PREDICTION OF SECONDARY STRUCTURES IN PROTEINS FROM THE *ESCHERICHIA COLI* 30 S RIBOSOMAL SUBUNIT

B. WITTMANN-LIEBOLD, S. M. L. ROBINSON and M. DZIONARA

Max-Planck-Institut für Molekulare Genetik, Abteilung Wittmann, 1 Berlin-Dahlem 33, Ihnestrasse 63–73, Germany

Received 21 March 1977

1. Introduction

A prerequisite for a detailed understanding of the ribosome at the molecular level is a knowledge of its complex topography and the tertiary structures of its numerous components.

At present, useful data from X-ray analysis are not available for the determination of the conformation of ribosomal proteins. Hence, it is hoped that predictive methods using the amino acid sequences will give at least some insight into the secondary structures of the protein chains. The complete amino acid sequences of more than 50% of the 54 proteins present in *Escherichia coli* ribosomes have so far been determined. This provides a sound basis for the possibility of predicting details of the secondary structures of these proteins.

In this paper we summarise the results obtained by predicting the secondary structures of eleven *E.coli* ribosomal proteins according to four different predictive methods.

2. Methods

The secondary structures of the ribosomal proteins were predicted according to the methods of Burgess, Ponnuswamy and Scheraga [1], Chou and Fasman [2–4], Nagano [5,6] and Robson and Suzuki [7]. Programmes for the calculation of the predictive algorithms were kindly supplied by Drs H. A. Scheraga, K. Nagano and B. Robson and were adapted to run on the DEC-10 computer of the Fritz-Haber-Institut der Max-Planck-Gesellschaft in Berlin-Dahlem. The results of the predictions according to Nagano and to Robson were slightly modified; this will be described elsewhere [8] as will details of the programme using the rules of Chou and Fasman.

Although the various authors have defined the conformational states (helix, extended structure and turn) in somewhat different ways, we have taken into account in our comparison the most general definitions and have assigned to a particular state a common symbol. Residues were assumed to be in a random coil conformation when they were not otherwise predicted. In the results obtained from the predictions of Chou and Fasman and of Nagano, in some instances amino acid residues were predicted to be in either of two conformational states and we have therefore taken both possibilities into consideration.

3. Results and discussion

The main purpose of this study was to examine various ribosomal proteins for regions with high probabilities for helices, extended structures and turns. The results which were obtained by applying the four different predictive methods to the ribosomal proteins were compared by plotting helical regions, extended structures, turns and random coil against the protein sequences. This is shown symbolically, for example, for protein S8 in fig.1. Similar figures for the other 26 ribosomal proteins investigated will be presented elsewhere ([8] and in preparation). It is convenient to summarise these comparisons as histograms which show the predicted occurrence of the individual residues in a particular conformational state. This has been done for eleven proteins from the small subunit (figs 2–5).

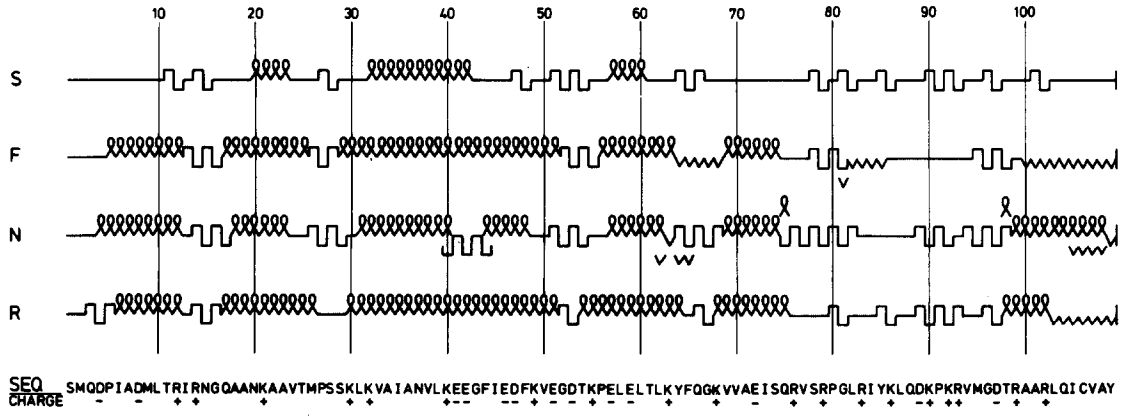

Fig.1. Secondary structure of ribosomal protein S8 predicted according to four different methods: '*S*', Burgess, Ponnuswamy and Scheraga [1]; '*F*', Chou and Fasman [2–4]; '*N*', Nagano [5,6]; '*R*', Robson and Suzuki [7]. The symbols represent residues in helical (ℓ), turn or bend (⊓), extended (V) and coil (–) conformational states, respectively. In cases of ambiguity, both predicted states are indicated.

