Protein Engineering vol.13 no.7 pp.519–526, 2000

Functional expression and affinity selection of single-chain Cro by phage display: isolation of novel DNA-binding proteins

Mikael T.I.Nilsson[1], Michael C.Mossing[2] and Mikael Widersten[1,3]

[1]Department of Biochemistry, Uppsala University, Biomedical Center, Box 576, SE-751 23 Uppsala, Sweden and [2]Department of Chemistry, University of Mississippi, University, MS 38677, USA

[3]To whom correspondence should be addressed. E-mail: mikael.widersten@biokem.uu.se

A robust selection system affording phage display of the DNA-binding helix–turn–helix protein Cro is presented. The aim of the work was to construct an experimental system allowing for the construction and isolation of Cro-derived protein with new DNA-binding properties. A derivative of the phage λ Cro repressor, scCro8, in which the protein subunits had been covalently connected via a peptide linker was expressed in fusion with the gene 3 protein of *Escherichia coli* filamentous phage. The phage-displayed single-chain Cro was shown to retain the DNA binding properties of its wild-type Cro counterpart regarding DNA sequence specificity and binding affinity. A kinetic analysis revealed the rate constant of dissociation of the single-chain Cro-phage/DNA complex to be indistinguishable from that of the free single-chain Cro. Affinity selection using a biotinylated DNA with a target consensus operator sequence allowed for a 3000-fold enrichment of phages displaying single-chain Cro over control phages. The selection was based on entrapment of phage/DNA complexes formed in solution on streptavidin-coated paramagnetic beads. The expression system was subsequently used to isolate variant scCro8 proteins, mutated in their DNA-binding residues, that specifically recognized new, unnatural target DNA ligands.

Keywords: DNA binding/helix–turn–helix/mutagenesis/phage lambda

Introduction

Interactions between proteins and DNA are essential for biological life processes. Specificity and avidity in the binding interactions are dependent on the complementarity of the three-dimensional structures of both protein and nucleic acid in the complex formed. A large structural class of DNA-binding proteins rely on the helix–turn–helix (HTH) motif for specificity and binding strength (Harrison and Aggarwal, 1990; Wintjens and Rooman, 1996). Proteins in this heterogeneous group include transcriptional regulators from both prokaryotic and eukaryotic sources. Although much information from structural and biochemical studies has been gained concerning the prerequisites of HTH protein–DNA interactions, the understanding of the structural determinants responsible for sequence-specific and high-affinity protein–DNA complex formation is still incomplete and additional tools for the study of these issues would be of great importance. An effective approach used earlier in structure–property relationships employs *in vitro* evolution-based functional selection of variant proteins from large libraries generated by mutagenesis. Hence an expression system allowing for functional selection of mutated forms of HTH-proteins with altered properties would be a powerful tool in the analysis of the structural requirements involved in sequence-specific and avid DNA binding.

The reported functional selection of small zinc finger containing DNA-binding proteins has addressed these issues by employing *in vitro* evolution techniques. Zinc finger proteins, mutated in their DNA-interacting amino acid residues, have been expressed by phage display (Rebar and Pabo, 1994; Choo and Klug, 1995; Wu *et al.*, 1995) and challenged with 'new' target DNA sequences. Complexes formed between zinc finger variants and target DNAs have been isolated and the protein structures analyzed. From such analyses, clues revealing the structural relationship between protein and their target DNA sequences have emerged (Choo and Klug, 1994; Rebar and Pabo, 1994; Isalan *et al.*, 1998). Applying the gathered information, recombinant zinc finger proteins approaching predefined DNA binding specificities have been generated (Choo *et al.*, 1994, 1997; Beerli *et al.*, 1998; Pomerantz *et al.*, 1998).

In this work, an experimental system for the functional selection of DNA binding HTH proteins, based on the Cro repressor of bacteriophage λ (Johnson *et al.*, 1981), is presented. The selection for site-specific DNA–protein interactions is based on surface display of Cro-derived proteins, expressed in fusion with the C-terminal domain of the gene 3 protein (g3p) of filamentous *Escherichia coli* phage (Smith, 1985).

Materials and methods

Construction of phage display expression systems

Wild-type λ *cro* gene was assembled through three sequential PCRs with six overlapping oligonucleotides, denoted Cro1–Cro6 (Table I). The resulting *cro* gene was subcloned into the *Xho*I/*Spe*I sites of pGTac5H (Widersten, 1998) and the resulting plasmid was denoted pGTacCro5H. The construction of a *cro* gene derivative, encoding a monomeric form of the protein, was performed by insertion of a 15 bp fragment in the wild-type gene between bp 168 and 169 according to Mossing and Sauer (1990). The insertion was performed by PCR using pGTacCro5H as template, together with primers Cro1 and mono6DG (Table I). The PCR product was subcloned into pC3ΔNX (Widersten and Mannervik, 1995), resulting in pC3ΔNXCro.mDG. A redundant *Acc*65I site in pC3ΔNXCro.mDG was removed by Klenow fill-in of the 5′ protruding ends after *Acc*65I digestion of pC3ΔNXCro.mDG. The resulting plasmid was denoted pC3ΔACro.mDG. Plasmid pscCro8 (Jana *et al.*, 1998) was digested with *Bam*HI. The vector fragment with the partially deleted single-chain *cro* fragment of 222 bp was religated to result in plasmid pscCroΔBB. *Xho*I and *Spe*I restriction sites were inserted at the 5′- and 3′-ends, respectively, of the *cro* gene by PCR using pscCroΔBB as template together with primers Cro1 and

Table I. Primers used to PCR amplify gene fragments, introduce mutations and restriction sites and generate dsDNA ligands

