Tertiary structure of tRNA^{Phe}

A possible correlation between the structural functional unit of this tRNA and its exonic sequence

R. MALATHI and N. YATHINDRA

Department of Crystallography and Biophysics, University of Madras, Guindy Campus, Madras-600 025, India

(Received 16 January 1984/Accepted 23 January 1984)

It has been shown recently [Go (1981) *Nature (London)* **291**, 90–92; Blake (1983) *Trends Biochem Sci.* **8**, 11–13] that the exonic regions of the genes of proteins haemoglobin, lysozyme and immunoglobulin correspond closely to the compactly folded structural units. Despite the absence of classical domain structures in tRNA compared with those found in several proteins, close inspection of certain features in the distance maps obtained for yeast tRNA^{Phe} using the conformationally equivalent heminucleotide scheme reveals that a similar situation might also be present in ribonucleic acids such as tRNA species and the exonic sequences of their genes. Also it seems possible that certain segments of yeast tRNA^{Phe} may be characterized as possessing a domain-like character, and this seems to provide stereochemical support for possible conservation of L-shape structure for tRNA species lacking the entire dihydrouridine arm such as those found in mitochondria.

The folding of a polymer chain, especially of one having secondary and tertiary structural features, can be represented on a two-dimensional plane by computation of distances of all the conformationally equivalent repeating units. Recently we obtained (Malathi & Yathindra, 1981, 1982*a,b*) such a distance map for the first time for a nucleic acid structure (yeast tRNA^{Phe}) and correlated the several patterns found in the map with the various secondary and tertiary structural features. In addition, aspects like the not-so-obvious approximate twofold symmetry between tertiary structural domains and the similar nature of interactions between a number of substructures, were also brought out.

To obtain further details from the distance map we have now extended these studies and looked for regions or zones that are greater than 2, 3, 4 and 5 nm (20, 30, 40 and 50 Å). Such large separations occurring close to the diagonal are particularly important, since they provide the most relevant information concerning separation between residues that are proximal in the primary structure but significantly separated in the tertiary structure, owing to chain folding. They also enable one to distinguish and characterize domains or folded units which may be regarded as segregated structural entities virtually independent from the rest of the tertiary structure. This information

might become useful in relating the structure with their gene sequences as found (Go, 1981; Blake, 1983) in certain proteins (haemoglobin and lysozyme).

Fig. 1 shows the constant value contours drawn from 1 to 4 nm (10 to 40 Å) on the distance map of yeast tRNA^{Phe} obtained by using the coordinates (Ladner *et al.*, 1977) of the monoclinic structure. Three distinct regions of larger separations marked A, B and C appear nearer to the diagonal. Region A encloses heminucleotide separations 3 nm (30 Å) < t_{ij} > 2 nm (20 Å) and occurs between the patterns 2a and 2, which characterize the P10 loop and dihydrouridine arm respectively. This region spans approximately the residue sequences 7–10 and 13–22, which interestingly are not displaced far apart in the primary structure. However, the tertiary folding renders them more or less isolated so as to be regarded almost as a subdomain or a segregated entity. That the above can be considered as a subdomain is also apparent from a close examination of the yeast tRNA^{Phe} model, which reveals that the interface area (a lower value for this is used as one of the criteria in characterizing domain behaviour) between the 13–22-residue stretch of the polynucleotide backbone and the other regions is quite small. Further supporting evidence for this can be seen in the recent quantitative interface-area calculations of

Fig. 1. Distance map for yeast $tRNA^{Phe}$ obtained by computing distances of P and C-4' atoms of each of the heminucleotides from every other in the sugar-phosphate backbone

