A Trade between Similar but Nonequivalent Intrasubunit and Intersubunit Contacts in Cro Dimer Evolution[†,‡]

Tracey Newlove, Kelly R. Atkinson, Laura O. Van Dorn, and Matthew H. J. Cordes*

Department of Biochemistry and Molecular Biophysics, University of Arizona, Tucson, Arizona 85721

Received December 13, 2005; Revised Manuscript Received March 23, 2006

ABSTRACT: The homodimeric λ Cro protein has a "ball-and-socket" interface that includes insertion of an aromatic side chain, Phe 58, from each subunit into a cavity in the hydrophobic core of the other subunit. This overlap between the subunit core and dimer interface hypothetically explains the strong dimerization and weak monomer stability of λ Cro in comparison to homologues. According to a model developed here and in a previous study [LeFevre, K. R., and Cordes, M. H. (2003) *Proc. Natl. Acad. Sci. U.S.A. 100,* 2345−2350], the socket cavity evolved in part by replacement of a buried tryptophan in an ancestral stable monomer with a smaller side chain (Ala 33 in λ Cro). The resulting core defect was in effect repaired by insertion of a different side chain (Phe 58) from a second subunit, generating the ball and socket. Consistent with such an evolutionary trade between intrasubunit and intersubunit interactions, we showed in the previous study that restoration of the ancestral Trp 33 in λ Cro stabilized the monomer and reduced the extent of dimerization. Here, we report the solution structure of a stable λ Cro monomer containing the Ala33Trp mutation, which confirms that the restored tryptophan fulfills its ancestral role as a core side chain, filling part of the socket cavity occupied by Phe 58 in the wild-type dimer. The structure also reveals, however, that the cavity is not completely filled by Trp 33, suggesting that its formation could have involved multiple mutations that reduced side chain volume. We offer suggestive evidence of a role of mutations at a second position.

Many proteins function as homooligomers. A number of mechanisms for oligomer evolution have been proposed, ranging from the formation of a protein−protein interface through surface mutations in an ancestral folded monomer (*1*, *2*) to domain swapping (*1*), involving the structural replacement of polypeptide segments in a monomer with equivalent segments from identical chains, to direct evolution of functional oligomers without folded monomeric ancestors (*2*). Comparative sequence and structure analyses in families with both homooligomeric and monomeric members provide a good opportunity to study these and other possible mechanisms.

The Cro family of bacteriophage transcription factors offers a complex instance of oligomer diversity among related proteins, including radical evolutionary changes both in the overall fold and in the dimer interface, along with changes in folding stability and dimerization strength. The Cro protein from bacteriophage λ binds DNA as a homodimer (*3*) and also dimerizes as a free protein (Figure 1B) (*4*). The monomer is populated only below low micromolar concentrations ($K_d = 0.3-7$ μM) (*5*−*7*) in vitro and is metastable toward unfolding, with a thermal denaturation midpoint (T_m) of ≤40 °C (*7*, *8*). The dimer interface includes an intermolecular β-sheet formed by part of a β-hairpin near the C-terminus, as well as a flexible "ball-and-socket" motif (*3*) that includes insertion of a Phe side chain from each subunit into a hydrophobic cavity in the other. The Cro protein from phage P22 (Figure 1A) also binds DNA as a homodimer (A. J. Miller, B. M. Hall, and M. H. J. Cordes, unpublished data) but in contrast to λ Cro does not dimerize measurably as free protein (*9*). The P22 Cro monomer is more stable than that of λ Cro ($T_m \sim 55$ °C) (*9*). The C-terminal β-hairpin which forms the dimer interface of λ Cro is replaced by a pair of α-helices in P22 Cro (Figure 1A) (*9*). On the basis of comparisons to distantly related transcription factors with similar structures in the CI family, these helices probably form weak intersubunit contacts (not shown) that are stable only when monomers bind adjacent sites on DNA (*10*−*13*).

The directionality of Cro structural evolution is known, as are some general aspects of the mutational mechanism (*9*). The all-α monomer fold of P22 Cro is common and is shared among the distantly related CI proteins (*14*−*18*), while the α+β fold thus far appears to be unique to λ Cro. Thus, the all-α monomer is probably ancestral, while the α+β fold and stronger oligomerization of λ Cro evolved from it. The sequences of λ Cro, P22 Cro, and other Cro proteins, though very diverse, are globally homologous and can be aligned in an evolutionarily relevant manner over nearly their entire lengths. Thus, the dramatic change from an all-α monomer to an α+β dimer was caused by relatively straightforward mutational events such as substitutions, rather than *en bloc* sequence changes such as heterologous recombinations or frameshifts (*9*).

In this study and in a previous paper (*7*), we have focused on specific mutational mechanisms for the evolution of λ

[†] Supported by NIH Grant RO1 GM066806.

[‡] Atomic coordinates have been deposited with the Protein Data Bank as entry 2A63.

* To whom correspondence should be addressed. Telephone: (520) 626-1175. Fax: (520) 626-9204. E-mail: cordes@email.arizona.edu.

10.1021/bi052541c CCC: $33.50
Published on Web 04/26/2006

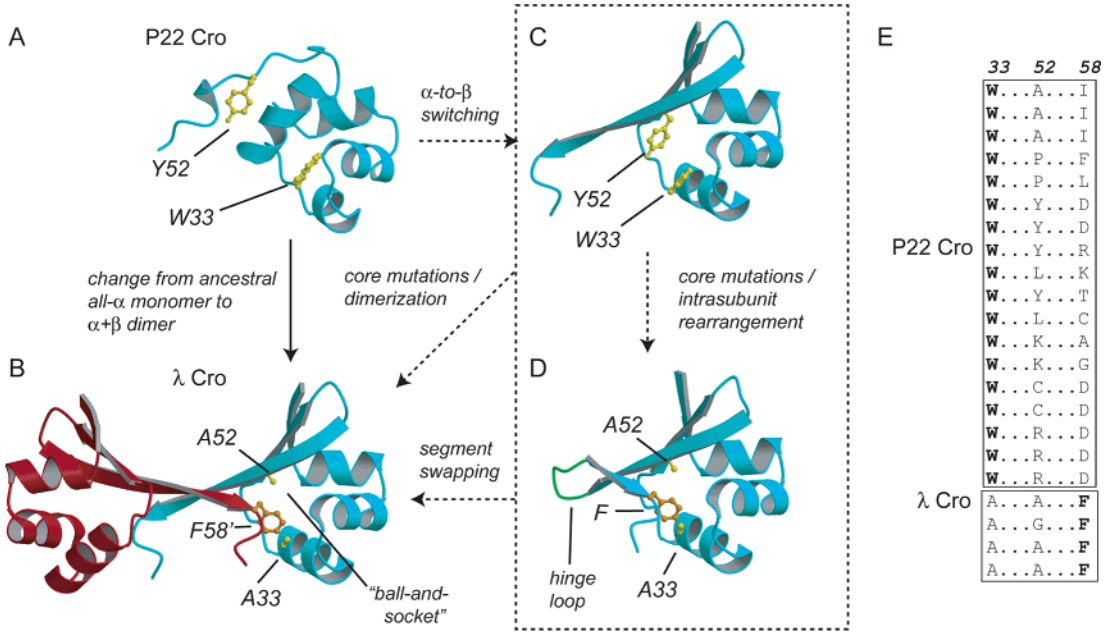

FIGURE 1: Hypothetical scenario of Cro structural evolution showing possible sequence determinants involved in changes in oligomerization strength and monomer stability. (A) The all-α structure of P22 Cro (1RZS) represents the ancestral Cro fold, while (B) the α+β fold and stronger dimer interface of λ Cro (6CRO) evolved in a Cro subfamily through structure switching in the C-terminal half of the domain. Key residues in the ball-and-socket dimer interface of λ Cro are shown, including two small side chains (Ala 33 and Ala 52, colored yellow) that line the socket and part of the ball (Phe 58) from a second subunit (side chain colored orange, second subunit backbone colored red). Counterparts of the socket side chains in P22 Cro are large aromatics (Trp 33 and Tyr 52, colored yellow; numbering is that of λ Cro). The two structures on the right show hypothetical evolutionary intermediates: (C) a model, generated using Deep View, of a hypothetical "filled-socket" monomer with Ala 33 and Ala 52 substituted with Trp and Tyr, respectively, from P22 Cro and (D) a domain-swapped monomer (1ORC) engineered through a hinge−loop insertion (*8*). Panel E shows part of a previously published alignment (*7*) highlighting sequence differences in residues corresponding to the λ ball and socket. Structures were drawn with MolScript (*42*) and Raster3d (*43*).

Cro's strong dimerization and weak monomer stability and the relationship of changes in these properties to the switch from α-helical to β-sheet secondary structure. Strong dimerization in λ Cro could conceivably result primarily from main chain hydrogen bonding interactions formed by the edges of its β-sheet (Figure 1B) (*19*) and might thus have arisen as a direct consequence of secondary structure switching. If, on the other hand, side chain interactions such as those in the ball and socket (Figure 1B) also stabilize the λ Cro dimer, specific sequence changes that contributed to dimer evolution may be separate from those which caused secondary structure evolution. If so, these two major events in Cro evolution may not have happened concurrently.

A comparison of λ Cro and P22 Cro supports the possible involvement of sequence substitutions in the ball and socket in dimer evolution, as well as the potential separability of secondary structure evolution and dimerization. Two small residues which line the socket cavity of λ Cro, Ala 33 and Ala 52, flank the Phe 58 side chain that protrudes from the other subunit (Figure 1B). The small size of these side chains may help maintain a cavity of sufficient size for favorable packing with Phe 58. In Cro sequence alignments (Figure 1E) (*7*), P22 Cro has large aromatic side chains at the corresponding positions (Trp 33 and Tyr 52 in λ numbering). These side chains appear to play stabilizing structural roles within (Trp 33) or on the periphery of (Tyr 52) the hydrophobic core of the all-α monomer (Figure 1A). If introduced into λ Cro, these aromatics would be expected to disrupt the dimer interface, but they can easily be modeled onto a single subunit (Figure 1C). In such an α+β λ Cro monomer model, both side chains are accommodated in the socket cavity and become partly buried in contact with other hydrophobic residues in the monomer interior. The Trp 33 and Tyr 52 side chains appear to occupy the same space as Phe 58 from the other subunit in the dimer and could be regarded as structural surrogates, trading interfacial hydrophobic interactions in a dimer for core interactions in a monomer. In short, it appears that substitution of large hydrophobic side chains from P22 Cro into λ Cro could convert it from a α+β dimer to an α+β monomer, and in the process might even stabilize the folded state of this monomer.

