CONFORMATIONAL PREDICTION AND CIRCULAR DICHROISM STUDIES ON RIBOSOMAL PROTEIN S4 FROM *ESCHERICHIA COLI*

SUSAN H. ALLEN AND KIN-PING WONG, *Department of Biochemistry, University of Kansas Medical Center, Kansas City, Kansas 66103 U.S.A.*

ABSTRACT The conformation of ribosomal protein S4 from *Escherichia coli* has been studied by circular dichroism (CD) and shown to possess unique conformation free in solution. The near ultraviolet spectrum suggests the existence of unique tertiary structural environment for the aromatic amino acid residues. The far ultraviolet spectrum gives an estimation of its secondary structure which is 32% α-helix and 14% β-structure in reconstitution buffer at 25°C. The conformation of S4 has been predicted from its sequence, and two models are presented here. An attempt is made to correlate these two molecular models with the available physicochemical data concerning the shape, conformation, and possible RNA binding site of protein S4.

INTRODUCTION

Protein S4 plays an important role in the assembly of the 30S subunit of the ribosome and is perhaps the most extensively studied ribosomal protein. Protein S4 binds early in the in vitro assembly process (Traub and Nomura, 1969; Mizushima and Nomura, 1970; Nashimoto et al., 1971; Hochkeppel et al., 1976) and the in vivo assembly (Nierhaus and Montejo, 1973; Pichon et al., 1975). In the absence of other ribosomal proteins, protein S4 is known to bind specifically to the 16S RNA (Schaup et al., 1970, 1971; Garrett et al., 1971; Schaup and Kurland, 1972; Zimmerman et al., 1972; Held et al., 1974) and its RNA binding sites have been characterized (Schaup et al., 1970, 1971; Schaup and Kurland, 1972; Ungewickell et al., 1975; Mackie and Zimmermann, 1975).

The interaction of 16S RNA with protein S4 to form a complex provides a model by which to study the nature of RNA-protein interaction in ribosome. Whether the 16S RNA undergoes a conformational change when S4 binds is being investigated by several laboratories (Bollen et al., 1970; Nanninga et al., 1972; Schulte et al., 1974; Folkhard et al., 1975). Efforts have also been made to identify the binding site(s) of S4 and to monitor possible changes in its conformation upon complex formation (Daya-Grosjean et al., 1972, 1974; Funatsu et al., 1972; Lemieux, 1974; Amons et al., 1974; Ehresmann et al., 1975; Ungewickell et al., 1977; Ehresmann et al., 1977; Barritault and Hayes, 1977). A rigorous approach to the study of protein-RNA interaction entails the characterization of the physical properties of the complex and its components under the conditions of reconstitution. Lemieux and co-workers (1973, 1974) have reported far ultraviolet circular dichroism and fluorescence studies as a function of tem-

perature in order to gain a better understanding of the secondary structures of protein S4. An elongated shape has been inferred from several different studies (Rhode et al., 1975; Tischendorf and Stöffler, 1975; Sommer and Traut, 1975; Paradies and Franz, 1976; Österberg et al., 1977). Paradies and Franz (1976) suggest from their small-angle X-ray scattering studies that protein S4 could be an ellipsoid with dimensions of 135 × 10.5 × 5 Å or a rod with a length of 140 Å and a diameter of 10 Å. However, other small-angle X-ray scattering studies (Österberg et al., 1977) indicate that protein S4 is a flat, elongated disk with dimensions of 180 × 50 × 8 Å. Fluorescence decay studies (Brochon et al., 1976) suggest that protein S4 may exist in two conformational states. In an effort to resolve the apparent difference in small-angle X-ray scattering results, we have predicted the secondary structure (α-helix, β-sheet, β-turn, coil) from the sequence of 203 residues by the methods of Chou and Fasman (1974*a*, *b*) and Maxfield and Scheraga (1976), and carefully considered whether two conformations with such different lengths could be apparent. We present here two possible models which differ only in the prediction of one β-turn as determined by the method of Chou and Fasman (1974*a*, *b*). These models are consistent with the shapes suggested from the small-angle X-ray scattering studies of Paradies and Franz (1976) and also from Österberg et al. (1977). We have also correlated these models with the nanosecond fluorometry study (Brochon et al., 1976) and suggest a model that may closely approximate the shape of protein S4 in the conformation within the 30S subunit. In addition, we have attempted to discuss the present available evidence concerning the binding site of S4 on 16S RNA with respect to this model. Finally, we have characterized the conformation of protein S4 under the conditions of reconstitution, and we report here its near and far ultraviolet circular dichroism (CD) and absorption spectra. The amount of secondary structure was estimated from the CD spectra by a recently revised method (Chen et al., 1974).

MATERIALS AND METHODS

The preparation of *Escherichia coli* ribosomes (Allen and Wong, 1978), the separation of 30S subunits (Allen and Wong, 1978), and chromatographic separation and purification of protein S4 (Hardy et al., 1969) have been reported previously.

The ultraviolet absorption spectrum was obtained with a Cary 118 CX double-beam recording spectrophotometer (Cary Instruments, Fairfield, N.J.) equipped with derivative mode. CD measurements were made with a JASCO J-20 spectropolarimeter (Jasco Inc., Easton, Md.) equipped with a temperature-regulated cell holder. The results were expressed in terms of mean residue ellipticity, $[\theta]_{MRW}$, in units of degrees per square centimeter per decimole using the equation $[\theta]_{MRW} = \theta/dc$ (MRW); where θ is the observed ellipticity in degrees, d and c are the optical path length in decimeters and concentration in grams per cubic centimeter, respectively and MRW is mean residue weight, which is 115 for protein S4. The spectropolarimeter was routinely calibrated with *d*-10, camphorsulfonic acid according to Cassim and Yang (1969). All experiments were performed in a reconstitution buffer: TMK_{360} (10 mM Tris, 20 mM $MgCl_2$, 360 mM KCl, pH 7.6). The analysis of the secondary structure was carried out by the method of Yang and co-workers (Chen et al., 1974), using a BMD073 computer program (Biomedical Computer Programs, University of California Press, 1973) and an IBM 370/145 computer (IBM Corp., White Plains, N.Y.).

The secondary structure of protein S4 was predicted by the method of Chou and Fasman (1974 *a,b*) with the revised β-turn potentials based on the crystallographic structures of 29 proteins, kindly provided to us by Professor Fasman. The computer method of Maxfield and Scheraga (1976) was also used to predict secondary structure with an IBM 370/145 computer.

