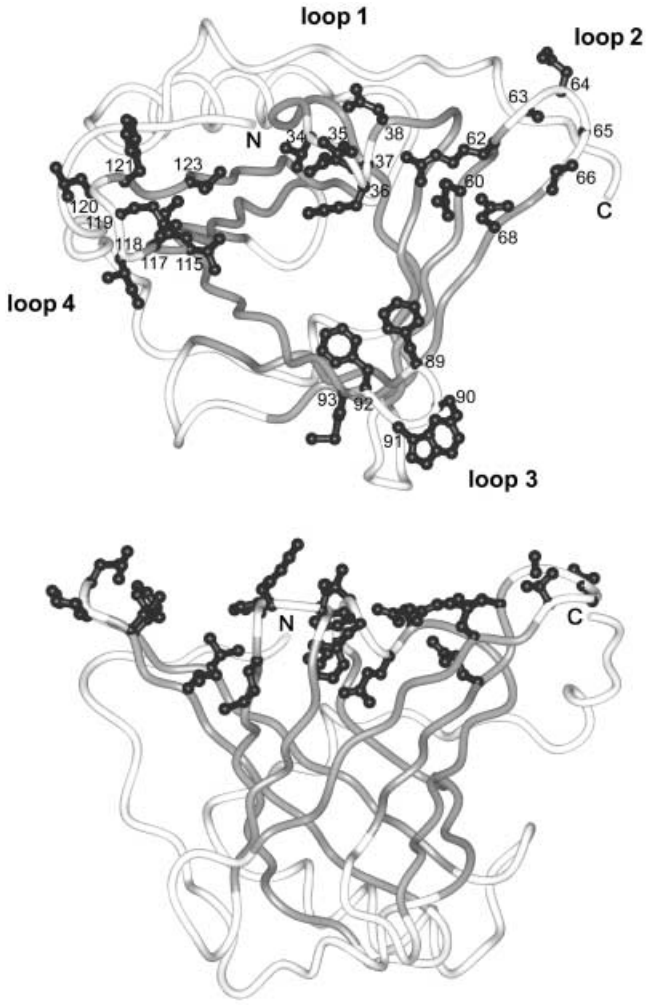

Figure 1. *Three-dimensional model of apolipoprotein D with the positions chosen for random mutagenesis labeled. The polypeptide Cα backbone is shown in a ribbon representation (InsightII molecular modeling software, MSI) based on the structural model of Peitsch and Boguski.[14] Top: view into the putative ligand pocket; bottom: side view. The part of the molecule that forms the β barrel is colored grey (residues 24 – 27, 43 – 48, 55 – 62, 67 – 75, 82 – 88, 93 – 103, 108 – 117, 121 – 130). The 24 amino acid positions in the four loops (marked 1 – 4) of the lipocalin that were chosen for random mutagenesis are depicted with their side chains in black and labeled with residue numbers.*

loop 1, which connects the β strands A and B; Glu60, Arg62, Ala63, Asp64, Gly65, Thr66, and Asn68 in loop 2, which connects β strands C and D; Phe89, Ser90, Trp91, Phe92, and Met93 in loop 3, which connects β strands E and F; Thr115, Ile117, Ile118, Gln119, Leu120, Phe121, and Val123 in loop 4, which connects β

