Biochem. J. (1976) **155**, 465–475
Printed in Great Britain

Spectroscopic Evidence for the Uneven Distribution of Adenine and Uracil Residues in Ribosomal Ribonucleic Acid of *Drosophila melanogaster* and of *Plasmodium knowlesi* and its Possible Evolutionary Significance

By ROBERT A. COX and ELIZABETH GODWIN
National Institute for Medical Research, Mill Hill, London NW7 1AA, U.K.

and JEREMY R. B. HASTINGS
The Chester Beatty Research Institute, Fulham Road, London SW3 6JB, U.K.

(*Received* 13 *October* 1975)

RNA was isolated from subribosomal particles of the malaria parasite *Plasmodium knowlesi*. The nucleotide composition (mole fraction) of the principal species was obtained (S-rRNA, 0.295A, 0.36U, 0.25G, 0.105C: L-rRNA, 0.326A, 0.31U, 0.228G, 0.144C). Ribosomal RNA was also isolated from *Drosophila melanogaster*. Optical properties of these A + U-rich species were measured. In all four cases analysis of the hypochromic effect revealed that adenine and uracil residues tended to form clusters along the polynucleotide chain. A substantial fraction of residues was located in bihelical regions of approx. 50% G·C base pairs or in regions of approx. 30–35% G·C base pairs. The possible evolutionary significance of these results was considered on the basis of comparison with properties of rRNA from bacteria (*Escherichia coli*) and a mammal (rabbit reticulocyte).

The nucleotide composition of the major RNA species (L-rRNA) of the larger subparticle of cytoplasmic ribosomes ranges from approx. 37% G + C (protozoa) to approx. 67% G + C for mammals (Lava-Sanchez *et al.*, 1972; Attardi & Amaldi, 1970). Rabbit L-rRNA (67% G + C) contains long tracts (approx. 0.4×10^6 daltons mass) that are G + C-rich (approx. 78% G + C) and which amount to approx. 40% of the total mass. This clustering of guanine and cytosine residues was revealed by spectrophotometric measurements (Cox, 1966*a*, 1970) and by partial nuclease-digestion studies (for review, see Cox *et al.*, 1973*a*).

The aim of the present work is to show that L-rRNA with a low percentage of G + C isolated from a protozoan (*Plasmodium knowlesi*) and from *Drosophila melanogaster* both have clusters of adenine and uracil residues, as judged by spectrophotometric measurements. A similar result was found for S-rRNA (the rRNA species of the smaller subparticle). Although rabbit L-rRNA differs appreciably from L-rRNA of *P. knowlesi* and *D. melanogaster* in percentage content of G + C (67, 37 and 42% respectively) all three species have a 'core' of 0.5×10^6–0.7×10^6 daltons mass with a nucleotide composition of 50–55% G + C.

The evolutionary significance of these findings is discussed on the basis of the notion that L-rRNA comprises a region of approx. 0.5×10^6–0.7×10^6 daltons mass that has been conserved throughout the evolution of eukaryotes, whereas the remainder of the molecule has diverged to a much greater extent (Cox *et al.*, 1973*a*; Godwin *et al.*, 1974).

Theoretical

In principle three kinds of information may be derived from 'melting' studies. These are (i) the fraction of residues involved in secondary structure; (ii) the nucleotide composition of the bihelical parts; (iii) an estimate of N, the number of base pairs per bihelical segment. The basis of the method has been described earlier (Cox, 1970).

Fraction of residues involved in secondary structure

The increase in extinction found for the transition from the native molecule to a single-stranded amorphous form depends mainly on the fraction of rA·rU and of rG·rC base pairs and to a lesser extent on the fraction of residues that 'stack' in single-stranded regions. On the basis of reference data for 'melting' rA·rU and rG·rC base pairs, the sum of the fraction of base-paired residues (f_{bh}) and the fraction of 'stacked' residues (f_s) may be calculated. The data used are $\Delta\varepsilon_{260} = 4500\ l \cdot mol^{-1} \cdot cm^{-1}$, $\Delta\varepsilon_{280} = 0\ l \cdot mol^{-1} \cdot cm^{-1}$ for the 'melting' of an A·U base pair, and $\Delta\varepsilon_{260} = 1500\ l \cdot mol^{-1} \cdot cm^{-1}$, $\Delta\varepsilon_{280} = 3800\ l \cdot mol^{-1} \cdot cm^{-1}$ for 'melting' a G·C base pair.

Estimation of the nucleotide composition of bihelical regions

The difference in the spectrum of native rRNA and the amorphous single-stranded form can be fitted by two standard spectra, one representing the 'melting' of an rA·rU base pair, and the other the 'melting' of an rG·rC base pair. There is an empirical relation between the increments found in E at 260 and 280 nm, ΔE_{260} and ΔE_{280} respectively, and the nucleotide composition of the base-paired regions 'melting' over the temperature range under study (see eqn. 1).

$$\Delta E_{260}/\Delta E_{280} = 1.38\,(\delta f_{A\cdot U}/\delta f_{G\cdot C}) + 0.45 \qquad (1)$$

In eqn. (1) $\delta f_{A\cdot U}$ and $\delta f_{G\cdot C}$ are the mole fractions of rA·rU and rG·rC base pairs respectively 'melting' over the temperature range over which E_{260} increased by ΔE_{260} and E_{280} increased by ΔE_{280}. The ratio $\Delta E_{260}/\Delta E_{280}$ provides a convenient index for the differential melting of rA·rU and rG·rC base pairs, because the melting of an rA·rU base pair has little effect on E_{280}, whereas the increment in E is maximum at this wavelength for 'melting' an rG·rC base pair; 'melting' of rA·rU and rG·rC base pairs both lead to an increase in E_{260}.