Sequence number of the residues from 5' to 3' along the abscissa and 3' to 5' along the ordinate are shown. Equivalence contours of 1, 1.5, 2, 3 and 4 nm (10, 15, 20, 30 and 40 Å) are shown. All patterns representing residue separations of less than 2 nm (20 Å) are marked by dots. These bring out nicely the approximate twofold symmetry present between the secondary as well as tertiary domains. Region marked A close to the diagonal encloses separations greater than 2 nm (20 Å). Dark regions marked B and C, again rather close to the diagonal, represent residue separations greater than 5.5 nm (55 Å). Regions A and C are interpreted as implying the segregated-domain-like behaviour of the variable dihydrouridine extra arms respectively. This would enable their easy insertion or deletion into the L-shape tertiary structure. Region B, together with the feature within the square of 35 residues near the origin suggesting the folding back of the polynucleotide chain, are identified as reflecting the presence of intron in the yeast $tRNA^{Phe}$ gene. Key to contour lines: ----, 1 nm separation; ---, 1.5 nm; —, 2 nm; ···, 3 nm; —○—, 4 nm.

Janin & Wodak (1983). On their interface area scan, two minima are found within the 10 nm^2 (1000 Å^2) contour. The lowest minimum of 9 nm^2 (900 Å^2) for the interface area occurs for the residue stretch 22–45, and the other minimum, which is obviously only slightly higher, corresponds to residue sequence 13–45. We find that the value for the stretch 13–22 also lies within the 10 nm^2 (1000 Å^2) region and therefore not significantly different from the other two minima. Thus the polynucleotide stretch 13–22 encompassing the dihydrouridine loop and part of the dihydrouridine DH stem warrants it to be characterized at least as a subdomain. Although the number of residues

involved in forming this domain is only ten, it should be realized that this is nearly equivalent to at least 20 peptide residues. The domain-like character of the dihydrouridine arm in the tertiary structure should facilitate insertion or deletion of nucleotide residues in the dihydrouridine region while still maintaining the basic L-framework (Robertus *et al.*, 1974; Quigley *et al.*, 1975; Sundaralingam *et al.*, 1976). The smallest separation between the 5'-terminus of the amino acid and the beginning of the 5'-anticodon stem is only about 1.5 nm (15 Å), which can be very easily linked by a stretch of three to four nucleotides that are always present, even in tRNA species that

Fig. 2. Distance map for yeast tRNA^{Phe} constructed by using separations of glycosyl nitrogen atoms of the bases. As in the backbone plot, three regions marked A, B and C, close to diagonal, represent base separations greater than 2 and 5.5 nm (20 and 55 Å) and lead to conclusions similar to those derived from the backbone plot. Key to contour lines: ----, 0.5 nm separation; —, 1 nm; ····, 1.5 nm; — — —, 2 nm.

completely lack the dihydrouridine arm, such as those found in mitochondria (Gauss & Sprinzl, 1983). Suggestions have also been made from sequence studies (de Bruin *et al.*, 1980) and stereochemical modelling (Sundaralingam, 1982) that it may be possible to preserve the L-shape for mitochondrial tRNA species. The tertiary interactions stabilizing the structure here may be different in nature and also less in number.

Dark regions B and C in Fig. 1 occurring rather close to diagonal represent distances of separation of more than 5.5 nm (55 Å). Region C suggests the segregated or domain-like character of the structure formed out of the extra arm and loop present in certain tRNA species such as *Escherichia coli* tRNA^{Ser} etc. This would facilitate the insertion or deletion of this structural unit without interfering with the L-shape tertiary structure. Molecular-modelling studies (Brennan & Sundaralingam, 1976) have indicated the non-interference of long variable loops with the L-shape of tRNA species.

