DOI: 10.1002/cbic.200800745

Disulfide Bond Substitution by Directed Evolution in an Engineered Binding Protein

Antoine Drevelle, Agathe Urvoas, Mériam Ben Hamida-Rebaï, Gérard Van Vooren, Magali Nicaise, Marie Valerio-Lepiniec, Michel Desmadril, Charles H. Robert, and Philippe Minard*[a]

The chromoprotein neocarzinostatin (NCS) has been intensively studied for its antitumour properties. It has recently been redesigned as a potential drug-carrying scaffold. A potential limit of this protein scaffold, especially for intracellular applications, is the presence of disulfide bonds. The objective of this work was to create a disulfide-free NCS-derived scaffold. A generic targeted approach was developed by using directed evolution methods. As a starting point we used a previously engineered NCS variant in which a hapten binding site had been created. A library was then generated in which cysteine Cys88 and Cys93 and neighbouring residues were randomly substituted. Variants that preserved the hapten binding function were selected by phage display and further screened by colony filtration methods. Several sequences with common features emerged from this process. The corresponding proteins were expressed, purified and their biophysical properties characterised. How these selected sequences rescued folding ability and stability of the disulfide-free protein was carefully examined by using calorimetry and the results were interpreted with molecular simulation techniques.

Introduction

Molecular recognition is undoubtedly one of the most important aspects of biological function, and correspondingly underlies much of the development of pharmaceuticals. Until recently, antibodies were the unique avenue leading to the creation of binding molecules with tailored specificity. However, the poor folding and expression properties of most antibody fragments still present an obstacle for many applications, including in the intracellular environment.

Alternatives to antibodies have been explored by applying directed evolution methods to other proteins.[1–4] The general idea is to use combinatorial approaches to create a new binding site on a protein scaffold displaying good expression and folding properties.[5,6] Very efficient binders in terms of affinity and specificity have been obtained, each scaffold being suited to a particular field of application: designed ankyrin repeat proteins (DARPins) for protein–protein interactions,[1,7–9] engineered lipocalin derivatives for binding various low-molecular-weight compounds[10] and now as specific drug carriers targeted to intracellular receptors.[4]

Further, polyvalent scaffolds permitting both protein–protein recognition and small-molecule binding simultaneously would unlock a wide range of applications, including vectors for targeting small-molecule pharmaceuticals to specific tissues. It is with this goal in mind that we have been exploring the neocarzinostatin (NCS) protein. Wild-type NCS is a 113 amino-acid enediyne-binding chromoprotein and is naturally secreted by *Streptomyces*. It is currently used as an antitumour agent and constitutes an advanced model for protein-assisted drug targeting.[11] Besides its clinical interest, NCS is an enticing candidate for a polyvalent scaffold, for several reasons. First, pure recombinant protein can be obtained from *E. coli* by periplasmic expression with a high yield,[12] and by cytoplasmic expression although much less efficiently.[13] Second, NCS can be engineered for the binding of both low-molecular-weight targets and proteins. The well-defined binding crevice for its native chromophore has already been engineered to specifically bind a chosen hapten, testosterone.[14,15] The protein also displays three surface loops that are topologically equivalent to the hypervariable loops of an immunoglobulin domain.[16] This property allowed Nicaise et al.[17] to engineer NCS for protein–protein interaction (antilysozyme). Nevertheless, the presence of two disulfide bonds in wild-type NCS especially limits its routine use in the reducing environment of the cell. Therefore, the aim of the present work was to design a cysteine-free NCS scaffold without altering its favourable expression properties and stability.

Disulfides often considerably complicate protein folding, and introduce a dependence on the redox conditions that might be undesirable. Yet, in small proteins disulfides can provide a means of stabilizing an otherwise labile structure. Finding a way to systematically replace disulfides in a targeted manner is thus a problem of general interest. However, there is no trivial approach to replacing disulfide bonds as their role in protein structure, stability and folding differs from protein to pro-

[a] *Dr. A. Drevelle,[+] Dr. A. Urvoas,[+] M. B. Hamida-Rebaï, G. Van Vooren, Dr. M. Nicaise, Dr. M. Valerio-Lepiniec, Dr. M. Desmadril, Dr. C. H. Robert, Prof. P. Minard*
Institut de Biochimie et Biophysique Moléculaire et Cellulaire
UMR 8619 CNRS, Université Paris Sud 11, 91405 Orsay cedex (France)
Fax: (+33) 1-69-85-37-15
E-mail: philippe.minard@u-psud.fr
[+] *These authors contributed equally to this work.*

tein.[18] In some cases, problems associated with disulfide bonds have been circumvented: several approaches have been used to develop intrabodies by using hyper stable immunoglobulin domains[19] or *Camelidae* immunoglobulin domains,[20] but these examples remain target- or cell specific.[21,22] It has been shown that in some cases disulfide bonds can be removed, provided the loss of folding stability is at least partly compensated by stabilizing amino acid substitutions elsewhere in the structure.[18,23] Systematic cysteine mutagenesis has been suggested to find the best substitutions for the replacement of a disulfide bond in BPTI.[24] However, it is unfeasible to apply such an exhaustive approach to each protein scaffold. The use of directed evolution presents the advantage of exploring an enormous number of diverse sequences, which opens the door to identifying novel solutions for disulfide bond replacement. A directed evolution approach has been reported in a limited number of cases: for an antibody fragment[25,26] and for the gene-3 protein of the filamentous phage, fd.[27] In both cases, consensus solutions that compensated the loss of the disulfide bond were found by using error-prone PCR, and thus positions throughout the protein sequence were randomised. However, it would be desirable to use a more targeted approach to finding disulfide bond replacement solutions.

NCS displays two disulfide bonds that are strictly conserved in the enediyne chromoprotein family:[11] the disulfide bond Cys37–Cys47, which is located in the chromophore binding crevice, and Cys88–Cys93, which is located in a short loop between two β strands (Figure 1).[28] Our previous attempts to

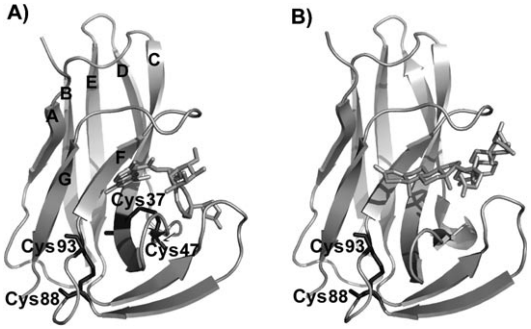

Figure 1. Crystal structures of NCS–ligand complexes. A) NCS-wt complexed with its natural enediyne chromophoric ligand (PDB ID: 1nco).[28] B) Variant NCS-3.24 complexed with two testosterone molecules (PDB ID: 2cbo).[15] Proteins and their ligands are represented in light grey; β strands A to G are indicated. The two disulfide bonds of NCS-wt are represented in black: Cys37–Cys47 in the crevice binding site and Cys88–Cys93 in a constrained loop before the F β strand. The variant NCS-3.24 displays only one disulfide bond (Cys88–Cys93). The figure was prepared with PyMOL (http://www.pymol.org).

remove these disulfide bonds from wild-type NCS (NCS-wt) by standard alanine mutations failed: the corresponding mutants were either unexpressed or poorly expressed by using either a periplasmic or cytoplasmic expression system in *E. coli*. This suggested that simple replacement of either disulfide bond by alanine residues destabilises the NCS structure. We thus decided to use directed evolution to remove the two disulfide bonds of NCS by randomizing residues local to the corresponding cysteines. This led to efficient and exhaustive exploration of the sequence space and revealed several stable cysteine-free NCS variants. A simple generic approach was developed for this purpose to select mutations that are tolerated by the protein architecture and which would not necessarily be predictable by a rational approach.