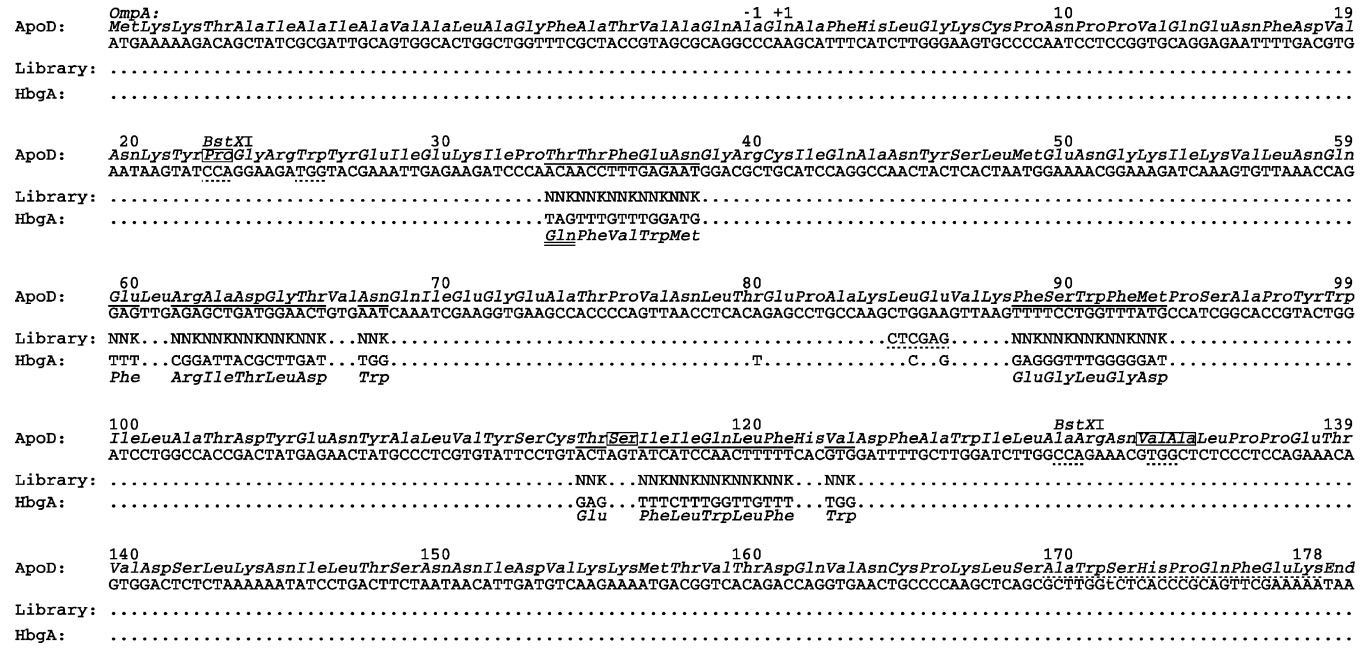

Figure 2. *DNA and corresponding amino acid sequences of wild-type ApoD, together with those of the ApoD random library and the selected variant HbgA. Sequences are shown as encoded on pApoD10, beginning with the bacterial OmpA signal peptide and ending with the* Strep*-tag II. Amino acid positions 34 – 38, 60, 62 – 66, 68, 89 – 93, 115, 117 – 121, and 123 were subjected to random mutagenesis employing NNK codons (N = any nucleotide; K = G or T) and are underlined. For the genetic library scheme and the single variant HbgA, only the bases and amino acids differing from those in the ApoD sequences are depicted. The recombinant ApoD gene carried the following amino acid exchanges with respect to its natural counterpart (see the text for details): Leu23Pro, Cys116Ser, Pro133Val, and Asn134Ala (boxed). Two restriction sites for* Bst*XI (CCAN$_6$TGG) flanking the whole loop region, and a site for* Xho*I (CTCGAG), which was introduced during PCR mutagenesis at amino acid positions 85/86 for diagnostic reasons, are marked with broken lines. The* Strep*-tag II is labeled with a broken line as well. An amber stop codon (TAG) that was found at amino acid position 34 in the variant HbgA and partially translated as Gln in the* E. coli sup*E strain XL1-Blue is doubly underlined. To enable efficient biosynthesis of this anticalin, the amber codon was later substituted by a Gln triplet (see the text for details).*

strands G and H (Figure 2). Residues in the exposed positions of each loop were chosen for random mutagenesis in order to ensure enhanced structural plasticity[10] and provide a large potential contact surface for the macromolecular target.

To randomize these residues in a concerted manner, a generic PCR assembly strategy was applied[5] that made use of four oligodeoxynucleotides altogether with degenerate (NNK) codons. The chosen kind of degeneracy allowed all 20 amino acids but avoided the formation of stop codons, except for the amber stop codon (TAG), which could be tolerated because the *sup*E *E. coli* strain XL1-Blue[18] was used for phagemid production (translating the TAG codon as Gln). In a first step, two separate PCR reactions were performed that introduced random mutations at the selected positions in either loops 1 and 2 or loops 3 and 4, respectively (Figure 2). In the second step, the two resulting DNA fragments were assembled by PCR splicing through overlap extension in the presence of two flanking primers and one oligodeoxynucleotide bridging the central gap.

In addition to the randomized base triplets at the chosen positions, the mutated genes encoded the replacement Cys116Ser, which has been previously introduced to avoid possible dimerization or disulfide isomerization effects due to the presence of an unpaired thiol side chain.[16] The substitutions Leu23Pro, Pro133Val, and Asn134Ala were introduced to generate two distinct *Bst*XI restriction sites that allowed efficient cloning of the PCR fragment. The resulting genetic library of ApoD variants was inserted into the phagemid vector pApoD19, which codes for a fusion between ApoD and the full-length pIII protein of filamentous phage M13,[19] with an amber stop codon in between (Figure 3). Expression studies with phagemids presenting ApoD as a fusion either with the entire phage coat protein or with fragments of pIII revealed a better yield for the full-length pIII chimeric protein, as determined by gel electrophoresis of phagemid proteins and detection with an anti-pIII antibody, while no significant effect on the achievable phagemid titer was observed (not shown).

Transformation of *E. coli* XL1-Blue yielded a genetic library with 1.7×10^9 independent transformands. Sequence analysis of 10 randomly chosen clones showed that the phasmid-encoded structural genes carried mutations at the desired positions while the original sequence of ApoD was otherwise preserved by the PCR strategy. However, in some of the clones additional mutations were detected: three variants carried single point mutations, presumably caused by the natural error rate of the polymerase, and two clones showed frameshift mutations. Nevertheless, the majority of the genes appeared to encode suitable mutants of ApoD. After superinfection with a helper phage of the whole *E. coli* population harboring the phasmid library, the phagemids displaying the ApoD variants as fusions with the minor coat protein pIII on their surfaces were prepared.