Estimation of N, the number of base pairs per bihelical segment

The 'melting' temperature, T_m, is defined as the midpoint of the transition from a bihelical to an amorphous form. When $N < 20$, the T_m is less than the T_m [$T_m(\infty)$] of a bihelix of equivalent nucleotide composition but of 'infinite' length. Ignoring imperfections in the short bihelical regions and the effect of non-base-paired residues, T_m, N and $T_m(\infty)$ are related by eqn. (2) (see Cox, 1970, for discussion).

$$T_m = T_m(\infty) - 172/N \qquad (2)$$

Experimental

Isolation of rRNA from D. melanogaster

Whole flies were homogenized in a phenol/water mixture (Hastings & Kirby, 1966) and the rRNA fraction was isolated by the phenol method as described by Hastings & Kirby (1966). The rRNA

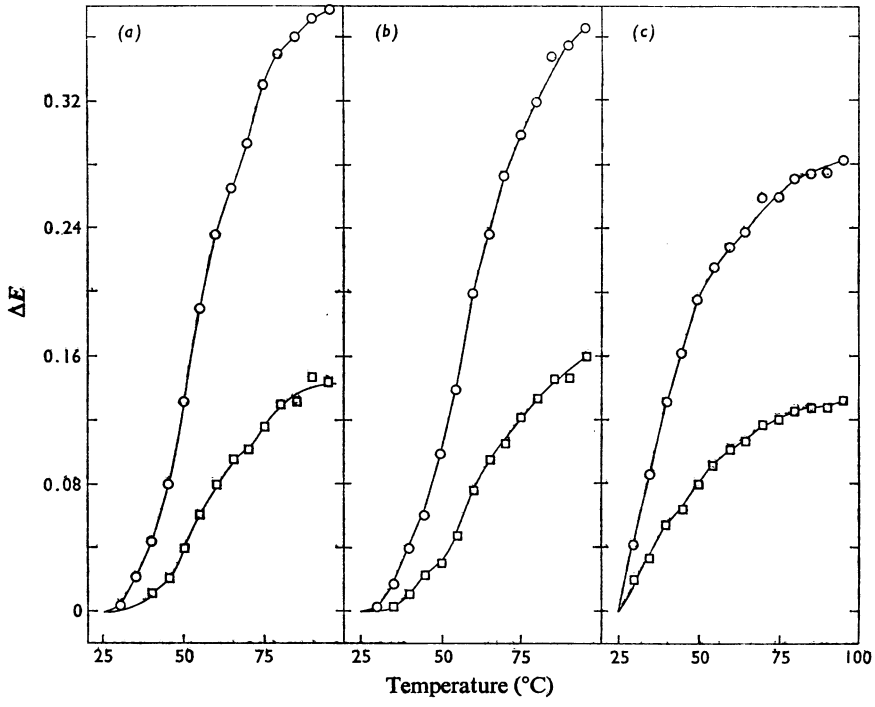

Fig. 1. *'Melting' profiles of D. melanogaster S-rRNA*

(*a*) Solvent 0.1 mM-$MgCl_2$/0.01 M-Tris/HCl, pH 7.6; (*b*) solvent A4 (0.25 M-sucrose, 0.025 M-KCl, 1 mM-$MgCl_2$, 0.05 M-Tris/HCl, pH 7.6). Ribosomes and subparticles retain their activity in both solvents (*a*) and (*b*) (cf. Cox *et al.*, 1973*b*); (*c*) solvent 0.01 M-potassium phosphate buffer, pH 7.0. ○, ΔE_{260}; □, ΔE_{280}. In each case $E_{260} = 1.0$ at 25°C. ΔE is the difference in the extinction of a treated sample and a sample kept at 25°C (see the Experimental section for details).

fraction was separated into S-rRNA and L-rRNA by zone centrifuging. Polyacrylamide-gel electrophoresis by the method of Loening (1968) confirmed that the fractions were homogeneous.

Isolation of rRNA from P. knowlesi.

The malaria parasite *P. knowlesi* was isolated from rhesus monkeys. The parasites were isolated free of host-cell material by a modification (Shakespeare & Trigg, 1973) of the method of Williamson & Cover (1967). Ribosomes were isolated by the method of Arnstein *et al.* (1964) and subparticles were prepared by treatment with heparin as described by Pratt & Cox (1971). Isolation of rRNA from other sources was described by Cox (1966*b*).

Nucleotide analysis

The nucleotide composition of *P. knowlesi* was measured by the method of Katz & Comb (1963).

Spectrophotometry

A Unicam SP.700 instrument was used. The sample-cell holder was an electrically heated copper block. In each case the difference in the spectrum between a heated sample and a sample kept at 25°C was measured over the range 210–330nm. The extinction measured above 25°C was corrected for changes in the concentration of RNA due to the thermal expansion of water.

Results

Both S-rRNA and L-rRNA are believed to comprise a large number of hairpin loops stabilized by base-paired residues (Cox, 1966*a*) and the 'melting' profile is the sum of the 'melting' of all these individual bihelical regions and also includes a contribution from residues that are 'stacked' in single-stranded regions. For L-rRNA of 1.3×10^6–1.7×10^6 daltons mass there could be 80–120 hairpin loops (Cox, 1970), each stabilized by at least one bihelical region. Each bihelical region has a characteristic T_m governed mainly by its nucleotide composition (i.e. by percentage of G·C base pairs in the bihelical segment) and by N, the number of base pairs/bihelical segment. The increments $\Delta\varepsilon_{260}$ and $\Delta\varepsilon_{280}$ found on heating over a particular 'melting' range reflects the fraction of base-paired residues 'melting' over this range, and