The most probable secondary structure for a given region is indicated in the lowermost block of these figures, denoted by 'PRE', in which only those secondary structures appear when at least three predictive methods were found to be in complete agreement. This leads, in some instances, to shorter regions of secondary structure than are conventionally necessary for the complete development of a helix, turn or extended structure. Thus, it appears that short regions indicate only the location of a structural feature.

The four different predictions provide a fairly consistent picture of the regions of secondary structure in the ribosomal proteins. As can be seen in the histograms, there is a relatively high content of helix predicted in many of the proteins; turns are also predicted rather frequently. On the other hand, this group of proteins appear to be lacking in many extended structures.

3.1. *Helical regions*

A remarkably high content of helix structure is predicted for proteins S20 and S21. In table 1, the percent of residues is given which are predicted to be located in helices: 64% of the residues in protein S20 and 57% in S21 are located in helical regions. For proteins S6 and S8 the values are 36%, whereas for proteins S4, S13, S15 and S16 they range between 20% and 30%. The lowest helix content (7–15%) is predicted for proteins S12 and S18.

There is good agreement among the different predictive methods about the positions at which helical regions start but more uncertainties exist about the lengths of these regions (see ref.[8]). In some of the

Table 1
Prediction of secondary structure content in ribosomal proteins

Protein	Sequence [Ref.]	Helix (%)	Turn (%)	Extended (%)
S4	[9]	28.1	17.7	3.4
S6	[10,11]	36.3	10.3	4.4
S8	[12]	36.7	16.5	4.6
S9	[13]	18.0	19.5	6.3
S12	[14]	7.3	16.3	13.0
S13	[15]	21.4	18.8	nc[a]
S15	[16]	29.9	13.8	nc[a]
S16	[17]	22.0	14.6	4.9
S18	[18]	14.9	13.5	5.4
S20	[19]	64.0	4.7	nc[a]
S21	[20]	57.1	4.3	nc[a]

[a]nc: Not calculated according to the above assumption

Average values are calculated when at least three predictions were in complete agreement about the conformational state of a given residue