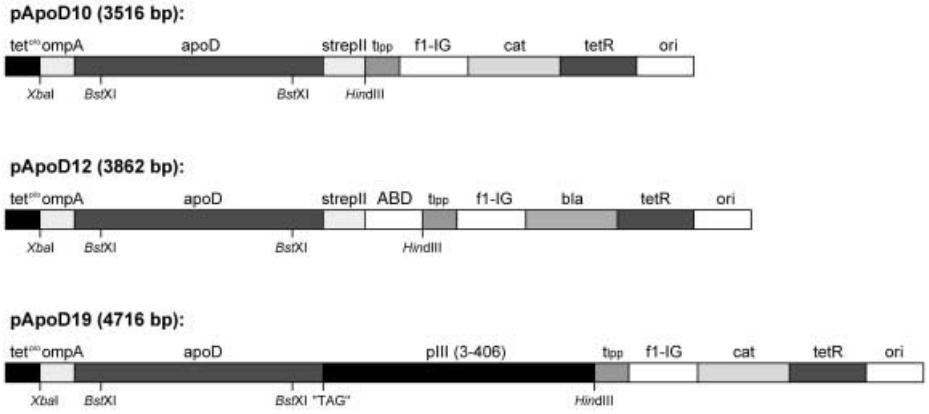

Figure 3. *Linearized representation of the vectors used in this study. The plasmids are based on the generic vectors pASK75 or pASK111, wherein the recombinant gene expression is tightly regulated by the tetracycline promoter/operator ($tet^{p/o}$). All the vectors used encode fusion proteins carrying the N-terminal OmpA signal peptide for translocation across the bacterial inner cell membrane. pApoD10 enables the production of soluble ApoD or its variants as fusions with the* Strep-*tag II (strepII) at the C terminus. pApoD12 encodes a fusion of ApoD with the* Strep-*tag II and the bacterial albumin-binding domain (ABD). pApoD19 was developed for phagemid display of the ApoD variants and encodes a fusion with the minor phage coat protein pIII, which comprises amino acids 3 – 406 of the mature polypeptide chain. In this case, transcription of the fusion gene is partially interrupted by a suppressible amber stop codon (TAG) introduced between the ApoD and pIII sequences. In all constructs, the ApoD gene carries two mutually incompatible restriction sites (*BstXI*) that permit the directional insertion and subcloning of the randomized gene cassette. The transcription ends with the lipoprotein terminator (t_{lpp}). The other elements of the vectors are the intergenic region of the phage f1 for single-stranded replication, an antibiotic resistance gene (cat, chloramphenicol acetyl transferase; bla, β-lactamase), the tetracycline repressor gene (tetR), and the origin of replication (ori).*

Selection of an ApoD mutant with binding activity towards hemoglobin

Human hemoglobin was chosen as an exemplary protein target because of its well-known biochemical and structural properties and because related proteins like myoglobin are available for comparison.[17] For the selection of cognate ApoD variants from the phagemid library, hemoglobin was adsorbed to the polystyrene surface of Nunc immunosticks. The freshly prepared phagemid solution was applied, incubated, and the sticks were washed several times. Remaining bound phagemids were eluted at pH 2.2, immediately neutralized, and used for amplification by reinfection of XL1-Blue cells.

Five cycles of repeated panning and amplification of selected phagemids were carried out. The phasmid DNA of the enriched particles obtained from the last cycle was isolated as a pool after infection of XL1-Blue cells. To qualitatively analyze the binding activities of individual ApoD variants towards hemoglobin, a filter sandwich colony screening assay was performed.[4] To this end, the gene cassette encoding the ApoD variants was subcloned—as a mixture from the phasmid pool—onto the vector pApoD12 (Figure 3), which codes for a fusion protein of ApoD with ABD.[20]

Transformed *E. coli* cells were plated on a hydrophilic membrane on top of an agar plate with selective medium. After visible colonies had formed, this membrane was laid on a hydrophobic capture membrane coated with human serum albumin (HSA) and placed on top of an agar plate containing the chemical inducer of gene expression. During the following phase of recombinant protein production, synthesized fusion proteins were secreted into the bacterial periplasm and partially released from the colonies. After diffusion through the hydrophilic support membrane, the proteins became fixed to the capture membrane by biomolecular complex formation between ABD and HSA. The two membranes were separated and the immobilized ApoD variants were tested for target-binding activity by incubation in the presence of hemoglobin labeled with digoxigenin groups. After washing, bound conjugate was detected in a chromogenic reaction with an antibody – alkaline phosphatase conjugate directed against digoxigenin. 16 colonies giving rise to distinct color signals were finally recovered from the first membrane, propagated, and their plasmid DNA was isolated.

Each mutagenized ApoD gene cassette was analyzed by DNA sequencing. An obviously functional gene product was detected in eight of the colonies. The remaining eight colonies carried either frameshifts ($2\times$), insertions or deletions ($3\times$), or additional Cys residues and/or multiple amber stop codons ($3\times$), which were undesirable for soluble protein production. Among the functional clones, one sequence, which corresponded to the most intense color signals in the colony assay, occurred twice. This variant was designated HbgA and its sequence was translated and compared with that of wild-type ApoD (Figure 2).

HbgA differs from ApoD in 21 of the 24 randomized amino acid positions. Arg62, Leu120, and Phe121 remained conserved, although different codons were detected for the first two of these residues. Only one silent point mutation was found outside the randomized regions. No additional Cys residue was observed, but an amber stop codon, which was partially translated as Gln in the *sup*E strain, was present in the first mutagenized loop.

Production and characterization of the soluble anticalin HbgA

After subcloning of the mutagenized gene cassette onto pApoD10 (Figure 3), HbgA was produced in *E. coli* JM83 by secretion into the periplasm in a similar fashion as described before for the production of the functional recombinant wild-type ApoD.[16] To ensure efficient biosynthesis in this *sup*E$^-$ strain, the amber codon (TAG) at amino acid position 34 (Figure 2) was replaced with a Gln codon (CAG) by site-directed mutagenesis. The ApoD variant was subsequently isolated as a soluble protein from the periplasmic cell fraction and purified in one step by streptavidin affinity chromatography using the C-terminal *Strep*-tag II.[21] The yield of purified protein was approximately 0.2 mg L^{-1} culture medium, which is similar to the yield obtained

for the recombinant ApoD.[16] SDS-PAGE analysis of HbgA revealed its purity, while its electrophoretic mobility was comparable to that of the recombinant wild-type ApoD, corresponding to an apparent molecular mass of 24 kDa (Figure 4).