All pH measurements were made with a Radiometer model 26 pH meter equipped with combined glass electrodes (GK2302C; Radiometer Co., Copenhagen, Denmark). The pH-meter was standardized against pH 4.01 and 7.00 standard buffers. The concentration of ribosomal protein S4 was determined by the Hartree modification of the Lowry method (Hartree, 1972). All chemicals used were reagent grade. The water used in all experiments was double-deionized and distilled.

RESULTS

The secondary structure of protein S4 was predicted from its amino acid sequence (Schiltz and Reinbolt, 1975) by the method of Chou and Fasman (1974*a, b*) using the revised average conformational potentials for α-helix, $< P_\alpha >$, and for β-structure, $< P_\beta >$ (Fasman et al., 1976); together with the most recent β-turn potential, $< P_t >$, for the prediction of β-turns (Chou and Fasman, personal communication). The conformational profile of the α-helical and β-structure regions are shown in Fig. 1, and the predicted α-helix and β-structure regions and their corresponding average conformational potentials are summarized in Table I. There are 45% α-helix, 18% β-structure, 19% β-turn, and 18% coil residues in this model. Although the present prediction of α-helix and β-structure is very similar to that proposed by Schiltz and Reinbolt (1975), an interesting feature of our model is the location of the chain reversals denoted by the β-turn tetrapeptides. Utilization of the revised β-turn potentials which were derived from 29 proteins resulted in prediction of the β-turns summarized in

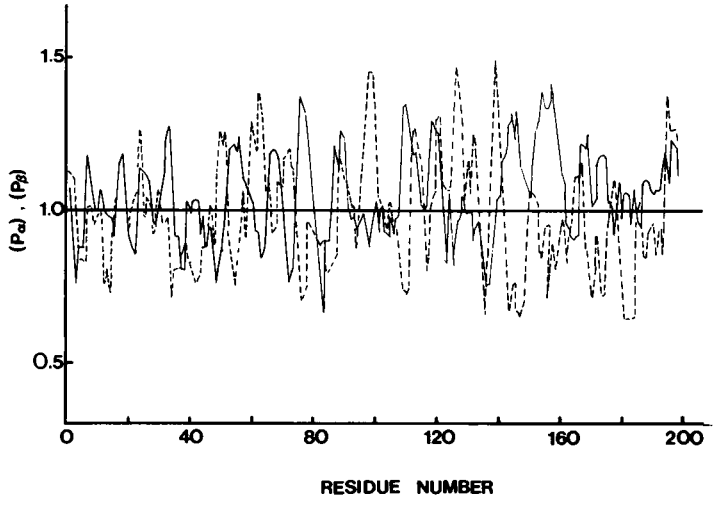

FIGURE 1 The conformational profile of protein S4. The helical potential (——), $< P_\alpha >$, at residue i is the average of the $< P_\alpha >$ values of residues i to $i + 3$. The β-structure potential (- - - -), $< P_\beta >$, at residue i is the average of $< P_\beta >$ values of residues i to $i + 3$.

TABLE I

PREDICTED α-HELIX AND β-STRUCTURE IN RIBOSOMAL PROTEIN S4*

	Average conformation potentials	
	$<P_\alpha>$	$<P_\beta>$
Predicted α-helix regions		
25–36	1.11	0.99
52–60	1.15	1.08
76–82	1.13	0.82
89–93	1.23	1.16
107–121	1.16	0.99
141–163	1.22	0.92
168–181	1.11	0.90
189–201	1.09	1.05
Predicted β-structure		
61–75	0.96	1.15
94–106	0.92	1.09
122–135	0.97	1.22

*The predicted α-helix and β-structure regions are based on the method of Chou and Fasman (1974*b*) using the revised conformation parameters (Fasman et al., 1976).

Table II. All the β-turn tetrapeptides have their average conformational potentials $<P_t>$ greater than those for α-helix ($<P_\alpha>$) and β-structure ($<P_\beta>$), i.e. $<P_\alpha> < <P_t> > <P_\beta>$, with the exception of the tetrapeptide 71–74 in which the average conformational potential for β-turn, $<P_t>$, is smaller than that for the β-structure, $<P_\beta>$, i.e. $<P_t> < <P_\beta>$. However, this turn has a β-turn probability, p_t, equal to 1.49×10^{-4} which is greater than the 1.0×10^{-4} cutoff point. The probabilities of the 10 β-turn tetrapeptides, p_t, are also shown in Table II.

We have also used the predictive algorithm of Maxfield and Scheraga (1976) to esti-

TABLE II

CONFORMATION PREDICTION OF β-TURN REGIONS FOR RIBOSOMAL PROTEIN S4: $<P_\alpha>$, $<P_\beta>$, AND $<P_t>$ VALUES*

β-Turn	Tetrapeptide	$p_t \times 10^4$‡	$<P_t>$	$<P_\alpha>$	$<P_\beta>$
5–8	Gly-Pro-Lys-Leu	1.65	1.17	0.88	0.84
15–18	Gly-Thr-Asp-Leu	1.38	1.14	0.91	0.95
36–38	Ala-Pro-Gly-Gln	3.60	1.18	0.92	0.81
48–51	Ser-Asp-Tyr-Gly	2.29	1.40	0.76	0.88
71–74	Phe-Arg-Asn-Tyr	1.49	1.06	0.87	1.17
84–87	Asn-Thr-Gly-Glu	2.11	1.21	0.97	0.80
136–139	Asp-Pro-Asn-Ser	9.60	1.49	0.68	0.68
164–167	Lys-Pro-Thr-Trp	1.90	1.11	0.91	0.96
182–185	Lys-Pro-Glu-Arg	1.16	1.06	1.06	0.65
185–188	Arg-Ser-Asp-Leu	1.11	1.11	0.99	0.88

*$<P_t>$, $<P_\alpha>$, and $<P_\beta>$ are the average conformational potentials for the tetrapeptide i to i + 3 to be in the β-turn, α-helix, and β-structure, respectively.

‡p_t is the product of the β-turn potentials for the tetrapeptide of residues i to i + 3.