Results

Design and construction of the library

As a starting point for the directed evolution approach, we used the single-disulfide NCS-3.24 previously described.[14] Residues Cys37 and Cys47, which lie in the binding crevice, were already substituted in this variant into tryptophan and tyrosine, respectively, in the course of a previous directed evolution study[14,15] (Figure 1). Moreover this mutant presents a distinct advantage over wild-type NCS for the selection of cysteine-free NCS variants—its engineered testosterone-binding function. This allows the selection of functional, and thus folded, proteins by phage display. On the contrary, the NCS-wt binding properties would not allow this approach because the natural chromophoric ligand is highly unstable when not bound to the protein and thus unsuitable for selection procedures.[29]

The NCS cysteine-free library (pHDiex-NCS-RdCys) was designed and generated from the variant NCS-3.24. The two remaining cysteine residues (Cys88 and Cys93) are located in a highly constrained six amino acid loop (CTTAAC). As the mutation of the cysteine residues was likely to change the rigidity and/or conformation of the short loop, we decided to randomize the six residues of the entire loop. Among the five known homologues of NCS, Cys88 and Cys93 are strictly conserved and the rest of loop 89–92 presents variability in sequence and length.[11] The six positions were randomised with degenerate codons VNN (V = A, C or G; N = A, C, G or T); this allowed the presence of all amino acids except cysteine and its neighbours TNN in the genetic code (Tyr, Trp, Phe). This mutagenesis scheme theoretically codes for 1.7×10^7 different protein sequences. A library of 10^7 variants was cloned in the pHDiex-0 vector (see the Experimental Section and ref. [14]) by using complementary degenerated primers, which created cohesive ends compatible with restriction sites previously introduced in the vector. Randomly picked clones from this naive library were sequenced. Sixty percent were in the correct reading frame with the expected variability, with no apparent bias in amino acid composition for the randomised residues (data not shown).

Selection and screening by phage display

General and direct selection for folded proteins remains an unsolved issue.[30,31] To bypass this problem, we chose to select the protein variants of our library for their function, that is, testosterone binding, assuming that specific interaction of NCS with testosterone only occurs when the protein is folded. Furthermore, the remaining disulfide bond in NCS-3.24 (Cys88–Cys93) is located relatively far from the testosterone binding

site; thus in most cases its mutation should not directly interfere with selection and the selected mutants should be both folded and functional.

Selection was performed by phage display by using biotinylated testosterone bound to streptavidin-coated magnetic beads. After three rounds of selection, phage ELISA was used to test the functionality of a sample of individual clones (36 clones for the second round and 36 clones for the third round). After the second round of selection, two clones were positive (2/36; 5.5%) and after the third round, eighteen clones were positive (18/36; 50%). The increasing proportion of positive clones by phage-ELISA was indicative of an efficient selection procedure for testosterone binding, and thus for folded proteins exposed on the phage surface.

Sequences of the selected clones

The sequences of the eighteen positive clones from the third round of selection screened by phage-ELISA were analysed (Table 1). None of the sequenced clones contained cysteine. One nucleotide sequence was found three times (clone 55). Finally, sixteen different sequences were identified. The results of the observed enrichment coupled with the sequence analysis indicated that testosterone binding of a cysteine-free NCS exposed on the phage surface was a functional selection powerful enough to select a variety of structural solutions among a vast number of unfoldable sequences. A detailed analysis of the sequences revealed consensus solutions for Cys88–Cys93 replacement (Table 1).

The redundant clone with the sequence Ala88 Thr Lys Thr Val Thr93, seems to be favoured and also contains motifs conserved in several other variants. Residues Ala88 or Thr89 are present in thirteen variants, eight of which contain both. Basic residues are present in eight sequences in position 90. The residue Thr91 is present in six sequences and is associated with Lys90 in four. The alternative for position 91 seems to be a histidine residue, which is present in six sequences. Position 92 is the most diversified. However, the presence of four sequences with proline residues and three with glycine residues at this position is to be noted as these amino acids are essentially absent in the other positions. These residues are expected to be favoured in constrained loop geometries. Finally, at position 93 three families can be distinguished, one with either Thr93 or Ser93 (nine variants), one with Met93 (four variants), and the last with Asn93 (two variants), which can be thought of as possessing the polar properties of serine and threonine and the steric properties of methionine. We also note that the cysteine pair Cys88–Cys93 is replaced by the exclusive associations of Met93 with Ala88 and Asn93 with Thr88, while Thr/Ser93 is associated with Val, Thr or Ala88.

Table 1. Sequences of the selected clones after functional selection and cytoplasmic expression screening.

Screen[a]	Clone	DNA sequence	Protein sequence[b]						Solubility[c]	Binding[d]
		VNN VNN VNN VNN VNN VNN	88	89	90	91	92	93		
Periplasmic screen	51	GCA ACA ACA AAT GAA ATG	**A**	**T**	T	N	E	M		
	52	GCG ACA AAA ACG GTA AGC	**A**	**T**	K	T	V	S	+	+
	53	GCA CGG AAA GTA CCC AGT	**A**	R	K	V	P	S	+	
	54	ACC GAG GCA GTG CCA AAC	T	E	A	V	P	N		
	55	GCC ACC AAA ACC GTG ACG	**A**	**T**	K	T	V	T	+	+
	56	GCA ACG CAC CAC AAA ATG	**A**	**T**	H	H	K	M		
	57	GCA ACT CGG CAC GTA AGT	**A**	**T**	R	H	V	S		
	60	GCC ACA ATC CAT CCA GTA	**A**	**T**	I	H	P	V	+	+
	61	GCG ACA CAA CAC CAG ATG	**A**	**T**	Q	H	Q	M	+	+
	62	ACA AGC AAG ACG GGA AGT	T	S	K	T	G	S	–	
	63	ACC ACA ACG ACA ATA AAC	T	**T**	T	T	I	N	+	+
	64	GTG ACA GAT CTG GTA ACC	V	**T**	D	L	V	T	+	+
	66	GTT GGG AAA ACA GGG AGC	V	G	K	T	G	S	–	
	67	GCT GAT ATA ACG GGG ACT	**A**	D	I	T	G	T	–	
	68	GCT GAG AAA CAT CCT ATG	**A**	E	K	H	P	M	+	+
	69	GCA ACA AAA CAC CAT ACC	**A**	**T**	K	H	H	T		
		Consensus	**A**	**T**	K/R	T/H	X	S/T		
Cytoplasmic screen	A6	GCT ACA AGG ACT CAA ACG	**A**	**T**	R	T	Q	T	+	+
	A7	GCG CGG CGA ACA GCG AAT	**A**	R	R	T	A	N	–	
	C7	GCA AAT AAG GTT CCC ATG	**A**	N	K	V	P	M	+	+
	C8	GCA ACC GAA AGT GAA ATG	**A**	**T**	E	S	E	M	+	+
	D7	GCG ACA ATA GTA CCG ACG	**A**	**T**	I	V	P	T	–	
	D10	GCG ACC CGG GTC GAA ACG	**A**	**T**	R	V	E	T	–	
	E8	GCA ACA GCA CAT GAG ATG	**A**	**T**	A	H	E	M	+	+
		Consensus	**A**	**T**	X	X	E/Q	M		
	NCS-3.24	TGC ACC ACC GCG GCA TGC	C	T	T	A	A	C	–	+

[a] The periplasmic screen was performed by phage-ELISA and the cytoplasmic screen was performed by CoFi blot on the ΔOmpT clones. [b] Amino acids found with a strong consensus (≥60% of the sequences) are indicated in bold; amino acids found with a moderate consensus (40–60%) are underlined. [c] Solubility was assessed by Western blot analysis after expression in liquid cultures at 30 °C for the periplasmic screen and at 37 °C for the cytoplasmic screen. [d] Binding of the purified proteins was tested by ELISA for their ability to bind immobilized testosterone.

Periplasmic expression of the selected variants

Certain proteins, such as antibody fragments, are occasionally expressed in folded and functional form at the low level required for display on a phage particle, but the same proteins can be unfolded or insoluble when expressed at high concentration in bacterial culture.[32,33] In order to directly evaluate the expression and folding properties of the selected proteins, eleven NCS variants were chosen with the different pairs of residues 88–93 identified through the sequence analysis. The same conditions used for the selection procedure were employed: production at 30 °C in the periplasmic space of *E. coli*. Among the eleven proteins produced, eight were expressed in soluble form as shown by Western blot analysis (Table 1) and seven could be purified with a yield of 2 to 5 mg per litre of culture. Among these seven variants, no apparent bias for preferred consensus mutations could be detected. An ELISA experiment performed with the seven purified proteins revealed that all seven variants were clearly still able to specifically bind testosterone (Table 1). The observed signal was similar for NCS-3.24 and for all mutants: soluble testosterone (100 μM or 1 mM THS) competed with immobilised testosterone in a dose-dependent response for binding the NCS variants. Clone 52 (named hereafter NCS-C52), which displays a sequence close to the consensus, was the best expressed in periplasm and was chosen for further characterisation along with certain other clones described below.