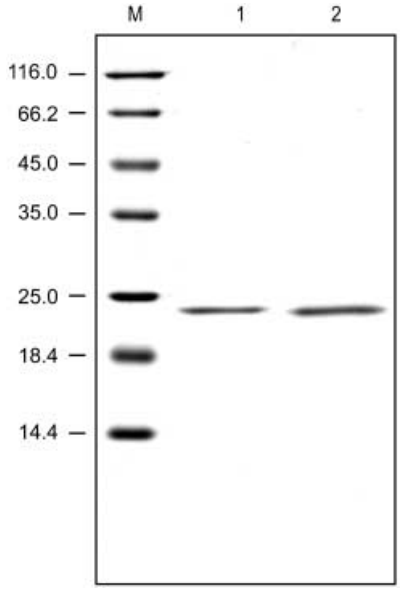

Figure 4. *SDS-PAGE analysis of the bacterially produced wild-type ApoD (lane 1) and HbgA (lane 2). Proteins were produced in* E. coli *JM83 harboring the expression vector pApoD10 and were purified from the periplasmic cell fraction by streptavidin affinity chromatography using the* Strep-*tag II.*

To investigate the structural properties of HbgA, circular dichroism (CD) measurements were carried out (Figure 5). The CD spectrum in the far-UV region revealed the typical features of a β-sheet protein, with a maximum at 195 nm and a local minimum at about 210 nm (see typical spectral features of BBP described in ref. [1]). The spectrum of HbgA shows considerable similarity with that of wild-type ApoD, which indicates that HbgA retained the secondary structure of the lipocalin even though in total 25 amino acid residues were exchanged.

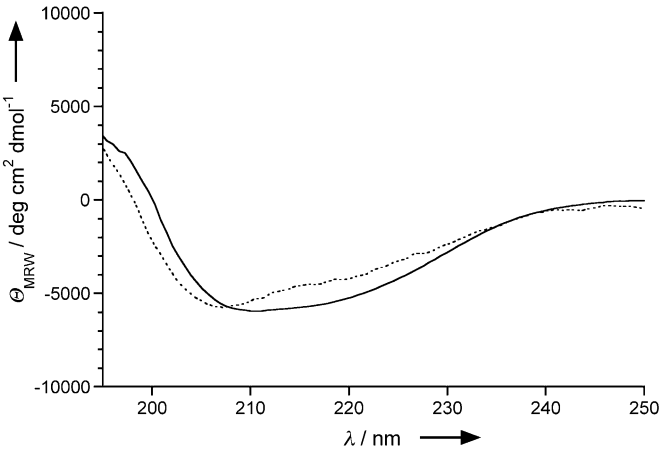

Figure 5. *Circular dichroism spectra of recombinant wild-type ApoD (broken line) and HbgA (solid line). The purified proteins (10 μM) were dissolved in sodium phosphate buffer (10 mM, pH 7.5). The spectra were recorded on a J-810 instrument in a quartz cuvette (2-mm path length) at 22 °C.*

Determination of the target-binding properties of HbgA

The binding of the selected ApoD variant to the prescribed macromolecular target hemoglobin was analyzed by surface plasmon resonance (SPR) spectroscopy. For this purpose, hemoglobin was covalently immobilized on a Biacore sensor chip. The anticalin HbgA was applied in various concentrations and the resonance unit values attained at equilibrium—after near completion of the association phase—were measured. When these data were plotted, a simple concentration-dependent interaction with the immobilized hemoglobin became apparent (Figure 6). A dissociation constant of 2.16 ± 0.36 μM

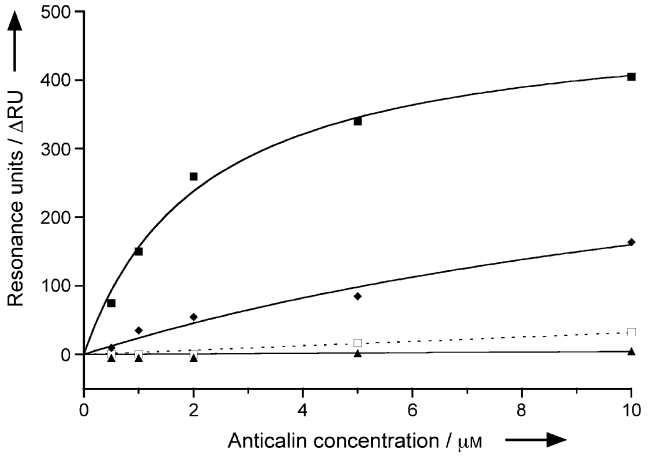

Figure 6. *Analysis of the ligand-binding activity of the engineered lipocalin HbgA by surface plasmon resonance spectroscopy. Human hemoglobin was covalently immobilized on a CM5 sensor chip as the target and changes in the resonance signal were detected upon application of ApoD or its variant at several concentrations. The steady-state equilibrium values obtained (HbgA, filled squares; recombinant ApoD, open squares) were measured and plotted versus the applied protein concentration. Nonlinear least squares regression of the data revealed a dissociation constant of 2.16 ± 0.36 μM for the complex between HbgA and hemoglobin. Similar experiments were performed for HbgA with BSA (triangles) or myoglobin (rhombus) as the immobilized target to assess the specificity of binding.*

for the complex formed between HbgA and hemoglobin was deduced by nonlinear least squares regression according to the equation of mass law for bimolecular complex formation. In contrast, the recombinant wild-type ApoD did not show detectable binding to hemoglobin when applied in a control experiment under identical conditions.

The surface concentrations of hemoglobin and HbgA, which were calculated according to the measured resonance units for the covalently immobilized target protein and the maximal amount of bound anticalin, respectively, revealed a ratio of 0.27

molecules HbgA per hemoglobin. This value indicates that only part of the sensor-coupled hemoglobin remained functionally accessible. To investigate the specificity of the interaction between HbgA and its target protein, two series of control experiments were performed. First, when a sensor chip coated with BSA was used, no cross-reactivity of HbgA could be detected. Second, when myoglobin was used as a control target, a moderate interaction with HbgA was observed (Figure 6). This weak complex formation may reflect the known structural similarity between myoglobin and hemoglobin. The observation also suggests that HbgA probably does not bind at a subunit interface but rather at the exposed surface of either the α or β subunit of hemoglobin.

