

Switching DNA-binding specificity by unnatural amino acid substitution

Atanu Maiti and Siddhartha Roy^{1,*}

Department of Biophysics, Bose Institute, P-1/12, C.I.T. Scheme VII M, Kolkata 700 054, India and
¹Indian Institute of Chemical Biology, 4, Raja Subodh Mullick Road, Kolkata 700 032, India

Received July 21, 2005; Revised September 17, 2005; Accepted September 28, 2005

ABSTRACT

The specificity of protein–nucleic acid recognition is believed to originate largely from hydrogen bonding between protein polar atoms, primarily side-chain and polar atoms of nucleic acid bases. One way to design new nucleic acid binding proteins of novel specificity is by structure-guided alterations of the hydrogen bonding patterns of a nucleic acid–protein complex. We have used λ repressor of bacteriophage λ as a model system. In the λ -repressor–DNA complex, the ϵ -NH₂ group (hydrogen bond donor) of lysine-4 of λ -repressor forms hydrogen bonds with the amide carbonyl atom of asparagine-55 (acceptor) and the O₆ (acceptor) of CG6 of operator site O_L1. Substitution of lysine-4 (two donors) by iso-steric S-(2-hydroxyethyl)-cysteine (one donor and one acceptor), by site-directed mutagenesis and chemical modification, leads to switch of binding specificity of λ -repressor from C:G to T:A at position 6 of O_L1. This suggests that unnatural amino acid substitutions could be a simple way of generating nucleic acid binding proteins of altered specificity.

INTRODUCTION

A structural understanding of protein–nucleic acid interaction has advanced sufficiently for creation of engineered and artificially evolved DNA-binding proteins for some DNA sequences (1). Designed zinc finger proteins are now being applied for cell engineering and therapeutic purposes in some cases (2). However, a simple strategy for creating new nucleic acid binding proteins, in general, is still beyond reach. For example, an understanding of RNA recognition has not yet reached the sophistication needed for engineering RNA-binding proteins of novel specificity. A simple stereochemical rule for creating new recognition points on an existing protein or peptide scaffold would be of great benefit for designing

RNA-binding proteins of novel specificity and allow creation of more diversified DNA-binding protein modules as well.

Several studies have looked into alterations and relaxations of specificity by mutating critical residues (substitution with natural amino acids). Some of these effects are weak (3–6). However, in other cases a specificity difference of several orders of magnitude is seen (7–13). Ebright *et al.* (7), Wharton and Ptashne (8) and Youderian *et al.* (9), reported that a single amino acid change in the recognition helix changes the binding specificity for a mutant operator by several orders of magnitude. Wharton and Ptashne (10) demonstrated that by changing the recognition helix of 434 repressor to that of the P22 repressor, by introducing several mutations, it is possible to completely switch the binding specificity from 434 operator to that of P22 operator. In many cases, recognition specificity is based on hydrogen bonding between amino acid side chains and nucleic acid bases, particularly with substituents in position 6 in purine and position 4 in pyrimidine. Base substitution within the same class (that is purine to purine and pyrimidine to pyrimidine) approximately preserves the size and shape of the base with concomitant switch of hydrogen bond donors to acceptors or vice versa in these positions. An example of such a switch is the position 6 of the purine ring which changes from a hydrogen bond donor ($-\text{NH}_2$) in adenine to a nearly iso-steric (sterically equivalent) hydrogen bond acceptor (6-Oxo) in guanine. Thus, if an adenine is recognized at the 6-amino group by an amino acid side-chain, which is a hydrogen bond acceptor, substitution of an iso-steric amino acid with hydrogen bond donor properties may recognize guanine in place of adenine or vice versa. Such a simple substitution rule may allow relatively simple engineering of novel recognition specificity.

As a test case, we chose λ -repressor–O_L1 complex to test the above-mentioned idea. Lysine-4 of λ -repressor is a crucial residue from point of view of operator recognition. In conjunction with Asn-55 it recognizes base pair CG6. Lysine-4 forms a crucial hydrogen bond with CG6, in which it recognizes the 6-Oxo group of the guanine by hydrogen bond donation from the ϵ -NH₂ group (Figures 1 and 2) while Asn-55 donates a hydrogen bond to purine N7 of CG6 base pair (14).

*To whom correspondence should be addressed. Tel: +91 33 2413 1157; Fax: +91 33 2473 5197; Email: siddhartharoy@iicb.res.in