TABLE III

CONFORMATION PREDICTION OF RIBOSOMAL PROTEIN S4 USING THE METHODS OF MAXFIELD AND SCHERAGA[c] AND CHOU AND FASMAN[d]

```
                    1                   2                   3                   4
a 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0
b A R Y L G P K L K L S R R E G T D L F L K S G V R A I D T K C K I E Q A P G Q H
c 1 1 1 6 1 6 2 1 1 1 1 6 1 6 1 1 2 2 1 1 1 2 4 1 1 1 6 2 2 2 1 2 2 2 1 1 2 2 1 2
d        [T U R N]H H H H H H[T U R N]            H H H H H H H H H H H[T U R N]

                    5                   6                   7                   8
a 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0
b G A R K P R I S D Y G V Q L R E K Q K V R R I Y G V L E R Q F R N Y Y K E A A R
c 2 6 6 1 1 1 1 6 1 6 4 1 1 1 2 2 2 2 2 2 1 1 1 1 4 1 1 2 2 2 1 1 3 6 2 2 2 2 2 2
d              [T U R N]H H H H H H H H H B B B B B B B B B B[T U R N]  H H H H H

                    9                   10                  11                  12
a 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0
b L K G N T G E N L A L L Q G R L D N V V Y R N G F G A T R A E A R Q I V S H K A
c 2 6 4 3 6 5 2 2 2 2 2 2 2 1 1 1 4 1 2 1 1 1 1 1 3 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2
d H H  [T U R N]  H H H H H B B B B B B B B B B B B H H H H H H H H H H H H H H H

                    13                  14                  15                  16
a 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0
b I M V N G R V V N I A S Y Q V D P N S V V I R E K A K K E S R V K A A L E L A E
c 2 1 1 4 4 1 1 1 1 1 1 1 1 1 1 1 6 6 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
d H B B B B B B B B B B B B B B[T U R N]  H H H H H H H H H H H H H H H H H H H H

                    17                  18                  19                  20
a 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0
b Q R Q K P T W L E V N A G K M E G T E K R K P E R S D L S A D I N E H L I V E L
c 2 2 6 1 1 6 1 2 2 2 2 2 5 2 1 2 2 2 2 1 1 1 6 6 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
d H H H[T U R N]H H H H H H H H H H H H H [ T U R N S ]   H H H H H H H H H H H H

a 1 2 3
b Y S K
c 2 6 3
d H
```

[a]Residue number.

[b]Amino acid sequence according to the following: Ala, A; Arg, R; Asn, N; Asp, D; Cys, C; Gln, Q; Glu, E; Gly, G; His, H; Ile, I; Leu, L; Lys, K; Met, M; Phe, F; Pro, P; Ser, S; Thr, T; Trp, W; Tyr, Y; Val, V.

[c]Results of Maxfield and Sheraga prediction (1976) according to the following code: 1, β-structure; 2, α-helix; 3, ζ-R; 4, α-L; 5, ζ-L; 6, α-R.

[d]Results of prediction using the method of Chou and Fasman (1974*a*, *b*). Helix is denoted by H, β-structure by B, and turns by [TURN].

mate helical and β-structure regions, and the results of this computer analysis are summarized in Table III, together with the results of the Chou and Fasman (1974*a*, *b*) prediction (Fasman et al., 1976). The prediction of the long sections of α-helix and β-structure are in good agreement for both methods. The predicted secondary structure of protein S4 derived from both methods is shown in Fig. 2. This model includes the β-turn 71–74.

Although this prediction of secondary structure results in only a schematic model, it does enable us to speculate on some of the tertiary structural features of S4, its shape, and its RNA-binding sites. Inclusion of chain reversals by the β-turn tetrapeptides allows us to predict possible tertiary interactions between α-helices as well as β-pleated sheets. For example, it is likely that the helix 76–82 is involved in a tertiary

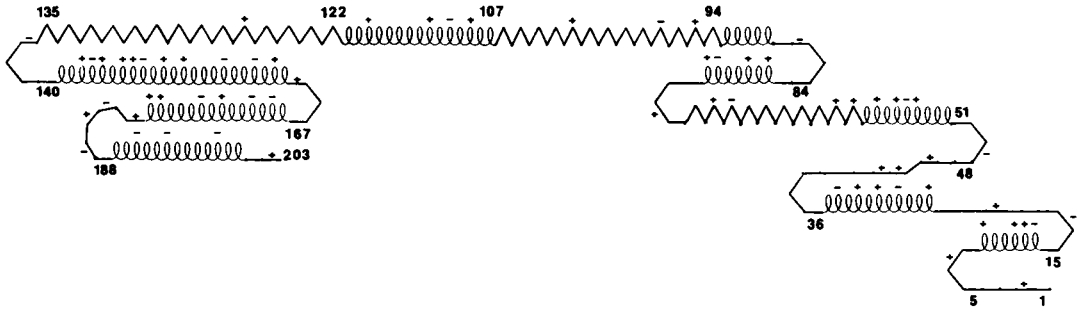

FIGURE 2 Schematic representation of the prediction of secondary structure of protein S4. The β-turn of 71-74 is included in this model. There are 45% α-helix ( ), 23% β-structure ( ), 16% β-turn ( ), 16% random-coil residues (——). The positions of the charged residues are indicated by + or −.

interaction with helix 89–93 because the direction of the chain is reversed by a β-turn strategically located between the two helices. A similar prediction can be made for the helices 141–163 and 168–181.

One unusual feature of this model is that long segments of the sequence are involved in α-helical and β-structure regions. If we assume that the rise of the α-helix along its helical axis per repeating unit is 1.5 Å, whereas that of the β-structure is 3.5 Å (Dickerson and Geis, 1969), the minimum lengths of these long segments can be calculated. For example, the region 87–136 is 135 Å long. The entire length of this predicted model of S4 is 170–180 Å. These estimates are consistent with an elongated structure observed by immunoelectron microscopy (Tischendorf and Stöffler, 1975) to have a length of 160–200 Å and that by small-angle X-ray scattering (Österberg et al., 1977) to have dimensions of 180 × 50 × 8 Å. These dimensions indicate that the thickness of protein S4 is limited to that of only one helical cylinder, and suggest that this schematic representation may truly reflect that the shape of protein S4 is a flattened dumbbell 180-Å long with two domains for the carboxy- and amino-terminal regions, each approximately 50-Å wide and 8-Å thick. This flattened dumbbell-shaped model of S4 could correspond to a recent postulate (Changchien and Craven, 1976) that the amino-terminal end of S4 is located in the "head" and the carboxy-terminal end is located in the "middle body" of the 30S subunit model proposed by Lake (1976).