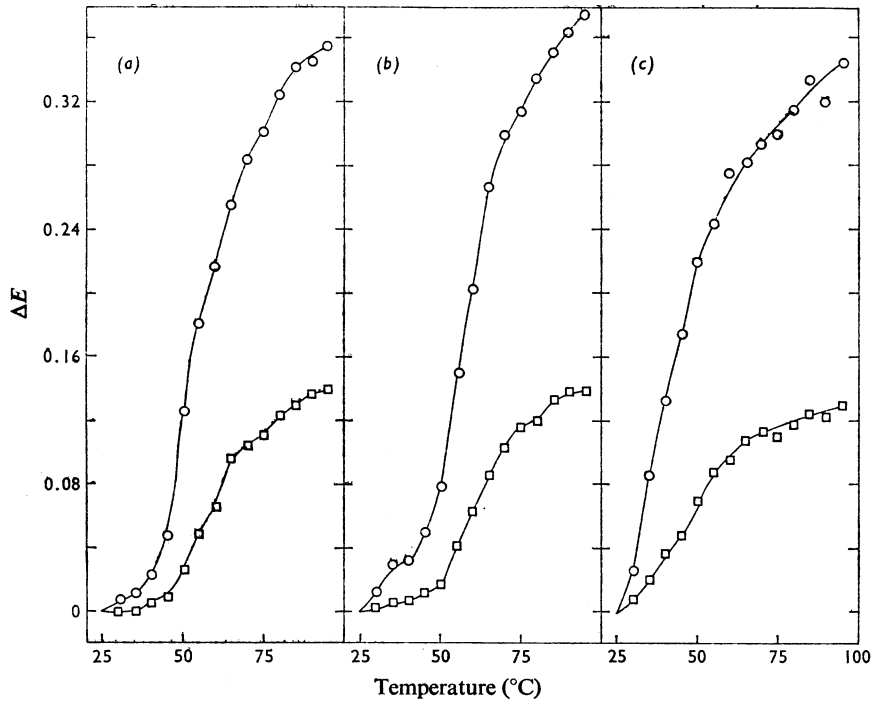

Fig. 2. *'Melting' profiles of D. melanogaster L-rRNA*

(*a*) Solvent 0.1mM-$MgCl_2$/0.01M-Tris/HCl, pH7.6; (*b*) solvent 0.25M-sucrose/0.025M-KCl/0.05M-Tris/HCl, pH7.6; (*c*) solvent 0.01M-potassium phosphate buffer, pH7.0. ○, ΔE_{260}; □, ΔE_{280}. In each case $E_{260} = 1.0$ at 25°C.

Table 1. *Summary of the 'melting' properties of D. melanogaster and P. knowlesi rRNA*

RNA species	Solvent	'Melting' range (°C)	T_m measured at 260 and 280 nm (°C)	% rG·rC base pairs of bihelical parts (eqn. 1)	Approx. f_{bh}	Approx. N (eqn. 2)	10^{-6} × Approx. mass of rRNA (daltons)
D. melanogaster S-rRNA	0.1 mM-$MgCl_2$/0.01 M-Tris/HCl, pH7.6	25–95	54–58	38	0.8		0.7
		25–50		33	0.41	6	0.36
		60–95		48	0.32	4–10	0.28
	A4 (see legend to Fig. 1)	25–90	58–60	38.5	0.8		0.7
		25–50		32	0.21	6	0.18
		60–90		48	0.39	4–10	0.34
	0.01 M-potassium phosphate buffer, pH7.0	35–80	41–41	40	0.62		0.7
		30–45		37	0.41	6	0.27
		45–80		51	0.27	4–10	0.28
D. melanogaster L-rRNA	0.1 mM-$MgCl_2$/0.01 M-Tris/HCl, pH7.6	25–95	55–60	38	0.77		1.4
		25–50		24	0.25	6	0.5
		60–95		50	0.35	4–10	0.6
	A4	25–95	58–62	39	0.81		1.4
		25–55		29	0.32	6	0.6
		65–95		51	0.33	4–10	0.6
	0.01 M-potassium phosphate buffer, pH7.0	25–70	46–49	39			1.4
		25–40		30	0.28	6	0.5
		45–70		53	0.29	4–10	0.5
P. knowlesi S-rRNA	0.01 M-potassium phosphate buffer, pH7.0	25–90	39–48	43	0.67		0.7
		25–40		25	0.3	6	0.26
		40–90		55	0.38	4–10	0.33
P. knowlesi L-rRNA	1 mM-$MgCl_2$/0.01 M potassium phosphate buffer, pH 7.0	25–95	58–62	42.5	0.85		1.4
		25–55		33	0.3	6	0.52
		60–95		49	0.4	4–10	0.7
	0.01 M-potassium phosphate buffer, pH7.0	25–85	40–44	45	0.64		1.4
		25–35		38	0.18	6	0.31
		40–85		49	0.31	4–10	0.54

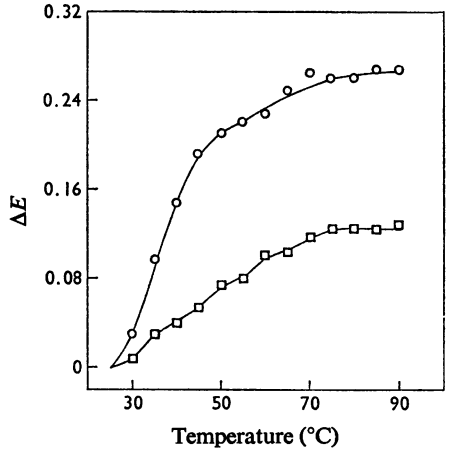

Fig. 3. *'Melting' properties of S-rRNA of P. knowlesi*

Solvent 0.01 M-potassium phosphate buffer, pH7.0. ○, ΔE_{260}; □, ΔE_{280}: $E_{260} = 1.0$ at 25°C.

the corresponding ratio $\Delta\varepsilon_{260}/\Delta\varepsilon_{280}$ (or $\Delta E_{260}/\Delta E_{280}$) is a measure of the nucleotide composition of all the bihelical regions 'melting' over this range. For rRNA of approx. 40% G + C a value of $\Delta E_{260}/\Delta E_{280} \simeq 2.5$ would be predicted for most of the 'melting' range provided that the nucleotide sequence is near-to-random. A marked deviation from this predicted behaviour indicates the presence of clusters of particular nucleotides. The major feature of the 'melting' of both S-rRNA and L-rRNA of *P. knowlesi* and *D. melanogaster* is the evidence that a high proportion of adenine and uracil residues are clustered together.

