RNA Molecular Weight Determinations by Gel Electrophoresis under Denaturing Conditions, a Critical Reexamination†

Hans Lehrach, Don Diamond, John M. Wozney, and Helga Boedtker*

ABSTRACT: RNA molecular weight measurements were carried out by gel electrophoresis under four different denaturing conditions including 99% formamide, 10 mM methyl mercury, 2.2 M formaldehyde, and 6 M urea at pH 3.8. Electrophoresis at a series of gel concentrations and at least two different voltage gradients resulted in some RNA species exhibiting apparent molecular weights that vary with both gel concentration and voltage gradient. Three different deviations from the requirement for hydrodynamically equivalent conformations were observed: (1) deformation of the random coil structure of very large RNAs at moderately high gel concentrations and voltage gradients resulting, in extreme cases, in a molecular weight independent migration of RNA molecules; (2) incomplete denaturation of RNA molecules with very GC rich helical regions; and (3) varying charge/mass ratio due to differential protonation at pH 3.8. Reliable molecular weight measurements of RNA molecules as large as 4.0×10^6 containing GC rich helical regions could only be made on dilute (0.5–1.0%) agarose gels after reaction with either 2.2 M formaldehyde or 10 mM methyl mercury hydroxide. A theoretical justification for the use of the empirical log molecular weight–mobility relation is presented. It is also demonstrated that the gel electrophoretic behavior of a homologous series of random coils can be approximated by that of a series of spheres with radii proportional to the square root of radius of gyration of a random coil. Consequently, molecular weight determinations of denatured RNAs, especially those obtained by extrapolation, are more reliable if the square root of the molecular weight is plotted vs. log mobility.

Gel electrophoresis has become the major analytical procedure for characterizing charged macromolecules both because of the high resolution it provides and the relative simplicity of the technology it requires. The applicability of this procedure to molecular weight determination of nucleic acids and $NaDodSO_4$–protein complexes is based both on the fact that these macromolecules are polymers with a constant charge-mass ratio and hydrodynamically equivalent conformations, and on the linearity of the empirical log molecular weight–mobility relation used to determine molecular weights from mobilities. When these assumptions are fulfilled, molecular weights can be measured in a single experiment using standards of known molecular weight. Moreover, since RNA molecules do not have hydrodynamically equivalent conformations in aqueous solutions (Boedtker, 1968; Groot et al., 1970; MacLeod, 1975), several methods were developed in which RNA molecular weights could be determined by gel electrophoresis under denaturing conditions. Denaturing conditions first used included reaction with formaldehyde (Boedtker, 1971), 8 M urea at 60 °C (Rejinders et al., 1973) and 99% formamide at room temperature (Pinder et al., 1974), the latter being by far the most widely used. In addition to the denaturing gel systems first developed, three others have been reported more recently: 6 M urea at pH 3.5 run at 2 °C (Rosen et al., 1975), 5 mM methyl mercury run at room temperature (Bailey and Davidson, 1976), and 99% formamide run at 45–55 °C (Spohr et al., 1976). This proliferation of denaturing gel systems arose because it was recognized that some of the denaturants first used were not able to denature very GC rich RNA and could not be used in the dilute agarose gels required for molecular weight determinations of very large RNA molecules.

We report here a comparison of the behavior of high molecular weight RNA molecules on four different denaturing gels: polyacrylamide in 99% formamide, agarose in 10 mM methyl mercury hydroxide, 2.2 M formaldehyde, and 6 M urea, pH 3.8. In addition, we offer a rationale for the use of the empirical log molecular weight–mobility relation and suggest an alternate method of obtaining RNA molecular weights from mobilities based on the finding that the migration of denatured RNA molecules depends on the square root of the radius of gyration, or the fourth root of the molecular weight.

Materials and Methods

RNA Samples. TMV and chicken ribosomal RNA were prepared as described previously (Boedtker, 1960; Boedtker et al., 1973). *E. coli* 23S and 16S rRNA was prepared by phenol–$CHCl_3$ extraction (Perry et al., 1972) followed by fractionation on linear 5–20% sucrose gradients.

Mouse 28S rRNA was donated by Jesse F. Scott of the Harvard Medical School. Sindbis virus, a gift from Michelene McCarthy of the Department of Biochemistry and Molecular Biology, was extracted with phenol–$CHCl_3$ (Perry et al., 1972).

Polyacrylamide Gel Electrophoresis in 99% Formamide. To achieve reproducible mobilities, pure reasonably dry deionized formamide is essential. 99% formamide (Eastman Chemical Co.) was deionized following the procedure of Pinder et al. (1974) by stirring with 40 g/L mixed bed resin (Bio-Rad AG501-X8, 20–50 mesh). After about 5 h, the conductivity should decrease to about 70 μmho. We have found, however, that some batches of formamide are not deionized under these conditions even when more resin is used. Such formamide is unsuitable for gel electrophoresis because irreproducible polymerization of gels occurs. The deionized formamide was filtered through a sintered-glass filter and then distilled under

† From the Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138. *Received March 30, 1977.* H. Lehrach was the recipient of a postdoctoral fellowship from the Jane Coffin Childs Memorial Fund for Medical Research. John M. Wozney was awarded a Camille and Henry Dreyfus Foundation Summer Research Fellowship. This investigation was supported by National Institutes of Health Grant HD-01229.

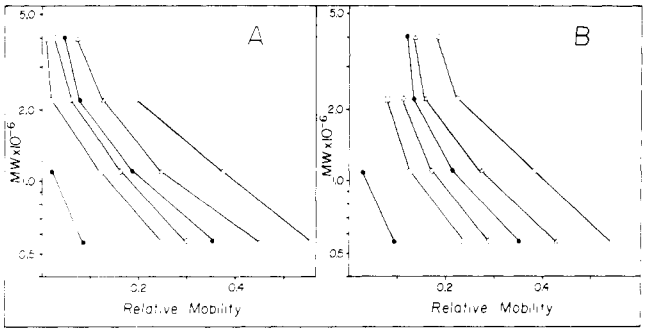

FIGURE 1: Log molecular weight–mobility relation of RNAs in polyacrylamide formamide gels as a function of gel concentration and voltage gradient. In order of decreasing molecular weight, RNAs analyzed were: Sindbis (4.0×10^6), TMV (2.2×10^6), 23S *E. coli* (1.1×10^6), and 16S *E. coli* (0.56×10^6). Acrylamide concentrations, from left to right: 5.0, 4.2, 3.8, 3.6, 3.4, and 3.2%. Relative mobilities were determined from the location of each band on photographs of the gels and then normalized to rate per voltage gradient given in arbitrary units. (A) Electrophoresis for 16 h at 0.8 mA/gel (2.2 V/cm). (B) Electrophoresis for 30 min at 1 mA/gel followed by 4.5 h at 2.5 mA/gel (7 V/cm).

a vacuum of 200 millitorr. Distilled formamide was stored at −20 °C in 50-mL aliquots.

