The T4 translational repressor RegA protein folds into two structural domains, as revealed by the crystal structure (Kang, C.-H., Chan, R., Berger, L., Lockshin, C., Green, L., Gold, L., and Rich, A. (1995) Science 268, 1170–1173). Domain I of the RegA protein contains a four-stranded β-sheet and two α-helices. Domain II contains a four-stranded β-sheet and an unusual 3/10 helix. Since β-sheet residues play a role in a number of protein-RNA interactions, one or both of the β-sheet regions in RegA protein may be involved in RNA binding. To test this possibility, mutagenesis of residues on both β-sheets was performed, and the effects on the RNA binding affinities of RegA protein were measured. Additional sites for mutagenesis were selected from molecular modeling of RegA protein. The RNA binding affinities of three purified mutant RegA proteins were evaluated by fluorescence quenching equilibrium binding assays. The activities of the remainder of the mutant proteins were evaluated by quantitative RNA gel mobility shift assays. The RNA binding site was found to be a surface pocket formed by residues on two loops and an α-helix. Thus, RegA protein appears to use a unique structural motif in binding RNA, which may be related to its unusual RNA recognition properties.

The bacteriophage T4 RegA protein is a unique translational repressor in that it is able to regulate the expression of 15–30 T4 genes, including its own. Previous studies have demonstrated that RegA protein binds to the translation initiation region of target mRNAs (1, 2) and inhibits the formation of ribosome-mRNA initiation complexes (1, 2). The crystal structure of RegA protein by Kang et al. (3) and the MS2 coat protein that participate in sequence-specific RNA binding lie on β-strands within an extended 10-stranded β-sheet in the protein dimer (4). A β-sheet also plays a critical role in the binding of U1 RNA to the small nuclear ribonucleoprotein U1A protein (5).

In considering where RNA might bind on RegA protein, Kang et al. (6) noted that β-sheets often are important contributors to RNA recognition. For example, a β-sheet in glutamyl-tRNA synthetase plays a role in binding the anti-codon region of tRNA (7). Five residues in two loops at the base of the β-sheet, referred to as “jaws,” which interact with the backbone of the U1 RNA hairpin, RegA protein also has two pairs of basic residues (Lys and Lys) in loops at the base of the β-sheet (see Fig. 1), which could be envisioned to function in binding RNA. Although domain II of RegA protein does not exhibit structural similarity with any known RNA binding protein, there is experimental evidence suggesting that domain II residues contribute significantly to RNA binding. For example, the site of

1 The abbreviations used are: RE, recognition element; WT, wild type.
photocross-linking of RegA protein to nucleic acid was found to be \( \text{Pfu} \), which lies within domain II. Also, removal of 13 or 17 residues from the C terminus of RegA (located in domain II) resulted in 100–1000-fold reductions in RNA binding affinity (9). Finally, partial proteolysis of RegA protein, which leads to cleavage at three sites in the C-terminal domain, is reduced by RNA binding (9).

In addition to the above experimental observations and the observed similarities of RegA protein with other RNA-binding proteins, the fact that RegA protein is small (14.6 kDa) and binds to a relatively large binding site on RNA \( (\approx 12 \text{ nucleotides}) \) (4) suggests that residues involved in RNA binding may span both domains of the protein. In fact, Kang et al. (6) have suggested that the two \( \beta \)-sheet regions, which are 25 Å apart, may both participate in RNA binding. To test this possibility, site-specific mutagenesis of surface residues in both domains I and II of RegA protein was performed. Basic and aromatic residues were particularly targeted, since they have been found to play important roles in a number of protein-RNA complexes (11, 15). Ten residues that lie within the central \( \beta \)-strands and \( \beta \)-loop regions of domain I were mutated, and the resulting proteins were assayed for RNA binding in vitro. In addition, residues within \( \beta \)-sheet A of domain II were mutated to test for their contribution to RNA binding. During the course of these studies, molecular models of RegA protein were generated to examine surface regions for potential RNA binding pockets and to determine whether residues to be mutated are located at the surface or in the core of the folded protein. These models revealed the presence of a pocket and a deep cleft at opposite sides of the interface between domains I and II. Subsequent mutagenesis was performed to test the potential role of these two regions in RNA binding. Taken together, the results of these mutagenesis and modeling studies have ruled out the participation of a number of proposed functional residues, revealing instead an unexpected site for RNA binding on RegA protein.

**MATERIALS AND METHODS**

**Reagents and Strains—**Oligodeoxiribonucleotides were synthesized on an Expedite (model 8909) nucleic acid synthesizer by the Medical University of South Carolina Oligosynthesis Facility. Oligoribonucleotides were synthesized by the W. M. Keck Foundation Biotechnology Resource Laboratory (Yale University) and were deprotected as described previously (16) and then purified by perfusion chromatography using Poros™ HQ and R1 columns (PerSeptive Biosciences), essentially as described (16). Poly(U) and poly(U)-agarose were purchased from Amersham Pharmacia Biotech. Escherichia coli AR120 (AcI, N+) was obtained from A. Shatzman (Smith, Klein and French; construction); pAS\(_{\text{regA}}\) was described previously (2).