Name	Deoxyoligonucleotide sequence
Cro1	5′ TTTTTTGAATTCCTCGAGATGGAACAACGCATAACCCTGAAAGATTATGCAATGCGCTTT
Cro2	5′ ATATACGCCGAGATCTTTAGCTGTCTTGGTTTGCCCAAAGCGCATTGCATAATCTTT
Cro3	5′ CTAAAGATCTCGGCGTATATCAAAGCGCGATCAACAAGGCCATTCATGCAGGCCGAAAGA
Cro4	5′ ATACAGATCCATCAGCGTTTATAGTTAAAAAAATCTTTCGGCCTGCATGAAT
Cro5	5′ AAACGCTGATGGATCTGTATATGCGGAAGAGGTAAAGCCCTTCCCGAGTAACAAAA
Cro6	5′ TTTTTTTCTAGATTAACTAGTTGCTGTTGTTTTTTTGTTACTCGGGAAGGG
mono6DG	5′ CCCCAAGCTTTTAACTAGTTGCTGTTGTTTTTTTGTTACTCGGGAAGGGTTTAACTTCACCATCCTTTACCTCTTCCGCATATA
BlnIfor	5′ GCTAAAGACCTAGGCGTGTATCAAAGCGCG
BlnIrev	5′ CACGCCTAGGTCTTTAGCTGTCTTGGTTTG
ORC+B	5′[a] B-GTGTGTGTATCACCGCGGGTGATAGT
ORC+	5′ TGTATCACCGCGGGTGATAGT
ORC–	5′ ACTATCACCCGCGGTGATACA
#11+B	5′ B-GTGTGTCCGTCACCGCCAGTTAATCT
#11+	5′ TCCGTCACCGCCAGTTAATCT
#11–	5′ AGATTAACTGGCGGTGACGGA
ORas1+B	5′ B-GTGTGTGATACCAAGCGGGTGATAGT
ORas1–1	5′ ACTATCACCCGCTTGGTATCA
ORas1–2	5′ ACTATCACCCGCTTII[b]TATCA
Cro8Lrev	5′ GCCGCCAGAGCCACC
Cro9mut	5′ TAAAGATCTCGGCGTGTATNNS[c]NNSNNSATCNNSNNSGCCATCCATGCCGGCC
pUCfor	5′ GTAAAACGACGGCCAGTG
pUCrev	5′ CAGGAAACAGCTATGACC
scCro-L	5′ TATGCGTTGTTCCATGCCGCC

[a]B denotes 5′ incorporated biotin.
[b]I denotes inosine.
[c]N = G, A, T or C; S = G or C.

Cro6 giving the PCR product CroΔBB. The *cro* fragment of CroΔBB was subcloned into the *Xho*I and *Spe*I sites of pC3ΔNX-A yielding pC3CroΔBB. The 222 bp *Bgl*II containing fragment from the *Bam*HI digested pscCro8 was inserted into pUC18. A consecutive overlapping PCR was performed deleting the *Bgl*II site and inserting a *Bln*I site into the gene. The PCR was divided into two steps using two mutation primers BlnIrev and BlnIfor (Table I). (i) Two parallel PCRs were performed both using pCC8 as template, where the first reaction contained BlnIrev primer together with the pUC18 specific pUCfor primer (Table I) and the second reaction contained BlnIfor primer together with pUCrev primer (Table I). (ii) A mixture of the two partially complementary PCR products was then used as template for the last amplification reaction together with primers pUCfor and pUCrev. The final product, containing the 325 bp single-chain *cro* fragment, was denoted C8Bln. *Bam*HI-digested C8Bln was ligated into *Bam*HI-digested and dephosphorylated pC3CroΔBB. The resulting pC3scCro8 plasmid (Figure 1a) was then transferred into *E.coli* XL1-Blue (Stratagene, La Jolla, CA). All polymerase chain reactions and cloning steps were performed according to standard procedures (Sambrook *et al.*, 1989). All described gene constructs were verified by sequencing (Sanger *et al.*, 1977).

DNA binding parameters for scCro8-g3p phage

Phages displaying scCro8 or Cro.mDG fused to g3p were prepared by culturing *E.coli* XL1-Blue cells containing the pC3scCro8 or pC3ΔNXCro.mDG phagemid, respectively, essentially as described earlier (Widersten and Mannervik, 1995). Phages harvested after proliferation were redissolved in phage resuspension buffer (PRB) [10 mM Tris–HCl, pH 7.4, 100 mM KCl, 1% (w/v) bovine serum albumin (BSA) and 0.02% (w/v) NaN_3]. A final centrifugation at 16 000 *g* was performed and the phage-containing supernatants were collected and used immediately or stored at 4°C.

DNA ligands were obtained by annealing the complementary oligonucleotides ORC+ with ORC– or #11+ with #11– (Table I). Single-chain Cro8 phages were incubated together with ^{32}P-labeled ORC DNA ligand in binding buffer (BB) (10 mM Tris–HCl, pH 7.4, 100 mM KCl) fortified with 5 mg/ml BSA and 0.8 mM EDTA on ice for 30 min in a volume of 100 µl. Samples (95 µl) of the mixture were removed and vacuum filtered through a nitrocellulose membrane using a dotblot minifold system (Kim *et al.*, 1987). The membrane was subsequently washed with 200 µl of ice-cold BB and captured radioactivity was detected by liquid scintillation counting and expressed as counts per minute (c.p.m.). The titrations were performed at both constant DNA concentration (1 nM) with a varied phage concentration (1.5 pM–1.3 nM) or at constant phage concentration [~1 nM colony-forming units (cfu)] with a varied DNA concentration (1.0 pM–2.0 nM).

