DEAE-Spherodex column with the use of a phosphate gradient. The purified M.Ab3c9 and the M.AbC (M.IgG, Nordic Immunology) were conjugated to Affi-Gel10 (Bio-Rad) columns (2.5 mg antibody per milliliter of swollen gel). The columns were then equilibrated in 50 mM tris-HCl (pH 7.8), glycerol (20%), 1 mM EDTA, 500 mM KCl, and NP-40 (0.01%).
14. Protein immunoblotting was performed as described (*15*), except that the proteins were detected with the Amersham chemiluminescence protein immunoblotting detection reagents.
15. N. Burton, B. Cavallini, M. Kanno, V. Moncollin, J. M. Egly, *Protein Expression Purif.* **2**, 432 (1991); V. Moncollin, L. Schaeffer, C. Chalut, J. M. Egly, *ibid.*, in press.
16. C. Chalut, and J. M. Egly, unpublished results.
17. O. Flores, E. Maldonado, D. Reinberg, *J. Biol. Chem.* **264**, 8915 (1989).
18. D. Reinberg, J. Conaway, R. Conaway, personal communications.
19. W. J. Feaver, O. Gileadi, R. G. Kornberg, *ibid.* **266**, 19000 (1991); O. Gileadi, W. J. Feaver, R. G. Kornberg, unpublished results.
20. S. Groth and F. S. Scheidegger, *J. Immunol. Methods* **35**, 1 (1980).
21. L. M. Reick *et al.*, *ibid.* **100**, 123 (1987).
22. We thank J. M. Chipoulet and A. Fery for technical assistance; A. Staub and D. Black for microsequencing and oligonucleotide synthesis; J. M. Garnier for the HeLa cDNA library; and S. Ali, H. Gronemeyer, V. Moncollin, and R. Roy for critical reading of the manuscript. We also thank the Reinberg laboratory for TFIIE antibody. This work was supported by grants from the INSERM, the CNRS, the Ministère de la Recherche et de l'Enseignement Supérieur, the Fondation pour la Recherche Médicale, the Association pour la Recherche sur le Cancer, and the Ligue Nationale contre le Cancer.

27 April 1992; accepted 22 June 1992

Modulation of the Dimerization of a Transcriptional Antiterminator Protein by Phosphorylation

Orna Amster-Choder and Andrew Wright

The transcriptional antiterminator protein BglG inhibits transcription termination of the *bgl* operon in *Escherichia coli* when it is in the nonphosphorylated state. The BglG protein is now shown to exist in two configurations, an active, dimeric nonphosphorylated form and an inactive, monomeric phosphorylated form. The migration of BglG on native polyacrylamide gels was consistent with it existing as a dimer when nonphosphorylated and as a monomer when phosphorylated. Only the nonphosphorylated dimer was found to bind to the target RNA. When the dimerization domain of the λ repressor was replaced with BglG, the resulting chimera behaved like an intact λ repressor in its ability to repress λ gene expression, which suggests that BglG dimerizes in vivo. Repression by the λ-BglG hybrid was significantly reduced by BglF, the BglG kinase, an effect that was relieved by conditions that stimulate dephosphorylation of BglG by BglF. These results suggest that the phosphorylation and the dephosphorylation of BglG regulate its activity by controlling its dimeric state.

Control of transcription involves the interaction between specific DNA or RNA sequence elements and protein factors (*1*, *2*), many of which have been shown to bind DNA as dimers (*3–5*). Changes in the oligomeric state of transcription factors, which can change their sequence specificity or activation function, increase the diversity of regulation by a limited number of regulatory molecules in the cell. The regulation of dimerization has not yet been shown to participate in the regulation of transcription by RNA binding proteins.

The processes that control the switch between dimers and monomers in the cell are poorly understood. These processes may involve interaction with other cellular accessory proteins or covalent modification of the proteins being dimerized. In the case of hepatocyte nuclear factor–1α, a mammalian homeodomain protein, a protein dimerization cofactor has been identified (6). In no case has a covalent modification been shown to be involved in the regulation of the dimerization process.