Gels were polymerized using from 2.5 to 5.0% acrylamide (Canalco, Prep grade) and bisacrylamide equal to 17.5% of the acrylamide used. Both of the former were dissolved in formamide buffered with 0.01 M Na_2HPO_4–0.01 M NaH_2PO_4. The solution was then made 0.24% (v/v) *N,N,N′,N′*-tetramethylethylenediamine (TEMED).[1] Aqueous ammonium persulfate (0.2 mL of 18% (w/v)) was added to 25 mL of this solution; the solution was mixed thoroughly and then poured into 10 × 0.6 cm plexiglass tubes to a height of 7 cm (about 2.5 mL solutions per tube). Fifty microliters of 70% formamide was layered on top of each gel and the polymerization allowed to proceed overnight. Before electrophoresis, the bottom of each tube was covered with a small piece of nylon stocking to prevent the gel from slipping out and 0.5 mL of 70% buffered formamide was layered over the gel.

RNA samples were dissolved in buffered formamide, heated at 65 °C for 20 s, and fast cooled to room temperature, and 25–50 μL of a 1:1 mixture of glycerol–formamide containing 0.004% bromophenol blue was added to provide both a marker dye and to increase the density of the sample before it was layered on the gels. The tubes were then carefully filled with buffer.

Electrophoresis was carried out for 4 h at 2.75 mA per tube, or as indicated. The electrophoresis buffer was 0.01 M Na_2HPO_4–0.01 M NaH_2PO_4.

Agarose Gel Electrophoresis in Methyl Mercury. Methyl mercury hydroxide (97%, Alfa Chemicals, Ventron Corporation, Beverly, Mass.) was opened in an operating flow hood and immediately dissolved in deionized water to a final concentration of 0.25 M. The suspension was stirred overnight in a glass-stoppered bottle using a magnetic stirrer. After stirring the solution was filtered through a Millipore filter and stored in a glass-stoppered bottle in the hood. The Millipore filter and the insoluble residue remaining in the original bottle were rinsed into a brown glass bottle that was sealed and later disposed of as a hazardous chemical as described below.

Agarose gels were prepared as described by Bailey and Davidson (1976). "*Seakem*" agarose powder, 0.5 to 3 g (Marine Colloids, Inc.), was suspended in 100 mL of the electrophoresis buffer (E buffer): 0.05 M boric acid, 0.005 M $Na_2B_4O_7 \cdot 10H_2O$, 0.01 M sodium sulfate, and 0.001 M Na_3EDTA, pH 8.2, without CH_3HgOH. The suspension was heated for 5 min in an autoclave and diluted to the appropriate concentration with hot E buffer. CH_3HgOH was added to the hot buffered agarose by syringe pipet and stirred rapidly at 60 °C. The gels were poured immediately using 2.5 mL per tube. Before samples were applied, the top of the gel was sliced off with a razor blade to provide a flat surface.

RNA samples were prepared by dissolving the RNA in E buffer containing 5 or 10 mM methyl mercury hydroxide as indicated. One-half volume of a 1:1 mixture of glycerol–H_2O solution containing 0.004% bromophenol blue was added and the sample applied to the gel. Electrophoresis was carried out for various times as indicated at 2.5 or 3 mA per tube, at room temperature inside the closed hood. The latter was clearly designated as being hazardous both because of high voltage and CH_3HgOH.

All operations involving methyl mercury hydroxide were carried out in the hood, with the operator wearing gloves. The tops of the gels, leftover agarose, used gels, and any material contaminated with CH_3HgOH (gloves, paper, disposable items) were placed in a disposable plastic bag and stored in the hood. CH_3HgOH waste and contaminated items were disposed of every 3 months by the Department of Chemistry's hazardous chemical disposal service, the Radiac Corp. The operator was monitored for CH_3HgOH accumulation every 6 months by the Harvard Health Services.

Agarose Gel Electrophoresis in 2.2 M Formaldehyde. RNA samples were heated in 2.2 M formaldehyde (prepared from 37% Mallinckrodt formaldehyde) in 50% formamide, 0.018 M Na_2HPO_4–0.002 M NaH_2PO_4, for 5 min at 60 °C. Agarose gels were prepared by heating a "3%" agarose–water suspension in the autoclave for 5 min and diluting with either 1 volume of 4.4 M formaldehyde in 0.036 M Na_2HPO_4–0.004 M NaH_2PO_4 to make 1.5% gels or with 2 volumes of the latter plus 1 volume of water to make 0.75% gels. The gels were poured immediately using 2.5 mL per tube. Electrophoresis was carried out at 2 mA per tube at room temperature for the times indicated. The electrophoresis buffer was 2.2 M formaldehyde–0.018 M Na_2HPO_4–0.002 M NaH_2PO_4.

Agarose Gel Electrophoresis in 6 M Urea at pH 3.8. Electrophoresis was performed following the procedure described by Rosen et al. (1975) with the following modifications. Three times agarose (usually 3 g/100 mL of H_2O) was heated in the autoclave and then diluted with 2 volumes of warm 1.5× buffer to give the final concentration of agarose and 6.0 M urea, 0.025 M citric acid (pH 3.8), heated at 65 °C for 20 s, and fast cooled. One-half volume of the glycerol dye was added as described above. Electrophoresis was carried out for 12 h at 0.38 mA per gel at 4 °C.

Staining Gels with Ethidium Bromide. Agarose and polyacrylamide gels were stained overnight in 1 μg/mL ethidium bromide in 0.1 M ammonium acetate (Bailey and Davidson, 1976). After staining, the gels were photographed under short-wave UV light with a Polaroid MP 3 Land camera and high-speed type 57 film (Polaroid Corp.) using a yellow filter.

Results

Polyacrylamide Gel Electrophoresis in 99% Formamide. Since analysis of RNA on gels in 99% formamide is the most commonly used procedure for the determination of RNA molecular weights under denaturing conditions, it is important to define the limitations of this system. The use of pure, dry formamide (see Materials and Methods) is essential for ob-

[1] Abbreviations used: TEMED, *N,N,N′,N′*-tetramethylethylenediamine; UV, ultraviolet.

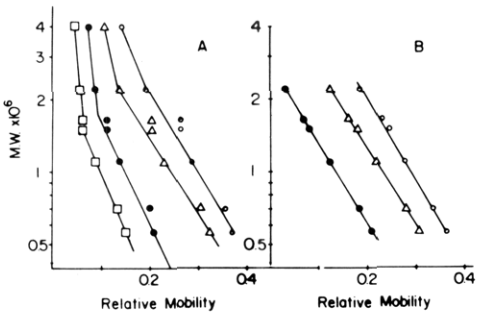

FIGURE 2: Log molecular weight–mobility relation of RNAs in polyacrylamide–formamide gels at two temperatures. Acrylamide concentrations: (□) 4.2%; (●) 3.8%; (△) 3.4%; (○) 3.2%. (A) Electrophoresis at room temperature for 5 h, 40 min at 2 mA/gel (5.6 V/cm). In order of decreasing molecular weight, RNA samples were: Sindbis, TMV, 28S mouse, 27S chick, 23S *E. coli*, 18S chick, and 16S *E. coli* RNA. (B) Electrophoresis at 58 °C for 5 h, 40 min at 2 mA/gel (5.6 V/cm). RNA samples as in A except for Sindbis RNA which could not be measured because it apparently degraded during the electrophoresis.

taining reproducible mobilities and reproducible gel polymerization. Even when reproducible mobilities were obtained, however, the apparent molecular weight of some RNA species varied with gel concentration. This aberrant behavior could be attributed to two different structural factors resulting in too rapid migration of RNA molecules examined.