Taken together, these results show that a range of different cysteine-free NCS variants, with distinct consensus sequence elements, can be obtained by directed evolution. After three rounds of selection, 50% of the clones were positive for testosterone binding by phage ELISA, and 63% of those variants could be exported to the periplasm of *E. coli* and obtained as pure, soluble and functional proteins. These new variants were further screened for their cytoplasmic expression properties.

Screening for cytoplasmic expression

As we have demonstrated that NCS is tolerant to the removal of its two disulfide bonds by directed evolution, the cysteine-free variants are potentially well suited for the development of intracellular applications. In order to find the most soluble variants expressed in the bacterial cytoplasm, the clones obtained after the third round of phage display selection were transferred into a cytoplasmic expression format.

The periplasmic export sequence (OmpT) was removed from the pool of phagemids obtained after the third round of selection. The resulting sublibrary contained 10^6 clones. We assumed that no bias was introduced in this step and that diversity of this sublibrary reflected diversity obtained after the selection process. No periplasmic export sequence could be detected by a restriction analysis of this DNA population. This sublibrary (pCytex-NCS-RdCys) was further screened for cytoplasmic expression at 37 °C in the *E. coli* strain BLR(DE3)pLysS. While periplasmic proteins were previously expressed at 30 °C, cytoplasmic expression was performed at 37 °C. As high temperature increases protein aggregation, this more stringent screening condition was chosen to isolate stable soluble clones. Furthermore, with the aim of developing intrabody applications we chose to select folded NCS scaffolds at the usual growth temperature (37 °C) for diverse cell types.

A sample of 60 individual clones was grown in a 96-well matrix plate. This matrix was replicated on agar plates, and the colonies were screened for total expression by a colony lift assay,[34] and for soluble expression by a colony filtration blot (CoFi blot; Figure 2).[35,36] For these two methods, the proteins

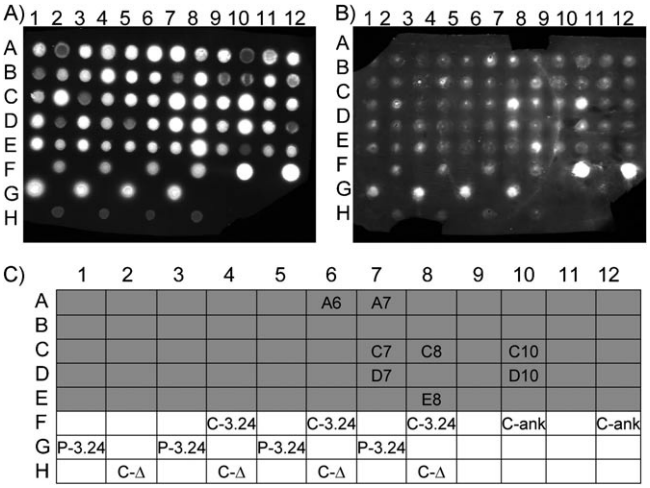

Figure 2. Screening for cytoplasmic expression of the cysteine-free NCS variants. Screening by: A) colony lift, and B) CoFi blot experiments followed by fluorescent detection of the His-tag. Colony lift detects total protein expression and CoFi blot detects soluble proteins expressed at 37 °C and filtered after a lysis step. C) Composition of the matrix plate used for the colony lift and CoFi blot experiments. The grey wells contained the screened clones of the library. The wells named "C-3.24" contained NCS-3.24 expressed in the cytoplasm, "P-3.24" contained NCS-3.24 expressed in the periplasmic compartment (soluble), "C-Ank" contained a highly soluble cytoplasmic protein as a positive control (ankyrin), "C-Δ" contained a deleted NCS-3.24 used as a negative control for cytoplasmic expression. Positive clones obtained from the CoFi blot experiment after analysis of the signal intensity are represented in the grey wells.

recovered on the nitrocellulose filter were revealed by fluorescent immunodetection of the C-terminal His-tag (see the Experimental Section). The colony lift experiment indicated that 96% (58/60) of the screened colonies expressed the cytoplasmic protein (Figure 2A). This technique, however, while allowing elimination of unproductive clones, does not discriminate between soluble and insoluble expression. With the CoFi blot, soluble expressed proteins are specifically revealed. The CoFi blot experiment revealed that 13% (8/60) of the sample were over-expressed as soluble proteins (Figure 2B). The eight positive clones (Figure 2C) from the CoFi blot were sequenced; two were identical (C7 and C10).

The sequence of these seven variants revealed a strong consensus with an alanine in position 88 for all variants, a threonine in position 89 for five variants, and in position 93 either a methionine or a threonine (Table 1). Positions 90 to 92 were more variable but charged residues were found six times in position 90. These consensus mutations (Ala88, Thr89, Thr93 or

Met93) found here after the cytoplasmic expression screen, had also been found after the periplasmic expression screen by phage display (Table 1). However, while the mutation Cys93Met was only found in 22% of the clones after the first screen, it represented 42% of the clones after the cytoplasmic screen.

In order to validate the CoFi blot results, the expression of the seven different positive clones was tested at 37 °C in liquid culture (Table 1). The total cellular extracts and soluble fractions were analysed by Western blot. One variant was weakly expressed (NCS-D10), two were expressed but weakly soluble (NCS-A7, NCS-D7) and four were expressed and mainly present in the soluble fraction (NCS-A6, NCS-C7, NCS-C8, NCS-E8; data not shown). Among these four soluble variants, the mutation Cys93Met was found three times. These four variants were further characterised in order to evaluate the effect of disulfide bond removal on the structure and stability of the NCS scaffold.

Functional and structural assessment of the selected mutants

The four soluble cytoplasmic variants (named hereafter NCS-A6, NCS-C7, NCS-C8, NCS-E8), the periplasmic variant (NCS-C52), the only rational mutant we could purify (Cys88Ala–Cys93Ala named hereafter NCS-wt-Δ2) and the NCS-wt and NCS-3.24 proteins were produced and purified for further characterisation.

An ELISA experiment clearly showed that all four cytoplasmic soluble variants were still able to bind testosterone specifically (Figure 3). No signal was observed in the absence of immobilised testosterone. Soluble testosterone competed with immobilised testosterone for the binding of each purified protein. The efficiency of the competitor was similar for each variant and for the protein NCS-3.24; this indicates that the affinities were in the same range.

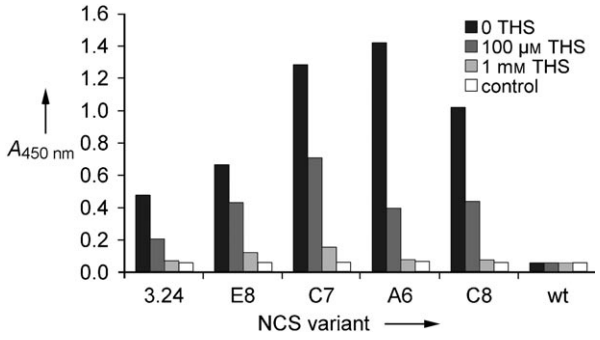

Figure 3. Testosterone binding of cytoplasmic NCS variants. The purified proteins (50 nM) were tested by ELISA for their ability to bind to immobilised testosterone in the absence (black) and in the presence of 100 μM (dark grey) and 1 mM (light grey) free testosterone hemisuccinate (THS), and to bind to the ELISA plate in the absence of immobilised testosterone (control; white). The fraction of residual binding signal is represented after background correction and normalisation. Average absorbance for two experiments is represented.