'Melting' properties of rRNA of D. melanogaster

The data of *D. melanogaster* S-rRNA and L-rRNA are summarized in Figs. 1 and 2 respectively. In 0.01 M-potassium phosphate buffer, pH7, part of the native structure had already 'melted' at 25°C. However, the complete transition was observed when

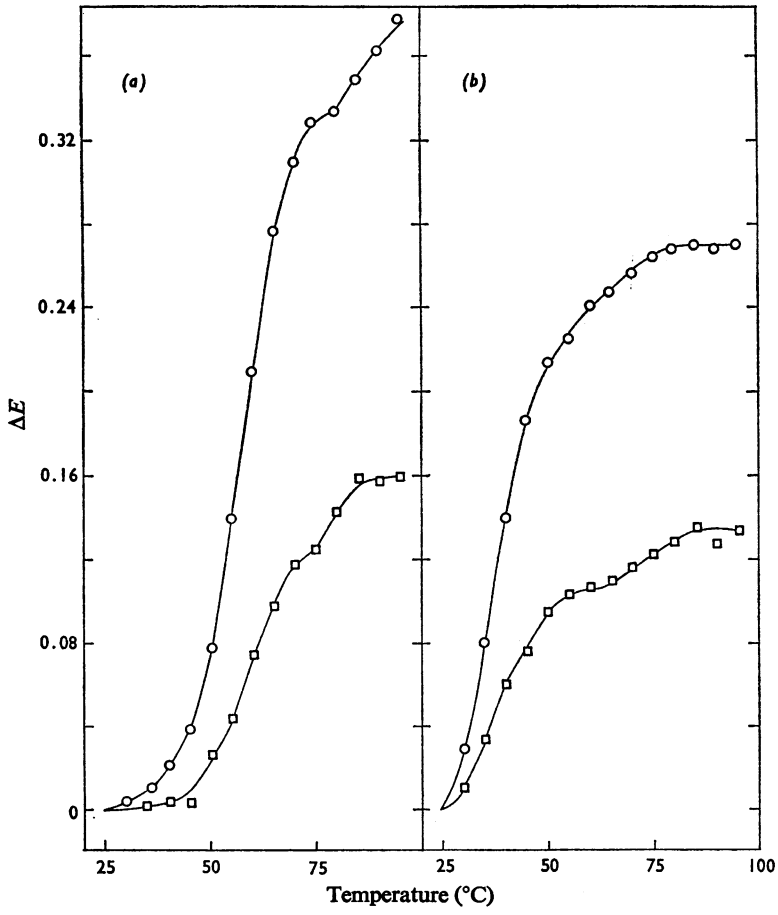

Fig. 4. *'Melting' properties of L-rRNA of P. knowlesi*

(*a*) Solvent 1 mM-$MgCl_2$/0.01 M-potassium phosphate buffer, pH 7.0; (*b*) solvent 0.01 M-potassium phosphate buffer, pH 7.0. ○, ΔE_{260}; □, ΔE_{280}. In both cases $E_{260} = 1.0$ at 25°C.

the solution was made 0.1 mM-$MgCl_2$ and then the ratio $\Delta E_{260}/\Delta E_{280} \simeq 2.5$ was found for the temperature range 25–94°C. This value is close to that predicted for 40% G·C base pairs of bihelical parts. However, the 'melting' profile of both S-rRNA and L-rRNA is biphasic. For L-rRNA, ΔE_{260} increased far more rapidly than ΔE_{280} over the range 25–50°C ($\Delta E_{260}/\Delta E_{280} = 4.7$), whereas above 60°C $\Delta E_{260}/\Delta E_{280}$ was 1.9. It was estimated by means of eqn. (1) and by using the data of the standard reference spectra for rA·rU and rG·rC base pairs (Cox, 1970) that at least 25% of residues are located in bihelical regions of approx. 24% G·C base pairs, and at least 35% of residues form bihelical regions of approx. 50% G·C base pairs. Biphasic 'melting' was also reported for A + U-rich rRNA of the mitochondria of fungi (Edelman *et al.*, 1970, 1971) and for G + C-rich mammalian L-rRNA (Cox, 1966*a*, 1970).

Fractions of residues forming base pairs in the native state

The total increment in E_{260} (approx. 36%) is approx. 80% of the increase noticed for bihelical RNA of comparable nucleotide composition (Cox *et al.*, 1970). This finding suggests that approx. 80% of residues participate in organized secondary structure comparable with estimates for *Escherichia coli* S-rRNA and L-rRNA. Analysis of the c.d. (circular-dichroism) spectrum (Cox *et al.*, 1976) showed that approx. 65% of residues appear to form base pairs and approx. 15% of residues tend to 'stack' in single-strand regions.

Length, N, of independently 'melting' bihelical regions

The A·U-rich bihelical regions 'melt' first at about 40°C below the T_m of their bihelical analogue.

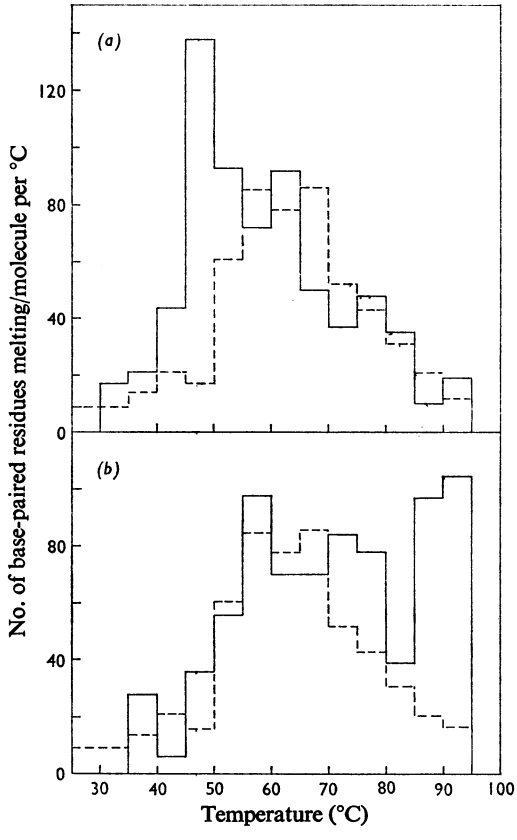

Fig. 5. *Comparison of the differential 'melting' properties of several L-rRNA species*