Department of Molecular Biology and Microbiology, Tufts University Health Sciences Campus, Boston, MA 02111.

Transcription of the *bgl* operon in *E. coli* is regulated by reversible protein phosphorylation (*7*, *8*). BglG is a sequence-specific RNA binding protein whose ability to inhibit termination of transcription of the *bgl* operon is modulated by BglF, a membrane-bound protein kinase–phosphatase that senses the presence of β glucosides (*7–10*).

To increase our understanding of how BglG functions as a transcriptional antiterminator, we constructed a series of hybrid proteins that contained various portions of BglG and complementary portions of Sac Y, a similar antiterminator protein from *Bacillus subtilis* (*11*). The hybrids (*12*) were tested for their ability to prevent termination of transcription of a *bgl-lacZ* fusion gene in *E. coli*. One of the hybrids failed to antiterminate, giving only 2 to 3 units of β-galactosidase activity compared to 60 units obtained with wild-type BglG. Moreover, this hybrid exerted a dominant negative effect on antitermination by wild-type BglG co-expressed in the same cell, reducing β galactosidase production from 60 units to 2 to 3 units. One possible explanation for this result is that the dominance resulted from the formation of an inactive hetero-oligomer between the hybrid protein and BglG, which implies that an oligomeric form of BglG is the active species in antitermination.

To test this hypothesis, we determined the molecular size of native BglG on a set of nondenaturing polyacrylamide gels (Fig. 1A). The molecular size of nonphosphorylated BglG, estimated by extrapolation of its coefficient of retardation (Fig. 1), was ~60,000 daltons. The molecular size of BglG, calculated from its amino acid sequence and confirmed by denaturing gels, is 32,067 daltons. On the basis of this result, we conclude that BglG exists as a dimer. The molecular size of the phosphorylated form of BglG, determined with the same method, was ~30,000 daltons (Fig. 1B), which suggests that phosphorylated BglG exists as a monomer.

We obtained direct evidence that dimeric BglG binds to its RNA target although monomeric phosphorylated BglG does not by probing the proteins in their native state with RNA (Fig. 2). Extracts prepared from bgl^0 and bgl^+ *E. coli* strains, enriched respectively for either the dimer or the phosphorylated monomer, were fractionated on native gels, blotted onto nitrocellulose filters, and probed with labeled target RNA. Conditions that allowed the separation of the two forms of BglG were chosen on the basis of their known relative mobilities (R_m) in native gels (Fig. 1). Our results indicate that the RNA probe bound to the dimeric nonphosphorylated form of BglG but did not bind to the monomeric phosphorylated form (Fig. 2A). By probing the extracts, after fractionation by SDS–polyacrylamide gel electrophoresis (SDS-PAGE), with anti-BglG antibodies (Fig. 2B), we were able to show that the two extracts contained the same total amount of phosphorylated and nonphosphorylated BglG.

To further establish that BglG functions as a dimer, we used the NH_2-terminal DNA binding domain of the bacteriophage λ repressor as a reporter for dimerization (*13*). The NH_2-terminal domain of the λ repressor does not dimerize and thus requires a dimerization domain to allow it to bind strongly to its operator and repress transcription. To test the ability of BglG to serve as a dimerization domain, we constructed a gene fusion between the NH_2-terminal coding sequence of the λ repressor and the complete *bglG* gene (λ-*bglG*). The regulatory properties of the chimeric λ-BglG protein were compared to the properties of intact λ repressor (*ind1*), λ-zip, a λ-leucine zipper chimera that behaves like an intact λ repressor (*13*), the DNA binding domain of the λ repressor (amino acids 1 to 131), and λ-zip interrupted by truncat-