**RegA Mutagenesis—**Site-directed mutagenesis of residues Lys\(_7\), Lys\(_8\), Lys\(_{35}\), Lys\(_{36}\), Lys\(_{41}\), Lys\(_{42}\), Tyr\(_{45}\), and Tyr\(_{46}\) was carried out by annealing mutagenic oligonucleotides (33–42 nucleotides in length) to a double-stranded plasmid carrying the wild-type (WT) regA gene (pAS\(_{\text{regA}}\)) (2) as described previously (7). The remainder of the mutations were introduced by the Quik-Change™ Sit-directed Mutagenesis protocol (Stratagene). Pfu DNA polymerase was used in a 16-cycle thermal cycling reaction to extend and incorporate the appropriate mutagenic primers, which resulted in nicked circular strands. The methylated, nonmethylated parental DNA was then digested with DpnI, and the DNA was transformed into E. coli AR120 cells. Plasmids were purified from overnight LB/amp cultures (25 ml) using the Qiagen miniprep system. Mutations were then confirmed by DNA sequence analysis, using an ABI 377 automated DNA sequence, by the Medical University of South Carolina Biotechnology Resource Laboratory.

To facilitate purification of selected mutant RegA proteins, the RegA protein expression vector was modified to code for a fusion protein containing four His residues at the COOH terminus of RegA protein. An oligonucleotide containing four CAT codons was inserted into the pAS\(_{\text{regA}}\) vector between the terminus codon (AAAT) and the stop codon (TAA) of regA, using the Quik-Change™ protocol. Insertion of the 12-nucleotide sequence was confirmed by DNA sequencing. Mutations K14A, T18A, R21A, and W81A were then introduced into the regA-His4 vector.

**RegA Protein Purification—**Wild-type (WT), K7L, K41L, and K42L RegA proteins were purified from AR120 cells containing WT or mutant pAS\(_{\text{regA}}\) plasmids following induction of transcription from the phage \( \lambda\) P promoter by naldixic acid treatment, as described previously (7, 17). Protein concentrations for fluorescence analysis were determined by polyacrylamide gel electrophoresis and then pooled and concentrated by centrifugation through Centriprep\(_{10}\)™ (Amicon, Inc.) filtration units.

**Fluorescence Quenching Assays—**Fluorescence quenching assays were performed at 25 °C using an SLM model 8000C spectrophotometer (4). Reverse titrations (addition of poly- or oligoribonucleotide lattice to protein ligand) were performed in 2-mL stirred cuvettes at 25 °C, with final concentrations of 0.1–2.5 \( \mu \text{M}\) in buffer C (10 mM HEPES, pH 7.2, 5 mM Mg\(_{\text{Cl}_2}\), 1 mM EDTA, and 1 mM \( \beta\text{-mercaptoethanol}\) (4) plus 0.150 mM NaCl for gene 44 RE RNA) or 20 mM NaCl for (poly(U)). Data were acquired at an excitation wavelength of 282 nm and an emission wavelength of 347 nm. The effects of photobleaching during titrations were corrected for by monitoring RegA protein fluorescence in a control cuvette. The average photobleaching control was 7.3%. Correction for absorption of incident light by oligonucleotides was made by performing a parallel titration of \( N\)-acetyl-t-tryptophanamide (Sigma) with nucleic acid.

The effect of salt on RNA affinity was determined by "salt-back" titration (18), in which NaCl was added stepwise to RegA protein-poly(U) complexes following reverse titrations. The apparent association constant \( (K_{\text{up}}) \) for poly(U) was calculated from the equivalence point of the titration (i.e., at the addition of an equimolar amount of RNA), assuming a binding site size of \( n = 9 \) nucleotides (4). A single binding site was assumed for RegA protein binding to gene 44 RE (16-mer) RNA (4, 7).

**Preparation of Induced Cell Supernatants—**Cell supernatants were prepared by the method of Johnson and Hecht (19), with slight modifications. E. coli AR120/pAS1regA cells were grown at 37 °C to an \( A_{590} \) of 0.8–0.9, and then WT or mutant RegA expression was induced for \( \sim 15 \) h by the addition of naldixic acid (80 \( \mu \text{g}/\text{ml} \)). After centrifugation of a 100-ml culture, cell pellets were frozen in an ethanoldry ice bath for 2 min, followed by thawing in an ice water bath for 8 min, repeated for a total of four cycles. The pellet was resuspended in 800 \( \mu \text{l} \) of 20 mM Tris-HCl, pH 8.0, 100 mM NaCl, and 1 mM EDTA and incubated for 30 min. The suspension was centrifuged in a microcentrifuge for 3 min; the supernatant was recovered, and the pellet was resuspended in 1 ml of 1× TE (10 mM Tris, 1 mM EDTA, \( pH \) 7.4). Cell supernatants were stored in 80-\( \mu \text{l} \) aliquots at \( -20 ^\circ\text{C} \).

Evaluation of the protein content of both the supernatant and the pellet resuspension was carried out by SDS-polyacrylamide gel electrophoresis. Some mutant RegA proteins were not completely soluble when produced at 37 °C, and in those cases, induction was repeated at 25 °C. In each case, solubility was increased sufficiently to allow gel shift assays to be performed. RegA protein concentrations in cell lysates were determined by quantitation of protein fluorescence in gels stained with SYPRO Orange (Molecular Probes, Inc., Eugene, OR) compared with a standard curve of known concentrations of purified RegA protein, using a Molecular Dynamics Storm Imager. Protein concentrations had an error range of 1–16% (average 7%).

