Phylogenetic analysis of the core histones H2A, H2B, H3, and H4

Thomas H. Thatcher⁺ and Martin A. Gorovsky^{*}

Department of Biology, University of Rochester, Rochester, NY 14627, USA

Received October 26, 1993; Revised and Accepted December 7, 1993

ABSTRACT

Despite the ubiquity of histones in eukaryotes and their important role in determining the structure and function of chromatin, no detailed studies of the evolution of the histones have been reported. We have constructed phylogenetic trees for the core histones H2A, H2B, H3, and H4. Histones which form dimers (H2A/H2B and H3/H4) have very similar trees and appear to have co-evolved, with the exception of the divergent sea urchin testis H2Bs, for which no corresponding divergent H2As have been identified. The trees for H2A and H2B also support the theory that animals and fungi have a common ancestor. H3 and H4 are 10-fold less divergent than H2A and H2B. Three evolutionary histories are observed for histone variants. H2A.F/Z-type variants arose once early in evolution, while H2A.X variants arose separately, during the evolution of multicellular animals. H3.3-type variants have arisen in multiple independent events.

INTRODUCTION

The nucleosome consists of DNA wound around a histone octamer containing two molecules each of the core histones H2A, H2B, H3, and H4. In contrast to early ideas that the histone octamer played a passive role in packaging DNA, it is now recognized that histones play a crucial role in DNA replication and transcription (see 1–4 for recent reviews). Transcriptionally active chromatin has a different structure than inactive chromatin. Secondary modifications of histones, especially acetylation, appear important for some types of regulation. In addition, primary sequence variants of some histones may also play a role in determining chromatin structure and accessibility of the DNA to transcription (5–9).

Ever since the first histone (H4) was sequenced in two diverse species (10), it has been apparent that histone genes evolve very slowly. Therefore, evolutionary analyses of histones should be informative regarding the phylogenetic relationships of distantly related organisms. Because of their importance in chromatin structure and function, their conservation (enabling cross-species hybridization with histone gene probes), and their small size,

numerous histone sequences are available, either as DNA or protein. Although compendia of histone sequences have been published (11,12), surprisingly no detailed analyses of histone evolution have been described.

Therefore, we have carried out a complete phylogenetic analysis of the core histones. Histone sequences were identified from the DNA database GenBank and the protein database PIR; a total of 265 non-redundant complete sequences were found encompassing 71 H2As, 58 H2Bs, 73 H3s, and 63 H4s. The sequences were used to construct phylogenetic trees using the neighbor-joining method. The histones which form dimers (H2A and H2B, H3 and H4) have similar reconstructed phylogenies and appear to co-evolve. H3 and H4 are much more conserved than H2A and H2B, probably because they form a tetramer with a critical role in nucleosome formation. Three distinct evolutionary histories are observed for the quantitatively minor basal/replacement variants. H2A variants of the H2A.F/Z class arose early and are more conserved between species than the major replication dependent H2As. H3.3 variants have arisen in multiple independent events. H2A.X variants appear to have arisen once late in the evolution of multicellular organisms. Excepting the H2A.F/Z variants, the topologies of the H2A and H2B trees are similar. Protist and plant histones branch first, followed by fungi and animals, which appear to have a common ancestor. Because H3 and H4 sequences evolve on the order of 10 times more slowly than H2A and H2B, the branching order of major phylogenetic groups in these evolutionary trees is less certain.

METHODS

Phylogenetic analysis of histone protein sequences

Histone protein sequences were obtained from the Protein Identification Resources database (PIR, release 33.0, June, 1992) and by translating histone DNA sequences available in GenBank (release 77.0, June, 1993) (13). The histone compilations of Wells (11,12) were used as a guide. Only complete protein sequences were used. Where DNA and direct protein sequence information was available for the same organism, the translated DNA sequence was used, as many protein sequences obtained directly are mixtures of primary sequence variants.

* To whom correspondence should be addressed

⁺ Present address: University of Rochester, School of Medicine and Dentistry, 601 Elmwood Avenue, PO Box 363, Rochester, NY 14642, USA

The compiled protein sequences were aligned using the program Pileup in the GCG Sequence Analysis package (14), version 7.1, which uses a simplified form of the progressive sequence alignment method (15). Evolutionary distances (d) were calculated from the proportion of amino acid identity (S) with a Poisson correction, $d = -\ln S$. These values were used to construct phylogenetic trees by the neighbor-joining method (16).