ed *lacZ* sequences (lacZ′) (λ-zip::*lacZ*′) (*14*). These proteins were all expressed at low concentrations from equivalent plasmid constructs; their properties are shown in Table 1. Bacterial cells that expressed λ-BglG, intact λ repressor, or λ-zip were immune to infection by bacteriophage λ, whereas those that contained the λ DNA binding domain alone or the λ-zip::*lacZ*′ were sensitive, as demonstrated by cross-streak, plaque formation, and colony-nibbling assays. Cells that contained λ-BglG exhibited 74% repression of expression from a λP_R-*lacZ* fusion, somewhat lower than the 93% repression given by the intact λ repressor and the 95% given by λ-zip, but significantly higher than the 0% given by the λ DNA binding domain alone or the 7% given by λ-zip::*lacZ*′ (Table 1). Thus, fusion of BglG to the DNA binding domain of the λ repressor resulted in a stable, biologically active dimer.

The chromosomal *bgl* operon in the strain used for the experiment described in Table 1 is cryptic (*bgl*^{0}), although a very small amount of its expression can be detected (*15*). The small amount of BglF produced in this strain was previously shown to phosphorylate ~20% of BglG overproduced from a plasmid (*8*). The native gel analysis suggested that phosphorylated BglG exists as a monomer. Thus, the lower repression exhibited by λ-BglG, as compared to the intact λ repressor and λ-zip, might be explained by phosphorylation leading to the monomerization of ~20% of the λ-BglG hybrid proteins. To assess the influence of phosphorylation of BglG on its ability to dimerize in vivo, we tested the ability of λ-BglG to repress expression of the λP_R-*lacZ* fusion in *bgl*$^+$ cells (Table 2, −Salicin). In this background, expression of the *bgl* operon is inducible, but even without an inducer, the basal amount of expression is higher than in the *bgl*^{0} background (*15*). The amount of BglF produced in the *bgl*$^+$ strain without induction was previously shown to be enough to lead to phosphorylation of 50 to 60% of the BglG being overproduced in these cells (*8*). Repression of *lacZ* expression by λ-BglG dropped from 74% in the *bgl*^{0} background to

