Single-Chain Lambda Cro Repressors Confirm High Intrinsic Dimer−DNA Affinity†

Rinku Jana, Tony R. Hazbun, Jeffrey D. Fields, and Michael C. Mossing*

Department of Biological Sciences, University of Notre Dame, Notre Dame, Indiana 46556

Received January 21, 1998; Revised Manuscript Received March 18, 1998

ABSTRACT: The overall affinity of the bacteriophage lambda Cro repressor for its operator DNA site is limited by dimer dissociation at submicromolar concentrations. Since Cro dimer−operator complexes form at nanomolar concentrations of Cro subunits where free dimers are rare, these dimers must bind with compensating high affinities. Previous studies of the covalent dimer Cro V55C suggest little change in DNA binding affinity even though the dimeric species is quantitatively populated; this is an apparent contradiction to the expectation of high intrinsic dimer−DNA affinity. In contrast to the disulfide linkage at the center of the dimer interface in Cro V55C, polypeptide linkers that join the two subunits allow single-chain Cro repressors to bind operator DNA with picomolar affinities. A series of five single-chain Cro repressors have been expressed from fused tandem *cro* genes. Each contains a peptide linker of 8−16 hydrophilic residues that connects the C-terminus of one subunit to the N-terminus of the next. All bind to operator DNA with at least 100-fold higher affinity than Cro V55C. Proteins containing the longest and shortest linkers have been purified and characterized in detail. Both exhibit similar CD spectra to wild-type Cro and enhanced thermal stability. Sedimentation equilibrium experiments show that single-chain Cro repressors do not associate at concentrations up to 30 μM. The rate of dissociation of Cro−DNA complexes is almost unchanged by covalent linkage. Biophysical characterization of Cro variants such as these, where DNA binding is uncoupled from subunit assembly, is necessary for a quantitative understanding of the structural and energetic determinants of DNA recognition in this simple model system.

Many DNA binding proteins function as oligomers. In cases where protein−protein association is weak in relation to DNA binding, the saturation of a DNA site depends on the coupling of protein−DNA and protein−protein association equilibria (*1*). These coupled equilibria modulate both the extent of binding at a given protein concentration and the shape of the regulatory response as a function of protein accumulation. Model systems where such coupling has been demonstrated include *Escherichia coli* gal and lac repressor systems (*2*), phage P22 Arc repressor (*3*), and phage lambda cI repressor systems (*4*, *5*). A primary reason to investigate systems such as these is to build a solid understanding of the physical principles which govern protein−DNA recognition. Here we use protein engineering to demonstrate that the DNA binding energetics of another model system, the lambda Cro repressor, have been seriously underestimated due to the neglect of the coupled assembly equilibria.

The Cro repressor of bacteriophage lambda binds to operator DNA as a dimer of two identical 66 residue subunits (*6*−*8*). Although many studies of Cro have been conducted under the assumption that dimers are the predominant species in solution (*9*−*11*), recent results indicate that dimers undergo dissociation into subunits at concentrations that are far in excess of those required for DNA binding (*8*). Direct measurements of dimerization in the absence of DNA suggest that dimers constitute substantially less than 1% of the free protein at the nanomolar concentrations that are required for half saturation of an operator site. These rare dimers must therefore bind to DNA with very high affinity.

One way to confirm the predicted high dimer−DNA affinity is to study DNA binding in dimers whose dissociation has been prevented. In the Cro V55C variant, the two Cro subunits are covalently linked by a disulfide bond at the center of the dimer interface. Even though the covalent linkage ensures that the dimeric species will be populated at high dilution, the apparent DNA binding affinity (measured as the total concentration of subunits required for half saturation of an operator site) of Cro V55C is similar to Cro wild-type. The observed differences range from small (3- to 8-fold) decreases (*11*−*13*) to a similar small increase (*8*) in operator affinity relative to wild-type. We propose that the apparent minimal effects on DNA binding are due to the fortuitous cancellation of two opposing effects. Covalent linkage results in a greater than 100-fold increase in the fraction of assembled dimers at nanomolar concentrations. To counterbalance this increase, the disulfide bond at the

† Supported by NIH Grant GM46513.

* Author to whom correspondence should be addressed. Phone: 219 631-8412. Fax: 219 631-7413. E-mail: mossing.1@nd.edu.

[1] Abbreviations: IPTG, isopropyl β-D-thiogalactoside; EDTA, ethylenediaminetetraacetic acid; BME, β-mercaptoethanol; SDS, sodium dodecyl sulfate; LB, Luria broth; MALDI-TOF, matrix-assisted laser desorption ionization−time of flight; CD, circular dichroism spectropolarimetry; PCR, polymerase chain reaction; EMS, electrophoretic mobility shift; MRE, mean residue ellipticity; F_u, fraction unfolded; T_m, temperature at which half of the protein has undergone the unfolding transition; scCro, single-chain Cro repressor. KP20 buffer contains 20 mM potassium phosphate, 1 mM EDTA, and 1.4 mM BME at pH 7.0. KP200 buffer is KP20 plus 200 mM KCl; TBE contains 90 mM Tris, 90 mM Borate, and 2 mM EDTA, pH 8.0.

S0006-2960(98)00152-4 CCC: $15.00
Published on Web 04/16/1998

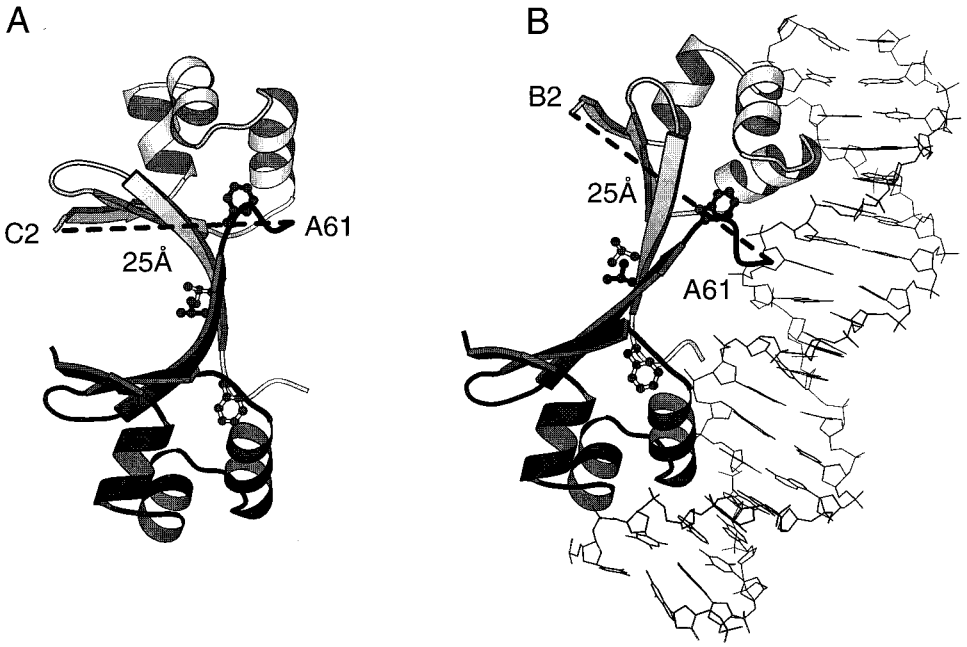

FIGURE 1: Cro and Cro-Operator complex structures. (A) Wild-type Cro protein, shown as a ribbon diagram. Subunit A is colored dark gray and subunit C is colored white. The side chains of V55 and F58 from each subunit are indicated in ball-and-stick representation. The distance from asparagine 61 of subunit A to glutamate 2 of subunit C is indicated by a dashed line. The coordinates of the dimer are from ref *31*. (B) The Cro dimer bound to operator DNA. Subunits A and B from ref *32* are drawn as in panel A. The A subunits, shown in dark gray in both the free and complex structures, were superimposed prior to lateral displacement to generate the figure. Even though a large reorientation of the upper subunit (subunit C in panel A and subunit B in panel B) is evident upon DNA binding, the distance from the C-terminal residue of one subunit to the N-terminus of the other is almost unchanged. Panels A and B were generated using MOLSCRIPT (*30*).

center of the dimer interface is predicted to interfere with DNA binding.