The histone protein alignments are available from the authors by electronic mail at marty@mag.biology.rochester.edu, or on computer diskette in Macintosh or PC format; the compiled protein sequences are available on disk in Macintosh format only. Please send a blank disk and your complete mailing address and phone number with your request.

RESULTS AND DISCUSSION

H2A and H2B have similar phylogenetic trees

Neighbor-joining trees were constructed from 71 H2A protein sequences and 58 H2B sequences (Figs. 1 and 2). The trees are drawn using *Leishmania* and *Trypanosome* histones as outgroups since phylogenies based on rRNA sequences show that these organisms diverged from the main eukaryotic line earlier than any other organisms represented on these trees (17,18). With the exception of the H2A.F/Z variants (see below), the tree topologies are very similar to each other and to a eukaryotic phylogeny based on rRNA sequences (19). *Tetrahymena* and plant histones branch first, followed by fungi and animals. As in the earlier study of rRNAs (19) as well as studies of alpha- and beta-tubulin (20,21), fungal and animal H2As and H2Bs are monophyletic; that is, there is a branch point from which all animal and fungal sequences diverge (filled circle in Figs. 1 and 2), implying a common ancestor. Within the fungal–animal lineage the fungi form a distinct branch, and invertebrates, vertebrates, and mammals tend to cluster together.

Although *Tetrahymena* and plant histones group closely on both trees, *Tetrahymena* H2A and H2B fall into somewhat different groupings. *Tetrahymena* H2A falls on a branch within the plant lineage that separates ancient (*Volvox*, *Picea*) and modern (Wheat, etc.) plants, suggesting that *Tetrahymena* and plants diverged at the same time. However, *Tetrahymena* H2B appears to have diverged prior to plant H2Bs. It should be noted that the distances separating ancient and modern plant branch points on the H2A tree (Fig. 1), and the distance between the *Tetrahymena* and plant H2B branch points (Fig. 2) are relatively small compared to the sequence branch lengths. The branching order of plant and *Tetrahymena* H2A and H2B could be clarified by additional plant H2B sequences as well as histone sequences from other protists.

H2As and H2Bs from the same species have similar rates of evolution

Since the H2A and H2B trees are drawn using the same species as an outgroup, the length of time between the outgroup branch and the end of each other branch is constant between the two trees. Thus, the length of each branch is proportional to the rate of evolution of that protein. In most cases, the rates of evolution for H2As and H2Bs from the same species are very similar. For example, the *S.cerevisiae* H2As and H2Bs have evolved faster than the *S.pombe* H2As and H2Bs. A more divergent rat testis H2A (Fig. 1, number 62) has a counterpart divergent testis H2B (Fig. 2, number 40). Since H2A and H2B interact as a dimer, it is likely that they constrain each other's freedom to evolve. However, the testis-specific H2Bs of the sea urchins (Fig. 2,

numbers 19–25) have evolved faster than other sea urchin H2Bs or H2As. No divergent sea urchin H2As have been identified. This suggests that sea urchin testis H2Bs perform an atypical function or assemble nucleosomes in such a way that they are not under the same constraint as other sea urchin H2Bs. Alternatively, divergent sea urchins H2As may exist but are yet to be discovered.

H2A replacement variants arose in a single early event and are conserved

The synthesis of most histones is coupled to the cell cycle and occurs only during DNA replication. Replication independent variants of H2A have been identified in many organisms including mammals (H2A.Z), birds (H2A.F), sea urchins (H2A.F/Z), *Drosophila* (H2A.vD) and *Tetrahymena* (hv1). The neighbor-joining analysis of H2As indicates that these replication-independent H2A.F/Z-type variants (Fig. 1, boxed) arose once very early in evolution, and are more closely related to each other than to the major H2As from the same species. This confirms an earlier statistical analysis (22) and suggests that these variants play an important and highly conserved role in chromatin structure or function. Consistent with this are recent studies that indicate that H2A.vD is an essential gene in *Drosophila* (23), that hv1 (but not the major replication dependent H2A genes) is an essential gene in *Tetrahymena* (X.Liu and M.Gorovsky, unpublished observations), and that hv1 is spatially and temporally correlated with the transcriptional competence of *Tetrahymena* chromatin (9). Note that the evolution of the H2A.F/Z variants is in contrast to the replication independent variants of H3, which arose in multiple independent events (24).