Finally, to determine whether the ApoD variant HbgA had retained binding activity towards the natural ligand progesterone, a fluorescence titration experiment was performed as previously described.[16] No binding was detected, which indicates that the ligand pocket was considerably reshaped by the mutations that were introduced into the loop region.

Discussion

In this study we were able to generate a variant of human ApoD that exhibits novel binding specificity towards the prescribed protein target hemoglobin. Our observations reveal that ApoD is another member of the lipocalin family that provides a stable scaffold for the generation of diverse binding sites. In the selected variant HbgA, this lipocalin tolerates the exchange of altogether 21 amino acids that are likely located within the set of four loops at the open end of the characteristic β barrel structure. In addition, this variant has four other side chains exchanged from those in the natural ApoD (see above) and it lacks the glycosylation of the human plasma protein. Nevertheless, HbgA is a soluble protein that can be isolated from the periplasmic cell fraction of the bacterial expression host and essentially retains the secondary structure of ApoD, as revealed by CD spectroscopy. Not unexpectedly, binding activity towards progesterone was lost as a consequence of mutating the ligand pocket.

The fact that hemoglobin successfully served as a protein target demonstrates that the loop region of lipocalins can in principle recognize a macromolecular "antigen" of considerable size. Up to now, ApoD was only known to form complexes with low-molecular-weight compounds. The dissociation constant between the variant HbgA and hemoglobin was measured as approximately 2 μM by using the Biacore system. Moreover, the binding of this anticalin to its protein target is rather specific and discrimination of the homologous protein myoglobin was observed.

The achieved target affinity is comparable to that between ApoD and its original hapten-like ligand progesterone at neutral pH, as determined by fluorescence titration.[16] It seems likely that the affinity of HbgA for hemoglobin can be increased further by in vitro affinity maturation, as was previously demonstrated in the case of a digoxigenin-binding anticalin derived from BBP.[4, 11] To this end, partial mutagenesis of HbgA could be employed, for example by randomization of one loop in the context of fixed sequences for the other three loops, followed by another phage display selection against the target protein.

Unfortunately, the lack of experimentally reliable structural information about the precise three-dimensional fold of ApoD (from X-ray crystallography or multidimensional NMR spectroscopy) prevented us from choosing a smaller set of side chains that are likely to provide effective contacts with the macromolecular target. Instead, a rather larger number of 24 amino acid residues was considered in order to take into account some structural uncertainty of the homology model in the exact positioning of the corresponding side chains. As a result of the randomization strategy, our library had a theoretical complexity of $32^{24} \approx 1.3 \times 10^{36}$ possible codons, which exceeds the experimentally achieved diversity of our phage display library by many orders of magnitude. However, for steric reasons (see Figure 1) only a fraction of the 24 side chains is likely to interact with the protein antigen and should thus contribute to the "functional" diversity of the library. In contrast, mutations in several of the remaining position must be expected to have a deleterious effect on the folding of the protein, especially if they turn out to be, for example, part of the conserved β barrel structure.

Consequently, the dissociation constant of 2 μM achieved in the example experiment described herein is in fact not far from the value that can be expected under these circumstances and compares nicely with the affinities described for earlier selection experiments with combinatorial libraries of antibody fragments[22, 23] or alternative scaffold proteins, such as affibodies.[24] Based on the results we have obtained so far, a more focused library of ApoD variants may be prepared in the future, and in this manner higher affinities should be attainable.

In conclusion, we have taken two important steps in the development of anticalin technology.[5] First, a human lipocalin was successfully recruited as a scaffold and shown to retain its protein fold even after extensive mutagenesis of its binding site. Second, the specific recognition of a prescribed protein target by an anticalin was demonstrated. Consequently, the way has been paved for the preparation of other anticalins, for example directed against extracellular domains of cell surface proteins, with potential application in medical therapy.

Experimental Section

Strains and vectors: *E. coli* strain XL1-Blue[18] was used for the production and amplification of phagemids. *E. coli* TG1[25] was employed for the biosynthesis of the ApoD – ABD fusion proteins in the colony screening assay. *E. coli* JM83[26] was used for production of the soluble recombinant ApoD and its engineered variants.

All plasmids used for phage display and soluble protein production were based on the generic expression vector pASK75,[27] in which recombinant gene expression is tightly regulated by the chemically inducible tetracycline promoter ($tet^{p/o}$). The previously constructed expression vector pApoD5,[16] which codes for a fusion protein between the bacterial OmpA signal sequence, mature ApoD (with the mutation Cys116Ser), and *Strep*-tag II[21] served as a starting point. For the subcloning of randomized gene cassettes encoding the four peptide loops of ApoD, two *Bst*XI recognition sites were introduced

by site-directed mutagenesis according to published procedures[28, 29] by using the oligodeoxynucleotides askD27 (5′-TTC GTA CCA TCT TCC TGG ATA CTT ATT CAC GT-3′) and askD28 (5′-GTT TCT GGA GGG AGA GCC ACG TTT CTG GCC AAG ATC CAA-3′). During this step, the amino acid substitutions Leu23Pro, Pro133Val, and Asn134Ala were introduced. The resulting vector was designated pApoD10 and was used for the periplasmic secretion of soluble "wild-type" ApoD.