This modeling yields a speculative scenario for Cro structural evolution. An all-α monomeric ancestor resembling P22 Cro (Figure 1A) undergoes secondary structure switching to a hypothetical α+β monomer intermediate (Figure 1C). The side chains at positions 33 and 52 play no active part in this change, yet easily permit it because they play similar structural roles in both folds. Mutations then occur which reduce their size, forming the socket cavity and destabilizing the α+β monomer. A hydrophobic side chain at a different sequence position, corresponding to Phe 58 in λ Cro, is introduced into the socket cavity, acting as a surrogate for the lost ancestral side chains (Figure 1B). This replacement could occur either directly between subunits to yield a stable dimer (Figure 1B) or could occur first as a rearrangement within a single subunit (Figure 1D) to restore monomer stability, followed by domain−segment swapping to give dimer. Swapped λ Cro monomers (Figure 1D) have been generated in vitro by loop insertions that allow Phe 58 to interact with the socket of the same subunit (*8*, *20*, *21*). If such monomers were evolutionary intermediates, the dimer would not result directly from introduction of the ball-and-socket residues but would require additional mutations to

favor the intersubunit interface over its intrasubunit equivalent. Regardless of whether any α+β monomer intermediates (panel C or D of Figure 1) intervened, however, the ultimate outcome of this scenario is a trade of monomer folding stability for dimer interface strength. As such, it has the virtue of simultaneously explaining both the lower monomer stability and greater dimer strength of λ Cro relative to P22 Cro.

On the whole, this scenario is speculative, but certain aspects are supported by detailed sequence and structure analysis. Neglecting the possible role of position 52 for the moment, we have found considerable evidence to support the idea that Phe 58 in the λ Cro dimer interface acts as a structural surrogate for an ancestral Trp 33 in the core of an ancient Cro monomer. First, sequence analyses, including more distant relatives in the CI family as an outgroup, suggested strongly that Trp is the ancestral Cro residue at position 33 (*7*). Second, its ancestral structural role was probably as a hydrophobic core residue. The Trp is almost entirely buried in P22 Cro (*9*), and a homologous Trp with a nearly identical rotamer and structural role also occurs in the CI from P22 (not shown) (*18*). Third, consistent with Phe 58 acting as a structural surrogate for the ancestral Trp, loss of the Trp by replacement with Ala in a small subgroup of the Cros correlates with the appearance of Phe 58 as a conserved residue (Figure 1E) (*7*). Fourth, Trp 33 in the P22 Cro monomer and Phe 58 in the λ Cro dimer are close to each other in space relative to conserved portions of the structure (Figure 1A,B).

It might be possible to reverse the structural replacement of Trp 33 with Phe 58 through mutagenesis and thus convert λ Cro to a stable α+β monomer similar to that shown in Figure 1C. We previously substituted Ala 33 in λ Cro with the ancestral Trp and removed Phe 58 by replacing it with Asp, a polar side chain which was the most frequently occurring residue at this position in the Cro alignment (Figure 1E) (*7*). As one might anticipate, λ-A33W/F58D exhibited weakening of dimerization to at least millimolar levels, more similar to that of P22 Cro and much weaker than the K_d of 3 μM for wild-type λ Cro. Monomer thermal stability was also raised to a T_m of 52 °C, close to that of P22 Cro (∼55 °C) (*9*). In a variant containing only the A33W substitution, the same monomer stabilization was observed, indicating that the restored Trp caused the increased stability, presumably through occupation of the socket (Figure 1C). Interestingly, A33W alone did not abolish λ Cro dimerization but only attenuated it by 10-fold. To the extent that Trp 33 and Phe 58 were perfect structural surrogates, one might have expected severe loss of dimerization in λ-A33W due to expulsion of Phe 58 from the interface by Trp 33. Indeed, mutant models predicted a steric clash between Trp 33 and Phe 58 in a λ-A33W dimer, but the modest reduction in the extent of dimerization observed suggested that both might somehow be accommodated within the socket. Finally, circular dichroism (CD) spectra suggested retention of the α+β fold in the λ mutant monomers, such that disruption of the dimerization interface did not change the secondary structure of λ Cro to all-α.

We now report the solution structure of a λ Cro monomer containing the A33W and F58D mutations. The structure confirms the restoration of Trp 33 in its putative ancestral conformation and role as a hydrophobic core residue within the space corresponding to λ Cro's socket. This finding provides a clear structural basis for the observed increase in λ Cro monomer stability. The stabilized monomer retains an α+β fold despite the absence of any of the contacts present in the wild-type dimer interface, and despite the introduction of a hydrophobic core residue from the all-α homologue P22 Cro. To some extent at least, dimerization and secondary structure evolution appear to be separable processes in Cro proteins, each with its own sequence determinants. Trp 33 in the α+β monomer and Phe 58 in the wild-type α+β dimer bury a common set of core side chains and occupy partly equivalent regions of space, lending weight to the view that Phe 58 in the λ Cro dimer acts as a partial surrogate for the ancestral Trp 33. The monomer structure also shows, however, that this overlap is limited and that restoration of Trp 33 does not completely fill the socket cavity. This finding hints that, as drawn in our scenario (Figure 1), Phe 58 may replace more than one ancestral core side chain, perhaps including position 52 as well as position 33. Mutational studies at position 52, reported at the end of the paper, yield results consistent with such a model.

EXPERIMENTAL PROCEDURES

Mutagenesis and Protein Purification. Mutations were introduced into the pET21b-based (Novagen) λ Cro expression vector pMC140 (*7*) using the QuikChange mutagenesis protocol (Stratagene). To obtain protein, the resulting expression constructs were transformed into chemically competent *Escherichia coli* BL21(λDE3) cells. For unlabeled samples, freshly transformed cells were grown in 2−4 L of Luria-Bertani (LB) broth. For isotopically labeled samples, freshly transformed cells were grown in 2−4 L of M9T medium supplemented with 0.8 g/L [^{15}N]ammonium chloride or 3 g/L [$^{13}C_6$]glucose as the sole nitrogen and/or carbon source. Protein expression was induced by addition of 100 μg/mL IPTG[1] at an A_{600} of 0.55, and growth was typically continued for 3−6 h, followed by harvesting of the cells by centrifugation. Histidine-tagged Cro variants were then purified by denaturing Ni affinity chromatography as described previously (*7*). In the case of tagged proteins to be used in sedimentation equilibrium studies, size exclusion chromatography on a Sephacryl S-100 26/60 column was also performed.

Purification of untagged λ-A33W/F58D/Y26Q (also termed λ-WDQ) for NMR studies was performed using the following protocol. The cell pellets from 4 L of culture were resuspended in a total of 50 mL of lysis buffer [100 mM Tris (pH 8.0), 500 mM KCl, 1 mM EDTA, and 10 mM $MgCl_2$]; 100 μL of 100 mM PMSF in ethanol was added to the resuspension, and the cells were lysed by sonication (4 × 1 min bursts). An additional 100 μL of 100 mM PMSF was then added, and the sonicated mixture was diluted to a

[1] Abbreviations: NMR, nuclear magnetic resonance; CD, circular dichroism; rmsd, root-mean-square deviation; TOCSY, total correlation spectroscopy; NOESY, nuclear Overhauser effect spectroscopy; DQF-COSY, double-quantum-filtered correlation spectroscopy; HSQC, heteronuclear single-quantum coherence; NOESY-HSQC, two-dimensional (2D) and three-dimensional (3D) heteronuclear NOESY; TOCSY-HSQC, 2D and 3D heteronuclear TOCSY; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; IPTG, isopropyl β-D-thiogalactoside; PMSF, phenylmethanesulfonyl fluoride; PEI, polyethyleneimine; TMSP, sodium 3-trimethylsilylpropionate.

volume of 400 mL in lysis buffer. Four milliliters of 10% (w/v) PEI was added dropwise with stirring at 4 °C, followed by stirring for an additional 30 min, followed by centrifugation for 30 min at 14000*g*. Ammonium sulfate was slowly added to the supernatant, with stirring at 4 °C, to the extent of 55% saturation. The mixture was stirred for 1 h at 4 °C, followed by centrifugation for 30 min at 20000*g*. After the pellet had been discarded, the ammonium sulfate concentration was brought to 90% saturation, and the mixture was stirred for 6 h or overnight, followed by centrifugation for 30 min at 20000*g*. The pellet was resuspended in a minimal amount of PC buffer [20 mM Tris (pH 8.0), 0.1 mM EDTA, 5% glycerol, and 1.4 mM β-mercaptoethanol] and dialyzed extensively against the same. The dialysate was centrifuged for 15 min at 12000*g* to remove precipitated proteins. The supernatant was loaded onto a Mono S HR 10/10 column equilibrated in PC buffer, and the proteins were separated using a sodium chloride gradient. λ-WDQ eluted at approximately 200 mM salt. The pooled Mono S fractions were concentrated to a volume of 5 mL and loaded onto a preparative Sephacryl S-100 26/60 size exclusion column equilibrated in PC buffer with 200 mM NaCl. The λ-WDQ-containing fractions were >95% pure as judged by Coomassie R-250 staining of SDS−PAGE gels. The pooled fractions were dialyzed extensively against water. The dialysate was frozen at −75 °C and lyophilized, and the lyophilized protein was stored at −20 °C.

Analytical Ultracentrifugation and Circular Dichroism. Unless otherwise noted, experiments in this section were performed in 50 mM Tris (pH 7.5), 250 mM KCl, and 0.2 mM EDTA. Sedimentation equilibrium experiments were performed on a Beckman Optima XL-I analytical ultracentrifuge. All proteins were dialyzed extensively before centrifugation and blanked directly against dialysis buffer. Sedimentation was performed at 20 °C, at three protein concentrations in the range 50−250 μM, and at rotor speeds of 23 000 and 30 000 rpm. The buffer used for λ-WDQ was 50 mM sodium phosphate (pH 6.1) and 200 mM KCl. Data were measured as averages of 10−25 replicate radial scans at wavelengths of 280−300 nm, with a radial spacing of 0.001 cm. Sedimentation curves were fit to standard single-species models using Kaleidagraph (Synergy Software, Reading, PA) to obtain apparent molecular weights. Relevant parameters including solvent densities and protein partial specific volumes were computed from buffer or amino acid composition using SEDNTERP (J. Philo, Thousand Oaks, CA, and RASMB). For λ-WDQ, all individual curves at three concentrations and two rotor speeds gave apparent molecular weights within 3% of the theoretical monomer value, suggesting that the protein was cleanly monomeric. Thermal denaturation curves were measured on an Aviv 62A DS CD spectrometer by monitoring the circular dichroism signal at 222 nm as a function of temperature. Protein concentrations of 5−10 μM were used. Thermal denaturation midpoints (T_m) were obtained using nonlinear least-squares fitting to the Gibbs−Helmholtz equation using Kaleidagraph.

