

---

# Independent tyrosyl contributions to the CD of Ff gene 5 protein and the distinctive effects of Y41H and Y41F mutants on protein–protein cooperative interactions

---

TUNG-CHUNG MOU,<sup>1</sup> NARASIMHA SREERAMA,<sup>2</sup> THOMAS C. TERWILLIGER,<sup>3</sup>  
ROBERT W. WOODY,<sup>2</sup> AND DONALD M. GRAY<sup>1</sup>

<sup>1</sup>Department of Molecular and Cell Biology, The University of Texas at Dallas, Richardson, Texas 75083-0688, USA

<sup>2</sup>Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, Colorado 80523, USA

<sup>3</sup>Biosciences Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

(RECEIVED July 24, 2001; FINAL REVISION November 22, 2001; ACCEPTED November 22, 2001)

## Abstract

The gene 5 protein (g5p) of the Ff virus contains five Tyr, individual mutants of which have now all been characterized by CD spectroscopy. The protein has a dominant tyrosyl 229-nm L<sub>a</sub> CD band that is shown to be approximately the sum of the five individual Tyr contributions. Tyr41 is particularly important in contributing to the high cooperativity with which the g5p binds to ssDNA, and Y41F and Y41H mutants are known to differ in dimer–dimer packing interactions in crystal structures. We compared the solution structures and binding properties of the Y41F and Y41H mutants using CD spectroscopy. Secondary structures of the mutants were similar by CD analyses and close to those derived from the crystal structures. However, there were significant differences in the binding properties of the two mutant proteins. The Y41H protein had an especially low binding affinity and perturbed the spectrum of poly[d(A)] in 2 mM Na<sup>+</sup> much less than did Y41F and the wild-type gene 5 proteins. Moreover, a change in the Tyr 229 nm band, assigned to the perturbation of Tyr34 at the dimer–dimer interface, was absent in titrations with the Y41H mutant under low salt conditions. In contrast, titrations with the Y41H mutant in 50 mM Na<sup>+</sup> exhibited typical CD changes of both the nucleic acid and the Tyr 229-nm band. Thus, protein–protein and g5p–ssDNA interactions appeared to be mutually influenced by ionic strength, indicative of correlated changes in the ssDNA binding and cooperativity loops of the protein or of indirect structural constraints.

**Keywords:** Circular dichroism; Ff gene 5 protein; tyrosine CD bands; protein–protein cooperative interactions; secondary structure analysis

---

Reprint requests to: Donald M. Gray, Department of Molecular and Cell Biology, Mail Stop FO31, The University of Texas at Dallas, P.O. Box 830688, Richardson, TX 75083-0688; e-mail: dongray@utdallas.edu; fax: (972) 883-2409.

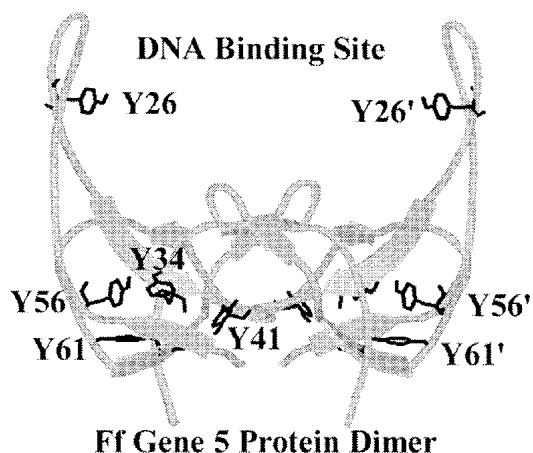
**Abbreviations:** CD, circular dichroism; DSSP, Define Secondary Structures of Proteins computer program (Kabsch and Sander, 1983);  $\Delta\epsilon$ ,  $\epsilon_L - \epsilon_R$  (or molar CD); Ff, filamentous phages (M13, fd, and f1) that require F pili to infect *E. coli*; g5p, gene 5 protein;  $K\omega$ , the intrinsic binding constant  $K$  times a cooperativity factor  $\omega$ ;  $n$ , the number of nucleotides bound per g5p monomer; OB fold, oligonucleotide/oligosaccharide binding fold; PDB, Protein Data Bank; [P]/[N], [protein monomer]/[nucleotide] molar ratio; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; ssDNA, single-stranded DNA; WT, wild-type protein.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.30002>.

Single-stranded DNA (ssDNA) binding proteins are essential for regulation of many biological processes such as DNA replication and gene expression. Among the ssDNA-binding proteins, the gene 5 protein (g5p) encoded by Ff viruses has been particularly amenable to biophysical and biochemical studies because of its small size, abundance, and solubility. Dimers of g5p bind with a cooperativity factor ( $\omega$ ) that could be as high as 5000 (Terwilliger 1996) to coat newly synthesized ssDNA genomes during viral morphogenesis. In addition, g5p binds to leader sequences and represses translation of five of the viral mRNAs (Zaman et al. 1990, 1991). Three binding modes have been observed

for the g5p, with stoichiometries of  $n = 4, 3,$  or  $2-2.5$  nucleotides per g5p monomer. The  $n = 2-2.5$  binding mode is weak, and is dissociated at salt concentrations of 0.1 M or less (Kansy et al. 1986; Sang and Gray 1989; Thompson et al. 1998). The binding affinity of g5p is dependent on the ssDNA sequence, and is inversely related to the nearest-neighbor base stacking tendencies in (A + C)-containing ssDNA sequences (Mou et al. 1999).

In the absence of nucleic acid, the three-dimensional structure of g5p has been determined by X-ray crystallography and NMR spectroscopy (Folkers et al. 1994; Skinner et al. 1994). The Ff g5p monomer is a relatively compact OB-fold  $\beta$ -structure of only 87 amino acids, and it exists as a stable dimer in solution even at a concentration as low as 1 nM (Murzin 1993; Terwilliger 1996). Aromatic residues of the protein consist of five Tyr, three Phe, and no Trp. The five Tyr are conserved or conservatively substituted in the related IKe and Pf3 ssDNA binding proteins (de Jong et al. 1989). Two Tyr (Tyr56 and Tyr61) are largely buried in the interior of the g5p. The other three Tyr (Tyr26, Tyr34, and Tyr41) are located on the protein surface. Tyr26 is directly involved in binding ssDNA, and Tyr34 and Tyr41 are involved in protein-protein cooperative interactions. Tyr41, located in a flexible dimer-dimer contact loop, is particularly important in stabilizing dimer-dimer interactions. The Y41H mutant protein has improved solubility, making possible high-resolution NMR structural studies (Folkers et al. 1994). The locations of the five Tyr within the g5p dimer are shown in Figure 1.

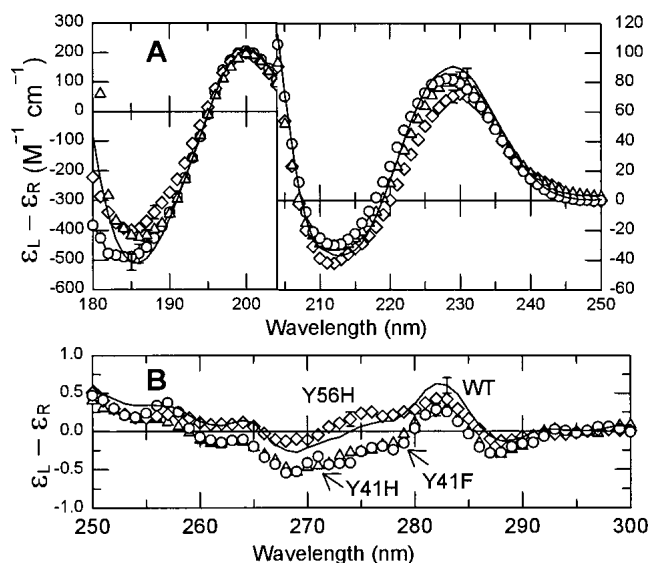


**Fig. 1.** Ribbon model of the Ff gene 5 protein dimer in the complex described by Guan et al. (1995). The view is perpendicular to the twofold axis of the dimer. Locations of the five tyrosines in each monomer are shown. Tyrosines in the second monomer of the dimer are marked with primes. The DNA-binding wings hold up to four base pairs on each of two segments of single-stranded DNA that are directed into and out of the plane of the paper. Tyr34 and Tyr41 are involved in cooperative dimer-dimer interactions. In this view, Tyr41 and Tyr41', respectively, extend toward and away from the viewer on the cooperativity loops of the protein.

The CD spectrum of the wild-type g5p contains an atypical positive CD band at 229 nm, due mainly to the  $L_a$  bands of the five tyrosyl residues (Day 1973; Woody 1978). Prior CD studies of four single Tyr→Phe mutants, Y26F, Y34F, Y41F, and Y61F, showed that each has a positive 229-nm CD band, but with a slightly (<10%) reduced magnitude compared with the wild-type g5p. Our initial inference was that Tyr56 might dominate this CD band, but this was without taking into account the CD contributions of the substituted Phe. It was also inferred that Tyr56 should have a strong positive  $L_b$  band at longer wavelengths, if tyrosyl contributions are additive (Mark et al. 1995; Thompson et al. 1998). In this paper we present new CD data for a Y56H mutant, and show that the  $L_a$  band of the g5p consists of approximately additive contributions of the five Tyr, while, at the same time, tyrosyl contributions to the  $L_b$  band are not additive. This insight has implications for the interpretation of the decreased  $L_a$  band intensity associated with the binding of g5p to ssDNA, as will be discussed.

