CHEMBIOCHEM

DOI: 10.1002/cbic.200700671

Rational Modification of Ligand-Binding Preference of Avidin by Circular Permutation and Mutagenesis

Juha A. E. Määttä,[a] Tomi T. Airenne,[b] Henri R. Nordlund,[a] Janne Jänis,[c] Tiina A. Paldanius,[a] Pirjo Vainiotalo,[c] Mark S. Johnson,[b] Markku S. Kulomaa,[a] and Vesa P. Hytönen*[a]

Chicken avidin is a key component used in a wide variety of biotechnological applications. Here we present a circularly permuted avidin (cpAvd4→3) that lacks the loop between β-strands 3 and 4. Importantly, the deletion of the loop has a positive effect on the binding of 4′-hydroxyazobenzene-2-carboxylic acid (HABA) to avidin. To increase the HABA affinity of cpAvd4→3 even further, we mutated asparagine 118 on the bottom of the ligand-binding pocket to methionine, which simultaneously caused a significant drop in biotin-binding affinity. The X-ray structure of cpAvd4→3(N118M) allows an understanding of the effect of mutation to biotin-binding, whereas isothermal titration calorimetry revealed that the relative binding affinity of biotin and HABA had changed by over one billion-fold between wild-type avidin and cpAvd4→3(N118M). To demonstrate the versatility of the cpAvd4→3 construct, we have shown that it is possible to link cpAvd4→3 and cpAvd5→4 to form the dual-chain avidin called dcAvd2. These novel avidins might serve as a basis for the further development of self-organising nanoscale avidin building blocks.

Introduction

Avidin is a tetrameric glycoprotein that has four identical subunits, each is composed of eight antiparallel β-strands.[1] Avidin and its prokaryotic analogues streptavidin,[2] bradavidin[3] and rhizavidin[4] bind D-biotin with extremely high affinity ($K_d \approx 10^{-15}$ M for avidin). This strong interaction is widely exploited in so-called avidin–biotin technology.[5,6]

Avidin and streptavidin, have previously been genetically engineered in various ways, and have been recently reviewed by Kulomaa and associates.[7,8] Recently, a circular permutation strategy was used to join the avidin subunits within a single polypeptide chain by forming both dual-chain avidin (dcAvd)[9] and single-chain avidin (scAvd),[10] which respectively display two and four ligand-binding sites, each of which can be independently modified to bind different ligands with varying affinities. A dual-affinity avidin has already demonstrated the feasibility of this concept.[11] Streptavidin has also been subjected to circular permutation[12] and subunit fusion.[13]

When D-biotin is bound to streptavidin, only 18 Å^2 of the ligand surface is accessible to a solvent. In a circularly permuted form of streptavidin in which the loop-connecting β-strands 3 and 4 were deleted, the solvent-accessible surface increased to 57 Å^2, and the carboxylate group of D-biotin became more exposed.[12] In avidin, proteinase K cleaves the analogous loop, and proteinase K-treated avidin has been reported to display enhanced affinity towards the azo-compound, HABA (4′-hydroxyazobenzene-2-carboxylic acid; the structure of HABA is shown in Figure S4 in the Supporting Information).[14]

Streptavidin has a lower affinity towards HABA ($K_d = 1.4 \times 10^{-4}$ M)[15] than avidin ($K_d = 6.0 \times 10^{-6}$ M);[16,17] this difference can at least partially be explained by the fact that the L3,4 loop, which is known to be important for ligand binding is shorter in streptavidin than in avidin.[18,19] Accordingly, it has been observed that AVR4, an avidin-related protein,[20,21] also exhibits lower affinity for HABA compared to avidin, most probably because AVR4 has a kinked and therefore shorter L3,4 loop.[22]

In the present study, we constructed a novel circularly permuted form of avidin, cpAvd4→3, to alter the ligand-binding preferences of avidin and to increase the potential of avidin to bind ligands that are larger than biotin. A point mutation, N118M (numbering according to wt avidin) was introduced into cpAvd4→3 to increase the affinity for HABA and to reduce that for biotin. Further, cpAvd4→3 was fused to the C terminus of cpAvd5→4, another circularly permuted form of avidin,[9] to create a novel fusion of avidin subunits; this was designated dual-chain avidin 2 (dcAvd2) and it possess advanced structural and functional characteristics.

Avidin and streptavidin are examples of widely used protein biotools. They are utilised in technologies that range from the biosciences and medicine to material science.[6,23–25] The extreme affinity for the small biotin ligand and the high stability of these proteins has encouraged scientists to develop them

[a] *J. A. E. Määttä, Dr. H. R. Nordlund,[+] T. A. Paldanius, Prof. Dr. M. S. Kulomaa, Dr. V. P. Hytönen*
Institute of Medical Technology, University of Tampere
33014 University of Tampere (Finland)
Fax: (+358) 3-3551-7710
E-mail: vesa.hytonen@uta.fi

[b] *Dr. T. T. Airenne, Prof. Dr. M. S. Johnson*
Department of Biochemistry and Pharmacy, Åbo Akademi University
20520 Turku (Finland)

[c] *Dr. J. Jänis, Prof. Dr. P. Vainiotalo*
Department of Chemistry, University of Joensuu
80101 Joensuu (Finland)

[+] *Present address: NEXT Biomed Technologies NBT Oy*
00790 Helsinki (Finland)

Supporting information for this article is available on the WWW under http://www.chembiochem.org or from the author.

further to meet the growing needs of bio- and nanotechnology.

Results and Discussion

Novel circularly permuted avidins

CpAvd4→3 and dcAvd2 were successfully purified by 2-iminobiotin, and cpAvd4→3(N118M) was purified by biotin affinity chromatography. The final products were highly pure and had the correct size according to SDS-PAGE analysis (Figure 1).

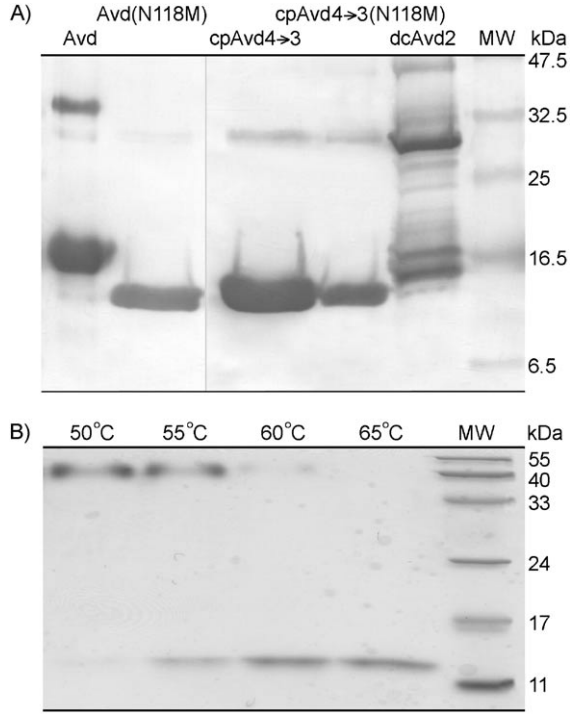

Figure 1. Western blot and SDS-PAGE-based thermal stability assay. A) Western blot analysis of the purified proteins, avidin, avidin(N118M) mutant, cpAvd4→3, cpAvd4→3(N118M), dcAvd2 and a molecular weight marker. B) Thermal stability assay of cpAvd4→3 protein at temperatures of 50, 55, 60, 65 °C and a molecular-weight marker. The loss of the quaternary structure of cpAvd4→3 was observed at 55 °C.

However, in the case of dcAvd2, a significant amount of a ~20 kDa form was observed in addition to the protein of the expected size (~30 kDa). The smaller polypeptide chain detected in the dcAvd2 samples was likely due to a proteolytic cleavage. A typical yield after purification was ~5 mg of protein per litre of bacterial culture. Gel filtration chromatography showed the homogeneity of the purified proteins; all were detected as single peaks in the analysis (not shown).

The measured molecular masses of the protein forms (cpAvd4→3, cpAvd4→3(N118M) and dcAvd2) produced in *E. coli* were 47–56 kDa, which are lower than the theoretical mass values of the corresponding tetramers (57–58 kDa; Table 1). However, this is a typical behaviour of avidin, and was also observed in our previous studies.[11]

Table 1. Oligomeric size of the proteins.[a]

Protein	Elution time[b]		M_W (kDa)	
	−BTN	+BTN	−BTN	+BTN
avidin	28.23	27.97	52.9	56.1
cpAvd4→3	29.03	28.78	44.3	46.9
cpAvd4→3(N118M)	28.69	28.70	47.8	47.7
dcAvd2	27.91	27.95	56.8	56.3

[a] The apparent molecular weights (MW) obtained by gel filtration analysis are indicated. [b] Elution time in minutes.

Mass spectrometry

ESI FT-ICR mass spectrometry was used to confirm the amino acid sequence and proper folding of the proteins (mass spectra data not presented). For some unknown reason, dcAvd2 was the only protein that could not be identified by this method. The most abundant isotopic masses were determined to be 14685.52 ± 0.07 Da for Avd(N118M), 14285.24 ± 0.01 Da for cpAvd4→3 and 14302.18 ± 0.03 Da for cpAvd4→3(N118M). The determined mass of Avd(N118M) was as expected (calcd 14685.50 Da), but the masses of both cpAvd4→3 and cpAvd4→3(N118M) were about 14 Da higher than the expected calculated values of 14271.18 and 14288.17 Da, respectively. This result would suggest that protein methylation (theor. + 14.016 Da) had occurred.