NMR Samples, Spectra, and Chemical Shift Assignments. Simple comparisons of ^{15}N−^{1}H correlation spectra for λ-A33W/F58D, λ-WDQ, and λ-A33W/F58D/A52Y (see Figure 2B and the text) were performed on a Varian Inova-600 spectrometer using 0.5 mM uniform ^{15}N-labeled, histidine-tagged samples in 50 mM sodium phosphate (pH 6) and 10% $^{2}H_2O$. Unless otherwise noted, all experiments used in the determination of the λ-WDQ solution structure were performed on a Bruker DRX-600 spectrometer using three untagged samples which contained either 5 mM unlabeled, 5 mM uniformly ^{15}N-labeled, or 2.5 mM uniformly ^{13}C-labeled λ-WDQ protein, in addition to 1 mM trimethylsilylpropionate (TMSP), 0.01% azide, 10% $^{2}H_2O$, and 50 mM sodium phosphate. The following spectra were obtained at 293 K and pH 5.3 (uncorrected meter reading): two-dimensional (2D) ^{1}H−^{1}H DQF-COSY, 2D ^{1}H−^{1}H NOESY (50, 100, and 150 ms mixing times), 2D ^{1}H−^{1}H TOCSY (80 ms mixing time), 2D ^{15}N−^{1}H HSQC, three-dimensional (3D) ^{15}N−^{1}H NOESY-HSQC (100 ms mixing time), 3D ^{15}N−^{1}H TOCSY-HSQC (80 ms mixing time), HNHA, HNHB (and reference spectrum), 2D ^{13}C−^{1}H HSQC (5−75 and 110−140 ppm), 2D constant-time ^{13}C−^{1}H HSQC (5−75 ppm), 3D ^{13}C−^{1}H NOESY-HSQC (7−47 ppm, 100 ms mixing time), and 2D ^{13}C−^{1}H TOCSY-HSQC (5−75 ppm, 35 and 70 ms mixing time; 105−145 ppm, 15 ms mixing time). 2D ^{1}H−^{1}H NOESY spectra were also collected at alternate temperatures (298 K) and pH values (pH 6.1) and in 100% $^{2}H_2O$, aiding in the assignment of some cross-peaks. Amide−hydrogen exchange experiments were performed at 293 K by resuspension of a lyophilized sample in pure $^{2}H_2O$, followed by periodic acquisition of ^{15}N−^{1}H HSQC spectra.

Spectra were processed using NMRPipe/NMRDraw (*22*) and analyzed using NMRView (*23*). Due to the small size of λ-WDQ, a complete set of ^{15}N, ^{13}C, and ^{1}H chemical shifts was obtained straightforwardly from the spectra listed above, with the exception of carbons and nitrogens without detectable protons attached. No stereospecific assignments were obtained directly, though a small number of β-methylene ^{1}H assignments were inferred during structural refinement (see below). ^{1}H chemical shifts were referenced to TMSP at 0.00 ppm, while ^{15}N and ^{13}C shift referencings were based on the ^{1}H values according to the method of Wishart et al. (*24*). We add a cautionary note that the TMSP-based referencing led to an unusually high value for the water ^{1}H signal (4.966 ppm at 293 K and pH 5.3). We suspect that the TMSP resonance may be shifted by as much as −0.15 ppm, possibly due to some transient interaction with the protein.

Restraint Assignments. Chemical shift degeneracies involving NOESY cross-peaks were resolved in part using an initial model of λ-WDQ based on the structure of one subunit from a λ Cro dimer. Atom pairs with distances of >8 Å in this model were eliminated as possible NOESY cross-peak assignments. Ambiguities were also resolved by examination of ^{15}N and ^{13}C isotope-edited NOESY experiments as well as 2D ^{1}H−^{1}H NOESY spectra obtained at different temperatures and pH values or in 100% $^{2}H_2O$. Cross-peak intensities from a 50 ms 2D ^{1}H−^{1}H NOESY spectrum were calibrated to interatomic distances using tyrosine and phenylalanine $d_{\delta\epsilon}$ NOEs and assuming a $1/r^6$ distance dependence of the intensity. Assigned cross-peaks were then binned into strong (1.8−2.8 Å), medium (1.8−3.4 Å), weak (1.8−4.5 Å), or very weak (1.8−6.0 Å) restraint classes. Some restraints in the very weak category were visible only in longer mixing time NOESY spectra. Restraints assignable only in 100 ms isotope-edited NOESY spectra were given bounds of 1.8−5.0 Å.

$^{3}J_{HN\alpha}$ values were obtained from analysis of HNHA spectra as described previously (*25*). A dihedral angle

restraint of $-90° \leq \phi \leq -40°$ was assigned to residues with $^3J_{HNH\alpha}$ values of <6.0 Hz. A dihedral angle restraint of $-160° \leq \phi \leq -80°$ was assigned to residues with $^3J_{HNH\alpha}$ values of >8.0 Hz. $^3J_{HN\beta}$ values were derived from HNHB spectra, with quantitative values obtained by comparison of volume integrals with those from a 2D reference spectrum (*26*). A dihedral angle restraint of $140° \leq \chi_1 \leq 220°$ was assigned to six residues for which both of two β-methylene protons had $^3J_{HN\beta}$ values of <2.0 Hz. For residues that had one β-methylene proton with a $^3J_{HN\beta}$ of <2.0 Hz and the other with a $^3J_{HN\beta}$ of >4.0 Hz, χ_1 angles near either 60° or −60° were possible. For four of these ambiguous cases, structure calculations during the late stages of refinement yielded no accepted structures with rotamers in which $\chi_1 \sim 60°$. In the final stages of refinement, we therefore assigned these residues dihedral angle restraints of $-20° \leq \chi_1 \leq -100°$. This also permitted stereospecific assignments of the two β-methylene protons for these four residues. Ambiguities involving three isoleucine and three threonine residues with a $^3J_{HN\beta}$ of <2.0 Hz for the single β resonance were also resolved iteratively, leading likewise to assignment of dihedral angle restraints of $-20° \leq \chi_1 \leq -100°$. For one additional residue, Asp 47, one large and one small $^3J_{HN\beta}$ were observed, and NOE data were more qualitatively consistent with a χ_1 of ∼60° thana χ_1 of −60°. We performed separate calculations in which restraints corresponding to each possibility were imposed and found that restricting the conformation to $-20° \leq \chi_1 \leq -100°$ led to small violations or near violations of the restraint in all calculated structures, while restricting the conformation to $20° \leq \chi_1 \leq 100°$ did not. Comparisons to the wild-type structure also favored a χ_1 of ∼60° for this residue. For these reasons, we decided to apply a $20° \leq \chi_1 \leq 100°$ restraint to the Asp 47 side chain in the final structure calculation.

Amide protons were considered to be hydrogen-bonded if they were visible in the first $^{15}N-^1H$ correlation spectrum following resuspension in ^2H_2O of a lyophilized H_2O sample. In α-helical regions, hydrogen bond acceptors were identified by observation of appropriate backbone $d_{\alpha N}$ NOEs (*27*). In β-sheet regions, acceptors were assigned on the basis of the known wild-type λ Cro sheet topology. Hydrogen bonds were imposed as a pair of distance restraints: 1.8−2.5 Å between the carbonyl oxygen and the amide proton and 1.8−3.5 Å between the carbonyl oxygen and the amide nitrogen.

Structure Calculations. Using the standard simulated annealing protocol in CNS 1.1 (*28*), 40 structures of λ-WDQ were calculated on the basis of an extended starting structure with random variation of initial velocity. A total of 919 experimental restraints were used, including 838 NOE distances, 14 hydrogen bond distances, 50 ϕ angles, and 17 χ_1 angles (Table 1). All 40 calculations converged to structures with no NOE violations of >0.5 Å and no dihedral angle restraint violations of $>5°$. Fourteen of these 40 structures were discarded on the basis of poor agreement of the observed rotamer of Val 55 with a small $^3J_{NH\beta}$ coupling constant, even though no explicit restraint on χ_1 for Val 55 had been included in the calculation. Of the 26 remaining structures, the 20 with the lowest energy were included in the ensemble reported in Table 1 and Figure 3B. The six discarded structures exhibited both higher energies and unusual conformations that were unreasonable and in strong disagreement with previously published structures of λ Cro variants. Ramachandran statistics (Table 1) were obtained using PROCHECK-NMR (*29*). Superpositions and rmsds were performed and calculated using MOLMOL (*30*).

^{15}N Relaxation Studies. ^{15}N relaxation experiments for λ-WDQ were carried out on a 5 mM uniformly ^{15}N-labeled sample at pH 5.3 and 293 K, using a Varian Inova-600 spectrometer and standard published methods (*31*, *32*). In measuring T_1 values, we carried out separate experiments with delays of 0, 30, 60, 120, 240, 480, and 750 ms. In measuring T_2 values, we carried out separate experiments with delays of 10, 30, 50, 70, 90, 130, and 190 ms. T_1 and T_2 values for each residue were computed by analyzing the decay in peak intensity as a function of the delay, using the Levenberg−Marquardt algorithm as implemented in the rate analysis functionality in NMRView (*23*). "Jitter" mode was used to allow for slight variations in peak location. The overall correlation time (τ_m) was computed from an adjusted average T_1/T_2 ratio as described in Results, using the -l option in the program *r2r1_tm*, obtained from the website of A. Palmer (Columbia University, New York; http://cpmcnet.columbia.edu/dept/gsas/biochem/labs/palmer/software/r2r1_tm.html). In measuring $^{15}N[^1H]NOE$ values, we acquired two spectra identically except for a 2.5 s recycle delay during which proton saturation was either present or absent. NOE values were calculated from ratios of peak heights between spectra measured with and without ^1H saturation.

Structural Analysis and Modeling. Accessible surface area calculations, used here to compare the structural roles of Phe 58 and Trp 33, were performed using MOLMOL (*30*). Models of steric clashes between Trp 33 and Phe 58 in λ-A33W dimers (Figure 7) were generated by in silico mutation of the Ala 33 side chain to Trp in the O monomer of the OB dimer of the free protein (PDB entry 5CRO) or in one subunit of the DNA-bound dimer (PDB entry 6CRO). Using MOLMOL, the side chain conformation of the Trp was fixed to the ensemble average value for λ-WDQ (−71, −13).