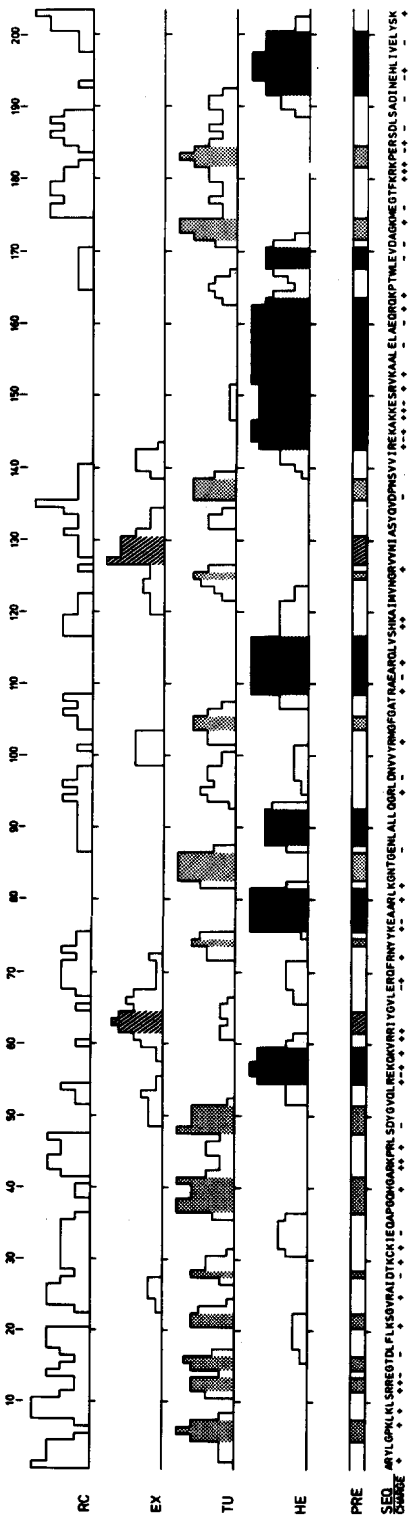

Fig.2. Histograms for the secondary structure of protein S4 after averaging the weighted results from the four predictions. A residue predicted to be in a unique state was given a weight of unity; a weight of a half was assigned to each state of an ambiguously predicted residue. 'PRE' Summarises the predictions when at least three of them are in complete agreement. *Abbreviations:* sequence (SEQ), random coil (RC), extended structure (EX), turn or bend (TU) and helical regions (HE).

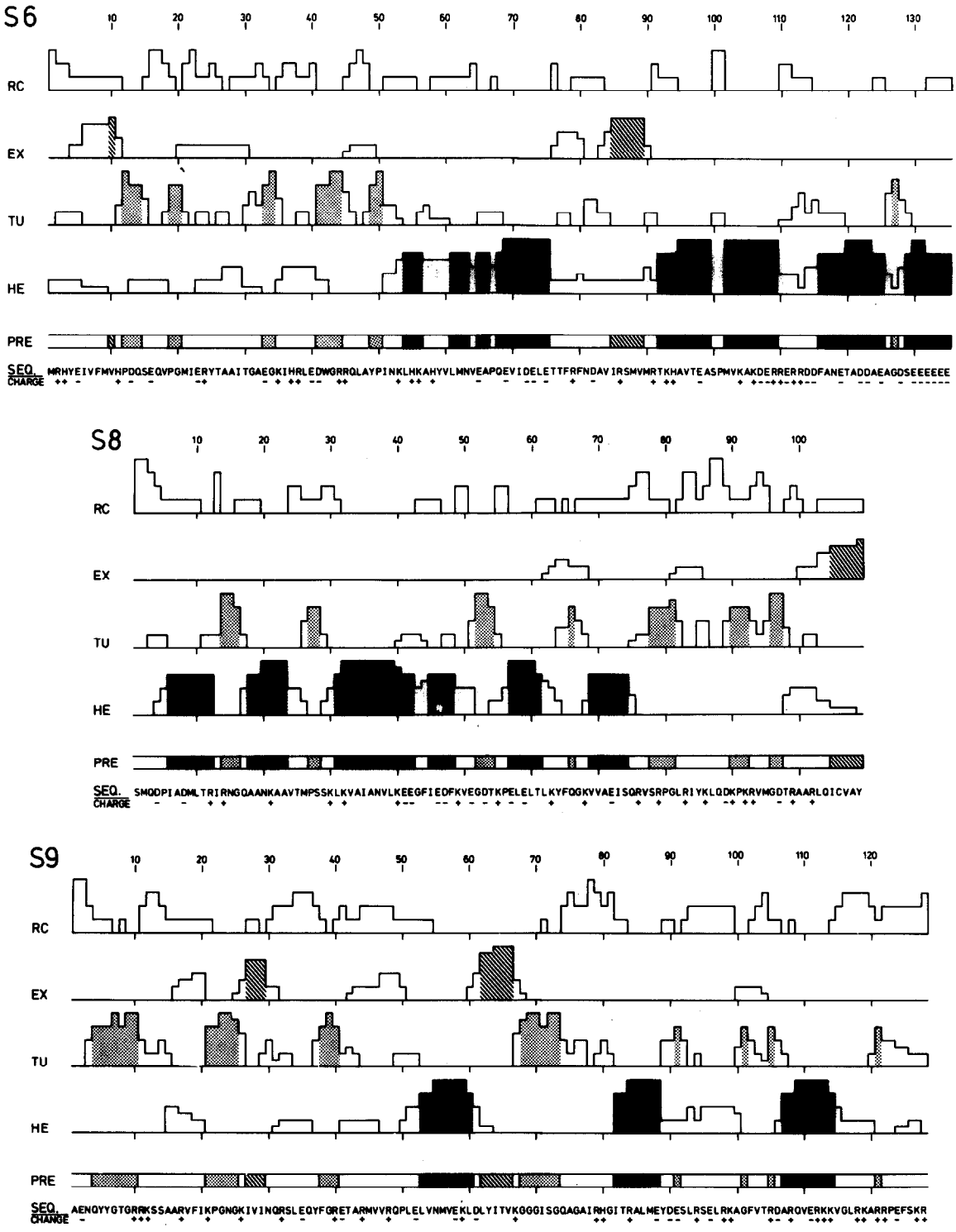

Fig.3. Histograms for the secondary structures of proteins S6, S8 and S9. Abbreviations as for fig.2.