The first, the deformation of large RNA molecules from the random coil conformation to more extended forms, is illustrated by the data obtained with Sindbis RNA and TMV RNA and *E. coli* rRNA at six different acrylamide concentrations and at two voltage gradients. The log molecular weight–mobility results obtained on gels run at low voltage gradients (2.2 V/cm, 0.8 mA/gel) and at high gradients (7.0 V/cm, 2.5 mA/gel) is shown in Figures 1A and 1B, respectively. The molecular weight of TMV RNA extrapolated from the straight line defined by 23S and 16S *E. coli* rRNA falls well below its correct value of 2.2×10^6 at all gel concentrations above 3.2% at both current densities. At low voltage gradients, however, this effect results from the inherent nonlinearity of log molecular weight–mobility plots (Neville, 1971) since it is possible to estimate the correct molecular weight of TMV RNA if the square root of the molecular weight is plotted vs. log mobility as suggested in the Appendix.

The use of low gel concentrations, low voltage gradients, and the square root molecular weight–log mobility plot permits valid RNA molecular weight determinations for RNA molecules as large as 2.0×10^6 on polyacrylamide gels in 99% formamide. The second physical constraint on RNA molecules resulting in a variation of apparent molecular weight with gel concentration cannot be resolved so readily. As predicted by Pinder et al. (1974), Wellauer and Dawid (1973), and Wellauer et al. (1974), 99% formamide does not fully denature the GC-rich helical regions of the large ribosomal RNA of vertebrates. As a result, these molecules have anomalously high mobilities on polyacrylamide gels in 99% formamide at room temperature (Figure 2A). Not only are the mobilities of 27S chick and 28S mouse rRNA higher than expected for their molecular weights determined on more fully denaturing gels (see below) but the deviation increases with decreasing acrylamide concentration. In addition, the two rRNAs comigrate on these gels as is illustrated in Figure 3.

If these rRNAs are electrophoresed at 58 °C, using a modification of the procedure of Spohr et al. (1976), the 27S and 28S rRNAs are well separated and the anomalous dependence of molecular weight on gel concentration is eliminated (Figure 2B). Since the GC-rich helical regions are known

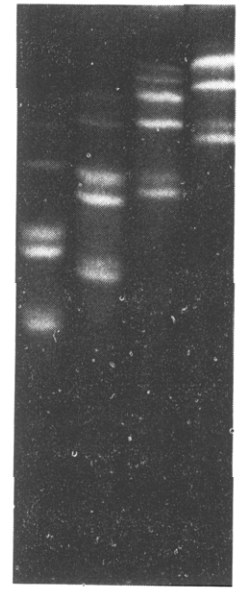

FIGURE 3: Polyacrylamide gel electrophoresis in 99% formamide. Electrophoresis from top to bottom. Sindbis and TMV RNA, 0.1 μg/band, others, 0.5 μg/band. Other details given in Figure 2A. Gel concentration increasing from left to right.

to melt at about 50 °C in 99% formamide in 0.03 M Na^+ (unpublished results), these data clearly indicate that the anomalous mobilities of 27S and 28S rRNA are due to incomplete denaturation of these RNAs. Electrophoresis at elevated temperature also appears to counteract in part the anomalous mobilities of TMV RNA either because the random coil conformation is less readily distorted at the higher temperature,[2] or because the properties of the gels are different.

Polyacrylamide gel electrophoresis in 99% formamide *at room temperature* is thus limited to RNA molecules having molecular weights of less than 4.0×10^6 which are fully denatured at room temperature. Full denaturation can be achieved at room temperature if Cu(II) (1.5 mM $CuCl_2$/mM nucleotide) is added to the sample, and if Cu(II) is electrophoresed into the gel after polymerization (Gratzer, 1976). Under normal conditions, that is without Cu(II), the applicability of this method to RNA molecules of unknown properties can be determined by measuring their molecular weights at several acrylamide concentrations and/or several voltage gradients.

Agarose Gel Electrophoresis in 10 mM Methyl Mercury Hydroxide. The limitations of acrylamide gel electrophoresis in 99% formamide led us to investigate other denaturing RNA gels, employing an agarose matrix in which high porosity gels can be formed. The use of agarose gels, however, precludes the use of simple hydrogen bond breaking agents such as formamide because they dissolve the gel, but does permit the use of reagents which specifically prevent base pairing. Methyl mercury hydroxide, first used by Bailey and Davidson (1975) provides such a denaturant. We have confirmed that RNA mobilities in agarose gels containing increasing concentrations of methyl mercury undergo the sharp twofold decrease at approximately 3 mM methyl mercury described by Bailey and Davidson. At methyl mercury concentrations greater than 10 mM, the mobilities of the various RNAs increased, however, reaching the mobilities of native RNAs. An explanation of this

[2] Less deformation of random coil molecules is expected at higher temperatures because this deformation involves a reduction in entropy. Since the negative ΔS is amplified as the temperature is raised, deformations are less favorable.

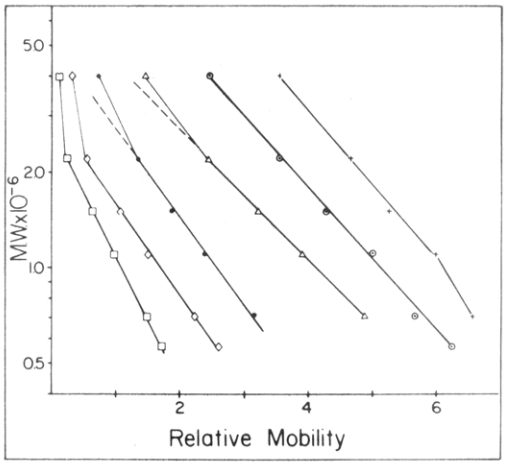

FIGURE 4: Log molecular weight–mobility relation for RNAs in agarose gels in 10 mM methyl mercury hydroxide. Agarose concentrations and time of electrophoresis: (□) 2.5%, 8 1/3 h; (◇) 2.0%, 6 5/6 h; (●) 1.5%, 5 1/3 h; (△) 1.0%, 4 h; ⊙, 0.75%, 2 1/2 h; (+) 0.5%, 1 1/3 h. RNA samples as in Figure 2A except that 28S mouse rRNA was not included in this set of measurements. Electrophoresis was carried out at room temperature at 2.5 mA/tube (3.3 V/cm).

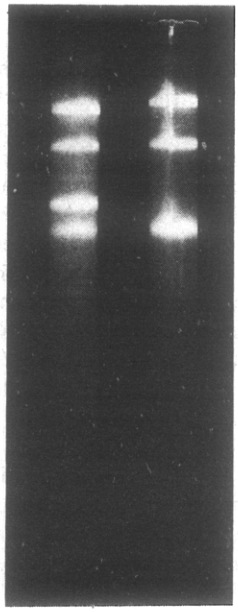

FIGURE 5: Two percent agarose gels in 10 mM CH_3HgOH. Electrophoresis from top to bottom at 3 mA/gel (3.5 V/cm) for 3 h. Left in descending order: TMV RNA (very faint band), 27S chick rRNA, 23S *E. coli* rRNA, 18S chick rRNA, and 16S *E. coli* rRNA. Right in descending order: TMV RNA, 28S mouse rRNA, 23S *E. coli* rRNA, and 16S *E. coli* rRNA. Each band except that of TMV RNA contains 2 μg; TMV RNA contains 0.5 μg.