The kinetic dissociation rate for the scCro8-g3p–ORC complex was determined as follows: scCro8 phages (~1 nM) and 0.4 nM ^{32}P-labeled ORC DNA in BB containing 5 mg/ml BSA and 0.1 mM EDTA were mixed and equilibrated on ice for 60 min in a volume of 2.5 ml. At different time points, 310 µl of the mixture were removed and mixed with 10 µl of 8 µM unlabeled ORC DNA giving a final molar excess of 600-fold of unlabeled over labeled ligand DNA. At different time points, samples of 95 µl were withdrawn and filtered through a nitrocellulose membrane by vacuum. The filter was washed and captured radioactivity was detected as described above. The dissociation rate constant, k_{off}, was estimated by fitting a single exponential decay equation, $<f(t) = A\exp(-k_{off}t) + C>$, to the collected data using SIMFIT (Bardsley *et al.*, 1995).

DNA affinity selection of scCro8-g3p phage

Biotinylated ligand DNA was obtained by annealing oligo ORC+B with ORC– and #11+B with #11– (Table I). Approximately 10^9 cfu of Cro.mDG-g3p or scCro8-g3p phages were

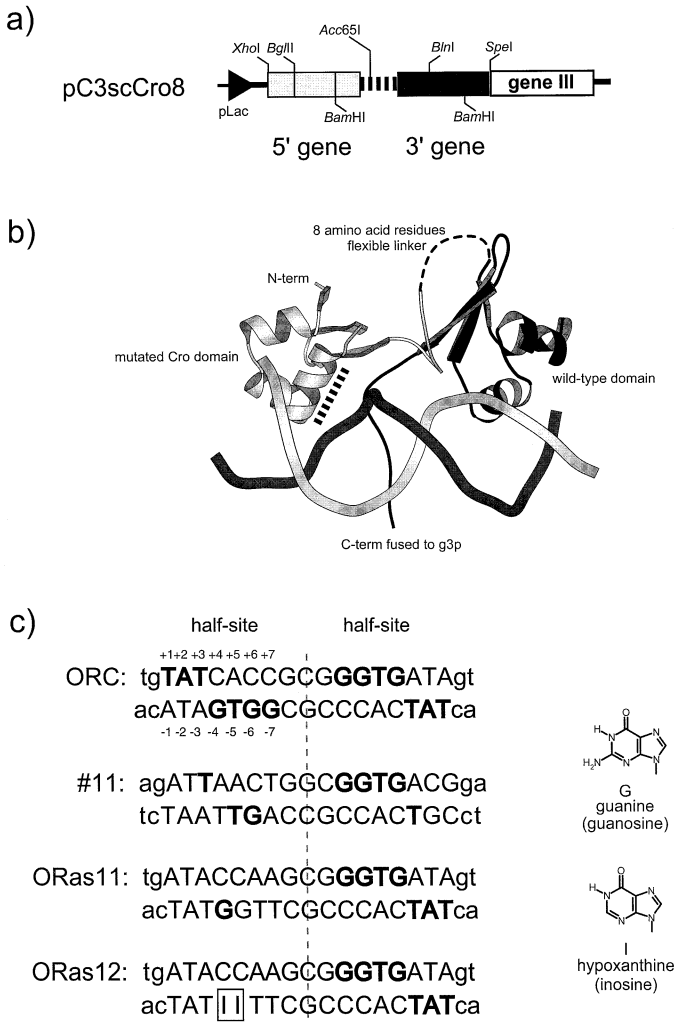

Fig. 1. Gene construct, protein and ligand structures. (**a**) Gene construct for phage display of scCro8. Two *cro* genes, light gray and black boxes, were inserted in tandem downstream of the *E.coli lac* operon promoter and in fusion with the 3′ portion of the phage M13 gene 3. The construct allowed for display of single-chain Cro8 protein on the phage surface. The *cro* genes were modified to facilitate cassette mutagenesis of codons for putative DNA binding amino acid residues. See the Materials and methods section for details on the construction. (**b**) Molecular model of single-chain Cro8. The N-terminal domain (light gray) is covalently linked through a short peptide linker (position indicated by a dashed line) to the C-terminal domain (black). The C-terminus of scCro8 is in turn coupled to the C-terminal fragment of the gene 3 protein of phage M13. The region of the third α-helix of the N-terminal domain (red), facing the major groove of DNA in the molecular model which was randomly mutated to construct a library of variant 'heterodimeric' scCro8 phages, is indicated by a dashed black line. The model was constructed with Molscript 2.1 (Kraulis, 1991) using atomic coordinates for wild-type Cro in complex with DNA, 4CRO (Brennan *et al.*, 1986). (**c**) DNA ligands used in affinity selection of phage-displayed scCro8 and constructed variants thereof. ORC is an almost perfect palindromic sequence bound tightly and specifically by the 'wild-type' scCro8. #11 is a DNA which is bound poorly by scCro8 (Takeda *et al.*, 1992). The ORas11 and ORas12 ligands were used to select for mutant 'heterodimeric' forms of scCro8. ORas12 contains guanosine to inosine replacements in positions –4 and –5. The structure of the relevant bases guanine and hypoxanthine is also shown.

incubated with 1 μM biotinylated DNA and 1 μM unbiotinylated DNA in BB with 0.5 mg/ml BSA in a final volume of 100 μl for 1 h at 10°C. The mixtures of phages and DNA ligands were added to 0.1 mg of streptavidin-coated paramagnetic beads (Promega, Madison, WI) previously washed three times in 500 μl of PRB. The phage–DNA–bead mixtures were incubated for an additional 30 min before the beads were collected and washed five times with 500 μl of ice-cold BB. In the last wash the beads were transferred to a clean tube and the supernatant was removed. Bound phages were eluted with 100 μl of elution buffer (EB) (90 mM glycine–HCl, pH 2.2, 3 mg/ml BSA) for 10 min at room temperature. Eluted phages were transferred to a 10 ml sterile plastic tube. The beads were then subjected to an additional 50 μl of EB and the two eluates were pooled and neutralized with Tris base. Eluted phages were allowed to infect 5 ml of log phase *E.coli* XL1-Blue, grown in 2TY with 10 μg/ml tetracycline, for 15 min at room temperature. Infected, phagemid-containing cells were titered on 100 μg/ml ampicillin-containing LB plates [1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, 1.5% (w/v) bacto agar].

