A COMPARISON OF THE STRUCTURES OF SEVERAL ACID-UREA EXTRACTED RIBOSOMAL PROTEINS FROM *ESCHERICHIA COLI* USING PROTON NMR

C. A. MORRISON and E. M. BRADBURY

Biophysics Laboratories, Portsmouth Polytechnic, St Michael's Building, White Swan Road, Portsmouth, Hampshire, England

and

R. A. GARRETT

Max-Planck-Institut für Molekulare Genetik, Abt. Wittmann, Ihnestrasse 63-73, 1 Berlin 33 (Dahlem), Germany

Received 27 June 1977

1. Introduction

The ribosome of *Escherichia coli* contains about 52 proteins most of which are basic and have molecular weights between 5000 and 18 000 and some bind directly to the ribosomal RNAs and facilitate their folding [1]. The functions of the proteins are uncertain, however, and although some have been assigned functions, indirectly, such as GTPase and peptidyl transferase, no such activity has been exclusively and unambiguously assigned to any single protein [2]. On the other hand the direct involvement of the RNA in the functioning of the ribosome has been established (e.g. [3]).

Ribosomal proteins are generally prepared under strongly denaturing conditions [4]. Most of the 30 S subunit proteins, however, are biologically viable since they can be subsequently reconstituted into active subunits [5]. Very little is known however about the protein structures, for two reasons:

1. Most of the proteins are difficult to dissolve in buffers at neutral pH.
2. They have not been systematically studied because of doubt as to whether they assume similar structures in the free state as in the ribosome.

In the present study, 24 ribosomal proteins were investigated to establish whether they exhibited tertiary structures or behaved as random coil polypeptides in solution. High resolution proton NMR was chosen since it yields direct information about (a) the presence of tertiary folding within the proteins and (b) the degree of flexibility of the structure. The occurrence of tertiary folding, involving aromatic residues, is indicated by the formation of ring-current shifted apolar methyl resonances and perturbation of aromatic residue resonances in the spectra. The degree of flexibility in the protein is reflected in the relative widths of the amino acid resonance peaks under 'native' and denatured conditions. Of the 24 proteins studied, six, namely S4, S15, S16, L7/12, L18 and L25 exhibited tertiary structure but none was detected in the remainder.

2. Materials and methods

Proteins were prepared by CM-cellulose chromatography and Sephadex G 100 gel-filtration in 6 M urea [4]. They were lyophilised and dissolved in D_2O containing 0.1 mM dithiothreitol or in 10 mM K phosphate, 0.1 mM dithiothreitol, p*D* 7.0. Denatured protein spectra were obtained at low pH or in 6 M urea. The proteins were provided by Dr H G Wittmann. Protein concentrations were determined by weighing or using the Folin reagent [6]. They were between 1 mg/ml and 5 mg/ml. Proton NMR spectra were obtained using a Bruker WH 270 spectrometer operating in the Fourier Transform mode. Typical observation conditions were 12 μs pulse length, data collection over 0.5 s and multiplication of the free induction decay by an exponential equivalent to line broadening by about 2 Hz.

3. Results

Protein NMR spectra were obtained for the following proteins: S3, S4, S5, S6, S7, S8, S9, S11, S12, S13, S15, S16, S18, S20, S21, L1, L2, L7/12, L18, L24, L25, L27, L29 and L30. They were examined in low salt at p*D* 7, because only under these conditions were they sufficiently soluble for NMR studies.

Only the spectra of proteins S4, S15, S16, L7/12, L18, and L25 were typical of structured proteins. They each exhibited ring-current shifted resonances in the high-field region of the spectra and perturbations in the low-field aromatic region. They are shown in fig.1