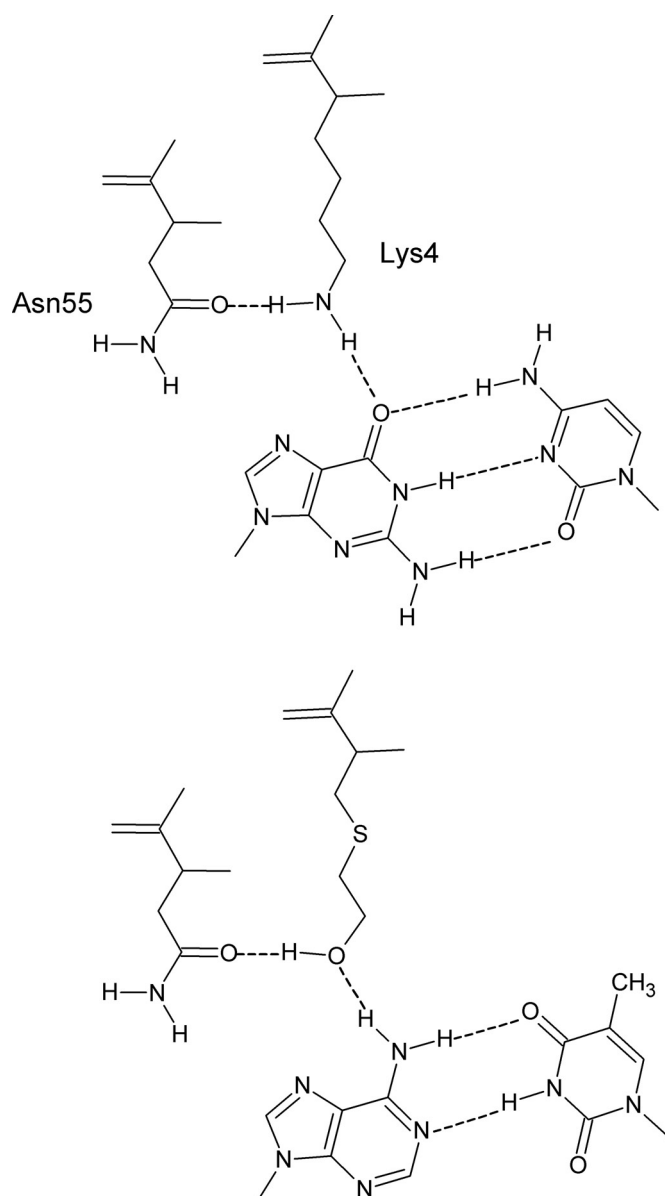


Figure 1. Upper panel shows the hydrogen bonding pattern of CG6 with lysine-4 and hydrogen bonding of lysine-4 with Asn-55. Lower panel is the hypothetical hydrogen bonding pattern of the mutant TA6 and S-(2-hydroxyethyl)-cysteine-4, along with Asn-55.

This interaction is further stabilized by hydrogen bond between lysine-4 amino group and the side-chain carbonyl of Asn-55. If a TA base pair is substituted for CG6, an amino group replaces the oxo group at the 6 position of the purine ring, while retaining other recognition points. The two hydrogen atoms of this amino group make this group a hydrogen bond donor and may be recognized by a hydrogen bond acceptor, otherwise iso-steric with lysine-4. An ideal iso-steric group with above-mentioned properties will be where the ϵ -NH₂ group of lysine is replaced by a hydroxyl group (Figure 1). This will also preserve the hydrogen bond between lysine-4 and Asn-55, thus preserving the recognizing pair. Site-specific incorporation of such a lysine analog would involve either chemical synthesis or nonsense codon suppression, both of which could be cumbersome (15,16).

Modification of a cysteine that has been introduced by site-directed mutagenesis has been used as a strategy to introduce unnatural amino acids. In several cases mutation to a chemically similar amino acid has been used to derive the stringent requirement of a particular amino acid (17–22). We developed a strategy to create an amino acid, which is iso-steric with lysine but chemically dissimilar, through side-chain modification of a reactive cysteine. In this paper, we demonstrate that such a semi-synthetic repressor indeed recognizes the mutated DNA sequence.

MATERIALS AND METHOD

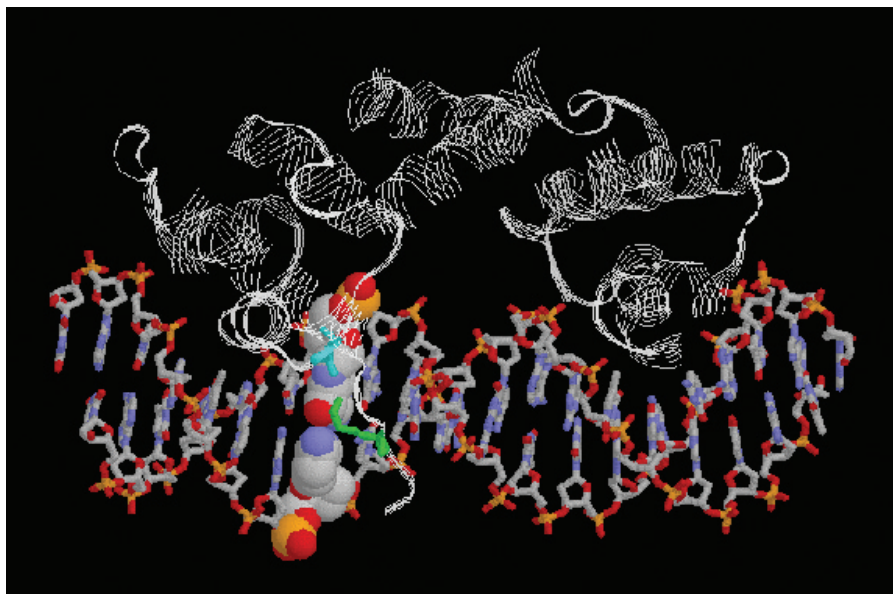
Materials

Acrylamide, bisacrylamide, TEMED, phenylmethylsulfonyl fluoride (PMSF), ethidium bromide, 2-iodoethanol, bromophenol blue, Commassie brilliant blue, hydroxylapatite, QAE-Sephadex A-50, Sephadex G-25, streptavidine agarose, EDTA, DTNB, DTT, BSA and ampicillin were purchased from Sigma chemical Co. Tryptone and yeast extract was purchased from Himedia Laboratories Pvt. Ltd (India). Magnesium chloride, sodium chloride, calcium chloride, 2-mercaptoethanol, glycerol and Tris were purchased from E-Merck (India). X-gal, isopropyl- β -D-thiogalactopyranoside (IPTG), sodium MOPS, agarose, triethanolamine and SDS were purchased from SRL (India). *N*-iodoacetyl-*N*-biotinyl-hexylenediamine was purchased from Pierce (USA). Restriction enzymes and T4 DNA ligase were purchased from Boehringer Mannheim and GIBCO BRL. *Taq* DNA polymerase primers and dNTPs were purchased from Genei (Bangalore, India). 5'-Amino linked oligonucleotides were from Microsynth (Switzerland).