To obtain covalent dimers which reveal the intrinsic affinity of the wild-type dimer for DNA, a linkage is required which does not perturb the structure or flexibility of the wild-type dimer. Mutagenesis studies on the Cro repressor have shown that deletion or substitution of the extreme C-terminal residues have only small effects on DNA affinity (*11*). Cro variants with C-terminal fusions of up to 30 residues have been isolated which exhibit DNA binding affinities similar to wild-type Cro (*14*). Both C- and N-terminal residues are found to be disordered in the structures of Cro that have been determined (*15*, *16*). Together these observations suggest that a flexible linkage between the C-terminus of one subunit and the N-terminus of the next might well be compatible with the structure and function of the Cro dimer.

The linkage of independent folding units into single polypeptide chains is common in evolution. Many natural proteins are constructed of repeated units. Separate genes encoding interacting subunits have been fused in vitro as a practical means to obtain recombinant proteins with desirable properties such as single-chain antibodies (*17*). Single-chain dimers of other nucleic acid binding proteins have been constructed to study fundamental issues in macromolecular recognition, protein folding, and stability (*18*−*21*).

Here we describe the design, construction, and characterization of single-chain Cro repressors. These unitized repressors meet our expectations for enhanced stability and DNA binding affinity and confirm the high intrinsic dimer affinity predicted by the analysis of the coupled dimerization and DNA binding equilibria of wild-type Cro (*8*).

EXPERIMENTAL PROCEDURES

Enzymes, Reagents, and Buffers. Restriction endonucleases and DNA-modifying enzymes were obtained from New England Biolabs, Promega, Fisher, Stratagene, and Invitrogen. Oligonucleotides were obtained from the Biocore facility at the University of Notre Dame. Radiochemicals were obtained from Amersham. All other chemicals and reagents were from Fisher or Sigma. Spectroscopy and ultracentrifugation experiments were conducted in KP200 buffer. Electrophoretic mobility shift (EMS) assay buffer is KP200 with 200 μg/mL bovine serum albumin and 5% glycerol.

Bacterial Strains and Plasmids. JM109(DE3), DH5α, and BL21(DE3) cells were obtained from laboratory stocks. The XL1-Blue and XL1mutS competent cells were from Stratagene. The pCR2.1 plasmid was from Invitrogen. The pET21a plasmid was from Novagen. The pAP114 (*22*), pUCroRS (*23*), and pUCroV55C (*8*) plasmids were obtained from laboratory stocks.

Design and Construction of Single Chain Cro Genes. The three step strategy for constructing single chain Cro dimers is outlined below and in Figure 2a. First a synthetic *cro* gene in the plasmid pUCroRS (*23*) was modified to remove the stop codon and introduce new restriction sites at the N- and C-termini. Second, this gene was cloned in tandem with the original gene in the T7 expression vector pET21a. Finally, synthetic linkers were introduced to fuse the two genes.

1. PCR Modification of Terminal Restriction Sites. In the first step a synthetic *cro* gene was modified by PCR using the oligonucleotide primers listed below. MM213: 5′-GGA