When the disulfide bridges were reduced, the most abundant isotopic masses of the proteins increased by two Daltons as expected; this indicates proper folding. This further supports the hypothesis that the mass shift of ~14 Da could be a result of methylation. Post-translational protein methylation is a common amino acid modification known to occur at the side chains of many amino acids.[26] We have further identified and localised post-translational methylation in cpAvd4→3(N118M) by using on-line digestion combined with ESI FT-ICR tandem mass spectrometry (Jänis, J. et al., unpublished results). A single methylation site in one of the serine residues in cpAvd4→3(N118M) was unambiguously observed; this revealed a previously unknown protein modification.

Deletion of critical residues in the L3,4 loop by circular permutation reduces biotin affinity

The biotin-binding properties of the engineered avidins were studied by measuring the dissociation of fluorescent biotin conjugate BF-560-biotin. At 25 °C, cpAvd4→3 showed an eight-fold higher dissociation rate compared to wt avidin. In comparison, dcAvd2, which has dissimilar subunit pairs showed a three-fold-increased dissociation rate compared to avidin.

Half of the binding sites were assumed to behave as in wt avidin (in dcAvd2 the circularly permuted subunit cpAvd5→4 should, according to our knowledge, have wt-like biotin-binding properties),[9] thus a four-fold-increased dissociation rate was attributed to the cpAvd4→3 subunit of dcAvd2.

At 50 °C, cpAvd4→3 showed the highest dissociation rate for biotin, evincing about 20-fold weaker binding compared to

avidin. DcAvd2 had an about five-fold faster dissociation rate compared to avidin. When the data were analysed as described above, after subtracting the contribution due to the wt-like subunit, the dissociation rate of biotin from the cpAvd4→3 subunit of dcAvd2 was found to be about ten times faster than avidin (Table 2).

Table 2. Dissociation analysis of fluorescent biotin.

Protein	k_{diss} [10^{-5} s^{-1} M^{-1}] at 25 °C	k_{diss} [10^{-5} s^{-1} M^{-1}] at 50 °C	Release [%] at 25 °C	Release [%] at 50 °C
avidin	2.3[a]	27.4[b]	14.1[a]	71.5
avidin(N118M)	>300[e]			
cpAvd4→3	18.7	449.9[c]	46.8	
cpAvd4→3(N118M)	>300[e]			
dcAvd2	5.3/*7.81*[d]	135.3[c]/*200*[d]	17.7	71.0
dcAvd	1.5	41	6.0	77.9

[a] Ref. [27]. [b] Ref. [28]. [c] Determined during the first 500 s of the measurement. [d] Half of the binding sites were assumed to behave like wt avidin. The fitting was performed only for the first 500 s. [e] The dissociation rate was already too fast to be measured accurately at 10 °C.

The dissociation rate of *d*-[8,9-^{3}H]biotin from the protein samples was measured at different temperatures (Figure 2). Avidin had the slowest dissociation rate of *d*-[8,9-^{3}H]biotin over

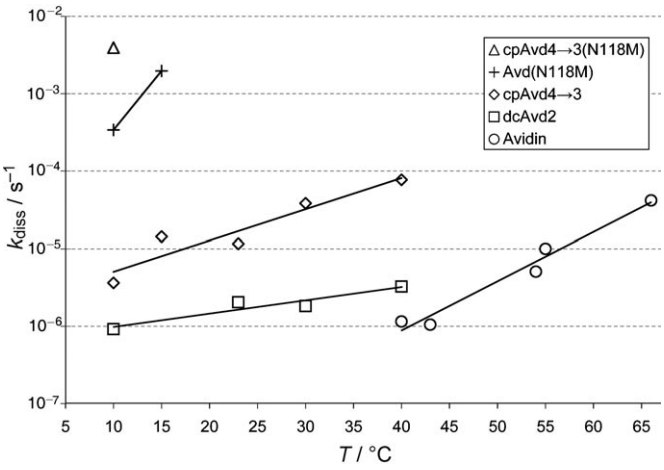

Figure 2. Radiobiotin assay. Dissociation rate constants that were determined by using *d*-[8,9-^{3}H]biotin at various temperatures are shown.

the entire temperature range analysed, and as assumed, dcAvd2 had a slower dissociation rate for biotin than cpAvd4→3.

The affinity of avidin for biotin was too tight to be measured by isothermal titration calorimetry (ITC), but the binding constant (K_b) and dissociation constant ($K_d = 1/K_b$) could be solved for the modified avidins (Figure 3). If the wild-type binding site (cpAvd5→4) of dcAvd2 is assumed to bind biotin with femtomolar (10^{-15} M; wt-like) affinity, the biotin affinity of the other circularly permuted binding site (cpAvd4→3) can be directly determined from the data. The cpAvd4→3 subunit in fusion with cpAvd5→4 of dcAvd2 seemed to have a slightly higher affinity for biotin than the tetrameric cpAvd4→3. Moreover, the N118M point mutation reduces the biotin-binding affinity of the cpAvd4→3 protein by about five-fold (Table 3).

Novel circularly permuted avidin has altered thermodynamics and increased affinity for HABA

Our ITC experiment revealed that avidin binds HABA with moderate affinity ($K_d = 7.9\times10^{-6}$ M), which is in good agreement with the previously published result of $K_d = 6\times10^{-6}$ M (Green 1975), whereas its bacterial analogue, streptavidin has a weaker HABA-binding affinity ($K_d = 1.4\times10^{-4}$ M).[15] The different HABA-binding affinities of avidin and streptavidin are thought to result from the varying architecture of the L3,4 loop, that is, its conformation and primary structure, as was also noted in the case of AVR4.[22] Our working hypothesis was that by deleting the L3,4 loop from wt avidin by using a circular permutation strategy, the HABA-binding affinity of avidin could be increased. This was tested by measuring the HABA-binding affinities of cpAvd4→3 and wt avidin by using ITC. It emerged that the cpAvd4→3 construct had indeed stronger affinity to HABA ($K_d = 5.4\times10^{-6}$ M) than wt avidin ($K_d = 7.9\times10^{-6}$ M).

The reaction between avidin and biotin is strongly exothermic ($\Delta H = -23.0 \pm 0.2$ kcal mol^{-1}). In contrast, that between HABA and avidin is endothermic ($\Delta H = 1.9 \pm 0.03$ kcal mol^{-1}), and thus the moderate binding affinity is entropy driven. Our ITC analysis showed the interaction of HABA with Avd(N118M), cpAvd4→3 or cpAvd4→3(N118M) to be exothermic. The most markedly exothermic reaction with HABA was measured for cpAvd4→3(N118M), which has about six-fold more favourable enthalpy than Avd(N118M) and one and a half-fold more favourable enthalpy than cpAvd4→3 (Table 3). The calculated entropy term was found to be beneficial for binding in every HABA-binding reaction that was analysed here.

A problem arises in calorimetric measurements in that the wt-biotin-binding site of dcAvd2, that is, the cpAvd5→4 subunit, binds HABA in an endothermic manner, whereas the cpAvd4→3 subunit binds HABA in an exothermic manner. Since it is possible to see only the sum of the energies, the net sum of the observed heat of the binding reaction for dcAvd2 was almost zero (Figure 4) and therefore the affinity of the two different kinds of binding sites of dcAvd2 to HABA could not be determined.

Mutation N118M increases HABA binding and reduces the biotin-binding affinity of avidin

The mutation N118M was designed based on the known X-ray structures of the avidin–biotin (PDB code: 2AVI)[1] and avidin–HABA complexes.[19] In the structure of the avidin–biotin complex, the side chain of Asn118 is hydrogen bonded to biotin (Figure 6A), whereas in the avidin–HABA structure there is no significant interaction between the side chain of Asn118 of the avidin and HABA (Figure 6B). Therefore, a mutation at Asn118 which eliminates the hydrogen bond should lead to a decrease in the biotin-binding affinity of avidin. Methionine was chosen to replace asparagine, because its long and flexible side chain