Site-directed mutagenesis

The mutation from lysine-4 to cysteine-4 was performed by site-directed mutagenesis of the *cl* gene using PCR. A mutant upstream primer containing EcoRI site (5'-GTA TAG AAT TCACGT ATG AGC ACA AAA TGC AAA CCA TT-3') and a downstream primer complementary to non-coding region containing a BamHI site (5'-ACC ATT CTT CAT AAT TGG ATC CAT TTA CTA TG-3') were used. The PCR was carried out in 0.5 ml Perkin-Elmer microcentrifuge tube. The reaction mixture was prepared by mixing template DNA, primers, dNTPs, *Taq* DNA polymerase and the reaction buffer (supplied by the manufacturer).

The reaction was carried out in Perkin-Elmer GeneAmp PCR System 2400 (USA) DNA thermal cycler. The PCR product was run in 1% agarose gel and the desired DNA band was cut. The PCR product was extracted using QIAGEN gel extraction kit (Qiagen, Germany). PCR product and plasmid vector were digested with restriction enzymes EcoRI and BamHI following GIBCO BRL protocol. Ligation was carried out using T4 DNA ligase. Digestion of DNA with restriction endonucleases was carried out at 37°C for 3 h, in the buffer supplied by the manufacturer. The ligation of the EcoRI and BamHI digested PCR product with the EcoRI and BamHI digested pUC-19 vector was carried out at 23°C for 3 h using T4 DNA ligase. The ligated plasmid was transformed into competent DH5- α cells and colonies selected by blue-white selection. Five colonies were picked. One of those five



TA-1
AT-2
TA-3
CG-4
AT-5
CG-6
CG-7
GC-8
<u>CG-9</u>
CG-10
AT-11
GC-12
TA-13
GC-14
GC-15
TA-16
AT-17

Figure 2. Crystal structure of λ -repressor N-terminal domain crystal structure complexed with O_L1 . Lysine-4 is shown in green and asparagine-55 is shown in cyan. The base pair CG6 is shown in space filling representation. Lower panel depicts the sequence of O_L1 . The top half is the consensus half. Of the two strands, the top of left strand is the 5' end.

plasmids was fully sequenced in an automatic DNA sequencer and the desired mutation at K4C was confirmed. This plasmid having K4C λ -repressor gene was named pAM 109.

pAM109 and pKK223-3 plasmids were digested with EcoRI and PstI restriction enzymes. The restriction fragment containing K4C gene was extracted from digested pAM109 using QIAGEN gel extraction kit following their protocol. The K4C gene was ligated into digested pKK 223-3 using Boehringer Mannheim's rapid ligation kit. The ligated product was transformed in *Escherichia coli* XL-1 blue lac I^q strain following the previous method. Cells were grown on ampicillin Luria Agar plates at 37°C. Clone bearing colonies were confirmed through (i) plasmid isolation followed by agarose gel electrophoresis, (ii) restriction digestion of the plasmid by EcoRI and PstI restriction enzymes and (iii) after IPTG

induction the over expression of 26 kDa protein in 15% SDS-polyacrylamide gel. Cells from one of the cloned colonies were stored in 20% glycerol at -80°C and the corresponding plasmid was named pAM209.

Isolation and characterization of K4C λ -repressor

K4C λ -repressor protein was isolated from *E.coli* XL-1 blue strain carrying the plasmid pAM209, which contains Lys to Cys mutation at the 4th amino acid under the control of *tac* promoter. The cells were grown at 37°C, upto $A_{590} = 0.6$, then IPTG at a final concentration of 1 mM was added and grown further for 2 h. The cells were then harvested by centrifugation at 12 000 r.p.m. for 15 min and were re-suspended using 1.7 ml lysis buffer per gram of wet packed cells (SLA-1500). The cell

pellet was kept at -20°C . K4C λ -repressor was purified according to Johnson *et al.* (23), with some modifications. All volumes and amounts in the following purification have been normalized to 5 g of the starting cells. All steps were carried out at $0-4^{\circ}\text{C}$ unless stated otherwise.

Frozen cells were thawed, suspended in 25 ml of lysis buffer containing 20 $\mu\text{g/ml}$ of PMSF and disrupted by sonication. After sonication, (200 W; 30 s each time; 5 times with 5 min gap) the crude lysate was diluted to 100 ml by SB (23) containing 200 mM KCl and cellular debris were removed by centrifugation at 12 000 r.p.m. for 50 min. The repressor was precipitated with gradual addition of 5 ml of 10% polyethyleneimine, stirred for 1 h and precipitate was collected by centrifugation at 10 000 r.p.m. for 10 min. The polyethyleneimine pellet was re-suspended in 75 ml of SB containing 200 mM KCl and centrifuged at 10 000 r.p.m. for 10 min. Then the pellet was again re-suspended in 50 ml of SB containing 600 mM KCl, stirred for 10 min and centrifuged at 10 000 r.p.m. for 10 min. The supernatant containing the repressor was saved.

Repressor in the decanted supernatant was precipitated by gradual addition of 40 g of solid ammonium sulfate per 100 ml of the supernatant (i.e. 80% saturated) and pH was maintained at 8.0 by adding 3N KOH. The ammonium sulfate suspension was stirred for 2 h and the precipitate was collected by centrifugation at 13 000 r.p.m. for 45 min. The ammonium sulfate pellet was re-suspended in 5 ml of SB containing 50 mM KCl and was dialyzed overnight against 500 ml of SB containing 50 mM KCl. The dialyzed material was centrifuged at 12 000 r.p.m. for 10 min to remove any precipitated protein.