Examination of this model reveals the existence of hydrophobic regions, in addition to domains where the basic residues are clustered. Most of the predicted β-structure contains few charged residues. On the other hand, many of the hydrophilic residues can be found in the predicted helices. It is possible to correlate the proposed RNA binding sites of protein S4 with our schematic model. The carboxy-terminal residues of S4 have been suggested to play a role in its binding to 16S RNA (Funatsu et al., 1972; Daya-Grosjean, 1972, 1974). Three S4 mutants with as many as 20 carboxy-terminal residues missing exhibited reduced binding (Funatsu et al., 1972). Similarly, carboxypeptidase treatment of wild-type S4 eliminated its binding capacity (Daya-Grosjean, 1972, 1974). Consistent with these findings are the ultraviolet irradiation

studies (Ehresmann et al., 1975) which suggest that regions 181–185 as well as 144–151 and 56–59 are involved in the RNA binding site. One can also speculate that the proposed binding site segment 181–185 may be in close proximity to the 144–151 binding site because the direction of the chain is reversed by a β-turn. Reductive methylation of the lysine residues of the complex and protein alone have provided additional information about the binding site of S4 (Amons et al., 1974). Residues lys 147, lys 30, and lys 119 cannot be modified in the 16S RNA-S4 protein complex. Because these lysines can be methylated in unbound protein S4, it has been suggested that these residues are located in the RNA binding region. However, it is still possible that these residues may not all actually mask the binding sites, but rather that some of them might be shielded from the solvent if a conformational change were induced upon binding to RNA. One common result of these binding studies is that the carboxy-terminal end of S4, involving almost one-third of the sequence, contains one or more RNA binding sites. Our model suggests that the secondary structure in this region is predominantly α-helical. This region also contains many charged residues, including clusters of basic residues which could be involved in electrostatic interactions with the polyanionic nature of the RNA. A recent hypothesis by Heléne (1977) suggests that guanine bases in a double-stranded nucleic acid or both guanine and cytosine bases in a single-stranded nucleic acid could specifically interact with an arginine-glutamic acid sequence of protein, especially if it has a basic amino acid neighbor. Several arginine-glutamic acid pairs are located in the sequence of protein S4: 13–14, 55–56, 68–69, 143–144, and 184–185. Pairs 55–56, 143–144, and 184–185 are also in regions that have been implicated to be involved in the RNA binding sites. Our model can also be used to correlate other studies of protein S4. For example, Changchien and Craven (1976) have determined that trypsin digestion of 16S RNA-protein S4 complex results in the loss of the amino-terminus of protein S4 up to residues 43–46, which does not affect the ability of protein S4 to bind to the 16S RNA but does affect the successful assembly of some of the 30S subunit proteins. Our model indicates that there are regions of random coil of protein S4 located in the amino-terminal domain. The basic residues at 42 and 43, which are the probable sites for trypsin digestion, are located in the middle of one random-coiled region, 40–47, These regions of random coil in the amino-terminal region may have interesting implications with respect to several protein-protein interactions, and also with respect to the evidence that suggests that this region may not be involved significantly in protein-RNA interactions (Changchein and Craven, 1976).

Exclusion of the β-turn tetrapeptide 71–74 results in a model represented in Fig. 3. The predominant unique feature of this model is the interaction of the α-helical region 52–60 with the α-helical region 107–121 and also the β-structure 61–75 with that of 94–106 to form a hydrophobic antiparallel β-ribbon. These interactions could also place Phe_{71} close to Tyr_{101}, Tyr_{50} to Trp_{167}, and Phe_{19} to Tyr_{201}. The length of this model is approximately 135 Å which corroborates the small-angle X-ray scattering studies of Paradies and Franz (1976). These workers suggest that protein S4 is either a rod with a length of 140 Å and a diameter of 10 Å or an ellipsoid with dimensions of

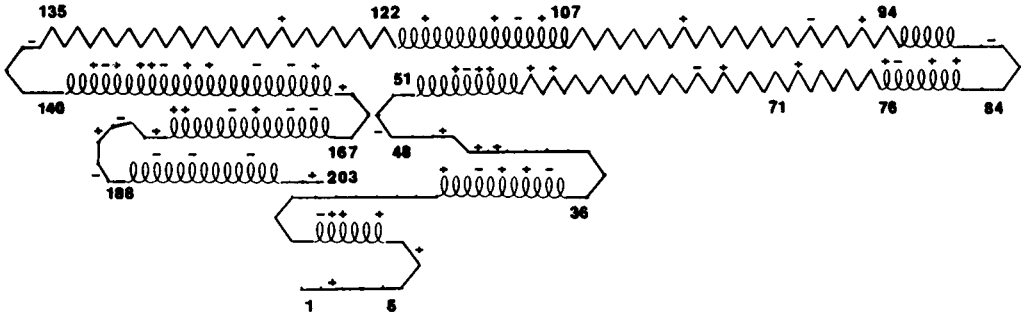

FIGURE 3 Schematic representation of the prediction of secondary structure of protein S4. The β-turn 71-74 is not included in this model. Symbols used are the same as in Fig. 2.

135 × 10.5 × 5 Å. However, it is difficult to envision how the smaller dimensions of our model could be as small at 10 Å.

A nanosecond-pulse fluorometric study of protein S4 (Brochon et al., 1976) indicated that the molecule may exist in two conformations in solution. Ribosomal protein S4 contains only one tryptophan at 167 which appears to undergo dynamic quenching in one conformation, but does not in the other. Presumably, the conformation characterized by dynamic quenching of the tryptophan fluorescence corres-

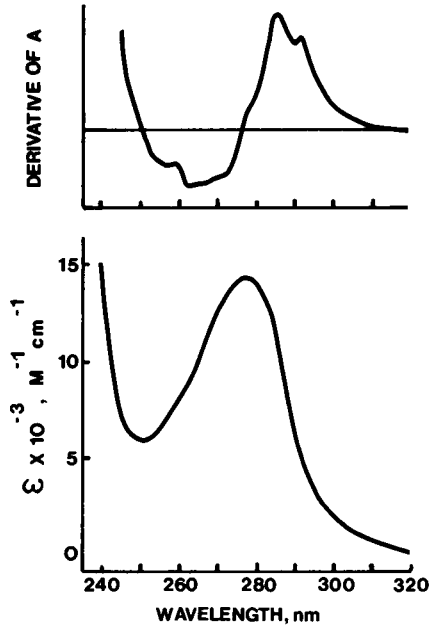

FIGURE 4 Ultraviolet absorption spectrum of protein S4 in TMK_{360} buffer at ambient temperature (lower panel). The corresponding first derivative spectrum is shown on the upper panel. The concentration was 0.25 mg/ml and the path length was 1.0 cm.

ponds to one in which the Trp_{167} is buried. Our second model, which is 135-Å long, is consistent with these results because the tryptophan could interact with groups in the amino-terminal domain. On the other hand, the conformation in which the tryptophan fluorescence does not undergo quenching could correspond to our flattened dumbbell-shaped model in which the tryptophan does not appear to be buried. Furthermore, this conformation appears to be favored at 42°C, the optimum condition for reconstitution.