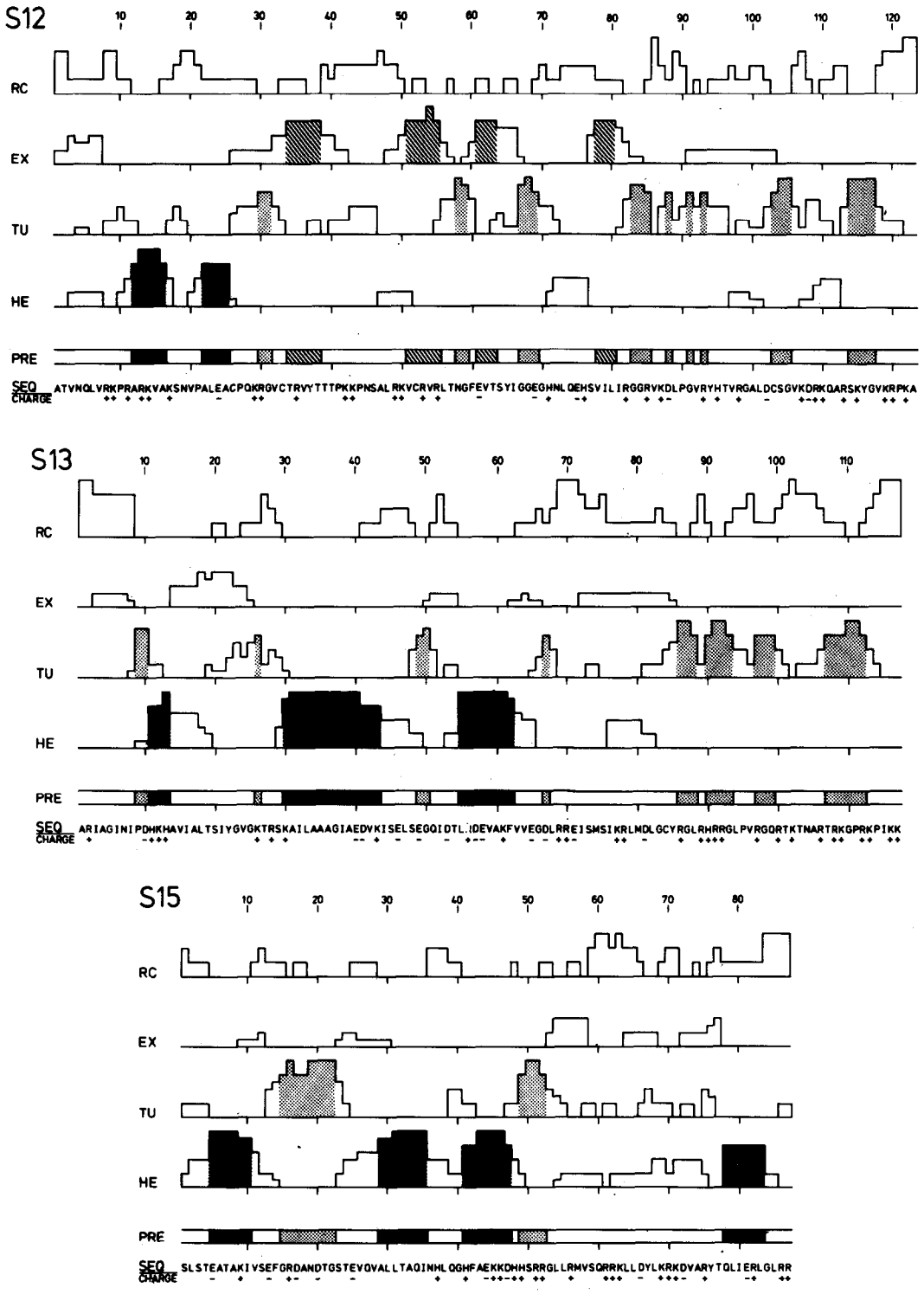

Fig.4. Histograms for the secondary structures of proteins S12, S13 and S15. Abbreviations as for fig.2.