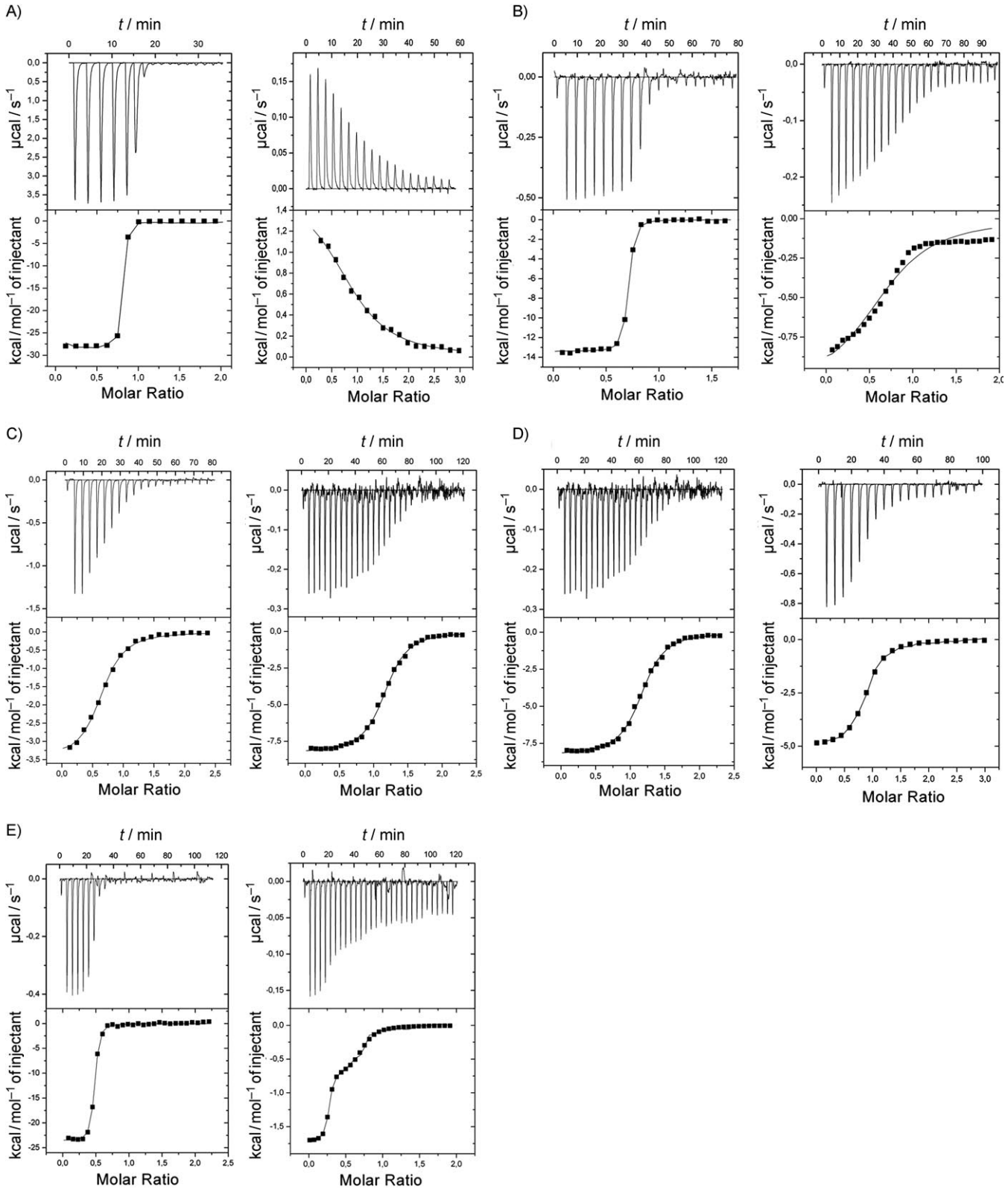

Figure 3. ITC diagrams. A) Avidin; B) Avd(N118M); C) cpAvd4→3; D) cpAvd4→3(N118M); E) dcAvd2. Biotin titrations are shown on the left, and HABA titrations are shown on the right panel. All measurements were made at 25 °C.

was assumed to constitute steric hindrance for biotin. Also, preliminary experiments between different kinds of mutations that were made by Paldanius et al. (unpublished results) revealed that methionine caused the largest changes in both biotin and HABA binding. Indeed, when using ITC, a significant decrease in the biotin binding affinity of avidin was observed for an Avd(N118M) mutant. As a result of the N118M mutation, the extremely strong femtomolar biotin-dissociation constant

Table 3. Thermodynamic analysis of ligand binding by using ITC.

Protein	ΔH [kcal mol^{-1}]	$-T\Delta S$ [kcal mol^{-1}]	ΔG [kcal mol^{-1}]	K_d [M]
BTN				
avidin	-23.0 ± 0.2	4.4[b]	-18.6[b]	1.3×10^{-15}[a]
avidin(N118M)	-16.4 ± 0.1	5.0	-11.4	$(4.2\pm0.2)\times10^{-9}$
cpAvd4→3	-17.6 ± 0.1	7.2	-10.4	$(1.4\pm0.3)\times10^{-8}$
cpAvd4→3-(N118M)	-8.3 ± 0.1	-0.7	-9.0	$(2.4\pm0.4)\times10^{-7}$
HABA				
Avidin	1.9 ± 0.03	-8.8	-7.0	$(7.9\pm0.1)\times10^{-6}$
avidin(N118M)	-0.8 ± 0.02	-6.4	-7.2	$(5.2\pm0.5)\times10^{-6}$
cpAvd4→3	-3.5 ± 0.06	-3.6	-7.1	$(5.4\pm0.7)\times10^{-6}$
cpAvd4→3-(N118M)	-5.1 ± 0.05	-3.0	-8.1	$(1.0\pm0.2)\times10^{-6}$

[a] Ref. [29]. [b] Calculated from the K_d value.

of wt avidin[29] was reduced to a nanomolar value ($K_d = 4.17\times10^{-9}$ M). Furthermore, the enthalpy of the Avd(N118M)–biotin binding ($\Delta H = -16.4\pm0.1$ kcal mol^{-1}) was only two-thirds that of the Avd–biotin binding ($\Delta H = -23.0\pm0.2$ kcal mol^{-1}; Table 3). The point mutation N118M seemed to have a clear effect on the biotin dissociation rate: Avd(N118M) has a *d*-[8,9-^{3}H]biotin dissociation rate several orders of magnitude higher than that of wt avidin (Figure 2). The N118M mutation was also introduced into the cpAvd4→3 construct (cpAvd4→3(N118M)). It had, in turn, a very fast dissociation rate ($k_{diss}\approx4\times10^{-3}$ s^{-1}) even at a low temperature of 10 °C. The corresponding dissociation rate of avidin would be about 10^{-8} s^{-1} if extrapolated from the measured data (Figure 2). At higher temperatures, the dissociation rate of cpAvd4→3(N118M) for biotin was too fast to be determined accurately (Figure 2).

In contrast to biotin, HABA has been observed to bind to Avd(N118M) slightly more strongly than to wt avidin: analysis by using fluorescence spectrometry (Figure S1) showed an apparent $K_d = 6.3\times10^{-6}$ M for the avidin–HABA complex and a $K_d = 0.6\times10^{-6}$ M for Avd(N118M)–HABA. Our ITC measurements were in agreement with these results; the HABA-binding affinity was almost the same for both Avd(N118M) ($K_d = 5.2\times10^{-6}$ M) and cpAvd4→3 ($K_d = 5.4\times10^{-6}$ M), that is, both have higher affinity for HABA than wt avidin ($K_d = 7.9\times10^{-6}$ M; Table 3).

Similarly to the result of mutation in the case of wt avidin, the N118M mutation in cpAvd4→3 exerted the following effects: 1) it reduced the affinity to biotin by about 16-fold and 2) it increased the affinity to HABA by about fivefold compared to cpAvd4→3. Thus, of all the avidin forms that have been examined to date, the cpAvd4→3(N118M) construct has the highest affinity for HABA ($K_d = 1.0\times10^{-6}$ M; Table 3).

These results were further confirmed by using a spectrometric method.[16] Scatchard analysis of the data, which was measured by UV/Vis spectroscopy showed $K_d = (6.2\pm0.2)$ μM for avidin–HABA, $K_d = (2.1\pm0.6)$ μM for Avd(N118M)–HABA and $K_d = (1.9\pm0.5)$ μM for cpAvd4→3(N118M)–HABA. Representative data are shown in Figure S2.