**RNA Gel Mobility Shift Assays—**Purified gene 44 RE RNA (5′-GAUAGGGAAGAU-3′) was 5′-32P-end-labeled by treatment with T4 polynucleotide kinase and \( \gamma\text{-32P}\text{-ATP} \) (20). Increasing volumes of induced cell supernatants were incubated with a constant amount of \( 3\text{2P-labeled gene 44 RE RNA (10 \mu l) to generate binding curves. Binding was carried out in 10 mM Tris-HCl, \( pH \) 7.5, 50 mM NaCl, and 1 mM EDTA at 4 °C for 15 min. Binding reactions were performed in a 20-\( \mu l \) final volume with final concentrations of 2.5–40 \( \mu \text{g}/\text{ml} \) RegA protein. A freshly thawed aliquot of supernatant was used for each experiment and then discarded. Reaction products were analyzed by electrophoresis on a native 8% polyacrylamide gel in 0.5 × TBE (89 mM Tris, 89 mM boric acid, 4 mM EDTA, \( pH \) 8.3) at 4 °C. To eliminate hydrolysis of RNA
RNA Binding Domain of RegA Protein

**RESULTS**

*Mutagenesis of Basic Loop Residues of RegA Protein*—From inspection of the structure of RegA protein, Kang et al. (6) noted that in domain I of RegA, basic residues in the loops at the base of the central two \( \beta \)-strands, *i.e.*, Lys\(^{50} \), Lys\(^{52} \), and Lys\(^{41} \) (Fig. 1) are located in positions similar to basic residues of U1A protein (residues Lys\(^{22} \), Lys\(^{23} \), Lys\(^{50} \), and Arg\(^{52} \)) that function in RNA binding (12). To test if these residues form electrostatic interactions with RNA, as they do in U1A, individual lysine to leucine substitutions were introduced at each of these residues to remove the basic charge and leave a similarly sized side chain. Mutant proteins K7L, K41L, and K42L were purified (see “Materials and Methods”), and their equilibrium RNA binding affinities were measured by fluorescence quenching assays (4), as shown in Fig. 2. Like many other nucleic acid-binding proteins, the intrinsic tryptophan fluorescence of RegA protein is quenched upon binding to nucleic acids (4). This quenching is due to a change in the environment of one or more of the three tryptophan residues in RegA, presumably reflecting either a direct interaction of the tryptophan(s) with RNA or a conformational change in RegA upon RNA binding.

Because the fraction of the maximal quenching obtained is directly proportional to the fraction of RegA protein bound to RNA (4), fluorescence quenching assays offer an accurate method of determining equilibrium binding constants.

The affinities of mutant proteins for a specific RNA (T4 gene 44 RE RNA (4)) and a nonspecific RNA (poly(U)) were assessed to evaluate the effect of these mutations on the RNA binding properties of RegA protein (Table I). As shown in Fig. 2A and Table I, all three purified mutant proteins bound 44 RE RNA with affinities similar to that of wild type (WT) RegA protein. In addition, K41L and K42L RegA proteins bound poly(U) with WT affinity (Table I); however, K7L bound poly(U) with approximately 10-fold lower affinity. Previous studies (4) have indicated that RegA protein binds polynucleotides with weak cooperativity (i.e. with a cooperativity parameter \( \omega \) of \( \sim 10–12 \)). Thus, the lower affinity of K7L for poly(U) could result from a decrease in RNA binding affinity or from alterations in protein-protein interactions leading to reduced cooperativity in RNA binding.

To assess the effects of the mutations on electrostatic interactions between RegA and RNA, salt-back titrations (18) were performed, in which NaCl was added incrementally to RegA protein-poly(U) complexes. As shown in Fig. 2C, RNA complexes formed with mutated proteins K41L and K42L exhibited salt sensitivities similar to that of WT RegA protein, while K7L-poly(U) complexes were considerably more sensitive to ionic strength. These data suggest the observed reduction in \( K_{\text{app}} \) for nonspecific poly(U) (Table II) may be due to decreased protein affinity for RNA, rather than decreased cooperativity, since RegA protein-protein interactions have been found to be unaffected by salt concentrations (150 mM NaCl) that disrupted the K7L RNA complexes (7).

**Evaluation of RNA Binding Affinities by Gel Shift Assays**—To eliminate the requirement for purification of individual mutant proteins, the activities of the remainder of the mutated RegA proteins were assessed by RNA gel mobility shift assays (21, 22) rather than fluorescence quenching as-
RNA Binding Domain of RegA Protein

**Table I**

| Protein      | $K_{app}^{a} \pm \text{S.D. for gene 44 RNA}^{b}$ | $K_{app}^{c} \pm \text{S.D. for poly(U)}^{d}$ |
|--------------|---------------------------------|---------------------------------|
| WT           | $1.9 \pm 0.8 \times 10^7$       | $4.6 \pm 0.8 \times 10^7$       |
| K7L          | $1.1 \pm 0.1 \times 10^7$       | $4.2 \pm 0.5 \times 10^7$       |
| K41L         | $1.6 \pm 0.3 \times 10^7$       | $1.4 \pm 0.5 \times 10^7$       |
| K42L         | $2.8 \times 10^7$              | $1.9 \times 10^7$              |

*Binding affinities were determined by equilibrium fluorescence quenching assays (4); S.D. values are given in those instances where multiple determinations were carried out.