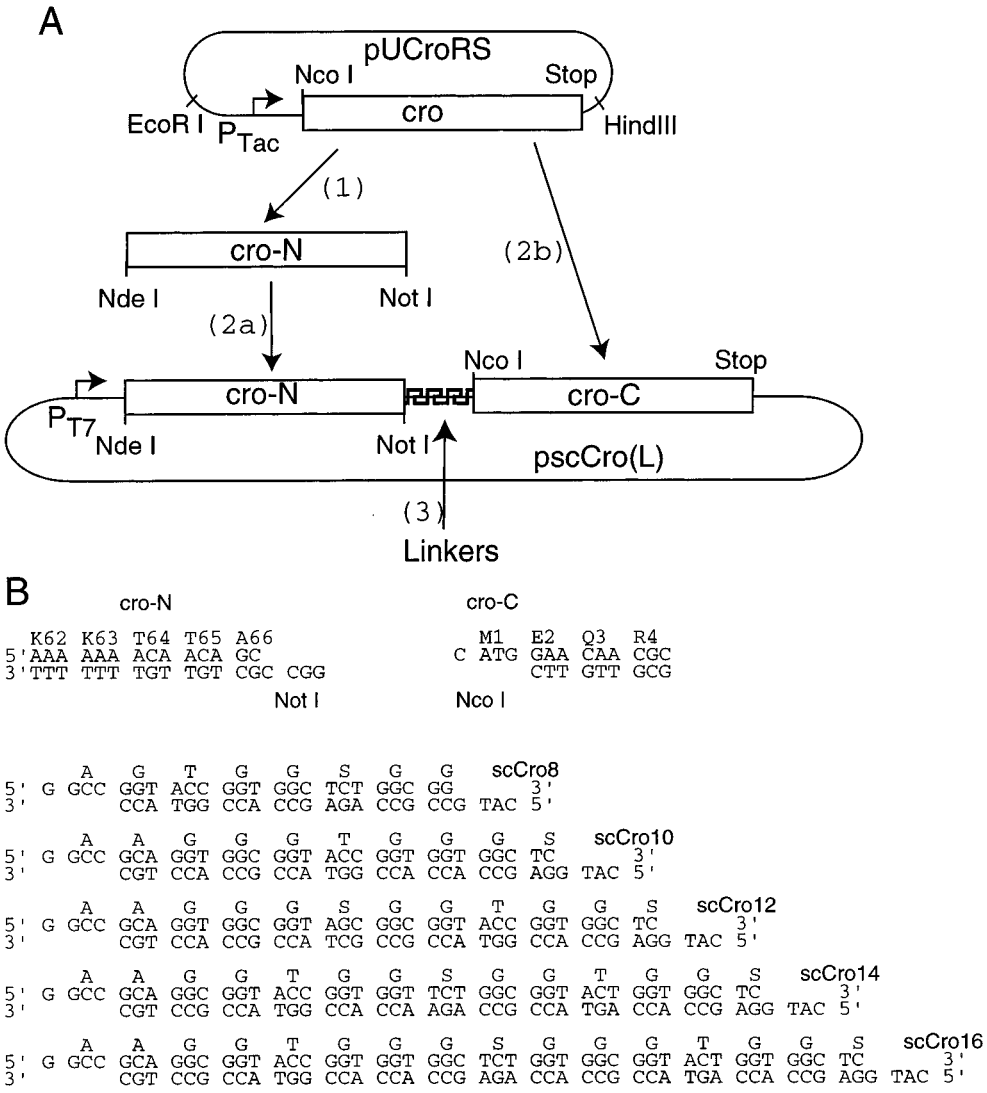

FIGURE 2: Plasmid construction and linker sequences. (A) The construction of the pscCro plasmids. (1) Denotes the introduction of new terminal restriction sites and elimination of the stop codon by PCR, (2a) the cloning of the N-terminal *cro* gene into pET21a, (2b) cloning of the C-terminal *cro* gene in tandem with the first and (3) fusion of the two genes by introduction of the linkers illustrated in panel B. Each pscCro plasmid contains tandem fused synthetic *cro* genes under the control of a T7 promoter. (B) DNA sequences of the ends of the two synthetic *cro* genes and the oligonucleotide linkers used to fuse them. DNA sequences are written as complementary duplexes. The bottom strand is the template strand for transcription. Amino acid residues encoded by each triplet are listed above. The five duplex oligonucleotides listed at the bottom of the figure encode linkers of 8, 10, 12, 14, and 16 amino acid residues.

GGT CAT ATG GAA CAA CGC ATA ACT-3′ and MM214: 5′-CTC GAG TGC GGC CGC TGT TGT TTT TTT −3′.

Oligonucleotide MM213 changes the *Nco*I site in the template gene to an *Nde*I site in the amplified product. MM214 introduces a C-terminal *Not*I site in the amplified product and eliminates the stop codon at the end of the *cro* gene. The PCR-amplified gene product was ligated between flanking *Eco*R1 sites into the vector pCR2.1.

2. Tandem Cro Genes under T7 Control. (a) The modified *cro* gene, designated *cro-N*, was excised by digestion with *Nde*I and *Eco*RI and ligated into the T7 promoter vector pET21a which had been digested with the same enzymes. In the resulting plasmid, pJF1a, *cro-N* is fused to the six-histidine tag sequence of pET21a. (b) The original (unmodified) *cro* gene, designated *cro-C*, was cloned in tandem with the modified gene. The *cro-C* gene and associated tac promoter sequences were excised from the plasmid pUCroRS by digestion with *Eco*RI and *Hind*III and ligated with similarly digested pJF1a. Subsequent elimination of the *Not*I and *Hind*III sites at the end of *cro-C* by primer-directed mutagenesis yielded the plasmid pCro2.

3. Selection for Gene Fusion. The third step was to replace the intervening Tac promoter fragment in pCro2 with an oligonucleotide linker. pCro2 was digested with *Not*I and *Nco*I and ligated to synthetic oligonucleotide linkers as illustrated in Figure 2b. After ligation, plasmids were transformed into XL-1 Blue cells and plated on LB plates containing ampicillin and IPTG. In this strain, IPTG results in the induction of transcription from the P_{Tac} promoter but not the P_{T7} promoter. Plasmids from Amp- and IPTG-resistant colonies were transferred to JM109(DE3), and protein expression was screened in crude lysates by SDS−PAGE (*24*). Plasmids expressing proteins of the appropriate sizes were designated pscCro(L) where L designates the number of amino acid residues in the linker peptide.

Protein Expression and Purification. The single-chain Cro repressors were expressed in JM109(DE3) cells under the

control of a T7 promoter. For protein expression, cells were grown with aeration at 37 °C in LB broth plus ampicillin (100 μg/mL) to an OD_{600} of ~1. IPTG was added to a final concentration of 1 mM. The cultures were allowed to grow for 3 h at 37 °C after induction. The cells were chilled to 4 °C and harvested by centrifugation at 3000*g*.

For initial characterization, each of the single-chain dimers scCro8, scCro10, scCro12, scCro14, and scCro16 were purified from 50 mL cultures. Cells were resuspended in 1 mL of KP20, lysed by sonication, and cleared by centrifugation. Supernatants were mixed with 1 mL of a DEAE-Sephacell suspension and centrifuged briefly to settle the resin. Supernatant containing unadsorbed scCro repressor was applied to a Mono S column and eluted in a gradient of 0−1 M KCl. Peak fractions (>80% pure) were used directly for EMS assays.

Cro wild-type and Cro V55C were expressed from tac promoter plasmids and purified as described (*8*). scCro8 and scCro16 were expressed in 4 L cultures and purified after lysis in urea as described (*24*). Single-chain Cro repressors behaved essentially identically to wild-type Cro at each step in the purification. Purified proteins were characterized by SDS gel electrophoresis and MALDI-TOF mass spectrometry (*8*). Extinction coefficients were determined by comparison of the absorbance of native and guanidine hydrochloride denatured samples (*25*).

Analytical Ultracentrifugation. Analytical ultracentrifugation experiments were performed as described (*8*) using a Beckman Optima XL-I analytical ultracentrifuge. High-speed sedimentation equilibrium experiments were conducted with protein concentrations of 1.5, 5, and 30 μM at 20 000, 24 000, and 30 000 rpm. Absorbance was monitored at 220 and 280 nm in four scans at each speed. The absorbance data at 280 nm were analyzed to generate equilibrium profiles for the single-chain repressor molecules.

Circular Dichroism. Circular dichroism spectra and denaturation profiles were obtained on an AVIV 62 DS spectropolarimeter. Spectra were acquired between 200 and 300 nm with a 0.5 nm bandwidth. CD signals were acquired at 1 nm intervals for 10 s. Each spectrum is an average of five scans. All data were collected at 25 °C using a 0.2 or 1 cm path length quartz cuvette. For thermal stability experiments, the CD signal at 222 nm was monitored at 2.5 °C intervals from 15 °C to 80−95 °C. Upon completion of the unfolding transition, temperatures were reduced according to the same schedule to allow the protein to refold. Unfolding and refolding transitions were identical within experimental error. Low- and high-temperature baselines were estimated by linear extrapolation, and T_ms were determined as previously described (*24*).

DNA Binding Assays. Equilibrium DNA binding titrations by Cro wild-type, Cro V55C, and the single-chain variants were performed as described previously (*8*). Cro proteins at concentrations of 1−10 μM in storage buffer (KP200 containing 50% glycerol) were diluted in EMS assay buffer. The initial dilution was followed by 3-fold serial dilutions to generate the desired range of protein concentrations. These protein solutions were allowed to equilibrate at room temperature for at least 5 min before mixing with equal volumes of 0.2 pM ^{32}P-labeled consensus operator DNA. After equilibration for another 30 min, samples were loaded onto a 12.5% acrylamide 0.5X TBE gel. Stoichiometric titrations of protein with 30 pM DNA indicated that the scCro8 and scCro16 proteins were greater than 80% active. Equilibrium binding constants for the single-chain repressors were estimated from the midpoint in the binding isotherm where the free DNA and Cro−DNA complex bands were of equal intensity.

Dissociation rates for Cro−DNA complexes were determined from EMS assay data. Sufficient protein (5.5 nM) was mixed with DNA (11 pM) and equilibrated for 30 min to give greater than 80% saturation of labeled consensus operator DNA. Protein−DNA complex dissociation was initiated by mixing with 1/10 volume of 2.5 μM unlabeled competitor consensus operator DNA to give final concentrations of 5 nM protein, 10 pM labeled DNA, and 250 nM unlabeled DNA. At the appropriate times (15−2400 s), 5 μL aliquots were loaded onto a gel running at 43 V/cm (*8*). Control reactions, where the protein was added to unlabeled competitor DNA before addition of labeled DNA, resulted in less that 4% of the labeled DNA in complex band. Radioactivity in the free and complex bands was quantified on a Molecular Dynamics Storm 840 instrument, and the fraction of bound DNA was determined for each time point. The data from 90 to 2400 s were fit to a single-exponential decay using MATLAB.