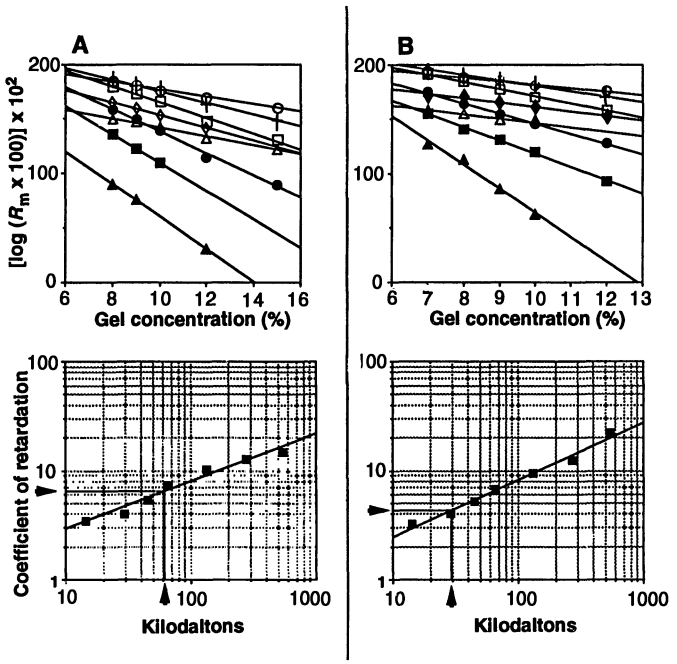

Fig. 1. Determination of molecular sizes of (**A**) BglG and (**B**) phosphorylated BglG on nondenaturing polyacrylamide gels. BglG was labeled either in vivo with [^{35}S]methionine or in vitro with [^{32}P]phosphate as described (*7*). The procedure used for determining the molecular size of BglG and phosphorylated BglG is a modification of Bryan (*18*) and involves the use of a nondenatured protein molecular size kit (Sigma). Aliquots of cell extracts and molecular size markers from the kit were subjected to electrophoresis on a set of native gels that contained various concentrations of acrylamide. The molecular size markers were revealed by Coomassie blue staining, and ^{35}S-labeled BglG and ^{32}P-labeled BglG were detected by autoradiography. A function of the relative mobilities (R_m) was plotted versus the concentration of acrylamide for each protein to construct Ferguson plots (upper graphs), and the coefficient of retardation (K_r) of each species was deduced from the slopes. The logarithm of the coefficient of retardation of the markers was then plotted versus the logarithm of their molecular sizes (lower graphs). The plots obtained were used to determine the molecular sizes of (A) BglG or (B) P-BglG by extrapolation of their respective K_r's (arrowheads). Open circles, α-lactalbumin; open triangles, carbonic anhydrase; daggers, chicken egg albumin; open squares, bovine serum albumin (monomer); closed circles, bovine serum albumin (dimer); closed squares, urease (trimer); closed triangles, urease (hexamer); open diamonds, BglG; and closed diamonds, [^{32}P]BglG.

Table 1. Properties of plasmids expressing hybrid proteins containing the λ repressor DNA binding domain. The replacement of the dimerization domain of the λ repressor with BglG results in a stable dimer that acts as an intact λ repressor. The indicated plasmids, which are derivatives of pZ150 (*20*), were introduced by transformation into the λ-sensitive *bgl*^{0} strain AG1688 (*21*) to test for sensitivity to the cI$^-$ phage λKH54 or into JH372 [AG1688(λ202)] (*13*) to measure in vivo binding to λO_R1. pJH157 expresses the intact λ repressor (*13*), pOAC100 expresses the DNA binding domain (amino acids 1 through 131) of the λ repressor (*22*), pOAC101 expresses the DNA binding domain fused to the entire *bglG* gene (*22*), and pJH370 expresses the DNA binding domain fused to the leucine zipper of GCN4 (*13*). pJH391 is a derivative of pJH370 containing a truncated *lacZ* insertion (*14*). We confirmed sensitivity to infection (i) by testing the ability of the phage to form plaques on the various strains (*23*), (ii) by testing the ability of the strains to grow on plates when streaked through a phage lysate, and (iii) by the colony nibbling assay (*24*). We assayed the binding to the λ operators by measuring β-galactosidase activity expressed from a P_R-*lacZ* fusion present on the λ202 prophage found in strain JH372. Because in this phage O_R contains an O_R2^- mutation, the intact λ repressor does not show cooperative binding; its activity can be readily compared with the activity of the chimeric proteins derived from it. Repression is calculated as: 1 − (β-galactosidase activity with repressor/β-galactosidase activity with no repressor).

Plasmid	Repressor	Sensitivity to λKH54	Repression of λP_R-*lacZ*	
			β-Gal units	Repression (%)
pZ150	None	Sensitive	2202	0
pOAC100	Amino acids 1–131	Sensitive	2287	0
pJH391	λ-zip::*lacZ*′	Sensitive	2048	7
pJH157	*ind1*	Immune	144	93
pOAC101	λ-BglG	Immune	564	74
pJH370	λ-zip	Immune	100	95

Table 2. A negative effect of BglF, the BglG kinase, on the repressor-like activity of the λ-BglG chimera that can be relieved by β glucosides. A spontaneous *bgl*$^+$ derivative of strain JH372 was isolated on MacConkey salicin medium (Difco, Detroit, Michigan) and transformed with the indicated plasmids. The strains were grown in minimal medium, and salicin was added to a final concentration of 0.05% where indicated. Repressor activity was measured as described (Table 1).

Plasmid	Repressor	Repression of λP_R-*lacZ* (%)	
		Salicin (−)	Salicin (+)
pOAC100	Amino acids 1–131	0	0
pJH391	λ-zip::*lacZ*′	6	7
pJH157	*ind1*	94	91
pOAC101	λ-BglG	22	54

22% in the *bgl*$^+$ background (Tables 1 and 2). Thus, the efficiency of repression by λ-BglG appears to be inversely proportional to the concentration of BglF in the cell. No differences were seen in the repression efficiency of the other λ repressor derivatives in the *bgl*^{0} and *bgl*$^+$ strains (Tables 1 and 2).