The ultraviolet absorption spectrum of protein S4 in reconstitution buffer is shown in Fig. 4. The first derivative of this spectrum accentuates the small shoulders of the absorption curve; thus, the assignment of the aromatic residues can be made easily. Tryptophan, found only at position 167 in protein S4, gives rise to the absorption maxima at 291.5 and 284 nm. Tyrosine absorption can also be assigned to the 284-nm maximum in addition to the shoulder at 278 nm appearing in the derivative curve. The characteristic phenylalanine fine structure appears in the region below 275 nm.

Fig. 5 shows the near ultraviolet CD spectrum of protein S4 in reconstitution buffer at 25°C and is characterized by a negative band with minima at 274, 282, and a shoulder at 290 nm. Overlap with the negative far ultraviolet bands undoubtedly diminishes the peak at 255 nm. Fig. 6 shows the far ultraviolet CD spectra of protein S4 in reconstitution buffer at 25° and 37°C. Both spectra show double troughs at 222 and 208 nm which are typical of proteins containing significant amounts of α-helical contents. The magnitude of these troughs is smaller for the spectrum at 37°C than that at 25°C. Estimation of the amount of secondary structure of protein S4 from its far ultraviolet CD spectrum using the method of Yang and co-workers (Chen et al., 1974) yielded 32% α-helix and 14% β-structure at 25°C. At 37°C the estimated α-helix is 21% and β-structure is 17%.

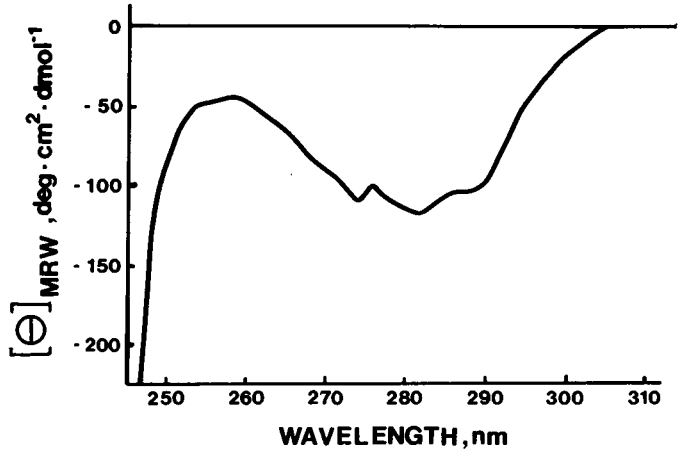

FIGURE 5 Near ultraviolet CD spectrum of protein S4 in TMK_{360} buffer at 25°C. The concentration was 0.35 mg/ml and the path length was 1.0 cm.

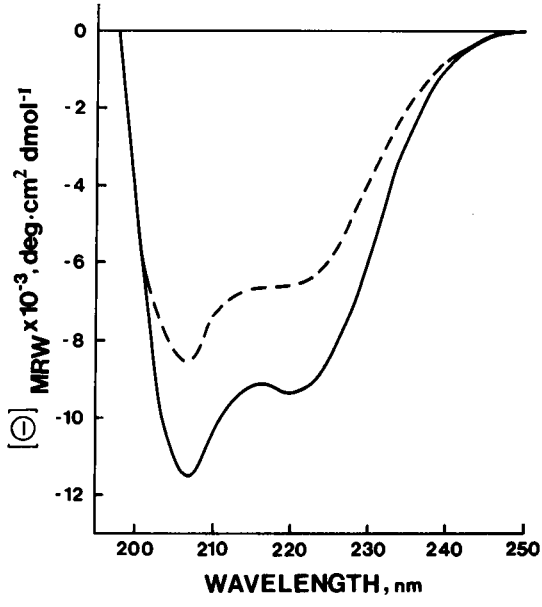

FIGURE 6 Far ultraviolet CD spectrum of protein S4 in TMK_{360}. The solid curve, at 25°C, and broken curve, at 37°C. The concentration was 0.35 mg/ml and the path length was 0.05 cm.

DISCUSSION

Near ultraviolet CD arises from cystine and aromatic amino acid residues located in a dissymmetric environment. Thus, near ultraviolet CD reflects the unique local conformation of these chromophores, i.e., the tertiary structure of the protein (for review, see Strickland, 1974). To our knowledge this is the first report of a near ultraviolet CD spectrum of a ribosomal protein. Ribosomal proteins typically contain few aromatic amino acids (Kaltschmidt et al., 1970); however, protein S4 contains one tryptophan, eight tyrosines, and four phenylalanines (Schiltz and Reinbolt, 1975). Because it contains only one cysteine residue, the possible contribution to the near ultraviolet CD by a disulfide bridge can be ruled out. These near ultraviolet CD bands arise from interaction of the aromatic amino acids with their surrounding environment, and provide evidence that protein S4 does possess unique tertiary structure, because extremely small or no near ultraviolet CD bands will be observed if the chromophores do not exist in an asymmetric or rigid environment such as the random-coil state of the protein (for review, see Adler et al., 1973). Many proteins possessing well-defined tertiary structures have characteristic near ultraviolet CD spectra which vanish when the native tertiary structures are disturbed or rendered structureless, such as by denaturation in 6 M guanidine hydrochloride (e.g. Wong and Tanford, 1973; Wong and Hamlin, 1975).

Tentative assignments of these CD bands can be made. Because the sole tryptophan residue in protein S4 absorbs at 291.5 and also at 284 nm (Fig. 4), it is likely that the CD bands at 289 and 282 nm can be attributed to the ^1L_b vibronic transition of trypto-

phan. The band at 274 nm probably arises from its ^1L_a transition (Strickland et al., 1969; Strickland, 1974). Ionized tyrosine residues may also give rise to the negative bands at 282 nm, as well as the shoulder at 289 nm. These bands are characteristic of insulin (Menendez and Herskovitz, 1970) and leutenizing hormone (Bewley et al., 1972), proteins that do not contain tryptophan. The 0 + 800 cm^{-1} vibronic transition has been associated with a 282-nm band and the 0–0 cm^{-1} transition with a band at 289 nm (Strickland, 1974). The peak at 255 nm may arise from tyrosine also because poly-L-tyrosine exhibits a corresponding band at 250 nm (Beychok and Fasman, 1964). Overlap with the large negative far ultraviolet band arising from the secondary structure of S4 probably accounts for the change in sign of this band. Inasmuch as phenylalanine usually exhibits very distinct fine structure at 255–270 nm, it appears that it may be masked by the large background CD band in that region. It is interesting that protein S4 exhibits a near ultraviolet CD spectrum similar to that of horse heart ferricytochrome *c* (Myer, 1968) which contains only one tryptophan, four tyrosines, four phenylalanines and no disulfides. Our assignments of near ultraviolet CD bands for S4 are in agreement with those of ferricytochrome *c* which have been rigorously determined, a fact that lends credence to our assignments.