The structure of the g5p-ssDNA complex, which forms a left-handed superhelix with two antiparallel binding channels for ssDNA, has been studied by electron microscopy, NMR, CD spectroscopy, solution scattering techniques, and molecular modeling (Gray 1989; Folmer et al. 1994; Skinner et al. 1994; Olah et al. 1995). Of particular importance to our current study, two models of the g5p-ssDNA complex have been proposed for constructing the polymeric protein shell of the g5p complex on the basis of the X-ray structures, in the absence of ssDNA, of the wild-type protein and two g5p mutants, Y41H and Y41F (Guan et al. 1994). A superhelical complex based on the crystal structure of the Y41H mutant was more satisfactory in its dimer-dimer contacts than that based on the wild-type or Y41F proteins. The model, based on the crystal structure of the Y41H protein, was similar to that independently derived from the solution structure of the Y41H protein (Folmer et al. 1994). Because NMR results showed that the solution structure of the Y41H mutant was very nearly identical to that of the wild-type or Y41F mutant proteins (Prompers et al. 1995), differences in the crystal structures of Y41H and the other two proteins were apparently the result of crystal packing forces. It has been puzzling why the crystal packing interactions of the mutant Y41H protein, which is defective in cooperative interactions in solution, appeared to mimic the dimer-dimer arrangement of wild-type g5p complexed with ssDNA in solution.

By simultaneously monitoring the CD changes of the nucleic acid CD at 270 nm and the protein tyrosyl CD band at 229 nm, we show that Y41H and Y41F mutants have significantly different binding properties when added to poly[d(A)] at low ionic strength, although they have similar CD spectral characteristics when free in solution. In the presence of 50 mM  $\text{Na}^+$ , the CD binding properties observed for Y41F and Y41H mutants were both like those of



**Fig. 2.** CD spectra of wild-type and three mutant gene 5 proteins. (A) Far-UV spectra at 180–250 nm. Wild-type (—), Y56H (open diamonds), Y41H (open triangles), and Y41F (open circles) gene 5 proteins. (B) Near-UV CD spectra of the wild-type and mutant Y56H, Y41H, and Y41F gene 5 proteins at 250–300 nm. Error bars giving the range of values from two or three independent measurements are shown at peak wavelengths; error bars were generally about the size of the symbols. CD data are plotted in units of  $M^{-1} \cdot cm^{-1}$ , per mole of protein monomer, in Figures 2 and 3.

the wild-type g5p. Our results suggest that these effects reflect different defects in the cooperative binding of the two Tyr41 mutants. Moreover, our results show that the Y41H protein can undergo coordinated changes in its DNA binding and dimer–dimer cooperative interactions with a change of salt concentration in solution. We conclude that it is not unexpected that dimer–dimer packing interactions of the wild-type and Y41F proteins in crystals without ssDNA do not represent those that occur in solution when the g5p is bound to ssDNA.

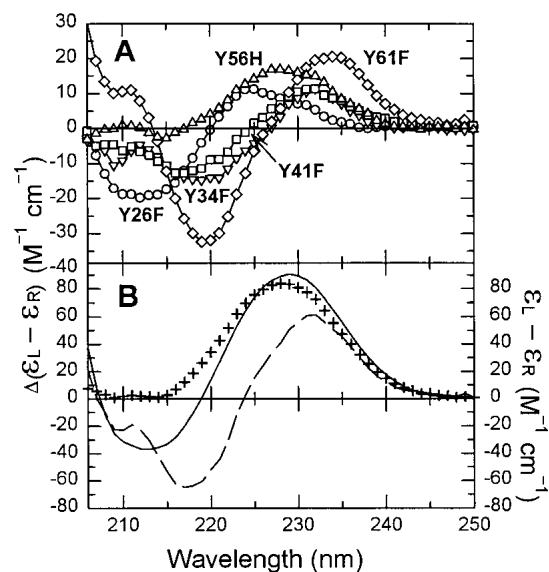
## Results

### CD spectra

The far-UV CD spectra (180–250 nm) of the Y56H, Y41H, and Y41F mutant proteins were qualitatively similar to that of wild-type g5p, as shown in Figure 2A. In this wavelength region, the g5p mutants retained the large tyrosyl  $L_a$  CD band of the wild-type g5p at 229 nm, but the magnitude at 229 nm was about 10% smaller for Y41H and Y41F and 19% smaller for Y56H. Because the 229-nm CD band is completely abolished by denaturation (Liang and Terwilliger 1991), and because there was an overall similarity in

the far-UV CD spectra of the mutant and wild-type gene 5 proteins, these data provided evidence that the mutant proteins maintained a native structure in solution.

The individual CD contributions from all five Tyr of g5p, derived from spectra in Figure 2 and previously published data (Mark et al. 1995; Thompson et al. 1998), are reflected in the difference spectra in Figure 3A. These spectra were obtained by subtracting the spectra of the Tyr→Phe/His mutants from the spectrum of wild-type g5p. The sum of these difference spectra is compared with the measured spectrum of the wild-type g5p in Figure 3B. The maximum in the summed difference spectra is considerably smaller than the maximum in the wild-type CD spectrum and is shifted to the red. These differences do not necessarily reflect deviations from additivity of the Tyr  $L_a$  CD bands, because difference spectra are being compared with an absolute spectrum. If Phe contributions to four of the difference spectra are estimated and subtracted as described in Materials and Methods, the sum of the contributions of the individual Tyr side chains is close to the experimental  $\Delta\epsilon(229)$  of the wild-type protein, as shown in Figure 3B and Table 1. All five Tyr residues make positive contributions to the  $L_a$  band, ranging in magnitude from  $8.5 M^{-1} cm^{-1}$  for Tyr26 to  $26.9 M^{-1} cm^{-1}$  for Tyr61. This analysis of the



**Fig. 3.** CD contribution of tyrosines of Ff gene 5 protein in the far-UV region. (A) Difference CD spectra of [CD(WT) – CD(Y26F)] (open circles), [CD(WT) – CD(Y34F)] (open down triangles), [CD(WT) – CD(Y41F)] (open squares), [CD(WT) – CD(Y56H)] (open up triangles), [CD(WT) – CD(Y61F)] (open diamonds). (B) Comparison of the measured CD spectrum of wild-type g5p (—), the sum of [CD(WT) – CD(mutant)] difference CD spectra (---) that are shown in (A), and the sum of the calculated contributions of individual Tyr (+++). The wild-type g5p CD spectrum and sum of Tyr contributions are plotted as  $\epsilon_L - \epsilon_R$  (right ticks), and the difference CD spectra are plotted as  $\Delta(\epsilon_L - \epsilon_R)$  (left ticks).

**Table 1.** Individual Tyr side-chain contributions to  $\Delta\epsilon(229)^a$  of the wild-type g5p

Tyrosine	$\Delta\epsilon(229)$ ( $M^{-1} cm^{-1}$ )
Y26	+8.5
Y34	+15.6
Y41	+15.6
Y56	+16.5
Y61	+26.9
Total	+83.2

<sup>a</sup> Calculated from difference CD spectra [ $CD(\text{wild-type}) - CD(\text{Tyr} \rightarrow \text{Phe/His mutant})$ ] by procedure described in Materials and Methods.

difference spectra therefore yields a value of  $+83 M^{-1} cm^{-1}$  for the sum of the individual Tyr contributions to  $\Delta\epsilon(229)$ , which is within 8% of the directly observed value of  $+90 M^{-1} cm^{-1}$  for the wild-type protein. Thus, the far-UV Tyr contributions are, to a good approximation, additive. There is no one Tyr that makes a dominant contribution.

Figure 2B shows CD spectra of the wild-type and mutant proteins in the near-UV region (300–250 nm). At wavelengths above 260 nm, where Phe and His make little or no CD contributions, the CD spectrum is dominated by tyrosyl  $L_b$  bands (Strickland 1974). CD spectra of the Y41H and Y41F mutants were almost identical, but were different from that of the wild-type g5p in having somewhat more negative magnitudes. This showed that the two mutant proteins were essentially identical, and their spectra in this region reflected only the removal of Tyr41. The CD spectrum of the Y56H mutant was only slightly changed from that of the wild-type protein, showing that Tyr56 does not

make a large contribution to the near-UV spectrum of the wild-type protein, contrary to expectations from the long-wavelength spectra of the other mutants (Thompson et al. 1998).

### Empirical CD analyses

As assigned by the DSSP method (Kabsch and Sander 1983), the secondary structures in the crystallographic structures of wild-type, Y41F, and Y41H proteins (Protein Data Bank [PDB] codes 1GVP, 1YHB, and 1YHA, respectively) are listed in Table 2A. The mutant proteins have structural features similar to those of wild-type g5p, with a predominant  $\beta$ -structure content (44–48%). Moreover, all three proteins have 7% of distorted  $\alpha$ -helix (two short segments of three residues in  $3_{10}$  helices), 15–17% of turn, and 30–33% of unordered structures. The experimental protein CD spectra were analyzed by three CD analysis programs (CONTIN [Provencher and Glöckner 1981], SELCON3 [Sreerama et al. 1999], and CDSSTR [Johnson 1999]) for secondary structure assignments, and the averaged results of these three analyses are shown in Table 2B. All three mutant proteins (Y41F, Y41H, and Y56H) had secondary structure fractions similar to those of the wild-type g5p with less than 1%  $\alpha$ -helix and significant percentages (38–41%) of  $\beta$ -structure. These results for wild-type and mutant (Y41F and Y41H) gene 5 proteins were comparable to the results derived from crystallographic data (Table 2A). For example, the fractions of regular  $\beta$ -sheet were 24–26% and 26–30% for data from CD and crystallographic structures, respectively. In general, the protein secondary structure assignments from the CONTIN and the CDSSTR analyses were more consistent with the crystal structures than were

**Table 2.** (A) Percentages of secondary structures assigned by the DSSP method from the crystallographic structures for wild-type and mutant gene 5 proteins

Protein	PDB code	Regular $\alpha$ -helix (%)	Distorted $\alpha$ -helix (%)	Regular $\beta$ -sheet (%)	Distorted $\beta$ -sheet (%)	Turn (%)	Unordered (%)
Wild-type	1GVP	0	7	28	18	17	30
Y41F	1YHB	0	7	30	18	15	30
Y41H	1YHA	0	7	26	18	15	33

**(B) Percentages of secondary structures for wild-type and mutant gene 5 proteins from the average of empirical analyses of CD spectra using CONTIN, SELCON3, and CDSSTR programs<sup>a</sup>**

Protein	Regular $\alpha$ -helix (%)	Distorted $\alpha$ -helix (%)	Regular $\beta$ -sheet (%)	Distorted $\beta$ -sheet (%)	Turn (%)	Unordered (%)
Wild-type	$-1 \pm 1^b$	$2 \pm 1$	$26 \pm 5$	$14 \pm 1$	$24 \pm 1$	$35 \pm 7$
Y41F	$-1 \pm 2$	$0 \pm 2$	$24 \pm 4$	$14 \pm 1$	$25 \pm 2$	$38 \pm 6$
Y41H	$-1 \pm 1$	$1 \pm 2$	$26 \pm 4$	$15 \pm 1$	$25 \pm 2$	$34 \pm 6$
Y56H	$-1 \pm 1$	$1 \pm 2$	$25 \pm 6$	$14 \pm 2$	$24 \pm 1$	$35 \pm 10$

<sup>a</sup> The sum of percentages determined by CDSSTR and SELCON3 analyses are restricted to the range of 95 to 105%.