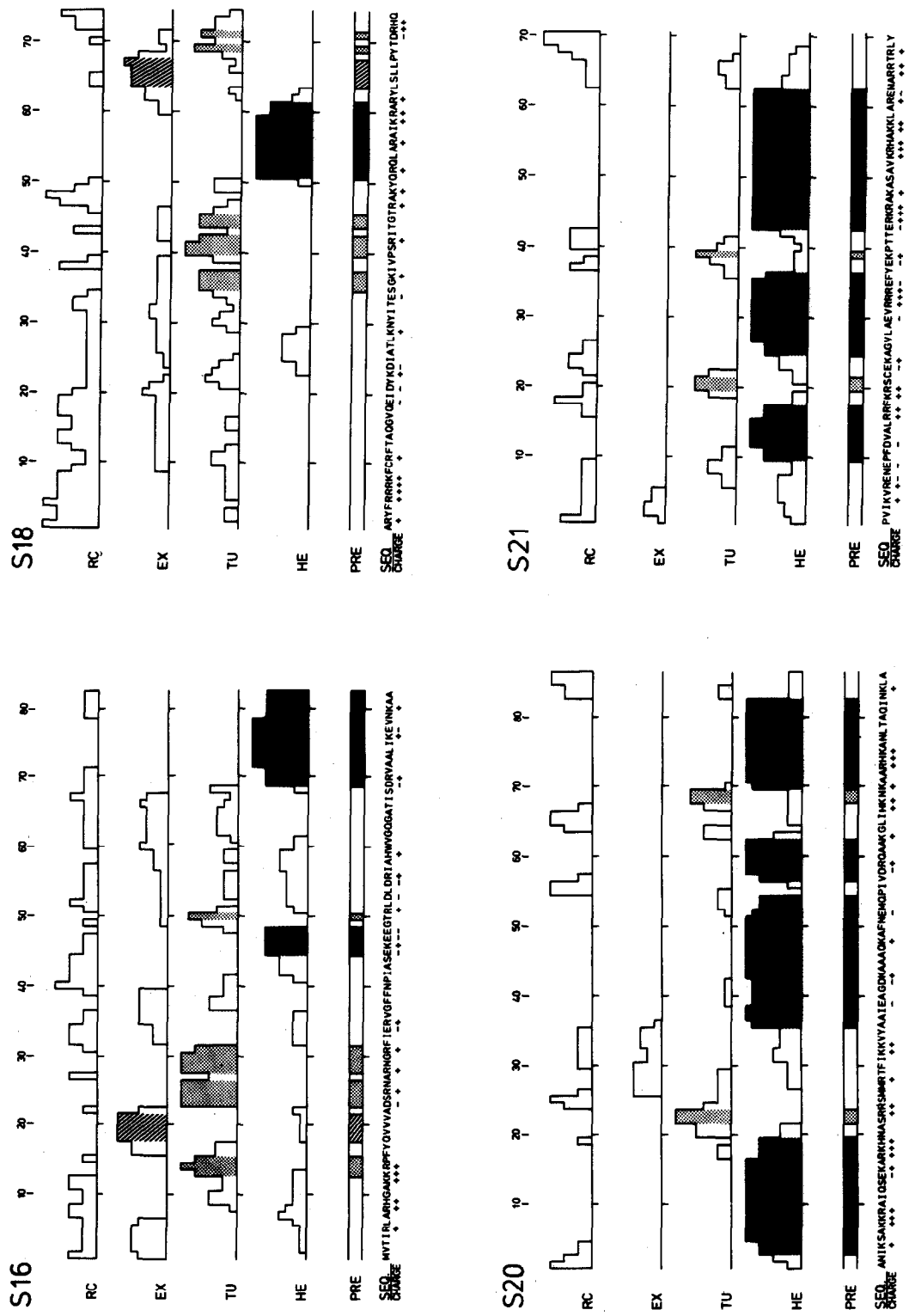

Fig.5. Histograms for the secondary structures of proteins S16, S18, S20 and S21.

ribosomal proteins the helix regions are distinctly clustered in certain parts of the protein chains (figs. 2–5).