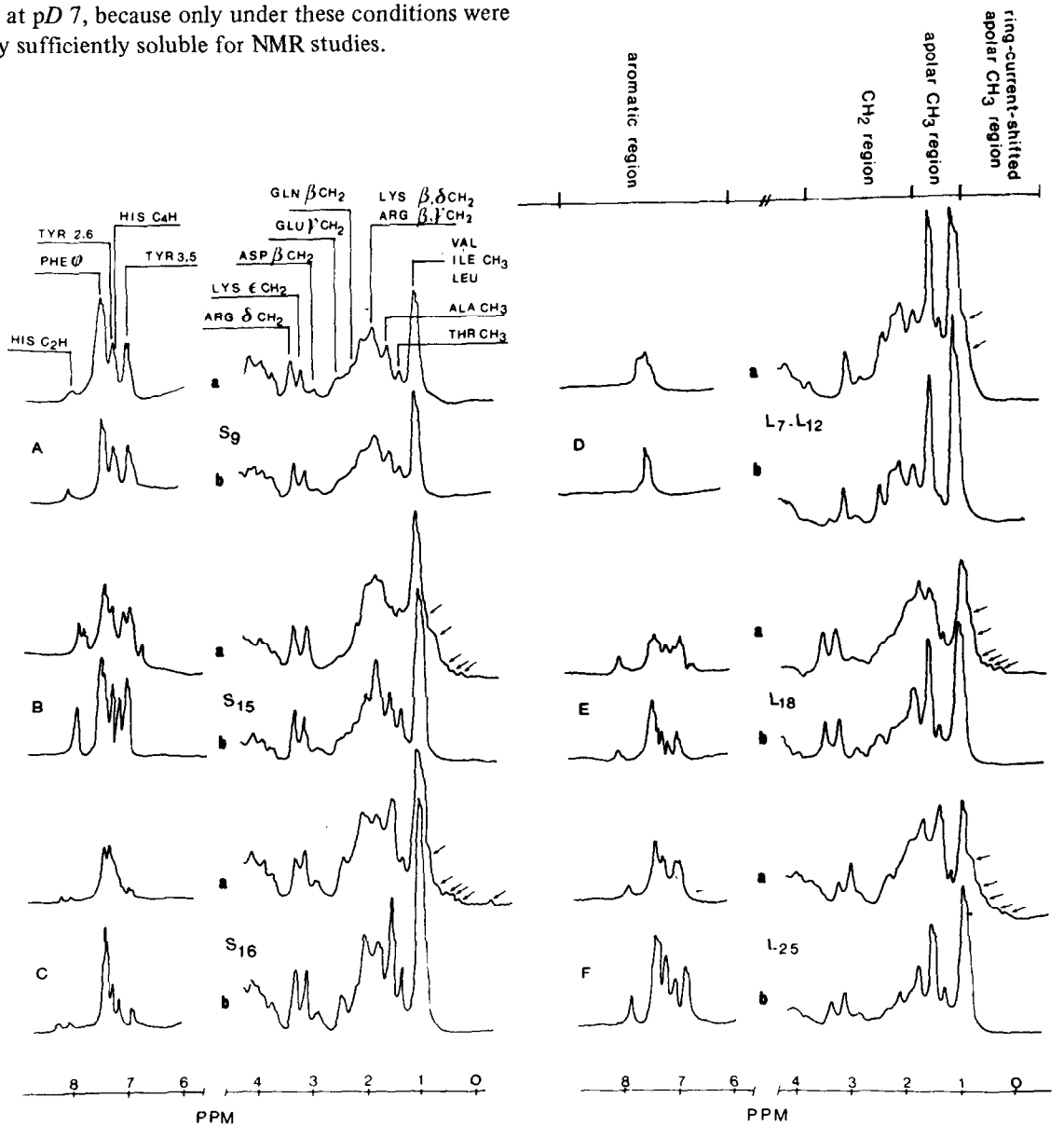

Fig.1. 270 MHz proton NMR spectra of proteins (A) S9, (B) S15, (C) S16. (D) L7/12, (E) L18 and (F) L25. For each protein spectra were obtained (a) in 10 mM K phosphate, 0.1 mM dithiothreitol, p*D* 7. Denatured protein spectra (b) were obtained in 6 M urea. Aromatic region is shown X4 for each spectrum.

together with the spectra of the denatured proteins in 6 M urea. The spectra of protein S4 have been published elsewhere [7]. The degree of tertiary folding in these proteins decreased from S16, with six well-resolved ring-current shifted peaks, to S4 and L7/12 with only one or two such peaks occurring as shoulders on the high-field side of the apolar methyl resonance position. Each of the proteins also yielded complex perturbations in the low-field aromatic region. The remainder of the proteins yielded spectra typified by S9 in fig.1A. No ring-current shifted peaks or low-field perturbations were seen and there was almost no dipolar broadening of the resonance line-widths: it was concluded that they exhibit very little or no tertiary structure (see [8] for a general review on protein NMR). Proteins S6, S8 and L1 were exceptional in that they tended to aggregate during the NMR run at neutral pH. This reduced the signal to noise ratio in the spectra. Nevertheless, there was no evidence for tertiary structure.