<sup>b</sup>  $\pm$  values are the ranges of values from the three analyses.

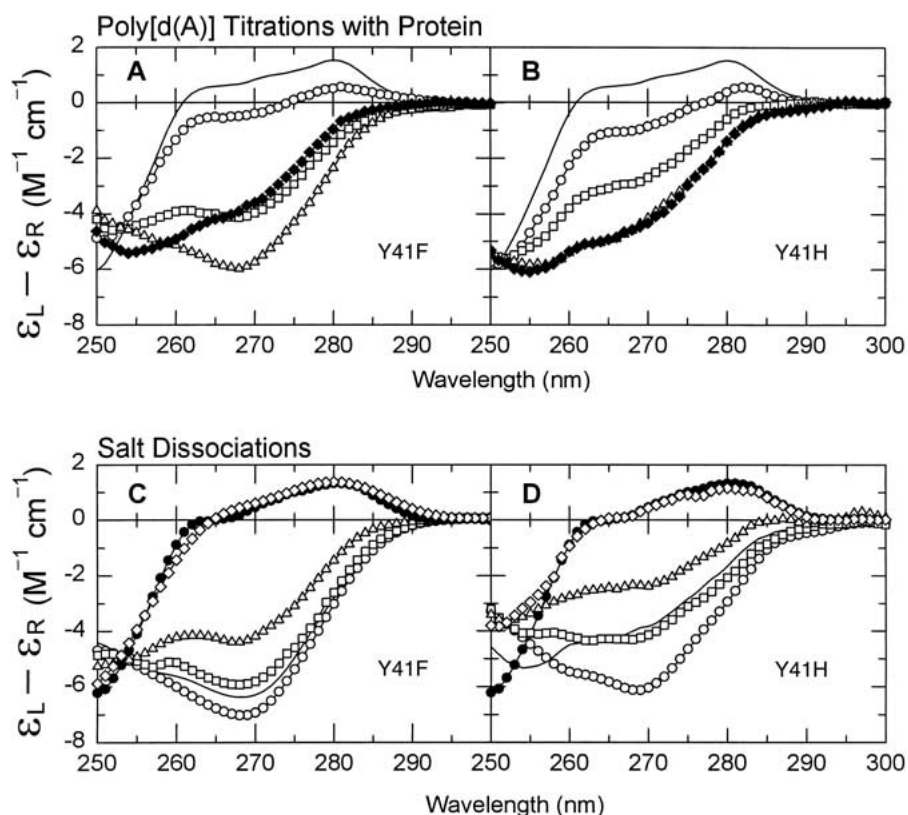
those from the SELCON3 analysis. Specifically, the SELCON3 analysis gave less  $\beta$ -structure (by 7–10%) and more unordered structure (by 7–17%) than given by the CONTIN and CDSSTR analyses (data not shown).

The crystal structure of the Y41F mutant is very similar to that of wild-type g5p. The structure of Y41H g5p is also similar to that of wild-type g5p but with small differences in the DNA binding hairpin region and in the monomer–monomer contact hairpin (Guan et al. 1994). The CD analyses gave secondary structures for the wild-type and all three mutant gene 5 proteins that were comparable with each other, and the CD results for the wild-type and mutant Y41F and Y41H proteins were in reasonable agreement with the crystallographic results. Therefore, we conclude that all the proteins, including Y56H, have similar native structures in solution.

#### Protein-poly[d(A)] complexes

##### CD titrations and nucleic acid CD changes at 270 nm.

Figure 4A and B shows the CD spectral changes in the near UV of poly[d(A)] upon additions of Y41F or Y41H mutant proteins in 2 mM Na<sup>+</sup> (phosphate buffer, pH 7.0). CD spectral changes during titrations of poly[d(A)] with the Y41F protein were qualitatively the same as those published for titrations of poly[d(A)] with wild-type, Y34F, and Y61F proteins (Mark et al. 1995; Thompson et al. 1998), as well as with the Y56H protein (not shown). Upon additions of g5p at [P]/[N] ratios <0.25, shown as open symbols, the positive CD bands of poly[d(A)] decreased in magnitude and became negative. Because the CD spectrum of g5p is negligible in comparison with that of the ssDNA in this



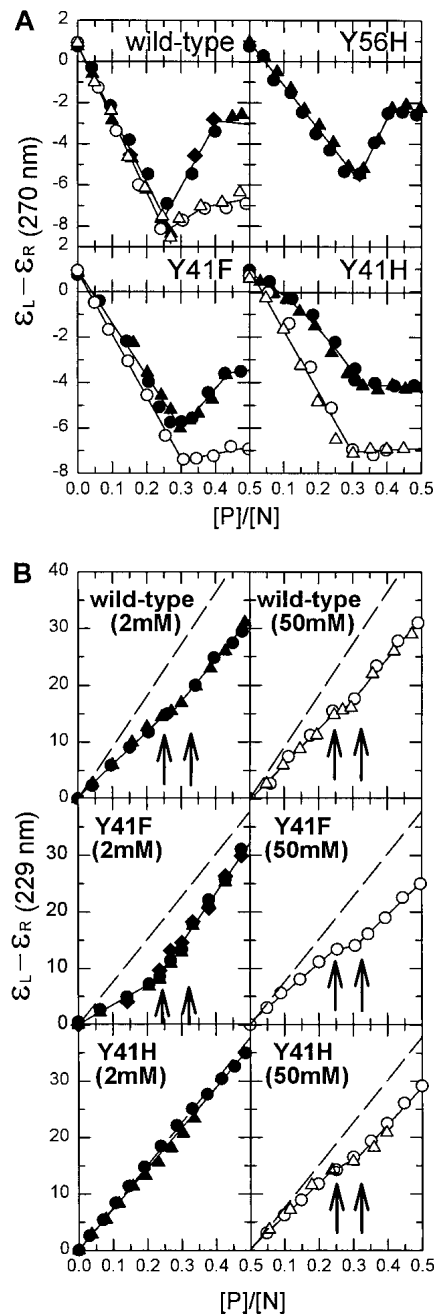
**Fig. 4.** Representative CD spectra showing changes during formation and NaCl dissociation of g5p-poly[d(A)] complexes. (A) Representative spectra from a CD titration of poly[d(A)] with Y41F g5p; [g5p]/[nucleotide] ([P]/[N]) ratios were 0.00 (—), 0.06 (open circles), 0.20 (open squares), 0.27 (open triangles), and 0.43 (filled diamonds). (B) Representative spectra from a CD titration of poly[d(A)] with Y41H g5p; [g5p]/[nucleotide] ([P]/[N]) ratios were 0.00 (—), 0.15 (open circles), 0.24 (open squares), 0.33 (open triangles), and 0.45 (filled diamonds). (C) Representative CD spectra from a salt dissociation experiment of a saturated Y41F-poly[d(A)] complex that was formed at a [P]/[N] ratio of 0.33; no added NaCl (—), and [NaCl] of 0.05 M (open circles), 0.08 M (open squares), 0.10 M (open triangles), and 0.15 M NaCl (open diamonds). (D) Representative CD spectra from a salt dissociation experiment of a saturated Y41H-poly[d(A)] complex that was formed at a [P]/[N] ratio of 0.33; no added NaCl (—), and [NaCl] of 0.03 M (open circles), 0.05 M (open squares), 0.07 M (open triangles), and 0.14 M NaCl (open diamonds). The CD spectrum of free poly[d(A)] at 1.0 M NaCl (filled circles) is shown for reference in both panels (C) and (D). The concentrations of g5p were 6.0 and 6.1  $\mu$ M in complexes with Y41F and Y41H, respectively. CD data are plotted in units of  $M^{-1}\cdot cm^{-1}$ , per mole of nucleotide, in Figures 4 and 5.

wavelength region (Thompson et al., 1998), this type of CD change is attributable to changes in the nucleic acid and includes the effects of dehydration, partial base–base unstacking, and the conformational alterations needed to form a left-handed superhelical structure when complexed with g5p (Gray 1996).

The four panels of Figure 5A show the reduction in CD magnitude at 270 nm as a function of the [P]/[N] molar ratio for titrations with the four proteins. Solid symbols show data for titrations in 2 mM Na<sup>+</sup> (phosphate, pH 7.0). Binding of the wild-type, Y41F, and Y56H gene 5 proteins to poly[d(A)] was essentially stoichiometric, and the first titration endpoints were reached at [P]/[N] ratios of 0.25 to 0.30, which indicated binding in  $n = 4$  or 3 binding modes, as summarized in Table 3. Also, at higher [P]/[N] ratios with wild-type, Y41F, or Y56H proteins, the CD at 270 nm became less negative until a second titration endpoint was reached at [P]/[N] ratios of 0.39 to 0.44, indicative of a weaker  $n \approx 2.5$  binding mode (Fig. 5A). Full spectra are shown in this [P]/[N] range for the Y41F protein in Figure 4A (solid symbols). This weak binding mode has also been observed when titrating poly[d(A)] with the three other Tyr mutants, Y26F, Y34F, and Y61F (Mark et al. 1995; Thompson et al. 1998). In such a binding mode where there are only a few nucleotides per protein monomer, structural considerations led to the inference that the ssDNA is probably partially released from the protein binding sites, which are held in juxtaposition by cooperative interactions, accounting for the reduced CD effect at 270 nm (Mark et al. 1995; Thompson et al. 1998).