scCro8 phage enrichment by affinity selection

Approximately 10^{10} cfu of Cro.mDG phages, diluted with scCro8 phages, mixed at a ratio of one scCro8 phage per 12 000 Cro.mDG phages, was subjected to a single selection cycle in the presence of 1 μM biotinylated ORC and 1 μM unbiotinylated #11 ligands. Selection was performed as described above. The number of phages in the population before and after the affinity selection with ORC DNA was titered and the identification of scCro8 phagemid containing bacterial clones was determined by colony hybridization (Sambrook *et al.*, 1989) using a ^{32}P-labeled scCro8-specific oligo, scCro-L (Table I) as probe.

Selection of scCro8 derivatives with new DNA binding properties

Library construction. A library of scCro8 mutant genes was constructed by PCR using oligos Cro9mut and Cro8Lrev (Table I). The PCR product was inserted into pC3scCro8 digested with *Bgl*II and *Acc*65I. The library ligation mixture was transferred into *E.coli* XL1-Blue by electroporation. Propagation and preparation of phages displaying scCro8 derivatives were performed as described previously (Widersten and Mannervik, 1995). The library size was estimated by titration of ampicillin-resistant transformants and the quality of the library was assessed by DNA sequencing of 40 randomly picked library clones.

Affinity selection. The phage scCro8 variant library (5×10^{11} cfu) was incubated with 1 μM biotinylated ORas11 or ORas12 (from annealing of ORas1+B and ORas1–1 or ORas1+B and ORas1–2, respectively, Table I) in 100 μg/ml sonicated salmon sperm DNA, 10 mM Tris–HCl, pH 7.4, 100 mM KCl, 0.5% (w/v) BSA and 0.01% (w/v) NaN_3, in a final volume of 100 μl for 1 h at 10°C. The phage–DNA mixtures were transferred to 0.1 mg of streptavidin-coated paramagnetic beads, prewashed with PRB and incubated for an additional 30 min at the same temperature. The beads were captured and washed twice with 500 μl of ice cold BB. Phage elution, titration and propagation were performed as described above. The selection procedure was repeated for three additional rounds with the modifications that captured phages were washed five times in subsequent rounds. The enrichment during selection was estimated from the ratio of the titered number of output phages divided by the number of input phages. After completed selection, the respective scCro8 genes of randomly picked phagemid-carrying colonies from round four were sequenced.

DNA binding properties.
A chosen number of identified scCro8-g3p variants were tested for their DNA binding properties in a single selection round as described above. The scCro8-g3p variant phages were mixed with 1 μM of biotinylated DNA (ORC or ORas11 or ORas12) in the presence of 100 μg/ml sonicated salmon sperm DNA. Wild-type scCro8-g3p phages were included for comparison. The amount of protein–DNA binding was estimated by dividing the phages out/phages in ratios for the different DNA ligands.

Results and discussion

Construction of expression system for phage display of Cro

The main aim of the present work was to construct a robust expression/selection system that allowed for an efficient construction and isolation of HTH proteins, homo- or heterodimeric, that possessed new DNA binding specificitites. The Cro repressor of phage λ binds its target operator DNA as a homodimer, where the dimeric form of the protein is crucial to ensure sequence specificity and tight binding (Jana *et al.*, 1997; Albright and Matthews, 1998). As a consequence of the homodimeric nature of Cro, target sequences that are efficiently recognized and bound by Cro display dyad symmetry. If given the possibility to construct stable heterodimeric Cro variants there is, however, no obvious reason why non-palindromic DNA sequences should not classify as good target ligands.

In order to achieve both an efficient selection of a large number of Cro mutants and the possibility to construct heterodimeric proteins, phage display expression of a single-chain derivative of Cro was utilized. In this Cro derivative (scCro8), an eight-residue peptide linker covalently connects the two Cro subunits (Jana *et al.*, 1998). The amino acid linkage between the two subunits prevents dimer dissociation, resulting in a reduction of the half-site saturation concentration by 250-fold as compared with the wild-type Cro (Jana *et al.*, 1998). The linkage also eliminates subunit interchange between Cro dimers, which is an advantage if aiming for the construction of synthetic Cro-like heterodimeric proteins. The expression plasmid and a model of the expressed fusion protein is shown in Figure 1a and b. To facilitate separate structural manipulation of the two HTH motifs in scCro8, a unique restriction site (*Bln*I) was inserted by silent mutagenesis, adjacent to the gene sequence encoding the DNA-binding α-helix of the HTH motif of the C-terminal domain of scCro8. This modification of the scCro8 gene and its fusion to gene 3 resulted in an expression system designed to allow for phage display of the native scCro8 protein, to facilitate cassette mutagenesis of scCro8, and the generation of heterodimeric proteins. Such heterodimeric Cro-like proteins where the HTH motifs of the subunits can be manipulated independently of each other may be employed to increase the recognition repertoire of possible target DNA sequences to include also non-palindromic target sequences.