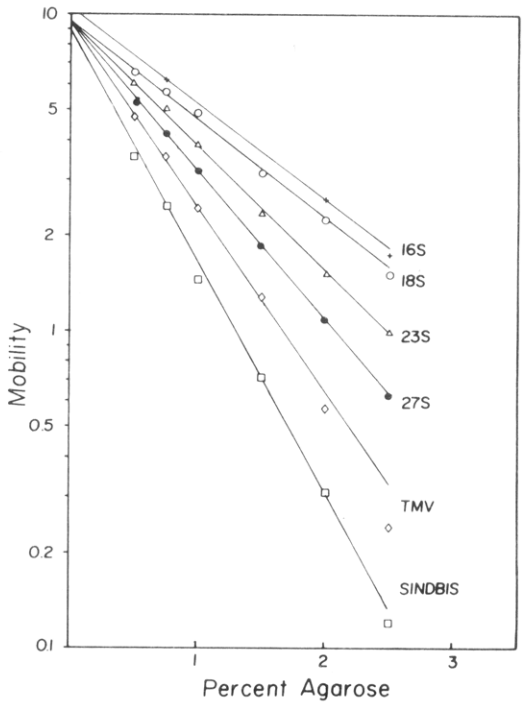

FIGURE 6: Ferguson plot of RNA mobilities from methyl mercury–agarose gels. Details of electrophoresis given in legend to Figure 4.

increase is that compact RNA structures are formed because significant amounts of methyl mercury are bound to RNA at these relatively high concentrations of methyl mercury (Fu and Gruenwedel, 1976; Gruenwedel and Davidson, 1967). Therefore, varying the methyl mercury concentration cannot be used to demonstrate complete denaturation of GC-rich regions of ribosomal RNA. To determine the extent to which these regions are denatured by methyl mercury, we measured the apparent molecular weight of chicken ribosomal RNA at different gel concentrations (Figure 4) and found it to be 1.5×10^6 at all gel concentrations measured. The molecular weight of 28S mouse rRNA was also measured on methyl mercury agarose gels (Figure 5) and found to be 1.6×10^6. While these results do not prove full denaturation, they at least show that 27S and 28S ribosomal RNA are as denatured as TMV RNA and *E. coli* rRNA, neither of which contains GC-rich helical regions.

Moreover, the log molecular weight–mobility plot is linear up to molecular weights of 4.0×10^6 at agarose concentrations of 0.75% or less at the voltage gradient used in these measurements (3.3 V/cm, 2.5 mA per gel). The use of lower voltage gradients should extend the range of applicability to higher molecular weight RNA molecules.

Additional verification of the validity of methyl mercury–agarose gels is the absence of any obvious anomalies if the log mobility is plotted as a function of gel concentration (Ferguson, 1964) as shown in Figure 6. Except for the expected slight deviation of the high molecular weight RNAs at the higher gel concentrations, all the data fall on straight lines intersecting at a common intercept at zero gel concentration.

Methyl mercury agarose gels appear to provide stringent denaturing conditions that still permit very dilute agarose to gel. The only disadvantage is the toxicity of this volatile denaturant making it hazardous for routine investigations.

Agarose Gel Electrophoresis in 2.2 M Formaldehyde. Another denaturant readily adaptable for use in agarose gels is formaldehyde, the first denaturant used for electrophoretic analysis of RNA under denaturing conditions (Boedtker, 1971). The conditions originally used for reacting RNA with formaldehyde are not sufficient to denature the GC-rich helical regions of some of the large eucaryotic RNAs. However, reaction in 2.2 M formaldehyde, 50% formamide for 5 min at 60 °C, a slight modification of the conditions used by Lizardi et al. (1975), successfully denatures the RNA without appreciable degradation. The molecular weights of 27S chicken and 28S mouse rRNA were determined by electrophoresis in 1.5% and 0.75% agarose gels in 2.2 M formaldehyde and found to be 1.5×10^6 and 1.65×10^6, respectively (Figure 7). These values are in excellent agreement with those obtained on both formamide gels at 58 °C and on methyl mercury agarose gels at 25 °C. Moreover, molecular weights of RNAs as large as

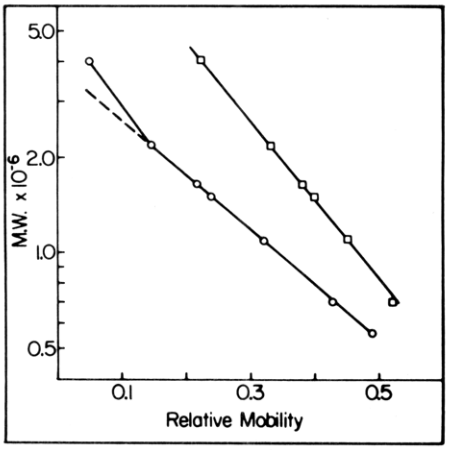

FIGURE 7: Log molecular weight–mobility relation of RNAs in agarose gels in 2.2 M formaldehyde. RNA samples as in Figure 2A. (○) with 1.5% agarose, electrophoresis for 6 h at 2 mA/gel (3 V/cm); (□) 0.75% agarose, electrophoresis for 3.5 h at 2 mA/gel (3 V/cm).

Sindbis RNA (4.0×10^6) can be measured on 0.75% agarose–formaldehyde gels. Lower agarose concentrations and lower voltage gradients should make this method applicable for molecular weight determinations of even larger RNAs.

One disadvantage of formaldehyde–agarose gels is that the formaldehyde reaction is reversed very slowly at room temperature after unreacted formaldehyde is removed (Boedtker, 1967). For this reason, it is difficult to recover biologically active RNA from gels. In experiments in which recovery of RNA is not required, this method has obvious advantages. Formaldehyde is commercially available, can be used with no further purification and is not toxic. Moreover, formaldehyde–agarose gels can be extended to molecular weight measurements of very large RNA molecules which would behave anomalously on polyacrylamide gels in either 99% formamide or 8 M urea run at 50–60 °C.

Agarose Gel Electrophoresis in 6 M Urea at pH 3.8. The application of polyacrylamide gel electrophoresis in 6 M urea, pH 3.5 to separate oligonucleotides prior to sequencing is well documented (DeWachter and Fiers, 1971; Lee and Wimmer, 1976). These conditions were first used to study high molecular weight RNAs by agarose gel electrophoresis by Rosen et al. (1975). It is difficult to estimate the extent of denaturation and the validity of this technique for the determination of RNA molecular weights because this procedure is based both on the destabilization of hydrogen bonds by urea and the specific protonation of adenine and cytosine. This partial protonation will also change the total charge. RNA molecules with different base compositions may therefore have different charge/mass ratios, invalidating the use of mobilities to determine molecular weights. Moreover, it is known that at neutral pH, 6 M urea does not fully denature most RNA molecules at or below room temperature (Boedtker, 1971; Rejinders et al., 1973). It is also known that both poly(A) and poly(C) form helices at acid pH. Hypochromicity measurements could not be used to study the extent of denaturation because the pH of the solution changes irreversibly from 3.8 to 7.0 upon prolonged heating. It was possible, however, to again estimate the relative denaturation of 27S chicken and 28S mouse rRNA by measuring the apparent molecular weight as a function of gel concentration. The two RNAs were well separated on these gels as shown in Figure 8. Their apparent molecular weights were independent of gel concentration (Figure 9) and were in excellent agreement with the values determined on methyl mercury and formaldehyde gels.