The nucleotide composition of bihelical regions 'melting' over a 5°C range was calculated by means of eqn. (1); $f_{A\cdot U}$ and $f_{G\cdot C}$ respectively are the mole fraction of A·U and G·C base pairs 'melting' over the temperature range, were calculated from the eqn.

$$f_{bh} = f_{A\cdot U} + f_{G\cdot C} = \varepsilon_{260} \times \Delta\varepsilon_{260}/(4500\, f_{A\cdot U} + 1500\, f_{G\cdot C})$$

Hence no. of base-paired residues 'melting'/molecule per °C is equal to

$$f_{bh} \times (MW/333)/\delta T(°C)$$

where MW is the molecular weight of the particular L-rRNA (see Table 2) and 333 is taken as the average mass of a nucleotide; δT(°C) is the temperature range studied (5°C in this case). The solvent was 0.1 mM-$MgCl_2$/0.01 M-Tris/HCl, pH7.6 (cf. Cox *et al.*, 1973*b*). The Figure illustrates the possibility of there being features of secondary structure (namely 50–55% G·C, T_m 50–75°C) common to all three species. (*a*) – – – –, *E. coli* L-rRNA; ——, *D. melanogaster* L-rRNA. (b) – – – –, *E. coli* L-rRNA; ——, rabbit reticulocyte L-rRNA.

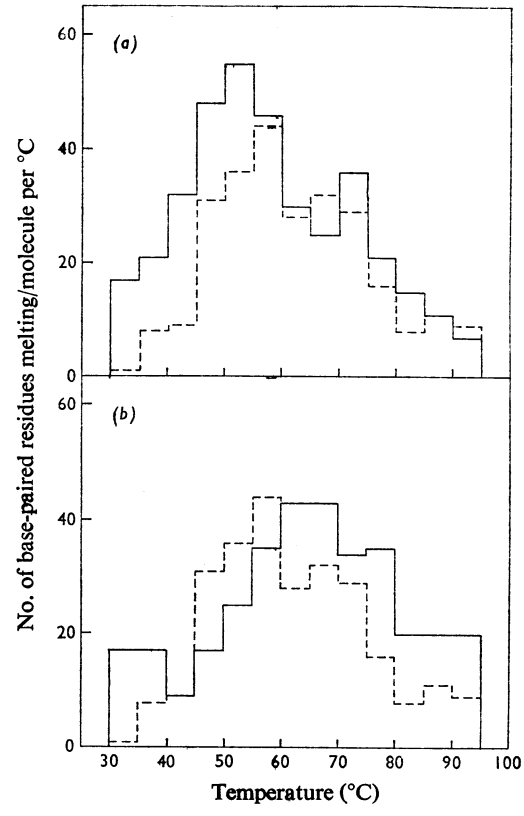

Fig. 6. *Comparison of the differential 'melting' profiles of several S-rRNA species*

The number of base-paired residues 'melting' in a single molecule/°C rise in temperature was calculated by means of the expression:

$$\frac{f_{bh}\cdot MW}{333\cdot\delta T}$$

where MW/333 = no. of nucleotides/molecule and f_{bh} = fraction of residues involved in base-pairing, 'melting' over an increase in temperature of δT. f_{bh} was calculated for 5°C increments in temperature from 25°C to 95°C. The solvent was 0.1 mM-$MgCl_2$/0.01 M-Tris/HCl, pH7.6. (*a*) – – – –, *E. coli* S-rRNA; ——, *D. melanogaster* S-rRNA. (*b*) – – – –, *E. coli* S-rRNA; ——, rabbit reticulocyte S-rRNA.

According to eqn. (2) this signifies that either N, the number of base pairs per bihelical segment, is unlikely to be much greater than six or so, or that the bihelical parts contain imperfections. Similar conclusions were reached for the value of N for the bihelical domains of approx. 50% G·C base pairs (see Table 1).

Melting properties of P. knowlesi rRNA

The properties of this species were not studied in the same detail as were those of *D. melanogaster*, but

evidence for A·U-rich bihelical regions was noted in both cases, although the phenomenon was less pronounced in the protozoa (see Figs. 3 and 4). This does not necessarily mean that there is less tendency for adenine and uracil residues to cluster in *P. knowlesi* than in *D. melanogaster* rRNA. The 'melting' properties of the two species are summarized in Table 1.

Comparison of the 'melting' of D. melanogaster and other rRNA species

The differential 'melting' properties of the L-rRNA species of *D. melanogaster* and rabbit reticulocytes are compared in Fig. 5 with those of *E. coli* L-rRNA. The principal result is that in each case there are a number of common features, thus: the number of base pairs/molecule 'melting' over the temperature range 50–75°C is comparable in each case, the percentage of G·C base pairs is approx. 50–55 over this same temperature range, and by inference *N* is likely to be comparable in each case. Although differential melting of A·U and G·C base pairs is apparent for prokaryotic L-rRNA, the trend is much more pronounced in the other species. The evidence reveals that approx. 0.5×10^6 daltons mass of these L-rRNA species comprise features of bihelical secondary structure that have much in common (cf. Godwin *et al.*, 1974). Similarly approx. 0.3×10^6 daltons mass of the S-rRNA species studied appear to have features of secondary structure in common (Fig. 6). This suggests that particular features of secondary structure have been conserved throughout evolution, possibly because they are important to ribosome function. This notion is elaborated in the Discussion section.

Discussion

The main result of this study is that the adenine and uracil residues of the A + U-rich rRNA species studied are unevenly distributed throughout the S-rRNA and L-rRNA molecules. The analysis of the 'melting' profiles of partly bihelical molecules the size of rRNA is a technique that is more akin to the bludgeon than the rapier because there are probably

Table 2. *Comparison of the properties of L-rRNA of different species illustrating evidence for sequence homologies as well as for evolutionary divergence*

References: [1] Lava-Sanchez *et al.* (1972); [2] Loening (1968); [3] Martin *et al.* (1970); [4] Sinclair & Brown (1971): these authors found that rDNA–rRNA hybrids were formed between *X. laevis* rDNA and rRNA of a wide range of eukaryotes including *D. melanogaster*.