**Table II**

| Mutant protein | $K_{app}^{a} \pm \text{S.D.}$ | Location of residue |
|----------------|--------------------------------|---------------------|
| Wild type      | $4.7 \pm 0.9$                  | $\beta_1$           |
| I4A            | $2.7 \pm 0.3$                  | $\beta_1$           |
| T6A            | $3.1 \pm 0.2$                  | $\beta_1$           |
| L6R            | $5.7 \pm 1.5$                  | $\beta_1$           |
| K7L            | $4.8 \pm 1.2$                  | Loop 1              |
| K8L            | $2.1 \pm 0.1$                  | Loop 1              |
| E10A           | $14.7 \pm 2.5$                 | $\alpha$-Helix A    |
| D11A           | $13.1 \pm 0.5$                 | $\alpha$-Helix A    |
| F12A           | $2.4 \pm 0.2$                  | $\alpha$-Helix A    |
| K71A           | $0.7 \pm 0.1$                  | $\alpha$-Helix A    |
| K16L           | $1.7 \pm 0.1$                  | $\alpha$-Helix A    |
| E17A           | $0.8 \pm 0.1$                  | $\alpha$-Helix A    |
| T15A           | $0.15 \pm 0.01$                | $\alpha$-Helix A    |
| R21A           | $0.5$                          | $\alpha$-Helix A    |
| K41L           | $4.9 \pm 0.5$                  | $\beta_4$           |
| K42L           | $2.6 \pm 0.1$                  | (Hairpin) loop 4    |
| L44K           | $3.2 \pm 0.2$                  | $\beta_5$           |
| Y45A           | $3.3 \pm 0.2$                  | $\beta_5$           |
| Y46A           | $1.3 \pm 0.1$                  | $\beta_5$           |
| W76A           | $4.8 \pm 0.1$                  | $\alpha$-Helix C    |
| W81A           | $0.7$                          | Loop 6              |
| L85A           | $7.4 \pm 0.8$                  | Loop 6              |
| F933Y          | $5.4 \pm 0.2$                  | $\beta_6$           |

* $K_{app}^{a}$ was determined by quantitative RNA gel shift analysis (see “Materials and Methods”). Binding was performed in 10 mM Tris (pH 7.5), 50 mM NaCl, and 1 mM EDTA. Values are the mean of 2–4 experiments except for WT (nine experiments).

** FIG. 2. **Reverse titration of RegA protein with synthetic RNAs. RNA binding was measured as quenching of RegA protein’s intrinsic tryptophan fluorescence (%Q) upon serial additions of RNA. A, specific binding of wild type RegA (0.2 $\mu$m) and K7L (0.2 $\mu$m), K41L (0.2 $\mu$m), and K42L (0.1 $\mu$m) mutants to gene 44 RE RNA in buffer C plus 150 mM NaCl. B, nonspecific binding of wild type RegA (0.08 $\mu$m) and K7L (0.2 $\mu$m), K41L (0.04 $\mu$m), and K42L (0.11 $\mu$m) mutants to poly(U) in buffer C plus 10 mM NaCl. Concentration of added poly(U) is plotted as [phosphate]/[protein], and a binding site size of 9 nucleotides/protein was used to calculate $K_{app}$ from the binding curves. Titrations were performed at 25 °C. C, salt-back titrations of RegA proteins bound to poly(U). Titrations were made by the addition of NaCl to samples saturated with poly(U). The percentage of initial fluorescence (% fluorescence) was estimated from comparison with a photobleach control of each protein, which was corrected for dilution effects.

say, as shown in Fig. 3, the addition of lysed cell supernatants containing WT RegA protein (induced AR120/pAS,regA cells) to $^{32}$P-labeled gene 44 RE RNA produced a shift in the electrophoretic mobility of the RNA, while cell supernatants that did not contain RegA protein (induced AR120/pAS, cells) did not alter RNA mobility. To determine the affinity of a mutant protein for gene 44 RE RNA, increasing volumes of cell supernatants were incubated with a constant amount of RNA. In these assays, final RegA protein concentrations (see “Materials and Methods”) ranged from 2.5 to 40 nM and [RNA] was 10 nM. The fraction of $^{32}$P-labeled RNA bound by RegA protein in each reaction was determined by PhosphorImager analysis (see “Materials and Methods”), to generate a titration binding curve. Cell supernatants containing K7L, K41L, and K42L proteins each produced shifts in RNA mobility, with $K_{app}$ values of $4.8 \times 10^7$ $\text{M}^{-1}$, $4.9 \times 10^7$ $\text{M}^{-1}$, and $2.6 \times 10^7$ $\text{M}^{-1}$, respectively, compared with the affinity of WT RegA protein of $4.7 \times 10^7$ $\text{M}^{-1}$ (Table II). The WT RegA protein affinity measured by gel shift assays ($4.7 \times 10^7$ $\text{M}^{-1}$ in 50 mM NaCl) agrees well with the value determined by fluorescence equilibrium binding assays ($4.0 \times 10^7$ $\text{M}^{-1}$, in 50 mM NaCl [Ref. 4, Fig. 8]). This supports the validity of the quantitative gel shift method...
Fig. 3. RNA binding by WT and mutant RegA proteins, determined by gel mobility shift assays. \(^\text{13}^3\text{P}-\text{Labeled gene 44 RE RNA (10 nM) was incubated with increasing volumes of E. coli cell supernatants (20-\mu l total volume in buffer A) for 15 min on ice and then loaded onto a native 8% polyacrylamide gel in 0.5 × TBE. The gel was electrophoresed at 5 V/cm for 4 h in a cold room. Lane 1, no supernatant; lanes 2 and 3, 1 and 2 \mu l, respectively, of supernatant from induced AR120/pAS\(_5\) cells, which do not contain RegA; lanes 4–7, 1, 2, 4, and 8 \mu l, respectively, of supernatants from induced AR120/pAS\(_5\),RegA cells containing WT RegA (final [RegA] was 5–40 nM); lanes 8–11, supernatants containing 5, 10, 20, and 40 nM RegA K7L, respectively; lanes 12 and 13, supernatants containing 20 and 40 nM RegA K41L, respectively; lanes 14 and 15, supernatant containing 20 and 40 nM RegA K42L, respectively. RegA protein concentrations in supernatants were determined by quantitation of Sypro Orange-stained SDS gels (see “Materials and Methods”).

to measure RNA binding affinities for mutant RegA proteins. Gel shift assays demonstrated that K85L RegA protein bound RNA with a \(K_\text{app}\) value of 2.1 \times 10^7 M\(^{-1}\), indicating this mutation, like the K42L mutation, has a small effect on RNA binding (2-fold reduction).