RESULTS

Initial NMR Studies of λ-A33W/F58D. For initial structural studies of λ-A33W/F58D, we prepared uniformly ^{15}N-labeled samples of histidine-tagged protein and readily assigned ∼90% of the backbone ^1H resonances from 3D ^{15}N-edited NOESY and TOCSY spectra. A 2D $^{15}N-^1H$ correlation spectrum of histidine-tagged λ-A33W/F58D (pH 6.1 and 293 K) is shown in Figure 2B. The resonances are well-dispersed and indicate a specific folded structure. However, examination of one-dimensional (1D) ^1H spectra (not shown) showed that λ-A33W/F58D had slightly broader than expected amide ^1H line widths. Some amide resonances, such as Gln 16, Gln 27, and Ile 30, were quite weak in intensity in the 2D $^{15}N-^1H$ correlation spectrum and difficult to assign unambiguously (Figure 2B). Optimization of solution conditions such as pH, temperature, and salt produced only minor improvements in spectral quality. Moreover, many weak peaks (∼5% of the volume of the major peaks) were observed at ^1H chemical shifts of 7.7−8.5 in the $^{15}N-^1H$ correlation spectrum (Figure 2B), suggestive of a small population of polypeptides with a random coil configuration. In addition, in 2D $^1H-^1H$ NOESY spectra, weak cross-peaks were

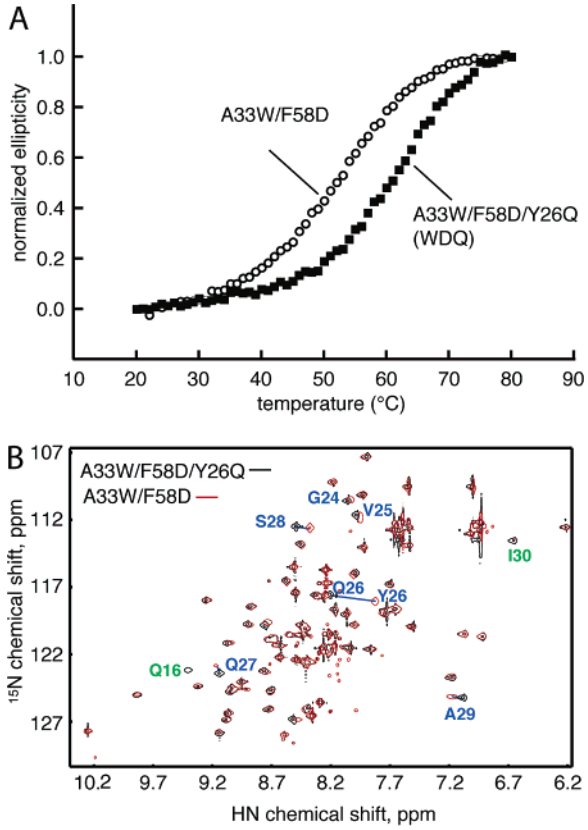

FIGURE 2: Structurally nonperturbing stabilization of the monomeric λ-A33W/F58D mutant by introduction of a Y26Q substitution. (A) Thermal denaturation curves of λ-A33W/F58D and λ-A33W/F58D/Y26Q, monitored by circular dichroism at 222 nm. Normalized ellipticities are shown. (B) $^{15}N-^{1}H$ correlation spectra of λ-A33W/F58D (red) and λ-A33W/F58D/Y26Q (black), showing differences in residue positions only near the site of mutation (blue labels). The comparison also shows more intense signals for certain residues in λ-A33W/F58D/Y26Q (green labels). The weak cluster of peaks in the center of the spectrum is also reduced in intensity upon introduction of the Y26Q mutation.

Table 1: Structural Statistics for the λ Cro A33W/F58D/Y26Q (λ-WDQ) Solution Structure

Distance Restraints	
total	838
intraresidue	285
sequential ($\|i-j\|=1$)	231
medium-range ($2 \le \|i-j\| \le 4$)	155
long-range ($\|i-j\|>4$)	167
hydrogen bond distances	14
Dihedral Restraints	
$\phi(C'_{i-1}-N_i-C_{\alpha i}-C'_i)$	50
$\chi(N_i-C_{\alpha i}-C_{\beta i}-C_\gamma)$	17
Stereospecific Assignments	
β-methylene group	4
Average rmsds	
distance restraints (Å)	0.017 ± 0.0004
dihedral angle restraints (deg)	0.077 ± 0.0169
idealized covalent geometry	
bonds (Å)	0.002 ± 0.00003
angles (deg)	0.310 ± 0.0035
impropers (deg)	0.128 ± 0.0078
Energies[a]	
total	61.39 ± 2.59 (79.22)
NOE	18.44 ± 0.99 (31.3)
dihedral	0.03 ± 0.01 (0.04)
bond	3.84 ± 0.14 (5.04)
angle	28.12 ± 0.63 (29.57)
van der Waals	9.60 ± 1.48 (12.26)

Ramachandran Plot[a,b]

	no. of residues	%
most favorable region	863 (42)	88.1 (85.7)
additionally allowed region	117 (7)	11.9 (14.3)
generously allowed region	0.0 (0)	0.0 (0)
disallowed region	0.0 (0)	0.0 (0)

Average rmsds of Atomic Coordinates of 20 Structures, Residues 3−55[c]

backbone heavy atoms (Å)	0.66 (0.45)
all heavy atoms (Å)	1.29 (0.89)

[a] Values in parentheses are for the minimized average of the 20 ensemble structures. [b] Region of the Ramachandran plot as defined by PROCHECK-NMR, for residues 3−55. [c] Values given are average pairwise rmsds. Values in parentheses are rmsds from the mean.

observed between certain aromatic side chain resonances and frequencies which did not correspond to any assigned aromatic or amide shift but which were consistent with random coil values for the correlated proton.

One possible source of some or all of these effects is a small population of unfolded protein (~5%) in slow or intermediate exchange with the native state. In fact, analysis of thermal denaturation curves of λ-A33W/F58D using the Gibbs−Helmholtz equation suggests that despite a T_m of 52 °C (Figure 2A), its free energy of unfolding near room temperature is not much greater than 2 kcal/mol. This allows for an equilibrium in which several percent of the protein molecules are unfolded. To test whether stability was related to the NMR effects described above, we sought to generate variants of λ-A33W/F58D which would exhibit increased thermodynamic stability toward unfolding but would otherwise be structurally unaffected.

On the basis of the work of Pakula and Sauer (*33*) on stabilizing replacements of the highly solvent-exposed residue Tyr 26, we decided to introduce a Y26Q substitution into the λ-A33W/F58D background. This sequence change had been found to raise the thermal denaturation midpoint of wild-type Cro by 8 °C (*33*). A 2D $^{15}N-^{1}H$ correlation spectrum of λ-A33W/F58D/Y26Q shows resonance positions almost identical to those of λ-A33W/F58D, with the exception of the sequence region immediately adjacent to the mutation (Figure 2B). The Y26Q mutation stabilizes λ-A33W/F58D by 9 °C toward thermal denaturation as monitored by far-ultraviolet CD (Figure 2A), similar to its effect on the wild-type protein. These findings suggest that the Y26Q substitution is both structurally nonperturbing and stabilizing. Consistent with the hypothesis that the unwanted NMR effects seen in λ-A33W/F58D resulted from a small population of unfolded protein, λ-A33W/F58D/Y26Q exhibited improved NMR spectral quality, including narrower ^{1}H line widths (not shown), strengthened resonance intensities (Figure 2B), and a strong lessening of small "random coil" peaks near the center of the spectrum (Figure 2B).

Like the parent variant λ-A33W/F58D, λ-A33W/F58D/Y26Q exhibited little propensity to form dimers in sedimentation equilibrium experiments (not shown). For histidine-tagged λ-A33W/F58D, a minor population of higher-order species was previously detected at the highest concentration that was examined [283 μM protein, 50 mM Tris (pH 7.5), 250 mM KCl, and 0.2 mM EDTA] but not at lower concentrations (30−100 μM protein) (*7*). The K_d for a putative monomer−dimer equilibrium was estimated to be

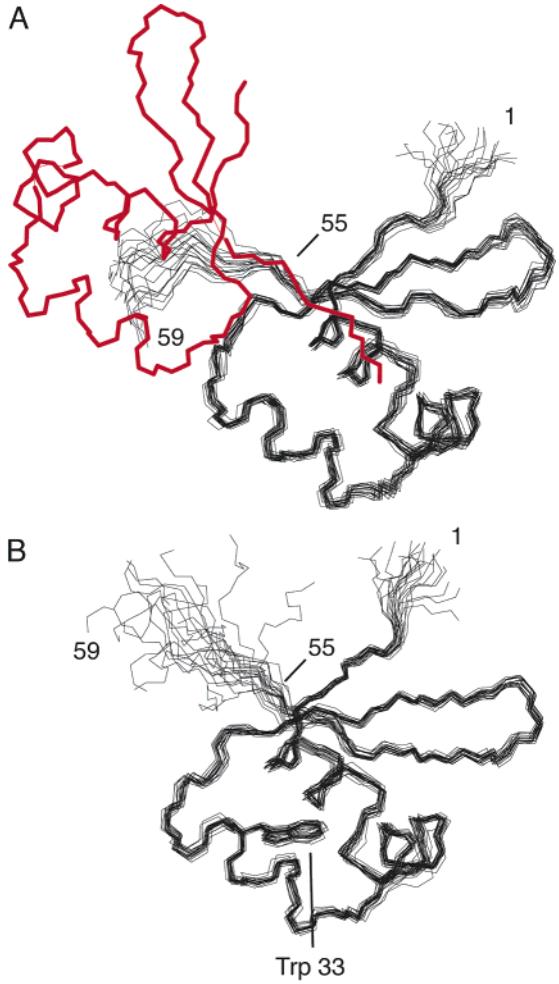

FIGURE 3: Comparison of (A) wild-type λ Cro dimer (1COP) (*34*) and (B) λ-WDQ monomer solution ensembles. Both ensembles contain 20 structures with residues 1−59 shown. Only one half of the wild-type dimer ensemble is shown, with one member of the ensemble for the second subunit shown (red) to indicate the location of the dimer interface. The comparison illustrates the greater disorder for residues 56−59 in λ-WDQ, due to the absence of interactions with a second subunit. The side chain of Trp 33 in λ-WDQ is also shown to illustrate its well-defined position. This figure was generated using MOLMOL (*30*).