Mammals have an additional H2A replacement variant, H2A.X (Fig. 1, numbers 53 and 54), which arose independently during vertebrate evolution. It should be noted that while the H2A.F/Z variants have core sequences which are different from the conserved cores of the major H2As, the H2A.X core region is nearly identical to that of the major vertebrate H2A. H2A.X has an unusual C-terminal region that distinguishes it from the major H2A and which contains a sequence element also found in the C-terminal regions of some H2As from lower eukaryotes (25). The significance of this element is unknown, and it is not clear whether it arose by chance or convergent evolution, or whether it is a truly ancient conserved sequence that has been lost from most other H2As.

H3 and H4 are highly conserved

The extreme conservation of H3 and H4 is reflected in their phylogenetic trees (Figs. 3 and 4). In both cases, a small number of sequences (30 for H3, 21 for H4) accounts for a large number of species (73 and 63, respectively). The sequences themselves are highly conserved, as indicated by the short branch lengths. In particular, there is very little diversity among plant and animal H3s and H4s. For example, the total sequence divergence between plant and animal H2A is about 10-fold as great as between plant and animal H3 (data not shown). *Tetrahymena* and fungal H3 and H4 are somewhat more diverse, but still much less so than their corresponding H2As and H2Bs. Note that the H3 variants (Fig. 3, boxed) appear to have arisen in independent events; this is discussed further elsewhere (24). We should note that while the H3 tree is drawn with an evolutionary ancient sequence from *Entamoeba* as an outgroup, no such ancient sequence has been determined for H4, so we have drawn it using as the outgroup the H4s from the protists, which are the most

Figure 1. Phylogenetic tree of histone H2A proteins. Seventy-one H2A protein sequences were obtained (see Materials and Methods), aligned, and used to construct a neighbor-joining tree. The total evolutionary distance between any two sequences is the sum of the horizontal branch lengths between them. Vertical distances are for illustration purposes only. The scale bar indicates the branch length that corresponds to 0.05 substitutions per position (the alignment contains 168 positions). The filled circle indicates the branch point at which animals and fungi are monophyletic. The boxed sequences are replacement variants. The *S.pombe* H2A variant sequence (#4) was obtained from P.Nurse (personal communication). GenBank and PIR accession numbers are as follows (PIR numbers are indicated by *): 1, X60054; 2, X61936; 3, X67287; 5, X14137; 6, X15549; 7, S07392*; 8, M37583; 9, X52318; 10, M37584; 11, J00868; 12, M86865; 13, X67819; 14, M31921; 15, M31922; 16, A02600*; 17, L18892; 18, L18893; 19, A02601*; 20, M64838; 21, X53831; 22, S48210; 23, S00623*; 24, A02602*; 25, J01325; 26, J01326; 27, M18258; 28, X05220; 29, M11500; 30, X15633; 31, A02593*; 32, M11083; 33, A02593*; 34, M14140; 35, M11085; 36, M14141; 37, L11067; 38, X56335; 39, X17072; 40, X14215; 41, M84797; 42, A02595*; 43, X53330; 44, X58895; 45, A02597*; 46, X06642; 47, A02596*; 48, M11084; 49, V01357; 50, M10559; 51, V01141; 52, M25281; 53, X14850; 54, X14850; 55, X01064; 56, I24510; 57, M21287; 58, X14730; 59, D11054; 60, J00867; 61, D11055; 62, X59962; 63, X16148; 64, A02591*; 65, X57138; 66, A02591*; 67, X16495; 68, X59961; 69, X00089; 70, M37736; 71, M33988

divergent in this data set. Nevertheless, the structure of the two trees is remarkably similar.

Tetrahymena and fungal H3 and H4 branch from these trees prior to the branching of plants and animals. This is in conflict with the branching order of trees for H2A and H2B (Figs. 1 and 2), alpha- and beta-tubulin (20,21) and eukaryotic rRNA (19). However, the high degree of conservation of plant and animal

H3 and H4, reflected in the extremely short distances between branch points, makes assignment of evolutionary order problematic. It is likely that the (relatively) large number of differences between *Tetrahymena* and fungal, and plant and animal H3 and H4 has caused the divergent proteins to cluster as outgroups to the more conserved proteins. In this regard, the H2A and H2B trees (and the tubulin and rRNA trees), with their