The results of CD titrations with the Y41H mutant were quite different from those with the wild-type and the other Tyr mutant proteins. Like the other mutants, Y41H saturated poly[d(A)] at a [P]/[N] ratio of about 0.3 in the 2 mM Na<sup>+</sup> (phosphate, pH 7.0) buffer. However, unlike the titration data for the other proteins, the data for Y41H were slightly curved before the breakpoint, indicative of a reduced binding affinity and nonstoichiometric binding at [P]/[N] ratios < 0.2 (Fig. 5A; solid symbols for Y41H). Moreover, there was only one titration breakpoint, so that the CD did not change at [P]/[N] molar ratios above this breakpoint (Figs. 4B,5A), which was consistent with a loss of binding cooperativity. Finally, the total CD change at 270 nm of poly[d(A)] was only about  $-4 \text{ M}^{-1} \cdot \text{cm}^{-1}$  when titrated with the Y41H mutant. Separate experiments were performed to test whether the data for the Y41H-poly[d(A)] complex were time-dependent, and no changes in measured CD values were found between those taken within minutes after mixing and those taken after 24 h (not shown).

When titrations were performed in 50 mM Na<sup>+</sup> (phosphate buffer, pH 7.0), the binding of the Y41H protein was stoichiometric, and titrations with the Y41H, Y41F, and the wild-type proteins were similar (Fig. 5A, open symbols). There were single breakpoints in the titrations in 50 mM



**Fig. 5.** Nucleic acid and protein CD titration curves during titrations of poly[d(A)] with g5p. (A) Nucleic acid CD changes at 270 nm of poly[d(A)] as a function of [P]/[N] ratios during titrations with wild-type, Y41F, Y41H, and Y56H gene 5 proteins. Filled symbols: titration data at 2 mM Na<sup>+</sup>. Open symbols: titration data at 50 mM Na<sup>+</sup> (phosphate, pH 7.0). Different symbol shapes represent data from individual titrations. (B) CD changes of the protein tyrosyl CD band at 229 nm during titrations of poly[d(A)] with wild-type and mutant Y41F and Y41H gene 5 proteins in 2 or 50 mM Na<sup>+</sup> (phosphate, pH 7.0). The nucleic acid contribution to the 229 nm CD band was subtracted as described in Materials and Methods. Dashed lines show the expected CD changes for additions of g5p in the absence of poly[d(A)]. Different symbol shapes represent data from individual titrations. Arrows at [P]/[N] ratios of 0.25 and 0.33 indicate where breakpoints indicated a shift between the  $n = 4$  and  $n = 3$  binding modes, especially at the higher salt concentration.

**Table 3.** Titration breakpoints in 2 mM Na<sup>+</sup> and salt dependence of binding constants for poly[d(A)] complexes formed with wild-type and mutant gene 5 proteins

	[P]/[N] breakpoints (270 nm)		Salt dependence of binding constant	
	<i>n</i> = 4 or 3 mode	<i>n</i> ≈ 2.5 mode	Slope (-∂log(K <sub>ω</sub> )/∂log[NaCl])	K <sub>ω</sub> at [NaCl] = 0.2 M (×10 <sup>-3</sup> M <sup>-1</sup> )
Wild-type	0.25 ± 0.01	0.39 ± 0.02	4.1 ± 0.6	175 ± 37
Y56H	0.30 ± 0.01	0.40 ± 0.01	3.6 ± 0.9	52.0 ± 12
Y41F	0.30 ± 0.01	0.44 ± 0.02	4.9 ± 0.6	5.0 ± 0.7
Y41H	0.31 ± 0.02	N/A	3.7 ± 0.8	1.9 ± 0.3

Na<sup>+</sup>, because the weaker binding mode of g5p is dissociated at the higher salt concentration (Mark et al. 1995).

#### Tyrosyl CD changes at 229 nm.

The panels of Figure 5B show the tyrosyl CD changes at 229 nm as a function of the [P]/[N] ratio during CD titrations of poly[d(A)] with the wild-type g5p and the Tyr41-substituted mutants. Dashed lines represent the linear increase in CD with increasing the [P]/[N] ratio that would be observed from adding the free proteins to buffer. Symbols show the experimental data obtained during titrations, after subtracting the contribution of the nucleic acid as described in Materials and Methods. In 2 mM Na<sup>+</sup> (phosphate, pH 7.0), the tyrosyl CD band of wild-type and Y41F proteins respectively decreased 37 and 46%, relative to the respective free proteins, as the proteins were added to poly[d(A)] up to the breakpoint. (The apparent decrease would be even greater without a correction for the decreased CD of the nucleic acid when bound by the protein.) Breakpoints in CD titration plots at 229 nm are more difficult to discern than in plots at 270 nm, but, for the wild-type and Y41F proteins, there was at least one breakpoint in a [P]/[N] range consistent with the first (*n* = 4 or 3) binding mode detected at 270 nm. A weak mode of binding at *n* ≈ 2.5 was not detected in the 229-nm CD titrations. Similar CD perturbations and breakpoints have been observed before in the 229-nm CD titrations of genomic ssDNA at 2 mM Na<sup>+</sup> with Y26F and Y61F, but not with Y34F, which led to the conclusion that Tyr34 is responsible for the 229-nm CD perturbation during cooperative binding (Mark et al. 1995; Thompson et al. 1998). Consistent with this conclusion, we found in the present work that the 229 nm band of the Y56H mutant also decreased by 31–32% during the titration of poly[d(A)] (not shown), close to the decrease found for the wild-type protein. That is, the CD change at 229 nm for the wild-type and Y41F proteins is most simply explained as being the result of cooperative interactions and a perturbation of Tyr34 at the dimer–dimer interface under these low salt conditions.

In the case of titrations of poly[d(A)] with Y41H at 2 mM Na<sup>+</sup>, there was no change in CD magnitude at 229 nm throughout the full range of protein additions (Fig. 5B),

although the nucleic acid perturbation at 270 nm (Fig. 5A) clearly showed that Y41H bound to, and perturbed, the ssDNA. This result can be interpreted as indicating that the Y41H mutation caused a more drastic reduction in the binding cooperativity than did the Y41F mutation so that, indirectly, the Tyr34 was not perturbed to the same extent in titrations of poly[d(A)] with the two mutant proteins.

On the other hand, when titrations with Y41H were performed in 50 mM Na<sup>+</sup> (phosphate, pH 7.0), the CD band at 229 nm did exhibit a decrease and showed two breakpoints at [P]/[N] ratios at which the poly[d(A)] became saturated in the *n* = 4 and *n* = 3 binding modes (Fig. 5B, open symbols). Below [P]/[N] ≈ 0.25, the 229-nm CD magnitude was reduced by about 23% from that expected for addition of free protein. Then, as the protein was added to give [P]/[N] values between 0.25 and 0.30, the tyrosyl band was more fully perturbed to be decreased a total of 29%. Above a [P]/[N] ratio of 0.3, the 229-nm CD band paralleled the slope expected for the addition of free protein. This pattern of change in the 229-nm CD band that was exhibited by the Y41H protein was very close to the patterns observed for the wild-type and Y41F proteins at this higher salt concentration. The two breakpoints could arise from different perturbations of Tyr34 when binding is in the *n* = 4 and *n* = 3 modes, or possibly from the added perturbation of another Tyr such as Tyr26 when binding is in the *n* = 3 mode. In any case, the CD titrations left no doubt that differences in the binding characteristics of the two Tyr41 mutants that were apparent in 2 mM Na<sup>+</sup> almost disappeared in 50 mM Na<sup>+</sup>.

#### Salt-dissociation of the complexes.

Figure 4C and D shows representative CD spectra during the dissociation of saturated Y41F-poly[d(A)] or Y41H-poly[d(A)] complexes by increasing the salt concentration. Spectra for complexes with wild-type or Y56H g5p at various salt concentrations (data not shown) were similar to those shown for the complex with Y41F. The complexes were formed in 2 mM Na<sup>+</sup> (phosphate buffer, pH 7.0) at [P]/[N] ratios of 0.25 or 0.33. As the NaCl concentration was increased to 0.03 M, the CD bands above 250 nm of the

Y41H-poly[d(A)] complex became more negative, indicating a greater perturbation of the poly[d(A)]. As the [NaCl] increased to >0.1 M, the complexes dissociated, as shown by the fact that the CD spectra above 260 nm became less negative and eventually were identical to the spectrum of free poly[d(A)] at a high NaCl concentration of 1.0 M.

Dissociation curves (not shown) were obtained by monitoring the CD change, relative to the maximum difference in CD value at 270 nm, as a function of NaCl concentration (see Materials and Methods).

#### *Dependence of binding constant on NaCl concentration.*

The binding affinity ( $K_{\omega}$ ) was determined from the free g5p dimer concentration required for 50% dissociation of each complex. The salt concentration needed for 50% dissociation increased as the concentration of the g5p-poly[d(A)] complex increased. Figure 6 shows that  $\log[K_{\omega}]$  was linearly related to the  $\log[\text{NaCl}]$  for each type of complex. The slopes of the  $\log[K_{\omega}]$  versus  $\log[\text{NaCl}]$  plots indicate the number of ions released per g5p dimer when forming a complex (Record et al. 1976; Terwilliger 1996) and are listed in Table 3. Binding of wild-type or Tyr→Phe/His mutant gene 5 proteins to poly[d(A)] generally resulted in the release of about four ions per g5p dimer, consistent with results of prior CD and fluorescence work (Bulsink et al. 1985; Stassen et al. 1992a; Mark et al. 1995; Thompson et al. 1998). Exceptions are the Y41H and Y34F mutants, for which the measured slopes are  $-4.9 \pm 0.6$  (this work) and  $-6.6 \pm 1.7$  (Mark et al. 1995), respectively, indicating the possible additional release of one to two ions per protein dimer-dimer interface for these mutants.