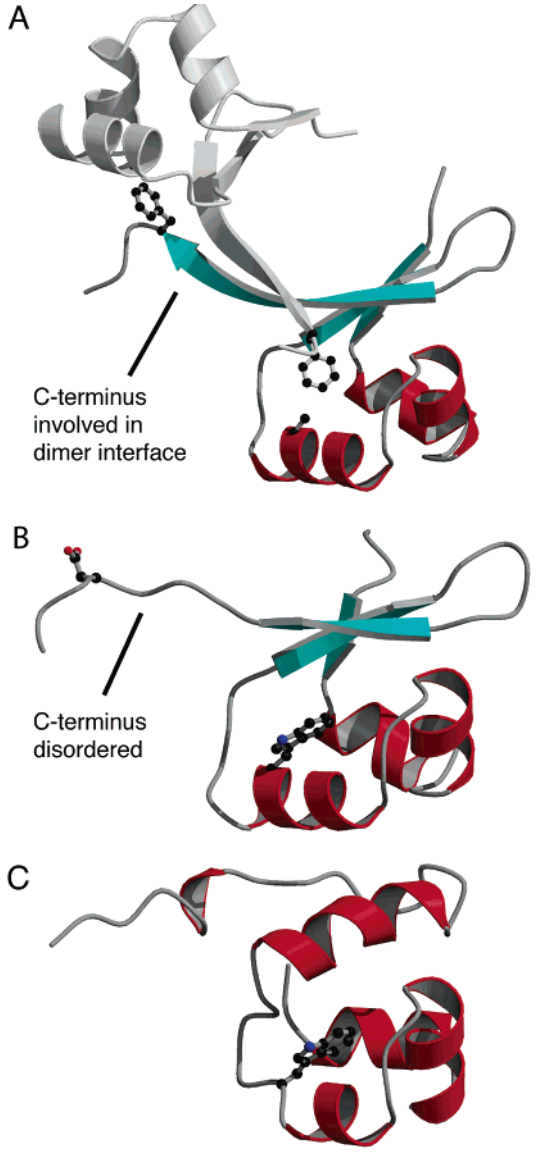

FIGURE 4: Ribbon diagrams of (A) wild-type λ Cro dimer crystal structure (5CRO, with chain O colored and chain B in gray) (*4*), (B) a representative structure from the λ-WDQ monomer solution ensemble, and (C) a minimized average solution structure of P22 Cro (1RZS) (*9*). For wild-type λ Cro, Phe 58 side chains from both subunits are shown, along with Ala 33 from the one subunit. For λ-WDQ, the mutated side chains of Trp 33 and Asp 58 are shown. For P22 Cro, Trp 33 (λ Cro numbering) is shown to illustrate the similarity of its side chain conformation to that of Trp 33 in λ-WDQ.

very weak, ≥2 mM. Precisely comparable experiments were not performed here for λ-A33W/F58D/Y26Q, but sedimentation curves of an untagged protein at similar salt concentrations [not shown; 50−250 μM protein, 50 mM sodium phosphate (pH 6.1), 200 mM KCl, and 23000−30000 rpm] showed no evidence of self-association at all, giving an average apparent molecular mass of 7355 Da as compared to the theoretical monomer molecular mass of 7411 Da. The very weak self-association detected for histidine-tagged λ-A33W/F58D previously (*7*) may have been an artifact of the tag, but in any case, it is clear that dimerization is essentially abolished in both the λ-A33W/F58D and λ-A33W/F58D/Y26Q variants.

Solution Structure of the λ-WDQ Monomer and Analysis of the General Fold. We determined the NMR solution structure of λ-A33W/F58D/Y26Q (henceforth termed λ-WDQ) rather than λ-A33W/F58D, reasoning that the atomic coordinates might be more precisely determined for λ-WDQ due to higher spectral quality. λ-WDQ was purified to >95% homogeneity in untagged form (see Experimental Procedures), yielding unlabeled and isotopically labeled (^{13}C and ^{15}N) samples with a concentration of up to 5 mM at pH 5.3. NMR spectra obtained for λ-WDQ, as well as assignment and restraint generation methods, are detailed in Experimental Procedures. Forty structures were calculated for λ-WDQ by simulated annealing from a combination of NOE distance restraints, coupling-derived angle restraints, and hydrogen bond distance restraints. Twenty structures were selected for the reported solution structure ensemble and are shown superimposed in Figure 3B. A ribbon diagram for a representative structure is shown in Figure 4B. Statistics for the structure determination are listed in Table 1.

Before analyzing the structure in detail, we make two important general observations. First, the structure of the λ-WDQ monomer is the first of a stable λ Cro monomer that is not domain swapped and thus lacks any interactions

analogous to those of the wild-type dimer interface (compare panels A and B of Figure 4). Second, the overall backbone structure in the λ-WDQ monomer is qualitatively similar to that of subunits in the wild-type Cro dimer solution and crystal structures (compare panels A and B of Figure 3 and panels A and B of Figure 4). For regions which are well-defined in λ-WDQ (residues 3−55; see Figure 3B), C_α rmsds between the λ-WDQ minimized average and each of the four monomers in the asymmetric unit of the wild-type Cro dimer crystal structure (*4*), as well as the single monomer in the asymmetric unit of the DNA-bound wild-type crystal structure (*3*), range from 1.03 to 1.23 Å. These are not major structural differences. As discussed below, the largest difference between λ-WDQ and wild-type λ Cro subunits is that the third β-strand is shorter in λ-WDQ, being disordered beyond residue 55. In sum, λ-WDQ confirms that the wild-type α+β fold can be retained and even stabilized in the absence of dimer interactions, consistent with the potential separability of secondary structure and oligomerization determinants in Cro evolution, and the plausibility of stable α+β monomer intermediates such as that shown in Figure 1C.

Comparison of the λ-WDQ Solution Structure to the Wild-Type Solution Structure. We next compared the general character, resolution, and quality of the λ-WDQ ensemble to those of the NMR-derived solution structure of wild-type λ Cro determined by Matsuo et al. (*34*) (Figure 3A, all structures from one of two subunits overlaid, with a representative second subunit in red). λ Cro is a 66-residue protein, and residues 62−66 are disordered in all Cro variant structures, whether determined by NMR or X-ray methods (*3*, *4*, *20*, *21*, *34*, *35*). The solution ensemble of the wild-type λ Cro homodimer shows some degree of order from approximately residue 3 to 59. For residues 56−59, which include Phe 58, some variation in the coordinates is observed relative to the rest of the subunit. However, the local backbone conformation is qualitatively consistent and resembles that in Cro dimer crystal structures, including a cis conformation for Pro 59, which is sometimes considered along with Phe 58 to be part of the "ball" in the ball and socket (*3*). The only long-range contacts for residues 56−59 are between subunits in the dimer, and the less precise coordinate definition in this region may result in part from a paucity of intersubunit NOEs (*34*).

In λ-WDQ, Phe 58 has been mutated to the polar Asp, destroying dimerization. The region of residues 56−59 might be expected to show greater disorder in the λ-WDQ monomer than in the wild-type dimer, because it is not anchored by any long-range intrasubunit contacts. Indeed, λ-WDQ NOESY spectra show no medium- or long-range NOEs involving this region. No qualitative backbone conformation is shared by the λ-WDQ ensemble members for residues 56−59, as illustrated by the more scattered nature of the backbone coordinates relative to the wild-type subunits (compare panels A and B of Figure 3). Asp 58 in particular shows a wide range of $\phi-\psi$ angle combinations in the ensemble. Both Pro 57 and Pro 59 appear to have trans peptide configurations as judged by the value of the chemical shift difference, $\delta[^{13}C_\beta] - \delta[^{13}C_\gamma]$ (5.7 and 3.9 ppm, respectively) (*36*). Wild-type λ Cro has a cis peptide bond for Pro 59. There is evidence of a minor conformer (~10%) in this region of λ-WDQ as judged from the presence of a weak set of peaks in the proline $^{13}C_\delta-^1H_\delta$ region of a $^{13}C-^1H$ correlation spectrum. This is probably a cis isomer of one of the prolines. The dominance of trans Pro peptide conformations with minor populations of cis isomers reinforces the impression of a heterogeneous, disordered, non-wild-type-like structure in this region.

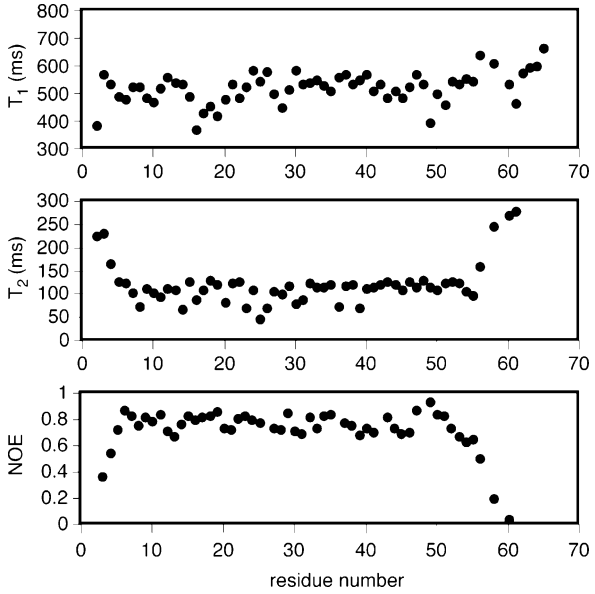

FIGURE 5: Backbone dynamics of the λ-WDQ monomer. T_1, T_2, and $^{15}N[^1H]$ NOE values were measured as described in Experimental Procedures. NOE values for residues 2 and 61−66 were negative and do not appear on the graph at the scale shown. Similarly, T_1 values for residue 66 and T_2 values for residues 62−66 are off-scale in the positive direction and are not shown. NOE values for several residues, e.g., Leu 42, were omitted from this analysis due to adventitious spectral streaks that inhibited quantitative intensity measurements.

As a consequence of this disorder, the λ-WDQ ensemble has poorer overall resolution than the individual wild-type subunits across the region of residues 3−59 (λ-WDQ, 1.07 Å rmsd to the mean for backbone, 1.45 Å for heavy atoms; wild type, 0.66 Å rmsd to the mean for backbone, 1.07 Å for heavy atoms). When only residues 3−55 are considered, however, the λ-WDQ and wild-type λ Cro ensembles have identical, fairly high resolution (0.45 Å rmsd for backbone atoms and 0.9 Å for all heavy atoms; Table 1). The stereochemical quality of this region for λ-WDQ is good, with almost 90% of residues in most favored regions of a Ramachandran plot (Table 1).