Dephosphorylation of BglG by BglF occurs in cells after the addition of β glucosides to the growth medium (*8*). To see if addition of β glucosides influenced the repressor activity of λ-BglG, we measured *lacZ* expression from the λP_R-*lacZ* fusion in *bgl*$^+$ cells, which contained the various λ repressor derivatives grown in the presence of the β glucoside salicin (Table 2, +Salicin). The addition of salicin did not affect the efficiency of repression by the intact λ repressor, the λ DNA binding domain, or the λ-zip::*lacZ*′ (Tables 1 and 2). In contrast, repression by λ-BglG increased from 22% in the absence of salicin to 54% in its presence (Tables 1 and 2). Thus, β glucosides relieved the negative effect of BglF on the ability of λ-BglG to repress λ gene expression. The amount of repression observed in this case (54%) was lower than that seen in the *bgl*^{0} strain (74%) and was not affected by variation of the concentration of salicin in the growth medium (*16*).

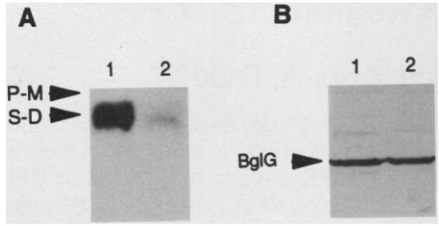

Fig. 2. Target RNA binds to dimeric BglG. Extracts from *E. coli* K38 *bgl*^{0} (lanes 1) and *E. coli* K38 *bgl*$^+$ (lanes 2), each containing plasmids pT7FH-G and pGP1-2, were prepared as described (*8*). Proteins were fractionated on 7% (**A**) and 12% (*16*) native gels or on a 12.5% SDS–polyacrylamide gel (**B**), transferred onto nitrocellulose filters (*19*), and probed either with ^{32}P-labeled RNA (A) (*10*) or with anti-BglG antibody (B). The RNA probe, R83, that contains the BglG binding site, was prepared as described (*10*). The positions of the two BglG forms on native gels, known from the analysis described in Fig. 1, were further established with the use of [^{32}P]BglG (monomer; P-M) and [^{35}S]BglG (dimer; S-D), labeled in vitro and in vivo, respectively, as described (*7*), and by the use of anti-BglG antibodies. [^{35}S]BglG together with molecular size standards were used to determine the position of BglG in the SDS-PAGE protein immunoblot analysis.