Experimental  $K_{\omega}$  values for the various proteins were extrapolated to 0.2 M NaCl for comparison; these are listed in the last column in Table 3. The mutant proteins all showed lower binding affinities compared with the wild-type protein. The Tyr41 mutants, which are defective in protein-protein interactions and have a lowered cooperativity factor, had binding affinities for poly[d(A)] much lower than those of the Y56H mutant or of the Y26F or Y61F mutants previously studied (Mark et al. 1995; Thompson et al. 1998). The binding affinities of Y41H and Y41F were, respectively, about 90 and 35 times lower than that of wild-type g5p. Binding affinities of the Y41F and Y34F mutants are almost identical at 0.2 M NaCl (Mark et al. 1995). These data further showed that Y41H has a more defective cooperative interface than do the other two mutants, Y41F and Y34F, that have a substitution at the protein-protein interface.

## Discussion

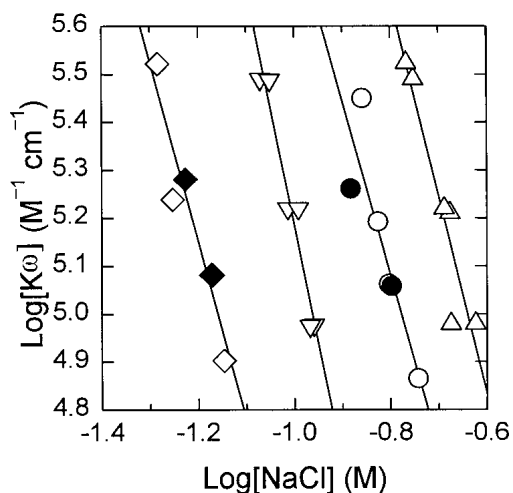
### *CD spectra of mutant gene 5 proteins*

The five Tyr in the gene 5 protein are distinguishable in their CD contributions in the far-UV and near-UV regions

(Mark et al. 1995; Thompson et al. 1998; this study). The unusually large positive 229-nm band in the CD spectrum of the g5p presumably is dominated by  $L_a$  bands of the five Tyr of the protein, with small contributions from the three Phe (Phe6, Phe13, and Phe73). Amide contributions are generally negative in this region (Sreerama and Woody 2000b). The far-UV CD spectra (180–250 nm) in Figure 2A showed that the three mutant gene 5 proteins in this study (Y41F, Y41H, and Y56H), like previously studied mutant proteins (Y26F, Y34F, and Y61F; Mark et al. 1995; Thompson et al. 1998), retained the unique positive band at 229 nm that is characteristic of the native OB-fold structure of the wild-type g5p.

For the first time, it is possible to consider whether the atypical CD spectrum of g5p, which has no Trp or disulfides, has Tyr contributions that are approximately additive. We found that the sum of the  $L_a$  bands of the individual Tyr side chains agrees well with the observed intensity of the 229-nm band for the wild-type protein (Fig. 3B, Table 1). If exciton interactions among the Tyr  $L_a$  excited states played a significant role in determining the CD of the 229 nm band, this band should exhibit deviations from additivity. For example, if two Tyr  $L_a$  transitions were coupled to give an exciton CD couplet, mutation of either side chain would eliminate the couplet. This contribution would then be counted twice when the spectra derived from the mutants are combined. The nearly additive Tyr  $L_a$  contributions demonstrate that exciton interactions among the Tyr  $L_a$  excited states play at most a minor role in determining the 229-nm CD. The amplitude of an exciton couplet is proportional to  $\epsilon_{\text{max}}^2$  and inversely proportional to the square of the interchromophoric distance (Woody and Dunker 1996). Whereas the fully allowed B bands of Tyr, Phe, and Trp have  $\epsilon_{\text{max}}^2 \geq 35,000 \text{ M}^{-1} \text{ cm}^{-1}$ , the  $L_a$  bands have  $\epsilon_{\text{max}}^2 \sim 5000\text{--}10,000 \text{ M}^{-1} \text{ cm}^{-1}$ . Exciton couplets can be expected for B bands at separations of 5–10 Å, but such a couplet in the Tyr  $L_a$  band is only likely to be observable if the rings are nearly in van der Waals contact. The closest pair of Tyr side chains in g5p is Tyr56 and Tyr61, separated by  $\sim 5.5 \text{ \AA}$ . Therefore, it is not surprising that exciton contributions to the Tyr  $L_a$  band of g5p are not significant. The CD of the  $L_a$  band for each Tyr derives primarily from coupling of the  $L_a$  transition with the  $\pi \rightarrow \pi^*$  transitions of nearby peptide groups and with the fully allowed, short-wavelength  $B_a$  and  $B_b$  transitions of nearby Tyr and Phe side chains (unpublished results).

The absence of a change in  $\Delta\epsilon(229)$  when Y34F binds to poly[d(A)] at  $[P]/[N] \leq 0.25$ , implies that Tyr34 plays a major role in this ligand-induced effect. The contribution of Tyr34 to the 229-nm band ( $\approx +16 \text{ M}^{-1} \text{ cm}^{-1}$ ) can be compared with the change in  $\Delta\epsilon(229)$  upon g5p binding to poly[d(A)], which is  $-37\%$  of  $+90 \text{ M}^{-1} \text{ cm}^{-1}$  or  $-33 \text{ M}^{-1} \text{ cm}^{-1}$  for the wild-type protein. The effect of binding to ssDNA is about twice as large as the Tyr34 contribution to



**Fig. 6.** Binding affinity of gene 5 proteins to poly[d(A)]. Plots showing the dependence of  $\log[K\omega]$  on  $\log[\text{NaCl}]$  for binding of wild-type (open triangles), Y56H (open circles), Y41F (open down triangles), and Y41H (open diamonds) gene 5 proteins to poly[d(A)]. Data were obtained from salt dissociation curves of g5p-ssDNA complexes formed at  $[\text{P}]/[\text{N}]$  ratios of 0.33 (open symbols) or 0.25 (closed symbols). The total g5p concentrations in these complexes ranged from 6 to 28  $\mu\text{M}$ . The binding affinity ( $K\omega$ ) was determined from the free protein concentration at 50% saturation of the poly[d(A)], and the salt concentration was that needed for 50% dissociation of complexes as monitored by the reversal of the maximum CD change at 270 nm (see Materials and Methods).

the wild-type CD spectrum. Therefore, the effect is not simply an abolition of the Tyr34 contribution. If the effect is solely due to Tyr34, the results imply that the Tyr34 contribution is reversed in sign upon binding to ssDNA, presumably due to conformational changes induced in the g5p. Alternative explanations include: (1) coupling of Tyr34 with chromophores in neighboring dimers upon binding to ssDNA, or (2) transmission by Tyr34 to other Tyr residues of conformational changes induced by ssDNA binding. In our previous paper (Thompson et al. 1998) it was suggested that coupling of Tyr34 with the Tyr56/Tyr61 pair may account for the absence of perturbation of the 229 nm band when Y34F binds ssDNA. Our present results show that Tyr56 does not dominate the 229-nm band, so although this coupling still could provide part of the explanation, it is no longer considered to be as probable.

Above 270 nm, where tyrosyl  $L_b$  bands dominate, Tyr34 and Tyr61 make significant negative contributions and the other three Tyr have nearly zero or small positive signals (Mark et al. 1995; Thompson et al. 1998). Comparison with the near-UV CD spectrum of the wild-type protein, which has positive bands at 276 and 283 nm (Fig. 2B), led to the prediction that Tyr56 should make a significant positive CD contribution near 280 nm to compensate for the negative CD contributions of Tyr34 and Tyr61. As seen in Figure 2B, this is not the case, and, therefore, the tyrosyl CD contributions are not additive in the near UV. The deviations from

additivity of the Tyr  $L_b$  CD bands may result from relatively minor conformational changes induced in the remaining Tyr side chains by replacement of a given Tyr by Phe (or His). Because the  $L_b$  transition is polarized perpendicular to the twofold axis of the phenolic ring, variations in  $\chi_2$  (the torsional angle about the  $C_\beta-C_\gamma$  bond) will affect the  $L_b$  rotational strength. By contrast, the  $L_a$  transition is polarized along the twofold axis and is insensitive to variations in  $\chi_2$ .

The very similar near-UV CD spectra of Y41H and Y41F in Figure 2B indicate that neither replacement of Tyr41 detectably changes the average CD signals of the remaining aromatic residues. These CD data are consistent with NMR work that shows Y41F and Y41H to have nearly identical structures to each other and to the wild-type g5p in solution (Prompers et al. 1995). In their crystal structures, Y41H differs from the Y41F and wild-type proteins in the DNA-binding loop and dimer-contact regions (Guan et al. 1994; Skinner et al. 1994). Calculations suggest that these aspects of the Y41H crystallographic structure are energetically unstable in solution, and, therefore, differences between the solution and crystal structures may well be due to crystal packing forces (Prompers et al. 1995).

#### *Empirical CD analyses of the $\beta$ -structure in gene 5 protein*

Because the CD spectra of g5p have a unique large positive 229-nm CD band resulting from tyrosyl contributions, it might have been difficult to perform a successful analysis of the g5p secondary structures (Woody 1978). However, we found that CD analyses of wild-type and mutant (Y41F and Y41H) gene 5 proteins were consistent with high-resolution crystallographic and NMR studies in showing that the protein has a predominant  $\beta$ -structure. Of the three programs used, CONTIN (Provencher and Glöckner 1981), SELCON3 (Sreerama et al. 1999), and CDSSTR (Johnson 1999), CDSSTR gave the best overall analysis in terms of the agreement between the secondary structure fractions estimated from CD and those from the X-ray structures. However, the results for a given gene 5 protein from all three programs were similar, all being within about 5% for a given secondary structure, except for the regular  $\beta$ -sheet, for which the results were all within 10% (not shown). The successful analysis of CD spectra of wild-type and mutant gene 5 proteins, despite substantial aromatic contributions, is attributable to the improvements in protein CD analysis achieved by the use of a larger reference protein set (Sreerama and Woody 2000a).

#### *Structural implications of differential binding properties of Y41H and Y41F*

The Y41F mutant has characteristics of binding to poly[d(A)] that are similar to those of the wild-type and four

other mutant proteins (Y26F, Y34F, Y56H, and Y61F) (Mark et al. 1995; Thompson et al. 1998; this work). These proteins exhibit two modes of binding to poly[d(A)] when the nucleic acid CD at 270 nm is monitored during titrations in 2 mM Na<sup>+</sup>. The Y41H protein, however, showed only one mode of binding, as monitored by the perturbation of the nucleic acid CD. Other differences from the Y41F protein were that the 270-nm binding curve for the Y41H mutant was nonstoichiometric (at low [P]/[N] ratios) and poly[d(A)] was less perturbed when it was complexed with Y41H than with the wild-type or any of the other mutant proteins studied to date, because the CD at 270 nm only reached a negative value of  $-4 \text{ M}^{-1} \text{ cm}^{-1}$  in the low salt buffer (Fig. 5A).