Functional expression of phage-displayed scCro8–DNA binding properties

It is well established that wild-type Cro and variants thereof have high affinity for the O_R operator consensus DNA (ORC) sequence, while binding to non-operator DNA sequences, such as the '#11' DNA (Figure 1c) (non-specific binding) occurs with approximately 10^5-fold lower affinity (cf. Harrison and Aggarwal, 1990; Takeda *et al.*, 1992). Titration experiments varying the phage concentration at constant DNA concentration (Figure 2a) or varying the DNA concentration at constant concentration of scCro8 phage (Figure 2b) indicated a stoichiometric titration of the non-varied component in the two binding reactions (Riggs *et al.*, 1970). This is probably due to a K_D of the scCro8-g3p phage–DNA complex considerably lower than 1 nM, the concentration of the non-varied binding component (i.e. DNA ligand or scCro8-g3p phage). The K_D for the complex between free scCro8 and a DNA closely similar to ORC has previously been determined from gel-shift assays to be 4 pM (Jana *et al.*, 1998). The high affinity in the complex between phage-fused scCro8 and ORC DNA complicates an accurate determination of K_D by the filter assay applied and hence a value of the dissociation constant was not determined. Alternative methods for K_D determination, such as electrophoretic mobility shift or surface plasmon resonance assays, are badly suited for proteins displayed on filamentous phage owing to the physico-chemical properties (e.g. size) of the phage particles. When the titration was repeated using a target DNA sequence, denoted #11 (Figure 1c), for which Cro displays a low affinity (Takeda *et al.*, 1992), no radioactivity was captured in the filter assay. In addition, no radioactivity was captured when titrations were repeated with either ORC or #11 DNA together with phages displaying the low affinity binder Cro.mDG (data not shown).

In a separate experiment, the kinetic dissociation rate constant, k_{off}, for the scCro8 phage–DNA complex was also determined (Figure 2c). The value obtained, $k_{off} = 0.029 \pm 0.018$ min^{-1}, translates to a scCro8 phage–DNA complex half-life ($t_{1/2}$) of 24 ± 14 min. This value is in good agreement with earlier published data for the half-life of scCro8 and its target DNA of 23 ± 4 min (Jana *et al.*, 1998). The identical dissociation rates of the scCro8-phage–DNA and the scCro8–DNA complexes suggest that the affinity between scCro8 and its ORC target DNA, and also the sequence recognition specificity, are unaltered when the protein is displayed on the phage surface.

DNA affinity selection of scCro8-g3p phage

In order to function in a phage display-based affinity selection, scCro8-g3p must retain its Cro-like DNA-binding properties regarding sequence specificity and avid binding of target DNA sequences under the special conditions used. To investigate this, complexes between biotinylated DNA and the Cro variants scCro8 and Cro.mDG displayed on phage, formed in solution, were then immobilized by capturing on streptavidin-coated paramagnetic beads. The beads were subsequently washed and bound phages were eluted and titered. The results showed that scCro8-g3p phages were trapped with 1700-fold higher efficiency by biotinylated ORC DNA ligand (1:3000 of added phages were recovered) than by biotinylated #11 (1:5×10^6 of added phages were recovered). In the case of the low-affinity Cro.mDG-g3p phage construct, the ratios were in the non-specific range with values of 1: 4×10^6 for ORC and 1:1×10^6 for #11. The observed efficiency by which scCro8-g3p phages were recovered by affinity selection with biotinylated ORC-DNA clearly demonstrates that neither fusion with g3p nor the surface of the magnetic beads used during the affinity selection interferes with the high affinity or selectivity of scCro8 towards its natural operator consensus target.

scCro8 phage enrichment by affinity selection

To test if the apparently efficient capturing of scCro8-g3p phages was sufficient to ensure strong clonal amplification, scCro8-g3p phages were diluted 1:12 000 with weakly binding

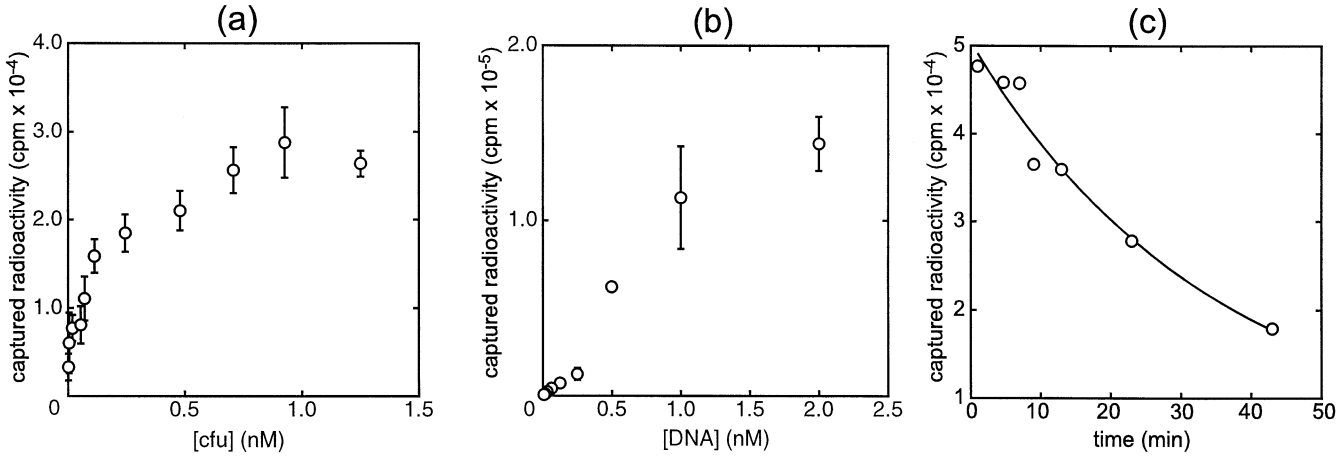

Fig. 2. DNA binding properties of scCro8-g3p phages. Formation of scCro8-g3p–[^{32}P]ORC complex was determined by filter assay and liquid scintillation counting. Titration of complex formation was performed under conditions of (**a**) constant DNA concentration (1 nM) and varying phage concentration from 1.5 pM to 1.3 nM or (**b**) constant phage concentration (~1 nM cfu) and varying DNA concentration from 1.0 pM to 2.0 nM. Bars indicate standard deviation, $n \geqslant 3$. (c) The decay in the concentration of the scCro8-g3p–[^{32}P]ORC complex was monitored by filter assay after addition of a 650-fold excess of unlabeled ORC to an equilibrium mixture of scCro8-g3p and [^{32}P]ORC. The data obtained were fitted to a single exponential decay equation to determine the dissociation rate constant k_{off}.