The supernatant containing the repressor was mixed with 50 ml of SB containing 100 mM KCl and 50 ml of pre-swollen QAE-Sephadex A-50 in SB containing 100 mM KCl and the mixture was stirred for 2 h at 4°C . It was then poured into a glass column and allowed to settle. The protein was eluted with a linear gradient of 150 ml of SB containing 100 mM KCl and 150 ml of SB containing 600 mM KCl. Flow rate was 0.5 ml/min and 6 ml fractions were collected. Fractions containing mutant repressor were identified by SDS-PAGE. The pooled fractions containing the mutant repressor were loaded on to a hydroxylapatite column and eluted by a gradient of 0.1–1.0 M potassium phosphate buffer (pH 6.8) containing 1.5 mM 2-mercaptoethanol (100 + 100) ml, flow rate was 0.5 ml/min and 3 ml fractions were collected. The mutant repressor pulled from the hydroxylapatite column was $\sim 98\%$ pure by criterion of electrophoresis on 15% SDS-polyacrylamide gel and gave dark single band (24) corresponding to a molecular weight of $\sim 26\,000$ Da. The native repressor concentration was determined by using $E^{1\%}_{280} = 11.3$ (25). Molar concentration was always calculated in terms of monomer subunit unless stated otherwise.

Determination of sulfhydryl content

K4C λ -repressor protein was mixed with 5 mM DTT and was kept for 1 h at 4°C and exhaustively dialyzed against degassed 0.1 M potassium phosphate buffer (pH 8.0) containing 1 mM EDTA. The protein solution was placed in a sample cuvette and the dialysate buffer was placed in a reference cuvette of a double-beam spectrophotometer. 5,5'-Dithiobis(2-nitrobenzoic acid) (DTNB) reacts with free sulfhydryl

groups of proteins to form thionitrobenzoate anion for each $-\text{SH}$ group. The concentration of this strongly colored anion is determined from its absorbance at 412 nm ($\epsilon = 13\,600\text{ M}^{-1}\text{ cm}^{-1}$ at pH 8). DTNB solution was added to each cuvette to a final concentration of 0.1 mM and the reaction was followed for 30 min. The determination of exposed cysteines of K4C λ -repressor proteins were followed at 6 μM protein by reaction with 0.1 mM DTNB.

Chemical modification of K4C by 2-iodoethanol

A total of 50 μl of 100 mM DTT (final concentration 3.3 mM) was added to 1.5 ml K4C λ -repressor protein (50 μM stock) and was kept for 2 h at 4°C . A 25 ml Sephadex G-25 column was equilibrated with the required amount of degassed reaction buffer [50 mM triethanolamine-HCl (pH 7.8) containing 50 mM NaCl, 1 mM EDTA and 10% glycerol]. Then the DTT and protein mixture was passed through a Sephadex G-25 column using the above-mentioned buffer to remove excess DTT and the reduced K4C λ -repressor was collected as 1.2 ml fractions. The fractions containing K4C λ -repressor were pooled in a screw cap glass vial (5 ml of 13 μM protein). The reaction between K4C λ -repressor and 2-iodoethanol was performed at 27°C . A total of 6 μl 2-iodoethanol was added into the reaction mixture under stirring conditions, three times with 2 h interval. The reaction was allowed to continue for 6 h and the final 2-iodoethanol concentration was 45 mM. After 6 h, iodoacetyl biotin derivative at a final concentration of 300 μM was added to the reaction mixture and the reaction was allowed to proceed. After 1 h, another fresh addition of iodoacetyl biotin to a final concentration of 150 μM (not counting the previous addition) was added and the reaction was allowed to proceed for another 1 h. Then excess DTT was added and was kept for 1 h. The reaction mixture was dialyzed against 1 litre degassed reaction buffer to remove excess DTT, 2-iodoethanol and the unreacted iodoacetyl biotin derivative followed by three dialysis against degassed 0.1 M potassium phosphate buffer (pH 8) containing 1 mM EDTA. A total of 500 μl streptavidine agarose was equilibrated with 0.1 M potassium phosphate buffer (pH 8) containing 1 mM EDTA. The dialyzed modified λ -repressor was mixed with 500 μl of pre-equilibrated streptavidine agarose and stirred for 10 min. Then the mixture was placed in a spin column and the iodoethanol modified K4C λ -repressor was separated from the mixture by centrifugation. The presence of free $-\text{SH}$ group was measured by DTNB reaction. The chemical modification at desired position was confirmed by DTNB test and N-terminal sequencing of the modified protein.

Circular dichroism (CD)

To see if any secondary structural elements change in the modified K4C λ -repressor, the far UV-CD spectra of modified K4C λ -repressor and wild λ -repressor were measured in a JASCO J-600 spectropolarimeter in a 1 mm path length cuvette at 25°C , between 250 and 200 nm. The bandwidth was 2 nm and scan speed was 20 nm/min. Samples were prepared at 50 mM Tris-HCl buffer (pH 7.5), containing 0.1 mM EDTA and 200 mM KCl and the protein concentration were modified 2.4 μM for K4C λ -repressor and 2 μM for wild-type λ -repressor. The results are reported as mean residue ellipticity, with the unit of $\text{deg cm}^2\text{ dmol}^{-1}$.