Moreover, the Y41H and Y41F mutants differed in that no 229-nm CD perturbation was found during titrations of poly[d(A)] with Y41H in 2 mM Na<sup>+</sup> (Fig. 5B). The titration of poly[d(A)] with Y41F at low ionic strength resulted in a decrease of  $\approx 46\%$  in the magnitude of the 229-nm tyrosyl CD band, similar to that observed during titrations with the wild-type g5p and three other mutant proteins (Y26F, Y56F, and Y61H). As described in Results, this effect on the Tyr band at low ionic strength is probably caused by the perturbation of Tyr34 that results from protein–protein interactions during cooperative binding at  $[P]/[N] \leq 0.25$ .

These observations were undoubtedly a reflection of the decreased cooperativity in Y41H dimer–dimer interactions (Folkers et al. 1991), and were in accord with the fact that the Y41H mutant differs from the Y41F and wild-type proteins in its crystal packing (Guan et al. 1994). The Y41H and Y41F proteins were originally described as both having a reduced tendency to aggregate, presumably because of reduced dimer–dimer interactions relative to the wild-type protein (Folkers et al. 1991). We now have shown that the binding affinity of Y41H for poly[d(A)] is significantly less than that of Y41F, by a factor of about 2.6 when extrapolated to 0.2 M NaCl (Fig. 6, Table 3).

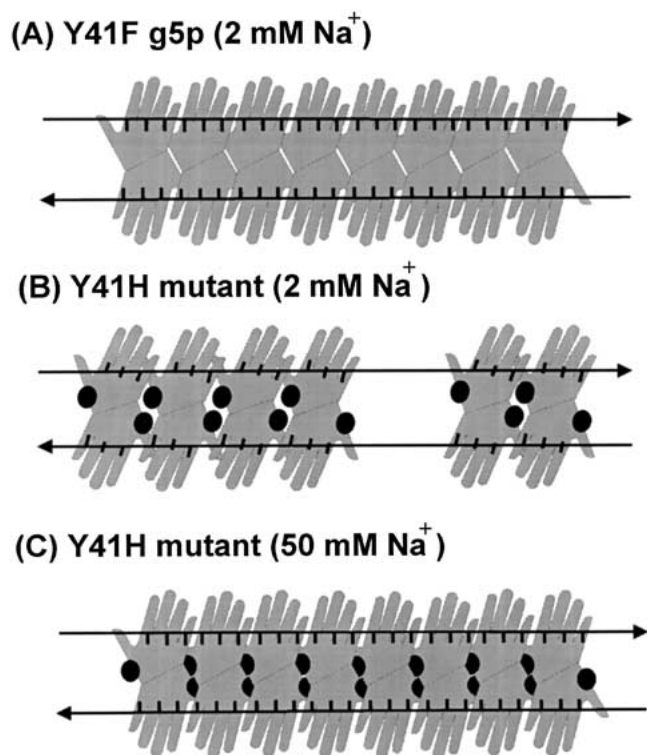
Although the Y41H protein has a reduced cooperativity of binding and reduced binding affinity in solution, Guan et al. (1994) proposed that the dimer–dimer packing of Y41H in crystals represents the interaction of neighboring dimers when the wild-type g5p is bound to ssDNA and that the contacts in the Y41F and wild-type crystals represent those of the protein when it aggregates in solution in the absence of ssDNA. These authors modeled the g5p–ssDNA superhelix on the basis of the Y41H crystal packing, and the neighboring dimers in their model were similar to those in a superhelical complex independently modeled by Folmer et al (1994). The success of this model, despite the effect of the Y41→H41 substitution on the solution binding properties of g5p, may be rationalized on the basis of an interplay between (1) the protein loops involved in dimer–dimer contacts and (2) the DNA-binding wings that directly contact the ssDNA. Protein–protein cooperative interactions could

indirectly constrain (or be constrained by) contacts with the ssDNA. Alternatively, there could be coupled structural changes in the DNA binding and cooperativity loops of the protein. Correlated motions of the DNA-binding wings and the dimer–dimer interaction loops have recently been described for the closely related Pf3 DNA-binding protein (Horstink et al. 1999) and for the dynamic ssDNA-binding mechanism described for human replication protein A (Bochkareva et al. 2001). That is, it is possible that the dimer–dimer crystal packing of the Y41F or wild-type proteins differs from that when these proteins are bound to ssDNA in solution, but the Y41H mutation permits “native” interdimer contacts in the crystal structure in the absence of ssDNA. Guan et al. (1994) observed that, in the crystal structure of the Y41H mutant, the ssDNA binding hairpins were affected and even adopted different conformations within the same g5p dimer. Structural perturbations permitted by the mutation need be only relatively small, given that a Raman study showed that the g5p secondary structure does not change significantly upon binding to poly[d(A)] (Benevides et al. 1996).

In solution, our combined CD titration data from monitoring the nucleic acid at 270 nm and the protein at 229 nm are supportive of correlated changes in the DNA-binding site and the cooperativity domain of the Y41H mutant during formation of a Y41H·poly[d(A)] complex. As shown in Figure 5A, in 50 mM Na<sup>+</sup>, the ssDNA binding of Y41H was stoichiometric throughout the titration, and the poly[d(A)] was as fully perturbed as with Y41F or the wild-type protein. In addition, Figure 5B shows that, at the higher salt concentration, the 229-nm tyrosyl CD band decreased to approximately the same extent and with the same dependence on the binding mode as for the Y41F and wild-type proteins. Figure 7 schematically illustrates the simultaneous change in ssDNA binding and protein–protein interactions with an increase in salt concentration, which conceivably helps overcome repulsion between positive charges on His41 on neighboring dimers. Because the pK<sub>a</sub> of His41 is 6.5 (Folkers et al. 1991), His41 side chains will be only partially positively charged at pH 7.0. Nevertheless, the charge on His41 appears to be important in the binding properties of Y41H to poly[d(A)], because CD measurements of Y41F·poly[d(A)] and Y41H·poly[d(A)] complexes ( $[P]/[N] = 0.25$ ; 2 mM Na<sup>+</sup>) as a function of pH showed that the CD (270 nm) of the Y41H·poly[d(A)] complex was even less negative at pH 6.1 than at pH 7.0 (Fig. 5A), while the CD of the Y41F·poly[d(A)] complex was unchanged (not shown).

#### *Biological implications*

Not only does the Ff g5p saturate newly synthesized phage ssDNA genomes, but g5p also binds to mRNA leader sequences to control the translation of gene 2 and other phage



**Fig. 7.** Illustration of the effect of salt concentration on the binding of Y41H. (A) The Y41F mutant g5p binds cooperatively to and can saturate poly[d(A)] at 2 mM Na<sup>+</sup>. Each pair of hands represents a g5p dimer with the required twofold symmetry to interact with two antiparallel strands of DNA. Fingers represent the DNA-binding loops interacting normally with DNA nucleotides. Thumbs represent dimer–dimer interaction loops, which interact cooperatively between dimers. (B) The Y41H mutant protein binds noncooperatively to poly[d(A)] at 2 mM Na<sup>+</sup> because of imperfect dimer–dimer contacts. The dark circle represents a perturbation of the cooperativity loops by the partially protonated His41. As a result of perturbed protein–protein interactions, the DNA-binding loops of the protein cannot interact normally with DNA nucleotides, either because of an indirect constraint of contacts with the ssDNA or because of correlated changes in the DNA-binding and cooperativity loops of the protein. (C) When the salt concentration is raised to 50 mM Na<sup>+</sup>, the Y41H protein can now form almost normal dimer–dimer contacts, despite the presence of His41. In this situation, the protein can again bind cooperatively to poly[d(A)] and the DNA-binding loops can interact normally with DNA nucleotides.

genes. The Y41H and Y41F mutant proteins do not inhibit the production of phagmid DNA transducing particles, and both mutant proteins were inactive in repressing the translation of a reporter gene whose translational initiation region was that of the gene 2 protein (Stassen et al. 1992b). Our data suggest that a reason why both mutants are defective in these properties is that these biological effects have a stringent requirement for an intact dimer–dimer cooperative interface. On the other hand, overexpression of the Y41F mutant or wild-type g5p, but not the Y41H mutant, inhibits *Escherichia coli* growth by indiscriminate nucleic acid binding (Terwilliger et al. 1994). The inability of Y41H to inhibit *E. coli* growth is consistent with the especially low

binding affinity of this protein at physiological ionic strength (Table 3).

## Materials and methods

### Proteins and nucleic acid

Ff gene 5 protein was isolated essentially as previously described from *E. coli* K561 cells containing plasmids encoding the wild-type and mutant g5p genes (Terwilliger et al. 1988; Mark et al. 1995; Thompson et al. 1998). Although the wild-type g5p was present mainly in the soluble fraction after cell lysis by blending, only about one-third of the Y56H protein was present in the soluble fraction. Most of the Y56H protein was present in the insoluble fraction, and was isolated by the procedure described for the Y61F mutant (Thompson et al. 1998), but with an added Sephadex G75 chromatography step to remove high molecular weight contaminants. The Y56H protein from this procedure gave a re-folded protein that had the same elution profile from an ssDNA-cellulose column and the same CD spectrum as the protein from the soluble fraction. Because Tyr56 and Tyr61 are associated with each other in the protein core, and because Y56H and Y61F mutants were both in cellular aggregates, the Y56H protein is possibly a folding mutant, like the Y61F mutant (Thompson et al. 1998). Protein contaminants of all the g5p preparations were less than 1%, as determined by 18% SDS-PAGE. Protein concentrations were determined from absorption measurements using protein monomer molar extinction coefficients  $\epsilon(276)$  of 7074 and 5660 M<sup>-1</sup>·cm<sup>-1</sup> for the wild-type and Y41F mutant proteins, respectively (Day 1973; Mark et al. 1995). Because His and Phe have essentially no absorption above 270 nm, the same  $\epsilon(270)$  was used for all three mutants. All g5p stocks were dialyzed into and stored in a buffer of 2 mM Na<sup>+</sup> (phosphate buffer, pH 7.0). Poly[d(A)] (Sigma) had an average length of 260 nucleotides according to the manufacturer. The poly[d(A)] was dissolved in 2 mM or 50 mM Na<sup>+</sup> (phosphate buffer, pH 7.0), and its concentration was determined from absorption spectra using an  $\epsilon(260)$  of 9650 M<sup>-1</sup>·cm<sup>-1</sup>, per mole of nucleotide (Bollum 1966).