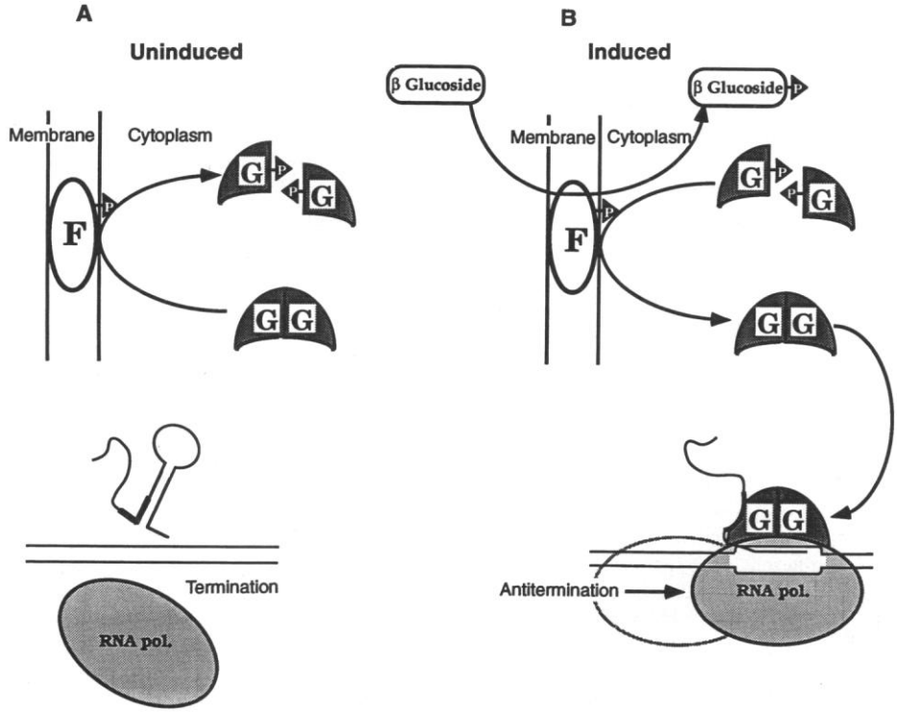

Fig. 3. Model for the mechanism of *bgl* operon regulation. (**A**) Uninduced. (**B**) Induced. The region in the *bgl* mRNA depicted by a heavy line represents the BglG binding site, which forms a secondary structure not illustrated in the model (*10, 17*). The graphic illustration showing the interaction between two BglG molecules and RNA polymerase under inducing conditions represents a more general idea of one or two BglG molecules interacting with either RNA polymerase or auxiliary proteins. F, BglF; G, BglG; RNA pol., RNA polymerase; and P, phosphate.

This observation may be explained by the ability of wild-type BglG protein, which is produced in stoichiometric amounts as a result of induction of the *bgl* operon in the *bgl*$^+$ strain, to titrate the λ-BglG through heterodimer formation.

We have proposed a mechanism for the regulation of the *bgl* operon in which BglG antitermination activity is regulated by reversible phosphorylation (*7*). The results presented herein suggest a possible model for *bgl* regulation in which the ability of BglG to bind RNA and to antiterminate depends on dimer formation, a process that is controlled by phosphorylation (Fig. 3). Under noninducing conditions, BglF transfers phosphate to BglG, interfering with its dimerization, preventing it from binding RNA, and thus blocking its action; a terminator structure is formed on the RNA transcript, and RNA polymerase dissociates from the template. The addition of β glucosides stimulates dephosphorylation of BglG by BglF; nonphosphorylated BglG can dimerize, bind its RNA target, and function as a transcriptional antiterminator, enabling RNA polymerase to transcribe through the operon. Implicit in the model is the idea that BglG must dimerize in order to function as an antiterminator.

Many dimeric proteins that regulate transcription bind to symmetric sequences on DNA. BglG recognizes and binds to a specific sequence of *bgl* mRNA that does not possess obvious twofold symmetry (*10, 17*). Thus, the need for BglG dimers to bring about antitermination might reflect a concomitant interaction with RNA and RNA polymerase or other cellular proteins. Alternatively, the dimer might be needed for interaction with RNA in a way that differs from dimer-DNA interactions.

REFERENCES AND NOTES

1. P. J. Mitchell and R. Tjian, *Science* **245**, 371 (1989).
2. A. D. Frankel, I. W. Mattaj, D. C. Rio, *Cell* **67**, 1041 (1991).
3. N. Jones, *ibid.* **61**, 9 (1990).
4. W. H. Landschulz, P. F. Johnson, S. L. McKnight, *Science* **240**, 1759 (1988).
5. C. Murre, P. S. McCaw, D. Baltimore, Cell **56**, 777 (1989).
6. D. B. Mendel *et al.*, *Science* **254**, 1762 (1991).
7. O. Amster-Choder, F. Houman, A. Wright, *Cell* **58**, 847 (1989).
8. O. Amster-Choder and A. Wright, *Science* **249**, 540 (1990).
9. S. Mahadevan and A. Wright, *Cell* **50**, 485 (1987).
10. F. Houman, M. R. Diaz-Torres, A. Wright, *ibid.* **62**, 1153 (1990).
11. M. Steinmetz, S. Aymerich, G. Gonzy-Treboul, D. Le Coq, in *Genetics and Biotechnology of Bacilli*, A. T. Ganesan and J. A. Hoch, Eds. (Academic Press, New York, 1988), vol. 2, pp. 11–15.
12. O. Amster-Choder, C. A. Bascom, A. Wright, unpublished data.
13. J. C. Hu, E. K. O'Shea, P. S. Kim, R. T. Sauer, *Science* **250**, 1400 (1990).
14. J. Hu, personal communication.
15. A. E. Reynolds, S. Mahadevan, S. F. J. LeGrice, A. Wright, *J. Mol. Biol.* **191**, 85 (1986).