Disulfide bond removal could induce a loss in ligand affinity as reported for the intrabody V_L12.3, which is a potential inhibitor of huntingtin aggregation in Huntington's disease.[26] To demonstrate quantitatively that affinity was maintained in our binding selections, we performed isothermal titration calorimetry (ITC) experiments for the protein variants NCS-wt, NCS-3.24 and NCS-E8 (Figure 4). These experiments were performed by

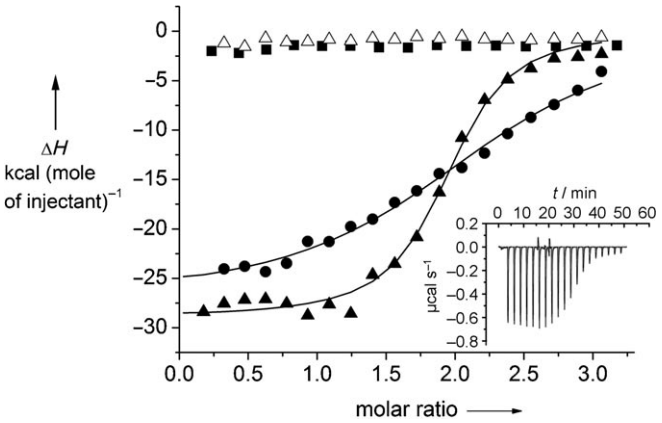

Figure 4. Microcalorimetric titration of testosterone binding with NCS-E8, NCS-3.24 and NCS-wt. The reaction cell was loaded with a solution containing 10 μM streptavidin tetramer complex saturated with biotinylated testosterone. Injections of 2 μL of 150 μM NCS-E8 (▲), NCS-3.24 (●) and NCS-wt (■) were carries out every 150 s from a 40 μL syringe. Control experiments were performed by injecting either NCS-3.24 (not shown) or NCS-E8 (△) into a cell containing streptavidin (10 μM) alone. The solid line is the curve fit to the data by the Origin program. This fit yields values for K_d and stoichiometry.

using the same biotinylated testosterone exposed on streptavidin as in the selections. No binding signal was observed for NCS-wt whereas binding signals were observed with NCS-3.24 and NCS-E8. Also, no binding signal was observed when these proteins were mixed with streptavidin alone; this indicates specificity for testosterone binding. The titration curve indicated a stoichiometry of two NCS molecules bound per streptavidin tetramer. This result is consistent with our previously published structural study showing that one NCS protein binds two testosterone molecules exposed on each side of the streptavidin tetramer.[15] Dissociation constants measured in solution were (1.9×10^{-6}) M and (2.6×10^{-7}) M for NCS-3.24 and NCS-E8, respectively; this indicates that the disulfide-free evolved mutant NCS-E8 preserved, and even improved, its ability to bind testosterone. This result is consistent with the selection procedure as disulfide-free mutants were selected for testosterone binding.

To ensure that the mutations did not significantly modify the protein structure, the recombinant proteins were further characterised by circular dichroism. CD spectra were recorded for NCS-wt, NCS-3.24 and different mutant proteins under identical conditions at 25 °C (Figure 5A, B). All spectra have similar shapes and display the characteristic signal of an all β protein with a maximum at 195 nm and minimum at 210 nm. The spectra also showed the positive contribution at approximately 223 nm, which is specific to apo-NCS and has

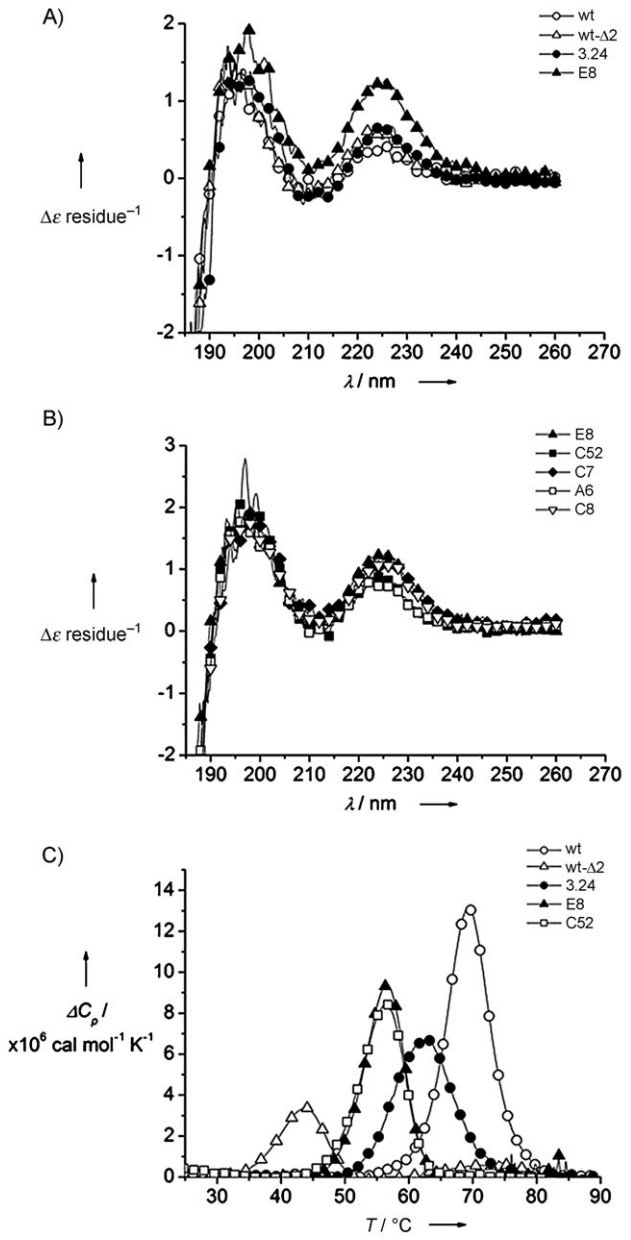

Figure 5. Biophysical properties of the NCS variants assessed by CD and DSC. A) CD spectra of the NCS variants -wt, -wt-Δ2, -3.24 and the mutant -E8 obtained by directed evolution; B) CD spectra of the NCS variants -E8, -A6, -C7, -C8 and -C52 obtained by directed evolution; C) DSC traces of the NCS variants -wt, -wt-Δ2, -3.24, -E8 and -C52.

been reported to represent "no typical" secondary structure.[17,37,38] These data indicate that the mutations induced little change in the conformation of the protein. For all mutants, the positive contribution at 223 nm is significantly higher than that observed with NCS-3.24 for which loop 88–93 was not modified and slight differences between the mutants were observed in this region. This strengthens the hypothesis according to which this positive contribution is related to this loop but not to the disulfide bond itself.

Thermal stability of selected mutants compared to NCS-3.24 and NCS-wt

Differential scanning calorimetry (DSC) was used to analyse the thermal stability of the selected mutants compared to that of NCS-wt, NCS-3.24 and the rational mutant NCS-wt-Δ2 (Figure 5C). The unfolding of the wild-type protein resulted in a transition peak centred at 69.3 °C (Table 2). When analysed by a single-transition model, a $\Delta H_{vH}/\Delta H_{cal}$ ratio of 0.96 was obtained; this suggests a two-state model[39] with a calorimetric enthalpy ΔH_{cal} of 111 kcal mol^{-1}. The unfolding of NCS-3.24 resulted in a transition peak centred at 62.4 °C. The loss of the first disulfide Cys37–Cys47 (along with accompanying hapten-binding mutations) was thus seen to decrease the melting temperature by approximately 7 °C. This shift was accompanied by a decrease in the denaturation enthalpy to 73.7 kcal mol^{-1}.

For the five disulfide-free mutants obtained from NCS-3.24 with our targeted directed evolution approach, an additional decrease in the melting temperature was observed, with a shift comprised between 6 and 11 °C compared to NCS-3.24. Except for the NCS-C7 mutant, the $\Delta H_{vH}/\Delta H_{cal}$ ratios were again near unity; this suggests a two-state model as for the wild-type protein. Moderate decreases in denaturation enthalpy were observed for these mutants compared to NCS-wt (74.5 to 101.4 kcal mol^{-1} vs. 111 kcal mol^{-1} for NCS-wt). A second scan of each sample performed after cooling showed that we recovered 10–40% of the enthalpy for all the tested disulfide-free mutants. No reversibility was observed for NCS-3.24.

For comparison, NCS-wt-Δ2, which was the only rational mutant that could be purified, was much more unstable than the mutants obtained by directed evolution. The decrease of melting temperature (T_m) was about 19 and 26 °C compared to NCS-3.24 and NCS-wt, respectively. An asymmetrical transition was also observed for this mutant, the denaturation enthalpy decreased significantly (46.5 kcal mol^{-1} vs. 73.7 kcal mol^{-1}) and the ratio $\Delta H_{vH}/\Delta H_{cal}$ was 1.9; this suggests that aggregation occurs. For this mutant 66% reversibility was obtained.