Table II. Affinity selection using ORC as biotinylated ligand in a mixture of scCro8-g3p and Cro.mDG-g3p phages

	scCro8 phages (cfu)	Total phages (cfu)	scCro8 ratio (scCro8/total cfu)	Positive hybridization/ total No. of clones	scCro8 phage enrichment (ratio out/ratio in)
In	1.9×10^{6}	2.3×10^{10}	8.3×10^{-5}	1/5009[c]	–
Out	3.4×10^{3}[a]	1.4×10^{4}	0.24[b]	332/1377	2921

[a]The number of scCro8 phages out was estimated by multiplying total number of phages out with the scCro8 ratio out.
[b]scCro8 ratio out was calculated by dividing the number of positively hybridized clones with the number of clones examined.
[c]The probability of occurrence of one or more scCro8 clones in a sample of this size is 0.34.

Cro.mDG-g3p phages. After a single selection cycle, the ratio of scCro8 phages had increased to 24% of all phages recovered from the selection (Table II). This corresponds to an almost 3000-fold amplification of a high-affinity binder such as scCro8 during a single selection cycle from a mixture containing a large excess of low-affinity binders. An amplification rate this high implies that it would be possible to increase the fraction of one high-affinity binding protein from a library of 5×10^{6} individual members to represent 40% of the total protein population within two rounds of affinity selection. This high amplification rate, however, may not extrapolate to a more complex situation where a library of mutant proteins would represent a mixture of low-, moderate- and high-affinity binders together. The law of mass action would interfere and the amplification rate of high-affinity binding protein would probably decrease with an increasing number of moderate binders. Nevertheless, the amplification rate observed here is expected to suffice well for affinity selection of high-affinity/specificity binders from a mixture of scCro8 variants.

Selection of scCro8 derivatives with new DNA binding properties

Library construction. For the isolation of variant Cro proteins with new DNA binding properties, a library of randomly mutated scCro8 genes was constructed by an NNS strategy. The five codons chosen for mutagenesis, 27–29 and 31–32, all encode amino acid residues situated in helix 3 of the N-terminal domain of the translated scCro8 protein. The residues chosen, Q27, S28, A29, N31 and K32, have been established by earlier studies to be important for sequence-specific recognition of the natural operator target (Eisenbeis *et al.*, 1985; Pakula *et al.*, 1986; Mossing and Sauer, 1990). Q27, S28, N31 and K32 all make base-specific contacts through their respective side-chain functional groups whereas A29 interacts with DNA through its backbone amide.

To facilitate the isolation of Cro proteins with new binding properties, a repertoire of variant proteins with a basal albeit low-affinity binding was considered to be a better starting point than to attempt the isolation of, presumably very few, functional variants hidden in a large pool of non-functional Cro mutants. Therefore, mutagenesis was restricted to only one of the two fused *cro* genes, resulting in a chimeric construct consisting of a 5′-gene carrying random mutations in codons 27–29 and 31–32, fused to a 3′-gene encoding a wild-type Cro subunit (Figure 1a). The proteins expressed from the chimeric genes were therefore expected to be heterodimeric in nature with the N-terminal domain carrying a mutated helix-3 and a wild-type C-terminal domain (Figure 1b). The final library of mutant scCro8-phage was titered to contain 5×10^{7} individual clones, which translates to ~1.5-fold over-representation of all possible gene constructs ($32^{5} \approx 3.3\times10^{7}$). A number of clones ($n = 40$) from the library were analyzed by sequencing and were all identified as distinct variants only carrying the expected mutations inserted by the mutagenesis.

Affinity selection. The scCro8 variant was challenged for binding to the DNA targets ORas11 and ORas12 (Figure 1c). These 21-mer dsDNA ligands consist of one preserved ORC-like half-site, expected to interact with the wild-type C-terminal domain. The other half-sites of the target DNA molecules

Table III. Ratios of ratios of output/input phages for different DNA ligands from single selection rounds

DNA ligands compared	Cro protein phage								
	scCro8	11:401	11:405	11:416	11:417	12:401	12:406	12:407	12:416
	Deduced amino acid sequence[a]								
	QSANK	EEEKW	EGTQN	KGTQK	RRCQR	TQCDY	RECTY	DQCTY	RECKY
ORas11/ORC	0.01	0.6	0.07	0.5	0.02	–	–	–	–
ORas11/ORas12	0.2	2.3	0.1	5.7	0.7	–	–	–	–
ORas12/ORC	0.06	–	–	–	–	2.2	0.4	3.8	0.5
ORas12/ORas11	5.6	–	–	–	–	5.9	2.0	3.6	2.0

[a]The sequence denotes the one-letter code of the residues mutated in scCro: Q, Gln27; S, Ser28; A, Ala29; N, Asn31; K, Lys32.