	Evidence for evolutionary divergence of L-rRNA						Evidence for conserved sequences	
	Nucleotide composition (mol %)[1]					10^{-6} × Mol. wt.[2]	Observed homology (%) with *S. cerevisiae*[3] L-rRNA	10^{-6} × Mass of L-rRNA homologous with *S. cerevisiae* L-rRNA (daltons)
Species	C	A	U	G	G+C			
Fungi								
S. cerevisiae	19.2	26.4	26.0	28.4	47.6	1.3	100	1.3
N. crassa	21.9	24.8	23.9	29.4	51.3	1.3	73	0.95[4]
Plants								
Pea	22.6	23.6	21.6	31.2	54.7	1.27–1.31	50	0.65
Wheat	28.0	21.2	16.9	33.7	61.7	1.27–1.31	48	0.63
Protozoa								
T. pyriformis	16.1	34.7	32.1	16.8	32.9	1.3	38	0.49[4]
P. knowlesi	14.4	32.6	30.9	22.8	37.2	1.3	*	*
Arthropoda								
D. melanogaster	19.6	30.8	27.1	22.5	42.1	1.4	†[4]	†[4]
Amphibia								
X. laevis	27.9	19.7	17.4	35.0	62.9	1.5	†[4]	†[4]
Mammals								
Mouse	27.8	19.5	18.2	34.5	62.3	1.7	40	0.52
Rabbit	31.7	16.9	15.5	35.9	67.6	1.7	‡	‡

* To be compared with data for *T. pyriformis*.
† rRNA forms hybrid with *X. laevis* rDNA.
‡ To be compared with data for mouse rRNA.

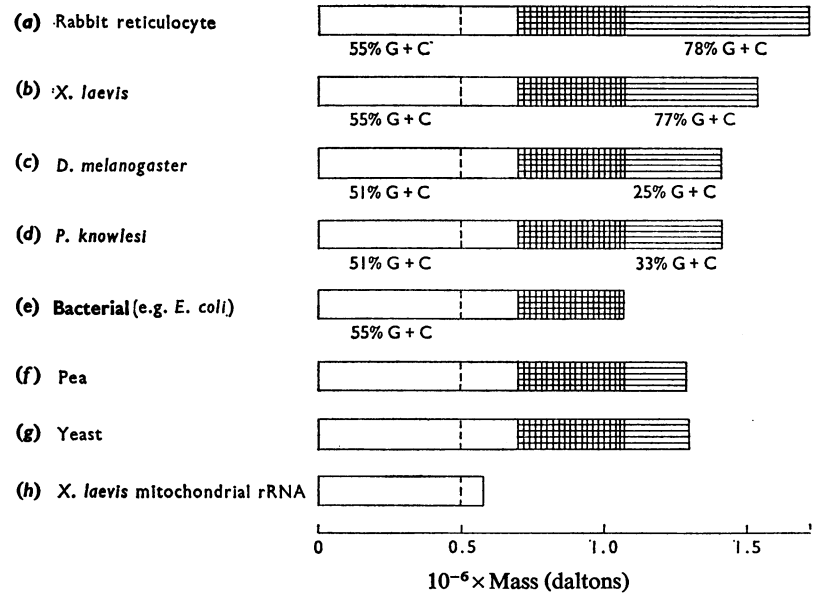

Fig. 7. *Summary of the properties of L-rRNA and their possible evolutionary significance*

(*a*) Schematic presentation of rabbit reticulocyte L-rRNA based on analysis of the 'melting' profile and on analysis of the products of partial digestion with T_1 ribonuclease. (*b*) Data for *X. laevis* L-rRNA based on analysis of the 'melting' profile and on analysis of the products of partial T_1 ribonuclease digestion. The scheme does not conflict with the secondary-structure maps of Wellauer & Dawid (1973). (*c*) Scheme for *D. melanogaster* based on analysis of the 'melting' profile. (*d*) Scheme for *P. knowelsi* based on analysis of the 'melting' profile. (e) Scheme for *E. coli* L-rRNA based on the analysis of the 'melting' profile. □, 'Conserved', indicates the region of L-rRNA that apparently has features of secondary structure in common irrespective of its origin. ▦, Indicates the uncertainty in the estimation of the size of the conserved regions. ▤, Non-conserved, this denotes that part of the RNA moiety that may vary from one class to another, yet may be similar for species within a single class. (*f*) and (*g*) Proposed scheme for pea and yeast L-rRNA, based on the general hypothesis outlined above. (*h*) *X. laevis* mitochondrial L-rRNA. This species is included for comparison. According to our tentative hypothesis most of this species should form secondary structure corresponding to the conserved regions in eukaryotic cytoplasmic L-rRNA and prokaryotic L-rRNA, although there are no data relating to the secondary structure of this species.

so many overlapping transitions that any differential 'melting' that is observed is likely to arise from major structural features. We conclude that rA + rU-rich regions are a major structural feature of rRNA of *P. knowlesi* and *D. melanogaster*. The results (e.g. Figs. 5 and 6) further support the notion that as both S-rRNA and L-rRNA evolved, a substantial part appears to have been conserved while another part has diverged more rapidly.

Possible evolutionary significance of the structure of L-rRNA revealed by spectrophotometric measurements

Table 2 summarizes the evidence that cytoplasmic L-rRNA has not only diverged in some respects (e.g. mass and overall nucleotide composition) during evolution but also has conserved sequences as judged by hybridization studies (Martin *et al.* 1970; Sinclair & Brown, 1971; Birnstiel & Grunstein, 1972). Sinclair & Brown (1971) found that *Xenopus laevis* rDNA would hybridize with the rRNA of rat, *Tetrahymena pyriformis* (33% G + C), *P. knowlesi*, *Neurospora crassa*, spinach and *D. melanogaster*. Birnstiel & Grunstein (1972) showed by hybridization with excess of DNA that highly conserved sequences in *X. laevis* rRNA and HeLa rRNA were also present in sea-urchin rRNA.