16. O. Amster-Choder and A. Wright, unpublished data.
17. M. R. Diaz-Torres and A. Wright, unpublished data.
18. J. K. Bryan, *Anal. Biochem.* **78**, 523 (1977).
19. H. Towbin, T. Staehlin, J. Gordon, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4350 (1979).
20. R. J. Zagursky and M. L. Berman, *Gene* **27**, 183 (1984).
21. AG1688 [*ara*D139 Δ(*ara-leu*)7697 Δ*lac*X74 *gal*E *gal*K *rps*L *hsd*R (F'128 *laci*^{q} *lacZ*::*Tn5*)] is the same as strain AG115 described by D. J. Rudner, J. R. LeDeaux, K. Ireton, A. D. Grossman, *J. Bacteriol.* **173**, 1388 (1991).
22. pOAC100 was constructed by the deletion of the *zip::lacZ'* from pJH391 and the introduction of a translational stop codon. We constructed pOAC101 by replacing the *zip::lacZ'* in pJH391 with *bglG*.
23. The number of plaques in the plaque assay ranged from 0 for strains immune to the phage to 90 to 110 for strains sensitive to the phage.
24. H. R. Revel, *Virology* 31, 688 (1967).
25. We thank M. Steinmetz for providing *sac Y* plasmids; J. Hu and R. T. Sauer for advice on the use of the λ repressor to study dimerization and for providing plasmids and bacterial strains; G. Tran Van Nhieu for advice on the native gel system; and A. L. Sonenshein and C. A. Kumamoto for critical reading of the manuscript.

16 March 1992; accepted 10 July 1992

MacroH2A, a Core Histone Containing a Large Nonhistone Region

John R. Pehrson*† and Victor A. Fried

A histone, macroH2A, nearly three times the size of conventional H2A histone, was found in rat liver nucleosomes. Its N-terminal third is 64 percent identical to a full-length mouse H2A. However, it also contains a large nonhistone region. This region has a segment that resembles a leucine zipper, a structure known to be involved in dimerization of some transcription factors. Nucleosomes containing macroH2A may have novel functions, possibly involving interactions with other nuclear proteins.

The nucleosome, as the basic structural unit of chromatin, must be adapted to diverse chromatin functions such as DNA packaging, transcription, replication, and mitotic chromosome segregation. Replication and transcription, for instance, require a disruption of nucleosomal structures to allow polymerases and accessory factors access to the DNA. The mechanisms used to alter nucleosomal structures are not well understood.

One way of varying nucleosome structure and function is to alter its protein composition. To identify proteins that may be involved in producing specialized nucleosome structures, we looked for proteins that remained bound to mononucleosomes during sedimentation in 0.5 M NaCl, conditions in which the core histones remain bound, but H1 and most nonhistone proteins are removed (*1*). Two proteins with a molecular mass of ~42 kD cosedimented with mononucleosomes under these conditions (Fig. 1A). These two proteins had similar electrophoretic mobilities in conventional SDS-polyacrylamide gels but were resolved in less cross-linked gels (Fig. 1B). Two proteins of identical electrophoretic properties cochromatographed with mononucleosomes and oligonucleosomes through molecular sieve columns (*2*); a preparation from oligonucleosomes is shown in Fig. 1B. The copurification of these two proteins with mononucleosomes during both sedimentation and gel filtration, which are oppositely affected by molecular shape, and also with oligonucleosomes, indicated that they were associated with the nucleosome.