RESULTS

Design of the Single-Chain Dimer. The structures of the Cro dimer in its free (*31*) and DNA-bound (*32*) forms, show that the distance between residue N61 at the C-terminus of one subunit and residue E2 at the N-terminus of the other subunit is approximately 25 Å (Figure 1). The large conformational change that is evident upon DNA binding does not appreciably affect this distance. Note that in neither structure are the five C-terminal residues or the N-terminal methionine observed. Since the goal of the present work is to determine the intrinsic Cro dimer affinity, we selected a range of peptide linker sizes. In this way we hoped to identify linkers that were neither so short as to interfere with the normal interaction of C-terminal residues with DNA nor so long as to interfere with DNA binding on their own. Taking 3.6 Å as the interresidue distance in a fully extended polypeptide chain, linkers of 8 to 16 amino acid residues should be able to span from 29 to 58 Å. Considering the mobile N- and C-terminal residues as part of the connecting loop would yield flexible connections of 14 to 22 residues.

Also shown in Figure 1 are the side chains of V55 and F58. V55 side chains from each subunit project away from the DNA at the center of the dimer interface. Substitution of this residue by a cysteine in the Cro V55C mutant results in the spontaneous formation of a disulfide bond which links the two subunits. The F58 side chains make the key hydrophobic contacts of the dimer interface as they extend from the C-termini to pack into the hydrophobic core of the opposite subunit. Changes in the packing of F58 relative to the recognition helix α3 in the free and DNA-bound forms of Cro may be important for the energetics of DNA recognition (*32*, *33*).

Figure 2A outlines the construction of single-chain dimer genes as described in Experimental Procedures. First, a previously cloned synthetic *cro* gene was modified by PCR

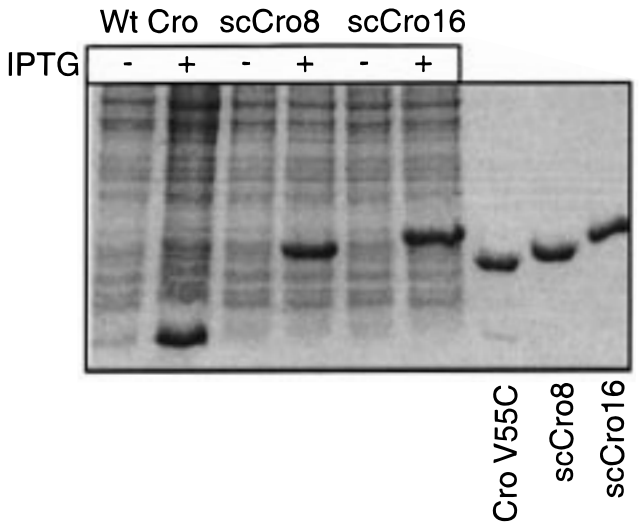

FIGURE 3: Expression levels and final purity of Cro variants. The nonreducing SDS−PAGE gel shows induced expression level of wild-type Cro, scCro8 and scCro16 and the purified covalent dimer proteins that were used for further analysis. Lanes 1−6 contain alternating uninduced and induced lysates of JM109(DE3) cells expressing Cro wild-type, scCro8, and scCro16 as indicated. Lanes 7−9 contain purified Cro V55C, scCro8, and scCro16 as indicated.

to remove the stop codon at the end of the gene and introduce two new restriction sites. After an intermediate cloning step, this modified *cro* gene (*cro-N*) was inserted into pET21a. The original gene including its Tac promoter was inserted downstream of *cro-N*. Fusion proteins were constructed by replacing the intervening Tac promoter fragment with the duplex oligonucleotides shown in Figure 2B. The final plasmids encoding fused tandem genes under the control of the T7 promoter were transferred to JM109(DE3) for protein expression.

The linker sequences depicted in Figure 2B were selected on the basis of earlier successful single-chain designs (*17*, *19*). Other than the alanines at the N-terminus which were required to maintain the *Not*I restriction site, all of the residues in the linkers are glycine, serine, and threonine. These residues were selected for maximal flexibility and hydrophilicity. Charged residues were avoided because of the importance of electrostatic interactions between proteins and the highly charged DNA polyelectrolyte.

Protein Expression and Purification. JM109(DE3) cells bearing candidate scCro plasmids were harvested after induction, lysed in SDS, and subjected to SDS−PAGE analysis. Clones which expressed proteins of the expected sizes were identified for each of the oligonucleotide linkers shown in Figure 2B. All were expressed at similar levels (scCro8 and scCro16 are shown in Figure 3). Representatives of each of the five scCro proteins were purified from candidate clones to ∼80% homogeneity by batch purification on DEAE Sephacell and chromatography on Mono S. EMS assays of these partially purified fractions showed that each of the linked dimers was competent for operator binding at concentrations less than 20 pM. In these preliminary experiments the highest affinity that we could distinguish was limited by the specific activity of our labeled operator. Within the parameters of the experiment each of the proteins showed equivalent affinity for DNA. Rather than continue with the analysis of all five proteins, whose behavior

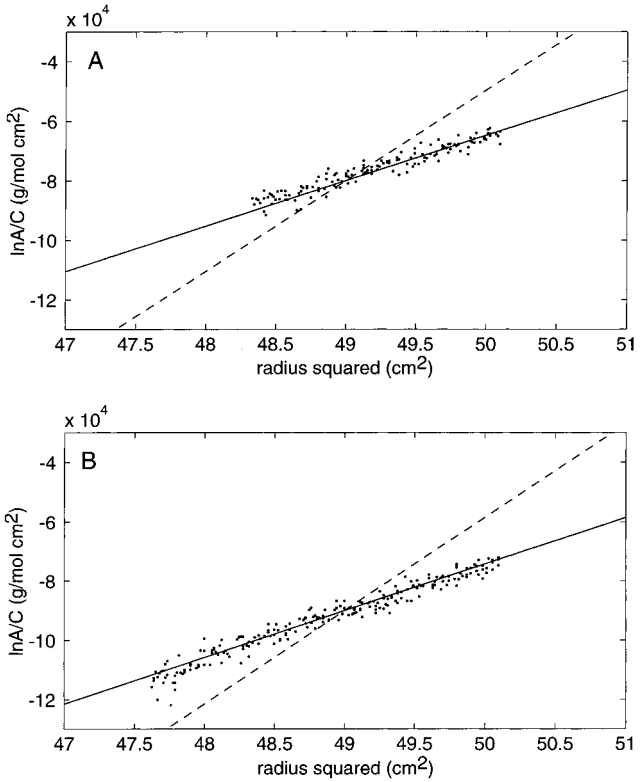

FIGURE 4: Sedimentation equilibrium analysis of the scCro repressors. Representative absorbance profiles from analytical ultracentrifugation experiments with (A) scCro8 and (B) scCro16. Absorbance at 280 nm was measured as a function of radial distance for ∼30 μm protein samples at 20,000 rpm, 25 °C. The log of the measured absorbance at 280 nm is scaled by a correction factor calculated from the equation: $c = (1 - \bar{\nu}\rho)\omega^2/2RT$ where > $\bar{\nu}$ is the partial specific volume calculated from the amino acid sequence (0.727 for scCro8 and 0.723 for scCro16), ρ is the density of the solution (1.04 g/cm^3), ω is the angular velocity in radians per s, R is the gas constant 8.314×10^7, and T is the temperature in Kelvin (293). The slopes of the theoretical lines give molecular mass directly. The solid lines represent the theoretical profile of a single ideal species with molecular mass of the single chain dimer. The dashed lines represent the profile expected for a single ideal species of twice the expected mass.