The data obtained from the spectrophotometric studies, which are summarized in Fig. 7, provide an estimate of the minimum size of the conserved regions (0.5×10^6–0.7×10^6 daltons mass) and also suggest that these sequences have an overall nucleotide composition of approx. 50–55% G + C. Among the species studied the wide variation in both the size and nucleotide composition of the non-conserved sequences (approx. 24% G + C in *D. melanogaster* and approx. 78% G + C in rabbit) is also revealed.

In mammals the G + C-rich sequences are clustered

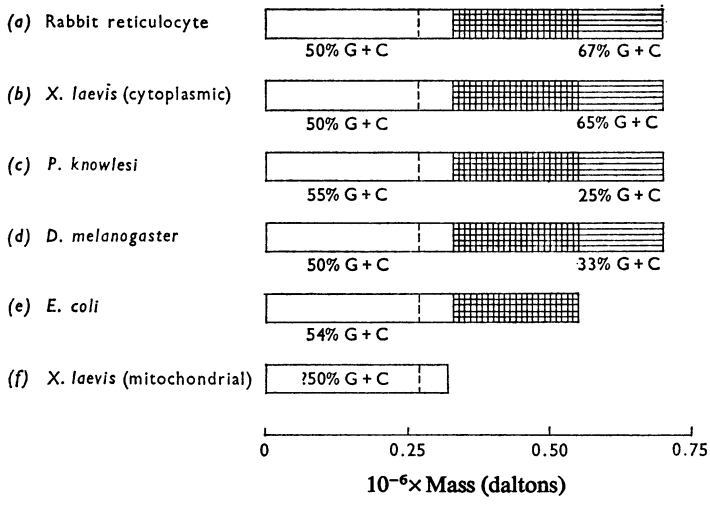

Fig. 8. *Summary of properties of S-rRNA and their possible evolutionary significance*

Schematic presentation based on analyses of 'melting' profiles for: (*a*) rabbit S-rRNA (e.g. Cox, 1970); (*b*) *X. laevis* S-rRNA; (*c*) *P. knowlesi* S-rRNA (see Table 1); (*d*) *D. melanogaster* S-rRNA (see Table 1); (*e*) *E. coli* S-rRNA (the analysis reveals that the percentage of G·C base pairs of bihelical region ranges from approx. 50–60% (cf. Cox, 1966*a*); (*f*) *X. laevis* mitochondrial rRNA. □, Conserved region; ▦, indicates uncertainty in the estimates of the conserved and non-conserved regions; ▤, non-conserved region.

in a few regions as judged by the products of T_1 ribonuclease partial digests (Cox *et al.*, 1973*a*) and by electron micrographs of partly denatured rRNA (Wellauer & Dawid, 1973; Schibler *et al.*, 1975). The tendency to form long stretches of A + U-rich sequences is also apparent in *D. melanogaster* rRNA. It was reported previously (Kirby *et al.*, 1962; Hastings, 1964) that fragments of *D. melanogaster* r-RNA (unfractionated), which sedimented in the range 4–15S (mainly 6–8S, approx. 100000 daltons mass), could be separated by counter-current distribution into fractions ranging in nucleotide composition from 28% G + C to 54% G + C. These data not only support our analysis of the 'melting' profiles but also suggest that the tracts of base compositions approx. 28% G + C and approx. 54% G + C extend to at least approx. 300 nucleotides.

It appears that the nucleotide composition of the 'non-conserved' region is similar to that of those regions of pre-rRNA that are lost during processing. The regions removed from mammalian and *X. laevis* pre-rRNA during the maturation process are G + C-rich (Birnstiel & Grunstein, 1972; Wellauer & Dawid, 1973; Schibler *et al.*, 1975) and are comparable with the G + C-rich regions of the mature L-rRNA. The evidence for A + U-rich tracts in *D. melanogaster* L-rRNA is in accord with Greenberg's (1969) observation that its pre-rRNA has a high A + U content.

The estimated mass of the conserved region approaches the mass of the RNA moieties of the smallest ribosomes, the ribosomes of vertebrate mitochondria (Bartoov *et al.*, 1970; Dawid & Chase, 1972). We infer that the 'conserved' sequences of L-rRNA are likely to be essential to the active centre of the subparticle because the same basic mechanism for the sequential formation of a peptide bond applies to both eukaryotes and prokaryotes (for review, see Cox & Godwin, 1975). The functions that appear common to all species include the binding of the smaller subparticle, peptidyl-tRNA, aminoacyl-tRNA, and peptidyl-transferase and -translocase activities. There is no evidence for extensive sequence homologies between L-rRNA of eukaryotes and bacteria or between mitochondrial L-rRNA and cytoplasmic L-rRNA. However, the possibility that these are homologous features of secondary structure remains.

Evolution of S-rRNA

Both S-rRNA and L-rRNA appear to have evolved in parallel (see, for example, Birnstiel & Grunstein, 1972) as may be expected because both species are derived from the same pre-rRNA (see, for example, Wellauer & Dawid, 1973). Although the features of a conserved part and non-conserved part are more pronounced for L-rRNA, the same trends are found

Table 3. *Comparison of the properties of S-rRNA of different species illustrating evidence for sequence homologies as well as for evolutionary divergence*

References: [1] data from review by Lava-Sanchez *et al.* (1972); [2] data from review by Cox (1968); [3] Martin *et al.* (1970); [4] Sinclair & Brown (1971). † The mol. wt. of these S-rRNA species is 0.7×10^6 as determined by polyacrylamide-gel electrophoresis (Loening, 1968).