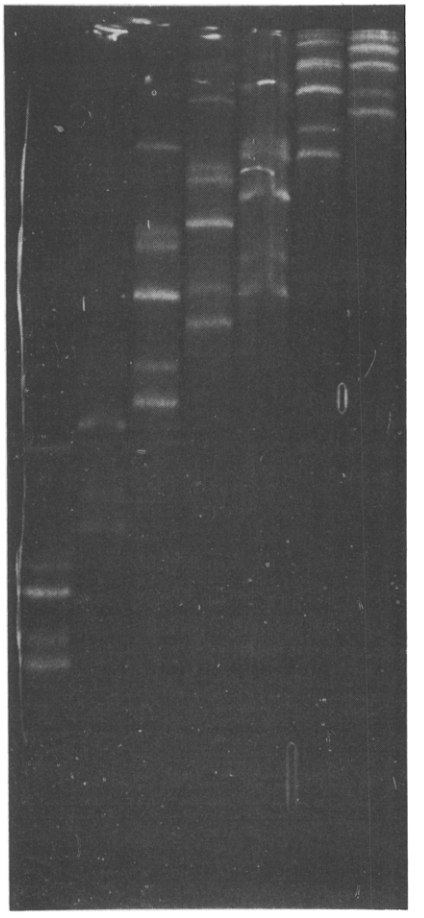

FIGURE 8: Agarose gels in 6 M urea, pH 3.8. Electrophoresis from top to bottom; 0.5 μg of RNA/band. Other details given in Figure 9. Gel concentration, from left to right: 0.5, 1, 1.5, 2, 2.5, 3, 3.5%.

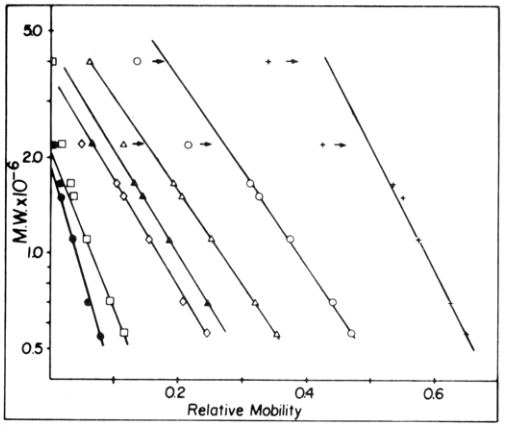

FIGURE 9: Log molecular weight–mobility relation for RNAs electrophoresed on acid urea–agarose gels. In order of decreasing molecular weight, the RNA samples are ones given in Figure 2A. Agarose concentrations: (●) 3.5%; (□) 3%; (◇) 2.5%; (▲) 2%; (△) 1.5%, (○)1%; and (+) 0.5%. Electrophoresis was carried out for 12 h at 4 °C at 0.38 mA/gel (2.0 V/cm).

In contrast to the large ribosomal RNAs, TMV RNA and Sindbis RNA exhibit anomalous mobilities at almost all gel concentrations used, resulting, for example, in an apparent molecular weight of 4.3×10^6 for TMV on the 0.5% gel, a value almost twice that of its true molecular weight. The abnormally low mobilities obtained with Sindbis RNA in 0.5 and 1.0% gels, and for TMV RNA at all gel concentrations were unexpected and could not be explained in terms of either incomplete de-

naturation or deformation of the random coil conformation since both of these result in higher rather than lower mobilities.

Analysis of the data on a Ferguson plot suggests that the anomalous mobilities relative to those obtained for ribosomal RNAs result from a considerably more expanded structure and from different protonation. The apparent failure of all RNA species to have a constant charge to mass ratio at pH 3.8 as well as the observed structural differences make the use of acid urea gels for RNA molecular weight determination highly questionable. At the very least, results obtained on these gels must be confirmed by analysis on formaldehyde or methyl mercury agarose gels. Nevertheless, acid urea gels are very easy to handle and have very high resolution. For preparative purposes, particularly of high molecular weight RNAs, these gels have a decided advantage over other RNA gels.

Discussion

A comparative study of RNA molecular weight determinations by gel electrophoresis under four different conditions, while failing to provide any ideal system, has enabled us to delineate the advantages and disadvantages of each. Anomalous mobilities resulting in erroneous molecular weight determination of some RNA molecules were revealed by analyzing RNAs at more than one gel concentration, voltage gradient, and temperature. Using 23S and 16S *E. coli* rRNA as the primary standards, we have been able to determine three different factors leading to apparent RNA molecular weights which varied with the electrophoretic conditions used: deformation of the random coil conformation of large RNAs, incomplete denaturation, and differences in the charge/mass ratios. The first effect, the "limiting mobility" exhibited by large RNA molecules under denaturing conditions has previously been observed for double-stranded DNA (Fisher and Dingman, 1971) and more recently for single-stranded DNA (McDonnell et al., 1977), but it is clearly a more general phenomenon which ultimately leads to the molecular weight independent migration of any large linear charged macromolecule at some finite gel concentration. The presumptive cause is the deformation of the random coil conformation during migration through the gel matrix. In extreme cases this results in an end on migration, or "snaking" of molecules through the gel with all molecules larger than a given size having the same or "limiting" mobility. Since the friction/force ratio of a molecule "snaking" through a gel is constant, its mobility must be independent of molecular weight. Failure to recognize this limiting mobility phenomenon not only can and has resulted in the underestimation of the molecular weight of large RNAs but can also erroneously suggest the existence of a single homogeneous RNA species when in fact a very heterogeneous population exists (Boedtker and Lehrach, 1976).

Incomplete denaturation inevitably causes various RNAs to have solution conformations which are not hydrodynamically equivalent and hence invalidates the use of the log molecular weight–mobility relation to obtain molecular weights from electrophoretic mobilities. In our limited experience the only clear example of incomplete denaturation occurred with the large chick and mouse ribosomal RNAs in formamide acrylamide gels run at room temperature. Although their molecular weights differ by 10%, these two RNAs appear to coelectrophorese on these gels. In addition, their apparent molecular weight varied from 1.3×10^6 to 1.6×10^6 as the acrylamide concentration was increased from 3.2 to 4.2%. This relatively rare denaturation resistant RNA species nevertheless poses a serious problem since vertebrate rRNAs and their precursors are commonly used as size markers in molecular weight determinations of both messenger RNA and messenger RNA precursors on formamide acrylamide gels run at room temperature. These difficulties can, of course, be avoided by running the electrophoresis at 50–60 °C as first suggested by Spohr et al. (1976) and confirmed by us. However, very large RNA molecules are not likely to remain intact after several hours at 60 °C (Eigner et al., 1961).