appeared identical, we concentrated on only the shortest and longest of the linked dimers.

scCro8 and scCro16 were purified from bacterial cell extracts as described previously (*24*). The proteins obtained were found to be >95% pure, and they migrate as expected on SDS−PAGE (Figure 3). Figure 3 also shows that more than 90% of Cro V55C is in the disulfide bonded form. Extinction coefficients at 280 nm for scCro8 and scCro16 were 7700 M^{-1} cm^{-1}, close to twice that of the wild-type Cro subunit. Cro V55C has a slightly higher extinction coefficient, 8400 M^{-1} cm^{-1}, as expected due to the presence of the cystine.

Structure, Association, and Stability. The masses of scCro8 (15 267) and scCro16 (15 811) were determined by MALDI-TOF mass spectrometry. These masses are within 0.1% of the expected masses of 15 254 for scCro8 and 15 798 for scCro16. Analytical ultracentrifugation shows the native molecular masses of both scCro8 and scCro16 to be those expected for a single unassociated polypeptide chain. The data shown in Figure 4 were obtained with protein samples at initial concentrations of 30 μM while at sedimentation equilibrium at 20 000 rpm. Concentration profiles were

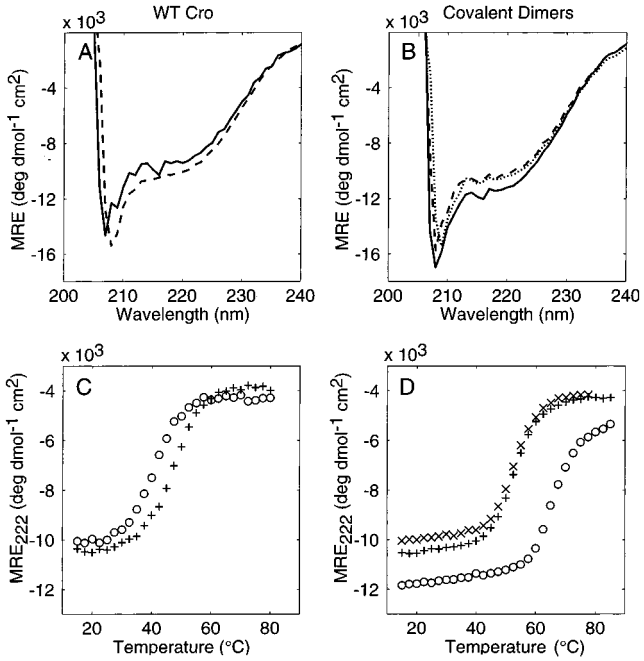

FIGURE 5: Circular dichroism studies. (A) Mean residue ellipticity of wild-type Cro at 4 μM (solid line) and 34 μM (dashed line) plotted as a function of wavelength. (B) Mean residue ellipticity of Cro V55C (solid line) and scCro8 (dashed line) and scCro16 (dotted line) plotted as a function of wavelength. All proteins were at 5 μM. (C) Mean residue ellipticity at 222 nm of wild-type Cro at 4 μM (o) and 58 μM (+) plotted as a function of temperature. (D) Mean residue ellipticity at 222 nm of Cro V55C (40 μM − o) and scCro8 (33 μM − +) and scCro16 (34 μM − x) plotted as a function of temperature.

obtained as a function of radial distance by scanning absorbance at 280 nm. Theoretical lines with ideal slopes for a single species of the expected molecular masses (continuous lines) and for twice the expected molecular masses (dashed lines) are superimposed on the data. Similar results are obtained at lower protein concentrations and higher rotor speeds.

Figure 5 compares circular dichroism spectra and thermal stabilities of wild-type Cro and covalent dimers. Wild-type Cro in panels A and C shows concentration dependence both in mean residue ellipticity and in thermal stability. The enhancement in mean residue ellipticity at high concentrations is subtle but reproducible. The thermal stability enhancement at high concentrations is more dramatic, resulting in melting temperatures of 42 °C at 4 μM and 46 °C at 58 μM as seen in panel C. In contrast, experiments over the range of 3 to 50 μM for Cro V55C, scCro8, and scCro16 showed no concentration-dependent changes in either mean residue ellipticity or T_m. Panels B and D show data acquired with protein concentrations from 5 to 35 μM.

Mean residue ellipticities for each of the covalent dimers are slightly higher than those measured for even the highest concentrations of wild-type Cro. Note that if the linker regions are presumed to adopt random coil configurations, the mean residue ellipticity scale which is normalized to the number of residues might be expected to give signals that are somewhat lower for scCro8 (130 residues) and scCro16 (138 residues) relative to Cro V55C or wild-type Cro (122 residues per dimer). Compared to each other, scCro8 and scCro16 are remarkably similar both in MRE and in thermal stability (T_m = 53 °C). Cro V55C gives a more negative MRE and a higher T_m (66 °C).

Circular dichroism spectra and stability of scCro repressors give no indication of structural perturbation due to the peptide linkage. CD spectra of the scCro proteins are similar to the spectrum of wild-type Cro, consistent with little change in secondary structure. The thermal denaturation profiles of both scCro8 and scCro16 show that both variants are stabilized compared to micromolar concentrations of wild-type Cro. Whereas wild-type Cro is stabilized at increasing concentrations (Figure 5c), the denaturation profiles for each of the covalent dimers exhibit no concentration dependence over a 10-fold range (not shown). The concentration dependence of the T_m of wild-type Cro was extrapolated as $1/T_m$ versus the log of concentration (*21*). This extrapolation suggests that millimolar concentrations of wild-type Cro would yield stability equivalent to the scCro proteins (T_m = 53 °C). A similar calculation for V55C (T_m = 66 °C) suggests that the effective concentration provided by this linkage is in the molar range.

scCro Repressors Exhibit Picomolar Affinity for Operator DNA. Equilibrium binding to a 25 bp consensus operator was determined by EMS assays for the scCro proteins, Cro wild-type, and Cro V55C. As described in more detail previously (*8*) Cro wild-type protein at 8×10^{-10} M was sufficient to bind half of the labeled operator DNA, while Cro V55C showed half-maximal binding at 4×10^{-10} M protein concentration. scCro8 and scCro16 bound with over 100-fold higher affinity at $(4 \pm 2) \times 10^{-12}$ M. These midpoint concentrations can be interpreted directly as K_ds only when the concentration of free dimers is equivalent to the total concentration. This is only the case for covalent dimers whose dissociation is prevented and only when dimers are in sufficient excess of labeled operator sites. The analysis of the wild-type Cro binding isotherm depends on an independent estimate of the dimerization constant as described previously (*8*). The mobility differences between the free and bound DNA bands are similar for each of the Cro variants, suggesting that the size of the complexes are similar as well.