Backbone Dynamics of the λ-WDQ Monomer. The global solution dynamics of λ-WDQ are as expected for a monomer of its size and shape. Relatively uniform ^{15}N T_1, ^{15}N T_2, and $^{15}N[^1H]$ NOE values were observed across residues 5−55 of λ-WDQ (Figure 5), corresponding closely to the well-defined region of the ensemble. For this region, NOE values averaged 0.77 ± 0.07, T_1 values averaged 514 ± 47 ms, and T_2 values averaged 108 ± 20 ms. An overall rotational correlation time (τ_m) for λ-WDQ was calculated by averaging the T_1/T_2 ratio across all residues and discarding those with values more than one standard deviation from the mean, including residues 2−4 and 56−66, along with some residues with short T_2 values in the region of residues 5−55. An adjusted average T_1/T_2 ratio (4.41 ± 0.09) was then figured and used to calculate a τ_m of 5.86 ± 0.08 ns (*37*). This value resembles the τ_m of 5.6 ns reported by Mossing for the

reverse domain-swapped λ Cro monomer shown in Figure 1D (*21*).

Residues 2−4 at the N-terminus and 56−66 at the C-terminus of λ-WDQ showed significant NOE reductions and increases in T_2, suggestive of increased mobility, in approximate agreement with the location of poorly defined regions in the ensemble. In these experiments, T_1 appears to be less sensitive to the increased internal motion than T_2 or the NOEs, showing only slight increases at the C-terminus (a larger increase in T_1 is seen for residue 66, but this is off-scale in Figure 5). In the wild-type λ Cro dimer, longer T_2 values and lower NOEs at the N-terminus are generally similar to those in λ-WDQ, but at the C-terminus, significant NOE reductions and increases in T_2 values begin to occur at approximately residue 61, not at residue 56 (not shown) (*38*). Consequently, residues between positions 56 and 60 are probably more mobile in λ-WDQ than in the wild-type dimer, in agreement with the appearance of the ensembles (Figure 3).

Conformation and Structural Role of Trp 33. The restored ancestral Trp 33 in λ-WDQ has a well-defined side chain conformation (see Figure 3B) nearly identical to the putative ancestral rotamer observed in P22 Cro (compare panels B and C of Figure 4) (*9*) and P22 CI (not shown) (*18*). Specifically, the ensemble average side chain angles (χ_1 and χ_2) observed for Trp 33 in λ-WDQ are −71° and −13°, respectively, while those for the homologous Trp residues in P22 Cro and P22 CI are −77° and −12°, and −68° and −17°, respectively. In the P22 CI and P22 Cro ensembles, an average of only 3−4% of the surface area of the Trp is accessible to solvent. Trp 33 in λ-WDQ is also largely buried, with an ensemble average of 14% of its surface area accessible to solvent. This burial is achieved by sequestering the Trp side chain within the socket, which in the dimer is occupied by Phe 58 of the other subunit (Figure 6A). Eight socket-lining side chains, most of them hydrophobic, are within 6 Å, heavy atom distance, of Trp 33: Leu 7, Leu 23, Val 25, Ala 29, Ile 30, Arg 38, Ile 40, and Ala 52. A total of 82 Å^2 of accessible surface area of these residues is buried through contact with Trp 33 (Figure 6B), as measured by comparisons with a model in which Trp 33 in λ-WDQ is changed back to Ala.

These interactions provide a structural basis for thermal stabilization of the λ Cro monomer upon mutation of Ala 33 to Trp. In effect, the restoration of Trp 33 restores its ancestral structural role as a stabilizing hydrophobic core residue in the monomeric state. In the forward time direction, the replacement of the ancestral Trp with Ala can be viewed as akin to removal of a critical hydrophobic core residue. Such a mutation results in a loss of the favorable folding free energy associated with the transfer of the Trp side chain from solvent to the hydrophobic interior of the protein and, in the absence of any compensatory structural adjustments, also creates an unfavorable cavity in the core. The ability of the Trp to play its ancestral role is independent of the overall fold (see panels B and C of Figure 4), and its destabilizing loss could easily have occurred either from stable all-α monomer ancestors or from stable α+β monomer intermediates.

Partial Surrogacy of Phe 58 for the Ancestral Trp 33 Side Chain. Comparisons of the λ-WDQ monomer with the wild-type dimer lend some support to the model that Phe 58 in wild-type Cro (Figure 6A) acts as a surrogate for the lost ancestral Trp 33. Like Trp 33 in the λ-WDQ monomer, Phe 58 in the wild-type dimer is largely buried, with ∼6% accessible surface area. Residues showing significant burial (>5 Å^2) by Trp 33 in λ-WDQ are also significantly buried by Phe 58 in the wild-type dimer (Figure 6B), and the total degree of burial for this set of six residues is fairly similar (78 Å^2 for Trp 33 and 102 Å^2 for Phe 58). Thus, in the sense of surface area burial, the interactions made by Trp 33 in its ancestral conformation are well covered by the intersubunit contacts made by Phe 58.

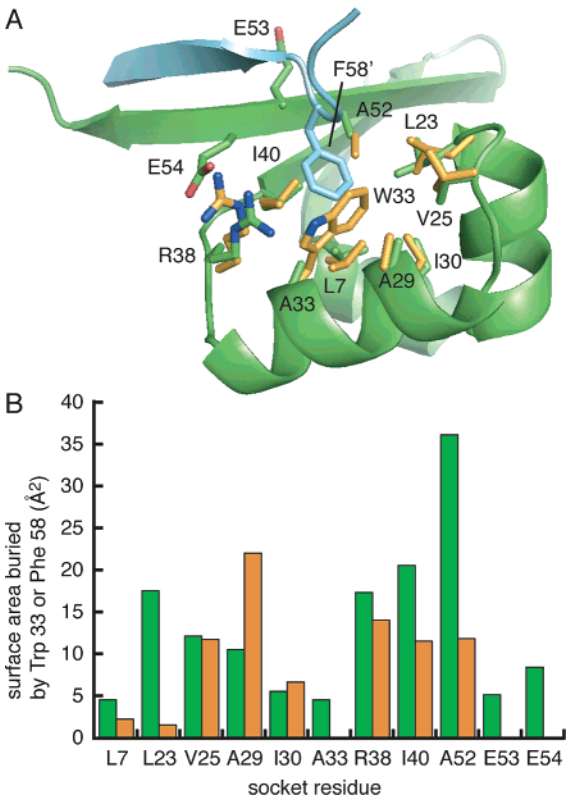

FIGURE 6: Partial surrogacy of Phe 58 for the ancestral Trp 33. (A) Superposition of socket regions from the wild-type Cro dimer and λ-WDQ monomer. The green cartoon shows backbone and socket side chains of one subunit (5CRO, chain O) in the wild-type dimer. The light blue cartoon shows part of the second wild-type dimer subunit (5CRO, chain B), including Phe 58. Orange side chains are socket residues from the λ-WDQ monomer (minimized average solution structure), superimposed on those of the wild-type protein. This panel was generated using PyMol (DeLano Scientific, San Carlos, CA). (B) Similarities and differences in surface area of socket residues buried by Phe 58 in the wild-type dimer and Trp 33 in the λ-WDQ monomer (see Experimental Procedures).

At the same time, this surrogacy is imperfect. Despite the qualitative similarities in burial, there are quantitative differences. For example, Ala 29 is considerably more buried by Trp 33 than by Phe 58, while Ala 52 is exactly the opposite (Figure 6B). The actual overlap in space covered by Phe 58 and Trp 33 is fairly small as judged by the overlay in Figure 6A. Moreover, there are residues such as Leu 23 and Glu 54 which clearly border the cavity into which Phe 58 fits but which are buried very little or not at all by Trp 33 in λ-WDQ (Figure 6B). As a consequence, the overall amount of burial of socket side chains by Phe 58 in the wild type is greater (125−136 Å^2) than the amount of burial by

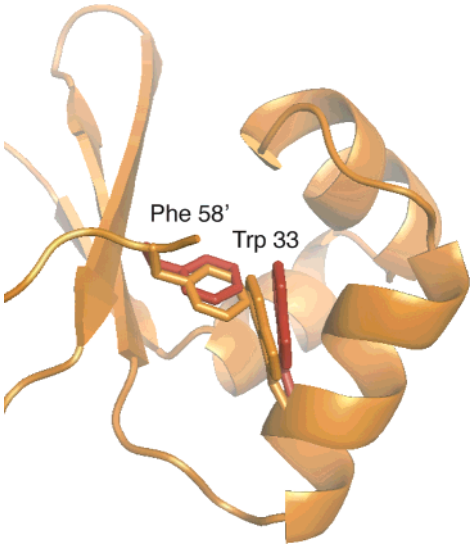

FIGURE 7: Models of an Ala33Trp mutation in free (orange; 5CRO, chains O and B) and DNA-bound (red; 6CRO) wild-type Cro dimer, showing differences in predicted steric clashes between Trp 33 and Phe 58 in the dimer interface. The ribbon diagram is from the structure of the free protein. The side chain conformation of Trp 33 is that found in the λ-WDQ minimized average solution structure. This figure was generated using PyMol.

Trp 33 in λ-WDQ (82 Å^2). This suggests that Phe 58's role may not be limited to simple replacement of the ancestral Trp 33, an idea which is discussed further below. The socket structure is also slightly different in this region between the two variants, with the backbone of Ala 52 being pushed back slightly in the dimer, perhaps to accommodate Phe 58, and with the wild type and λ-WDQ having different rotamers for Leu 23 (Figure 6A). On the other side of the socket cavity, Arg 38 appears to be displaced slightly away from the socket in λ-WDQ relative to the wild type, possibly due to a steric conflict with the Trp 33 side chain.

Had the above comparisons suggested a more perfect overlap between Trp 33 in λ-WDQ and Phe 58 in the wild type, one might have expected the A33W mutation by itself to abolish dimerization in λ Cro, which it does not. λ-A33W still dimerizes fairly strongly, with a K_d of 28 μM, approximately an order of magnitude weaker than the wild-type protein. Only when the F58D mutation is also incorporated is dimerization destroyed (*7*). Thus, Trp 33 and Phe 58 can coexist within the dimer interface at some level, though the close approach seen in the overlay of Figure 6A suggests that structural adjustments might be required to accommodate both side chains. A model of a Trp 33 mutation in a wild-type λ Cro dimer structure (5CRO, chains O and B; same Trp rotamer as in λ-WDQ) shows severe steric clashes (Figure 7) (*4*). Twenty-four heavy atom distances under 3.0 Å are seen, some as short as 0.8 Å. These clashes involve every carbon of the Phe 58 ring and every carbon in the six-membered ring of Trp 33. However, it has been noted (*4*) that the dimer linkage of λ Cro is flexible, on the basis of comparisons between free and DNA-bound forms. Indeed, a λ-A33W dimer model based on the DNA-bound wild-type structure (6CRO) (*3*) predicts somewhat milder conflicts, with the closest heavy atom approach of Trp 33 and Phe 58 being ~2.0 Å (Figure 7). The flexibility of the ball and socket may underlie the relatively small experimentally observed effects of Trp 33 on dimerization strength.