Table IV. Enrichment of variant scCro8-g3p phages during affinity selection with ORas11 or ORas12 DNA

Selection round	Input phages (cfu)		Output phages (cfu)		Ratio (cfu out/cfu in)		Enrichment (-fold)	
	ORas 11	ORas 12	ORas 11	ORas 12	ORas 11	ORas 12	ORas 11	ORas 12
1	2.4×10^{12}	2.4×10^{12}	2.4×10^{4}	1.9×10^{4}	9.9×10^{-9}	8.0×10^{-9}	(1)	(1)
2	5.4×10^{11}	7.1×10^{11}	2.7×10^{5}	2.1×10^{5}	5.1×10^{-7}	2.9×10^{-7}	52	36
3	4.2×10^{11}	1.1×10^{11}	5.8×10^{5}	3.0×10^{5}	1.4×10^{-6}	2.8×10^{-6}	141	350
4	1.2×10^{10}	9.0×10^{9}	2.1×10^{5}	2.1×10^{5}	1.7×10^{-5}	2.3×10^{-5}	1717	2875

differ in six out of seven (ORas11) or seven out of seven (ORas12) positions of the operator half-site known to affect Cro binding. These 'unnatural' half-sites were aimed to select for mutated subunits of the proteins expressed in the library, subunits promoting the selection for high-affinity binding of the phage-displayed heterodimeric variants.

To ensure that non-mutated scCro8 phages bound the ORas molecules with a lower affinity as compared with ORC, wild-type scCro8-g3p phages were assayed for binding to biotinylated ORas11 and ORas12 by one round of affinity selection. The ratios observed between output/input phages dropped about 100-fold (ORas11) and 17-fold (ORas12), respectively, as compared with the ratio obtained when using biotinylated ORC as the target ligand (Table III), demonstrating an anticipated lower affinity in the scCro8-g3p phage–ORas DNA complexes. The only structural difference between ORas11 and ORas12 is the substitution of guanosine for inosine in positions –4 and –5 (Figure 1c), removing the amino group bound to C-6 of guanine. This amino group is not, as deduced from structural data, expected to interact with residues of bound Cro. The change from guanosine to inosine could, however, influence the dynamic properties of the target DNA duplex since the hypoxanthine base only allows for two Watson–Crick base pairs with the complementary cytidine. The affinity selection was continued for four rounds and the clonal enrichment was assayed by titration of output phages/input phages. After four selection rounds, an approximately 2000-fold enrichment was reached with either ligand (Table IV). An increase in the output/input ratio is an indication of successful selection, since library members with adequate binding properties are amplified in the phage population.

Structural analysis. When randomly sampled clones from round four were analyzed at gene level, a selection for certain structural motifs was observed (Table V). The majority of clones selected for with the ORas11 ligand showed an over-representation of glutamic acid/glutamate residues in positions 27–29, with the dominating variant having the deduced amino acid sequence E27, E28, E29, K31, W32. In some instances clones selected for binding to the ORas12 ligand also contained glutamic acid/glutamate in either position 27, 28 or 29. The dominating feature of these variants, however, was the selection for Cys in position 29 accompanied by a Tyr residue at position 32. This X27, X28, C29, X31, Y32 motif was present in 11 out of 22 clones analyzed. Interestingly, in spite of the high similarity between the two ORas ligands, at no time was the same clone found to be selected for simultaneously with the different ligands. This may be an indication that the inosine-containing ORas12 is not equally well recognized by proteins with affinity for the guanosine-containing ORas11.

DNA binding properties. The ability of the isolated variant scCro8 proteins to recognize selectively and bind the target DNA to which it had been selected for was assayed by one round of affinity selection against either biotinylated ORas11, ORas12 or ORC DNA. The results are shown in Table III. Since one half-site of the target DNA duplexes used is identical in ORas11, ORas12 and ORC a certain amount of binding affinity from the interactions made between this half-site and the wild-type C-terminal domain of the scCro8 mutants was expected. Hence a strong preference for one ligand over the other in the isolated proteins was unexpected and not found. All variants tested, except mutant Mutant 11:417 showed improved binding towards the respective ORas ligand as compared with the parental scCro8 protein-phage, demonstrating a successful selection for affinity. Mutant 11:417 contained three arginine residues which may be responsible for tight binding through ionic interactions with negatively charged phosphate groups of the DNA backbone. Such interactions are independent of base composition, which may explain the absence of improved sequence specificity in this variant. However, mutants 12:401 and, in particular 12:407, both

Table V. Deduced amino acid sequence of mutated residues in scCro8-g3p phage variants isolated after four rounds of selection with ORas11 or ORas12 DNA

DNA		Residue position in scCro8 variant				
	Clone No.	27	28	29	31	32
	wt	Q	S	A	N	K
ORas11	401, 402, 404, 406, 408, 409, 410, 412, 413, 414	E	E	E	K	W
	405	E	G	T	Q	N
	411	E	M	Q	S	K
	407	R	R	E	?	H
	417	R	R	C	Q	R
	416	K	G	T	Q	K
	403	P	T	Q	A	R
ORas12	407, 408, 413, 414	D	Q	C	T	Y
	401, 425	T	Q	C	D	Y
	405	A	Q	C	Q	Y
	403, 416	R	E	C	K	Y
	406	R	E	C	T	Y
	421	S	G	C	R	Y
	404	R	?	C	K	G
	402	R	C	S	A	A
	410	P	C	R	K	R
	411	E	N	T	T	Q
	418	E	N	T	A	R
	415	D	W	Q	S	R
	419	P	W	S	R	C
	420	R	T	D	G	K
	422	D	T	D	R	R
	423	K	I	Q	T	C
	424	R	E	E	N	N

selected for with the ORas12 ligand, had shifted selectivity to recognize ORas12 preferentially over ORC. This is a major alteration in specificity and implies that the mutated protein domain contributes directly to recognition and binding affinity. Interestingly, these particular variants were highly related in their deduced primary structures (Table V), sharing a T/D27, Q28, C29, T/D31, Y32 motif.