Agarose–urea gels at pH 3.8 provide good separation of RNA molecules from 0.5×10^6 to 4.0×10^6 but are unreliable for molecular weight determinations because differences in structure and protonation result in anomalous mobilities for some RNA molecules.

Agarose gels in which either formaldehyde or methyl mercury serves as the denaturant provide both the advantage of high porosity gels permitting good separation of RNA molecules up to at least 4.0×10^6 and reliable denaturation of the GC rich-hairpin regions in eucaryotic ribosomal RNA. Using these gel systems, we have been able to obtain reasonably reliable molecular weight estimates of 1.5×10^6 and 1.65×10^6 for the large avian and mammalian rRNA, and of 4.0×10^6 for Sindbis RNA. A previous estimate of the avian 27S RNA was as low as 1.3×10^6 (Spohr et al., 1976) while estimates of the mammalian 28S rRNA ranged from as high as 2.0×10^6 (Bailey and Davidson, 1976) to as low as 1.44×10^6 (Rejinders et al., 1973). Our intermediate value is also in good agreement with the value of 1.74×10^6 obtained for mouse L cell 28S by Wellauer et al. (1974).

Finally, a theoretical basis for the log molecular weight–mobility relation was derived (see Appendix) and the limits of its linearity defined. For denatured RNA molecules, whose molecular weight can only be determined by extrapolation, an alternative semiempirical relation is suggested which clearly offers a more reliable method for molecular weight determination than those available previously.

Acknowledgments

We want to express our appreciation of the gifts of 28S mouse rRNA by Jesse F. Scott of the Harvard Medical School and of Sindbis virus by Michelene McCarthy of the Department of Biochemistry and Molecular Biology. We also want to thank both Stephen Henikoff and Peter Upcroft for the valuable discussion and exchange of information while this work was being carried out, and Lynn Klotz and William McClure for a critical reading of the manuscript. We are indebted to F. William Studier for sending us a copy of his manuscript prior to publication and to Walter Gratzer for sending us a copy of his method of denaturing GC rich polynucleotides using $CuCl_2$ in 99% formamide.

Appendix

Evaluation of Procedures Used in RNA Molecular Weight Determinations by Gel Electrophoresis. Molecular weights of nucleic acids or sodium dodecyl sulfate–protein complexes are usually obtained from a plot of log molecular weight vs. mobility. This procedure has been empirically developed and has no direct theoretical justification. Various groups (Neville, 1971; Andersson et al., 1975) have pointed out that the linear relationship used only constitutes a segment of a nonlinear, sigmoidal function. These groups have suggested an alternative plot of molecular weight vs. log mobility. This plot can be derived from the theory of gel electrophoresis developed by Ogston (1958), Morris (1967), and Rodbard and Chrambach (1970) using certain assumptions. We will show that relationship between plots of different powers of the molecular weight vs. log mobility provides the basis of the applicability

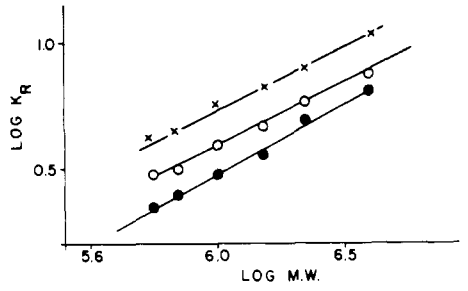

FIGURE 10: Log retardation coefficient-log molecular weight relation. RNAs analyzed in order of increasing molecular weight: 16S *E. coli*, 18S chick, 23S *E. coli*, 27S chick, TMV, Sindbis, mRNA. (X) polyacrylamide gel in 99% formamide, slope = 0.5; (○) agarose gels in 10 mM methyl mercury, slope = 0.5; (●) agarose gels in 6 M urea, pH 3.8, slope = 0.55. Details of electrophoresis are given in Figures 2A, 4, and 9.

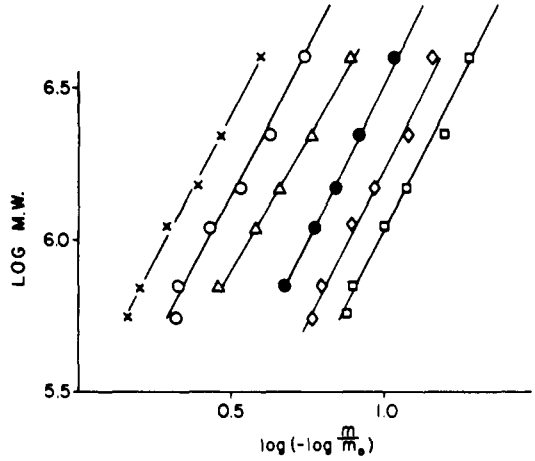

FIGURE 11: Log molecular weight-log ($-\log(m/m_0)$) relation for RNAs on methyl mercury-agarose gels. Details of electrophoresis given in Figure 4. Agarose concentrations from (left to right): 0.5, 0.75, 1.0, 1.5, 2.0, 2.5%.

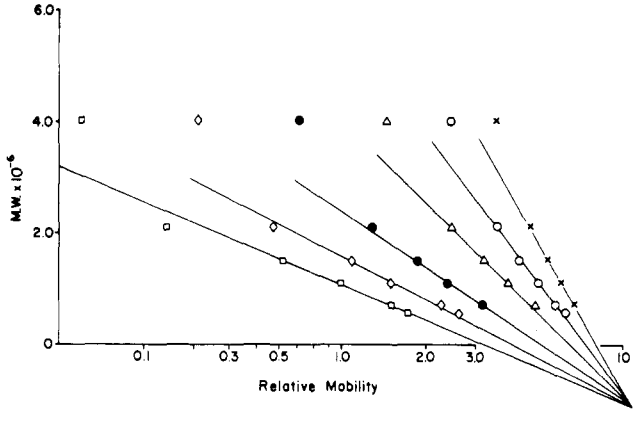

FIGURE 12: Molecular weight-log mobility relation for RNAs, analyzed on methyl mercury-agarose gels. Details of electrophoresis given in Figure 4. $st = 1$.

of the more commonly used log molecular weight-mobility plot. We can also show that, for random coil molecules, a plot of the square root of the molecular weight against the log of the mobility appears to result in fewer systematic deviations than either of the two procedures mentioned above. Therefore, the migration of denatured RNA is a function of an effective radius that is proportional to the square root of the radius of gyration of a random coil molecule. Ogstons model for the structure of gels predicts that a plot of the log mobility, m, of different molecules as a function of the gel concentration, c, should result in straight lines with different slopes (retardation coefficients, K_R) and intercepts ("free mobilities", m_0) for different molecules. In a homologous series of macromolecules with identical charge/mass ratios, these lines intersect the y axis at a common intercept m_0, the extrapolated mobility at zero gel concentration (see Figure 6, for example). This may be expressed mathematically as

$$m = m_0 e^{-K_R c} \tag{1}$$

In addition, for spherical molecules, the relationship between retardation coefficient and molecular weight can be predicted from the following equation (Rodbard and Chrambach, 1971):