Dissociation Kinetics for Covalent Dimer and Cro Wild-Type Operator Complexes Are Similar. Dissociation kinetics (Figure 7) of protein−DNA complexes for scCro, Cro V55C, and Cro wild-type are similar. Each of the protein−DNA complexes investigated shows some evidence of a rapid dissociation within the first minute. This may reflect some heterogeneity in the population of protein−DNA complexes or an intrinsic difference between dissociation that occurs for the bulk of the incubation time versus the few seconds during which the protein−DNA complexes are loaded onto the running gel. Regardless of its origin, the rates and amplitudes of the early phase are comparable for all four complexes. Fitting the time points from 90 to 2400 s to a single-exponential decay yielded rate constants which are all within 2-fold of 4×10^{-4} s^{-1}. Half-lives for each of the complexes were 23 ± 5 min for scCro8, 29 ± 12 min for scCro16, 38 ± 14 min for Cro V55C, and 17 ± 3 min for Cro wild-type. The similar dissociation rates suggest that the differences seen in equilibrium DNA binding of Cro and its variants must be in the association rates of protein with DNA.

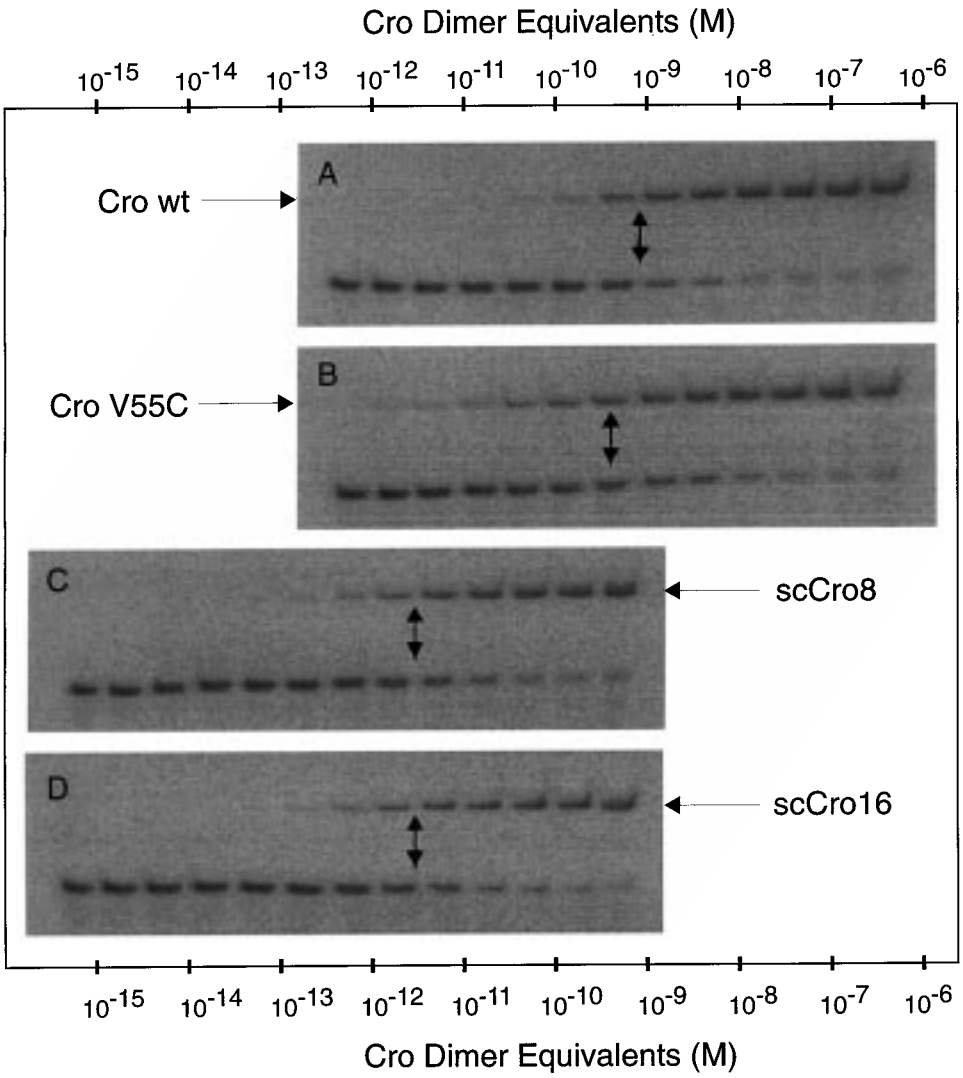

FIGURE 6: Equilibrium DNA binding titrations of Cro wild-type, V55C, scCro8 and scCro16. The position of each protein-DNA complex is indicated by the single arrowhead. The point of half-maximal saturation of the DNA operator by each protein is indicated by the double arrowhead. The total concentration of ^{32}P labeled operator DNA was 2×10^{-13} M in each case. (A) Titration of Cro wild-type starts with a protein concentration of 364.5×10^{-9} M in the right most lane. Lanes to the left represent successive 3-fold dilutions. The left most lane in each panel contains only DNA. The DNA half-saturation point is 0.8×10^{-9} M. (B). Cro V55C was titrated as in (A). The DNA half saturation point is 0.4×10^{-9} M. (C) Titration of scCro8 as in (A) except the highest protein concentration was 0.5×10^{-9} M. The DNA half-saturation point is 4 ($\pm$ 2) $\times 10^{-12}$ M. (D) scCro16 was titrated as in (C). The DNA half-saturation point is 4 ($\pm$ 2) $\times 10^{-12}$ M.

DISCUSSION

The goal of the studies described here is to confirm that weak dimer assembly limits Cro−operator interactions. Previous work (*8*) showed that in the absence of DNA, wild-type Cro dimers dissociate in the micromolar concentration range. The observed dimer dissociation constant [(7 $\pm$ 4) $\times 10^{-6}$ M)] dictates that at a total Cro subunit concentration of 1×10^{-9} M only two subunits in 7000 will form dimers at equilibrium. This free monomer−dimer distribution has important implications both for the intrinsic equilibrium DNA affinity and for the kinetics of DNA complex formation. The most straightforward way to reconcile the observed dimer−DNA complex formation at nanomolar total concentrations of Cro with the predicted sub picomolar concentration of free dimers is to postulate that these rare dimers must bind with very high intrinsic affinity. The predicted large pool of free subunits and very limited population of free assembled dimers also suggests that dimer assembly might be a kinetically limiting process in Cro dimer−DNA association. The wide span between the dimer dissociation constant and the dimer−DNA complex dissociation constant makes an independent determination of the intrinsic dimer affinity problematic (*8*). The experimentally accessible parameter is the product of the dimer dissociation and the dimer−DNA binding constants. Here we have attempted to quantitatively populate the normally rare dimeric species at low concentrations by engineering a covalent linkage between the subunits.