The far ultraviolet CD spectrum indicates that protein S4 possesses a significant amount of ordered structure. Lemiuex and co-workers (1974) reported 31% α-helix and 19% β-structure for protein S4 in their CD studies. Although we used the revised method for estimation of secondary structure (Chen et al., 1974), this probably does not account for the smaller amount of β-structure we predicted. A qualitative inspection of the two CD spectra indicates that Lemieux reports two equivalent troughs at 208 and 222 nm, whereas our study shows that the trough at 222 nm is approximately 80% of the magnitude at 208 nm. We cannot explain the difference at this time.

The present study does, however, substantiate the report by Lemieux et al. (1974) and Brochon et al (1976) that the conformation of S4 is highly temperature dependent. The amount of α-helix decreases from 32 to 21%, whereas the amount of β-structure slightly increases from 14 to 17% at 37°C as compared to 25°C. However, we also observed the tendency of protein S4 to precipitate at this temperature, and these differences in secondary structural determination at 37°C as compared with 25°C may arise from aggregation. The nanosecond fluorescence study (Brochon et al., 1976) suggests that S4 may exist in more than one temperature-dependent conformation. Furthermore, one conformation is favored in the optimal conditions of reconstitution. These results suggest that protein S4 may be responsible, in part, for the temperature-dependent rate-limiting step in the assembly of the ribosomal subparticle (Traub and Nomura, 1969; Held and Nomura, 1973). Garrett and co-workers (Schulte et al., 1974) reported that binding of protein S4 to 16S RNA is temperature dependent and indicated that a heat-induced conformational change of the RNA is responsible for this change in binding behavior. Whether or not a temperature-dependent conformational change in protein S4 facilitates this structural change in 16S RNA remains to be ascertained. The question of whether 16S RNA induces a conformational change in protein S4 or stabilizes its structure is also unanswered. However, it is possible that

the conformation of free protein S4 at 25°C is not the same as that in the ribosome. Other studies in our laboratory suggest that the conformations of ribosomal proteins free in solution are not identical to those in the ribosome (Allen and Wong, 1978).

The 16S RNA-S4 complex formation is highly dependent upon Mg (II) concentration (Schulte et al., 1974). Although this has been attributed to a Mg(II)-dependent conformational change in the RNA, our proposed model of S4, in correlation with the binding studies, suggests that Mg(II) may mediate the electrostatic interaction between the carboxylic residues and the polyanionic phosphate sugar backbone of the RNA. This would be feasible for the α-helix 189–201 which contains four carboxylic residues and no basic residues.

Although the complete three-dimensional structure of protein S4 cannot be elucidated until the X-ray diffraction studies are complete, it is possible to gain a better understanding of its shape and conformation by a variety of physicochemical techniques. The spectroscopic studies presented here indicate that protein S4 free in solution does have conformational integrity and, in fact, extensive secondary structure and unique tertiary structure. The proposed conformations for S4 have provided a model by which the available structural and binding data can be correlated, and further studies designed to test their validity.

It has become increasingly apparent that many of the ribosomal proteins are not globular, but rather elongated. Furthermore, it appears that some proteins may even be flattened, as in the case of protein S4. It is important to revise our concept of the shapes of ribosomal proteins especially in light of the large number of protein-protein interactions that have been implicated from structural studies. For example, cross-linking studies show that protein S4 interacts with eight other proteins. These include S13 (Sommer and Traut, 1974), S6, S8, S9, S12 (Sommer and Traut, 1975), S3, S5, and S17 (Sommer and Traut, 1976). It is difficult to envision how a globular protein of 203 residues can interact with eight proteins, in addition to the 16S RNA. The elongated, flattened shape provides a greater surface area to allow for such interaction.

Dr. Charles Kurland's encouragement and support are gratefully acknowledged. The new β-turn potentials used here were kindly provided to us by Professor G. D. Fasman before their publication. We are grateful to Professor Harold A. Scheraga for sending us the computer program used for prediction, and to Dr. Gary Clark of the Computer Center at the University of Kansas Medical Center for his help in adapting the computer programs.

This work has been supported by grant GM 22962, and in part by grant HL 18905, and a Research Career Development Award (GM 70628) from the National Institutes of Health. One of us (K.P.W.) was an EMBO fellow of the European Molecular Biology Organization in Professor Charles Kurland's Laboratory in Uppsala, Sweden.

Dr. Wong was also a Research Career Development Awardee of the National Institute of General Medical Sciences (GM 70628).

Received for publication 14 December 1977 and in revised form 21 April 1978.

REFERENCES

ADLER, A. J., N. J. GREENFIELD, and G. D. FASMAN. 1973. Circular dichroism and optical rotatory dispersion of proteins and polypeptides. *Methods Enzymol.* **27**:675.

ALLEN, S. H., and K-P. WONG. 1978. The conformation of ribosomal proteins free in solution and bound to ribosomal RNA. *Biochemistry*. In press.

AMONS, R., W. MÖLLER, E. SCHILTZ, and J. REINBOLT. 1974. Studies on the binding sites of protein S4 to 16 S RNA in *Escherichia coli* ribosomes. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* **41**:135.

BARRITAULT, D., and D. H. HAYES. 1977. Localization of part of the binding sites of 30S ribosomal proteins S4 and S20 in a small uninterrupted fragment of 16S RNA. *Biochimie. (Paris)*. **59**:463.

BEWLEY, T. A., M. R. SAIRAM, and C. H. L. 1972. Circular dichroism of ovine interstitial cell stimulating hormone and its subunits. *Biochemistry*. **11**:932.

BEYCHOK, S., and G. D. FASMAN. 1964. Circular dichroism of poly-L-tyrosine. *Biochemistry*. **3**:1675.