	Evidence for evolutionary divergence of S-rRNA*					Evidence for conserved sequences	
	Nucleotide composition (mol %)[1,2]					Observed homology (%) with *S. cereviseae* S-rRNA	$10^{-6} \times$ Mass of S-rRNA homologous with *S. cereviseae* S-rRNA (daltons)
Species	C	A	U	G	G+C		
Fungi							
S. cereviseae	26.5	23.5	25.0	25.0	51.5	100	0.7
Plants							
Pea	20.1	23.7	25.1	31.1	51.2	63	0.44
Wheat	25.0	22.0	20.2	32.8	57.8	61	0.43
Protozoa							
T. pyriformis	15.8	31.9	31.7	20.5	36.3	48	0.34
P. knowlesi	12.5	29.5	36	25	37.5	†	†
Arthropoda							
D. melanogaster	20.3	28.8	27.4	23.5	43.8	*[4]	*[4]
Amphibians							
X. laevis	24.1	24.1	22.9	28.9	53	*[4]	*[4]
Mammals							
Mouse	24.7	23.5	29.9	21.9	54.8	42	0.29
Rabbit	28.8	20.5	30.7	20.0	59.3	‡	‡

* Evidence for formation of rRNA·rDNA hybrid between *X. laevis* rDNA and *D. melanogaster* rRNA.
† Cf. data for *T. pyriformis*, which may be regarded as an analogue.
‡ Cf. data for mouse S-rRNA, which is likely to be closely related to rabbit S-rRNA.

in S-rRNA (for summary, see Fig. 8). The evidence for evolutionary divergence and for conserved sequences among eukaryotes based on hybridization studies is summarized in Table 3. The sequence close to the 3′-hydroxyl terminus of eukaryotic S-rRNA is apparently conserved (Dalgarno & Shine, 1973), e.g. yeast, *D. melanogaster* and rabbit S-rRNA all terminate with the sequence pGpApUpCpApUpAOH. The estimate of the mass (approx. 0.3×10^6 daltons) of the conserved sequences approaches that of the *X. laevis* mitochondrial S-rRNA. There are no data relating to the secondary structure of *X. laevis* mitochondrial S-rRNA, but the known nucleotide composition allows the possibility that most of the molecule has a nucleotide composition of approx. 50% G + C. Thus there remains the possibility that there are common features of secondary structure between the mitochondrial and conserved parts of cytoplasmic rRNA. This is to be expected if certain regions of rRNA are of particular importance and so would be more highly conserved than others. The function of the non-conserved sequences remains a mystery.

We are most grateful to Dr. O. G. Fahmy and Dr. M. J. Fahmy, Pollards Wood Research Station, Chester Beatty Research Institute, for kindly supplying the *Drosophila* flies.

References

Arnstein, H. R. V., Cox, R. A. & Hunt, J. A. (1964) *Biochem. J.* **92**, 648–661
Attardi, G. & Amaldi, F. (1970) *Annu. Rev. Biochem.* **39**, 183–226
Bartoov, B., Mitra, R. S. & Freeman, K. B. (1970) *Biochem. J.* **120**, 455–466
Birnstiel, M. L. & Grunstein, M. (1972) *FEBS Symp.* **23**, 349–365
Cox, R. A. (1966*a*) *Biochem. J.* **98**, 841–857
Cox, R. A. (1966*b*) *Biochem. Prep.* **11**, 104–109
Cox, R. A. (1968) *Q. Rev. Chem. Soc.* **22**, 499–525
Cox, R. A. (1970) *Biochem. J.* **117**, 101–118
Cox, R. A. & Godwin, E. A. (1975) *Biochem. Ser. One* **7**, 179–253
Cox, R. A., Kanagalingam, K. & Sutherland, E. (1970) *Biochem. J.* **120**, 549–558
Cox, R. A., Huvos, P. & Godwin, E. A. (1973*a*) *Isr. J. Chem.* **11**, 407–422
Cox, R. A., Pratt, H., Huvos, P., Higginson, B. & Hirst, W. (1973*b*) *Biochem. J.* **134**, 775–793
Cox, R. A., Hirst, W., Godwin, E. A. & Kaiser, I. (1976) *Biochem. J.* **155**, 279–295
Dalgarno, L. & Shine, J. (1973) *Nature (London) New Biol.* **245**, 261–262
Dawid, I. B. & Chase, J. W. (1972) *J. Mol. Biol.* **63**, 217–231
Edelman, M., Verma, I. M. & Littauer, U. Z. (1970) *J. Mol. Biol.* **49**, 67–83

Edelman, M., Verma, I. M., Saya, D. & Littauer, U. Z. (1971) *Biochem. Biophys. Res. Commun.* **42**, 208–213
Godwin, E. A., Cox, R. A. & Huvos, P. (1974) *Acta Biol. Med. Germ.* **33**, 733–752
Greenberg, J. R. (1969) *J. Mol. Biol.* **46**, 85–98
Hastings, J. R. B. (1964) Ph.D. Thesis, University of London
Hastings, J. R. B. & Kirby, K. S. (1966) *Biochem. J.* **100**, 532–539
Katz, S. & Comb, D. G. (1963) *J. Biol. Chem.* **238**, 3065–3067
Kirby, K. S., Hastings, J. R. B. & O'Sullivan, M. A. (1962) *Biochim. Biophys Acta* **61**, 978–979
Lava-Sanchez, P. A., Amaldi, F. & La Posta, A. (1972) *J. Mol. Ecol.* **2**, 44–45
Loening, U. E. (1968) *J. Mol. Biol.* **38**, 355–365
Martin, T. E., Bicknell, J. N. & Kumar, A. (1970) *Biochem. Genet.* **4**, 603–615
Pratt, H. & Cox, R. A. (1971) *Biochem. J.* **124**, 897–903
Schibler, U., Wyler, T. & Hagensbüchle, O. (1975) *J. Mol. Biol.* **94**, 503–517
Shakespeare, P. & Trigg, P. (1973) *Nature (London)* **241**, 213–233
Sinclair, J. H. & Brown, D. D. (1971) *Biochemistry* **10**, 2761–2769
Wellauer, P. K. & Dawid, I. B. (1973) *Proc. Natl. Acad. Sci. U.S.A.* **70**, 2827–2831
Williamson, J. & Cover, B. (1967) *Trans. R. Soc. Trop. Med. Hyg.* **60**, 425–427