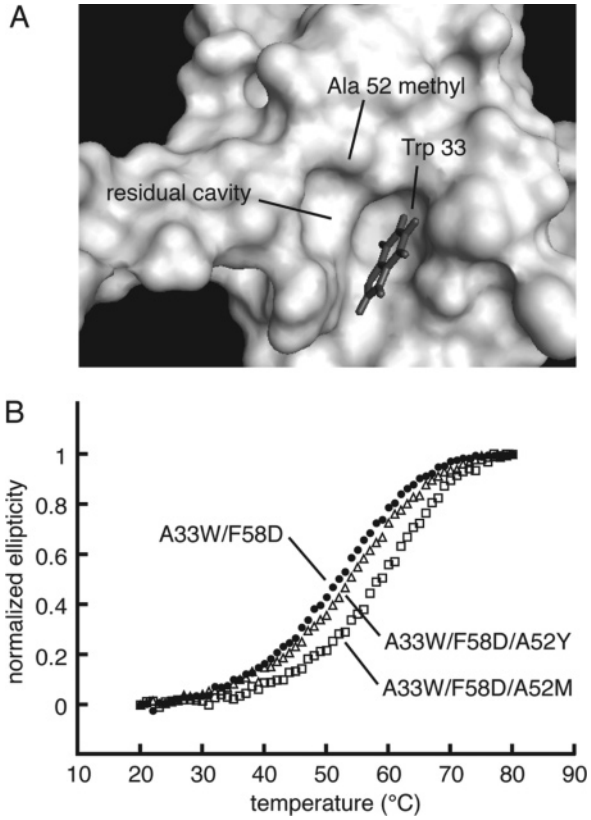

FIGURE 8: Incomplete filling of the socket cavity by Trp 33 allows for further monomer stabilization by mutation of Ala 52 to larger side chains. (A) Surface diagram of λ-WDQ minimized average solution structure generated using PyMol. The portion of the surface corresponding to Trp 33 has been omitted to illustrate the surface of a full socket cavity similar to that present in one subunit of a wild-type dimer. The Trp 33 side chain of λ-WDQ is shown filling part of this depression, but a substantial residual cavity remains adjacent to Ala 52. (B) Thermal denaturation curves, monitored by circular dichroism at 222 nm, for position 52 mutants introduced into the λ-A33W/F58D sequence. Both A52Y and A52M mutations lead to thermal stabilization, presumably due to filling of the residual cavity.

Effect of Mutations at Position 52. The socket cavity may be too large to be filled by restoring the ancestral Trp 33. The surface of λ-WDQ shows a residual cavity remaining after introduction of Trp 33 (Figure 8A). This depression is adjacent to Ala 52 (Figure 8A), a small side chain lining the socket cavity like Ala 33. In our introductory scenario (Figure 1), we proposed a possible evolutionary role for position 52 similar to that for position 33, with an ancestral buried or partly buried hydrophobic side chain being mutated to an alanine to modify the socket cavity. Models (Figure 1C) suggested that Tyr 52 of P22 Cro could fit into the socket cavity along with Trp 33. However, we had originally neglected position 52 in our experimental studies (*7*), because a number of factors rendered its evolutionary history and ancestral role more ambiguous than those of position 33. Position 52 shows poor conservation in Cro alignments (Figure 1E) (*7*), making a specific ancestral residue type difficult or impossible to identify. It is partially rather than completely buried in P22 Cro (*9*), and some Cros have polar residues at this position, casting some doubt on a putative ancestral role as a buried hydrophobic residue in an all-α monomer. Here, we study mutations at position 52 but limit ourselves to addressing the basic plausibility of its proposed role in the scenario of Figure 1.

An A52Y mutation introduced into the λ-A33W/F58D background increases monomer thermal stability by 2 °C (Figure 8B) to 54 °C, showing that a large side chain at position 52 can reinforce the effect of Trp 33 and further stabilize the monomer (Figure 1C). Introduction of Met, a more flexible hydrophobic chain, at position 52 gives a larger increase in thermal stability to 59 °C (Figure 8B), suggesting that other hydrophobics may be more optimal than Tyr for packing into the residual cavity. $^{15}N-^{1}H$ HSQC spectra of λ-A33W/F58D and λ-A33W/F58D/A52Y (not shown) exhibited only minor differences, confirming that the A52Y mutation does not convert the monomer to a grossly different structure but instead acts to slightly stabilize the α+β fold.

The A52Y mutation also reinforces the effect of the A33W mutation on dimerization strength. An A52Y mutation introduced into the λ-A33W background reduces the level of dimerization from a K_d of 28 μM to at least millimolar levels as measured by sedimentation equilibrium (data not shown). At 50 and 100 μM concentrations of λ-A33W/A52Y average apparent molecular masses from single-ideal species fits at 23 000 and 30 000 rpm were 8550 and 8548 Da, respectively. Both numbers are within ∼1% of the theoretical monomer molecular mass of 8635 Da. The average apparent molecular mass of λ-A33W/A52Y at 200 μM was 9041 Da, slightly higher. Although it is not possible to accurately determine a dimerization constant from these data, we estimate that the value for K_d must be ≥2 mM, based on monomer−dimer fits at 200 μM. Thus, λ-A33W/A52Y exhibits dimerization that is at least 2 orders of magnitude weaker than that of λ-A33W.

These results show that two hybrid substitutions of P22 Cro residues into λ Cro convert it from a ball-and-socket dimer to a monomer which is similar to P22 Cro in terms of its thermal stability and tendency to dimerize. However, the double mutant retains λ Cro's α+β fold rather than converting to the all-α secondary structure of P22 Cro. The effects are consistent with a model of λ-A33W/A52Y like that proposed in our introductory scenario, in which the aromatic side chains fill the socket cavity (Figure 1C). A larger side chain at position 52 stabilizes the monomer by occupying part of the socket much in the way that Trp 33 confers stability by occupying the remainder. The larger side chains combine to exclude Phe 58 from the socket, inhibiting dimerization.

DISCUSSION

Our findings support the plausibility of two aspects of the scenario presented in Figure 1. First, restoration of an ancestral tryptophan (Trp 33) to the subunit core of λ Cro stabilizes the monomer by partly filling a hydrophobic cavity which is normally a key part of the wild-type dimer interface. In so doing, Trp 33 occupies a region of space partly equivalent to that occupied by Phe 58 from a second subunit in the wild-type homodimer. The restored ancestral Trp 33 in λ-WDQ and Phe 58 in the wild type contact and bury many of the same side chains in the cavity, suggesting that they play partly surrogate stabilizing roles in the monomer and dimer, respectively. Restoration of the ancestral Trp 33 and removal of Phe 58 convert λ Cro to a stable monomer which dimerizes weakly, similar to the putative ancestral Cro protein. These results support a model in which the strong dimerization and low monomer stability of λ Cro are in part the result of an evolutionary trade, in which intrasubunit hydrophobic core interactions made by the ancestral Trp 33 are exchanged for nonequivalent intersubunit contacts made by Phe 58. We have also suggested that Phe 58 may replace interactions made by two ancestral side chains in the subunit core, rather than just Trp 33 (see Figure 1).

Our second major finding relates to the plausibility of intermediates in Cro structural evolution. The evolution of strong dimerization and low monomer stability in λ Cro is one component of an elaborate structural transformation that also includes a change from the ancestral all-α to an α+β fold (Figure 1). Our results suggest some separability of the changes in oligomerization and secondary structure. For example, the structure of λ-WDQ clearly shows that λ Cro can retain its α+β fold in the absence of dimer interface interactions, though part of the third β-strand becomes disordered. Since the structure of λ Cro had been determined previously only as a dimer (*3*, *4*) (Figure 1B) or a reverse domain-swapped monomer (*20*, *21*, *35*) (Figure 1D), one might have hypothesized that a λ Cro monomer could adopt an all-α fold like that of P22 Cro. If this were true, Cro secondary structure differences might be more the result of a conformational switching event associated with oligomerization and less the product of mutationally induced evolution of individual subunit folds. Our results show that there must be important sequence determinants of Cro secondary structure distinct from those which govern dimerization.

To further emphasize this point, substitution of two interfacial alanines in the α+β, dimeric λ Cro with aligned aromatic residues from the all-α, monomeric P22 Cro converts the former not to an all-α monomer or all-α dimer but to a more stable α+β monomer (modeled in Figure 1C). λ-A33W/A52Y shows no significant dimerization and has a thermal stability essentially equal to that of P22 Cro, much higher than that of the wild-type λ Cro monomer. The restored ancestral tryptophan at position 33 has a similar conformation and plays a similar structural role in the α+β Cro monomer as in the all-α P22 Cro. The role of the introduced tyrosine at position 52 is less certain but is presumed to involve further monomer stabilization by filling the residual cavity (Figure 8A) left after the introduction of Trp 33. In such a role, its conformation and level of burial would be similar to those of Tyr 52 in P22 Cro (see Figure 1). Consequently, it is unlikely that sequence changes at positions 33 and 52 would have been key direct determinants of secondary structure evolution in Cros. It would have been structurally plausible for a stable all-α monomer to switch to a stable α+β monomer intermediate (Figure 1C) while maintaining the same aromatic/aliphatic residues at these two positions. Introduction of the λ Cro alanines at positions 33 and 52 likewise would not be expected to strongly perturb the equilibrium between α+β and all-α monomers, destabilizing both. Instead, the alanine side chains probably act directly as determinants of dimerization and low monomer stability, by promoting formation of functionally competent dimers within the α+β scaffold.

The structure presented here is the first of a stable monomeric λ Cro variant completely lacking the interactions present in the wild-type dimer interface. Because this monomer was generated solely by mutations in the ball and socket, it may be tempting to negate the importance of other

interactions in dimerization, including main chain hydrogen bonding as well as cross-strand hydrophobic and ionic interactions in the intersubunit β-sheet. In our view, such a conclusion would not be justified by our findings. For example, it is important to remember that the F58D mutation present in λ-WDQ goes beyond simple removal of Phe 58 and introduces unfavorable burial of a polar residue in the dimer interface. A λ Cro variant without the ball and socket, but also without disruptive polar side chain burial, might therefore still dimerize detectably. We did, however, report previously that λ-A33W/F58A also exhibits very weak dimerization (*7*), which seems to downplay the importance of the polar disruption. But even if other interactions in the interface are not sufficient to significantly populate dimers at micromolar concentrations in the absence of a properly formed ball and socket, they may still contribute favorable free energy of self-association. In other words, our results highlight the ball and socket as a determinant of dimerization in Cro, but the relative importance of other possible determinants has yet to be studied fully.