Each protein was examined under higher ionic strength conditions in order to establish whether extra structure could be induced by salt. This was achieved by slowly titrating with 1 M KCl. Many of the proteins precipitated and no spectra were obtained. A few proteins S4, S18, S20, L2 and L24, appeared soluble but showed a loss in the NMR signal area indicating that they were partially aggregated; however, the general features of the spectra remained unaltered. Only proteins S9, S21 and L7/12 were very soluble at higher ionic strengths, but under these conditions they yielded spectra identical to those in low salt. There was no evidence, therefore, for salt-induced tertiary structure in the latter proteins.

4. Discussion

Very little is known about the secondary and tertiary structures of ribosomal proteins. The α-helix contents of proteins S3, S4, S6, S7 and S8 were determined at neutral pH where they exhibit the most structure and they fell in the range 18–50% [9]. Theoretical secondary structure estimations have also been made (reviewed in [1]) but the extent to which they apply to the exceptional and complex structures of the ribosomal proteins is uncertain. Shape determinations have been performed on proteins S4, L7/12, L18 and L25 and they were each shown to have asymmetric and extended structures (e.g. [10–12]).

Our NMR results demonstrate that proteins S4, S15, S16, L7/12, L18 and L25 contain some tertiary structure, although less than expected for fully globular proteins. No evidence of such structure was detected in the other 18 proteins. Moreover, no salt-induced structure was observed for those proteins that were soluble in high salt:

Two inferences can be drawn from these results:

1. The spectra reflect the 'native' protein structures in solution and the six proteins are more structured than the others.
2. All proteins are to some degree denatured after the fractionation procedure, and the group of structured proteins simply have more stable, or more easily renaturable, structures than the rest.

In an attempt to distinguish between these two possible inferences we examined further the known primary sequences to see whether the structured proteins exhibited any properties that distinguished them from the others. Of the proteins examined by NMR, S4, S6, S8, S9, S12, S13, S15, S16, S18, S20, S21, L7/12, L18, L25, L27, L29 and L30 have been sequenced (reviewed in [1]). The six structured proteins are included in this group. In fig.2 the proteins are arranged in histograms according to (A) charged residue content (B) charged/apolar residue ratio and (C) basic/acidic residue ratio. For comparison, data are included for globular proteins of known tertiary structure, fibrous proteins and histone proteins [13].

In fig.2A it can be seen that the ribosomal proteins are exceptional in that their percentage content of charged residues is highly conserved; 11 of the 21 sequenced proteins have a charged residue content of 31 (±1)%.

The distribution of the same proteins according to their charge/apolar residue ratios (fig.2B) shows that the values for the ribosomal proteins are higher than for the globular proteins. Since high charge/apolar ratios would be compatible with more extended protein structures in which there is a maximum exposure of polar residues to the solvent, the high ratio values for the ribosomal proteins supports the NMR data which indicated little tertiary structure.

The arrangement of the proteins according to their basic/acidic residue ratios (fig.2C) indicates that the proteins which exhibited structure in their NMR