Fluorescence spectroscopy

The wavelength maximum of tryptophan fluorescence of modified K4C λ -repressor (concentration 2.4 μ M) and wild λ -repressor (concentration 2 μ M) were measured in a Hitachi F3010 spectrofluorometer at 25°C and the experiments were carried out in 1 cm pathlength quartz cuvette. The buffer was 50 mM Tris-HCl (pH 7.5), containing 0.1 mM EDTA and 200 mM KCl. The excitation wavelength was at 295 nm, the excitation and emission bandpasses were 10 nm each.

Fluorescein isothiocyanate (FITC) labeling of oligonucleotide

Oligonucleotides with 5' amino link (X = hexyl amino group) were purchased from Microsynth (Switzerland). Sequence of oligonucleotides were as follows: Operator O_L1: 5'-XTCGACATACCACTGGCGGTGATATGGATCC-3' and 5'-XGGATCCATATCACCGCCAGTGGTATGTCTGA-3'; Mutated O_L1 (CG6 to TA at the consensus half site): 5'-XTCGACATACCACTGGCGATGATATGGATCC-3' and 5'-XGGATCCATATCATCGCCAGTGGTATGTCTGA-3'.

For preparation of FITC labeled, oligonucleotides were labeled in 500 μ l solution containing 1 M sodium carbonate/bicarbonate buffer (pH 9); DMF:water [high-performance liquid chromatography (HPLC) grade] = 5:2:3. Reaction was carried out for 20 h at 25°C. After incubation, the reaction was loaded onto a Sephadex G-25 column [pre-equilibrated in 0.1 M phosphate buffer (pH 8.0)] and eluted with same buffer. The oligonucleotides and their complementary strands were annealed by heating at 80°C followed by cooling slowly at room temperature.

Operator titration of wild λ -repressor and its iso-steric analogue

All fluorescence studies were done in a Hitachi F3010 spectrofluorometer at 25°C and the experiments were carried out in 1 cm pathlength quartz cuvette. Anisotropy experiments were performed using Hitachi (650-10/40) polarization accessory. The excitation wavelength was at 480 nm and emission was at 530 nm, the excitation and emission bandpasses were 10 nm each. Fluorescein labeled wild O_L1 and mutated O_L1 (the CG6 of consensus site has replaced by TA6) were titrated with increasing concentrations of both wild-type λ -repressor and *S*-(2-hydroxyethyl)-cysteine containing λ -repressor separately upto 350 nM protein concentrations. Anisotropy values were calculated at each point according to the equation and plotted against different protein concentrations. The titrations were performed in 50 mM Tris-HCl buffer (pH 7.5), containing 1 mg/ml BSA, 0.1 mM EDTA and 250 mM KCl. The binding isotherms obtained were fitted to a single site binding equation of the dimer coupled with monomer-dimer equilibrium using a MATLAB program written by the authors. The dimer-monomer dissociation constant was taken as 20 nM.

RESULTS AND DISCUSSION

The strategy to create an *S*-(2-hydroxyethyl)-cysteine containing λ -repressor at position 4 relies on the fact that λ -repressor has no reactive cysteines (26). Thus, site-directed mutation of K4 to C, followed by reaction with 2-iodoethanol may create

an amino acid, which is almost iso-steric with lysine but having an OH instead of the ϵ -NH₂. K4C mutant cI gene was created by PCR using one mutagenic end primer and another end primer. The mutation was confirmed by sequencing of the whole gene. The mutant gene expressed well and was purified as described previously (25).

The mutant protein showed one reactive sulfhydryl group per subunit upon reaction with DTNB (Table 1). The reaction kinetics of C4 was rapid, indicating exposed nature of the sulfhydryl group. The K4C λ -repressor was labeled with 2-iodoethanol, followed by reaction with an iodoacetamido derivative of biotin to block the unreacted sulfhydryl groups. The reaction mixture was then passed through a column of streptavidin-agarose to remove all biotin-modified protein. This leads to isolation of λ -repressor that has *S*-(2-hydroxyethyl) cysteine in both subunits. The purified *S*-(2-hydroxyethyl)-cysteine-4 containing λ -repressor was characterized by ten rounds of N-terminal Edman sequencing, CD and fluorescence spectroscopy. N-terminal sequencing revealed an unknown amino acid at position 4, while the other nine are consistent with the wild-type sequence. λ -repressor has three tryptophans at 129, 142 and 230. They are sensitive probes of the tertiary structure of the protein. Figure 3 shows the tryptophan emission spectra of the wild-type and *S*-(2-hydroxyethyl)-cysteine-4 λ -repressor. The emission maximum and fluorescence intensities (after corrections for concentration differences) are very similar (emission maximum of 339 nm in both cases), indicating no significant changes in the tertiary structure due to modification. The tryptophans in λ -repressor are all situated in or around the

Table 1. Reaction of K4C and wild-type λ -repressor with DTNB

Type of protein	Protein concentration (μ M)	Observed OD ₄₁₂	No. of exposed -SH group/subunit
K4C	6	0.083	1.04
Wild type	6	0.019	0.23

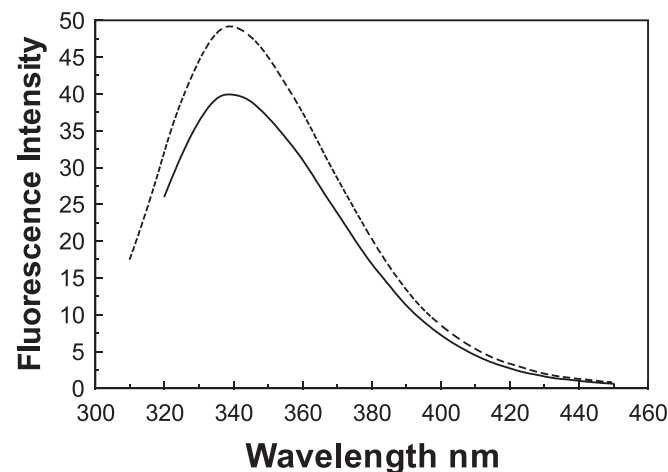


Figure 3. Fluorescence emission spectra of wild-type (solid line; 2 μ M) and *S*-(2-hydroxyethyl)-cysteine-4 λ -repressor (dashed line; 2.4 μ M). Experimental details are in Materials and Methods.