The vector pApoD19 was constructed for the production of a phagemid random library of ApoD. pApoD19 codes for a fusion protein between the OmpA signal peptide, the mutated ApoD encoded on pApoD10, a suppressible amber stop codon followed by a short linker sequence (Ala-Gly-Gly-Ala), and the minor coat protein pIII of filamentous bacteriophage M13,[19] (amino acids 3 – 406 of the mature polypeptide). For the colony screening assay of individual ApoD variants the vector pApoD12 was prepared. pApoD12 encodes a fusion protein between the OmpA signal peptide, the mutated ApoD described above, the *Strep*-tag II, and ABD.[20]

Recombinant DNA techniques and library construction: For the concerted mutagenesis of 24 predefined amino acid positions distributed across all four peptide loops of ApoD, a two-step gene assembly strategy involving the polymerase chain reaction (PCR) was applied.[5] In the randomization step, two separate amplification reactions (20 cycles) were performed with pApoD10 plasmid DNA as template and two appropriate pairs of primers: MV-4 (5′-GAT GGT ACG AAA TTG AGA AGA TCC CAN NKN NKN NKN NKN NKG GAC GCT GCA TCC AGG-3′) and MV-5 (5′-GGC TTC ACC TTC GAT TTG MNN CAC MNN MNN MNN MNN MNN CAA MNN CTG GTT TAA CAC TTT GAT C-3′) in one reaction, MV-6 (5′-GAG CCT GCC AAG CTC GAG GTT AAG NNK NNK NNK NNK NNK CCA TCG GCA CCG TAC TG-3′) and MV-7 (5′-CCA AGA TCC AAG CAA AAT CMN NGT GMN NMN NMN NMN NMN NAC TMN NAC AGG AAT ACA CGA GGG C-3′) in the other. These primers carried degenerate codons at the desired positions and each covered one loop region of ApoD. The reactions were performed in a total volume of 50 μL, with *Taq* DNA polymerase (2.5 units; Promega, Madison, USA) and buffer according to the manufacturer's recommendation, plus template DNA (10 ng) and the two primers (25 pmol each).

The resulting amplification products with sizes of 149 bp and 147 bp, respectively, were assembled in a second PCR (20 cycles) that was carried out on the two fragments (3 ng each, total volume 100 μL) by adding the flanking primers (50 pmol each) MV-8 (5′-GTG AAT AAG TAT CCA GGA AGA TGG TAC GAA ATT GAG AAG-3′) and MV-10 (5′-CTG GAG GGA GAG CCA CGT TTC TGG CCA AGA TCC AAG CAA AAT C-3′) as well as a third oligodeoxynucleotide (1 pmol) to bridge the gap between the two fragments, MV-9 (5′-CCT CGA GCT TGG CAG GCT CAG TGA GGT TAA CTG GGG TGG CTT CAC CTT CGA TTT G-3′). A prominent PCR product (358 bp) was obtained, isolated by agarose gel electrophoresis, and cut at both ends with *Bst*XI. The resulting fragment (322 bp) was inserted into the similarly cut vector pApoD19 by ligation of the PCR product (3.25 μg) with the digested vector (43 μg) as previously described.[1] After ethanol precipitation, the DNA was used for transformation of electrocompetent *E. coli* XL1-Blue cells[30, 31] by means of a MicroPulser (Bio-Rad Laboratories, Munich, Germany).

DNA sequencing was performed from plasmid DNA with appropriate oligodeoxynucleotides as primers on a Genetic Analyzer 310 with the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit (both from Applied Biosystems, Weiterstadt, Germany).

Phage display and affinity selection of ApoD variants: 46 separate electroporations were carried out in cuvettes (2-mm gap width, Bio-Rad) containing the library DNA (1 μg per reaction) and electrocompetent cells (100 μL per reaction). The suspension was then shaken (1 h at 37 °C) and afterwards diluted in 2xYT medium[32] (1.5 L) containing chloramphenicol (Cam; 35 μg mL^{-1}), followed by cultivation until the optical density at 550 nm (OD_{550}) had increased by 0.5 units. Phagemid particles were produced from this culture (200 mL) after superinfection with VCS-M13 helper phage (Stratagene, Heidelberg, Germany) according to published procedures.[4]

For affinity selection toward hemoglobin as target, approximately 3×10^{12} phagemids displaying the mutated ApoD were subjected to panning with immunosticks (Nunc, Wiesbaden, Germany) that had been coated overnight with a solution of human hemoglobin (Sigma, Deisenhofen, Germany; 800 μL, 500 μg mL^{-1}) in phosphate-buffered saline (PBS; 4 mM KH_2PO_4, 16 mM Na_2HPO_4, 115 mM NaCl, pH 7.4) and blocked (1.2 mL, 2 h) with bovine serum albumin (BSA; 2% (w/v)) in PBS/T (PBS containing 0.1% Tween 20). The immunosticks were incubated (1 h) with a mixture of the phagemid solution (250 μL) and PBS/T containing 3% (w/v) BSA (500 μL) and then washed with PBS/T (eight times, each with 950 μL for 2 min). Bound phagemids were finally eluted in the presence of glycine/HCl (0.1 M, pH 2.2, 950 μL; 15 min) and immediately neutralized by addition of tris(hydroxymethyl)aminomethane (0.5 M, 150 μL). The solution was directly used for the infection of *E. coli* XL1-Blue cells to determine the phagemid titer and either amplify the phagemids for repeated selection cycles or prepare the phasmid DNA.[4] For phagemid amplification, an exponentially growing culture of XL1-Blue (3 mL) was infected (30 min at 37 °C) with the elution fraction (1090 μL). After sedimentation, the cells were resuspended in LB/Cam (800 μL), plated on four LB agar plates (∅ = 14 cm) containing Cam (35 μg mL^{-1}), and incubated (16 h at 32 °C). The dense lawn of colonies was then resuspended in LB/Cam (50 mL) and used for the inoculation of prewarmed LB/Cam (200 mL) to $OD_{550} = 0.08$, followed by the production of fresh phagemids (see above).