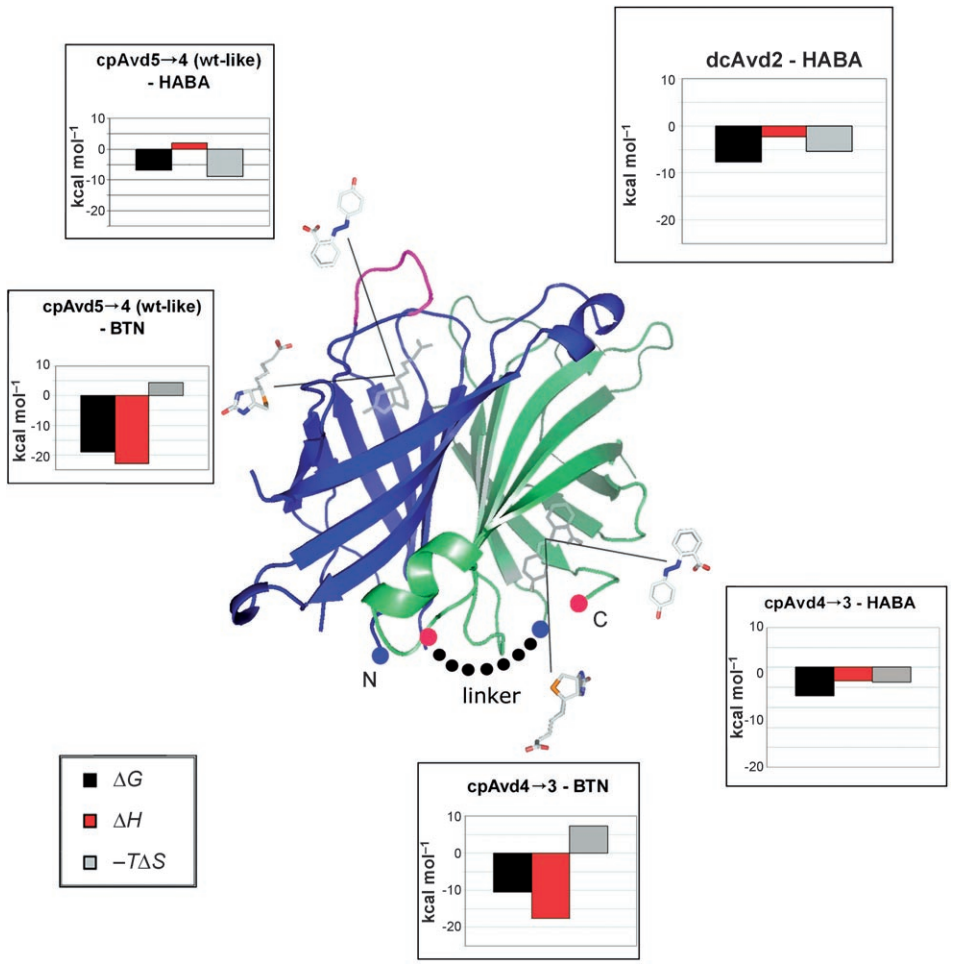

Figure 4. The ligand-binding sites of dcAvd2. A modelled structure of dcAvd2 with two different kinds of binding sites is presented. The cpAvd5→4 (blue) of dcAvd2 has a wt-like binding site. It binds biotin with tight affinity ($K_d\approx10^{-15}$ M) and HABA with moderate affinity ($K_d\approx10^{-6}$ M). The cpAvd4→3 subunit (green) of dcAvd2 lacks the L3,4 loop (pink) and binds both HABA and biotin with moderate affinity. The new N and C termini of cpAvd4→3 and cpAvd5→4 are marked with blue and red dots, respectively. Diagrams represent thermodynamic values that were obtained from ITC measurements—Gibbs free energy (ΔG, black bar), enthalpy (ΔH, red bar) and the entropy term in the measurement temperature ($-T\Delta S$, grey bar). DcAvd2–HABA diagram shows the combined energies of both cpAvd5→4 and cpAvd4→3–HABA binding events. The structural model was prepared by manually deleting the L3,4 loop from two of subunits from the X-ray structure of avidin (PDB code: 2AVI). The linker that connects cpAvd4→3 and cpAvd5→4 in dcAvd2 is shown schematically with black dots.

Overall X-ray structure of cpAvd4→3(N118M)

The crystal structure of cpAvd4→3(N118M) in complex with D-biotin was determined at 1.9 Å resolution. The structure determination statistics are summarised in Table 4. Judging from the 3D structure, the circular permutation strategy had not affected the folding of cpAvd4→3(N118M). The overall structure of cpAvd4→3(N118M) was tetrameric and highly similar to that of chicken avidin[1] and bacterial streptavidin.[18,30] However, in each of the four β-barrel monomers of cpAvd4→3(N118M), the biotin-binding pocket was less deeply buried and much more exposed to solvent in comparison to wt avidin.[1,31] This is due to the lack of residues 39–44 of the L3,4 loop (numbering according to wt avidin[1]) in the cpAvd4→3 construct, and the artificial locations of the N and C termini (Figure 5). In fact, the biotin-binding pocket of cpAvd4→3(N118M) was even more open than the biotin-binding site of streptavidin.[30] The average solvent-accessible surface area of the bound biotin molecules in the avidin–biotin (PDB code: 2AVI), streptavidin–biotin (PDB code: 1MK5) and cpAvd4→3(N118M)–biotin complexes were calculated to be 7, 18 and 22 Å^2, respectively.

The Cα atoms of the four subunits of cpAvd4→3(N118M) were superimposed (subunits B, C and D

Table 4. Data collection and structure determination statistics for cpAvd4→3(N118M).

Data Collection[a]	
wavelength [Å]	0.900
beamline	ID29 (ESRF)
detector	ADSC
resolution [Å]	25–1.9 (2.0–1.9)
unique observations	36425 (5172)
I/σ	12.3 (3.2)
R_{factor} [%][b]	7.8 (48.2)
completeness	99.8 (99.9)
redundancy	4.9 (5.0)
Refinement	
space group	$P2_1$
unit cell:	
a, *b*, *c* [Å]	41.8, 79.3, 71.7
α, β, γ [°]	90, 98.7, 90
monomers (asymmetric unit)	4
resolution [Å]	25–1.9
R_{work} [%][c]	20.3
R_{free} [%][c]	25.1
protein atoms	3556
heterogen atoms	64
solvent atoms	172
rmsd:	
bond lengths [Å]	0.013
bond angles [°]	1.6

[a] The numbers in parenthesis refer to the highest resolution bin. [b] Observed *R* factor from XDS.[33] [c] From Refmac 5 (TLS and restrain refinement).[34]

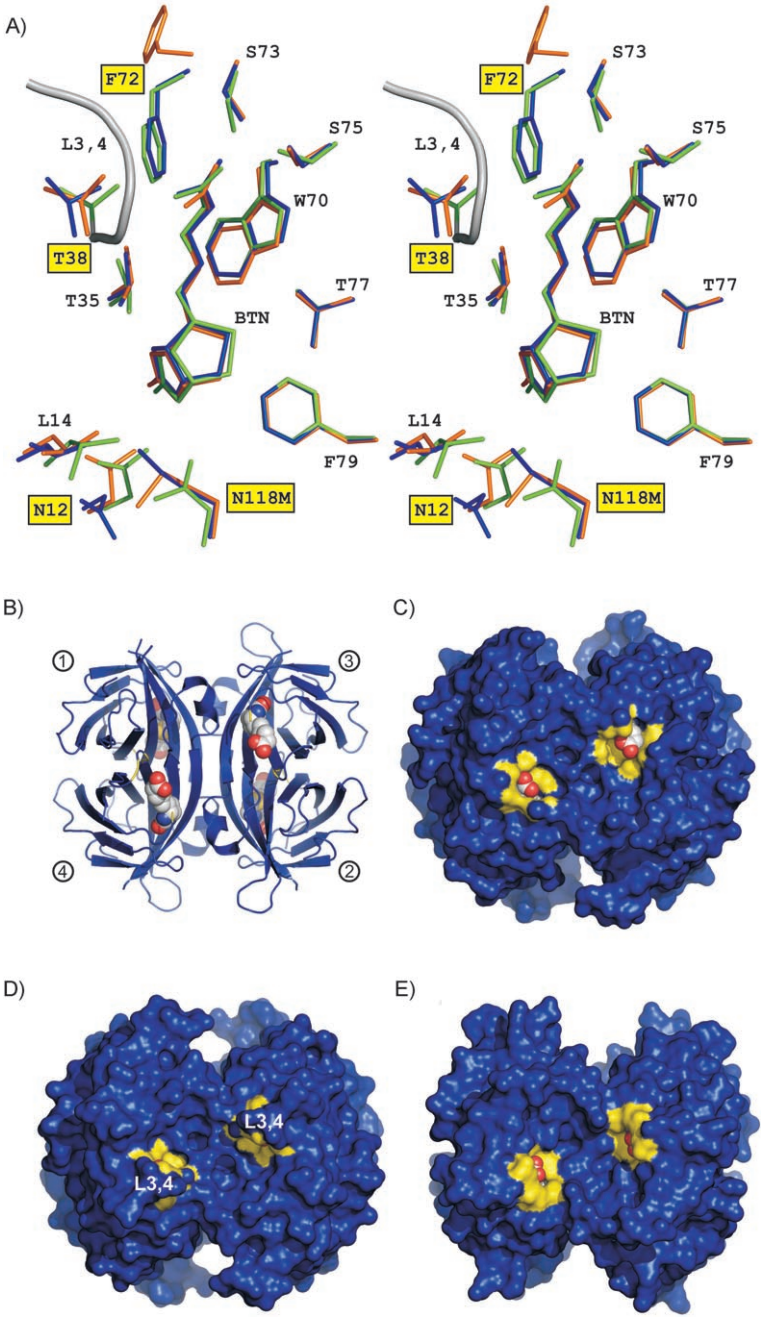

Figure 5. Structural characteristics of cpAvd4→3(N118M). A) Stereo view of the biotin-binding site of cpAvd4→3(N118M). Selected residues at the biotin-binding site of cpAvd4→3(N118M) (chain C, orange; chain D, blue) and avidin–biotin complex structure (PDB code 2AVI; chain A; green) are shown as stick models. Only the side-chain atoms are shown. The amino acids are numbered according to ref. [1]. Bound biotin (BTN) ligands are shown as sticks. The labels for residues N12 and F72, which have clearly different conformations in chains C and D of cpAvd4→3(N118M), as well as the N118M mutation and the C-terminal residue (T38) of cpAvd4→3 are boxed and have a yellow background. A part of the L3,4-loop of avidin that extends from T38 and is not present in cpAvd4→3(N118M) is drawn as a grey ribbon. B) structure of cpAvd4→3(N118M) that shows the secondary structural elements (ribbon model) and connecting loops. The subunits are numbered according to ref. [1]. C) Molecular surface of cpAvd4→3(N118M); D) avidin–biotin complex (PDB code 2AVI); and E) streptavidin–biotin complex (PDB code: 1MK5). The surfaces are coloured blue except for areas that correspond to atoms 6 Å from the C10 atom (named according to ref. [32]) of biotin. These atoms are coloured yellow to pinpoint the entrances to the biotin-binding sites. The bound biotin molecules are shown as spheres (oxygen atoms, red; carbon atoms, white; nitrogen atoms, blue; sulfur atoms, yellow).