The thermal stability of the four soluble mutants identified after expression in the most stringent conditions (cytoplasmic expression) is in the same range as the thermal stability of the mutant NCS-C52 obtained after periplasmic expression (T_m 55.9 °C for NCS-C52 and 56.2 °C for NCS-E8). It is remarkable to note that all the mutants expressed as soluble proteins also presented partial reversibility upon thermal denaturation, contrary to NCS-3.24 for which thermal denaturation was fully irreversible. The propensity of the selected proteins to resist aggregation during denaturation suggests that the procedure used here has selected against aggregation-prone folding intermediates.

Molecular simulations of NCS-E8 compared to NCS-3.24

We selected the NCS-E8 mutant, which exhibited a high stability and degree of cytoplasmic expression, for interpretation using molecular modelling and simulation. A homology-mod-

Table 2. Thermodynamic parameters of the heat denaturation of the NCS variants monitored by DSC.

NCS variant[a]	No. of disulfide bonds	ΔH_{cal} [kcal mol^{-1}]	ΔH_{vH} [kcal mol^{-1}]	T_m [°C]	$\Delta H_{vH}/\Delta H_{cal}$	Recovery in 2nd scan [%]
wt	2	111	107	69.3	0.96	97
3.24	1	73.7	79.4	62.4	1.07	0
E8	0	88.4	99.1	56.2	1.12	40
C7	0	56	109	55.2	1.94	26
A6	0	94.4	88.6	52.3	0.94	10
C8	0	101.4	92.9	51.8	0.92	36
C52	0	74.5	57.3	55.9	0.77	24
wt-Δ2	1	46.5	88.2	43.2	1.90	66

[a] Variants E8, C7, A6 and C8 were screened for their cytoplasmic soluble expression. The variant NCS-C52 was screened by phage-ELISA for testosterone binding. The variant wt-Δ2 (Cys88Ala–Cys93Ala) is the only rationally engineered mutant presenting enough stability to be expressed and purified.

elled structure of NCS-E8 was obtained by using NCS-3.24 (PDB ID: 2cbm) as a template (95% identity). Molecular dynamics (MD) simulations were used to validate this model and to obtain structural insights into this example of a solution to disulfide replacement, and compared with results obtained for NCS-3.24 with the original disulfide bond.

MD of the NCS-3.24 structure provided residue mobilities that correlated well with measured NMR spectroscopy order parameters for NCS-wt;[40] in particular, the region of highest mobility was confirmed to be the FG loop bordering the binding cavity.[16] The modelled protein NCS-E8 behaved very similar to NCS-3.24, judging by root mean square (rms) C_α distances from the starting structure of (1.4 ± 0.2) Å in both cases and nearly identical C_α fluctuation profiles (data not shown). Overall, these results suggest that the modelled NCS-E8 mutant has dynamic properties that are similar to those of the starting NCS-3.24 and the wild-type protein.

In NCS-E8, the mutated loop region exhibited no additional fluctuation compared to NCS-3.24 in which the loop is constrained by the disulfide bond. In the modelled structure, Met93 appears to act as a flexible anchor for beta strand F. Throughout the simulations, the methionine side chain was found principally in an extended conformation, which closely approximated the geometry of the original disulfide bond. In this conformation Met93 atoms C_β, C_γ and S_δ in NCS-E8 occupy positions that superpose nearly exactly on their respective counterparts Cys93-C_β, Cys93-S_γ and Cys88-S_γ in NCS-3.24, such that the S_δ of Met93 is in van der Waals contact with the C_β of Ala88 (Figure 6). The geometry of the arrangement is also consistent with favourable interactions of the Met S_δ with nucleophiles and electrophiles highlighted in statistical analyses of protein structures.[41]

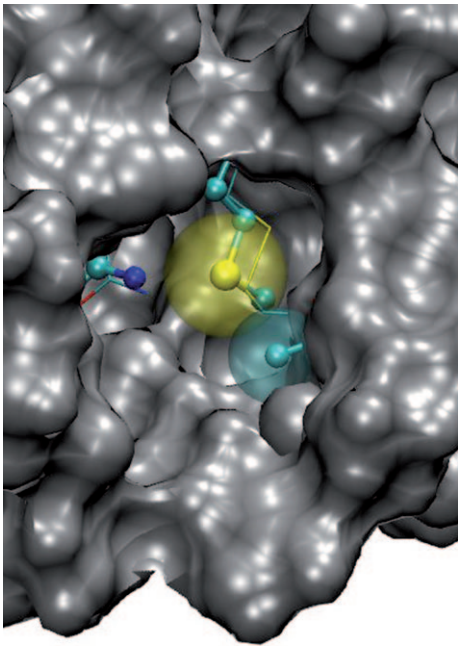

Figure 6. Structure of NCS-3.24 superimposed on that of the modelled mutant NCS-E8, showing the mutated disulfide bond. Structures are displayed in a line representation for NCS-4.24 and ball-and-stick representation for NCS-E8. Mutated residues Ala88 and Met93 are shown in the principal conformation observed in the molecular dynamics simulations. The binding cleft is at the top of the figure, with the immunoglobulin-like domain to the left. The side chain of Gln36 is indicated for both structures. The van der Waals spheres of the Met93 δ sulfur and the Ala88 β carbon are shown as semitransparent surfaces. The last two residues of the protein are omitted for clarity; the figure was prepared with VMD.[51]

Discussion

The role of disulfide bonds in protein structure, stability and folding is often crucial but differs from protein to protein depending on chain orientation, accessibility and electrostatic environment.[18] In addition, disulfide bonds complicate the design of protein scaffolds intended to be functional in various redox environments. However, the payoff for developing such scaffolds is substantial, including in particular the possibility of developing targeting vectors for therapeutic applications or intrabodies for studying fundamental molecular processes in the cell. In this work our goal was to create such a scaffold, with the specific objective of obtaining soluble and stable NCS mutants in the reducing environment of the cytoplasm. Wild-type NCS binds its natural ligand very tightly and some other compounds with lower affinity;[11,42] the range of compounds potentially bound by the protein can also be extended by engineering its natural crevice.[14] Indeed, NCS variant 3.24 has already been engineered to bind a new ligand (testosterone)[14] and is well-suited to screening for folded proteins. Only one disulfide bond (Cys88–Cys93) remains in this variant; we have used directed evolution to find viable solutions for a cysteine-free scaffold.

Rational engineering versus directed evolution for disulfide bond replacement

Disulfide bond replacement can be approached either by rational engineering or by directed evolution techniques. In practice, the rational approach often defaults to substituting one or both cysteines by small hydrophobic residues, as it is not a trivial problem to predict mutations that allow compensation of the energetic and structural properties of disulfide bonds. When this approach was applied to the NCS-wt and NCS-3.24 scaffolds, the mutants Cys88Ala–Cys93Ala were either not or poorly expressed. The only mutant we could purify (NCS-wt-Δ2) was destabilised by more than 25 °C in terms of T_m. Residues Cys37–Cys47, located in the ligand-binding crevice, could not be replaced by alanines at all: all rational mutants carrying these mutations were unexpressed. These results highlight the context dependent role of disulfide bonds.

Directed evolution, on the other hand, allows a large region of the sequence space to be explored in parallel and favours the discovery of novel solutions. A key point in a directed evolution approach to disulfide bond replacement is the necessity to select folded variants from a random library. While in general such a folding selection is still an unsolved problem, functional selection (e.g., binding or catalysis) is much easier. The natural binding function of NCS-wt was useless for such a purpose, because of the high instability of the natural ligand; so we exploited the variant NCS-3.24, which had been evolved to bind a more convenient ligand.[14] The location of the disulfide bond targeted for replacement (Cys88–Cys93) allowed randomisation of the disulfide bond and the selection of folded proteins with binding functions essentially intact.

Stringency of cytoplasmic expression

The selection process used in our work combines a function-plus-folding selection of disulfide-free mutants based on periplasmic expression (phage display) and a screening step for cytoplasmic expression. Three rounds of functional selection led to a diverse population of well-folded variants with good expression properties in the periplasm. After transfer to a cytoplasmic expression system, the percentage of soluble clones and the diversity of adapted sequences decreased dramatically. These results suggest that only a subset of the mutants selected on phage particles are also able to fold efficiently in the cytoplasm. Thermal denaturation of all soluble clones from the periplasmic and cytoplasmic screens is partially reversible, and their thermal stabilities are in the same range. This suggests that additional selection pressures could be acting in cytoplasmic expression that would be not directly related to thermal stability but rather to folding and misfolding pathways.[43] Indeed, cytoplasmic expression can lead to very high expressed protein concentrations, while in periplasmic expression NCS mutants are exported to the periplasm for folding and then secreted into the extracellular medium where protein concentration is very low.