An intertwined dimer interface like that in λ Cro (Figure 1B) is sometimes interpreted as suggesting evolution by 3D domain swapping (*1*). Domain swapping involves an exchange of intrasubunit interactions for equivalent contacts in a multimer and has been termed segment swapping in cases where a very limited region is exchanged (*2*). The case for domain swapping in the evolution of an oligomer is somewhat strengthened when monomeric homologues are known that appear swapped relative to the oligomer. No known naturally occurring monomeric Cros have this relationship to the λ Cro dimer. However, stable segment-swapped monomers of λ Cro (Figure 1D) can be engineered through insertion mutations that introduce a hinge loop (*8*, *20*, *21*). Mossing and co-workers have also reported evidence suggesting that some version of the ball and socket may even form weakly in the metastable λ Cro wild-type monomer (*6*).

Our hypothetical scenario for Cro dimer evolution (Figure 1) does not negate the possible involvement of domain−segment swapping but does suggest that there is no necessity of invoking it. Rather than being formed from the direct swapping of equivalent, homologous intra- and intersubunit contacts, induced for example by a hinge loop deletion (Figure 1D to Figure 1B), the dimer interface of λ Cro may have arisen (Figure 1C to Figure 1B) from the direct exchange of similar but *nonequivalent* contacts, induced by side chain mutations. If generalized, the latter pathway for λ Cro oligomer interface evolution might be termed a "surrogacy" mechanism (or perhaps a "patching" mechanism where a significant destabilization of the monomer is involved). It will be interesting to see if other dimer interfaces, especially those containing a ball-and-socket motif (*39*−*41*), show evidence of having evolved in this fashion.

ACKNOWLEDGMENT

We thank Vahe Bandavian for helpful comments on the manuscript and Neil Jacobsen for technical support with NMR experiments.

REFERENCES

1. Bennett, M. J., Schlunegger, M. P., and Eisenberg, D. (1995) 3D domain swapping: A mechanism for oligomer assembly, *Protein Sci. 4*, 2455−68.
2. Xu, D., Tsai, C. J., and Nussinov, R. (1998) Mechanism and evolution of protein dimerization, *Protein Sci. 7*, 533−44.
3. Albright, R. A., and Matthews, B. W. (1998) Crystal structure of lambda-Cro bound to a consensus operator at 3.0 Å resolution, *J. Mol. Biol. 280*, 137−51.
4. Ohlendorf, D. H., Tronrud, D. E., and Matthews, B. W. (1998) Refined structure of Cro repressor protein from bacteriophage lambda suggests both flexibility and plasticity, *J. Mol. Biol. 280*, 129−36.
5. Darling, P. J., Holt, J. M., and Ackers, G. K. (2000) Coupled energetics of lambda cro repressor self-assembly and site- specific DNA operator binding I: Analysis of cro dimerization from nanomolar to micromolar concentrations, *Biochemistry 39*, 11500−7.
6. Jana, R., Hazbun, T. R., Mollah, A. K., and Mossing, M. C. (1997) A folded monomeric intermediate in the formation of lambda Cro dimer-DNA complexes, *J. Mol. Biol. 273*, 402−16.
7. LeFevre, K. R., and Cordes, M. H. (2003) Retroevolution of lambda Cro toward a stable monomer, *Proc. Natl. Acad. Sci. U.S.A. 100*, 2345−50.
8. Mossing, M. C., and Sauer, R. T. (1990) Stable, monomeric variants of λ Cro obtained by insertion of a designed β-hairpin sequence, *Science 250*, 1712−5.
9. Newlove, T., Konieczka, J. H., and Cordes, M. H. (2004) Secondary structure switching in Cro protein evolution, *Structure 12*, 569−81.
10. Aggarwal, A. K., Rodgers, D. W., Drottar, M., Ptashne, M., and Harrison, S. C. (1988) Recognition of a DNA operator by the repressor of phage 434: A view at high resolution, *Science 242*, 899−907.
11. Shimon, L. J., and Harrison, S. C. (1993) The phage 434 OR2/R1-69 complex at 2.5 Å resolution, *J. Mol. Biol. 232*, 826−38.
12. Rodgers, D. W., and Harrison, S. C. (1993) The complex between phage 434 repressor DNA-binding domain and operator site OR3: Structural differences between consensus and non-consensus half-sites, *Structure 1*, 227−40.
13. Beamer, L. J., and Pabo, C. O. (1992) Refined 1.8 Å crystal structure of the lambda repressor-operator complex, *J. Mol. Biol. 227*, 177−96.
14. Huang, G. S., and Oas, T. G. (1995) Structure and stability of monomeric lambda repressor: NMR evidence for two-state folding, *Biochemistry 34*, 3884−92.
15. Mondragon, A., Subbiah, S., Almo, S. C., Drottar, M., and Harrison, S. C. (1989) Structure of the amino-terminal domain of phage 434 repressor at 2.0 Å resolution, *J. Mol. Biol. 205*, 189−200.
16. Neri, D., Billeter, M., and Wuthrich, K. (1992) Determination of the nuclear magnetic resonance solution structure of the DNA-binding domain (residues 1 to 69) of the 434 repressor and comparison with the X-ray crystal structure, *J. Mol. Biol. 223*, 743−67.
17. Pabo, C. O., and Lewis, M. (1982) The operator-binding domain of lambda repressor: Structure and DNA recognition, *Nature 298*, 443−7.
18. Sevilla-Sierra, P., Otting, G., and Wuthrich, K. (1994) Determination of the nuclear magnetic resonance structure of the DNA-binding domain of the P22 c2 repressor (1 to 76) in solution and comparison with the DNA-binding domain of the 434 repressor, *J. Mol. Biol. 235*, 1003−20.
19. Richardson, J. S., and Richardson, D. C. (2002) Natural β-sheet proteins use negative design to avoid edge-to-edge aggregation, *Proc. Natl. Acad. Sci. U.S.A. 99*, 2754−9.
20. Albright, R. A., Mossing, M. C., and Matthews, B. W. (1996) High-resolution structure of an engineered Cro monomer shows changes in conformation relative to the native dimer, *Biochemistry 35*, 735−42.
21. Mossing, M. C. (1998) Solution structure and dynamics of a designed monomeric variant of the lambda Cro repressor, *Protein Sci. 7*, 983−93.
22. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR 6*, 277−93.
23. Johnson, B. A., and Blevins, R. (1994) NMRView. A computer program for the visualization and analysis of NMR data, *J. Biomol. NMR 4*, 603−14.
24. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) ^{1}H, ^{13}C

and ^{15}N chemical shift referencing in biomolecular NMR, *J. Biomol. NMR 6*, 135−40.
25. Vuister, G. W., and Bax, A. (1993) Quantitative *J* correlation: A new approach for measuring homonuclear three-bond $J(H^NH^\alpha)$ coupling constants in ^{15}N-enriched proteins, *J. Am. Chem. Soc. 115*, 7772−7.
26. Bax, A., Vuister, G. W., Grzesiek, S., Delaglio, F., Wang, A. C., Tschudin, R., and Zhu, G. (1994) Measurement of homo- and heteronuclear J couplings from quantitative J correlation, *Methods Enzymol. 239*, 79−105.
27. Arseniev, A., Schultze, P., Worgotter, E., Braun, W., Wagner, G., Vasak, M., Kagi, J. H., and Wuthrich, K. (1988) Three-dimensional structure of rabbit liver [Cd7]metallothionein-2a in aqueous solution determined by nuclear magnetic resonance, *J. Mol. Biol. 201*, 637−57.
28. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D54*, 905−21.
29. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR, *J. Biomol. NMR 8*, 477−86.
30. Koradi, R., Billeter, M., and Wuthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures, *J. Mol. Graphics 14*, 51−5, 29−32.
31. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation, *Biochemistry 33*, 5984−6003.
32. Kay, L. E., Nicholson, L. K., Delaglio, F., and Bax, A. (1992) The effects of cross-correlation between dipolar and chemical shift anisotropy relaxation mechanisms on the measurement of heteronuclear T_1 and T_2 values in proteins: Pulse sequences for the removal of such effects, *J. Magn. Reson. 97*, 359−75.
33. Pakula, A. A., and Sauer, R. T. (1990) Reverse hydrophobic effects relieved by amino-acid substitutions at a protein surface, *Nature 344*, 363−4.
34. Matsuo, H., Shirakawa, M., and Kyogoku, Y. (1995) Three-dimensional dimer structure of the lambda-Cro repressor in solution as determined by heteronuclear multidimensional NMR, *J. Mol. Biol. 254*, 668−80.
35. Albright, R. A., Mossing, M. C., and Matthews, B. W. (1998) Crystal structure of an engineered Cro monomer bound nonspecifically to DNA: Possible implications for nonspecific binding by the wild-type protein, *Protein Sci. 7*, 1485−94.
36. Schubert, M., Labudde, D., Oschkinat, H., and Schmieder, P. (2002) A software tool for the prediction of Xaa-Pro peptide bond conformations in proteins based on ^{13}C chemical shift statistics, *J. Biomol. NMR 24*, 149−54.
37. Kay, L. E., Torchia, D. A., and Bax, A. (1989) Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: Application to staphylococcal nuclease, *Biochemistry 28*, 8972−9.
38. Matsuo, H., Sugeta, H., Shirakawa, M., and Kyogoku, Y. (1996) Backbone dynamics of the λ-Cro repressor protein determined by ^{15}N relaxation measurements. Application of an efficient method for calculation of dynamics parameters, *J. Mol. Struct. 379*, 143−50.
39. Lesk, A. M., and Chothia, C. (1988) Elbow motion in the immunoglobulins involves a molecular ball-and-socket joint, *Nature 335*, 188−90.
40. Vargo, M. A., Nguyen, L., and Colman, R. F. (2004) Subunit interface residues of glutathione S-transferase A1-1 that are important in the monomer−dimer equilibrium, *Biochemistry 43*, 3327−35.
41. Knowlton, J. R., Bubunenko, M., Andrykovitch, M., Guo, W., Routzahn, K. M., Waugh, D. S., Court, D. L., and Ji, X. (2003) A spring-loaded state of NusG in its functional cycle is suggested by X-ray crystallography and supported by site-directed mutants, *Biochemistry 42*, 2275−81.
42. Kraulis, P. J. (1991) MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures, *J. Appl. Crystallogr. 24*, 946−50.
43. Merritt, E. A., and Bacon, D. J. (1997) Raster3D: Photorealistic Molecular Graphics, *Methods Enzymol. 277*, 505−24.

BI052541C