A similar large separation of more than 5.5 nm

denoted by 'B' involves the anticodon-loop region. Most interestingly this is also the region where the intron is found in the gene of yeast tRNA^{Phe}. Since there is no possibility of variability either in the anticodon stem or loop, the large separations in region B might actually be reflecting the presence of an intron found to occur in the yeast tRNA^{Phe} gene (Gauss & Sprinzl, 1983). Close observation of patterns, especially those within the square of 35 residues near the origin, clearly reflect that the polynucleotide chain is folded back, with the approximate stretch of residues comprising the 5'-terminus to the anticodon forming one fold, and the remaining portion of the tRNA chain constituting the other fold. It is interesting that the intron occurs in this region separating these folded units, each of which comprises the most important functional component of tRNA, namely the CCA end and the anticodon. Distance plots (Fig. 2) obtained by considering base separations involving successive glycosyl nitrogen atoms have also yielded three regions of large separations, all of

which can be interpreted similarly to those described above.

The results of the present analysis, leading to the characterization of the variable dihydrouridine and extra arms as segregated structural units probably indicate their limited influence in stabilizing the L-shape tRNA structure as well as in tRNA function. The tertiary structure of tRNA devoid of these folds and arms may actually represent the compact ancestral basic structural functional unit of tRNA species in general, at least with regard to biosynthesis of proteins. It might be that tRNA species with the full complement of dihydrouridine and variable arms may be a subsequent development in their evolution, enabling them to perform other functions in which they are presumably involved. Probably mitochondrial tRNA species (Gauss & Sprinzl, 1983), which lack the dihydrouridine arm, may be ancestral in origin (Deno *et al.*, 1982). The occurrence of longer introns in these tRNA genes compared with nuclear tRNA genes also lend support to the notion that mitochondrial tRNAs may be evolutionarily ancient.

N. Y. thanks the International Centre for Theoretical Physics, Trieste, Italy, for a Fellowship, and Professor Quadrifoglio for hospitality.

This work is supported through a fellowship grant from the Council of Scientific and Industrial Research to R. M. This paper is contribution 633 from this Department.

References

- Blake, C. C. F. (1983) *Trends Biochem. Sci.* **8**, 11–13
- Brennan, T. & Sundaralingam, M. (1976) *Nucleic Acids Res.* **3**, 3235–3251
- de Bruin, M. H. L., Schreier, P. H., Eperon, I. C., Barrel, B. G., Chen, E. Y., Armstrong, P. W., Wong, J. F. H. & Roe, B. A. (1980) *Nucleic Acids Res.* **8**, 5213–5218
- Deno, H., Kato, A., Shinozak, K. & Sugiura, M. (1982) *Nucleic Acids Res.* **10**, 7511–7520
- Gauss, D. H. & Sprinzl, M. (1983) *Nucleic Acids Res.* **11**, r1–r103
- Go, M. (1981) *Nature (London)* **291**, 90–92
- Janin, J. & Wodak, S. J. (1983) *Prog. Biophys. Mol. Biol.* **42**, 21–77
- Ladner, J. E., Jack, A., Robertus, J. D., Brown, R. S., Rhodes, D., Clark, B. F. C. & Klug, A. (1977) *Labquip: Nicholson Molecular Model Kit*, Wokingham, Berks., U.K.
- Malathi, R. & Yathindra, N. (1981) *Curr. Sci.* **50**, 1053–1055
- Malathi, R. & Yathindra, N. (1982a) *Biochem. J.* **205**, 457–460
- Malathi, R. & Yathindra, N. (1982b) *Biopolymers* **21**, 2033–2047
- Quigley, G. J., Seeman, N. C., Wang, A. H. J., Suddath, F. L. & Rich, A. (1975) *Nucleic Acids Res.* **2**, 2329–2341
- Robertus, J. D., Ladner, J. E., Finch, J. T., Rhodes, D., Brown, R. S., Clark, B. F. C. & Klug, A. (1974) *Nature (London)* **250**, 546–551
- Sundaralingam, M. (1982) in *Conformation in Biology* (Srinivasan, R. & Sarma, R. H., eds.), pp. 191–225, Adenine Press, New York
- Sundaralingam, M., Mizuno, H., Stout, C. D., Rao, S. T., Leibman, M. & Yathindra, N. (1976) *Nucleic Acids Res.* **3**, 2471–2481