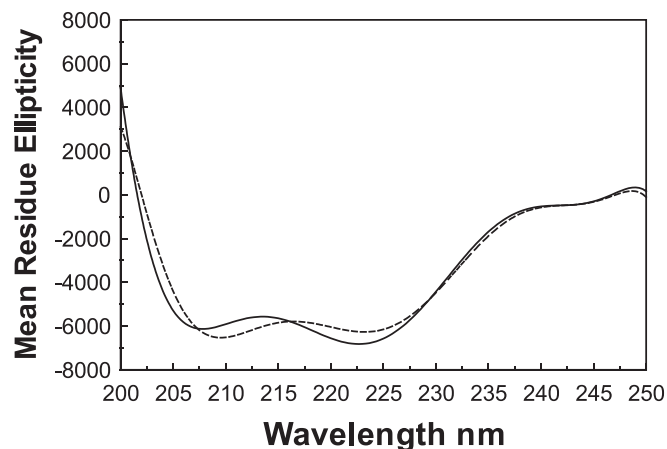


Figure 4. CD spectra of wild-type (solid line; 2 μ M) and *S*-(2-hydroxyethyl)-cysteine-4 λ -repressor (dashed line; 2.4 μ M). Experimental details are in Materials and Methods.

C-terminal domain and thus report the integrity of the structure there. The DNA-binding N-terminal domain is devoid of tryptophan but consists of several helices. The far UV-CD spectrum of the whole protein is dominated by the N-terminal helices (26) and thus the CD spectra is a sensitive monitor for the conformation of the N-terminal domain. Figure 4 shows the CD spectra of the wild-type and the *S*-(2-hydroxyethyl)-cysteine-4 λ -repressor. The spectra are similar indicating that there is no significant conformational perturbation in the N-terminal domain due to mutation and chemical modification. Taken together, we may conclude that the protein remains in native-like conformation.

Fluorescence anisotropy is now widely used to study protein–DNA interaction (27). Figure 5 shows binding of wild-type λ -repressor with wild-type O_L1 operator site containing oligonucleotide duplex. The binding was measured by fluorescence anisotropy of an attached fluorescein probe to the 5' ends of the duplex oligonucleotide. The binding was conducted at somewhat elevated ionic strengths to offset any possible non-specific binding. The ionic strength used is not incompatible with physiological conditions. The binding isotherm was fitted to a binding equation as described in Koblan and Ackers (28,29). The initial anisotropy of the duplex oligonucleotide and amplitude of increase is consistent with those obtained previously (30). The dissociation constant thus obtained was 56.2 nM for wild-type repressor dimer– O_L1 complex. The dissociation constant of a similar synthetic O_R1 containing oligonucleotide–wild-type λ -repressor complex was determined by isothermal titration microcalorimetry and found to be 4.2 nM at 210 mM monovalent salt and pH 7.0 (31). Koblan and Ackers (28,29) have determined dissociation constant of O_R1 –wild-type λ -repressor complex as a function of monovalent salt concentration and pH. After extrapolation of dissociation constant obtained by microcalorimetry to 300 mM monovalent cation (including tris) concentration and pH 7.5, a value of \sim 35 nM was predicted for the dissociation constant of O_L1 –wild-type λ -repressor complex at pH 7.5 and 300 mM monovalent cation (linear extrapolation of data presented in Koblan and Ackers to 300 mM monovalent salt concentration yielded a factor of