**Mutagenesis of Residues on \(\beta\)-Sheet B of Domain I**—In the U1A-RNA complex, three residues within the RNP-2 segment (\(\beta\)-strand 1) (Tyr\(^{118}\), Asn\(^{15}\), and Asn\(^{16}\)) form bonds with RNA (12). In addition, three residues in the RNP-1 segment (on \(\beta\)-strand 3) (Arg\(^{52}\), Gin\(^{54}\), and Phe\(^{56}\)) form critical interactions with RNA, which include formation of base-specific hydrogen bonds (Arg\(^{52}\)) and stacking with RNA bases (Phe\(^{56}\)). To see if residues on the central \(\beta\)-strands of \(\beta\)-sheet B (in domain I) play a role in RegA-RNA complex formation, mutations were introduced at Ile\(^{10}\), Thr\(^{12}\), and Leu\(^{4}\), on \(\beta\)-strand 1 and at Leu\(^{41}\), Tyr\(^{45}\), and Tyr\(^{46}\) on \(\beta\)-strand 5. Ala substitutions were made at Ile\(^{4}\) and Thr\(^{12}\) to eliminate potential hydrophobic and hydrogen bond interactions, respectively; Leu\(^{4}\) was mutated to Arg, and Leu\(^{44}\) was mutated to Lys to introduce basic charges at those hydrophobic sites, and Tyr\(^{45}\) and Tyr\(^{46}\) were mutated to Ala to eliminate potential base-stacking interactions with RNA. Of these six mutations, only the Tyr\(^{46}\) substitution significantly affected RNA binding, producing a 3-fold reduction. Thus, none of these mutations eliminated or severely (>10-fold) reduced RNA binding. In contrast, mutations in RNP-1 and RNP-2 residues in U1A protein reduced affinity 15–140-fold or eliminated RNA binding (23). Taken together with the effects of the basic loop mutations, these results suggest that \(\beta\)-sheet B and adjacent loop residues of domain I do not contribute significantly to RNA binding.

**Mutagenesis of Domain II of RegA**—To determine if residues on \(\beta\)-sheet A play a role in RNA binding, mutations were introduced at surface residues Asn\(^{26}\), Lys\(^{30}\), Tyr\(^{31}\), Gin\(^{34}\), and Lys\(^{113}\). Mutations were also introduced at nearby basic residues on the hairpin loop (Lys\(^{26}\)), loop 7 (Lys\(^{157}\)), helix D (Lys\(^{199}\)), and loop 8 (Lys\(^{117}\)). As shown in Table III, none of the nine mutations significantly reduced RNA affinity. Thus, \(\beta\)-sheet A and basic residues flanking it do not appear to directly participate in RNA binding. It is interesting to note that for the three mutations in this region and the six mutations in domain I where Leu was substituted for Lys (Tables II and III), the altered proteins remained soluble and retained WT RNA binding affinity, indicating that the introduction of a hydrophobic residue on the surface of RegA does not lead to misfolding of the protein. Overall, the lack of effect on RNA affinity of the 19 mutations described above suggests that when mutations are restricted to surface residues, as deduced from crystal structure data, RegA protein retains its native structure.

**Modeling of RegA Protein**—To examine the surface topography of RegA protein for potential RNA binding pockets or platforms, space filling and ribbon models were generated from the RegA protein crystal structure coordinates (Protein Data Bank [Rutgers University], accession number 1REG), using the program SYBYL (Tripos, Inc.). These models revealed the presence of a pocket (loosely defined by residues Phe\(^{106}\), Lys\(^{14}\), Arg\(^{31}\), and Trp\(^{35}\)) that is formed at the interface of domains I and II and a cleft, formed between the C-terminal 5 residues and residues on \(\alpha\)-helix B. As shown in Fig. 4, the cleft is narrow (10 \AA) and deep, and there is an aromatic side chain (Tyr\(^{119}\)) at the top of the cleft. The pocket region is shallow and contains two unusually exposed aromatic side chains (Fig. 4). The fact that this pocket contains Phe\(^{106}\), which was previously found to be the site of photocross-linking to nucleic acid (8), as well as two basic residues, makes this region a strong candidate for the site of RNA binding.

**Mutagenesis of the Aromatic Pocket**—To test the role of the aromatic pocket region in RNA binding, individual substitutions were made at Thr\(^{31}\), Lys\(^{34}\), and Arg\(^{21}\). As shown in Fig. 5, the K14A mutation greatly reduced the RNA binding affinity of RegA protein, while the R21A mutation abolished RNA binding. Substitution of Thr\(^{31}\) with Ala also abolished RNA binding (see Table II). Previous studies have shown that mutations at Phe\(^{106}\) do not significantly affect RNA binding affinity (9), which suggested that although this residue lies at the interface of RegA-nucleic acid complexes (8), it does not form a bond to RNA.

| Residue | Affinity of domain II mutant RegA proteins for gene 44 RE RNA |  |
|---------|-------------------------------------------------|---|
| WT      | 4.7 \pm 0.9                                      |  |
| N26V    | 4.2 \pm 1.2                                      | \beta_2 |
| K28L    | 5.0 \pm 0.4                                      | (Hairpin) loop 3 |
| K30A    | 2.5 \pm 0.4                                      | \beta_3 |
| Y33A    | 5.5 \pm 0.8                                      | \beta_3 |
| Q34A    | 4.1 \pm 1.2                                      | \beta_3 |
| F106V   | 3.6                                             | Loop 7 |
| K107L   | 3.7 \pm 0.1                                      | Loop 7 |
| K109L   | 4.1 \pm 0.1                                      | Helix D |
| W112L   | 4.3                                             | Turn |
| K113L   | 4.1 \pm 0.7                                      | \beta_8 |
| K117L   | 4.9 \pm 1.1                                      | Loop 9 |
| Y118A   | 3.1 \pm 0.1                                      | Loop 9 |
| T119    | 2.5 \pm 0.6                                      |  |
protein and were not mutated. As shown in Table II, the T18A substitution had an even greater effect than the K14A substitution, reducing affinity by approximately 30-fold. The Phe81, Lys107, and Glu17 mutations produced only moderate or no reductions in RNA affinity, while the Glu10 and Asp11 mutations significantly altered the RNA binding affinity of RegA proteins with mutations in the aromatic pocket could be due to loss of a protein-RNA interaction or due to protein misfolding.