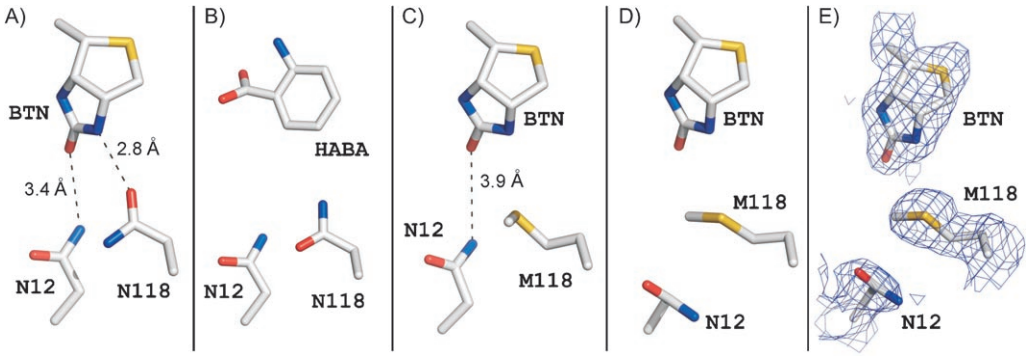

Figure 6. Mutation N118M. The putative hydrogen bonding of amino acid residues N12 and N118 with biotin (BTN; in A, C–E) or HABA (in B) are illustrated with dotted lines. A) avidin–biotin complex (chain A; PDB code 2AVI); B) avidin–HABA complex;[19] C) subunit B of cpAvd4→3(N118M); and D) subunit D of cpAvd4→3(N118M). Only the side-chain atoms of residues N12 and N118, or a portion of the BTN and HABA structures near these residues, are shown as sticks (C, grey; N, blue; O, red; S, yellow). E) $2F_o-F_c$ electron density map (blue) that was calculated in the absence of biotin and contoured at 1σ around the residues shown in D is drawn.

on A) with an rmsd (root mean squared deviation) of 0.3 Å. Only slight variations were detected in the fine architecture of the different chains, mainly within the loop regions, and these were probably caused by crystal packing effects. Out of the 130 residues that comprise cpAvd4→3(N118M), residues 1–3 (chains A, C, D), 45–47 (A), 81 (A, C) and 82–95 (A–D) could not be included in the final X-ray model due to the lack of interpretable electron density around these residues (high thermal motion). Residues 82–95 correspond to a linker region, which was introduced to connect the N and C termini of wt avidin in forming the cpAvd4→3(N118M) construct. In addition, the subunit interfaces seen in cpAvd4→3(N118M) were highly similar to those present in wt avidin.

The biotin-binding site

All four subunits of cpAvd4→3(N118M) had biotin bound at the ligand-binding site. However, the electron density around the ligand of subunit C (BTN-C) was clearly weaker (data not shown), and the *B*-factors for the ligand were higher in comparison to biotin of the other subunits; this indicates the higher thermal motion and conformational flexibility of BTN-C. Phe72 at the biotin-binding site of subunit C of cpAvd4→3(N118M) also exhibited a conformation that was different from that seen in the other subunits; this further indicates the lower occupancy of BTN-C at subunit C. Compared to wt avidin (PDB code: 2AVI)[1] and streptavidin (PDB code: 1MK5),[30] the overall biotin-binding mode of each subunit of cpAvd4→3(N118M), as well as the conformations of the bound biotin were well conserved (Figure 5).

The hydrophobic side chain of Met118 of cpAvd4→3(N118M) did, as expected, lead to different interactions with biotin in comparison to wt avidin. The hydrogen bonds seen in the avidin–biotin complex structure[1] between the side chains of Asn12 and Asn118 of the protein, and the oxygen (O2′) and the nitrogen (N3′) atom of the ureido ring of biotin, respectively (naming according to DeTitta et al.)[32] were substituted either partially or completely by weak (multi)polar and van der Waals interactions in the cpAvd4→3(N118M) structure (Figure 6). In the ligand-binding sites of subunits A–C of cpAvd4→3(N118M) (see, for example, subunit B in Figure 6C), a weak hydrogen bond between the O2′ atom of biotin and the side-chain nitrogen atom of Asn12 might exist as in the wt avidin structure (Figure 6A), but in subunit D the hydrogen bond was missing (Figure 6D). Moreover, the hydrophobic side chain of Met118 could not form a strong hydrogen bond with the N3′ atom of biotin in any of the subunits of cpAvd4→3(N118M). This interaction is known to be important for biotin binding in wt avidin, where the Asn118 side chain is hydrogen-bonded to N3′.[1] In addition, the lack of the L3,4 loop residues Ala39 and Thr40 of wt avidin at the biotin-binding site of cpAvd4→3(N118M) made the site clearly more accessible to solvent (see above; Figure 5C–E), while the biotin-binding mode was nonetheless very similar to that in both streptavidin and wt avidin (Figure 5A). Some rearrangement of water molecules due to the deletion of the L3,4 loop and the artificial C terminus was, however, detectable in the X-ray structure of cpAvd4→3(N118M). For example, in chains C and D only, a water molecule was seen between the C-terminal threonine residue and the carboxylic acid group of biotin in the cpAvd4→3(N118M) structure.

Thermal stability of the novel proteins

The thermostability of cpAvd4→3, cpAvd4→3(N118M) and dcAvd2 (Table 5) was analysed by SDS-PAGE.[35] Tetramers of

Table 5. Thermal stability of the proteins.[a]

Protein	T_r [°C] −BTN	T_r [°C] +BTN
avidin	60	95
cpAvd4→3	55	75
cpAvd4→3(N118M)	50	75
dcAvd2	55	80

[a] Transition midpoint temperatures (T_r) of the oligomeric disassembly measured by an SDS-PAGE-based method are shown in the presence (+BTN) and absence (−BTN) of biotin.

cpAvd4→3 without biotin were found to breakdown into subunits at 55 °C, and in the presence of biotin, at 75 °C. Similarly, cpAvd4→3(N118M) lost its quaternary structure at 50 °C without biotin, and at 75 °C with biotin. dcAvd2, in turn, was found to breakdown into subunits at less than 60 °C, and with biotin the breakdown temperature was found to be between 75 and 80 °C. By comparison, wt avidin breaks down in the presence and absence of biotin at 95 and 60 °C, respectively (Table 5).

Conclusions

Since the cloning of the genes, numerous modified (strept)avidins have been reported in which the pI, glycosylation, stability and the ligand binding of these proteins has been altered, as has been recently reviewed by Kulomaa and associates.[7] However, two relatively recent innovations have had a major impact on the engineering of (strept)avidin beyond the original products. First, the creation of circularly permuted forms has made it possible to create mutant proteins that possess new N and C termini at artificial locations. A circular permutation was originally observed in nature, and has been reviewed by Grishin,[36] but was also applied to streptavidin[12] and avidin.[9] Second, the introduction of the dual-chain methodology[9] that is based on circular permutation has made it possible to produce rational designs, and to generate chimeric avidins and, more recently, streptavidins.[13]

In this study, our working hypothesis was that removal of the L3,4 loop that is located between β-strands 3 and 4 of wt avidin would alter its ligand-binding preferences, that is, it would reduce the affinity for biotin and increase the affinity for HABA. This hypothesis was made based on results from Ellison and colleagues[14] and comparisons of several crystallographic structures of (strept)avidin. Furthermore, we sought to apply the recently introduced dual-chain technology to generate chimeric avidins that would have a wt-like ligand-binding site and a binding site that lacks the L3,4 loop.

Our topologically novel cpAvd4→3 construct was created as follows: 1) six amino acids, residues 39–44 of the original L3,4 loop of wt avidin, were deleted; 2) the original N and C termini of wt avidin were connected with a Gly-Gly-Ser-Gly-Gly-Ser (GGSGGS) linker and 3) novel N and C termini were introduced—the translation of cpAvd4→3 starts before the β4-strand and stops after the β3-strand. The protein was successfully produced in *E. coli*, and spectroscopic analyses showed that the affinity for HABA was increased in comparison to the wt protein, albeit only a little. In order to improve the HABA-binding properties of cpAvd4→3, we introduced, based on the X-ray structures of the avidin–biotin and avidin–HABA complexes, a point mutation at the bottom of the ligand-binding site. This mutation changed the polar side chain of asparagine (Asn118) to the hydrophobic side chain of methionine, thus it directly modulates the physicochemical properties of the binding site. In the X-ray structure of cpAvd4→3(N118M) (Figure 6), it can be seen that, depending on the conformation of Met118, either one or two hydrogen bonds are eliminated. This is in good agreement with our thermodynamic measurements (see below).

Loss of the hydrogen bond between the ureido nitrogen and streptavidin reduces the biotin-binding free energy by approximately 4.3 kcal mol^{-1}.[30] Although the free energy of a hydrogen bond can vary over −0.2 to −40 kcal mol^{-1}, as calculated and reviewed by Stainer,[37] the hydrogen bonds in avidin–biotin binding are thought to represent strong bonds with a free energy of about 4 kcal mol^{-1}. Due to the lack of the amino acid residues Ala39 and Thr40, which form hydrogen bonds with the carboxylic tail of biotin,[38] cpAvd4→3 has two hydrogen bonds less than wt avidin. The observed decrease in free energy ($\Delta\Delta G = 8.2$ kcal mol^{-1}) can be satisfactorily explained by the loss of these hydrogen bonds. Our thermodynamic measurements revealed that the difference in Gibb's free energy ($\Delta\Delta G$) between wt avidin and Avd(N118M) was 7.2 kcal mol^{-1}; this indicates that the N118M mutation affected more than one hydrogen bond, or alternatively, the bond is co-operative as in streptavidin.[30] Indeed, the bond between biotin and the Asn12 side chain was found to be disturbed due to the mutation of N118M (Figure 6). These results could suggest that cpAvd4→3(N118M) has significantly lower affinity for biotin in comparison to either Avd(N118M) or cpAvd4→3. The ITC experiments revealed that the biotin binding of cpAvd4→3(N118M) was weakened by 21% and 13% in terms of ΔG when compared to Avd(N118M) and cpAvd4→3, respectively. The ITC experiments also revealed that cpAvd4→3(N118M) had increased affinity for HABA in comparison to wt avidin, Avd(N118M) or cpAvd4→3. Although we were not able to convert the ligand-binding preference of avidin to favour HABA over its natural ligand, biotin, the relative binding affinity was modified over one billion-fold. The relative dissociation constant $K_d(\text{BTN})/K_d(\text{HABA}) = 1.64 \times 10^{-10}$ for wt avidin was changed to $K_d(\text{BTN})/K_d(\text{HABA}) = 0.23$ for cpAvd4→3(N118M).