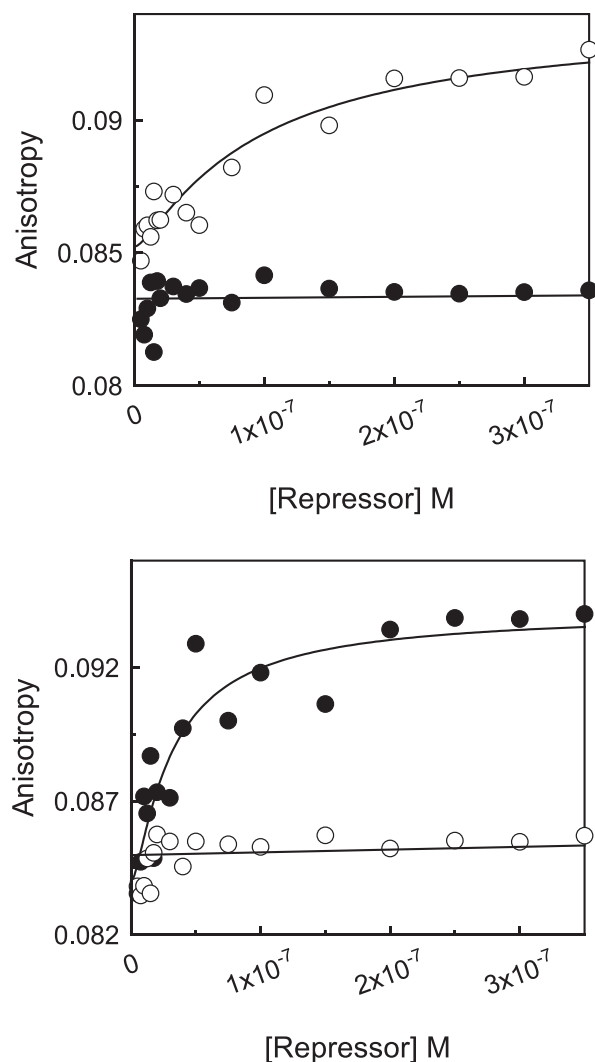


Figure 5. Upper panel reflects the anisotropy increase of 5' fluorescein labeled synthetic duplex oligonucleotide containing either wild-type O_L1 operator site (open circles) or TA6 mutant (solid circles), with wild-type λ -repressor. Lower panel reflects the anisotropy increase with *S*-(2-hydroxyethyl)-cysteine-4 λ -repressor. The symbols have the same meaning as above. The oligonucleotides were end-labeled with fluorescein as described. The titration procedure is also given in the Supplementary Data. Each point is an average of eight independent titrations.

three weakening when compared to that at 210 mM. Similar extrapolation from pH 7.5 yields a factor of about four weakening when compared to pH 7. This 12-fold difference is then corrected for the small difference in affinity for O_L1 and O_R1). This compares well with the dissociation constant for O_L1 –wild-type λ -repressor obtained (56.2 nM) by fitting the data shown in the upper panel of Figure 5.

Under identical conditions, wild-type λ -repressor does not bind the CG6→TA6 mutant operator (Figure 5). This is consistent with base pair substitution experiments (32). Figure 5 also shows the binding of *S*-(2-hydroxyethyl)-cysteine-4 containing λ -repressor with wild-type O_L1 . There is little binding in the concentration range tested indicating importance of having a lysine ϵ -NH₂ in position 4. The deletion of first six amino acids of λ -repressor abolishes operator

binding indicating importance of lysine-4 (33). The *S*-(2-hydroxyethyl)-cysteine-4 modified λ -repressor however binds tightly to CG6→TA6 mutant O_L1 as indicated by increased anisotropy. When the binding isotherm was fitted to the binding equation, a dissociation constant of 9.82 nM was obtained for *S*-(2-hydroxyethyl)-cysteine-4 modified λ -repressor complex. This indicates that *S*-(2-hydroxyethyl)-cysteine-4 modified λ -repressor regains all the binding energy that is lost upon TA6 substitution.

Since both the dimers contain *S*-(2-hydroxyethyl)-cysteine-4 modifications, while the mutant operator carries only a mutation in the consensus half site, it is likely that putative K4 interaction with the symmetrically related CG pair in the non-consensus half site does not contribute significantly to the binding energy. This is consistent with the fact that the N-terminal arm of the subunit that contacts the non-consensus half site remains disordered even at low temperatures in the crystal and substitution of a TA base pair in the non-consensus half have little effect on the binding energy (see Figure 2) (10,32).

Although the sterically equivalent substitution made here is incorporated by chemical modification, the enormous progress that is being made in recent years in incorporating artificial amino acids *in vivo* (34) should allow us to soon make non-natural iso-steric amino acid substitutions, *in vivo*, with ease. This may open up new possibilities of engineering new DNA-binding proteins and engineering of the genetic programs, not hitherto accessible.

ACKNOWLEDGEMENTS

This work was funded by a grant from Department of Science and Technology, Government of India. We also acknowledge CSIR (India) fellowship to A.M. Funding to pay the Open Access publication charges for this article was provided by CSIR, New Delhi, India.

Conflict of interest statement. None declared.