Colony screening assay: The target-binding properties of the enriched ApoD variants from the last cycle of phagemid selection were analyzed in an individual manner by means of the filter-sandwich colony screening assay.[4] The gene cassette flanked by the two *Bst*XI cleavage sites was subcloned from pApoD19, corresponding to the pooled phasmid DNA obtained from the elution step of the fifth enrichment cycle, onto pApoD12. Briefly, $CaCl_2$-competent *E. coli* TG1/F$^-$ cells were transformed[32] with the ligation mixture and plated onto a hydrophilic membrane (GVWP, 0.22 μm; Millipore, Eschborn, Germany) placed on a petri dish with LB/ampicillin (Amp; 100 μg L^{-1}) agar such that after incubation (7 h at 37 °C) approximately 500 colonies were obtained. The membrane was then placed onto a hydrophobic membrane (Immobilon-P, 0.45 μm, Millipore) that had been coated with HSA (Sigma) on an LB/Amp agar plate containing anhydrotetracycline (200 μg L^{-1}). During further incubation (15 h at 22 °C) the secreted ApoD – ABD fusion proteins were partially released from the colonies and were captured on the hydrophobic membrane through interaction between ABD and HSA. The lower membrane was then removed and washed with PBS/T (3 times).

Finally, the immobilized ApoD variants were probed for target-binding activity by incubation (1 h) with a solution of digoxigenin-labeled hemoglobin (1 μM), which was prepared by following a published protocol.[4] The bound conjugate was detected by treatment with an antidigoxigenin F_{ab} fragment/alkaline phosphatase conjugate (Roche Diagnostics, Mannheim, Germany; dilution 1:1000 in PBS/T), followed by staining with standard chromogenic substrates until distinct color signals appeared at the positions of some of the colonies. The colonies corresponding to the most intense signals were identified, recovered from the first membrane, and propagated for plasmid DNA sequencing and further investigation.

Protein production and analysis: For the preparative production of ApoD variants, the gene cassette from the corresponding clone was transferred from pApoD12 onto the expression vector pApoD10 by using the *Bst*XI restriction sites. In the case of HbgA, the amber stop codon at amino acid position 34 was replaced with a Gln codon by site-directed mutagenesis with the oligodeoxynucleotide MV-21 (5′-CCC ATC CAA ACA AAC TGT GGG ATC TTC T-3′) to allow functional production in the *sup*E$^-$ strain JM83.[16] Purification from the periplasmic cell fraction by use of the C-terminal *Strep*-tag II typically yielded 0.4 mg HbgA from a 2-L shaker flask of culture. SDS-PAGE was performed by using standard slab gel methodology with the buffer system of Fling and Gregerson[33] and Coomassie Brilliant Blue R 250 (Carl Roth, Karlsruhe, Germany) for staining.

Protein concentration was determined by measuring the absorption at 280 nm. A calculated extinction coefficient[34] of 54600 $M^{-1}cm^{-1}$ was used for HbgA and an experimentally corrected extinction coefficient[34] of 34150 $M^{-1}cm^{-1}$ for the recombinant ApoD. Secondary structure was analyzed by CD spectroscopy on a J-810 instrument (Jasco, Tokyo, Japan). For this measurement, the protein solution was concentrated to approximately 10 μM by means of Ultrafree-4 centrifugation units (Millipore) and dialyzed against sodium phosphate buffer (10 mM, pH 7.5). Spectra were recorded between 195 and 240 nm at 22 °C by accumulating 16 runs and were normalized for the amino acid content of the respective protein according to Equation (1):

$$\theta_{MRW} = \frac{\theta_\lambda M_R}{10cdN_A} \quad (1)$$

where θ_λ denotes the measured ellipticity, c the protein concentration [mg mL^{-1}], d the path length of the quartz cuvette (0.2 cm), N_A the number of amino acids (178, including the *Strep*-tag II), and M_R the molecular mass (20608 Da for HbgA and 20341 Da for recombinant ApoD).

Real-time protein interaction analysis: Surface plasmon resonance measurements were carried out on a Biacore X system (Biacore, Freiburg, Germany). Hemoglobin was immobilized as the target protein at approximately 5000 resonance units (RU) by amine coupling to the carboxylated dextrane layer in one of two flow channels of a CM5 sensor chip (Biacore) according to the manufacturer's instructions. 2-[4-(2-hydroxyethyl)-1-piperazinyl]ethanesulfonic acid (HEPES)-buffered saline (HBS; 10 mM HEPES/HCl, 150 mM NaCl, 3.4 mM ethylenediaminetetraacetate (EDTA), pH 7.4) containing surfactant P20 (0.005% (v/v); Biacore) was used as running buffer (flow rate 10 μL min^{-1}). Purified recombinant ApoD or its variant HbgA was dialyzed against HBS, concentrated as described above, and filter sterilized (0.45 μM, Spin X, Costar, Corning, USA). A dilution series with protein concentrations between 10 μM and 0.5 μM in HBS/P20 was prepared and interaction with the immobilized hemoglobin was measured after application of each protein solution (75 μL). Regeneration of the sensor chip was achieved by extensive washing with running buffer.