Mutations selected in the disulfide-free NCS mutants

Using three rounds of phage display selection, we identified several solutions to replacing the Cys88–Cys93 disulfide bond on the NCS scaffold. The constrained structure of the loop defined by this bridge is probably important for folding but viable solutions can be obtained without the associated disulfide bond. However, as seen by the need for several selection rounds, clearly not all amino acid sequences are compatible with the disulfide-free loop. The solutions for the replacement of disulfide 88–93 exhibited consensus signals, which are, as expected, stronger for buried positions (88 and 93) than for solvent-exposed positions (89 to 92). Solvent exposed positions 89, 90 and 92 accept mainly polar or charged residues. The NMR structure of NCS-wt (PDB ID: 2g0k)[42] shows H-bond networks between residues in this loop and their neighbours. Concerning position 91, no charged residue was observed in sixteen sequences; this suggests that such a residue destabilises the structure. According to the modelled structure of NCS-E8, the conformation of the side chain at this position (His91) masks a buried hydrophobic side chain (Val86) while interacting with polar groups of the main chain. Such a conformation would not be adopted by a charged side chain, which would be most favoured when extended into the high dielectric medium of the solvent. Other residues seen at this position (Thr, His, Val, Asn, Leu) in the mutant proteins also suggest the importance of masking a hydrophobic patch that would otherwise be exposed.

In the two cysteine positions 88 and 93, we observed a strong consensus for Ala88 combined with a polar residue or a methionine in position 93. Alanine is a small, surface-compatible residue without the high flexibility that would be conferred by a glycine. In position 93, the model of NCS-E8 suggests that the methionine acts as an anchor for the long β strand, F, in the absence of the wild-type disulfide link. The anchoring of strand F, one end of which participates in the "thumb" structure and the other in the β barrel, could be essential to maintaining the structural integrity of the NCS scaffold itself. An additional role for such an anchor might be to simply fix the geometry of the binding pocket—if this strand were too flexible in the unbound state, ligand binding would be correspondingly burdened by an unfavourable entropic contribution. Of the naturally occurring amino acids, methionine is well-adapted to playing the role of such an anchor thanks to its linear topology, which allows it to maintain packing in the protein interior. Further, the Met sulfur would appear to satisfy electrostatic interactions in the protein interior made by the disulfide that could play a role in stability.[41] The asymmetric pair Met–Ala thus appears to mimic the disulfide bond and could constitute a generic suggestion for disulfide replacement.

Conclusions

A major goal of the current work was to evolve NCS into a disulfide-free polyvalent scaffold suitable for cytoplasmic expression. A key aspect of our approach was to exploit an evolved binding function in this protein to allow us to select well-ex-

pressed and stable molecules. The strategy we have demonstrated, while not a folding selection per se, could nevertheless be used to select for structure- and stability-related targets for any protein presenting a natural or engineered binding function.

Experimental Section

Construction of the library: All constructs were prepared by using the phage display and expression vector pHDiex previously described.[14] The NCS-3.24 gene was cloned between the restriction sites NotI and HindIII. A MluI restriction site is located in position 272 of the gene. An NCS-deleted vector pHDiex-0 was obtained by replacing the fragment of the NCS-3.24 gene MluI–HindIII by a small cassette containing the restriction site SmaI. This vector pHDiex-0 was further used as a negative control for protein expression.

The library was constructed by cloning a degenerated cassette with cohesive ends compatible with the MluI and HindIII restriction sites into pHDiex-0. The cassette was generated with three oligonucleotides: a degenerate oligonucleotide randomizing the protein sequence Cys88 Thr89 Thr90 Ala91 Ala92 Cys93 "NCS-Δcys-rd" (CGC GCT GGG GAA CTG TGG ACV NNV NNV NNV NNV NNV NNC AGG TGG GCT TGA GTG ATG CAG CAG GCA ACG GAC CCG AAG GTG TGG CAA TC) and two nondegenerate oligonucleotides generating the cohesive extremities of the cassette "O1" (AGC TGA TTG CCA CAC CTT CGG GTC CGT TGC CTG CTG CAT CAC TCA AGC CCA CCT G) and "O2" (GTC CAC AGT TCC CCA G). The oligonucleotides were hybridised and ligated in linearised pHDiex-0 vector (1 μg). The resulting ligation product was digested with SmaI to eliminate background arising from the initial pHDiex-0 vector, purified on a Qiagen column and finally electroporated into the *E. coli* XI1 Blue MRF′ strain; this yielded a library of 10^7 clones (pHDiex-NCS-RdCys). The randomized region remained single stranded in vitro but was corrected into double-stranded DNA in vivo by bacterial polymerases after the electroporation step.

Phage display selection: The XL1 Blue MRF′ bacteria from the phagemid library were infected with the helper phage M13KO7 (New England Biolabs) and phages were allowed to replicate, overnight, at 30 °C. Culture supernatants were precipitated with PEG (20% PEG 6000, 2.5 M NaCl) and solubilised in TBS (20 mM Tris-HCl, pH 7.5, 150 mM NaCl). The phage library was panned by using streptavidin magnetic beads (Roche Diagnostics) coated with biotinylated testosterone (Sigma) essentially as described.[14] For the three rounds of selection, 10^{13} phages were incubated for 1 h in TBST (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% Tween 20) with the bead suspension (50 μL) in a total volume of 1 mL. Unbound phages were removed by washing eight times in TBST. Bound phages were eluted with glycine (0.1 M, pH 2.4) and neutralised with Tris-HCl (1 M, pH 7.4). Exponentially growing XL1 Blue MRF′ cells were infected with the eluted phages and spread onto 2YT plates containing glucose (1%, *w/v*) and ampicillin (200 μg mL^{-1}). The recovered bacteria were further used for the following selection round.

Screening for testosterone binding by phage ELISA: Individual clones obtained after the second and third selection rounds were screened for testosterone binding by phage ELISA. Individual colonies from these two selection rounds were randomly picked and grown, overnight, at 37 °C in a 96-well microplate in 2YT (150 μL) containing ampicillin (200 μg mL^{-1}), tetracycline (12 μg mL^{-1}) and glucose (1%). This master plate of clonal culture in the *E. coli* strain XL1 Blue MRF′ was used as a preculture plate for phage production and as a matrix stored at −80 °C in the presence of glycerol (20%). Exponentially growing cells (100 μL) were infected for 1 h at 37 °C with 10^{10} particles of helper phage M13KO7 and transferred into 2YT (1.5 mL) containing ampicillin (200 μg mL^{-1}) and kanamycin (50 μg mL^{-1}) in a deep-well culture plate (ABgene). The phage particles were produced, overnight, at 30 °C. Phage ELISA was performed as described by using the phage supernatant for each well (100 μL).[14] A Maxisorp ELISA plate (Nunc) was coated with streptavidin–biotinylated testosterone complex.[14] Bound phages were revealed with a horseradish peroxidase conjugated anti-M13 monoclonal antibody (Amersham) and detected at 450 nm by using BM Blue POD as a substrate (Roche Diagnostics) after the addition of H_2SO_4.

Deletion of the periplasmic export sequence: The vector pHDiex[14] contains the periplasmic export presequence OmpT, which allows phage surface exposure and periplasmic expression of the proteins. For the cytoplasmic expression screening, the presequence OmpT was deleted from the pHDiex vector library obtained after the third round of selection. The pool of phagemids was digested with NdeI and NotI successively and purified on a Qiagen column. A cassette with cohesive ends compatible with the digested phagemid was generated by annealing the two primers "strep-rev" (GGC CGC TTT CTC GAA CTG AGG ATG AGA CCA CGC CA) and "strep-sens" (TAT GGC GTG GTC TCA TCC TCA GTT CGA GAA AGC). The cassette was ligated into the vector by using T4 DNA ligase. The ligation product was digested by AvrII, purified and electroporated into the *E. coli* XI1 Blue MRF′ strain to yield the pCytex-NCS-RdCys sublibrary for cytoplasmic expression.