Figure 2. Phylogenetic tree of histone H2B proteins. Fifty-eight H2B proteins sequences were obtained, aligned, and used to construct a neighbor-joining tree. The scale bar indicates the branch length that corresponds to 0.05 substitutions per position (the alignment contains 162 positions). The filled circle indicates the branch point at which animals and fungi are monophyletic. GenBank and PIR accession numbers are as follows (PIR numbers are indicated by *): 1, M38125; 2, X60982; 3, M31333; 4, M31332; 5, M31921; 6, M31922; 7, S00622*; 8, X59873; 9, M11494; 10, X05220; 11, X55547; 12, J01328; 13, J01327; 14, L11067; 15, A02615*; 16, X01344; 17, X06640; 18, M11088; 19, A02620*; 20, A02619*; 21, X04384; 22, M13634; 23, A02618*; 24, M11086; 25, X04681; 26, M14143; 27, M14142; 28, A02613*; 29, A02612*; 30, X15633; 31, X58895; 32, X53330; 33, A02611*; 34, X56335; 35, X17072; 36, X14125; 37, M84797; 38, M84798; 39, M84799; 40, X59962; 41, X00088; 42, X57138; 43, X57127; 44, X59961; 45, X02621; 46, M60751; 47, X05862; 48, A02605*; 49, M60750; 50, X02916; 51, X03017; 52, X03018; 53, X05099; 54, A30221*; 55, M57901; 56, X05096; 57, X14731; 58, X57263;

longer branch lengths, are more reliable predictors of true phylogeny. Why fungi and *Tetrahymena* have divergent H3 and H4 proteins is unknown. It is worth noting that the more divergent H3 and H4 in yeast and *Tetrahymena* could reflect reduced constraints due to reduced interactions between H1 and these core proteins. Fungi lack histone H1 and *Tetrahymena* have an unusual H1 lacking a central globular core.

Histone evolution in multigene families

The organization of histone genes varies widely among eukaryotes. At one extreme, *Drosophila* and sea urchins have multiple tandem repeats of hundreds of histone gene clusters. At the other extreme, yeast and *Tetrahymena* have two or three genes encoding each histone with little clustering and no tandem

repetition. Our analysis suggests that some histone proteins, such as the H2A.F/Z variants, sea urchin testis-specific H2Bs, and rat testis-specific H2A and H2B, have evolved differently than the major histones, are under different selective pressures, and therefore have different functions. In other cases where more than one histone gene has been sequenced from a given organism, the proteins are identical or nearly identical. This could be due to functional constraints in the protein or to concerted evolution, a process by which genes in large multigene families tend to become homogeneous (26).

There is no evidence that different types of histone gene organization have influenced the evolution of histone proteins either within or among species. The two major H2As from *Tetrahymena*, for example, are no more different from each other

Figure 3. Phylogenetic tree of histone H3 proteins. Neighbor-joining tree drawn from 73 aligned H3 protein sequences. The scale bar indicates the branch length that corresponds to 0.05 substitutions per position (the alignment contains 137 positions). Boxed sequences indicate replication-independent H3 variants. Plant H3 nomenclature is according to (30). GenBank and PIR accession numbers are given below (PIR indicated by *). Tetrahymena major H3 includes *T. thermophila* HHT1 and HHT2, M87304 and M87504, and *T. pyriformis* H3, A28852*. Plant H3.III is *Arabidopsis* H3, X60429. Plant H3.II includes Alfalfa, X13673; Pea, A02631*; Wheat, A26014*. Plant H3.I includes *Arabidopsis*, M35387; Maize, M36658; Parsley, M77493; Rice, X13678. Animal H3.3 includes Chicken H3.3A, M11392, M11667; Chicken H3.3B, M11393; Clam, M17876; *D. melanogaster* H3.3q, X53822; Human H3.3, X05855, X05856, X05857; Mouse H3.3, X13605; Rabbit H3-3A, X51897. Sea urchin major H3 includes *D. imbricatus*, X07505; *L. pictus* late H3, X00593; *L. pictus* early H3, X00628; *P. brevispinus*, X54112; *P. helianthoides*, X54114; *P. lividus*, M25281; *P. mliaris*, V01140; *P. ochraceus*, X54113; *S. drobachiensis*, M36921; *S. stimpsoni*, X54115. Animal major H3 includes Buffalo fish, A02627*; Cat shark, A02626*; Chicken, X02218; Chicken H3-IV, X62291; Chicken H3-V, X62292; *Chironomus thummi*, X56335; Cow H3.1, A02624*; *D. hydei*, X17072; Duck H3.1, X14732; Mouse H3.2, M33989; *Platyneris*, X53330; Trout, X01064; *Tigriopus californicus*, X52393; *Urechis caupo*, X58895; *Xenopus* H3.1, A24279*. Others are 1, L02418; 2, B28852*; 3, M87305; 7, X14230; 8, X00724; 9, X00725; 10, X55548; 11, X01612; 12, X05222; 13, X05223; 14, X05224; 22, X06963; 23, X06964; 25, A23604*; 26, M15664; 31, M35867; 35, X15634; 36, M26150; 37, A02624*; 38, X57128; 39, M32460; 40, J01175; 41, X03952; 42, M36919; 53, J00984; 54, L11067; 55, M14396; 56, X14215; 57, X16148; 58, M61155.