$$K_R{}^{1/s} = a(r + R) \tag{2}$$

with r the radius of the gel fiber, R the molecular radius, and $s = 3, 2, 1$ depending on whether the gel is 0, 1, or 2 dimensional. For agarose gels, idealized as a random network of fibers, one expects s to be 2. For nonspherical molecules, R will be equal to the radius of a sphere showing equivalent electrophoretic behavior and will therefore be a general function of the molecular weight, M, which can usually be approximated by a power function

$$R = bM^t \tag{3}$$

When $r \ll bM^t$, the product, st, can be determined from the slope of a plot of the log retardation coefficient against the log molecular weight. As shown in Figure 10, a slope of approximately 0.5 is obtained for all three gel systems. To determine if there is a systematic trend in this value as the gel concentration is varied, we plotted the $\log[-\log(m/m_0)]$ against log M for each concentration of the methyl mercury gels (Figure 11). Again the slopes of the individual lines are close to 0.5. Therefore, over the concentration range examined, the electrophoretic behavior of denatured RNA molecules can be approximated by the behavior of a set of equivalent spheres. If we accept the theoretical value of $s = 2$ for these gels, we can conclude that the radius of an equivalent sphere, R_e, is not proportional to the calculated radius of gyration of the random coil as one might expect but appears to vary approximately with the 4th root of the molecular weight (the square root of the radius of gyration).

If s and t are constant, and $r \ll bM^t$, eq 2 can be approximated by

$$K_R = a^*(r^* + b^*M^{st}) \tag{4}$$

Combining eq 1 and 4 results in eq 5, which describes the dependence of the mobility on the molecular weight.

$$\ln m = \ln m_0 - ca^*(r^* + b^*M^{st}) \tag{5}$$

$$M^{st} = -\frac{1}{ca^*b^*}(\ln m - \ln m_0) - (r^*/b^*) \tag{6}$$

Plotting M^{st} against ln m at different gel concentrations should therefore result in a set of straight lines with slopes proportional to $1/c$, intersecting at the point $(\ln m_0, -(r^*/b^*))$. The plot of M against log m suggested by Neville (1971) and Andersson et al. (1975) is obviously a special case of $st = 1$. The results obtained with the methyl mercury gels for $st = 1$ and $st = 0.5$ are shown in Figures 12 and 13. For st equal to 0.5, the relation is linear up to molecular weights of 4.0×10^6. We suggest, therefore, that this relation be used for molecular weight determination of large denatured RNA molecules.

The linearity of the square root molecular weight-log mobility relation makes possible the prediction of limited linearity

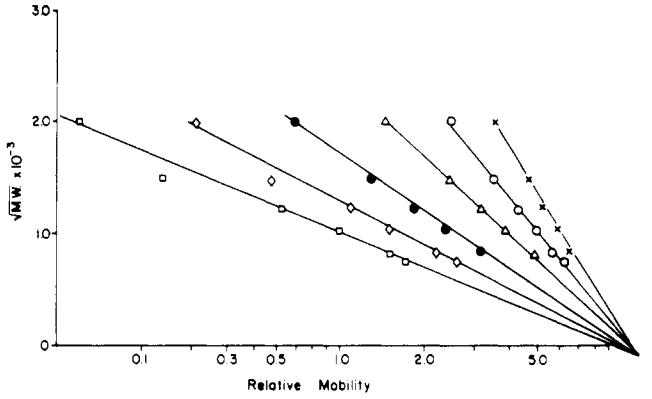

FIGURE 13: The square root of the molecular weight–log mobility relation for RNAs analyzed on methyl mercury–agarose gels. Details of electrophoresis given in Figure 4. $st = 0.5$.

of the log molecular weight–mobility relation generally used to determine molecular weights. To simplify eq 5 we define

$$m^* = m_0\left(\frac{1}{e^{cr^*a^*}}\right)$$

$$u = \frac{m}{m^*}$$

$$M_0 = \left(\frac{2}{ca^*b^*}\right)^{1/st}$$

$$v = \left(\frac{M}{M_0}\right)^{st}$$

giving

$$\ln u = -2v \tag{7}$$

Expanding the logarithm into a series we get:

$$\ln u = 2\left\{\frac{(u-1)}{(u+1)} + \frac{1}{3}\left(\frac{(u-1)}{(u+1)}\right)^3 + \frac{1}{5}\left(\frac{(u-1)}{(u+1)}\right)^5 + \ldots\right\}$$

$$= 2\left(\frac{u-1}{u+1}\right) + R_u$$

Substituting into eq 7 and rearranging results in:

$$\left(\frac{u-1}{u+1}\right) = -v - \frac{1}{2}R_u$$

which can be rearranged to:

$$\left(\frac{v-1}{v+1}\right) = -u - \frac{1}{2}R_u\left(\frac{u+1}{v+1}\right)$$

and with

$$\ln v = 2\left(\frac{v-1}{v+1}\right) + R_v$$

we obtain:

$$\ln v = -2u + \left\{R_v - R_u\left(\frac{u+1}{v+1}\right)\right\} \tag{8}$$

or substituting back, we get:

$$\ln\frac{M}{M_0} = -\frac{2}{st}\frac{m}{m^*} + \frac{1}{st}\left\{R_v - R_u\left(\frac{u+1}{v+1}\right)\right\} \tag{9}$$

If we ignore the correction terms, R_v and R_u, we get first approximations for the intercept and slope.

$$\text{intercept: } M_i \approx M_0 = \left(\frac{2}{ca^*b^*}\right)^{1/st} \tag{10}$$

$$\text{slope: } S \approx -\frac{2.303}{st}\frac{2}{m^*} = -\frac{4.606}{st}\left(\frac{e^{cr^*a^*}}{m_0}\right) \tag{11}$$

By developing the higher order terms, R_u and R_v, into a Taylor series around the point $(1/e, 1/2)$ and terminating after two terms, we can show that the inclusion of these terms will increase the slope by a factor of $e/2$. The same result can also be obtained by differentiation of the exact equation,

$$\ln v = \ln\left(-\frac{\ln u}{2}\right) \tag{12}$$

resulting in:

$$\text{slope} = \frac{1}{u\ln u} \tag{13}$$

$$\text{intercept } v_i = e^{[\ln(-\ln u/2) - (1/\ln u)]} \tag{14}$$

Both these functions have extrema at $m/m^* = e^{-1}$ with the slope $= -e$ and the intercept, $v_i = e/2$ so that $M_{i,\min} = M_0(e/2)^{1/st}$.

The exact expression for the slope and intercept go to plus or minus infinity for values of $m/m^* = 0$ or 1 and are relatively constant over a wide range around the extremum at $m/m^* = e^{-1}$. Furthermore, the location of the maximum of the negative slope also identifies the largest relative separation, $\mathrm{d}\ln M/\mathrm{d}m$. From that, it follows that the best relative separations will be reached at:

$$\frac{m}{m^*} = e^{-1} \text{ and } \frac{M}{M_i} = e^{-1/st}$$

The corrected equations for intercept and slope

$$M_i = M_0\left(\frac{e}{2}\right)^{1/st} = \left(\frac{e}{ca^*b^*}\right)^{1/st} \tag{15}$$

$$\text{slope} = \frac{-2.303e}{st}\left(\frac{e^{cr^*a^*}}{m_0}\right) \tag{16}$$

allow the determination of these parameters as a function of the physical variables, gel concentration, charge, and type of structure. From eq 15 and 16, we can predict how log M vs. m plots will vary as a function of the gel concentration, c. At lower gel concentration ($cr^*a^* \ll 1$), the lines will be almost parallel; at high gel concentrations ($cr^*a^* > 1$), the slopes will start to increase. In addition, we can predict that the intercepts will be roughly proportional to $(1/c)^{1/st}$. Finally, changes in charge/mass ratio will only influence the slope by changing m_0, while changes in structure will influence both the intercept and the slope.