BOLLEN, A., A. HERZOG, A. FAVRE, J. THIBAULT, and F. GROS. 1970. Fluorescence studies on the 30S ribosome assembly process. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* **11**:49.

BROCHON, J-C., P. WAHL, P. VACHETTE, and M. P. DAUNE. 1976. Nanosecond-pulse fluorometry study of S4 ribosomal protein. *Eur. J. Biochem.* **65**:35.

CASSIM, J. Y., and J. T. YANG. 1969. A computerized calibration of the circular dichrometer. *Biochemistry*. **8**:1947.

CHANGCHIEN, L-M., and G. R. Craven. 1976. The function of the N-terminal region of ribosomal protein S4. *J. Mol. Biol.* **108**:381.

CHEN, Y-H., J. T. YANG, and K. H. CHAU. 1974. Determination of the helix and β form of proteins in aqueous solution by circular dichroism. *Biochemistry*. **13**:3350.

CHOU, P. Y., and G. D. FASMAN. 1974*a*. Conformational parameters for amino acids in helical, β-sheet and random coil regions calculated from proteins. *Biochemistry*. **13**:211.

CHOU, P. Y., and G. D. FASMAN. 1974*b*. Prediction of protein conformation. *Biochemistry*. **13**:222.

DAYA-GROSJEAN, L., R. A. GARRETT, O. PONGS, G. STÖFFLER, and H. G. WITTMANN. 1972. Properties of the interaction of ribosomal protein S4 and 16 S RNA in *Escherichia coli*. Revertants from streptomycin dependence to independence. *Mol. & Gen. Genet.* **119**:277.

DAYA-GROSJEAN, L., J. REINBOLT, O. PONGS, and R. A. GARRETT. 1974. A study of the regions of ribosomal proteins S4, S8, S15 and S20 that interact with 16S RNA of *Escherichia coli*. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* **44**:253.

DICKERSON, R. E., and I. GEIS. 1969. The structure and action of proteins. Harper & Row, Publishers. New York. 120.

EHRESMANN, C., J. REINBOLT, and J. P. EBEL. 1975. Studies of the binding of *E. coli* ribosomal protein S4 to 16S r-RNA after UV irradiation of the S4-16S RNA complex. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* **58**:106.

EHRESMANN, C., P. STIEGLER, P. CARBON, E. UNGEWICKELL and R. A. GARRETT. 1977. A topographical study of the 5'-region of 16S RNA of *Escherichia coli* in the presence and absence of protein S4. *FEBS Letters*. **81**:188.

FASMAN, G. D., P. Y. CHOU, and A. J. ADLER. 1976. Prediction of the conformation of the histones. *Biophys. J.* **16**:1201.

FOLKHARD, W., I. PILZ, O. KRATKY, R. GARRETT, and G. STÖFFLER. 1975. Small-angle X-ray studies on the structure of 16S ribosomal RNA and of a complex of ribosomal RNA from *Escherichia coli*. *Eur. J. Biochem.* **59**:63.

FUNATSU, G., W. PULS, E. SCHILTZ, J. REINBOLT, and H. G. WITTMANN. 1972. Ribosomal proteins. XXXI. Comparative studies on altered proteins of S4 of six *Escherichia coli* revertants from streptomycin dependence. *Mol. & Gen. Genet.* **115**:131.

GARRETT, R. A., K. H. RAK, L. DAYA, and G. STÖFFLER. 1971. Ribosomal proteins. XXIX. Specific protein binding sites on 16S RNA of *Escherichia coli*. *Mol. Gen. Genet.* **114**:112.

GERARD, D., J. P. SCHREIBER, and G. LAUSTRAIT. 1975. Fluorescence study of S4 and S7 ribosomal proteins complexed to 16 S RNA. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* **50**:250.

HARDY, S. J. S., C. G. KURLAND, P. VOYNOW, and G. MORA. 1969. The ribosomal proteins of *Escherichia coli*. I. Purification of the 30 S ribosomal proteins. *Biochemistry*. **8**:2897.

HARTREE, E. F. 1972. Determination of protein: a modification of the Lowry method that gives a linear photometric response. *Anal. Biochem.* **48**:422.

HELD, W. A., S. MIZUSHIMA, and M. NOMURA. 1973. Reconstitution of *Escherichia coli* 30S ribosomal subunits from purified molecular components. J. Biol. Chem. **248**:5720.

HELD, W. A., and M. NOMURA. 1973. Rate-determining step in the reconstitution of *Escherichia coli* 30S ribosomal subunits. *Biochemistry*. **12**:3273.

HELÈNE, C. 1977. Specific recognition of guanine bases in protein-nucleic acid complexes. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* **74**:10.

HOCHKEPPEL, H-K., E. SPICER, and G. R. CRAVEN. 1976. A method of preparing 16 S RNA possessing previously unobserved 30 S ribosomal protein binding sites. *J. Mol. Biol.* **101**:155.

KALTSCHMIDT, E., M. DZIONARA, and H. G. WITTMANN. 1970. Ribosomal proteins. XV. Amino acid compositions of isolated proteins from 30 S and 50 S subunits of *Escherichia coli*. *Mol. & Gen. Genet.* **109**:292.

LAKE, J. A. 1976. Ribosome structure determination by electron microscopy of *Escherichia coli* subunits, large subunits and monomeric ribosomes. *J. Mol. Biol.* **105**:131.

LEMIEUX, G. 1974. Chemical modification of *Escherichia coli* ribosomal proteins S4, S7, and S8 in regard to their capacity for binding 16 S RNA. *Can. J. Biochem.* **52**:1038.

LEMIEUX, G., and D. GERARD. 1973. Fluorescence studies of *Escherichia coli* ribosomal proteins S4 and S7 in regard to reconstitution conditions. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* **37**:234.

LEMIEUX, G., J-F. LEFEVRE, and M. DAUNE. 1974. Effect of reconstitution conditions on the structure of *Escherichia coli* 30 S ribosomal subunit components. *Eur. J. Biochem.* **49**:185.

MACKIE, G. A., and R. A. ZIMMERMANN. 1975. Characterization of fragments of 16 S ribonucleic acid protected against pancreatic ribonuclease digestion by ribosomal protein S4. *J. Biol. Chem.* **250**:4100.

MAXFIELD, F. R., and H. A. SCHERAGA. 1976. Status of empirical methods for the prediction of protein backbone topography. *Biochemistry.* **15**:5138.

MENENDEZ, C. J., and T. T. HERSKOVITZ. 1970. Optical rotatory dispersion and circular dichroism studies on insulin and its trypsin-modified derivatives. *Arch. Biochem. Biophys.* **140**:286.