than the H2As of human or mouse, and there are no consistent differences in the rates of histone protein evolution between *Drosophila*, sea urchins, yeast, and *Tetrahymena*. Because of our interest in the function of the histones per se, we have limited our analysis to the evolution of the proteins themselves. Therefore we can not comment on the effects of different patterns of gene organization on intra-species homogeneity or on the evolution of silent substitutions in histone genes. However, the functional

Figure 4. Phylogenetic tree of histone H4 proteins. Sixty-three H4 protein sequences were obtained, aligned, and used to construct a neighbor-joining tree. The scale bar indicates the branch length that corresponds to 0.05 substitutions per position (the alignment contains 104 positions). GenBank and PIR accession numbers are given below (PIR indicated by *). Yeast H4 includes *S. cerevisiae* H4.1, X00724; *S. cerevisiae* H4.2, X00725; *S. carlsbergensis*, K03154. Plant major H4 includes *Arabidopsis* H4ga, M17132; *Arabidopsis* H4gb, M17133; Maize H4c7, M13377; Maize H4c14, M13370; Pea, A02646*; Wheat, X00043. Invertebrate H4 includes *Chironomus thummi*, X56335; *D. hydei*, X52576; *D. melanogaster*, X14215; *Tigriopus californicus*, X52393. Sea urchin H4 includes *C. elegans*, X15634; *L. pictus* late H4, X00593; *P. brevispinus*, X54112; *P. helianthoides*, X54114; *P. mliaris* late H4, M10556; *P. ochraceus*, X54113; *S. purpuratus* early H4, J01172; *S. purpuratus* late H4.1, X06639; *S. purpuratus* late H4.2, X06641. Animal major H4 includes Cow, A02637*; Chicken H4A, M74533; Chicken H4B, M74534; Chicken H4L, X02218; Duck, X14732; Human, X67081; Mouse H4a, X13235; Pig, A02638*; *Platyneris*, X53330; Rat testes H4, M28409; *Tilapia nilotica*, X54078; Trout, X02916; *Urechis caupo*, X58895; *Xenopus borealis*, J00985; *Xenopus laevis* H4, J00987; *Xenopus laevis* H4xlh1, X03017; *Xenopus laevis* H4xlh3, X03018. Others are 1, A02648*; 2, A02649*; 3, X04756; 4, X04755; 5, M24411; 6, X16019; 7, X16018; 8, Z15134; 9, Z15135; 13, X05222; 14, X05223; 15, X05224; 16, X55550; 17, X01611; 18, X55549; 19, X15142; 20, X15141; 21, X06963; 22, X06964; 23, X69179; 24, M12277; 31, L11067; 36, X54115; 46, X13236.

constraint that histone proteins appear to be under might be expected to limit the allowable DNA sequence heterogeneity even without mechanisms of concerted evolution acting directly on the DNA sequences.

Histone diversity and the role of the nucleosome

The differences in the rates of evolution between H3/H4 and H2A/H2B can be rationalized in terms of current models of chromatin assembly and function (see 1,2 for reviews). During DNA replication, an (H3/H4)₂ tetramer is the first particle to assemble on newly synthesized DNA, and this tetramer has been shown to position itself at the same sites as whole nucleosomes (27). Two H2A/H2B dimers assemble later, appear to interact with transcription factors and may release or unfold from the nucleosome upon transcription (28,29). Thus, it is likely that H3 and H4 play critical roles in DNA binding and chromatin assembly, roles that have changed little over the course of eukaryotic evolution and which impose strict constraints on their

structure. H2A and H2B, which may interact with non-histone chromatin proteins, transcription factors, and the transcription apparatus, are probably not only permitted but required to evolve to optimize interactions with other proteins as they have evolved along different eukaryotic lineages.