Protein–DNA recognition is dependent on structural complementarity in the formed binding complex. To an extent, HTH protein–DNA recognition is expected to rely on a 'linear' relationship between the primary structures of protein and DNA. A particular amino acid residue, at a specific position in the protein, interacts (preferentially) with a particular nucleotide, in a particular position in the target sequence. In addition to such a simplistic modeling of recognition, however, the tertiary structure of both protein and DNA in the formed complex (Albright and Matthews, 1998) together with entropic terms (Takeda *et al.*, 1992) are of importance for favorable specific binding. By enforcing *in vitro* evolution principles for the selection of HTH proteins with new DNA sequence specificities one cannot directly deduce which of these factors had driven the amplification of a particular protein. The variants isolated, however, will provide information about the rules for HTH protein recognition of DNA. These recognition rules probably involve both amino acid–base contacts and surface shape complementarity.

Evaluation of scCro8-g3p phage display system

The method of displaying proteins on phage particles by fusing them to the truncated minor coat protein 3 have been shown to be a powerful strategy for functional *in vitro* selection in large-sized protein libraries (for a review, see Dunn, 1996). The system presented here, employing the scCro8 protein, is the first to be described for functional phage display expression of an HTH DNA-binding protein. Two of the most important features regarding functional affinity selection have been verified: a scrupulous DNA sequence selectivity (1700-fold) for its high-affinity ligand and a 3000-fold clonal amplification during a single cycle of affinity selection. These properties are essential for successful scanning of libraries containing mutant forms displaying a range of different binding affinities for a specific DNA ligand. This assumption was verified by the isolation of variant scCro8 proteins mutated in the putative DNA-binding residues, that specifically recognized new target DNA ligands.

In conclusion, phage display of scCro8 is expected to be a suitable molecular model to explore the underlying mechanisms of interactions between the Cro protein and its natural target operator and also a powerful tool in the construction and isolation of HTH proteins with new and redesigned DNA binding properties.

Acknowledgements

We thank Dr Per Jemth, Uppsala University, for valuable aid in critically reviewing the manuscript. This work was supported by grants to M.W. from the Swedish Research Council for Engineering Science.

References

Albright,R.A. and Matthews,B.W. (1998) *J. Mol. Biol.*, **280**, 137–151.

Bardsley,W.G., Bukhari,N.A.J., Ferguson,M.W.J., Cachaza,J.A. and Burguillo,F.J. (1995) *Comput. Chem.*, **19**, 75–84.

Beerli,R.R., Segal,D.J., Dreier,B. and Barbas,C.F.,III (1998) *Proc. Natl Acad. Sci. USA*, **95**, 14628–14633.

Brennan,R.G., Takeda,Y., Kim,J., Anderson,W.F. and Matthews,B.W. (1986) *J. Mol. Biol.*, **188,** 115–118.

Choo,Y. and Klug,A. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 11163–11167.
Choo,Y. and Klug,A. (1995) *Curr. Opin. Biotechnol.*, **6**, 431–436.
Choo,Y., Sánchez-García,I. and Klug,A. (1994) *Nature*, **372**, 642–645.
Choo,Y., Castellanos,A., García-Hernández,B., Sánchez-García,I. and Klug,A. (1997) *J. Mol. Biol*., **273**, 525–532.
Dunn,I.S. (1996) *Curr. Opin. Biotechnol.*, **7**, 547–553.
Eisenbeis,S.J., Nasoff,M.S., Noble,S.A., Bracco,L.P., Dodds,D.R. and Caruthers,M.H. (1985) *Proc. Natl Acad. Sci. USA*, **82**, 1084–1088.
Harrison,S.C. and Aggarwal,A.K. (1990) *Annu. Rev. Biochem.*, **59**, 933–969.
Isalan,M., Klug,A. and Choo,Y. (1998) *Biochemistry*, **37**, 12026–12033.
Jana,R., Hazbun,T.R., Mollah,A.K.M.M. and Mossing,M.C. (1997) *J. Mol. Biol.*, **273**, 402–416.
Jana,R., Hazbun,T. R, Fields,J.D. and Mossing,M.C. (1998) *Biochemistry*, **37**, 6446–6455.
Johnson,A.D., Poteete,A.R., Lauer,G., Sauer,R.T., Ackers,G.K. and Ptashne,M. (1981) *Nature*, **294**, 217–223.
Kim,J.G., Takeda,Y., Matthews B.W. and Anderson,W.F. (1987) *J. Mol. Biol.*, **196**, 149–158.
Kraulis,P.J. (1991) *J. Appl. Crystallogr.*, **24**, 946–950.
Mossing,M.C. and Sauer,R.T. (1990) *Science*, **250**, 1712–1715.
Pakula,A.A., Young,V.B. and Sauer,R.T. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 8829–8833.
Pomerantz,J.L., Wolfe,S.A. and Pabo,C.O. (1998) *Biochemistry*, **37**, 995–970.
Rebar,E.J. and Pabo,C.O. (1994) *Science*, **263**, 671–673.
Riggs,A.D., Suzuki,H. and Bourgeois,S. (1970) *J. Mol. Biol*., **48**, 67–83.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning. A Laboratory Manual.* 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.
Smith,G.P. (1985) *Science*, **228**, 1315–1317.
Takeda,Y., Ross,P.D. and Mudd,C.P. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 8180–8184.
Widersten,M. (1998) *Protein Express. Purif*., **13**, 389–395.
Widersten,M. and Mannervik,B. (1995) *J. Mol. Biol*., **250**, 115–122.
Wintjens,R. and Rooman M. (1996) *J. Mol. Biol*., **262**, 294–313.
Wu,H., Yang,W.-P. and Barbas,C.F., III (1995) *Proc. Natl Acad. Sci. USA*, **92**, 344–348.

Received December 24, 1999; revised May 12, 2000; accepted May 17, 2000