### UV absorption and CD measurements

UV absorption spectra were measured with an Olis-modified Cary 14 spectrophotometer (On-Line Instrument Co.). CD spectra were measured with a Jasco Model J715 spectropolarimeter (Jasco Incorporated). Far-UV CD spectra (250–180 nm) of free g5p at concentrations of 90–150  $\mu$ M were obtained using water-jacketed cylindrical quartz cells having a light path of 0.1 mm. Near-UV CD spectra (320–250 nm) of free g5p and spectra during CD titrations and salt dissociations were taken in 10- or 20-mm cuvettes. Spectropolarimeter parameters and processing were as in earlier work (Mou et al. 1999), except that the near-UV spectra, collected at 0.1-nm resolution, were smoothed over 13 points instead of 99 points to preserve the resolution of vibronic bands. All spectra were taken at 20  $\pm$  0.5°C. The smoothed CD data were plotted at 1-nm intervals as  $\Delta\epsilon$  in units of M<sup>-1</sup>·cm<sup>-1</sup>, per mole of g5p monomer for free protein spectra or per mole of nucleotide for CD titrations and salt dissociations.

### Calculation of Tyr contributions to the 229-nm CD band

To calculate the contribution of a given Tyr side chain from the difference CD spectrum [CD(wild-type) – CD(Tyr→Phe mutant)],

different procedures are required, depending on the shape of the difference spectrum. The difference spectra for the five mutants exhibit three different shapes. Those for Y34F, Y41F, and Y61F proteins show the classic first-derivative shape expected for difference spectra resulting from a shift in wavelength with little or no change in intensity (Donovan et al. 1961). If the Tyr→Phe substitution does not lead to a conformational change, the  $L_a$  band of the Phe is expected to have a rotational strength of the same sign and comparable magnitude to that of the Tyr that it replaces. However, the Phe  $L_a$  band is shifted by about 10 nm to shorter wavelengths, that is, from ~230 to ~220 nm. Therefore, when the CD difference spectrum of the mutant is subtracted from that of the wild-type, the  $L_a$  band of the replacement Phe will contribute a band of opposite sign, centered at ~220 nm. The resultant CD difference spectrum will have the appearance of a CD couplet (Schellman 1968; Bayley 1973) with a crossover wavelength at the average of the  $L_a$  wavelengths for Phe and Tyr (~225 nm). The long-wavelength maximum will be red-shifted relative to  $\lambda_{\max}$  for the Tyr  $L_a$  band, and the short-wavelength maximum will be blue-shifted relative to  $\lambda_{\max}$  for the Phe  $L_a$  band. Moreover, the amplitude for the long-wavelength branch of the couplet will be smaller than that of the Tyr contribution.

For this type of difference spectrum, the value of  $\Delta\epsilon_{\max}$  for the 229 nm band can be estimated from the maximum in the difference spectrum,  $\Delta\Delta\epsilon_{\max}$ , assuming that the substitution of Phe for Tyr leads to a shift of  $\lambda_{\max}$  without a change in amplitude:

$$\Delta\Delta\epsilon_{\max}(\lambda_{\max}) = \Delta\epsilon_{\max} \left\{ \frac{\exp[-(\lambda_{\max} - \lambda_{\text{Tyr}})^2/\Delta^2]}{\exp[-(\lambda_{\max} - \lambda_{\text{Phe}})^2/\Delta^2]} \right\},$$

where  $\lambda_{\max}$  is the wavelength of the maximum in the difference spectrum (232 nm),  $\lambda_{\text{Tyr}} = 229$  nm,  $\lambda_{\text{Phe}} = 220$  nm, and  $\Delta$  is the Gaussian band width for the  $L_a$  bands of Tyr and Phe, assumed to be the same and taken to be 9 nm. This leads to the relationship:

$$\Delta\epsilon_{\max}(229) = 1.38 \Delta\Delta\epsilon_{\max}(232).$$

The mutant Y26F also gives a CD difference spectrum with a first-derivative shape, but this is shifted about 7 nm to the blue, with a positive maximum at 225 nm, a crossover near 220 nm, and a negative maximum at 212 nm. Because of this shift in wavelength, it is not clear that a procedure analogous to that used for the previous mutants should be used. Instead, the observed  $\Delta\Delta\epsilon(229)$  value for the Y26F mutant has been corrected for the contribution of a Gaussian band of 9 nm bandwidth, with the sign, magnitude, and wavelength of the negative maximum at 212 nm. Finally, the CD difference spectrum for Y56H has the shape of a simple Gaussian band centered at 228 nm, indicating that the substituted His makes little contribution to the CD in this region. Therefore, the value of  $\Delta\Delta\epsilon(229)$  from the difference CD spectrum for this mutant was used for the  $\Delta\epsilon(229)$  of Tyr56.

#### Empirical CD analyses for protein secondary structure

Protein CD spectra were analyzed by CDPro software (CONTIN, SELCON3, and CDSSTR, available at the website <http://lamar.colostate.edu/~sreeram/CDPro>) for determining the fractions of secondary structure in the individual proteins. The input experimental CD data ranged from 240 to 190 nm at 1-nm intervals, and were in units of  $\Delta\epsilon$  per mole of amino acid residue. The largest reference protein set, containing 48 proteins, was selected for all three CD analyses (Sreerama and Woody 2000a). Reference protein sets that include the poly(Pro)II fraction were not used for the analyses because g5p has no poly(Pro)II structure. CD analysis

results were compared with the fractions of secondary structures in the crystallographic structures as assigned by the DSSP method (Kabsch and Sander 1983). These secondary structures were the regular  $\alpha$ -helix, distorted  $\alpha$ -helix, regular  $\beta$ -sheet, distorted  $\beta$ -sheet, turn, and unordered structures.

#### Titration of poly[d(A)] with g5p

CD titrations were performed by adding weighed volumes of concentrated g5p ( $\approx 1 \times 10^{-4}$  M) to poly[d(A)] ( $\approx 60 \mu\text{M}$ ) in buffers of 2 mM or 50 mM  $\text{Na}^+$  (phosphate, pH 7.0). For plots of protein CD changes at 229 nm during the titrations, the nucleic acid contribution was subtracted from the measured CD. The CD value at 229 nm of poly[d(A)] when saturated with g5p (at a [protein monomer]/[nucleotide] molar ratio, [P]/[N], of 0.25) has been shown to be close to that of free poly[d(A)] at about 76°C (Mark and Gray 1997). Thus, at [P]/[N]  $\geq 0.25$ , this constant value for the poly[d(A)] component was subtracted from the measured CD. For samples with [P]/[N] molar ratios between 0 and 0.25, the measured CD at 229 nm was corrected by subtracting a proportional contribution due to the poly[d(A)] (Thompson et al. 1998).

#### Determination of binding affinities

The binding affinities ( $K\omega$ ) of individual proteins for single-stranded poly[d(A)] were determined by the NaCl-induced dissociation of g5p-poly[d(A)] complexes. Saturated g5p-poly[d(A)] complexes were formed at [P]/[N] molar ratios of 0.25 or 0.33 in a buffer of 2 mM  $\text{Na}^+$  (phosphate, pH 7.0). The NaCl concentration was increased by adding weighed aliquots of a 4.0 M NaCl solution. Light scattering was minimal, as in previous work (Mou et al. 1999). Dissociation of the complexes as the salt concentration was increased was determined by the change in the CD at 270 nm of poly[d(A)], calculated as: % CD change =  $100 \times [\text{CD}(\text{ssDNA}) - \text{CD}(\text{complex at given } [\text{NaCl}])]/[\text{CD}(\text{ssDNA}) - \text{CD}(\text{complex with maximum negative CD})]$ , where the CD(ssDNA) was the value at 270 nm for poly[d(A)] at 1 M NaCl. The salt concentrations required for 50% dissociation of the complex were determined by fitting the dissociation curves with a nonlinear, rational, four-parameter regression function (SigmaPlot 5.0, SPSS Inc.). The binding affinities ( $K\omega$ ) for each complex were determined as  $K\omega = 1/(2L)$ , where L was the free dimer concentration (taken to be one-half of the total protein concentration) at 50% dissociation (Thompson et al. 1998).

#### Acknowledgments

We appreciate a critical reading of the manuscript by Dr. Carla W. Gray (Department of Molecular and Cell Biology, The University of Texas at Dallas). This work was performed by T.-C.M. in partial fulfillment of the requirements for the Ph.D. degree in the Department of Molecular and Cell Biology, The University of Texas at Dallas. This work was supported by a Robert A Welch Foundation Grant AT-503 to D.M.G., an NIH Grant RO1 38714 to T.C.T., and an NIH Grant GM-22994 to R.W.W.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

#### References

- Bayley, P.M. 1973. The analysis of circular dichroism of biomolecules. *Prog. Biophys. Mol. Biol.* **27**: 1–76.