REFERENCES

- Jantz,D., Amann,B.T., Gatto,G.J., Jr and Berg,J.M. (2004) The design of functional DNA-binding proteins based on zinc finger domains. *Chem. Rev.*, **104**, 789–799.
- Klug,A. (2005) Towards therapeutic applications of engineered zinc finger proteins. *FEBS Lett.*, **579**, 892–894.
- Schildbach,J.F., Karzai,A.W., Raumann,B.E. and Sauer,R.T. (1999) Origins of DNA-binding specificity: role of protein contacts with the DNA backbone. *Proc. Natl Acad. Sci. USA*, **96**, 811–817.
- Botfield,M.C., Jancso,A. and Weiss,M.A. (1994) An invariant asparagine in the POU-specific homeodomain regulates the specificity of the Oct-2 POU motif. *Biochemistry*, **33**, 8113–8121.
- Nilsson,M.T. and Widersten,M. (2004) Repertoire selection of variant single-chain Cro: toward directed DNA-binding specificity of helix–turn–helix proteins. *Biochemistry*, **43**, 12038–12047.
- Cowley,S.M., Kang,R.S., Frangioni,J.V., Yada,J.J., DeGrand,A.M., Radhakrishnan,I. and Eisenman,R.N. (2004) Functional analysis of the Mad1-mSin3A repressor–corepressor interaction reveals determinants of specificity, affinity, and transcriptional response. *Mol. Cell. Biol.*, **24**, 2698–2709.
- Ebright,R.H., Cossart,P., Gicquel-Sanzey,B. and Beckwith,J. (1984) Mutations that alter the DNA sequence specificity of the catabolite gene activator protein of *E.coli*. *Nature*, **311**, 232–235.
- Wharton,R.P. and Ptashne,M. (1987) A new specificity mutant of 434 repressor that defines an amino acid–base pair contact. *Nature*, **326**, 888–891.
- Youderian,P., Vershon,A., Bouvier,S., Sauer,R.T. and Susskind,M.M. (1983) Changing the DNA-binding specificity of a repressor. *Cell*, **35**, 777–783.
- Wharton,R.P. and Ptashne,M. (1985) Changing the DNA-binding specificity of a repressor by redesigning an α -helix. *Nature*, **316**, 601–605.
- Kopke Salinas,R., Folkers,G.E., Bonvin,A.M., Das,D., Boelens,R. and Kaptein,R. (2005) Altered specificity in DNA-binding by the lac repressor: a mutant lac headpiece that mimics the gal repressor. *ChemBiochem*, **6**, 1628–1637.
- Lehming,N., Sartorius,J., Kisters-Woike,B., von Wilcken-Bergmann,B. and Muller-Hill,B. (1990) Mutant lac repressors with new specificities hint at rules for protein–DNA recognition. *EMBO J.*, **9**, 615–621.
- Sartorius,J., Lehming,N., Kisters,B., von Wilcken-Bergmann,B. and Muller-Hill,B. (1989) lac repressor mutants with double or triple exchanges in the recognition helix bind specifically to lac operator variants with multiple exchanges. *EMBO J.*, **8**, 1265–1270.
- Beamer,L.J. and Pabo,C.O. (1992) Refined 1.8 Å crystal structure of the lambda repressor–operator complex. *J. Mol. Biol.*, **227**, 177–196.
- Mendel,D., Cornish,V.W. and Schultz,P.G. (1995) Site-directed mutagenesis with an expanded genetic code. *Annu. Rev. Biophys. Biomol. Struct.*, **24**, 435–462.
- Wang,L. and Schultz,P.G. (2004) Expanding the genetic code. *Angew Chem. Int. Ed. Engl.*, **44**, 34–66.
- Smith,H.B., Larimer,F.W. and Hartman,F.C. (1988) Subtle alteration of the active site of ribulose biphosphate carboxylase/oxygenase by concerted site-directed mutagenesis and chemical modification. *Biochem. Biophys. Res. Commun.*, **152**, 579–584.
- Planas,A. and Kirsch,J.F. (1991) Reengineering the catalytic lysine of aspartate aminotransferase by chemical elaboration of a genetically introduced cysteine. *Biochemistry*, **30**, 8268–8276.
- Smith,H.B. and Hartman,F.C. (1988) Restoration of activity to catalytically deficient mutants of ribulosebiphosphate carboxylase/oxygenase by aminoethylation. *J. Biol. Chem.*, **263**, 4921–4925.
- Lukac,M. and Collier,R.J. (1988) Restoration of enzymic activity and cytotoxicity of mutant, E553C, *Pseudomonas aeruginosa* exotoxin A by reaction with iodoacetic acid. *J. Biol. Chem.*, **263**, 6146–6149.
- Bednar,R.A., McCaffrey,C. and Shan,K. (1991) Introduction of unnatural amino acids into chalcone isomerase. *Bioconjug. Chem.*, **2**, 211–216.
- Dhalla,A.M., Li,B., Alibhai,M.F., Yost,K.J., Hemmingsen,J.M., Atkins,W.M., Schineller,J. and Villafranca,J.J. (1994) Regeneration of catalytic activity of glutamine synthetase mutants by chemical activation: exploration of the role of arginines 339 and 359 in activity. *Protein Sci.*, **3**, 476–481.
- Johnson,A.D., Pabo,C.O. and Sauer,R.T. (1980) Bacteriophage lambda repressor and cro protein: interactions with operator DNA. *Methods Enzymol.*, **65**, 839–856.
- Laemmli,U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.
- Saha,R., Banik,U., Bandopadhyay,S., Mandal,N.C., Bhattacharyya,B. and Roy,S. (1992) An operator-induced conformational change in the C-terminal domain of the lambda repressor. *J. Biol. Chem.*, **267**, 5862–5867.
- Banik,U., Saha,R., Mandal,N.C., Bhattacharyya,B. and Roy,S. (1992) Multiphasic denaturation of the lambda repressor by urea and its implications for the repressor structure. *Eur. J. Biochem.*, **206**, 15–21.
- Lundblad,J.R., Laurance,M. and Goodman,R.H. (1996) Fluorescence polarization analysis of protein–DNA and protein–protein interactions. *Mol. Endocrinol.*, **10**, 607–612.
- Koblan,K.S. and Ackers,G.K. (1991) Cooperative protein–DNA interactions: effects of KCl on lambda cI binding to OR. *Biochemistry*, **30**, 7822–7827.
- Koblan,K.S. and Ackers,G.K. (1991) Energetics of subunit dimerization in bacteriophage lambda cI repressor: linkage to protons, temperature and KCl. *Biochemistry*, **30**, 7817–7821.
- Reedstrom,R.J., Brown,M.P., Grillo,A., Roen,D. and Royer,C.A. (1997) Affinity and specificity of trp repressor–DNA interactions studied with fluorescent oligonucleotides. *J. Mol. Biol.*, **273**, 572–585.

31. Merabet,E. and Ackers,G.K. (1995) Calorimetric analysis of lambda cI repressor binding to DNA operator sites. *Biochemistry*, **34**, 8554–8563.
32. Sarai,A. and Takeda,Y. (1989) Lambda repressor recognizes the approximately 2-fold symmetric half-operator sequences asymmetrically. *Proc. Natl Acad. Sci. USA*, **86**, 6513–6517.
33. Eliason,J.L., Weiss,M.A. and Ptashne,M. (1985) NH₂-terminal arm of phage lambda repressor contributes energy and specificity to repressor binding and determines the effects of operator mutations. *Proc. Natl Acad. Sci. USA*, **82**, 2339–2343.
34. Wang,L. and Schultz,P.G. (2002) Expanding the genetic code. *Chem. Commun. (Camb.)*, **1**, 1–11.