To determine if mutant proteins are properly folded, three proteins with mutations in the aromatic pocket could be due to loss of a protein-RNA interaction or due to protein misfolding. The observed reduction in the RNA binding affinity of RegA proteins further supports the conclusion that the K14A, T18A, and W81A mutations do not significantly affect the secondary structure of RegA protein. Gel shift assays indicate that WT-His<sub>4</sub> RegA protein binds gene 44 RE RNA with an affinity similar to that of WT RegA (data not shown). This is consistent with the report of Allen and Miller (25) that the addition of six His to the COOH terminus of RegA protein does not dramatically alter the RNA binding properties of RegA protein. The fact that the overall shape and intensities of the CD spectra for the three mutant proteins are the same as that of WT RegA protein suggests that there is little difference in the proteins’ secondary structures. The observation that the spectra crossover points (−198 nm), which are independent of concentration effects, are also similar for the mutant and WT proteins further supports the conclusion that the K14A, T18A, and W81A mutations do not significantly affect the secondary structure of RegA protein.

As noted above, the quenching of the intrinsic fluorescence of RegA that occurs upon RNA binding is due a change in the environment of one or more of the three tryptophan residues. The effect of the W81A mutation on RNA affinity suggests that Trp<sup>81</sup> contributes to the observed fluorescence change. To determine if either of the other two tryptophan residues are involved in RNA binding, Ala substitutions were introduced at Trp<sup>76</sup> and Trp<sup>112</sup>. As shown in Tables II and III, mutagenesis of Trp<sup>76</sup> did not affect RNA affinity, while mutation of Trp<sup>112</sup> eliminated RNA binding. Jozwik and Miller (24) saw a similar effect in vivo for a W112C mutation, suggesting that Trp<sup>112</sup> plays a role in RNA binding. However, the crystal structure of RegA shows that the side chain of Trp<sup>112</sup> is predominately buried, suggesting that the observed loss of RNA binding could be due to structural alterations rather than functional alterations. Thus, structural studies would need to be performed before a conclusion can be made about the role of Trp<sup>112</sup> in RNA binding.

Assessment of the Effect of Mutations on RegA Structure—

The observed reduction in the RNA binding affinity of RegA proteins with mutations in the aromatic pocket could be due to loss of a protein-RNA interaction or due to protein misfolding. To determine if mutant proteins are properly folded, three mutant proteins were purified, and their conformations were assessed by CD spectroscopy. To facilitate purification, mutations were introduced into an expression vector coding for RegA protein with four His residues fused to the COOH terminus. The His<sub>4</sub> tag was added to the COOH terminus because the NH<sub>2</sub> terminus is in the core of the protein, whereas the COOH terminus is on the surface of RegA protein. WT-His<sub>4</sub>, K14A-His<sub>4</sub>, T18A-His<sub>4</sub>, and W81A-His<sub>4</sub> proteins were purified, and their ultraviolet CD spectra were examined. A fourth mutant, R21A-His<sub>4</sub>, was expressed but was not purified because only a fraction of the recombinant protein was present in cell extract supernatants.

As shown in Fig. 6A, RegA proteins K14A-His<sub>4</sub>, T18A-His<sub>4</sub>, and W81A-His<sub>4</sub> exhibit spectra with a single minimum at 208 nm and a shoulder at 220 nm, which is characteristic of the WT RegA protein spectrum (Ref. 9 and Fig. 6). Comparison of the CD spectra of WT and WT-His<sub>4</sub> RegA proteins showed no detectable differences (Fig. 6B), confirming that the His<sub>4</sub> tag does not alter the secondary structure of RegA protein. Gel shift assays indicate that WT-His<sub>4</sub> RegA protein binds gene 44 RE RNA with an affinity similar to that of WT RegA (data not shown). This is consistent with the report of Allen and Miller (25) that the addition of six His to the COOH terminus of RegA protein does not dramatically alter the RNA binding properties of RegA protein. The fact that the overall shape and intensities of the CD spectra for the three mutant proteins are the same as that of WT RegA protein suggests that there is little difference in the proteins’ secondary structures. The observation that the spectra crossover points (~198 nm), which are independent of concentration effects, are also similar for the mutant and WT proteins further supports the conclusion that the K14A, T18A, and W81A mutations do not significantly affect the secondary structure of RegA protein.
DISCUSSION

In this report, the functional activity of surface residues within both domains I and II of RegA protein have been explored. Ten residues in or near $\beta$-sheet B of domain I were assayed by site-specific mutagenesis. Substitutions of residues in the central two $\beta$-strands and the basic loop regions of domain I of RegA had little (2–3-fold) or no effect on RNA binding. In contrast, mutagenesis of residues in the central $\beta$-strands (containing the RNP-1 and RNP-2 residues) of U1A protein (e.g. residues Arg$^{52}$, Gln$^{54}$, and Phe$^{56}$) reduced affinity.
10–100-fold or completely abolishment of RNA binding (23, 26). The minor effects of mutations in potentially analogous residues in RegA (i.e., residues Lys42, Leu44, and Tyr46) suggest that \( \beta \)-sheet B does not function as an RNA binding platform. Similarly, mutation of basic loop residues of U1A protein (e.g., Lys22 and Lys23) had large effects on RNA binding affinity (26), while mutation of basic loop residues in RegA protein had only minor effects. In addition, five surface residues on \( \beta \)-sheet A of domain II do not appear to contribute to RNA binding. This is an important finding, because \( \beta \)-sheet residues have been found to play critical roles in protein-RNA complexes, even when different overall structural motifs are involved (e.g., the MS2 coat protein versus U1A protein).