Thermodynamic ITC measurements showed that the endothermic binding of HABA to wt avidin was changed dramatically to an exothermic binding reaction in the case of cpAvd4→3 and cpAvd4→3(N118M). However, the net entropy term ($T\Delta S$) for both decreased at the same time, which is thermodynamically the reason why even higher affinities were not observed. It has been reported that the binding of HABA to avidin leads to the displacement of five water molecules from the binding site.[19] This displacement is favourable in terms of entropy, and although the binding of HABA is stabilised by hydrogen bonds, binding does not take place without energy contributions from the surrounding solution; this can be detected as a positive net enthalpy in ITC measurements.

These previous reasons might also clarify the observed, enhanced HABA binding and the reduced biotin-binding properties of cpAvd4→3. The avidin–biotin complex, which is one of the tightest noncovalent ligand–protein complexes, is known to be well optimised; it has a large number of hydrogen bonds and van der Waals interactions.[1] Because cpAvd4→3 lacks the L3,4 loop, two hydrogen bonds that stabilise the interactions between wt avidin and the valeryl moiety of biotin are lost, this at least partially explains the decreased affinity of cpAvd4→3 for biotin (Figure 5, Figure S3). On the other hand, the deletion of the L3,4 loop "opens the lid" of the ligand-binding pocket and leaves the pocket more exposed to sol-

vent. Interestingly, deletion of the L3,4 loop or the N118M mutation alone does not cause a clear change in the entropy term when compared to wt avidin, whereas the point mutation N118M and loop deletion together affect the entropy term such that it favours biotin binding. Favourable entropy is usually observed when water molecules are liberated upon burial of a hydrophobic surface.[39] In the case of the avidin–biotin complex, ligand binding stabilises the structure of avidin and reduces the number of internal degrees of freedom in the system that lead to the unfavourable contribution of the free energy (negative ΔS).[40] In the case of cpAvd4→3(N118M), the entropy term of the biotin binding changed to a positive sign, which favours binding; this is totally opposite to any other cases of avidins or streptavidins.[30] The effect of the positive entropy on binding at room temperature is over 5 kcal mol^{-1} when compared to biotin binding of wt avidin and cpAvd4→3(N118M).

We successfully created a novel subunit fusion form of avidin, dcAvd2, by fusing cpAvd4→3 to the C terminus of cpAvd5→4. Judging from our ITC measurements, the fusion of cpAvd4→3 to cpAvd5→4 did not seem to affect the biotin-binding properties of the subunit; cpAvd4→3 of dcAvd2 binds biotin in a manner similar to the "non-fused" form. Importantly, analysis of dcAvd2 by using ITC indicated a HABA-binding enthalpy of −2.3 kcal mol^{-1}, which agrees with the theoretical value of −1.6 kcal mol^{-1} that was obtained by summing the separately determined enthalpies for the wt and cpAvd4→3 proteins (Figure 4).

What is the added value of the "next-generation" avidins presented here? There is a significant demand for advanced tools in biotechnology and nanotechnology, especially as more experimental systems are miniaturised on the nanometre scale. Furthermore, strept(avidin) with novel properties, especially multiple binding sites, which can be specifically and separately targeted for binding, would facilitate the development of more complicated assay systems or materials than those that are now available.

One demand in both biotechnology and nanotechnology is for new binders for new ligands. The biotin-binding site of strept(avidin) is well known and fairly flexible physically; this offers an excellent starting material that can be modified to bind new or existing ligands with good specificity and affinity. Although the avidin forms that are presented here do not yet completely fulfil these requirements with regard to biotin and HABA binding, the conceptual steps that are made here bring us closer to more sophisticated protein engineering for that purpose.

Experimental Section

D-Biotin was obtained from Sigma Chemicals. Chicken avidin, which was a generous gift from Belovo S.A. (Bastogne, Belgium) was used as a control protein in all analyses. Molecular weight markers were obtained from New England BioLabs and Fermentas (Burlington, Canada). The concentration of the avidin solution was calculated from the absorbance at 280 nm by using an extinction coefficient of 24 280 M^{-1} cm^{-1} per binding site.

The nucleotide and amino acid sequences of the studied proteins are available at EMBL (http://www.ebi.ac.uk/embl/) with the following access numbers: cpAvd4→3, AM779410; cpAvd4→3(N118M), AM779411 (named in the database cp43N76M); dcAvd2, AM779412; and dcAvd, AJ616762.

Design and mutagenesis of the expression constructs: In order to construct a DNA that encodes for circularly permuted cpAvd4→3, PCR reactions were performed by using oligonucleotides 5.4N2 and 3.4C1 in the first reaction, oligonucleotides 3.4N1 and 5.4C1 in the second (Supporting Information, Table S1) and the cDNA of wt avidin as a template. The resulting DNA fragments were connected together by using another round of PCR, by using primers that overlapped the PCR products from the first and second PCR reactions, together with a DNA sequence that encoded the polypeptide linker that connected the original termini. The resulting DNA (cpAvd4→3) was cloned into a pGemTeasy vector that contained attL recombination sites and a DNA sequence that encoded for the ompA signal peptide.[41] Mutagenesis of cpAvd4→3 was carried out with the QuikChange mutagenesis kit (Stratagene, La Jolla, CA, USA). To create the expression construct for dcAvd2, cpAvd4→3 was digested with the restriction enzymes BamHI and HindIII and subcloned into a pGemTeasy vector that already contained the coding sequence for cpAvd5→4.[9]

Production and purification of the expressed recombinant proteins: To achieve bacterial secretion for cpAvd4→3 and dcAvd2, the signal peptide from the OmpA-protein was used at the N terminus of the expressed proteins as previously described.[41] For expression, the DNA constructs were transferred into a pBVboostFG plasmid[42] by using the Gateway LR-cloning reaction (Invitrogen). *E. coli* BL21-AI cells (Invitrogen) were used for protein production. The expression constructs were validated by using DNA sequencing (ABI PRISM 3100 Genetic Analyzer, Applied Biosystems).

Protein purification was accomplished in a single step by using either 2-iminobiotin- or biotin-Sepharose 4 Fast flow (Affiland) affinity chromatography columns as has been previously described.[43] The protein concentration was determined by a UV/Vis spectrophotometer (ND-1000 NanoDrop Spectrophotometer, Wilmington, DE, USA) by measuring the absorbance at 280 nm and by using an extinction coefficient of 23 615 M^{-1} cm^{-1} for cpAvd4→3 and 48 320 M^{-1} cm^{-1} for dcAvd2.

Isothermal titration calorimetry: High-sensitivity VP-ITC titration calorimetry (Microcal Inc., Northampton, MA) was used to measure heat release or uptake, which is mostly associated with ligand binding, but might also reflect conformational changes. All protein samples were dialysed against a buffer that contained 50 mM sodium phosphate, 100 mM NaCl (pH 7.0) or 0.1 mM potassium acetate (pH 5.0); solid biotin and HABA were dissolved in the same dialysis buffer as protein. In order to eliminate air bubbles and long equilibration periods, prior to VP-ITC analysis, the sample solutions were degassed and preheated near the measuring temperature (25 °C). Biotin (0.1–0.4 mM) or HABA (1.0–1.2 mM) was injected into the calorimeter cell that contained protein samples (0.01–0.1 mM) at 4–5 min intervals. The titration cell was stirred continuously at 440–510 rpm. Both ligand-to-buffer and buffer-to-protein titrations were made in order to detect and correct the heat effect of dilution.

The heat of binding after each injection was calculated by integrating the area of the measured peak by using the ORIGIN software suite (v 7.0 SR4, Originlab Corporation, Northampton, MA, USA) that was tuned to the ITC instrument. All data except those related to dcAvd2 were analysed with the "One Set of Sites" fitting proce-

dure. The "Two Sets of Sites" procedure, which was designed for cases where two different kinds of binding sites exist, was used for dcAvd2. The heat (Q) that evolved from each titration can be represented by Equation (1):

$$\Delta Q = \frac{V_0 \Delta H [M]_t K_a [L]}{1 + K_a [L]} \quad (1)$$

where V_0 is the volume of the calorimeter cell, ΔH is the enthalpy of binding per mole of ligand, $[M]_t$ is the total macromolecule concentration, K_a is the association constant, and [L] is the free ligand concentration. The values of ΔH and K_a can be determined directly from the chromatogram for a titration, whereas the change in Gibb's free energy term ΔG and its entropic component $T\Delta S$ can be calculated from Equations (2) and (3):

$$\Delta G = -RT \ln K_a \quad (2)$$

$$\Delta G = \Delta H - T\Delta S \quad (3)$$

where R is the gas constant and T is the temperature in kelvin.