Proteins other than the core histones and 42-kD proteins were also present in the mononucleosome region of the gradient (Fig. 1A). The proteins at the top of the gel probably sediment in this region because of their large size rather than because of their association with nucleosomes. However, some of the smaller proteins present in this region may be associated with nucleosomes.

To examine the association of the 42-kD proteins with the nucleosome more closely, we used a stepwise salt elution procedure developed by Simon and Felsenfeld (*3*) to resolve different core histones. The 42-kD proteins were present in the H2A and H2B fraction. Relative to H2A and H2B, the 42-kD proteins were diminished at the beginning of the elution of H2A and H2B and enriched at the end (Fig. 1C, lanes 3 to 5). Most importantly, the later fractions contained less H2A than H2B (Fig. 1C, lanes 5 and 6); scans of lane 6 indicate the deficit of H2A to be approximately 25% (*4*). In the early fractions, H2A and H2B are present in essentially equal amounts (Fig. 1C, lane 7), consistent with observations that they are eluted from the chromatin as a heterodimer. The gel scans indicated that the deficit of H2A in later fractions could be accounted for by the presence of the 42-kD proteins on a molar basis. The simplest explanation for the deficit of H2A, therefore, is that the 42-kD proteins form dimers with H2B. To confirm that this band contained the two 42-kD proteins that copurified with mononucleosomes and oligonucleosomes, we ran samples from this experiment in an SDS gel with low cross-linking (Fig. 1C, lanes 8 to 10).

The two 42-kD proteins were excised from SDS gels (Fig. 1B) and digested with trypsin or V8 protease for peptide mapping (*5*). These maps indicated that the two proteins have related primary structures. Of five peptides sequenced (*6*), three proved to be similar to regions of histone H2A, and the other two were unlike any protein in the SwissProt database. A polymerase chain reaction (PCR) primed with degenerate oligonucleotides based on the peptide sequences ARDNKK and KEFVEA obtained from the 42-kD proteins amplified a DNA fragment of ~550 bp from a rat lymphoma cDNA library (*7*). This fragment encoded sequences identical to the H2A-like peptides obtained from the 42-kD proteins (*8*).

The cloned PCR product was used to screen a rat liver cDNA library (*9*). Sequencing of a positive clone (*10*) revealed a single large open reading frame with an ATG context conforming to the consensus sequence of Kozak (*11*), indicating a probable site for initiation. This ATG is preceded by two in-frame stop codons. The open reading frame encodes a 39-kD protein consisting of 367 amino acids and contains sequences identical to all five peptides obtained from the 42-kD proteins (Fig. 2A). The NH_2-terminal third (amino acids 1 to 122) of the encoded protein is typical of full-length H2As, diverging mainly in regions known to differ between H2A subtypes within an organism and between species (*12*) (Fig. 2B).

The similarity of the NH_2-terminal region of this clone to H2A, together with the close association of the 42-kD proteins with nucleosomes and their apparent association with H2B, indicates that the 42-kD proteins are forms of H2A that occupy the place of conventional H2A in a subset of nucleosomes. We have accordingly named these proteins macroH2A (mH2A) and the subtype encoded by the cloned cDNA mH2A.1. The exact relation between mH2A.1 and the two mH2A bands seen in an SDS gel (Fig. 1B) is not known. On the basis of the relative staining intensity of gel bands, we estimate that there is about one mH2A for every 30 nucleosomes in rat liver.

Although the data presented above in-

J. R. Pehrson, Fox Chase Cancer Center, Institute for Cancer Research, Philadelphia, PA 19111.
V. A. Fried, Department of Cell Biology and Anatomy, Basic Sciences Building, New York Medical College, Valhalla, NY 10595.

*To whom correspondence should be addressed.
†Present address: Department of Animal Biology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA 19104.