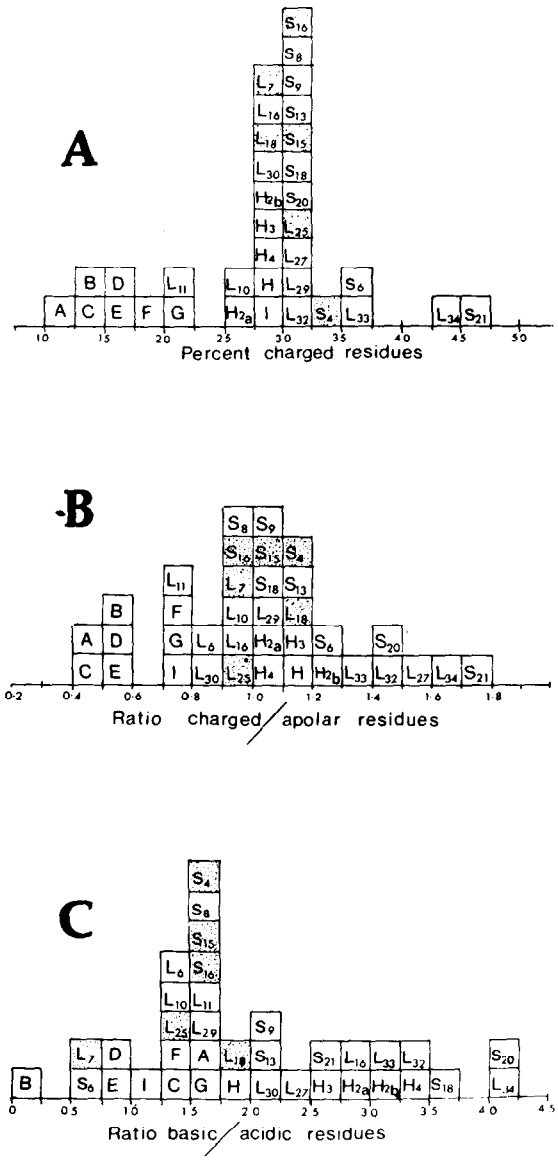

Fig.2. Histograms of ribosomal proteins, histones and other proteins of known sequence arranged according to (A) charged residue content (B) charged/apolar residue ratio and (C) basic/acidic residue ratio. The following abbreviations are used: A=bovine trypsinogen, B=ribonuclease T1, C=bacteriophage F2 coat protein, D=repressor protein, LAC operon, E=TMV (vulgare) coat protein, F=bovine RNAase, G=HEW lysozyme, H=cytochrome *c* (k angaroo), I=Myoglobin (sperm-whale). H2A, H2B, H3 and H4 denote the histones from calf thymus. The basic residues do not include histidines. The apolar residues do not include alanine, which because of its small size is assumed to be readily incorporated into any structured region. 'Structured' proteins are shown stippled.

spectra occur in the same region as the globular proteins. Proteins S6 and S8, that were difficult to study at neutral pH (see Results section), and L29, also occur in this region but showed no structure. Apart from these exceptions, however, the grouping of the proteins correlates with the NMR results. The relatively low basicities of 'structured' ribosomal proteins would lead to less intra-molecular repulsions between positively charged amino acids, and this in turn would produce stable (and possibly more easily renaturable) tertiary structures in those proteins. It also suggests the possibility that the more basic proteins may assume tertiary structure when their positive charges are partially neutralised in the environment of either ribosomal RNA or the acidic regions of other ribosomal proteins.

A further correlation was also possible between the distributions of basic, acidic and apolar residues in the sequences and the NMR data. For most proteins the distributions are fairly random. However, a few proteins are very polarised and they include five of the six structured proteins. S4 has a highly apolar central region. In S15 the basic residues are concentrated in the C-terminal region and in S16 and L18 they are concentrated in the N-terminal region. L7/12 is rich in apolar residues. This concentration of basic residues at the ends of four of these proteins will render the composition of the other region of the protein more suitable for forming a tertiary structure. Moreover, the highly basic regions will be too highly charged to assume a globular structure, and this could explain why none of the spectra observed for the 'structured' proteins resemble those of globular proteins in all aspects.

There are two criteria that can be invoked to show that many of the proteins under study have retained some native structure and are therefore not irreversibly denatured, namely the capacities of the acid–urea extracted proteins (1) to reconstitute into subunits and (2) to bind specifically to ribosomal RNAs. The former has been established for the 30 S subunit proteins [5]. We have tested each of the proteins, prepared as described above, for binding to the RNAs (reviewed in [14]). Of the proteins examined by NMR, ten bind specifically to the RNAs, namely S4, S7, S8, S15, S20, L1, L2, L18, L24 and L25. Four of these, only, fall in our group of structured proteins. This result, together with the observation that proteins

prepared by procedures that minimise protein denaturation show very similar RNA binding properties to the acid–urea extracted proteins [15], suggests that even in the 'native' state these proteins exhibit very little tertiary structure.