Biotin dissociation assays: The biotin-binding properties of avidin, cpAvd4→3, cpAvd4→3(N118M) and dcAvd2 were studied by measuring the dissociation of *d*-[8,9-^{3}H]biotin and fluorescent biotin from these proteins. The dissociation of *d*-[8,9-^{3}H]biotin (Amersham) was determined at different temperatures as described by Klumb and colleagues.[44] The dissociation of fluorescent biotin was detected by measuring the reversal of quenching of the biotin-coupled fluorescent probe ArcDia™ BF560 (ArcDia Ltd., Turku, Finland) after addition of a 100-fold molar excess of free biotin; analyses were made in a 50 mM sodium phosphate (pH 7.0) buffer that contained NaCl (650 mM) at two different temperatures (25±1 °C and 50±1 °C), and by using a PerkinElmer LS55 luminometer as previously described.[41]

The fluorescence data were interpreted by using the "single-phase dissociation model", as described elsewhere.[41] The dissociation rate constants (k_{diss}) were determined by using Equation (4):

$$-k_{diss} t = \ln (B/B_o) \quad (4)$$

where B_o is the maximum binding (100%) that was measured (the difference between the fluorescence of the free dye and that of the protein–dye complex) and B is the amount of complex that was measured as a function of time. The first 500 s of measurements were omitted from the analyses to abolish the "fast initial phase" effect, which is characteristic of the avidin–BF560–biotin interaction.[11] The release of the fluorescent biotin was determined after 1 h of measurements. In order to determine the dissociation rate for the cpAvd4→3 subunit of dcAvd2, half of the four biotin-binding sites were hypothesised to behave as in wt avidin, and a ligand dissociation model was fitted to the subtracted data as previously described for other dual-chain avidins.[11]

Size exclusion chromatography: The oligomeric state of the proteins was assayed at +4 °C by using the Superose 6 10/300 GL (Tricorn) gel filtration column and the ÄKTA Purifier 10 Chromatography system as previously described.[27] A buffer that contained 50 mM sodium phosphate buffer (pH 7.0) and 650 mM NaCl was used as the running buffer, and gel filtration was performed with a flow rate of 0.5 mL min^{-1}. The high concentration of NaCl reduced the binding of the basic avidin protein (pI ≈ 10.5)[29] to the chromatography column. The gel filtration MW standards were bovine thyroglobulin, bovine gamma-globulin, chicken ovalbumin, horse myoglobin and vitamin B_{12} (Gel Filtration Standard, Bio-Rad).

SDS-PAGE-based thermostability assay: Protein samples with biotin, HABA or both ligands were acetylated in vitro and the temperature-dependent dissociation of the subunits was monitored. Samples were heated to a given temperature between 25 °C and 80 °C for 20 min. The oligomeric state of the protein was analysed by SDS-PAGE in the presence of 2-mercaptoethanol and the gels were stained with Coomassie blue.[35]

Disulfide bridge reduction: Disulfide reduction was performed by using the protocol of Scigelova and colleagues.[45] Briefly, the protein sample with an excess amount of DL-dithiothreitol (Sigma–Aldrich) was incubated at 70 °C for 5 min, after which the sample tube was placed on ice and diluted with an appropriate solvent for mass spectrometry.

Mass spectrometry: All mass spectrometric experiments were performed on a 4.7-T Bruker APEX-Qe hybrid Fourier transform ion cyclotron resonance (FT-ICR) instrument (Bruker Daltonics, Billerica, MA, USA) that had a mass-selective quadrupole interface. Ions were produced in an external Apollo II dual-ion–funnel electrospray ionisation (ESI) source. For broadband spectra (*m/z* 400–4000), the quadrapole was operated in an RF-only mode, and ions were accumulated in the second hexapole (collision cell) for 50–100 ms. Ions were then transmitted through a high-voltage ion optics region, prior to "Sidekick" trapping inside the Infinity ICR cell. Ions were excited by a conventional "RF-chirp", and detected in a direct broadband mode. A total of eight co-added, 512-kWord time-domain transients were acquired, zero-filled twice, followed by magnitude calculation, fast Fourier transform and external frequency-to-*m/z* calibration with respect to the ions of an ES Tuning Mix (Agilent Technologies, Santa Clara, CA). All data were processed by using the Bruker XMASS 7.0.8 software.

Determination of HABA affinity by fluorescence and absorbance spectroscopy: The HABA-binding affinity of Avd(N118M), cpAvd4→3, cpAvd4→3(N118M) and dcAvd2 was measured by using fluorescence spectroscopy. The protein samples (50 nM) in neutral 50 mM phosphate buffer that contained 650 mM NaCl were titrated with increasing concentrations of 4′-hydroxyazobenzene-2-carboxylic acid (HABA). The samples were excited by using a wavelength of 280 nm, emission was measured at 350 nm, and the decrease in intrinsic fluorescence of avidin as a result of HABA binding was monitored. However, because HABA absorbs light also when not bound to protein, there is an error component that cannot be fully subtracted from the data. Thus only the apparent dissociation constants could be measured by using this method.

The affinity of HABA to different avidin forms was determined by using UV/Vis spectrometer. The assay was based on the colour change of the HABA due to complexation with avidin.[16] The protein samples in neutral 50 mM phosphate buffer that contained 100 mM NaCl were titrated with increasing HABA concentrations. The concentration of the avidin–HABA complex was determined by using extinction coefficient of 35 500 cm^{-1} at 500 nm wavelength. The binding constants were determined by using a Scatchard analysis (Figure S2).

X-ray structure determination: The Classics™ (Nextal Biotechnology, (Qiagen), Valencia, CA, USA) screen, the vapour diffusion method and sitting drops on 96-well plates (Corning Inc.) were used to search for suitable conditions for crystallisation of cpAvd4→3(N118M). By using a 10:1 (*v/v*) ratio of protein and ligand, a protein solution of cpAvd4→3(N118M) (1.7 mg mL^{-1}) that contained sodium acetate (50 mM; pH 4) and NaCl (100 mM) was mixed with a D-biotin (Sigma) solution (1 mg mL^{-1}) that contained 5 mM Tris (pH 8.8) and CHES (8 mM; pH 9.5). After optimisation,

crystals formed at 22 °C under conditions in which 0.6–1.2 μL of the sample solution and 0.5–0.7 μL of a well solution that contained $MgCl_2$ 0.18–0.20 M, Tris–HCl (0.09–0.10 M, pH 8.5) and PEG 4000 (27–30%, *w/v*) were used. For data collection, two "cubic"-looking crystals with dimensions less than 0.2×0.2×0.2 mm were used. The first data set (2.6 Å resolution) was collected from a flash-frozen crystal by using MPD as a cryoprotectant (0.7 μL of 100% MPD was added to the crystallisation drop immediately prior to freezing) at the MAX-lab beam line I911_2 (Lund, Sweden) at 100 K by using a MarCCD detector. The second, higher resolution (1.9 Å) data set was later collected at the ESRF beam line ID-29 (Grenoble, France) by using a 100 K liquid nitrogen stream (Oxford Cryosystems, Devens, MA, USA) and an ADSC detector. Sodium formate was used as cryoprotectant (1 μL of 4 M sodium formate was added to the crystallisation drop immediately prior to freezing).

Both data sets were processed with programs from the XDS program package.[33] The 2.6 Å data set was used to solve the structure of cpAvd4→3(N118M) by molecular replacement with the program Phaser[46] from the CCP4i suite;[47,48] four copies of the polypeptide chain A of a known avidin structure (PDB code 2AVI)[1] were used as search models. After molecular replacement, the initial 2.6 Å X-ray model of cpAvd4→3(N118M) was extended to higher resolution and completed by using the 1.9 Å diffraction data: 1) the initial model was selected as input for automatic model building with ARP/wARP,[49] 2) the structure was refined with Refmac5[34] (TLS and restrained refinement)[50] and 3) modified and rebuilt with Coot.[51] Solvent atoms and other non-protein atoms were added to the model either by an automatic procedure in ARP/wARP,[52] or in Coot, and the final structure was analysed by using the inbuilt tools of Coot. The data collection and structural determination statistics for the 1.9 Å structure of cpAvd4→3(N118M) are summarised in Table 4, and the coordinates and structure factors of it have been deposited in the Protein Data Bank with entry codes 2JGS and 2JGSSF, respectively.

Miscellaneous methods: The solvent-accessible surface areas were calculated with the program Areaimol[53] of the CCP4i suite.[47,48] Figures 5 and 6 were created with the PyMol Molecular Graphics System.[54] CorelDRAW X3 was used to edit Figures 5 and 6. Figures S3 and S4 were generated by using VMD 1.8.6.[55]

Acknowledgements

This work was supported by the ISB (National Graduate School in Informational and Structural Biology), the Academy of Finland, the Sigrid Jusélius Foundation, and the Foundation of Åbo Akademi (Center of Excellence in Cell Stress). We would like to thank the staff at the MAX-lab beam line I911 and ESRF beam line ID-29 for excellent support. We acknowledge support by the European Community - Research Infrastructure Action under the FP6 "Structuring the European Research Area" Programme. We thank Ulla Kiiskinen for excellent technical assistance and Jarkko Valjakka for valuable discussions.