3.2. *Extended structures*

There appears to be a relatively poor agreement among the four methods for the prediction of extended structure. Nevertheless, at least three of them agree for several of such regions in the following proteins: S4 (in positions 62–64 and 127–130); S6 (85–89); S8 (105–109); S9 (27–29 and 62–66); S12 (34–38, 51–55, 61–63 and 78–80); S16 (18–21) and S18 (64–67). Protein S12 has an exceptionally high content of extended structure (table 1). Based on the consensus of three predictions, no extended structure was likely to occur in proteins S13, S15, S20 and S21.

3.3. *Turns*

There is considerable variation in the predictions of turns. In general, turns are more predominant than extended regions although their positions are not so well defined. High potentials for turns are predicted for proteins S9, S13, S8, S12, S4, S16, S15 and S18, as shown in figs 2–5 and in table 1. Turns are clustered in the N-terminal parts of proteins S6 and S9 and in the C-terminal regions of proteins S8, S12 and S13 (figs 2–5).

The results of this investigation show that the ribosomal proteins investigated are predicted to differ quite markedly in their secondary structures. This finding is analogous to conclusions drawn from a comparison of their primary structures [21]. Similarities at this level of conformation are shown by proteins S20 and S21 and to a certain extent also by proteins S13 and S15 and by proteins S16 and S18.

It is needless to emphasise that a prediction of the secondary structure of protein based on its amino acid sequence has to be supplemented by experimental results, such as those given by optical methods. CD-Studies of most of the ribosomal proteins are in progress and will enable further structural comparisons to be presented in the near future.

References

[1] Burgess, A. W., Ponnuswamy, P. K. and Scheraga, H. A. (1974) Israel J. Chem. 12, 239–286.
[2] Chou, P. Y. and Fasman, G. D. (1974) Biochemistry 13, 211–222.
[3] Chou, P. Y. and Fasman, G. D. (1974) Biochemistry 13, 222–249.
[4] Chou, P. Y., Adler, A. J. and Fasman, G. D. (1975) J. Mol. Biol. 96, 29–45.
[5] Nagano, K. (1977) J. Mol. Biol. 109, 235–250.
[6] Nagano, K. (1977) J. Mol. Biol. 109, 251–274.
[7] Robson, B. and Suzuki, E. (1976) J. Mol. Biol. 107, 327–356.
[8] Dzionara, M., Robinson, S. M. L. and Wittmann-Liebold, B. (1977) Hoppe-Seyler's Z. Physiol. Chem., submitted.
[9] Schiltz, E. and Reinbolt, J. (1975) Eur. J. Biochem. 56, 467–481.
[10] Hitz, H., Schäfer, D. and Wittmann-Liebold, B. (1975) FEBS Lett. 56, 259–262.
[11] Hitz, H., Schäfer, D. and Wittmann-Liebold, B. (1977) Eur. J. Biochem. in press.
[12] Stadler, H. and Wittmann-Liebold, B. (1976) Eur. J. Biochem. 66, 49–56.
[13] Chen, R. and Wittmann-Liebold, B. (1975) FEBS Lett. 52, 139–140.
[14] Funatsu, G., Yaguchi, M. and Wittmann-Liebold, B. (1977) FEBS Lett. 73, 12–17.
[15] Lindemann, H. and Wittmann-Liebold, B. (1976) FEBS Lett. 71, 251–255.
[16] Morinaga, T., Funatsu, G., Funatsu, M. and Wittman, H. G. (1976) FEBS Lett. 64, 307–309.
[17] Vandekerckhove, J., Rombauts, W. and Wittmann-Liebold, B. (1977) FEBS Lett. 73, 18–21.
[18] Yaguchi, M. (1975) FEBS Lett. 59, 217–220.
[19] Wittmann-Liebold, B., Marzinzig, E. and Lehmann, A. (1976) FEBS Lett. 68, 110–114.
[20] Vanderkerckhove, J., Rombauts, W., Peeters, P. and Wittmann-Liebold, B. (1975) Hoppe-Seyler's Z. Physiol. Chem. 356, 1955–1976.
[21] Wittmann-Liebold, B. and Dzionara, M. (1976) FEBS Lett. 65, 281–283.