ACKNOWLEDGEMENTS

The authors are grateful to Drs Howard Ochman and Ulfar Bergthorsson for advice on carrying out the phylogenetic analysis, to Norman Murray for computer support, and to Jody Bowen for critically reading the manuscript. This research was supported by Public Health Service Grant GM-21793 from the National Institutes of Health and grant NP-717A from the American Cancer Society.

REFERENCES

1. Ausio, J. (1992) *J. Cell Sci.*, **102**, 1–5.
2. Wolffe, A.P. (1991) *J. Cell Sci.*, **99**, 201–206.
3. Felsenfeld, G. (1992) *Nature*, **355**, 219–224.
4. Svaren, J. and Htmrz, W. (1993) *Curr. Opin. Genet. Dev.*, **3**, 219–225.
5. Zweidler, A. (1984) In Stein, G.S., Stein, J.L. and Marzluff, W.F. (eds.), *Histone Genes*. John Wiley & Sons, New York, pp. 339–369.
6. Ferrari, N., Pfeffer, U., Profumo, A. and Vidali, G. (1992) *Biochem. Int.*, **28**, 239–248.
7. Waterborg, J.H. (1993) *J. Biol. Chem.*, **268**, 4912–4917.
8. Wunsch, A.M. and Lough, J. (1990) *Cell Differ. Dev.*, **30**, 19–25.
9. Stargell, L.A., Bowen, J., Dadd, C.A., Dedon, P.C., Davis, M., Cook, R.G., Allis, C.D. and Gorovsky, M.A. (1994) *Genes & Dev.*, in press.
10. DeLange, R.J., Fambrough, D.M., Smith, E.L. and Bonner, J. (1969) *J. Biol. Chem.*, **244**, 5669–5679.
11. Wells, D. and McBride, C. (1989) *Nucleic Acids Res.*, **17 Suppl.**, r311–r346.
12. Wells, D. and Brown, D. (1991) *Nucleic Acids Res.*, **19 Suppl.**, 2173–2188.
13. Benson, D., Lipman, D.J. and Ostell, J. (1993) *Nucleic Acids Res.*, **21**, 2963–2965.
14. Devereux, J., Haerberli, P. and Smithies, O. (1984) *Nucleic Acids Res.*, **12**, 387–395.
15. Feng, D-F. and Doolittle, R.F. (1987) *J. Mol. Evol.*, **25**, 351–360.
16. Saitou, N. and Nei, M. (1987) *Mol. Biol. Evol.*, **4**, 406–425.
17. Sogin, M.L. (1989) *Amer. Zool.*, **29**, 487–499.
18. Woese, C.R. (1987) *Microbiological Reviews*, **51**, 221–271.
19. Wainright, P.O., Hinkle, G., Sogin, M.L. and Stickel, S.K. (1993) *Science*, **260**, 340–342.
20. Gaertig, J., Thatcher, T.H., McGrath, K.E., Callahan, R.C. and Gorovsky, M.A. (1993) *Cell Motil. Cytoskel.*, **25**, 243–253.
21. McGrath, K.E., Yu, S.-M., Heruth, D.P., Kelly, A.A. and Gorovsky, M.A. (1994) *Cell Motil. Cytoskel.*, in press.
22. Van Daal, A., White, E.M., Elgin, S.C.R. and Gorovsky, M.A. (1990) *J. Mol. Evol.*, **30**, 449–455.
23. Van Daal, A. and Elgin, S.C.R. (1992) *Mol. Biol. Cell*, **3**, 593–602.
24. Thatcher, T.H., MacGaffey, J., Horowitz, S.B., Shapiro, D.L. and Gorovsky, M.A. (1994) *Nucleic Acids Res.*, in press.
25. Mannironi, C., Bonner, W.M. and Hatch, C.L. (1989) *Nucleic Acids Res.*, **17**, 9113–9126.
26. Matsuo, Y. and Yamazaki, T. (1989) *Genetics*, **122**, 87–97.
27. Hayes, J.J., Clark, D.J. and Wolffe, A.P. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 6829–6833.
28. Baer, B.W. and Rhodes, D. (1983) *Nature*, **301**, 482–488.
29. Louters, L. and Chalkley, R. (1985) *Biochem.*, **24**, 3080–3085.
30. Waterborg, J.H. (1991) *Plant Physiol.*, **96**, 453–458.