Steady-state resonance values were determined at the end of each association phase for the channel with the immobilized antigen and corrected for the buffer effect measured in the second channel. The resulting equilibrium values were plotted against each applied anticalin concentration. The data were fitted by nonlinear least squares regression by using Equation (2):

$$[P \cdot L] = [L]_t \frac{[P]_t}{[P]_t + K_D} \quad (2)$$

with $[L]_t$ and K_D as parameters. $[L]_t$ is the total concentration of functionally immobilized hemoglobin (equivalent to the ΔRU value obtained for the bound anticalin at saturation), $[P]_t$ is the applied anticalin concentration, $[P \cdot L]$ corresponds to the concentration of the complex after equilibration (i.e. the measured ΔRU value), and K_D is the apparent dissociation constant. The interaction between HbgA and myoglobin (horse sceletal muscle, Sigma) or BSA was analyzed as control. For this purpose, approximately 5000 RU protein was immobilized in each case to the sensor chip as described above.

Acknowledgements

The authors wish to thank Ina Theobald for help in the construction of phagemid vectors. This work was supported by the Fonds der Chemischen Industrie.

Keywords: combinatorial biochemistry · ligand binding · lipocalins · phage display · protein design

[1] G. Beste, F. S. Schmidt, T. Stibora, A. Skerra, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 1898 – 1903.
[2] A. Skerra, *Biochim. Biophys. Acta* **2000**, *1482*, 337 – 350.
[3] D. R. Flower, *Biochem. J.* **1996**, *318*, 1 – 14.
[4] S. Schlehuber, G. Beste, A. Skerra, *J. Mol. Biol.* **2000**, *297*, 1105 – 1120.
[5] A. Skerra, *J. Biotechnol.* **2001**, *74*, 257 – 275.
[6] F. S. Schmidt, A. Skerra, *Eur. J. Biochem.* **1994**, *219*, 855 – 863.
[7] R. Huber, M. Schneider, O. Epp, I. Mayr, A. Messerschmidt, J. Pflugrath, H. Kayser, *J. Mol. Biol.* **1987**, *195*, 423 – 434.
[8] R. Huber, M. Schneider, I. Mayr, R. Müller, R. Deutzmann, F. Suter, H. Zuber, H. Falk, H. Kayser, *J. Mol. Biol.* **1987**, *198*, 499 – 513.
[9] I. P. Korndörfer, G. Beste, A. Skerra, *Proteins Struct. Funct. Genet.* **2003**, *52*, 121 – 129.
[10] I. P. Korndörfer, S. Schlehuber, A. Skerra, *J. Mol. Biol.* **2003**, *330*, 385 – 396.
[11] S. Schlehuber, A. Skerra, *Biophys. Chem.* **2002**, *96*, 213 – 228.
[12] S. Schlehuber, A. Skerra, *Biol. Chem.* **2001**, *382*, 1335 – 1342.
[13] J. Ross, D. Schenkein, B. Greene, G. S. Gray, J. Shulok, P. J. Worland, A. Celniker, M. Rolfe, *Expert Rev. Anticancer Ther.* **2003**, *3*, 107 – 121.
[14] M. C. Peitsch, M. S. Boguski, *New Biol.* **1990**, *2*, 197 – 206.
[15] E. Rassart, A. Bedirian, S. Do Carmo, O. Guinard, J. Sirois, L. Terrisse, R. Milne, *Biochim. Biophys. Acta* **2000**, *1482*, 185 – 198.
[16] M. Vogt, A. Skerra, *J. Mol. Recognit.* **2001**, *14*, 79 – 86.
[17] R. E. Dickerson, I. Geis, *Hemoglobin: Structure, Function, Evolution, and Pathology*, Menlo Park, California, USA, **1983**.
[18] W. O. Bullock, J. M. Fernandez, J. M. Short, *BioTechniques* **1987**, *5*, 376 – 378.
[19] B. K. Kay, J. Winter, J. McCafferty, *Phage Display of Peptides and Proteins: A Laboratory Manual*, Academic Press, San Diego, California, USA, **1996**.
[20] T. König, A. Skerra, *J. Immunol. Methods* **1998**, *218*, 73 – 83.
[21] A. Skerra, T. G. M. Schmidt, *Methods Enzymol.* **2000**, *326*, 271 – 304.
[22] A. Nissim, H. R. Hoogenboom, I. M. Tomlinson, G. Flynn, C. Midgley, D. Lane, G. Winter, *EMBO J.* **1994**, *13*, 692 – 698.
[23] J. D. Marks, H. R. Hoogenboom, T. P. Bonnert, J. McCafferty, A. D. Griffiths, G. Winter, *J. Mol. Biol.* **1991**, *222*, 581 – 597.
[24] K. Nord, E. Gunneriusson, J. Ringdahl, S. Ståhl, M. Uhlén, P.-Å. Nygren, *Nat. Biotechnol.* **1997**, *15*, 772 – 777.
[25] T. J. Gibson, *Studies on the Eppstein-Barr Virus Genome*, PhD thesis, Cambridge University, England, **1984**.
[26] C. Yanisch-Perron, J. Vieira, J. Messing, *Gene* **1985**, *33*, 103 – 119.
[27] A. Skerra, *Gene* **1994**, *151*, 131 – 135.
[28] J. Geisselsoder, F. Witney, P. Yuckenberg, *BioTechniques*, **1987**, *5*, 786 – 791.

[29] T. A. Kunkel, J. D. Roberts, R. A. Zakour, *Methods Enzymol.* **1987**, *154*, 367 – 382.
[30] W. L. Tung, K. C. Chow, *Trends Genet.* **1995**, *11*, 128 – 129.
[31] P. N. Hengen, *Trends Biochem. Sci.* **1996**, *21*, 75 – 76.
[32] J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA, **1989**.
[33] S. P. Fling, D. S. Gregerson, *Anal. Biochem.* **1986**, *155*, 83 – 88.
[34] S. C. Gill, P. H. von Hippel, *Anal. Biochem.* **1989**, *182*, 319 – 326.

Received: June 30, 2003 [F 703]