Keywords: avidin · circular permutation · protein engineering · thermodynamics · X-ray structure

[1] O. Livnah, E. A. Bayer, M. Wilchek, J. L. Sussman, *Proc. Natl. Acad. Sci. USA* **1993**, *90297*, 5076–5080.
[2] L. Chaiet, F. J. Wolf, *Arch. Biochem. Biophys.* **1964**, *106*, 1–5.
[3] H. R. Nordlund, V. P. Hytönen, O. H. Laitinen, M. S. Kulomaa, *J. Biol. Chem.* **2005**, *280*, 13250–13255.
[4] S. H. Helppolainen, K. P. Nurminen, J. A. Määttä, K. K. Halling, J. P. Slotte, T. Huhtala, T. Liimatainen, S. Ylä-Herttuala, K. J. Airenne, A. Närvänen, J. Jänis, P. Vainiotalo, J. Valjakka, M. S. Kulomaa, H. R. Nordlund, *Biochem. J.* **2007**, *405*, 397–405.
[5] M. Wilchek, E. A. Bayer, *Methods Enzymol.* **1990**, *18429*, 5–13.
[6] M. Wilchek, E. A. Bayer, *Biomol. Eng.* **1999**, *16*, 1–4.
[7] O. H. Laitinen, V. P. Hytönen, H. R. Nordlund, M. S. Kulomaa, *Cell. Mol. Life Sci.* **2006**, *63*, 2992–3017.
[8] O. H. Laitinen, H. R. Nordlund, V. P. Hytönen, M. S. Kulomaa, *Trends Biotechnol.* **2007**, *25*, 269–277.
[9] H. R. Nordlund, O. H. Laitinen, V. P. Hytönen, S. T. Uotila, E. Porkka, M. S. Kulomaa, *J. Biol. Chem.* **2004**, *279*, 36715–36719.
[10] H. R. Nordlund, V. P. Hytönen, J. Hörhä, J. A. Määttä, D. J. White, K. Halling, E. J. Porkka, J. P. Slotte, O. H. Laitinen, M. S. Kulomaa, *Biochem. J.* **2005**, *392*, 485–491.
[11] V. P. Hytönen, H. R. Nordlund, J. Hörhä, T. K. Nyholm, D. E. Hyre, T. Kulomaa, E. J. Porkka, A. T. Marttila, P. S. Stayton, O. H. Laitinen, M. S. Kulomaa, *Proteins* **2005**, *61*, 597–607.
[12] V. Chu, S. Freitag, I. Le Trong, R. E. Stenkamp, P. S. Stayton, *Protein Sci.* **1998**, *7*, 848–859.
[13] F. M. Aslan, Y. Yu, S. Mohr, C. R. Cantor, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 8507–8512.
[14] D. Ellison, J. Hinton, S. J. Hubbard, R. J. Beynon, *Protein Sci.* **1995**, *4*, 1337–1345.
[15] P. C. Weber, M. W. Pantoliano, D. M. Simons, F. R. Salemme, *J. Am. Chem. Soc.* **1994**, *116*, 2717–2724.
[16] N. M. Green, *Biochem. J.* **1965**, *94*, 23C–24C.
[17] N. M. Green, *Meth. Enzymol.* **1970**, *18*, 418–424.
[18] W. A. Hendrickson, A. Pahler, J. L. Smith, Y. Satow, E. A. Merritt, R. P. Phizackerley, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 2190–2194.
[19] O. Livnah, E. A. Bayer, M. Wilchek, J. L. Sussman, *FEBS Lett.* **1993**, *328165*, 165–168.
[20] R. A. Keinänen, M. J. Wallen, P. A. Kristo, M. O. Laukkanen, T. A. Toimela, M. A. Helenius, M. S. Kulomaa, *Eur. J. Biochem.* **1994**, *220*, 615–621.
[21] V. P. Hytönen, T. K. Nyholm, O. T. Pentikäinen, J. Vaarno, E. J. Porkka, H. R. Nordlund, M. S. Johnson, J. P. Slotte, O. H. Laitinen, M. S. Kulomaa, *J. Biol. Chem.* **2004**, *279*, 9337–9343.
[22] S. Repo, T. A. Paldanius, V. P. Hytönen, T. K. Nyholm, K. K. Halling, J. Huuskonen, O. T. Pentikäinen, K. Rissanen, J. P. Slotte, T. T. Airenne, T. A. Salminen, M. S. Kulomaa, M. S. Johnson, *Chem. Biol.* **2006**, *13*, 1029–1039.
[23] B. J. Hinds, N. Chopra, T. Rantell, R. Andrews, V. Gavalas, L. G. Bachas, *Science* **2004**, *303*, 62–65.
[24] K. Keren, R. S. Berman, E. Buchstab, U. Sivan, E. Braun, *Science* **2003**, *302*, 1380–1382.
[25] G. Paganelli, M. Bartolomei, C. Grana, M. Ferrari, P. Rocca, M. Chinol, *Neurol. Res.* **2006**, *28*, 518–522.
[26] C. T. Walsh, *Posttranslational Modification of Proteins: Expanding Nature's Inventory* (Ed.: J. Murdzek), Roberts, Greenwood Village, **2005** pp. 121–149.
[27] V. P. Hytönen, O. H. Laitinen, A. Grapputo, A. Kettunen, J. Savolainen, N. Kalkkinen, A. T. Marttila, H. R. Nordlund, T. K. Nyholm, G. Paganelli, M. S. Kulomaa, *Biochem. J.* **2003**, *372*, 219–225.
[28] V. P. Hytönen, J. A. Määttä, T. K. Nyholm, O. Livnah, Y. Eisenberg-Domovich, D. Hyre, H. R. Nordlund, J. Hörhä, E. A. Niskanen, T. Paldanius, T. Kulomaa, E. J. Porkka, P. S. Stayton, O. H. Laitinen, M. S. Kulomaa, *J. Biol. Chem.* **2005**, *280*, 10228–10233.
[29] N. M. Green, *Adv. Protein Chem.* **1975**, *295*, 85–133.
[30] D. E. Hyre, I. Le Trong, E. A. Merritt, J. F. Eccleston, N. M. Green, R. E. Stenkamp, P. S. Stayton, *Protein Sci.* **2006**, *15*, 459–467.
[31] L. Pugliese, A. Coda, M. Malcovati, M. Bolognesi, *J. Mol. Biol.* **1993**, *231*, 698–710.
[32] G. T. DeTitta, J. W. Edmonds, W. Stallings, J. Donohue, *J. Am. Chem. Soc.* **1976**, *98*, 1920–1926.
[33] W. J. Kabsch, *Appl. Crystallogr.* **1993**, *26*, 795–800.
[34] G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1997**, *53*, 240–255.
[35] E. A. Bayer, S. Ehrlich-Rogozinski, M. Wilchek, *Electrophoresis* **1996**, *17*, 1319–1324.

[36] N. V. Grishin, *J. Struct. Biol.* **2001**, *134*, 167–185.
[37] T. Steiner, *Angew. Chem.* **2002**, *114*, 50–80; *Angew. Chem. Int. Ed.* **2002**, *41*, 48–76.
[38] Y. Pazy, T. Kulik, E. A. Bayer, M. Wilchek, O. Livnah, *J. Biol. Chem.* **2002**, *277*, 30892–30900.
[39] D. B. Smithrud, T. B. Wyman, F. Diederich, *J. Am. Chem. Soc.* **1991**, *113*, 5420–5426.
[40] D. H. Williams, E. Stephens, D. P. O'Brien, M. Zhou, *Angew. Chem.* **2004**, *116*, 6760–6782; *Angew. Chem. Int. Ed.* **2004**, *43*, 6596–6616.
[41] V. P. Hytönen, O. H. Laitinen, T. T. Airenne, H. Kidron, N. J. Meltola, E. Porkka, J. Hörhä, T. Paldanius, J. A. Määttä, H. R. Nordlund, M. S. Johnson, T. A. Salminen, K. J. Airenne, S. Ylä-Herttuala, M. S. Kulomaa, *Biochem. J.* **2004**, *384*, 385–390.
[42] O. H. Laitinen, K. J. Airenne, V. P. Hytönen, E. Peltomaa, A. J. Mähönen, T. Wirth, M. M. Lind, K. A. Mäkelä, P. I. Toivanen, D. Schenkwein, T. Heikura, H. R. Nordlund, M. S. Kulomaa, S. Ylä-Herttuala, *Nucleic Acids Res.* **2005**, *33*, e42 1–10.
[43] O. H. Laitinen, A. T. Marttila, K. J. Airenne, T. Kulik, O. Livnah, E. A. Bayer, M. Wilchek, M. S. Kulomaa, *J. Biol. Chem.* **2001**, *276*, 8219–8224.
[44] L. A. Klumb, V. Chu, P. S. Stayton, *Biochemistry* **1998**, *37*, 7657–7663.
[45] M. Scigelova, P. S. Green, A. E. Giannokopulos, A. Rodger, D. H. G. Crout, P. J. Derrick, *Eur. J. Mass Spectrom.* **2001**, *7*, 29–34.
[46] A. J. McCoy, R. W. Grosse-Kunstleve, L. C. Storoni, R. J. Read, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2005**, *61*, 458–464.
[47] Collaborative Computational Project, Number 4, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1994**, *50*, 760–763.
[48] E. Potterton, P. Briggs, M. Turkenburg, E. Dodson, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2003**, *59*, 1131–1137.
[49] A. Perrakis, R. Morris, V. S. Lamzin, *Nat. Struct. Biol.* **1999**, *6*, 458–463.
[50] M. D. Winn, M. N. Isupov, G. N. Murshudov, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2001**, *57*, 122–133.
[51] P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 2126–2132.
[52] V. S. Lamzin, K. S. Wilson, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1993**, *49*, 129–147.
[53] B. Lee, F. M. Richards, *J. Mol. Biol.* **1971**, *55*, 379–400.
[54] W. L. DeLano, *Curr. Opin. Struct. Biol.* **2002**, *12*, 14–20.
[55] W. Humphrey, A. Dalke, K. J. Schulten, *J. Mol. Graphics* **1996**, *14*, 33–38.

Received: November 7, 2007
Published online on April 1, 2008