The same procedure was performed with individual pHDiex phagemids (pHDiex-0 and NCS-3.24) to result in the pCytex-0 and pCytex-NCS-3.24 vectors used for the cytoplasmic expression tests.

Screening for cytoplasmic expression: The pCytex-NCS-RdCys library was transformed in the *E. coli* strain BL21-DE3-plysS. Individual colonies were randomly picked and grown at 37 °C, overnight, in a 96-well microplate in 2YT (100 μL) containing ampicillin (200 μg mL^{-1}) and glucose (1%). Glycerol (20%) was added in each well and the plate was stored at −80 °C. This master plate was used as a matrix for the cytoplasmic expression screening experiments (colony lift and CoFi blot).

Colony lift: The master plate was replicated on a 2YT agar plate containing ampicillin (200 μg mL^{-1}) and glucose (1%). After incubation at 37 °C, overnight, colonies were transferred onto a nitrocellulose membrane. The membrane was placed, with colonies facing up, on a 2YT agar plate containing ampicillin (200 μg mL^{-1}) and IPTG (500 μM) and bacteria were grown at 37 °C for 4 h. The nitrocellulose membrane was recovered, washed with water to eliminate the bacteria blocked with TBS-Tween (0.1%) before the detection step (see below). No lysis step was performed at this stage; the NCS mutants are probably toxic in the cytoplasm at high concentration and might trigger bacterial lysis.

CoFi blot: This was performed as described.[35,36] The master plate was replicated on a 2YT agar plate containing ampicillin (200 μg mL^{-1}) and glucose (1%). After incubation at 37 °C, overnight, colonies were transferred onto a Durapore filter (0.45 μm; Millipore). The filter was placed, with the colonies facing up, on a 2YT agar plate containing ampicillin (200 μg mL^{-1}) and IPTG (500 μM) and protein expression was induced at 37 °C for 4 h. The filter was placed on top of a nitrocellulose filter and a sheet of 3 MM Whatman paper presoaked in lysis buffer (20 mM Tris-HCl. pH 7.5, 150 mM NaCl, 0.2 mg mL^{-1} lysozyme, 100 U benzonase (No-

vagen), and one complete EDTA-free protease inhibitor cocktail tablet (Roche) in 50 mL buffer). The "filter sandwich" was incubated at 25 °C for 30 min and lysis was continued with three freeze-thaw cycles (10 min at −80 °C and 10 min at 37 °C). The nitrocellulose membrane was removed from the sandwich and treated as with the colony lift membrane before analysis.

Protein detection: Expressed proteins were immunodetected on nitrocellulose filter with a mouse anti-His antibody (Qiagen) followed by an AlexaFluor 680 labelled anti-mouse antibody (Molecular Probes). Fluorescence detection was performed by using the Odyssey® Infrared Imaging System (Li-Cor) with excitation at 680 nm and emission at 700 nm. Positive clones were detected by using a quantitative fluorescence signal analysis.

Expression and purification of the selected variants: *E. coli* BLR-(DE3)pLysS strain (Novagen) was used as expression host for NCS mutants. Periplasmic expression was performed at 30 °C for 36 to 48 h in 2YT medium in the presence of ampicillin (200 μg mL^{-1}). Under these conditions the proteins were efficiently secreted and purified as previously described.[15]

For cytoplasmic expression, *E. coli* cells transformed with the pCytex-NCS vectors were grown at 37 °C in 2YT medium (1 L) to which ampicillin was added (200 μg mL^{-1}). When the OD_{600} reached 0.7, protein expression was induced with IPTG (0.5 mM) for 4 h. Bacterial pellets were suspended in phosphate buffer (50 mM, pH 7.9) submitted to three freezing/defreezing cycles and treated with lysozyme and benzonase for 30 min. After centrifugation, the His-tagged protein was purified from supernatant by using nickel affinity chromatography (Ni-NTA-agarose, Qiagen) followed by size-exclusion chromatography (Hiload 16/60, Superdex 75).[15] For each mutant protein, purity of the final sample was checked with an overloaded SDS-PAGE gel, which had to show one well-resolved band with no visible contamination.

ELISA analysis: The functionality of the purified proteins was checked by ELISA as described[14] with 50 nM protein. A Maxisorp ELISA plate (Nunc) was coated with streptavidin–biotinylated testosterone complex. Bound proteins were detected with an anti-His_6-tag horseradish peroxidase conjugated monoclonal antibody (Roche Diagnostics). TBS, streptavidin and BSA coated wells were used as negative controls. Proteins were also incubated in the presence of a free competitor (100 μM and 1 mM testosterone hemisuccinate (THS)) to check the specificity of the testosterone binding.

Isothermal titration microcalorimetry: ITC experiments were performed at 25 °C with an ITC 200 isothermal titration calorimeter (Microcal). Aliquots (2 μL) of NCS variant (150 μM; NCS-wt, NCS-3.24 or NCS-E8) were injected from a computer-controlled microsyringe (40 μL) at intervals of 150 s into a solution of streptavidin–biotinylated testosterone complex (10 μM streptavidin tetramer, 50 μM biotinylated testosterone, cell volume 200 μL) that was stirred at 1000 rpm. All solutions were prepared in the same buffer containing sodium phosphate (50 mM) and ethanol (0.83 %). Control experiments were performed by injecting either NCS-3.24 or NCS-E8 into a cell containing streptavidin alone (10 μM). Experimental data were fitted to a theoretical titration curve by using Origin® software (MicroCal®). This fitting uses the relationship between the heat generated by each injection and ΔH (enthalpy change in kcal mol^{-1}), K_a (association binding constant in M^{-1}), n (number of NCS binding sites per streptavidin–testosterone tetramer), total protein concentration and free and total ligand concentrations.[44]

Circular dichroism: CD spectra were recorded from 185 to 250 nm on a Jasco dichrograph equipped with a thermostatically-controlled cell holder and connected to a computer for data acquisition. Data were acquired from 10 to 13 μM samples in phosphate buffer (20 mM, pH 7.5) in quartz cells with a 1 mm path length.

Differential scanning calorimetry: Thermal stability was studied by DSC by using a differential scanning calorimeter VP-DSC (Microcal). DSC measurements were made with an apo-NCS solution (0.15 to 0.3 mg mL^{-1}) in phosphate buffer (20 mM, pH 7.5). The buffer from the dialysis bath was used as a reference. All solutions were degassed just before loading into the calorimeter. Scanning was performed at 1 K min^{-1}. The percentage of recovery of the native protein after heat denaturation was evaluated by rescanning after the denatured sample had cooled.

The heat capacity of the solvent alone was subtracted from that of the protein sample. These corrected data were analysed by using a cubic spline as a baseline in the transition. Thermodynamic parameters, calorimetric enthalpy (ΔH_{cal}) and van't Hoff enthalpy (ΔH_{vH}), were determined according to the standard relationship provided by Privalov and Potekhin.[45] All calculations were performed with ORIGIN software (Microcal).

Molecular modelling and simulations: Molecular simulations were performed by using the crystal structure of the apo form of the starting protein NCS-3.24 (PDB ID: 2cbm)[15] as well as a modelled structure of the disulfide-free mutant NCS-E8. Structural models of the NCS-E8 mutant were generated by homology modelling (>95 % sequence identity) with Modeller[46] by using 2cbm as template and allowing modifications to residues within 5 Å of the mutated region. The model with the lowest objective function was retained. Molecular dynamics simulations (1 ns heating and equilibration plus 3 ns production for each molecule) were used to study the NCS-3.24 structure and to validate the NCS-E8 model, similar to the approach taken by Schmidt Am Busch et al.[47] Simulations were carried out with Charmm[48] by using the param27 all-atom force-field with CMAP.[49] After standard energy minimisation and heating stages,[50] isothermal–isobaric (NPT) dynamics were performed at 300 K and 1 atm pressure, with approximately 10 600 explicit TIP3 water molecules in a truncated icosahedral unit cell with periodic boundary conditions, 10 Å nonbonded-energy-term cutoff with force shifting, SHAKE constraints on bonds involving hydrogen atoms, and a 2 fs integration step. Structural variations throughout the simulations were measured by calculation of root mean square (rms) C_α distance of the structure with respect to the minimised structure used at the beginning of the heating and equilibration phase and by rms C_α fluctuations.