There is one reservation, however, about the acid–urea extracted proteins, in that little quantitative RNA-binding data is available. There is the possibility, for example, that, say, a large percentage of the protein molecules are denatured and only a small percentage are not; the NMR signals from the 'native' and presumably more structured protein molecules, might then be too weak relative to those from the denatured material. There are three lines of evidence to suggest that this may be so at least for some acid-urea extracted proteins.

1. Limited proteolysis of protein S4 when free and complexed with 16 S RNA gives higher yields of a protected fragment from the complex; this together with the observation that the protein fragment binds more efficiently to the 16 S RNA than the free protein suggests that the free S4 protein contains some denatured molecules [16].
2. Fluorescence spectroscopic studies on protein S4 indicated the presence of two different conformations [17].
3. When one of the 'structured' proteins, namely S 16, was dissolved in 8 M urea and dialysed into high salt buffer at neutral pH, there was great enhancement of the ring-current shifted peaks in the NMR spectrum. The procedure selectively enriched the structured form of the protein by causing precipitation of denatured material. The results of similar experiments performed on other proteins, including some of the 'non-structured' ones, suggest that renaturation of already denatured material does not occur.
4. NMR experiments currently in progress on ribosomal proteins prepared under non-denaturing conditions suggest that some of them contain significantly more structure than those prepared under the denaturing conditions described in this paper [18].

Acknowledgements

We thank Dr H G Wittmann for his support and encouragement. Dr Anders Liljas of the Wallenberg Laboratory, Uppsala, is sincerely thanked for a stimulating discussion based on some of his own unpublished observations relevant to the data in fig.2.

References

[1] Stöffler, G. and Wittmann, H. G. (1977) in: Protein Biosynthesis (Weissbach, H. ed.) in press.
[2] Brimacombe, R., Nierhaus, K. H., Garrett, R. A. and Wittmann, H. G. (1976) in: Prog. Nucl. Acid. Res. Mol. Biol. 18, 1–44.
[3] Noller, H. F. and Chaires, J. B. (1972) Proc. Natl. Acad. Sci. USA 69, 3115–3118.
[4] Wittmann, H. G. (1974) in: Ribosomes (Nomura, M., Lengyel, P. and Tissières, A. eds) p.93, Cold Spring Harbor Press, New York.
[5] Mizushima, S. and Nomura, M. (1970) Nature 226, 1214–1218.
[6] Lowry, O. H., Rosebrough, N. J., Farr, A. L. and Randall, R. J. (1951) J. Biol. Chem. 193, 265–275.
[7] Morrison, C. A., Garrett, R. A. and Bradbury, E. M. (1977) Eur. J. Biochem. in press.
[8] McDonald, C. C. and Phillips, W. D. (1970) in: Fine Structure of Proteins and Nucleic Acids (Fasman, G. D. and Timasheff, N. eds) Vol. 4, p. 1, Dekker, New York.
[9] Lemieux, G., Lefevre, J. F. and Daune, M. (1974) Eur. J. Biochem. 49, 185–194.
[10] Rohde, M. F., O'Brien, S., Cooper, S. and Aune, K. C. (1975) Biochemistry 14, 1079–1087.
[11] Österberg, R., Sjöberg, B., Liljas, A. and Pettersson, I. (1976) FEBS Lett. 66, 48–51.
[12] Österberg, R., Sjöberg, B. and Garrett, R. A. (1976) FEBS Lett. 65, 73–76.
[13] Dayhoff, M. O. (ed) (1972, 1976 supp.) Atlas of Protein Sequence and Structure, Vol. 5, National Biomedical Research Foundation, Silver Spring, Maryland.
[14] Garrett, R. A. and Wittmann, H. G. (1973) Advan. Prot. Chem. 27, 277–347.
[15] Littlechild, J., Dijk, J. and Garrett, R. A. (1977) FEBS Lett. 74, 293–294.
[16] Newberry, V., Yaguchi, M. and Garrett, R. A. (1977) Eur. J. Biochem. 76, 51–61.
[17] Brochou, J-C., Wahl, P., Vachette, P. and Daune, M. P. (1976) Eur. J. Biochem. 65, 35–39.
[18] Morrison, C. A., Littechild, J., Dijk, J. and Bradbury, E. M. (1977) manuscript in preparation.