- Benevides, J.M., Terwilliger, T.C., Vohnik, S., and Thomas, Jr., G.J. 1996. Raman spectroscopy of the Ff gene V protein and complexes with poly(dA): Nonspecific DNA recognition and binding. *Biochemistry* **35**: 9603–9609.
- Bochkareva, E., Belegu, V., Korolev, S., and Bochkarev, A. 2001. Structure of the major single-stranded DNA-binding domain of replication protein A suggests a dynamic mechanism for DNA binding. *EMBO J.* **20**: 612–618.
- Bollum, F.J. 1966. Biosynthetic polydeoxynucleotides. In *Procedures in nucleic acids research* (Eds. G.L. Cantoni and D.R. Davies), pp. 577–583. Harper and Row, New York.
- Bulsink, H., Harmsen, B.J., and Hilbers, C.W. 1985. Specificity of the binding of bacteriophage M13 encoded gene-5 protein to DNA and RNA studied by means of fluorescence titrations. *J. Biomol. Struct. Dyn.* **3**: 227–247.
- Day, L.A. 1973. Circular dichroism and ultraviolet absorption of a deoxyribonucleic acid binding protein of filamentous bacteriophage. *Biochemistry* **12**: 5329–5339.
- de Jong, E.A., van Duynhoven, J.P., Harmsen, B.J., Konings, R.N., and Hilbers, C.W. 1989. Two-dimensional <sup>1</sup>H nuclear magnetic resonance studies on the gene V-encoded single-stranded DNA-binding protein of the filamentous bacteriophage IKe. I. Structure elucidation of the DNA-binding wing. *J. Mol. Biol.* **206**: 119–132.
- Donovan, J.W., Laskowski, Jr., M., and Scheraga, H.A. 1961. Effects of charged groups on chromophores of lysozyme and of amino acids. *J. Am. Chem. Soc.* **83**: 2686–2694.
- Folkers, P.J., Nilges, M., Folmer, R.H., Konings, R.N., and Hilbers, C.W. 1994. The solution structure of the Tyr41→His mutant of the single-stranded DNA binding protein encoded by gene V of the filamentous bacteriophage M13. *J. Mol. Biol.* **236**: 229–246.
- Folkers, P.J., Stassen, A.P., van Duynhoven, J.P., Harmsen, B.J., Konings, R.N., and Hilbers, C.W. 1991. Characterization of wild-type and mutant M13 gene V proteins by means of <sup>1</sup>H-NMR. *Eur. J. Biochem.* **200**: 139–148.
- Folmer, R.H., Nilges, M., Folkers, P.J., Konings, R.N., and Hilbers, C.W. 1994. A model of the complex between single-stranded DNA and the single-stranded DNA binding protein encoded by gene V of filamentous bacteriophage M13. *J. Mol. Biol.* **240**: 341–357.
- Gray, C.W. 1989. Three-dimensional structure of complexes of single-stranded DNA-binding proteins with DNA. Ike and fd gene 5 proteins form left-handed helices with single-stranded DNA. *J. Mol. Biol.* **208**: 57–64.
- Gray, D.M. 1996. Circular dichroism of protein–nucleic acid interactions. In *Circular dichroism and the conformational analysis of biomolecules* (Ed. G.D. Fasman), pp. 469–500. Plenum Press, New York.
- Guan, Y., Zhang, H., Konings, R.N., Hilbers, C.W., Terwilliger, T.C., and Wang, A.H. 1994. Crystal structures of Y41H and Y41F mutants of gene V protein from Ff phage suggest possible protein–protein interactions in the GVP–ssDNA complex. *Biochemistry* **33**: 7768–7778.
- Guan, Y., Zhang, H., and Wang, A.H.-J. 1995. Electrostatic potential distribution of the gene V protein from Ff phage facilitates cooperative DNA binding: A model of the GVP–ssDNA complex. *Protein Sci.* **4**: 187–197.
- Horstink, L.M., Abseher, R., Nilges, M., and Hilbers, C.W. 1999. Functionally important correlated motions in the single-stranded DNA-binding protein encoded by filamentous phage Pf3. *J. Mol. Biol.* **287**: 569–577.
- Johnson, W.C. 1999. Analyzing protein circular dichroism spectra for accurate secondary structures. *Proteins* **35**: 307–312.
- Kabsch, W. and Sander, C. 1983. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**: 2577–2637.
- Kansy, J.W., Clack, B.A., and Gray, D.M. 1986. The binding of fd gene 5 protein to polydeoxynucleotides: Evidence from CD measurements for two binding modes. *J. Biomol. Struct. Dyn.* **3**: 1079–1110.
- Liang, H. and Terwilliger, T.C. 1991. Reversible denaturation of the gene V protein of bacteriophage f1. *Biochemistry* **30**: 2772–2782.
- Mark, B.L. and Gray, D.M. 1997. Tyrosine mutant helps define overlapping CD bands from fd gene 5 protein–nucleic acid complexes. *Biopolymers* **42**: 337–348.
- Mark, B.L., Terwilliger, T.C., Vaughan, M.R., and Gray, D.M. 1995. Circular dichroism spectroscopy of three tyrosine-to-phenylalanine substitutions of fd gene 5 protein. *Biochemistry* **34**: 12854–12865.
- Mou, T.C., Gray, C.W., and Gray, D.M. 1999. The binding affinity of Ff gene 5 protein depends on the nearest-neighbor composition of the ssDNA substrate. *Biophys. J.* **76**: 1537–1551.
- Murzin, A.G. 1993. OB(oligonucleotide/oligosaccharide binding)-fold: Common structural and functional solution for non-homologous sequences. *EMBO J.* **12**: 861–867.
- Olah, G.A., Gray, D.M., Gray, C.W., Kergil, D.L., Sosnick, T.R., Mark, B.L., Vaughan, M.R., and Trewella, J. 1995. Structures of fd gene 5 protein–nucleic acid complexes: A combined solution scattering and electron microscopy study. *J. Mol. Biol.* **249**: 576–594.
- Prompers, J.J., Folmer, R.H., Nilges, M., Folkers, P.J., Konings, R.N., and Hilbers, C.W. 1995. Refined solution structure of the Tyr41→His mutant of the M13 gene V protein. A comparison with the crystal structure. *Eur. J. Biochem.* **232**: 506–514.
- Provencher, S.W. and Glöckner, J. 1981. Estimation of globular protein secondary structure from circular dichroism. *Biochemistry* **20**: 33–37.
- Record, Jr., M.T., Lohman, T.M., and Dehaseth, P. 1976. Ion effects on ligand–nucleic acid interactions. *J. Mol. Biol.* **107**: 145–158.
- Sang, B.C. and Gray, D.M. 1989. CD measurements show that fd and Ike gene 5 proteins undergo minimal conformational changes upon binding to poly(rA). *Biochemistry* **28**: 9502–9507.
- Schellman, J.A. 1968. Symmetry rules for optical rotation. *Acc. Chem. Res.* **1**: 144–151.
- Skinner, M.M., Zhang, H., Leschnitzer, D.H., Guan, Y., Bellamy, H., Sweet, R.M., Gray, C.W., Konings, R.N., Wang, A.H., and Terwilliger, T.C. 1994. Structure of the gene V protein of bacteriophage f1 determined by multi-wavelength X-ray diffraction on the selenomethionyl protein. *Proc. Natl. Acad. Sci.* **91**: 2071–2075.
- Sreerama, N. and Woody, R.W. 2000a. Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **287**: 252–260.
- Sreerama, N. and Woody, R.W. 2000b. Circular dichroism of peptides and proteins. In *Circular dichroism: Principles and applications* (Eds. N. Berova, K. Nakanishi, and R.W. Woody), Wiley-VCH Inc., New York, pp. 601–620.
- Sreerama, N., Venyaminov, S.Y., and Woody, R.W. 1999. Estimation of the number of  $\alpha$ -helical and  $\beta$ -strand segments in proteins using circular dichroism spectroscopy. *Protein Sci.* **8**: 370–380.
- Stassen, A.P., Harmsen, B.J., Schoenmakers, J.G., Hilbers, C.W., and Konings, R.N. 1992a. Fluorescence studies of the binding of bacteriophage M13 gene V mutant proteins to polynucleotides. *Eur. J. Biochem.* **206**: 605–612.
- Stassen, A.P., Zaman, G.J., van Deursen, J.M., Schoenmakers, J.G., and Konings, R.N. 1992b. Selection and characterization of randomly produced mutants of gene V protein of bacteriophage M13. *Eur. J. Biochem.* **204**: 1003–1014.
- Strickland, E.H. 1974. Aromatic contributions to circular dichroism spectra of proteins. *CRC Crit. Rev. Biochem.* **2**: 113–175.
- Terwilliger, T.C. 1996. Gene V protein dimerization and cooperativity of binding of poly(dA). *Biochemistry* **35**: 16652–16664.
- Terwilliger, T.C., Fulford, W.D., and Zabin, H.B. 1988. A genetic selection for temperature-sensitive variants of the gene V protein of bacteriophage f1. *Nucleic Acids Res.* **16**: 9027–9039.
- Terwilliger, T.C., Zabin, H.B., Horvath, M.P., Sandberg, W.S., and Schlunk, P.M. 1994. In vivo characterization of mutants of the bacteriophage f1 gene V protein isolated by saturation mutagenesis. *J. Mol. Biol.* **236**: 556–571.
- Thompson, T.M., Mark, B.L., Gray, C.W., Terwilliger, T.C., Sreerama, N., Woody, R.W., and Gray, D.M. 1998. Circular dichroism and electron microscopy of a core Y61F mutant of the F1 gene 5 single-stranded DNA-binding protein and theoretical analysis of CD spectra of four Tyr →Phe substitutions. *Biochemistry* **37**: 7463–7477.
- Woody, R.W. 1978. Aromatic side-chain contributions to the far ultraviolet circular dichroism of peptides and proteins. *Biopolymers* **17**: 1451–1467.
- Woody, R.W. and Dunker, A.K. 1996. Aromatic and cystine side-chain circular dichroism in proteins. In *Circular dichroism and the conformational analysis of biomolecules* (ed. G.D. Fasman), pp. 109–157. Plenum Press, New York.
- Zaman, G., Smetsers, A., Kaan, A., Schoenmakers, J., and Konings, R. 1991. Regulation of expression of the genome of bacteriophage M13. Gene V protein regulated translation of the mRNAs encoded by genes I, III, V and X. *Biochim. Biophys. Acta* **1089**: 183–192.
- Zaman, G.J., Schoenmakers, J.G., and Konings, R.N. 1990. Translational regulation of M13 gene II protein by its cognate single-stranded DNA binding protein. *Eur. J. Biochem.* **189**: 119–124.