Ideally we would like to construct a Cro dimer in which nothing was changed except the ability of its subunits to dissociate from each other. If dimer dissociation could be prevented, while leaving the structure and function of the protein unperturbed, the concentration of protein required for saturation of an operator site should reveal the intrinsic dimer affinity. By making a series of single-chain dimers with different linker segments, we hoped to find one which approximated this ideal. In fact, from the longest to the shortest of the engineered linkers, the single-chain repressors described here behave similarly. Although several observa-

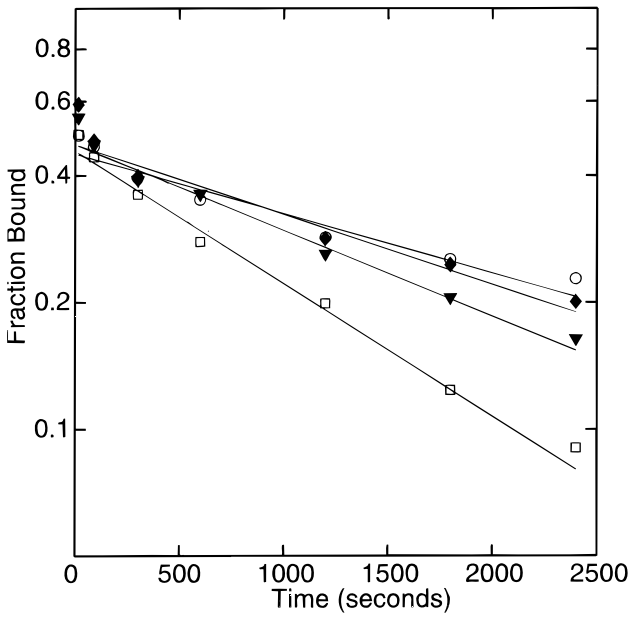

FIGURE 7: Cro-Operator complex dissociation kinetics. EMS assays were used to monitor the dissociation of Cro-Operator complexes. Fraction bound was determined by dividing the integrated phosphorimager counts from the complex band by the sum of the counts in the complex and free DNA bands at each time point. Cro V55C (open circles), scCro8 (filled triangles), scCro 16 (filled diamonds) and wild-type Cro (open squares) are plotted as a function of time after addition of unlabeled competitor DNA. Lines represent best fit parameters for a single-exponential decay of data points between 90 and 2400 s.

tions suggest that the perturbations to the structure and function of the Cro dimer are small, a quantitative discrepancy persists between the observed DNA binding affinity of the single-chain repressors and that predicted for wild-type Cro.

scCro Repressors Confirm That Dimer Assembly Limits DNA Binding. The general prediction that prevention of dimer dissociation would enhance DNA binding affinity has been found to be correct. Both 8 and 16 residue linkers reduce the concentration of Cro subunits required for half saturation of an operator site by approximately 250-fold (Figure 6). The thermodynamic origins of this effect is likely largely entropic since linkage of the two subunits eliminates the costs in cratic entropy of assembly of the dimer from free subunits. Since dissociation rates for each of the dimer−DNA complexes are very similar (Figure 7), the overall enhancement of affinity can be attributed almost entirely to an increased rate of Cro−DNA complex formation. Again, this result is predicted if the primary effect of covalent linkage is to populate the ordinarily rare dimer species at low concentrations.

Do the DNA Binding Affinities of Single-Chain Cro Repressors Reflect the Intrinsic Affinity of the Wild-Type Cro Dimer? The fact that equivalent affinities and thermodynamic stabilities are observed for scCro8 and scCro16 suggests that even the shortest of the linkers constructed (eight residues) is not so restrictive as to drastically disrupt structure or function. Even though the five C-terminal residues of Cro are disordered in both of the crystal structures depicted in Figure 1, at least some of these residues are important for DNA binding. In particular, K62 and K63 have been implicated in specific and nonspecific binding (*11*). Restrictions on the mobility of this region could hinder binding if the active conformation were disfavored by the linker. Alternatively, if the restrictions to mobility were primarily in the free protein, the linker could provide an entropic benefit for binding. No such systematic effects are evident so far.

The fact that the dissociation behavior of the scCro−operator complexes so closely parallels the behavior of the wild-type Cro−operator complex provides another indication that perturbations induced by the peptide linkers are small. For many DNA-binding proteins including Cro (*26*), changes in equilibrium binding affinity due to disruption of protein−DNA contacts are manifest primarily in the dissociation rate of the complex. These effects have been observed either through alterations in the sequence of the protein or the sequence of the DNA site. This generalization is not always true. The dissociation behavior of V55C described here appears to be a case where the difference in affinity is reflected primarily in the association rate (see below).

The equilibrium Cro-operator dissociation constant (K_d) of 4×10^{-12} M that we measured for both scCro8 and scCro16 still falls short of the predicted K_d for wild-type Cro dimer, 6×10^{-14} M (*8*). The remaining 66-fold discrepancy may arise from imperfections of the polypeptide linkers that we have introduced or in the simplified model with which we analyze the behavior of wild-type Cro.

In contrast to the uniform behavior of the single-chain dimers, Cro V55C exhibits dramatically reduced DNA binding affinity and enhanced thermodynamic stability. We believe that the disulfide at the center of the dimer interface significantly interferes with DNA recognition in this molecule. Indications of structural perturbations of V55C have been observed before (*12*, *13*, *27*). The primary locus of the DNA-induced conformational change in wild-type Cro is in the residues immediately flanking V55 (Figure 1, ref *32*). Surprisingly, the dissociation rate of Cro V55C−DNA complexes is not accelerated with respect to those observed for the more tightly binding scCro repressors. This result suggests that the distortion induced by the disulfide bond in this variant does not affect the stability but rather the rate of assembly of the Cro V55C−operator complex. It may be that the primary effect of the disulfide is to bias the ensemble of Cro dimer conformations in solution away from those conformations required for DNA recognition. Apparently once the complexes are formed, the energetic barriers to dissociation are unchanged or even slightly enhanced by the disulfide bond.

It is certainly possible that direct negative (*21*) or positive effects (*19*) of the inserted linker sequence could be operating. One test for direct effects would be to construct a single-chain−V55C double mutant. If the effect of the single-chain linker is purely on assembly, the V55C−scCro double mutant should have similar affinity for DNA as the V55C single mutant. Dimers which are pre-assembled by virtue of the disulfide bond should show no further enhancement as a result of the second covalent connection. If the single-chain peptide linkage were to have intrinsic positive or negative interactions with DNA, a higher or lower affinity for operator DNA might be expected for the double mutant compared to the V55C parent.

Implications for Previous Work. The fact that the assembly of Cro−DNA complexes proceeds through a rare free dimer intermediate may help explain some of the discrepancies in the Cro − operator binding affinities that have been reported (*8*−*11*, *22*, *28*). If the first-order dimer dissociation rate were substantially slower than the second-order dimer−DNA association process, rapid dilution of the protein from a high concentration directly into a solution of operator DNA may allow one to trap dimers in DNA complexes before they have a chance to dissociate. While technically difficult (the wild-type Cro dimer half-life is less than 10 s − T.R.H. and M.C.M., unpublished) preliminary data indicates that substantially higher operator occupancy can be obtained by rapid direct dilution than by preequilibration of protein in advance of mixing with DNA. These excess complexes are expected to decay over time to yield the true equilibrium distribution, but the complex dissociation times can be on the order of hours at low salt (*26*). Unless protein samples are preequilibrated at appropriate concentrations in advance of addition of the DNA, the level of complex formation may not reflect the true equilibrium value but rather will vary with the details of the dilution protocol.