These extensive mutagenesis studies of RegA protein have shown that there are four residues (Lys14, Thr18, Arg21, and Trp81) where substitutions significantly impair or abolish RNA binding. Since three of the four mutations do not affect the secondary structure of RegA protein, it is likely that the observed effects on RNA binding affinity are due to loss of RNA-protein interactions. The clustered location of these residues near the site of photocross-linking to nucleic acid (8) suggests that RNA binds to a surface pocket on RegA protein formed by residues on loop, helix, and loop secondary structures. As shown in Fig. 7, this arrangement positions aromatic residues on the loops and basic and polar groups on the \( \alpha \)-helix. The helix residues offer a number of side chains that could form hydrogen bonds to RNA bases or the ribose-phosphate backbone. As shown in Fig. 7, the three helix residues that are critical for RNA binding are clustered on one face of helix A, spanning approximately 90° of the helix. Within this helix are two acidic side chains (Asp11 and Glu10) (Fig. 7), where mutations that eliminate the negative charge slightly increase RNA binding affinity. This may indicate that these residues lie near the binding site of the ribose-phosphate backbone and that some charge repulsion may occur in the WT protein. It is interesting to note that two of the three helix residues that potentially participate in RNA binding are charged and that hydrogen bonds involving a charged residue contribute more to the free energy of complex formation than hydrogen bonds between uncharged groups (27). In studies with tyrosyl-tRNA synthetase, Fersht et al. (27) have shown that a hydrogen bond to a charged group can contribute up to 3–4 kcal/mol to complex stability, which provides a factor of about 10 in specificity. In the case of RegA protein, mutagenesis studies have demonstrated that Arg21 may make a high energy hydrogen bond to RNA, with smaller energy contributions coming from hydrogen bonds between RNA and Lys14 and Thr18. This is consistent with an earlier estimate that the number of ionic interactions involved in RegA-RNA binding is between two and three, based on the effect of [NaCl] on RegA-RNA affinity (4). Additional stability (and specificity) to the complex apparently comes from interactions between RNA and Trp81, which may include base-stacking interactions. Bonds formed between RNA and main chain atoms are also likely to make contributions to complex stability, although these bonds would not be detected by the mutagenesis approach used here. Small contributions also may come from side chains of adjacent residues in the two loop regions; however, the bulk of the energy of complex formation apparently comes from bonds to Lys14, Thr18, Trp81, and potentially Arg21.

Prior to solution of the crystal structure of RegA protein, Jozwik and Miller (24) used a genetic approach to identify functionally important residues in RegA. From examination of the \textit{in vitro} translational repression properties of 15 regA mutants, they identified five mutants that exhibited altered repression activity, where some target genes were repressed, but not others. Interestingly, two of the mutants with altered repression specificity had substitutions in residues on the RNA binding helix, i.e., mutants T18A and E17K. The reduced RNA affinity of the T18A mutation observed in our \textit{in vitro} studies is consistent with the reported repression of high affinity target RNAs and the loss of repression of the low affinity targets in vivo (24). The roles of the other residues where mutation affected repression specificity, i.e., Ala25, Asp72, and Ser73, are not as clear, in that Ala25 is located some distance from the binding pocket, and Asp72 and Ser73 are buried in the core of the protein.

Are there other examples of protein helix-RNA interactions? \( \alpha \)-Helices frequently play a role in protein binding to double-stranded DNA, where both helix-turn-helix and zinc finger motifs use \( \alpha \)-helices as recognition elements in binding to the major groove of DNA (for discussion, see Ref. 28). Although fewer examples of helices being involved in protein-RNA interactions are known, one example is the binding of human immunodeficiency virus REV protein to REV response element RNA. A 19-amino acid peptide corresponding to the RNA binding site of REV forms an \( \alpha \)-helix in solution and has been found to bind specifically to the REV response element hairpin IIB RNA (29). Genetic studies have demonstrated the importance of specific residues on the helix in RNA recognition and have led to development of a model of how the REV helix might dock in the major groove of double-stranded RNA (30). Presumably, the helix projects from the surface of REV, where it can position side chains into the groove of the RNA. In RegA protein, the helix lies within a recessed pocket, which potentially offers target single-stranded RNAs multiple opportunities for interaction with residues in the sides and trough of the pocket. The formation of an RNA binding pocket by residues in helix and loop regions of the protein appears to be unique to RegA protein and may be related to the protein’s unusual RNA recognition properties.

Two remaining questions are how RegA protein discriminates between target and non-target RNAs and what features contribute to the observed differences in affinities for target RNAs. A recurrent theme in protein-RNA recognition is that RNA structure is often altered by protein binding (for discussion, see Ref. 31). Examples in which RNA distortion occurs upon protein binding include the binding of tRNA\textsuperscript{Glu} to glutaminyl tRNA synthetase (10), the binding to U1A protein to U1 RNA (32, 33), and binding of the REV peptide to the REV response element (34). In the case of RegA protein, RNA discrimination may well depend on deformability of the RNA as well as nucleotide sequence recognition. In fact, since deformability can depend on nucleotide sequence (10), the ability of an RNA to alter its structure to fit into a pocket on RegA protein may be a form of nucleotide discrimination. The previous finding that two dodecamer RNAs with a single base difference have different affinities for RegA protein but have the same structure in solution (35) is consistent with RNA deformability being a critical element in RegA-RNA recognition. Now that the unique binding site on RegA protein has been deciphered, future experiments can be directed at determining how different RNAs might be accommodated in the binding pocket.