MIZUSHIMA, S., and M. NOMURA. 1970. Assembly mapping of 30 S ribosomal proteins from *Escherichia coli*. *Nature (Lond.).* **226**:1214.

MYER, Y. P. 1968. Conformation of cytochromes. III. Effect of urea, temperature, extrinsic ligands, and pH variation on the conformation of horse heart ferricytochrome *c*. *Biochemistry.* **7**:765.

NANNINGA, N., R. A. GARRETT, G. STÖFFLER, and G. KOTZ. 1972. Ribosomal proteins. XXXVIII. Electron microscopy of ribosomal protein S4-16 S RNA complexes of *Escherichia coli*. *Mol. & Gen. Genet.* **119**:175.

NASHIMOTO, H., W. HELD, E. KALTSCHMIDT, and M. NOMURA. 1971. Structure and function of bacterial ribosomes. XII. Accumulation of 21S particles by some cold-sensitive mutants of *Escherichia coli*. *J. Mol. Biol.* **62**:121.

NIERHAUS, K. H., and J. MONTEJO. 1973. A protein involved in the peptidyltransferase activity of *Escherichia coli* ribosomes. *Proc. n atl. Acad. Sci. U.S.A.* **70**:1931.

ÖSTERBERG, R., B. SJÖBERG, R. A. GARRETT, and J. LITTLECHILD. 1977. Small-angle X-ray scattering study of the 16 S RNA binding protein S4 from *Escherichia coli* ribosomes. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* **73**:25.

PARADIES, H. H., and A. FRANZ. 1976. Geometry of the protein S4 from *Escherichia coli* ribosomes. *Eur. J. Biochem.* **67**:23.

PICHON, J., J. MARVALDI, and G. MARCHIS-MOUREN. 1975. The *in vivo* order of protein addition in the course of *Escherichia coli* 30 S and 50 S biogenesis. *J. Mol. Biol.* **96**:125.

RHODE, M. F., S. O'BRIEN, S. COOPER, and K. C. AUNE. 1975. Physical properties of some ribosomal proteins in solution and evidence for molecular interactions between isolated ribosomal proteins. *Biochemistry.* **29**:1079.

SCHAUP, H. W., M. GREEN, and C. G. KURLAND. 1970. Molecular interactions or ribosomal components. I. Identification of RNA binding sites for individual 30 S ribosomal proteins. *Mol. & Gen. Genet.* **109**:193.

SCHAUP, H. W., M. GREEN, and C. G. KURLAND. 1971. Molecular interactions of ribosomal components. II. Site-specific complex formation between 30 S proteins and ribosomal RNA. *Mol. & Gen. Genet.* **112**:1.

SCHAUP, H. W., and C. G. KURLAND. 1972. Molecular interactions of ribosomal components. III. Isolation of the RNA binding site for a ribosomal protein. *Mol. & Gen. Genet.* **114**:350.

SCHULTZ, E., and J. REINBOLT. 1975. Determination of the complete amino acid sequence of protein S4 from *Escherichia coli* ribosomes. *Eur. J. Biochem.* **56**:467.

SCHULTE, C., C. A. MORRISON, and R. A. GARRETT. 1974. Protein-ribonucleic acid interactions in *Escherichia coli* ribosomes: Solutions studies on S4-16 S ribonucleic acid and L24-23 S ribonucleic acid binding. *Biochemistry.* **13**:1032.

SOMMER, A., and R. R. TRAUT. 1974. Diagonal polyacrylamide-dodecyl sulfate gel electrophoresis for the identification of ribosomal proteins crosslinked with methyl-4-mercaptobutyrimidate. *Proc. Natl. Acad. Sci. U.S.A.* **71**:3946.

SOMMER, A., and R. R. TRAUT. 1975. Identification by diagonal gel electrophoresis of nine neighboring protein pairs in the *Escherichia coli* 30 S ribosome crosslinked with methyl-4-mercaptobutyrimidate. *J. Mol. Biol.* **97**:471.

SOMMER, A., and R. R. TRAUT. 1976. Identification of neighboring protein pairs in the *Escherichia coli* 30 S ribosomal subunit by crosslinking with methyl-4-mercaptobutyrimidate. *J. Mol. Biol.* **106**:995.

STRICKLAND, E. H. 1974. Aromatic contributions to circular dichroism spectra of proteins. *CRC Crit. Rev. Biochem.* **3**:113.

STRICKLAND, E. H., J. HORIWITZ, and C. BILLUPS. 1969. Fine structure in the near-ultraviolet circular dichroism and absorption spectra of tryptophan derivatives and chymotrysinogen A at 77°K. *Biochemistry.* **8**:3205.

TISCHENDORF, G. W., and G. STÖFFLER. 1975. Localization of *Escherichia coli* ribosomal protein S4 on the surface of the 30 S ribosomal subunit by immunoelectron microscopy. I. Distribution of antibody-binding sites as obtained with immunoglobulins and monovalent antibody fragments from various S4-specific antisera. *Mol. Gen. Genet.* **142**:193.

TRAUB, P., and M. NOMURA. 1969. Structure and function of *Escherichia coli* ribosomes. VI. Mechanism of assembly of 30 S ribosomes studied *in vitro*. *J. Mol. Biol.* **49**:391.

UNGEWICKELL, E., C. EHRESMANN, P. STIEGLER, and R. GARRETT. 1975. Evidence for tertiary structural RNA-RNA interactions within the protein S4 binding site at the 5'-end of 16 S ribosomal RNA of *Escherichia coli*. *Nucleic Acids Res.* **2**:1867.

UNGEWICKELL, E., R. A. GARRETT, C. EHRESMANN, P. STIEGLER, and P. CARBON. 1977. Further characterization of the RNA structure in the binding region of protein S4 and 16 S ribosomal RNA of *Escherichia coli*. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* **81**:193.

WONG, K. P., and L. M. HAMLIN. 1975. The role of Zn(II) on the folding of bovine carbonic anhydrase B. *Arch. Biochem. Biophys.* **170**:12.

WONG, K. P., and C. TANFORD. 1973. Denaturation of bovine carbonic anhydrase B by guanidine hydrochloride. *J. Biol. Chem.* **248**:8518.

ZIMMERMANN, R. A., A. MUTO, P. FELLNER, C. EHRESMANN, and C. BRANLANT. 1972. Location of ribosomal protein binding sites on 16 S ribosomal RNA. *Proc. Natl. Acad. Sci. U.S.A.* **69**:1282.