References

Andersson, K., Fagerlind, M., and Daneholt, B. (1975), *Mol. Biol. Rep. 2,* 195–201.

Bailey, J. M., and Davidson, N. (1976), *Anal. Biochem. 70,* 75–85.

Boedtker, H. (1960), *J. Mol. Biol. 2,* 171–188.

Boedtker, H. (1967), *Biochemistry 6,* 2718–2727.

Boedtker, H. (1968), *Methods Enzymol. 12,* 429–458.

Boedtker, H. (1971), *Biochim. Biophys. Acta 240,* 448–453.

Boedtker, H., Crkvenjakov, R. B., Dewey, K. F., and Lanks, K. (1973), *Biochemistry 12,* 4356–4360.

Boedtker, H., and Lehrach, H. (1976), *Prog. Nucleic Acid Res. Mol. Biol. 19,* 253–260.

DeWachter, R., and Fiers, W. (1971), *Methods Enzymol. 21,* 167–187.

Eigner, J., Boedtker, H., and Michaels, G. (1961), *Biochim. Biophys. Acta 51,* 165-168.
Ferguson, K. A. (1964), *Metabolism 13,* 985-1002.
Fisher, M. P., and Dingman, C. W. (1971), *Biochemistry 10,* 1895-1899.
Fu, J. C. C., and Gruenwedel, D. W. (1976), *Arch. Biochem. Biophys. 174,* 402-413.
Gratzer, W. B. (1976), private communication.
Groot, P. H. E., Aij, C., and Borst, P. (1970), *Biochem. Biophys. Res. Commun. 41,* 1321-1327.
Gruenwedel, D. W., and Davidson, N. (1967), *Biopolymers 5,* 847-861.
Lee, Y. F., and Wimmer, E. (1976), *Nucleic Acid Res. 3,* 1647-1658.
Lizardi, P. M., Williamson, R., and Brown, D. D. (1975), *Cell 7,* 239-245.
MacLeod, M. C. (1975), *Anal. Biochem. 68,* 299-310.
McDonell, M. W., Simon, M. N., and Studier, F. W. (1977), *J. Mol. Biol. 110,* 119-146.
Morris, C. J. O. R. (1967), Protides of the Biological Fluids, 14th Colloquim, Peeters, H., Ed., New York, N.Y., American Elsevier, p 543.
Neville, D. M., Jr. (1971), *J. Biol. Chem. 246,* 6328-6334.
Ogston, A. G. (1958), *Trans. Faraday Soc. 54,* 1754-1757.
Perry, R. P., LaTorre, J., Kelley, D. E., and Greenberg, J. R. (1972), *Biochim. Biophys. Acta 262,* 220-226.
Pinder, J. C., Staynor, D. Z., and Gratzer, W. B. (1974), *Biochemistry 13,* 5373-5378.
Rejinders, L., Sloof, P., Sival, J., and Borst, P. (1973), *Biochim. Biophys. Acta 324,* 320-333.
Rodbard, D., and Chrambach, A. (1970), *Proc. Natl. Acad. Sci. U.S.A. 65,* 970-977.
Rosen, J. M., Woo, S. L. C., Holder, J. W., Means, A. R., and O'Malley, B. W. (1975), *Biochemistry 14,* 69-78.
Spohr, G., Mirault, M., Imaizumi, T., and Scherrer, K. (1976), *Eur. J. Biochem. 62,* 313-322.
Wellauer, P. K., and Dawid, I. B. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 2827-2831.
Wellauer, P. K., Dawid, I. B., Kelley, D. E., and Perry, R. P. (1974), *J. Mol. Biol. 89,* 397-407.

Resonance Raman Spectra of Lobster Shell Carotenoproteins and a Model Astaxanthin Aggregate. A Possible Photobiological Function for the Yellow Protein†

V. R. Salares,* N. M. Young, H. J. Bernstein, and P. R. Carey

ABSTRACT: A yellow protein, isolated from lobster shells and containing ~20 astaxanthin molecules per unit of protein, has absorption and resonance Raman properties identical with those of aggregates of astaxanthin. The astaxanthin electronic absorption maximum shifts from ~480 to 410 nm upon aggregation or binding to the yellow protein. Taken with the resonance Raman data on the frequency of the C=C stretching vibration, this indicates that in both systems a large perturbation of the electronic excited state takes place while only a minimal perturbation of the ground state occurs. Low temperature (−95 °C) absorption measurements of the yellow protein and the astaxanthin aggregate show no vibronic structure and suggest that the electronic absorption is solely due to the 0-0 or vibrationless transition. The molecular exciton model satisfactorily explains the spectral data in terms of interactions of chromophore molecules in the excited state. The altered absorption spectrum of the yellow protein compared to that of free astaxanthin is thus attributed to chromophore-chromophore (exciton coupling) rather than protein-chromophore interactions. Resonance Raman experiments on live lobsters and reflectance measurements on shell fragments reveal a preponderance of the yellow protein at the surface of the shell. The pigment distribution in the lobster shell together with several known physiological observations and the energy-transfer capability of excitons suggest a possible photobiological function for the yellow protein, viz. that the yellow proteins act as bulk light harvesters for photons in the deep blue spectral region.

Two types of proteins, both containing astaxanthin as the chromophore, are responsible for the shell color of the lobster *Homarus americanus*. These are the three blue crustacyanins which absorb in the 600-nm region (Wald et al., 1958; Jencks and Buten, 1964; Cheeseman et al., 1966, 1967; Buchwald and Jencks, 1968b) and a yellow protein of λ_{max} 410 nm (Buchwald and Jencks, 1968a). Since free astaxanthin, in organic solvents, absorbs at 480 nm, binding in both types of protein causes large shifts in absorption (Figure 1). The mechanism of these shifts is of interest in its own right and in relation to studies on the visual pigments (Ebrey and Honig, 1975; Sulkes et al., 1976) and enzyme-substrate reactions (Carey et al., 1976). On the basis of mainly absorption and circular dichroic (CD) data, different and often conflicting mechanisms for the spectral shifts have been proposed for the astaxanthin-containing proteins. One problem with absorption and CD data is that both techniques are sensitive to changes in the ground and the excited electronic states. In contrast, Raman and resonance Raman peak positions are a property solely of the ground state. The present study therefore uses resonance Raman peak positions together with absorption data to discriminate between ground- and excited-state effects for astaxanthin in the yellow

† From the Divisions of Chemistry (V.R.S., H.J.B.) and Biological Sciences (N.M.Y., P.R.C.), National Research Council of Canada, Ottawa, Ontario, Canada, K1A 0R6. *Received December 3, 1976; revised manuscript received June 13, 1977.* This paper is NRCC Publication No. 16276.