\textbf{Acknowledgments—}We thank Chul-Hee Kang (Washington State University) and Alexander Rich (Massachusetts Institute of Technology) for many helpful discussions. We are also grateful to Erika Bullesbach and George Fullbright (Medical University of South Carolina) for performing the CD spectroscopy and to Shelley Elvington for excellent technical assistance. We thank the Medical University of South Carolina Biotechnology Resource Laboratory for excellent DNA sequence analysis and synthesis of oligoribonucleotides and the W. M. Keck Foundation Biotechnology Resource Laboratory (Yale University) for amino acid analyses. J. G. thanks Starr Hazard of the Medical Univer-
REFERENCES

1. Winter, R. B., Morrisey, L., Gauss, P., Gold, L., Hsu, T., and Karam, J. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 7822–7826
2. Webster, K. W., Adari, H. Y., and Spicer, E. K. (1989) Nucleic Acids Res. 17, 10947–10968
3. Unnithan, S., Green, L., Morrisey, L., Binkley, J., Singer, B., Karam, J., and Gold, L. (1990) Nucleic Acids Res. 18, 7083–7092
4. Webster, K. R., and Spicer, E. K. (1990) J. Biol. Chem. 265, 19007–19014
5. Karam, J., Gold, L., Singer, B. S., and Dawson, M. (1981) Proc. Natl. Acad. Sci. U. S. A. 78, 4669–4673
6. Kang, C.-H., Chan, R., Berger, I., Lockshin, C., Green, L., Gold, L., and Rich, A. (1995) Science 268, 1170–1173
7. Phillips, C. A., Gordon, J., and Spicer, E. K. (1996) Nucleic Acids Res. 24, 4319–4326
8. Webster, K. R., Keill, S., Konigsberg, W., Williams, K. R., and Spicer, E. K. (1992) J. Biol. Chem. 267, 26097–26103
9. O'Malley, S. M., Sattar, A. K. M., Williams, K. R., and Spicer, E. K. (1995) J. Biol. Chem. 270, 5107–5114
10. Rould, M. A., Perona, J. J., Still, D., and Steitz, T. A. (1989) Science 246, 1135–1142
11. Valesgard, K., Murray, J. B., Stockley, P. G., Stonehouse, N. J., and Liljas, L. (1994) Nature 371, 623–626
12. Oubridge, C., Ito, N., Evans, P. R., Teo, C-H., and Nagai, K. (1994) Nature 372, 432–438
13. Birney, E., Kumar, S., and Krainer, A. R. (1993) Nucleic Acids Res. 21, 5803–5816
14. Burd, C. G., and Dreyfus, G. (1994) Science 265, 615–621
15. Shamo, Y., Krueger, U., Rice, L. M., Williams, K. R., and Steitz, T. A. (1997) Nat. Struct. Biol. 4, 215–222
16. Webster, K. W., Shamo, Y., Konigsberg, W. H., and Spicer, E. K. (1991) BioTechniques 11, 658–661
17. Adari, H. Y., Rose, K., Williams, K. W., Konigsberg, W. H., Lin, T.-C., and Spicer, E. K. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 1901–1905
18. MeSwiggen, J. A., Bear, D. G., and Von Hippel, P. H. (1988) J. Mol. Biol. 199, 609–622
19. Johnson, B. H., and Hecht, M. H. (1994) BioTechniques 12, 114–117
20. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., p. 5.86, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
21. Rebar, E. J., and Pabo, C. O. (1994) Science 263, 671–673
22. Fried, M. G., and Crothers, D. M. (1983) Nucleic Acids Res. 11, 141–158
23. Jessen, T. H., Oubridge, C., Teo, C. H., Pritchard, C., and Nagai, K. (1991) EMBO J. 10, 3447–3456
24. Jozwik, C. E., and Miller, E. S. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 5053–5057
25. Allen, S. V., and Miller, E. S. (1999) Anal. Biochem. 269, 32–37
26. Nagai, K., Oubridge, C., Jessen, T. H., Li, J., and Evans, P. R. (1990) Nature 348, 515–520
27. Fersht, A. R., Shi, J.-P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. Y., and Winter, G. (1985) Nature 314, 235–238
28. Brandon, C., and Tooze, J. (1998) Introduction to Protein Structure, 2nd Ed., pp. 129–147, Garland Publishing, New York
29. Tan, R., Chen, L., Buettnner, J. A., Hudson, D., and Frankel, A. D. (1993) Cell 73, 1031–1040
30. Jain, C., and Belasco, J. G. (1996) Cell 87, 115–125
31. Draper, D. E. (1995) Annu. Rev. Biochem. 64, 583–620
32. Hall, K. B. (1994) Biochemistry 33, 10976–10988
33. Allain, P.-T.-H., Gubser, C. C., Howe, P. W. A., Nagai, K., Neuhaus, D., and Varani, G. (1996) Nature 380, 646–650
34. Battiste, J. L., Mao, H., Rao, N. S., Tan, R., Muhandiram, D. R., Kay, L. E., Frankel, A. D., and Williams, J. R. (1996) Science 273, 1547–1551
35. Szewczak, A. A., Webster, K. R., Spicer, E. K., and Moore, P. B. (1991) J. Biol. Chem. 266, 17832–17837