Abbreviations: NCS: neocarzinostatin; DSC: differential scanning calorimetry; MD: molecular dynamics; CD: circular dichroism; IPTG: isopropyl β-D-thiogalactopyranoside; THS: testosterone hemisuccinate; TBS: Tris buffer saline; BSA: bovine serum albumin.

Acknowledgements

Funding for A.D. was provided by CNRS-BDI program and Association pour la Recherche sur le Cancer. Funding for M.B.H.-R. was provided by the French Ministry of Education. Many of the calculations reported in this study were carried out at the IDRIS (Institute for Development and Resources in Intensive Scientific Computing) supercomputing facility of the CNRS.

Keywords: calorimetry · directed evolution · disulfide bonds · molecular dynamics · neocarzinostatin

[1] H. K. Binz, P. Amstutz, A. Plückthun, *Nat. Biotechnol.* **2005**, *23*, 1257.
[2] T. Hey, E. Fiedler, R. Rudolph, M. Fiedler, *Trends Biotechnol.* **2005**, *23*, 514.
[3] R. J. Hosse, A. Rothe, B. E. Power, *Protein Sci.* **2006**, *15*, 14.
[4] A. Skerra, *Curr. Opin. Biotechnol.* **2007**, *18*, 295.
[5] H. Leemhuis, V. Stein, A. D. Griffiths, F. Hollfelder, *Curr. Opin. Struct. Biol.* **2005**, *15*, 472.
[6] P. Soumillion, J. Fastrez, *Curr. Opin. Biotechnol.* **2001**, *12*, 387.
[7] H. K. Binz, A. Plückthun, *Curr. Opin. Biotechnol.* **2005**, *16*, 459.
[8] M. T. Stumpp, P. Amstutz, *Curr. Opin. Drug Discov. Devel.* **2007**, *10*, 153.
[9] C. Zahnd, E. Wyler, J. M. Schwenk, D. Steiner, M. C. Lawrence, N. M. McKern, F. Pecorari, C. W. Ward, T. O. Joos, A. Plückthun, *J. Mol. Biol.* **2007**, *369*, 1015.
[10] S. Schlehuber, A. Skerra, *Expert Opin. Biol. Ther.* **2005**, *5*, 1453.
[11] J. R. Baker, D. N. Woolfson, F. W. Muskett, R. G. Stoneman, M. D. Urbaniak, S. Caddick, *ChemBioChem* **2007**, *8*, 704.
[12] B. Heyd, G. Lerat, E. Adjadj, P. Minard, M. Desmadril, *J. Bacteriol.* **2000**, *182*, 1812.
[13] S. Nozaki, Y. Tomioka, T. Hishinuma, M. Inoue, Y. Nagumo, L. R. Tsuruta, K. Hayashi, T. Matsumoto, Y. Kato, S. Ishiwata, K. Itoh, T. Suzuki, M. Hirama, M. Mizugaki, *J. Biochem.* **2002**, *131*, 729.
[14] B. Heyd, F. Pecorari, B. Collinet, E. Adjadj, M. Desmadril, P. Minard, *Biochemistry* **2003**, *42*, 5674.
[15] A. Drevelle, M. Graille, B. Heyd, I. Sorel, N. Ulryck, F. Pecorari, M. Desmadril, H. van Tilbeurgh, P. Minard, *J. Mol. Biol.* **2006**, *358*, 455.
[16] E. Adjadj, E. Quiniou, J. Mispelter, V. Favaudon, J. M. Lhoste, *Eur. J. Biochem.* **1992**, *203*, 505.
[17] M. Nicaise, M. Valerio-Lepiniec, P. Minard, M. Desmadril, *Protein Sci.* **2004**, *13*, 1882.
[18] S. F. Betz, *Protein Sci.* **1993**, *2*, 1551.
[19] E. Ohage, B. Steipe, *J. Mol. Biol.* **1999**, *291*, 1119.
[20] U. Rothbauer, K. Zolghadr, S. Tillib, D. Nowak, L. Schermelleh, A. Gahl, N. Backmann, K. Conrath, S. Muyldermans, M. C. Cardoso, H. Leonhardt, *Nat. Methods* **2006**, *3*, 887.
[21] Y. Y. Wheeler, S. Y. Chen, D. C. Sane, *Mol. Ther.* **2003**, *8*, 355.
[22] R. E. Kontermann, *Methods* **2004**, *34*, 163.
[23] C. Frisch, H. Kolmar, A. Schmidt, G. Kleemann, A. Reinhardt, E. Pohl, I. Uson, T. R. Schneider, H. J. Fritz, *Fold Des.* **1996**, *1*, 431.
[24] Y. Liu, K. Breslauer, S. Anderson, *Biochemistry* **1997**, *36*, 5323.
[25] K. Proba, A. Worn, A. Honegger, A. Plückthun, *J. Mol. Biol.* **1998**, *275*, 245.
[26] D. W. Colby, Y. Chu, J. P. Cassady, M. Duennwald, H. Zazulak, J. M. Webster, A. Messer, S. Lindquist, V. M. Ingram, K. D. Wittrup, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 17616.
[27] I. Kather, C. A. Bippes, F. X. Schmid, *J. Mol. Biol.* **2005**, *354*, 666.
[28] K. H. Kim, B. M. Kwon, A. G. Myers, D. C. Rees, *Science* **1993**, *262*, 1042.
[29] L. S. Kappen, I. H. Goldberg, *Biochemistry* **1980**, *19*, 4786.
[30] Y. Bai, H. Feng, *Eur. J. Biochem.* **2004**, *271*, 1609.
[31] G. S. Waldo, *Curr. Opin. Chem. Biol.* **2003**, *7*, 33.
[32] H. Bothmann, A. Plückthun, *Nat. Biotechnol.* **1998**, *16*, 376.
[33] H. Bothmann, A. Plückthun, *J. Biol. Chem.* **2000**, *275*, 17100.
[34] R. H. Kennett, R. Leunk, B. Meyer, V. Silenzio, *J. Immunol. Methods* **1985**, *85*, 169.
[35] T. Cornvik, S. L. Dahlroth, A. Magnusdottir, M. D. Herman, R. Knaust, M. Ekberg, P. Nordlund, *Nat. Methods* **2005**, *2*, 507.
[36] S. L. Dahlroth, P. Nordlund, T. Cornvik, *Nat. Protoc.* **2006**, *1*, 253.
[37] A. Perczel, I. Jakli, B. M. Foxman, G. D. Fasman, *Biopolymers* **1996**, *38*, 723.
[38] G. C. Sudhahar, K. Balamurugan, D. H. Chin, *J. Biol. Chem.* **2000**, *275*, 39900.
[39] E. Freire, *Methods Enzymol.* **1995**, *259*, 144.
[40] N. Izadi-Pruneyre, E. Quiniou, Y. Blouquit, J. Perez, P. Minard, M. Desmadril, J. Mispelter, E. Adjadj, *Protein Sci.* **2001**, *10*, 2228.
[41] D. Pal, P. Chakrabarti, *J. Biomol. Struct. Dyn.* **2001**, *19*, 115.
[42] S. Caddick, F. W. Muskett, R. G. Stoneman, D. N. Woolfson, *J. Am. Chem. Soc.* **2006**, *128*, 4204.
[43] F. Baneyx, M. Mujacic, *Nat. Biotechnol.* **2004**, *22*, 1399.
[44] T. Wiseman, S. Williston, J. F. Brandts, L. N. Lin, *Anal. Biochem.* **1989**, *179*, 131.
[45] P. L. Privalov, S. A. Potekhin, *Methods Enzymol.* **1986**, *131*, 4.
[46] A. Sali, T. L. Blundell, *J. Mol. Biol.* **1993**, *234*, 779.
[47] M. Schmidt Am Busch, A. Lopes, D. Mignon, T. Simonson, *J. Comput. Chem.* **2008**, *29*, 1092.
[48] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, M. Karplus, *J. Comput. Chem.* **1983**, *4*, 187.
[49] A. D. MacKerell, Jr., M. Feig, C. L. Brooks, 3rd, *J. Am. Chem. Soc.* **2004**, *126*, 698.
[50] E. Hajjar, D. Perahia, H. Debat, C. Nespoulous, C. H. Robert, *J. Biol. Chem.* **2006**, *281*, 29929.
[51] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, *14*, 33.

Received: November 12, 2008
Published online on May 4, 2009