As intensively as protein−DNA interactions have been studied and as numerous as their cocrystal structures are, we are still at a primitive stage in understanding how structures can be interpreted in energetic terms. A common feature of protein−DNA recognition that has been proposed on the basis of thermodynamic similarities to protein folding reactions is the coupling of protein conformational change to DNA recognition (*29*). Conformational changes in the Cro dimer as documented in the crystal structures of the free (*15*) and DNA-bound protein (*7*) certainly must enter into the energetics of DNA recognition. Perhaps even more importantly the large entropic cost in assembling the dimer from free subunits consumes a large part of the favorable free energy of complex formation. The studies described here are a crucial step in clarifying the relationship of protein sequence and structure to DNA binding energetics in the lambda Cro−operator DNA interaction. By systematically identifying and experimentally manipulating each of the sources of energetic coupling in this system, we can begin to attribute binding free energies to specific molecular interactions.

If, for example, one seeks a specific attribution of binding free energy to the helix-turn-helix recognition unit of Cro, its solvent accessible surface changes or its network of hydrogen bonding interactions with DNA, it is obviously critical to first know whether the intrinsic dimer−DNA affinity is 10^{-9} M or 10^{-13} M. Even for this very simple homodimer−DNA complex, the coupling of DNA recognition to protein folding and assembly leads to a complicated network of energetic effects. Now that dimer assembly can be uncoupled from DNA binding, it should be more straightforward to interpret the energetic consequences of changes to the sequence of the Cro dimer interface. Alterations to the structure or flexibility of this crucial element of protein architecture will allow us to explore the energetic coupling of DNA recognition by the two helix-turn-helix motifs in the Cro dimer.

A hallmark of transcriptional regulation is the utilization of different constellations of transcription factors and DNA sites to achieve distinct regulatory characteristics. A first step in the characterization of such a regulatory complex is a study of the interaction of individual protein factors with their cognate sites in vitro. Often these protein−DNA complex affinities are characterized simply by the protein concentration necessary to achieve half saturation of a DNA site. As seen here, similar apparent affinities such as those observed for Cro wild-type and Cro V55C can mask fundamental differences in the energetics of DNA recognition. To extend the analysis of individual protein−DNA interaction to the behavior of multiprotein−DNA complexes, it is necessary to come to grips with the energetic contributions of protein−protein interactions to DNA recognition. A thorough understanding of the thermodynamic coupling in the relatively simple lambda Cro system will help to guide the analysis of more complex systems. Hopefully, studies such as these will serve as a foundation for the understanding of the affinity and specificity of multiprotein−DNA complexes.

ACKNOWLEDGMENT

We thank Dr. Paul Huber for comments on the manuscript. Dr. Ron Albright, Dr. Dale Tronrud, and Dr. Brian Matthews at the University of Oregon provided coordinates of Cro and its variants. Dr. Francis Castellino and Dr. Mary Prorok in the Department of Chemistry and Biochemistry provided invaluable assistance in the use of their XL-I Analytical Ultracentrifuge. We thank Dr. Bill Boggess in the Department of Chemistry and Biochemistry for Mass Spectrometry and Dr. Graham Lappin for use of the CD instrument.

REFERENCES

1. Wong, I., and Lohman, T. M. (1995) *Methods Enzymol. 259*, 95.
2. Brenowitz, M., Mandal, N., Pickar, A., Jamison, E., and Adhya, S. (1991) *J. Biol. Chem. 266*, 1281.
3. Brown, B. M., Bowie, J. U., and Sauer, R. T. (1990) *Biochemistry 29*, 11189.
4. Sauer, R. T., Hehir, K., Stearman, R. S., Weiss, M. A., Jeitler-Nilsson, A., Suchanek, E. G., and Pabo, C. O. (1986) *Biochemistry 25*, 5992.
5. Rusinova, E., Ross, J. B. A., Laue, T. M., Sowers, L. C., and Senear, D. F. (1997) *Biochemistry 36*, 12994.
6. Takeda, Y., Folkmanis, A., and Echols, H. (1977) *J. Biol. Chem. 252*, 6177.
7. Brennan, R. G., Roderick, S. L., Takeda, Y., and Matthews, B. W. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8165.
8. Jana, R., Hazbun, T. R., Mollah, A., and Mossing, M. C. (1997) *J. Mol. Biol. 273*, 402.
9. Hochschild, A., Douhan, J. D., and Ptashne, M. (1986) *Cell 47*, 807.
10. Takeda, Y., Sarai, A., and Rivera, V. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 439.
11. Hubbard, A. J., Bracco, L. P., Eisenbeis, S. J., Gayle, R. B., Beaton, G., and Caruthers, M. H. (1990) *Biochemistry 29*, 9241.
12. Shirakawa, M., Matsuo, H., and Kyogoku, Y. (1991) *Protein Eng* 4, 545.
13. Baleja, J. D., and Sykes, B. D. (1994) *Biochem Cell Biol* 72, 95.
14. Mossing, M. C., Bowie, J. U., and Sauer, R. T. (1991) *Methods Enzymol. 208*, 604.
15. Anderson, W. F., Ohlendorf, D. H., Takeda, Y., and Matthews, B. W. (1981) *Nature 290*, 754.
16. Matsuo, H., Shirakawa, M., and Kyogoku, Y. (1995) *J. Mol. Biol. 254*, 668.

17. Huston, J. S., Levinson, D., Mudgett-Hunter, M., Tai, M. S., Novotny, J., Margolies, M. N., Ridge, R. J., Bruccoleri, R. E., Haber, E., Crea, R., et al. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 5879.
18. Simoncsits, A., Chen, J., Percipalle, P., Wang, S., Törö, I., and Pongor, S. (1997) *J. Mol. Biol. 267*, 118.
19. Robinson, C. R., and Sauer, R. T. (1996) *Biochemistry 35*, 109.
20. Robinson, C. R., and Sauer, R. T. (1996) *Biochemistry 35*, 13878.
21. Predki, P. F., and Regan, L. (1995) *Biochemistry 34*, 9834.
22. Pakula, A. A., and Sauer, R. T. (1989) *Proteins 5*, 202.
23. Mossing, M. C., and Sauer, R. T. (1990) *Science 250*, 1712.
24. Mollah, A. K., Aleman, M. A., Albright, R. A., and Mossing, M. C. (1996) *Biochemistry 35*, 743.
25. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) *Protein Sci. 4*, 2411.
26. Kim, J. G., Takeda, Y., Matthews, B. W., and Anderson, W. F. (1987) *J. Mol. Biol. 196*, 149.
27. Griko, Y. V., Rogov, V. V., and Privalov, P. L. (1992) *Biochemistry 31*, 12701.
28. Johnson, A., Meyer, B. J., and Ptashne, M. (1978) *Proc. Natl. Acad. Sci. U.S.A.* 75, 1783.
29. Spolar, R. S., and Record, M. T., Jr. (1994) *Science 263*, 777.
30. Kraulis, P. (1991) *J. Appl. Crystallography 24*, 946.
31. Ohlendorf, D. H., Tronrud, D. E., and Matthews, B. W. *J. Mol. Biol.* (in press).
32. Albright, R. A., and Matthews, B. W. *J. Mol. Biol.* (in press).
33. Albright, R. A., Mossing, M. C., and Matthews, B. W. (1996) *Biochemistry 35*, 734.

BI980152V